Towards New Hybrid Networks for Industrial Automation

Thilo Sauter
Austrian Academy of Sciences
Viktor Kaplan Str. 2
A-2700 Wiener Neustadt
thilo.sauter@oeaw.ac.at

Jürgen Jasperneite
inIT - Institut Industrial IT
Liebigstrasse 87
D-32657 Lemgo
juergen.jasperneite@hs-owl.de

Lucia Lo Bello
University of Catania
Viale Andrea Doria, 6
95125 Catania
lucia.lobello@diit.unict.it

Abstract

In factory communication systems, wireless networks are likely to complement wired automation networks, rather than replacing them. There is therefore an increasing interest in the way wireless and wired systems will interact and form a hybrid network. The hybrid network architecture proposed in this paper specifically addresses WLAN infrastructures in industrial environments. The paper discusses design challenges and describes the proposed system architecture and its main components.

1. Introduction

Communication systems for factory automation have seen a remarkable evolution during the last 30 years from dedicated, application-specific fieldbus systems to solutions based on standard Ethernet as well as Ethernet-inspired real-time networks. These modern concepts together with the trend of using Internet technologies in automation (both on communication and application level) allows for a simpler integration of field-level networks and the office domain inside a company [1]. In recent years, as a next step in the evolution, wireless networks – also borrowed from the general IT world – have received increasing attention also for factory automation, and wireless field-level networks have become a wide and challenging research topic. The benefits of wireless technologies on the factory floor are obvious:

- Wiring can be reduced, which is advantageous in areas where nodes are widely scattered or where wires can be installed only with great difficulties due to hostile environments.
- Nodes can be mobile, which is hardly possible if a node has to drag wires. Mobility in this respect does not necessarily mean that nodes may go wherever they want; they can as well stay within a strictly confined area.

- Installations can become flexible, which supports the trend towards reconfigurable production systems that need to be adaptable to changing needs of the market [2,3]. Parallel to flexible automation concepts on application level, wireless networks provide utmost flexibility on the communication side.

Nonetheless, industry is somewhat reluctant to apply purely wireless networks on the factory floor for dependability reasons. Contrary to Industrial Ethernet, which is widely seen as a long-term replacement of conventional fieldbus systems, it is currently not to be expected that wireless networks will completely supersede wired automation networks. Rather, they will complement them. The usual case considered by industry is therefore that wireless or wired systems will interact and form a hybrid network.

1.1. Hybrid automation networks

The typical structure of a future automation system exhibits a two-level hierarchy with a wireless lower level and a wired field-level backbone network likely based on Ethernet (Fig. 1). The wireless segments are organized in peer-to-peer fashion, but each of them will have a central access point connected to the backbone network, which in turn is connected to the higher levels of the company hierarchy.

Figure 1. Two-level hierarchy of a hybrid wired/wireless factory automation network.

The work presented in this paper is funded in part by the European Commission in the 7th Framework Program under contract ICT-224350, FlexWare – Flexible Wireless Automation in Real-Time environments.
Typically, the wireless clusters will form independent automation islands. The higher communication levels are based on Internet technologies, so it is reasonable to use compatible technologies (like IP-based networks, web standards for data handling and engineering, and communication technologies proven in the office world) also in the lower layers.

From a functional viewpoint, there are essentially two ways how wireless and wired segments can be integrated. The first and simpler option is that wireless segments are independent autonomous islands without the need to exchange data – at least not with real-time requirements (Fig. 1a). In this case, the access points interconnecting the wired and wireless segments work as bridges transparent for higher protocol layers or simply as gateways. The wireless cluster may then use a completely different, possibly much simpler protocol than the wired backbone, and delays introduced by the protocol and data translation in the gateways are not relevant.

To speed up the data exchange and better separate the data exchange mechanisms on either side of the gateway, the access point application may also feature a caching strategy. This scenario is useful for subordinate sensor-actuator networks [4,5,6] that are spatially concentrated and functionally independent.

A by far more complex scenario can be found in large installations where several access points are needed to cover the entire area and where the automation application cannot be divided into independent islands (Fig. 1b). Here, multiple wireless segments need seamless integration. The situation becomes even more complicated if mobile nodes need to be considered that may roam between access points. In such a case, a relatively simple gateway approach is no longer feasible, and the access points need proper synchronization to retain consistency throughout the network. The wireless parts together with the backbone and possibly also conventional wired field devices form a uniform domain with stringent timing requirements. This may also include small remote segments which are connected via a wireless link (like in the left corner of Fig. 1).

