Summary  The enormous growth of bandwidth needs and the constant revenues of the carriers at the same time require the introduction of packet technologies in transport networks. The extension of Ethernet with carrier-grade features is a promising approach to provide necessary characteristics like scalability, quality of service, and operations, administration, and maintenance. Interesting technological aspects are multi-layer and multi-domain operation of these packet transport networks and their control via Generalized MPLS. A possible solution can be seen in a function split where IP stays the convergence platform for applications and services and Ethernet becomes the convergence platform for transport.

1 Introduction  Today’s traffic predictions for the Internet range between growth rates of 40–100% per year. This is supported by observations at major internet exchange points (IXPs) as shown in Fig. 1 for the locations Frankfurt (DE-CIX), Stockholm (netnod), London (LINX), and Amsterdam (AMX).

The main reason for these enormous growth rates are the upcoming Web 2.0 driven end user applications like peer-to-peer and video file sharing.

All this growth has one thing in common: The expectation of the Internet users that most information, services, and applications offered in the web come for free or are financed by advertising. It is commonly understood that the resulting steep decline in revenue per bandwidth unit can only be absorbed via introducing new technologies into the transport networks that serve as the common basis for fixed and mobile networks and also the Internet. The introduction of packet technologies in transport networks is expected to provide the desired cost advantages that allow the carriers to keep pace with the bandwidth growth.

In this paper we describe the changes needed in backbone/transport networks to reduce total cost of ownership. Section 2 shows the increasing importance of the Ethernet protocol for this purpose and the carrier-grade features needed. Section 3 then explains multi-layer optimization between Ethernet and the optical layer and Section 4 details on control plane and management.
plane issues for future Ethernet carrier network. Finally, Section 5 discusses the current and future relation between Ethernet and IP layer.

2 Ethernet

In the emerging packet transport paradigm, especially the Ethernet protocol is considered as the enabler of more cost-efficient backbone networks, as it is characterized by simplicity, flexibility, inter-operability and, low cost.

2.1 Ethernet Economics

Being a permanent success story since nearly two decades, scale effects have lowered the cost for enterprise Ethernet equipment down to very promising levels. Already in 2002, 300 million switched Ethernet ports were installed. Nowadays, nearly all major data traffic streams originate from Ethernet interfaces [1].

Also in carrier equipment, Ethernet starts already to show now huge cost advantages: For example, optical 10 Gbit/s interface cards for IP routers with Ethernet interfaces they are around ten times less expensive than those having classical SONET/SDH interfaces [2].

Also important aspects are Ethernet end-to-end virtual private network (VPN) services (E-LAN and E-Line) that increasingly become an important source of revenue for the network operators – also due to the ubiquity of Ethernet interfaces in customers’ equipment.

The introduction of carrier-grade features into Ethernet is only the logical consequence – combining inexpensive hardware with intelligent software to provide necessary features like scalability, quality of service including reliability and availability, and operations, administration, and maintenance [3].

2.2 Ethernet Services and Ethernet Transport

Offered services have to be differentiated from enabling technologies of the underlying transport network. However, for both services and transport technology Ethernet has become a solution of choice in the last years.

Ethernet Services

Ethernet services can be divided into two main groups. E-Line services are often used for connecting headquarters to data-centers or other offices in a topology formed of point-to-point links via supplying layer 2 point-to-point virtual circuits to connect one customer edge device (CE) to exactly one other remote CE.

E-LAN services, which are often also referred to as Virtual Private LAN Services (VPLS [4]), provide multipoint to multipoint connectivity. A LAN is emulated between the provider edge devices (PEs) and the network core appears as one large switch. From the CE perspective the virtual backbone is the set of PE bridges that evaluate layer 2 information to forward the packet.

In both cases two E-service variants can be differentiated: Port-based services in that the user network interfaces (UNIs) do not support multiplexing and tag-based services in that UNIs support multiplexing and multiple connections per port and bandwidth.

