Quality of Service Parameter Tracking and Transformation in Industrial Applications

György Kálmán

Centre for Cyber and Information Security
Critical Infrastructure Protection Group
Norwegian University of Science and Technology
mnemonic AS
Email: gyorgy.kalman@ntnu.no

***Abstract*—Quality of Service (QoS) is a key property to deliver communication services in automation environments. Machine to machine communication offers both an opportunity and poses a challenge for communication networks. In this paper, an overview of typical QoS metrics is given and their relation to automation metrics is analysed. The paper recommends the use of formalized methods from industrial safety to introduce formalized management of communication network requirements in an industrial scenario.**

Keywords–critical infrastructure; QoS; metrics; automation; operational envelope

I. Introduction

Since the introduction of packet switched networks, questions and analyses around the possible service level have been a hot topic. In current networks, the use of best-effort forwarding is dominating. Although it is very efficient, guaranteeing end-to-end connection parameters is a challenge.

The technology landscape is similar in both office or communication and industrial networks: on the Local Area Network (LAN) field, Ethernet is dominating, on the WAN side, standard telecommunication solutions are used also for industrial applications.

Since its introduction in industrial automation, Ethernet's determinism has been a returning concern, mainly because of both outdated information (use of, e.g., 10-Base2) and bus-like topologies [2] with long chains of switches.

Most of the bandwidth-related problems were solved with the introduction of Gigabit Ethernet and for the most demanding applications, technologies like EtherCAT, with intrinsic QoS are available. For traditional switched networks, there are efforts for the inclusion of a resource management plane in the IEEE 802.1 Time-Sensitive Networking Task Group (TSN).

The paper is structured as follows: the second section gives an overview of different QoS metrics. Section 3 provides an overview of Distributed Control System (DCS) structures, Section 4 provides an analysis of how formal methods from safety development could be adapted in QoS requirement specification. Section 5 analyses the need for requirements tracking. The sixth section presents parameters of a control loop and how QoS parameters can be converted between the industrial and communication metrics. Section 7 draws the conclusion and provides an outlook on future work.

II. Quality of Service

QoS is the measure of transmission quality and service availability of a network [3], thus not only limited to actual forwarding parameters like bandwidth and delay, but also, e.g., availability, reconfiguration time and reliability.

Keeping a certain service level was a requirement in telecommunication networks and it was a natural decision to have features to support service level definition when packet switched networks were introduced in the telecom networks.

Providing QoS in Local Area Networks (LANs) networks was focused on services, where at least one of the communicating parties was a human. The services could range from web browsing through VoIP to multi-party video conferencing. The parameters were adopted to the human perception and also tolerance for disturbances was adapted to the human users. The metrics for service quality were not new either at that time; telecommunication networks had service levels defined already and since those were also technical and focused on human users, metrics introduced there were also adapted to computer networks, like Ethernet or more generally, Internet Protocol (IP). In current industrial applications, IPv4 is generally used, if needed, then as IPv4 islands interconnected with tunnels over IPv6 networks. In Internet of Things (IoT) installations, the use of IPv6 is expected as a result of the large number of connected devices.

The evolution of technology showed, that in the vast majority of cases, an over dimensioning of the network resources is both the cheapest and easiest to manage.

A. Telecommunication metrics

As an example, Asynchronous Transfer Mode (ATM) metrics for traffic contracts are composed from traffic parameters such as:

- *Peak Cell Rate (PCR)* The maximum allowable rate at which cells can be transported along a connection in the ATM network. The PCR is the determining factor in how often cells are sent in relation to time in an effort to minimize jitter.
- *Sustainable Cell Rate (SCR)* A calculation of the average allowable, long-term cell transfer rate on a specific connection.
- *Maximum Burst Size (MBS)* The maximum allowable burst size of cells that can be transmitted contiguously on a particular connection.

and QoS parameters,

- *Cell Transfer Delay (CTD)* The delay experienced by a cell between the time it takes for the first bit of the cell

 ISBN: 978-1-61208-473-2

to be transmitted by the source and the last bit of the cell to be received by the destination. Maximum Cell Transfer Delay (Max CTD) and Mean Cell Transfer Delay (Mean CTD) are used.

