Carrier Grade Metro Ethernet Networks

Beitrag der ITG-Fachgruppe 5.3.3 „Photonische Netze“
Achim Autenrieth, Andreas Kirstädter, Siemens Networks GmbH&Co KG; Bernhard Edmaier, Clemens Epple, BT (Germany) GmbH&Co oHG; Gert Eilenberger, Gert Grammel, Alcatel-Lucent Deutschland AG; Andreas Gladisch, Fritz-Joachim Westphal, T-Systems; Klaus Grobe, ADVA AG Optical Networking; Kristof Obermann, FH Regensburg; Erwin Patzak, Michael Schlosser, Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut; Jan Späth, Ericsson GmbH
Kontaktadresse: Dr. Gert J. Eilenberger, Alcatel-Lucent Deutschland AG, Lorenzstr. 10, 70435 Stuttgart
Tel.: +49.711.821-32166, Fax: +49.711.821-32457, E-Mail: Gert.Eilenberger@alcatel-lucent.de

Abstract

This paper gives an overview on actual trends and deployments of carrier-grade Ethernet in metro, access, and core networks. This includes the related motivation, concepts, and technologies as well as open issues regarding research, development, and standardization.

Ethernet as a packet-based, connection-oriented technology is deployed for metro networks worldwide today. This is driven by the massive increase of (IP-based) data traffic and the related applications. The Ethernet deployments aim at most cost-efficient data service provisioning and the migration of all legacy Layer-2 services towards a unified platform. The goal is a massive reduction of both, CapEx and OpEx.

Network operators and service providers impose increased requirements regarding scalability, quality of service including reliability and availability, and Operations, Administration, and Maintenance (OAM) features on their metro Ethernet solutions. These requirements are usually referred to as carrier-grade or transport Ethernet.

Metro Ethernet services as deployed today mainly consist of Ethernet Private Lines (EPL) or Ethernet Virtual Private LANs (EVPLAN). These can provide dedicated LAN extension or LAN-like connectivity via IP/MPLS, respectively. A different approach is Pseudowire Emulation Edge-to-Edge (PWE3) which allows MPLS transport of Ethernet and other packet services as well as synchronous TDM services.

Various network architecture and protocol options exist to migrate from metro SONET/SDH and WDM networks towards even more Ethernet-centric and -optimized networks. These include Layer-2 transport like Transport MPLS (T-MPLS), Provider Backbone Transport (PBT), and Ethernet-over-SONET/SDH/WDM/OTH. These services are currently under investigation or being standardized, and they will also migrate into long-haul and backbone networks. These approaches have common requirements regarding network and control planes (e.g., ASON/GLMPS, GELS, T-MPLS). Thus, the corresponding management and control mechanisms have to have an integrated view on the lower 3 network layers.

Further challenges for transport Ethernet result from upcoming technology steps like 100 Gbps Ethernet. Again, carrier-grade requirements and interworking aspects with transport networks have to be taken into account.

1 Introduction and Market Drivers

The last years have seen an enormous growth of bandwidth needs in metro networks of around 40% per year. Recently, this growth has been sharply accelerated especially from the residential side by the upcoming Web 2.0 driven end user applications like peer-to-peer and video file sharing. Traffic growth rates up to 100% per year are anticipated for the backbone networks.

At the same time, the revenues of the carriers stay pretty constant or are only rising in the order of a few percent. It is commonly understood that the resulting steep decline in revenue per bandwidth unit can only be absorbed via technological discontinuities: Only the introduction of new technologies provides the chance to substantially improve the cost situation of operators.

In this situation, Ethernet gets right in the focus of attention. Being a permanent success story since nearly two decades, scale effects have lowered the cost for enterprise Ethernet equipment down to such levels that it appears the suitable salvation for the bandwidth-revenue dilemma that the network operators and their customers are facing. Consequently, we currently see Ethernet appearing in many locations throughout carrier networks – both as a service and an infrastructure: E.g., Ethernet-based LAN interconnects are replacing classical Layer-2 services like ATM and Frame Relay in the form of transparent LAN point-to-point services. Their extension towards virtual private

LAN services is currently finding more and more customers. Many new applications in the sectors of health and education are based on Ethernet technology right from the start as are new concepts like Services grid computing.

At the same time the current residential access infrastructure is being migrated to architectures based on IP DSLAMs leading to a pure Ethernet connectivity throughout the access networks. By the way, this development will provide even more bandwidth to the end users – additionally fuelling the bandwidth explosion in the backbone networks. As it seems, Ethernet has the potential to become the new convergence platform for packet transport in the same way as IP has become the convergence platform for applications and services in the Internet.

However, not only the side of the capital expenditures has to be considered. Ethernet technology can only provide the correct solution for packet transport if operational expenditures in the network can be reduced at the same time. Therefore, special attention has to be concentrated on the carrier-grade design of those features of Ethernet that support the operations, administration, and management of networks.