1.2. Related work

There have been attempts to achieve integration of wired and wireless automation networks before. One of the earliest was the RFieldbus project, which aimed at extending Profibus networks with wireless links [7,8]. To gain optimal performance, a dedicated physical layer implementation had to be developed, and the MAC protocol was identical with that of Profibus DP, hence there was no compatibility with modern wireless technologies.

A more recent project is VAN, developing concepts for industrial communication via heterogeneous networks, including also wireless state-of-the-art networks [9,10]. The VAN project is mainly focused on including wide area networks into future automation systems. Essentially, VAN uses Web Services to establish end-to-end VPN-based tunnels between communication endpoints (automation devices) for the exchange of cyclic process data. The devices themselves use these links and are unaware of the underlying network technologies. The core of this approach is the IEC standard 61158 - Type 10 (PROFINET). A third project is RAVE, which focuses on the improvement of available wireless technologies for mobile and reliable real-time services in automation [11]. The project develops a convergence layer to enable the generic application of different wireless networks, with Profinet IO functionality on top of it.

Apart from such R&D projects, there have been many research contributions addressing the integration of wireless and wired networks in industry [12,13,14], possible architectures for such hybrid approaches [15,16], as well as case studies using various field-level networks and wireless communication standards [17].

The hybrid network architecture proposed in this paper is being developed within the scope of the European research project FlexWare [18,19]. This project specifically addresses WLAN infrastructures in industrial environments, but the concept is more generic. The infrastructure relies on a real-time backbone network which will be used by different applications spread over the entire factory floor. An essential feature is that the network infrastructure can transparently switch between access points. This is evident if nodes are mobile, but roaming may also be induced by changing conditions especially in harsh industrial environments, where the quality of communication paths may change with time.

To implement sufficient flexibility, the access points have to interact in order to organize this roaming between clusters. Roaming within real-time wireless networks can be further supported by location awareness of the nodes, such that the handover from one access point to another can be pre-scheduled by making appropriate bandwidth reservations when a node approaches a cell border.

2. FlexWare System Concept

Starting from a system architecture as depicted in Fig. 1, the setup chosen for the FlexWare project foresees a centralized access point cluster (Fig. 2). It introduces a central controller to which the access points (or a group of access points) are attached. The nodes then communicate directly with this controller, and the access points are transparent. All tasks which are real-time critical, in particular scheduling between access points, are accomplished by the controller, and the backbone is relieved.

As a matter fact, the controller introduces an additional level in the data processing hierarchy appropriate for keeping local traffic confined. A part of this task can also be transferred to the access points that can manage traffic inside their own cells and help in ensuring real-time communication guarantees. Handover between
access points is facilitated as long as they are attached to the same controller. This three-tier centralized system consists of three essential components.

2.1. System components

The FlexWare Node (FN) is the wireless component on the lowest hierarchy level. Its role is to provide a communication interface to field level devices inside a factory. A FN has to manage its energy resources and has to ensure safety of and security for the end devices.

The FlexWare Access Point (FAP) coordinates traffic with its associated FNs and guarantees timely data transportation from end devices to the control room. FAPs are heavily involved in traffic scheduling within their respective cells, so that real-time guarantees are met and QoS is not compromised. They also are involved in the localization of all the (moving) nodes on the factory floor and have to accommodate resources in case of seamless handover, when a FN moves from one cell to another. It is clear that this functionality cannot be achieved by means of conventional legacy wireless access points. To support localization, the FAPs need precise synchronization. Furthermore, they require the implementation of advanced scheduling techniques. There have already been prototype realizations of such feature-enhanced access points [20].

The FlexWare Controller (FC) is the central entity coordinating all FAPs belonging to its automation island in the system to provide real-time communication guarantees between itself and the FNs. The connection between FCs and FAPs is based on wired interfaces. The actual protocol is currently under discussion, but a time-slot-based approach seems best suited for the purpose. With the help of high accurate clock synchronization and device management, the FC also provides localization information to all FAPs. Furthermore, it coordinates handover of a node from one AP to another and governs power management. The FC is also responsible for providing the interface between the FlexWare system and the real-time backbone, which is considered to lie outside the system boundaries.