- *Peak-to-peak Cell Delay Variation (CDV)* The difference between the maximum and minimum CTD experienced during the connection. Peak-to-peak CDV and Instantaneous CDV are used.
- *Cell Loss Ratio (CLR)* The percentage of cells that are lost in the network due to error or congestion and are not received by the destination.

The list shows the focus areas of QoS already in the 90s: bandwidth (in bits per second), burstiness and parameters related to disturbances in forwarding.

In addition to these connection-related parameters, the communication network had also network-wide parameters in other relations, like redundancy with, e.g., reconfiguration time in case of link loss or routing alternatives.

ATM is raised as an example, since it offers one of the widest range of possibilities for QoS. It also introduced a couple of concepts, which, although ATM was later deemed as a failure, do a comeback in today's QoS networks.

B. Metrics on packet switched networks

On packet switched networks, initially the focus was on efficient forwarding. Efficiency and simple network operation lead to cheaper devices and ultimately to today's technology landscape with the domination of Ethernet and IP.

While there were different approaches for QoS (integrated and differentiated services), the main QoS metrics were bandwidth, loss, delay and jitter [3]. In future installations with IPv6 it is expected that the use of differentiated services will be more widespread, as after RFC 2460/3697, the properties of Traffic Class and Flow Label can be used to select flows of the aggregated traffic and grant priority. The 20 bit field of Flow Label also allows a large number of flows to be present concurrently which would fit even a large industrial deployment. The impact of this feature however depends on the timing of tasks running on the network and also how this field could be used for other properties important in the automation applications: redundancy and reconfiguration time in case of link loss.

An effort to include some of the traffic engineering possibilities of ATM for LANs is the IEEE Shortest Path Bridging (SPB). This standard is being developed by the TSN working group and allows, amongst others call admission, resource reservation over the whole path. SPB has raised a high interest in the automation field and most of the industry is either contributing directly or closely following the development.

C. Automation

QoS requirements of an automation system tend to be very different than those of an office network. The protocol set used is different and the typical communication inside an automation system runs on Layer 2 [13]. Sources and sinks of traffic streams are typically machines with little tolerance on disturbances, but good predictability in communication.

The network topology of automation networks is often contributing to the challenges around QoS [5]. Networks are built with low port count switches. This typically results in an infrastructure that has more devices than an office network. A bigger refinery can have several hundreds of switches with a typical branching factor of 4-7. The still widely used bus-topology leads to even longer forwarding chain, introducing delay and jitter, which only exists in considerably larger networks in the office/telecommunication scenarios.

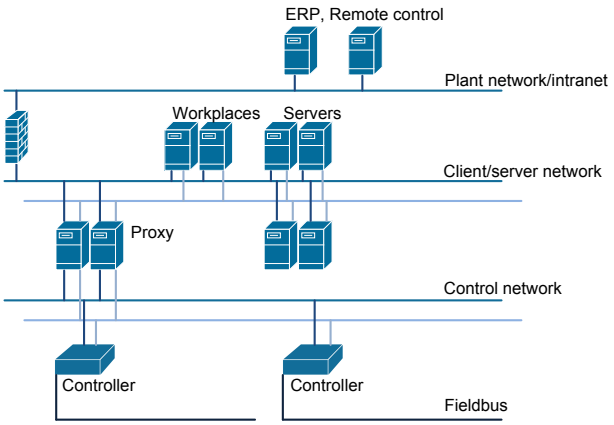

Figure 1. Traditional DCS network architecture

III. DCS ARCHITECTURE

Control systems are traditionally built using a three network levels (Figure 1). The plant, the client-server and the control network. These levels might have different names, but they share the following characteristics:

- *Plant network* is home of the traditional IT systems, like Enterprise Resource Planning (ERP), office services and other support applications. It is typically under the control of the IT department.
- Client-server network is the non-time critical part of the automation system, where the process-related workplaces, servers and other support entities are located. It is firewalled from the plant network and is under the control of Operations.
- Control network includes everything close to the actual process: controllers, sensors, actuators and other automation components. Typically it follows a strict time synchronization regime and contains the parts of the network with time-critical components. It is accessible through proxies from the client-server network and under the control of Operations.

Remote monitoring was introduced to industrial applications decades ago with the different Supervisory Control and Data Acquisition (SCADA) systems. These used various communication technologies (leased lines, radio links, etc.) to feed in status data to a central monitoring entity. Typically remote control was not available.

