Possibilities of 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*—Machine to machine communication offers both an opportunity and poses a challenge for communication networks. In this paper, after an introduction to industrial control networks, an overview of typical QoS metrics is given and their relation to automation metrics is analyzed, current industrial networking technologies and their QoS possibilities are presented. Conversion or mapping of QoS metrics between communication and control systems is evaluated. As a possible direction, use of formal methods and procedures known from industrial safety systems are recommended.**

***Keywords*—critical infrastructure; QoS; metrics; automation; control networks; PROFINET; EtherCAT; Ethernet.**

I. Introduction

This paper is an extension of [1], "Quality of Service Parameter Tracking and Transformation in Industrial Applications," published at IARIA AICT 2016.

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 and currently mostly done by overprovisioning.

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 Wide Area Network (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its introduction in industrial automation, Ethernet's determinism has been a returning concern, mainly because of both outdated information (behavior of 10-Base2) and bus-like topologies [2] with long chains of switches.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 introduction to industrial control networks, the third section gives an overview of QoS. Section four gives an overview on QoS features of current networking technologies used, section five provides an overview of Distributed Control System (DCS) structures. Section VI presents how requirements may be specified in a structured way, Section VII explains the importance of requirements tracking. Section VIII presents parameters of a control loop and how QoS parameters can be converted between the industrial and communication metrics. Section IX draws the conclusion and provides an outlook on future work.

II. Industrial Control Networks

In a historical perspective, control of manufacturing and process plants was done mechanically: the transmission of signals were done by some physical mean, like pneumatics, hydraulics or manual force. These mechanical structures were replaced by electric solutions in parts of the systems. Electric control had been successful and mechanical control systems were replaced by electronics, mostly employing hardwired circuits [3].

The hardwired circuits were both prone to errors and consumed large amounts of space and money. A similar evolution has happened as with the telecommunication lines: a digital, interleaved solution was needed [4].

With the introduction of microelectronics and digital bus systems, it became possible to exchange the long and expensive dedicated wiring with bus systems, commonly called as fieldbus. We can date the birth of QoS in industrial environments to this step of the evolution: in case of direct wiring, there was no question on access to the transmission media. Delay or jitter were not applicable, the signal propagated with nearly the speed of light and had dedicated media (bandwidth) to the controller.

The use of digital communication solutions has spread on all levels of automation and resulted in the current state,

where Ethernet is used in both industrial and connected corporate networks. The main difference remaining in these, similar looking networks is the requirements posed by the communication parties. An industrial network is connected to the physical world and events on communicated on this network have a physical dimension. This connection results in different priorities for QoS [5], [6].

Practically all new industrial deployments will use a communication technology, in the vast majority of cases, Ethernet. Each industry has its own set of different, but similar requirements. On the timescale basis, we typically distinguish between: bus bar protection/motion control, manufacturing and process control.

Bus bar protection and motion control require the most stringent achieved QoS levels: the information must reach its destination with great precision and low latency.

Manufacturing has a typical requirement time scale of tens of milliseconds. In the view of corporate networks these requirements are still hard to keep, but in an industrial environment, on this level standard equipment is used. The main support property for the QoS parameter calculation is that the typical source-destination of a control loop with hard requirements is in most of the cases close to each other. As a result, the network as a whole does not have to adhere for the requirements, but paths on the network might be involved.

Process control is the most relaxed of the three and typically poses no strict requirements. In case of a process control loop, the typical timing possibilities are in the second range rather than milliseconds, so delays on current Ethernet networks (which are in the microsecond range) are mostly not noticeable in such installations.

Due to that industrial networks do have a connection to the physical world, failure of such a system has the potential for much more severe impact than that of a corporate system. Effects of failure can be, e.g., damage to equipment, production loss, environmental damage or the worst: injury or loss of life. The connection to physical processes also introduces the need for real-time behavior. The expression real-time is often used as a synonym to *fast*, but the more precise definition is that the network has to give an answer within a specified (real/physical) time slot and if it fails, the data is or nearly is worthless [7], [8].