Ethernet Transport

Generally, there are different ways of enabling layer-2 services and connecting CEs: Today, a layer-2.5 packet-forwarding technique (described in IETF-RFC4664) based on IP/MPLS (Multi-Protocol Label Switching) is most commonly used. However, other technologies have recently been developed that provide end-to-end layer-2 services using native layer-2 concepts.

• Approaches based on MPLS concepts: Today, IP/MPLS technology is often used to enable multipoint connectivity between all edge devices by setting up a full mesh of transparent MPLS tunnels between all provider edge (PE) routers using control plane protocols like LDP or RSVP-TE and distributing interdomain reachability information via the Border Gateway Protocol (BGP). For every pair or group of PE devices that belong to the same VPN an additional MPLS path or multicast tree is created on top of that mesh. These tunnels, which are called ‘Pseudo Wires’ or ‘Martini Tunnels’ [4], are seen by customers as if all PE devices of a VPN would be connected by (single-hop) wires. Hence, a PE router can act as a virtual bridge – including learning the incoming client MAC addresses and broadcasting of packets with unknown addresses.

In order to reduce the complexity of having an integrated layer-3 control plane to establish layer-2 VPNs, adaptations
of MPLS are currently in development that are more suited for transport networks. Transport MPLS (T-MPLS, ITU-T G8110.1/Y.1370.1) defines operational, administration and maintenance characteristics for MPLS networks. Furthermore, joint standardization activities between IETF and ITU-T have been recently started to align the existing IP/MPLS standards with transport requirements (MPLS Transport Profile, MPLS-TP [5]).

- **Approaches based on Ethernet Transport concepts:** Provider Backbone Bridging (PBB, IEEE 802.1ag) and Provider Backbone Bridging with Traffic Engineering (PBB-TE, IEEE 802.1ah) are extensions of the classical Ethernet protocol to enable native Ethernet transport. Both technologies provide an additional transport hierarchy by encapsulating client Ethernet traffic by another Ethernet header to facilitate a transparent transport of client traffic. While PBB still relies on connection-less forwarding based on tree-structures (spanning tree protocol), PBB-TE establishes pre-defined connection-oriented tunnels between edge switches of a network. With this concept, similar traffic engineering and resilience mechanisms known from MPLS can be achieved. However, instead of adding a MPLS label to the header packet encapsulation is performed at the edge switches and an additional MAC header is stacked on top of the client traffic. The tunnels are encoded by the destination MAC address of the backbone egress switch (B-DA) as well as by a 12 bit VLAN-tag (backbone tag, B-VID). Hence, PBB-TE forms a topology of B-DA rooted trees and an independent sink-tree is configured for each <B-DA, B-VID> pair. Similar to MPLS-TP, PBB-TE replaces the Layer-3-based control plane with either a management plane or an external control plane. Since all backbone destination addresses are managed by the provider even a centralized configuration of backbone switches can be performed relatively easily.

2.3 Carrier-Grade Requirements for Ethernet

The Metro Ethernet Forum (MEF) refers to Carrier Ethernet as a ubiquitous, standardized, carrier-class service characterized by five attributes that distinguish it from LAN-based Ethernet [6]:

- **Standardized services:** Deployment of Carrier Ethernet requires standardization of provider equipment, and a standardized set of building blocks for definition of Carrier Ethernet services. A defined choice regarding service type (E-Line/E-LAN) and a finite set of service attributes – related to physical interface, traffic parameters, performance parameters, class of service, service frame delivery, VLAN tag support, service multiplexing, bundling, security filters – enables service offerings individually tailored to the customer needs, while simplifying their implementation on different infrastructures.

- **Quality of service (QoS):** Additionally to the aforementioned standardization activities, the enforcing and verifying of SLA conformance requires additional functionality for traffic management (admission control, policing/shaping and especially service differentiation) and performance management, as non-conformance is typically penalized.

