UNIVERSITY OF PIRAEUS

Department of Informatics

Adaptive Policy-based Security Management

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF INFORMATICS

OF THE UNIVERSITY OF PIRAEUS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Georgios A. Katsikogiannis

Supervisor: Christos Douligeris

Piraeus, 2018

“We can only see a short distance ahead, but we can see plenty there that needs to be done.”

— Alan Turing

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΔΙΑΤΡΙΒΗ

ΠΡΟΣΑΡΜΟΣΤΙΚΗ ΔΙΑΧΕΙΡΙΣΗ ΑΣΦΑΛΕΙΑΣ ΒΑΣΕΙ ΠΟΛΙΤΙΚΩΝ
για την απόκτηση διδακτορικού διπλώματος του Τμήματος
Πληροφορικής του Γεωργίου Α. Κατσικογιάννη

ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΕΠΙΤΡΟΠΗ

ΕΠΙΒΛΕΠΩΝ:
Χρήστος Δουληγέρης
Καθηγητής
Πανεπιστήμιο Πειραιώς

ΜΕΛΗ:
Θεμιστοκλής Παναγιωτόπουλος
Καθηγητής
Πανεπιστήμιο Πειραιώς

Σαράντης Μητρόπουλος
Αναπληρωτής Καθηγητής
Τ.Ε.Ι. Ιονίων Νήσων

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Χρήστος Δουληγέρης
Καθηγητής
Πανεπιστήμιο Πειραιώς

Θεμιστοκλής Παναγιωτόπουλος
Καθηγητής
Πανεπιστήμιο Πειραιώς

Σαράντης Μητρόπουλος
Αναπληρωτής Καθηγητής
Τ.Ε.Ι. Ιονίων Νήσων

Συμεών Παπαβασιλείου
Καθηγητής
Εθνικό Μετσόβιο Πολυτεχνείο

Δημήτριος Μάγος
Καθηγητής
ΤΕΙ Αθήνας

Δημήτριος Βέργαδος
Αναπληρωτής Καθηγητής
Πανεπιστήμιο Πειραιώς

Παναγιώτης Κοτζανικολάου
Επίκουρος Καθηγητής
Πανεπιστήμιο Πειραιώς

This page intentionally left blank.

Abstract

Machine-to-machine (M2M) communications must cope with the growing needs for robust and reliable connections and secure data exchange of the connected devices along with the distributed data volumes of M2M devices as well as play a significant role in the service delivery. Various communication protocols, network technologies and security mechanisms are expected to intensify as well as thrive in the coming years in order to accommodate the needs for a higher offered value of the M2M solutions. The solutions need to support the increasing volume of connected M2M devices, more complex gateways and various business applications with several functions. The technology varies in terms of the adopted devices, application platforms and domains. This variety allows several approaches and architectures ranging from monolithic applications to Service Oriented Architecture (SOA)-based macro-services, emerging lightweight and agile microservices.

Due to the high number of connected devices, the distributed and resource-constrained nature of the devices, the criticality of the M2M applications (e.g. patient monitoring and operation of critical infrastructures), the changing conditions of the environment and the network stability weaknesses, several security aspects arise and should be addressed efficiently. Appropriate security rules need to be enforced to mitigate the security risks and threats in order to ensure the appropriate levels of runtime security maturity. In this context, numerous access control and dynamic authorisations schemes have been proposed to this date with respect to an effective security management for the M2M devices, the M2M communication and the services. Nevertheless, although strengthening network and data security is a primary concern for the solution providers, these efforts entail significant challenges in order to control access to sensitive information and keep critical data secure. Moreover, in the cloud computing era, there is also an increasing interest in advanced and secure access control services that will improve security and compliance. Hence, various security concerns arise on how to protect the resources in a changing environment while supporting the extended functionalities and sharing the data in a secure way, where and when appropriate.

This doctoral dissertation presents a policy-aware Service-Oriented Architecture (SOA) capable of dealing with these emerging challenges while considering the relevant architectural challenges. Policy-based management capabilities are incorporated to improve the security of the M2M components and services, thereby enabling an adaptive policy enforcement of the suitable security controls, ensuring increased agility and delivering better service levels in the field of M2M communications. This approach allows dynamic and fine-grained access controls in order to access protected resources and proposes the integration of the authentication with the authorisation module by using a policy engine.

Lastly, a prototype has been developed to analyse the characteristics of the ASPIDA framework, aiming to illustrate the improvements in performance in comparison with other solutions. This prototype implements the main building blocks in order to evaluate the effectiveness of the proposed approach.

Λέξεις Κλειδιά: Διαχείριση Δικτύων με βάση την πολιτική, επιβολή πολιτικής, μηχανή πολιτικής, πολιτικές ασφάλειας, έλεγχος πρόσβασης, διαχείριση πρόσβασης, αυθεντικοποίηση, δυναμικές εξουσιοδοτήσεις, SLA, ενσωματωμένος διαχειριστής συμβάντων, προσαρμοστική δρομολόγηση, δρομολόγηση QoS, επικοινωνίες M2M, σύννεφο υπολογιστών, αρχιτεκτονικές προσανατολισμένες προς υπηρεσίες, δυνατότητες προσανατολισμένες προς υπηρεσίες, τομέας υπηρεσιών

Keywords: Policy-based Network Management; policy enforcement; policy engine; security policies; access control; access management; authentication; dynamic authorisations; SLA; Embedded Event Manager; adaptive QoS routing; M2M communications; cloud computing; Service Oriented Architecture; Service Oriented Capabilities; service domain

Acknowledgements

I wish foremost to thank Prof. Christos Douligeris for his guidance, allowing me to conduct my research and providing valuable comments with his thoughtful insights and constructive suggestions for improving my work. I am grateful to Assoc. Prof. Sarandis Mitropoulos for his precious advices and the committee members for their valuable expertise. I would also thank the co-authors, Dimitrios Kallergis and Zacharenia Garofalaki, who offered their support and commitment and what we have achieved together and Prof. Gilbert Leppelmeier for his friendship and all the useful discussions.

Finally, I would like to extend my deepest gratitude towards my wife and children, my family and all my friends for encouraging me throughout all these years. I would not have been able to complete this PhD course without them.

Table of Contents

List of Tables

List of Figures

List of Equations

List of Abbreviations

ABAC	**A**ttribute **B**ased **A**ccess **C**ontrol
ABS	**A**nti-Lock **B**raking **S**ystems
AMQP	**A**dvanced **M**essage **Q**ueuing **P**rotocol
AOAC	**A**lways **O**n, **A**lways **C**onnected
aPaaS	**a**pplication **P**latform **a**s **a** **S**ervice
APE	**A**ccess **P**olicy **E**ngine
APIs	**A**pplication **P**rogramming **I**nterfaces
AR	**A**daptive **R**outing
ASN	**A**d-hoc and **S**ensor **N**etworks
ASR	**A**pplication **S**ecurity **R**isk
AuthNSE	**A**uthentication **S**ervice **E**ngine
AuthZSE	**A**uthorisation **S**ervice **E**ngine
BAN	**B**ody **A**rea **N**etwork
BCM	**B**usiness **C**ontinuity **M**anagement
BTG-AC	**B**reak-**T**he-**G**lass **A**ccess **C**ontrol
CapBAC	**Cap**ability-**B**ased **A**ccess **C**ontrol system
CBOR	**C**oncise **B**inary **O**bject **R**epresentation
CBWFQ	**C**lass-**B**ased **W**eighted **F**air **Q**ueuing
CCAAC	**C**apability-based **C**ontext-**A**ware **A**ccess **C**ontrol
CIR	**C**ommitted **I**nformation **R**ate
CLI	**C**ommand **L**ine **I**nterface
CMP	**C**ertificate **M**anagement **P**rotocol
CoAP	**C**onstrained **A**pplication **P**rotocol
CoS	**C**lass **o**f **S**ervice
COSE	**C**BOR **O**bject **S**igning and **E**ncryption
CPS	**C**yber-**P**hysical **S**ystems
CRAC	**C**yber-**P**hysical **A**ccess **C**ontrol solution
CRLs	**C**ertificate **R**evocation **L**ists
CU	**C**ontrol **U**nit
CBWFQ	**C**lass-**B**ased **W**eighted **F**air **Q**ueuing
CWT	**C**BOR **W**eb **T**okens
DAC	**D**iscretionary **A**ccess **C**ontrol
DCapBAC	**D**istributed **Cap**ability-**B**ased **A**ccess **C**ontrol mechanism
DEN	**D**irectory-**E**nabled **N**etwork
dIa	**D**evice **I**nterface **a**pplication
DiffServ	**Diff**erentiated **Serv**ices
DPWS	**D**evices **P**rofile for **W**eb **S**ervices
DTLS	**D**atagram **T**ransport **L**ayer **S**ecurity
ECC	**E**lliptic **C**urve **C**ryptography
EEM	**E**mbedded **E**vent **M**anagers
EIGRP	**E**nhanced **I**nterior **G**ateway **R**outing **P**rotocol
GPS	**G**lobal **P**ositioning **S**ystem
GTRBAC	**G**eneralized **TRBAC**
GUI	**G**raphical **U**ser **I**nterface

IAB	**I**nternet **A**rchitecture **B**oard
IAM	**I**dentity and **A**ccess **M**anagement
IBC	**I**dentity **B**ased **C**ryptography
IdaaS	**Id**entity **a**s **a** **S**ervice
IDP	**Id**entity **P**rovider
IGP	**I**nterior **G**ateway **P**rotocol
IoT	**I**nternet **o**f **T**hings
IPFRR	**IP** **F**ast **R**e**R**oute
IPFIX	**I**nternet **P**rotocol **F**low **I**nformation e**X**port
IPSec	**I**nternet **P**rotocol **S**ecurity
JML	**J**oin/**M**ove/**L**eave
JSON	**J**ava**S**cript **O**bject **N**otation
JWT	**J**SON **W**eb-**T**oken
KLOC	**T**housand **L**ines **O**f **C**ode
KPI	**K**ey **P**erformance **I**ndicator
LFA	Remote **L**oop-**F**ree **A**lternate
M2H	**M**achine to (**2**) **H**uman
M2M	**M**achine to (**2**) **M**achine
MAC	**M**andatory **A**ccess **C**ontrol
MANET	**M**obile **A**d-hoc **Net**works
MCC	**M**obile **C**loud **C**omputing
MEC	**M**obile **E**dge **C**omputing
mIa	**M**2M **I**nterface **a**pplication
MQTT	**M**essage **Q**ueue **T**elemetry **T**ransport
µServices	micro (**µ**) **S**ervices
MRBAC	**M**ulti-**R**ole **B**ased **A**ccess **C**ontrol
NFC	**N**ear **F**ield **C**ommunication
NSCL	**N**etwork **S**ervice **C**apabilities **L**ayer
OAuth	**O**pen standard for **Auth**entication
OCSP	**O**nline **C**ertificate **S**tatus **P**rotocol
OWASP	**O**pen **W**eb **A**pplication **S**ecurity **P**roject
P2P	**P**eer-to- (**2**) **P**eer computing
PAP	**P**olicy **A**dministration **P**oint
PAS	**Policy**-based **A**ssistance **S**ystem
PBM	**Policy**-**B**ased **M**anagement
PCIM	**P**olicy **C**ore **I**nformation **M**odel
PDA	**P**olicy-**D**riven **A**uthorisation
PDP	**P**olicy **D**ecision **P**oint
PEP	**P**olicy **E**nforcement **P**oint
PHB	**P**er-**H**op **B**ehaviour
PIP	**P**olicy **I**nformation **P**oint
PLC	**P**rogrammable **L**ogic **C**ontrollers
PoLP	**P**rinciple **o**f **L**east **P**rivilege
POP	**P**oint **O**f **P**resence
pps	**p**ackets **p**er **s**econd
PRP	**P**olicy **R**etrieval **P**oint
PU	**P**rocessing **U**nit

RA	**R**egistration **A**uthority
RBAC	**R**ole **B**ased **A**ccess **C**ontrol
REST	**RE**presentational **S**tate Transfer
RFID	**R**adio **F**requency **ID**entification
RME	**R**ole **M**apping **E**ngine
RTT	**R**ound **T**rip **T**ime
RTC	**R**eal-**T**ime **C**ommunication
SCADA	**S**upervisory **C**ontrol **a**nd **D**ata **A**cquisition
SCP	**S**ocial-**C**yber-**P**hysical
SCVP	**S**imple **C**ertificate **V**alidation **P**rotocol
SDLC	**S**oftware **D**evelopment **L**ife **C**ycle
SDN	**S**oftware-**D**efined **N**etworking
SIEM	**S**ecur**I**ty and **E**vent **M**anagement
SLS	**S**ervice **L**evel **S**pecification
SOA	**S**ervice-**O**riented **A**rchitecture
SOC	**S**ervice-**O**riented **C**omputing
SoD	**S**egregation **o**f **D**uty
SOAP	**S**imple **O**bject **A**ccess **P**rotocol
SPF	**S**hortest **P**ath **F**irst
SSE-CMM	**S**ystems **S**ecurity **E**ngineering **C**apability **M**aturity **M**odel
SSO	**S**ingle-**S**ign-**O**n
TCA	**T**rusted **C**ertificate **A**uthority
TLS	**T**ransportation **L**ayer **S**ecurity
ToS	**T**ype **o**f **S**ervice
TRBAC	**T**emporal **R**ole-**B**ased **A**ccess **C**ontrol
trTCM	**T**wo **R**ate **T**hree **C**olour **P**olicer
UMA	**U**ser-**M**anaged **A**ccess
UML	**U**nified **M**odelling **L**anguage
V2I	**V**ehicle-to-(**2**) **I**nfrastructure
V2V	**V**ehicle-to-(**2**) **V**ehicle
VA	**V**alidation **A**uthority
VPN	**V**irtual **P**rivate **N**etwork
WSDL	**W**eb **S**ervices **D**escription **L**anguage
WBAN	**W**ireless **B**ody **A**rea **N**etworks
WS	**W**eb-**S**ervice
WSN	**W**ireless **S**ensor **N**etwork
XACML	e**X**tensible **A**ccess **C**ontrol **M**arkup **L**anguage
XML	e**X**tensible **M**arkup **L**anguage
XMPP	e**X**tensible **M**essaging and **P**resence **P**rotocol

Γεώργιος Α. Κατσικογιάννης

Μηχανικός Ηλεκτρονικών Υπολογιστών και Πληροφορικής

Εκτενής περίληψη

Οι επικοινωνίες μηχανής με μηχανή (M2M) παρέχουν εξειδικευμένες υπηρεσίες για την αντιμετώπιση των αυξανόμενων αναγκών και του μεγάλου όγκου δεδομένων των κατανεμημένων συσκευών. Οι υπηρεσίες αυτές διαδραματίζουν σημαντικό ρόλο στον έλεγχο των ροών και των δεδομένων M2M. Τα διάφορα πρωτόκολλα επικοινωνίας, οι τεχνολογίες δικτύων και οι μηχανισμοί ασφαλείας αναμένεται να εξελιχθούν ταχύτατα τα επόμενα χρόνια προκειμένου να ικανοποιήσουν τις αυξανόμενες ανάγκες. Παράλληλα, αναμένεται να προσφέρουν υψηλότερη αξία στις λύσεις M2M που εξυπηρετούν ολοένα μεγαλύτερη φορτίο στις επικοινωνίες M2M, όπως και πιο σύνθετες εφαρμογές. Η τεχνολογία ποικίλλει ανάλογα με τις συσκευές που χρησιμοποιούνται, τις μεθοδολογίες και τις αρχιτεκτονικές που κυμαίνονται από μονολιθικές εφαρμογές έως αρχιτεκτονικές προσανατολισμού υπηρεσιών (Service Oriented Architecture/SOA) καθώς και αναδυόμενες υπηρεσίες υπολογιστικού νέφους και μικρο-υπηρεσίες (microservices). Τέτοιου είδους αρχιτεκτονικές βασισμένες σε μικρο-υπηρεσίες για την επικοινωνία M2M αποκαλύπτουν τις δυναμικές εξουσιοδοτήσεις και ενισχύουν τον σχεδιασμό των αποκεντρωμένων δομικών στοιχείων, την ευκολότερη ανάπτυξη εφαρμογών, την καλύτερη διαχείριση των ταυτοτήτων και τη βέλτιστη χρήση των πόρων.

Λόγω του μεγάλου αριθμού των συνδεδεμένων συσκευών, της κατανεμημένης και περιορισμένης από τους πόρους φύσης των συσκευών, της κρισιμότητας των εφαρμογών M2M (π.χ. παρακολούθηση ασθενών και λειτουργία κρίσιμων υποδομών), τις μεταβαλλόμενες συνθήκες του περιβάλλοντος και τις αδυναμίες στη σταθερότητα του δικτύου, προκύπτουν διάφορες πτυχές ασφάλειας και ιδιαιτερότητες που θα πρέπει να ληφθούν υπόψη. Συγκεκριμένα, θα πρέπει να εφαρμοστούν οι κατάλληλοι κανόνες ασφάλειας για τον μετριασμό των κινδύνων και των απειλών, όπως και να εξασφαλιστούν τα απαραίτητα επίπεδα ασφάλειας κατά την διάρκεια του χρόνου εκτέλεσης. Στο πλαίσιο αυτό, έχουν προταθεί μέχρι στιγμής πολυάριθμα συστήματα ελέγχου πρόσβασης και δυναμικές εξουσιοδοτήσεις για την αποτελεσματική διαχείριση της ασφάλειας των συσκευών M2M, της επικοινωνίας M2M και των υπηρεσιών στο σύννεφο. Παρόλο που η ενίσχυση της ασφάλειας των δικτύων και των δεδομένων αποτελεί πρωταρχικό μέλημα για τους περισσότερους παρόχους λύσεων, οι προσπάθειες αυτές συνεπάγονται σημαντικές προκλήσεις για τον έλεγχο της πρόσβασης σε ευαίσθητες πληροφορίες και τη διατήρηση ασφαλών κρίσιμων δεδομένων. Επιπλέον, στην εποχή της υπολογιστικής νέφους, υπάρχει ένα αυξανόμενο ενδιαφέρον για προηγμένες και ασφαλείς υπηρεσίες ελέγχου πρόσβασης που θα βελτιώσουν την ασφάλεια, όπως και τις απαιτήσεις για διαλειτουργικότητα. Ως εκ τούτου, θα πρέπει να προστατευθούν και να διασφαλιστούν οι απαραίτητοι πόροι, όπως στο υπολογιστικό σύννεφο, που υποστηρίζουν τις διευρυμένες λειτουργίες και τα κοινόχρηστα δεδομένα.

Η συγκεκριμένη διατριβή παρουσιάζει μια αρχιτεκτονική (εφεξής ASPIDA) προσανατολισμένη στην υπηρεσία για το δίκτυο επικοινωνίας M2M με γνώμονα τις πολιτικές ασφαλείας, ικανή να αντιμετωπίσει τις αναδυόμενες προκλήσεις και λαμβάνοντας υπόψη τις ειδικές σχεδιαστικές ανάγκες. Οι δυνατότητες διαχείρισης με πολιτικές θα δύνανται να βελτιώσουν την ασφάλεια των υπηρεσιών M2M, επιτρέποντας μια προσαρμοστική πολιτική με την επιβολή των κατάλληλων ελέγχων ασφάλειας, ενώ παράλληλα θα προσφέρει

αυξημένη ευελιξία και υψηλότερα επίπεδα υπηρεσιών στον τομέα των επικοινωνιών M2M. Η αρχιτεκτονική επιτρέπει δυναμικά επίπεδα πρόσβασης σε προστατευμένους πόρους με την ενοποίηση των ενοτήτων ελέγχου ταυτότητας και τις εξουσιοδοτήσεις με τις πολιτικές ελέγχου πρόσβασης. Τέλος, αναλύονται τα χαρακτηριστικά της αρχιτεκτονικής ASPIDA και παρουσιάζεται η απόδοση του πρωτοτύπου βασισμένου σε αυτή προκειμένου να γίνουν συγκριτικές μελέτες με άλλες καθιερωμένες λύσεις, προκειμένου να αξιολογηθεί η αποτελεσματικότητα αυτής.

Συγκρίνοντας την αρχιτεκτονική ASPIDA με την υπάρχουσα λειτουργική αρχιτεκτονική επικοινωνίας ETSI M2M (ETSI TS 102 690), προτείνεται και περιλαμβάνεται ένας πρόσθετος τομέας. Η αρχιτεκτονική ASPIDA υποστηρίζει τέσσερις τομείς και είναι προσανατολισμένη στις υπηρεσίες. Οι τρεις πρώτοι τομείς, η συσκευή, το δίκτυο και η εφαρμογή, είναι κοινοί σε άλλα καθιερωμένα και αναγνωρισμένα μοντέλα αναφοράς. Ο νέος τομέας υπηρεσίας (service domain) καταργεί κάποιες από τις διαδικασίες που χειρίζονταν προηγουμένως οι άλλοι τομείς, όπως η καθιέρωση της σύνδεσης μιας συσκευής στο δίκτυο και η καταχώριση υπηρεσιών από τον τομέα δικτύου για ασφαλή πρόσβαση με επίγνωση της πολιτικής ασφαλείας που επιβάλλει το δίκτυο, μειώνει το υπολογιστικό φορτίο των τριών προαναφερθέντων τομέων, και επιβάλλει την ασφάλεια που απαιτείται από ορισμένες εφαρμογές M2M. Ένα πλαίσιο βασισμένο στην πολιτική για το στρώμα υπηρεσιών μπορεί να εποπτεύει όλες τις πτυχές ασφάλειας και να αναφέρει παράτυπες συμπεριφορές (π.χ. μια συσκευή που επικοινωνεί με διαφορετική πύλη). Σε σύγκριση με τις υπάρχουσες αρχιτεκτονικές επικοινωνίας M2M, η συγκεκριμένη επικεντρώνεται στην παροχή πρόσθετης επιβολής ασφαλείας στις υπηρεσίες που διακινούνται από τα σημεία αλληλεπίδρασης στο δίκτυο. Συγκεντρωτικά, τα *κύρια χαρακτηριστικά* της αρχιτεκτονικής είναι τα εξής:

- Υπηρεσιοστρεφής αρχιτεκτονική
- Προσέγγιση βασισμένη σε πολιτικές ασφαλείας
- Εφαρμογή της κατάλληλης πολιτικής ασφάλειας από τον διαχειριστή πολιτικής
- Για την εκτέλεση των απαραίτητων ενεργειών λαμβάνονται υπόψη τα διάφορα γεγονότα (event-management)

Η συγκεκριμένη λύση παρέχει πρόσθετα χαρακτηριστικά στην αρχιτεκτονική ETSI M2M. Για παράδειγμα, όσον αφορά την προσαρμοστική φύση του τομέα του δικτύου, υποστηρίζεται η προσαρμοστική δρομολόγηση με γνώμονα τα χαρακτηριστικά κίνησης και τις συμφωνίες σε επίπεδο υπηρεσιών (Service-Level Agreements/SLA) μέσω ενός συστήματος που χρησιμοποιεί διαχείριση γεγονότων και ενεργειών. Επιπλέον, μία από τις πιο πρόσφατες τάσεις στο λογισμικό σύννεφου είναι η αρχιτεκτονική βασισμένη σε μικρο-υπηρεσίες, η οποία αποτελεί το κλειδί για την ανάπτυξη σύγχρονων αρχιτεκτονικών εφαρμογών και αποδοτικών λύσεων. Οι συγκεκριμένες προσεγγίσεις απαιτούν την καλύτερη ενορχήστρωση των μικρο-υπηρεσιών που θα επιτρέψει την αυτοματοποιημένη ρύθμιση και διαχείριση των πολλαπλών υπηρεσιών ως μία ενιαία υπηρεσία αθροιστικών υπηρεσιών. Επιπλέον συστατικά που εξυπηρετούν και εφαρμόζουν την πολιτική διαχείρισης (π.χ. σημείο επιβολής πολιτικής / Policy-Enforcement Point, σημείο απόφασης πολιτικής / Policy Decision Point) περιλαμβάνονται στην αρχιτεκτονική και ενσωματώνονται σε διάφορα επίπεδα των μικρο-υπηρεσιών. Η πολιτική διαχείρισης αποφάσεων και επιβολής βάσει χαρακτηριστικών επιτρέπει την ταχεία παράδοση αλλαγών και την αυξημένη ευελιξία της τεχνολογίας. Τα σημαντικά πλεονεκτήματα της προσέγγισης διαχείρισης με πολιτικές αυξάνονται, καθώς οι επικοινωνίες M2M εξελίσσονται και οι πόροι γίνονται περισσότερο περίπλοκοι. Οι πόροι αποδεικνύονται περισσότερο διαθέσιμοι και αξιοποιήσιμοι για τα διασυνδεδεμένα στοιχεία

μέσω διαλειτουργικών υπηρεσιών. Τα μοντέλα εφαρμόζονται σε σενάρια που βασίζονται στο σύννεφο και περιπτώσεις που απαιτείται πρόσβαση στους πόρους.

Συγκεντρωτικά, οι *κύριες συμβολές* της παρούσης διατριβής είναι οι ακόλουθες:

- Το μοντέλο M2M διαχειρίζεται αποτελεσματικά τις δυναμικές εξουσιοδοτήσεις και προσαρμόζεται στις μεταβαλλόμενες συνθήκες, εξασφαλίζοντας ταυτόχρονα την ασφαλή επικοινωνία και υποστηρίζοντας ασφαλείς ανταλλαγές δεδομένων μέσω των προγραμματιζόμενων διασυνδέσεων εφαρμογών (Application Programmable Interface / API) που εκθέτουν οι μικρο-υπηρεσίες. Το μοντέλο εφαρμόζεται σε σενάρια που βασίζονται σε σύννεφο και χρησιμοποιούν ανταλλακτικές με βάση το διακριτικό σήμα για να αποκτήσουν πρόσβαση στους πόρους και να επιβάλουν πολιτικές ή επιχειρηματικούς κανόνες.
- Στηριζόμενοι σε μικρο-υπηρεσίες προσφέρεται η ποικιλομορφία και η ανεξαρτησία της υψηλότερης τεχνολογίας μεταξύ της δυναμικής εξουσιοδότησης, της διαχείρισης με βάση την πολιτική και των υπολογιστικών προσανατολισμένων σε υπηρεσίες (Service Oriented Computing / SOC). Ταυτόχρονα, υποστηρίζεται η δημιουργία και η διαχείριση των πολιτικών M2M για τη βελτίωση της ασφάλειας των ροών εργασίας των δεδομένων και των ανταλλαγών μηνυμάτων, καθώς πραγματοποιούνται πολλές αλληλεπιδράσεις επιπέδου εφαρμογής μεταξύ των συσκευών.
- Διεξάγεται ένα πρωτότυπο για την εκτίμηση και αξιολόγηση των χαρακτηριστικών του προτεινόμενου πλαισίου. Αναπτύσσεται μια ανάλυση και σύγκριση της προτεινόμενης λύσης με άλλα πλαίσια και επεξηγούνται οι διαφορετικοί τύποι εγκρίσεων (OAuth grant types) που υποστηρίζονται. Αξίζει να σημειωθεί ότι σε σύγκριση με άλλες τεχνικές επιλογές, η αξιολόγηση καταδεικνύει ότι το πρωτότυπο έχει καλύτερες επιδόσεις όσον αφορά τον αριθμό των ταυτόχρονων συνδέσεων, τη διακίνηση και το ποσοστό μεταφοράς.
- Αναλύονται διάφορα σενάρια για την αποκάλυψη των δυνατοτήτων της προτεινόμενης αρχιτεκτονικής, για την ανάπτυξη των βελτιώσεων, για την αντιμετώπιση των προκλήσεων της υπηρεσίας στις επικοινωνίες M2M και για την αντιμετώπιση των πτυχών ασφάλειας με τους δυναμικούς μηχανισμούς ελέγχου πρόσβασης.
- Τέλος, αναδεικνύονται διάφορες προκλήσεις για την αντιμετώπιση και την υποστήριξη των κατάλληλων επιπέδων ελέγχου πρόσβασης, όπως και θέματα που αφορούν τις δυνατότητες διαχείρισης των πολιτικών.

Προφανώς, οι υπηρεσίες ασφάλειας και ελέγχου πρόσβασης είναι απαραίτητες για την πραγματοποίηση μιας επιτυχημένης αρχιτεκτονικής χρησιμοποιώντας το πρότυπο SOA. Στη βιβλιογραφία αρκετές ερευνητικές και βιομηχανικές δραστηριότητες ασχολήθηκαν με τη διαχείριση ταυτότητας, τον έλεγχο ταυτότητας, τον ρόλο και την εξουσιοδότηση. Στη βιβλιογραφία, σημειώνεται έντονο ερευνητικό ενδιαφέρον για την ανάπτυξη μηχανισμών ελέγχου πρόσβασης και εξουσιοδοτήσεων. Ωστόσο, υπάρχει ανάγκη ενίσχυσης των δυναμικών εξουσιοδοτήσεων με δυναμική ανάλυση εξουσιοδότησης με ανεξάρτητες μικρο-υπηρεσίες. Η αρχιτεκτονική ASPIDA υποστηρίζει την ικανότητα κλιμάκωσης των λειτουργιών και κάλυψης διαφορετικών επιπέδων ζήτησης ή ικανοτήτων μεταξύ ενός μεγάλου πληθυσμού ετερογενών έξυπνων αντικειμένων, εφαρμογών, και υπηρεσιών. Η απόδοση αυξάνεται λόγω της υψηλότερης τεχνολογικής πολυμορφίας και της ανεξαρτησίας μεταξύ της δυναμικής εξουσιοδότησης, της διαχείρισης βάσει πολιτικής και των υπολογιστικών προσανατολισμένων σε υπηρεσίες. Η αρχιτεκτονική δημιουργεί επίσης ένα ολοκληρωμένο

μοντέλο ελέγχου πρόσβασης για να εξασφαλίσει υψηλό επίπεδο βεβαιότητας και να αποφύγει συγκρούσεις, ασυνέπειες ή τυχόν σαφώς καθορισμένες πολιτικές (κανονιστικές ή συστημικές).

Προκειμένου να αξιοποιηθεί καλύτερα και αποτελεσματικότερα η χρήση των πόρων του δικτύου και της απόδοσης, όταν οι πόροι υπερβαίνουν τα επιτρεπτά όρια SLA, τότε οι κατάλληλοι κανόνες πολιτικής πρέπει να ενεργοποιηθούν για την επίλυση προβλημάτων ή βελτίωση της συμπεριφοράς συνδεσιμότητας. Για παράδειγμα, σε περίπτωση συμφόρησης των συνδέσεων, η δρομολόγηση πρέπει να επιλέξει άλλες διαδρομές. Η αρχιτεκτονική διαμορφώνει τις συνθήκες διαχείρισης για την επιβολή των κατάλληλων πολιτικών στο δίκτυο. Ενδεικτικά, με τη χρήση της τεχνολογίας της συμφωνίας παροχής υπηρεσιών του πρωτοκόλλου Internet (Internet protocol service level agreement / IP SLA), οι ειδοποιήσεις SNMP στο σύστημα καταγραφής των προκαθορισμένων ορίων μπορούν να ενεργοποιήσουν τις απαραίτητες αλλαγές διαμόρφωσης. Σύμφωνα με αυτά, είναι δυνατόν να εξαχθούν στατιστικά στοιχεία SLA και να επιβεβαιωθεί αν πληρούνται τα κριτήρια SLA ή υπάρχει παραβίαση και επομένως πρέπει να ληφθούν συγκεκριμένες ενέργειες. Η διαχείριση του ελέγχου πολιτικής και οι συμφωνίες του επίπεδου υπηρεσιών ενσωματώνονται στην αρχιτεκτονική για να παρέχουν λειτουργίες παρακολούθησης και εφαρμογής των κατάλληλων μέτρων. Επιπρόσθετα, η εφαρμογή των λειτουργιών διαχείρισης πολιτικών περιλαμβάνει τη σχετική διαμόρφωση στις οντότητες δικτύου (δηλαδή στους δρομολογητές και τους μεταγωγείς) με το κατάλληλο λογισμικό διαμόρφωσης. Ενδεικτικά, οι αποτυχίες των κόμβων και των συνδέσμων που υπερβαίνουν τους πόρους που έχουν διατεθεί στο εύρος ζώνης μπορούν να ενεργοποιήσουν τις αλλαγές διαμόρφωσης που καθορίζονται από την πολιτική. Τέτοια ζητήματα μπορεί να προσδιοριστούν και να επιλυθούν προληπτικά, θέτοντας τους ανιχνευτές συμβάντων για να παρακολουθήσουν συγκεκριμένους τύπους καταστάσεων, επιτρεπτών ορίων, ή να εκτελούν περιοδικά ένα σύνολο ενεργειών.

Οι μηχανισμοί ποιότητας υπηρεσίας (Quality of Service / QoS) δύναται να χρησιμοποιηθούν ως ρυθμιστές κυκλοφορίας στην αρχιτεκτονική ASPIDA. Οι κλάσεις ποιότητας υπηρεσίας χρησιμοποιούνται από το σύστημα διαχείρισης που επιτρέπει την επιβολή των κατάλληλων τεχνικών που γνωρίζουν την ποιότητα υπηρεσίας. Για να επιτύχει αποτελεσματικά αυτή την επιβολή, διαμορφώνεται στο PBNM μια χαρτογράφηση μεταξύ τεχνικών ποιότητας υπηρεσίας και συμβάντων που οδηγούνται από SLA. Ένα τυπικό παράδειγμα αυτής της αρχιτεκτονικής είναι ένα γεγονός πλημμύρας ICMP που παρακολουθείται με κριτήρια SLA. Εάν τα κριτήρια SLA παραβιαστούν και ένας έλεγχος πολιτικής για αυτό έχει ρυθμιστεί στο σύστημα διαχείρισης, τότε μπορούν να ενεργοποιηθούν προσαρμοστικές αλλαγές διαμόρφωσης και παραμετροποίησης. Αυτές οι αλλαγές στοχεύουν είτε για την αντιμετώπιση της κατάστασης είτε για την ανακατανομή των πόρων του δικτύου. Τελικά, οι αλλαγές μπορούν να επηρεάσουν τον μηχανισμό επιλογής της διαδρομής δρομολόγησης.

Οι λειτουργίες ελέγχου πρόσβασης της αρχιτεκτονικής ASPIDA απεικονίζονται μέσω ενός πρωτότυπου (intelligent Bus of Campus / iBuC) για να προσφέρουν μια σύγχρονη υπηρεσία μεταφοράς που υποστηρίζει πολλαπλές μονάδες αυτόνομων οχημάτων, υποστηριζόμενη από τις υπάρχουσες ασύρματη υποδομή και υπηρεσίες M2M. Στη συνέχεια, η μονάδα ελέγχου υπολογίζει τόσο την εκτιμώμενη ώρα άφιξης του αυτόνομου οχήματος όσο και του επιβάτη στη στάση του λεωφορείου προκειμένου να επιλέγει ένα επιλέξιμο όχημα για τη λειτουργία ενός δρομολογίου για τους επιβάτες. Η μονάδα ελέγχου συλλέγει διάφορα σύνολα δεδομένων σχετικά με την κατάσταση των οχημάτων και των επιβατών

χρησιμοποιώντας ασύρματες συσκευές αισθητήρων, δυνατότητες επικοινωνίας μέσω του συστήματος γεωγραφικής θέσης (Geographical position system / GPS) και ασύρματης επικοινωνίας (Wi-Fi). Οι πτυχές ασφάλειας για τις επικοινωνίες αντιπαραβάλλονται μεταξύ των ακόλουθων δύο σεναρίων, όπου εξετάζονται διεξοδικά οι αλληλεπιδράσεις επικοινωνίας και οι ροές των μηνυμάτων που ανταλλάσσονται μεταξύ των οντοτήτων

(α) εγγραφής / απόσυρσης των πόρων και
(β) αιτήματος υπηρεσίας σύλληψης του αιτήματος

Στην παρούσα διατριβή, η αρχιτεκτονική ASPIDA υποστηρίζει ένα σύστημα πολιτικών διαχείρισης και ελέγχου πρόσβασης. Εξετάζονται οι προκλήσεις αυθεντικοποίησης ταυτότητας υποστηριζόμενες με την ανάκτηση πολιτικών πρόσβασης και επανεξετάζονται οι συνθήκες πολιτικής κατά την εκτέλεση που επιδιώκουν την επιβολή της αντίστοιχης αποτελεσματικής πολιτικής απόφασης. Επιπλέον, εξετάζεται σε βάθος η υιοθέτηση μιας αρχιτεκτονικής βασισμένη σε μικρο-υπηρεσίες, έτσι ώστε κάθε μία από τις υπηρεσίες να μπορεί να αναπτυχθεί ανεξάρτητα και να κατανεμηθεί σε ανεξάρτητες μονάδες, η εξισορρόπηση, η κλιμάκωση και η δημοσίευση / εγγραφή υπηρεσιών. Στην ASPIDA, τα αιτήματα πρόσβασης μπορεί να εξυπηρετηθούν μέσω διακριτών μικρο-υπηρεσιών, οι οποίοι να διευκολύνουν την ανάπτυξη υπηρεσιών API σε σύγκριση με σύνθετα μοντέλα ανάπτυξης (π.χ. Java EE και Spring). Από την πλευρά του υποσυστήματος διαχείρισης των πολιτικών, το σημείο αποφάσεων (PDP) λαμβάνει πληροφορίες από το περιβάλλον και τα χαρακτηριστικά (attributes) που κρατούνται στη βάση που διατηρούνται οι πολιτικές για να ανακτήσει τυχόν ενημερωμένες πολιτικές, προκειμένου να αποφασίσει με βάση τις πληροφορίες που ανακτήθηκαν. Τέλος, το σημείο αποφάσεων αποκρίνεται και εφαρμόζει την σχετική απόφαση μέσω του σημείου εφαρμογής.

Με λίγα λόγια, η αρχιτεκτονική ASPIDA προσφέρει σημαντικά αυξημένες δυνατότητες απόδοσης και εκμάθησης, ενώ το σύστημα διαχείρισης πολιτικών επιτυγχάνει προσαρμοστική δρομολόγηση QoS μέσω αυτοματοποιημένης διαμόρφωσης. Ταυτόχρονα, τα ιδιαίτερα χαρακτηριστικά για την αποφυγή ζητημάτων δρομολόγησης ή συνθηκών χαμηλής απόδοσης των οντοτήτων του δικτύου. Οι πόροι δικτύου μπορεί να αξιοποιηθούν αποτελεσματικότερα μέσω της συγκεκριμένης διαδικασίας προσαρμοστικής διαδρομής δρομολόγησης. Η απόδοση βελτιώνεται σημαντικά, καθώς η κυκλοφορία μπορεί να επαναπροσανατολιστεί σε άλλες διαδρομές σε περίπτωση συμφόρησης, ενώ οι ανιχνευτές συμβάντων μπορούν να ενεργοποιήσουν τις κατάλληλες διορθωτικές ενέργειες (δηλαδή παραβίαση της κυκλοφορίας). Εάν τα επιτρεπτά όρια SLA ξεπεραστούν, τότε το σημείου εφαρμογής (PDP) μπορεί να επιβάλει τους απαραίτητους κανόνες πολιτικής.

Τα *πλεονεκτήματα* της αρχιτεκτονικής ASPIDA είναι η υποστήριξη των δυνατοτήτων βάσει πολιτικών. Προσφέρει αυτοματοποιημένη συντήρηση των συνόλων πολιτικής, βελτιωμένη αποτελεσματικότητα, απλοποιημένη διαχείριση και υποστήριξη διαφόρων τύπων περιβάλλοντος (δηλ. Επιχείρηση, Πάροχος υπηρεσιών). Το σύστημα ASPIDA είναι επεκτάσιμο για να υποστηρίζει περαιτέρω πολύπλοκες επεκτάσεις διαχείρισης πόρων (π.χ. βέλτιστη κατανομή πόρων, χρήση πόρων, απόδοση SOA, δυναμική μετεγκατάσταση φόρτου εργασίας), δυνατότητες εξουσιοδότησης (π.χ. σενάρια εξουσιοδότησης), προσθήκες διαχείρισης πολιτικής και περιβάλλοντος σε περιβάλλον SOA. Στην ASPIDA εξετάζονται διάφορες προκλήσεις για την υποστήριξη των κατάλληλων επιπέδων ελέγχου πρόσβασης με τις δυνατότητες διαχείρισης πολιτικής. Στον πυρήνα της, η αρχιτεκτονική ASPIDA επιτυγχάνει καλύτερη προσαρμοστική διαχείριση ασφάλειας με τη χρήση των PDP, PIP, PAP μαζί με την επιβολή των κατάλληλων ελέγχων ασφαλείας. Οι εξουσιοδοτήσεις που βασίζονται στην

πολιτική που υποστηρίζονται από τους ανιχνευτές συμβάντων μπορούν να μετριάσουν τις παραβιάσεις της ασφάλειας και να εφαρμόσουν την κατάλληλη διαμόρφωση ελέγχου πρόσβασης, εάν η ταυτότητα επαληθευτεί και η εξουσιοδότηση είναι έγκυρη.

Η απόδοση των πολιτικών συμβάντων εκτέλεσης εξαρτάται από τα χαρακτηριστικά του υλικού. Στην περίπτωση των συμβάντων και των μικρο-υπηρεσιών, η απόδοση εξαρτάται από την πλατφόρμα, το επίπεδο φόρτωσης και το συνολικό ποσό της διαθέσιμης μνήμης. Καθώς μπορεί να υπάρξουν περιορισμοί υλικού στον ταυτόχρονο αριθμό σεναρίων πολιτικής, ορισμένες ενέργειες συμβάντων πιθανόν να χαθούν σε περίπτωση υπερκάλυψης του συστήματος. Νέα συμβάντα μπορεί να αποσυρθούν, καθώς οι πολιτικές έχουν ομάδες αποδόσεων καθορισμένου μεγέθους. Επομένως, απαιτείται ένας χρονοπρογραμματιστής γεγονότων για την εκχώρηση και την εκτέλεση αποφάσεων, επιλογής συμβάντων, βελτίωσης της ελαστικότητας, και παρακολούθησης των συμβάντων. Όσον αφορά την επεκτασιμότητα, ο χρονοπρογραμματιστής συμβάντων επιτρέπει την ενεργοποίηση ή την αναστολή των σεναρίων και των αντίστοιχων πολιτικών βάσει των συνθηκών του δικτύου και των δυναμικών κριτηρίων SLA. Καθώς οι εφαρμογές που βασίζονται σε νέφος αναμένεται να αυξηθούν περαιτέρω, ο προγραμματιστής συμβάντων μπορεί να διαχειριστεί πολιτικές πολλών τομέων και να ασχοληθεί με την ικανότητα, την πολυπλοκότητα και τα προβλήματα απόδοσης που μπορεί να προκύψουν.

Οι εφαρμογές εξουσιοδότησης που βασίζονται στην πολιτική απαιτούν πολύπλοκες αλληλεπιδράσεις για να επιτρέπουν την διαχείριση βάσει πολιτικών. Εκτός από την πολυπλοκότητα των πολιτικών εξουσιοδότησης, τα χαρακτηριστικά του χρόνου εκτέλεσης της πολιτικής, οι μηχανισμοί επικύρωσης πολιτικής και η εκτέλεση πολιτικής επηρεάζουν τις υψηλές ετερογενείς υπηρεσίες M2M με πολύπλοκες απαιτήσεις (π.χ. μη επιτηρούμενες συσκευές, δίκτυο εύρους ζώνης, ανάγκες εφαρμογής για μηδενική καθυστέρηση, ελάχιστες απώλειες πακέτων και εφαρμογές χαμηλού ποσοστού σφάλματος μεταφοράς). Όσον αφορά τις μικρο-υπηρεσίες, η υπολογιστική ισχύς, η σύνθεση, οι δυνατότητες διαχωρισμού, η ενσωμάτωση τελικών σημείων, η αποκεντρωμένη διαχείριση δεδομένων και τα στοιχεία της υποδομής επηρεάζουν την απόδοση του συστήματος. Συγκριτικά, το πρωτότυπο βασισμένο στην αρχιτεκτονική ASPIDA υποστηρίζει υψηλότερες διεργασίες συγκριτικά με άλλες λύσεις που αντιπαραβάλλονται, παρόλο που η απόδοση του συστήματος εξαρτάται από διάφορους παράγοντες. Η απόδοση των εγγενών εργαλείων ομαδοποίησης και η λεπτομερής ενορχήστρωση, οι προκλήσεις των κατανεμημένων συστημάτων, η ανοχή για την αποτυχία των υπηρεσιών και τα όρια των υπηρεσιών μπορούν να επηρεάσουν σοβαρά το σχεδιασμό και τις λειτουργίες, όπως η αποτελεσματικότητα της παρακολούθησης σε πραγματικό χρόνο και ανίχνευση ανωμαλιών.

Πέραν κάθε αμφιβολίας, απαιτείται να δοθεί μεγαλύτερη προσοχή στην εξασφάλιση των κατάλληλων ελέγχων και των επιπέδων ασφαλείας σε διάφορους προσωπικούς, διακυβερνητικούς, υπηρεσιακούς, εμπορικούς και βιομηχανικούς τομείς για τις επικοινωνίας M2M. Λόγω των αυξανόμενων ανησυχιών όσον αφορά τις τεχνικές διαχείρισης πρόσβασης και των πολυάριθμων ελέγχων πρόσβασης που έχουν προταθεί τα τελευταία χρόνια, καθώς των αναδυόμενων λύσεων σύννεφου, αυξάνεται το ενδιαφέρον για προηγμένες και ασφαλείς υπηρεσίες ελέγχου πρόσβασης για πρόσβαση σε προστατευόμενους πόρους με βάση πολιτικές αποφάσεις. Εντούτοις, η προσέγγιση διαχείρισης μέσω πολιτικών ασφαλείας ενέχει διάφορες προκλήσεις, όπως θέματα που αφορούν την εξέλιξη της πολιτικής, η ασφαλής διαλειτουργικότητα, η σημασιολογική διαφοροποίηση και οι αποτελεσματικοί μηχανισμοί παρακολούθησης των γεγονότων.

Η διατριβή, που εκπονήθηκε, απαρτίζεται από επτά κεφάλαια που οργανώνονται ως εξής:

- Το *κεφάλαιο 1* περιγράφει τον τρόπο και την συμβολή της διατριβής στο πλαίσιο της διαχείρισης των πολιτικών αλλά και στόχων της συγκεκριμένης έρευνας. Πιο συγκεκριμένα, αυτή η ενότητα εξηγεί τη σημασία των ασφαλών M2M συνδέσεων αλλά και των ιδιαιρετοτήτων που θα πρέπει να ληφθούν υπόψη
- Το *κεφάλαιο 2* παρέχει μια λεπτομερή επισκόπηση των τεχνολογιών, όπως για παράδειγμα τα πρωτόκολλα και τα δίκτυα που χρησιμοποιούνται στις M2M επικοινωνίες, τα μεθοδολογικά πλαίσια διαχείρισης και θέματα που άπτονται της ασφάλειας
- Το *κεφάλαιο 3* αναλύει τις συγκεκριμένες τεχνολογίες που παρουσιάστηκαν στο προηγούμενο κεφάλαιο και αναλύει τις σχετικές προκλήσεις που αναδύονται από τη σχετική εργασία. Επίσης, ερευνά τις μεθόδους που έχουν συμπεριληφθεί στην αρχιτεκτονική ASPIDA και αναλύει γιατί είναι οι καταλληλότερες. Περιγράφει τις ερευνητικές δραστηριότητες, όπως η προσαρμοστική δρομολόγηση διαχείρισης δικτύου, η δρομολόγηση QoS και η παροχή της κυκλοφορίας, η διαχείριση βάσει συμβάντων και ο ενοποιημένος έλεγχος πρόσβασης μαζί με τον μηχανισμό πολιτικής πρόσβασης
- Το *κεφάλαιο 4* περιγράφει την αρχιτεκτονική ASPIDA που εκπορεύεται από τις βασικές οντότητες που παρουσιάζονται στα προηγούμενα κεφάλαια. Αυτή η ενότητα περιγράφει τα χαρακτηριστικά της αρχιτεκτονικής ASPIDA συνδυασμένα με την ανάλυση σεναρίων σχετικά με τις ροές των μηνυμάτων που ανταλλάσσονται μεταξύ των οντοτήτων της αρχιτεκτονικής
- Το *κεφάλαιο 5* παρουσιάζει τις απαιτήσεις διαχείρισης ελέγχου πρόσβασης, τις εξουσιοδοτήσεις M2M και το πρωτότυπο που αναπτύχθηκε για την ανάλυση της απόδοσης της αρχιτεκτονικής ASPIDA. Σε αυτό το κεφάλαιο αναλύονται τα χαρακτηριστικά ασφάλειας και ελέγχου της αρχιτεκτονικής ASPIDA. Πιο συγκεκριμένα, παρουσιάζονται θέματα που αφορούν την ταυτοποίηση, την αυθεντικοποίηση, τα επίπεδα ελέγχου εξουσιοδότησης πρόσβασης και τους μηχανισμούς επικύρωσης των κανόνων, μαζί με την ενσωμάτωση των αντίστοιχων ενοτήτων με την προτεινόμενη αρχιτεκτονική
- Το *κεφάλαιο 6* παρουσιάζει το πλαίσιο ελέγχου της αξιολόγησης των μετρήσεων με τα σχετικά αποτελέσματα. Οι μέθοδοι συλλογής δεδομένων εξηγούνται ακολουθούμενες από δεδομένα, ενώ οι μετρήσεις απόδοσης συγκρίνονται με άλλα παρόμοια μεθοδολογικά πλαίσια. Πιο αναλυτικά, το συγκεκριμένο κεφάλαιο εξετάζει τα εξής:
  - Το πλαίσιο αξιολόγησης και τις μεθόδους για την προσέγγιση διαχείρισης βάσει πολιτικής σύμφωνα με το οποίο ελέγχεται η συμπεριφορά του δικτύου και του τομέα υπηρεσιών στο πλαίσιο πολιτικών, SLA, QoS, ασφάλειας, επαλήθευσης ταυτότητας και εξουσιοδότησης
  - Περιγραφή του σχεδιασμού των πειραμάτων για τη διεξαγωγή των απαραίτητων δοκιμών
  - Περιγραφή της κατανομής, ανάλυσης και αιτιολόγησης των παρατηρήσεων. Επιπλέον, περιλαμβάνεται η περιγραφή της ανάλυσης των δεδομένων και των δοκιμών

- Το *κεφάλαιο 7* εξετάζει τα συμπεράσματα, τις διαπιστώσεις και τις μελλοντικές βελτιώσεις. Η διδακτορική διατριβή ολοκληρώνεται με το συγκεκριμένο κεφάλαιο στο οποίο αναφέρονται τα γενικά συμπεράσματα της διατριβής, οι στόχοι που καλύφθηκαν από την συνολική έρευνα, οι επιπτώσεις της έρευνας καθώς και οι κρίσιμοι παράγοντες που επηρεάζουν την απόδοση της αρχιτεκτονικής και του συστήματος. Παρουσιάζεται η ex post facto ανάλυση της ASPIDA σε σύγκριση με άλλες ερευνητικές δραστηριότητες, περιγράφοντας τις περιγραφικές, διερευνητικές και επιβεβαιωτικές αναλύσεις που έχουν διεξαχθεί στα προηγούμενα κεφάλαια και η περίληψη των αποτελεσμάτων για την απεικόνιση των αριθμητικών πληροφοριών. Τέλος, παρατίθενται προτάσεις για μελλοντικές προεκτάσεις της παρούσας έρευνας.

Με κριτήριο όλα τα παραπάνω, η συγκεκριμένη διατριβή αντιμετωπίζει τα παρακάτω:

- Πληροί την προσαρμοστική φύση του τομέα του δικτύου για την υποστήριξη της επιλεκτικής δρομολόγησης με γνώμονα τις απαιτήσεις και συμφωνίες σε επίπεδο υπηρεσιών (Service Level Agreements / SLAs) με την ανάπτυξη ενός συστήματος διαχείρισης με χαρακτηριστικά ποιότητας εξυπηρέτησης (Quality of Services / QoS) με βάση την πολιτική για προσαρμοσμένη δρομολόγηση
- Καλύπτει τις ανάγκες για προσαρμοστικές πολιτικές της αρχιτεκτονικής που να διευκολύνουν τις δυνατότητες εξυπηρέτησης με τα διάφορα μέσα ελέγχου πρόσβασης. Ιδιαιτέρως, περιλαμβάνεται η ενσωμάτωση των ενοτήτων ελέγχου ταυτότητας και εξουσιοδότησης με τις πολιτικές ελέγχου πρόσβασης που χρησιμοποιούνται από την πλατφόρμα παράδοσης εφαρμογών και διαχείρισης των υπηρεσιών
- Επιτυγχάνει την εκτέλεση των σχετικών εντολών βασισμένων σε γεγονότα (event based management) τα οποία ενεργοποιούνται υπό συγκεκριμένες συνθήκες (για παράδειγμα υπέρβαση αποδεκτών ορίων SLAs), επιτρέποντας τον κατάλληλο χειρισμό των συμβάντων και την βελτίωση της δυναμικότητας και προσαρμοστικότητας της αρχιτεκτονικής ASPIDA. Για παράδειγμα, ενδεχόμενες βλάβες/αστοχίες των κόμβων και των συνδέσμων ή υπερβάσεις των πόρων που έχουν διατεθεί στο εύρος ζώνης (όπως περιπτώσεις συμφόρησης), μπορούν να ενεργοποιήσουν τις απαραίτητες αλλαγές διαμόρφωσης που καθορίζονται από την πολιτική. Τέτοια ζητήματα μπορούν να αναγνωριστούν και να επιλυθούν προληπτικά, ρυθμίζοντας τους ανιχνευτές συμβάντων να παρακολουθούν συγκεκριμένους τύπους καταστάσεων, επιτρεπτών ορίων ή να εκτελούν περιοδικά ένα σύνολο ενεργειών
- Πραγματοποιεί ένα σύστημα με υψηλότερες αποδόσεις σε σύγκριση με άλλες λύσεις και μεθοδολογικά πλαίσια με σύγχρονες μεθόδους υλοποίησης, όπως μικρο-υπηρεσιών, που προσφέρουν αυξημένα πλεονεκτήματα στην ανάπτυξη, επεκτασιμότητα και βιωσιμότητα της μεθοδολογίας
- Πετυχαίνει την ενίσχυση των μεθόδων SOA για τις υπηρεσίες ελέγχου πρόσβασης, ταυτοποίησης και αυθεντικοποίησης για την προστασία και βελτίωση της αξιοποίησης των δεδομένων περιορισμένων πόρων του περιβάλλοντος (M2M resource constrained environment)

To Evi, Anastasios and Marios

Chapter 1

Introduction

1.1 Outline

This chapter describes how this dissertation makes an original contribution to the body of knowledge in the discipline of policy-based management and addresses the significance of the study. More specifically, this section explains the significance of secure M2M connections, which are being used in a number of emerging consumers and industry segments, and across a broad spectrum of data sources. The problem statement provides the context of this dissertation to enforce a security policy-based management approach.