Besides the basic components, the FlexWare Management Appliance (FMA) performs management, visualization and engineering functions. The management of the FlexWare network is divided between the FlexWare Controller and the FlexWare Access Points.

The hierarchical architecture also fits for some of the essential services (e.g., seamless handover, clock synchronization based on IEEE1588).

The backbone network is not the focus of the project, nevertheless there are some requirements it must meet. It should provide real-time support to accommodate traffic to and from external entities, such as a remote PLC interacting with a FN. Hence, the backbone network should offer deterministic cycle times and delivery delays. Different topologies of potentially large numbers of FCs may exist inside FlexWare, and therefore it can fit in nicely with those backbone networks that can support different topologies and support a large number of nodes. An additional requirement is that the backbone supports the safety and security focus of the system.
2.2. Scalability issues

An essential point for system architecture design is the number of access points that can be supported concurrently by a single controller. In large factory environments the cabling from the FC to the FAPs may be constrained by topological boundary conditions. Hence the usage of so-called star-point distribution units is reasonable. These units must be transparent to the FlexWare system functionality without affecting or impairing the system performance. The number of FAPs connected to an FC is influenced by the number of FNs that are being assigned to a single FAP. Considering that in a large factory the number of field devices can be thousands, the number of FlexWare Nodes can be high even if they act as data concentrators and can have multiple automation devices connected. The assumption for the FlexWare project is that the overall network consists of several clusters with a maximum of 100 nodes each. Not all of these nodes need to be mobile, though. The clusters are under the control of one (but replicated) FC.

If the wireless network must cover a large area, it may happen that the number of FAPs exceeds the maximum that could be assigned to an FC. Therefore, multiple FCs will have to be used, as shown in Fig. 2. Since any FC to FC communication would need to go through the backbone network (which is out of the scope of the project), no real time guarantees can be assured. This also applies to the case when an FN is moving from the area of coverage of one FC to some other. Seamless roaming is not guaranteed, which also imposes some limitations to the network layout.

The approach of using a central controller naturally raises dependability issues. However, individual automation islands can have multiple replicated controllers. There are many such concepts known, and the general system concept does not a priori prohibit their application to achieve better availability of the entire system. Meeting such demands does of course also require a careful network planning on the lower levels, in particular regarding the access points. It would, e.g., be necessary to cover any location in the system by more than one FAP.

3. Quality of Service management

The user requirement analysis performed in the reference scenarios targeted by the FlexWARE project makes it possible to define several traffic classes with different QoS requirements. A distinction is also made between negotiable and non-negotiable traffic flows. The negotiable traffic flows allow for a controlled reduction of the QoS received within specified performance bounds. More specifically, they make it possible to define a range of acceptable QoS values between a minimum QoS level and a target (desired) value. On the contrary, non-negotiable flows have to be always granted the target QoS level.

The requirement analysis indicated that the most relevant QoS parameters to be addressed are:

- Latency, defined as the transmission delay from sending a packet to receiving it,
- Relative jitter, defined as the delay variation between two consecutive packets of the same flow,
- Packet loss rate,
- Data throughput.

All of these parameters are defined at the application layer (i.e., they are end-to-end parameters). For traffic featuring real-time constraints (e.g., periodic traffic flows), some statistical guarantees on timely packet delivery or deadline meeting have to be provided. In order to achieve this target, resource allocation combined with utilization-based admission control will be applied, relying on the knowledge of the real-time traffic flow characteristics.

At the MAC layer the HCCA (Hybrid coordination function Controlled Channel Access) mechanism as a IEEE802.11e standard and further TDMA approaches will be first investigated by simulation to check the fulfillment of FlexWAREs QoS requirements as defined by the application designer.

However, in a hybrid system like FlexWare, the QoS offered to traffic flows can be affected by many changing factors, such as node movement, node removal or addition, system updates or reconfiguration (due to change in the network topology) and by the intrinsic non-determinism of the wireless channel. Therefore, QoS management mechanisms have to be provided with resource monitoring functions in order to be flexible and adaptive to the changing network conditions. Fig. 3 outlines the QoS architecture envisaged for the FlexWare project together with the relevant modules.