The strict requirements on delay or jitter are in a reverse order compared to closeness to the physical process: fieldbus tend to have strong requirements, especially if functions, like motion control is executed. The communication after the controller level is less critical, as here mainly the communication parties are the historians and the Human Machine Interface (HMI) units. As in these upper levels, human operators are the typical recipient of information, the expected delays due to traffic and non-determinism are magnitudes smaller than the reaction time of the employees.

Determinism is a property, which, beside telecommunication networks, is not typically used as a measure of QoS. In an industrial environment, it can be one of the questions which need to be answered. Here again, in parallel with jitter

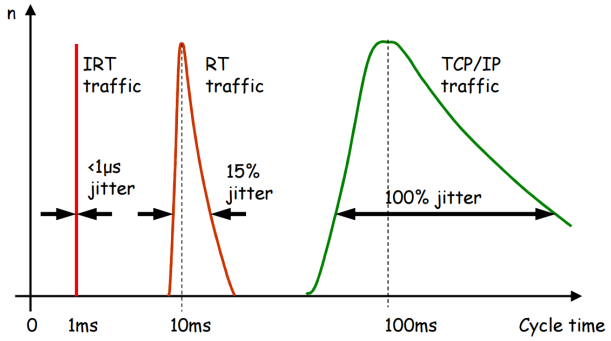

Fig. 1. Delay and jitter ranges in PROFINET [9]

and delay, the typical determinism requirements are more strict closer to the physical process. To call a network *deterministic*, it must be possible to give an upper bound for delivering a chunk of data (Fig. 1). In dedicated wire solutions or slotted technologies, like serial lines or bus like Profibus, the upper bound could be calculated from the network setup. In particular, early Ethernet is not a good solution to provide upper bounds. Half-duplex implementations suffered the well-known loss of throughput in high-traffic situations, because of the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) medium access protocol. This history still limits the acceptance of Ethernet in automation environments, however, in current networks with full-duplex switched topologies, CSMA/CD is not needed as there are no collisions. Still, if not media access, traffic situations can lead to queueing. In a typical situation, where the two end-hops are running on 100Mbps, while the backbone network is running on 1Gbps, the accumulated queueing delay after several core hops will be still magnitude lower than the propagation delay of the 100Mbps (non-congested) hops.

Jitter is the variance of delay. In a network, where a control loop is run, it is typically a requirement to have low jitter, thus the periods of sampling will be uniformly distributed over time. Here the network has to provide a jitter below the upper bound, which is acceptable for the control loop. Time synchronization is one of the features, which provide the connection point towards the physical world. Temporal consistency and the ability to record the events in correct order are important both for supervisory tasks but also for troubleshooting.

An important measure of the network equipment is the throughput. Industrial applications here also have different emphasis areas: in office or telecom, a typical frame has a long payload compared to the header (with the one famous exception being Asynchronous Transfer Mode (ATM)). Data frames transmitted on industrial networks are typically short, especially close to the physical process, where the other QoS requirements are high. This property makes fulfilling the QoS parameters more problematic, as it is easier to utilize the full bandwidth for network equipment if the frames are long, thus fewer forwarding decisions need to be taken and the overhead

is also smaller.

Composition of traffic in control networks, especially on the field level, differs considerably from the typical office environment. As the communicating parties are near exclusively machines and the operation of control systems is very often periodic, it is very typical to have a nearly static traffic picture with mostly very stable packet streams. Aperiodic events like state changes or alarm conditions compose a growing part of the traffic starting from being nearly negligible on the field level to being a considerable part on the client-server or plant level. Periodic traffic on the field level was expected to be be problematic, if strict real time requirements are extended with high data speed, like in the case of IEC 61850-sampled values. Experience shows that in most cases, the best effort forwarding works without causing problems, as in typical cases, the offered bandwidth is well above the requirements of the control loop. For special requirements, industrial Ethernet variants, like EtherCAT and Profinet IRT were developed. These offer intrinsic QoS witch scheduling functions supported by the special hardware implementations.

