Type of the Paper (Article, Communication)

Horizontal Solutions for Cyber-Physical Systems evaluated in Energy Flexibility and Traffic Accident cases

Juhani Latvakoski[1*], Marc Roelands[2], Marko Tilvis[3], Laura Genga[4], Gabriel Santos[5], Goreti Marreiros[5], Zita Vale[6], Lode Hoste[7], Wolfgang Van Raemdonck[8], Nicola Zannone[9]

1 VTT Technical Research Centre of Finland; Juhani.Latvakoski@vtt.fi
2 Nokia Bell Labs; marc.roelands@nokia-bell-labs.com
3 Polar Electro; marko.tilvis@polar.com
4 Eindhoven University of Technology; l.genga@tue.nl
5 ISEP/GEGAD; gajls@isep.ipp.pt
5 ISEP/GEGAD; mgt@isep.ipp.pt
6 ISEP/GEGAD; zav@isep.ipp.pt
7 Nokia Bell Labs; lode.hoste@nokia-bell-labs.com
8 Nokia Bell Labs; wolfgang.van_raemdonck@nokia-bell-labs.com
9 Eindhoven University of Technology; n.zannone@tue.nl
* Correspondence: E-Mail: Juhani.Latvakoski@vtt.fi; Tel.: +358-40-5200-149; Fax: +358-20-722-2320

Finalized: 20th Dec 2019; Published: 1st Jan 2020

Abstract: The motivation for this article arises from the security, complexity and interoperability challenges in cyber-physical systems (CPS) especially in energy and mobility domains. Handling peak consumption hours and balancing power levels in the energy grids are becoming more and more expensive for the energy sector, energy intensive industry and consumers. Moreover, mobile appliances with the related CPS service providers (SPs) do not serve properly the needs of the owners due to the vertical silo type of operating model in CPS industries.

In this work, we propose three horizontal solutions for cyber-physical systems for more smart and trustworthy collaborations within multisector and vendor systems. In particular, we enable automatic energy flexibility trading required by real-time interactions between energy sector stakeholders and multiple sets of energy sensitive distribute energy resources (DERs). In addition, our solutions allow for trustworthy information sharing between physical devices and related CPS SP, and between multiple systems to enable real-time situation and location awareness services. Real-time applications and demonstrations of these results were made possible by the provided horizontal solutions for M2M service platform, communication spaces and policy-based authorization.

The evaluation shows that the streams based M2M service platform can make development of new services smoother. The communication spaces solution can enable controllable information exchange between physical CPS resources owned by different stakeholders. The policy-based authorization can enable consideration of the owners' policies in the referred information exchange process. The demonstrations indicate that the provided horizontal solutions can enable trustworthy collaborations of multisector and -vendor systems in the energy flexibility and traffic accident cases, and they are estimated to be applicable also to other respective cases. In addition, a number of directions for future research were identified related e.g. to information models of different verticals, trustworthy of heterogeneous CPS assets and especially secure and trust required processes of cyber-physical systems.

Keywords: cyber-physical systems; machine-to-machine systems; Internet of Things; horizontal service platforms; authorization

1. Introduction

Today, industries and consumer markets are increasingly using services exposed from wireless sensor and actuator networks. Such systems are also referred as machine-to-machine (M2M), Internet of Things (IoT) and cyber-physical systems (CPS). M2M typically highlights direct communication between devices; however, it is also used to refer to the exposed services, e.g. telematics, smart metering, remote maintenance of machines etc. IoT usually includes more abstract things, living objects, such as animals and humans, but also devices and machines and related (IoT) infrastructure. CPS usually refers to combined cyber-physical information-based operation loop with physical devices and back-office services, including communication, computation, monitoring and control. Examples of such CPS systems are smart grids, robotic systems, and automatic traffic systems. These terms all refer to entities and capabilities enabling the physical world to collaborate with the cyber-world, and therefore they are all used interchangeably in this article. The motivation of this article arises from on the challenges in such cyber-physical systems especially in the energy and mobility domains.

The energy domain comprises multiple cyber-physical entities, such as e.g. solar power plants, windmills, electric vehicles, buildings, energy sensitive household appliances, and other kinds of distributed energy resources (DERs) consuming, producing or storing energy. This complex environment requires distribution network operators to balance power levels in the energy grid. In fact, peak consumption hours are becoming more and more expensive for the energy sector, energy intensive industries and consumers. The industry is thus looking for more flexible and smarter operational models for the energy grid. Therefore, an essential industrial objective has been to reduce peak energy loads to lower the cost of energy and its' distribution. Such flexibility capabilities require interoperability between the cyber-physical entities from multiple energy sensitive domains and energy domain systems and stakeholders.

There are huge number of appliances in the mobility sector, such as smart watches, smart phones, sensors, tracking devices, vehicles, and multiple stakeholders, which are often not capable to interact with each other. This is because, in M2M business sectors, service providers (SP) usually host specialized physical resources (e.g., products of the SP) and the related exposed data/information in their own service cloud as a kind of vertical silo. However, the owners of the referred physical resources may not be the SPs themselves, but instead their customers, which may be individuals (private persons), companies, organizations, or their combinations. Today such stakeholders require ever more trustworthy, smart interoperable operation from their SPs, which is in-line with their own collaboration agreements and privacy regulations[1]. For example, if some emergency type event occurs, it is challenging to know the situation, even if the information could be available. These situations require smart operation capabilities and interoperability between the appliances from multiple sectors, hosted by multiple SPs, and SP service systems/clouds.

Thus, the core reasons for the challenges demonstrated by both energy flexibility and traffic accident related cases are the need for smartness, trustworthiness and interoperability. These needs are visible also in larger contexts in the computer world, where billions of sensors interact with billions of things worldwide and operate with millions of applications and services. In such a future world, networking, distributed operation and virtualized computing in a distributed worldwide infrastructure and its' smart design is no longer a "nice thing to have" but a necessity. The key enabler

---

[1] GDPR, General Data Protection Regulation (EU) 2016/679

towards such infrastructure is here envisioned to be horizontal solutions, which can be applied in multiple domains/sectors and multiple stakeholder systems. As the results of this research, carried out within European research collaboration project[2], we propose such horizontal solutions for cyber-physical systems and evaluate them in the energy flexibility and traffic accident cases with real-time demonstrators executed in laboratory environments.

The remainder of the paper is organized as follows: Related work of the cyber-physical M2M systems is provided in Section 2. The challenges and requirements of the focused cases and transfer towards horizontal architectures are discussed in Section 3. The provided horizontal solutions for such cyber-physical systems are explained in Section 4, and evaluation results are provided in Section 5. Finally, concluding remarks are presented in Section 6.

2. Related work

There are several specifications, industrial forums and even standards targeting industrially relevant cyber-physical M2M systems. Separate specifications have been developed for each sector/vertical domain, such as e.g. home/buildings, manufacturing/industry automation, vehicular/transportation, healthcare, energy, cities, wearables etc. However, there are also initiatives aiming at cross-domain, horizontal type of IoT platforms [1, 2, 3, 4, 5, 6, 7, 8]. In addition, several standardization bodies like IEEE, OMG, W3C, OpenFog, AllSeen alliance, NGMN and ISO/IEC JTC1 WG41 have been working in the area. There are also other recent related actions ongoing such as e.g. securing IoT products with blockchain [9], and comparisons and related studies of IoT Platforms. For example, Guth *et al.* compare OpenMTC, Fiware, SiteWhere, AWSIoT and provide IoT specification with IoT integration middleware, Gateway and Devices as basic building blocks [10]. Burg *et al.* focuses to review wireless communications and security technologies for cyber-physical systems and conclude that security is an essential challenge for wireless cyber-physical systems operating in horizontal way across multiple domains [11].

Essential solutions in the referred horizontal IoT specifications are related to the edge system and IoT platform. The edge system usually comprises identifiable physical entities, which can be connected to IoT infrastructures and platforms either directly or via some sort of gateway [12, 13, 14, 15]. An IoT platform is typically an integrated physical/virtual entity system capable of controlling, monitoring, information processing and application execution. There are also typically different kinds of tiers defined according to the accessibility of the entities, platform and enterprise systems. The information models are used to define the properties of IoT information content. In addition, several studies have investigated how virtualization capabilities of IoT systems can be deployed at the edges of the network.

We apply in this work most of these elements in our solutions. However, there are a number of challenges that are not addressed by existing IoT specifications. A main challenge concerns the collaboration between consumer and industrial endpoints, which is not supported properly by any of the existing specifications. There are also fundamental requirements towards crosscutting solutions in the areas of security, safety, interoperability, composability, data management, analytics, resilience, composability, virtualization, and regulation. We rely on the principles for creating horizontal solutions, which can be deployed in multiple domain scenarios, and minimizes the need for application domain specific solutions [16]. This is especially challenging when speaking about the information and service level, which easily mix the domain and potentially horizontal generic services when handling services related to information streams.

The idea of using graphs of operators for stream processing queries has historically been explored in systems like Borealis [17] and STREAM [18], and more recently in frameworks like Millwheel [19], Storm [20], Heron [21], Spark Streaming [22] and Samza [23]. Most of these frameworks focus on data-parallel processing of partitionable streams and strong fault-tolerance guarantees within a *single* data centre. In contrast, we focus on *distributed and edge computing across*

---

[2] ITEA3 M2MGrids project, https://itea3.org/project/m2mgrids.html

wide-area cloud-integrated networks, modelled as a federated set of geographically widespread interconnected compute nodes, stream sources and sinks. While other stream processing platforms address edge locality and distribution within a confined machine cluster, like Quarks [24] and System S [25, 26], our solutions are few versatile and address wide-area cyber-physical systems. Further, 'Big Data' processing frameworks like Hadoop [27] and Dryad [28, 29] also model queries as data flows. Systems like Hive [30] and Sawzall [31] work with a query language and query plan optimization on top of the MapReduce paradigm [32]. However, they do not consider the wide-area setting, or, like CHive [33], do consider it exclusively from a data analytics perspective, which does not support the open-ended sink distribution, as e.g. needed in energy grid automation scenarios. Therefore, we adopted in this work the Nokia World Wide Streams (WWS) platform [34, 35], a horizontal, stream-processing-based M2M service platform, as a suitable starting point for our solution.

The dynamic changes in IoT systems, such as continual adding of physical entities involving OEMs and related SPs, heterogeneous sensors and actuators and other devices, have proved to be challenging especially from a security point of view. Industrial IoT devices are often used in physically protected and isolated environments; however, today there is the need to enhance the operation also with many other devices (e.g. for energy flexibility). State-of-the-art IoT specifications lack proper solutions for solving this challenge. Proper identification, authentication and authorization capabilities seem to be missing for dynamic IoT environments, which prevents establishment of trust relationships. Therefore, uncertainty exists in information ownership and validity, and to remote management of the physical assets including reconfiguration and reprogramming on the fly. Therefore, we contribute to increase the level of trust in communications between physical cyber-physical resources owned by different stakeholders.

The challenges of communication with CPS entities arise from accessibility, trustworthiness, heterogeneity, mobility, and ownership of physical entities [36]. The interaction with physical resources (devices) over the communication networks shall concern especially the ownership, communications, and trust aspects. Our solution relies on the application of communication spaces concept, which provides a virtual home for people and their resources [37]. The respective idea has also been applied, for example, in the concept of virtual home environment for supporting roaming ecosystems in the mobile context for the mobile device of a user [38], and for sensor devices [39]. CPS systems usually require capability to deliver information from one source to one (1–1) or multiple destinations (1–N) according to the needs of the destinations, which refers to application of publish-subscribe paradigm [40]. For example, Extensible Messaging and Presence Protocol (XMPP) [41, 42, 43], message queue based telemetry transport (MQTT) [44, 45], data distribution service (DDS), advanced message queuing protocol (AMQP), and simple text oriented messaging protocol (STOMP)) or broker-less (e.g., zero message queue protocol (ZeroMQ)) has such capabilities. Naming and identification of owners are provided e.g. by electronic email systems (SMTP, RFC 5321 and 5322), Post Office Protocol (POP), and Internet Message Access Protocol (IMAP)), session initiation protocol (SIP/SIMPLE) [46] and XMPP, which support also management of relationships between owners. Presence management of mobile CPS resources could be supported by e.g. SIP/SIMPLE and XMPP, and e.g. the smart instant messaging (SIM) system provide solutions for presence management, user centric configuration and adaptive grouping [47]. The CPS information content heterogeneity requires multiple data sharing methods ranging from constrained systems (e.g., constrained application protocol (CoAP) for supporting REST-like applications), messaging with known topic names (e.g., MQTT), multimedia content (e.g., SIP) and domain specific content (e.g., energy domain [63]). The capabilities to control who can use the resources could be supported by XMPP "Buddy list" service; however, there is lack of means for defining more specific rules and conditions for the information that can be shared considering e.g. aspects of privacy. Based on the analysis, none of the existing technologies is capable to fulfil all the required aspects. In this research, we selected the specific features from XMPP and MQTT, and combine with the trust solutions to operate with dynamic resources owned by different stakeholders when exchanging information, and adopted the VTT CPS hub solutions (VTT CPS hub) as the basis for development [48]. Thus, the referred communication spaces concept, and its reference realization (VTT CPS hub) are in this article elaborated and validated in the real-time energy flexibility and traffic accident cases.

To date, trust, security and privacy are among the primary concerns that limit the widespread adoption of IoT. Roman et al. [49] analyse the features and security challenges such as interoperability and management of access rights with respect to various IoT architectures. This study shows the need to integrate centralized and distributed architectural approaches to provide the foundation of full-fledged IoT. However, this integration has an impact on the efficacy of authorization mechanisms. Ouaddah et al. [50] perform a detailed analysis of existing access control solutions for IoT with respect to domain specific IoT requirements. Alonso et al. [51] identify a number of requirements (related to application-scoped, client-independent, flexible, delegated and configurable) that need to be addressed in order to devise more effective access control mechanisms tailored to IoT ecosystems. Ho et al. [52] examine the security mechanisms employed in home smart lock solutions. Their study unveils flaws in the architectural design and communication models of existing locks, which an attacker can exploit to learn confidential information about users and even gain unauthorized access to the house.

3. Towards Horizontal Architectures of Cyber-Physical Systems

3.1 Challenges and requirements of targeted use cases

The starting point for this research has been the challenges and requirements of the focused industrial applications related to energy flexibility and traffic accident use cases. The major requirements, identified in a European research collaboration project, were arising from the needs of energy and mobility sector stakeholders, to exchange information and provide services in cost efficient way, and so that policies of resource owners are respected. Next, we provide an overview of the main requirements arising from the targeted CPS related use cases.