1.2 M2M computing

In the recent years, Machine-to-Machine (M2M) communications and applications have increased dramatically, comprising several devices and nodes with distinctive features and diverse technical capabilities. The selection of the appropriate components (e.g. water-resistance/anti-vibration/dust-proof/GPS-enabled nodes) plays a significant role in the operations and performance of the final M2M solution in a resource-constrained environment and needs to address the inclusion of fundamental factors like low-power consumption, device battery life, available connection speed and localized management. In terms of the M2M communication protocols, the M2M environments need to support cross-domain information exchanges among several smart interconnected nodes. These exchanges complicate the operations of the routing, the management protocols, the M2M communication services and the device reachability, which all of them result in various security and performance issues [1].

Several M2M applications such as office security and automation, personal area and home networking, automotive, transportation, human/inventory tracking, water/energy distribution, quality monitoring, habitat monitoring, data centre monitoring, disaster avoidance and recovery, military surveillance, industry operations, medical/healthcare monitoring, process monitoring and more smart spaces rely on M2M communication to improve business. Most of these applications collect large volumes of heterogeneous data in real-time in physical conditions that change over time. The M2M communications are becoming more prevalent and require the ensuring of data transmission and accuracy in a secure, scalable and reliable manner.

M2M communications have attracted considerable research and industry attention and a noteworthy development is the availability and accessibility of operational information and continuously streamed real-time M2M data. Data delivered via Internet of Things (IoT) technologies can be pervasive and enable new business models and opportunities. According to IDC[1], worldwide spending on the IoT is forecast to reach $772.5B in 2018. That represents an increase of 15% over the $674B that was spent on IoT in 2017. The IoT market is expected to reach $267B by 2020, according to Boston Consulting Group. By the end of 2020, close to 50% of new IoT applications built by enterprises will leverage an IoT platform that offers

[1] IDC FutureScape: Worldwide IoT 2018 Predictions

outcome-focused functionality based on comprehensive analytics capabilities. The global IoT market will grow from $170B in 2017 to $561B by 2022, attaining a Compound Annual Growth Rate (CAGR) of 26.9%. Gartner Inc.[2] forecasts that 11.2 billion connected things will be in use worldwide in 2018 and will reach 20.4 billion by 2020. Therefore, it is reasonable to address the issue of managing efficiently the connections and the secure data distribution. Cisco [2] projects that M2M connections will grow from 1.5 billion in 2018 to 2.6 billion by 2020 with the evolution of mobile technologies and IoT capabilities. Moreover, software and services are expected to be a $600B market by 2019, attaining a 44% CAGR (Compounded Average Growth Rate) from 2015 to 2019. A large number of the M2M devices utilise these rapidly evolving technologies to connect the physical entities with the applications often hosting the information in the cloud. Hence, in order to design scalable models, expandable structures and efficient methodologies, there is a need to expand the M2M solutions in order to accommodate the expected exponential increase of billions of connected M2M devices along with the respective business services.

To shape the architecture, it is vital to provide an exhaustive analysis of design decisions and explore various technology combinations and security options for relevant use cases and research activities. In this context, one should carefully review the device capabilities, the functions, the processes and the information flows along with the communication design, all of which affect the diverse edge architectures. The Gartner IoT Reference Model (Figure 1) illustrates the three basic components for the creation of an IoT architecture and details the layers, the tiers coupled with the interfaces.

The design decisions for each one of the five layers affect the edge performance, storage, power and communication design. More specifically, the edge is where the physical devices, processes and events exist, the platform and enterprise tiers define where components, functions and processes operate in the architecture. In the device layer, there is a need to manage the processing and storage requirements. While the communication layer outlines the communication services, the information layer describes the data models, the data flows and how data is transformed and stored. The function layer includes the analytics combined with the processing and learning capabilities, and the process layer describes the device management and how to integrate the activities with the governance and management processes.

Various architecture and design challenges arise when it is necessary to improve the operational efficiency and cope with the technical diversity and the integration complexity. For instance, diverse edge architectures can evolve based on the enterprise challenges and the application like the fleet management and health care demands. The devices (sensors, readers, etc.) need to capture the event (temperature, pressure, change in the status, etc.) and interoperate in the technology domain through a network (wireless, wired or hybrid) and gateways. Moreover, the design decisions are often affected by the physical environment constraints and need to ensure adherence to IT standards and comply with safety, regulatory and legal requirements. Finally, the enterprise's vision, its strategic plans and its performance objectives can all influence the development and implementation of the solution and the target architecture for realising the business requirements.

[2] https://www.gartner.com/newsroom/id/3598917

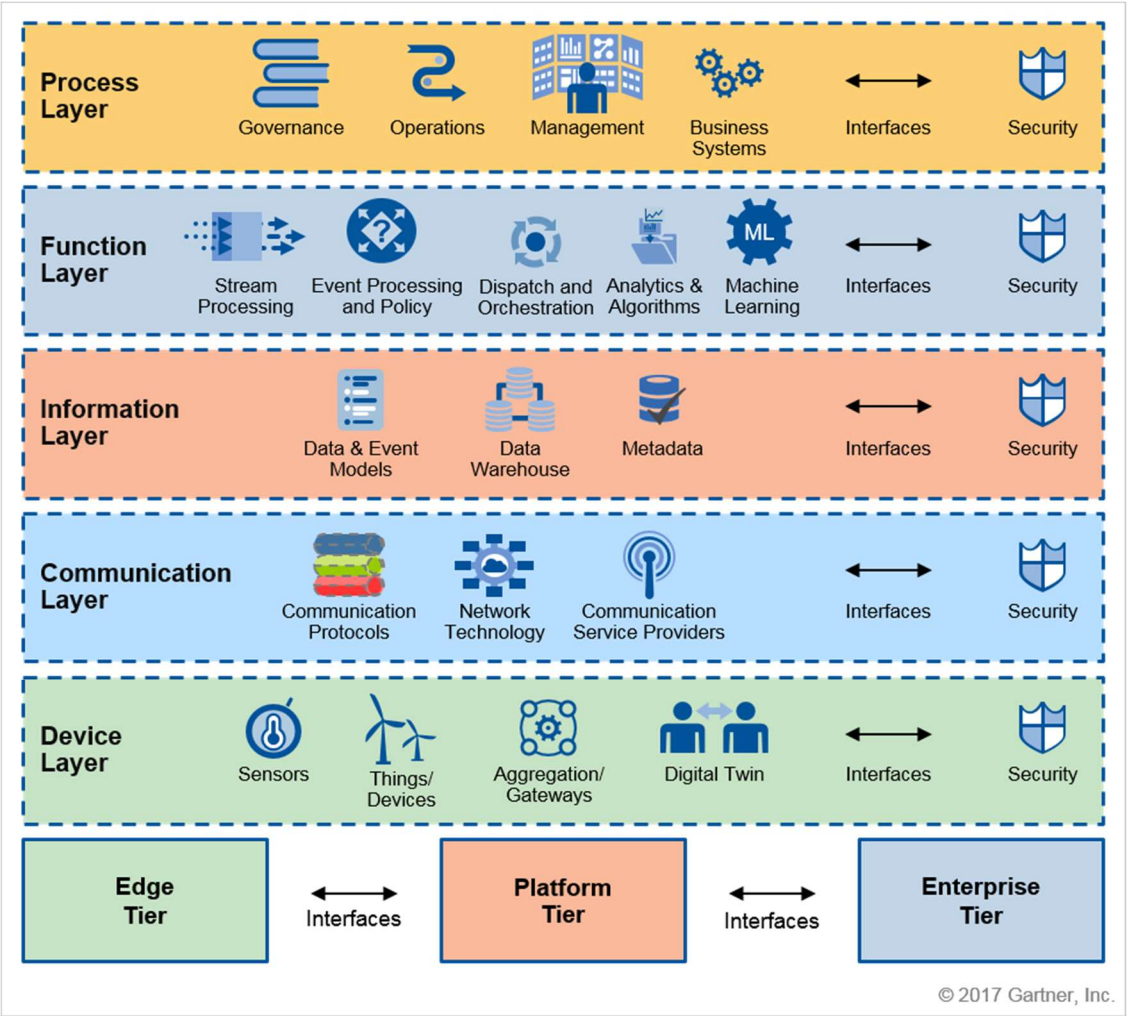


Figure 1. The Gartner IoT Reference Model at a glance (March 2017)

Figure 2 illustrates a 5-five step approach on how to develop and execute an IoT technical strategy.

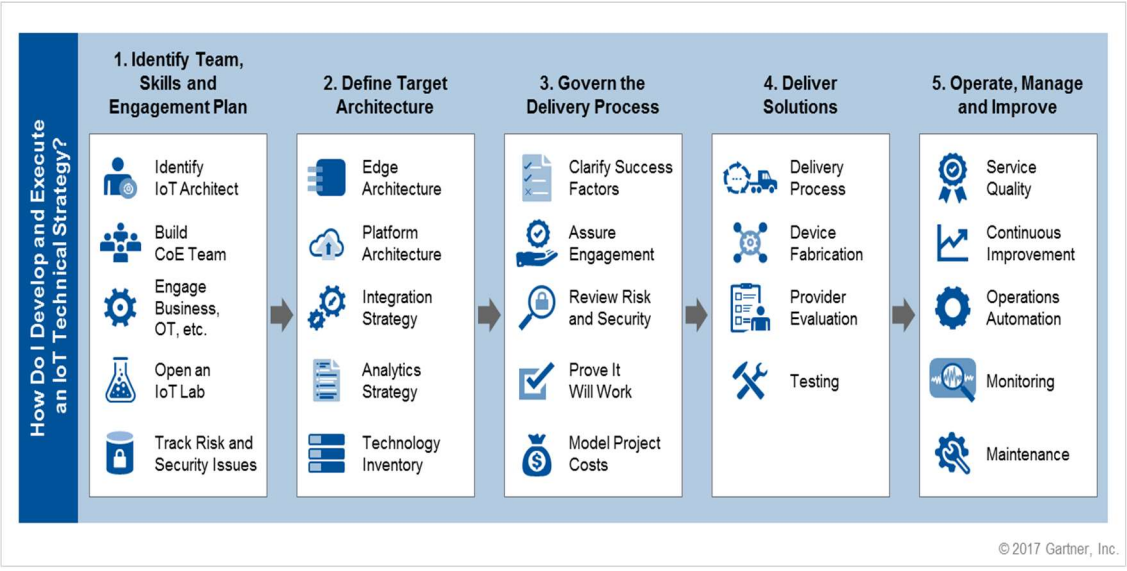


Figure 2. Gartner Five Steps to an IoT Technical Strategy (December 2017)

There is an increasing need to improve the performance, the availability and the expandability of the relevant applications and the endpoints by considering:

- The evolution of the emerging related technologies
- The design decisions in order to fulfil the requirements
- The management framework with the aim of ensuring vitality over time
- The factors affecting the target architecture

Hence, the technology standardisation is a key-factor for faster development, adoption of high-quality M2M modules and value-added services. The service disruptions and the overwhelming system complexity can be minimized through proper planning and design of the M2M service operations like monitoring, security network management, data processing and operational performance. The system complexity should be decomposed into layers and event-driven architectures to address the messaging requirements and deliver efficient and effective complex solutions on a broad range of used technologies.

1.3 Problem statement

Although the existing Policy-based Management (PBM) models provide viable and feasible solutions, there is a growing need to deliver an adaptive behaviour to the offered services. PBM allows the creation of certain expressions, which enable policy enforcement on the interconnected components and the target resources in order to apply the appropriate resource access and usage. Nevertheless, the current solutions interpret the policy rules and update the configuration of the network elements without considering the dynamic nature of the environment and the changing circumstances, such as the network conditions, the SLA metrics and the conformance needs. Moreover, the static models based on discrete management deployments raise scalability issues. Common approaches translate the policies and interpret the authorisation decisions in a static manner. Most of the proposed solutions consider the enforcement of static policy rules. Although augmented Temporal Role-Based Access Control (TRBAC), Generalized TRBAC (GTRBAC) and XML-RBAC (X-RBAC) models have been proposed in the literature [3] in order to incorporate temporal capabilities for pervasive computing applications, there are still various challenges due to the multi-tenancy models of cloud computing, resource-sharing, virtualization and microservices (μServices). Due to the static and context insensitive nature for the majority of the traditional access control approaches, Atlam et al. [4] conduct a risk analysis to estimate the security risk associated with each access request with an Adaptive Risk-Based Access Control (AdRBAC) model aiming to provide adaptive and real-time features. The authors use a qualitative approach and validate the model through an expert review. Still, there is a need to understand the interactions of real-time user interactions based on the context, understand how to automatically enforce access control decisions and enable them by using the proper triggers. Hence, adaptive controls on the access decisions need to adapt to the varied and changing circumstances at runtime to grant permissions that can be exercised. The hierarchical access control models [5] also require an efficient key management scheme to provide access in hierarchical sensor networks and privileged hierarchy. However, different issues with existing key management techniques can arise in conjunction with the availability of the limited resources of M2M networks.

Aiming to acquire an adaptive policy-based architecture, the authorisations should be managed dynamically based on policies and context information such as the user, the user's attributes and the conditions. Vincent et al. [6] provide techniques for creating a policy-

compliant service composition through a graph and She et. al [7] present the evaluation of policies during service compositions. The existence of a minimum policy model to integrate various access control constraints with policy-based compliant services needs to be depicted. To be able to present changes in the status of the entities and the conditions, the system operations can be improved with event-driven models [8] in order to capture the data of the M2M connected devices. Sinjilawi et al. [9] demonstrate multi-level security classes of information flows for cloud-based solutions where there is a need to identify and present the access control limitations, and then propose a unified access control model with integrated components. Moreover, the integrated access control model needs to ensure a prominent level of assurance and avoid conflicts, inconsistencies or any ambiguously specified policies (i.e. regulatory or systemic) [10].

At the same time, in order to support further complex resource management extensions (i.e. optimal predictive resource allocation, resource usage, Service-Oriented Architecture (SOA) performance, dynamic relocation of workloads), it is mandatory to define improved authorisation capabilities (i.e. permission classes, task flows, Single-Sign-On (SSO) functionality, Security Assertion Markup Language (SAML) and eXtensible Access Control Markup Language (XACML) [11][12] uses in complex authorisation scenarios), as well as policy and context management additions in SOA environment. Some of the objective grading criteria to improve the quality of the offered services are:

- Better resource utilisation
- Secure data distribution
- Efficient monitoring of the services
- Metrics for performance and availability

In terms of path selection optimization and intelligent path control, the SLA adaptive routing decisions need to cope with the fast-growing needs and the required loads as well as with the application performance and event-based management options for adaptive routing. For instance, in the case of a failure or a network topology change, it should be feasible for the policy engine to trigger an event-based policy action in order to reconfigure the network immediately and automatically [13].

In more detail, there is a need to integrate QoS management even on a well-behaved PBM system, so that to be able to provide better performance and service-levels. Chen et al. [14] analyse the security attacks and performance issues of the Choudhury et al. [15] framework, which is based on a two-step verification for user authentication in cloud computing using either passwords or smart-cards. However, the traditional identity and access management controls are no longer sufficient for the current cloud computing and distributed systems [16]. Khan [17] categorizes the security threats and performs a comparative analysis of security issues for secure mobile cloud computing infrastructures to identify potential problems such as web services and protocol based attacks. The technologies and the operational models in cloud computing create additional security risks because of the outsourcing to third parties and the diverging legal and compliance frameworks [18]. These security risks include data security, privacy, service availability, confidentiality, data integrity, data segregation, privileged user access and accountability.

Due to the high number of interconnected heterogeneous smart resources, the use of various networking technologies for M2M communications and the distributed nature of the

smart applications, several security and service quality issues may arise. Concerning the interconnected objects, most of the times the M2M nodes are resource-constrained with respect to [19]:

- Power-consumption
- Computational capabilities
- Bandwidth
- Storage capacity

Therefore, the traditional security mechanisms are not always applicable for M2M implementations, making the M2M security management extremely challenging. For instance, the M2M wireless communications inherit additional security threats and operational constraints compared to their wired counterparts, as well as additional and exceptional characteristics. Because of the distributed and dynamic nature of the environment, the M2M nodes are free to move, thus triggering a high rate of physical topology changes and they often operate in an unattended fashion with very limited maintenance support. Several access control and dynamic authorisations schemes have been proposed to this date to enforce an effective security management of the M2M devices. Ngo et al. [20] extend the access control model for Intercloud scenarios [21] by exchanging tokens and transforming complex logical expressions in policies to compact decision diagrams aiming to simplify the attribute-based policy evaluation. Acquiring a security policy-based management approach introduces several new challenges such as the need for policy evolution, secure interoperability, semantic diversification and efficient event monitoring mechanisms. Numerous other challenges should also be addressed, such as the need for better policy-based management including advanced decision-making criteria, complex conditions, cross-domain policies and dependencies.

1.4 Proposed solution

To address the aforementioned issues, an **A**daptive **S**ecure **P**ol**I**cy-base**D** **A**rchitecture (ASPIDA) is introduced that incorporates novel and state-of-the-art technologies in order to facilitate the access control needs and the M2M service capabilities. As depicted in Figure 3, ASPIDA is a service-oriented M2M architecture that establishes four domains, as an expansion of the ETSI M2M functional architecture [22], which includes the first three domains, device, network and application domain.

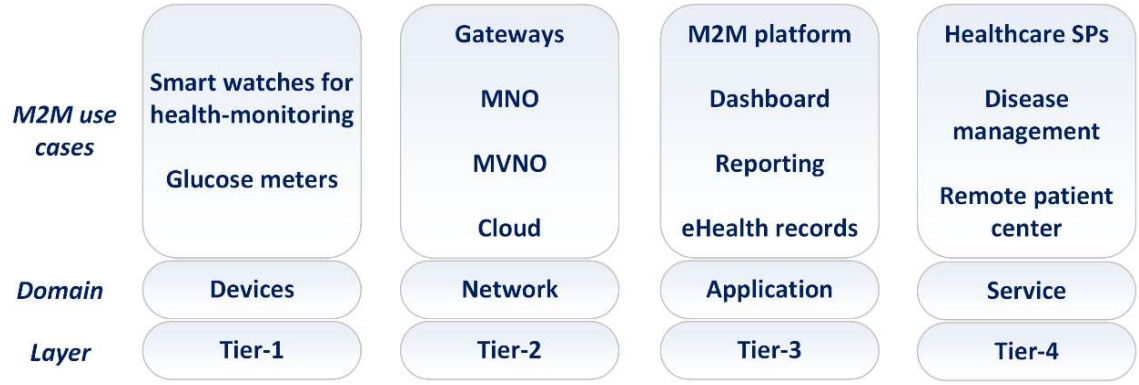


Figure 3. Domains for M2M computing

This dissertation also approaches a series of issues related to access control models and SOA, which necessitates common terminology and semantic, syntactic and technical

interoperability. The orchestration of the access control components for the deployment of SOA is depicted in Figure 4, where the Identity and Access Management (IAM) is illustrated with the Service Delivery solution and the administration modules.

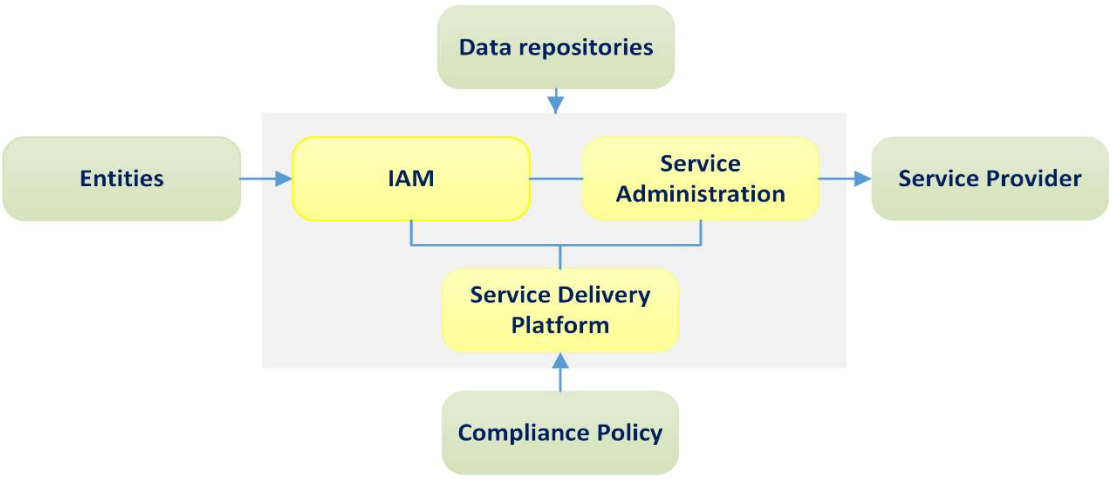


Figure 4. Access control model for SOA

Furthermore, the architecture supports policy-aware M2M and dynamic authorisations for µServices by utilising capability-based exchanges to gain access to the resources. By managing the access policies and enforcing the business rules accurately and efficiently, the effectiveness of the policy-decision service solutions for M2M entities is improved. For instance, the elderly may consent to allow access to the data retrieved for remote patient monitoring according to certain conditions and policies, such as their physical aging, the type of disease, special treatments, or any other interventions. Localization and temporal access can be added to support different cases, such as “break-the-glass” cases, granting access to health-care records that was not originally allowed in case of an emergency (i.e. accident). Eventually, higher technology independence for the dynamic authorisations is accomplished by utilising µServices. Finally, the components of the conceptual model are detailed by means of Unified Modelling Language (UML) diagrams that include the entities, the relationships and the flows. The detailed reference architecture aims to effectively manage the security aspects and adapt to the changing conditions whilst ensuring secure communications and supporting secure data exchanges through the appropriate APIs. The architecture description is divided into domains, layers and views, so that each one of them contains components that execute processes and offer services to the next one like the dynamic authorisation process and the policy-based management. In order to validate the proposed architecture and its supporting protocol, a corresponding prototype is developed and implemented. This prototype based on Java and µServices technologies is submitted to simulations to verify the functional characteristics and collect performance measurements. The security analysis and performance benchmarking in comparison with other similar approaches are also discussed. The approach is sketched in Figure 5.

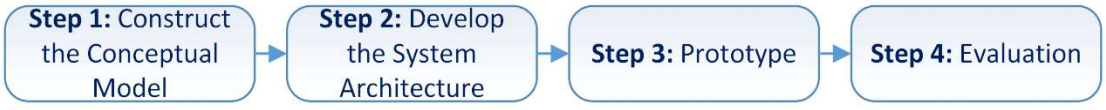


Figure 5. On the modelling and analysis of the architecture

1.5 Contributions and outline

M2M communications are based on specialized services in order to cope with the growing need for a larger number of connected objects, the use of resource-constrained devices and high distributed data volumes. The focus is put on metering efficiently the technology capabilities and performance of containerized M2M applications with respect to better security and quality levels. The µServices based architecture for M2M communication uncovers the dynamic authorisations and fine-grained µServices for designing decentralized service building blocks, making application development easier, achieving greater flexibility in consuming the authentication tokens and utilising the resources optimally.

One of the more contemporary trends in cloud software is the microservice architecture, which is a key to develop creative application architectures and delivery solutions. Thus, the microservice architectures, as a more granular way of implementing SOA, are included and analysed in this work in order to enable a series of better service orchestration for M2M computing. The existence of a service orchestration enables the automated arrangement and management of the multiple services exposed as a single aggregate service. Additional policy-based components (e.g. PEP, PDP) are included in the architecture to decide how to incorporate these with other endpoints in the µServices. The proposed attribute-based policy enforcement enables the rapid delivery of changes and increased technology flexibility. The considerable benefits of the PBM approach grow as the M2M communications evolve and the resources become more complex. The resources turn out to be available to the interconnected components and can be accessed using interoperable services. Finally, the proposed architecture is also applicable for cloud-based scenarios and for capability-based exchanges to gain access to the resources.

Beyond any doubt, greater attention should be paid to ensure the appropriate security controls and the security service levels across various personal, cross-governmental/inter-agency, commercial and industrial sectors. Due to the growing concerns in access management techniques and the numerous access controls that have been proposed in the recent years, as well as because of the emerging cloud computing era, there is an increasing interest in advanced and secure access control services in order to access protected resources based on policy-driven decisions. Acquiring a security policy-based management approach poses several challenges such as policy evolution, secure interoperability, semantic diversification and efficient event-monitoring mechanisms.

The proposed architecture supports the ability to scale operations and meet different demand or capacity levels among a large population of heterogeneous smart objects, applications and services. The performance is increased with higher technology diversity and independence between the dynamic authorisation, the policy-based management and the Service-Oriented Computing (SOC) services. In this work, various scenarios are demonstrated to uncover the capabilities of the proposed architecture, unfold the improvements, develop and evaluate a policy-based QoS management model for adaptive routing, reinforce the SOA policy-based methods for access control services, address the service challenges in M2M communications and tackle the security aspects with dynamic access control mechanisms. Finally, the architecture also establishes an integrated access control model for adaptive security management.

Aiming to increase the utilisation of the network resources and the performance, when SLA thresholds are exceeded, the proper policy rules need to be triggered to resolve

connectivity problems. Furthermore, in the case of congested links, the traffic should be rerouted to other paths. The architecture shapes the most influential ideas of policy-based management at the network to enforce the proper policies.

The main innovative contribution concerns the implementation of a cross-domain and lightweight system architecture conceived for M2M applications. The following chapters are organised as follows:

- Chapter 2 provides a detailed overview of the state of the art on M2M communication and the technologies in use to derive the architecture.
- Chapter 3 analyses in detail the methods, the challenges and the solutions addressed by the related work and included in the proposed architecture
- Chapter 4 describes the policy-based access control architecture for adaptive security management
- Chapter 5 portrays the access control management requirements, the M2M authorisations and various other security aspects related to the proposed architecture. The prototype is also explained
- Chapter 6 demonstrates the validation and evaluation framework, the design and the outcome
- Chapter 7 discusses the conclusions, the findings and some future directions. The implications of the findings, such as other critical success factors that affect the system architecture, and the summary of the results to illustrate the numeric information are also presented in this chapter

Chapter 2

State of the Art

2.1 Outline

This chapter presents the most current research in the given area, describes the differences between various concepts (e.g. fog, cloud, pervasive computing) and summarises the emerging trends. Apart from the literature review, which analyses previous research, the modules and the components of the proposed architecture like the policy-based QoS management model for adaptive network management decisions, are described in detail.

2.2 X-computing

We witness a fast development of the cloud-computing environment together with the continuous evolution of resource-hungry applications and the explosion of M2M as well of as multi-sensory applications in the last few years. These factors result in a growing demand for improving the connected devices and the network resource usage. By incorporating the recent technology advancements like cloud computing services and application virtualization, selecting the proper type of network and communication protocols becomes a stimulating task. This can develop more innovative services, build versatile applications and enforce enterprise application delivery strategies. Moreover, this can result in imposing concise, coherent and consistent security and network policies.

The continued interest in *X-computing* (i.e. cloud, pervasive, wearable, Peer-to-Peer, ubiquitous, mobile-edge, mobile-cloud, fog computing) has resulted in several variations of the computing services for the connected devices. First, the peers communicating with the nearby participants exchange data with the applications by partitioning tasks or workloads; this decentralized model is known as *Peer-to-Peer (P2P) computing.* In general, the peers need to enable resource sharing (e.g. processing power, disk storage, network bandwidth) directly with other network participants without the need for central coordination by servers or by stable hosts. Recent advancements [23] allow resource sharing with diverse peers in order to manage resource allocation mechanisms more efficiently and enable capabilities to achieve better results with file sharing, content distribution and collaborative applications.

In the context of software engineering and computer science, the *ubiquitous computing* can occur using any device, in any location and in any format. More specifically, a user may interact with the computer with everyday objects such as with wearable devices on humans, vehicles and buildings. Ubiquitous computing can be supported by various devices (i.e. smart objects, sensors, microprocessors, tags, pads, boards), networks (i.e. Internet, mobile networks, network robot systems, networked vehicles, service location and positioning protocols) and applications (i.e. mobile code, advanced middleware). When the execution of the services occurs at the edge of the network (i.e. 5G mobile environment), then it is possible to optimize the service and cloud-computing capabilities. By utilising *Mobile Edge Computing (MEC)*, the platform offers cloud-computing capabilities and an IT service environment at the very edge of the mobile network, which results in higher bandwidth, low latency and real-time access to the radio network due to the resulting proximity. Alternatively, in the case of resource-constrained devices and limited capabilities of the edge devices, the *Mobile Cloud*

Computing (MCC) can offload storage and processing from the devices to the cloud. The devices can be augmented with cloud-based resources, such as offloading computation. The MCC allows the execution of computationally intensive tasks to external services and cloud-based nearby computing systems. Although this provides several advantages (i.e. the storage of bulk data is feasible), there are certain challenges that need to be taken into consideration on data security, identity privacy, location privacy, secure routing, reliability, QoS service degradation and potential delays in the execution of the application need to be addressed properly. In the scientific literature, one can find various proposals that deal with remote mobile agents to process the information and enable the offloading of the mobile applications by providing the necessary additional resources and execution environment. In this vein, Magurawalage et al. [24] propose the *Aqua computing* architecture for mobile edge networks where computing and wireless networking resources are allocated jointly or cooperatively by a Mobile Cloud Controller. By mixing the notion of mobile agents and clones, the clones placed at the edge of the network can serve various scenarios, as well as act as a buffer to migrate tasks (including memory image, CPU states) and provide computing services to the connected devices.

In recent years, other novel paradigms have emerged such as *fog computing* [25]. Contrariwise, fog computing aims to bring the storage needs and computational capabilities at the edge of the network. It can use collaborative end-user clients or near-user edge devices to carry out the storage, communication, control, configuration, measurement and management. This reduces the centralized communication overheads, but it requires a stronger coordination between the near-user edge devices and the intermediary network entities with a layered approach. Fog computing was initially defined as a platform with the purpose of allowing the creation of new applications and services in the context of IoT. However, this paradigm has been recently extended to implement other types of services such as low-latency augmented interfaces for constrained devices, Cyber-Physical Systems (CPS) [26] and various Vehicle-to-Vehicle (V2V) along with Vehicle-to-Infrastructure (V2I) services. For instance, in the case of Mobile Ad-hoc Networks (MANET), where each node acts as a router, the communication can take place in smaller local groups. This enables scheduling the tasks with data locality, preserving data privacy and improving the performance with faster information exchanges plus reconfigurations. In these cases, the formation of densely populated networks does not require fixed and costly infrastructures to be available beforehand, which also has a positive impact on privacy due to the data locality.

The Service metrics have a profound impact in the realisation of the benefits, the improvement of the service quality and the better monitoring of the full-service lifecycle. Among other metrics, the following play a significant role:

- Reachability
- Availability of connections
- Response time
- Routing changes
- First-time repair rate
- Other repair operations.

To ensure a clear separation between the specific infrastructure solutions and the necessary monitoring capabilities, a general-purpose monitoring architecture is required. De Chaves et al. [27] provide the architectural views of a private cloud monitoring system and

implement an open source configurator tool. As X-computing services are usually exposed and vulnerable to potential security threats, various security metrics need to be established and the necessary effective controls need to be identified. The formulation of the essential security metrics is a key factor in the road of X-computing maturity. Roman et al. [28] analyse the security threats in all the edge paradigms, mobile edge computing, mobile cloud computing and fog computing in heterogeneous environments where high-speed links and access technologies coexist.

2.3 Connecting everything

This section presents the preliminaries of M2M, IoT, Wireless Sensor Networks (WSN), Wireless Body Area Networks (WBAN) and performs the comparison between M2M and CPS. Even though M2M, WSN, WBAN and CPS are quite similar in many networking aspects, there are major differences from an architecture and design perspective.

2.3.1 Devices

IoT is defined as a collection of physical devices, vehicles, home appliances or smart objects that connect with others within the Internet infrastructure. The *IoT devices* establish the communication among various smart devices and exchange information to achieve automation, monitoring, surveillance, tracking and location identification objectives. Employing the IoT devices in the environment creates new needs for the business processes and necessitates more secure operations in order to support the production automation and enhanced resilience. Sensors, actuators, embedded processors, computers, smart meter readers, assets, devices, RFID tags, mobile terminal devices, surveillance cameras, laser printers, GPS devices and NFC devices can be utilised to transfer data in several business processes.

Under the architecture of IoT, the *M2M devices* are usually embedded in real-life objects (i.e. monitoring the health conditions of patients, controlling the home/office automation like lighting/HVAC/security surveillance systems) and concentrate on the direct communication between machines without or with limited human intervention. The industrial M2M market undergoes a fast development of new business models and develops M2M devices with enhanced functionalities (i.e. learning capabilities and contextual information intelligence). Notably, the characteristics and the network capabilities of the M2M devices play a significant role in the interactions and the performance of the solution such as high-speed data transmission, highly interconnected machine networks and feature extraction methods to convert the raw data into useful information. The data is transmitted over distinct types of networks, diverse transmission medium and security characteristics, all of which may introduce several new threats and challenges. A SOA based security governance [29] is vital to accommodate the security requirements especially in dynamic environments with changing conditions. Moreover, maintaining data privacy to the highest degree possible, gaining control of the data resources and forming the appropriate security controls for secure and structured access are crucial factors in the architectural design choices.

CPS consists of computation, communication and control components tightly coupled with the processes as well as the physical components aiming to enable more intelligent and interactive operations. Often, CPS is considered to be related to the critical infrastructures, the physical and the engineered systems. In comparison with IoT, CPS requires a deeper knowledge of the environment and the physical objects. It also entails higher risks and critical infrastructures, whereas IoT can rely on broader control systems and wider embedded

systems. The operations of the connected objects in the cyber space are closely-monitored, controlled and managed by a computing core.

The wireless technology innovations, the recent developments in embedded computing along with the X-computing evolution enable the realisation of higher level systems for IoT (e.g. CPS). Nevertheless, CPS emphasizes on distributed real-time control and cross-domain optimization. Such optimization is achieved by involving multiple sensor networks, various control functions and intelligence across multiple domains. The widespread adoption of CPS applications requires breakthroughs in the research of theoretical and technical issues. Chen et al. [30] present advanced network techniques for the emerging CPS in M2M research. CPS requires stronger cyber capabilities in any physical or resource-constrained connected device, complex multiple temporal and spatial scales, dynamic reconfiguration capabilities, higher degrees of automation and more intelligent plus interactive control of the operations [31]. Moreover, the CPS implementations need to be closely integrated at multiple and extreme scales. The large-scale CPS installations supported by M2M communications include critical and life-threatening applications (i.e. wearable devices on humans, vehicles, buildings) equipped with a higher volume of sensor inputs and a richer network connectivity. Due to the increased needs for more intelligent interactions, cross-layer and cross-domain optimizations, and distributed real-time controls, the CPS methodologies need to be designed to meet the higher requirements in terms of reliability, security, privacy and real-time performance. Wan et al. [32] describe some of these aspects and challenges that need to be solved, such as energy management, network security, data transmission along with management, model-based design, control technique, system resource allocation, services and applications. Table 1 summarises the differences between the main characteristics of M2M and CPS.

Table 1. Comparison between M2M and CPS

Characteristics	M2M	CPS
Autonomous communications of all intelligent nodes	√	√
Connectivity without or with limited human intervention	√	
Cross-domain intelligence		√
Optimization from multiple WSN		√
Closed-loop/real-time controls		√
High degrees of automation (e.g. unmanned vehicles with WSN navigation)		√
Criticality in the research of theoretical and technical support		√

2.3.2 Networks

M2M communications utilise both wireless and wired systems to communicate end-to-end with a large number of nodes supporting seamless domain interoperability, autonomous operations and self-organisation capabilities to provision intelligent applications. M2M communications need to support mobility, collision detection, intermittent connectivity resolution, topology control, QoS levels and data aggregation with secure services. It is crucial to design properly the network operations to ensure resilient communications, high availability with optimized performance among heterogeneous nodes and different ad-hoc clusters in order to accommodate the traffic load fluctuations and the frequent physical topology changes. There is an increasing number of connections that are controlled by present-day technological developments (i.e. smart phones, workstations over the internet), which reveal behavioural patterns of the monitored objects. Moreover, due to the large diversities of the M2M devices and the associated business cases (e.g. wireless sensor networks, vehicle ad hoc networks, smart grids, etc.), the devices can traverse distinct types of networks (i.e. MANETs, WSN, WMNs, VANETs) with diverse distribution and capabilities of

the nodes. For instance, the nodes in ad-hoc and sensor networks (ASN) have no central communication point allowing the traffic flows between one or more sink nodes, which are often fixed and mesh. An ASN cluster may communicate with other clusters either via each cluster head or via a node acting as the gateway. Focusing on the potential of these technologies and the business cases, the advancements of ASN have unfolded a variety of trends and initiatives that are widely used in energy distribution systems, industrial productions systems, transportation systems and several other application domains. Figure 6 summarises a classification of the several types of networks and the respective subcategories of the Wireless ad-hoc networks (WANET). The classification of the network categories is based on the type of the participating devices, the physical medium, the physical topology, the collection of the adopted technologies and the architectural model. For instance, on the grounds of shaping a dialectic basis towards a comparison between the ad hoc and the sensor nodes, Figure 7 summarises their characteristics. In general, the sensor networks manage more network nodes in comparison to the ad-hoc networks, which creates an increased administrative overhead and network congestion because of the high volume of sensors. Moreover, the ad-hoc nodes are relatively dispersed, self-configuring and include self-restoring capabilities. Often, the sensor nodes once installed are not movable any longer. Several sensor nodes operate in an unattended mode without any constant supervision by a trusted sink and dispersed in large sensor networks that may consist of a few to thousands of nodes. For instance, in harsh and unknown environments such as in the ones found in forest fire detection systems, unattended sensor nodes can be deployed for continuous monitoring. However, multiple sensors are needed for an event-detection, as there is a need for several sensor data from different sensor sources to define and recognise the patterns corresponding to specific events recognition and result in better accuracy. Similarly, in industrial cases there is a need to analyse large volumes of multi-sensor data from vehicles and engines [33].

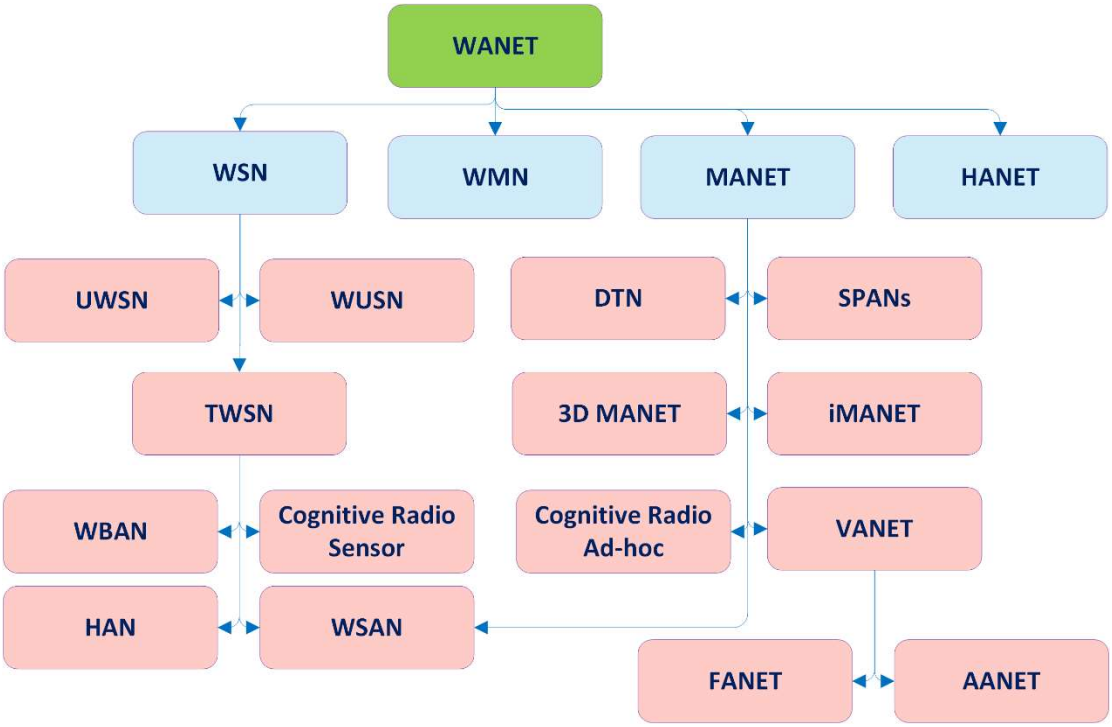


Figure 6. WANET classification

The sensor nodes are often small, homogeneous, inexpensive, with small batteries but longer lifetime and in some cases, they are also embedded with power generators like small

solar plates. They are more hardware adaptive, less hardware complicated, tiny disposable and low-power devices. Contrary to the sensor nodes, the ad-hoc nodes have larger and replaceable batteries with AC-DC charger/adapter with higher levels of power-consumption and they are equipped with wireless receivers along with transceivers using several types of antennas (i.e. omnidirectional, directional).

Figure 7. Characteristics of ad-hoc and sensor nodes

Table 2 illustrates the variety of applications in the scope of WSN and MANETs.

Table 2. WSN, MANETs and their application types

	TWSN [34]	WUSN [35]	UWSN [36]	WSAN [37]	HAN [38]	CRN [39]	WBAN [40]	DTNs [41]	SPANs [42]	3D [43]	iMANET [44]	VANET [45]	AANET [46]	FANET [47]
Disaster relief operation (e.g. drop sensors over a wildfire, observe wildlife)	✓	✓	✓	✓	✓			✓					✓	✓
Military applications (e.g. intrusion detection, battlefield surveillance)	✓	✓	✓	✓		✓	✓	✓		✓		✓	✓	✓
Environmental applications (e.g. sea erosion, air-pollution, water-quality)	✓	✓	✓	✓				✓		✓			✓	✓
Medical & healthcare applications (e.g. long-term surveillance of chronically ill patients)	✓			✓	✓		✓							
Smart-homes & buildings (e.g. HVAC)	✓				✓				✓		✓			
Industrial control, supply chain, machine surveillance, preventive maintenance	✓	✓	✓				✓					✓	✓	✓
Public tracking (e.g. transportation, precision agriculture)	✓			✓		✓		✓	✓	✓		✓		✓
Human tracking (e.g. human presence with optical sensors)	✓	✓	✓	✓	✓	✓			✓	✓		✓		✓

The software used in ad-hoc nodes has more requirements than the ones used in sensor nodes. Although security is a concern for the ASN nodes because of the wireless medium, the sensor nodes have very limited security capabilities and are exposed to numerous security threats. The nodes move independently and autonomously in any direction resulting in frequent changes to the links with other devices and the topology. Apart from the dynamic nature of the topology and the frequent changes, the transmission is performed with multi-hop routing in less redundant networks and there is low node-density.

2.3.3 Communication models

The Internet Architecture Board (IAB) [1] describes the four communication patterns that can be used in the smart object environment. The IoT implementations use different technical communications models, each with its own characteristics. The four common communications models described by the Internet Architecture Board include: *Device-to-Device*, *Device-to-Cloud*, *Device-to-Gateway* and *Back-End Data-Sharing*.

- *Device-to-Device* communication characterizes two or more devices that communicate with each other. In this case, short-range radio or RFID technologies can be used to support the communication with smart-wearable objects, RFID tags, readers, etc. This type of communication is often used to transfer a low volume of data packets between the devices at low data rates.
- *Device-to-Cloud* communication refers to one or more devices that connect to a cloud-based service for data exchange. This type of connection includes the integration of the connected devices and the cloud-based services, so both interoperability and security (e.g. data privacy, strong authentication and authorisation methods) are important aspects in such communication. For instance, a forest fire modelling and early detection systems based on WSN can utilise a cloud service provider to support the data exchange and control centres for detecting fires.
- *Device-to-Gateway* communication implies an intermediary to bridge the devices and the application service provider. The gateway enables the communication, the protocol translation, the security features and the data transcoding for the connected devices. The application solution can be thus simplified, as the complexity as well as the re-usability of the repeated functions are provided by the gateway.
- *Back-end Data Sharing* outspreads the single-device to cloud-solution provider model to benefit from the integration and data sharing with various authorised third-parties. For example, the Uber API [48] allows customers to access real-time customized content by combining data from various sources. The sensor data can be sent to authorised application service providers for predictive data visualization and analytics, aggregation, or statistical analysis. The RESTful APIs enable the secure and user-friendly communication between the clients and the server that employs representational state transfer (REST) constraints.

2.3.4 Protocols

Communication is central to the M2M devices and needs to occur securely and reliably. A variety of standardised protocols specify the formats and the methods in regard to the communication, the messaging, the device management and the information exchange. As there are numerous legacy and emerging communication protocols and management frameworks [49], the entities and the devices can be interconnected for data exchange and information sharing in several ways. Nevertheless, despite the substantial number of M2M protocols currently available, the major hindrances are the limited data rates, the low-power

consumption requirements, the interoperability needs among several heterogeneous devices and infrastructure, and the security sophistication. The technical report for cellular and non-cellular communication technologies in the M2M domain [50] summarises and categorizes them depending upon the coverage distance, as shown in Figure 8.

The selection of the optimal and more appropriate protocol depends on the:

- Use cases
- Network coverage and range capabilities
- Required data bit rate and latency
- Network topology
- Scalability
- Payload length
- Quality-of service
- Deployment implications
- Environmental conditions
- Costs

For instance, the *identification* used (i.e. URIs, IPv6, Electronic Product Code/EPC, uCode), the *discovery* used (i.e physical Web, mDNS, DNS-SD), the *messaging protocols* (i.e. MQTT, CoAP, STOMP, AMQP, Websocket), the type of the *transport layer network* (i.e. TCP, UDP, DTLS, TLS, DPWS), the *network* (i.e. IPv4/IPv6, Routing Protocol for Low power & Lossy Networks, 6LowWPAN), the *infrastructure* that is employed (i.e. 3GPP, IEEE 802.15.4, RFID, NFC, Bluetooth 4.0 Low Energy) and the *type of sensors* (i.e. Infra-red, Electrochemical, Analog to Digital Converter, Thermal Conductivity) affect the setup and the performance of the communication. There is a plethora of short-range/local area wireless technologies available such as Bluetooth (including Bluetooth Low Energy), NFC, RFID, Wi-Fi, Zigbee, Z-Wave and Wireless M-Bus. For long range/wide-area links there are mobile networks using among others GSM, GPRS, 3G, and LTE connections. Most of all these technologies are summarised in Figure 9.

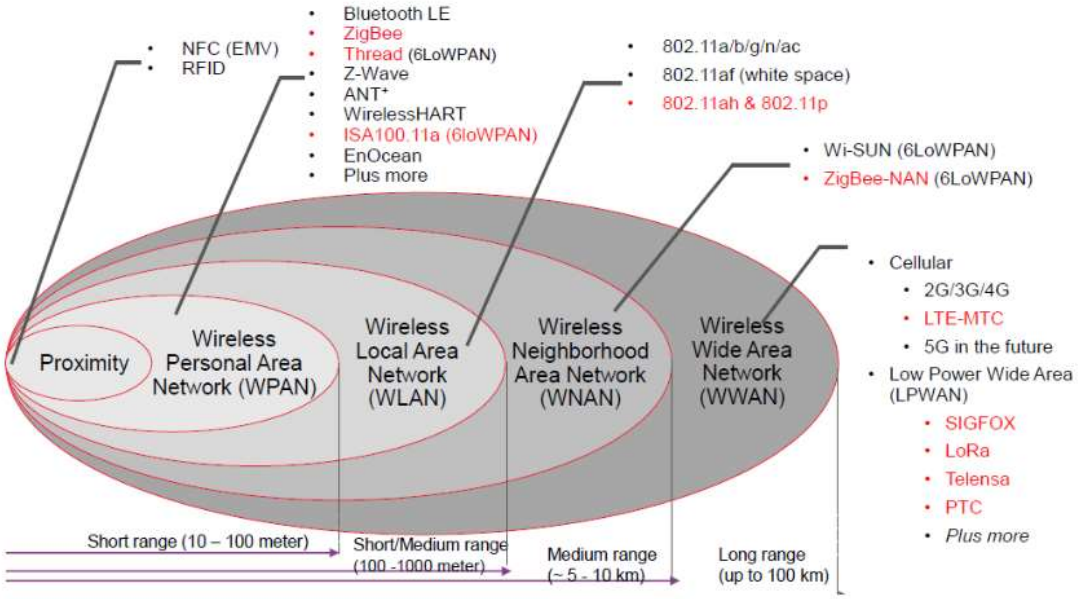


Figure 8. Key-enabling cellular & non-cellular technologies for M2M

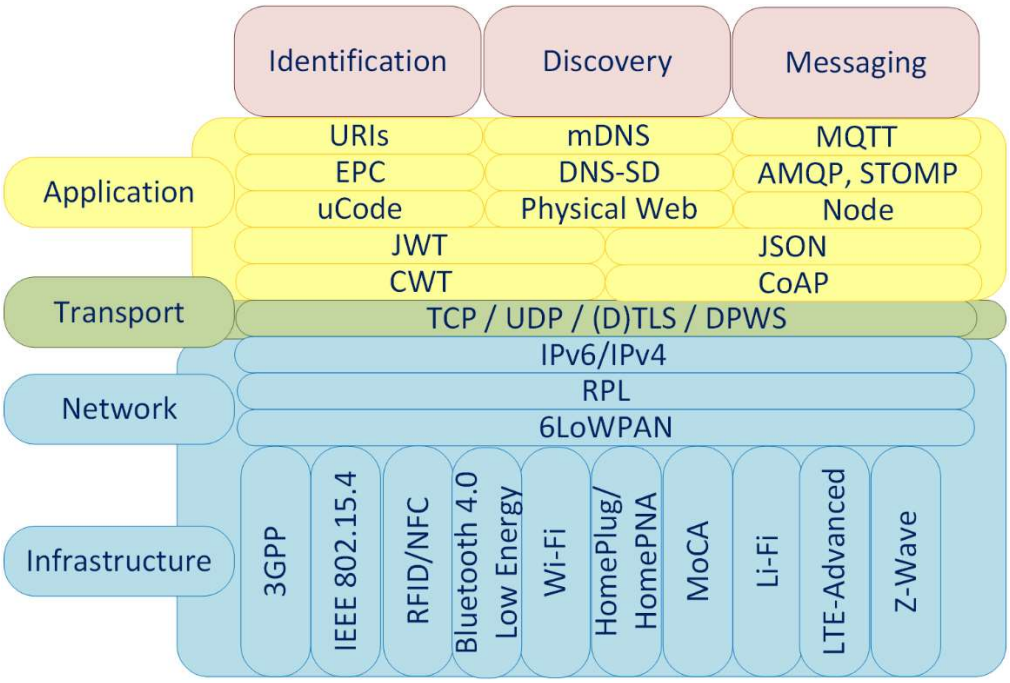


Figure 9. IoT protocols

In parallel, various M2M device management protocols have been proposed and developed by the industry and the academia for device-to-device (D2D) application communication, server-to-device (S2D) and device-to-server (D2S) management. Table 3 summarises most of the M2M management protocols according to their usage, their security and policy-based characteristics.

Table 3. M2M Device Management protocols

	M2M Device Management Protocols	Usage	Features	Security				Policy Management
				Device Discovery	Device Registration	Authentication	Authorisation policies	
D2D	CWMP (TR-069) [51]	CPE WAN management protocol	▪ Vendor Agnostic ▪ Event Management System ▪ QoS at the subscriber endpoint	O		O	O	O
S2D	OMA DM [52]	Designed for management of mobile devices	▪ Request response protocol ▪ Server and client are stateful ▪ Authentication challenge sent by the server or the client	O	O	O		
S2D	OMA LWM2M [53]	Manage remotely constrained devices, service enablement	▪ CRUD on the resources ▪ Resource management ▪ Diagnostics and Monitoring ▪ Multiple LWM2M Server support	O	O	O		
D2S	SNMPv3	Manage network devices and configure networks	▪ User-based Security Model ▪ Traps and Notifications ▪ Authentication and privacy			O		
Dev ID	IEEE 802.1AR [54]	Locally significant identities with initial manufacturer-DevID	▪ Authentication of the device's identity ▪ Establishment of the trustworthiness of devices			O		

Table 4 illustrates a taxonomy of the communication protocols utilised in ASN and addresses their strengths and shortcomings [55].

