Dependability Demands and State of the Art in the Internet of Things

Thomas Frühwirth, Lukas Krammer, Wolfgang Kastner
Vienna University of Technology / Automation Systems Group
{tfruehwirth, lkrammer, k}@auto.tuwien.ac.at

***Abstract*—The number of devices connected to the Internet of Things (IoT) has steadily been increasing over the last years and so has the variety of applications. There is no sign for this trend to weaken. A broad subset of them imposes requirements on attributes, such as availability, reliability, safety and many more. These attributes are commonly subsumed under the term dependability. In this paper we first give a brief introduction to dependability. Three application scenarios typical for the IoT are presented and their demands for each dependability attribute are discussed. Furthermore, an overview of the current state of the art is presented. Last, a summary of required technologies and future research topics to enable dependability in the IoT is given.**

I. Introduction

Pushed by recent developments, such as IPv6, and the availability of cheap while powerful embedded systems, the number of devices connected to the Internet is increasing rapidly. The complexity of the resulting system is far too high to be handled by human individuals. An infrastructure where smart devices communicate among each other and fulfill their tasks autonomously is required. This is the Internet of Things (IoT). In the near future most of the devices connected to the Internet will be operated without any human intervention. Their types and functionality vary from simple temperature sensors over smoke detectors to more complex systems, including smart metering, health monitoring and others. Many of these applications place high demands on dependability. There are several sophisticated technologies and protocols to provide dependable communication in Local Area Networks (LANs) [1] and control networks [2], [3]. But to the best knowledge of the authors there is yet no comprehensive approach addressing the use cases and special requirements of dependable distributed systems w.r.t. the IoT.

II. Dependability

Traditional fields of application where dependability is of particular interest are X-by-wire systems, e.g., steer-by-wire in the automotive industry or fly-by-wire in the aerospace industry, operation of industrial plants, and basically any other surroundings where human health is threatened if the computer system fails. All these systems are employed in well controlled environments with clear limits regarding their physical expansion and the number and types of interconnected devices. These assumptions are clearly not applicable to dependable distributed computing systems in the IoT. Therefore, new ideas are required to bring the know-how gathered from years of research and development in the field of dependability to a larger scale. Addressing dependability in a structured and comprehensive manner is achieved on the basis of the dependability tree [4].

978-1-4673-7929-8/15/$31.00 © 2015 IEEE

A. Dependability Tree

The dependability tree illustrated in Fig. 1 is divided into threats, attributes and means.

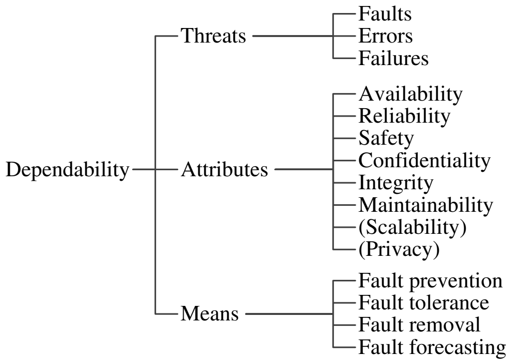

Fig. 1. Dependability tree (adapted from [4])

1) Threats: In any computing system the root cause of a malfunction is called **fault**. Faults are either caused by physical effects, such as electromigration [5], interference in Wireless Sensor Networks (WSNs) [6], etc. or human-made. Human-made faults may be intentional (e.g., security attacks) or unintentional (e.g., design faults, improper operation or incorrect maintenance). A fault is active if it produces an error; otherwise, it is dormant. Furthermore, it can be distinguished between transient and permanent faults. In contrast to permanent faults, which are easier to find and repair, transient faults appear and disappear spontaneously. An **error** is the part of the system state that might lead to a failure. Depending on the type of system, an error may be detected and corrected before it manifests as a failure. If a fault is activated and no measures are taken to correct the subsequent error, the system might deviate from its specified or intended behavior. This is then called a **failure**. In any dependable system special precautions are taken to avoid this situation. A failure of a subsystem (e.g., a temperature sensor) becomes a fault to the supersystem (e.g., the controller), where it again might cause an error. This causal relationship is known as *fundamental chain* [4].