The recent changes in the energy domain have led to the need to reduce energy peak loads to lower the cost of energy and energy distribution. Obtaining such flexibility capabilities requires *interoperability* between multiple energy-related domains, systems and stakeholders (Requirement 1, R1). Consider a collection of buildings, each of which has a specific set of energy resources (white goods, boilers, heaters, and so on), and electric vehicles. Such distributed energy resources (DER) need to be controlled by at least one energy *aggregator* (R2), which shall be able to operate in a local energy *market* (R3). There can be *multiple* such aggregators operating on the referred local market (R4). A distribution service operator (DSO) need to be able to *monitor the load balance* in the electrical grid (R5), be able to make a *forecast* of balance for the next day (R6), and compare the monitored balance with the forecast in *real time* (R7). When a DSO detects discrepancies in the current and predicted load balance, he must be able to ask for *flexibility* on the local market, i.e. a capacity to change or shift the energy production or consumption for a certain amount of power in particular time intervals (e.g. in 15-minute slots on a 24-hours horizon) (R8). The energy aggregators on the local market shall be able to *offer* flexibility (R9a), e.g. by means of time shifting energy consumption or production of the DER they control, considering energy market prizes (R9b) and restrictions as set by the DER owners (R10). For example, the energy consumption of white goods, boilers, heaters or electric vehicle charging may well happen earlier or later than a presumed time, or may happen at a lower or higher pace, i.e. power, within the bounds set by the asset owner. After completing the transaction on the local market, awarded aggregators shall be able to *fulfil* the required flexibility with the related DERs, by *instructing* the individual devices (R11). As a result, the local electrical grid balance is maintained better, with *flattened, lowered discrepancies* between energy production and consumption, in effect having the finer-grain consumers/producers *follow the fluctuations of (renewable) energy production* (R12). The DSOs can save on energy storage investments (R13a), new aggregator businesses can conduct sound, revenue-generating business cases (R13b), and households and vehicle owners get energy bill rebates for their flexibility contribution (R13c). Enabling energy flexibility automatically and in real-time is needed for fulfilling the real needs of the energy sector,

that is challenged by the introduction of ever more *renewable but intermittent* energy sources (R14a), and the *ever more distributed* nature of both productions and consumption (R14b).
Electric vehicles (EV) can be considered a specific category of DERs, which can be used as an energy consumption point (charging of the EV batteries), production point (discharging of EV batteries) or energy storage (time-shifting of consumption or production). In addition, EVs represent an essential asset to the mobility sector, targeted in the second use case in this research, the traffic accident case. In such a case, it is required to have means for surrounding physical asset devices to interact with each other in *interoperable* way (R15). However, the *security, trust and privacy policies* and agreements of the device owners shall be taken into concern in the referred interactions (R16). When a traffic accident happens, smart interoperation capabilities and interoperability between the devices from multiple sectors, hosted by multiple SPs, and SP service systems/clouds is required (R17). The information services exposed from the referred devices shall operate according to the requirements of their *owners* (R18), without compromising the business values of related *service providers* (R19). Consider people having smart appliances such as a smart watch, vehicles on the road, lighting infrastructure installed beside the road, hunting dogs with trackers running and various sensors mounted more or less statically in the environment. Let's assume that the dog caused a traffic accident, and the person that happens to be on the road calls to alarm centre. There is need for the authorities such as the police and medicals to receive information from the situation. However, the policies of the information owners shall be considered (R20). When an alarm happens, it shall be able to change the operation policies of the owners (R21). If the owners want, the information delivery after the authorized alarm event in a specific geographical area shall be possible to be changed so that the authorities will get the information from the devices (R22). Based on this information sharing process, a situation aware view what is happening in the emergency area shall be possible to be visualized for authorities (R23).

The requirements of the energy flexibility case (R1-14) and traffic accident case (R15-R23) summarise at a high level what came out of an extensive requirements analysis for the targeted use cases[3]. From these, the project analysed business and technical horizontal platform requirements that are *common* for the involved use cases. A common understanding of *information meaning and exchange format* is needed for interoperability between resources in the system (R24). However, smart information-based algorithms and operations need to apply detailed sector/application/case specific data (R25). The information service system shall be able to handle resources dynamically (R26), support customization according to business application and personalization (R27), enable execution of services in distributed manner (R28) and scale to needs of many use cases and businesses (R29). The communication system shall be able to support communications over heterogeneous accesses (R30), heterogeneous device types (R31) and scale to needs of multiple use cases and business applications (R32). The system shall be secure and reliable, comply with privacy and legal restrictions, and scale to needs of multiple use case and businesses (R33). This summarises the common requirements[4] for all the targeted use cases establish the framework of the concept development in this research.

3.2 From requirements towards horizontal architectures and solutions

In the requirement analysis, significant commonalities were identified, justifying the potential for horizontal solutions and applicability of horizontal architectures. Such horizontal architectures are expected to be applicable in multiple domains/sectors and multiple stakeholder systems [16]. An analysis of the position of identified common requirements (R24-33) in horizontal architectural levels,

---

[3] 116 detailed business and technical requirements were detected in the targeted use cases in the M2MGrids project.

[4] 46 detailed business and technical requirements were estimated to be common for all the targeted use cases in the M2MGrids project.

with main actor types of the CPS system, are presented in Figure 1. A CPS system can be divided in an application-domain-dependent and an application-domain-agnostic part. The vertical services offered by a specific vertical business actor, a vertical sector service provider such as e.g. a DSO or an energy aggregator in the energy sector, are determining the service logic and requirements. These actors request their services to be performed by the application-domain-agnostic layers of the system, by passing a programmatic description for the services to the horizontal platform. As parts of the horizontal solution, we distinguish the M2M service platform, the horizontal communications and the horizontal trust and security sub-solutions.

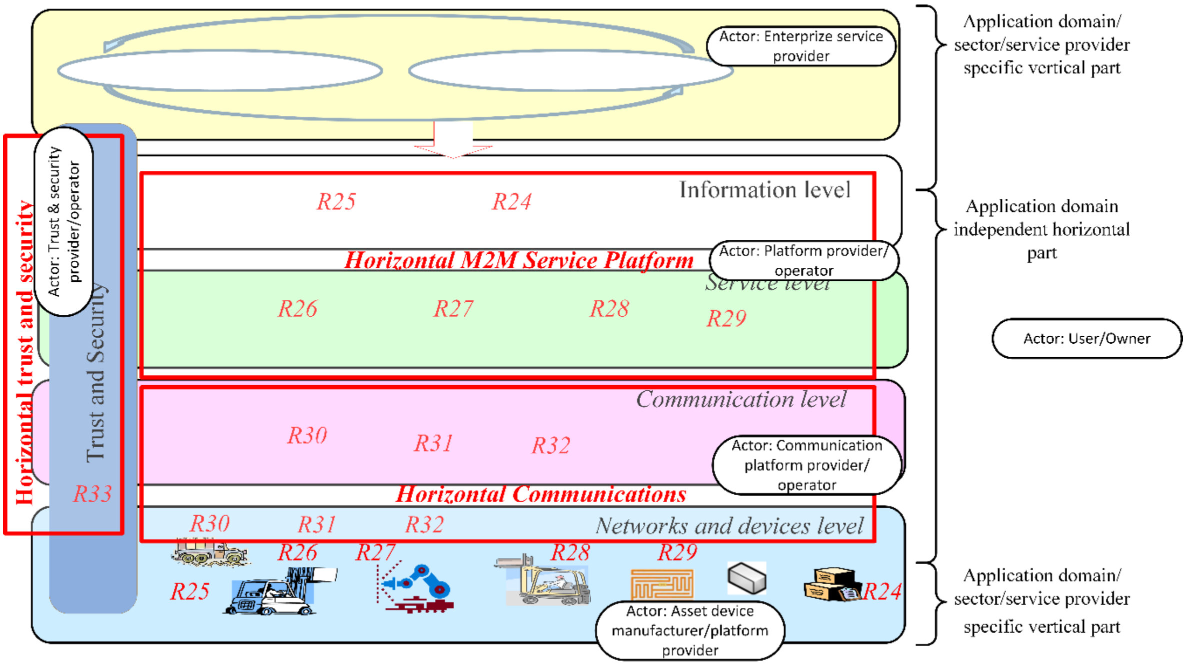

Figure 1. Identified common requirements, distinguished horizontal solutions (red rectanges) in the architectural levels of the CPS system [16].

The M2M service platform presumes at least four detailed actor roles. First, the actors *building the M2M service platform technology* either sell or rent out the platform as solution equipment directly to a customer who then hosts it, or they offer it *as-a-Service (*aaS)*, in a PaaS (Platform-as-a-Service), IaaS (Infrastructure-as-a-Service) or XaaS (Everything -as-a-Service) model. Secondly, the *platform operator* operates the M2M service platform as a vertical-business-agnostic service infrastructure. This role may be taken up by traditional telecom service providers, but also by large Internet data players and vendors, or large verticals, such as conglomerates of large cities or multi-national utilities. Thirdly, the vertical business actors (enterprise service provider in the Figure 1), as *platform users*, are conducting their vertical business by launching services on the platform, either directly or via a dedicated vertical application front-end. Vertical sector service providers, such as energy sector stakeholders, typically take up this role, but also new actors can enter the ecosystem in this role, like third parties with a specific cross-sector business model and services. Finally, *service users*, consumers or other businesses, are the customers using the offered services. They consume the services offered by the vertical players but can also add their own service configuration logic to the platform. The M2M service platform brings key advantages to all the detailed actor roles.

A key advantage to platform users - and to platform operators offering a vertical, domain-specific framework on the platform - is that the *platform hides the complexities of distributed deployment of a collection of data-stream intensive services*. Service programming, and ultimately even high-level specification, can be done in an infrastructure-agnostic way, enabling vertical service providers to build services just in terms of the concepts, conventions and algorithms from their own business domain. But at the same time, the automation brought by the platform and its distributed stream processing capability, also benefitting from emerging new underlying 5G and edge cloud network technologies, supports their business without needing them to care about the technology.

This separation of concerns is also an important advantage to the *platform operator*, in this way providing the platform operator more control over platform execution resource efficiencies, for which the platform supports various optimization options. This becomes most apparent when multiple platform users launch different application cases on the shared service infrastructure of the platform, not necessarily knowing each other's objectives. E.g., if two service instances require the same pre-processing of the same stream, there is a stream processing reuse opportunity. E.g., the real-time evolving sum of values in energy consumption of a set of devices, as appearing in the programmatic description of multiple services, may be a reused data stream.

Ensuring that cyber-physical systems offer an acceptable *security level* requires addressing challenges related to privacy, confidentiality, end-to-end security, secure data management, and access control. The horizontal trust and security and communication solutions should ensure that the communication and exchange of information is allowed only according to the privacy and security policies of the owners and business actors.

Human owners of physical M2M resources today usually employ services of multiple M2M SPs and cloud storage services according to their needs. However, cloud storage systems are often tightly bounded with specific cloud service providers. For example, Apple’s iPhone is tightly integrated with its cloud storage service iCloud, Android phones with Google Drive, and Windows phones with OneDrive [53]. Physical M2M resources such as e.g. wearables, home automation devices, machines etc. are tightly bounded to operate with specific M2M SPs.

Therefore, owners have the additional challenge of managing access rights across several SPs that host their data. Currently, there is a lack of proper means for owners to monitor and control the policies used to regulate access to their data. Furthermore, each SP usually uses its own specific solution for access control. This requires owners to redefine their policies for each SP system, which causes difficulty in synchronizing them across multiple SPs. The situation becomes even more challenging when owners want to share their data with other owners who are not registered to the same SP, and SPs do not have any means of verifying authorizations for unknown users.

The successful application of federated identity services in cloud environments [54], where services allow their users to login based on credentials provided by third-party identity providers such as Google and Facebook, indicates that similar approaches in the context of authorization may be able to solve the issue of data sharing in a multiple clouds setting. Existing access control standards such as XACML [55] already provide a baseline for the development of authorization services for cloud environments [56]. However, current authorization solutions, even those based on XACML; suffer severe limitations that hamper their adoption in these cloud environments.

A high-level functional overview, positioning the three provided horizontal solutions, is shown in Figure 2Figure 2. Domain knowledge and requested service specifications are provided by vertical sector service providers, as platform users of the horizontal M2M service platform. The M2M service platform operation is based on stream processing with specific compilers and tools working with world model, resource topologies and policies, run-time registers, and distributed execution engines. The M2M service platform concept has been realized in the form of the Nokia World Wide Streams (WWS) platform, which serves as a reference implementation of the concept [34, 35].

The concept of communication spaces applies the owner-centric approach for communication with physical CPS resources (CPS devices) [37]. According to it, each owner has a virtual communication space into which the presence of CPS devices is registered. Resource owners can make agreements with each other to enable information exchanges between their resources. When the resources want to share information, the trust and policies between the owners need to be checked before the information can be delivered. The communication spaces concept has been realized in the form of horizontal VTT CPS communication hub, which is thus a reference implementation of the concept [48].

The concept of policy-based authorization is targeted to enabling owners to define and manage access control policies to protect their resources. These policies are then evaluated at runtime by the authorization service to determine whether access to the resources should be allowed or not. The policy-based authorization concept was originally realized in the form of SAFAX [53], an XACML-based authorization service framework developed by Eindhoven University of Technology (TU/e), and it is enhanced and applied here as a reference implementation of the concept.

