Standard Ethernet in real-time, industrial applications

Claudio Salati, Daniele Fiorini

The paper describes the characteristics of today's Ethernet and shows that plain Ethernet and the higher level protocols related to TCP/IP, without ad hoc real-time enhancements, are often adequate to support industrial applications. Ethernet can yield additional advantages, simplifying the design of these systems and supporting a degree of resiliency that previously required a significant redundancy. Finally, Ethernet provides a clear path to future bandwidth increases.

Though Ethernet has been used in real-time and industrial applications since it first appeared, there have always been claims that it isn't suited for this because of its intrinsic non-determinism [1][4]. This has generated several enhancement attempts (collectively dubbed industrial Ethernet), conflicting with each other and often incompatible with the standard [24][26], with which they share sometimes only one characteristic, wiring; even when coexistence and cooperation are possible, the problem remains that a major advantage of Ethernet, the wide availability of low cost interface circuits and connectivity devices, is lost [16].

No convenient alternative to plain Ethernet has ever emerged, neither an Ethernet based field bus, nor an alternative LAN approach (Token Bus [11], Token Ring [12]), and, eventually, the introduction of switched Ethernet has significantly improved its real-time characteristics [1][4][5], thus relieving some of the previous concerns: so we think that plain Ethernet might become a dominant LAN also in the industrial world, independent of the emphasis field bus families are putting now on their Ethernet variant (e.g. [23][24]).

Another reason that makes Ethernet appealing is the straightforward integration with Internet via the use of TCP/IP, and the immediate availability of the related application services, e.g. DHCP for automatic network setup, FTP and HTTP for file transfer, and Telnet and HTTP for remote interactive access (the first, and still widespread, Ethernet based field busses were implemented on top of TCP/IP [27][28]).

Industrial applications require robustness and the failure of the communication network is often considered unacceptable. Redundant bus schemes have been devised to deal with this problem (e.g. figure 1), but they have a relevant cost since components and interfaces must be duplicated: today's Ethernet helps also with this issue.

C. Salati, Computer Networks, II Facoltà di Ingegneria, Università di Bologna; D. Fiorini, Platforms and Technologies, Datalogic Automation S.r.l.

Finally, Ethernet versions with different speeds are available to match the needs of different applications.

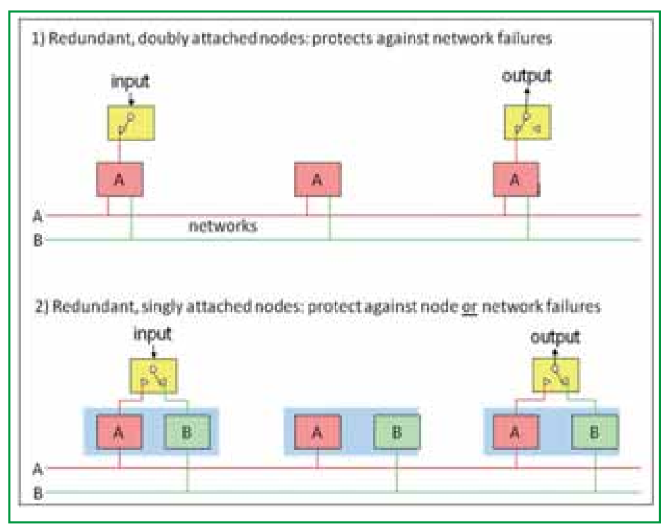

Figure 1 - Device and network redundancy (from [18])

Standard Ethernet: characteristics

Point-to-point, full-duplex, switched Ethernet

Ethernet [7] (or IEEE 802.3 [6], differences are negligible) has changed a lot from its first definition. Vintage Ethernet started as a real ether: all stations share the same communication medium, every station hears what another station emits and several effects are possible such as frames collision and loss, random delays, starvation [1]. Today's Ethernet is completely different: it consists of a mesh of switches interconnected with each other and with end-systems by point-to-point, (normally) full-duplex links [8][9]: in the office environment 1 Gbps is typical, but in industrial applications 10/100 Mbps is still the rule.