Table 4. Taxonomy of communication protocols for ASN

Technology	Description	Spectrum	Range	(+) Strengths / (-) Shortcomings
ↄ 6LoWPan [55]	IPv6 over Low Power Wireless Personal Area Networks, which defines encapsulation and header compression mechanisms	Unlicensed	Short	(+) Designed to send IPv6 packets over IEEE 802.15.4-based networks, robust, scalable, self-healing (-) Lack of application that utilises 6LoWPAN, requires extensive knowledge of stack and the workability of IPv6
ↄ Bluetooth & Bluetooth Low-Energy (BLE)/Smart (Wibree) [57]	Exchanges data over a small distance using UHF radio waves in the ISM band from 2.4 to 2.485 GHz	Unlicensed	Short	(+) Simple to use, free, low-power consumption, no line-of-sight (-) Range is power-class-dependent, a quasi-optical wireless path must be viable
ↄ Cellular such as GSM/GPRS/EDGE (2G), UMTS/HSPA+ (3G), HSDPA (High-Speed Downlink Packet Access), LTE (4G)/ LTE Advanced [58]	Wireless mobile telecommunications and each generation is characterized by new frequency bands, higher data rates and non–backward-compatible transmission technology	Licensed	Long	(+) Utilise advanced wireless technologies, multiple uses, users access the radio spectrum on a similar frequency band (-) Expenses, high power-consumption, authentication centres can be breached
ↄ 5G [59]	A consolidation of 2G, 3G, 4G, Wi-fi and other innovations	Licensed	Long	(+) Far greater coverage, always-on reliability, high download speeds, low Latency rate (-) Unresolved security issues
ↄ THREAD, based on IEEE 802.15.4 and 6LowPAN[3]	IPv6 networking protocol aimed at the home automation environment	Unlicensed	Short	(+) No new hardware is needed, handles several 250 nodes with prominent levels of authentication and encryption, designed as a complement to Wi-Fi (-) Based on "hub-and-spoke" hierarchical communications which means that if the hub device fails the whole network fails

[3] https://www.threadgroup.org/

☉ WCDMA/3G[4]	Used for third generation, or 3G, mobile communication networks, spread spectrum modulation technique supporting high-speed multimedia services	Unlicensed	Short	(+) Supports high-speed data services, better suited for deployment in densely populated areas, reduced multi path effects (-) Complexity, versatility
☉ Wi-Fi (802.11x)	Used vastly with radio signal communications	Unlicensed	Short	(+) Good coverage, wide existing infrastructure, offer fast data transfer, handle high quantities of data (-) High power consumption, costly
☉ Zigbee RF4CE/PRO/Smart Energy/Light Link/Home	The IEEE 802.15.4 is an industry-standard specification that requires infrequent data exchanges at low data-rates over a restricted area and within a 100m range	Unlicensed	Short	(+) Simpler, less expensive, very low-power consumption, AES-128 encryption (-) Every network needs at least one coordinator device
☉ Z-Wave Alliance ZAD12837 / ITU-T G.9959 [60]	Used primarily for home automation and provides reliable, low-latency transmission of small data packets	Unlicensed	Short	(+) Low-latency transmission, power-save mode (-) Proprietary design, short-range about 30 meters
☉ NFC (ISO/IEC 18000-3)	Extends the capability of contactless card technology and enables devices to share information	Unlicensed	Short	(+) Simple and safe two-way interactions between electronic devices, smartphone usage (-) Distance range 10cm, security
☉ Ultra-Wideband	Uses short-duration pulses and low energy level, high-bandwidth communications over a large portion of the radio spectrum	Unlicensed	Short	(+) Large channel capacity, ability to share the frequency spectrum, ability to work with low signal-to-noise ratios, low probability of intercept and detection, resistance to jamming, high performance in multipath channels (-) Low performance using classical matched filter receivers, high-frequency synchronization and very fast

[4] https://en.wikipedia.org/wiki/UMTS#W-CDMA_(UTRA-FDD)

				analogue-to-digital converters are needed, multiple-access interference, difficulty predicting the template signals
ↄ WiMAX	Provides portable mobile broadband connectivity across cities and countries	Licensed	Long	(+) Single station can serve hundreds of users, faster deployment, speed of 10 Mbps at 10 kilometres with Line-of-Site (LoS) (-) LoS is needed for longer connections affected by weather conditions, multiplied frequencies, very power intensive technology that requires strong electrical support, installation and operational costs
ↄ Sigfox [61]	Uses Ultra Narrow Band (UNB) technology designed to handle low data-transfer speeds across areas of several square kilometres	Licensed	Long	(+) ultra-narrow-band, industry-leading wireless performance, extended range, ultra-low power consumption, low cost (-) no collision-avoidance techniques, mandates the use of a very precise crystal, interference with any other wideband system
ↄ Neul [61]	Weightless communication, uses the white space radio to access the high-quality UHF spectrum	Licensed	Long	(+) high scalability, high coverage, low power and low-cost wireless networks (-) temporal variation as an available spectrum might be occupied later by a primary user
ↄ LoRa by Semtech [61]	Provide low-power WANs with features specifically needed to support low-cost mobile secure bi-directional communication. Suited for short and periodical communications	Unlicensed	Long	(+) supports large networks, optimized for low-power consumption (-) low bandwidth does not allow continuous sending

As regards the messaging protocols, there is a plethora and a continuous evolution for this type of protocols. They allow devices to communicate from a single constrained device to complex cross-domain and cloud-based solutions to ease the service delivery. Fysarakis et al. [62] analyse the characteristics and evaluate the capabilities of three dominant IoT messaging application protocols concluding that the Constrained Application Protocol (CoAP) [63] is more suitable for the lightweight M2M interactions, the Message Queuing Telemetry Transport (MQTT) [64] protocol facilitates the cross-domain communications utilising an extremely lightweight publish/subscribe messaging transport model and the Devices Profile for Web Services (DPWS) [65] protocol is more appropriate for SOA-based Machine-to-Human (M2H) exchanges. For example, various protocols are oriented for use in constrained (i.e. low-power, limited storage) devices, while there is also a shift from standardised transport protocols (i.e. HTTP) to lightweight protocols. For instance, the CoAP can be used in order to support the high volume of devices and the enormous number of sessions, as CoAP is explicitly designed and more suitable for constrained environments. Alternative realisations can employ the lightweight protocol MQTT with the purpose of allowing the connection of a single constrained device with the systems and facilitating near real-time communications. In terms of security, integrity, confidentiality and protection of the communication, either asymmetric or symmetric encryption can be used depending on the constraints and the specific use cases. Hernandez-Ramos et al. [66] demonstrate a set of Elliptic Curve Cryptography (ECC) optimizations in the design and implementation of a capability-based access control mechanism (DCapBAC) on smart objects. The authors propose a mechanism to compute a session key that establishes a secure channel, so that the capability token is used to get access to the resource(s). However, the distributed approach does not address the impact in the authentication and authorisation modules, as this requires a complex management of authorisation policies across several segments. With these aspects in mind, an integrated access control model is proposed in [67]. Those results are presented in the following sections for motivating the options and choices in Chapter 5 and 6. Table 5 summarises the usage and the features of the most common M2M messaging web-transfer protocols.

By comparing the two lightweight communication protocols CoAP and MQTT, which have gained considerable attention and are the most commonly used, several similarities and differences can be easily identified. Both protocols address the communication needs through the use of small message sizes, message management and lightweight message overhead. However, the MQTT protocol is client-server oriented where every sensor is a client that connects over TCP to a MQTT server named broker. The TCP connection is encrypted with MQTT SSL/TLS to ensure privacy. On the contrary, the CoAP protocol is based on the REST architecture, utilises UDP for group communication and message exchanges, relies on the request/response using methods and responses codes to facilitate the communication needs, and provides a one-to-one “request/report” interaction model with accommodations for multicast to achieve group communication. The CoAP specification suggests either the use of Datagram Transport Layer Security (DTLS) as the recommended security mechanism, or Internet Protocol Security (IPsec) to achieve the security assurance for CoAP messages. In the case of resource-constrained devices, because the HTTP protocol provides low data rates, high computation complexity and high energy consumption, the lightweight CoAP can be used as a replacement of HTTP to overcome the limitations offering low overhead, simplicity and IP multicast support.

Table 5. M2M Message-Oriented protocols

Request / Response architecture	***CoAP*** (Constrained Application Protocol) [63]	Application layer protocol for use in resource-constrained internet devices	▪ RESTful protocol ▪ URI and content-type support ▪ Asynchronous message exchanges ▪ Low header overhead ▪ Parsing complexity ▪ Security bindings with DTLS
	HTTP	Application layer protocol for distributed information systems	▪ Broker-less ▪ TCP bindings ▪ Security bindings with SSL/TLS
Publish / Subscribe architecture	***DDS*** (Data-Distribution Service for Real-Time Systems)[5]	Target devices that directly use device data	▪ Request response protocol ▪ TCP and UDP bindings ▪ Detailed QoS control ▪ Multicast ▪ Configurable reliability ▪ Pervasive redundancy
	DPWS (Devices Profile for Web Services) [65]	Enable secure Web Service messaging on resource-constrained devices	▪ Discovery services ▪ Metadata exchange service ▪ Eventing
	MQTT (Message Queuing Telemetry Transport) [64]	Telemetry and remote monitoring for large networks of small devices	▪ Bandwidth efficiency ▪ Data agnostic nature ▪ Continuous session awareness ▪ Quality of Service (QoS)
	UPnP (Universal Plug and Play) [49]	Discover network devices' presence for data sharing, & communication	▪ Zero configuration networking ▪ D2D networking ▪ Control messages are expressed in SOAP ▪ Event notification
	XMPP (Extensible Messaging and Presence Protocol) [49]	Provide text communication between points	▪ XML text format as its native type ▪ Request response protocol ▪ Lightweight middleware ▪ Content syndication ▪ Generalized routing of XML data
	ZeroMQ [49]	High-performance asynchronous messaging library	▪ Request–reply ▪ Publish–subscribe ▪ Extensible security mechanisms ▪ Command and message framing, ▪ Connection metadata
	AMQP (Advanced Message Queueing Protocol) [49]	An open standard application layer protocol for message-oriented middleware	▪ P2P mode ▪ TCP bindings ▪ P2P, C2B and B2B communication ▪ Message orientation ▪ Queuing & routing
	LLAP (Lightweight Local Automation Protocol) [49]	Short message that is sent between intelligent objects using normal text	▪ Over any communication medium
	STOMP (Simple Text Oriented Messaging Protocol) [49]	STOMP clients talk with any message broker supporting the protocol	▪ Language-agnostic ▪ Interoperable wire format
	Web-socket [68]	Web browsers and web servers communicate continuously	▪ Push/pull communications to a web server ▪ Web Socket JavaScript interface ▪ Full-duplex single socket connections ▪ Bi-directional channel ▪ Message-based ▪ Server-side library support

2.4 Reference architectures

Aiming to improve the interoperability, there is a need to establish reference architectures with standardised methods and approaches to simplify development and ease implementation. The IoT security is highly fragmented using proprietary implementations,

[5] Data Distribution Service (DDS) https://www.omg.org/omg-dds-portal/

differences in methods and properties, and varying sets of supported events. This results in increased complexity and sustainability difficulties aiming at long-term solutions and stability. The reference architecture needs to provide rigorous designs and long-term implementation to overcome several limitations of the computing-X services, the business processes and the endpoints. With these in mind, a reference architecture standardises concrete architectures of systems and enables the systematic reuse of common functionalities and configuration implying increased efficiency and reduced costs. Additionally, a reference architecture reduces the risks through the use of proven elements, enables better interoperability of different systems and achieves increased quality levels with the use of the proper quality attributes.

The standardisation activities for M2M communication are proceeding in different organisations such as in ETSI, OMA, 3GPP, IEEE and TIA [69]. Over the last few years, these organisations have defined various network architectures and functions to support the unique features of M2M communications in their standardisation bodies. Nonetheless, several impeding factors may have a significant impact on the operations and performance of the connected devices such as the competing standardisation activities, the proliferating increase of industry devices lacking the proper quality and safety controls, the rising complexity in the communications and performance issues. In the literature, several architecture models akin to the OSI model have been proposed for the communication among interconnected nodes like:

- The ETSI M2M [ETSI TR 102 690] communication functional architecture [22]
- The Internet of Things—Architecture (IoT-A) [70]
- The CISCO IoT Reference Model [71]
- The Industrial Internet Reference Architecture (IIRA) [72]
- The Reference Architecture Model Industrie (RAMI) 4.0 [73]
- The IBM Internet of Things Architecture [74]
- The WSO2 reference architecture for the Internet of Things [75]

In more detail, the ETSI M2M communication functional architecture standardises the procedures for handling the M2M resources and the information exchange over the reference points. This RESTful architecture provides certain communication mechanisms (i.e. asynchronous, synchronous, store and forward) and supports standardised security mechanisms (i.e. mutual end-point authentication, optional secure sessions, REST APIs). However, the ETSI M2M does not incorporate any policy-based management to enforce the appropriate security policies in a dynamic M2M environment. The Services Capabilities Layer (SCL), which exists both in the second as well as in the third layer in the ETSI architecture, enables a service layer for the M2M applications by abstracting the complexity of heterogeneous devices and thus enhancing M2M interoperability. Nevertheless, the ETSI M2M architecture lacks a service-oriented approach, which allows the integration and development of different services in highly heterogeneous environments. Moreover, the ETSI M2M [ETSI TS 102 689] service requirements [76] provide the service primitives needed for service request/response and the capabilities that support the communication among the devices in heterogeneous environments. Apart from the 3-tiered ETSI M2M architecture, other service primitives and communication frameworks for smart appliances [77] have also been developed to cater for smooth communications and for sharing information among smart appliances combined with remote applications. Such primitives can be found in the Home Gateway Initiative [78] that portrays the business requirements and the architecture,

whereas the Wireless Home Area Networks (WHANs) [79] standard depicts how to connect the devices and gateways for home automation and service systems. Aiming to leverage the implementation of these evolving technologies, the ETSI HGI open platform [80] successfully addresses all the requirements for the service delivery, such as the software and firmware management. Also, the TIA - 4940.005 [81] reference architecture advocates the authentication, the authorisation and the accounting services to other network entities in order to establish the appropriate security policies. The oneM2M specification [82] focuses on the operations of the M2M applications supported by the policy enforcement and security solutions [83]. The Open Mobile Alliance Lightweight Machine-to-Machine (OMA LWM2M) architecture enables the enforcement of the access control decisions [84].

Several cloud-enabled IoT platforms and middleware have emerged to provide a secure framework for personal devices and things to communicate. These solutions facilitate the information exchange with third parties and other individuals. The connection of further M2M applications to the IoT platform is possible via the respective REST APIs. Table 6 presents some of the most known open-source and proprietary IoT platforms.

Table 6. IoT Platforms

Open-source	Proprietary platforms
FIWARE [6]	AWS IoT [7]
Mainflux [8]	CISCO IoT Cloud Connect [9]
OpenMTC [10]	IBM's Watson IoT [11]
SiteWhere [12]	Microsoft's Azure IoT Hub [13]
Webinos [14]	Samsung's SmartThings [15]

IoT-A [85] describes an architectural reference model (ARM) that establishes a common ground for IoT architectures and systems. This consists of several sub-models, which conceptually contain the basic aspects (i.e. information flow, functional structure, communication means, security, privacy) for each sub-model. IoT-A focuses on the interactions of the communicating systems between different stacks and key elements, such as devices, applications, end-users and network. CISCO has proposed the IoT Reference model [71], which consists of seven layers based on bidirectional control information flows. The CISCO IoT reference model addresses the security challenges related to the authentication, the authorisation, the network enforced policy and the security analytics. This model defines the functions required for an IoT system as well as the tasks management per level. For instance, the connectivity layer supports the transmissions between the device and the network domain. Additionally, the model supports the communication across the networks and the low-level information processing with the upper next layer, which manages the data element analysis and the transformation. Nevertheless, this model does not address the integration with any other technology in the IoT, since the overall complexity is based on a single technology stack. A summary of the reference models is depicted in Table 7.

[6] https://www.fiware.org/
[7] https://aws.amazon.com/en/iot/
[8] https://www.mainflux.com/
[9] https://www.cisco.com/c/en/us/solutions/service-provider/iot-cloud-connect/index.html
[10] http://www.open-mtc.org/
[11] http://www.ibm.com/internet-of-things/
[12] http://www.sitewhere.org/
[13] https://azure.microsoft.com/de-de/suites/iot-suite/
[14] http://www.webinos.org/
[15] http://www.smartthings.com/

Table 7. M2M references models and supported capabilities

Reference models	M2M/IoT	Security Access control	Policy decision-making
ETSI M2M	O	O	
ETSI HGI / WHANs / Smart Appliances	O		
TIA - 4940.005	O	O	
oneM2M v2.10	O	O	O
OMA L2M2M	O	O	O
IoT-A	O	O	
CISCO IoT	O	O	
RAMI	O	O	
IIRA	O	O	
IBM	O	O	
WSO2	O	O	O

The reference architectures have a noteworthy place in the analysis of the relevant literature. In regard to these architectures, RAMI 4.0 goes beyond the IoT adding manufacturing and logistics details. The IIRA model is focused on the Industrial data, the respective components and how these influence each other. The viewpoints are classified in five functional domains:

- Control
- Operations
- Information
- Application
- Business

RAMI supplements the model with the values of types, instances and hierarchical levels. Guth et al. [86] analyse several state-of-the-art platforms such as *OpenMTC* an IoT middleware consisting of gateway and backend functionalities, *FIWARE* an IoT and Smart Cities platform, *SiteWhere* an open platform for monitoring and controlling IoT devices based on µServices, and the *Amazon Web Services IoT* cloud computing services to allow connected devices easily and securely interact with cloud applications and other devices. Additionally, the authors introduce an IoT reference architecture, which is compared to the IoT ARM [87], IoT-A [85], Cisco and WSO2 reference architectures, and the Zheng et al. proposed architecture [88]. The review highlights that these approaches need to provide common definitions and components of the architectural design capable of addressing all the relevant industrial needs and the evolving technology innovations. IoT-A perspective is semantic oriented where the interpretation of data and information covers the modelling and structuring of IoT business process management. Moreover, the architecture analyses virtual entities, IoT services and cross-service organisation from the functional, information and domain viewpoints. From the architecture viewpoint, the agents and the code are defined on domain-specific devices tailored to the implementation. IIRA focuses on these aspects as well but remains closer to the business and use cases.

Mehmood et al. [89] deliver a detailed survey of M2M communications and an overview of the M2M services. ETSI TR 102 691 [90] analyses the M2M service requirements and the related challenges, while other surveys are quite focused on mobile networks, such as the Long-Term Evolution-Advanced (LTE-A), the 5G networks and the 3GPP Service accessibility [91] [3GPP TS 22.011], the architecture description for group radio access networks [92], the architecture enhancements for packet data networks [93] and the 3GPP system

improvements for machine-type communications [94] [3GPP TR 23.888], which also lack of policy enforcement and service orientation.

In addition to these research activities, Alshuqayran et al. [95] perform a literature review into the available studies on µServices and the relevant architectural challenges. Although they underline the key-role of service discovery, registration, service registry, performance, fault-tolerance, tracing and logging, communication and integration, these conclude with the open architectural issues and the gaps in µServices research. As a consequence, the dominant security controls required by the µServices to handle the advanced M2M communication requirements with policy-based management techniques need to be further analysed. The future networks require from the architectures to satisfy a large number of connected devices and manage the heterogeneity of the devices besides the technology diversity. The survey work done by Lin et al. [96] presents the enabling technologies and presents various security and privacy issues in IoT and fog/edge computing that have been proposed to be implemented in real-world applications.

2.4.1 Data management

There are several data management risks related to the IoT data model efficacy. First, the industrial Internet devices generate a massive amount of data that need to be processed and stored, a situation that certainly affects the data storage capabilities and the capacity in most organisations. For instance, moving the data to fog/edge computing and providing an application platform as a service (aPaaS) can be an option to lean toward more distributed solutions and architectures with tiered mini centres. Aiming to alleviate the increased patterns of data traffic and provide services closer to the originators and the clients, it would be interesting to test this scenario, which still remains an open challenge.

Focusing on the IoT data security, most of the data is unstructured with myriads of formats to accomplish the business-related tasks. The increasing volume, the variety and the necessary higher velocity of data processing to recover/discover information and identify patterns related to location, environmental conditions, health-related and medical status, monitoring of vibrations and material conditions in buildings can affect the design of the data centres. This data influx imposes the necessity to develop the appropriate data management models with advanced data mining tools. The work in [97] discusses the managerial and technical challenges that enterprises face in adopting data management, data mining, privacy and security. Regarding security, the connected devices can be exploited by various data security vulnerabilities such as lack of transport encryption, insecure web interfaces, inadequate software protection and insufficient authorisation. As breaches may have devastating consequences and negative ramifications for the enterprises, the organisations and the consumers, it is imperative to protect and secure the IoT data. For instance, the Meltdown, a well-known processor vulnerability, was published in January 2018 affecting several modern processors. A successful exploitation of this vulnerability could allow an attacker to get access to sensitive information (i.e. passwords) inside protected memory regions. Hence, there is a need to maintain data privacy and form the necessary security controls for secure and structured access in the architectural design choices, which requires the proper identification, authentication and encryption means.

The handling and sharing of sensitive information within special interest applications and industries (e.g. hospitals, military facilities) is an area that is expected to proliferate widely. Likewise, real-time and authorised access as well as data availability are becoming critical

aspects for access control policy models. In this scenario, spatio-temporal (time and location) factors can be checked in case of break-the-glass access control (BTG-AC) cases [98] along with emergencies to grant access in a controlled and exceptional manner [99]. This could result in policy extensions tailored to these application-specific requirements. The context-aware applications and services can also be used via the application programming interfaces (APIs) to enable the policy extension to meet the intrinsic needs. For instance, like in [100] the context-aware transportation parameters can offer numerous services in vehicular ad hoc networks (VANET) based on the parking policies, traffic policies and fuel-type policies. Maw et al. [101] propose the extension of BTG-AC model to address the data availability and detect the security policy violations in a hospital environment that utilises a WSN. Additionally, the authors contrast and evaluate their proposal that introduces a behaviour-based trust model aiming to manage effectively data access decisions and consequently increase data availability at any time. The proposed lightweight BTG-AC incorporates the policy enforcement component and the authentication service between the users and the sensor nodes to accommodate the access control decisions using the users' location and privileges. The BTG-AC model supports the detection of security violations by examining the respective audit record with the hosted prevention and detection mechanisms.

2.4.2 Policy-based management models

The considerable benefits of the PBM approach grow as the M2M communications evolve and the resources become more complex. The resources are available to the interconnected components and can be accessed using interoperable services, while the existence of a service orchestration can automate the configuration of the entities and controls the allocation of resources. The use of SOA allows the loose coupling of the web services that can be combined and reused contrary to other component models and architectures. By using the PBM methods, the operations of the resources can follow certain rules, bring together the conditions on how to use their outcome and simplify the enforcement of the relative actions.

Policy based management frameworks need to consider the most significant challenges and address efficiently the limitations of M2M communications such as the limited policy storage and the hardware computational restrictions of the connected devices. Owasmi et al. [102] demonstrate a shared policy to store, locate, access and execute policies for WSN in which the policies are distributed among the memory of the nodes in the WSN. In the literature, several distributed and cross-domain resource sharing and access control models have been proposed [103] [104], which take advantage of the distributed systems across application domains to accommodate shared policy storage and enforce the appropriate event-based actions. The events trigger the enforcement of the appropriate usage and access control policies including the authorisation decisions. PBM can be used to manage the network and the access control policies in a simplified manner. The PBM key-elements are defined [105] as follows:

- *Policy Enforcement Point* (PEP): The logical entity that performs the decision requests, receives the policy updates, translates the updates appropriately and enforces the policy decisions.
- *Policy Decision Point* (PDP): The logical entity that evaluates the applicable policy against relevant policies and attributes, makes admission policy decisions and is also responsible for relaying the information to the PEPs.

- *Policy Information Point* (PIP): The logical entity that acts as a source of attribute values to make a policy decision.
- *Policy Administration Point* (PAP): The logical entity that provides the authoring and maintenance of a policy or policy set(s). This includes a policy store, which is a repository for the policies.
- *Policy Retrieval Point* (PRP): The database where all the XACML access authorisation policies are kept.
- *Policy management service*: This is a graphical user interface (GUI) for defining, changing and managing policies.
- *Dedicated policy repository*: This is a place to store and retrieve policy information, such as an LDAP server or a Directory-Enabled Network (DEN) device.

The aforementioned building blocks and their interactions are illustrated in Figure 10.

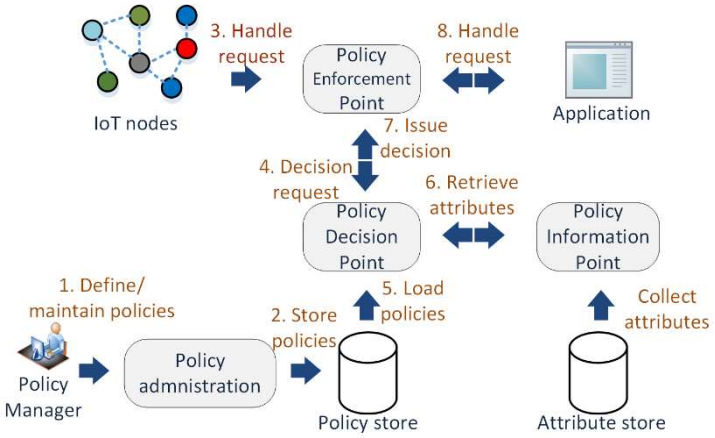


Figure 10: Policy-based enforcement system of ASPIDA

In the scientific literature, there are several policy specification frameworks have been established and evaluated extensively [105][106], such as Ponder, KAoS, WS-Policy and the IETF policy framework [107], all of which have been considered in the proposed architecture. The ultimate aim of PBM is to manage the business goals and deploy a set of policies that govern its behaviour. Therefore, the PBM systems need to utilize a policy language with the purpose of expressing the authorisation policies utilised in various access control implementations and authoring the policies, which are composed of entities, sets, rules and obligations [108]. Several approaches describing the policy management specifications to account for network and security policies written with policy expression languages have been proposed in the literature [109] [110]. Some service providers [111] have developed their own language for proprietary usage that pose maintainability and longevity problems along with support issues. On the contrary, a formal access control policy language can ease the semantic integration, convey the policy requirements and express complex policies in a secure and interoperable way to produce admission policy decisions such as by using the eXtensible Access Control Markup Language formalisation [112]. The authors in [113] address the semantic gap between the policies and the low-level mechanisms by forming a simulation apparatus with various high-level policy languages for cross-domain policy enforcement (i.e. XACML, WS-Policy). A formal access-control policy language should be utilised to describe how to evaluate access requests according to the policy rules and then provide real-time admission policy decisions [114]. Mazzoleni et al. [115] express the need to introduce a policy language for policy integration and for transferring dynamic security context, as a standardised policy

language can ease the interoperability between policy control implementations. Dell' Amico et al. [116] introduce a hierarchical security policy language for distributed environments and the policy enforcement with the use of distributed reference monitors. Nevertheless, several challenges emerge regarding the security policy approach, which needs to face policy composition issues, as well as to enable the policy enforcement, validation and conflict resolution. Further implications may arise in the case of distributed, multi-domain and often unattended connected devices environments such as the semantic interoperability, the definition of concrete policies, the multi-domain policy consistency, the security policy refinement and the policy completeness. For example, Neisse et al. [117] define a metamodel for policy specification using the interaction system design language to generate and refine concrete usage control policies.

Other relevant research activities present PBM approaches, so that networks can autonomously manage themselves. Sicari et al. [118] integrate the policy-based framework for security policies with networked smart objects by implementing a distributed IoT middleware platform, while Ferraiolo et al. [119] present an innovative policy machine that manages access control policies independently of the hardware and software configuration. Even though there is no common set of hardware and software capabilities among the smart nodes, the nodes need to interoperate by defining and establishing the proper common sets of security policies and services. By analysing the outcomes on the heterogeneity of networks in the IoT in [120], the authors address the difficulties and the key-issues with respect to self-organising protocols and modelling. They propose a distributed learning method that can be reinforced by setting the policy-based QoS requirements. The results of the policy-based system to achieve adaptive QoS routing with increased performance and learning capabilities are published in [121]. The main goal is to provide improved performance and minimize the overhead, while other research activities incorporate SLA capabilities in the policy-based system [122]. The increasingly voluminous interconnected devices raise various scalability, interoperability and serviceability, while satisfying the resource constraints of the environment. To tackle this challenge, the policy tool can mitigate network failures and security attacks without necessitating the development of sophisticated protocols and mechanisms. For instance, in the case of a failure or a physical topology change, it should be feasible for the policy engine to trigger an event-based policy action and reconfigure the network.

By evolving from ASN to M2M communications and IoT ecosystems to CPS, the access control models need to be enriched to efficiently handle attacks and violation attempts. In this direction, Um et al. [123] propose a Social-Cyber-Physical (SCP) infrastructure, which is the adhesive entity of the three distinct worlds (i.e. of society, cyber-space and physical world). Like any other ICT infrastructure, SCP suffers from *trust* problems that is of paramount importance when it comes to specific socio-economic systems (e.g. government, economy). Akin to a policy-based enforcement system, the authors present an ITU-T trust framework, which consists of the following major parts.

- The *trust agent* (TA): The logical entity that collects data from CPS environments
- The *trust analysis and management platform* (TAMP): The logical entity that models, reasons and manages the data collected from TA
- The *trust service enabler* (TSE): The logical entity that provides (a) trust knowledge of objects that relate to the service and (b) capabilities of adapting new and legacy services

- The *trust service broker* (TSB): The logical entity that shares and disseminates domain-specific trust knowledge across domains via TAMPs.

The ITU-T trust architectural framework (Figure 11) aims to strengthen SCP and stimulate activities towards a standardised development.

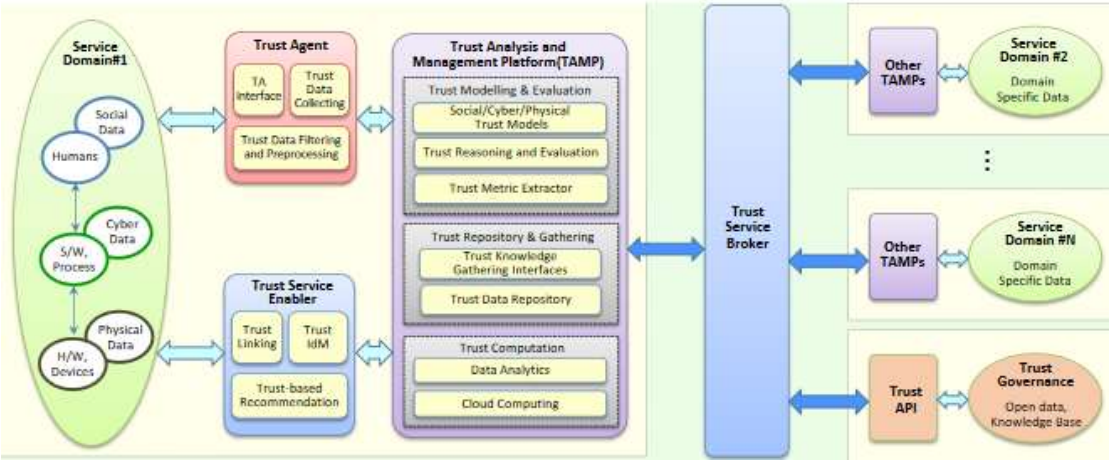


Figure 11. ITU-T trust architectural framework

Um et al. address the absence of a unified approach of trust, security and privacy that should be coupled for CPS. The modern CPS solutions need to overcome not only the technical, but also the economic, social and political challenges, all of which illustrate the necessity for adaptability and increased flexibility in the dynamic environment with changing conditions. As regards policy-based sensor network management, Li et al. [124] propose a policy-based assistance system (PAS), which addresses the issue of resource limitation in WSN by encouraging collaborations of sensors. PAS comprises of four main components:

- *Communication interface* to exchange data with other sensors
- *Clustering module* to cooperate with the sensors in groups
- *Policy management* to allocate and reserve policies
- *Policy evaluation* to assess and direct sensor requests for optimizing the availability

The authors key idea for this optimization is to form virtual groups of sensors of similar capabilities and then perform the policy-based request assignment. Nonetheless, various challenges still exist on how to orchestrate the components across the services to optimize the workflows and cut down on load time due to downstream service dependencies. In [67], the orchestration of different authentication and authorisation processes is presented with formal policy-based methods to provide secure access to the resources. More specifically, this work provides the authorisation needs and the architectural artifacts how to facilitate the development, the enforcement and the management of the access control policies.

2.4.3 Policy-based Service-Oriented Architectures

Most of the pioneering advanced business applications can benefit from solution frameworks based on the service-oriented architecture (SOA). SOA captures several architectural best practices, such as providing higher flexibility by enabling the development of the business applications more quickly and coping with the emerging web activities besides the changing demands. SOA is also deemed to be dominant in distributed, decentralized and grid-based heterogeneous environments, like cloud computing, by providing the resources as

services. However, given that data owners and cloud service providers are not in the same trusted domain, several security and access management risks arise. Several studies [125] have reported access management issues and inconsistencies in managing identities and services, such as duplication of identity and role information across the business processes, lack of identity data aggregation and increased complexity in the SOA design [126]. Complexity produces configuration and management overhead (i.e. repetitive tasks for the account and authorisation management), and results in higher administrative costs along with reduced usability. These are common complications for the deployment of SOA applications. In order to ensure a high assurance in the provision of the IAM implementations, the related access control policies need to be validated and utilised properly. In [127], several frameworks have been analysed and the authors propose to enforce an end-to-end access control model for web applications. In this proposal, the access model decision depends not only on the user intentions, the rule sets, the group memberships and the user attributes, but also on the enterprise architecture, the cross-domain deployments, the trust model and the authorisation challenges for each access control system. Among many others, granularity, manageability, delegation, revocation and composition are critical success access control factors [128].

2.5 M2M security solutions

Security by design for M2M communications should be implemented to eliminate, minimize, or mitigate various threats. The systems can be protected by network security principles (i.e. ISO/IEC 27033), storage security (i.e. ISO/IEC 27040) and other infrastructure guidelines (i.e. ISO/IEC 27003), whereas the application security needs to be realised during the SDLC of the applications and especially the design, the development and testing phases. In this direction, the relevant security controls (e.g. secure HTTP headers, JSON/CBOR web-tokens and cross-site scripting protection) [129] need to be measured and controlled with the proper security mechanisms and frameworks, as the applications need to continuously evolve to meet the business needs, adapt to improvements in the processes and securely facilitate the technology upgrades. For instance, ISO/IEC 27034 offers guidance on application security, whereas the Open Web Application Security Project (OWASP) testing framework (ISO/IEC 27034 for application security) is widely used to evaluate the application security within SDLC. To measure the occurrences of specific vulnerabilities, a technique is to measure the vulnerabilities as software defects per thousand lines of code (KLOC) so as to provide an indication of the code quality and important security flaws. To avoid security attacks on the applications, the application security risk model (ASRM) may be used to the determine the risk levels in the applications and the proper design of cost-effective solutions based on the appropriate mitigation approaches.

Aiming to increase the maturity of the security engineering processes, the Systems Security Engineering Capability Maturity Model (SSE-CMM/ ISO/IEC 21827) captures several practices in use by the industry and describes the characteristics of the security processes. For instance, the **c**ritical infrastructures (e.g. industrial control systems, water and power plants and defence bases) are considered as complex systems with deep interdependencies and highly sensitive systems. These dependencies can be dynamically quantified [130] to prepare for failures (i.e. the critical infrastructures are not operational), otherwise the failure duration may increase exponentially. For instance, a supervisory control and data acquisition (SCADA) is a critical infrastructure industrial control system. In regard to this type of cases, Alcaraz and Zeadally [131] consider the need for prioritising the respective security requirements such as availability, integrity and confidentiality of the sensitive information. They present the

importance of the security, the reliability of the information and the communication. The authors identify the security requirements in the field of business continuity in critical contexts, which include *Performability*, *Interoperability*, *Scalability*, *Extensibility, Availability, Reliability, Resilience, Safety, Criticality, Autonomy and self-healing*, *Usability, and Trust with collaboration.*

These can be used between the heterogeneous objects in order to address anomalous and threatening situations, while maintaining *fault tolerance* and a stronger sense of *security*. Knowles et al. [132] pose the challenge of balancing safety and security requirements by engaging the concept of *functional assurance*. The functional assurance goals include *confidentiality*, *integrity* and *availability.* The assurance is achieved by assessing the performance metrics of an industrial control system to determine if failure conditions (i.e. security and safety states) exist. A fundamental issue is to obviate either injuring humans or destroying control systems. To maintain the integrity of the data of the connected devices and the services, requires secure preventive mechanisms in terms of shielding the systems from physical attacks (e.g. sabotage). The systems need to resist and recover from malicious attacks or system corruptions. It should also be decided whether to place preventive and defensive mechanisms before further propagating important assets to the interests of the enterprise. The organisations invest significantly in reducing the response time and remediating the impact of any potential security incident. The Security and Event Management (SIEM) solution can be enhanced with advanced statistical analysis and automated security events supported by the appropriate security policies to improve the breadth and depth of the access management coupled with security capabilities. Therefore, it is crucial to design, define and ensure the proper adequate security policies that can be improved by dynamic access control mechanisms to support the changing conditions (i.e. time-ranges, significant variations in weather conditions, climatic changes, health deterioration, domain and runtime identities). The inability to react and adapt quickly to the changing conditions may result in dire consequences for the object (i.e. health of the patient) or for the environment (i.e. a fire in the woods).

Regarding the industrial IoT, new security and privacy challenges arise, because of its increased diversity and large volume of nodes that have dramatically limited resources. Within the industrial IoT ecosystem various security goals exist, which require certain planning to prevent the illegitimate access to the infrastructures and services. With the aim to avoid unnecessary delays in production and loss of revenues, Sadeghi et al. [133] argue that IoT systems require protection against DoS attacks and higher availability. They report that several attack surfaces of IoT currently exist in electronics, software and also human interaction. The main security objectives of industrial IoT include:

- Availability
- Prevention of any system physical damage or harm to humans
- Authenticity and integrity of any production-related information
- Confidentiality of code/data/configuration information

In [134], the full-threat landscape of ad-hoc and sensor networking for M2M Communications is explained and analysed. A full taxonomy map of ASN threats is presented uncovering physical (i.e. disaster, natural outages, hardware failures/malfunctions), cyber-security (i.e. espionage, interception, hijacking), cross-domain (i.e. device, network, application) and non-IT related (i.e. regulation, violation of law, contractual implications) security issues aiming to support decision makers to take informed decisions regarding cyber-

security, M2M communication security and privacy-friendly design of systems and services. Additionally, Wang et al. [135] survey and evaluate the potential data privacy threats related to the M2M data transmissions, such as the unintentional damage and loss of information, the impersonation attacks, the eavesdropping of user data and control signalling, the manipulation of exchanged data and the user privacy disclosure (i.e. sensitive data on location). Barki et al. [136] propose that several design challenges exist in securing the M2M infrastructure:

- Scalability (i.e. scalable authentication mechanisms for real-time applications and scalable key management [137] for the large volume of devices)
- Device heterogeneity (i.e. distinct security mechanisms)
- Resource constraints (i.e. utilisation of lightweight symmetric cryptography)
- Several types of end-to-end communications (i.e. security mechanisms for group and peer-to-peer communication)
- Delay constraints/real-time communication (i.e. quality of service must not be affected by the overhead of security mechanisms)
- Robustness (i.e. to employ tamper resistance for avoiding physical attacks)

Fadel et al. [138] summarise the security challenges at various levels and focus on the design objectives for WSN:

- Reliability
- Memory management of the nodes
- QoS-aware application protocols
- Low-power consumption
- Security

Additionally, Barki et al. [136] categorise the security issues of M2M communications into six groups:

- Key management
- Data-origin authentication
- Entity authentication
- Privacy
- Data integrity
- Device integrity

The authors illustrate the paramount importance of the Identity Based Cryptography (IBC) and the ephemeral identity (pseudonym) in terms of privacy. In order to protect the confidentiality in M2M communications, various transport layer security protocols are used to establish a reliable and secure end-to-end communication path. Within the context of M2M communications, the handling of sensitive data in business-critical environments (e.g. hospitals, industry and military facilities) poses certain security questions in the communication process. For instance, the M2M communications are widely used in improving the functionality of critical infrastructures (e.g. industrial control systems, water and power plants and defence bases), which are complex systems with various interdependencies and highly sensitive information. The addition of computing elements to traditional physical components results in increased complexity and hampers the insight into how the elements of the system interact with each other. The increased complexity may trigger human mistakes and operational errors allowing an attack or an intrusion. Considering these issues, Etigowni

et al. [139] propose the cyber-physical access control solution (CRAC). The authors focus on a smart grid that comprises of sensors and programmable logic controllers (PLCs). The logical policy enforcement is achieved by utilising information flow tracking and logic-based context-aware policies to block arbitrary operations that can impair the entire system or may allow illegitimate access to sensitive data.

One of the major security requirements in policy enforcement is the detection of policy violations. Maw et al. [140] propose a model that supports the detection of security violations by examining the respective audit record in the hosted prevention and detection mechanisms. The authors also introduce access decisions for authorisation, operational and obligation policies. Singh et al. [141] enhance the policy enforcement with event-based systems and address cross-domain policy issues, whereas Sicari et al. [118] propose an enforcement engine, which supports the management of new policies at runtime without any service disruptions. Other studies [142], [143] also demonstrate how to enforce security policies and integrate the defined security capabilities (e.g. identity management, authorisation service) to present more secure communication models. The above-mentioned security requirements in M2M communications need to be considered and addressed efficiently.

2.5.1 M2M Working Groups and research activities

Various standards [144] developed by ETSI, IEEE, 3GPP, NIST, Open Mobile Alliance (OMA), Mobile and Wireless Communication Enablers for twenty-twenty (METIS2020), oneM2M and Expanding LTE for Devices (EXALTED) are presented in Figure 12.

Figure 12. Standardisation bodies and WGs for M2M communications

The acquired knowledge needs to be shared with all interested security actors (ISPs, CERTs, security vendors, etc.), enabling them to make sound security investment decisions and focus on the most dangerous activities first. Exceptional care can be devoted to impact the level of confidence of the European citizens in the net economy by leveraging security awareness in Europe thanks to the gained expertise. Most of the research and project activities have been conducted to achieve better quality, application and novelty levels. Based on their relevance to the security and operational aspects of ad-hoc and sensor networks, a reduced subset out of a large number of projects is presented below providing the short details on the respective projects.

- *SemSorGrid4Env* [168]: Specify, design, implement, evaluate and deploy a service-oriented architecture and middleware, which allows application developers to build open large-scale semantic-based sensor network grids for environmental management.
- *SENSEI* [169]: Heterogeneous wireless sensor and actuator networks (WS&AN) should be integrated into a common framework of global scale and made available to services and applications via universal service interfaces.
- *SERVFACE* [170]: Service-oriented Architectures - User interfaces together with complex control logic must be developed as an additional layer on top of services.
- *SHAPE* [171]: Semantically-enabled Heterogeneous Service Architecture (SHA). SHA extends Service Oriented Architectures (SOA) with semantics and heterogeneous infrastructures (Web services, Agents, Semantic Web Services, P2P and Grid) under a unified service oriented approach.
- *SMARTSANTANDER* [172]: Exhibit the diversity, dynamics and scale that are essential in advanced protocol and algorithmic solution for IoTs.
- *SMART-Net* [173]: Comprehensive and configurable experimental facilities for investigation of future wireless networks.
- *SOA4All* [174]: Provide a comprehensive framework that integrates complementary and evolutionary technical advances (i.e., SOA, context management, Web principles, Web 2.0 and semantic technologies) into a coherent and domain-independent service delivery platform.
- *SOCRATES* [175]: (Self-Optimization and self-ConfiguRATion in wirelEss networkS) The project aims to develop self-organisation methods in order to enhance the operations of wireless access networks by integrating network planning, configuration and optimization into a single mostly automated process requiring minimal manual intervention.
- *SPITFIRE* [176]: Semantic-Service Provisioning for the Internet of Things using Future Internet Research by Experimentation) project is to reduce the effort required for development of robust and interoperable applications in the Internet of Things. This facilitates the building of new kinds of applications and services that were not possible before, thus having an impact on research, industry and private households.
- *TAMPRES* [177]: Improve the trustworthiness of WSN (i.e. prevention of side-channel and fault-injection attacks, provision of flawless implementation of lightweight cryptographic cores, attack resistant system architecture).
- *TECOM* [178]: The Trusted Embedded Computing (TECOM) project aims at developing trusted computing solutions for embedded platforms, which means, that TECOM ensures especially the security and safety of embedded computing systems and infrastructures.

- *VITRO* [179]: Develop architectures, algorithms and engineering methods, which enables the realisation of scalable, flexible, adaptive, energy-efficient and trust-aware Virtual Sensor Networking (VSN) platforms. Simplify the discovery and management of the underlying hardware and software resources of large collections of heterogeneous smart objects and scalable inter-objects collaboration. Although VITRO aims to be application-neutral, the architecture and protocol toolbox is validated through extensive simulation testing and furthermore implemented in a large trial of more than 500 sensor nodes.
- *WISEBED* [180]: (Wireless Sensor Network Testbeds) The project provides a multi-level infrastructure of interconnected testbeds of largescale wireless sensor networks for research purposes, pursuing an interdisciplinary approach that integrates the aspects of hardware, software, algorithms and data. This demonstrates how heterogeneous small-scale devices and testbeds can be brought together to form well-organised, large-scale structures, rather than just some large network; it allows research not only at a much larger scale, but also in different quality due to heterogeneous structure and the ability to deal with dynamic scenarios, both in membership and location.
- *WOMBAT* [181]: The WOMBAT project aims at providing new means to understand the existing and emerging threats that are targeting the Internet economy and the net citizens. To reach this goal, the proposal includes three key work packages:
    - real time gathering of a diverse set of security related raw data
    - enrichment of this input by means of various analysis techniques
    - root cause identification and understanding of the phenomena under scrutiny

2.6 Open challenges

As we witness a rapid evolution of X-computing along with the widespread availability of wireless and mobile networking technologies such as the always on, always connected (AOAC) usage model, the technological advancements bring forth the various challenges for well-established, secure and trusted communications. Due to the heterogeneous characteristics, the restricted nature along with the limited storage and processing capabilities of the nodes/devices, several challenges remain to be addressed. X-computing requires the devices to operate efficiently with low-latency, high resilience and high service reliability considering the critical nature of a number of M2M applications (e.g. in healthcare, critical infrastructures, transportation etc.). The M2M design should deliver low-power consumption solutions with small-sized data transmissions at irregular intervals and be able prolong the network lifetime. The type of the M2M service (i.e. health application, Intelligent transportation system), the random and nonrecurrent events for a variety of reasons (i.e. traffic congestion, parking difficulties and free space, weather conditions such as heavy rainfalls/storms, urban or rural contexts, location, density changes of population) and the condition details of the objects (i.e. speed of travel, full-deliveries, humidity levels, temperature, shock or damage suffered, high blood pressure, cardiac arrhythmia, cardio/vascular problems) highlight the requirements for increased flexibility and adaptability. Therefore, the shifting dynamics of the environment i.e. mobility, dynamic conditions, failures impose the need to adapt and respond to changes based on the given security and network policies and take all the proper event-based actions.

The conventional application security solutions and the built-in security need to be enhanced with policy-based management, event-based management options for adaptive routing and secure access controls. In more detail, there is a need for a policy enforcement to

apply the appropriate security policies and ensure data privacy and effective countermeasures in a dynamic environment. The proper security policies should be enforced in the M2M services effectively and dissuade from any security policy violations. Any potential service inconsistencies and policies need to be detected and avoided quickly. Among several other security solutions, the authentication and authorisation of network access, the access control mechanisms, the data confidentiality and integrity, and the M2M service availability are among the critical success factors for the delivery of secure M2M services.

The rapid development of the M2M communication technology stack and the changes in various applications including firmware upgrades require the detection of the emerging threats and mitigate them, assess the unknown vulnerabilities and addressing the issues in security-threatening environments with M2M sensing devices. The security controls need to protect efficiently the connected devices and the respective applications. A typical example the IoTroop/Reaper malware that refers to a botnet in September 2017. The attackers can manipulate to get dynamic control of the infected devices and then target anything with an attack code. The proper evaluation of the security controls and as the policy-based administration can mitigate security risks to avoid software-related vulnerabilities and remedy the security weaknesses. Additionally, the imminent quantum computing advancements can reveal new security challenges for common legacy public-key cryptography. Doubtless, a greater consideration should be given in order to ensure that the appropriate security policies and controls provide the adequate security service levels (in terms of security, confidentiality, authentication, access control, non-repudiation, audit, integrity, authenticity of data, availability, accountability, privacy, trust) at every (local, national, global) level across multiple (personal, cross-governmental/inter-agencies, commercial, industrial) domains.

The underlying network, the protocols and the topology (i.e. the heterogeneous ad-hoc sensor nodes, the M2M clusters) in use play a significant role in the performance, deployment and operations capabilities. For instance, the non-position oriented protocols do not require localization mechanisms or GPS for the location knowledge, thus they are more energy efficient and can prolong the lifetime of the networks and the nodes. From another perspective, the nodes can be deployed in remote locations with very little or no human intervention contingent upon the network topology and the dispersion of the nodes. The unattended nature of the devices also raises certain management and security issues. Additionally, the restrictions of the computational constraints require allocating more efficiently and dynamically the energy resources (i.e. consider the power-consumption) to extend the lifetime of the nodes and manage the growing distributed data volumes of the M2M devices.

Finally, metering efficiently the technology capabilities of containerized cloud-based applications requires advanced service models for M2M systems with high quality and performance service levels. Hence, emphasis is put on implementing one function or feature into smaller independent granular modules with μServices aiming at a rapid delivery of changes and increased technology flexibility compared to SOA services, which are more coarse-grained, request-driven and logically coupled.

Chapter 3

Core entities

3.1 Outline

This chapter provides the technologies and the methods that have been included in the target architecture and also explains why these are the most appropriate. It describes in more detail the research activities and the reference literature about these methods, such the policy-based network management adaptive routing, QoS routing and traffic provisioning, the event-based management and the unified access control along with the access policy engine.

3.2 Adaptive routing based on Policy-based QoS Management

3.2.1 Demystifying PBMN

In a policy-based network it is crucial to identify the appropriate policies required to deliver the desired performance and quality of service. A Policy-based Network Management (PBNM) system enables the mapping and the enforcement of policy rules to the network elements. With respect to policies, the Policy Core Information Model (PCIM) [182] defines that policies need to be consistent, reusable, clear, well-defined and manageable by the administrators of the PBNM system and the network components. Several PBNM frameworks and research activities have been published to date. Doherty et al. [183] presented a distributed data layer of increased scalability of the client applications by dynamically altering the replication schemes based on the network state. Raymer et al. [184] incorporated additional reinforcement and concept learning capabilities. Ozianyi et al. [185] demonstrated an XML-driven architecture for policy control of Differentiated Services (DiffServ) networks. These research efforts show the necessity to elaborate further upon the dynamic capabilities of a policy-based network management framework.

The PBNM system maintains the policy rules and makes configuration changes into the network via policy decisions and enforcement points. In addition to the policy enforcement, the system can also validate the policy implementation and check the policy consistency with the functional and resource constraints within the administration domain under consideration. The appropriate enforcement of policies on the network components can be validated through a systematic monitoring of the network resources. Figure 13 illustrates a typical scheme of a policy-based network lifecycle management based on the IETF policy framework specifications. Pertaining to cope with the diverse conditions of the user applications and needs as well as with the heterogeneity of the network elements with their varying functionalities in the administration domain, a network management framework for grouping objects and retrieving the fitting policy service needs to be established.

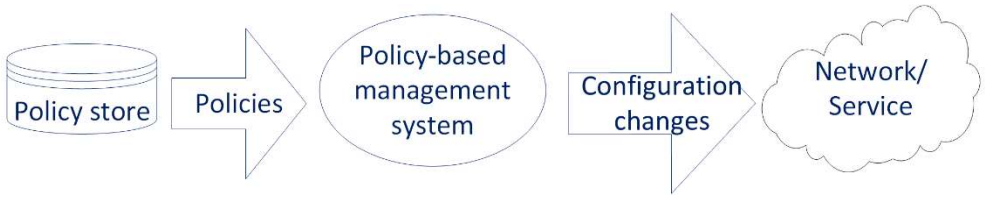


Figure 13: A typical scheme for policy-based management lifecycle

In more detail, the ASPIDA proposed framework facilitates the definition and deployment of policy rules in order to apply the appropriate resource access and usage. Nevertheless, the proposed solutions are restricted to static condition-action rules. This framework addresses the integration of SLA-driven conditions through QoS-aware mechanisms to accomplish adaptive IGP routing. SLA is integrated with QoS, forming the dynamic policy control. The performance degradation and SLA conformance are achieved by using event detectors (EDs) with the appropriate applets. These detectors are utilized to enforce the necessary actions to influence QoS adaptive routing through the monitoring of the conditions and the collection of information from the PEPs. In this context, the policy control components (e.g. PDP) can analyse the weights of the links, apply changes to the QoS service and influence adaptive routing. If the network conditions change (e.g. traffic increases), then the SLA monitoring ensures that the appropriate actions take place, when appropriate. A dedicated policy repository keeps the relevant information and provides the policy information on-demand. and the information needed to apply the policy control efficiently. The policy rules may affect the QoS adaptive routing. The QoS adaptive routing framework ASPIDA enhanced with SLA-driven conditions is presented in Figure 14.

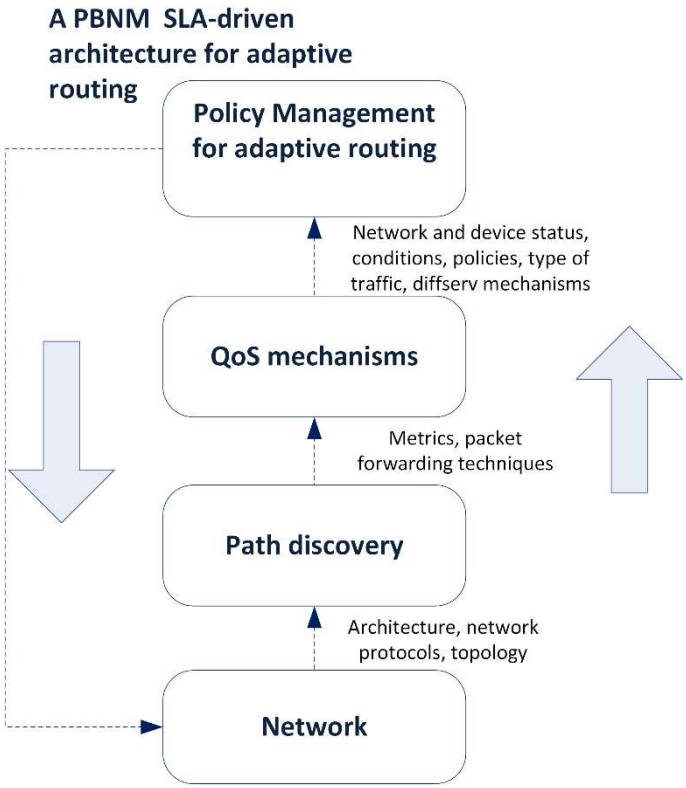


Figure 14: The ASPIDA framework

In ASPIDA, the PBNM model integrates SLA monitoring, QoS-aware and path discovery mechanisms ending up with a better network performance. The model identifies the needs through policy services, such as:

- *Device configuration* (e.g. access filter lists)
- *Queuing mechanisms* (e.g. drop strategies)
- *Traffic classification/marking/policing* (e.g. Class of Service and DiffServ marking)
- *Resource reservation* and *admission control* (e.g. bandwidth allocation, thresholds)
- *Service Level Agreements* between two adjacent domains

3.2.2 SLA monitoring

A system eligible to monitor and analyse the performance of distinct administrative domains can incorporate various SLAs to report the availability of the resources as adequate or not. The SLA records a collective understanding of a negotiated agreement about the requirements that have been established. In case of the existence of various SLA metrics, the SLA needs to report a violation whenever any of the resources fails to meet the acceptable availability thresholds. In APIDA, the SLA metrics include the followings:

- Traffic conditioners (classification, marking, metering, shaping, dropping)
- Uptime and path measurements
- Service performance indicators

In terms of traffic conditions, the majority of the IP routing protocols do not include QoS and SLA attributes in the routing path selection process. Additionally, due to the computational needs of the routing protocols regarding the path selection process, the responsibility for adapting the routing paths to the prevailing traffic falls on the network operators and the management systems. In ASPIDA, by monitoring and analysing the utilisation levels, the uptime and path measurements, the resources and the service performance indicators, the effective routing path is recalculated in order to optimally distribute traffic across the links.

3.2.3 QoS and traffic provisioning

Most QoS techniques focus on particular SLA indicators (e.g. delay, jitter, packet loss, bandwidth), because it is often difficult to deal simultaneously with several indicators and manage possible contradicting outcomes. For instance, minimizing the end-to-end delay may result in higher packet losses. Granville et al. [186] showed that policies are not sufficient to provide QoS-guaranteed services, but a global QoS management approach must also be provided. ASPIDA allows the execution of QoS management in an integrated fashion; routing computations, QoS and SLA mechanisms. In ASPIDA, the proposed model handles traffic policies to manage the bandwidth allocation. By incorporating a QoS packet scheduler, the packets can be moved dynamically at different rates in accordance with the QoS policy. Using packet classifiers, markers and policers, the model ensures the appropriate levels of traffic controls and adjusts the bandwidth consumption per classes of traffic. Moreover, this model addresses the dynamic adaptations of the QoS requirements and the policies to alleviate under-performing entities and inappropriate resource usage.