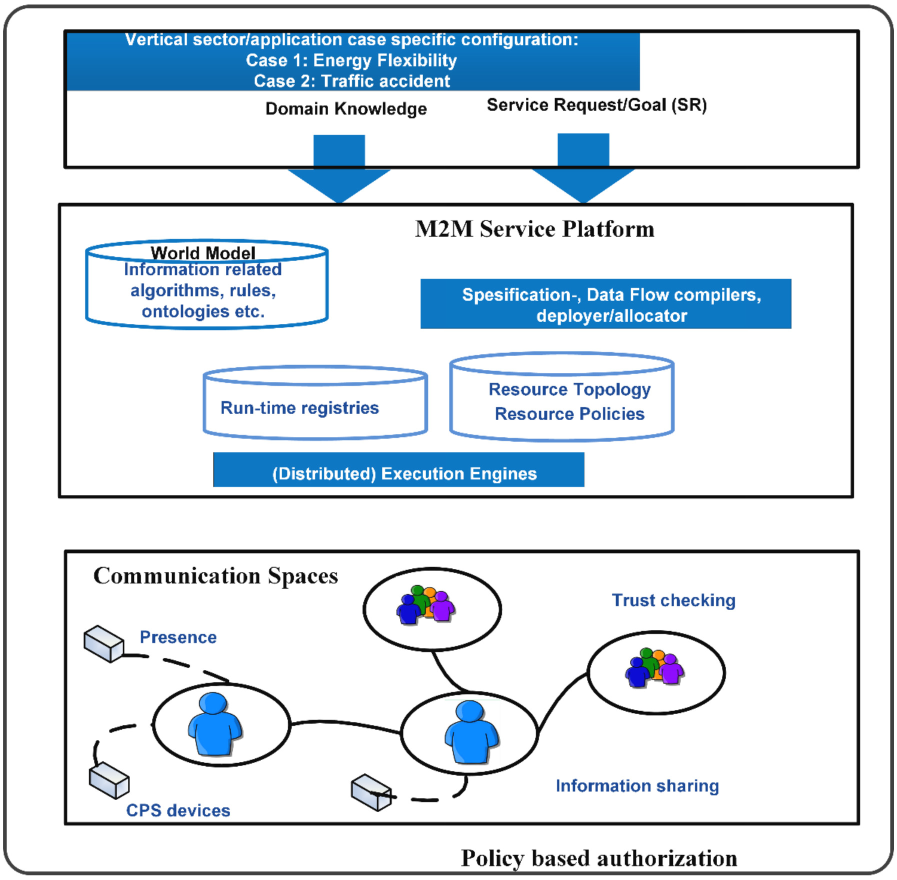

Figure 2. Conceptual solutions for horizontal architectures of cyber-physical systems.

4. Horizontal Solutions for Cyber-Physical Systems

4.1 M2M service platform

In this section, we discuss the main technical design aspects of the M2M service platform, as shown in Figure 3. After a brief report of the experiments done concerning World Model and Specification Compiler, we highlight the main features of key components of the Nokia World Wide Streams (WWS) platform [34,35] implementing the M2M service platform.

Platform users can specify and launch services directly on the platform, as so-called *Service Requests* (SRs), assuming the *World Model* as the domain of discourse, and a declarative domain-specific language (DSL) called *XStream* as a means to specify the actual service logic. Assuming a set of given execution primitives of the underlying execution platform, a *Specification Compiler* stage translates the SRs into *Service Execution Requests* (SERs), by reasoning over the desired SR goal in the context of the World Model's domain knowledge, possibly in combination with already expressed, outstanding goals. As a result, the Specification Compiler expresses a discovered set of possible stream processing dataflows enriched with candidate optimization transformations, as a set of implementation options in terms of the given execution primitives for the requested service.

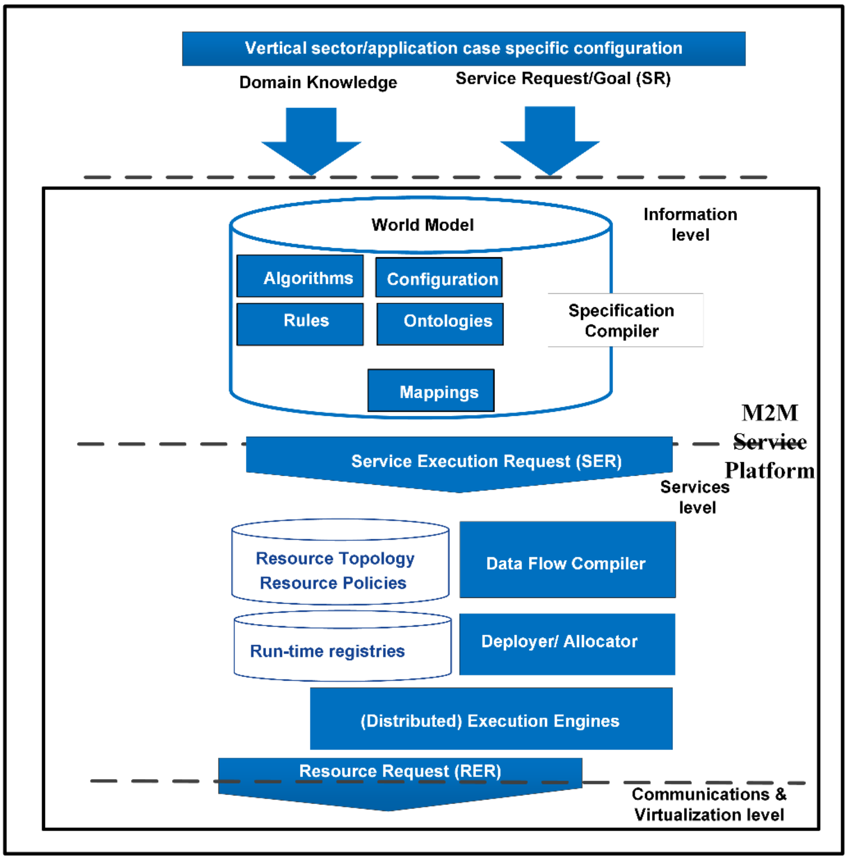

Figure 3. M2M service platform architecture based on the Nokia WWS platform [34, 35].

The flexible service interpretation in the SER is passed on to the actual M2M service platform, in which the *Dataflow Compiler* and *Deployer* can now decide which service implementation options to select, to minimize compute and transport resource spending for the service. The *Deployer/Allocator* maintains the lifecycle of the requested service execution after distributing the dataflow physically across the chosen execution engines, that can reside in central or edge data centres as well as in device gateways and programmable devices. *Run-time registries* keep track of which resource allocations correspond to which service instances and data streams, including sources and sinks of data at system boundaries connecting physical devices and human interfaces. Possibly dynamically appearing and disappearing streams, from devices or the web, are registered as data stream *sources* (platform ingress) or *sinks* (platform egress).

4.1.1 World Model and Specification Compiler

The World Model provides a way to represent the domain knowledge context of requested services in the considered business domain. Beyond ontological concepts and their attributes, described in standard RDF/OWL notation, the model may contain more complex relations and behavioural correlations in the domain at hand. For example, an energy domain world model is built up from the base energy concepts and their physical properties, as well as the business and physical rules they obey, including the specification of metering and control interface protocols. The World Model may grow to support concepts needed in ever-smarter services involving more actors. For example, during the lifespan of the platform in the energy domain, services can evolve over time, originally being just basic smart metering services, later becoming full energy flexibility ecosystems of many cooperating actors, and services involving behavioural or business strategies as well as, e.g., real-time energy consumption and production predictions.

In the energy domain, we aimed to leverage as much as possible *existing* ontologies and application-level standards. We considered ontologies like SEAS [57, 58] and SAREF [59], and schemas provided by electrical (pre)standards, such as EFI [60, 61, 62] and CIM [63]. Figure 3 introduces snippets of the M2MGrids energy domain World Model.

Further attributes and relations were added, extending the base M2M energy domain World Model, as needed to specify and implement the energy services demonstrated in the project. A common semantic base, including the alignment of equivalent concepts from different ontologies, is also aimed at, for allowing automating the translation of messages assumed by different services, both among services authored on the platform by different service providers, and with externally authored or pre-existing ones. As an example of that, Figure 4 and Figure 5 illustrate the semantic model for the request and response messages of the FlexOptimizer service, which was implemented on top of WWS in a multi-service energy scenario in which the flexibility of energy consumption of household devices was matched against an overall flexibility need on an energy market. The messages allow negotiating a time shift of the energy consumption of individual devices.

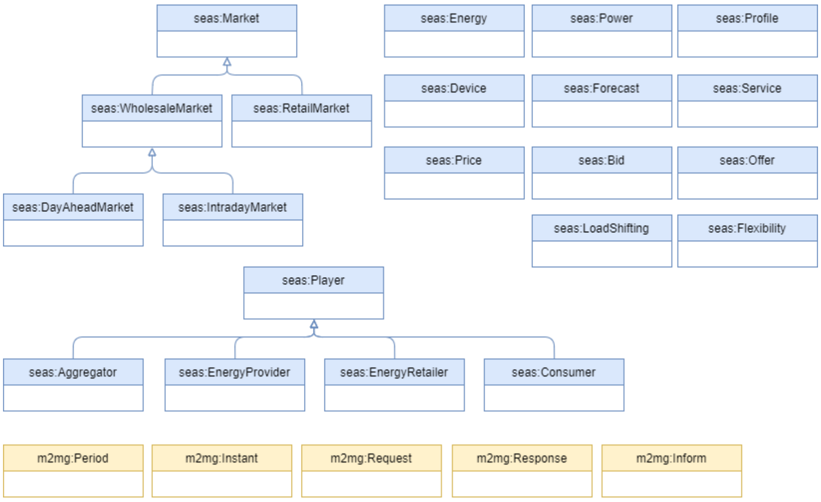

Figure 3. M2MGrids energy domain World Model snippets

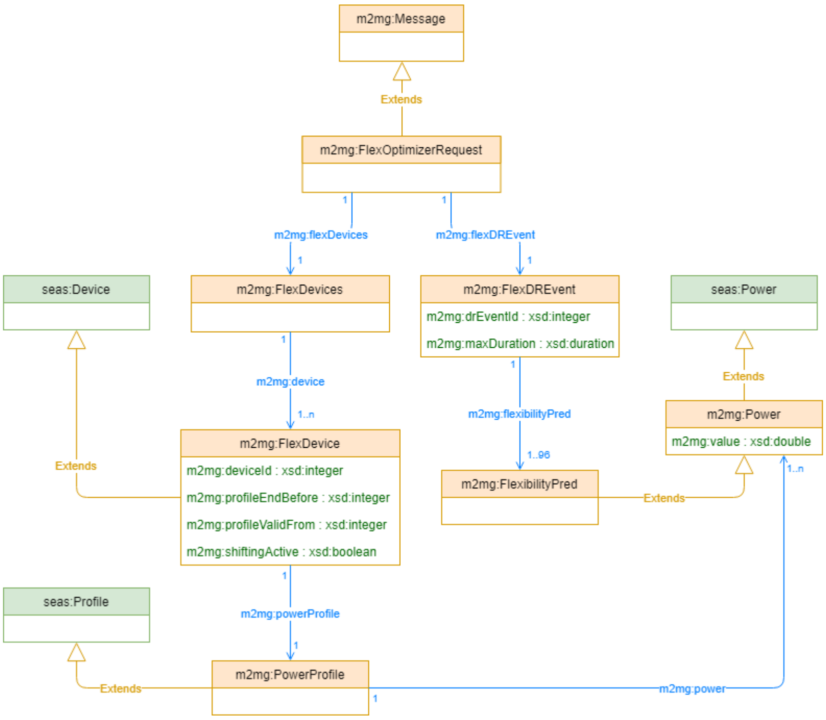

Figure 4. Semantic model of the FlexOptimizer request message.

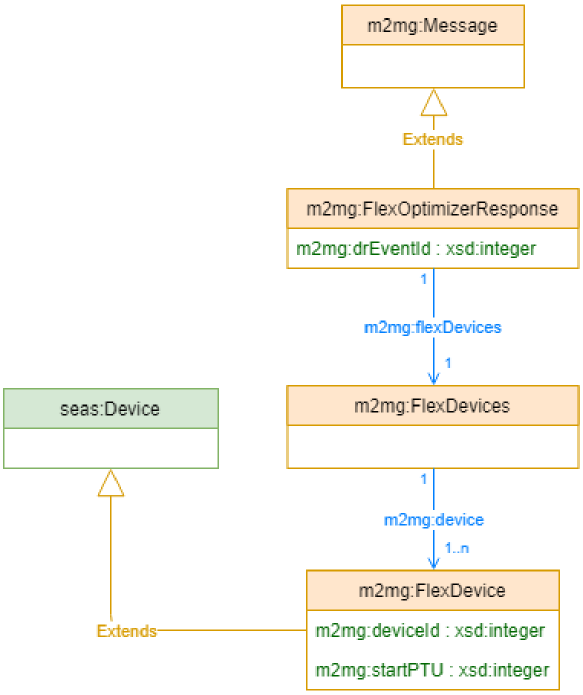

Figure 5. Semantic model of the FlexOptimizer response message.

The extension of the World Model with such aligned application ontologies makes the system 'aware' of the fact that different concepts and protocols have the same meaning and serve the same purpose, thus facilitating the translation of messages and protocol states across services from different service providers. Once aligned, one can say the platform offers a single 'lingua franca' among platform users launching services that need to cooperate on the platform. This especially holds for the semantics of the data streams entering or leaving the system, interfacing with legacy and other externally realized services.

The World Model thus also contains the observation and actuation concepts for the sensors and actuators registered in the system, like the measurement of real-time power consumption values of an energy-consuming device. Beyond that, the World Model can also encompass behavioural patterns of the real world, e.g. in the form of (pre-trained) predictive models for physical state, e.g. the battery level of an electrical vehicle under observation can be assumed to behave according to a particular model according to the vehicle's movement and acceleration. A service provider can include such models in the World Model for reuse as common domain knowledge.

When platform users issue Service Requests (SR), these requests can be considered as *service specifications*, that declaratively expresses exactly *what* is the requested effect of the platform on the outside world, as opposed to *how* it will eventually be implemented and deployed on the platform. These *increments* to the World Model, also based on the base concepts already in the model, need to be processed by *the Specification Compiler* of the system, to translate them to actual data flows to be executed on the platform. Early experiments concerning such *Specification Compiler* have been conducted based on a Drools-style [64] declarative rule-based language. We succeeded in generating candidate data flows as Service Execution Requests (SERs) for the underlying WWS XStream Dataflow Compiler, for basic service examples.

4.1.2 Dataflow Compiler

The Service Execution Request (SER) interface is the main interface of the Dataflow Compiler of the M2M service platform, exposing a range of SER execution primitives (stream processing operator specifications, in the case of WWS). The Dataflow Compiler transforms the received *logical* data flows, into a *physical* dataflow ready for deployment by the Deployer/Allocator.

In WWS, the Dataflow Compiler exposes the *XStream language* as a SER interface. It extends TypeScript, which is a typed JavaScript extension. Services can thus be authored directly on the Dataflow Compiler as *XStream scripts*, which describe how *stream processing operators* are to be wired up in a logical graph, as a *dataflow*. Figure 6 presents a generic representation of a WWS XStream dataflow. To WWS programmers, XStream is available as a *TypeScript* library, and so can easily leverage regular integrated development environments (IDEs).

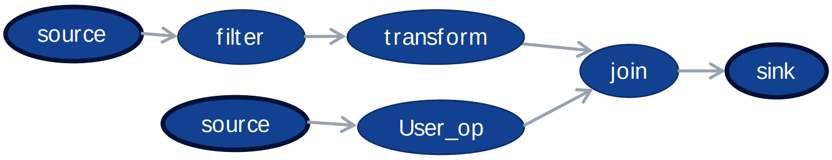

Figure 6. Generic representation of a WWS XStream dataflow

As execution primitives are natively part of the language, all classical stream processing operators, such as *join*, *transform* and *filter*, can be applied to *source* (ingress) streams, to produce *sink* (egress) data streams. Those XStream-native operators are implemented in the XStream compiler as dynamically loadable JavaScript functions, assuming a NodeJS execution environment.