Compared to office networks, a distinct physical feature of the industrial deployments is the ruggedness of devices. A typical device has to withstand vibration, shock, has to accept wider operational and storage temperature ranges and might even need to withstand moisture. From the operational viewpoint, however, these properties have little impact. On the performance side, current chipsets are providing adequate resources even with only using passive cooling and if needed, special connectors are used to avoid physical damage on the connecting wires. The designed lifetime of the devices is much higher than in the office environment: a representative life expectation is around 20 years.

A. Types of Information

On a typical industrial network, there are a handful types of data: control information, related to keeping the process under control, the sampled data, which connects the control system with the physical process, diagnostics and management and auxiliary functions like technical safety.

Control information and process data is the communication between instruments and controller and are the main connection of the control system to the physical world. In most cases, the majority of traffic is part of this category and often the only considered part of the traffic mix. The hard QoS parameters are typically property of traffic in this category, as this layer of the network is the closest of the actual process. Example traffic on this level is sampled values from some instrument, like an Intelligent Electronic Device (IED) or temperature/pressure sensor. Event-controlled traffic is also present, for example, valve status changes.

Diagnostics and management are important auxiliaries: diagnostics is an integral part of creating a reliable control system. Errors of various causes can happen in the system and an effective diagnostics can predict or identify failed components. The detail of diagnostics and the capabilities provided by this subsystem depend on the reliability and redundancy requirements. Diagnostics is a type of data, which is collected by the system, but by default, for expected values, there is no reaction. High coverage diagnostics is also an enabler, for example, Safety Instrumented Systems (SIS). Management of the system is necessary above the very basic level. The overview of current system status is important information for both operators and engineers.

Safety is the most important auxiliary function. A dimension of running a SIS is to have adequate diagnostics. Categories of safety levels are defined by IEC 61508. Safety information is carried in parallel to control information. The different Safety Integrity Levels (SIL) have different implications on redundancy of the safety system and the coverage of diagnostics.

III. Quality of Service

QoS is the measure of transmission quality and service availability of a network [11], 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) 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, the introduced metrics 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, 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 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 [11]. 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 [12]. 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 [13]. 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.

IV. QoS features of Industrial Ethernet

Industrial Ethernet variants are mostly building their QoS features on the existing traffic prioritization services of Ethernet. While not directly a QoS feature, the most important step towards the usability of Ethernet in industrial applications was the introduction of full duplex networks and Ethernet switches.

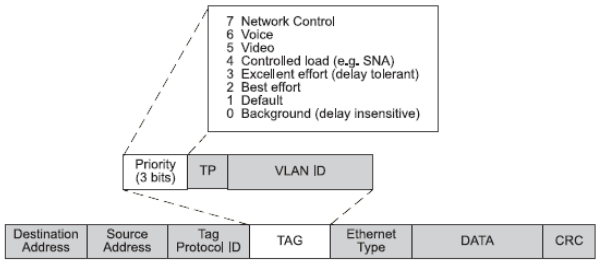

Fig. 2. Ethernet priority field [28]

- *IEEE 802.1p* As the most important traffic management feature, inbuilt in Ethernet. .1p offers the possibility to assign. Priorities have been defined in the project and the switches typically implement a solution with multiple queues and round-robin scheduling with ageing. There are 8 different traffic classes (Fig. 2), from background and best effort up to network control. Unfortunately, in industrial applications, 7 (the highest) is often used. Although this will give these frames priority over all other frames, but since nearly all automation traffic is in the same traffic class, delays might occur. Also, it is not practical that all automation traffic is getting the same priority as the network control, as signaling traffic for the network infrastructure might have much larger impact on the system as the loss of a couple of automation frames.
- *IEEE 802.1 Time Sensitive Networks* TSN is a group of standards, which provides a real-time Ethernet implementation, where deterministic transport of data is possible. In addition to implementing call admission control to guarantee that the communication requirements are fulfilled

through the path, it also introduces a global reference to physical time. The standard has emerged from the IEEE 802.1AVB, Audio-Video Bridging (Fig. 3) proposal and widened the possible field of use with automation and especially the use of Ethernet in the Internet of Things (IoT).