A switch is a MAC bridge [21] that, instead of bridging between two Ethernet segments, bridges between several point-to-point ports: it selects the output port where a frame must be forwarded based on its knowledge of the network; if the output port is busy with the transmission of a previous frame, the frame is queued and will be transmitted when the port gets free.
This implies that switches must know the topology of the network: they learn it by themselves: an address can be reached through a link because a frame with that sender address came in from that link. Only at the beginning of operations, when they know nothing about the network, switches use flooding (they forward a frame to all ports except the one the frame came in).
This full-duplex architecture avoids collisions since there is never contention for the use of the medium, but it can't guarantee loss-free operations since overflows could still occur in switches if the traffic pattern exceeds the buffer capacity of an egress port [5]. To minimize losses, a back-pressure mechanism has been defined: a switch that is getting congested issues a PAUSE frame against the sender and forces it to suspend the emission of new frames for a given duration [8]. Most literature on real-time communications assumes an absence of transmission errors [1]: in the context of today's Ethernet this assumption looks quite fair.
In a mesh network there may be several paths between 2 nodes and this creates loops; loops must be avoided because they result in flooding the network, so switches implement a spanning tree algorithm that disables redundant links [17]: the spanning tree protocol (STP) "allows a network design to include spare links to provide automatic backup paths if an active link fails, without the danger of bridge loops, or the need for manual enabling/disabling of these backup links" [19].
The current version of STP (Rapid STP, RSTP) provides for faster reconfiguration after a topology change: while STP could take 30 to 50 s, RSTP can typically react within at most few seconds (see section *Network topology*).
Some residual non-determinism of the transmission time is present due to switch traversal, including queuing time. Switches come in two main flavors [1]:

- Store-and-forward switches don't start forwarding a frame before it has been completely received.
- Cut-through switches start forwarding a frame as soon as there is enough information to identify the output port (i.e. the frame's header has been received: see figure 2), no collision has been detected and the output port is free.

Store-and-forward switches introduce much larger transmission delays, especially if many of them must be traversed: in a 100 Mbps network it takes more than 0.1 ms to a frame of maximum size to traverse such a switch. Cut-through switches can be 100 times faster, though the switch latency may not be null (addresses must be read from the frame's header and the egress port computed); for small frames the difference is less relevant.
Switches are generally built with a non-blocking switch fabric: this means that the bandwidth of the fabric equals the total bandwidth of all connected ports, and frames can be switched at line speed.

VLAN and traffic priority
In 2003 IEEE 802.1Q introduced VLANs (Virtual LANs) [20]: this concept allows separating the logical from the physical topology of a LAN. A physical LAN can be partitioned by assigning different ports of a switch to different logical LANs that don't belong to a same broadcast domain (port based VLAN: each ingress port is associated with a restricted set of egress ports). Multiple logical LANs can also run on a same physical infrastructure, but to support this frames belonging to different VLANs must be identified: IEEE 802.1Q changed the header of the Ethernet frame!

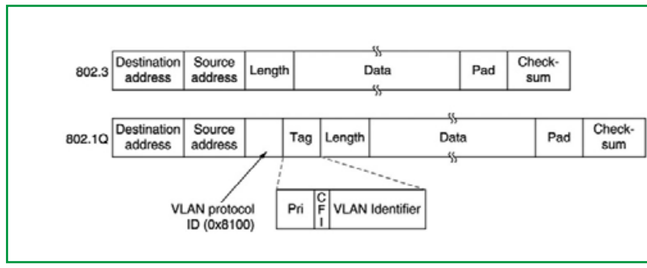

Figure 2 - IEEE 802.3 and 802.1Q Ethernet frame formats (from [21])

A specific value (0x8100) of the type/length field marks the frame as being 802.1Q and a new field is introduced immediately after, which yields the identifier of the VLAN the frame belongs to.
Because a switch may interconnect 802.3 and 802.1Q LANs, it must be able to mark traffic incoming from an 802.3 port as belonging to the appropriate VLAN (traffic classification): in practice the switch must add/remove the VLAN tag from the header when it bridges a frame from one LAN type to the other.