3.2.4 Adaptive QoS routing

Due to the dynamic multi-condition changes in the network environment and the rapid growth of applications demanding high bandwidth, there is an increasing need for adaptive routing capabilities for improved performance. ASPIDA runs a policy management service, which adheres to the policies and weights affecting the QoS service, and implements a policy-based adaptive routing by combining different traffic conditioning mechanisms (i.e. QoS routing). The network entities are continuously monitored to ensure that the appropriate policy rules have been applied successfully. When the network conditions change (e.g. traffic increases), then SLA monitoring and the PBNM system ensure that the appropriate actions take place.

It is vital to determine the specific node attributes and the lower-level link parameters on QoS routing, since QoS routing encompasses the collection and the maintenance of the latest

state information about the network conditions and, utilises QoS requirements in the path-finding mechanism. An adaptive QoS routing mechanism can be divided into three distinct functions. The first function provides the dynamic routing algorithm for route discovery and collects the local QoS-related information enabling routing cost optimization based on different QoS metrics. The route discovery provides a path-finding mechanism. The second function uses a local routing table, which is created with QoS related references for each node. In order to accomplish this function, a local link monitoring function is used. Typically, if a link fails, then the best route in the new topology needs to be recomputed.

Nonetheless, the applications and the business objectives requirements (i.e. time-sensitive applications with relative high importance) pose certain requirements along with the environment (i.e. the network topology), the conditions (i.e. congestion) and the state of the managed systems. All these can deteriorate the resulting routing performance. For instance, in the case of suboptimal routing, the policy-based system need to revert back the routing changes to the original configuration through the evaluation of the appropriate key performance indicators. Aiming to address this shortcoming, ASPIDA introduces a third function that utilises a final decision-making system. This system identifies and ranks alternative routing paths based on QoS constraints such as the traffic distribution rates, bandwidth reservation and the number of hops. The routing metrics are measurable values that can be changed based on the decision-making system policy rules, which influence the new path selection targeting the integration of different modules and characteristics (e.g. QoS, network diameter, transmission rates, load sharing capabilities) with the changing network conditions (e.g. traffic load, remaining bandwidth, reserved bandwidth, error rates). In the case of a link or relay node failure, the algorithm redirects traffic to other paths by supporting IGP tuning and QoS configuration changes to modify routing towards the destination. The flows are sent to the destination endpoint, if there is an alternate path. Load sharing techniques across the redundant paths can also provide increased network efficiency and performance. Other design challenges involve an intra-domain policy management for intra-domain routing or an inter-domain policy management for inter-domain routing. Finally, the model keeps a dedicated policy repository in order to store and retrieve policy information and the information needed to apply the policy control efficiently to improve the QoS adaptive routing of positive affect.

3.2.5 Event-Management Routing

Monitoring the network speed, the network connectivity, the throughput, the delay, the jitter and the packet loss are basic metrics to detect problems, reinforce a new path selection and configure changes. By observing the network and application event sequences, various event detectors can handle abnormal and irregular traffic patterns (i.e. bursts, explosion of streams and messages). Based on the capacity levels/thresholds, the utilisation levels of the links and the network components, the detectors may trigger the corrective policy-based routing decisions. If required, the SLA metrics can protect critical traffic, reinforce advanced routing capabilities (i.e. unequal cost load-balancing) and detect network outages. Therefore, the network components can be utilised to gather network performance information through SLA probes and flow export events (i.e. IPFIX, sFlow). These mechanisms can be used explicitly for application performance monitoring [187]. The information is collected by the centralized *controller*, which also performs policy checking and take informational or corrective actions to keep the traffic classes in policy or improve the resource usage. The appropriate event sequences are associated with pattern definitions and policy configuration sets to define the

methods (i.e. applets) that enable the selection of the suitable resources accompanied by the multi-site aware path control optimization.

In more detail, the *event subscribers* define the corrective actions associated with the event. The actions are administered by the appropriate *event handlers* (SNMP/XML/CLI agents, or TCL shell methods) [188]. The event manager [189] of the controller can analyse the events sent by the *event detectors* as soon the monitored events occur or the set conditions are met. Several different core components (i.e. security manager, watchdog system monitoring) publish the events following the policies set by the event subscribers. Finally, the event classifiers store the event subscribers in the database and they behave as meta-search engines to retrieve the effective policy. The integration of these components is depicted in Figure 15. First, the event detectors notify the Embedded Event Managers (EEM) and next the predefined policy deploys the configuration statements either via applets or scripts. For example, a Java applet that listens on the ports of a network component can detect a failure or service outage based on the event criteria. After that, the applet can be triggered and the remediation policy can be applied (i.e. forward the traffic within a defined number of seconds). Each policy can be compiled by different applets and is stored in the policy repository, which supports the creation, modification and editing of policies.

Chang et al. [190] presented how an event-driven policy distribution can be achieved using an adaptive policy procedure. Examples of event-driven based policy distributions include the backup schedules, the recovery processes, the database replications, the file system mirroring and the data/time based configuration of network access. These components have different functionalities, but they can be integrated into a policy-based management model. By using this model, the policies can be evaluated, stored, interpreted and translated into an understandable form to be enforced on the network and system configuration.

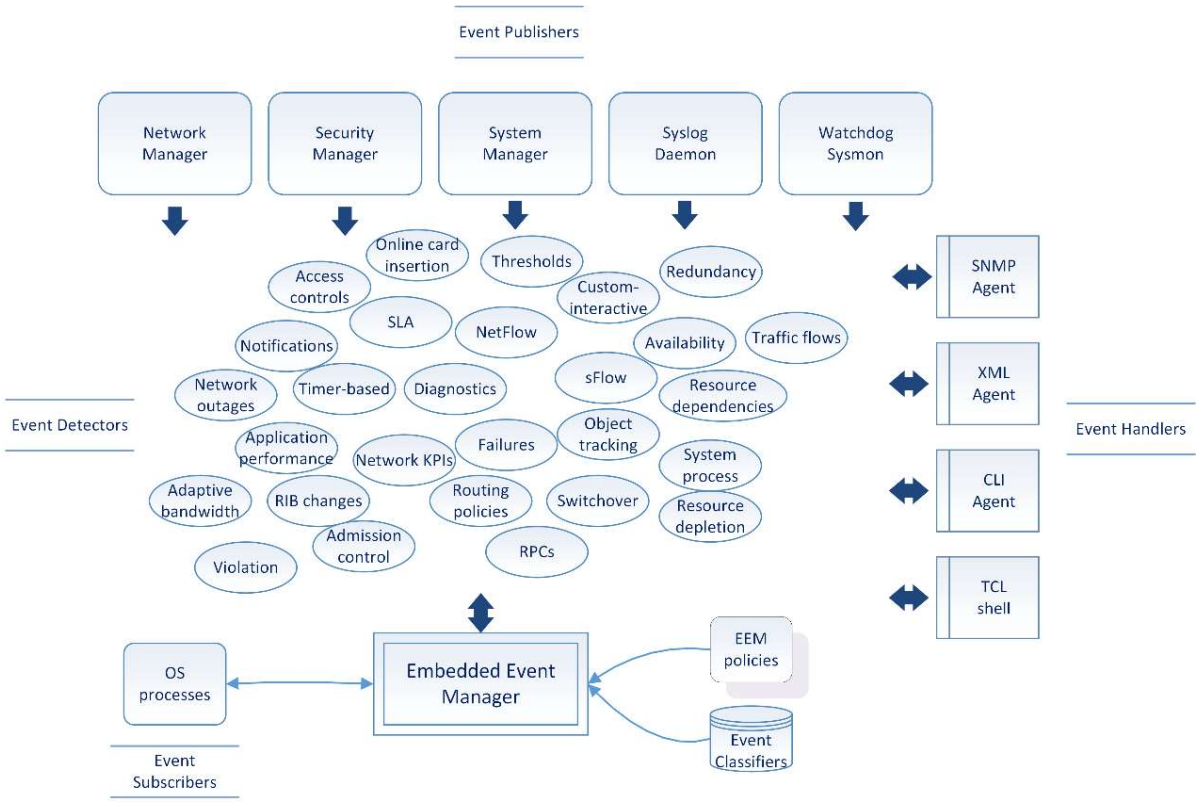


Figure 15: Relationships of event management for adaptive routing

Being able to capture the state of the network components and any excess of the resource thresholds, then immediate remediation and/or improvement actions can be easily taken.

3.3 Unified Access Control Services

As a result of the growing concerns in access management techniques, numerous access controls have been proposed recently. Additionally, in light of the emerging cloud computing era, there is an increasing interest in advanced and secure access control services in X-Computing that will improve security and compliance. In order to meet both the application access control and the X-computing requirements, there is a need to safeguard data integrity, protect data privacy, meet the appropriate audit requirements with on-demand reporting, ensure effective access certifications, support data and service availability, and increase the efficiency in policy-driven enforcement of security and compliance rules. In parallel, the security breaches should be eliminated, while the administrative and operating costs for access management should be reduced. In addition to the aforementioned policy-based capabilities for SOA, ASPIDA addresses the access control needs by integrating the IAM methods supporting the IAM lifecycle processes. The outcome of this integration has been published in [67].

ASPIDA enables the access to protected resources based on policy-driven decisions and realises the integration of the authentication and authorisation modules with the access control policies used by the application delivery platform and the service management disciplines. The policy-driven access controls and validation rules achieve increased automation, ensure compliance and deliver better access control service efficiency. The model is based on a set of standards, such as the policy language, the event monitors and the policy-based management components. The integrated access control model needs to ensure a prominent level of assurance aiming to avoid conflicts, inconsistencies and any unambiguously specified policies. The security and policy assurance, the conformance to compliance and the enforcement of the policies are achieved by using standard policy-based management components [10].

The Identity Management Engine (IDE) maintains the identities across multiple domains, distributed environments, or cloud systems. The Authentication Service Engine (AuthNSE) verifies the identities based on the identity data stored across the domains and forwards the access requests to the Authorisation Service Engine (AuthZSE) to provide the appropriate authorisation levels. Policy-driven controls and validations are utilised to ensure the consistency and the integrity of the rules and the access control services, whereas the validation rules for the policy-based management model ensure that the appropriate policies and authorisation decisions have been accurately enforced.

3.3.1 Access Control Policies

The access policy rules define the restrictions and the criteria to specify the actions and the privileges on resources; they allow or deny access to the resources. The definition of the format and the management of the access policies need to be formulated in multi-level security classes along with the enforcement of these policies. Therefore, the security and functional requirements need to be translated into access control policies by using a formal language to express and ensure the restrictions before granting permissions.

In more detail, the access control policies are composed of rules to be examined while processing the access requests. A *rule* is the most elementary unit of *policy*, which is typically

encapsulated within a policy. The rules are not exchanged amongst system entities, but PAP and PDP combines rules in a policy. More precisely, a rule includes a target, which refers to the set of subjects, the conditions (rule applicability) and the effect, which expresses the intended consequence (permit or deny) of the rule. The obligation and the advice expressions define specific actions that should be performed, when enforcing a decision. They can be associated with either a rule or a policy to return a decision provided that the applicable attributes (e.g. FullfillOn) are enabled. Table 8 provides the rule attributes and their respective values, whereas Table 9 depicts the counterpart for the policy attributes and their respective values regarding the implementation of the policies.

Table 8. Rule attribute template

Rule attribute	Value
Rule ID	A unique identifier that allows the rule to be referenced within a policy set.
Description	A textual description of the purpose of the rule. It typically provides information on most of the attributes found in this template.
Rule Target	A description of the kinds of request to which a particular rule applies. If a Rule Target is not present, the Policy Target is used to determine whether the Rule is applicable to an incoming request. When a policy target exists, it is applicable to every rule in the policy which does not have its own Target. In practice, a rule target is often more constrained than the associated policy target, fine tuning to specific Subject/Resource/Action match criteria that are in the context of the particular rule.
Effect	The intended consequence of a satisfied rule. It can take the values "Permit" and "Deny". The authorisation decision returned by the PDP to the PEP can also include the values “Indeterminate” or “NotApplicable" and (optionally) a set of *obligations and advices*
Condition (optional)	Represents a Boolean expression that refines the applicability of the rule
Obligations Expressions (optional)	Operation that should be performed by the PEP in conjunction with the enforcement of an authorisation decision.
Advice expressions (optional)	A supplementary piece of information in a policy or policy set which is provided to the PEP with the decision of the PDP.

Table 9. Policy attribute template

Policy attribute	Value
Policy ID	A unique identifier that allows the policy to be referenced within a policy set.
Description	A textual description of the purpose of the policy. It typically provides information on most of the attributes found in this template.
Rule Combining Algorithm	The procedure for combining decisions from multiple rules. Valid values for this attribute are defined below.
Policy Target	The part of a policy that specifies matching criteria for figuring out whether a particular policy is applicable to an incoming service request. Contains three basic "matching" components: **Subjects** (An actor*)*, **Actions** (An operation on a *resource)* and **Resources** (Data, service or system component)**.** Attributes of the subjects, resource, action, environment and other categories are included in the request sent by the PEP to the PDP.
Set of rules	A policy set about a specific target consists of a number of applicable policies
Obligations expressions (optional)	Operation that should be performed by the PEP in conjunction with the enforcement of an authorisation decision.
Advice expressions (optional)	A supplementary piece of information in a policy or policy set which is provided to the PEP with the decision of the PDP.

Aside from the integration challenges, the implementation of access policies can be complicated, as they need to be validated and be compliant with the business needs. Ledru et al. [191] demonstrate the validation activities in such security policies featuring

authorisations. As cloud computing evolves and delivers innovative technologies along with opportunities in reducing costs and ensuring better results in the policy compliance, further research activities can reveal how to deploy the access control policies in X-computing applications.

3.3.2 Access Policy Engine

Concerning the fine-grained authorisation, the access control mechanisms attributes and classes can be stored in the LDAP server and exploited in the Access Policy Engine (APE) for the authorisation decision process. The purpose of the authorisation decision process is to provide the appropriate level of access only to authorised users to perform specific tasks according to their role and responsibilities. APE utilises PDP to evaluate the policy rules and to determine the authorisation decisions and responses. For multi-domain and distributed environments, this can be more complicated utilising distributed decisions points, as the policy decision engine needs to be unified and consistent across all the systems, when the policy-making process relies on the conditional criteria and the priorities. Next, PDP evaluates the access requests against the authorisation restrictions, which have been defined before in PAP such as the hierarchical, fixed and variable access policies. The appropriate definition of the access policies is critical for the policy exchange messages, so the PAP provides the appropriate format of the authorisation policies and ensures interoperability and scalability. The policy makers need to consider the requirements and the needs of the role memberships and access to the resources to provision the appropriate policy rules. The PEP receives the access requests to a resource and grants/prevents access based on the policy decisions.

PIP is used to provide attribute values upon request from the PDP context to assess and evaluate the appropriate action. It retrieves the information from relational databases through SQL queries and from directories through LDAP queries. Additionally, the environmental information can also be used to provide the required attributes such as the time zone, the language settings and the last login details. Based on the requests from PEP, which can be encoded in the Simple Object Access Protocol (SOAP) messages with SAML assertions, PDP consults the attributes (PIP) and policy repositories (PRP) to provide a formal response back to PEP, as shown in Figure 16. The policy management is performed via PAP. PIP and PRP are the authoritative sources of the authorisation policy criteria and values, as the information needs to be validated for consistency checks and compliance tests. For instance, there is a need to automate the definition of policies based on the toxicity criteria in order to identify Segregation of Duties (SoD) combinations and separate tasks with role responsibilities.

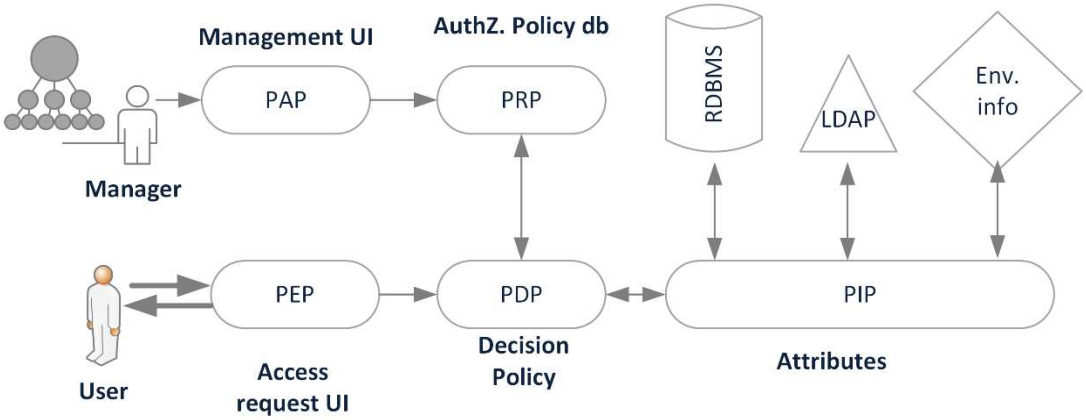


Figure 16: Access Policy Engine

Another reason to automate the access control processes and compliance checks is to avoid the duplication of role assignments. Due to the potential numerous roles, when the subjects request access to several processes, the provisioning and the assigning of duplicate roles and permissions to allow access to the resources should be avoided.

The APE can be either centralized in Identity as a Service (IdaaS) cases or distributed when trust relationships have been established among the external trusted Identity Providers (IDPs) solutions. Policy decisions are also simplified and more efficient, as the PDP takes real-time decisions based on the defined policies and attributes. Therefore, better fine-grained permissions can be achieved and any change in the attribute data set can be populated in the configured access-rights and permissions. A typical example for this case is the delegation of access-rights and the assignment of responsibilities and tasks, when there is a need to assign the appropriate levels of delegated rights to another subject. The automated policy decision workflows reduce the administration tasks to delegate certain responsibilities and verify the access policies, which empower the subject to access the resource(s). Any further automation in the policy and SoD enforcement to align access privileges with job tasks can result in improvements to the periodic access reviews, reporting and certifications to meet the demanding audit requirements easier.

3.3.3 SOA access solutions

The most common web service (WS) architectures in use are SOAP and REST. Both SOAP and REST are used to exchange information and implement web-services with the service-specific APIs, when deploying web applications. The XML-based messages are often used for information exchange between different services. The SOAP messages can be enhanced to enforce access restrictions to web services with WS-Security. By using either the WS-Security or the SAML approach, open-standards can be utilised for authentication and authorisation controls. Moreover, open standards for federation, such as SAML and the Open standard for Authentication (OAuth), are widely supported by the identity providers and the cloud service providers to offer access to cloud applications. Concerning the authorisation needs, policy-based authorisation services can be used to define the appropriate permissions and access rights to provide the appropriate access levels to the resources. SOAP is utilised for the exchange of structured information for message negotiation and transmission. Furthermore, it is used for the setup of the trust relationship between the Service Provider (SP) and the IDPs [192]. Moreover, the policy engine provides the authentication policies, constraints and actions to realise the authentication needs for the access requests. The access requests are then validated against the configured policies for any constraints or actions. During the authentication phase, the subject needs to be authenticated before access is granted to any resources.

As SOA necessitates common terminology and semantic, syntactic, technical and legal interoperability [193], the ASPIDA model incorporates the idea of policy-based management in access control rules. In this direction, several IAM architectural challenges are emerging for the effective deployment of applications in SOA context. For instance, Nguyen et al. [194] propose a methodology for developing security schemes for REST-based (IoT) systems of any kind. Although they propose an increased adoption of REST results in the evolution of μServices, there is no analysis and inclusion of dynamic authorisation and policies, which are gaining significant attention.

SOA solutions need to incorporate efficient access control techniques and adopt the optimal setup among countless approaches in providing access control services [195]. Web Single Sign-on (SSO), federated identities, password synchronization and service granularity can be accomplished through the IAM capabilities, so that SOA is able to address and fulfil most of the contemporary access management challenges. ASPIDA is an innovative model to manage the multilevel integration of identity, authentication and authorisation modules based on formal policy-based methods and access control mechanisms in order to provide secure access to the resources.

Furthermore, parallel to the continuous developments and enhancements in technology, organisations continually explore and deploy sophisticated identity, authentication and authorisation capabilities (i.e. trust and federation, delegated identity management, password management, multi-domain access) to provide secure access to network and system resources. Until recently, in light of the increased overhead and security complexity in decision-making for user access management solutions, most of the IAM tasks were assigned to the application developers, to the network engineers and/or to the security officers, who most often had little knowledge about the service processes and needs. As a typical example, some of these tasks can be related to access requests to grant or revoke access to protected resources during any Join/Move/Leave (JML) activity in the Organisation. Due to the higher integration of SOA combined with the increased interest in cloud infrastructure solutions (i.e. SaaS) and mobility needs, the access control services need to adopt the latest IAM advancements, such as the utilisation of federated identity management to establish logical links between the identities and the services in a secure way.

Even though the Role-Based Access Control (RBAC) models have been utilised extensively as the dominant access control models until recently, they require an efficient and constructive organisational policy and a hierarchical role-based approach. RBAC lacks temporal capabilities, dynamic key management and consequently hierarchical access control models have evolved [196]. Moreover, as many of the models can handle the files but not the content, Multi-Role Based Access Control (MRBAC) for distributed multimedia systems [197] has been proposed. Recent works [6] provide techniques for creating a policy-compliant service composition through a graph, as well as the evaluation of policies during service compositions and multi-level security classes of information flows [198] for cloud-based solutions. In the current evaluation process, the necessity for the existence of a minimum policy model is depicted that can integrate various access control constraints with policy-based compliant services.

Complexity produces configuration and management overhead (i.e. repetitive tasks for the account and authorisation management), and results in higher administrative costs along with reduced usability. These are common complications for the deployment of SOA applications. Therefore, there is a need to incorporate common terminology along with semantic, syntactic, technical and legal interoperability. The integrated components should be common management services, independent of the service delivery platform and the service administration.

Chapter 4

The architecture

4.1 Outline

The purpose of this chapter is to describe the reference system architecture derived from the core entities presented in the preceding chapters. This section details the policy-based characteristics along with the event-based monitoring capabilities of ASPIDA and few scenarios are analysed to demonstrate the flows of the messages exchanged between the entities in the target architecture.

4.2 ASPIDA architecture

Even though some studies [199] propose the enforcement of security policies and access controls to secure heterogeneous M2M networks, there is an imperative need to address the service-based development, the service orchestration and the outcomes of services capabilities in order to protect data exchange (i.e. authenticated calls). The existing research works [200], [201] neither exploit the SOC capabilities, nor do they apply the enforcement of security policies across different entities. There is a need to manage efficiently the massive number of smart objects using automated policy-decision making, orchestrate the services and improve the secure data exchange in the M2M communication era. A SOA approach can provide the application of the management policy, optimises the service composition and enables increased interoperability as well as better message control and improved information flows. The ability to scale operations, develop new functions rapidly and meet different demand or capacity levels among a large population of distributed entities are also important key factors that can be realised with the combination of SOA and policy-based management methods.

The ETSI M2M functional architecture identifies the functional entities and the related reference points required to allow the communication of the M2M service capabilities by including three domains; *device*, *network* and *application*. ASPIDA has been designed to provide an additional fourth domain (4-tier architecture), the *service domain*, which concentrates the services in a loosely coupled way. Hence, the services can be reused by any of the applications for purposes of integration and realisation into an end-to-end SOA. Thus, the architecture consists of the following four domains:

- *M2M Device domain*, which includes the M2M devices
- *M2M Network domain*, which enables the communication with the core network and provides either the direct connectivity between the M2M devices, or the connectivity with the network domain via M2M gateways acting as proxies with authentication, authorisation, management and provisioning capabilities
- *M2M Application domain*, which enables the data manipulation and usage by the business-applications
- *M2M service domain*, which supports distinct service capabilities

The architecture and the governance of the security and service policies are presented in Figure 17.

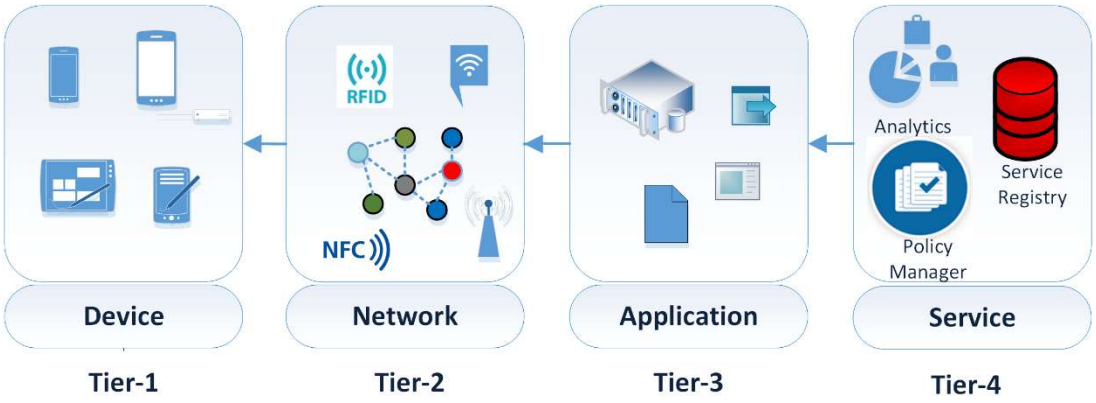


Figure 17. Four-tier ASPIDA architecture

Monitoring the levels of the unauthorised access requests, the error rate, the out-of-order delivery and other security attributes could trigger a spontaneous proper policy action in order to remedy the illegitimate activities, mitigate security issues and handle resource exhaustion and unreachable nodes. By utilising the service location registry, the service consumers, also in other clusters, can identify the PEPs in a multi-domain environment and negotiate the enforcement of the appropriate policies. For instance, if insufficient or no policies are identified in the visiting cluster, then the visitor PEP consumer can negotiate and exchange the policies with the home PEP server in the original cluster. It is vital that the service location discovery supports the policy exchange between the nodes either within the same cluster or in different clusters in order to enable the policy enforcement requested by the PEP consumer. Furthermore, in several M2M use cases, the appropriate policies and actions need to be employed aiming to avoid the needless depletion of the limited bandwidth, storage and power resources by setting a set of rules, obligations and policies.

The key functions and M2M service connection procedures from the device and the network domains in the ETSI M2M reference architecture [22] are detached to create the basis of the service domain. In more detail, the network domain entities communicate with the M2M Network SCL (Service Capabilities Layer) and access specific M2M services, and various functional components coupled with processes (i.e. the M2M SCL, the M2M management functions, the event-management) reside in the service domain to enable secure data transportation over the reference points.

Additionally, the exchange of data between the entities and the resources is supported by the respective reference points; the device application interface (dIa) between a device/gateway application and the respective device/gateway SCL, and the M2M application Interface (mIa) between a network application and the Network SCL. These services are included in the new service domain consolidating the M2M services that can be reused by the M2M applications, as depicted in Figure 18.

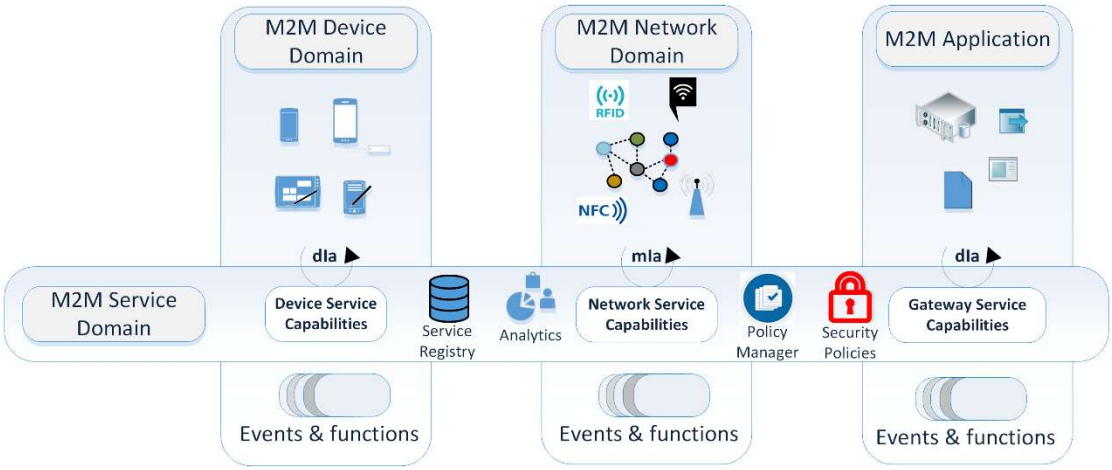


Figure 18. SOC capabilities of ASPIDA

4.2.1 Architecture domains

Device domain

The connected devices have an immense potential, but also consist of a heterogeneous collection of communication models, diverging and often competing technologies, numerous standards, restrictions due to limited hardware and software capabilities, and finally distributed trust-less environments. Moreover, the continuously increasing volume of devices and traffic requires additional resources and improvements in security and processing capabilities [202] [ETSI TR 102 691], in several ways:

- Augmented security, privacy and safety needs because of emerging threats (e.g. cyber-attacks, industrial espionage, potential destruction of national infrastructure) and potential vulnerabilities
- Larger population and usage coverage of the connected devices
- "Smarter" objects, communities and Internet of Everything[16] – IPv6 ready to support the high number of objects

The devices need to gather and/or disseminate the data, such as the health and mental status, the driving/working conditions, the location and the environmental settings/values. The devices are vulnerable, as security is not typically part of their design and they usually have very few security capabilities. For instance, the firmware and software update process may have serious security risks that could put the communication or the data in danger (i.e. insecure or lack of software updates as the devices are difficult to update, the devices are rarely updated, leaving the devices open to be updated with new software results in high risks of malicious activities). Since the smart interconnected objects (i.e. sensors and RFID tags) are often unattended, severe security consequences may be suffered (i.e. capturing and compromising the objects) that may have a range of impacts on the service levels. Finally, monitoring, controlling and managing the devices is not always feasible in devices that further deteriorates by the limited security capabilities.

[16] The Internet of Everything (IoE), http://www.cisco.com/web/about/ac79/docs/innov/IoE-Value-Index_External.pdf.

Network domain

Multiple factors can affect the security of the network and the communication between the devices and the applications, as presented by Pathan [203]. Certain policies may be required to control the data at specified intervals or dynamically allot network resources to support an increasing number of IoT devices. Moreover, the network needs to cover the data transmission requirements despite the finite network resources. The transmitting ranges should be defined for optimal network setup with the purpose of prevailing capacity, performance and scalability problems. The sensor nodes are equipped with sensing capabilities and the sensor networks are mostly focused on sensing environmental data (i.e. patient's state of health) and conditions (i.e. monitoring a room to keep track of environmental such as temperature and humidity) in real-time.

The ad-hoc networks are widely used in emergency situations due to human-activities (accidental, on-purpose) and natural disasters. Hence, the ad-hoc nodes can establish links among the participants such as multimedia sharing for mobile users in a localized network and in short-range communications. In the case of sensor nodes, the data is being transferred to the computational centre via the sink node, while in ad-hoc based scenarios the data can be shared on peer to peer applications (i.e. PDAs/tablets connected in a single-hop ad-hoc network during a healthcare conference providing increased mobility to the participants and sharing data among themselves). To deal with the IoT requirements, such as latency, energy efficiency and mobility, fog computing has been presented as a better alternative for IoT compared to cloud-computing [204]. The solution can be amplified with service health checkers and off-loading with load-balancers in order to enable the seamless operational controls and services. Among various network-related security concerns, the wireless medium is highly prone to security threats and only a few gateways can support encrypted communication channels.

Application domain

The application domain enables the data manipulation and usage by the business applications. Given the wide range of these applications in various environments (i.e. healthcare, industrial, home automation), various techniques can be used. In principle, the software used in sensor networks is simpler with fewer hardware requirements, but the sensory data and applications are evolving rapidly. Among several areas of sensor network applications and monitoring, the healthcare can be cited (i.e. wearable devices specially adapted to be attached to the surroundings of the users or worn on the body surface), as well as fitness, area and region, air-pollution (i.e. concentration of malicious gases), military applications such as battlefield surveillance, industrial (i.e. vehicular, machine-health and preventive maintenance, data-logging, data-centre), supply-chain and management, facility management (i.e. control of leakages in nuclear plants, intrusion detection with camera-equipped sensors), disaster relief operations, landslide detection, forest fire detection, water quality, waste, natural disaster prevention, biodiversity and wildlife observations, intelligent buildings (i.e. ventilation, air-conditioning control), telematics, transportation and traffic management, weather forecast (i.e. ocean temperature monitoring), avalanche detection (i.e. sensors with location-based capabilities) and others. Concerning the ad-hoc networking applications, typical examples are the military cases, which allow the maintenance of an updated status about the equipment and the exchange of real-time information between the soldiers and the headquarters.

The heterogeneous devices connect to each other for data exchange and information sharing. The network should facilitate the service discovery and registry on the given constraints of the infrastructure such as on limited resources, mobility and dynamic context [205]. Although conventional approaches utilise the proper middleware technologies to ease the growth of advanced sensor services like the use of virtual sensors in Smart Environments [206], others [207] propose a SOA-based approach for the wireless networks without the need of a middleware advocated by standardised and open solutions. Nevertheless, various security and governance issues arise to ease the administration of the applications in integrated or cross-domain environments.

Service domain

The proposed service domain removes procedures handled previously by other domains aiming to increase abstraction, reduce the computational load of the first three domains and enforce better security controls for the M2M applications. This results in higher availability, re-usability and the capability to combine the offered web services irrespectively of the underlying technologies (i.e. Java, .NET) as well as it achieves a greater level of flexibility and independence between the infrastructure and the interfaces in use.

There is a clear evolution in the enterprise architectures, which raises the need for better orchestration of the integrated services to ensure efficient, reliable and secure data exchange between multiple devices and the applications. In this direction, ASPIDA is an integrated architecture reinforcing policy-based controls and accomplishes the realisation of a higher scalability due to the inherent necessity to represent diverse types of context information within multiple services. This modular architecture focuses on a strong decoupling of the processes and simplification improvements by providing higher level abstractions, and alleviates security threats on event-based M2M communications.

4.2.2 SOC capabilities

The service domain in the policy-based ASPIDA architecture incorporates SOC capabilities (i.e. service registry, service requester, service provider, service analytics) and security functions such as the generation of keys, generation and validation of certificates, validation of signatures and other security measures to fulfil the service challenges, design constraints, security requirements and policy enforcement goals. The analysis and the development of the appropriate rules are incorporated in the service domain. New policy servers can be discovered automatically with a service location discovery functionality.

This novel domain employs SOC capabilities in order to enable the service and security mechanisms with the appropriate policy-decision and policy enforcement engine. The service registry improves the performance and efficiency of the model by offloading computationally demanding tasks of the service capabilities to the service domain, which provides the respective policies for the device, network and application domains. The web services registry includes the web-service details, which is a service broker for both clients and providers. These services can also be automatically detected and offered by the service discovery component. In such cases, the services are accompanied by the necessary meta data to be successfully registered and located. With the convergence of data, the service analytics captures the data and offers an improved extended insight of the services. ASPIDA collects the environmental information and data by incorporating such a service analytics component. The service analytics component enhances the event management with the proper alerts and data analysis, and extracts the data elements from the service components. Consequently, the

service reports (i.e. related to performance, quality) can be produced and potential issues can be identified with the M2M resources or configuration. The proper reference interconnection points need to accommodate the related interactions (i.e. registration, request to access, event and configuration management).

The service domain capabilities can simplify the dynamic service redundancy and the sufficient capacity of the policy-based enforcement system due to the agile nature of the interconnected objects (i.e. nodes moving from cluster to cluster). Using a hierarchical control of the policy-based components, the control can be transferred and delegated seamlessly from the service to other infrastructure components. Simultaneously, service discovery functionalities need to manage efficiently and effectively the enrolment requests regardless of the number of services and devices. Other SOA-based solutions can enhance the performance also improve the security and the operations, such as data analytics handling with a service analytics component for event management in order to produce the proper alerts and service reports (i.e. related to performance, quality). Hence, the service domain employs an integrated policy-based service along with security management capabilities including interactions with the nodes in the device domain, the access and core network in the network domain, the data and various applications in the application domain. The policy manager orchestrates the security policy management and optimizes the automated interactions with the other components. Figure 19 depicts the building blocks and the interactions concerning the message exchange and the communication flows.

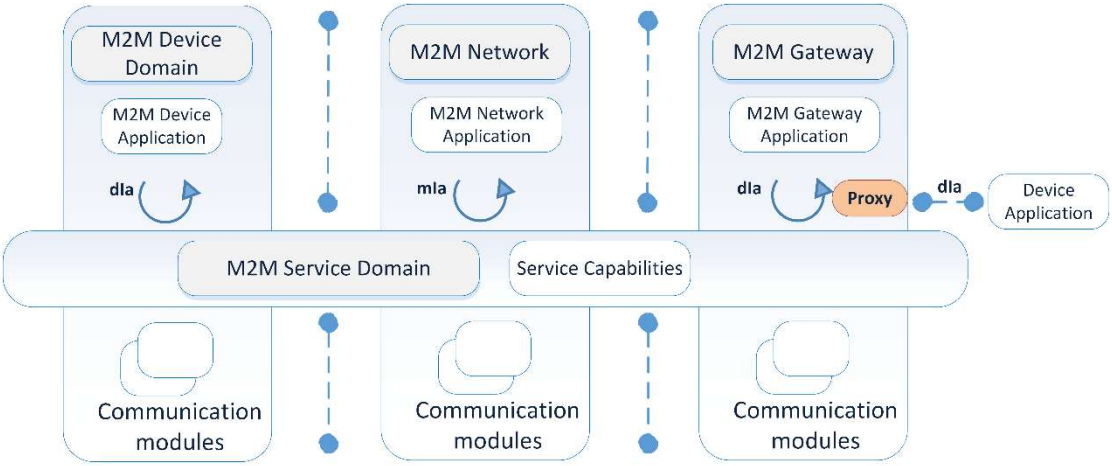


Figure 19. Interactions & service capabilities of M2M Service Domain

The SCL of the M2M service domain runs across all the other M2M domains and exposes the functionalities (service registry, management objects, etc.). The mIa is used for the interactions with the M2M network application and the network SCL, whereas the dIa is used for the interactions with either the device application or the gateway application.

4.2.3 Policy enforcement in the service domain

In this section, the analysis of the security policy enforcement of ASPIDA is presented. New services can be established to utilise the existing code for re-usable services. The policy engine needs to allocate and manage the resources in an efficient way, and mitigate any security risks such as vulnerabilities, nefarious activities, abuse, eavesdropping, interception, hijacking, loss of information, attacks and media failures. It is expected, though, that a SOA approach can simplify the administrative management and the overhead by establishing appropriate security policies to monitor the events and conditions, the interconnected objects and the

services. For instance, in the case of an ASN in which the devices are resource-constrained in terms of computation and power capabilities, the policy engine needs to allocate the resources in a more efficient way during the deployment of the relevant policies in order to ensure the operations of the specific target in question.

In the context of SOA, the service provider initially publishes the services in the registry (i.e. an XML-based registry to describe, publish and utilise the services). New services are built to establish, or leverage existing code for re-usable services. The service registry maintains a set of published services with their associated service properties with the purpose of facilitating any subsequent queries to the service registry originated by the service requester. When the client queries the services, the relevant registered service provides the requested information. The requester searches for a specific service in the registry and in the case of an existing and valid service forwards the request to PEP to intercept the request into the appropriate format (i.e. using a SOAP transport message). Then, PEP invokes and binds the methods of the requested service to the service provider that provides back to PEP the outcome of the service request. Finally, PEP acts on the received policy decision and translates the updates into the appropriate format for usage, which is then communicated to the publisher component (i.e. notice-board, APIs, text-messages and notifications). Upon receiving a request by the service-provider, PDP evaluates the service request based on the managed policies already defined by the policy manager in PAP. PAP feeds the policy store repository with the policies. The relevant policies to the specific service request are retrieved and consulted by PDP to reach a decision. If there is no blocking or negative policy decision, PDP retrieves the necessary additional service data from PIP, which offers the interface to collect the appropriate service data related to the service request. The policy decision, supported by the service attributes, is forwarded to the service provider to resolve the request. The detailed structure of the policy enforcement methods and of the corresponding abstract interactions is shown in Figure 20. Figure 21 illustrates the interactions of the entities by using a UML sequence diagram that depicts the detailed messages exchange in a time-order manner. Upon receiving a request by the service provider, PDP evaluates the service request based on the managed policies already defined by the policy manager.

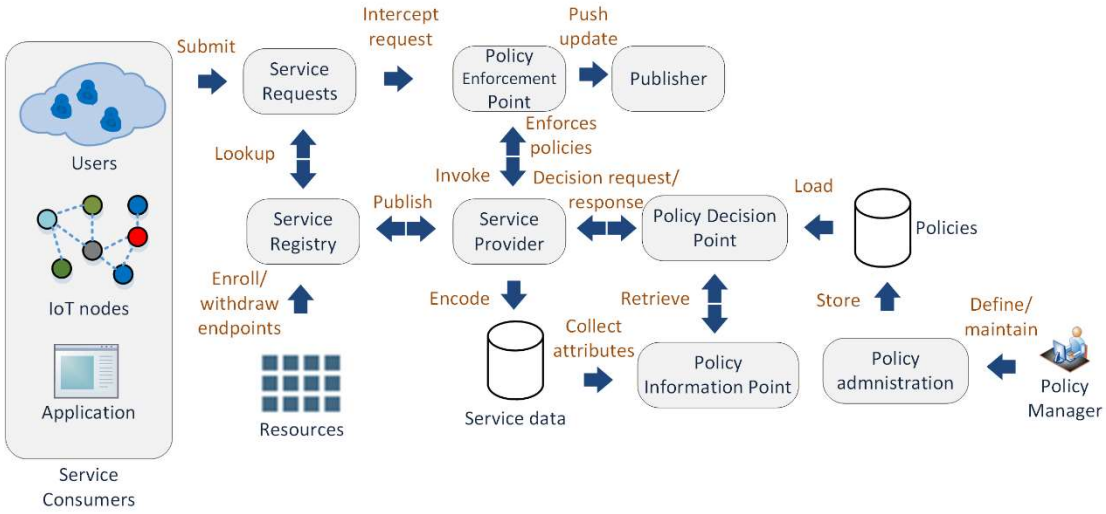


Figure 20. Policy enforcement in the service domain

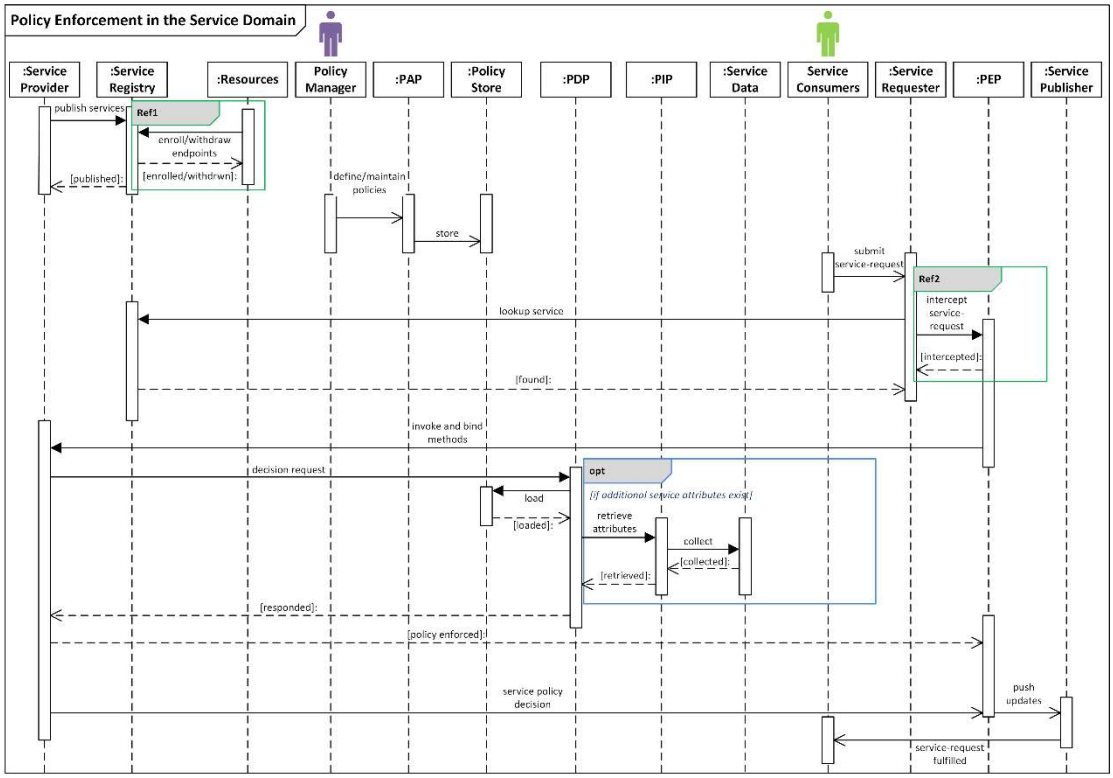


Figure 21. Entities' interactions for policy enforcement in the service domain

Next, PAP feeds the policy store repository with the policies. The relevant policies to the specific service request are retrieved and consulted by PDP to reach a decision. Provided that there is no blocking or negative policy decision, then PDP retrieves the necessary additional service data from PIP, which offers the appropriate service data related to the service request.

4.3 Summary

Considering the recent rapid advances and the respective security challenges, ASPIDA introduces a policy based management model to facilitate the enforcement of the security policies to comply with the security objectives. One of the major challenges for M2M communications is to support a large device volume with various heterogeneous characteristics and different mobility profiles by meeting the resource constraints requirements. For instance, higher energy efficiency and reliability are among others the main design objectives for the M2M devices. These requirements result in compact physical size and lower power consumption in a resource constrained environment. Hence, these should also be considered in the design of low latency, high resilience and high reliability services considering the critical nature of several M2M applications (e.g. in healthcare, critical infrastructures and transportation). Additionally, by using the proper SLA related metrics, the operations and performance efficiency can be controlled and further analysed. An SLA-driven adaptive model is presented in [122] to depict how to employ policy based decisions.

In the context of distributed, multi-domain and often unattended M2M networks, the multi-level policy approach can provide improvements in policy composition, enforcement, validation and conflict resolution. Nevertheless, various complicated issues may arise such as the semantic interoperability, the definition of concrete policies, the multi-domain policy consistency, the policy refinement and the completeness.

Chapter 5

Security management

5.1 Outline

From a distinct perspective, ASPIDA can also address the access control challenges of the M2M communications in the SOA policy based management processes. In this chapter, the security and access control characteristics of ASPIDA are analysed. More specifically, the identity, the authentication, the user roles, the authorisation access control levels and the rule validation mechanisms along with the integration of the respective modules with the proposed architecture are presented.

5.2 Incentive

In the last few years, several applications (e.g. transportation, human/inventory tracking, water/energy distribution and quality monitoring, personal area and home and habitat networking and monitoring, data centre monitoring, disaster avoidance and recovery, military surveillance and industry operations, medical/healthcare monitoring, process monitoring and smart spaces) rely on the capabilities of the M2M communications. These applications collect the environmental conditions along with the physical information from the M2M measurement nodes in order to process, analyse and present the measurement data. Several security and service quality issues arise due to the:

- high number of interconnected heterogeneous smart objects
- use of various networking technologies for the M2M communications
- distributed nature of the smart applications into consideration

In respect to the network communication protocols, the M2M environments need to support secure cross domain information exchanges among several smart interconnected nodes. Because of the distributed and dynamic nature of the environment, the M2M nodes are free to move, trigger a high rate of physical topology changes and operate in an unattended fashion with very limited maintenance support. These fundamental M2M characteristics complicate the operations of the routing and management protocols, increase the security risks and raise a greater potential for security issues [1]. Hence, it is important to integrate the IAM methods with the policy based SOA management and the service management disciplines, when accessing identity data and supporting identity life cycle processes.

5.3 Security challenges

The traditional security mechanisms are not always applicable for M2M implementations making the M2M security management extremely difficult. For instance, the M2M wireless communications inherit additional security threats and operational constraints compared to their wired counterparts, albeit additional and exceptional characteristics. Taking into account the M2M characteristics, there is an imperative need to meet the security requirements and to address the appropriate access control services on the identification and privacy protection technologies. Among others, the distributed nature of the pervasive device domain, the resource constrained devices, the interconnection requirements of the network domain (i.e. connectivity with other networks, roaming, control functions, scalability and resiliency) and

the various M2M traffic patterns in diverse application domains play a significant role in M2M communications.

Although the M2M networks may not be accessible from the public networks, there are numerous security threats as the M2M networks need to be integrated with other networks for business development, maintenance and troubleshooting purposes. Since the classic security mechanisms consume significant amount of energy and the connected devices have limited cryptography and access control capabilities, the current security models are not always a suitable option for the resource constrained M2M environments such as Virtual Private Network (VPN) and Transportation Layer Security (TLS) for the network and transport layers. Hence, new security controls and light weight mechanisms need to evolve (i.e. hash and XOR computations, password hash values and mutual authentication by using pre-shared keys). The security solutions for the endpoints should be used in accordance with the engaged protocols and the security management solutions aiming to achieve secure communication and manage efficiently threats, vulnerabilities and breaches. In order to mitigate the security threats in M2M communications, various methods have evolved like intrusion detection systems and anomaly detection technologies for detecting zero-day attacks. In the cases where immediate detection is required (i.e. critical infrastructure), apart from the conventional cybersecurity countermeasures, some solutions are based on the high sensitivity detection of behaviours of information networks in order to detect suspicious activities. In this vein, Tanaka et al. [209] have developed a layered architecture with relay devices between the sensors and the cloud. These devices collect the massive amount of sensory data and detect likely sensor problems.

In order to achieve an effective use of the M2M applications and improved service levels, the service needs combined with the security requirements should be identified properly and the available constrained resources should be managed efficiently [210]. In this vein, the policy-based management allows the creation of the necessary security policy expressions and afterwards their deployment. The benefits of this approach arise as the M2M communications develop, the criticality of the service and the information availability increases and the resources become more complex. Hence, a principal factor is that these resources can be accessed and become available to the interconnected M2M components using interoperable services in a secure manner. The existence of a service orchestration is crucial to improve the resources management, whereas SOA allows the reusability of the web services contrary to other architectural models. Considering the numerous security threats, the rapidly increasing traffic volume, the computational limitations of the connected objects and the distributed nature of the objects, the reference architecture models need to focus on the adequacy of in-place security controls and the performance indicators. The appropriate security rules need to be enforced in order to mitigate the security risks/threats and ensure the appropriate levels of runtime security maturity.

Furthermore, special attention needs to be paid to develop and select the appropriate M2M device management and communication protocols, as the M2M devices may work under severe resource constraints and have limited processing capabilities. There is a need for appropriate data and messaging protocols to achieve the reliable and secure communication between the M2M devices and the applications. As mentioned in Chapter 2, various lightweight open standard protocols have been proposed for use in real-time communication, such as the Advanced Message Queuing Protocol (AMQP), the Extensible Messaging and Presence Protocol (XMPP), the OMA LWM2M and the MQTT protocols

enabling the communication and message exchange, where a small code footprint is required and/or network bandwidth is at a premium. Some of the M2M transport protocols like MQTT, CoAP, XMPP and AMQP are not suitable for real-time applications with high sampling rates due to the high latency. These protocols lack open-source libraries for constrained devices and do not provide automatic discovery. [211] presents measurement frameworks and metrics for resilient networks, runtime reliability, automated recovery of the services and how to extend the use of metrics for policies and procedures. Thus, the creation of certain condition expressions enables the realisation of the benefits of the policy-based management approach, which arises as the M2M communications evolves and the resources become more complex. Others [212] developed a tool that provides a benchmarking analysis of the most relevant M2M transport protocols showing that many of those protocols are not suitable for real-time applications with high sampling rate due to the high latency for eHealth solutions [213].

With respect to the end-to-end security and resiliency, various challenges are spread across the domains and trust boundaries. For instance, several application level interactions take place between the devices regarding the data workflows and the message exchanges. Therefore, it is crucial to identify the entities to be granted access to the resources across multiple areas and to standardise the access requests, so that M2M data can be distributed, searched and retrieved by authorised clients. The devices and the clients need to be properly identified, while logical links must be established between the identities and the services in a secure way. Considering the higher integration needs with SOA, the growing interest in cloud infrastructure solutions (i.e. SaaS) and the increasing focus on increased mobility, the access control services must utilise the proper extensions to carry identity data as dictated by the authentication and authorisation protocol. For instance, XACML can be utilised to define, evaluate and process the access control and context-aware policies. The system needs to secure the data exchange in the information retrieval, while ensuring a prominent level of assurance and preventing any ambiguous policies. All these can be achieved by using standard policy-based management methods.

Given the widespread distribution of the devices, the solutions should facilitate a sustainable, scalable and secure management of the devices and the applications should allow the extension of the controlled objects to meet the increased demands for more condensed device density or to provide wider area coverage. Moreover, the solution should be topology-agnostic and resilient by design for fast recovery. The systems should be designed to account a fast deployment. New challenges emerge in terms of interoperability due to the heterogeneity and diversity of the technology in use, the fragmented and proprietary environments, integration inflexibility, the lack of management control of the full list of devices, the ownership complexity and vendor lock-in.

Intended for securing the M2M communication, the Datagram Transport Layer Security (DTLS) [214] can be used to provide channel security and TLS support for COAP. Keoh et al. [215] demonstrate a review of the communication security solutions for IoT using public-key in DTLS negotiating DTLS extensions such as the fragment length negotiation extension to reduce the amount of fragmentation and fine-tuning DTLS functionality. Additionally, in the cases where a pre-compiled list of identifiers is used, the M2M data collection server keeping all the identifiers can associate the endpoint with additional information for the originator and instantiate the relevant DTLS sessions for access control needs.

Focusing on the potential service inconsistencies and the required policies across the heterogeneous smart M2M nodes and clusters, there is a growing need to manage efficiently the software-related issues of the M2M communication components and remedy any existing security weaknesses. Typical examples are the non-invasive attacks on cyber-physical systems, which pose a considerable threat to sensitive devices by potentially impairing them. In [216], the authors describe the development of a prototype to allow an Anti-Lock Braking Systems (ABS) spoofer to initiate such attacks and evaluate the potential consequences to these attacks. Among others, there are serious concerns regarding the corruption of the sensor measurements in industrial plants [217], such as the case of the Stuxnet malware over specific SCADA components (i.e. Programmable Logic Controllers – PLC).