With current developments in the smart grid and IoT the possibilities for remote operations is being extended by taking current communication solutions in use. The extension of the features also requires a well-defined network infrastructure [9].

 ISBN: 978-1-61208-473-2

A. QoS in automation

Traffic flows in automation typically are machine to machine (M2M). This property and the systems connectivity to the physical world require both different tolerances for disturbances and potentially different metrics [7].

An automation system somewhere in the process is connected to the physical world. This means, that amongst others, it has to refer to real time. Forwarding disturbances might lead to potentially dangerous situations with implications far beyond a dropped Voice over Internet Protocol (VoIP) call.

The definition of QoS requirements in the automation world has its roots in the definition of control loops. In control of the early DCSs bus and serial links were used, which typically operated in a slotted or polled way. This allowed the automation engineers to exactly set the communication parameters to meet the requirements of the control system in a deterministic way.

For special applications, technologies with intrinsic QoS are used, e.g., Ethernet for Control Automation Technology (EtherCAT), which allows deterministic communication, but represents a minority of installations. In the following, focus will be on solutions, where no intrinsic QoS is available.

The physical world connection also has an influence on the used QoS metrics. In automation, beside bandwidth, time and availability related metrics are more emphasized, like delay and jitter or availability (redundancy, reconfiguration time). A special aspect is also the quality of time synchronization. The importance and weighting of these metrics is different compared to the telecommunication or other communication operations. One of the most important differences is, that at the moment there is no protocol which would bridge the gap between requirements specification in automation terms and network operations, which results in extended engineering work and challenging life-cycle support. This is in contrast with, e.g., VoIP, where protocols like the Resource Reservation Protocol (RSVP) can be used to reserve resources on the communication path.

IV. Requirements specification

Defining requirements and keeping the original intention in complex systems is a problematic task. In automation, the main challenge is, that the requirements are defined in the automation context, but the bearer network uses by default different metrics for expressing forwarding parameters.

In a control loop, typical parameters are control frequency (how often the data is refreshed or modified), maximum tolerable delay, jitter and availability parameters. One of the most demanding applications, where no technology with intrinsic QoS is used is substation automation with IEC 61850 [6].

IEC 61850 is a standard for communication networks and systems for power utility automation. This protocol is a great step forward for substation automation, as it, amongst others translates all information into data models, which is supported by the application focused architecture. This speeds up the engineering process both in planning and integration [4].

However, also IEC 61850 is not defining exact QoS requirements for the network infrastructure. Although the Specific Communication Service Mapping (SCSM) feature allows the definition of communication links inside the IEC 61850 world, the translation of requirements is not included.

When the control loops are defined, the current process is based on individual mapping of automation requirements to network QoS parameters. This process, although not efficient, can and is working for smaller installations, but suffers from scalability problems.

The lack of direct coupling between the automation and communication parameters typically leads to very pessimistic QoS requirements and over dimensioning the network capacity which leads to excess cost.

In the Internet of Things (IoT) scenario, where the automation networks are extended behind the LAN [8], tracking requirements is becoming more important. Very strict parameters of the automation system on the LAN can be mixed into the WAN requirements, which might lead to prohibitive cost on communication. Validity of requirements for each flow has to be analysed to ensure an efficient fit. The efforts for keeping the QoS parameters as close to the requirements as possible can lead to more efficient and cheaper operation.

A. Industrial safety

Conversations on Safety Integrated Systems (SIS) mainly include questions on QoS. The cause is that these installations share the communication network between the automation task and the safety function (as they can also share infrastructure with the fire alarm system). In a safety sense, SIS have no QoS requirements. The safety logic is built in a way, that a communication error is interpreted as a dangerous situation and the safety function will trip. So the system avoids dangerous situations at the expense of lower productivity and availability.

Safety as such is an availability question and through availability, it implies QoS requirements on the automation system as any other communication task. Special treatment is not required.

Although a solution like this does not exists for communication QoS, but the industry has a field, where a similar challenge was solved with structured approach and formal methods: safety. Safety is already considered as a process, which is present for the whole life cycle of the product.

Safety systems are classified into 4 levels, Safety Integrity Level (SIL) 1 to 4. The different levels pose well-defined requirements towards the system. These integrity levels cover all aspects of the system, including hardware, software, communication solution and seen in contrast with the application. A similar approach could be also beneficial for formalizing the relationship between the automation application and the bearer network.