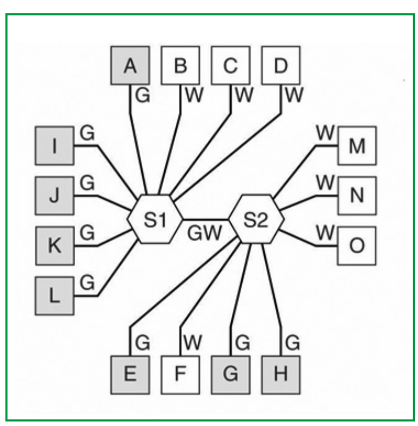

Figure 3 - VLANs (from [21])

2 VLANs, G and W, are defined in figure 3: some hosts belong to G, some to W (e.g. A to G and B to W), based on the assignment of a switch port to one VLAN or the other. Frames of both VLANs are transmitted on the trunk link connecting the 2 switches: on their trunk ports switches S1 and S2 support 802.1Q frames. Tributary ports may be plain 802.3 ports, depending on hosts capabilities: if so, traffic classification is required.
Besides introducing VLANs, IEEE 802.1Q adds a priority

field in the frame header (figure 1: a 3 bit field, thus up to 8 priority levels): frames of a same priority are queued in FIFO order at the egress port, but frames of a different priority are queued separately. When the egress link is free, the switch selects the oldest frame from the nonempty, highest priority queue. Traffic priorities allow separating real-time from best effort traffic. Traffic classification must be performed also with respect to priority.

Performance

[1] reports that, when frames are long and even with a large number of nodes, a vintage Ethernet can operate close to 100% utilization of the bandwidth (no time wasted in collisions and retransmissions). In industrial applications the size of real-time frames is generally small, and with smaller frames the maximum utilization drops, though it remains much higher than 37% [1]; but bandwidth requirements are often limited. Problems, if any, are due mostly to transmission delay and jitter: [1] reports also that very large link utilization can be obtained while keeping the delay below 1 ms on a 100 Mbps Ethernet.
The full-duplex switched architecture improves on this, except for the delay introduced by the traversal of switches: [5] shows that deterministic performances can be achieved also in a linear topology where jitter is amplified and that the support of traffic priority reduces significantly both maximum transmission time and jitter. We can use 70% of the available bandwidth without affecting real-time characteristics.
All performance characteristics are inversely related to the network load: by increasing the speed, they improve. Brute force is the strategy that supported the growth of the Internet: 1 Gbps Ethernet is now appearing on the factory floor and we may expect to see 10 Gbps Ethernet there in some time. There is a clear path to future bandwidth increase and technological development.
Several Ethernet based field busses guarantee determinism by slotting the time and assigning the LAN to a single node for each slot [16]: since in today's Ethernet collisions are intrinsically eliminated and all segments can be used in parallel without conflicts, this is largely useless and wastes resources.

Network topology

Ethernet allows to set up arbitrary mesh networks but in industrial applications only few topologies are relevant [1][18]. Our network model assumes that all nodes are connected to a same LAN, but one node is the functional master (controller) of the application. The same structure applies recursively if our distributed application is part of a larger system, with the master device being the gateway between the external and the internal worlds.
VLANs allow enforcing the distinction between internal and external communications, even if both are based on a same physical infrastructure. Consider figure 4 to figure 6: two VLANs can be defined over the internal LAN, one for internal and one for external communications; a single Ethernet port of the controller (or of any slave node) can be used to connect to both VLANs. Frames of the internal VLAN won't interact with the external world since they won't exit the internal LAN.
The star, or tree (figure 4), is a centralized topology where a set of switches, interconnected in a tree, is located in a cabinet and nodes are connected to this centralized infrastructure as the leaves of the tree (a single switch is enough if all nodes can be connected to it).
This topology is inconvenient if the number of nodes is large because of the amount of wires that must exit the centralized cabinet. Additionally, the largest distance that can be reached is the maximum length of a single Ethernet wire.
Two other drawbacks are the cost of the switches, including spare parts, and the fact that the centralized switch represents an additional single point of failure.
Theoretically this topology has an advantage, it supports Power on Ethernet (PoE), a technology that allows to pass electrical power, along with data, on Ethernet cabling [22]. But in industrial applications most nodes require more power than PoE can currently provide [22]; moreover, providing a centralized power source for all nodes may be a problem. PoE, therefore, isn't relevant yet.