2) Attributes: Only two of the dependability attributes can be defined in a quantitative manner: **availability** and **reliability**. Availability refers to the up- and downtime of a system. It gives a metric for the readiness of usage for a service. Thus, availability is a key attribute in regard to dependability. On the other hand, reliability gives the probability that a system will conform to its specification after a given time t. **Safety** is defined as the absence of catastrophic consequences on the user and the environment. **Confidentiality** refers to preventing information from unauthorized disclosure. The absence of improper system state modifications is called **integrity**. Confidentiality, integrity and availability often are subsumed under the term **security**. **Maintainability** is the system's capability to be repaired and modified. Further attributes, which are not present in the classical dependability tree but of particular interest in the IoT, are **scalability** and **privacy**. Several definitions of scalability exist, but none of them is particularly precise [7]. For the purpose of this paper, we take scalability as the ability of a system to be extended by further components. Privacy is often defined as ”the interest an individual has to control how information about them is collected, used, and shared” [8].

3) Means: The way today's dependable systems are designed is mainly motivated by four techniques. Development guidelines, international standards and other procedures are followed to enhance **fault prevention**. However, the existence of faults can never be ruled out with absolute certainty. The key to achieve high dependable systems is **fault tolerance**, which is most often deployed by redundancy. The number of replica required to tolerate the existence of a single or even multiple faults varies depending on the failure-mode [9]. Once a fault is detected, **fault removal** may be applied. Finally, **fault forecasting** techniques allow to make statements about to which extent certain dependability attributes are fulfilled in a particular system.

III. Application Scenarios

In this section various application scenarios are presented to show that the demand for dependability features discussed in the previous section indeed exists in the IoT. The scenarios differ regarding their application domain, their complexity as well as the dependability attributes of interest. Table I summarizes the relevance of each dependability attribute for the investigated applications. However, the result of this evaluation is not predetermined but arguable. It may vary according to the needs of user of the specific system. Nevertheless, the examples are suitable to provide a more profound intuition of the properties and the importance of each dependability attribute.

TABLE I. Evaluation of dependability attributes in various application scenarios. 1 .. minor, 2 .. medium, 3 .. high priority

Domain / Appl. Scenario	**Availability**	**Reliability**	**Safety**	**Confidentiality**	**Integrity**	**Maintainability**	**(Scalability)**	**(Privacy)**
Smart Grid / Smart Meter	2	2	1	2	3	3	3	2
ITS / Variable Speed Limit	1	3	3	1	3	2	3	1
AAL / Emergency Call System	3	3	3	2	3	1	2	1

A. Smart Grid

A key factor in reducing greenhouse gas emission and limiting global warming is the electricity industry, as it accounts for a considerable fraction of the world's carbon-dioxide emissions. 25% of the worlds anthropogenic greenhouse gas emission is caused by the energy sector, subsuming electricity as well as heat production [10]. Furthermore, due to the extension of renewable energies, the power grid will become more diverse in the future. A way to support the change in infrastructure, while increasing its efficiency, is the smart grid enabling projects like *demand response* [11], [12]. In demand response applications the consumers' electricity consumption patterns are time-shifted to match with the available resources. A typical example is recharging electric vehicles when the general demand for electricity is low. A major building block of any smart grid is smart metering, which can be used to generate energy consumption profiles for single buildings and even single devices.

A smart metering system does not necessarily require high availability, but collecting data should be possible at regular intervals. The same argument holds true for reliability. As the system is only used for monitoring, safety is only of minor importance. However, the data should be protected from unauthorized access, as it contains information about personal behaviors of the system's users. Smart meters are generally mounted in people's homes. Smart metering is thus a rare example of applications where a particular part of the system's critical infrastructure is not in the sphere of its operator. Therefore, special precautions have to be taken to guarantee the physical integrity of the device itself as well as the integrity of the transmitted data. As the devices are typically not accessible by their operators, maintainability is of special interest. Naturally, smart grid applications are large scale systems with regard to the number of devices and their physical extension. Therefore, scalability is an important attribute. The information provided by a smart meter is not highly sensitive but might be used by various different actors, e.g., by the customer to monitor its own energy profile, by the energy operator to calculate the electricity bill and by public or private organizations to perform statistical analysis. Thus, privacy is of concern when designing a smart grid.

B. Intelligent Transportation Systems

The next example application presented is situated in the field of Intelligent Transportation Systems (ITS) [13]. ITS are targeted to increase the efficiency of the traffic system as a whole and to make the driving experience less stressful and safer by collecting, fusing and using live data about road and environmental conditions. A very basic but, nevertheless, particularly useful application in ITS is the Variable Speed Limit (VSL) [14] system. Thereby, electronic speed limit signs are used to reduce the risk for accidents, traffic jams and limit value violation for air pollution.