The IEC 61508 standard requires that each risk posed by the components of the safety system is identified and analysed. The result of the risk analysis should be evaluated against tolerability criteria.

Key processes of a safety development are risk analysis and risk reduction. These are executed in an iterative manner until the acceptable risk level is achieved. A possible method for risk classification is shown on figure 2 from the United Kingdom Health and Safety Executive.

Analogue to this, a similar approach could be used for defining an operational envelope for the communication infrastructure. All possible flows of data should be identified (analogue with identifying risk), which is possible with high

 ISBN: 978-1-61208-473-2

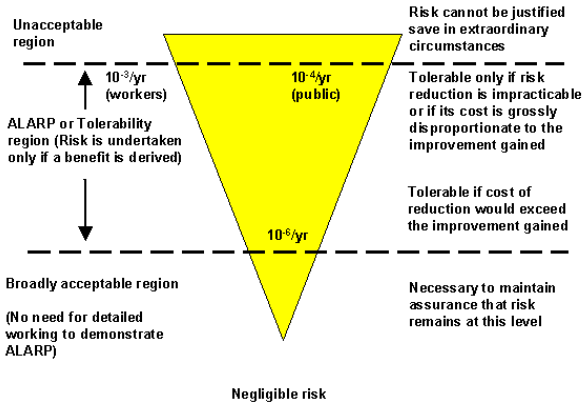

Figure 2. The Health and Safety Executive's Risk criteria

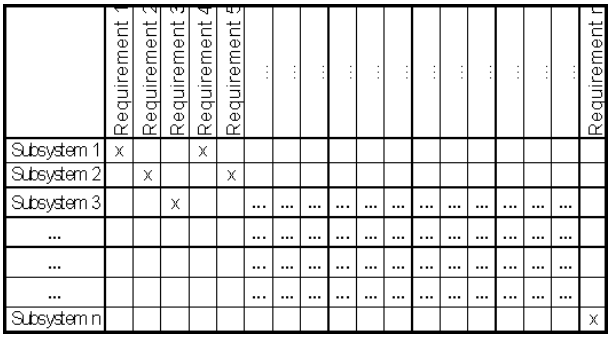

Figure 3. Requirements traceability matrix by the U.S. Department of Transportation

confidence on a mostly machine to machine (M2M) communication system. Then these flows should be analysed and as a result, QoS requirements for the flows should be identified. As these are identified, the aggregated results should be evaluated against the possibilities of the underlying infrastructure [14].

The analysis will result in a range, stating the minimum QoS requirement (with a certain confidence) and the preferred QoS requirement. If the expected QoS after taking communication flows into account is inside the operational envelope, the system can deliver with the defined confidentiality level.

The operational envelope will be larger than zero (not just forming a baseline composed from the single QoS requirements) because of the stochastic nature of best-effort forwarding and large networks. Also, an analogy with the different SIL can be drawn with comparing them to the confidentiality level of keeping the Service Level Agreement (SLA) [11].

The approach taken for safety can be a solution for other properties of the industrial communication system, e.g., QoS for transport or security [15].

V. REQUIREMENTS TRACKING

One of the key aspects missing in engineering work today is the follow-up of requirements stated against the communication infrastructure.

On the LAN level, the lack of tracking only results in minor problems, as network resources are typically not problematic. Even not on the redundancy requirements, since most of the critical network will have approximately the same reliability requirements. As an example, a current IEC 61850 substation will have tens of devices connected to the network.

The local communication of IEC 61850 is composed from horizontal and vertical flows, where horizontal flows tend to use more resources, as Sampled Values (SV) traffic is sent this way. SV is the continuous stream of sampled input or output values, which is sent to a controller for processing. The stream can fill 10s of Mbps. On a network with a gigabit backhaul, conveying traffic in several 100 Mbps range is not problematic. Redundancy is typically covered by either a secondary network or redundant links.

Already in the horizontal-vertical split of flows, different requirements are valid against the network infrastructure. As the automation task gets more far away from the fieldbus level (direct contact with the physical world), so are the deadlines for communication and processing more relaxed.

Requirements tracking is becoming key as the automation system passes the LAN boundary. Costs associated to network communication are becoming more expensive and obeying QoS parameters increasingly problematic.

Several well-known approaches can help the aggregation and validation of the QoS parameters during the life cycle of the project. One of these solutions is the requirements traceability matrix.