5.4 M2M authorisations

M2M authorisation services play an increasingly vital role in the development of the current connected applications and have attracted considerable interest. Considering the M2M constraints and the distributed nature of the devices, the M2M devices and the communication modules should enable the transmission of context and/or authenticated information to the application servers. Additional services can be supported such as data collection, remote control operations, real-time monitoring, malfunction detection and reporting functionalities. Highly demanding scenarios can be revolutionized when realised with the use of policy-based management for M2M devices in the case of critical applications and dynamic environments (e.g. changing conditions, user/object attributes) [218]. For instance, in the case of Intelligent Transportation Services and in the need of coordinating the demand-response of the vehicle charging stations information, the distribution network needs to provide responsive, reliable, low latency and real-time communication services with localized control [219]. The entities can be authenticated according to their attributes (e.g. electric vehicles, natural gas vehicles) and the changing environmental conditions (e.g. location-based services, localization services, service availability, traffic models, queuing mechanisms), which can affect the access policy decision making. The collected data can be utilised to manage these devices, create alarms and notifications, trigger the required actions and optimize the service efficiency [76] [ETSI TS 102 689].

The unique characteristics of the M2M communications require a new set of approaches to address the technical and security challenges that arise. The M2M devices run on batteries and have limited storage and computational capabilities [220], which influences the choice of the appropriate authorisation solution. For instance, the candidate solutions may not support a full HTTP stack or a device discovery/registration process, a lack which impacts the selection and enforcement of an effective security model and the resulting veracity. Hence, access delegation must be enforced in a controlled manner in order not to jeopardize the security of the system, such as the integration of the access control scheme with controlled access delegation capabilities. Moreover, the M2M ecosystem needs to support security services like validation, authentication and authorisation between the originators and the applications.

Recent research and industry activities focus on the M2M dynamic authorisations with policy-based capabilities in pursuit of a systematic empirical research to address an extensible scheme approach to support the secure message-exchange, the dynamic authorisations for µServices and the access control requirements. Preuveneers et al. [221] demonstrate the enforcement of the authorisation policies based on the Universal Scalability Law to their throughput measurements. The efficiency curve demonstrates the latency overhead of

delegating the policy to another one along with the performance trade-offs of policy delegation. The delegated policy evaluation leverages feature toggles that are managed at runtime by software circuit breakers in order to secure the distributed data processing workflows. They propose the development of the µServices framework using the Spring Cloud, the Netflix Hystrix library and ForgeRock's OpenAM deployed in a docker container. However, there is no mechanism for controlling and evaluating the access delegations given the complexity of the authorisation policies and the computational characteristics of the solution.

Aiming to support the policy-based capabilities and accommodate the dynamic nature by collecting the event-driven updates, proposed architecture in [222] introduces additional SOC blocks in the typical architecture of ETSI [22] [ETSI TR 102 690]. The architecture also facilitates a series of security enhancements for M2M computing and dynamic authorisations. Through the security policy enforcement, the architecture addresses a SOA based on security policies, which alleviates security threats on event-based M2M communications. Addressing and managing efficiently the security for each domain of the proposed architecture with a policy engine and incorporating the M2M service requirements, the security service levels of M2M communications can be improved. The architecture incorporates a standardised horizontal M2M service domain with the required capabilities, facilitates the evolution of the appropriate security policies and ensures the interoperability between the device, the network, the application and the service domain. The attribute-based policy evaluation enables the rapid delivery of changes and increased technology flexibility. Finally, the model is also applicable for cloud-based scenarios and for capability-based exchanges to gain access to the resources, as presented in the following sections.

5.4.1 Policy-driven authorisation management

The most usual examples of changing operational conditions can be found in the cases of healthcare monitoring, military activities and extreme event responses (i.e. hurricanes, storms, floods, droughts, forest fires, earthquakes, volcanic eruptions, landslides). For example, a common scenario includes the protection of critical infrastructures, the critical military/civil surveillance applications used for forest fire monitoring and building security by unattended and resource constrained-devices often located in areas with low or poor network coverage. This requires a secure framework to accommodate the dynamic authorisations and the technological innovations in X-computing. The convergence of these approaches is also vital toward a standardised architecture enabling the inclusion of the technological changes and towards the facilitation of the simplification besides the explosion of new opportunities. For instance, in general the remote devices used in surveillance are assumed to be restricted in resources (in terms of storage, processing) and with limited computational capabilities.

In [223], an innovative policy-driven authorisation (PDA) for µServices scheme is proposed to enable valid access token exchanges with respect to the enhancement of the overall security (authenticity, confidentiality, integrity, privacy). In terms of policy management, the policy enablers (PEP, PDP and PIP) reside centrally on independent µServices to avoid consuming computational resources from the constrained devices. Event-based changes enable the enforcement of the updated and applicable policies. They also empower PEP to publish the corresponding effective policy decision. New protected resources can be registered and added spontaneously at runtime on an as-needed basis such as the User-Managed Access (UMA) protocol [224] where the resource owner can control the protected resource access and the authorisations. Nevertheless, the approach raises various implications regarding the architecture and the capacity management, because of the

introduction of new additional resource servers required to manage the real-time changes. The PDA scheme accommodates the dynamic nature by collecting the event-driven updates, which present changes in the status of the entities and the conditions. Event-based changes enable the enforcement of the updated and applicable policies and empower PEP to publish the corresponding effective policy decision. The event management capabilities address the dynamic authorisation challenges that cannot be performed with static access control policies. In addition, the M2M solution provider needs to control and implement the respective security functionalities by ensuring the efficient communication of the micro-segmented services.

The authorisation provided to access a resource relies on various checks (i.e. platform, reputation, risk, payment-based) and can be either static or dynamic. Dynamic authorisations allow a formerly restricted M2M device to access and perform newer operations on resource(s) hosted at the M2M Gateway [225]. A service layer dynamic authorisation function ensures the policy provisioning and configuration along with the policy evaluation and enforcement to process and resolve the access requests. Other research activities [226] describe the implementation and the design of the capability token-based system for secure access control in M2M and, more specifically, in Internet-of-Things (IoT) environments to improve the process time of access control by using capability tokens. Conclusively, there is a necessity to develop dominant security controls required by µServices to handle the advanced M2M communication requirements with policy-based management techniques. In this context, S. Cirani et al. [227] propose an architecture targeting HTTP/CoAP services to provide an authorisation framework, which can be integrated by invoking an external OAuth-based authorisation service (OAS).

5.4.2 Capability-based access control

In order to provide a very high level of security assurance, advanced data protection and increased security, the applications can generate a capability (i.e. an access token) that provides access to the protected service(s)/resource(s). Multiple methods can be utilised to generate a capability and, thus, to increase the application security protection. The secured claims can be expressed by Concise Binary Object Representation (CBOR) [228] and utilise CBOR Web Tokens (CWT), which are derived from a JSON Web-Token (JWT) [229] for secure and stateless authorisations. This approach enables the access control systems to manage the creation, the delegation and revocation of the tokens. CWTs can be used as authentication credentials to control access for IoT systems that use low power technologies. In [230], Gusmerolia et al. describe the Capability-Based Access Control system (CapBAC) that utilises PDP to process an access request. The system retrieves the relevant access control policies from the policy repository and allows PEP to delegate/enforce the access decisions. The distributed nature of the IoT solutions often requires the certificate management, the authentication and the authorisation processes to be independent. The distributed management of the authentication requests without a centralized control entity is demonstrated in [231]. Other proposals [232] support the authorisation decisions based on local conditions offering Capability-based Context-Aware Access Control (CCAAC) where the contextual information is of very high importance for critical operations networks.

5.5 Securing with ASPIDA

Motivated by the above, there is a need to address the hinderances to grant or deny access to the resources, uncover the open issues with policy-driven decisions to enforce the proper

rules and orchestrate the key-components to manage efficiently the access control services. The access control services are the essential building blocks in realising SOA. The authentication and authorisation of network access, the access control mechanisms, the data confidentiality and integrity and the smart M2M communication service availability can be imposed with policy-based methods ensuring the secure delivery of the smart M2M services. In [233], a consolidated model is presented to integrate these components based on policy-based access controls that can be enforced to ensure the conformity of policy decisions. ASPIDA is an inclusive PBM access control architecture, which supports the policy-based and integration capabilities. The architecture offers automated maintenance of the policy sets and controls, improved efficiency, simplified management and support of several types of environment (i.e. enterprise, service provider). ASPIDA introduces the integration and deployment of the proper security policies in the M2M communications services to strengthen and improve a robust security management. The architecture also manages the multilevel integration of identity, authentication and authorisation modules based on formal policy-based methods. Access control mechanisms can be provided for secure access to the resources.

5.5.1 Integrated access control model

As there is an increasing interest in advanced and secure access control services with a delegated authorisation policy to mitigate security threats and improve compliance, ASPIDA consolidates and integrates the authentication, the roles and the authorisation policies providing a holistic policy-based access control model. The model incorporates an *Identity Service Engine* (ISE) module, an *Authentication Service Engine* module (AuthNSE), a *Role Mapping Engine* (RME) and an *Authorisation Engine* module (AuthZSE).

ISE stores the identity data and provides the necessary information for the subject to be authenticated and authorised. The subject can be a user or a group of users and the information needs to be accessible by other entities, such as by other users, resources and objects. Upon receiving a request from an entity, ISE examines the identity data against a predefined set of identity attributes of the subject, the profile types and the resources to be accessed. The profiles associate the users and the groups with a set of data that describe their characteristics. For instance, in an attribute-based access control environment, the access controls facilitate the initiation and the efficient termination of a user access request. Identity management is very challenging as it may associate hardware devices (i.e. Secure Signature Creation Devices – SSCDs), software tokens and security/network equipment with the subject to be uniquely identified. Additionally, the maintenance of identity information in multiple identity systems increases the risks of identity data exposure. In several cases, the identity management solutions support not only the standard identity management capabilities (i.e. provisioning, de-provisioning, recovery, updates, suspending and unsuspending), but also additional functionalities, such as Single-Sign-On (SSO) and Single-Logout (SLO) across providers in a circle of trust. ISE also manages identity policies in the service domain and facilitates the identity data flows required for authentication, authorisation and accounting operations. In the case of multi-domain and distributed environments or trusted multiple systems, there is a need to allow access to objects and resources outside the local domain of control. Thus, it is necessary to formulate, standardise and regulate not only cross-domain access and authorisation controls, but also the attributes of the subject, the identity policies and the data manipulation for a seamless cross-domain attribute exchange. As the identity information can be distributed and replicated across multiple identity systems, there is a risk

that the data may end-up out-of-sync and inconsistent. The role of ISE is crucial, as there is a continuous need to cope with security threats, such as data loss, data leakage and unauthorised access to protected resources. Moreover, several risks in identity and account management emerge related to the legal ownership of data complications, the lack of interoperability among the systems, the existence of multiple proprietary protocols and the complicated application of specific identity management work streams. Therefore, it is imperative to identify the risk event types and mitigate them by enforcing a strong authentication framework against stolen identities, leaked credentials, or infected devices.

AuthNSE is used for authentication purposes that may provide various API interactions with the applications. AuthNSE receives the access requests from ISE and realises the configured authentication policies including the policy subjects, constraints and actions. The access request is subsequently redirected to IDP for validating the credentials. This is achieved with the use of secure communication protocols along with secure directory information sharing. The solution can also be enhanced with strong and multi-factor authentication options. Special restrictions may be applicable (i.e. external third-party or cloud service IDP authentication are not allowed) in order to conform to the exceptional and specific regulatory requirements. Additional privacy and security constraints (i.e. mandatory protected transport-layer connections, data integrity, auditability for fine-grained user access, separation of user management authentication with application access) also affect the authentication service implementation and integration. At a minimum, the providers need to offer user authentication and authorisation enforcement. In ASPIDA, ASE verifies the subject's identities (i.e. username/password, digital certificates, biometrics) against the identity data which are kept in the identity data stores (i.e. in the Active Directory Domain Services - AD DS) via the Lightweight Directory Access Protocol (LDAP) message exchange. The identity data repositories keep the information about the users and the application permissions, while LDAP is used for the queries and account synchronization between the IDP and the AD DS. The original request is initiated by the client to the SP in order to gain access to the web components or to the cloud computing resources. In Figure 22, the client challenges the IDP and the Security Token Service (STS), which is provided through the use of the appropriate APIs. STS issues and provides the security tokens in order to establish a trusted relationship between SP and STS.

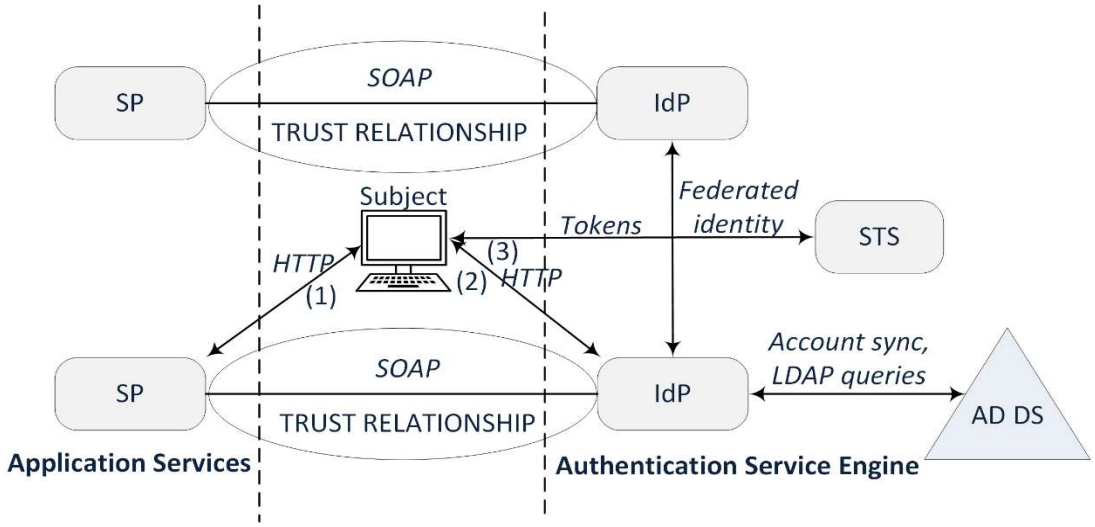


Figure 22. Authentication Service Engine

SOAP is utilized for the exchange of structured information for message negotiation and for transmission. It can also be used for the setup of the trust relationship between the SP and the IDPs. The engine provides the authentication policies, constraints and actions in order to realise the authentication needs for the access requests. The access requests are then validated against the configured policies for any constraints or actions. During the authentication phase, the subject must be authenticated before access is granted to any resources.

In order to define the roles that provide the appropriate access rights and permissions, RME defines a set of permissions, roles and policy criteria. The RME also maintains the role memberships to achieve consistent information protection, while the proper controls ensure the appropriate access levels. RME maintains the role policy and determines the entitlements and the role memberships. The policy exchange messages between the policy makers and the role managers can be formulated by using a standardised policy language. In this case, the policies need to address the association of roles with the resources, the hierarchy and the management of resources with the access control management. The policy engine is integrated with RME and the role manager of RME controls the assigned roles to the subjects based on specific attributes (i.e. organisation user manager) and conditions (i.e. expiry date). For instance, RME determines the contractor's temporary data analysis access to a system component for a limited time and restricts access to another resource. In this scenario, the policy defines the role and the combined permissions of the subject, the authorised sessions to access the resources and the access control conditions in the enterprise or in the organisational policy. The RME determines whether the role policy is evaluated and applied according to the role memberships, the entitlements and the decisions in each context. Various other complications can be encountered during the definition of the policies in terms of the semantic and syntactic consistency, as well. Thus, the validity of the policies should be verified to ensure the policy enforcement and the conformance with the policy constraints

AuthZSE defines the authorisation decisions, the resource authorisations, the authorisation hierarchy delegation, the secure access management actions and the changing access rights without modifying roles. Upon receiving a request, AuthZSE applies the appropriate access controls, which incorporate the policy subjects, the resources, the actions and the authorisations. The authorisation policy sets define the actions and the privileges of an entity on the resources. This brings forth numerous security and operational issues related to the delegation procedures, the secure access management of shared resources, data confidentiality, system dependencies, integration limitations with SOA and many other obstacles. The goal is to provide access to the target resource or service only to authorised entities that undertake specific duties. The access control policy management models may have different policy support capabilities and administration mechanisms. The access control service should define the application-level operations with the access policy rules based on the access criteria (i.e. sensitivity classification) and the infrastructure constraints (i.e. heterogeneity, cloud computing platforms, technology limitations), while the operational constraints are governed by the access admission policy decisions rules. AuthZSE utilizes various mechanisms and coordinates with different components to prevent unauthorised access. In principle, the SOA context should be decoupled from the authorisation policy and the associated decisions. The authorisation policy should be defined in the XACML policy document and retrieved on-demand during an authorisation request. The authorisations on protected resources can be granted to the entity and revoked either statically or dynamically. The appropriate role patterns need to be defined in the security authorisation restrictions,

although in the case of many policies for various providers there might be performance and scalability issues. Then, the access policy engine maintains the specified roles and uses a role repository to store the mapping between the roles and the permissions for a predefined period and, optionally, under certain conditions or special restrictions (i.e. exceptional access-rights or restrictions to access specific objects and resources). The role memberships are also deployed and maintained according to the enterprise administration along with the resource policies.

The respective policies from multiple sources along with the SoD definitions can be authorised by the policy rule validator. Special attention should be paid to the existence of a large number of policies from multiple sources, as it is likely that there are difficulties in deploying and implementing the appropriate policy decisions. In order to prevent any contradicting requirements, detect access rights conflicts and resolve any policy violations, ASPIDA introduces a policy rule validator. The validator can detect non-conformities and violations to ensure policy consistency and checks the data quality of the configured policies for missing/invalid data information. Any conflict and violation determined under the review process can be resolved by the policy rule validator utilising a set of service policies and hierarchical resources to detect and manage the policy violations. The policy rule validator monitors the system and assesses the privileges, the roles and the resources, and reports any deficiencies besides violations. The access control remediation mechanism resolves any conflicts or violations because of misconfiguration and not because of applicable policies, fraudulent or negligent activities. When a policy violation or non-conformity occurs, a policy review process is triggered for the conditions and the policy data sets [234][233]. Vance et al. [235] manage to reduce access policy violations through user-interface artefacts. During the review process, the policy checker validates the updated configuration, whether the violation has been resolved, so that the appropriate policy remediation is applied. The policy checker reviews ensure the existence of the appropriate level of access controls to prevent from malicious, careless or exploited threats to eliminate the disclosure of data and unrestricted access to protected resources. Apart from the policy checker, the APE should monitor for unattended or unauthorised usage of the resources. The system introduces a remediation indicator to provide a better suggestion, whether this is a real conflict or not. If the problem is real, then the validator is expected to provide a high value to the indicator. Otherwise, the indicator gets a low value. In order to execute the conformity checks, the compliance policy conflicts are configured in the validator (i.e. regulatory constraints on data types, security and network management, access-control, business needs, work flows). The policy verification tools can provide the deviations and the conformance to the specified policy rules, and integrate with other security systems such as the Security Information Management System, and the Log and Embedded Event Managers. This allows the facilitation of the use of event detectors for parser-based events. Based on the event criteria, an action is executed during which the detector waits the policy to exit.

The access control model offers the interconnection interfaces for external IDPs through delegation authentication, federation services or other security token services, when there is a trusted relationship with other IDPs. The authentication module requires no changes of the communication protocol, the SPs or the external IDPs. This functionality is transparent and the SPs are not aware of the external trusted IDPs. The SAML assertions are forwarded to the external IDPs to avoid copying code and functionalities. The integration with other trusted external IDPs provides a secure means to improve user experience, minimize the user

administration overhead and federate with other systems that have shared or merged resources. Hence, the policy rule validator ensures policy conformance by performing the appropriate validation controls and policy rule checks and resolving any policy conflict.

The consolidation of the policies and the integration of the aforementioned modules facilitates a holistic policy-based access control approach. The integrated components (Figure 23) are common management services, independent of the service delivery platform and the service administration.

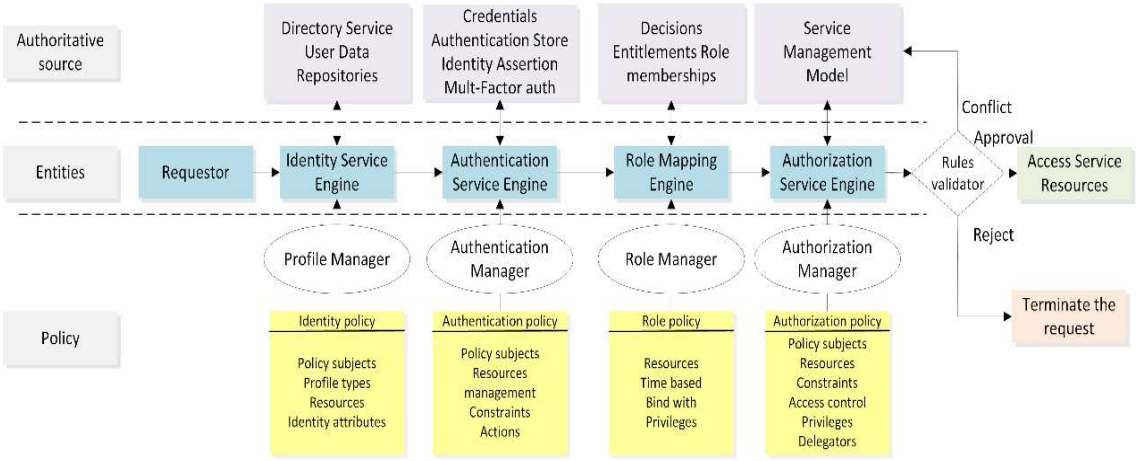


Figure 23: Integrated Identity-Management model for SOA

System Model and Definitions

The model enables an authentication decision to be requested from PDP. In this subsection a paradigm is presented where a SAML protocol response for an authenticated client is analysed. In more detail, PDP considers that the assertion *subject* (S) and the assertion *resource* (R) are both associated with the specified *attributes* response (A) to construct the assertions from PIP and PRP based on the given *constraints* (N) and on the *context* (C). Next, the authorisation function for each policy in the set of this context and permissions (C, P) returns a value to the *EventMonitor* that triggers an action upon the specific subject, the resource and the context (S, R, C), if necessary. Eventually, the authentication assertion is generated and consumed by the downstream web-services. By monitoring the appropriate sets with these objects, the proper policies for the appropriate targets can be handled to generate the necessary assertions, as shown in Figure 24.

```
1:  S←subject(assertion); R←resource(assertion);
2:  C←context(request); A←attribute(PIP);
3:  N←constraints(conditions); P←permissions(assertion);
4:  PDP←PIP(S,R,A) ∧ PRP(S,R,N);
5:  for policy in PDP(C,P) do
6:          EventMonitor←authorisation(policy)
7:          if EventMonitor = '0' then exit
8:                  else action(S,R,C) return;
9:  exit
```

Figure 24: Access control policy algorithm

Reviewing the location of the rules, the targeted objects and the operational failures are key factors with a view to improve the model and avoid any inconsistencies. Additionally, in

the interest of ensuring consistency, improving the performance and fulfilling the security requirements for unified access control services, ASPIDA amplifies the policy capabilities by utilising a set of standards, such as XACML, SOAP envelopes, messages and SAML assertions to integrate the modules. Figure 25 presents the architectural overview of the ASPIDA and the access request flows from the subject to the protected resources.

In more detail, the access request is redirected from SP back to the subject and then the request is encoded in a SAML authentication request. Once the subject has been authenticated, a SAML authorisation assertion is inserted into the SOAP message. Provided that the assertion format and the processing requirements are met [236], and the security filters are successful, then the assertion can be consumed by a downstream web service. The SOAP messages with SAML assertions can provide PDP with the necessary information by consulting the attribute and policy repository to collect some more attributes from the subject, the environment or the resource store (i.e. the file system database). Finally, PDP provides a formal response back to PEP to enforce the policy decision and grant access to the subject.

5.5.2 ASPIDA authorisations

The software developers need to face several challenges in the design and in the implementation of new devices, gateways, applications and services. Mobility, mission-critical devices, needs for real-time updates, gateways in use for military purposes and data privacy for e-health cases pose software and configuration requirements in order to implement reliable and secure solutions. This section describes the ASPIDA in terms of the policy-driven authorisations by utilising token-based exchanges for gaining access to the resources and to the µServices in the resource-constrained M2M ecosystems. The system architecture enables higher technology diversity and independence between the dynamic authorisation and the policy-based management by dividing the architecture description into domains, layers and views. By utilising µServices technologies, the architecture enables higher technology diversity and independence between the dynamic authorisation, the policy-based management and the SOC services. Additionally, the reference architectures can be used to meet the challenges in the simplification of the development of the appropriate solutions. A proof-of-concept is developed to evaluate the characteristics, perform a comparison analysis of the proposed solution with other frameworks and illustrate the diverse types of authorisations supported by the proposed framework. The proposed scheme and the authentication flows facilitating the message exchanges between the involved entities provide a series of security enhancements for M2M computing.

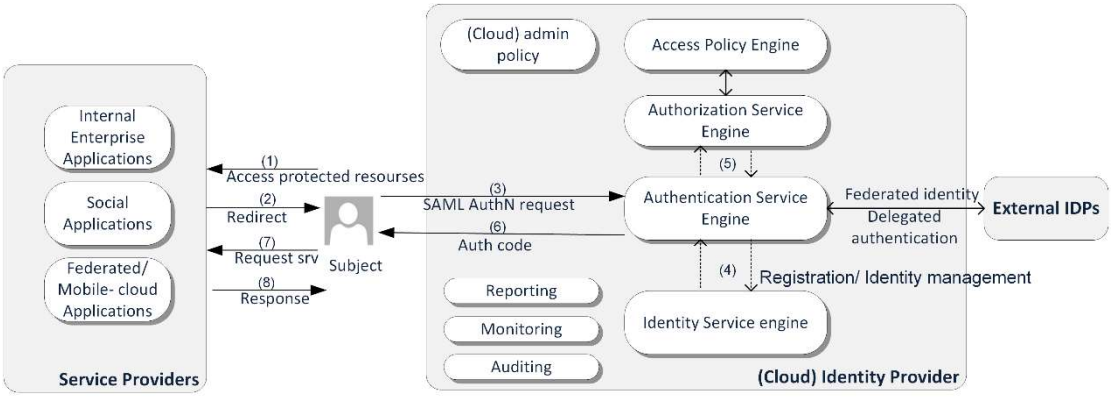


Figure 25: Overview of the model

The key entities involved in the authentication flows facilitating the message exchanges are summarised below.

- *Client,* which often represents the application to gain access to the protected resources.
- *Authorisation Services,* which negotiate access to the resource and verify the access rights of the originators.
- *Resource server,* which provides the protected resources and receives the requests from the clients.
- *Originator,* which refers to the M2M device or the end-user application (i.e. mobile/smartphone application) that initiates the requests to the resource server.
- *Policy Information Point* (PIP), which acts as a source of attribute values needed for authorisation policies.
- *Registrar,* which stores the created capability tokens and responds to token authentication challenges by the resource server.
- *Policy Administration Point* (PAP), which includes a policy store (a repository with the policies) and provides the authoring and maintenance of the policy sets.
- *Policy Enforcement Point (PEP*), which intercepts the resource access requests, submits the resource decision request and enforces the authorisation decisions.
- *Policy Decision Point (PDP),* which interacts with PIP to retrieve the relevant authorisation policies and attributes and then evaluates the access request providing an authorisation decision outcome to PEP.

Figure 26 illustrates the full interactions between the entities in the policy decision authorisations in the case of M2M resources. The originator (i.e. the M2M device) needs to communicate over constrained networks to send the access request to the resource server, which provides the protected resources and receives the requests from the client (i.e. the application). A capability token is generated and exchanged with the originator to get access to the protected resources or the restricted service. The resource server utilizes the authorisation services to validate the capability token and determine whether to process or deny the request. The resource server may utilize an external authorisation server for the verification of the capability tokens (i.e. a database lookup in the token table) and uses the dynamically established keys to protect the resources. The authorisation services interact with the Registrar, which keeps the created capability tokens and provide the successful/error capability token response to the client with a set of claims / context / attribute values to the resource server about the authorisation. For instance, the authorisation services need to verify the validity of the capability token and that the access rights can be exercised. The keys need to be provisioned for generating and verifying the tokens with the resource server in advance, whereas the client needs to be trusted by the registered services before initiating the request. Additionally, the revocation of the tokens ensures that there are no stale requests. This capability is achieved by adding a lifetime to the tokens.

In order to facilitate and resolve authorisation policies at runtime, the capability token passes to the API endpoint and calls the relevant policy method. Then, PEP intercepts the access request to determine whether to entitle and authorise the client to the resource. PDP evaluates the relevant policies so as to accept or reject the request. Upon receiving all the relevant attributes from PIP along with the access control policies from the policy store, PDP generates the policy decision. Provided that all the authorisation checks (i.e. subscription, payment-based and platform validation checks) have been completed, PDP updates PEP with

the policy decision granting access to the protected resource based on short-lived capability tokens issued by the token authority for a finite amount of time and without sharing the originator's credentials. Access can be granted temporarily given the related attributes, the conditions and the originator's needs. Eventually, PEP enforces the appropriate authorisation attributes decisions and delegates the access decision.

The decoupled µServices realise the processing of the requests where the communication is widely based on COAP methods (i.e. GET, POST, PUT and DELETE). The binary encoding of the COAP message header and a small subset of options are required making them suitable for IoT systems that use low power technologies. More precisely, the originator needs to communicate over constrained networks to send the access request to the resource server asynchronously over COAP messages. A capability token is generated by the originator and the resource server utilises the authorisation services for the validation of the capability token. Afterwards, the resource server consumes the capability token for matching the responses with the requests independently from the underlying messages. The resource server utilises an external authorisation server for the verification of the capability tokens (i.e. a database lookup in the token table) and uses the dynamically established keys to protect the resources.

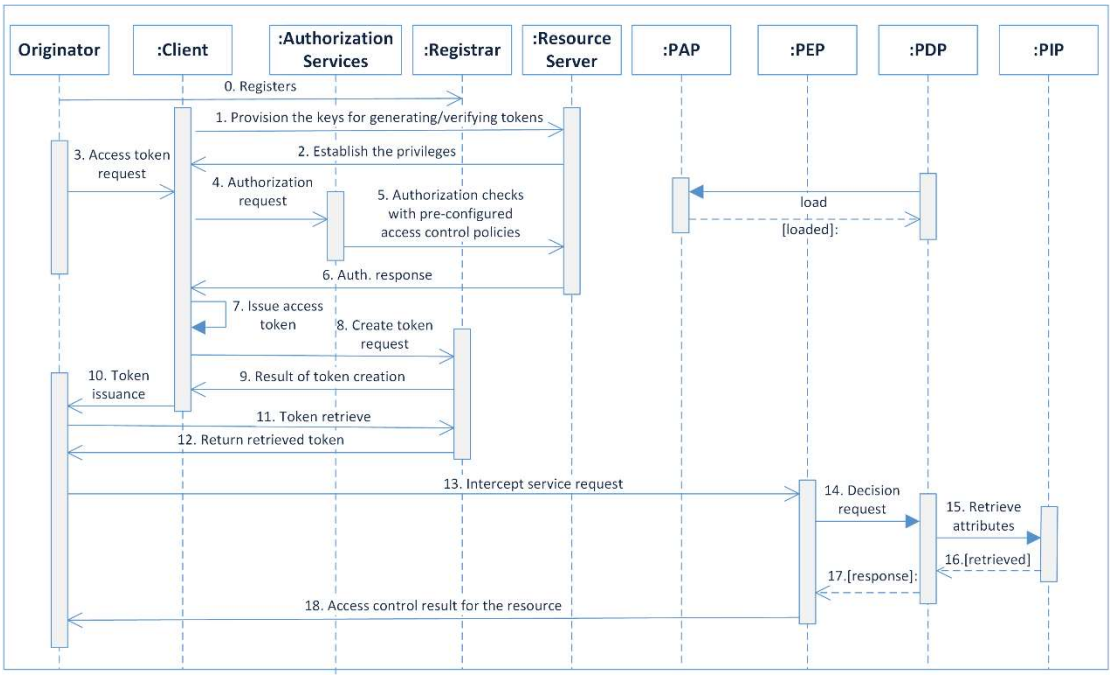


Figure 26. Policy-driven authorisations

5.5.3 Relationships

ISE provides the identity and policy management procedures along with control functions for identity and policy definition. The *users* (U) have access to a given *set of services* (S) and they have been assigned *a set of roles* (R) in order to perform specific *actions* (Act) with respect to the users' requests, their duties and the compliance with the policies (Pol_i) and the business needs. Thus, each user may be assigned to one or more roles and each role (r_i) permits one or more actions (Act_i). Furthermore, each action should follow the appropriate policies and implement the necessary controls. In terms of the associations, the ASPIDA model defines the following relationships:

$$\forall u \in U \rightarrow \langle r_1, \dots r_n \rangle \quad (5.5.3.1)$$

$$\forall r \in R \rightarrow \langle Act_1, \dots Act_n \rangle \quad (5.5.3.2)$$

$$\forall act \in Act \rightarrow \langle Pol_1, \dots Pol_n \rangle \quad (5.5.3.3)$$

AuthNSE receives the access request from ISE and performs the authentication policies, constraints and actions. The users' entitlements are managed and enforced across a variety of resources through *a set of privileges* (P) across *the resources* (Res). Each action is related to one or more privileges (p_i), which may provide access permissions and assignments to resources (res_i, i.e. systems, data) across multiple locations to the users.

$$\forall act \in Act \rightarrow \langle p_1, \dots p_n \rangle \quad (5.5.3.4)$$

$$\forall p \in P \rightarrow \langle res_1, \dots res_n \rangle \quad (5.5.3.5)$$

An identity assertion is created during the authentication request. Notably, there can also exist authentication constraints when trying to access the protected resources. For instance, an access request may be protected with SSL, may imply multi-factor authentication needs or enforce additional proxy authentication constraints.

$$u_i \in U; res_i, res_j \in Res \rightarrow u_i \; may \; access \; (res_i, res_j) \quad (5.5.3.6)$$

RME defines the role policy and determines the entitlements and the role memberships. Apart from the relationships (5.5.3.1), (5.5.3.2) and (5.5.3.3), the RME may also incorporate a role policy making process to specify the appropriate roles and restrictions. For instance, the role lifecycle management, the role of review process and the role of retention policies are typical examples to be undertaken by the RME.

$$\forall r \in R \rightarrow \langle pol_1, \dots pol_n \rangle \quad (5.5.3.7)$$

The AuthZSE includes the delegation and secure access management actions, the resource authorisations and the access control operational issues. More specifically, the delegation actions can be the carry-over of responsibilities to other people to undertake their duties, while secure access management options could be the classification of data sensitivity and confidentiality, the special treatment of the resources and authorisations, but also the management of user roles, operations and permissions.

$$\begin{gathered} u_i \in U; r_i \in R; act_i \in Act \; \rightarrow \\ u_i, u_j \in U; r_i, r_j \in R; act_i, act_j \; \in Act \end{gathered} \quad (5.5.3.8)$$

- The *principle of minimal privilege* or of least privilege (POLP) giving a user account only those privileges, which are essential to that user's work is enforced.

$$\begin{gathered} u_i, u_j, u_a \in U; r_i, r_j \in R \rightarrow \\ The \; approver \; u_a \; assigns \; only \; r_i \; to \; u_i \; and \; not \; r_j \\ The \; approver \; u_a \; assigns \; r_i \; only \; to \; u_i \; and \; not \; to \; u_j \end{gathered} \quad (5.5.3.9)$$

- The *privilege revocation* is giving up some or all the privileges the user possesses, or takes the privileged rights away.

$$\begin{gathered} u_i, u_a \in U; r_i, r_j \in R; act_i \in Act; p_i, p_j \in P \\ \rightarrow u_a \; removes \; r_j \; from \; u_i; \\ u_i \; cannot \; accomplish \; acti \; anymore \end{gathered} \quad (5.5.3.10)$$

- The *privilege escalation* is the act of exploiting access and gaining elevated access to resources that are normally protected from an application or user, usually for unauthorised actions.

$$u_i \in U; r_i, r_j \in R; act_j \in Act \rightarrow r_i \subseteq r_j; \\ u_i \; gains \; elevated \; access \; r_j \; to \; perform \; act_j \tag{5.5.3.11}$$

- The *privilege bracketing* is a temporary increase in software privilege within a process to perform a specific function.

$$u_i, u_a \in U; r_i, r_j \in R; act_j \in Act \rightarrow r_i \subseteq r_j; \\ u_a \; assignes \; r_j \; to \; u_i \; to \; perform \; act_j; \\ u_a \; revokes \; elevated \; privilege \; r_j \; from \; u_i \tag{5.5.3.12}$$

- The *privilege separation* is necessary when a program is divided into parts which are limited to the specific privileges they require to perform a specific task.

$$u_i, u_j \in U; r_i, r_j \in R; act_i, act_j \in Act \rightarrow \\ act_i, act_j \neq \emptyset; act_i \cup act_j = act_{ij} \tag{5.5.3.13}$$

The privilege separation may also separate the privileged parts of the system from its unprivileged parts by putting them into different processes, as opposed to switching between them within a single process.

$$u_i, u_j \in U; p_i, p_j \in P; res_i, res_j \in Res; act_i, act_j \in Act \\ \rightarrow \; create \; new \; priv.parts \; res_{i1}, res_{j1}; \\ new \; unprv.parts \; res_{i2}, res_{j2}; \\ res_{i1} \cap res_{i2} = \emptyset; res_{j1} \cap res_{j2} = \emptyset; \\ res_{i1} \cup res_{i2} = res_i; res_{j1} \cup res_{j2} = res_j; \\ p_1 \rightarrow \langle res_{i1}, res_{j1} \rangle, p_2 \rightarrow \langle res_{i2}, res_{j2} \rangle \tag{5.5.3.14}$$

The Rule Validator authorises numerous policies from multiple sources. Various contradicting requirements, *access rights conflicts* and *violations* need to be resolved efficiently.

$$u_i, u_j \in U; r_i \in R; act_i \in Act \tag{5.5.3.15}$$

Finally, there is a variety of SoDs that need to be managed either at the service management level or at the application level before the approvers authorise access to the resources. These SoDs should be defined in advance through a set of policies and internal controls (Pol) in order to prevent any 'toxic combination' of roles and privileges to the users.

$$\forall sod \in Sod \rightarrow \langle pol_1, \dots pol_n \rangle \tag{5.5.3.16}$$

A SoD policy prevents the assignment of roles that would allow performing contradicting duties that may result in accidental or deliberate fraud.

$$u_i, u_a \in U; r_i, r_j \in R; act_i, act_j \in Act \rightarrow \\ u_a \; requests \; assignement \; of \; r_j \; to \; u_i; \\ pol_j \; is \; triggered \; and \; verifies \; if \; r_i \neq r_j; act_i \neq act_j \\ if \; not, the \; request \; is \; rejected \tag{5.5.3.17}$$

The framework also covers conformity checks for compliance policy conflicts, which should be configured in the policy rule validator. Policy conflicts may result into *violations* (Vio).

$$u_a\ assigns\ r_i, r_j\ to\ u_i\ \rightarrow$$
$$r_i \in act_i\ and\ r_j \in act_j\ then \quad (5.5.3.18)$$
$$act_i \cap act_j \neq 0\ or\ \{act_j, act_i\}\ tigger\ Vio_{ij}$$

5.5.4 An applicable case type

In this dissertation, a family locator application that lets the family members accurately locate each other is considered. In case of emergency, the authorised security guards should be notified and be able to access the precise location of the person whose life or health is in danger. The policy-based methods allow access based on the device location (e.g. area, region), the time and the date of the request. Additionally, various premium subscriptions can be used in the policy definitions offering better and higher quality services like who has access, which case data, the valid period of the request, the given purpose and the type of the request (e.g. health emergency, life-threatening and violent events). Hence, the guards can be granted access by enabling temporal delegations based on highly dynamic conditions and environment attributes (e.g. system time, device type), despite having the same access rights and roles. The dynamic authorisations support the complex policies to allow access to the patrol company based on the dynamic environmental settings. In order to implement our distributed capability-based access control approach, CBOR is used for the representation and the format of the capability token because of its suitability in constrained environments, such as those expected in IoT scenarios. When CBOR data protection is needed, CBOR Object Signing and Encryption (COSE) ensures the necessary security levels (i.e. encryption). The separation of the request and the access tokens also helps to reduce fragmentation. The COSE Web Tokens (CWTs) support built-in expiry mechanisms, they are designed for small code and message size and they can be used in the COAP responses enabling fine-grained access control information within the token. The business policies can also meet the growing number of regulations (e.g. EU 2016/679 [237] on the protection of personal data to protect patients' privacy during patient's data sharing for healthcare and research purposes). Figure 27 presents the authorisation after a successful login in the application.


```
{
 "_id": "login_1",
 "_version": 1,
 "_source": {
 "log_format": "http",
 "docker_swarm_service_name": "public-access",
 "http_authorisation": "Token eaJhbGciOiJIUzI1NiJ9.eyJqdGkiOiI4k1OS06M",
 "docker_container_id": "7f2fffbfd4ffe23620bfa56f4b76131a8b5792c395f",
 "http_x_forwarded_proto": "https",
 "http_duration": "0.385",
 "http_sent_body_bytes": "117",
 "tag": "7f2fffbfd4ff",
 "docker_image_id": "sha256:9811fc4f48dabffdc6e9ef3c497fdc33308a904a5b2",
 "http_sent_content_type": "application/json",
 "http_duration_numeric": 0.385,
 "jwt_claims": {
 "sub": " Locator_App_LogIN",
 "scopes": [
 "access"
 ],
 "jti": "8edaf959-7361-48ec-af29-eceda53d965a"
 },
}
```


Figure 27. Authorisation token to get access

Nevertheless, the authorisations can be revoked either on-demand, or after a definite period. The access control policies can be created with a view to facilitate and improve the family locator application interactions such as the dynamic authorisations to premier transportation services (e.g. taxi, shuttle bus, rental and leasing) based on a subscription and payment model as depicted in Figure 28. Alternatively, access can be granted to security wardens, emergency and rescue services for life-threatening cases (i.e. panic-mode) based on the enumerated risk management and threat levels, as shown in Figure 29.

```
{
 "_id": "Premier_Services_policy_1",
 "_Version": 1,
 "_Statement": {
    "Resource": [
  "subscribedservices::${fam:memberid}/",
  "subscribedservices::${fam:memberid}/*"
    ],
    "Action": [
  "fam:GetSubPolicy",
  "fam:GetSubStatus",
  "fam:ListSubServices"
    ],
    "Effect": ["Allow"]
 },
}
```

Figure 28. Dynamic authorisation policy for premier transportation services

```
{
 "_id": "Panic_Mode_policy_1",
 "_Version": 1,
 "_Statement": {
    "Resource": [
  "subscribedservices::${fam:memberid}/",
  "subscribedservices::${fam:memberid}/*"
    ]
    "Action": [
  {"fam:memberid":"Reason","value":["*","Alarm","Extreme conditions!"
    ]
    "Principal": "*",
    "Effect": ["Allow"],
    "Conditions": {
  "any": [
  { "field":"threat_level","operator":"greater_than","value":"normal"},
  { "field":"extr_condition","operator":"greater_than","value": "normal"}
  { "field":"high_risk_events","operator":"is","value":"present"}
  ],
    },
}
```

Figure 29. Dynamic authorisation policy rule for panic-mode

Each authorisation policy rule embraces a unique policy identifier, the version and the statement that expresses the relevant permissions, the *resources* (required) describe the object(s) referred to the applicable µServices, the *effect* (required) to allow or deny access, the *actions* (required) describe the type of access that should be allowed or denied, the *conditions* (optional) when the policy is in effect and the *principal* (optional) specifying the object(s) allowed or denied to access the resource(s). The dynamic authorisation policies

allow/deny access to the protected resources (i.e. services and operations) based on the access control decisions.

5.5.5 Additional case studies

Two simplified case studies are demonstrated here on how to enforce the SOA policy-based access control model by controlling and restricting access to the protected resources. In the first scenario, access to project management files is analysed. When a project manager requests to access the details of a specific project, she needs to be granted the appropriate permissions for each object. A project can be publicly accessible, accessible to the organisational units and the project managers groups, or accessible to specific individuals and classified access groups (i.e. for restricted, confidential, military projects). Thus, the authorisation policy, the permissions and the entitlements need to be specified in advance through PAP. The information is kept in the policy database and the attributes are kept in PIP (i.e. consultant/employee, types of sensitive/confidential resources). The initial request of the project manager is sent to the application that keeps the files, which can be stored locally or in the cloud. Then, the request is redirected to the authentication services. Now, there is a need to validate the project manager's identity in the directories or in the case of delegated authentication in the external IDPs. However, the subject is granted access to the resources based on the assigned roles and permissions to individual objects or specific groups. When the subject requests to modify a resource (i.e. the master project management file), then the subject needs to be authenticated against her credentials and user attributes to get access to the protected resources. If authenticated, PDP evaluates the authorisation request upon the PIP's and the policy database's information, and grants or rejects access. In the case of cloud computing or distributed administrative domains and federated identities, an interconnection is also required. The APE controls the rules and the policy governance for any access requests as well as the validation response for authorisation requests. The enforcement of the appropriate access control policies is performed upon a subject's access requests and by conducting conformance checking with verification tools for XACML policies [238].

In the second scenario, the provision of public administration services and the production of copies of birth certificates is analysed in the context of dynamic authorisations. Typically, the individuals can order a copy of a birth certificate through the portal of the Register Office ordering service. The services communicate by passing identity and personal data between the local and the central register offices. SOA is the model, which is delivered by central and local/regional governmental authorities (i.e. local register offices). The use of SOA in the implementation of e-government services aims to improve the interoperability and the efficiency of services used by the governmental agencies. The model can also maximize the integration and the orchestration accomplishments for these services. In terms of user access management services, the SOA based solutions need to review the integration and implementation with access control services and the respective business process modelling flows. The flows should define the appropriate access controls to secure the online resources, restrict access to sensitive personal information (i.e. Social Security Numbers, driver's license information, medical history and birth certificate details) and refrain from abusing the resources, avert masquerade attacks, or getting unauthorised access into e-government user data repositories. Although the access controls can be implemented based on Discretionary Access Control (DAC) or Mandatory Access Control (MAC), Temporal-RBAC is the most attractive solution to provide such e-government services. Temporal-RBAC supports information hiding with time varying nature, SoD and least privileges principles, which are

critical for the provision of e-government services. Each performing public servant in the Register Office is given the least privileges required to perform specific operations. The servant can access specific resources and execute the appointed tasks. Access to the resources is granted based on the servant's enterprise role. The type of the contract and the employment status determine the corresponding access control level for each public servant, while strict auditing infrastructure services such as transaction monitoring, event handling and error logging ensure the identity data integrity, accountability, confidentiality and access control requirements.

As the service governance model defines how to administer and maintain the services, role engineering and modelling should be defined and linked with dynamic human resource management. The administration of these services is achieved with policy-based mechanisms needed to enforce the governance decisions that reflect the structure and rules to be supported. Policy constraints or conflicts concerning the service orchestration and routing need to be resolved efficiently. For instance, a birth certificate request can be forwarded to another region, where the requester is registered. An authorised/delegated person ordering the certificate can be different with the name on the certificate. Special conditions may also be applicable for citizenship or conflicting laws regarding nationality, which need to be resolved based on formal procedures (i.e. policy language). Therefore, certain types of policy testing need to be conducted. The more extensive the policy tests and the evaluation of the policy sets are, the greater the confidence in the service assurance and the smaller the risk in policy conflict.

From an administrative point of view, the access rights need to be up-to-date and monitored regularly. The access can be revoked due to administration policies, employment status changes, or security reasons. For instance, an employee may have to move from the Registry Office to another public administration body, as there are changing capacity assets and human-resource needs. Consequently, the ex-Register Office employee's access-rights must be revoked. The access decision rules can be formulated according to the regulatory compliance. For instance, the Register office activities and tasks need to be compliant with the legal framework to offer these services (i.e. protection of personal data, privacy laws, EU regulatory framework). For instance, compensations, expenses and allowances should be reimbursed and allocated, provided that no individual public servant handles any combination of the record keeping, the authorisation and the reconciliation roles. In the absence of these controls, there can be potential fraud risks for unchecked faults or abuse. Therefore, not only there can be policy constraints, but there can also exist remediation policies to resolve any exceptions or to certify the exceptions as permitted.

In the 'read' access-level, the public servants use the values or the data set of the authoritative service data source (i.e. the tax clearance) as the requisites for multiple input administrative procedures by having access to the set S-read, but they cannot manipulate the user data records. In other cases, (i.e. e-Administrative fee payments, social security allowances and unemployment benefits), the appointed servants may execute transactions at 'execute' access-level S-exec. Additional verifications and validations can be imposed to access the user attributes or modify the user data. The workflow sequence is depicted in Figure 30.

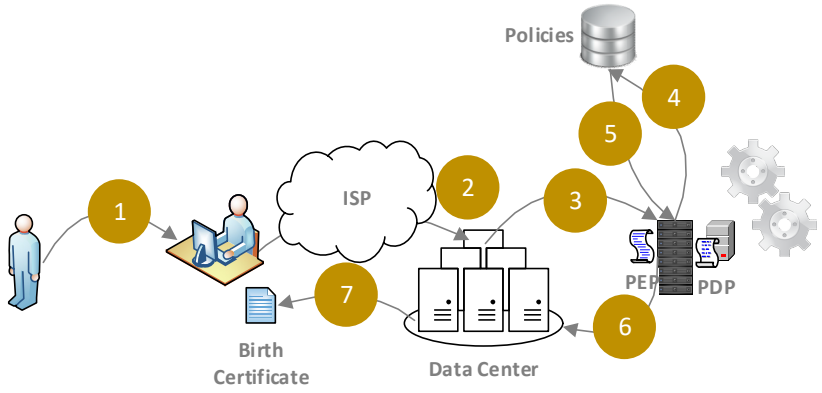


Figure 30. Access control scenario for the CSC case study

Chapter 6

Validation

6.1 Outline

This chapter describes the design of the evaluation, the validation methods and the respective outcomes of the experiments in the context of network and security policy-based management. The data gathering methods are explained followed by the data and statistical analysis, and the performance measurements are compared with other similar frameworks. More specifically, this chapter considers the:

- Evaluation framework and the methods for the policy-based management approach to formalise and control the behaviour of the network and the service domain in the context of policies, SLA, QoS, security, authentication and delegation
- Description of the experiments design to run the necessary tests
- Description of the distribution and the justification for the type of samples of the methods for making observations. Moreover, the description of the data analysis and tests are also included

6.2 Evaluation framework

6.2.1 Network domain

SLA metrics

ASPIDA includes various functionalities, as described in the preceding chapters, with respect to deploy the network and security policies. At the network level, the network resources and the policy rules are examined for the selection of a dynamic path depending on the current network state and conditions. For instance, the traffic is rerouted to other paths in the case of congested links and event thresholds (SLA) so that the obligations are met. Although the adaptive routing increases the complexity and the integration needs with policy-based and event-based components, they may increase connection blocking and underperforming cases. Hence, ASPIDA introduces SLA metrics into QoS-aware PBNM system in order to improve the adaptive routing performance significantly and the allocation of network resources.

Most QoS techniques focus only on one SLA indicator (e.g. one among delay, jitter, packet loss, and bandwidth), since it is often difficult to deal simultaneously probably with several contradicting indicators. For instance, minimizing the end-to-end delay may result in higher packet losses. Various QoS features can be enforced through traffic policies, using packet classifiers, markers and policers in order to provide the appropriate levels of traffic controls.

QoS and traffic provisioning

ASPIDA addresses the dynamic adaptations of the QoS requirements and the policies to alleviate under-performing entities and inappropriate resource usage. Although the vast literature on PBNM frameworks propose a solid base in this regard, the addition of QoS, context-aware and dynamic capabilities increase the scalability of the client applications by dynamically altering the replication schemes based on the network state, such as in [239][238]

for Software-Defined Networking. Hence, it is vital to determine the specific node attributes and the lower-level link parameters on QoS routing for both approaches, since QoS routing encompasses the collection and the maintenance of the latest state information about the network conditions and it utilises QoS requirements in the path-finding mechanism.

Adaptive QoS Routing

An *adaptive QoS routing* mechanism can be divided into three distinct functions. The first function provides the dynamic routing algorithm for route discovery and collects the local QoS-related information enabling routing cost optimization based on different QoS metrics. The route discovery provides a path-finding mechanism. The second function uses a local routing table, which is created with QoS related references for each node. In order to accomplish this function, a local link monitoring function is used. Typically, if a link fails, then the best route in the new topology can be recomputed. However, the aforementioned process does not implement a policy-based adaptive routing. Thus, in ASPIDA a third function is introduced that utilises a final decision-making system. This system identifies and ranks alternative routing paths based on QoS constraints, such as the traffic distribution rates, bandwidth reservation and the number of hops. The routing metrics are measurable values that can be changed based on the decision-making system policy rules, which influence the new path selection. Through this approach, the integration of different modules and characteristics (e.g. QoS, network diameter is accomplished, transmission rates, load sharing capabilities) with the changing network conditions (e.g. traffic load, remaining bandwidth, reserved bandwidth, error rates). In the case of a link or relay node failure, the algorithm redirects traffic to other paths. The flows are sent to the destination endpoint, if there is an alternate path. Load sharing techniques across the redundant paths can also provide increased network efficiency and performance. Another technique is to use Policy-based Routing (PBR) with IP precedence using a class-based marking. The predefined classes provide the route path selection mechanism.

The IGP tuning problem

Chang et al. [240] analyse the event-driven policy distribution in regard to adaptive policy procedures, nevertheless QoS adaptive routing should be used with discretion, because it can lead to sub-optimal routing, while re-computing and changing the path selection may lead to instability [241]. Apart from the routing protocol attributes, there are several other factors that impact the path selection process. The network diameter and the heterogeneity of the network entities result in distinctive features and capabilities that play a significant role in the diversity and the analysis of the selection process of the routing protocol. The IGP characteristics influence the path-finding operation and the enforcement of the appropriate routing protocol metrics. The path-finding mechanism for the provision of centralized or distributed models is responsible for providing updates about the availability of alternative paths. Pham et al. [242] analysed the performance of the reactive shortest path algorithm integrated with load balancing algorithms. The experiments showed that the amount of overhead increases with the number of multiple paths and if the number of paths exceeds three, then the overhead increases significantly. The path computation server can be either discovered dynamically or statically configured. Moreover, it can be used by other nodes to perform the route calculation. Additionally, there exist many other techniques for path recovery and new path selection. For instance, IP Fast Reroute (IPFRR), which can be enhanced with Loop-Free Alternates (LFAs) by tunnelling the packets to provide additional logical links [243], remedies network failures and reroutes traffic from one path to another when a link or

a node failure occurs. This multipath routing operation plays a significant role in the performance and the convergence of the routing protocol.