Beyond this, XStream allows working with *any user-defined operators/functions (UDF)*, if an implementation for the operator has been *onboarded* in the WWS system. This powerful feature allows working with any type of operator implementation technology from the XStream code. Next to NodeJS, a range of technology-specific execution environments - called processors in WWS - are commonplace in the system, with operators already implemented in Java, Python and C++. Many pre-built operators are also available in WWS out of the box, e.g. for advanced video analytics and TensorFlow-based online machine learning operations.

The XStream service author can use all these operators in a service design, aggregating, transforming and generating data streams, mixing various stream types, without needing to know the internal implementation details of the used operators. A technology expert can build the dedicated operators independently, outside the service design process. While not as 'decoupled' from the implementation as is ultimately possible using a Specification Compiler on top of it, versatile XStream scripts can be designed flexibly in this way, without burdening the designer with the many technological details beyond the data stream formats used, and without the need to worry about how and where to deploy the individual pieces, i.e. in a true 'serverless' way.

The XStream scripts are compiled by the Dataflow Compiler into a deployable form. By this, the task of *optimizing* the distributed deployment and other *lifecycle concerns* are hidden from the XStream programmer. Thanks to the underlying event brokers, and media servers for multimedia streams, the actual source streams can be bound to a service instance dynamically at run-time. The XStream code can express WWS Registry topic name subscription to a single or group of streams for this. This again alleviates the service designer from everything but composing the logical operators and streams.

4.1.3 Deployer/Allocator

The XStream Dataflow Compiler is translating XStream code into a collection of physical operator instances and wiring code, as an explicit deployment instruction for the Deployer/Allocator. The Deployer/Allocator is an extensive subsystem that oversees the complete run-time of the running services. It takes as input Dataflow Compiler deployment instructions and, based on the requested stream processing operator instantiations and wiring among them, decides on what computing resources (i.e. on which machines) to deploy operator instances as used by the services. In the process, the Deployer checks whether already running operator instances can be reused or clustered on the same machine and decides on the eventual distributed placement across multiple cloud sites and edge equipment. A *placement algorithm* (for an introduction on the state of the art see [65, 66]) can calculate, given a set of conditions (including privacy constraints), what is the most cost-effective placement of the operators, as part of the dataflow(s) they belong to, taking into account the cost of occupying transport (network) and computing resources. The Deployer uses the (network and computing) resource topology and potential resource policies in the process.

Key elements of the overall Deployer-managed process are the *Run-time registries* that keep track of which operators are deployed where, what their deployment state is, and eventually what data streams they produce. The registries also persist the *WWS site's configuration*, contain a record of all active services, and maintain a range of meta-data for each stream instance's context of use.

The deployment lifecycle of operators is illustrated in Figure 7. It comprises five different services, of which three have been successfully deployed completely (all operators are green), one service has a failed operator deployment (red), and one service has an operator that is still being deployed (staging phase, yellow). When a problematic operator state occurs, the Deployer attempts to automatically redeploy the operator.

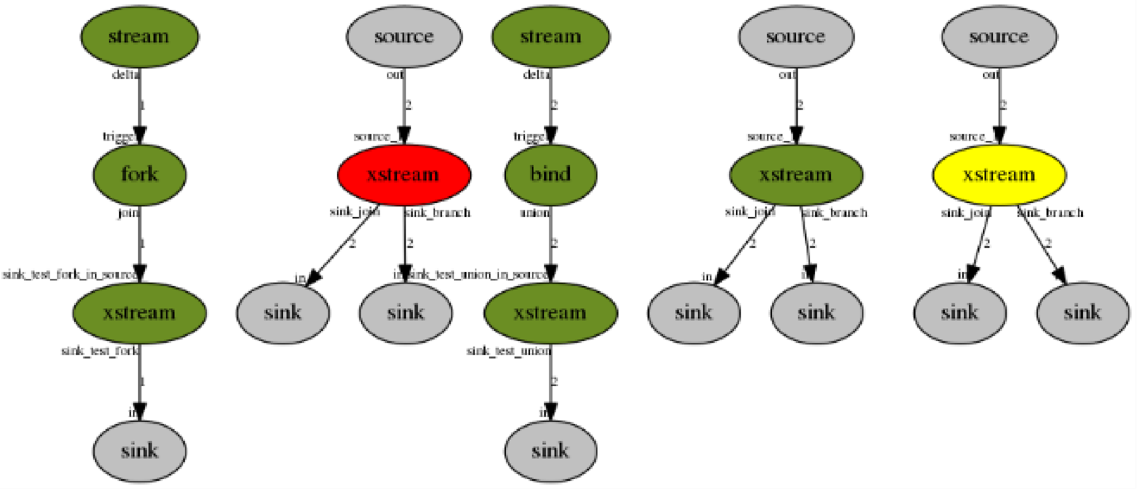

Figure 7. Operator deployment lifecycle states in services

As the execution environment for specific types of operators, the global WWS Deployer considers *'processors'* as the entity for executing the specific operator types (e.g. a Node.js engine). The processors are packaged into Docker images for flexible deployment on cloud machine clusters. The dynamic scheduling, both at the level of processors and at the level of operators leverages state-of-the-art scheduling frameworks like Marathon/Mesos or Kubernetes, which allow for dynamically scaling out operators, processors and any data brokers (RabbitMQ or Kafka) between them. A *Dispatcher*, locally on each WWS site, does the actual launching and removal of operators on processors and the setup of data stream channels between them. Like the Deployer, this function is stateless and will scale up with the amount of operator state changes.

4.1.4 Onboarding streams with WWS Stream Bridge

As is the case with any IoT platform from previous generations, an important platform capacity is to be able to flexibly onboard external streams. As illustrated in Figure 8, WWS has a dedicated interface for this purpose, called *Stream Bridge*.

The Stream Bridge defines WWS ingress streams as *source streams* and WWS egress streams as *sink streams*. Both source and sink streams should be registered via a straightforward REST *Registration API*, to describe the stream type, its data structure, chosen transport protocol, a name and any optional tag annotations. An additional control stream, to send stream control data in reverse direction, can also be registered if needed. Successful registration returns a unique reference for publishing (source), respectively receiving (sink), and the streaming data over the chosen transport protocol to/from WWS.

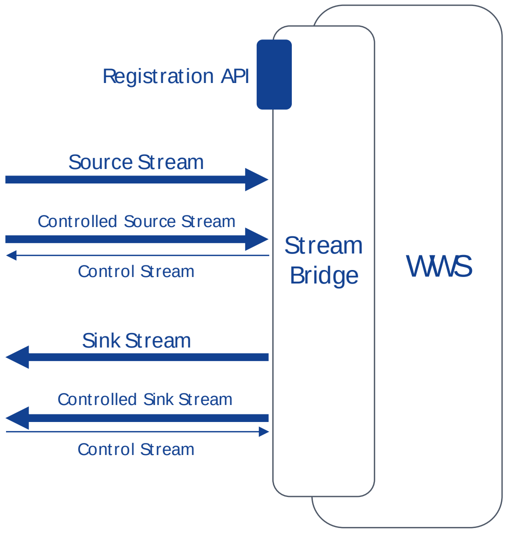

Figure 8. Stream Bridge overview

Stream Bridge provides support for JSON formatted transport over AMQP (RabbitMQ's native broker protocol), MQTT, STOMP, Kafka, and even HTTP/REST and is flexibly extendible beyond that. Stream Bridge leverages *protocol adapters* for these transports, for republishing the streams on the WWS-internal RabbitMQ broker.

Beyond the data transport and stream encoding itself, *application-level protocols* can be supported by designing a protocol-terminating proxy service at the application level - these can be regularly authored as any other XStream script on WWS. By its flexible multi-protocol and lightweight registration concepts, communication with practically any external entity can be modelled. Streams originating from (or delivered to) legacy or application-domain-specific systems can be connected, such as from/to:

- web systems, interactive user interfaces (web-based or other), that can use their assumed lightweight connection paradigm,
- IoT devices, such as wearables, smartphones and small sensor devices, that can preserve their small-footprint communication requirements, and
- even full-fledged industrial automation equipment that can be connected by means of their standardized industry protocols.

Thanks to the possibility to trigger instantiation of an application protocol-terminating device proxy service from Stream Bridge, any undesired complexity implied by the protocol can be kept isolated to (potentially auto-scaled-out) instances of the (typically single-operator) proxy service. Thus, while the services running on WWS can work according to representations implied by the semantic stream concepts in the World Model, with a uniform WWS cloud data transport, external entities do not need to be 'aware' of WWS and can talk to the proxy service as if it were their regular stateful protocol peer.

4.1.5 Onboarding compute resources to WWS

A WWS platform can also onboard new computing (or network) resources into the Resource Topology upon which the Deployer can request the deployment of a processor on that processing resource (machine or virtualized machine cluster), through a Resource Request (RER). This request is equivalent to the registration of a new Docker image containing a WWS processor on the machine. It allows the WWS system to launch the processor as soon as an operator, needed by a requested service, requires its execution. Ultimately, processing resources may also reside on the far edge cloud, i.e. on customer-premises gateways, mobile base stations or even small-footprint devices behind those, such as mobile phones, small computing devices like a Raspberry Pi, a smart electricity meter or an electrical vehicle charging control unit. Even wearables or smart cameras can be considered, whenever they expose a programmable compute capability on which a (potentially small) process can be executed. Once onboarded, all resources invariantly can be programmed by WWS to execute one or more operators processing the locally available data streams**.**

4.2 Communication spaces

The concept of communication spaces rely on an owner-centric approach for communications (Figure 3). This allows capturing that physical CPS device resources usually belong to an owner, who can be an individual person, a group of people, company or an organization. Each owner has its own communication space (CS) in the virtual world; the green lines in the **Error! Reference source not found.** represent such ownership relationships. The physical CPS device resources of a specific owner may register their presence into the CS of the referred owner, the blue lines in the **Error! Reference source not found.**. Such physical CPS device resources can be e.g. servers, computers, vehicles, buildings, smart phones, consumer electronic devices, sensors and actuators, which typically belong to a human/organizational owner.

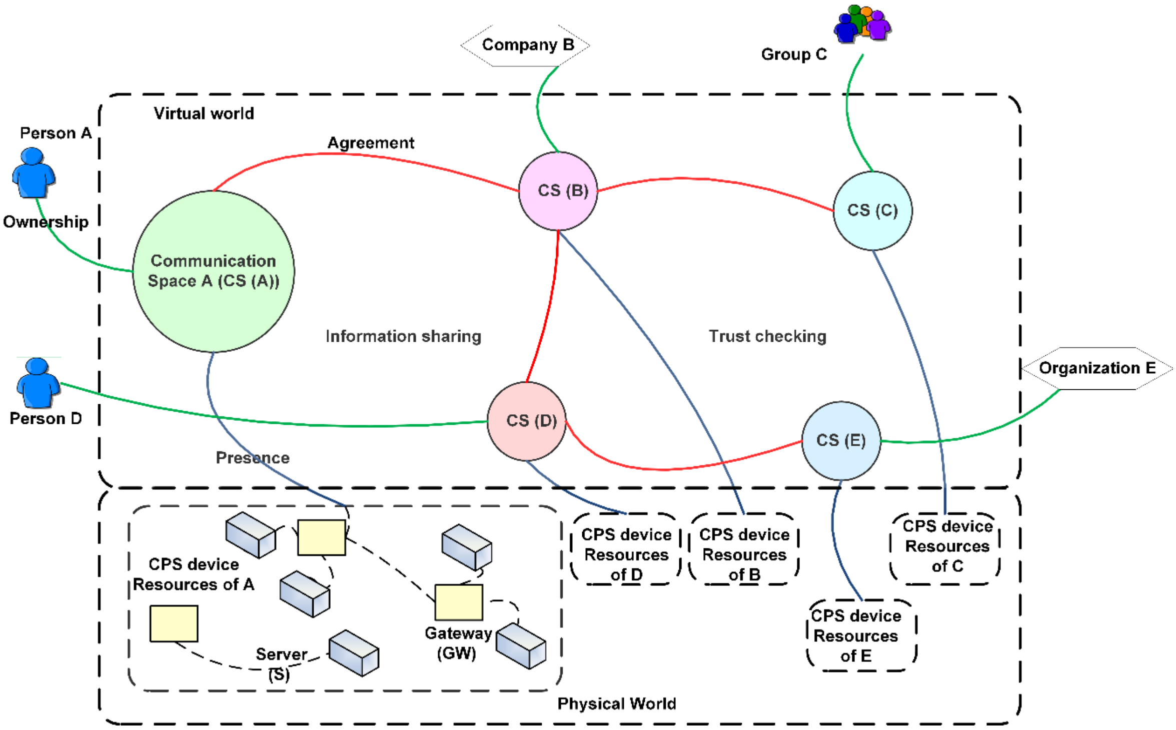

Figure 9. A view to the communication spaces concept based on the VTT CPS hub design [48].

The red lines connecting virtual communication spaces represent agreements between the owners. The referred agreements are negotiated online between the resources of the owners via the communication spaces. Such an agreement is required in order to enable information sharing between resources of different owners. In addition, there is need for trust checking before any information can be delivered, because of the situation, level of trust and detailed security policies of the owner may have been changed.

The key solutions for virtual communication spaces concept are presence management, trustworthy sharing of information and data plane operation [37, 48]. Next, we present these solutions.

4.2.1 Presence management

The need for presence management arises from the inherent characteristics of physical CPS asset devices and services (resources) to be online or offline at any given time. The reasons for being offline may be related e.g. to power sources, unreliable communication links, mobility or due to the user/owner needs. When such a resource needs to be reachable via communication networks, it should register its presence to a known place. According to our approach, such a place is the virtual communication space of the resource owner. When a physical resource accidentally will lose connection, it is obvious that there can be some level gap between the physical resource and its' state in the virtual communication space for some time [67]. The virtual space can operate towards solving this gap by acting on behalf of the physical resource when it is offline, and making more or less regular poll procedure to check the status of the physical resource.

When speaking about resources, an essential problem to be solved is related to identities and addressing. There are several solutions for identities in different levels of the systems. For example, MAC addresses are in use in the radio access level, IP addresses and URLs in Internet/Web contexts, service provider specific addresses in enterprise systems, some universal identities UUIDs have been proposed etc. Here, identities and addressing refer to the ones used at the communication level, more specifically in the communication overlay level and especially related to the concept of virtual communication space.