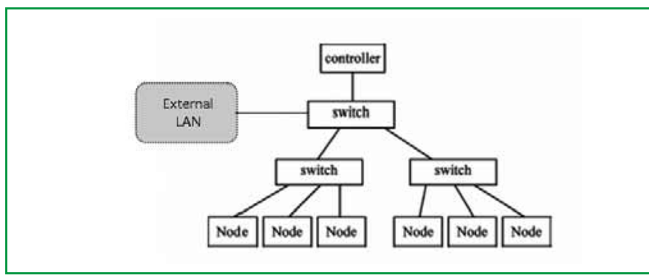

Figure 4 - Tree network architecture (from [1])

In conclusion, tree based networks are not considered the first choice in industrial applications except for specific cases, e.g. if very high levels of resiliency (and redundancy) are required [18] or switch traversal must be minimized.
A topology that doesn't require a centralized infrastructure and allows a convenient linear wiring scheme similar to that of busses like vintage Ethernet, LonWorks [2] and CAN [3], is daisy-chain: each node is connected to a local 3-port switch whose other two ports are linked to the switches of adjacent nodes.

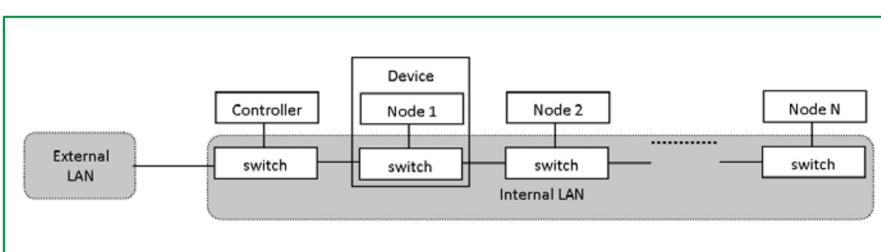

Figure 5 - Daisy-chain network architecture

There must be a switch in each device, but embedded switches are available at very low cost. This topology allows also to span large distances, since it is only the distance between two adjacent nodes that may not exceed the maximum length of a wire.
This topology is prone to single-fault failure but it can be made fault-tolerant by adding a redundant link between the spare port of the last switch of the chain and the first switch: this creates a ring, whose robustness has been known for a while (e.g. [29][30]).
RSTP is first used to decide what links will be active and what link will be disabled to avoid loops. In case of a topology change, e.g. if an active link fails, RSTP will reconfigure the ring and re-establish the connectivity between all nodes: because the topology is simple, RSTP is expected to complete the reconfiguration very fast ([18] reports that recovery may take less than 100 ms).

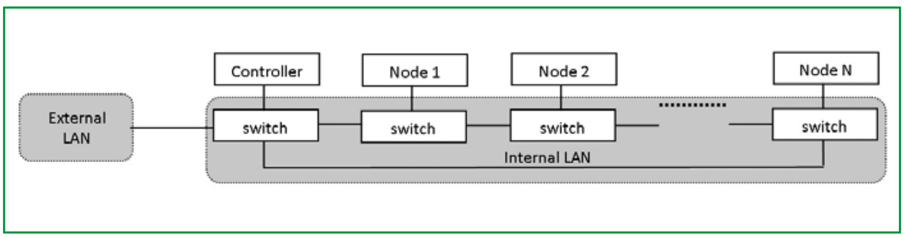

Figure 6 - Ring network architecture

In fact RSTP is considered adequate by IEC 62439 for network reconfiguration in general automation systems and also in some non-critical applications [18].
Also this topology has drawbacks (e.g. it doesn't support PoE, but this is irrelevant in our context): in a large application a frame may have to traverse tens of switches to go from the furthest node to the controller, and this implies a significant transmission delay, especially if switches are store-and-forward. Moreover, [5] shows that "due to the high number of switches the jitter is already high even if system loads are small", but it shows also that deterministic performances can still be achieved; additionally, RSTP can be used to minimize the maximum distance from a node to the controller [19].
Other topologies can be considered, e.g. see figure 7. These solutions can be easily devised since the network is based on standard components; a similar argument applies to the use of fiber or wireless connections.