In such a matrix, requirements posed by different automation tasks towards the infrastructure can be gathered (figure 3). To allow both aggregation of parameters and identification of the source of a specific requirement.

Source identification is key for long-life installations, where extensions and updates can be expected during the lifetime of the system.

Evaluation if a requirement is still valid in different parts or domains of the system has also a key importance in efficient deployments. It is important to set up an iterative process for QoS parameter evaluation. Here, a possible solution could be to follow the V-model used in, amongst others, software development and safety development. Figure 4 shows the iterative development process. The QoS requirements should be evaluated at each step and their fulfilment validated after each step. With using such a model, the bearer infrastructure would be more integrated into the development process. Integration can lead to more optimized QoS requirements. Current practice results more in a worst-case requirement list.

For Wide Area Network (WAN) situations, tracking requirement validity has key importance. The validity area of the respective QoS parameters has to be limited to cover only the necessary parts. As part of an iterative process, when the communication scope is getting wider (e.g., the data is being passed upward in a hierarchical network architecture), validity of the QoS parameters has to be checked. An example is that if there is a strict time synchronization requirement with IEEE 1588, but there is no such requirement for the WAN section, nor is a loop covering two endpoints in different networks, then the 1588 requirement should not be taken over to the SLA definition of the WAN interface.

 ISBN: 978-1-61208-473-2

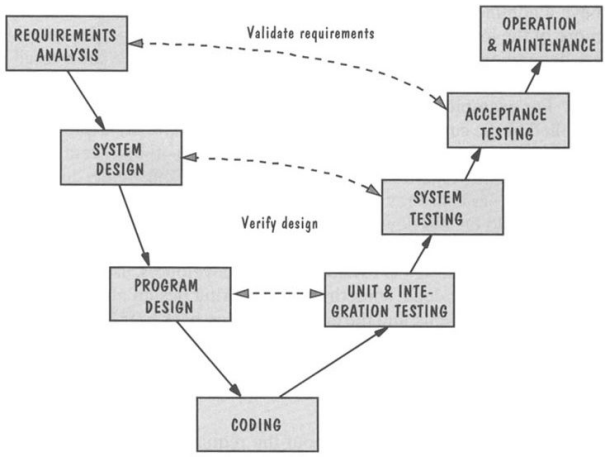

Figure 4. V-model [10]

VI. Control loop parameters

Requirements definition for the communication network is one of the actual challenges in automation. The challenge in this task is, that the automation flows are defined using different metrics than the communication links. An example IEC 61850 control loop would be defined as: having a sampling rate of 80 samples per cycle (4800 Hz for 60 Hz networks), with sampling 16 inputs, 16 bit per sample. Event-based traffic is negligible compared to the periodic traffic. If there is a requirement for synchronous operation, time precision (quality) can also be a QoS metric. Redundancy requirements can lead to topologies, which are unusual in a normal network infrastructure: first, the use of Rapid Spanning Tree Protocol (RSTP) to disable redundant links, second the general use of loops (rings) in the network to ensure that all nodes are dual-homed. With dual-homing, the network can survive the loss of one communication link without degradation in the service level. From the network viewpoint, this control loop will introduce a traffic flow, with a net ingress payload stream of approx. 98Mbps. The sampling will generate 2560 bytes of traffic each second, which can be carried by at least two Ethernet frames, thus the system can expect at least approx. 10000 frames per second. The traffic will be forwarded on a horizontal path to the controller. On the ingress port to the backbone, it will enter with approx. 110 Mbps (header+payload). The traffic flow will be consumed at the egress port to the controller.

Due to the stochastic nature of Ethernet, there will be jitter between the frames transmitted over the network. The maximum jitter is defined by the maximum delay variation tolerance of the control loop (typically, every second frame must arrive in good time). This requirement can then be calculated with either the length of the typical frame of the flow or with a maximum length frame. In both cases, the allowed jitter will be considerably longer than the expected disturbances on the LAN. Precision requirement on the time synchronization implies two choices: the choice of protocol and time source. The choice of protocol is generally IEEE 1588v2, which allows high precision time synchronization and GPS as a time source. The choice of GPS is actually an input to the risk analysis of the whole project, as then the time reference will depend on a network controlled by a third party.