The next section of this paper will explain in more detail these carrier-grade requirements for OAM, hierarchical layering and resilience. Section 3 then explains the most important Ethernet services and applications while section 4 describes Ethernet-based network architectures and technology trends. Finally, section 5 draws some conclusions.

2 Carrier-grade requirements on Ethernet transport

Carriers have different requirements with respect to their networks as compared to enterprise networks. These requirements reflect the necessity to operate and manage complex networks and to guarantee certain Service Level Agreements (SLAs). Consequently, the carrier requirements in particular apply to the areas of Operations, Administration, and Maintenance (OAM), layered network architectures, and mechanisms for resilience. These are briefly discussed hereinafter, with respect to the Ethernet protocol.

2.1 OAM

Next to the transport, the supervision of (Ethernet) signals is most relevant. Carrier Ethernet networks must provide OAM functionality similar to SONET/SDH. The basic Ethernet OAM mechanisms are described in the ITU-T Y.1730 and Y.1731 standards. These are related to the EPL, EVPL, EPLAN, and EVPLAN reference models and are aligned with the ITU-T SG15 (G.8010, G.8011).

In principle, OAM functions have to provide and support FCAPS management (Fault, Configuration, Accounting, Performance, and Security Management). The basic functions which are necessary include (continuous) Connectivity Check (CC), Loopback (LB), Trace Route (TR), and functions for alarm suppression, discovery, performance monitoring, and survivability (protection switching, restoration). This is described in Y.1731 (ex-Y.17ethoam) in more detail.

The OAM functions CC, AIS/RDI signalling, ping (LB) und Trace Route are schematically shown in Fig. 1.

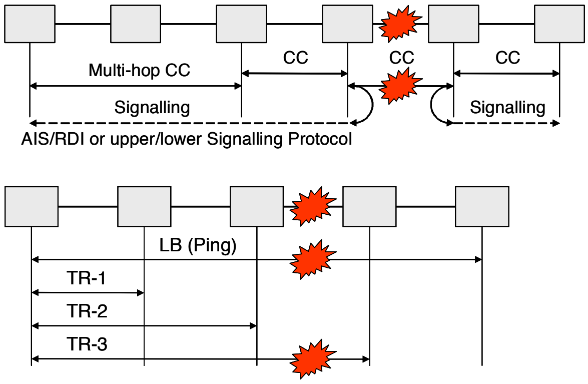

Fig. 1: Ethernet OAM: Connectivity Check (CC), AIS/RDI signalling, Trace Route (TR), and ping (LB)

End-to-end management is a major requirement for carrier-grade Ethernet. (The same is true for every other transport technology, and end-to-end management was one of the main drivers behind OTH, for example refer to TCM, Tandem Connection Monitoring.) In the Ethernet context, all network layers (core, access) and all technologies (e.g., MPLS, EFM) have to support the related basic OAM functions. This includes interworking over several carrier domains. This scenario is shown in Fig. 2, together with the most relevant OAM functions. The Ethernet standards 802.1ag and 802.3ah as mentioned in Fig. 2 are described in more detail hereinafter.

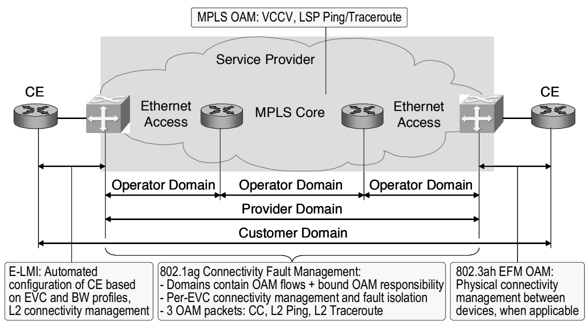

Fig. 2: Ethernet end-to-end OAM (E-LMI: Eth Line Management I/F)

2.2 Layered Network Architecture

Ethernet networks must be able to provide transparent interconnection of all sites of given customers A, B, C etc. while maintaining complete isolation between these customers. The corresponding function is known as VLAN tagging and is standardized in IEEE 802.1ad. The major disadvantage of this standard is its lack of scalability, or the lack of providing a hierarchically layered network architecture. 802.1ad is still limited to 4096 VLAN addresses. VLAN tagging (also referred to as Q-in-Q) was hence complemented by a fully recursive, layered architecture which is referred to as M-in-M (Mac-in-Mac), or Provider Backbone Bridge (PBB), and which is described in 802.1ah. The corresponding end-to-end network concept is shown in Fig. 3.