A VSL system does not impose high demands on availability. It is merely an additional feature in traffic control but car drivers will fall back to their normal behavior if the system is non-operational. Nevertheless, reliability and safety are extremely critical attributes. Displaying an inappropriate low speed limit may cause rear-end collisions, while unintended high speed limits might lull the drivers into a false sense of security. Confidentiality can be neglected as the information is intended to be publicly available. Integrity, however, is of highest priority for the same reasons as for reliability and safety. The devices are accessible by their operators but VSL systems might be large. Thus, maintainability should be considered during system design. Special attention has to be paid on scalability, as ITS systems quickly reach countrywide and even international scale. Privacy is of no particular interest for this use-case, but it might be important for other applications in the ITS domain.

C. Ambient Assisted Living

Ambient Assisted Living (AAL) subsumes any system that allows elderly or handicapped people, who would otherwise have to reside in a hospital or an assisted living facility, to stay in their private homes. It is a collection of applications targeted to support humans in their everyday lives. AAL reduces the cost of the healthcare system while increasing the quality of life for its users. A typical application in the field of AAL is an Emergency Call System (ECS). The users carry some kind of wireless device that allows them to send an alert signal on the press of a single button in case of an emergency. Such devices can also be equipped with additional sensors for monitoring vital parameters or detecting accidents.

Availability, reliability and safety are key factors in an ECS, as any outage or malfunction of the service would have serious consequences on the health of its user in case of an emergency. Confidentiality is less important than correct functionality. Nevertheless, health data is very sensitive data and protection might be desirable. The integrity of the transmitted data (name and address / ID of the patient) has to be protected as any manipulation renders the information useless. Maintainability is not a concern as the functionality is very simple and faulty devices can easily be replaced. Adding new ECS devices shall be possible, but the complexity of the overall system is rather low. Therefore, scalability is of medium interest. Finally, if the ECS is an isolated system, privacy is not an issue. However, privacy comes into play if the data is shared, e.g., with insurance companies to create statistics.

IV. State of the Art

The communication infrastructure is a major building block to fulfill the dependability requirements of IoT applications identified so far. Devices may be connected to the Internet directly via wireless communication, such as digital cellular networks including GSM, UMTS and LTE or a wired connection, such as DSL, ADSL, cable networks, etc. But more commonly the devices are connected to some local infrastructure (PAN, LAN, etc.) and linked to the Internet via one or several gateways. While the Time Division Multiple Access (TDMA) scheme is well understood and used in various wired protocols to guarantee dependable communication [15], the situation turns out to be of much higher complexity in the wireless world. The openness of the media and physical limitations regarding the node's memory, energy, processing and transmission power are major problems. Various approaches have been taken to improve reliability in WSN [16] [17].

A promising protocol enhancing safety in control networks is openSAFETY. The IEC 61784-3-13 [18] standard is compatible with basically any control network and currently supports PROFINET, SERCOS III, EtherNet/IP, Modbus-TCP and POWERLINK. A main feature of openSAFETY is its time-stamping functionality used to detect and correct errors including message duplication, unintentional chronological sequences and transmission delays. Furthermore, an openSAFETY frame always contains two identical subframes protected by individual checksums. The protocol is certified up to Safety Integrity Level 3 (SIL 3) defined in IEC 61508 [19].

To establish dependable connections between devices in the IoT, routing is of particular interest. First, routing within the local infrastructure may be error prone due to unreliable links particularly in wireless infrastructures. Multipath routing may be used to improve reliability [20]. Second, according to [21], linking the local infrastructure to the Internet via one Border Router (BR) (i.e., gateway) is in general not applicable for dependable applications because this creates a single point of failure. Thus, if redundancy is not addressed at the application level, e.g., by requesting the same service from multiple providers, multiple BRs are required. Third, more sophisticated routing strategies, like the IP Fast Reroute Framework defined in [22], are necessary to guarantee timely and reliable message transmission in the Internet, even in the presence of a faulty routing node. Furthermore, for many IoT applications multicast routing in combination with a publisher/subscriber paradigm may be beneficial over a client/server based communication scheme. This way a service provider does not need to handle multiple requests simultaneously and can even remain in a power-saving sleep mode most of the time.