VII. Conclusion and future work

With communicating automation systems covering large geographical areas and also expanding in logical complexity, current, non-scalable solutions for performance definition and evaluation are getting outdated.

Introduction of the structured approach used in safety development can both enhance the quality of deployments and also allow easier communication between the parties. One of the main advantages of the safety-approach is, that it is widely known and accepted in the industry, so the two worlds of operations and IT could work better together.

Future work will focus on how the transformation of QoS parameters can be formalized and which modifications are needed in the safety processes to suit the QoS process and possibly the security process in an effective manner. Also protocol development or adaptation for resource reservation for automation applications in both LAN and WAN environments is an important field of study, including the use of Software Defined Networking (SDN) in automation [1], [12].

As an outlook, future hot spots of research could be automatic parameter tracking through the design process and real time monitoring of deployments also during their operation. Automation and smart grids are an important field of 5G efforts and it is expected to utilize the existing telecommunication protocols with applying industry-specific profiles. Developing these profiles which will not only define the infrastructure requirements, but also interfaces towards other systems is one of the interesting areas for the success of 5G.

References

[1] Gy. Kálmán, "Applicability of Software Defined Networking in Industrial Ethernet", in Proceedings of IEEE Telfor 2015, pp. 340-343, Belgrade, Serbia

[2] Gy. Kálmán, D. Orfanus, and R. Hussain, "An Overview of Switching Solutions for Wired Industrial Ethernet", The Thirteenth International Conference on Networks ICN 2014, pp. 131-136, Nice

[3] Cisco, "End-to-End QoS Network Design: Quality of Service for Rich-Media & Cloud Networks", Cisco Press, 2013.

[4] M. Rensburg, D. Dolezilek, and J. Dearien, "Case Study: Using IEC 61850 Network Engineering Guideline Test Procedures to Diagnose and Analyze Ethernet Network Installations", in proceedings of PAC World Africa 2015, November 12-13., Johannesburg, South Africa

[5] L. Sheng, "QoS Design and Its Implementation for Intelligent Industrial Ethernet", International Journal of Materials, Mechanics and Manufacturing, Vol. 4, No. 1, 2016., pp. 40-45.

[6] V. Skendzic, I. Ender, and G. Zweigle, "IEC 61850-9-2 Process Bus and Its Impact on Power System Protection and Control Reliability", in proceedings of the 9th Annual Western Power Delivery Automation Conference, April 3-5, 2007, Spokane, USA

[7] J. Bilbao, C. Cruces, and I. Armendariz, "Methodology for the QoS Characterization in High Constraints Industrial Networks", Open Journal of Communications and Software, Volume 1, Number 1, 2014., pp. 30-41

[8] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications", IEEE Communication Surveys and Tutorials, Vol. 17, No. 4, 2015., pp. 2347-2376

[9] N. Barkakati and G. C. Wilshusen, "Deficient ICT Controls Jeopardize Systems Supporting the Electric Grid: A Case Study", Securing Electricity Supply in the Cyber Age, Springer, 2009, pp. 129-142

[10] G. Blank, "Object-oriented Software Engineering", http://www.cse.lehigh.edu/~glennb/oose/figs/pfleeger/Vmodel.jpg, Accessed 18.03.2016.

[11] P. Blanco, G. A. Lewis, and P. Merson, "Service Level Agreements in Service-Oriented Architecture Environments", Technical Note, Software Engineering Institute, CMU/SEI-2008-TN-021

 ISBN: 978-1-61208-473-2

[12] D. Cronberger, "The software-defined Industrial Network", The Industrial Ethernet Book, Issue 84, 2014., pp. 8-13

[13] C. Alcaraz, G. Fernandez, and F. Carvajal, "Security Aspects of SCADA and DCS Environments", In Critical Infrastructure Protection: Information Infrastructure Models, Analysis, and Defense, LNCS 7130., Springer, 2012., pp. 120-149

[14] Open Networking Foundation, "Software-Defined Networking: The New Norm for Networks white paper", https://www.opennetworking.org/images/stories/downloads/sdn-resources/white-papers/wp-sdn-newnorm.pdf, Accessed 28.01.2016.

[15] R.C. Parks and E. Rogers, "Best practices in automation security", Security & Privacy, IEEE (Volume:6 , Issue: 6), 2009., pp 37-43

 ISBN: 978-1-61208-473-2