The problems for the identities and addressing in the communication overlay level arises from the mobility, use of local identities, dynamic presence and topics of shared information. Local identities and addressing, such as IP and physical level addresses, may be temporal. Therefore, more reliable ways for identifying resource owners and the resources themselves are needed. Basically, the generally applied "scheme:[//host[:port]][/path]" could be applied; however, it lacks references to the ownership of resources. Therefore, identities and addressing for virtual communication spaces at the communication overlay level are specified using the following notation:

<owner>@<domain.server>/<physical resource>/<logical resource>

where

- <*owner*> denotes the resource owner.
- The home domain of the owner is <*domain.server*>.
- <*physical resource*> refers to a physical resource, which can typically be e.g. embedded devices, gateways or servers.
- The optional part <*logical resource*> can refer to any specific information content, or to logical resource tree of the information or e.g. semantic Web type of resource.

4.2.2 Sharing of information - Publish/Subscribe

The sharing of information with the communication overlay contains the following phases, Figure 10. The first phase is connecting of the owner into the CPS communication hub, authentication of the owner and registration of the presence of the physical resources of the owner into the home domain of the owner. As the result of this step, the presence of the physical resources is registered as resources of the communication spaces of the owner in the CPS communication hub. The referred resources can be identified by <owner>@<domain>/<resource> notation.

In the second phase, some resource may subscribe to receive information published by some other resources. The publish/subscribe service requires a separate authentication done using the same owner identity to the CPS information sharing service than applied in step 1. The published/subscribed information is identified by notation <owner>@<domain>/<resource>/<topic>. Any resource can publish information according to its own strategy, if the authentication has been done in successful manner. However, when a resource subscribes for information published by other owners/resources, then the trust service checks whether it is allowed or not.

In the third phase, the CPS Trust service checks the status related to the agreements between the owner (A), which is publishing the information, and the other owner (B), which wants to subscribe the information published by A. If there is an agreement between A and B, then the subscribed information is delivered to the subscriber B. However, before the delivery, A' security policies are checked by the CPS trust service using an external policy-based authorization service.

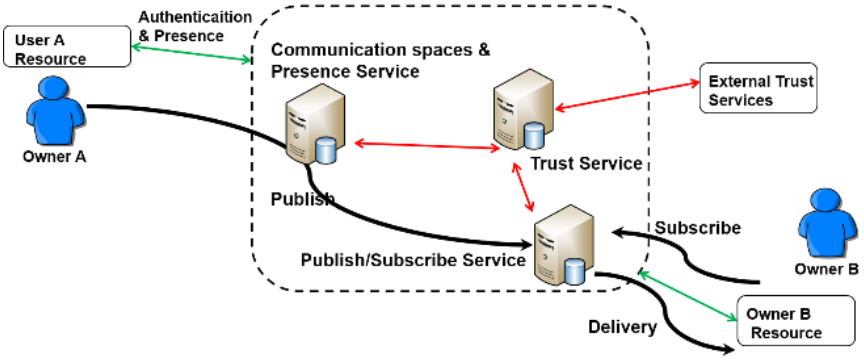

Figure 10. Information sharing of CPS communication hub.

4.2.3 Data plane operation

The sharing of information with the communication overlay using data plane operation capability is clarified in this section, Figure 11. The basic operation is divided to control plane and data plane functions. It is expected that the physical resources execute first the registration of their presence into the virtual communication space of the owner as a control plane function. After that the owners need to create first a mutual agreement ("trust" relationship), otherwise no interaction cannot happen between the physical resources of the referred owners. This is also a control plane function. After such an agreement has been created, then any resource of the owners can interact/exchange messages between each other as clarified earlier.

If the information exchange via the CPS communication spaces is not enough efficient, e.g. in the case of multimedia delivery or enterprise specific means is required to be used, then data plane operation can be a possibility. The data plane operation can be activated by executing so-called "M2MGrids negotiation" where the applied data plane communication channel is negotiated as the control plane action first. This negotiation can happen between any two individual physical resources that have passed all the control plane actions clarified earlier. Based on the "M2MGrids negotiation", the used end-to-end data plane format/protocol can be defined by the resources themselves. After the "M2MGrids negotiation" has happened, then the interaction between referred physical resources can happen as the data plane functions directly without any involvement of the communication spaces related network/control plane processing using the agreed data plane means.

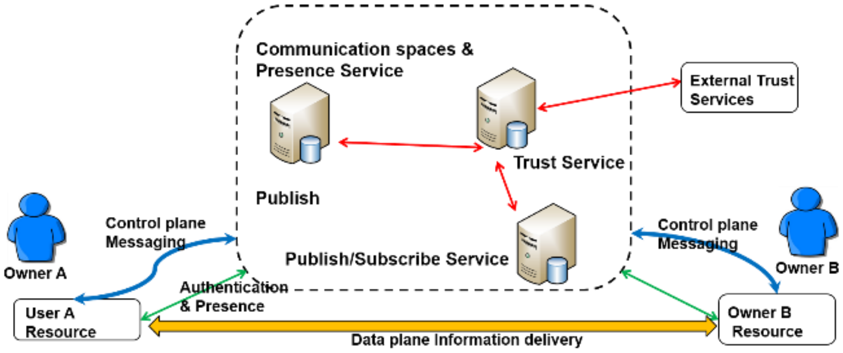

Figure 11. Data plane of CPS communication hub.

4.3 Policy based authorization

The eXtensible Authorization Framework as a Service (SAFAX) [53] is a XACML-based authorization framework that allows data owners to create and enforce policies that regulate the access to their data. SAFAX can be used to define access control policies for data residing in multiple domains, under different authorities. The process of defining access control policies is simplified with SAFAX, as data owners can manage their policies from a single administrative point, rather than having to create and maintain different policies within each data domain. SAFAX components are designed as loosely coupled services. An overview of the SAFAX framework is shown in Figure 12. The *domain controller* (DC) is the entity that host the data domains where resources are shared. These domains belong to *domain owners* (DOs), e.g. energy service providers. The DC is responsible for providing application specific components (the "Policy Enforcement Point" PEP, the "Policy Information Point" PIP, and the "Context Handler" CH), which generate a valid XACML access request to the SAFAX-CORE service for application specific access requests from the *data consumer*, the entity requesting access to the data.

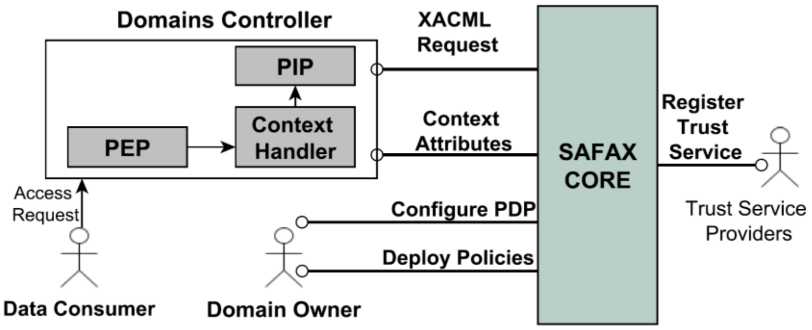

Figure 12. SAFAX Framework

The data owner can deploy access control policies expressed in XACML with SAFAX. These policies can include different trust dimensions such as reputation of the data consumer or the credentials a consumer should have. The evaluation of trust constraints is decoupled from the policy evaluation engine. In particular, trust service providers can register their services with SAFAX. This makes the approach *extensible*, since SAFAX can be extended without any changes to the underlying infrastructure, and *flexible*, since the data owner can choose trust services based on their requirements and needs.

4.3.1 SAFAX-CORE

The main component of SAFAX is SAFAX-CORE (Figure 12). Each component of SAFAX-CORE is designed as a service, thus breaking away from the monolithic component structure seen in existing XACML implementations. Each component are shortly described below, see Figure 13.

Router: Given that SAFAX serves multiple DOs, the router forwards the request to the PDP assigned to the DO whose data are requested.

Service Repository: A service repository contains records of external services that can be plugged to the PDP as User-Defined Functions (UDFs). External service providers can register their services with the SAFAX framework; however, a strict requirement is that the services must meet the interface specifications, which are described in the next section.

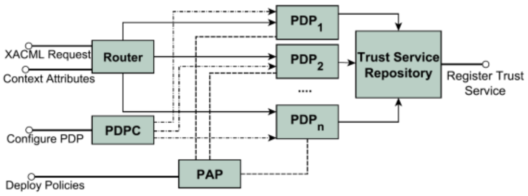

Figure 13. SAFAX-CORE Architecture

Policy Administration Point (PAP): This component facilitates DOs with the management of their policies.

Policy Decision Point (PDP): For each data domain, SAFAX assigns a dedicated PDP to handle requests pertaining to data within. The PDP fetches the relevant policies from the PAP. In SAFAX, we decouple UDFs from the XACML PDP module to allow the framework to be extensible with new functions without disrupting the existing components.

PDP Configuration (PDPC): Since SAFAX is extensible through external trust services; we provide DOs the control over the external services that they would like to use within their PDPs. This can be done through the PDPC. The configuration is used to initialize the PDP with respect to the trust services selected by the corresponding DO.

The overall view of the SAFAX architecture and its usage is shortly clarified in the following.

Authorities use the PAP to deploy their policies and the PDPC to configure their PDP, the PDP service to be used for policy evaluation, etc. SAFAX provides data owners a unique PDP-URL to invoke the assigned PDP. All access requests for resources belonging to the data owner are forwarded to the assigned PDP through the unique PDP-URL.

Authorities should implement the domain-specific components (i.e., PEP, CH and PIP) and register the Context Handler service with the SAFAX framework. This is necessary since during policy evaluation the PDP might need additional attributes from the CH of the data owner (e.g. resource type, classification level). In particular, the CH interface to fetch the attributes should be registered with the service registry of the SAFAX framework. Figure 14 shows the interaction of services within SAFAX.

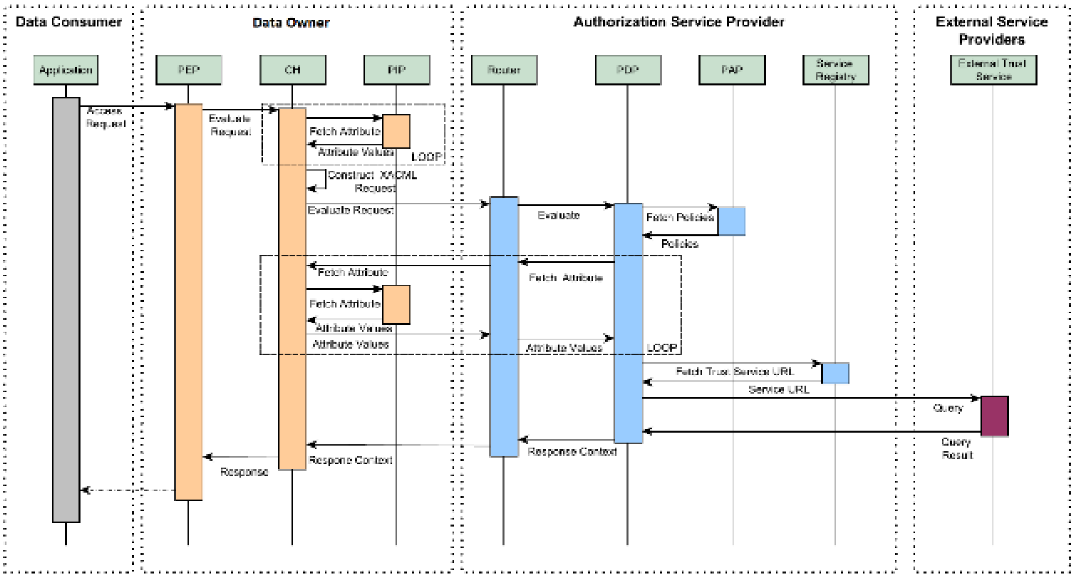

Figure 14. Message flow for policy evaluation using SAFAX

The request sent by the data consumer to access data items belonging to a data owner is intercepted by the PEP of the data owner. The PEP forwards the request to the CH service, which first fetches the default additional information from its PIP (e.g. type of resource), enriching the original request, and then constructs a valid XACML request. The CH forwards the XACML request to the PDP by invoking the SAFAX service. The Router intercepts the access request coming from the CH and forwards it to the proper PDP based on the PDP-URL. During policy evaluation, it is possible that the PDP needs additional attributes, besides the ones provided by default, case in which it contacts back the CH service. If the policy specified by the data owner requires additional (external) trust information, the PDP service contacts the external trust services. After receiving the required information, the PDP computes the decision and informs the CH, which then parses the XACML response and sends the decision to the PEP. The PEP enforces the access decision by allowing or denying access to the data item to the consumer.

We have realized the SAFAX architecture by implementing it as a web service. SAFAX is implemented in Java running on an Apache Tomcat server and using Jersey as the service framework. MySQL is used as a backend persistent data storage. SAFAX exposes interfaces that can be invoked by PEPs implemented by the DCs. The SAFAX GUI is implemented using HTML, CSS and AJAX to consume SAFAX services.

4.3.2 Policy Alignment Service

M2M systems are typically dynamic and open systems wherein parties may not know each other a priori. To deal with this challenge, we have extended SAFAX with Trust Management (TM). In TM, access decisions are based on credentials and policies issued by multiple authorities. However, one of the main problems of existing TM systems is that they implicitly assume a complete agreement among authorities on the vocabulary used for the specification of credentials, which is a too restrictive assumption in dynamic coalitions like M2M systems. A partial solution to these problems is offered by the use of ontologies, which provide a valuable means for enabling interoperability across heterogeneous systems [68]. However, it is often unrealistic to assume that all parties in the coalition reach a complete and precise semantic alignment before becoming operative.

To enable semantic alignment between the domain models of the parties in a coalition, we leverage the work in [69], which present a reputation system based on the notion of similarity [70, 71, and 72]. Similarity represents the semantic resemblance between two concepts. In our solution, similarity is asserted in the form of credentials. Each entity can specify similarity assertions between two concepts independently from the ontology in which those concepts have been defined. Based on these assertions, we use a reputation metric to assess the similarity between concepts during policy evaluation.

We have implemented and deployed the semantic alignment service within SAFAX. This service has been implemented as any other SAFAX service. Moreover, the SAFAX GUI has been extended to manage semantic alignment service configurations. Through the GUI, a user can specify their similarity credentials. These credentials are stored in a persistent database and used by the semantic alignment service to assess the similarity between concepts. During policy evaluation, the similarity constraints encoded in the XACML policy are resolved by invoking the semantic alignment service, which returns a response according to the UDF used to encode the constraints.