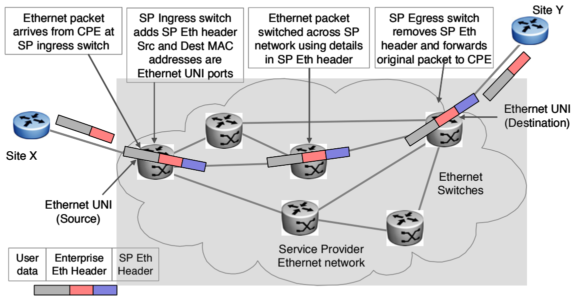

Fig. 3: IEEE 802.1ah MinM Principle

The layering (i.e. the encapsulation of client Ethernet frames into carrier frames) is provided in a shim header. Basic shim functions are mapping of 802.1ad S-VIDs (Service VLAN ID) into Extended Service VIDs (I-SIDs), encap/decap of 802.1ad frames, learning and correlation of backbone POP and customer MAC addresses, and filtering of L2 control packets sourced by core relays or by provider bridge relays (divides spanning trees). The PBB shim functions are shown in Fig. 4.

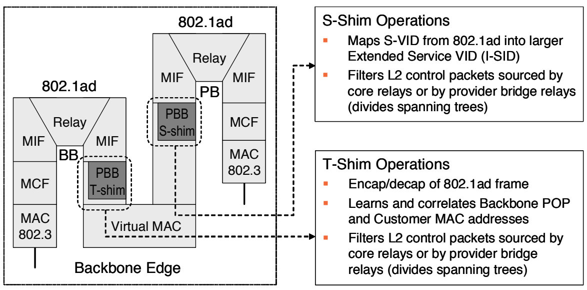

Fig. 4: PBB Shim Functions (MIF: Media Independent Function, MCF: MAC Convergence Function)

With M-in-M, each B-VLAN (Backbone VLAN) carries many S-VLANs (Service VLANs, i.e. 802.1ad VLANs). S-VLANs may be carried on a subset of a B-VLAN (i.e. all P-P S-VLANs could be carried on a single multipoint B-VLAN providing connection to all end points). An I-SID uniquely identifies an S-VLAN within the backbone. B-VLANs are addressed like regular VLANs with a 12 bit B-VID. B-VID and I-SID need to be separate ID spaces to allow many S-VLANs to be carried in a single B-VLAN. The resulting MinM data plane frame format is shown in Fig. 5.

M-in-M obviously is a complete recursion instead of just adding further tags as is the case with Q-in-Q.

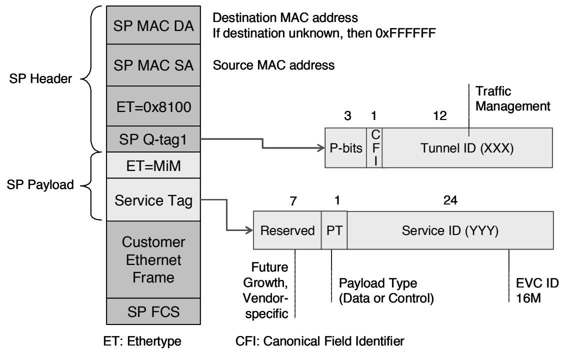

Fig. 5: MinM Data Plane Frame Format

2.3 Resilience

Resilience (protection, restoration) is necessary to enable a certain availability (AV) of services. This is necessary because AV is part of Service Level Agreements (SLAs). Today, SONET/SDH and WDM (ring) protection is used in most metro networks. These provide high (service) availabilities and fast switch-over. Disadvantages include high cost due to redundant capacity, and the fact that more and more services – in particular Ethernet – are not to be transported over SONET/SDH networks anymore. This leads to the requirement for additional resilience mechanisms in the Ethernet layer. These can be complemented on demand by protection in the SONET/SDH, OTH, and IP/MPLS layers, and later by GMPLS restoration (ITU-T ASON).

Typical examples of service (i.e. path) AVs are 98.5% for unprotected services, and up to 99.995% for highly available, protected services (e.g., in the SAN context). Here, path AV includes fibers and equipment. Hence, high AV can be achieved by providing redundancy with respect to fibers and transport equipment. Path AV is then influenced by the fiber downtime, together with the availabilities of the service-affecting components.

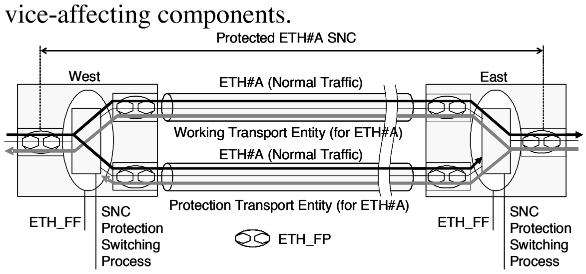

Fig. 6: Ethernet 1+1 protection (ETH_FF; Ethernet Flow Function, ETH_FP: Ethernet Flow Point)

Ethernet 1+1 and 1:1 point-to-point Sub-Network Connection (SNC) protection is currently standardized in ITU-T G.8031 "Ethernet Protection Switching" (Y.1342, ex. Y.17ethps). This standard describes SNC

protection for sub-networks constructed from point-to-point Ethernet VLANs. The basic functionality is shown in Fig. 6.
Next versions of the G.8031 will consider sub-networks constructed from multi-point-to-multi-point Ethernet VLANs. Potentially, this will include enhancements to IEEE 802.17 (Resilient Packet Ring, RPR) in order to provide ring protection for specific Ethernet VLANs.