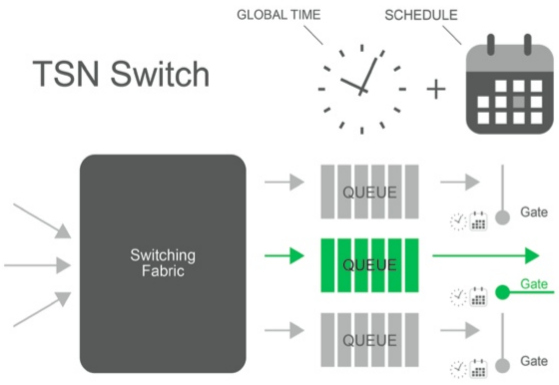

Fig. 3. Connection of physical time and traffic in TSN [14]

A. PROFINET

PROFINET was developed and is the preferred Industrial Ethernet variant of Siemens. The main QoS feature is that by default, PROFINET provides three different traffic classes (Fig. 4: the first one provides a framing service for legacy PROFIBUS and also carries non-critical data with cycle times of around or above 100ms. The traffic can be run on normal TCP/IP. The second traffic class is, what the protocol calls, Real Time (RT), which is supporting IO applications with a cycle time of around 1ms to 100ms [15]. The third traffic

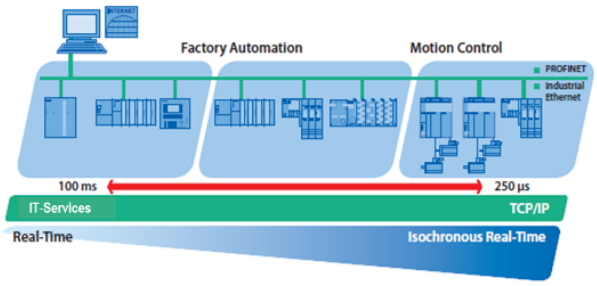

Fig. 4. Traffic classes of PROFINET [16]

class is Isochronous Real Time (IRT). With using special hardware, IRT provides a communication solution for low-latency applications.

B. EtherCAT

EtherCAT was developed to provide a deterministic network solution for devices on a local ring. It is a technology with an intrinsic QoS solution, as the processing of the data on the ring is done on the fly, as the frame travels through the

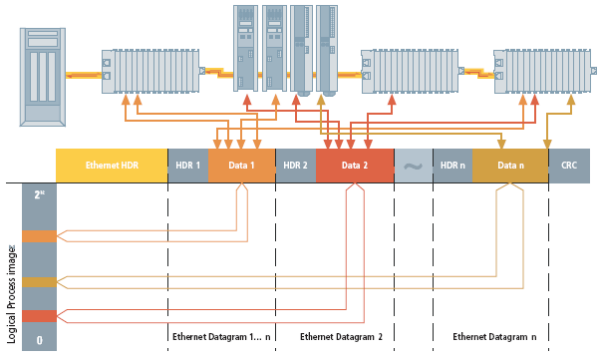

Fig. 5. EtherCAT master-slave example [17]

Application Specific Integrated Circuits (ASICs) of the slave machines (Fig. 5). A representative example of the cycle times is that the master is beginning to receive the frame at its input, before the sending is finished on the output. The simple topology and the call admission control are both enablers to allow the short cycle times. The most important property is the possibility to calculate the cycle time with high precision. An additional flexible feature is that in-between the periodic frames, it is compatible with normal traffic and it is possible to send out these also on the automation loop.

C. SERCOS III

SERCOS III combines a solution resembling EtherCAT for on-the-fly processing of the frames and the possibility to use L3 communication for non-critical information exchange. The slave processing is done as the frame traverses the Ethernet interface, but SERCOS splits the in- and output to different frames as compared to the single frame sent in the case of EtherCAT.

D. IEC 61850

IEC 61850 is implemented directly over L2. This industrial Ethernet type is used mainly in electric substation automation. The traffic itself is typically composed of multicast and unicast frames, the delay and jitter depends on the QoS functions of standard Ethernet. The lack of L3 in the communication stack enables potentially faster communication, reaching the level of the best case Ethernet.