Complementary to the path-finding operation, the routing protocol metrics assist in the calculation of the optimum routing path. Several routing protocol metrics can be utilised to calculate the total path cost, a choice which affects the final path determination. In order to be able to perform the path-finding process, IGP routing protocols associate a weight (cost) with each link. Each IGP routing protocol uses different metrics. According to RFC1195 [244], IS-IS provides for optional QoS routing based on throughput (the default metric), delay, expense, or residual error probability. OSPF uses an arbitrary cost as the metric for each link allowing the network designer to force the choice of routes. Even though OSPF is not suitable for adaptive routing, Apostolopoulos et al. [245] describe some QoS routing extensions to OSPF proposing that the processing cost of QoS routing is not excessive to support the QoS adaptive routing process. IGP link weights can significantly affect the performance of intra-domain link failures.

The available bandwidth of the link is the metric extension in the implementation of QoS routing in OSPF extensions. Thus, link resource reservation can be considered for QoS routing. If none of the extensions are configured, then OSPF performs equal cost load balancing over the paths in case there exist differing paths with an equal number of hops and equal cost links. The total cost from the source to the destination can be the same. Still, in the case of differing unequal cost paths, the lower cost path is used to send the packets toward the destination. Figure 31 illustrates that if r_i is the reserved bandwidth and c_i the capacity of the link, then the path computation at each router (using a modified Dijkstra algorithm) is the path with the largest value of $\min(c_i-r_i)$.

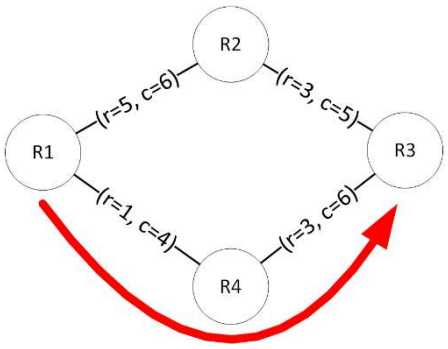


Figure 31. OSPF extensions for QoS routing

To tune the IGP performance, the available bandwidth, the link capacity, the reliability of the uplinks, the traffic levels and the routing protocol weights can also be considered by the adaptive routing process. The local IGP tuning may cause routing changes at the egress points for specific traffic flows. Additionally, it is very important to tune the Shortest Path First (SPF) delay value, because the SPF value can be appropriately set to adapt to the changes caused by a failure in the network. If the SPF delay is set to a high value, this may result in low convergence times. To the contrary, when set to a rather low value, there can be excessive next-hop recalculations and suboptimal best-path selection. Further, it is possible to consider both the SLA requirements and the link failures. In order to enhance IP resilience, the IGP tuning based on link weight optimization has been studied in the literature with techniques based on heuristic and genetic algorithms. For example, Salles et al. [246] have provided heuristic algorithms for weight assignment optimization.

Apart from these standard operations, PBNM can manage and handle SLA-driven events through applets. Depending on the design, there can be either an intra-domain policy management for intra-domain routing or an inter-domain policy management for inter-domain routing. This model can handle policies and incorporate a QoS packet scheduler, as different QoS classes can be implemented. Thus, the evaluation of an operational QoS model is possible, which is applicable to different traffic patterns. This model can deal with adaptive network conditions and various requirements.

The network topology is represented as a directed graph G=<N, L>, where N and L denote the sets of nodes and links respectively in a routing domain R. Each link $\ell_m \in L$ is associated with a bandwidth capacity $c(\ell_m)$, which is the maximum sustainable usage of ℓ_m, a reserved bandwidth $r(\ell_m)$ and a link utilisation $u(\ell_m)$. In the following sections, the Adaptive Routing (AR) indicators are examined between any pair of nodes (n_i, n_j) in a routing topology r, where r ∈ R to find the optimum path across the nodes according to the IGP metric $m(\ell_m)$. This metric is dependent on the routing protocol attributes and the dynamic traffic conditions. The traffic between nodes n_1 and n_k is $t(n_1, n_k)$. In order to conform traffic, the SLA policy $s(\ell_m)$ has been set in the PBNM system, which is the SLA threshold for link ℓ_m. This policy is checked against $t(n_i, n_j)$ and $u(\ell_m)$ and it will trigger the appropriate actions dynamically. The following notation is used:

- The symbol I (vertical bar) denotes a disjunction and indicates a logical OR
- The function left arrow Y ← X denotes that the set X is mapped into the set Y

The function, where each link $\ell_m \in L$ has an IGP metric $m(\ell_m)$ from source to destination, is based on the IGP metric $m(\ell_m)$ as the minimum sum of the accumulated costs of the routing path selection.

$$\sum_{\ell m \epsilon L} m(\ell_m) \tag{6.2.1.1}$$

$s_{min}(\ell_m)$ is the defined lowest SLA threshold for link ℓ_m
$s_{max}(\ell_m)$ is the defined maximum SLA threshold for link ℓ_m

Under normal conditions:

$$s_{min}(\ell_m) \leq u(\ell_m) \leq s_{max}(\ell_m), \text{ where } m \in L, \text{ and} \tag{6.2.1.2}$$

$$s_{min}(n_i, n_j) \leq t(n_i, n_j) \leq s_{max}(n_i, n_j), \text{ where } i,j \in N \tag{6.2.1.3}$$

When there is excess traffic, saturated links or an SLA violation:

$$s_{max}(l_m) < u(\ell_m) \mid s_{max}(n_i,n_j) < t(n_i,n_j) \tag{6.2.1.4}$$

In this case, a traffic alleviation policy can be applied to remedy this undesirable situation. The action policy is given below:

$$t'(n_1, n_k) \leftarrow \lambda\, t(n_1, n_k) \mid u'(\ell_m) \leftarrow \lambda\, u(\ell_m) \tag{6.2.1.5}$$

Where λ is the derived deviation ratio, 0 ≤ λ ≤ 1 and t'(n_1, n_k) is the new conformed traffic after applying the appropriate action policy. In the following section, an analysis on the values of λ is presented to extract safe and useful conclusions on the policy enforcement. The higher the value of λ, the more efficient and adaptive the policy enforcement achieved. In contrast, the lower the value of λ is, the smaller the improvement on the utilisation of the network

resources accomplished. This can be further extended to under-utilisation conditions for better resource re-allocation and performance enhancements:

$$u(\ell_m) < s_{min}(\ell_m) \mid t(n_i,n_j) < \text{smin(ni,nj)} \quad (6.2.1.6)$$

By enforcing the appropriate action and reallocating the network resources appropriately, then the action policy becomes:

$$t'(n_1, n_k) \leftarrow (1/\lambda)\ t(n_1, n_k) \mid u'(\ell_m) \leftarrow (1/\lambda)\ u(\ell_m),\ 0 \leq \lambda \leq 1 \quad (6.2.1.7)$$

In order to apply the appropriate deviation ratio, the AR indicator needs to be evaluated in advance and then the ratio to end up with an improved AR indicator must be evaluated.

$$\text{AR} = \sum_{\ell m \epsilon \text{L}} m(\ell_m) \quad (6.2.1.8)$$

Towards this, $AR_{(LS)}$ is the adaptive routing indicator for the link state routing protocol and $AR_{(CM)}$ the corresponding indicator for routing protocols with a composite routing metric. Each AR indicates the minimum sum of costs between a source-destination pair (n_i, n_j) in all the routing topologies. The algorithm consists of the following steps

Step-1: Identify the optimum link provided by the routing protocol and calculate min{$AR_{(LS)}$, $AR_{(CM)}$}.

Step-2: For the set of traffic flows that are routed through min{$AR_{(LS)}$, $AR_{(CM)}$} in the routing topologies check the SLA conformance:

$$s_{min}(\ell_m) \leq u(\ell_m) \leq s_{max}(\ell_m) \quad (6.2.1.9)$$

***Step-3*:** If the set of traffic flows does not conform to the SLA through the EDs, e.g.

$$s_{max}(\ell_m) < u(\ell_m) \mid s_{max}(n_i, n_j) < t(n_i, n_j) \quad (6.2.1.10)$$

then the relevant PBNM policy must be triggered.

Step-4: The appropriate applet is triggered and the necessary amendments are made. This action policy can be enforced by using rate limiters or congestion avoidance mechanisms to alleviate the loads of the link:

$$t'(n_1, n_k) \leftarrow \lambda\ t(n_1, n_k) \mid u'(\ell_m) \leftarrow \lambda\ u(\ell_m), \text{ where } 0 \leq \lambda \leq 1 \quad (6.2.1.11)$$

Thus, the action policy reallocates network resources, as follows:

$$c'(\ell_m) \leftarrow (1/\lambda)\ c(\ell_m), \text{ where } 0 \leq \lambda \leq 1 \quad (6.2.1.12)$$

6.2.2 Service domain

A prototype has been developed based on Java and µServices technologies to facilitate the performance analysis. In order to implement the distributed capability-based access control approach, CBOR is used for the representation and the formatting. The CBOR data item is structured, encoded in a fairly small message size and transmitted in constrained environments, such as those expected in the IoT scenarios. For the capability tokens, the CWTs can be used in the COAP responses. The calling clients can get access to the endpoints and the services according to the access policies, and it must be ensured that both the authorised calls and the successful authorisation allow the call to process the request. In the same way, the authors in [247] propose a policy-based management of resources in fog computing,

expanding the current fog computing platform to support secure collaboration and interoperability between different user-requested resources in fog computing.

6.3 Design of experiments

6.3.1 Setup for adaptive routing

The performance of the PBM SLA-driven ASPIDA is demonstrated by conducting experiments in a real environment, the Hellenic Public Administration Network, hereafter named SYZEFXIS. This network provides the core and access infrastructure for 4,598 Public Administration bodies[17] aiming to satisfy all their communication needs. The design of this network is hierarchical and it is based on a three-tier architecture: access, distribution and core layer. At the access layer, the network provides dual connections for redundancy purposes to the small-sized, medium-sized and large-sized nodes. The nodes are interconnected at the distribution layer Point of Presence (POPs), which are further interconnected with 6 high-end core layer POPs of the backbone network. The Service Provider of SYZEFXIS limits the Committed Information Rate (CIR) and sets the bandwidth of each link to 2 Mbps, 8 Mbps or 34 Mbps depending on the type of the connected node. More specifically, small-sized nodes have 2 Mbps links; medium-sized nodes have 8 Mbps links, whereas the large-sized nodes use 34 Mbps links.

The system architecture (Figure 32) permits the assignment of link weights based on the IGP routing protocol behaviour, as well as the setting of SLA-driven link weights. The Service Level Specification (SLS) must indicate the conformance values of the network resources, such as bandwidth utilisation, network node performance, traffic rate thresholds, delay, jitter, packet loss and others. The architecture fulfils the SLA constraints, while minimizing the maximum link utilisation across the network, because traffic can be bursty and may increase over time. It is important to keep the worst link utilisation as low as possible to improve the overall network performance figures. The SLA conformance through the proactive monitoring of the network components has been considered to make the appropriate amendments in the path selection.

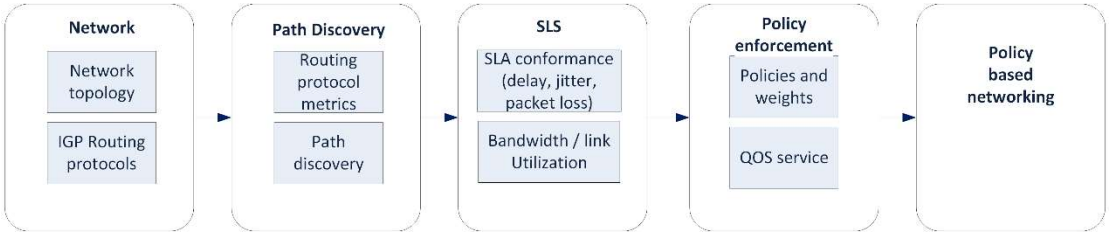


Figure 32. Policy-based networking with IP SLA

Since the SLA operations have been defined by setting thresholds, if the values exceed the defined thresholds, then the event detectors trigger the relevant applet to be executed, as a determiner of the actions to follow. The experiments combine different traffic conditioning mechanisms with the event detectors. For example, the IP SLA detector publishes an event, when an IP SLA reaction mechanism is triggered to enforce the related IGP and QoS configuration changes. With these in mind, there is a need to monitor the performance and the operational conditions, when events and faults occur. This will allow to take the proper corrective or any other desired actions.

[17] http://www.syzefxis.gov.gr, data as of the end of May 2018 including the nodes for VPN-1, VPN-2, VPN-3 VPN-4, VPN-21, VPN-22, VPN-23, VPN-24 and VPN-25.

Two scenarios are analysed in terms of path discovery and how the respective routing protocols both adapt to the dynamic network conditions with event-based adaptive access control decisions. The first scenario is a typical case of using a link state (LS) routing protocol and QoS definitions. The second scenario uses a routing protocol with a composite metric (CM) functionality, QoS definitions and an SLA-driven PBNM for adaptive routing. The former uses the OSPF protocol, while the latter utilises the Cisco proprietary Enhanced Interior Gateway Routing Protocol (EIGRP) protocol. In more detail, the packet loss thresholds for SLA conformance are defined and Class-Based Weighted Fair Queuing (CBWFQ) with policing actions for excess traffic is implemented. It is obvious that this architecture also applies to any kind of QoS aware actions, such as marking, classification, shaping and even resource allocation rescheduling. The goal of the setup is to evaluate an improved path selection process that meets the SLA requirements and enforce the appropriate policies, so that the adaptive routing can allow the network to intelligently choose an optimal path. The experimental results show that ASPIDA accomplishes a better AR, let AR' be the new effective adaptive routing indicator both for link state (LS) and composite metric (CM) based routing protocols in comparison to the AR indicator without the enhancements, i.e.:

$$AR'_{(LS)} \leq AR_{(LS)} \mid AR'_{(CM)} \leq AR_{(CM)} \qquad (6.3.1.1)$$

In order to improve the accuracy of the experiments, the IP SLA functionality on the target end-devices at the far-end remote sites is used. With IP SLA enabled to reverse the delays, the measurements can reflect the true network delays in place more accurately. Each node can be configured to support the IP SLA functionality either as the SLA agent or as the SLA responder. The SLA agent role is assigned to the router serving the PBNM functionality, while the IP SLA router responder feature is used in the remote network devices. Thus, the results can be evaluated by swapping the SLA roles and the PBNM functionality between nodes. PEP initiates the IP SLA tests including information about jitter, latency and packet loss to several remote responders. The topology illustrated in Figure 33 represents the simulated multi-path topology with redundant paths.

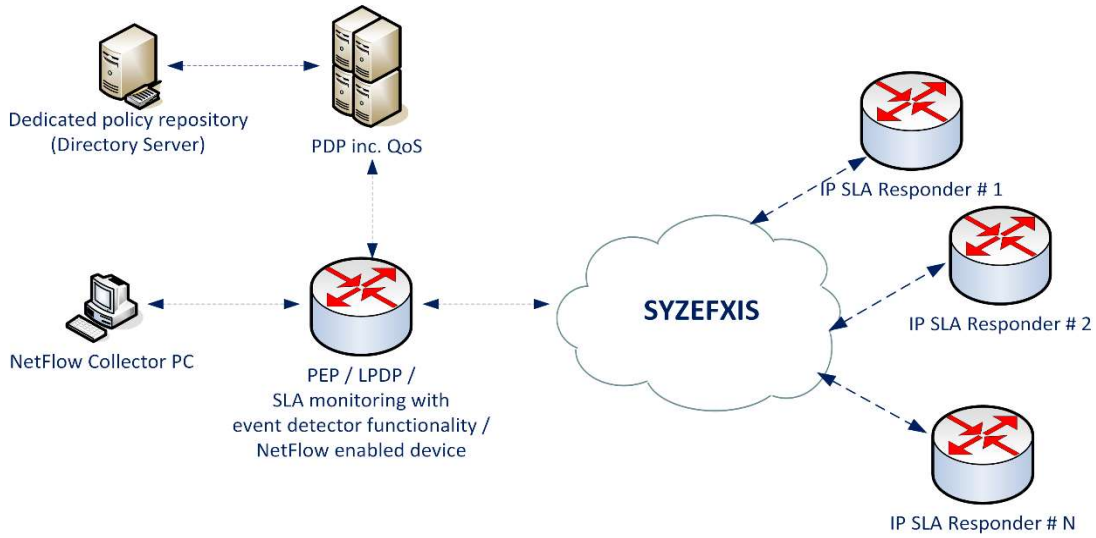


Figure 33. PBNM system – case study

With respect to the network management techniques, not only SNMP-based polling techniques, but also NetFlow techniques have been utilised that provide the beginning along

with the end times for each flow. The Internet Protocol Flow Information eXport (IPFIX) IETF NetFlow v10 [248], which is widely used and accepted by industry for real time monitoring, defines how IP flows are exported to a collector [249]. For the evaluation measurements, the NetFlow v9 technology is used for validation purposes in order to enable the monitoring and the accounting processes for the flow statistics. The flows at the packet level are examined in detail using an Intel-based PC as the data collector running a Real-Time NetFlow application at the PEP side. As NetFlow can be used either for ingress or for egress traffic, all the necessary statistical data can be collected. All data crossing the network nodes are cached and grouped together into *NetFlow export* UDP datagrams for export to the *NetFlow collector*. The NetFlow Collector assembles the exported flows and combines them to produce reports used for traffic and security analysis. The collector is located on a dedicated out-of-band management network, so that the NetFlow traffic is forwarded through the management interface without affecting the data traffic volume. With regard to avoid collecting duplicate flows, NetFlow is activated on the key-aggregation routers where traffic originates or terminates and not on the backbone routers that could provide duplicate views of the same flow information. The NetFlow records contain the Type of Service (ToS) field in the IP header as well as application ports, traffic volumes and timestamps. Hence, the NetFlow technology is used to observe the traffic profiles, verify the QoS levels achieved and optimize bandwidth for specific classes of service.

From a different viewpoint, the SNMP notifications can be triggered with system logging based on predefined thresholds to apply configuration changes through the PBNM system by utilising the IP SLA feature. With the proposed experimental setup, it is possible to export SLA statistics and define whether the SLA criteria are met via the command line interface (CLI) of the PEPs, or there is a violation. Therefore, the SLA and the policy control management are integrated in the architecture to provide better monitoring and policy controls. PBNM functionalities implementation is accomplished with the relevant configuration on the network entities (i.e. routers, switches). For the integration of these modules, EEM policies [250] using the router configuration software have been implemented. EEM uses applets, which allow the relevant commands to run when certain SLA conditions have occurred, thus enabling the appropriate handling of the events. For instance, the node and link failures exceeding allocated bandwidth resources can trigger policy-driven configuration changes. Such issues can be identified and resolved proactively by setting event detectors to monitor for specific types of situations or thresholds, or to run a set of actions periodically. EEM is used as the policy transaction protocol, since it allows automated capabilities inside both the PDP and the PEPs for specific user requirements and traffic conditions.

With the intention of validating ASPIDA, a testbed has been constructed based on Cisco Internetwork Operating System (IOS) on Unix (IOU) for the evaluation of the event-driven policies on the traffic flows. The network consists of the master controller (MC) acting as the Policy Decision Point (PDP), which is responsible for taking all the necessary decisions how to route traffic based on various criteria. The border routers (BRs) are also part of the testbed implementation. They collect the performance metrics and inform the MC about the network conditions. BRs also constitute the Policy Enforcement Points (PEPs) to enforce the policy decisions by the MC. Figure 34 illustrates the configuration of this testbed topology.

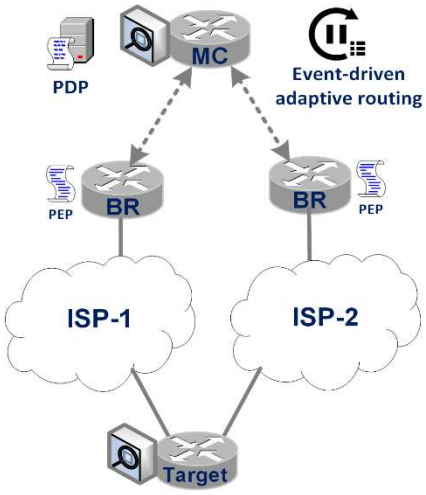


Figure 34. Optimizing SLA-driven adaptive routing

In this scenario, the end-to-end QoS configuration is tested based on the respective SLA criteria, and compared the improvements in the traffic flows before and after enforcing the event management model. In particular, the utilization, the delay, the loss and the throughput are analysed with event based routing criteria. Based on the conditions of the links, the appropriate policy is selected, which changes the routing control state for the predefined traffic classes. These classes can be defined by prefixes (i.e. destination/source IP prefix), DSCP values or the application (i.e. LDAP, POP3 over TLS/SSL, IMAP, HTTP). The delay, the hold-down timer parameters, the range and timer expired reasons, the out-of-policy events and the necessary thresholds (i.e. unreachable, delay, loss, egress/ingress bandwidth) can trigger the necessary actions and apply the appropriate routing changes.

Finally, QoS mechanisms are considered and configured as traffic conditioners. QoS service classes are used by the PBNM system allowing the enforcement of the appropriate QoS-aware techniques. In order to accomplish this enforcement efficiently, a mapping between QoS techniques and SLA-driven events is configured in the PBNM system. A typical example is the Smurf attack that sends a slew of ICMP Echo request packets, which can be monitored with SLA criteria. If the SLA thresholds are violated and a policy control for this has been configured in the PBNM, then adaptive configuration changes can be triggered. These changes refer to QoS policing actions either to remedy the situation/take preventive actions or to re-allocate the network resources. Eventually, the changes can affect the routing path selection mechanism, as well.

6.3.2 Access control analysis

The access control functionalities of the ASPIDA are depicted through a use-case. In more detail, the intelligent Bus on Campus (iBuC) prototype [219] offers a state-of-the-art transportation service within a university campus supporting multiple AVs and smart devices, also supported by the existing wireless infrastructure and M2M services. The security policies and the workflows for iBuC are further analysed in [222] where the iBuC service requests are forwarded via a booking application to the *control unit* (CU). Then, CU calculates both the AV's and the passenger's *estimated time of arrival* (ETA) to the bus stop and elects an eligible vehicle to operate an itinerary for passengers. CU collects various datasets regarding the vehicles' and the passengers' status by using wireless sensor devices, GPS and Wi-Fi communication capabilities. The security aspects for the communications are contrasted

between two functions; *Enrol/withdraw endpoints* and *Intercept service request*. These functions focus on the communication interactions and the flows of the messages exchanged between the entities.

Function 1: Enrol/withdraw endpoints

The registration with the appropriate security means that relying on digital identity management can be reinforced with the aim to uniquely identify the endpoints (i.e. resources) in the service registry. Thus, a Trusted Certificate Authority (TCA) is incorporated in the proposed solution for issuing digital certificates, a Registration Authority (RA) for verifying the endpoint request for digital identities and a Validation Authority (VA) verifying the validity of the digital certificates with the appropriate technical means, such as for:

- CA-issued full-Certificate Revocation Lists (CRLs)
- CA-signed delta CRLs, Compact CRLs
- VA-manufactured CRLs
- Online Certificate Status Protocol (OCSP)
- Simple Certificate Validation Protocol (SCVP)
- Certificate Management Protocol (CMP)

The entire certificate validation operation, including the path construction and the intermediate CA validation, is delegated to the VA service. Upon receiving the validation of the digital signature certificate, the validation application records the results in the registration reporting system. The reporting system audits the activity in the registration system and supports any required queries about the processes.

The service registry operator allows the endpoints to enrol into/withdraw from the service registry, if the digital identity assertion is successful and authorised. The endpoints are equipped with suitable hardware tokens that safely store the private/public keys and the process allows the legitimate access to the service registry. Finally, the brokered authentication with an IDP removes the need for a direct relationship with the service registry and allows the use of security token services that can also be used for multiple calls. Figure 35 depicts how an AV can be enrolled to or withdrawn from the transportation service registry based on specific workflows and sequence of activities.

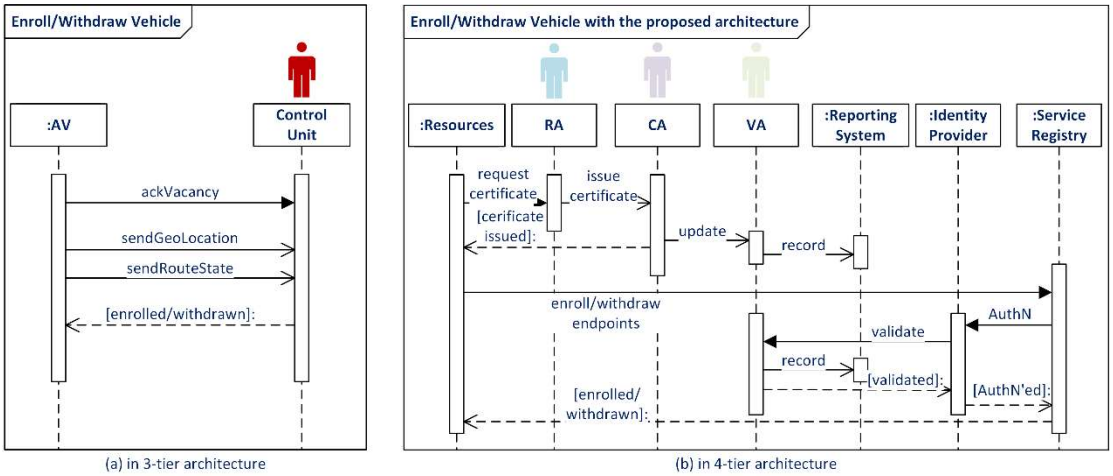


Figure 35. Enrol or withdraw an AV

Case A presents the workflows based on the original iBuC work and case B shows the messages exchanged using the ASPIDA approach. In case A, the AV provides real-time data and the state to the CU. Along with the location, route-paths, service-bulletins, stop-list and the current route state (e.g. running state, idle state, near finish state) that can be collected by the M2M devices through the M2M network (i.e. wireless sensor network), the CU records entering-AVs (i.e. operate the next itinerary) and leaving-AVs (i.e. to be substituted by another one) to leverage efficient AV services.

In case B utilising the policy-aware SOA ASPIDA, multiple entities enable the secure identification of the AVs and enhance the security based on digital identity assertions. The AVs exploit digital certificates in order to be authenticated and securely registered into the service. The message exchanges are securely protected with the encryption security mechanisms and the solution enables to securely store, retrieve, analyse and integrate data.

Function 2: Intercept service request

The service consumer submits a service request once authenticated and logged in an application user-interface. Then, a processing unit (PU) computes the necessary service levels and communicates the service delivery details back to the application. The authorised users are able to securely address the service request to the PU and enforce the new values/policies to the system-parameters according to the policy dataset. More precisely, the authorised service consumers connect to the service request module and submit the request over a secure channel. Then, the request is intercepted by the context handler to create a native XACML context object suitable for the PDP to be processed by converting the request from the native canonical representation into the appropriate XACML format. After the service request has been successfully intercepted, the request is forwarded to the service provider and processed accordingly. Typically, PEP is an integral part of the requester in terms of the XACML glossary. However, given the sensitive nature of the access control systems and the requests, it is vital that the PEP receives requests only from authenticated identities. Figure 36 illustrates how a transit passenger (i.e. a person with special needs) can submit a service request in a campus by utilising M2M location and communication technologies.

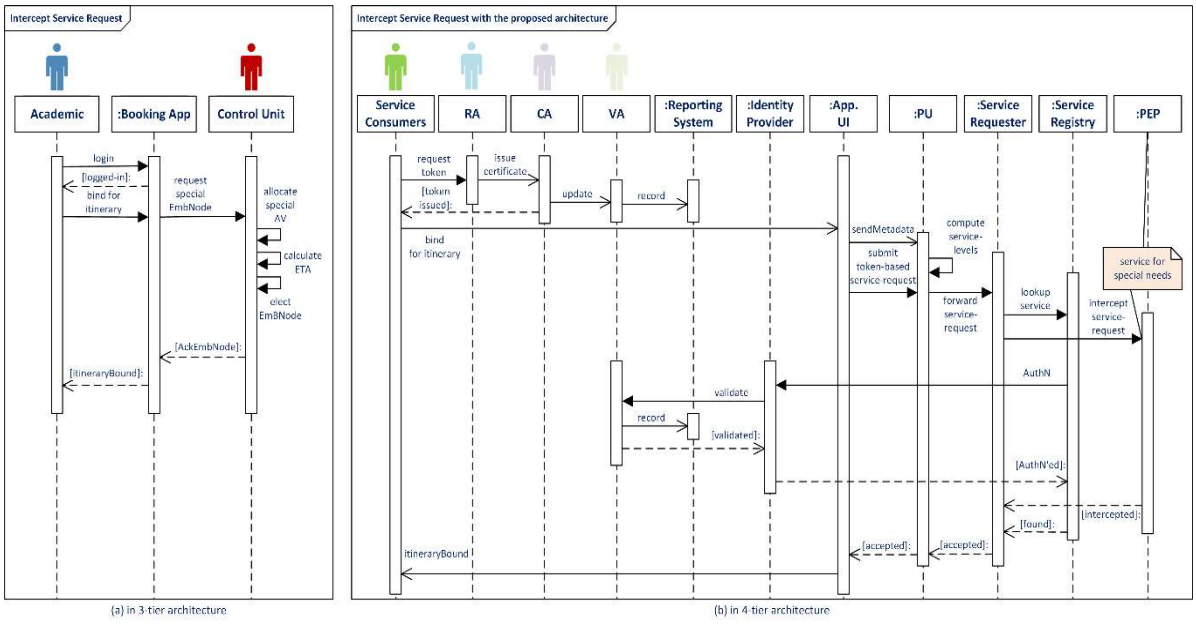


Figure 36. Intercept a service-request

The ASPIDA model can be further reinforced with stronger security controls in the context of secure communications by using X.509 certificates, Kerberos tickets or SAML assertions between the web-services.

In case A, the academic passenger logs into the information system (i.e. a booking application) and binds for a fitting itinerary. The application collects the sensory data (i.e. consider passenger's special needs) and forwards the request (i.e. for a special AV with a mechanical wheelchair ramp). For this purpose, the CU incorporates in the solution positioning M2M devices and applications for M2M communication with regard to allocate the suitable AV. Additionally, the CU performs calculations for the expected arrival time, while the node is elected according to the passenger's location. Finally, the user receives a confirmation about the embarking process and the estimated arrival time of the suitable AV to meet the service needs.

In case B, following the successful passenger's authentication, a security token mechanism is utilised in order to enable the policy-driven secure identity propagation and the token exchange between the respective web-services. This enables the secure communication between the web service client and the web services using X.509 certificates. The service consumer signs the request with the sender's private key and later the XML signature is verified with the sender's public key, which is available in public repositories.

Comparison

The standard iBuC operational scenario is compared with an extended scenario enhanced with SOA and policy enforcement aiming at improving the security, information flows and the service composition. In overall, the extended policy-aware SOA approach offers several enhancements and improved information flows. First, the model-driven development facilitates the business intelligence and optimizes performance functionality. With the policy-aware model, the certificate issuance policies can be used for a qualified subordination between different cryptography hierarchies (i.e. recognise certificates based on Elliptical Curve Cryptography algorithms by another CA that meet the certificate issuance requirements). Along with the ability to expose the functionalities, the extended model provides additional levels of security and robustness. The *audit records* are kept in the reporting system, which allows the construction of a comprehensive range of *traceability* and *serviceability* queries. The distribution of the public key objects in public certificate repositories enables the authenticity of the digital identities by other web services and results in increased data integrity besides privacy. For example, the certificate trust lists (CTLs) and the root certificates in the Trusted Root Certificate Store are needed to call a security-enhanced web service to authenticate the applications (i.e. the booking/reservation applications). Eventually, a verified and unique digital identity ensures that legitimate access is granted in the service registry.

In order to evaluate ASPIDA as presented above with the iBuC scenarios, the 3-tier ETSI functional architecture and the ASPIDA approaches are contrasted for the different sets of the aforementioned functions. The ASPIDA approach achieves higher *manageability* by enabling an easier management of the existing functionalities along with an overall increase in the development speed. The ASPIDA optimizes the development efforts of different use case configurations (i.e. AV initial setup) and complex event processing to adapt to a range of factors and environments (i.e. moving to another place with diverging climatic and environmental conditions). Additionally, the ASPIDA supports *message transformation* with

data translation from the canonical representation into the technical form, which simplifies the software structure needed for the implementation of the model. Not only the combination of various *service metadata* is supported, but also the *flexibility* and the *integration capabilities* can be reinforced with other service modules with data sharing. By incorporating a policy rule validator, the contradicting and conflicts can be resolved to avoid *policy violations*. The validator can detect non-conformities and violations in order to ensure policy consistency and checks the data quality of the configured policies for missing and invalid data information. Various requirements and service levels can be encoded with clear service descriptions for *service reusability*.

All the above enhancements along with the application security and management policies show the added-value of the ASPIDA architecture. In terms of security improvements, there is enhanced message security with respect to data *confidentiality* and *integrity*. The *data-origin authentication* guarantees that the message has a distinct origin entity whilst entity authentication facilitates the communication entity to prove its legitimate identity. The M2M communications rarely rely on static node infrastructures and, thus, these communications require secure transport layer protocols (e.g. TLS over TCP, SOAP over UDP, SOAP over http/TCP) in regard to reliable and secure end-to-end communication paths. The access to *sensitive data*, the *arbitrary operations* and the entity authentication can be imposed with policy-based methods to impose a more secure and strict framework for the delivery of the M2M services. Altogether, Table 10 depicts a comparison between 3-tier ETSI architecture and service-oriented SeMMA regarding their policies, mechanisms and concerns in the field of security.

Table 10. Comparison of security aspects, mechanisms

	Function 1		***Function 2***	
Security aspects	ETSI M2M	ASPIDA	ETSI M2M	ASPIDA
Policy violations	×	√	×	√
Access to sensitive data	×	√	√	√
Arbitrary operations	√	√	√	√
Integrity	×	√	×	√
Confidentiality	×	√	×	√
Trust	×	×	×	×
Key management	×	√	×	√
Data-origin	×	√	×	√
Entity authentication	×	√	√	√
Privacy	×	√	√	√
Data integrity	×	√	×	√
Device integrity	×	√	×	√
Secure communication protocols	√	√	√	√
Policies and low-level mechanisms	×	√	×	√
Cross-domain policy issues	×	×	×	×
Security Mechanisms	ETSI M2M	ASPIDA	ETSI M2M	ASPIDA
Audit records	×	√	×	√
Logical policy enforcement	×	√	×	√
Information flow tracking	×	√	×	√
Asymmetric/Symmetric keys	×	√	×	√
Authentication at physical layer	×	×	×	×
Identity Based Cryptography	×	×	×	×
Ephemeral identity (pseudonym)	×	×	×	×
Event-based policy enforcement	×	×	×	×
Security Policies	ETSI M2M	ASPIDA	ETSI M2M	ASPIDA
Decisions for authorisation	×	√	×	√
Operational and obligation policies	×	√	×	√
Logic-based context-aware policies	×	√	×	√

The use of cryptography and encryption methods improves the *data confidentiality* and *integrity*, the *data-origin*, the *privacy*, the *data integrity* and *trust*. The *key management* also enables more *secure communication protocols* and the *device integrity*.

6.3.3 Setup of authorisations

In contrast to Fysarakis et al. [103] where the authors describe the information flows among PDP, PIP, PEP and the protected resources for a cross-domain resource sharing for smart environments, ASPIDA presents an integration between the policy-based management services and the capability-based access control system. A dynamic policy-driven authorisation scheme and prototype are presented in [223] where capabilities tokens are issued to improve the security. In more detail, the authors present the token authentication challenges supported with access policy retrievals and revaluates the policy conditions at runtime aspiring to enforce the corresponding effective policy decision. Additionally, an in-depth look into adopting a µSservices-based architecture is taken, so that each one of the services can be deployed independently. The services can be deployed by independent components such as M2M discovery, lightweight messaging (i.e. Kafka), caching, load-balancing/scaling and publish/subscribe µSservices. In ASPIDA, the access requests can be served flowing through decoupled µServices, as light weight containers are getting traction to develop API services compared to the heavy Java EE and Spring frameworks. On the policy-domain side, PDP listens the environment for any applicable updated policies based on the information retrieved by the PIP/PAP. Finally, the PDP invokes the response and publishes the result via PEP.

In order to implement such a distributed capability-based access control approach, CBOR can be used for the representation and the formatting to represent the capability token, because of its suitability in constrained environments. The CWTs can be used in the COAP responses enabling fine-grained access control information within the token. By utilising CWTs in the ASPIDA scheme, the information is exchanged along with the identities in a secure and protected way that is not supported in other cases. For instance, when the TLS mutual authentication is used, then the originator's identity passes at the application level. The CWTs also support a scalable and stateless authentication method where the validity period of the short-lived access certificates is examined carefully considering numerous factors, such as the *data classification*, the *risk associated with the data*, the *data exposure* to the public and the *environmental conditions*. Even though the devices can work with the same topic space and among various OAuth grant types [227], the ASPIDA scheme supports the password, the symmetric-key grant, the token grant type as the client may have limited interactions and the client credentials grant. In comparison to these types, using the HTTP basic authentication to protect the API, a key and a capability token are created during the generation of a new API. After sending the request, the authorisation of the capability token from the authorisation server is allowed. Next, the capability token is returned in the response, if the application identity is authenticated and the authorisation grant is valid. For instance, the sample payload of a JWT is shown in Figure 37.

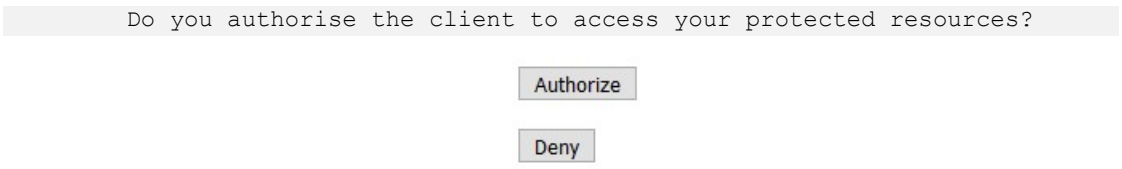


Figure 37. OAuth approval request

Afterwards, the capability token is returned in the response, if the application identity is authenticated and the authorisation grant is valid. The sample payload of the JWT is shown in Figure 38.

JWT response

```
{
 "aud": "Sensor01"
 "user_name": "user",
 "scope": [
 "read",
 "trust"
 ],
 "exp": 1518074605,
 "authorities": "ROLE_USER",
 "jti": "1d3b89..",
 "client_id": "ASPIDA-client "
}
```

Figure 38. The JWT registered claims set of a PDA request

In contrast, in the case of CWT, the client credentials grant uses client-id and client secret in the request payload. Nonetheless, the scheme can be further extended with certificates or Datagram Transport Layer Security (DTLS) pre-shared keys. Selecting the proper grant type depends upon the use case and the specific application type given that the client and the resource server support the proper security encoding and attributes. Other parameters weigh in as well, like the level of trust for the application, or what the users should be able to experience. In Figure 39, the payloads for diverse grant types using the REST APIs are presented; first the response containing an access token bound to symmetric key where transport layer security is with CBOR encoding, second the token request and response using client credentials with CBOR encoding and last the token refresh grant type grant where COSE is used to provide object-security.

CWT - Symmetric key

```
Header: Created
 Content-Type: "application/cbor"
 Payload:
 {
 "access_token": b64'eyGhbJciOi..'
 "profile": "coap_dtls",
 "expires_in": "3600",
 "cnf": {
 "COSE_Key": {
 "kty": "Symmetric",
 "kid": b64'44Gkam',
 "k": b64'JSU0ExXzUi..'
 }
 }
 }
```

CWT- Client credentials

```
Header: POST
 Uri-Path: "token"
 Content-Type: "application/cbor"
 Payload:
 {
 "grant_type": "client_credentials",
 "client_id": "ASPIDA-client",
 "aud" : "Sensor01"
  }
```

```
Request-Payload:
{
 "grant_type": "client_credentials",
 "aud": "Sensor01",
 "client_id": "ASPIDA-client ",
 "client_secret": "secret"
}

Response-Payload:
{
 "access_token": b64'eyGhbJciOi..'
 "token_type": "Bearer",
 "csp": "DTLS",
 "cnf": {
 "COSE_Key": {
 "kid": b64'd30tZS..',
 "kty": "oct",
 "alg": "HS256",
 "k": b64'JSU0ExXzUi..
 }
 }
 }
```

CWT- Token refresh grant type

```
Header: POST
 Uri-Path: "token"
 Content-Type: "application/cose"
 Decrypted payload:
 {
 "grant_type": "client_credentials",
 "client_id": "ASPIDA-client",
 "cnf": {
 "COSE_Key": {
 "kty": "EC",
 "kid": h'11',
 "crv": "P-256",
 "x": b64'trxcq..',
 "y": b64'Qwebq..'
 }
 }
}
```

Figure 39. CWT Requests for an authorisation

6.4 Performance evaluation

6.4.1 PBMN evaluation

The objective of the experimental setup is to provide a deterministic best path selection by applying routing adaptations based on the dynamic changing conditions of the network entities. The system architecture can provide advanced monitoring options such as CPU utilisation, memory thresholds, neighbour discovery, routing changes, identity events (i.e. failed authentications through 802.1), MAC address table events as well as several other options. Even for these advanced options, the appropriate action routines can be enforced through automated scripts loaded in the PBNM system. In more detail, when certain thresholds are exceeded (i.e. latency, throughput), the SLA monitor needs to monitor the status according to the SLA Key Performance Indicators (KPIs). Based on the decisions by the PDP to adapt the data flows and update routing, the related subsequent actions can either improve the performance, or resolve the connectivity problems. However, in case the

proposed adaptive routing deteriorates the overall performance due to suboptimal routing, then PDP reverts the original configuration through the evaluation of the appropriate KPIs.

First, the IGP routing protocol performance is examined by analysing the link state routing protocol. OSPF without routing protocol extensions is used within the large autonomous system network of SYZEFXIS. The test results using OSPF show that the path selection is indifferent to the traffic load and the network conditions. The enhancements to the routing protocol extensions are achieved through the proposed PBNM approach between the bandwidth capacity (c_i) and the reserved bandwidth of the link (r_i) by achieving the minimum value min(c_i-r_i). The improvement is proportional to the c_i/r_i ratio during the path computation process. In ASPIDA, it is possible to trigger applets and the PBNM can apply a new policy which influences the routing path selection. In this case, a different path can be selected by configuring the routing protocol metric accordingly. The optimal path can be selected based on the traffic load. This is achieved by applying a policy control which modifies the bandwidth of the links, so that the best or an alternate path towards the destination is selected. Nonetheless, this is likely to result in unequal routing costs towards the destination. For instance, assuming the two endpoints (n_i, n_k), there is an initial path computation $n_i \rightarrow n_j \rightarrow n_k$. By increasing the traffic rate between the two endpoints gradually, once the SLA criteria are violated (e.g. the packet drop exceeds 2%), the event detectors trigger the appropriate applets and the reserved bandwidth r_m of an alternative path is increased. This results in a new optimum path between the two endpoints (n_i, n_k) and the routing protocol re-computes the path $n_i \rightarrow n_m \rightarrow n_k$ and a new path is selected. The traffic is now directed over the new selected path, on the condition that the SLA thresholds are met and the monitored metric values are below the defined thresholds.

Notably, before carrying out the experiments, several tests are required to identify the proper settings for the SLA criteria and thresholds. If the threshold limits are set too high, the event detectors cannot trigger the policy rule and QoS will not have the expected effect on the adaptive routing. On the contrary, if the values are set too low, this can lead to under-performing conditions and frequent adaptive routing changes, which can lead to routing instability and connectivity problems, longer delays and a number of other anomalies in the exchange of the routing information. Once the traffic monitoring criteria meet the SLA thresholds for a sustainable period, the configuration changes can be rolled back enabling the accomplishment of the optimum path selection. Furthermore, more sophisticated adaptations of the path computation process can be provided, as the SLA thresholds can be adapted based on the traffic patterns (e.g. if the traffic exceeds 80% of the access rate, then the packet drop rate threshold can be easily adapted to 4% from 2% being considered as a saturated link). This illustration demonstrates the dynamicity of the proposed framework and the capability to influence the path selection process via the policy rules enforced by the PBNM system.

In the experimental setup, the analysis of the IGP routing conditions with varying traffic patterns, link loads and traffic rates over the 2 Mbps, 8 Mbps and 34 Mbps links of SYZEFXIS network is performed. When the SLA thresholds are violated, then EEM notifies PBNM, which instructs PEP to apply a policy to remedy this violation. Apart from traffic measurements, accounting and network data planning, the architecture can also dynamically alleviate and adapt to certain security incidents or threats. For instance, in the case of an ICMP flooding scenario, whenever a timeout occurs due to congestion reasons or due to an inability to

response within the timeout period, an operation return code reports *no connection*, *busy* or *timeout* conditions. After that, EEM can trigger the actions to enforce the appropriate policy to control and rate-limit the ICMP traffic.

During the experimental setup, a Two Rate Three Color Policer (trTCM) [251] is defined. The trTCM is used as a component in the traffic conditioner and the traffic is coloured depending on whether it exceeds the CIR. By configuring the SLA monitoring with the use of IP SLA responders, the UDP echo operation can be used to monitor the end-to-end response times between the network devices. This results in active traffic monitoring, while generating traffic in a continuous, reliable and predictable manner for measuring the network performance. In order to measure the response time between the source IP SLA device and the destination IP devices, the maximum time for SLA operation to complete (e.g. the timeout waiting for probe response), the boundary value measured over the operation, (e.g. the Round trip time (RTT), the jitter value collected during the operation) and the frequency of the operations are all set to two seconds. An example of SLA conformance statistics and operation return code values is given in Table 11.

Table 11. SLA conformance example

Round Trip Time (RTT)
Latest RTT: 1 millisecond
Latest operation return code: OK
Number of successes: 10
Number of failures: 0
Operation time to live: Forever

Table 12 presents the relevant statistics, when the SLA thresholds have been violated.

Table 12. SLA violation example

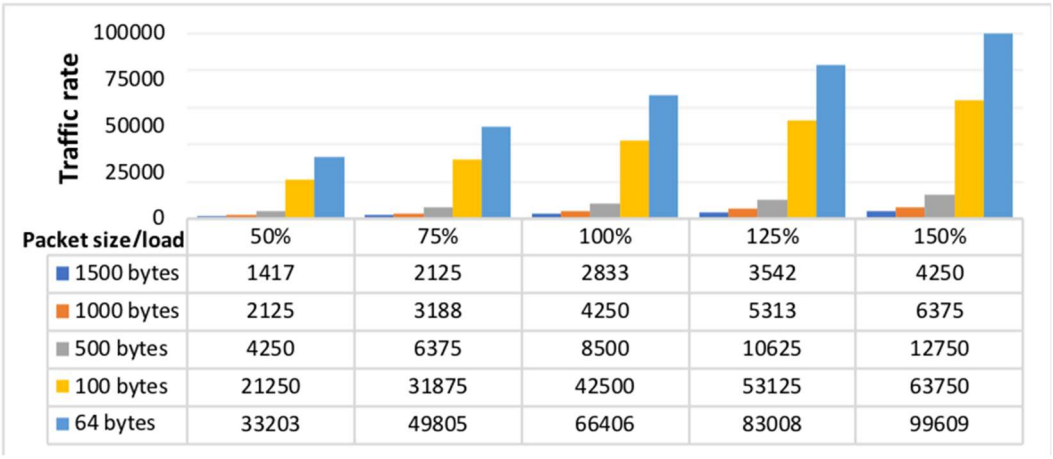
Round trip time (RTT)
Latest RTT: 36 ms
Latest operation return code: Over threshold
Number of successes: 25
Number of failures: 6

With regard to perform these measurements, the traffic generator sends streams of traffic to exceed the specified traffic thresholds. The data equivalent is generated to a percentage of the traffic thresholds for all diverse types of circuits. The data rate is configured via an inter-packet gap adjustment set to 1, while incorporating EDs through EEM applets. These applets access and configure global variables, so that a new policy rule can be applied to the PEPs. The applet introduces QoS traffic shaping to conform to the rate expected by the network based on the dynamic and changing conditions. When the transmitted traffic exceeds 10% of the traffic thresholds, then the transmitted traffic is associated with a specific policy profile to eliminate bottlenecks and avoid any packet drops.

Table 13. Sample applet for traffic monitoring

applet Action 1
event interface name "FastEthernet0/0" *parameter* txload *entry-val* 25 *entry-op* gt *entry-val* is *increment* false *poll-interval* 60
command "service policy output EXTRA CONTROL"

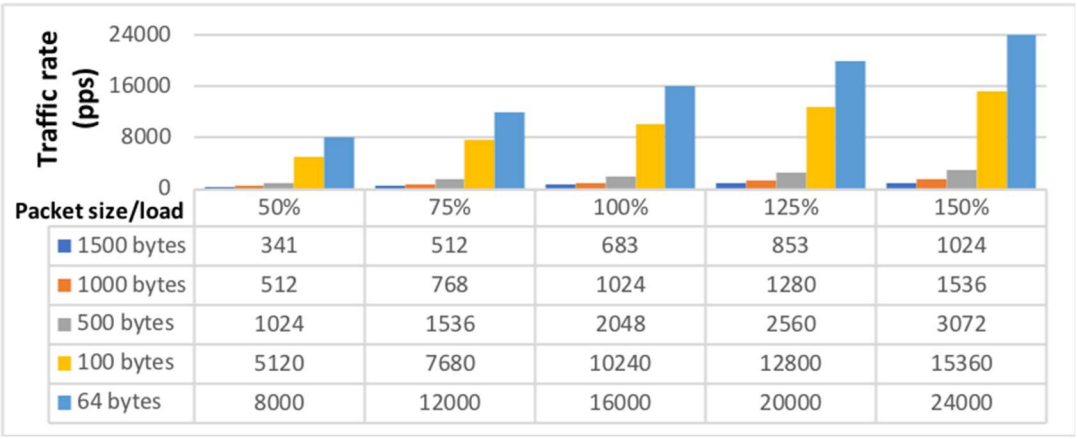
The experiments are repeated for varying packet sizes and different percent of traffic load values in order to analyse the IGP routing conditions. Thus, an analysis of the average traffic rate of the circuit bandwidth is conducted over the 34 Mbps, 8 Mbps and 2 Mbps links. The transfer and the drop rate were examined, when transmitting at the maximum sustainable traffic rate. The vertical axis of Figure 40 presents the respective traffic rate, when UDP echo packets are used, whereas the horizontal axis of the same figure presents the traffic load for varying packet size from 64 bytes to 1500 bytes over 34 Mbps links.



Packet size/load	50%	75%	100%	125%	150%
1500 bytes	1417	2125	2833	3542	4250
1000 bytes	2125	3188	4250	5313	6375
500 bytes	4250	6375	8500	10625	12750
100 bytes	21250	31875	42500	53125	63750
64 bytes	33203	49805	66406	83008	99609

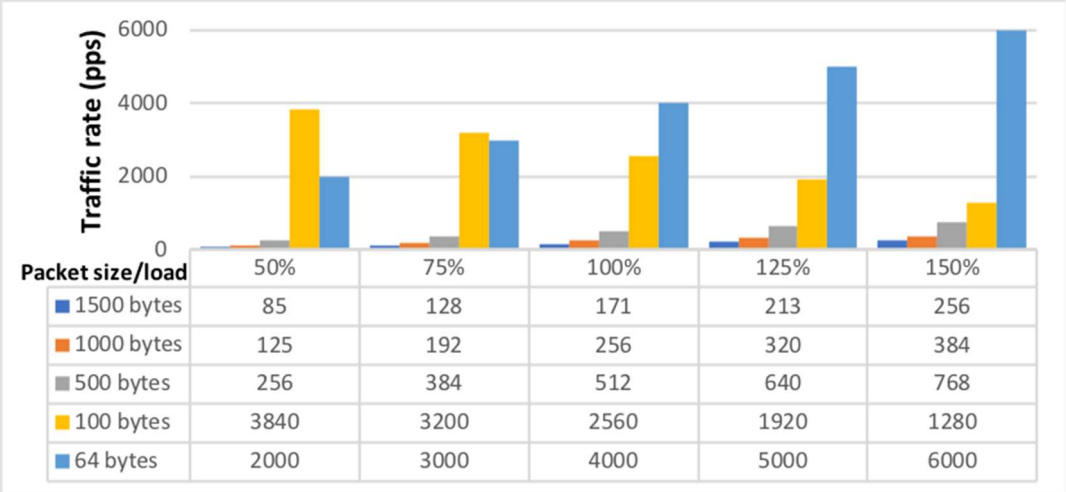
Figure 40. Traffic rate for UDP echo varying packet size with differing loads-34Mbps

Figure 41 presents the traffic rate over 8 Mbps links, while Figure 42 depicts the traffic rate over 2 Mbps links. The experiments are repeated for these three distinct types of links (2 Mbps, 8 Mbps and 34 Mbps) with varying packet sizes. During these tests, increasing traffic is gradually generated and sent over the links. The goal is to increase the load of the link to identify the maximum sustainable transfer rate limit (e.g. 100%) of the link, before violating the established SLA thresholds and the corresponding timeouts. The generated traffic is measured in packets per second (pps).



Packet size/load	50%	75%	100%	125%	150%
1500 bytes	341	512	683	853	1024
1000 bytes	512	768	1024	1280	1536
500 bytes	1024	1536	2048	2560	3072
100 bytes	5120	7680	10240	12800	15360
64 bytes	8000	12000	16000	20000	24000

Figure 41. Traffic rate for UDP echo varying packet size with differing loads - 8Mbps



Packet size/load	50%	75%	100%	125%	150%
1500 bytes	85	128	171	213	256
1000 bytes	125	192	256	320	384
500 bytes	256	384	512	640	768
100 bytes	3840	3200	2560	1920	1280
64 bytes	2000	3000	4000	5000	6000

Figure 42. Traffic rate for UDP echo varying packet size with differing loads - 2 Mbps

Regarding the low speed serial 2 Mbps links representing the majority of the SYZEFXIS connections, when the links are saturated, there are excessive packet drops, because any type of packet the traffic exceeds the nominal access rate. In order to alleviate the heavy traffic conditions, there should be either additional link aggregation or advanced QoS mechanisms for excess traffic implementation. The experiments for the traffic rate for low speed serial links are presented in Figure 42. Moreover, the transfer and the drop rates are both analysed when transmitting at the maximum sustainable traffic rate for examining and analysing the non-conforming traffic.