The main design principles behind the overall QoS management architecture are the separation of concerns and the minimization of the overhead due to management traffic. On the highest level, the FMA contains the definition of:

- the FN movement patterns and traffic specification,
- the position and transmission settings of the FAPs,
- the attenuation properties of the environment.

The FMA provides the FC with a model of the resource at design time. During run-time, however, interactions between different components are needed to enable them to efficiently accomplish their task, as explained below.

3.1. QoS management in the FlexWare Controller

The FC plays a major role in the QoS management of the automation island under its control. The FC implements run-time control functions in strict cooperation with the FAP to handle the shared resource,

---

2 Hard real-time bounds cannot be guaranteed on wireless networks, due to the various sources of unpredictable transmission errors and packet loss typical of the wireless channel.
i.e., the wireless medium. In FlexWare the resource is specified as the available time (e.g., time slots) for using the medium and the achievable data rate (in Mbps). The FC obtains from the FAPs the knowledge of the available resource (in terms of lower bound and upper bound) thus building a timely snapshot of the available resource in every cell of the whole network. The FC will then update the model of the resource provided by the FMA at design time in such a way to match the actual run-time resource behaviour.

The FC is also in charge of the inter-cell scheduling in order to avoid collisions in case of overlapping cells. This scheduling is performed both in frequency (assigning a suitable channel to each FAP) and time (assigning transmission time, e.g. time slots, on the channel). The FC will also maintain aggregated information about the cells (e.g., covered area, available resource and resource utilization) with the purpose of load balancing and further optimization.

The FC manages handoff to coordinate the resource reservation process in case of a mobile FN passing across several cells. Based on the traffic specification and the aggregated information about the cells, the FC will forward the request and negotiate it with one or multiple FAPs across the wired connection. The FAPs will run admission control algorithms within their cells, as explained below.

### 3.2. QoS management in the FlexWare Access Point

Based on the available resource granted by the FC, the FAP runs an admission control function to deal with any request for supporting a new traffic flow issued either by a FN already in its cell or by a new mobile FN trying to join the cell. As input to the admission control module, together with the request for supporting a traffic flow, the relevant FN will provide the traffic flow specifications which, together with the QoS requirements for traffic flows, will be stored in the QoS attributes image.

The FAP rules the transmissions of the FNs within its cell and performs intra-cell scheduling so that real-time guarantees and QoS requirements are met. The FAP is also responsible for run-time re-scheduling to accommodate either retransmissions (in case of faults) or regular transmissions of a newly-admitted traffic flow.

The FAP also perform error handling to adapt the schedule to guarantee the requirements of non-negotiable traffic flows, by re-negotiating the QoS for traffic streams which allow for a reduction of the QoS provided.
To monitor the channel state, the FAP periodically senses the channel and logs RSSI measurements to derive the signal strength profile of the channel. This information is locally exploited and is also provided to the FC in terms of upper/lower bound. This information allows to detect, for example, poor network coverage or potential interference to radio communication.

3.3. QoS management in the FlexWare node

The FN issues a resource request to let the FAP know its resource requirements. These requirements depend on the type and number of Automation Devices (ADs) connected to the FN. The request is fed into the admission control function run by the FAP at the system start-up or at run-time when a roaming FN enters the area of the FAP.

According to the resource granted by the FAP, the FN prioritizes the data to be transmitted to the FAP according to the QoS requirements of each flow. The FN performs traffic shaping to provide a better QoS to the highest priority traffic.

The FN is also in charge of the assessments of the wireless channel and of gathering resource statistics not only to support the handover process (as explained in the Sect. 5), but also to provide information to be exploited for increasing the network performance through the proper usage of the available resources (e.g., through load balancing).

3.4. Real-time traffic management

Admission control and traffic scheduling will be performed by the FMA to accommodate the traffic flows already known at design time. The schedulability will be assessed based on the resource model.

On the other hand, run-time resource re-allocation will be performed by the FC to deal with dynamically activated real-time traffic flows, while on-line admission control and intra-cell scheduling will be run by the FAP to handle mobility of real-time traffic across multiple FAPs.

Once a real-time data flow is admitted it is essential to provide it with the relevant guarantees also when changes or faults in the channel occur. Therefore, continuous monitoring is applied to ensure that the model of the resource always corresponds to the actual available resource. In case of mismatches, i.e., if the available resource decreases, the FC will inform the FAPs and coordinate the necessary corrective actions (e.g., QoS renegotiations of negotiable traffic flows to leave more resource to real-time ones).