E. Ethernet/Industrial Protocol

Ethernet/IP is a protocol developed by Open DeviceNet Vendors Association (ODVA), led by Rockwell Automation. The Ethernet/IP protocol is implemented on the application layer and provides an encapsulation service for Common Industrial Protocol (CIP) data (Fig. 6). Implementing a protocol on the application layer has both its positive and negative implications. From the QoS viewpoint, the use of application layer allows the utilization of the features offered by lower layers: prioritization on L2 (Ethernet) and IntServ or DiffServ features (if implemented) on L3. The shortcoming of the application layer is that strict, low latency control loops are

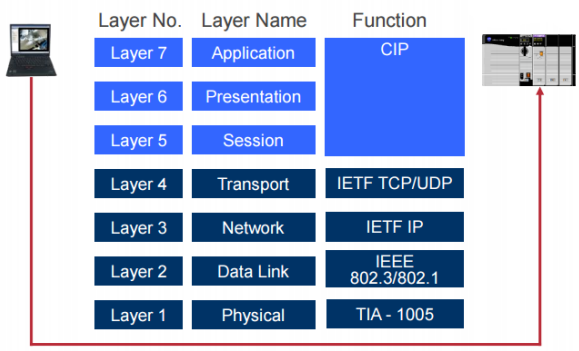

Fig. 6. Protocol stack of Ethernet/IP [18]

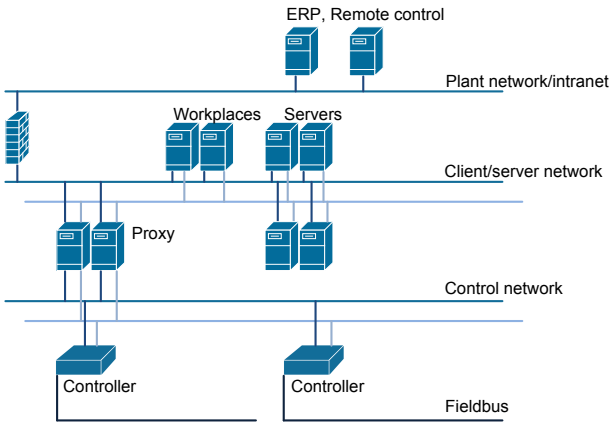

Fig. 7. Traditional DCS network architecture

in practice not feasible. This is a result of primarily the delay added by the travel through the protocol stack. In an optimal situation, the timing properties can be very close to one of standard Ethernet.

F. Foundation Fieldbus High Speed Ethernet

Foundation Fieldbus HSE is implemented as an application layer protocol and has similar properties and limitations as Ethernet/IP: in the best case, the delay and jitter can be close to the L2-based implementations, but the additional software layers will introduce some uncertainty.

V. DCS ARCHITECTURE

Control systems are traditionally built using a three network levels (Fig. 7.). 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.

The most important component from the control systems viewpoint is the Programmable Logic Controller (PLC) or controller. These are specialized industrial computers implemented as solid-state electronic devices, they replaced hard-wired relay-based circuits. Current devices, beside offering the traditional groups of digital and analogue circuit interfaces towards instrumentation, also offer a wide range of other services, like logging, status report over SNMP or security features like a firewall.

TABLE I
COMPARISON OF DCS AND SCADA PROPERTIES

SCADA	*DCS*
Small physical distances	Large distances
Independent system	Interconnects/monitors several systems
Full local network control	Typically uses third parties
Data driven	Event driven

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. SCADA operations can be very much represented as a software-only entity. Since SCADA only has the task to supervise the selected systems, the communication is less critical and mostly event-driven. Remote Terminal Units (RTUs) are used to feed the data to the central entity, the Master Terminal Unit (MTU). Remote monitoring is not a replacement for functionality on site: operations shall be possible to maintain also in island mode. SCADA is not expected to lower the site's reliability or security [19], [20]. The typical long physical extent requires the use of communication infrastructure delivered of third parties. The cost pressure on the communication costs typically led also in these operations to move from leased lines or other high QoS-high cost solutions to packet switched solutions.