A discovery mechanism supports the implementation of the proposed, rather simple, example systems and might also be useful to achieve redundancy on application level. In [23] a comprehensive approach to service discovery is presented. Additionally, the Real-World Service Discovery and Provisioning Process (RSDPP) is proposed. It supports an operator in finding the desired services by making use of ”lightweight ontologies” extracted from web search engines and encyclopaedias. In [24], a Role Based Access Control (RBAC) inspired approach for authentication and access control in the IoT is presented.

Web Services (WSs) [25] provide access to resources in an abstract, machine interpretable form using Web-related technologies, such as SOAP messages in combination with HTTP and XML documents to specify and offer services. WSs are specified by the World Wide Web Consortium (W3C). The W3C furthermore defined a number of protocols for service agreement, discovery, notification, secure communication, reliable messaging and other services in the world of WSs. A restriction to standard WSs is imposed by the REpresentational State Transfer (REST) approach, which limits the available interface essentially to standard HTTP methods and, furthermore, prescribes a stateless communication paradigm. WSs are implemented following the traditional architecture of the World Wide Web, i.e., a server (agent, in WS terminology) offers its services to an unknown number of clients. It is questionable whether a traditional client/server based approach is well suited for dependable applications in the IoT because of the often very limited processing capabilities and constraint energy resources of nodes. It may be preferable to use a gateway instead, which is connected to a set of local devices and offers the WS interface to the WWW.

The RELYonIT project [3] is an international research project addressing dependability for the IoT. It analyzes the impact of environmental factors, such as temperature, interference and others on WSNs. The project aims to increase dependability, mainly focused on reliability, by designing protocols that are aware of these environmental conditions. RELYonIT, however, does not address all attributes of dependability and exclusively covers dependability in WSNs.

The above mentioned protocols, technologies and projects make important contributions to realize dependable distributed applications in the IoT, but none of them addresses all attributes of dependability in a complete and comprehensive manner. Furthermore, there has not yet been taken particular effort to analyze the potential of redundancy on application level.

V. Current Research

The only feasible way to realize dependable systems with reasonable costs and in the short or medium term is to build on the existing infrastructure also used for the conventional Internet. This enables applications with huge flexibility but also raises new problems, e.g., exposing them to literally all types of **security attacks** known in the world. Furthermore, the **communication channels** used are usually **unpredictable** in terms of performance and latency. End points in IoT applications are often battery powered and, thus, have **scarce energy resources** and **limited processing capabilities**.

Building on top of existing infrastructure comes along with using existing technologies and protocols supported by the core components of the conventional Internet. The objective is not to develop an entirely new framework but to identify, choose and combine existing **building blocks** to realize dependable distributed systems in the IoT. A selection of these building blocks, including **communication technologies**, **safety mechanisms**, **routing**, **discovery**, **access control** and **Web Services** was already presented in the previous section. But there are certainly others, e.g., **encryption mechanisms** and **naming services**, which have to be identified and analyzed for their suitability in the IoT.

A systematic guideline is desired that supports a system designer in choosing a subset of the available building blocks and combining them in a way that best fits the application requirements. The proposed process for creating dependable distributed systems in the IoT is composed of three steps. First, the system designers will have to analyze and rate the intended application w.r.t. the dependability attributes it shall provide as suggested in Section III. This step allows to abstract from the specific implementation details of an application and instead focus on its properties particularly relevant to realize dependability. Thus, a generic guideline may be defined solely based on well defined attributes. Second, the system designers will follow the guideline to choose an appropriate subset of building blocks that satisfies the dependability attributes to the levels determined in the previous step. Defining this guideline is in the focus of current and future research. Finally, in step three, a simulation environment will allow the user to combine the suggested building blocks. This provides the user with the means to (1) verify whether the architecture, technologies and protocols suggested by the guideline can be seamlessly integrated, (2) perform quantitative measurements of availability and reliability, and (3) gather a more detailed understanding of the resulting system before actually implementing it.

References

[1] TTSAC, SAE AS2D, *Time-Triggered Ethernet (AS 6802)*, 2011.

[2] L. Krammer, "Dependability in building automation networks," Ph.D. dissertation, Vienna University of Technology, 2014.

[3] RELYonIT Consortium, "RELYonIT - Research by Experimentation for Dependability on the Internet of Things."

[4] A. Avizienis, J.-C. Laprie, B. Randell *et al.*, *Fundamental concepts of dependability*. University of Newcastle upon Tyne, Computing Science, 2001.