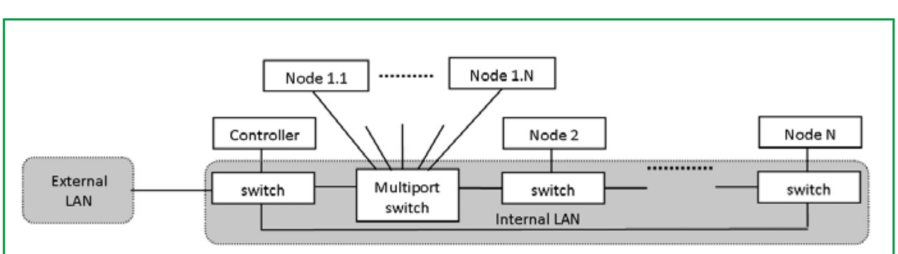

Figure 7 - Ring architecture including devices with a single Ethernet interface

Higher level protocols

Figure 8 shows several ways Ethernet can be used as the communication bus of industrial applications.
We advocate that several industrial applications

- don't need to be implemented on a modified real-time Ethernet,
- don't even need to be implemented directly on the data link interface,
- can be implemented on top of the TCP/UDP/IP stack, both for their real-time and the non-real-time part.

A first advantage is that programs based on the standard transport layer interface are easier to write and to port.
Second, there are many application services already available on TCP and UDP. Beside those already mentioned, NTP (Network Time Protocol) [13][15], that supports distributed clock synchronization, can play a major role in industrial applications where, for instance, data that are sampled by different sensors must be related to each other: to do it, thanks to NTP, they can be time-stamped consistently.
NTP runs on UDP, so it is easily portable even when not already available: the NTP Public Services Project distributes a reference implementation that can maintain time to within 10 ms over the public Internet, but in a constrained topology like an Ethernet ring the expected accuracy is around 1 ms [25].

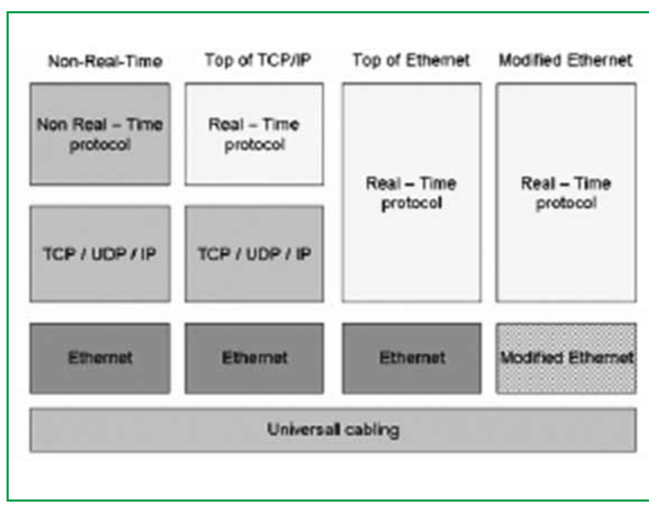

Figure 8 - Ethernet in industrial applications (from [16])

The Precision Time Protocol (PTP, IEEE 1588 [10]) has been designed to do even better: hardware based implementations achieve sub-microsecond accuracy [14]. PTP can be implemented on Ethernet and some embedded switches support it in hardware.
Time/space-stamping is the strategy adopted in Ethernet/IP to overcome

problems related to communications with a large transmission delay and jitter: it works as long as the system reaction time is not too small. This strategy can be adopted also in a plain Ethernet network thanks to NTP/PTP.
Some caveat:
- IP looks irrelevant since we are working within a single sub-network, but it is not: with TCP and UDP it is responsible to translate the quality of service (or an equivalent) parameter used within the TCP/IP stack to 802.1Q traffic priority.
- TCP must also support the TCP_NODELAY option that disables Nagel's algorithm [21] (not suitable to real-time applications).

The availability of a large and growing bandwidth may lead to introduce functions that were previously unconceivable, e.g. on-line image saving in vision systems. Strategies like traffic shaping can be adopted to avoid that these functions interfere with the real-time application.

Conclusions

Other characteristics of Ethernet make it particularly suited for the installation and maintenance of industrial systems, e.g. autosensing and auto-negotiation: one doesn't have to worry about wires, speed, duplex mode and use of Pause frames: two connected devices choose the optimal transmission mode they both support. Additionally, most switches support port monitoring and protocol analyzers are easily available.
Embedded switches with the characteristics described in this paper are actually available at very low cost (with some limits: all of them are store-and-forward and, less relevant, they support fewer priorities than specified by IEEE 802.1Q). The pervasive diffusion of Ethernet is certainly a reason for this and is also a guarantee of further future enhancements.
Ethernet has changed so much in time that some of the issues that raised discussions about it being suited to industrial applications have largely lost their meaning: it even allows implementing network topologies that, without additional redundancy, guarantee that a single fault won't affect the communication infrastructure.
The plain Ethernet solution is in practice strictly related to TCP/IP: this yields a convenient programming interface, availability of relevant network services and straight integration with Internet.