- **Service Management:** Standardized vendor-independent implementations of operations, administration, and maintenance (OAM) features providing fault, configuration, accounting, performance and security (FCAPS) management functions, e.g., for verification of service connectivity or proper operation of links and nodes, are a key requirement for dynamic, flexible and rapid service provisioning and recovery, as well as for traffic engineering.

- **Reliability:** For rapid detection, isolation, and recovery from node, link or service failures with virtually no user impact, proactive (i.e., protection-based) fault management functions are needed to maintain acceptable end-to-end service availabilities. Restoration, a more cost-efficient but slower variant, demands reactive (i.e., rerouting-based) fault management functions. These approaches complement each other in achieving service recovery also enabling service providers to perform service differentiation in terms of availability for generating additional revenues.

- **Scalability:** Besides the ability to scale in terms of service bit-rate profile (up to the physical bit-rate) and the ability to scale the network and related operational processes with increasing number of users, scalability is also needed in terms of geographical reach. Delivery must not be limited to LAN-based Ethernet networks, but can also happen via other transport technologies (multi-layer) and service provider networks (multi-domain).

3 Multi-Layer Networking

Figure 2 shows as an example (other options are, e.g., IP over WDM) the realization of Ethernet over WDM with optionally IP/MPLS running over it while. Cost minimization is achieved since traffic is forwarded on the layer which is most suited for it. For example, highly aggregated traffic can be carried by optical transport services using, e.g., Reconfigurable Optical Add/Drop Multiplexer (ROADM) technology which can switch transparently and thus avoid optical/electrical/optical (OEO) conversion. Less aggregated traffic flows and native Ethernet services can be carried by on the Carrier Ethernet Transport (CET) layer and IP/MPLS is responsible for low aggregated traffic flows.

The ability to select a layer for a service provides further ad-

---

Brought to you by | Universitätsbibliothek Stuttgart
Authenticated
Download Date | 5/17/18 11:21 AM
Choosing the layer enables it to make best use of resources in terms of network element selection and economy-of-scale, thus, it reduces the likelihood of service blocking and it can increase utilization.

- Multi-layer traffic engineering can (re-)optimize resources with high degree of freedom, considering metrics for the complete network.
- If services are differentiated by their time to recover from failures, the best supporting layer for the service can be selected.
- Multi-layer architectures are able to realize short service establishment times.
- As higher layers can directly request resources from lower layers (e.g., using control and management plane technology), a fast establishing process can be achieved and cumbersome inter-layer service requests are avoided.

Multi-layer nodes integrate functionalities and entities from different layers and the use of shared hardware (shelves, power supplies, etc.) reduces cost. As multi-layer nodes have reduced floor space requirements, they also lower Operational Expenditures (OPEX). Moreover, integration simplifies multi-layer management and control, since inter-layer interfaces can be realized within a node. Finally, multi-layer nodes can be deployed spatially different, locating functions where needed, e.g., IP/Ethernet/WDM nodes can be placed at the edge and simpler Ethernet/WDM nodes can be placed at the inner part of the network.

Many of the multi-layer networking benefits are obtained by conscious network planning supported by network planning and configuration tools. First and foremost, network planning optimizes the Total Cost of Ownership (TCO) and can be applied to both greenfield scenarios and extension scenarios. Furthermore, it can ensure service level agreements (SLAs) by considering parameters such as availability and delay. Multi-layer routing can, for example, decide the route through multiple elements of different layers such that a given availability is met as specified by an SLA. Another example is that delay restrictions can be properly considered, e.g., for IP-over-WDM traffic, since routing is aware of both the IP hops (contributing queuing delay) and the physical length (contributing propagation delay). Multi-layer network planning can also assign traffic to a layer, if the layers differ in their wattage for processing the traffic, to achieve minimization of overall power consumption.

**4 Control and Management of Ethernet Networks**

4.1 GMPLS, GELS

Next to centralized network management solutions and solutions based on external provisioning tools, control planes will be introduced more and more in transport networks. Control planes evolve from the current state of single and independent control planes, such as implemented in MPLS, to unified multi-layer and multi-domain control planes.