Developments in the smart grid and IoT extend the possibilities for remote operations is by taking current communication solutions in use. The extension of the features also requires a well-defined network infrastructure [21]. An interesting aspect from the interconnecting task of SCADA systems is that the typical hardware/software platform used for the SCADA system will be obsolete in a couple of years, and will be needed to be upgraded, the systems the SCADA is reporting abbot will still have the same, relatively old DSC as clients.

A. QoS in automation

Traffic flows in automation typically are M2M. This property and the systems connectivity to the physical world require both different tolerances for disturbances and potentially different metrics [22].

An automation system somewhere in the process is connected to the physical world even if some of the functions can be virtualized [23]. 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., 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.

VI. 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 [24].

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 [25].

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.

In the Internet of Things (IoT) scenario, where the automation networks are extended behind the LAN [26], 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 analyzed 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, so 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 exist 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 analyzed. 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

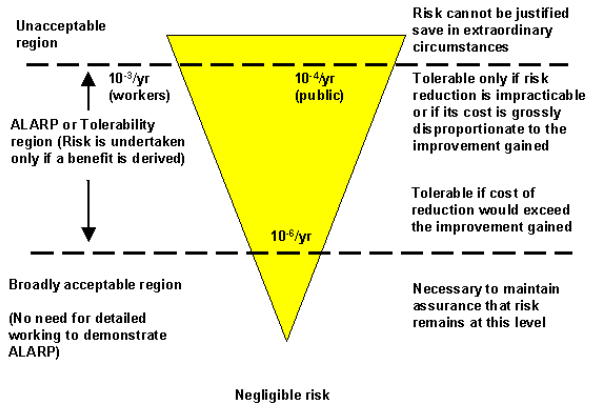

Fig. 8. The Health and Safety Executive's Risk criteria

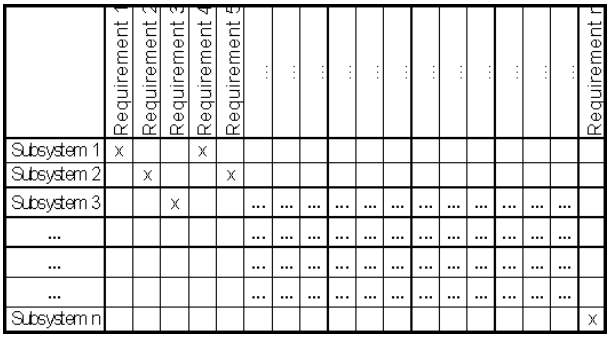

Fig. 9. Requirements traceability matrix by the U.S. Department of Transportation

until the acceptable risk level is achieved. A possible method for risk classification is shown on Fig. 8. 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 confidence on a mostly M2M communication system. Then these flows should be analyzed 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 [27].

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) [29].

The approach taken for safety can be a solution for other properties of the industrial communication system, e.g., QoS for transport or security [30].

VII. 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 (Fig. 9.). 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. Fig. 10. 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.

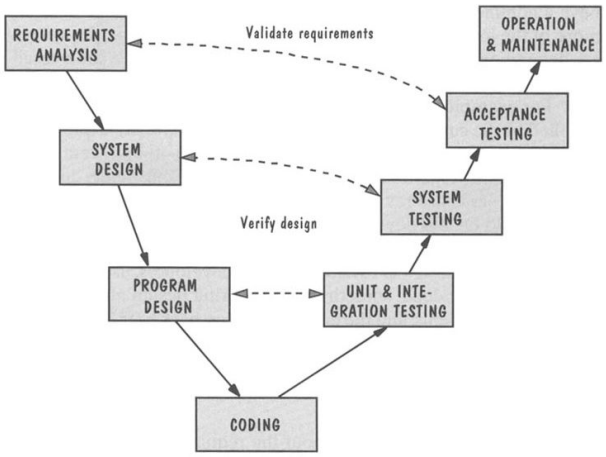

Fig. 10. V-model [31]

For 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.

VIII. 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.

IX. 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. Deterministic mapping of control-related parameters to QoS parameters of the used networking technologies supported with requirements tracking can be a way to go.