References

[1] J-D. Decotignie, "Ethernet-based real-time and industrial communications," *Proc. IEEE*, vol. 93, n. 6, 2005.
[2] *Control Network Specification*, Electron. Ind. Alliance Std. EIA-709.1, 1998.
[3] *Road Vehicles - Exchange of Digital Information - Controller Area Network (CAN) for High-Speed Communication*, ISO Std. 11898, 1993.
[4] M. Alves et Alii, *Ethernet goes real-time: A survey on research and technological developments*, Polytech. Univ. Porto, PT, 2002.
[5] J. Jasperneite et Alii, "Deterministic real-time communication with switched Ethernet," *Proc. 4th WFCS*, 2002.
[6] *Local Area Networks: Carrier Sense Multiple Access With Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications*, ANSI/IEEE Std. 802.3-1985.
[7] *The Ethernet - A Local Area Network - Data Link Layer and Physical Layer specifications*, v2.0, 1982.
[8] *Local and metropolitan area networks - Specific requirements - Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications*, IEEE Std 802.3, 1998.
[9] *Local Area Networks - Part 3: Carrier Sense Multiple Access With Collision Detect on (CSMA/CD) Access Method and Physical Layer Specifications*, IEEE Std. 802.3, 2002.
[10] *IEEE Std. for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems*, IEEE Std. 1588, 2002.
[11] *Local Area Networks - Part 4: Token-Passing Bus Access Method and Physical Layer Specifications*, ANSI/IEEE Std. 802.4-1990.
[12] *Local and Metropolitan Area Networks - Specific Requirements - Part 5: Token Ring Access Method and Physical Layer Specifications*, ANSI/IEEE Std. 802.5, 1995.
[13] D. L. Mills, "Internet time synchronization: The network time protocol," *IEEE Trans. Comm.*, vol. 39, n. 10, 1991.
[14] J. Edison, L. Kang, "Sharing a common sense of time," *IEEE Instrum. Meas. Mag.*, vol. 6, n. 1, 2003.
[15] "Network Time Protocol (Version 3) - Specification, Implementation and Analysis," *RFC 1305*, 1992.
[16] M. Felser, "Real-time Ethernet - Industry perspective," *Proc. IEEE*, vol. 93, n. 6, 2005.
[17] R. Perlman, "An Algorithm for Distributed Computation of a Spanning Tree in an Extended LAN," *ACM SIGCOMM CCR*, vol. 15, n. 4, 1985.
[18] H. Kirrmann, "Standard redundancy methods for highly available automation networks," *Proc. 11th IEEE ETFA*, 2006.
[19] "Spanning tree protocol", *Wikipedia*, http://en.wikipedia.org/wiki/Spanning_tree_protocol.
[20] *Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks*, IEEE Std. 802.1Q-2003.
[21] A.S. Tanenbaum, "Computer networks," *4th Ed., Pearson Edu. Int.*, 2003.
[22] "Power over Ethernet", *Wikipedia*, http://en.wikipedia.org/wiki/Power_over_Ethernet.
[23] J.Feld, "ProfiNet - Scalable Factory Communication for all Applications," *IEEE WFCS*, Sep. 22-24, 2004, Vienna, AT.
[24] E. Schemm, "SERCOS to link with Ethernet for its third generation," *Comput. Control Eng. J.*, Vol. 15, n. 2, 2004.
[25] http://www.ietf.org/proceedings/72/slides/tictoc-1.pdf.
[26] D. Jansen, H. Buttner, "Real-time Ethernet the EtherCAT solution," *Comput. Control Eng. J.*, Vol. 15 , n. 1, 2004.
[27] "Modbus messaging on TCP/IP implementation guide", *Schneider Automation*, 2002.
[28] ODVA, "The EtherNet/IPTM Specification," 2006.
[29] *Resilient Packet Ring*, IEEE Std. 802.17, 2004.
[30] M. Felser, "Media redundancy for ProfiNet IO," *IEEE WFCS*, Dresden, DE, 20-23 May, 2008.