The first step of the control plane evolution is the merging of the MPLS concept with the concepts of different transport networks so that a unified control plane can be applied to any transport technology. The protocol suite of choice for this control plane implementation is called GMPLS [7] and is currently in standardization, implementation, and optimization phase.

The next step in the control plane evolution will be the usage of GMPLS for Ethernet Label Switching (GELS) which is envisaged as an important issue to be solved in the future as Ethernet has to be controlled and managed from end-to-end together with all the traffic granularities present in the network, ranging from switching packets up to switching fibers. The layer-2 label switching concept is thus mixing the advantages of a well-known Ethernet data plane and the capabilities of the GMPLS control plane.

The last evolution step for the future will be the specification of one unified multi-layer and multi-domain control plane independent from the network layering and the transport technology. This control plane will provide an interface for application and/or service overlays to configure the high-speed packet transport network (see Fig. 3). It will automatically optimize the signaling and routing process within the future multi-layer transport network and will realize a multi-vendor and multi-domain interoperability (User Network Interface (UNI)/Network Network Interface (NNI)). This multi-layer and multi-domain interoperability requires an interaction of control plane, management plane and a computation element for computationally intensive algorithms. For example, the path computation can be done by a centralized or distributed path.

![Routing options in a multi-layer network.](image-url)
computational element (PCE) approach. This new control plane architecture will provide a multi-layer and multi-domain traffic engineering that guarantees an optimized routing, network utilization and a reduced demand blocking. The interoperability will enable an end-to-end quality of service provisioning. The reliability will be improved by automated multi-layer protection and restoration mechanisms (self-healing network). The multi-layer optimization requires a high visibility of the network equipment of the involved layers. This visibility facilitates an optimized inventory management. The automation of the control plane enables a fast service provisioning. The major objective of all these topics is the very important minimization of TCO in metro and core networks.

4.2 Standardized Interfaces for OAM and Control

Different OAM protocols are in the standardization to enable an end-to-end service management for Carrier-Grade Ethernet. Three main protocols in this area are Service OAM (IEEE 802.1ag), Link OAM (IEEE 802.3ah), and Ethernet Local Management Interface (E-LMI) [8]. Ethernet OAM provides the possibility to detect and repair failures and degradation in the network after they occur and help to increase the availability of transport networks. Service OAM allows monitoring of end-to-end Ethernet services instances, while Link OAM allows the monitoring of single Ethernet links. Four different messages are defined in Service OAM to monitor and debug Ethernet networks. These messages are continuity check messages, link trace messages, loopback messages, and alarm indication signal messages [9]. Another concept of Service OAM is the existence of maintenance domains and how different maintenance domains are related to each other. A maintenance domain is used to administrate a network and it is defined by maintenance intermediate points and maintenance endpoints. The relationship between different maintenance domains is that the maintenance endpoints of a domain are intermediate points for higher level domain. Different maintenance domains are defined to differentiate between the different scopes of management.

A further important aspect of Ethernet management is the monitoring of individual physical links. The Link OAM protocol allows putting a device in loopback mode to test a link if a critical event occurred. Four different functionalities are provided: discovery, link monitoring, remote failure indication, and remote loopback [9]. Discovery identifies the device at each end of the link. Link monitoring is used to detect link faults and to provide statistics on the number of frame errors. Remote failure indication allows communicating failures like loss of signal or power failure and remote loopback is responsible to put a peer in loopback mode to troubleshoot a link.