The same type of packets is used to obtain comparable results for both OSPF and EIGRP experiments. On the whole, namely the 34 Mbps links perform well, as there are very low probabilities for any drops, delays and underperformed conditions, because the traffic over-exceeds the access rate. Precisely, the traffic does not saturate the bandwidth of the high-speed links. Likewise, no packet drops can be found for the 8 Mbps links, except when transmitting 1500-byte packets and the utilisation of the 8 Mbps link exceeded 100% of the access rate. In the case of 100% utilisation, the offered rate is higher than the sustainable traffic rate, when the average drop rate is of 269 kbps. The drop rates increase for the 8 Mbps at 1500-byte packets and for 2 Mbps link for 1500, 1000 and 500-byte packets, which indicates that the packet quantization size is significantly important. The necessary traffic is generated based on various load utilisation of each circuit access rate. Remarkably, the impact on the drop rate is based on different traffic rates, instead of the total number of packets. Fewer packets may not saturate the links provided that the necessary conditions are met (e.g. ingress buffering, the specific integrated circuits in use, the network entities feature and capabilities). This approach can be used to simplify the network operations by selecting the appropriate set of service levels. Several more tests are conducted with a varying packet size for UDP-echo packets up to the 150% over the 2 Mbps links to analyse the performance and the packet drop rate. In this context, Figure 43 presents the drop rate, when OSPF is used.

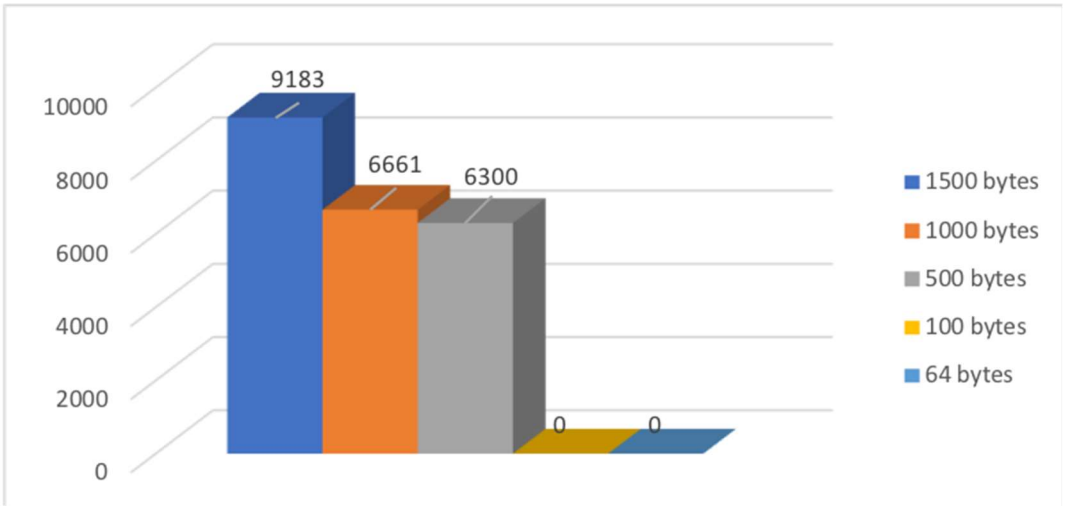


Figure 43. Drop packet rate for OSPF for varying packet size over 2 Mbps links

The transfer rate can be increased by matching the access rate with the actual traffic needs and thresholds. The algorithm uses incremental steps for the bandwidth increase. If the drop rates are not zero, then the bandwidth is further increased till all the drop rates become zero. The bandwidth increase steps are incorporated into the PBNM with the appropriate EDs. The following tests present the performance analysis of distributing traffic over two unequal cost links of an access layer node with a varying packet size. The cost of the link is calculated using the routing protocol metrics for each case. In the first scenario, when OSPF is used as the IGP to connect the network entities, the total OSPF cost is the accumulated associated cost of each interface from the source to the destination network. The total OSPF cost is indifferent to the network conditions and there is no change of the traffic rate or any alternate best path computation towards the destination. Upon the completion of the initial best path selection, no further routing protocol computation takes place, even during periods of saturated links and transmission over the maximum sustainable transfer rate. Contrary to the first scenario where OSPF is used, the second scenario (EIGRP) has different results with various utilisations and loads over the links. The composite metric used by EIGRP considers the load of the link and the configured bandwidth, which can influence the routing path selection. Each time the metric is re-calculated resulting in a new optimal path. If the output drops are a consequence of short bursts of data, then the line speed increases and the output drops can be avoided. The QoS adaptive routing can preserve the necessary bandwidth or even provide a better alternate path that improves the overall performance by forwarding traffic more efficiently. The deviation ratio λ is derived by:

$$\lambda = 1 - d(n_1, n_k) \,/\, t(n_1, n_k) \qquad (6.4.1.1)$$

$$t'(n_1, n_k) \leftarrow \lambda \; t(n_1, n_k) \qquad (6.4.1.2)$$

where $d(n_1, n_k)$ is the drop rate between the two endpoints n_1 and n_k. The goal of the algorithm is to find the largest value λ that complies with the SLA criteria and thresholds. Given that SYZEFXIS network supports only OSPF as the IGP routing protocol, the EIGRP performance analysis is simulated under the same experimental conditions as for the OSPF's measurements. The outcome of the deviation ratio λ for varying packet sizes for EIGRP is shown in Figure 44.

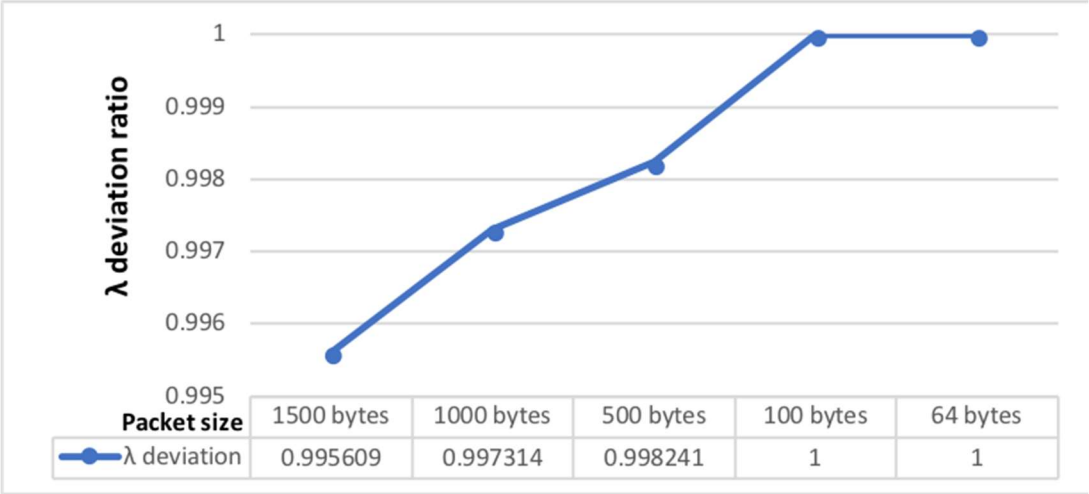


Figure 44. Deviation ratio λ for varying packet sizes for EIGRP over 2 Mbps links

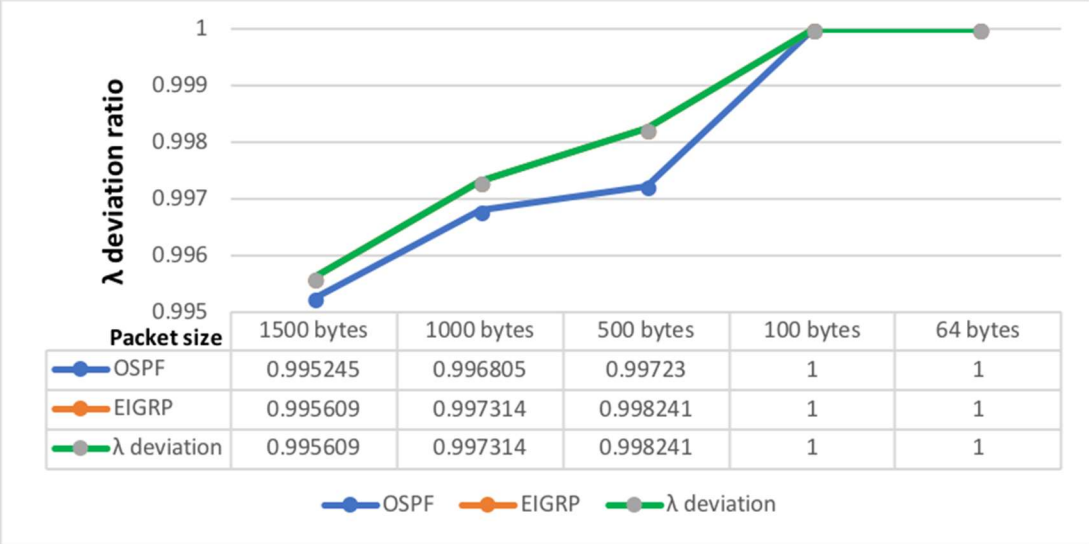


Figure 45. The deviation ratio λ between EIGRP and OSPF over 2 Mbps links

The improvement of the network resources reallocation when using a composite routing protocol metric instead of the link state metric is depicted in Figure 45. Hence, it can be derived that EIGRP has an improved performance compared to OSPF since it supports load sharing over unequal cost paths based on the composite metric for each link and the dynamic nature of the EIGRP composite metric. Therefore, EIGRP has two significant advantages compared to OSPF with regards to adaptive routing. EIGRP is more adaptive to the dynamic conditions of the network and supports a composite routing metric, in contrast with OSPF. EIGRP also supports unequal cost load balancing and, two different unequal cost paths can be tracked between two endpoints. OSPF does not support this functionality, whereas EIGRP achieves an improved utilisation of the network resources. Considering the amount of time to calculate the route, this impacts the routing convergence time and it is beneficial to avoid frequent and unnecessary recalculations, because of the topology changes. For this reason, EIGRP uses an algorithm based on the feasible successors in order to avoid any unnecessary recalculations. Furthermore, the scheme allows multiple network layer protocols to carry independent routing information.

Finally, several other tests are conducted with the EEM functionality by monitoring the interface threshold of the edge routers for a wide range of traffic patterns. Upon exceeding the defined thresholds, the EDs trigger the enforcement of the suitable policy providing alleviating actions for excess traffic and enforcing actions through applets. For instance, if the threshold value is greater than 5 (max. value is 255), then an extra service policy is applied through the appropriate actions. The EEM policy applies a traffic control service policy in the outgoing interface for Performance-based and Event-Management Routing (PEMR). It can either police traffic or provide QoS-aware mechanisms to cope with the excess traffic. During the first experimental (w/o PEMR), the TCP throughput is measured. Then, the tests are repeated using PEMR and the network performance is measured again. Figure 46 shows the improvements in TCP throughput by utilizing the adaptive event based routing with PEMR.

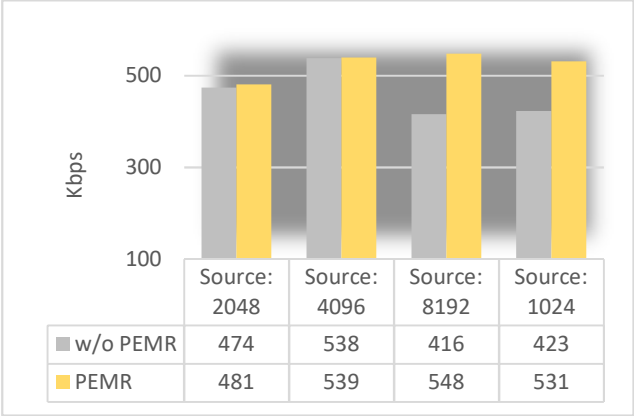


Figure 46. TCP throuhput for adaptive event-based routing

Figure 47 depicts the improvement by using a variable source and length of buffers read and written to the network.

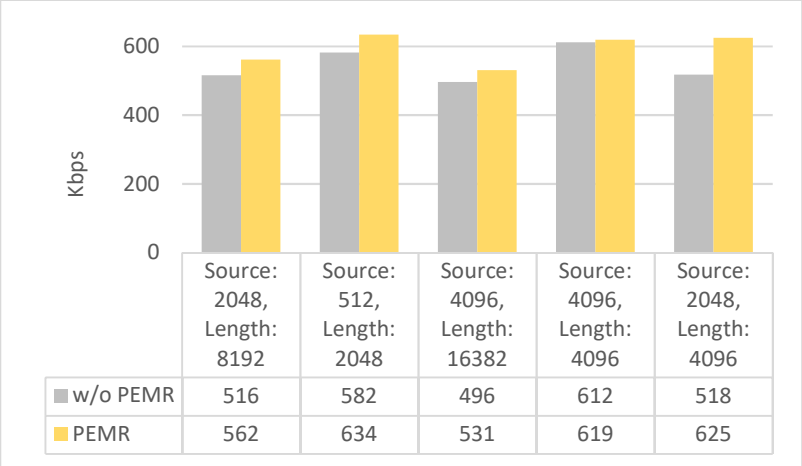


Figure 47. TCP throuhput for adaptive event-based routing with variable buffers

6.4.2 Access control evaluation

The evaluation of the ASPIDA model is based on the access control authorisation accuracy. The authorisation request results can vary depending on the authorisations. When granting the appropriate permissions and assigning the appropriate levels of authorisations, there are *Correct Authorisations* (COR) and *Mistaken Authorisations* (MIS) in the case of inappropriate access levels or mistaken policy decisions. Furthermore, *Lack of Authorisations* (LCK) exist when there is a lack of automation ending up in manual intervention and provisioning. Apart

from the above types of authorisations, the dynamic *Running costs* (RNG) along with the fixed *Setup costs* (STP) affecting the outcome have been considered in the model simulations.

Repeated random sampling is used to obtain numerical results and feed a Monte Carlo method [252], as the repeated simulations generate relevant outcomes. The method analyses the uncertainty propagation and evaluates iteratively a deterministic model using sets of numbers as inputs. The inputs are randomly generated from probability distributions to simulate the process of sampling from an actual population. Statistical sampling and simulations are performed to estimate the uncertainties and support the policy-based access control processes. In this context, Casassa et al. [253] explore the associated policy decision processes for user account provisioning and demonstrate how the system modelling and simulation activities can predict the impact of specific policies. Therefore, the modelling and simulation are complemented regarding the policy decisions processes for authorisations instead.

In the following simulations, Monte Carlo cases are tested with nominal and parameterized input values for policy-based authorisations for two cases. The first provides the modelling and prediction statistics for a standard access control system *Case #1* and the second for the ASPIDA model *Case #2*. Table 14 provides the simulation data points for the accuracy efficiency of the correct authorisations over the total number of authorisations.

Table 14. Nominal authorisation & cost values

Auth.types/costs	Case #1 – Std model	Case #2 – ASPIDA
COR authZ	100	100
MIS authZ	10	5
LCK authZ	5	0
RNG costs	10	2
STP costs	40	180

Aiming at determining the probability and stochastic distributions that may affect the processes and the outcome, the cumulative probability is estimated, as shown in Figure 48, despite the accuracy efficiency variance because of the uncertainty of values.

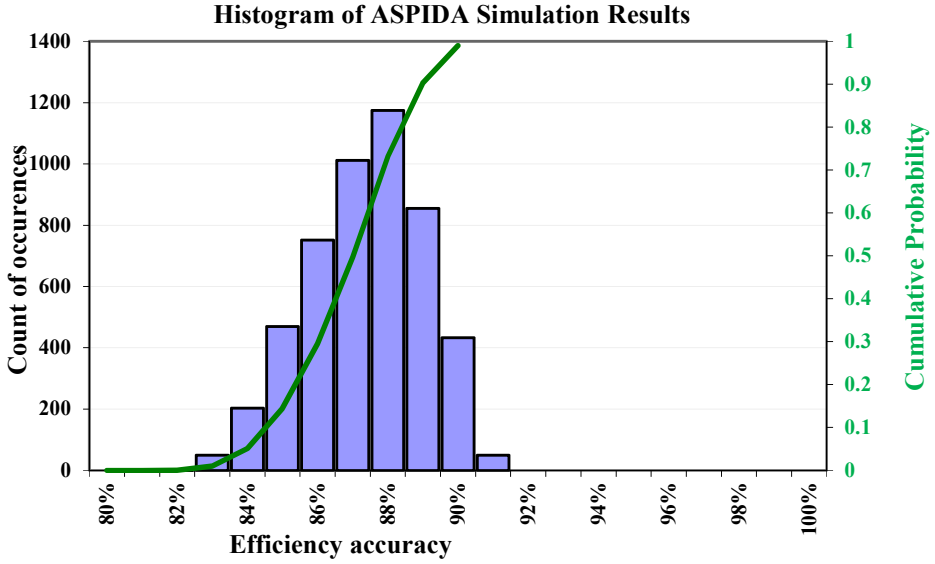


Figure 48. Statistical analysis of case #1

During the simulations, 5000 simulated observations are obtained to compare the two models and their respective statistics. The experiments can be reconfigured in a straightforward way by changing the simulated time-frame and/or the number of times a model needs to be executed. The analysis of the accuracy efficiency of introducing policies with the integrated access control modules of ASPIDA is depicted in Figure 49.

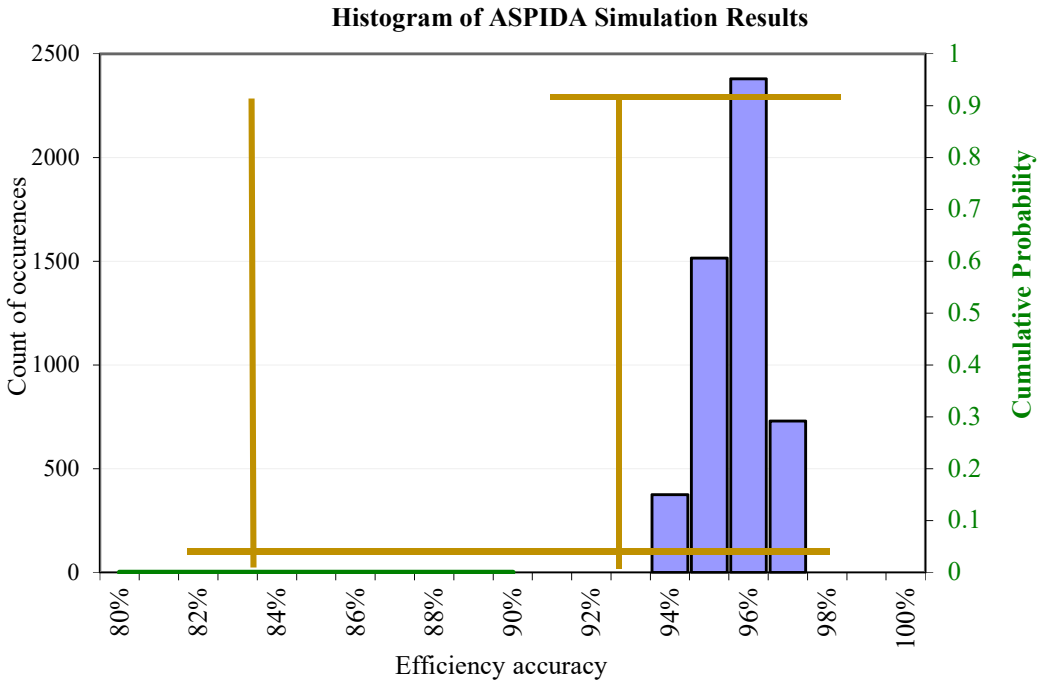


Figure 49. Statistical analysis of case #2

These simulations produce significant low-level measures and related high-level metrics. By introducing access control policies in the policy engine and by considering different automation cases, the access control authorisations play a significant role. The efficiency and the performance improvement for *Case #2* are presented in Figure 50.

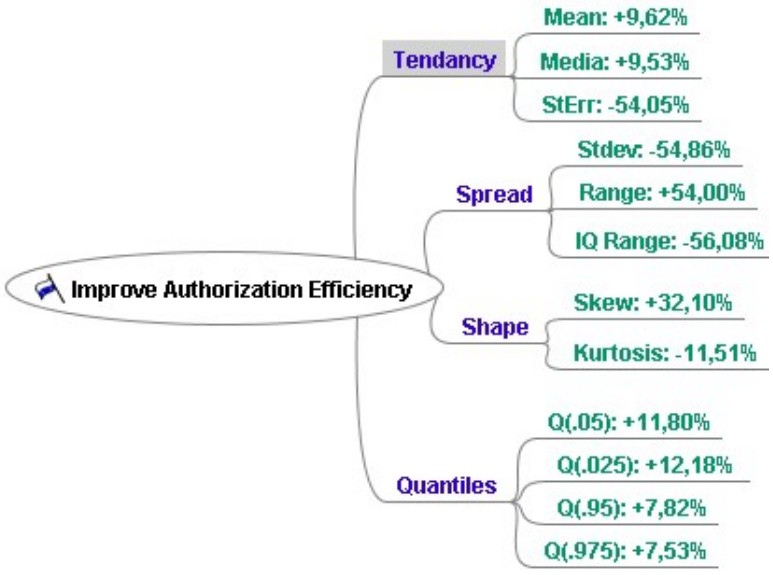


Figure 50. Comparison of statistics between *Case #1* and *Case #2*

6.4.3 Policy-driven authorisations evaluation

In ASPIDA, the access requests flowing through decoupled µServices can be served, as light weight containers are getting attention to develop API services compared to the heavy Java EE and Spring frameworks. In order to understand better the position of ASPIDA in comparison to other solutions, a comparative analysis of the ASPIDA and others is presented below. Notably, the developers are moving away from monoliths in favour of µServices that support continuous delivery and provide reusable, scalable, flexible and independent services. Various RESTful µServices frameworks have been proposed in Java and other languages [21], such as Wildfly-Swarm [254], Spark [255] and Spring-boot Tomcat [256], which bring several improvements like smaller start-up times and offer faster development along with built-in capabilities, as the µServices enable faster services and more efficient usage of the resources.

A prototype based on the open source µServices framework light-4j [257] has been developed to evaluate the characteristics of the ASPIDA model. A comparison analysis of the ASPIDA solution with the previous frameworks is performed to illustrate a performance comparison and depict the distinct types of authorisations. Various challenges that are faced in the effort to support the appropriate access control levels along with policy management capabilities are also presented. In the experimental setup, the authorisation µServices are deployed on a workstation with Intel Core i5-3230M CPU cores running at 2.60 GHz and 8 GB of memory with a 1 Gigabit network interface. With respect to evaluate the policy-driven authorisations, the proof of concept is based on µServices running in a Windows docker container using Docker v17.09.0-ce with the standalone Hystrix dashboard 1.6.0[18] to control the interactions between the dependencies and isolate the failures and latencies.

An increasing number of connections and threads are simulated by generating a significant load from the workstation with the aim to analyse the performance of ASPIDA. Figure 51 shows that APSIDA is the fastest among the other Java µServices frameworks, as ASPIDA requires the least time to process the number of requests and handle the threads accordingly.

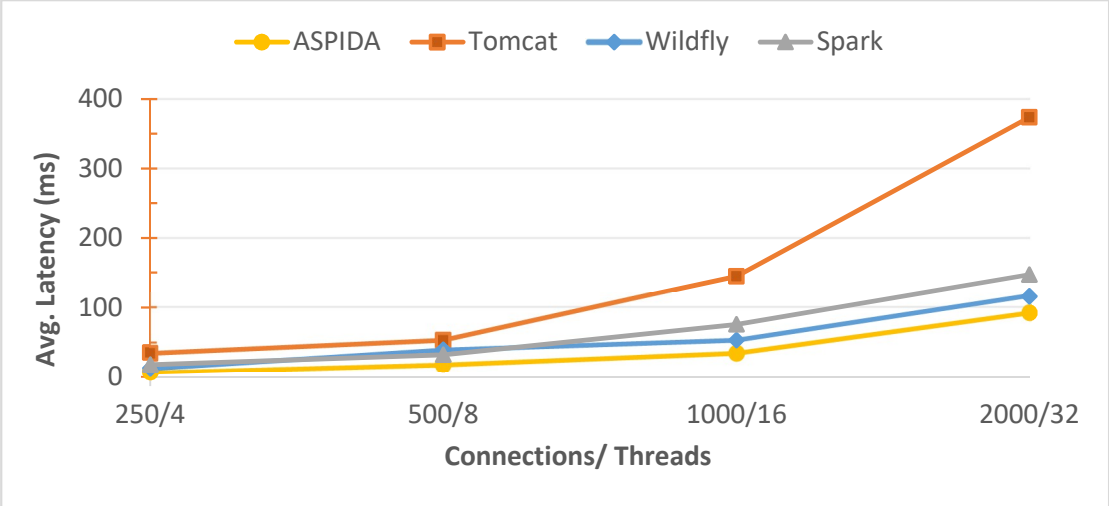


Figure 51. Performance evaluation of µService frameworks #1

Additionally, the respective data transfer rate is benchmarked for each of the four frameworks. Figure 52 illustrates that ASPIDA outweighs all others providing a higher data rate for a number of connections and threads.

[18] https://github.com/kennedyoliveira/standalone-hystrix-dashboard

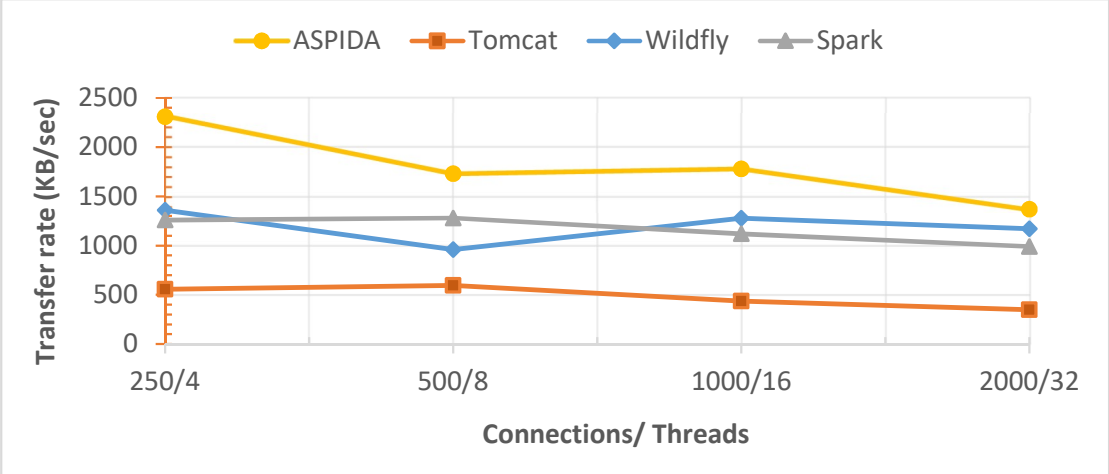
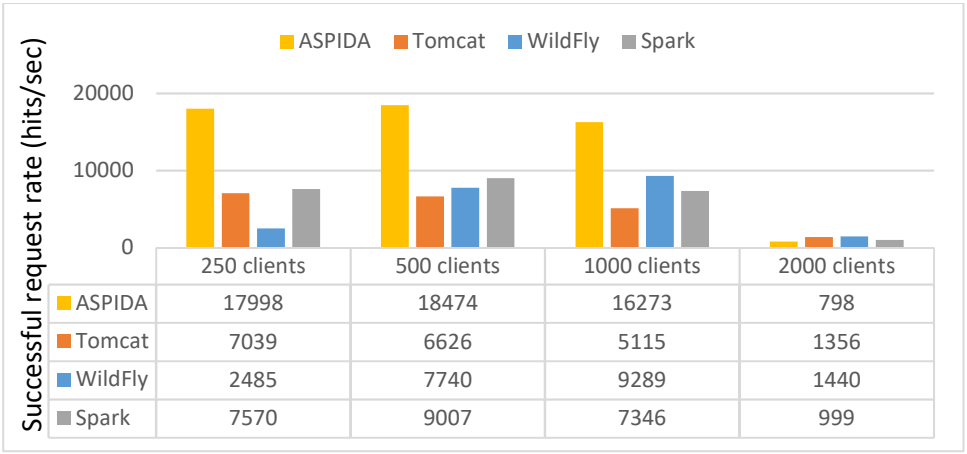


Figure 52. Performance evaluation of µService frameworks #2

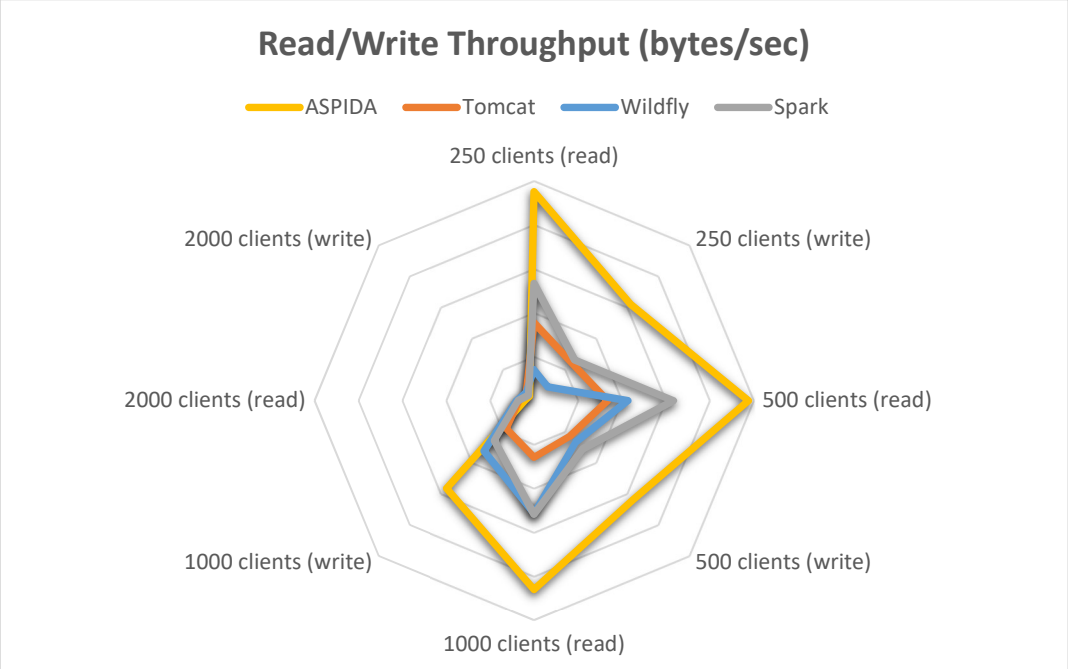
Next, the successful requests and the requests rate are measured as depicted in Figure 53. This presents the number of the executed requests produced per unit of time.



	250 clients	500 clients	1000 clients	2000 clients
ASPIDA	17998	18474	16273	798
Tomcat	7039	6626	5115	1356
WildFly	2485	7740	9289	1440
Spark	7570	9007	7346	999

Figure 53. Performance evaluation of µService frameworks #3

In order to be able to perform more detailed measurements related to the duration and the latency of the requests, the read and the write throughputs are measured for each framework as illustrated in Figure 54, which also depict better rate for the 250, 500 and 1000 clients, whereas the rate significantly drops for a higher number of concurrent connections for the ASPIDA scheme.



clients	*Op.*	*Wildfly*	*Tomcat*	*Spark*	*ASPIDA*
250	Read	352,838	888,601	1,334,201	2,379,022
	Write	222,095	608,028	654,159	1,552,325
500	Read	1,068,754	836,481	1,586,280	2,440,825
	Write	670,395	574,643	779,323	1,594,870
1000	Read	1,282,806	645,233	1,294,000	2,149,923
	Write	808,064	449,006	640,983	1,410,933
2000	Read	198,749	171,108	181,970	106,673
	Write	130,124	124,762	89,801	75,358

Figure 54. Performance evaluation of µService frameworks #4

Moreover, a histogram is created with requests duration times for analysing the number of user requests containing at least 1 event that exceeded the provisioned throughput in the selected period. The tests are performed using 2 CPUs and persistent connections for the duration of the tests, which last for 10 secs. This covers about 1 second of a 10 second test at 1000/requests per second and illustrates a load test with ~3.18% 'throttled' requests.

The histogram comparison in Figure 55 plots the frequency of the latency buckets in 7 classes, where each one has a class width of 1msec and each bucket upper bound is non-inclusive. Additionally, the results are redirected to an interactive plot.

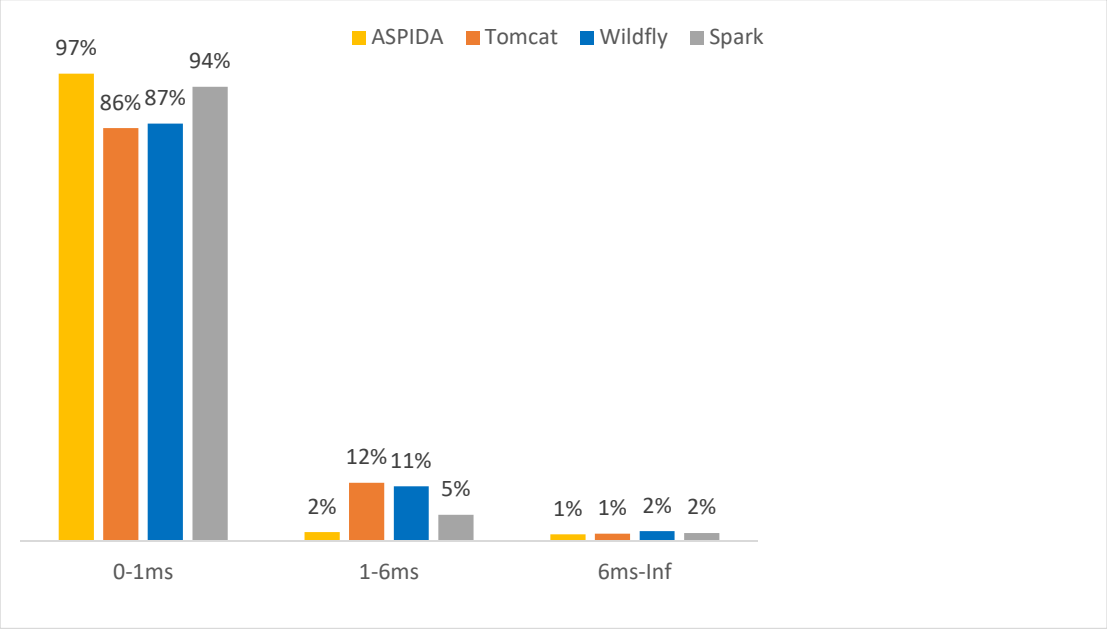


Figure 55. Histogram of the PDA for 1000 requests/sec rate

Figure 56 illustrates the analysis and visualization of the latency as a function of the requests using the ASPIDA scheme. Each point on the plot shows a request, the X axis represents the time at the start of the request and the Y axis represents the time taken to complete that request.

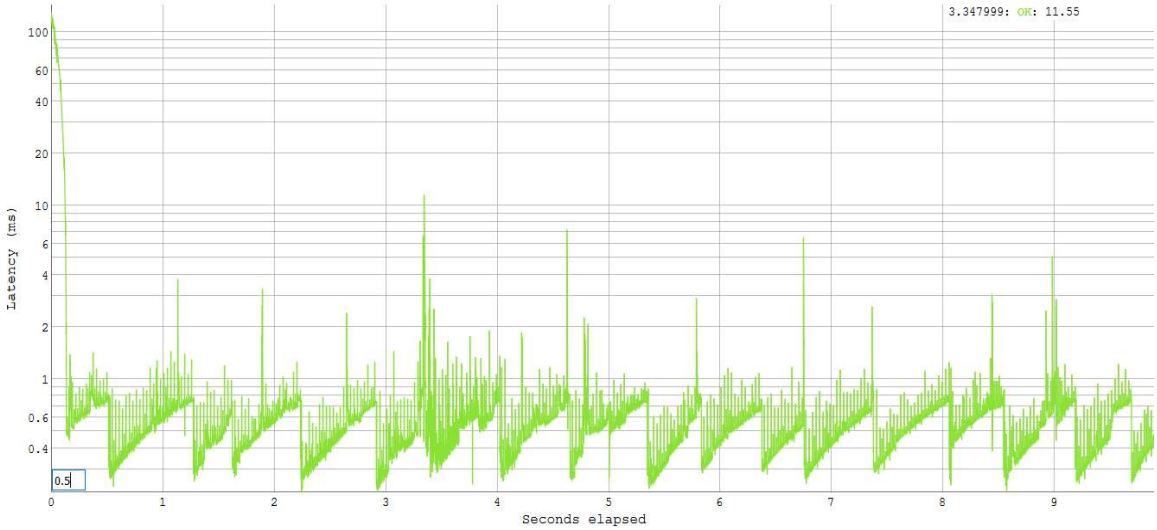


Figure 56. Latency as a function of the requests with 1000 requests/sec rate

To conclude, in comparison with other frameworks, the evaluation demonstrates that ASPIDA shows an improved performance in terms of the number of the concurrent connections, the throughput and the transfer rate. Figure 55 and Figure 56 illustrate that in ASPIDA with 1000 requests/sec rate most of the requests return at around or even under 1ms, which means that the throttle feature is successful. Even though the tests are performed with a high request rate, the request handler manages the requests efficiently and generates the responses timely, so that the threads are freed up faster to avoid any potential bottlenecks.

Chapter 7

Conclusions

7.1 Outline

This chapter concludes the doctoral dissertation and presents an *ex post facto* analysis of ASPIDA in comparison with other research activities by outlining the descriptive, exploratory and confirmatory analyses that have been conducted in the preceding chapters. The implications of the findings and the summary of the results to illustrate the numeric information are also presented in this chapter.

7.2 Overview

M2M communications can unlock opportunities in remote areas and generate new revenue, notwithstanding the fact that the privacy and security requirements may hinder the evolution and adoption of the related services. Thus, it is crucial to ensure trust and confidence in order to give access control to resources and secure the data efficiently. Nevertheless, the high number of interconnected heterogeneous devices raises several security and scalability issues, as the appropriate security countermeasures and policies need to be employed to achieve the appropriate levels of secure services. In order to tackle these challenges, the main goal of this doctoral dissertation is to propose, design and evaluate ASPIDA in the interest of an adaptive policy-based security management for M2M communications. Hence, several questions arose how to fulfil the security and the network requirements as well as enable the authorisation requests with policy decision criteria. This work has addressed some of these questions, as expressed below:

- As far as adaptive nature of the network domain to support SLA-driven adaptive routing selection is concerned, the PBMN system addresses the need
- The adaptive policy-driven architecture requirements are met to facilitate the service capabilities with capability-based access control means. The integration of the authentication and authorisation modules with the access control policies used by the application delivery platform and the service management disciplines is presented
- The representation, the relationships and the flows of the entities have been explained in a conceptual model by means of UML diagrams
- The ASPIDA scheme achieves higher throughputs compared with other contrasted solutions
- The granularity of implementing SOA solutions with the use of µServices architecture is also accomplished and increased

Vis-à-vis adaptive security management, this work demonstrates how to support the authorisation requests, the policy decision making, the policy enforcement and improve the security traits of the µServices to orchestrate the components across the services. The fundamental goal is to establish a general-purpose framework for M2M dynamic authorisations with policy-based access control capabilities that can be re-used to ensure the secure token-based authentications, even though most of the implementations may be dependent on the M2M resource-constrained devices and the µServices components, which

require complex interactions. Some of the foremost challenges in the context of adaptive security management are:

- A better adaptive security management with the use of the PDP, PIP, PAP along with the various PEPs to enforce the proper security controls
- A policy engine that combines the M2M service requirements and improves the security service levels of M2M communications
- The M2M network types with their diverging operational and architectural capabilities lead have introduced additional SOC blocks in the typical architecture
- The additional service domain facilitates procedures handled previously by the other domains to increase abstraction
- The access requests are processed through independent µServices with light-weight containers and data-driven µServices workflows

First, the use of policy-based management has several advantages in implementing adaptive M2M communications, where the physical topology can dynamically change in response to the mobility of the interconnected objects. A policy-based management approach ensures the enforcement of the security policies to mitigate the security issues, the risks and improves the security controls and data privacy.

The policy engine has access to the security policies plus the additional information and then applies the respective policies at fixed points in the device, network and application domains. Consequently, a policy engine can be used to address the security challenges of the M2M communications by relating the service requirements to provide the service primitives needed for service request/response. Additionally, the policy engine provides the capabilities that support the communication among the devices in heterogeneous environments, as predefined policies can be used for decision making in case of any service requests. For instance, in the case of a failure or a physical topology change, the policy engine can trigger an event-based policy action and immediately reconfigure the network or service components.

The establishment and the enforcement of clear and effective security and privacy policies are key issues to improving the offered service levels in M2M communications with the use of service-orientation principles. Therefore, policy-based management fosters the enforcement of service-oriented cognitive technologies in M2M communications. In this regard, a number of research activities focuses on the integration capabilities of the dynamic service registry, the service location discovery and the policy manager as the key functionalities of an integrated and secure policy-based SOA in M2M communications.

Compared with existing ETSI M2M communication functional architecture, this architecture proposes an additional service domain. The first three domains, device, network and application, are common to the other existing and established reference models. The proposed service domain removes procedures handled previously by the other domains such as establishing the connection of a device to the network to increase abstraction, reducing the computational load of the three domains and enforcing security required by some M2M applications. A policy based framework on the service layer can oversee all the security aspects and report any irregular behaviours (e.g. a device communicating to a different gateway). Compared with existing M2M communication architectures, the proposed solution focuses on providing additional security enforcement on the services handled by the

interaction points in the network. Through security policy enforcement by the policy engine, the service-oriented architecture can also alleviate security threats on event-based M2M communications by enforcing the appropriate policy controls. A case study using intelligent bus system is presented in order to demonstrate the applicability of the architecture.

Concisely, the architecture offers significantly increased performance and learning capabilities, while the PBNM system achieves adaptive QoS routing through automated configuration. The network resources can be utilised more efficiently through the adaptive routing path selection process. The performance also improves significantly, since the traffic can be rerouted to other paths in the case of congested links, while the EDs can trigger the appropriate corrective actions (i.e. traffic violation). Should the acceptable SLA thresholds be exceeded, the PDP deploys the necessary policies to take the proper corrective actions ensuring that any suboptimal routing changes or under-performance conditions of the network entities are mitigated.

The security and access control services are essential in realising a successful architecture using the SOA paradigm. Several research and industry activities reported in the literature have addressed the identification, the authentication and the authorisation management challenges, but without proposing a consolidated model to integrate these components. Apart from the QoS routing capabilities, ASPIDA aims to provide an adaptive PBM security management with appropriate indicators and controls to model identity, authentication, user roles, authorisation access control levels and to incorporate rule validation mechanisms for each policy set. The support of policy-based and integration capabilities offers automated maintenance of the policy sets, improved efficiency, simplified management and support of several types of environment (i.e. enterprise, service provider). The ASPIDA scheme is extensible to support further complex resource management extensions (i.e. optimal predictive resource allocation, resource usage, dynamic relocation of workloads). It also provides authorisation capabilities such as permission classes, task flows, SSO functionality and SAML along with XACML uses in complex authorisation scenarios. Further policy and context management additions in SOA environment can be supported. In ASPIDA, various challenges that are faced in the effort to support the appropriate access control levels along with policy management capabilities are considered, such as the access requests processed through independent µServices with light-weight containers. At the core of ASPIDA is the fact that the architecture achieves better adaptive security management with the use of the PDP, PIP and PAP along with the various PEPs to enforce the proper security controls. The policy-driven authorisations supported by the EDs can mitigate the security violations and apply the adequate access control configuration, if the application identity is authenticated and the authorisation grant is valid.

Many key functions and M2M service connection procedures from the device and the network domains can be embraced by the service domain in the proposed architecture. For instance, some of the M2M management functions and the event management processes enable secure data transportation over the reference points that reside in the service domain. Such services can be included in the service domain with the use of the relevant service capabilities. These can facilitate the evolution of the appropriate security policies, provide a secure interoperability between the device, the network, the application along with the service domains and enhance the security mechanisms as well as the efficiency.

As presented in the preceding chapters, the system architecture is evaluated with independent µServices and light-weight containers. Still, the data-driven µServices workflows

entail the access requests to be regulated by valid access tokens. Hence, there is a need to define, create, modify and validate the policy rules that express the dynamic constraints and dynamic system states. For instance, the resources can be generated and released during runtime to accommodate unforeseen demands in a scalable and elastic way with µServices, or the nodes can be even temporarily unavailable in the physical space. All the proposed solutions are evaluated by means of an ASPIDA prototype, various simulations, but also real use cases. The experiments are followed by the numeric data analysis to evaluate the performance and the effectiveness of ASPIDA. Nevertheless, the adaptive security management is expected to play a more integral role to establish the proper set of security and policy mechanisms all the way across the complete service life cycle.

7.3 Critical success factors

The performance of the execution event policies depends on the hardware characteristics. In the case of event scripting and applets, the performance depends on the *platform*, the *load-level* and the *total amount of available memory*. As there can be hardware constraints on the concurrent number of applets and policy scripts, some event actions are likely to be lost in the event of oversubscribing the system. New events might be dropped, as the policies have fixed-size event pools. Therefore, an event scheduler is needed to allocate and execute event selection decisions pre-emptively in order to improve the elasticity of event monitoring. In terms of *scalability* and *expandability*, the event scheduler allows for the activation or the suspension of the applets and scripts and the respective policies based on the network conditions and dynamically matching SLA criteria. As cloud-based applications are expected to grow further, the event scheduler can manage multi-domain policies and deal with the capacity, complexity and the performance issues that may arise. The event-based policies can remediate network failures (i.e. asymmetric routing, BGP dampening) and improve application performance by including aggregated bandwidth, low-latency links and SLA-driven routing. The policy driven authorisation implementations require complex interactions in order to enable policy based management (i.e. policy retrievals). Apart from the complexity of the authorisation policies, the runtime characteristics of the policy enforcement, the policy validation mechanisms and the policy execution (Table 15) all affect the high heterogeneous M2M services with complex requirements (i.e. unattended devices, energy restricted devices, low-bandwidth network, needs for zero latency and low bit-error rate applications).

Table 15. Other factors for policy-based management

Policy-based management	Factors
Policy definition	Complexity Detection of updates difficult to update Vulnerabilities Outdates rules
Policy execution	Load/retrieval Performance Difficult to manage Frequency of resource request Difficult to measure
Policy enforcement	Risks Urgency
Run time policy validation	Accuracy Correctness Performance Security violations Flaws Data quality

The outcome of the evaluation along with the metrics for dynamic authorisation frameworks for µServices rely on the interconnectivity, the constrained resources (i.e. data storage backend), the workload orchestration, the resilience and the existing dependencies. With regard to µServices, the computational power along with other factors such as the componentization, the segregating capabilities, the bulkheads, the circuit breakers, the decentralized data management and the infrastructure components impact the performance of the system. In [258], the authors show that there are certain trade-offs in the network performance evaluation of µServices against bare metal and regular containers to be considered in the design of the µServices components such as M2M discovery and lightweight messaging (i.e. Kafka) services. D. Shadija et al. [259] present the characteristics of µServices versus a SOA-approach aiming to assist the architects and the developers in selecting the most appropriate architecture.

The ASPIDA scheme supports higher throughputs compared to other contrasted solutions, even though the performance of the system depends on a large number of factors. The performance of the native clustering tools and the fine-grained orchestration, the distributed system challenges, the tolerance for failure of the services and the service boundaries can impact severely the design and the operations, such as the efficiency in real-time monitoring and detecting anomalies.

7.4 Future directions

This dissertation also leaves several open issues. This section outlines some of the limitations of the study and proposes some areas for future research. First, the NetFlow policy-based routing is not verified, which enables that traffic classification and the dynamic routing protocols with supported extensions are not tested; these ideas can be further studied and analysed. Subsequent tests can also study with the QoS classes to analyse per-hop behaviour (PHB) flows with changing traffic patterns.

In this work, the traffic is normally sent on a first-come-first-serve basis and all traffic is treated equally. However, future studies could analyse latency and bandwidth-sensitive, non-critical and non-interactive applications. Elaborating further on these experiments, a more detailed analysis of high priority traffic with security management aspects could provide useful results in relation to efficient bandwidth consumption, lower latencies and bottlenecks avoidance. The prioritisation and differential treatment of traffic could be provided based on the appropriate policy rules.

Furthermore, the performance issues of the policy management components have not been considered. The analysis of the adaptive policy-based security management in distributed environments and better exploitation of adaptive design for dynamic security policies are significant and could be taken into account for a more detailed examination. Additionally, the optimisation, the validation and security policy checks could be further analysed. With regard to the policy controls and policy engine, future research work could analyse the enforcement of dynamic policy-driven access controls. Moreover, as the policy engine can be either centralized or distributed, further analysis is required to evaluate the architectural and performance aspects of the contextual constraints setup.

The prototype illustrated in this dissertation could be expanded to incorporate the deployment of a policy service including event-based management and triggering mechanisms. The overall prototype could assess the high applicability and reinforce the modelling processes of the ASPIDA architecture.

Finally, the access controls and the authorisation mechanisms with dynamic contexts implemented with μServices remain an attractive research area. Considering the high technological diversity and the numerous side conditions of M2M protocols and communications, more complicated solutions and use cases are anticipated to evolve. Nevertheless, with a certainty, there is an increasing need to improve the adaptive policy-based security management.

References

[1] Tschofenig, J. Arkko D. Thaler D. McPherson, “RFC 7452 - Architectural Considerations in Smart Object Networking,” IETF, 2015. [Online]. Available: https://tools.ietf.org/html/rfc7452, 2015.

[2] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper, March 2017 [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/ service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

[3] N. Meghanathan, “Review of access control models for cloud computing,” Computer Science & Information Science, vol.3, no. 1, 2013, pp.77-85.

[4] H. F. Atlam, A. Alenezi, R. K. Hussein, G. B. Wills, “Validation of an Adaptive Risk-based Access Control Model for the Internet of Things," Computer Network and Information Security, vol. 1, 2018, pp. 26-35.

[5] M. Atallah, M. Blanton, N. Fazio, K. Frikken, “Dynamic and Efficient Key Management for Access Hierarchies,” ACM Transactions on Information and System Security, vol. 12, no. 3, 2009, p.18.

[6] C. Vincent, D. Ferraiolo, R. Kuhn, A. Schnitzer, K. Sandlin, K. Scarfone, “Guide to Attribute Based Access Control (ABAC) Definition and Considerations,” National Institute of Standards and Technology (NIST Special Publication 800-162), 2014.

[7] W. She, I.-L. Yen, B. Thuraisingham, E. Bertino, “Policy-Driven Service Composition with Information Flow Control,” 8th IEEE International Conference on Web Services, 2010, pp. 50-57.

[8] F. Paige, N. Matragkas, L.M. Rose, “Evolving models in model-driven engineering: State-of-the-art and future challenges,” Elsevier Journal of Systems and Software, vol.111, 2016, pp. 272-280.

[9] Y. Sinjilawi, M. Al-Nabhan, E. Abu-Shanab, “Addressing Security and Privacy Issues in Cloud Computing,” Journal of Emerging Technologies in Web Intelligence, vol. 6, no. 2, 2014, pp. 192-199.

[10] J. Vollbrecht, P. Calhoun, S. Farrell, L. Gommans, G. Gross, B. de Bruijn, C. de Laat, M. Holdrege, D. Spence, RFC 2904, “AAA Authorisation Framework,” IETF, 2000. [Online]. Available: http://tools.ietf.org/html/rfc2904.

[11] “Sun Microsystems Laboratories, XACML.” [Online]. Available: http://sunxacml.sourceforge.net.

[12] “PicketBox XACML.” [Online]. Available: https://community.jboss.org/wiki/PicketBox XACMLJBossXACML.

[13] G. Katsikogiannis, S. Mitropoulos, C. Douligeris, “Policy-based QoS management for SLA-driven adaptive routing,” Journal of Communications and Networks, vol.15, no.3, 2013, pp.301-311.

[14] N. Chen, R. Jiang, “Analysis and Improvement of User Authentication Framework for Cloud Computing,” IEEE Conference on Computer Science and Communication Technology vol. 756, 2013, pp. 3482-3486.

[15] A. Choudhury, P. Kumar, M. Sain, H. Lim, J. Hoon, “A Strong User Authentication Framework for Cloud Computing,” IEEE Services Computing Conference, 2011, pp. 110-115.

[16] W. Li and L. Ping, “Trust model to enhance Security and interoperability of Cloud environment,” 1st International conference on Cloud Computing, vol. 5931, 2009, pp. 69-79.

[17] M. A. Khan, “A survey of security issues for cloud computing,” Journal of Network and Computer Applications, vol.71, 2016, pp.11–29.

[18] K.o Hashizume, D.G Rosado, E. Fernández-Medina and E. B Fernandez, “An analysis of security issues for cloud computing,” Journal of Internet Services and Applications, vol.4, no.1, 2013, pp. 5.

[19] T. Qiu, N. Chen, K. Li, D. Qiao, Z. Fu, “Heterogeneous ad hoc networks: Architectures, advances and challenges”. Ad Hoc Networks, vol.2017, no.55, 2016, pp.143-152.

[20] C. Ngo, Y. Demchenko, C. Laat, “Multi-tenant attribute-based access control for cloud infrastructure services,” Journal of information security and applications, 2016, pp. 65–84.

[21] K. Vandikas, V. Tsiatsis, “Microservices in IoT clouds,” IEEE In Cloudification of the Internet of Things, 2016, pp. 1-6.

[22] European Telecommunications Standards Institute (ETSI), “Machine-to-Machine communications (M2M); Functional architecture,” Technical Specification (ETSI TS 102 690 v2.1.1, 2013-10). [Online]. Available: http://www.etsi.org/deliver/etsi_ts/102600_102699/102690/ 02.01.01_60/ts_102690v020101p.pdf

[23] H. Liu, F. Eldarrat, H. Alqahtani, A. Reznik, A., X. de Foy, Y. Zhang, “Mobile edge cloud system: Architectures, challenges, and approaches,” IEEE Systems Journal, 2017.

[24] C. S. Magurawalage, K. Yang, K. Wang, “Aqua Computing: Coupling Computing and Communications,” Preprint, available online arXiv:1510.0725, 2015.