To show a similar process in the engineering of automation systems, examples from safety integrated systems are shown. Introduction of the structured approach used in safety development can both enhance the quality of deployments and also allow easier communication between the parties. The main gain with using a process built on the safety development is, that the safety process (like the V-model) is already known and accepted. Networking and QoS is, as safety, not a single delivery, but a process and follows the life-cycle of the product.

Future work will focus on, how QoS requirements can be formalized in a technologically neutral way and mapped into actual solutions. 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 SDN in automation [10], [32].

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, including protocols like Resource reSerVation Protocol (RSVP). Developing these profiles which will not only define the infrastructure requirements, but also interfaces towards other systems.

REFERENCES

[1] Gy. Kálmán, "Quality of Service Parameter Tracking and Transformation in Industrial Applications," in Proceedings of IARIA AICT 2016, St. Julians, Malta

[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] B. Galloway, and G. Hancke, "Introduction to Industrial Control Networks," IEEE Communications Surveys and Tutorials, Volume 15, Issue 2, Pages: 860-880, 2013 Q2

[4] Alcate-Lucent, "Transformation of mission-critical communication networks," Alcatel-Lucent White Paper, 2015

[5] H. Bernhard, and J. Mottok, "Real-time behavior of Ethernet on the example of PROFINET," https://www.hs-regensburg.de/fileadmin/media/fakultaeten/ei/forschung_projekte/MAPR_Ver%C3%B6ffentlichungen/ARC_Heitzer.pdf, accessed: 08.09.2016

[6] Y. Jeon, "QoS Requirements for the Smart Grid Communications System," IJCSNS International Journal of Computer Science and Network Security, Volume 11, Issue 3, 2011

[7] I. Dominguez-Jaimes, L. Wisniewski, and H. Trsek, "Identification of Traffic Flows in Ethernet-based Industrial Fieldbuses," IEEE Emerging Technolgoies and Factory Automation (ETFA), 2010

[8] M. Yaghmaee, Z. Yousefi, M. Zabhi, and S. Alishahi, "Quality of Service Guarantee in Smart Grid Infrastructure Communication Using Traffic Classification," 22nd International Conference on Electricity Distribution, Stockholm, 2013

[9] A. Verwer, "Overview and Applications of PROFINET," PROFIBUS and PROFINET International, http://www.profibus.com/uploads/media/profinet_overview.pdf.pdf, accessed: 08.09.2016

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

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

[12] 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

[13] 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.

[14] TTTech, "Deterministic Ethernet and TSN: automotive and industrial IoT," Industrial Ethernet Book, Issue 89/8, 2016

[15] P. Neumann, and A. Pöschmann, "Ethernet-based Real-Time Communications with PROFINET IO," ACMOS'05 Proceedings of the 7th WSEAS international conference on Automatic control, modeling and simulation, Pages 54-61, 2005

[16] Siemens, "Profinet Answers for Industry," https://w3.siemens.com/mcms/water-industry/en/Documents/PROFINET.pdf, 2010, Accessed 28.01.2016

[17] EtherCAT Technology Group, "Moving up to Industrial Ethernet," Industrial Ethernet Book, Issue 45/35, 2016

[18] Rockwell Automation, "Fundamentals of Ethernet/IP Network Technology," Rockwell Automation presentation, TechED 2015

[19] Jari Ahokas, "Secure and Reliable Communications Solution for SCADA and PPDR Use," Master's Thesis, Laurea University of Applied Sciences, 2013

[20] C. Hauser, D. Bakken, I. Dionysiou, K. Gjermundřd, V. Irava, and A. Bose, "Security, Trust and QoS in next-generation control and communication for large power systems," International Journal of Critical Infrastructures, Volume 4, Issue 1-2, 2008

[21] 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

[22] 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

[23] J. Beran, and F. Zezulka, "Evaluation of Real-Time Behavior in Virtual Automation Networks," Proceedings of the 17th World Congress of The International Federation of Automatic Control, Seoul, Korea, 2008

[24] 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

[25] 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

[26] 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

[27] 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.

[28] Industrial Ethernet Book, "Quality of Service for high priority networks," http://www.iebmedia.com/index.php?id=5594&parentid=63&themeid=255&showdetail=true, accessed: 08.09.2016

[29] 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

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

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

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