E-LMI has been defined by the Metro Ethernet Forum [8] and is a protocol in which the provider edge (PE) is communicating with the customer edge (CE) to enable automatic configuration of the CE. This reduces the configuration errors and simplifies the addition and deletion of Ethernet services and leads to lower operations costs for the service provider. To enable the described benefits the LMI provides the following information to the CE: Notification of the addition and deletion of an Ethernet Virtual Connection (EVC), notification of the availability state of a configured EVC, and the communication of UNI and EVC attributes.
5 Co-Existence of Packet-Transport and IP-Networks

The Internet will continue to drive the service and application innovations by offering a ubiquitous connectivity shared with a world wide application developer community. The future of the Internet layer is shaped by the advancements of the computing technology and its applications such as mobile computing and multimedia that assume always-on connectivity with ever increasing high bandwidth needs. The main asset of the Internet protocol suite is that it isolates the application technologies from the networking infrastructure and by so doing it enables communication over very distinct technology layers. It offers inter-domain naming and routing services with very limited traffic engineering capabilities. However, when that traffic volume increases, efficient multi-layer traffic engineering requires that end-to-end congestion and quality-of-service mechanisms would interact with the rest of the transport system.

The TCP flow controls that functions on the end-to-end basis need to be complemented with mechanisms that deal with congestion control inside the network and that are not limited to packet streams only. Explicit Congestion Notification [10] and Pre-Congestion Notification [11] are at the moment the most developed proposals for combining end-to-end flow control and network resource usage. With these mechanisms the congested network elements can notify the network edges and end devices that the traffic load is approaching a level where packets will be dropped. The network devices could then act on behalf of the end user devices to protect them from malicious unwanted traffic. Based on such mechanisms the network edge devices could adapt the network routes and resources to circumvent potential bottlenecks using the control mechanisms of the transport networks.

Site multi-homing that is currently used to offer additional redundancy and independence from the upstream providers has resulted into scalability concerns with the current BGP based routing system [12] in terms of breaking the natural topology based address aggregation. In order to achieve multi-homing same address prefixes need to be announced through multiple providers which is exhausting the routing tables. When the BGP based inter-domain routing system is used for IP-level traffic engineering by announcing more specific routes to hosts through preferred routes, further de-aggregation occurs. Resilient inter-domain routing could clearly benefit from path selection that takes into account traffic needs. By combining routing with the connection-oriented approaches desired carrier-grade end-to-end quality could be achieved. The inter-domain routing system should leverage more efficiently new interconnection protocols such as GMPLS and path computational element (PCE) concept instead of trying to duplicate the functionality. Combining these connection-oriented technologies with the IP-layer inter-domain routing would off-load the exhausted BGP routing tables and offer better performance, faster route restoration times and use of resources. But much remains to be developed to reach this goal!

Because most of the Internet usage is about accessing information, instead of accessing a specific server, it seems likely that a new information-centric abstraction layer is needed. This abstraction layer should support content or information level connectivity, instead of the current host interface centric model. Naturally this would impact how addressing would be arranged and how content would be delivered. Ideally the right content should be delivered from anywhere where it exists meeting the cost constraints of the request. Naturally, this has an impact how the inter-domain connectivity will be arranged and what kind of functionality would be part of the delivery system.

6 Conclusion

The enormous growth of bandwidth demand and the continuing cost pressure lead to the introduction of packet technologies in transport networks. As an example, the well known Ethernet protocol will be adapted to the needs of carrier backbone networks.

Via control plane technologies and multi-domain solutions the transport network is more and more evolving into a flexible and fully-featured packet network layer beneath the Internet as we know it today.

Cost-efficiency requirements, however, render the co-existence of two fully-fledged packet layers, which are operated concurrently, an unlikely solution. Therefore, the future functional split between the IP layer and the packet transport network becomes a crucial topic. Those solutions appear to be appropriate, where IP provides the convergence platform for applications and services, while technologies like Ethernet provide the convergence platform for data transport. Inter-domain connectivity remains to be a challenge for both approaches as it is impacted by business practices and policies.