3 Ethernet Applications

3.1 Virtual Private Networks

The work on VPNs in IETF essentially started with BGP/MPLS VPNs leading to a basic requirements document in 1999 (RFC2547). Later work was based on this standard and split up into activities related to Layer 3 (IP), Layer 2 (Ethernet) and Layer 1 (SDH, OTH). As a consequence also RFC2547 got several updates and is now replaced by RFC4364. The L3VPN working group which initiated the work on VPNs is responsible for defining provider-provisioned Layer-3 (routed) Virtual Private Networks (L3VPNs). Ethernet as service or transport technology is not considered there but in L2VPN and L1VPN activities.

L2VPN

The L2VPN activity deals with the question on how to create and transport Ethernet services over an IP/MPLS network providing the following services:

- Virtual Private LAN Service (VPLS): A L2 service that emulates LAN across an IP and an MPLS-enabled IP network, allowing standard Ethernet devices to communicate with each other as if they were connected to a common LAN segment.
- Virtual Private Wire Service (VPWS): A L2 service that provides L2 point-to-point connectivity (e.g. Frame Relay DLCI, ATM VPI/VCI, point-to-point Ethernet) across an IP and an MPLS-enabled IP network.
- IP-only VPNs: A L2 service across an IP and MPLS-enabled IP network, allowing standard IP devices to communicate with each other as if they were connected to a common LAN or with some mesh of point-to-point circuits (not necessarily fully meshed).

L2 interworking is not in the current scope. Overall, the work on above subjects is well advanced and can be considered as stable from a standardization point of view.

L1VPN

In contrast to the L2VPN activity which is using IP/MPLS as a server layer technology for Ethernet services, the L1VPN Working Group specifies mechanisms necessary for providing layer-1 VPN over a GMPLS-enabled transport service-provider network. This work essentially deals with the question on how to provide non-IP VPNs directly on non-IP/MPLS transport technology such as SDH, OTH and in future Ethernet. Unfortunately, Ethernet is yet not addressed, because as a pre-condition a GMPLS control Plane for Ethernet needs to be defined. As long as IEEE, which is the 'owner' of Ethernet standardization, does not assign an appropriate label space which can be used for control, further work on this subject is blocked. A first attempt to create a working group dealing with GMPLS for Ethernet was made to achieve the following goals:

- Control of Ethernet switches using GMPLS protocols in support of point-to-point paths.
- It is a non-objective of the IETF to initiate any Ethernet data plane work

PWE3

One key Element to provide VPN Services is the capability to transport non-IP traffic over an IP/MPLS network. This requires communication services that can emulate the essential properties of traditional communication links over a PSN. A pseudowire emulates a point-to-point link, and provides a single service which is perceived by its user as an unshared link or circuit of the chosen service. It is not intended that an emulated service will be indistinguishable from the service that is being emulated. The emulation needs only be sufficient for the satisfactory operation of the service. Emulation necessarily involves a degree of cost-performance trade-off. Switching, multiplexing, modification or other operation on the traditional service, unless required as part of the emulation, is out of the scope of the PWE3 WG.
A PW operating over a shared PSN does not necessarily have the same intrinsic security as a dedicated, purpose built network. In some cases this is satisfactory, while in other cases it will be necessary to enhance the security of the PW to emulate the intrinsic security of the emulated service. PWE3 will work closely with the L2VPN WG to ensure that a clear demarcation is defined for where PWE3 stops and L2VPN starts.
WG Objectives are to specify the following PW types:

- Ethernet, Frame Relay, PPP, HDLC, ATM, low-rate TDM, SONET/SDH and Fibre Channel.
- PWE3 will not specify mechanisms by which a PW connects two different access services.
- Specify the control and management functions of chartered PW types, to include PW setup, configuration, maintenance and tear-down
- Specify Operation and Management (OAM) mechanisms for all PW types, suitable for operation over both IP/L2TPv3 and MPLS PSNs, and capable of providing the necessary interworking with the OAM mechanisms of the emulated service.
- Define requirements for and mechanisms to provide protection and restoration of PWs.