4. Handover concept

Special attention has to be paid to the handover procedure between different FAPs, because a connection disruption might be not tolerable for specific industrial time-critical applications.

Usually, the handover process consists of four different phases, (i) the roaming-trigger/search phase, (ii) open authentication, (iii) association, and (iv) robust security network association (RSNA).

The current IEEE 802.11 specification does not provide fast handover. Table 1 shows the handover times of an industrial WLAN client and two APs using the standard roaming procedure [21]. The results show, that the mean handover time is above 2.3 seconds. Using the standard roaming procedure of WLAN a quite long time is needed to detect the loss of the currently used connection at the client side. After detecting this event all 13 channels must be scanned for new APs, whereby the waiting time per channel is approximately 120 milliseconds. In order to minimize the overall handover duration, either the roaming trigger/search phase or the RSNA can be optimized.

Table 1. Measurement of the WLAN-based handover time (mean values) with standard roaming.

<table>
<thead>
<tr>
<th>Type of RSNA</th>
<th>Phase 1 Roaming Trigger/ AP Search $T_{\text{search}}$ [ms]</th>
<th>Phase 2 Open Systems Authentication $T_{\text{Auth}}$ [ms]</th>
<th>Phase 3 Association $T_{\text{Asso}}$ [ms]</th>
<th>Phase 4 Robust Security Network Association $T_{\text{SecAsso}}$ [ms]</th>
<th>Total Handover Time $T_{\text{HandOver}}$ [ms]</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSK</td>
<td>2300</td>
<td>14.4</td>
<td>15.23</td>
<td>31.93</td>
<td>2343</td>
</tr>
<tr>
<td>802.1X</td>
<td>2300</td>
<td>13.87</td>
<td>16.77</td>
<td>67.67</td>
<td>2370</td>
</tr>
</tbody>
</table>

To support fast handover in IEEE 802.11 networks, a number of fast schemes have been proposed in the literature, which focus mainly on the reduction of the probe delay and the authentication/reassociation delay by using centralized or decentralized techniques [22]. Furthermore in [22] an adaptive fast handover framework based on a cross-layer approach is introduced, which will be investigated for usability in FlexWare.

One of the FlexWare features enabling the potential for optimizing the handover procedure is the localization ability of mobile nodes. In such a scenario the FC acts as a “Handover Coordinator” having knowledge about the position of each FN and FAPs’ coverage. Based on this information it can trigger the pre-handover process (e.g., if a FN reaches a specific area) which will prepare and improve the following handover (Figs. 3 and 4).

After the pre-handover procedure has been triggered, the FC first checks to which FAPs the FN could roam with a statistics request frame. Based on the information provided by different FAPs (statistics response frame that includes, e.g., FAPs load) the most suitable FAP is chosen. This first step can be omitted if the FAPs are able to send the statistical information automatically on a regular basis to keep the FC updated.
5. Conclusions and Future Work

Hybrid wired/wireless networks are becoming increasingly relevant for factory automation – not just from a research viewpoint, but also for concrete industrial use. Nonetheless, hybrid network architectures pose extraordinary challenges to system design and traffic management, in particular when multiple wireless segments are required to cover large areas and real-time communication needs to be extended across the borders of the wired or wireless network segments.

This paper outlines on-going work on the design and implementation of such hybrid network concepts and the handling of QoS therein. The chosen three-level architecture with clusters of wireless access points, each coordinated by a central controller, and several controllers being connected by a real-time wired backbone is a pragmatic approach that allows coping with complex issues like seamless handover of mobile wireless nodes between neighboring cells.

The architecture presented in this paper also leaves room for further improvements, which are under study. For example, the knowledge of the flow QoS requirements makes the controller aware of the load distribution over all the cells, which might be exploited for load sharing purposes. Furthermore the controller may select the most suitable access point (among the eligible candidates) for a mobile node based on both the node traffic QoS requirements and a policy which takes into account the cell utilization or other parameters (e.g., expected interference, etc.). The FlexWare project will also address such advanced aspects – not only in theory, but also in practical implementations demonstrating that hybrid flexible real-time networks are indeed possible.

References


[18] http://www.flexware.at