[25] I. Stojmenovic, S. Wen, X. Huang, H. Luan, "An overview of Fog computing and its security issues," Journal Concurrency and Computation: Practice & Experience archive, vol. 28, issue 10, 2016, pp. 2991-3005.
[26] I. Stojmenovic, "Fog Computing: A Cloud to the Ground Support for Smart Things and Machine-to-Machine Networks," Australasian Telecommunication Networks and Applications Conference, 2014, pp. 117–12.
[27] S. Chaves, R. Uriarte, C. Westphall, "Toward an architecture for monitoring private clouds," Communications Magazine, IEEE, vol. 49, no. 12, pp. 130–137, 2011.
[28] R. Roman, J. Lopez, M. Mambo, "Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges," Future Generation Computer Systems, vol.78, 2018, pp.680-698.
[29] P. de Leusse, T. Dimitrakos, "SOA-based security governance middleware," 4th IEEE International Conference on Emerging Security Information, Systems and Technologies, 2010, pp.57-62.
[30] M. Chen, J. Wan, F. Li, "Machine-to-machine communications: Architectures, standards and applications," Transactions on internet & information systems, vol.6, no.2, 2012.
[31] D. Chen, G. Chang, "A survey on security issues of M2M communications in cyber-physical systems," KSII Transactions on Internet and Information Systems (TIIS), Vol.6, No.1, 2012, pp.24-45.
[32] J. Wan, M. Chen, F. Xia, L. Di, K. Zhou, "From machine-to-machine communications towards cyber-physical systems," Computer Science and Information Systems, Vol.10, no.3, June 2013, pp.1105-1128.
[33] M. Yadav, E. Hassan, G. Shroff, P. Agarwal, A. Srinivasan, "Searching for logical patterns in multi-sensor data from the industrial internet," In Machine Intelligence and Big Data in Industry, Springer International Publishing, vol.19, 2016, pp. 217-233.
[34] R. Helal, A. ElMougy, "An energy-efficient Service Discovery protocol for the IoT based on a multi-tier WSN architecture," IEEE 40th Local Computer Networks Conference Workshops, 2015, pp.862-869.
[35] S. Kisseleff, X. Chen, I.F. Akyildiz, W.H. Gerstacker, "Efficient Charging of Access Limited Wireless Underground Sensor Networks," IEEE Transactions on Communications, vol.64, no.5, 2016, pp. 2130-2142.
[36] J. Liu, Z. Wang, J. Cui, S. Zhou, B. Yang, "A Joint Time Synchronization and Localization Design for Mobile Underwater Sensor Networks," IEEE Transactions on Mobile Computing, vol.15, no.3, 2016, pp. 530-543.
[37] F. Hu, W. Siddiqui, K. Sankar, "Scalable security in Wireless Sensor and Actuator Networks (WSANs): integration re-keying with routing," Computer Networks, vol.51, no.1, 2007, pp.285-308.
[38] G. Toschi, L. Campos, C. Cugnasca, "Home automation networks: A survey," Computer Standards & Interfaces, vol.50, 2017, pp. 42-54.
[39] M. Zareei, E. Mahmoud Mohamed, M. Anisi, C. Vargas Rosales, K. Tsukamoto, M. Khurram Khan, "On-Demand Hybrid Routing for Cognitive Radio Ad-Hoc Network," IEEE Access, vol.4, 2016, pp. 8294-8302.
[40] K. Zhang, X. Liang, M. Baura, R. Lu, X. Shen, "PHDA: A priority based health data aggregation with privacy preservation for cloud assisted WBAN," Information Sciences, vol.284, 2014, pp. 130-141.
[41] C. C. Sobin, V. Raychoudhury, G. Marfia, A. Singla, "A survey of routing and data dissemination in Delay Tolerant Networks,", Journal of Network and Computer Applications, vol.67, 2016, pp. 128-146.
[42] A. Sher, N. Javaid, G. Ahmed, S. Islam, U. Qasim, Z. Khan, "MC: Maximum Coverage Routing Protocol for Underwater Wireless Sensor Networks," IEEE 19th International Conference on Network-Based Information Systems, 2016, pp. 91-98.
[43] P. Li, M. Pan, Y. Fang, "Capacity bounds of three-dimensional wireless ad hoc networks," IEEE/ACM Transactions on Networking, vol.20, no.4, 2012, pp. 1304-1315.
[44] S. Lim, W.C. Lee, G. Cao, C.R. Das, "Cache invalidation strategies for internet-based mobile ad hoc networks," Computer Communications, vol.30, no.8, 2007, pp. 1854-1869.
[45] J. Radak, B. Ducourthial, V. Cherfaoui, S. Bonnet, "Detecting road events using distributed data fusion: experimental evaluation for the icy roads case," IEEE Transactions on Intelligent Transportation Systems, vol.17, no.1, 2016, pp. 184-194.
[46] Q. Vey, S. Puechmorel, A. Pirovano, J. Radzik, "Routing in aeronautical ad-hoc networks," IEEE/AIAA 35th Digital Avionics Systems Conference, 2016, pp. 1-10.
[47] N. Islam, M. Hossain, G. Ali, P. Chong, "An expedite group key establishment protocol for Flying Ad-Hoc Network (FANET)," IEEE 5th International Conference on Informatics, Electronics and Vision, 2016, pp. 312-315.

[48] "UBER developers." [Online]. Available: https://developer.uber.com/docs/trip-experiences/introduction/
[49] "IoT Standards and Protocols," [Online]. Available: http://www.postscapes.com/internet-of-things-protocols/
[50] Technical Report for Machine-to-Machine Communication (M2M) / Internet of Things (IoT), v12.0, 2017. [Online]. Available: http://www.iotitaly.net/wp-content/uploads/2017/07/TEC_Communication_Technologies_M2M_IoT_Ver_12_0_-3rd-July-2017.pdf
[51] "Broadband Forum, Technical Report-069 CPE WAN Management Protocol," version: 1.4, 2018, [Online]. Available: https://www.broadband-forum.org/technical/download/TR-069.pdf
[52] "OMA Device Management Protocol," version 1.3, 2016, [Online]. Available: http://www.openmobilealliance.org/release/DM/V1_3-20160524-A/OMA-TS-DM_Protocol-V1_3-20160524-A.pdf
[53] "Open Mobile Alliance (OMA) LightweightM2M (LwM2M) Protocol," 2017. [Online]. Available: http://openmobilealliance.org/iot/lightweight-m2m-lwm2m/lightweightm2m-1-1-preview-3
[54] "IEEE 802.1AR: Secure Device Identity," 2009. [Online]. Available: https://1.ieee802.org/security/802-1ar/
[55] "Sensors4bins," [Online]. Available: http://www.mobicycle.co.uk/pilots-sensors4bins.html
[56] N. Kushalnagar, G. Montenegro, C. Schumacher, "RFC 4919, IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals," 2007. [Online]. Available: https://tools.ietf.org/html/rfc4919
[57] J. Nieminen, T. Savolainen, M. Isomaki, B. Patil, Z. Shelby, C. Gomez, "IPv6 over BLUETOOTH(R) Low Energy," 2015. [Online]. Available: https://tools.ietf.org/html/rfc7668
[58] "3GPP LTE-Advanced," [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced
[59] "IETF news, 5G and Internet Technology," 2017. [Online]. Available: https://www.ietf.org/blog/5g-and-internet-technology/
[60] A. Brandt, J. Buron, Transmission of IPv6 Packets over ITU-T G.9959, 2015, [Online]. Available: Networks https://tools.ietf.org/html/rfc7428
[61] S. Farrell, RFC 8376 Low-Power Wide Area Network (LPWAN) Overview, 2018, [Online]. Available: https://tools.ietf.org/html/rfc8376
[62] K. Fysarakis, I. Askoxylakis, C. Manifavas, O. Soultanos, I. Papaefstathiou and V. Katos, "Which IoT protocol? Comparing standarized approaches over a common M2M application," IEEE Global Communications Conference, 2016, pp. 1-7.
[63] Z. Shelby, K. Hartke, C. Bormann, "The constrained application protocol (CoAP)," IETF, 2014. [Online]. Available: https://tools.ietf.org/html/rfc7252
[64] A. Banks, R. Gupta, "OASIS Message Queuing Telemetry Transport (MQTT)," version 3.1.1, OASIS, 2014. [Online]. Available: http://docs.oasisopen.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.pdf
[65] I.K. Samaras, G.D. Hassapis, J.V. Gialelis, "A modified DPWS protocol stack for 6LoWPAN-based wireless sensor networks," IEEE Transactions on Industrial Informatics, vol.9, issue 1, 2013, pp.209-217.
[66] J.L. Hernández-Ramos, A.J. Jara, L. Marín, A.F.S. Gómez, "DCapBAC: embedding authorisation logic into smart things through ECC optimizations," International Journal of Computer Mathematics vol. 93, issue 2, 2016, pp. 345-366.
[67] G. Katsikogiannis, S. Mitropoulos, C. Douligeris, "An Identity and Access Management approach for SOA," IEEE Symposium on Signal Processing and Information Technology 2016, pp. 126-131.
[68] I. Fette, A. Melnikov, "RFC 6455 - The WebSocket Protocol," 2011. [Online]. Available: https://tools.ietf.org/html/rfc6455
[69] G. Wu, S. Talwar, K. Johnsson, N. Himayat, K.D. Johnson, "M2M: From Mobile to Embedded Internet," IEEE Communications Magazine, vol. 49, No. 4, 2011, pp.36-43.
[70] "Internet of Things – Architecture, Final architectural reference model for the IoT v3.0", [Online]. Available: http://www.iot-a.eu/public/public-documents/d1.5/at_download/file
[71] Cisco, "Cisco - The Internet of Things Reference Model" White paper, 2014. [Online]. Available: http://cdn.iotwf.com/resources/71/IoT_Reference_Model_White_Paper_June_4_2014.pdf
[72] Industrial Internet Reference Architecture, [Online]. Available: http://www.iiconsortium.org/IIRA.htm
[73] Reference Architecture Model Industrie (RAMI), https://www.phoenixcontact.com/assets/downloads_ed/local_gb/web_dwl_promotion/6788.pdf
[74] IBM Internet of Things Architecture, [Online]. Available: https://developer.ibm.com/architecture/iot

[75] P. Fremantle, “A reference architecture for the internet of things,” WSO2 White paper, 2015. [Online]. Available: https://wso2.com/download/getfile/wso2_whitepaper_a-reference-architecture-for-the-internet-of-things.pdf
[76] European Telecommunications Standards Institute (ETSI), “Machine-to-machine communications (M2M); M2M service requirements,” Technical Specification (ETSI TS 102 689 v2.1.1 2013-07). [Online]. Available: http://www.etsi.org/deliver/etsi_ts/102600_102699/102689/02.01.01_60/ts_102689v020101p.pdf
[77] European Telecommunications Standards Institute (ETSI), “SmartM2M, Smart Appliances, Communication Framework,” Technical Specification. [Online]. Available: http://www.etsi.org/deliver/etsi_ts/103200_103299/103267/01.01.01_60/ts_103267v010101p.pdf
[78] European Telecommunications Standards Institute (ETSI), “Smart Machine-to-Machine communications (SmartM2M) Home Gateway Initiative, RD036-Smart Home architecture and system requirements,” Technical Specification (ETSI TS 103 424 V1.1.1 2016-11). [Online]. Available: http://www.etsi.org/deliver/etsi_ts/103400_103499/103424/01.01.01_60/ts_103424v010101p.pdf
[79] European Telecommunications Standards Institute (ETSI), “Smart Machine-to-Machine communications (SmartM2M) Home Gateway Initiative, RD039-Requirements for Wireless Home Area Networks (WHANs) Supporting Smart Home Services” Technical Specification (ETSI TS 103 425 V1.1.1 2016-11). [Online]. Available: http://www.etsi.org/deliver/etsi_ts/103400_103499/103425/01.01.01_60/ts_103425v010101p.pdf
[80] European Telecommunications Standards Institute (ETSI), “Smart Machine-to-Machine communications (SmartM2M) Home Gateway Initiative RD048-HG Requirements for HGI Open Platform 2.1,” Technical Specification (ETSI TS 103 426 V1.1.1 2016-11). [Online]. Available: http://www.etsi.org/deliver/etsi_ts/103400_103499/103426/01.01.01_60/ts_103426v010101p.pdf
[81] Telecommunications Industry Association (TIA) - 4940.005, “Smart Device Communications Reference Architecture,” 2011
[82] European Telecommunications Standards Institute (ETSI), “oneM2M; Functional Architecture,” Technical Specification (oneM2M TS-0001 version 2.10.0 Release 2 2016). [Online]. Available: http://www.etsi.org/deliver/etsi_ts/118100_118199/118101/02.10.00_60/ts_118101v021000p.pdf
[83] oneM2M, “Security Solutions,” Technical Specification (TS-0003-V2.12.1, March 2018). [Online]. Available: http://www.onem2m.org/images/files/deliverables/Release2A/TS-0003-Security_Solutions-v_2_12_1-.pdf
[84] Open Mobile Alliance (OMA), Lightweight Machine to Machine Architecture, 2017. [Online]. Available: http://www.openmobilealliance.org/release/LightweightM2M/V1_0-20170208-A/OMA-AD-LightweightM2M-V1_0-20170208-A.pdf
[85] M. Bauer, M. Boussard, N. Bui, F. Carrez, C. Jardak, “Final architectural reference model for the IoT v3. 0. Internet of Things Architecture loT-A,” Techincal Report, http://www.meet-iot.eu/deliverables-IOTA/D1_5.pdf, 2013.
[86] J. Guth, U. Breitenbücher, M. Falkenthal, F. Leymann, L. Reinfurt, “Comparison of IoT platform architectures: A field study based on a reference architecture,” IEEE Conference in Cloudification of the Internet of Things, 2016, pp. 1-6.
[87] M. Bauer, N. Bui, J. De Loof, C. Magerkurth, A. Nettsträter, J. Stefa, J. W. Walewski, “IoT Reference Architecture,” Springer Berlin Heidelberg, 2013, pp. 163-211.
[88] L. Zheng, H. Zhang, W. Han, X. Zhou, J. He, Z. Zhang, Y. Gu, J. Wang, “Technologies, Applications, and Governance in the Internet and of Things,” in Internet of Things - Global Technological and Societal Trends. River Publishers, 2009.
[89] Y. Mehmood, C. Görg, M. Muehleisen, A. Timm-Giel, “Mobile M2M communication architectures, upcoming challenges, applications, and future directions,” EURASIP Journal on Wireless Communications and Networking, 2015, pp. 250-287.
[90] European Telecommunications Standards Institute (ETSI), “Machine-to-Machine communications (M2M); Smart Metering Use Cases,” Technical report (ETSI TR 102 691 v1.1.1, 2010-05. [Online]. Available: http://www.etsi.org/deliver/etsi_tr/102600_102699/102691/01.01.01_60/tr_102691v010101p.pdf
[91] European Telecommunications Standards Institute (ETSI), “Digital cellular telecommunications system (Phase 2+), Universal Mobile Telecommunications System (UMTS),” LTE, Service accessibility, (3GPP TS 22.011 version 12.2.0 Release 12), Technical Specification, 2014 [Online]. Available: http://www.etsi.org/deliver/etsi_ts/122000_122099/122011/12.02.00_60/ts_122011v120200p.pdf
[92] 3rd Generation Partnership Project (3GPP TM), “System improvements for Machine-Type Communications (MTC),” 3rd Generation Partnership Project, Technical Specification Group

Services and System Aspects, (Release 11), Technical Report, 2012. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=968
[93] European Telecommunications Standards Institute (ETSI), "ETSI, Technical specification: Smart Appliances; Communication Framework, ETSI TS 103 267 V1.1.1 (2015-12)," [Online]. Available: http://www.etsi.org/deliver/etsi_ts/103200_103299/103267/01.01.01_60/ts_103267v010101p.pdf
[94] 3rd Generation Partnership Project (3GPP), "System Improvements for Machine-Type Communications (MTC)," 3GPP TR 23.888. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=968
[95] N. Alshuqayran, N. Ali, R. Evans, "A Systematic Mapping Study in Microservice Architecture," IEEE Service-Oriented Computing and Applications, 2016, pp. 44-51.
[96] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, W. Zhao, "A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications," IEEE Internet of Things Journal, vol. 4, issue 5, 2017, pp. 1125-1142.
[97] I. Lee, K. Lee, "The Internet of Things (IoT): Applications, investments and challenges for enterprises," Business Horizons, vol. 58, no.4, 2015, pp. 431-440.
[98] Georgakakis, S. A. Nikolidakis, D. D. Vergados, C. Douligeris, "Spatio temporal emergency role based access control (STEM-RBAC): A time and location aware role based access control model with a break the glass mechanism," IEEE Symposium on Computers and Communications, 2011, pp. 764–770.
[99] P. Kotzanikolaou, E. Magkos, N. Petrakos, C. Douligeris, V. Chrissikopoulos, "Fair Anonymous Authentication for Location Based Services," Data Privacy Management and Autonomous Spontaneous Security, Springer, 2013, pp. 1–14.
[100] H. Vahdat-Nejad, A. Ramazani, T. Mohammadi, W. Mansoor, "A survey on context-aware vehicular network applications," Vehicular Communications, vol. 3, 2016, pp.43-57.
[101] H. Maw, H. Xiao, B. Christianson, J. Malcolm, "BTG-AC: Break-the-Glass Access Control Model for Medical Data in Wireless Sensor Networks," IEEE Journal of Biomedical and Health Informatics, vol. 20, issue 3, 2016, pp.763-774.
[102] N. Qwasmi, R. Liscano, "TinyPolicy: A Distributed Policy-based Management Framework for Wireless Sensor Networks," IFIP/IEEE International Symposium on Integrated Network Management, 2015, pp. 918-921.
[103] K. Fysarakis, O. Soultatos, C. Manifavas, I. Papaefstathiou, I. Askoxylakis, "XSACd—Cross-domain resource sharing & access control for smart environments," Future Generation Computer Systems, vol.80, 2018, pp.572-582.
[104] R. Neisse, G. Steri, I. N. Fovino, G. Baldini, "SecKit: a model-based security toolkit for the internet of things." Computers and Security, vol.54, 2015, pp.60-76.
[105] A. Westerinen, J. Schnizlein, J. Strassner, M. Scherling, B. Quinn, S. Herzog, A. Huynh, M. Carlson, J. Perry, S. Waldbusser, "RFC 3198 - Terminology for Policy-based Management," IETF, 2001. [Online]. Available: https://www.ietf.org/rfc/rfc3198.txt
[106] T. Phan, J. Han, J.-G. Schneider, T. Ebringer and T. Rogers, "A Survey of Policy-based Management Approaches for Service Oriented Systems," IEEE Australian Software Engineering. Conference, 2008, pp. 392-401.
[107] R. Yavatkar, D. Pendarakis, R. Guerin, "A Framework for Policy-based Admission Control,'" IETF RFC2753, 2000. [Online]. Available: http://www.rfc-base.org/rfc-2753.html
[108] K. C. Feeney, S. N. Foley, R. Brennan, "A Trust Model for Capability Delegation in Federated Policy Systems," 6th International Conference on Risk and Security of Internet and Systems, 2011, pp. 1–8.
[109] Open Mobile Alliance (OMA), "Policy Evaluation, Enforcement and Management Interface (PEM-2)," Technical Specification, version 1.0, 2012.
[110] Open Mobile Alliance (OMA), "Policy Evaluation, Enforcement and Management Architecture," version 1.0, 2012.
[111] Amazon Web Services: AWS Identity and Access Management(IAM) User Guide, [Online]. Available: http://docs.aws.amazon.com/IAM/latest/UserGuide/introduction.html
[112] M. Mejri, H. Yahyaoui, "Formal specification and integration of distributed security policies," Computer Languages, Systems & Structures, vol. 49, 2017, pp. 1-35.
[113] Z. Wu, L. Wang, "An innovative simulation environment for cross-domain policy enforcement," Simulation Model, Practice Theory, vol.19, no.7, 2011, pp.1558–1583.
[114] K. C. Feeney, S. N. Foley, R. Brennan, "A Trust Model for Capability Delegation in Federated Policy Systems," 6th International Conference on Risk and Security of Internet and Systems, 2011, pp. 1–8.

[115] P. Mazzoleni, B. Crispo, S. Sivasubramanian, E. Bertino, "XACML Policy Integration Algorithms," ACM Transactions on Information System Security, 2008, pp. 852-869.
[116] M. Dell'Amico, G. Serme, M. Idrees, A. Santana de Oliveira, Y. Roudier, "HiPoLDS: A Hierarchical Security Policy Language for Distributed Systems," Information Security Technical Report, vol. 17, issue 3, 2013, pp.81-92.
[117] R. Neisse, J. Doerr, "Model-based specification and refinement of usage control policies," IEEE Eleventh Annual International Conference on Privacy, Security and Trust, 2013, pp. 169-176.
[118] S. Sicari, A. Rizzardi, D. Miorandi, C. Cappiello, A. Coen-Porisini, "Security policy enforcement for networked smart objects," Elsevier Computer Networks, vol.108, 2016, pp. 133-147.
[119] D. Ferraiolo, V. Atluri, S. Gavrila, "The Policy Machine: A novel architecture and framework for access control policy specification and enforcement," Journal of Systems Architecture, vol.57, issue 4, 2011, pp. 412-24.
[120] D. Qiao, T. Qiu, H. Kim, "Self-organising and smart protocols for heterogeneous ad hoc networks in the Internet of Things," Ad Hoc Networks, vol.55, 2017, pp. 1-2.
[121] G. Katsikogiannis, S. Mitropoulos, C. Douligeris, "Policy-based QoS management for SLA-driven adaptive routing," Journal of Communications and Networks, vol.15, issue 3, 2013, pp. 301-311.
[122] G. Katsikogiannis, S. Mitropoulos, C. Douligeris, "Optimizing SLA-driven adaptive routing," IEEE Symposium on Computers and Communication, 2016, pp. 627-632.
[123] T. Um, G. Lee, J. Choi, "Strengthening trust in the future social-cyber-physical infrastructure: an ITU-T perspective," IEEE Communications Magazine, vol.54, issue 9, 2016, pp. 36-42.
[124] J. Li, A. Bowers, D. Lin, P. Jiang, W. Jiang, "PAS: policy-based assistance in sensor networks," Computing, 2016, pp. 1-14.
[125] Microsoft Developer Network, Identity and Access Management," [Online]. Available: http://msdn.microsoft.com/en-us/library/aa480030.aspx.
[126] "Gartner, Predicts 2014: Identity and access management," [Online]. Available: http://www.gartner.com/.
[127] N. Li, Q. Wang, W. Qardaji, E. Bertino, P. Rao, J. Lobo, D. Lin, "Access control policy combining: theory meets practice," 14th ACM Symposium on Access Control Models and Technologies, 2009, pp. 135–144.
[128] A. Karp, H. Haury, M. Davis, "From ABAC to ZBAC: The Evolution of Access Control Models," HP Laboratories, HPL-2009-30.
[129] T. Bray, "RFC7159: The JavaScript Object Notation (JSON) Data Interchange Format," Internet Engineering Task Force (IETF) Request for Comments, 2014. [Online]. Available: http://tools.ietf.org/html/rfc7159.
[130] A. Laugé, J. Hernantes, J. Sarriegi, "Critical infrastructure dependencies: A holistic, dynamic and quantitative approach," International Journal of Critical Infrastructure Protection, vol.8, 2015, pp. 16-23.
[131] C. Alcaraz, S. Zeadally, "Critical infrastructure protection: Requirements and challenges for the 21st century," International Journal of Critical Infrastructure Protection, vol.8, 2015, pp. 53-66.
[132] W. Knowles, D. Prince, D. Hutchison, J. Disso, K. Jones, "A survey of cyber security management in industrial control systems," International Journal of Critical Infrastructure Protection, vol.9, 2015, pp. 52-80.
[133] A. Sadeghi, C. Wachsmann, M. Waidner, "Security and privacy challenges in industrial internet of things," 52nd ACM/EDAC/IEEE Design Automation Conference, 2015, pp. 1-6.
[134] ENISA, "Ad-hoc & sensor networking for M2M Communications Threat Landscape and Good Practice Guide," ENISA publications, 2017. [Online]. Available: https://www.enisa.europa.eu/publications/m2m-communications-threat-landscape/at_download/fullReport
[135] M. Wang, Z. Yan, "A Survey on Security in D2D Communications," Mobile Network Application, vol. 22, issue 2, 2017, pp. 195–208.
[136] A. Barki, A. Bouabdallah, S. Gharout, J. Traore, "M2M Security: Challenges and Solutions," IEEE Communications Surveys & Tutorials, vol.18, issue 2, 2016, pp. 1241-1254.
[137] R. Amin, G.P. Biswas, "A secure light weight scheme for user authentication and key agreement in multi-gateway based wireless sensor networks," Elsevier Ad Hoc Networks, 2016, vol.36, pp. 58-80.
[138] E. Fadel, V. Gungor, L. Nassef, N. Akkari, M. Malik, S. Almasri, I. Akyildiz, "A survey on wireless sensor networks for smart grid," Computer Communications, vol.71, 2015, pp. 22-33.

[139] S. Etigowni, D. Tian, G. Hernandez, S. Zonouz, K. Butler, "CPAC: Securing Critical Infrastructure with Cyber-Physical Access Control," Conference on Computer Security Applications, 2016, pp. 139-152.
[140] H. Maw, H. Xiao, B. Christianson, J. Malcolm, "BTG-AC: Break-the-Glass Access Control Model for Medical Data in Wireless Sensor Networks," IEEE Journal of Biomedical and Health Informatics, vol.20, no.3, 2016, pp. 763-774.
[141] J. Singh, J. Bacon, D. Eyers, "Policy enforcement within emerging distributed, event-based systems," in: DEBS 2014 - Proceedings of the 8th ACM International Conference on Distributed Event-Based Systems, 2014, pp. 246–255.
[142] S. Sicari, A. Rizzardi, D. Miorandi, C. Cappiello, A. Coen-Porisini, "A secure and quality-aware prototypical architecture for the Internet of Things," Elsevier Information Systems, vol.58, 2016, pp. 43-55.
[143] J. Suarez, J. Quevedo, I. Vidal, D. Corujo, J. Garcia-Reinoso, R.L. Aguiar, "A secure IoT management architecture based on Information-Centric Networking," Journal of Network and Computer Applications, vol. 63, 2016, pp. 190-204.
[144] M2M device connections forecast 2014-2024. [Online]. Available: http://www.machinetomachinemagazine.com/2014/08/20/report-m2m-device-connections-forecast-2014-2024/
[145] "ETSI, Telecommunications and Internet Converged Services and Protocols for Advanced Networking (TISPAN); NGN Security; Security Architecture, ETSI TS 187 003 V2.3.2 (2011-03)," [Online]. Available: http://www.etsi.org/deliver/etsi_ts/187000_187099/187003/02.03.02_60/ts_187003v020302p.pdf
[146] "ETSI, Machine-to-Machine Communications (M2M); Study on Semantic support for M2M Data. ETSI TR 101 584 V2.1.1 (2013-12)," [Online]. Available: http://www.etsi.org/deliver/etsi_tr/101500_101599/101584/02.01.01_60/tr_101584v020101p.pdf
[147] "IEEE 802.16p-10/0005 Machine to Machine (M2M) Communications Technical Report: Describes the usage of M2M communication and various standard modifications which are required to support future M2M communications," [Online]. Available: http://wirelessman.org/m2m/index.html#10_0005
[148] "IEEE 802.16p-10/0014 provides various methodologies which seem to be helpful for M2M communications," [Online]. Available: http://wirelessman.org/m2m/index.html#11_0014
[149] "IEEE 802.16ppc-10/0003r9: Machine to Machine (M2M) Communication PAR Form and Five Criteria," [Online]. Available: http://www.ieee802.org/16/ppc/docs/80216ppc-10_0003r9.doc
[150] "IEEE 802.16ppc-10/0002r7: Machine to Machine (M2M) Communication Study Report," [Online]. Available: http://www.ieee802.org/16/ppc/docs/80216ppc-10_0002r7.doc
[151] "IEEE P802.16m System Requirements Document (SRD)," [Online]. Available: http://www.ieee802.org/16/m2m/docs/80216p-10_0004r3.doc
[152] "IEEE 802.16p-11/0014. IEEE 802.16p Machine to Machine (M2M) Evaluation Methodology Document (EMD). IEEE 802.16 Broadband Wireless Access Working Group, (2010)," [Online]. Available: http://wirelessman.org/m2m/index.html#11_0014
[153] "3GPP. Service requirements for machine-type communications (MTC); stage 1 (Release 12), 3GPP TS 22.368, ETSI TS122 368 V13.1.0, (2016-03)," http://www.etsi.org/deliver/etsi_ts/122300_122399/122368/13.01.00_60/ts_122368v130100p.pdf
[154] "3GPP. Technical Specification Group Services and System Aspects, Architecture enhancements to facilitate communications with packet data networks and application, 3GPP TS 23.682 V13.1.0, (2012)," [Online]. Available: http://www.arib.or.jp/english/html/overview/doc/STD-T63v10_00/5_Appendix/Rel11/23/23682-b30.pdf
[155] "NIST framework and roadmap for Smart Grid interoperability standards, v1.0," [Online]. Available: https://www.nist.gov/sites/default/files/documents/public_affairs/releases/smartgrid_interoperability_ final.pdf
[156] "Open Mobile Alliance: M2M enablers for IoTs," [Online]. Available: http://openmobilealliance.org/about-oma/work-program/m2m-enablers/
[157] "Open Mobile Alliance and Machine-to-Machine (M2M) communication," [Online]. Available: http://openmobilealliance.org/static/oma-annual-reports/documents/oma%20collateral%20m2m%205-11.pdf
[158] "METIS2020: Mobile and Wireless Communications Enablers for Twenty-Twenty (2020)," [Online]. Available: https://www.metis2020.com/?doing_wp_cron=1440943491553276062011 7187500000_links.html
[159] "METIS2020: Mobile and wireless communications Enablers for Twenty-twenty (2020) Information Society, Novel radio link concepts and state of the art analysis. Deliverable Number

ICT-317669-METIS/D2.2, 2013," [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D2.2_v1.pdf
[160] "METIS2020: Mobile and wireless communications Enablers for Twenty-twenty (2020) Information Society, Scenarios, requirements and KPIs for 5G mobile and wireless system (Deliverable Number ICT-317669-METIS/D1.1, 2013)," [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D1.1_v1.pdf
[161] "METIS2020: Mobile and wireless communications Enablers for Twenty-twenty (2020) Information Society, Initial report on horizontal topics, first results and 5G system concept (METIS Deliverable Number ICT-317669-METIS/D2.4, 2015)," [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D2.4_v1.pdf
[162] "oneM2M: Functional architecture. Technical Specification TS-0001-V1.6.1, (2015)," [Online]. Available: http://www.onem2m.org/images/files/deliverables/TS-0001-Functional_Architecture-V1_6_1.pdf
[163] "oneM2M: Security, Technical Report TR-0008-V2.0.0, (20165)," [Online]. Available: http://www.onem2m.org/images/files/deliverables/Release2/TR-0008-Security-V2_0_0.pdf
[164] "EXALTED: EXpAnding LTE for Devices, The EXALTED system architecture (Final). EXALTED Deliverable Number D2.3, (2012)," [Online]. Available: http://cordis.europa.eu/docs/projects/cnect/2/258512/080/deliverables/001-EXALTEDWP2D23.pdf
[165] "EXALTED: EXpAnding LTE for Devices, End-to-end (E2E) M2M system-device management. EXALTED Deliverable Number D4.3 (2012)," [Online]. Available: http://cordis.europa.eu/docs/projects/cnect/2/258512/080/deliverables/001-EXALTEDWP4D43.pdf
[166] "EXALTED: EXpAnding LTE for Devices, Security, Authentication and Provisioning Solutions (2012). Deliverable Number D5.1, (2012)," [Online]. Available: http://cordis.europa.eu/docs/projects/cnect/2/258512/080/deliverables/001-EXALTEDWP5D51.pdf
[167] "EXALTED: Security Solutions for P2P Relaying (2013). Deliverable Number D5.3," [Online]. Available: http://cordis.europa.eu/docs/projects/cnect/2/258512/080/deliverables/001-EXALTED WP5D53.pdf
[168] SemSorGrid4Env, [Online]. Available: http://www.semsorgrid4env.eu/home.jsp
[169] SemSorGrid4Env, [Online]. Available: http://www.ict-sensei.org
[170] SERVFACE, [Online]. Available: http://www.servface.org
[171] SHAPE, [Online]. Available: http://www.shape-project.eu
[172] SMARTSANTANDER, [Online]. Available: http://www.smartsantander.eu
[173] SMART-Net, [Online]. Available: http://www.ict-smartnet.eu
[174] SOA4All, [Online]. Available: http://www.soa4all.eu
[175] SOCRATES, [Online]. Available: http://www.fp7-socrates.org
[176] SPITFIRE, [Online]. Available: http://www.spitfire-project.eu
[177] TAMPRES, [Online]. Available: http://www.tampres.eu
[178] TECOM, [Online]. Available: http:// www.tecom-project.eu
[179] VITRO, [Online]. Available: http://www.vitro-fp7.eu
[180] WISEBED, [Online]. Available: http://www.wisebed.eu
[181] WOMBAT, [Online]. Available: http://www.wombat-project.eu
[182] B. Moore, E. Elleson, J. Strasser, A. Weterinen, "Policy Core Information Model," IETF RFC 3060, 2001. [Online]. Available: https://tools.ietf.org/html/rfc3060
[183] C. Doherty, N. Hurley, "Policy-based autonomic replication for next generation network management systems," 1st Conference on Distributed Autonomous Network Management Systems, 2006.
[184] D. Raymer, J. Strassner, E. Lehtihet, "End-to-End Model Driven Policy-based Network Management," 7th IEEE Policies for Distributed Systems and Networks, 2006, pp. 4-pp.
[185] V. Ozianyi, R. Good, N. Carrilho and N. Ventura, "XML-Driven Framework for Policy-Based QoS Management of IMS Networks," IEEE Global Telecommunications Conference, 2008, pp. 1685-1690.
[186] L. Granville, L. Tarouco, "QAME - QoS-Aware Management Environment," 25th Computer Software and Applications Conference, 2001, pp. 269-274.
[187] Cisco, Intelligent WAN: Use Path Control to Solve the Challenges of Application Performance, [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/enterprise-networks/intelligent-wan/whitepaper_C11-732389.html
[188] Cisco, Intelligent WAN: Technology Design Guide, January 2015, [Online]. Available: http://www.cisco.com/c/dam/en/us/td/docs/solutions/CVD/Jan2015/CVD-IWANDesignGuide-JAN15.pdf

[189] Cisco, Generic Performance of Cisco's Embedded Event Manager (EEM), [Online]. Available: https://supportforums.cisco.com/document/48891/cisco-eem-best-practices#Generic_Performance
[190] B. Chang, R. Hwang, "Distributed Cost-based Update Policies for QoS Routing on Hierarchical Networks," Information Sciences, vol.159, 2004 pp. 87-108.
[191] Y. Ledru, A. Idani, J. Milhau, N. Qamar, R. Laleau, J.-L. Richier, M.-A. Labiadh, "Validation of IS Security Policies Featuring Authorisation Constraints," International Journal of Information System Modeling and Design, 2015, pp. 24-46.
[192] "STORK 2.0, Secure Identity Across Borders Linked 2.0," D4.1 First version of process flows, 2012. [Online]. Available: https://www.eid-stork2.eu/.
[193] G. Di Modica, O. Tomarchio, "Matchmaking semantic security policies in heterogeneous clouds," Future Generation Computer Systems, vol.55, 2016, pp.176-185.
[194] H.V. Nguyen, L. Lo Iacono, "RESTful IoT Authentication Protocols," Mobile Security and Privacy book, 2016, pp. 217-234.
[195] S. Mitropoulos, C. Douligeris, "The impact of Service-Oriented Architecture (SOA) technologies in global market-oriented enterprises," International Journal of Applied Systemic Studies, vol. 4, no. 1-2, 2011, p. 106.
[196] M. Atallah, M. Blanton, N. Fazio, K. Frikken, "Dynamic and Efficient Key Management for Access Hierarchies," ACM Transactions on Information and System Security, vol. 12, no. 3, 2009, p.18.
[197] N. Zhao, M. Chen, S. Chen, M. Shyu, "MRBAC: Hierarchical Role Management and Security Access Control for Distributed Multimedia Systems," Proceedings of IEEE International Symposium on Object/Component/Service-oriented Real-time Distributed Computing, 2008, pp. 76-82.
[198] Y. Sinjilawi, M. Al-Nabhan, E. Abu-Shanab, "Addressing Security and Privacy Issues in Cloud Computing," Journal of Emerging Technologies in Web Intelligence, vol. 6, no. 2, 2014, pp. 192-199.
[199] K. Rantos, A. Papanikolaou, K. Fysarakis, C. Manifavas, "Secure policy-based management solutions in heterogeneous embedded systems networks," IEEE International Conference on Telecommunications and Multimedia, 2012, pp. 227–232.
[200] D. Ameller, X. Burgués, O. Collell, D. Costal, X. Franch, M. Papazoglou, "Development of service-oriented architectures using model-driven development: A mapping study," Information and Software Technology, vol.62, 2015, pp. 42-66.
[201] G. Alférez, V. Pelechano, R. Mazo, C. Salinesi, D. Diaz, "Dynamic adaptation of service compositions with variability models," Journal of Systems and Software, vol.91, 2014, pp. 24-47.
[202] European Telecommunications Standards Institute (ETSI), "Machine-to-machine communications (M2M), Smart Metering Use Cases," Technical Specification, [Online]. Available: http://www.etsi.org/deliver/etsi_tr/102600_102699/102691/01.01.01_60/tr_102691v010101p.pdf, 2010.
[203] AS Pathan, "Security of self-organising networks: MANET, WSN, WMN, VANET," CRC press, 2016.
[204] J. Santos, T. Wauters, B. Volckaert, F. Turck, "Fog Computing: Enabling the Management and Orchestration of Smart City Applications in 5G Networks," Entropy 20, no. 1, 2017, p.4.
[205] D.A. Egbe, M. B. Mutanga, M. O. Adigun, "Service discovery in ad-hoc mobile cloud: contemporary approaches and future direction," Journal of Theoretical and Applied Information Technology, vol.90, no.1, 2016, pp. 101.
[206] M. Familiar, J. Martínez, I. Corredor, C. García-Rubio, "Building service-oriented smart infrastructures over wireless ad hoc sensor networks: A middleware perspective," Computer Networks, vol. 56, no.4, 2012, pp. 1303-1328.
[207] J. Gustafsson, R. Kyusakov, H. Mäkitaavola, J. Delsing, "Application of Service Oriented Architecture for Sensors and Actuators in District Heating Substations," Sensors (Basel) Journal, vol. 14, no.8, 2014, pp. 15553–15572.
[208] Ponder2 Project, 2013. [Online]. Available: http://www.ponder2.net/
[209] S. Tanaka, K. Fujishima, N. Mimura, T. Ohashi, M. Tanaka, "IoT System Security Issues and Solution Approaches," Hitachi Review, vol. 65, no. 8, 2016.
[210] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," IEEE Communication Survey Tutorials, vol. 17, no. 4, 2015, pp. 2347–2376.
[211] ENISA, "Measurement Frameworks and Metrics for Resilient Networks and Services," Technical report, 2011.

[212] A. Talaminos-Barroso, M.A. Estudillo-Valderrama, L. Roa, J. Reina-Tosina, F. Ortega-Ruiz, "Machine-to-Machine protocol benchmark for eHealth applications – Use case: Respiratory rehabilitation," Elsevier Computer methods and programs in biomedicine, vol.129, 2016, pp.1-11.
[213] H.S.G. Pussewalage, VA Oleshchuk, "Attribute Based Access Control Scheme with Controlled Access Delegation for Collaborative E-health Environments," Elsevier Journal of Information Security and Applications, vol. 37C, 2017, pp. 50-64.
[214] E. Rescorla, N. Modadugu, "RFC 6347, Datagram Transport Layer Security," 2012.
[215] S.L. Keoh, S.S. Kumar, H. Tschofenig, "Securing the internet of things: A standardization perspective," IEEE Internet of Things Journal, vol. 1, no. 3, 2014, pp.265-275.
[216] Y. Shoukry, P. Martin, P. Tabuada, M. Srivastava, "Non-invasive spoofing attacks for anti-lock braking systems," In International Workshop on Cryptographic Hardware and Embedded Systems, Springer, 2013, pp. 55-72.
[217] M. Krotofil, D. Gollmann, "Industrial control systems security: What is happening?" 11th IEEE International Conference in Industrial Informatics, 2013, pp. 670-675.
[218] European Telecommunications Standards Institute (ETSI), "Machine to Machine communications (M2M); Use cases of Automotive Applications in M2M capable networks," ETSI TR 102 898 V1.1.1 2013.
[219] Z. Garofalaki, D. Kallergis, G. Katsikogiannis, I. Ellinas, C. Douligeris, "Transport services within the IoT ecosystem using localisation parameters," In 2016 IEEE International Symposium on Signal Processing and Information Technology, 2016, pp. 87-92.
[220] J. Kim, J. Lee, J. Kim, J. Yun, "M2M Service Platforms: Survey, Issues, and Enabling Technologies, IEEE Communications Surveys & Tutorials," vol. 16, no. 1, 2014, pp. 61-76.
[221] D. Preuveneers, W. Joosen, "Access Control with Delegated Authorisation Policy Evaluation for Data-Driven Microservice Workflows," Future Internet Journal, vol. 9, issue 4, 2017, pp. 1-21.
[222] G. Katsikogiannis, D. Kallergis, Z. Garofalaki, S. Mitropoulos, C. Douligeris, A Policy-aware Service Oriented Architecture for Secure Machine-to-Machine Communications, Elsevier Ad Hoc Networks Journal, April 2018.
[223] G. Katsikogiannis, Z. Garofalaki, D. Kallergis, C. Douligeris, "PDA: A Policy-driven for Authorisations Scheme with μServices," 2nd International Balkan Conference on Communications and Networking, June 2018.
[224] T. Hardjono, E. Maler, M. Machulak, "User-Managed Access (UMA) Profile of OAuth 2.0,". [Online]. Available: https://tools.ietf.org/html/draft-hardjono-oauth-umacore-13
[225] OneM2M technical report, "Dynamic Authorisation for IoT," TR-0019 version 2.0.0, 2016.
[226] B. K. Lee, M. S. Kim, J.H. Seo, "Design and Implementation of The Capability Token based Access Control System in the Internet of Things," Journal of The Korea Institute of Information Security & Cryptology, vol.25, no.2, 2015, pp. 439-448.
[227] S. Cirani, M. Picone, P. Gonizzi, L. Veltri, G. Ferrari, "IoT-OAS: An OAuth-Based Authorisation Service Architecture for Secure Services in IoT Scenarios," IEEE Sensors Journal, vol. 15, no. 2, 2015, pp. 1224-1234.
[228] C. Bormann, P. Hoffman, IETF RFC 7049, "Concise Binary Object Representation (CBOR)," 2013. [Online]. Available: https://tools.ietf.org/html/rfc7049
[229] O. Ethelbert, F. Moghaddam, P. Wieder, R. Yahyapour, "A JSON Token - Based Authentication and Access Management Schema for Cloud SaaS Applications", IEEE 5th International Conference in Future Internet of Things and Cloud, 2017, pp. 47-53.
[230] S. Gusmerolia, S. Piccionea, D. Rotondi, "A capability-based security approach to manage access control in the Internet of Things", Mathematical and Computer Modelling, vol. 58, Issues 5–6, 2013, pp. 1189-1205.
[231] J. L. Hernandez-Ramos, A. J. Jara, L. Marın, A.F. Skarmeta, "Distributed Capability-based Access Control for the Internet of Things", Journal of Internet Services and Information Security, vol. 3, no. 3/4, 2013, pp. 1-16.
[232] J. L. Hernandez-Ramos, V. Moreno, J. B. Bernabé, D. G. Carrillo, A. F. Skarmeta, "SAFIR: Secure access framework for IoT-enabled services on smart buildings", Journal of Computer and System Sciences, vol. 81, issue 8, 2015, pp. 1452-1463.
[233] G. Katsikogiannis, S. Mitropoulos, C. Douligeris, "UACS: Towards Unified Access Control Services," IEEE Symposium on Signal Processing and Information Technology, 2015, pp. 127-132.
[234] A. Mourad, H. Jebbaoui, "Towards efficient evaluation of XACML policies," 12th Annual International Conference on Privacy, Security and Trust, 2014, pp. 164-171.
[235] A. Vance, P. Lowry, D. Eggett, "A new approach to the problem of access policy violations: Increasing perceptions of accountability through the user interface," MIS Quarterly, vol. 39, no. 2, 2015, pp. 345-366.

[236] B. Campbell, C. Mortimore, M. Jones, "IETF RFC 7522, Security Assertion Markup Language (SAML) 2.0 Profile for OAuth 2.0 Client Authentication and Authorization Grants," [Online]. Available: https://tools.ietf.org/html/rfc7522
[237] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons. [Online]. Available: http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32016R0679.
[238] V. C. Hu, E. Martin, J. Hwang, T. Xie, "Conformance Checking of Access Control Policies Specified in XACML," IEEE 37th Annual Computer Software and Applications Conference, 2007, pp. 275-280.
[239] J. A. Wickboldt, W. P. De Jesus, P. H. Isolani, C. B. Both, J. Rochol, L. Z. Granville, "Software-defined networking: management requirements and challenges," IEEE Communications Magazine, vol. 53, issue. 1, 2015, pp. 278-285.
[240] B. Chang, R. Hwang, "Distributed Cost-based Update Policies for QoS Routing on Hierarchical Networks," Information Sciences, vol. 159, 2004, pp.87-108.
[241] S. Agarwal, A. Nucci, S. Bhattacharyya, "Measuring the Shared Fate of IGP Engineering and Interdomain Traffic," IEEE International Conference on Network Protocols, 2005, pp.236-245.
[242] P. Pham, S. Perreau, "Performance analysis of reactive shortest path and multipath routing mechanism with load balance," IEEE Conference in Computer and Communications Societies, vol.1, 2003, pp.251-259.
[243] S. Bryant, C. Filsfils, S. Previdi, M. Shand, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", IETF RFC 7490, 2015. [Online]. Available: https://tools.ietf.org/html/rfc7490
[244] R. Callon, "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments," IETF RFC 1195, 1990. [Online]. Available: http://www.rfc-base.org/rfc-1195.html
[245] G. Apostolopoulos, R. Guerin, S. Kamat, "Implementation and Performance Measurements of QoS Routing Extensions to OSPF," IEEE Conference in Computer and Communication Societies, vol.2, 1999, pp.680-688.
[246] R. Salles, V. Rolla, "Efficient Routing Heuristics for Internet Traffic Engineering," Computer Communications, vol.30, 2007, pp.1942-1952.
[247] C. Dsouza, G. J. Ahn, M. Taguinod, "Policy-driven security management for fog computing: Preliminary framework and a case study," IEEE 15th International Conference in Information Reuse and Integration, 2014, pp. 16-23.
[248] B. Claise, B. Trammell, Information Model for IP Flow Information Export (IPFIX), 2013, [Online]. Available: https://tools.ietf.org/html/rfc7012
[249] B. Claise, B. Trammell, P. Aitken, Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information, 2013, [Online]. Available: https://tools.ietf.org/html/rfc7011
[250] Cisco IOS Embedded Event Manager, 2013, [Online]. Available: http://www.cisco.com/
[251] J. Heinanen, T. Finland, R. Guerin, "A Two Rate Three Colour Meter," IETF RFC 2698, 1999, [Online]. Available: https://tools.ietf.org/html/rfc2698
[252] "Monte Carlo method," [Online]. Available: https://en.wikipedia.org/wiki/Monte_Carlo_ method/
[253] M. Casassa Mont, A. Baldwin, S. Shiu, "Identity Analytics - User Provisioning, Case Study: Using Modelling and Simulation for Policy Decision Support," HP Laboratories, HPL-2009-57.
[254] Wildfly Swarm. [Online]. Available: http://wildfly-swarm.io
[255] Spark. [Online]. Available: http://sparkjava.com
[256] Spring boot. [Online]. Available: http://projects.spring.io/spring-boot
[257] Light-4j. [Online]. Available: https://libraries.io/github/networknt/light-rest-4j
[258] M. Amaral, J. Polo, D. Carrera, I. Mohomed, M. Unuvar, M. Steinder, "Performance Evaluation of Microservices Architectures Using Container," IEEE Symposium on Network Computing and Applications, 2015, pp.27-34.
[259] D. Shadija, M. Rezai, R. Hill, "Towards an understanding of microservices," 23rd International Conference on Automation and Computing, 2017, pp. 1-6.
[260] M. Yadav, E. Hassan, G. Shroff, P. Agarwal, A. Srinivasan, "Searching for logical patterns in multi-sensor data from the industrial internet," In Machine Intelligence and Big Data in Industry, Springer International Publishing, vol.19, 2016, pp. 217-233.
[261] A. Tajalli-Yazdi, H. Lutfiyya, D. Kidston, "MANET security through a distributed policy-based evaluation of node behaviour," 2015 International Wireless Communications and Mobile Computing Conference, 2015, pp. 923-928.

[262] C. C. Sobin, V. Raychoudhury, G. Marfia, A. Singla, "A survey of routing and data dissemination in Delay Tolerant Networks,", Journal of Network and Computer Applications, vol.67, 2016, pp. 128-146.
[263] R. Di Pietro, S. Guarino, N. Verde, J. Domingo-Ferrer, "Security in wireless ad-hoc networks – A survey", Computer Communications, vol.51, 2014, pp. 1-20.
[264] P. Li, M. Pan, Y. Fang, "Capacity bounds of three-dimensional wireless ad hoc networks," IEEE/ACM Transactions on Networking, vol.20, no.4, 2012, pp. 1304-1315.
[265] S. Lim, W.C. Lee, G. Cao, C.R. Das, "Cache invalidation strategies for internet-based mobile ad hoc networks," Computer Communications, vol.30, no.8, 2007, pp. 1854-1869.
[266] J. Radak, B. Ducourthial, V. Cherfaoui, S. Bonnet, "Detecting road events using distributed data fusion: experimental evaluation for the icy roads case," IEEE Transactions on Intelligent Transportation Systems, vol.17, no.1, 2016, pp. 184-194.
[267] Q. Vey, S. Puechmorel, A. Pirovano, J. Radzik, "Routing in aeronautical ad-hoc networks," IEEE/AIAA 35th Digital Avionics Systems Conference, 2016, pp. 1-10.
[268] N. Islam, M. Hossain, G. Ali, P. Chong, "An expedite group key establishment protocol for Flying Ad-Hoc Network (FANET)," IEEE 5th International Conference on Informatics, Electronics and Vision, 2016, pp. 312-315.
[269] R. Helal, A. ElMougy, "An energy-efficient Service Discovery protocol for the IoT based on a multi-tier WSN architecture," IEEE 40th Local Computer Networks Conference Workshops, 2015, pp.862-869.
[270] A. Sher, N. Javaid, G. Ahmed, S. Islam, U. Qasim, Z. Khan, "MC: Maximum Coverage Routing Protocol for Underwater Wireless Sensor Networks," IEEE 19th International Conference on Network-Based Information Systems, 2016, pp. 91-98.
[271] J. Liu, Z. Wang, J. Cui, S. Zhou, B. Yang, "A Joint Time Synchronization and Localization Design for Mobile Underwater Sensor Networks," IEEE Transactions on Mobile Computing, vol.15, no.3, 2016, pp. 530-543.
[272] S. Kisseleff, X. Chen, I.F. Akyildiz, W.H. Gerstacker, "Efficient Charging of Access Limited Wireless Underground Sensor Networks," IEEE Transactions on Communications, vol.64, no.5, 2016, pp. 2130-2142.
[273] G. Toschi, L. Campos, C. Cugnasca, "Home automation networks: A survey," Computer Standards & Interfaces, vol.50, 2017, pp. 42-54.
[274] K. Zhang, X. Liang, M. Baura, R. Lu, X. Shen, "PHDA: A priority based health data aggregation with privacy preservation for cloud assisted WBAN," Information Sciences, vol.284, 2014, pp. 130-141.
[275] F. Hu, W. Siddiqui, K. Sankar, "Scalable security in Wireless Sensor and Actuator Networks (WSANs): integration re-keying with routing," Computer Networks, vol.51, no.1, 2007, pp.285-308.
[276] A. Barolli, T. Oda, I. Shinko, L. Barolli, F. Xhafa, M. Takizawa, "Performance Analysis of WMN-GA System for Different WMN Architectures and TCP Congestion-Avoidance Algorithms," IEEE 10th International Conference on Complex, Intelligent, and Software Intensive Systems, 2016, pp. 238-245.
[277] M. Zareei, E. Mahmoud Mohamed, M. Anisi, C. Vargas Rosales, K. Tsukamoto, M. Khurram Khan, "On-Demand Hybrid Routing for Cognitive Radio Ad-Hoc Network," IEEE Access, vol.4, 2016, pp. 8294-8302.
[278] H. Tran, G. Kaddoum, F. Gagnon, L. Sibomana, "Cognitive radio network with secrecy and interference constraints," Physical Communication, vol.22, 2017, pp. 32-41.
[279] A. Bicen, V. Gungor, O. Akan, "Delay-sensitive and multimedia communication in cognitive radio sensor networks," Ad Hoc Networks, vol.10, no.5, 2012, pp. 816-830.
[280] A. Nair, S. Asmi, A. Gopakumar, "Analysis of Physical Layer Security via Co-operative Communication in Internet of Things," Procedia Technology, vol.24, 2016, pp. 896-903.
[281] G. Zhao, Q. Liang, "On the uplink outage throughput capacity of hybrid wireless networks with Massive MIMO," Ad Hoc Networks, vol. 58, 2017, pp. 1-8.

Publications

The following journal and conference publications stem from the research work presented in this dissertation.

JOURNAL PUBLICATIONS

▪ G. Katsikogiannis, D. Kallergis, Z. Garofalaki, S. Mitropoulos, C. Douligeris, A Policy-aware Service Oriented Architecture for Secure Machine-to-Machine Communications, Elsevier Ad hoc Networks, July 2018.

▪ G. Katsikogiannis, S. Mitropoulos, C. Douligeris, "Policy-based QoS management for SLA-driven adaptive routing," Journal of Communications and Networks, Vol.15, issue 3, 2013, pp.301-311.

CONFERENCE PUBLICATIONS

▪ G. Katsikogiannis, Z. Garofalaki, D. Kallergis, Christos Douligeris, "PDA: A Policy-driven for authorisations scheme with μServices," 2018 International Balkan Conference on Communications and Networking, June 2018.

▪ Z. Garofalaki, D. Kallergis, G. Katsikogiannis, Ioannis Ellinas, Christos Douligeris, "A DSS Model for IoT-based Intelligent Transportation Systems," 2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'17), Bilbao, Spain, December 2017, pp. 276-281.

▪ Z. Garofalaki, D. Kallergis, G. Katsikogiannis, Christos Douligeris, "A Policy-aware Model for Intelligent Transportation Systems," 2017 International Balkan Conference on Communications and Networking, June 2017, arXiv preprint arXiv:1706.04803.

▪ G. Katsikogiannis, S. Mitropoulos, C. Douligeris, "An Identity and Access Management approach for SOA," 2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'16), December 2016, pp. 126-131.

▪ Z. Garofalaki, D. Kallergis, G. Katsikogiannis, Ioannis Ellinas, Christos Douligeris, "Transport Services within the IoT Ecosystem using Localisation Parameters," 2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'16), December 2016, pp. 87-92.

▪ G. Katsikogiannis, S. Mitropoulos, C. Douligeris, "Optimizing SLA-driven adaptive routing," 2016 IEEE International Symposium on Computers and Communications (ISCC'16), June 2016, pp.627-632.

▪ G. Katsikogiannis, S. Mitropoulos, C. Douligeris, "UACS: Towards Unified Access Control Services," IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'15), October 2015, pp. 127-132.

▪ G. Katsikogiannis, Mitropoulos S., Douligeris C. 2011, "A PBNM System for Adaptive Routing", Proceedings of the Eureka, 2011 Pan-Hellenic Conference, Kastoria, Greece, September 2011.