3.2 Ethernet Carrier Internal Use

From carriers' perspective, the main driver for Ethernet services will be external customers, but also for internal demands Ethernet will become more and more an effective and attractive solution.
The driver for the use of Ethernet will be the same, as described in chapter 1 of this document, namely reduced ports costs and scalability options. During the following years, Ethernet ports will be a standard product, which will be produced in a huge amount. For that reason, Ethernet interfaces will be much cheaper than e.g. SDH Interfaces. The price for a STM-16 SR (=Short Reach) SFP will be approximately equivalent to 2 x Gigabit Ethernet SFP ports (1000 Base LX) or 4 x Gigabit Ethernet ports (1000 Base SX). This small example shows what impact to the costs of a network operator can happen if a change from SDH / POS interfaces on equipment to Ethernet interfaces will be done. Also if the carriers will use for their internal platform one common "protocol", Ethernet, it is possible to standardize the in-house infrastructure (e.g. cabling). These savings for interfaces and infrastructure will become much larger if a complete network will be possibly changed. The benefit will increase if more new platforms like Voice over IP or VDSL or Layer 2 networks will be rolled out.

The second advantage of Ethernet will be the scalability effects of Ethernet. With SDH you will have a very rough granularity (E1, E3, STM-1, STM-4, STM-16). For low bitrates SDH has a very flexible and fine granularity, but for higher bandwidth (STM-4 → STM-16 → STM-64) the capacity increases by the factor of 4. For Ethernet it is possible to increase the capacity with two mechanisms. One is the increase on additional ports (e.g. n x 10 Mbps) or also the required bandwidth can be defined by software on a certain level (e.g. 23 Mbps). The advantage of this mechanism is that several platforms can share one big data pipe and every service will have a guaranteed bandwidth.

The change to Ethernet was also driven by the change of the method of transport of Ethernet or data traffic. In the beginning, backbone networks were based on SDH and there were only very difficult ways of mapping Ethernet into SDH. The at this time available solutions were low cost, not managed converter boxes or routers in bridge mode or the POS interfaces, which do some kind of a mapping of IP data into SDH on the card. Now functionalities like GFP or G.709 are available, which offer the opportunity for mapping Ethernet traffic directly into SDH or wavelength/sub wavelength. Also, DWDM systems in the Wide Area are no more only focused on SDH interfaces (e.g.STM-16/STM-64) and nowadays offer for example several Gigabit Ethernet and 10 Gigabit Ethernet interfaces on one wavelength transponder.

4 Network Architectures

4.1 Architectures for the First Mile and aggregation networks

Ethernet in the First Mile (EFM)
Ethernet in the First Mile (EFM) is a relevant standard in the Metro Ethernet context. It is described in the IEEE 802.3ah standard and promoted in the market by the Ethernet in the First Mile Alliance (EFMA). An overview is given in [1].
EFM describes native Ethernet access in metro networks. Design goals were on consolidation of the access with respect to the dominance of the Ethernet protocol, and hence lowest cost for voice, data, and video access. It is believed that EFM will replace other access technologies (E1/T1, E3/T3, STM-1/OC-3) over time. Using EFM, expensive protocol conversions in the access can be avoided. In addition, it is possible to use single-ended demarcation units which can be managed directly via the service providers' edge routers (no unit necessary in the service providers' PoPs).
In IEEE 802.3ah, three access topologies are defined – copper-based (EFMC), SSMF-basiert (EFMF), and based on a passive point-to-multipoint topology (EFMP, the EFM version of EPON). Hybrid solutions (EFMH) are also possible. For these topologies, IEEE 802.3ah defines the OAM methods, i.e. performance monitoring, loopback, and fault detection and isolation.
For fiber access, EFMF is the relevant substandard. It defines full duplex with 1 Gbps GbE via SSMF over at least 10 km distance. It also describes single- and dual fiber access for point-to-point at 100 Mbps. EFMC defines access via Cat3 copper cables at 10 Mbps over 750 m.

Ethernet Passive Optical Network (EPON)
The IEEE 802.3ah EFM standard also introduces the concept of Ethernet Passive Optical Networks (E-PONs), in which a Point-to-Multipoint (P2MP) network topology is implemented with passive optical splitters, along with optical fiber Physical Medium Dependent sublayers (PMDs) that support this topology. In addition, a mechanism for network Operations, Administration and Maintenance (OAM) is included to facilitate network operation and troubleshooting.
EPONs (also known as EFMPs) are supported in the market by the Ethernet First Mile Alliance (EFMA) which became part of the Metro Ethernet Forum (MEF) [2].
EPONs enable IP-based P2MP connections using passive fiber infrastructure. Up- and downstream (US, DS) are controlled using the Multi Point Control Protocol (MPCP). The US makes use of TDMA.
EPON was mainly motivated by the disadvantages of ATM (APON). These include the facts that dropped cells invalidate entire IP datagrams, that ATM imposes

a cell tax on variable-length IP packets, and that ATM in general did not live up to its promise of becoming an inexpensive technology.
EPON on the other hand provides an IP data-optimized access network, considering the fact that Ethernet is by far the most relevant protocol in the access. It provides EPON encapsulation of all data in Ethernet fames. The EPON layer stack is shown in Fig. 7, in comparison to APON and GPON.