References


Dr. Claus G. Gruber received his Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering and information technology in 2001 and 2007 from Technische Universität München (TUM), Germany. In 2006, he joined Siemens Corporate Technology and Nokia Siemens Networks in 2007 where he is project manager for Carrier-Grade Ethernet and packet networks. He is mainly interested in networking concepts, multi-layer traffic engineering and resilience, control plane, network management and configuration of ubiquitous communication technologies. Address: Nokia Siemens Networks GmbH & Co. KG, COO RTP NT NCT, St.-Martin-Str. 53, 81541 München, Germany, E-Mail: claus.gruber@nsn.com

Dr. Marco Hoffmann studied computer science at Munich University of Technology (TUM), Germany. He received his Dipl.-Inform. degree in 2001. In 2001 he joined a research and development group for GPRS base station Software at Siemens AG in Munich, Germany. He received the Dr. rer. nat. degree from TUM in 2005. In May 2004 he joined the research and development department for optical networks of Siemens AG, Corporate Technology, in Munich, Germany, where he became Senior Research Scientist. In October 2006, the department was transferred to Siemens Networks, which then has become part of the joint-venture Nokia Siemens Networks (NSN) in April 2007. At the NSN research part he became project manager and the responsible person for control plane topics (layer2–layer4). Address: Nokia Siemens Networks GmbH & Co. KG, COO RTP NT NCT, St.-Martin-Str. 53, 81541 München, Germany, E-Mail: marco.hoffmann@nsn.com

Dr. Dominic A. Schupke received his Dipl.-Ing. degree from RWTH Aachen in 1998 and his Dr.-Ing. degree from Technische Universität München (TUM), Germany in 2004. He is with Nokia Siemens Networks (NSN) in Munich, Germany and manages the Multi-Layer Networks and Resilience Group in the Research, Technology and Platforms Unit. Prior to NSN he was with Siemens and the Institute of Communication Networks at TUM. Address: Nokia Siemens Networks GmbH & Co. KG, COO RTP NT NCT, St.-Martin-Str. 53, 81541 München, Germany, E-Mail: dominic.schupke@nsn.com

Dr. Thomas Michaelis received his Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering and information technology in 1999 and 2007 from Technische Universität München (TUM), Germany. In late 1999, he joined the Institute of Communication Networks (LKN) at TUM for conducting research on physical impairment-aware routing in transparent DWDM networks with sparse regeneration. In early 2005, he joined Siemens Corporate Technology (now Nokia Siemens Networks), Munich, Germany, where he is managing a BMBF-funded research project on 100G transmission and networking (100GET-E3). Address: Nokia Siemens Networks GmbH & Co. KG, COO RTP NT NCT, St.-Martin-Str. 53, 81541 München, Germany, E-Mail: thomas.fischer@nsn.com

Lic. Tech. Hannu Flinck holds a M.Sc. (1986) and a Lic. of Tech. degree (1993) in Computer Science and Communication Systems from the Helsinki University of Technology. While in Nokia he research topics covered various communication and networking technologies with main focus on IP technologies. In 2007 he joined Nokia Siemens Networks where he conducts research related to the Future Internet theme. Address: Nokia Siemens Networks, COO RTP NT NCT, Linnoitustie 6, 02600 Espoo, Finland, E-Mail: hannu.flinck@nsn.com

Dr. Andreas Kirstädter studied electrical and industrial engineering and holds a PhD in communication networks. In 1997, he joined Siemens Corporate Technology in Munich and 2007 Nokia Siemens Networks where he is responsible for the "Network Control-Plane and Transport" department. Address: Nokia Siemens Networks GmbH & Co. KG, COO RTP NT NCT, St.-Martin-Str. 53, 81541 München, Germany, E-Mail: andreas.kirstaedter@nsn.com

Dipl.-Ing. Christian Merkle received the Dipl.-Ing. degree in Electrical Engineering from Munich University of Technology (TUM), Germany in 2006. Since 2006 he is a member of the research and teaching staff at the Institute for Communication Networks at Munich University of Technology, where he works on research topics in the area of next-generation packet networks like Carrier-Grade Ethernet.

Address: Technische Universität München, Lehrstuhl für Kommunikationstechnik, Arcisstr. 21, 80333 München, Germany, E-Mail: christian.merkle@tum.de