4.3.3 Data Governance and Transparency

Existing access control mechanisms typically assume that data objects are under the control of a single entity. However, this is often not the case within the M2MGrids where several users can contribute to the creation and management of data. This opens new challenges in the secure management of data. First, not all stakeholders might have the same level of authority. In particular, the degree of authority each stakeholder has over the data depends on its role with respect to the data. Stakeholders can define their own authorization requirements for the protection of data. These requirements should be combined to define an enforceable policy in such a way that the level of authority of each stakeholder is accounted for. Moreover, users can define conflicting access requirements. While existing access control mechanisms can solve policy conflicts, they usually fail to make users aware about the actual enforcement of their policies.

To address the first challenge, we introduce a data governance model that allows the integration of access requirements specified by different users based on their relationship with the data object. The governance model provides a general framework to reason on the level of authority that stakeholders have over shared data and allows the use of policy combination strategies to resolve policy conflicts. The relation between stakeholders and shared data is characterized using the notion of archetype introduced in [73]. Intuitively, the archetypes for a shared resource capture the roles that stakeholders can have with respect to the resource.

Archetypes are organized in a hierarchical structure to reflect the degree of authority that stakeholders have on shared resources [74]. Intuitively, each level in the hierarchy groups archetypes that have the same degree of authority on the shared data. The requirements of the stakeholders associated to those archetypes are combined with an intra-level aggregator that specifies how the requirements should be evaluated. In an archetype hierarchy, levels are ordered with respect to the degree of authority that the archetypes within a level have. We distinguish three types of priorities between levels: total, positive and negative. A total priority indicates that the access requirements defined by an archetype at a higher level always override the ones of archetypes at a lower level. However, in some cases it is desirable that only the positive authorizations take precedence. This is achieved using the positive priority. Similarly, negative is used when only negative requirements from the higher level take precedence.

Ideally, the authorization system should enforce the access requirements of all stakeholders. However, this is not always possible due to policy conflicts arising from stakeholders' conflicting access requirements. Existing access control mechanisms often provide a mean to resolve policy conflicts automatically. However, although resolving policy conflicts is necessary to make a conclusive decision, decision making often becomes non-transparent to users. The main problem lies in the fact that, in existing access control mechanisms, policy conflict resolution is embedded in the policy evaluation process and, thus, policy conflicts are not identified and/or recorded. This makes users unaware of whether their policies have actually been enforced.

To enable transparency in access control, we exploit the notion of policy mismatch introduced in [73]. Intuitively, a policy mismatch occurs when the decision enforced by the authorization system differs from the one obtained evaluating stakeholders' policies individually. Based on this notion, we have designed a Transparency Service that aims to detect mismatches between the decision enforced by the authorization system and the access requirements as specified by stakeholders. The Transparency Service reports the identified policy mismatch to the stakeholders whose decision was not enforced. By doing so, the service aims to make the stakeholders engaged in a collaborative resource management aware of possible conflicts with the authorization requirements of other stakeholders and how these conflicts were resolved by the authorization system.

5. Evaluation

5.1 Evaluation scenarios

The evaluation of the horizontal solutions for cyber-physical systems has been carried out by making an experimental system, consisting of energy flexibility and traffic accident use cases represented in Figure 15 and Figure 16 (see also section 3.1). The energy flexibility case focused on the automatic negotiation of the usage of energy sensitive resources in the coming 24 hours for lowering the peak loads, the cost of energy, its' distribution, and enabling trading of flexibility. The validation system consisted a set of simulated energy sensitive resources (white goods, boilers, heating resources), buildings and electric vehicles, which were controlled by multiple energy aggregators (4), acting on local markets, operating in the electric grid of a distribution service operator (DSO). The traffic accident case focused on information based interaction between multiple physical real world devices and services of companies (e.g. Polar smart watch, Bittium EEG sensor, Tracker dog collars, Valopaa's illuminator, LiveU announcement service, IMEC CO_2 sensor), causing alarm, considering security policies of the owners when sharing the information for authorities and visualizing situation in the emergency area.

The evaluation of the capabilities of provided M2M service platform (Nokia WWS as a reference implementation) has been carried out mainly in the energy flexibility case. The provided communication spaces (VTT CPS hub as a reference implementation) have been evaluated in energy flexibility and traffic accident cases so that the validations support each other. The policy-based authorization (TU/e eXtensible Authorization Framework as a Service (SAFAX) as a reference implementation) capability is mainly evaluated in the traffic accident case. The evaluations considered real-time operation with real physical CPS resources as well as simulated resources in both energy flexibility and traffic accident cases.

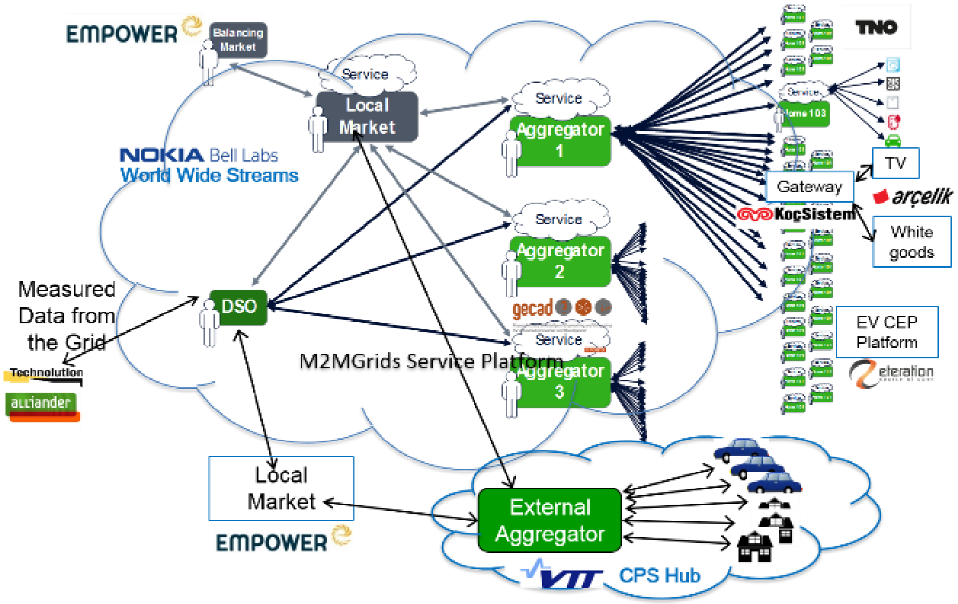

Figure 15. Structure of the energy flexibility case.

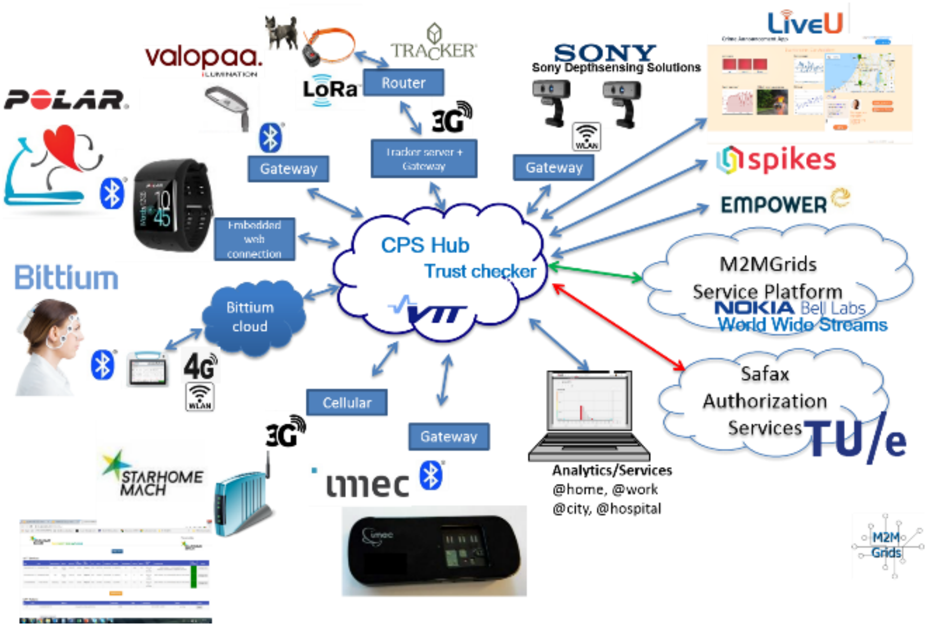

Figure 16. Structure of the traffic accident case.

5.2 Evaluations of the horizontal solutions

5.2.1. Evaluation of the WWS

Nokia World Wide Streams, as the reference implementation of an M2MGrids M2M service platform, was validated in the extensive, multi-stakeholder energy flexibility ecosystem case depicted in Figure 15. At the level of the business validation for the energy domain, the viability of new real-time energy flexibility trading scenarios was demonstrated, with (new third party as well as traditional) aggregators leveraging energy flexibility of household devices and cars in a new flexibility market ecosystem. A triple-win can be seen in the demonstrated case. Next to rebates for participating households and a sound business return for creative aggregator parties, the incumbent producers and distributers could save on energy storage investments dramatically. In the example setup, using realistic solar and wind power production simulation on one side, and fine-grained household device consumption and flexibility simulation on the other side, we showed that over 50% of the Gigawatt-Hours of required battery storage could be divested due to the real-time flexibility coordination.

While this demonstrated a high impact on how the energy grid will be able to operate in the near future, it also proved the WWS platform capable of supporting such cases. Realizing the energy flexibility case on the WWS platform indeed gave valuable insights on the stream handling and processing, and functional and non-functional requirements for the energy domain. This has led to several WWS improvements with respect to ease of programmability and debugging, as well as with respect to performance and scalability of the platform, making it now well fit for these and similar cases.

Below, we shortly highlight the main aspects and WWS features that we constructed and validated in the course of realizing this use case.

- **Domain alignment through the World Model**: The semantic alignment of the domain information, underpinned by multiple standards and their application, as was done for the energy domain use case at hand, showed to be essential for ensuring interoperability of energy services owned by different ecosystem actors. Next to its other technical goals, the World Model has shown to provide a systemic way to obtain this, and is recommended as a way forward for industry cooperation, as shown for the energy domain.
- **WWS onboarding of any device**: The Nokia WWS platform has an open support for onboarding any device type, by modelling any external entity as a set of streams. WWS can proxy any application-level protocol in a proxy service of choice for the device type. Stream Bridge, on which streams can be registered via a REST API, supports any practical data transport protocol for interfacing incoming and outgoing streaming data. This was validated by onboarding many devices and external systems via Stream Bridge, as Distributed Energy Resources (DER) as part of the energy flexibility use case. An application-level device proxy service running on WWS was generically handling seamless connection of EFI-compliant devices for time-shift flexibility control (Energy Flexibility Interface, EFI, is a Cenelec pre-standard). Next to various simulated household device types (dishwashers, washing machines, heaters, electrical vehicles, lighting), also physical devices where onboarded via EFI-compliant gateways, the gateway being either as customer-premises equipment (CPE) or directly integrated in the white goods.
- **Real-time energy device control on WWS**: Once onboarded, energy services running on the platform were shown to be able to control device energy consumption (or production) in real-time. As the energy devices are represented as a set of data streams within WWS, and the EFI protocol and XML formatting was terminated by a device proxy service, the energy control business logic of services could make abstraction of this.
- **WWS smart service pattern programming**: Many services were built on WWS using the XStream language for consecutive iterations of the energy flexibility case. Each iteration incorporated new patterns of processing logic and new types of streaming data. For example, beyond basic energy measurement monitoring, various automatic demand-response procedures, energy flexibility notions and market mechanisms were tackled, thus involving also transactional patterns between energy actors' services, also needed to comply to their inherent timing requirements (even at 3000 times faster than real-time operation).
- **WWS multi-actor service interworking**: Also inherent to the chosen smart grid ecosystem use case, multiple business actors' services needed to interact continuously. For example, household services are controlling devices according to a demand-response dialog with an aggregator, and aggregators are competing on an energy flexibility market where distributors and producers ask for energy flexibility. WWS has been shown to support such concurrently running set of interacting services, leveraging the loose data stream-based coupling among them as provided by the system. Beyond this evaluation, one can also imagine different security and privacy requirements to hold for different parties. For example, households could require that detailed

measurement data does not leave the home equipment, or aggregators could require their services running in isolation from competitors. The WWS platform was designed to be able to fulfil such placement constraints and provides multi-tenancy support, thus obtaining that, e.g., household data is (pre-)processed locally, or aggregator services do not share cloud resources.

- **WWS performance at scale**: Designing and running services on a cloud platform is not enough for large-scale deployments, such as would occur in real-world deployments of an energy grid. Indeed, thousands of devices can be expected to be involved, that may be geographically spread across a wide area, e.g. country- or pan-country-wide. In addition, many small service players, like households and other micro-actors in an energy grid, are expected to be involved. Furthermore, the processing (and transport) capacity needs may vary dynamically with the situation occurring in the large-scale physical setting that is controlled. Finally, once services are running in cyberspace, some use cases will benefit from having the services exchange data at higher data rates than what was possible in traditional systems. For example, many more - and more frequently trading - participants in an evolved smart energy grid scenario may leverage a real-time energy flexibility market. Through simulation of large amounts of energy devices from geographically widespread Internet locations, we could load WWS with the high-load processing corresponding to such near-future scenarios. To make the evaluation still stronger, we distributed a reference clock signal among the processing services and data simulations, such that a uniform speed-up of all data streams could be done, up to 3000 times faster than real-time operation. This showed the feasibility of scaling to even higher data rates. This also showed that ultra-short market cycles can be obtained, far below the 15-minute cycle considered in today's most progressive scenarios.
- **WWS onboarding of external algorithms**: A real-world service platform cannot assume a Greenfield starting point. This is also the case in the energy smart grid world, where pre-existing legacy solutions, and business and algorithmic logic already encapsulated in software implementations, need to be incorporated. To avoid ineffective re-implementation of existing software on WWS, the service platform can onboard existing code and services in varying degrees, all of which were evaluated in the M2MGrids energy case. Ideally, existing code can be encapsulated into a stream processing operator. This is the approach taken with many existing open source data processing libraries in WWS (Python libraries, Gstreamer, Tensorflow, Torch, and more). The approach has the advantage that, once available as an operator, WWS can manage, reuse and distribute its instances at scale. If, however, a solution owner does not want to expose source code, which is e.g. often the case for competitive algorithmic solutions, then it can be operated entirely outside of the WWS system, e.g. as a web service. For web services, proxy operators have been built on WWS, representing the external service in regular XStream dataflows by proxy.
- **Dynamically adding services to WWS**: As different stakeholders, like in the considered energy ecosystem case, can launch services, there is no fixed amount of services, no fixed starting order, nor a coordinated lifecycle among them. WWS can handle this situation smoothly, thanks to the loose intercoupling of services and operators, by dynamically query-able stream registration and broker topic-based streaming. In the energy use case, e.g., energy flexibility aggregator services could be added or removed from the market on the fly, and households could announce themselves as new aggregator customers, or step out, under the hood leveraging this WWS capability.
- **WWS service to legacy service interworking**: Combining the aspects of the two previous sections provides for an even stronger proposition. If a legacy service is running outside of WWS, and other services with a similar ecosystem role are executed natively on WWS, like different aggregators acting on a single energy flexibility market in the use case, then these services need to be 'compatible', in the sense that the market service should not be able to distinguish the fact that a service is running in or outside the platform. The WWS architecture was shown to be able to indeed hide this aspect. In a demonstration, internally running services

(representing aggregators, EFI devices, and more) could be swapped or work concurrently with an externally running instance in the same role, that had registered its streams via Stream Bridge.