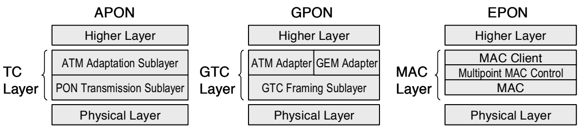

Fig. 7: EPON layers compared to APON and GPON. (G)TC: (GPON) Transmission Convergence layer.

Single-mode fibers are used for EPON. Single-Fiber Working is enabled by using 1300 nm for the US and 1500 nm for the DS, respectively. Splitting ratios of 4:1 to 64:1 are supported (typically 16:1). The maximum optical power budget is 20 dB, enabling maximum link lengths of 10…20 km.
EPON provides a symmetrical bit rate of 1.25 Gbps for Ethernet transport only. In the DS, Ethernet frames transmitted (broadcast) by the OLT pass through the N:1 passive splitter and reach each ONU (with own MAC addresses). This is similar to a shared-media network. Almost 50% of the available bandwidth is required for the protocol overhead, leaving only ~600 Mbps for revenue use.
In the US, data frames from any ONU will only reach the OLT due to the directional properties of the passive splitter/combiner. This is similar to an Ethernet P2P architecture. However, EPON frames from different ONUs transmitted simultaneously can still collide. Hence, ONUs need to share the trunk fiber channel capacity and resources.
The EPON system provides a very basic transport solution where cost-effective data-only services are the primary focus. EPONs are receiving a lot of attention in the Far East where missing pieces of the 802.3ah standard are being driven by NTT. There is not much interest in the U.S. and parts of Europe.
EPON as a protocol is still under work within the IEEE EFM group. In the 802.3av Task Force (10GEPON) the physical layer is extended to 10 Gbps.

Ethernet Aggregation Networks

Besides the last mile and PON structures, Ethernet is also starting to spread out in the aggregation and metro area – often denoted as “second mile”. Here, currently a replacement of traditional ATM-based aggregation structures is taking place: One or more Ethernet switching stages aggregate the traffic of the residential customers which is, at least in Europe, often provided by xDSL techniques on the lasst mile and thereby utilizing the existing copper-based-infrastructure. Beside the standard IP services, the Ethernet aggregation platforms are also used to offer IPTV, video on demand and voice-over-IP services. Moreover, Ethernet techniques in the aggregation domain are also a very promising candidate to be used as common production platform for the different service portfolios offered to residential and business customers (Figure 8) and thereby also providing business customers with dedicated IP- and Ethernet based services (e.g. LAN interconnection) over the same aggregation infrastructure.

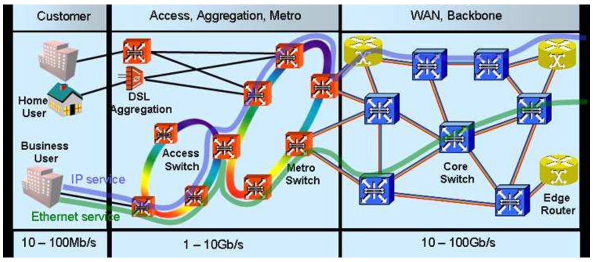

Fig. 8: Ethernet-based aggregation networks for residential and business customers [3].

4.2 Backbone Ethernet Networks (connection oriented forwarding)

Currently, transport networks mainly use SDH-based framing architectures like GFP for transferring Ethernet traffic over transport networks. However, novel concepts arise that use packet techniques directly above the WDM layer. However, the unmodified usage of end-to-end Ethernet network concepts in general is limited by scalability issues: Based on configurable IDs of switches, configurable port weights, and priorities, the Spanning Tree Protocol (STP) calculates a single tree-structure to connect any switch with each other. Although loop-less forwarding is guaranteed with this mechanism, STP provides only one path between two locations and a MAC address learning of any equipment is performed at the switches.
When combining large networks and adding hundreds of customer networks with an Ethernet-based core network, the number of MAC addresses grows rapidly. Thus, scalability cannot be provided and a separation of networks or an additional hierarchy between them has to be introduced to allow a scalable forwarding of data.
Also, the use of only one tree structure and with it the possibility to use only one path between two locations hamper the use of efficient traffic engineering and resilience mechanisms. Thus, three connection-oriented forwarding technologies are currently discussed at standardization bodies for Carrier-Grade Ethernet transport networks: VLAN Cross-Connect (VLAN-XC), Provider Backbone Transport (PBT), and Transport Multi-Protocol Label Switching (T-MPLS). Scalability is provided by all three proposed forwarding technologies via the introduction of a backbone-

network hierarchy for the forwarding of traffic. Edge switches manipulate the incoming Ethernet packets and add tunnel information. Instead of MAC learning (which is disabled inside the core in all the technologies) the forwarding is performed along pre-defined tunnels. The number of tunnels that have to be provided depend on the number of edge-switches, supported types of services, and the number of distinguishable networks (VLANs).