[5] T. Adler, H. Brocke, L. Hedrich, and E. Barke, "A Current Driven Routing and Verification Methodology for Analog Applications," in *Proceedings of the 37th Annual Design Automation Conference*, ser. DAC '00. New York, NY, USA: ACM, 2000, pp. 385–389.

[6] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of Interference on Multi-hop Wireless Network Performance," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '03. New York, NY, USA: ACM, 2003, pp. 66–80.

[7] D. Law, "Scalable means more than more: a unifying definition of simulation scalability," in *Simulation Conference Proceedings*, vol. 1, Dec 1998, pp. 781–788 vol.1.

[8] P. Madsen, N. M. C. Mont, and R. Wilton, "A Privacy Policy Framework–A position paper for the W3C Workshop of Privacy Policy Negotiation," 2006.

[9] H. Kopetz, "On the Fault Hypothesis for a Safety-Critical Real-Time System," in *Automotive Software-Connected Services in Mobile Networks*, ser. Lecture Notes in Computer Science, M. Broy, I. H. Krueger, and M. Meisinger, Eds. Springer Berlin Heidelberg, 2006, vol. 4147, pp. 31–42.

[10] L. M. IPCC Core Writing Team, Rajendra K.Pachari, "Climate change 2014 synthesis report," IPPC, Tech. Rep., 2014.

[11] P. Varaiya, F. Wu, and J. Bialek, "Smart Operation of Smart Grid: Risk-Limiting Dispatch," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 40–57, Jan 2011.

[12] W. Kastner, M. Jung, and L. Krammer, "Future Trends in Smart Homes and Buildings," in *Industrial Communication Technology Handbook, Second Edition*, R. Zurawski, Ed. CRC Press, Inc., 2014, ch. 59.

[13] T. Vaa, M. Penttinen, and I. Spyropoulou, "Intelligent transport systems and effects on road traffic accidents: state of the art," *Intelligent Transport Systems, IET*, vol. 1, no. 2, pp. 81–88, June 2007.

[14] P. Allaby, B. Hellinga, and M. Bullock, "Variable Speed Limits: Safety and Operational Impacts of a Candidate Control Strategy for Freeway Applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 4, pp. 671–680, Dec 2007.

[15] R. Obermaisser, *Time-Triggered Communication*, 1st ed. Boca Raton, FL, USA: CRC Press, Inc., 2011.

[16] N. Maalel, E. Natalizio, A. Bouabdallah, P. Roux, and M. Kellil, "Reliability for Emergency Applications in Internet of Things," in *Proceedings of the IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)*, May 2013, pp. 361–366.

[17] M.-A. Nef, L. Perlepes, S. Karagiorgou, G. I. Stamoulis, and P. K. Kikiras, "Enabling QoS in the Internet of Things," in *Proceedings of the 5th International Conference on Communication Theory, Reliability, and Quality of Service (CTRQ)*, 2012, pp. 33–38.

[18] International Electrotechnical Commission (IEC), "61784-3-13 – Industrial communication networks - Profiles - Part 3-13: Functional safety fieldbuses - Additional specifications for CPF 13," 2010.

[19] ——, "61508 – Functional Safety of Electrical / Electronic / Programmable Electronic Safety-related Systems," 2011.

[20] G. Wagenknecht, M. Anwander, M. Brogle, and T. Braun, "Reliable multicast in wireless sensor networks," *FGSN'08*, pp. 69–72, 2008.

[21] L. Deru, S. Dawans, M. Ocaña, B. Quoitin, and O. Bonaventure, "Redundant Border Routers for Mission-Critical 6LoWPAN Networks," in *Real-World Wireless Sensor Networks*. Springer, 2014, pp. 195–203.

[22] M. Shand and S. Bryant, "RFC5714: IP Fast Reroute Framework," *IETF Request for Comments*, 2010.

[23] D. Guinard, V. Trifa, S. Karnouskos, P. Spiess, and D. Savio, "Interacting with the SOA-Based Internet of Things: Discovery, Query, Selection, and On-Demand Provisioning of Web Services," *IEEE Transactions on Services Computing*, vol. 3, no. 3, pp. 223–235, July 2010.

[24] J. Liu, Y. Xiao, and C. Chen, "Authentication and Access Control in the Internet of Things," in *32nd International Conference on Distributed Computing Systems Workshops (ICDCSW)*, June 2012, pp. 588–592.

[25] K. Mockford, "Web services architecture," *BT Technology Journal*, vol. 22, no. 1, pp. 19–26, 2004.