5.2.2. Evaluation of the Communication spaces

The evaluation of the concepts of communication spaces realized with VTT CPS hub as a reference implementation has been done in several steps with both energy flexibility and traffic accident cases. The main issues evaluated in the first step related to energy flexibility case are shortly clarified in the following. Registration of the presence of the distributed resource components into the CPS hub: simulated electric vehicles simulated charging spots, charging manager, simulation model of the local distribution grid, user interface tool, visualizer tool, and energy market platform of Empower. Establishing simple trust relationship between Empower system and all the other components controlled by VTT. Enabling the resources to communicate with each other to collaborate to reach negotiated energy consumption/production flexibility in day ahead/intraday energy market case.

The main issues evaluated in the parallel first step in the traffic accident case are clarified in the following. Increasing the number of owners and physical resources such as smart watches, hunting dog collars & specific server, gateways, and street lamps and specific servers. Establishing more trust relationships between owners. Application of publish/subscribe information sharing between physical resources owned by different stakeholders.

After these parallel steps, the combined experimental system including CPS hub and WWS was evaluated. During this step, a new duplicated distributed energy resource set was established under control of small-scale sub-aggregator (VTT) interacting with the local energy market of the WWS. This small-scale aggregator of VTT (FlexEntities) aggregated a set of simulated distributed energy resources: electric vehicles and charging stations, and some buildings. This set of DER resources was controlled via the WWS Stream Bridge by higher-level aggregator service/local market, whereas the previous set of DER resources was controlled via Empower system aggregator.

The main evaluation points of the combined experimental system are clarified in the following. Establishing the required streams towards WWS for reaching the topic names for communication with the local market entity from the CPS hub (VTT stream proxy). Creation of the trust relationship with WWS by applying a VTT stream proxy to simplify the test case. Evaluation of the interoperability of the CPS hub with WWS. In addition, evaluation of the scalability of the CPS hub by execution of multiple instances of the simulation system owned by different stakeholders.

5.2.3 Evaluation of the Policy based authorization

The policy-based authorization, solution (SAFAX), was evaluated together with the traffic accident case. The policy-based authorization service was executed in close collaboration with the CPS hub services, because it was related strongly to the process for information delivery.

Functionality of the build system was evaluated incrementally during the project. The main evaluation points were following: Evaluation of the scalability of the CPS hub to support presence management of larger set of resources owned by different stakeholders. Evaluation of the trust checker of CPS hub for taking care of the dynamic grid of agreements dynamically negotiated between the parties/resources. Evaluation of the trust checker negotiation with SAFAX in a security policies case dealing with alarm type of event. Evaluation of the trusted publish/subscribe type of information sharing between resources so that the delivery of subscribed content happens only when authorization conditions are met.

The incremental development of the CPS hub with policy-based authorization and multiple evaluation steps proved their purpose. All the evaluations were passed successfully. Especially the capabilities to manage presence, dynamic grid of agreements, trust and publish/subscribe type of information sharing between heterogeneous resources/services of multiple stakeholders, and especially the policy-based access control are seen to be essential for the M2M/IoT/CPS architectures.

5.2 Discussion on evaluation results against the requirements

5.2.1 Energy flexibility case

The energy flexibility case focused on the real needs of the energy sector stakeholders to cut down the peak loads of a day to lower the cost of energy and its distribution. As the results of this work, we achieved demonstration of automatic energy flexibility trading required interaction between DSOs, balancing and local energy market stakeholders, aggregators, and multiple sets of energy sensitive DERs (boilers, heating resources, household appliances such as TVs, white goods, electric vehicles/charging stations, buildings). This successful flexibility trading demonstration shows that acceptable level interoperability between multiple energy sensitive domains (electricity grids, building automation, consumer households, electric vehicle charging), physical devices and systems (white goods, boilers, heaters, electric vehicles) and related stakeholders (DSO, energy market, energy aggregators) has been achieved (R1). The DERs were controlled by multiple energy aggregators operating in the local energy market (R2, R3, R4). Monitoring the load balance in the electric grid was successfully demonstrated (R5), forecast of the balance for the next day was done (R6), the balance was compared with the forecast (R7), and flexibility (for each 15-minute slots on a 24-hours horizon) was asked on the local market (R8). The energy aggregators offered flexibility (R9a) considering energy market prizes (R9b) and owners restrictions (R10). The awarded aggregators were able to provide the required flexibility instructing the individual devices (R11). As a result, the balance of the local electrical grid was maintained better, with flattened, lowered discrepancies between energy production and consumption (R12). It is estimated based on the successful energy flexibility demonstration, that DSOs can save in energy storage investments (R13a), new aggregator role can be possible and feasible (R13b), and households and vehicle owners get energy bill could get rebates for their flexibility contribution (R13c). In addition, the automated negotiation of energy flexibility is estimated to answer the real needs of the energy sector, because it is challenged by more *renewable but intermittent* energy sources (R14a), and the *ever more distributed* nature of both productions and consumption (R14b).

The essential novelty of the successful demonstration of energy flexibility was related to capabilities to include also DERs outside traditional energy sector, from consumer, buildings and mobility sectors, into the flexibility trading process. In addition, also production of energy via discharging of electric vehicle batteries was included into the trading process. All this was made possible by application WWS and CPS hub solutions contributed in this article. However, the evaluations were limited in the sense that most of the applied DERs were simulated for cost and practical reasons.

5.2.2 Traffic accident case

The traffic accident case focused on combining information exposed from multiple CPS of multiple SPs to enable real-time situation and location awareness for authorities. As the results of this work, we achieved demonstration of trustworthy information sharing between real world devices and services of companies (e.g. Polar smart watch, Bittium EEG sensor, Tracker dog collars, Valopaa's illuminator, LiveU announcement service, IMEC CO_2 sensor). This was made possible via application of CPS hub with SAFAX services, which together enabled trustworthy information sharing considering privacy policies of the owners in a limited scenario. In the traffic accident demonstration, the physical asset devices from multiple sectors, hosted by multiple SPs, and SP service systems/clouds were able to share information and interact in *interoperable* way (R15, R17). The security policies and agreements between device owners were considered in specific limited interaction case (R16). The information services exposed from the referred devices can operate according to the requirements of their owners using the referred security policies (R18). It can also estimated that the business values of related service providers are not compromised (R19). In the demonstration case, the smart watch and CO_2 sensor, dog tracker and street illuminator were able to interact with each other considering the policies of their owners (R20). When an alarm event was caused by (simulated) authorized stakeholder, it was able to trigger situation status change in the security policies of the owners so that the system behaviour was changed accordingly (R21). When an authorized alarm event occurred in a specific geographical area, the delivery of the information exposed from the devices in the area were changed so that the authorities were able to receive the information according to the policies of the device owners (R22). As the result, a situation aware view what is happening in the emergency area was visualized for authorities on a screen (R23).

5.2.3 Horizontal solutions

The successful demonstration of energy flexibility and traffic accident cases indicated that the provided horizontal solutions enabled cyber-physical system, in which the physical resources, communication system, information services and authorization are operating in multisector and multivendor environment in interoperable and horizontal way. The understanding of information meaning was enabled by the energy world model of the WWS for interoperability between different ecosystem actors and resources of the system in the energy flexibility case (R24). The importance of the common world model was highlighted also in the interations between multiple energy sensitive sectors, stakeholders and resources for enabling smart energy consumption/production scheduling and optimization algorithms (R25). However, the evaluation was limited in the sense that the only energy information model inherited from SEAS project and enhanced for the use in energy flexibility case. Therefore, the issues related to the required world (information) models of different verticals is seen to be critical future research item. Onboarding heterogeneous DERs operating with EFI/XML compliant devices for time-shift flexibility control via gateways, and application of stream processing method using the XStream language for consecutive iterations helped WWS operation. Simulations of DERs and speed-up of the time in simulations enabled evaluation of scalability of the WWS. Evaluation showed possibility to apply external algorithms execution with the WWS via proxy. WWS enables dynamic adding of new services, e.g. new aggregators and interoperation with real world energy sector services via stream bridge. For example, small-scale aggregator of VTT aggregated a set of simulated distributed energy resources under control of local market in WWS, and another set of DERs via Empower EMSP aggregator capable for operation with real world energy markets. This showed the interoperability of WWS with CPS hub in a practical case. The constructed information service system was able to handle resources dynamically (R26), configuration features (R27), distributed service execution (R28) was supported by WWS. The scalability of the solutions was showed by the demonstrations in both energy flexibility and traffic accident cases (R29).

The heterogeneous physical device resources in the traffic accident case applied several heterogeneous accesses (short range, long range, wired) and attachment methods (direct, indirect) in communications (R30, R31). The communication systems was able to scale to be applied in both energy flexibility and traffic accident cases, which have multiple business logics (energy flexibility, EV charging/discharing, sport -, wellness -, health -, and environment monitoring, street lightning, tracking hunting dogs, emergency situation monitoring) operating in interaction with each other (R32, R33). The CPS hub combined trust checking and information sharing so that delivery of information is possible only when the owners of the resources so allowed. This was demonstrated in the traffic accident case so that the security policies of the owners (SAFAX) were considered in the delivery of the privacy sensitive information exposed from the physical asset devices located in the emergency area. The demonstration considered security, privacy and legal restriction only in limited way, and scalability of the solution to multiple use cases and businesses in still not known (R33).

The evaluations of horizontal solutions were limited to specific energy flexibility and traffic accident scenarios, and all the horizontal solutions were not evaluated simultaneously in a single scenario, but instead in separate demonstration steps of the M2MGrids project story. In addition, a number of future research items were found out during the evaluation process, such as e.g. information models of different verticals, trustworthy of heterogeneous CPS assets and especially secure and trust required processes of cyber-physical systems.

6. Concluding remarks

As the results of this work, we achieved demonstration of automatic energy flexibility trading required interaction between energy sector stakeholders and multiple sets of energy sensitive DERs. The essential novelty is related to including also DERs outside traditional energy sector, from consumer, buildings and mobility sectors, into the energy flexibility trading process. In the traffic accident case, we achieved demonstration of trustworthy information sharing between real world devices and services of companies, and combining information exposed from multiple CPS of multiple SPs to enable real-time situation and location awareness services for authorities. These results were made possible by the horizontal M2M service platform (Nokia WWS), communication spaces (VTT CPS hub) and policy-based authorization (TU/e SAFAX) solutions, and their deployment combinations in the energy flexibility and traffic accident cases.

The evaluation results indicate that the streams based M2M service platform can make development of new services smoother. The communication spaces solution can enable controllable information exchange between physical CPS resources owned by different stakeholders. The policy-based authorization can enable consideration of the owners' policies in the referred information exchange process. The demonstrations indicate that the provided horizontal solutions can enable trustworthy collaborations of multisector and -vendor systems in the energy flexibility and traffic accident cases. However, they were not evaluated in a single integrated scenario, but instead as the slices of the M2MGrids project story demonstration. In addition, a number of future research items were identified related e.g. to information models of different verticals, trustworthy of heterogeneous CPS assets and especially secure and trust required processes of cyber-physical systems.

Author Contributions: Marc Roelands initiated the main ideas for the conceptual architectural adoption of an M2M service platform and a World Model in M2MGrids, positioning it against related work and articulating the proof points in the project's energy case evaluation. He authored the corresponding parts of Sections 2 and 3, many parts in Sections 4.1, and Section 5.2.1. Lode Hoste is the main designer of all the energy services on top of WWS and of various aspects of both the WWS specification and dataflow compilers. He co-authored Section 4.1.2, and part of Section 4.1.1. Wolfgang Van Raemdonck is the main architect of the WWS deployer/allocator subsystem, the various onboarding mechanisms below it, and a configurable energy case production & household devices simulation. He co-authored Sections 4.1.3, 4.1.4 and 4.1.5. Gabriel Santos has contributed into the information models and ontologies of energy domain and co-authored a part of Section 4.1.1. Goreti Marreiros has been focused into the requirement analysis process in the project, and reviewed especially the respective sections 3.1 and 5.2.3. Marko Tilvis has contributed into evaluation of the traffic accident case in Sections 5.1 and 5.2.2-3. Laura Genga and Nicola Zannone has contributed into the policy-based authorization and authored respective parts of sections 2, 4.3, and co-authored Section 5.2.3. Juhani Latvakoski has contributed especially into the description of the horizontal solutions, communication spaces and evaluations. In addition, he has combined and edited all contributes into a single harmonized article. He has authored especially Sections 2., 3., 4.2, and 5.1, 5.2.2-3, 5.3, 5.2.3 and contributed as the main editor of Sections 1 and 6.

Funding: This article is based on the research actions executed within ITEA 2 call 8 project 13011 M2MGrids, Smart M2MGrids - M2M Internet for dynamic M2M information business ecosystem. The M2MGrids project was made possible by ITEA framework and was funded by public authorities in Belgium, Finland, Netherlands, Portugal, Turkey, and Israel.