VLAN Cross-Connect (VLAN-XC):
The main idea of VLAN-XC is to establish pre-defined tunnels between edge switches of a network and to use these tunnels to route and differentiate traffic from each other. Instead of using a destination MAC address for the forwarding decision, a label (VLAN-XC Tag) is encoded in the Ethernet header to determine the appropriate tunnel. Ingress edge-switches have to analyze incoming packets, chose one of the pre-defined tunnels, and label an Ethernet packet accordingly. Intermediate switches route the traffic according to the given tunnel label and are able to swap the label. Finally, the tunnel label is removed at an egress switch to allow the transparent transportation of customer data. With this functionality, multiple paths between two edge-switches are supported. Traffic can be separated and distributed in the network and traffic engineering is facilitated. To avoid changing the Ethernet header structure, VLAN-XC uses the bits reserved for VLAN-IDs of IEEE 802.1Q and IEEE 802.1ad to encode the tunnels.

Provider Backbone Transport (PBT):
Similar to the VLAN-XC, Provider Backbone Transport establishes pre-defined tunnels between edge switches. However, instead of adding a label to the header, a MAC encapsulation is performed at the edge switches (Figure 8).

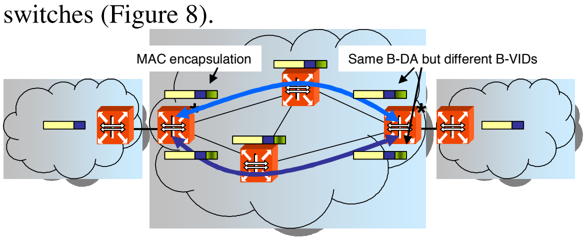

Fig. 8: MAC encapsulation in the Core network in PBT.

Transport Multi Protocol Label Switching (T-MPLS):
Transport Multi-Protocol Label Switching (T-MPLS) is an adaptation of MPLS and is defined in ITU-T G..8110.1. The main idea is to use the well established MPLS concept known from IP routing and adapt it for transport forwarding issues. As with VLAN-XC and PBT, T-MPLS establishes pre-defined tunnels. In T-MPLS an additional MPLS header is pushed in front of the client traffic that is transported transparently inside the backbone network. Similarly to VLAN-XC the 20bit label is used to encode the backbone tunnel and is removed at the egress backbone switch. Figure 9 illustrates the frame structure of T-MPLS.

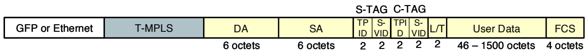

Fig. 9: T-MPLS frame structure.

To use the MPLS concepts also in transport environments, some changes were necessary. E.g., the control planes are separated, i.e. T-MPLS operates independently of its clients and its associated control networks (management and signalling network). The use of Penultimate Hop Popping is prohibited as are the merging of tunnels as well as the equal distribution of traffic onto paths (ECMP).

4.4 Control/Management aspects

Besides already well elaborated and widely used Control Plane protocols based on MPLS and GMPLS, new approaches specifically taylored for Ethernet are currently under discussion.
In November 2005 an initiative in the Internet Engineering Task Force (IETF) was started to use the GMPLS Control Plane for Ethernet switches in order to scale Ethernet solutions beyond the limitations of a LAN service. This initiative called GELS (GMPLS-controlled Ethernet Label Switching) intended to dynamically manage the Ethernet resources. The idea was to advertise the aggregate available bandwidth on each wavelength-link together with the set of available Ethernet VLAN tags via OSPF-TE. Provisioning actions could be instantiated using RSVP-TE signalling in order to set up Label Switched Paths (LSP) with the requested bandwidth and a proper VLAN tag. Each Ethernet switch would then translate RSVP-TE signalling messages into local switch commands to create the desired VLAN-ports associations along with the requested bandwidth guarantees. Whenever an Ethernet circuit (or LSP) is set up or torn down, the bandwidth and VLAN tag information would be updated via distribution of OSPF-TE Link State Advertisements (LSAs) in order to maintain proper link states across the network. This way, a scalable Ethernet network for a Wide Area Network could be achieved including all defined resilience and maintenance mechanisms currently available on GMPLS implementations for SDH/SONET networks.
While the underlying idea was appealing, no progress was made so far since backwards compatibility with existing Ethernet switches is of major concern for operators as well as for vendors. It is an ongoing activity to scope the GELS activity such that compatibility issues are covered sufficiently. At this point in time the following issues need to be resolved:

1. Ethernet VLANs have no bandwidth assigned, while in GMPLS bandwidth assignment would be used to improve scaling and allow traffic engineering.

2. Ethernet VLAN labels are not switchable entities while in GMPLS an addressing entity is required to be switched on a per port basis.
3. Alternative approaches using special identifiers or MAC addresses are also under consideration but raise concerns about interoperability and scalability of the overall solution.

Further work is necessary to identify in collaboration with IEEE the required identifiers and switching entities which allow the implementation of a GMPLS based control plane.