Acknowledgements are offerred for all ITEA M2MGrids project partners for their active contributions and discussions during the project. Especially, acknowledgements are offered to Juha Kemppainen and Jari Jauhiainen from Bittium, Erdem Ergen from KocSistem, Sibel Malkos from Arcelik, Pertti Puolakanaho and Niclas Granqvist from Polar, and Jouni Heikkinen from VTT.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alliance for Internet of Things Innovation. Online: https://ec.europa.eu/digital-single-market/en/alliance-internet-things-innovation-aioti accessed 21st Feb 2019.
2. IoT 2020: Smart and secure IoT platform. 181p. Available online (accessed 10th Jan 2018) http://www.iec.ch/whitepaper/pdf/iecWP-loT2020-LR.pdf
3. ITU-T Study Group 13, Next Generation Networks – Frameworks and Functional Models: Overview of the Internet of Things, International Telecommunication Union, Geneva, 2012.
4. ZVEI – German Electrical and Electronic Manufacturers' Association, Industrie 4.0: The Reference Architectural Model Industrie 4.0 (RAMI 4.0), Frankfurt am Main, 2015.
5. Industrial Internet Consortium, Industrial Internet Reference Architecture (Version 1.7), Object Management Group, Needham, MA, US, 2015.
6. Internet of Things – Architecture Consortium, The IoT Architectural Reference Model (ARM) - D1.3, European Commission, Luxembourg, 2012
7. Fiware open specifications. Online documents (Accessed 13th Jan 2018). https://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/Summary_of_FIWARE_Open_Specifications
8. SmartM2M Virtualized IoT architectures with Cloud Back-ends. ETSI TR 103.527 V0.2.1 (2018-01)
9. https://www.trusted-iot.org/
10. Amazon Web Services. Online documents (Accessed 13th Jan 2018). https://aws.amazon.com/documentation/iot/
11. Guth J., Breitenbucher U., Falkenthal M., Leumann F., Reinfurt L. Comparison of IoT Platform Architectures: A field study based on a reference architecture. CIoT'16.
12. IPSO Alliance, later OMA SpecWorks. Available online: https://www.omaspecworks.org/ (accessed on 21 Feb 2019)
13. The Internet Engineering Task Force (IETF). Available online: http://www.ietf.org/ (accessed on 21 Feb 2019)
14. ETSI M2M/Smart M2M. Available online: http://www.etsi.org/ (accessed on 21 Feb 2019).
15. One M2M forum. Available online: http://www.onem2m.org/ (accessed on 21 Feb 2019)
16. Latvakoski, J.; Alaya, M.B.; Ganem, H.; Jubeh, B.; Iivari, A.; Leguay, J.; Bosch, J.M.; Granqvist, N. Towards Horizontal Architecture for Autonomic M2M Service Networks. *Future Internet* **2014**, *6*, 261-301.
17. Y. Ahmad, B. Berg, U. Cetintemel, M. Humphrey, J. Hwang, A. Jhingran, A. Maskey, O. Papaemmanouil, A. Rasin, N. Tatbul, W. Xing, Y. Xing, and S. Zdonik. Distributed operation in the Borealis stream processing engine. In SIGMOD `05, pages 882-884, New York, NY, USA, 2005. ACM. ISBN 1-59593-060-4.
18. A. Arasu, B. Babcock, S. Babu, M. Datar, K. Ito, I. Nishizawa, J. Rosenstein, and J. Widom. STREAM: The Stanford stream data manager. In SIGMOD `03, pages 665–665, New York, NY, USA, 2003. ACM. ISBN 1-58113-634-X.
19. T. Akidau, A. Balikov, K. Bekiroglu, S. Chernyak, J. Haberman, R. Lax, S. McVeety, D. Mills, P. Nordstrom, and S. Whittle. Millwheel: Fault-tolerant stream processing at internet scale. Proc. VLDB Endow., 6(11):1033-1044, Aug. 2013. ISSN 2150-8097.
20. Apache Storm. http://storm.apache.org/.
21. S. Kulkarni, N. Bhagat, M. Fu, V. Kedigehalli, C. Kellogg, S. Mittal, J. M. Patel, K. Ramasamy, and S. Taneja. Twitter Heron: Stream processing at scale. In SIGMOD `15, pages 239-250, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-2758-9.
22. M. Zaharia, T. Das, H. Li, T. Hunter, S. Shenker, and I. Stoica. Discretized streams: Fault-tolerant streaming computation at scale. In SOSP `13, pages 423-438. ACM, 2013. ISBN 978-1-4503-2388-8.
23. Apache Samza. http://samza.apache.org/.
24. Apache Quarks. http://quarks.incubator.apache.org/
25. L. Amini, H. Andrade, R. Bhagwan, F. Eskesen, R. King, P. Selo, Y. Park, and C. Venkatramani. Spc: A distributed, scalable platform for data mining. In Proceedings of the 4th International Workshop on Data Mining Standards, Services and Platforms, DMSSP `06, pages 27-37, New York, NY, USA, 2006. ACM. ISBN 1-59593-443-X. URL http://doi.acm.org/10.1145/1289612.1289615.
26. B. Gedik, H. Andrade, K.-L. Wu, P. S. Yu, and M. Doo. Spade: The System S declarative stream processing engine. In Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data,

SIGMOD `08, pages 1123-1134, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-102-6. URL http://doi.acm.org/10.1145/1376616.1376729.

27. Apache Hadoop. http://hadoop.apache.org/.
28. M. Isard, M. Budiu, Y. Yu, A. Birrell, and D. Fetterly. Dryad: Distributed data-parallel programs from sequential building blocks. In EuroSys `07, pages 59-72. ACM, 2007. ISBN 978-1-59593-636-3.
29. Y. Yu, M. Isard, D. Fetterly, M. Budiu, U. Erlingsson, P. K. Gunda, and J. Currey. DryadLINQ: A system for general purpose distributed data-parallel computing using a high-level language. In OSDI'08, pages 1-14, 2008.
30. A. Thusoo, J. S. Sarma, N. Jain, Z. Shao, P. Chakka, N. Zhang, S. Anthony, H. Liu, and R. Murthy. Hive - a petabyte scale data warehouse using hadoop. In ICDE, pages 996-1005. IEEE, 2010. ISBN 978-1-4244-5444-0.
31. R. Pike, S. Dorward, R. Griesemer, and S. Quinlan. Interpreting the data: Parallel analysis with sawzall. Sci. Program., 13 (4):277-298, Oct. 2005. ISSN 1058-9244.
32. J. Dean and S. Ghemawat. Mapreduce: Simplified data processing on large clusters. In OSDI'04, pages 10-10, Berkeley, CA, USA, 2004. USENIX Association.
33. B. Theeten and N. Janssens. Chive: Bandwidth optimized continuous querying in distributed clouds. IEEE Trans. Cloud Computing, 3(2):219-232, 2015.
34. Wolfgang Van Raemdonck, Tom Van Cutsem, Kyumars Sheykh Esmaili, Mauricio Cortes, Philippe Dobbelaere, Lode Hoste, Eline Philips, Marc Roelands, Lieven Trappeniers. Building Connected Car Applications on Top of the World-Wide Streams Platform: Demo, Proc. of the 11th ACM International Conference on Distributed and Event-based Systems (DEBS'17), Barcelona, June 2017.
35. World Wide Streams website, http://worldwidestreams.io, consulted on January 2018
36. Latvakoski, J.; Iivari, A.; Vitic, P.; Jubeh, B.; Alaya, M.B.; Monteil, T.; Lopez, Y.; Talavera, G.; Gonzalez, J.; Granqvist, N.; Kellil, M.; Ganem, H.; Väisänen, T. A Survey on M2M Service Networks. *Computers* 2014, *3*, 130-173
37. Latvakoski, J. Small world for dynamic wireless cyber-physical systems. Academic Dissertation publication VTT Science 142. December 2016. 102p. + app. 163p. ISBN 978-951-38-8477-2. ISBN 978-951-38-8476-5.
38. Roussaki, I., Chantzara, M., Xynogalas, S., Anagnostou, M. The virtual home environment roaming perspective. Proceedings of the IEEE International Conference on Communications, 2003. ICC '03. Pp. 774-778.
39. Familiar, M.S., Martínez, J.F., Corredor, I., García-Rubio, C. Building service-oriented Smart Infrastructures over Wireless Ad Hoc Sensor Networks: A middleware perspective. Computer Networks, 2012, Vol. 56, No. 4, pp. 1303-1328.
40. Eugster, P.T.; Felber, P.A.; Guerraoui, R.; Kermarrec, A.-M. The many faces of publish/subscribe. *ACM Comput. Surv.* **2003**, *35*, 114–131, doi:10.1145/857076.857078.
41. Saint-Andre, P. (Ed.) Extensible Messaging and Presence Protocol (XMPP) Core. IETF RFC3920, October 2004. Available online: https://www.ietf.org/rfc/rfc3920.txt (accessed on 2 January 2018).
42. Saint-Andre, P. (Ed.) Extensible Messaging and Presence Protocol (XMPP): Instant Messaging and Presence. IETF RFC3921, October 2004. Available online: https://www.ietf.org/rfc/rfc3921.txt (accessed on 2 January 2018).
43. Millard, P.; Saint-Andre, P.; Meijer, R. *XEP-0060: Publish-Subscribe*; XMPP Standards Foundation. Draft Standard, version 1.16, 11 September 2019. Available online: https://xmpp.org/extensions/xep-0060.html (accesses on 24th Aug 2019)
44. *ISO/IEC 20922: 2016 Information technology Message Queuing Telemetry Transport (MQTT) v3.1.1; Springer: Cham, Germany, 2016.*
45. *OASIS. MQTT 3.1.1 Specification. 10 December 2015.* Available online: http://mqtt.org, and http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.html (accessed on 24 August 2019).
46. Rosenberg, J., Schulzrinne, E., Camarillo, G., Johnston, A., Peterson, J., Spark,s R., Handley, M., Schooler, E. Jun 2002. SIP: Session Initiation Protocol, IETF RFC 3261. Available online: https://www.ietf.org/rfc/rfc3261.txt (accessed on 30 November 2015).
47. Zhang, X., Law, C., Wang, C., Lau, F.C.M. Towards pervasive instant messaging and presence awareness. International Journal of Pervasive Computing and Communications, 2009,Vol. 5, No. 1, pp. 42-60.

48. Latvakoski J., Heikkinen J. A Trustworthy Communication Hub for Cyber-Physical Systems. MDPI Future Internet Journal. *Future Internet* **2019**, *11*(10), 211; https://doi.org/10.3390/fi11100211 (registering DOI) 38p.
49. Rodrigo Roman, Jianying Zhou, and Javier Lopez. 2013. On the features and challenges of security and privacy in distributed Internet of Things. Computer Networks 57, 10 (2013), 2266–2279
50. Aafaf Ouaddah, Hajar Mousannif, Anas Abou Elkalam, and Abdellah Ait Ouahman. 2017. Access control in The Internet of Things: Big challenges and new opportunities. Computer Networks 112 (2017), 237–262.
51. Álvaro Alonso, Federico Fernández, Lourdes Marco, and Joaquín Salvachúa. 2017. IAACaaS: IoT Application-Scoped Access Control as a Service. Future Internet 9, 4 (2017), 64.
52. Grant Ho, Derek Leung, Pratyush Mishra, Ashkan Hosseini, Dawn Song, and David Wagner. 2016. Smart locks: Lessons for securing commodity Internet of Things devices. In Proceedings of Asia Conference on Computer and Communications Security. ACM, 461–472.
53. Samuel Paul Kaluvuri, Alexandru Ionut Egner, Jerry den Hartog, Nicola Zannone: SAFAX - An Extensible Authorization Service for Cloud Environments. Front. ICT 2015 (2015)
54. Huhnlein D., Rossnagel H., Zibuschka J. 2010. Diffusion of federated identity management. in Sicherheit. P-170, 25-36.
55. eXtensible Access Control Markup Language (XACML). Version 3.0. Oasis Standard. 2013.
56. Takabi H., Joshi J., Ahn G-J. Security and privacy challenges in cloud computing environments. IEEE Secur. Priv. 8, 24-31. 2010.
57. SEAS Ontology. https://w3id.org/seas, consulted on January 2018
58. Lefrançois, Maxime. "Planned ETSI SAREF Extensions based on the W3C&OGC SOSA/SSN-compatible SEAS Ontology Paaerns." Workshop on Semantic Interoperability and Standardization in the IoT, SIS-IoT. 2017
59. Daniele L., den Hartog F., Roes J. (2015) Created in Close Interaction with the Industry: The Smart Appliances REFerence (SAREF) Ontology. In: Cuel R., Young R. (eds) Formal Ontologies Meet Industry. FOMI 2015. Lecture Notes in Business Information Processing, vol 225. Springer, Cham
60. Flexiblepower Alliance Network (FAN), Energy flexibility interface. https://github.com/flexiblepower/efi
61. Flexiblepower Alliance Network (FAN), Energy Flexibility Platform & Interface (EF-Pi). http://flexible-energy.eu/ef-pi/
62. Energy flexibility interface, http://flexiblepower.github.io/technology/efpi/, consulted on January 2018
63. Common Information Model (CIM) standards, http://wiki.cimtool.org/, consulted on January 2018
64. Drools Language documentation, https://docs.jboss.org/drools/release/6.5.0.Final/drools-docs/html_single/index.html#DroolsLanguageReferenceChapter, consulted on January 2018
65. Geetika T. Lakshmanan, Ying Li, and Rob Strom. Placement Strategies for Internet-Scale Data Stream Systems. IEEE Internet Computing, November/December 2008.
66. Valeria Cardellini, Vincenzo Grassi, Francesco Lo Presti, Matteo Nardelli. Optimal Operator Placement for Distributed Stream Processing Applications, Proc. of the 10th ACM International Conference on Distributed and Event-based Systems (DEBS'16), Irvine, CA, USA, June 2016.
67. Mordecai Y., Dori D. Minding the Cyber-Physical Gap: Model-Based analysis and mitigation of syetemic perception-induced failure. Sensors 2017 jul. 17(7): 1644. Published online 2017 Jul 17. doi: 10.3390/s17071644
68. N. Choi, I.-Y. Song, and H. Han. A survey on ontology mapping. SIGMOD Rec., 35(3):34–41, 2006.
69. D. Trivellato, N. Zannone, M. Glaundrup, J. Skowronek, and S. Etalle. A semantic security framework for systems of systems. Int. J. Cooperative Inf. Syst., 22(1), 2013.
70. R. Culmone, G. Rossi, and E. Merelli. An Ontology Similarity Algorithm for Bioagent. In Proc. of NETTAB'02, 2002.
71. L. D. Ngan, T. M. Hang, and A. E. S. Goh. Semantic Similarity between Concepts from Different OWL Ontologies. In Proc. of IEEE International Conference on Industrial Informatics, pages 618–623, 2006.
72. H. Nguyen and H. Al-Mubaid. A Combination-based Semantic Similarity Measure using Multiple Information Sources. In Proc. of IRI'06, pages 617–621. IEEE Press, 2006.
73. Damen, S., Zannone, N.: Privacy implications of privacy settings and tagging in facebook. In: Secure Data Management. pp. 121–138. LNCS 8425, Springer (2013).

74. Rauf Mahmudlu, Jerry den Hartog, Nicola Zannone: Data Governance and Transparency for Collaborative Systems. DBSec 2016: 199-216.