4. 5 Technology trends

In general, Ethernet operation at speeds of 100 Gbps is very desirable in terms of architecture-related network cost [4]. The transmission of high speed data rates above 100 Gbps itself is well understood and can be managed. As a consequence, the knowledge to realize the transmission of a 100 Gbps Ethernet signal is present. The problem still to solve is to find efficient electro-optical and opto-electrical conversion techniques. Electrical solutions are preferable to handle the data at the transmitter and receiver since OTDM techniques are still too complex and difficult to implement in commercial products.

By using ultra-fast electronic circuits instead of elaborating optical methods in high-capacity optical transmission systems cost per transmitted bit per second and kilometer can be reduced. Electronic circuitry for 40 Gbps is already commercially available. To really exploit the cost advantage of an electrical receiver compared to optical solutions a compact integrated device is needed - preferably a single chip.

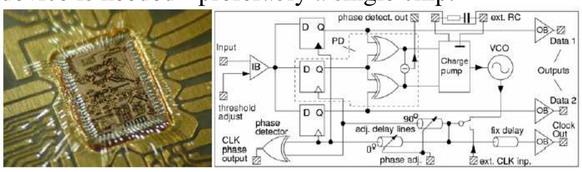

Fig. 11: Photo (left) and block diagram (right) of the integrated ETDM receiver chip [5].

Recently, as an important step towards 100 Gbps Ethernet an integrated ETDM receiver comprising 1:2-demultiplexing (DEMUX) and clock & data recovery (CDR) on a single chip was presented [5]. This receiver was tested in a 100 Gbps transmission experiment. Error-free performance (BER < 10-9) was obtained back-to-back and after transmission over 480 km of dispersion managed fiber. The ETDM receiver was initially designed for 80 Gbps operations. A redesign of the receiver chip is expected to enable an even better performance and operation at even higher bit rates.

The IEEE 802.3 Higher Speed Study Group was established in 2006 to evaluate the extension of the present Ethernet standards to interface speeds of 100 Gbps. An objective is the support of 100 Gbps over 40 km on standard single mode fibers and 100 meters on OM3 multi mode fibers. The Ethernet frame format and size will be unchanged. Besides the serial transmission of 100 Gbps currently the following WDM options are under discussion for the realization of the physical layer: 10x10 Gbps, 5x20 Gbps, 4x25 Gbps and 2x50 Gbps. Up to know it is not finally clear, which versions will go into the final standardization process.

Within the ITU there are activities concerning 100 Gigabit Ethernet in the Study Group 15. One objective is the support of OTN interworking, another is the investigation of parallel interfaces (WDM) or serial interfaces. A further concern is the support of already installed fibre infrastructure.

5 Conclusion and outlook

We describe the advantages of Ethernet for customer and carriers. Furthermore, the paper gives an overview of several tendencies in the development and future of Ethernet. But at the moment it is not completely decided, what will be "the" solution for Carrier Grade Ethernet.

A common understanding of the technology and interworking options of the different solutions (e.g. between carriers or vendors) should be available. Otherwise only island solutions for Ethernet networks will be available, like it is today.

Today there are different solutions for offering Ethernet Services, like e.g. conversion on a fibre or SDH bandwidth, an SDH/GFP solution, Switched Ethernet platforms or an MPLS based VPLS solution.

The problems occur, when the different solutions will be connected to one service. This can happen due to different possible frame sizes (e.g. from 64 Bytes to up to Jumbo Frames with 9028 Bytes), transparency (e.g. only data transparency to transparency of VLAN-IDs, Mac-in-Mac, Q-in Q, customer specific or vendor specific signalling information, Link Aggregation, EFM, Fault Management and Multicast Frames or protocols like Spanning Tree) and alarming status (e.g. Link Loss Forwarding, switch off of the port, or not defined status).

Nevertheless, Ethernet will be "the" transport protocol for the future and will lead us to Ethernet based networks.

6 References

[1] Ethernet in the First Mile. Making Universal Broadband Access a Reality. A White Paper presented by Ethernet in the First Mile Alliance (EFMA), 2003

[2] www.metroethernetforum.org

[3] Lehmann / Autenrieth / Derksen / Leisching: "Die neue Ethernet-Generation: 100-Gigabit-Ethernet mit integrierten elektrisch-optischen Hochgeschwindigkeitsschaltkreisen", Photonik, 2/2006.

[4] A. Schmid-Egger / A. Kirstädter: "Ethernet in Core Networks: A Technical and Economical Analysis", HPSR2006 - 2006 IEEE Workshop on High Performance Switching and Routing, Poznan, Poland, June 7-9, 2006..

[5] R.H. Derksen, G. Lehmann, C.-J. Weiske, C. Schubert, R. Ludwig, S. Ferber, C. Schmidt-Langhorst, M. Möller, J. Lutz, „Integrated 100 Gbit/s ETDM Receiver in a Transmission Experiment over 480 km DMF", Proc. of OFC 2006, PDP37.