ASTN enhanced Ethernet Transport Services

Dieter Stoll, Hermann Dentler, Gabor Greif, Peter Hessler, Ralph Schlenk
Lucent Technologies
Thurn-und-Taxis-Strasse 10
D-90411 Nürnberg
{dieterstoll,hdentler,ggreif,phessler,schlenkr}@lucent.com

Abstract
This work addresses one inherent dichotomy between the desire for efficient assignment of transport resources on the one hand and multipoint data connectivity on the other which comes with Ethernet transport over the WAN. In a lab-demonstrator a new concept is evaluated which targets at controlling transport bandwidth with an ASTN, optimized for Ethernet in the WAN. It is shown by means of several example cases how overcapacity can be reduced in principle and how actual data flow topologies and their changes can be supported by the ASTN.

Introduction

Ethernet as the ubiquitous technology in today's landscape of LAN technologies has also been extended into the WAN over the past years. It has evolved to become a viable means to replace existing frame relay and ATM technology in private line applications and the required transport technologies like GFP and LCAS are widely deployed in modern multi service SONET/SDH transport equipment. The focus for Ethernet deployment in the transport network has thus moved on to switched Ethernet service offerings, UNI, and carrier grade service definitions.

Regarding switched Ethernet services in the wide area network a disadvantage arises from an inherent dichotomy. On the one hand there is the desire to connect clients to a transport network in a most economical fashion, i.e. with as little transmission bandwidth and equipment overhead as possible and satisfying as many customers as possible, implicitly or explicitly utilizing statistical multiplexing gains. On the other hand there is the Ethernet intrinsic constraint that all frames must be transported along the leaves of a topological tree. Thus if the nodes are connected by a tree as a result of provisioning by the operator then data will in general not follow the shortest possible path through the network. This obviously causes cost for the operator in terms of additional transport connections. If the bridging nodes are connected by a mesh, which permits in principle optimized adequate data flows, then the inefficiency arises again, but this time from (the traffic agnostic) spanning tree which will disable part of the transport connections.

The purpose of this lab-demonstrator based study was to show that ASTN [G.841, G.8080] features may be used to overcome some of the aforementioned shortcomings of native Ethernet over SDH transport. The background and concepts applied will be outlined in the following section "Concept Outline". A description of the experimental setup and the limitations set by the demonstrator scope is given in the subsequent section "Scope of Experiments" which is followed by the discussion of some example "Tests and Results". Finally a "Summary" and an "Outlook" following these results is given.

Concept Outline

Today's Ethernet over SDH technologies provide point-to-point or point-to-multipoint connectivity on a static infrastructure. I.e. changes to the infrastructure are only foreseen on time

scales which are either long compared to client traffic changes and are e.g. dealt with by adjusting the size of LCAS groups, or they are so short that client traffic can be assumed unaffected, e.g. in case of an MS-SPRing protection switch event. In reality there is, however, also traffic variations to be observed on wide area connections of time scales of minutes to hours [1-3]. Referring to the topological constraints imposed by Ethernet which were explained in the introduction, overcapacity must be provisioned in the transport network to be able to accommodate these traffic variations in all possible topological scenarios, i.e. in all (by failure condition caused) active topologies. Since this issue cannot be solved entirely, this study focused on possibilities to minimize the over-provisioning needs.
To that extent the transport network underlying the WAN Ethernet is enriched by ASTN capabilities. The ASTN control plane is used to adapt bandwidth needs to real traffic demands and thus provides traffic aware transport connectivity. The goal is to provide traffic awareness for two major cases

- to size transport links on the active topology according the actually realized traffic load
- to adjust the transport bandwidth to the active topology, e.g. after a spanning tree reconfiguration

and thus follows the spirit of the work of [4] although the goals differ in the respective details.
Since these aspects cannot be realized over UNI, the proposed methodology uses internal traffic counters to measure traffic load and derives the necessary activities regarding resizing of LCAS groups from a comparison with the provisioned bandwidth. Triggered by the results from the measurements, bandwidth adjustments are executed over the ASTN control plane.

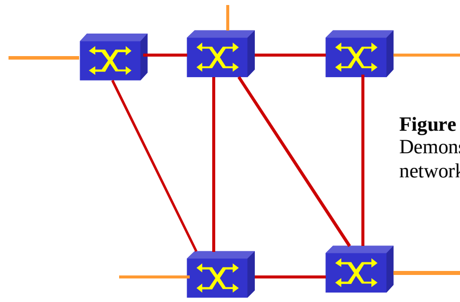


Figure 1: Demonstrator network

Scope of Experiments

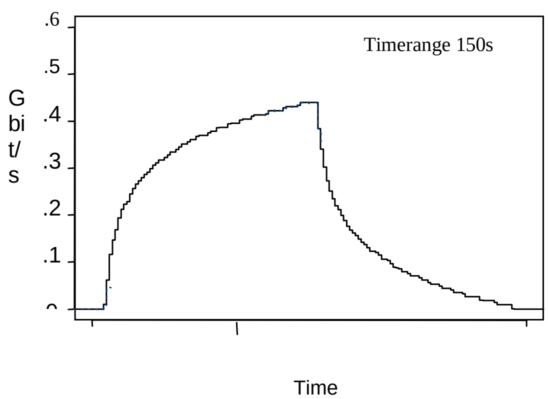
All experiments have been executed using LambdaUnite Multi-Service-Switch™ nodes. The network was set up as shown in figure 1 above. It consisted of five ASTN capable nodes interconnected via STM16 and STM64 interfaces on the WAN side and connected to traffic generators via 1GbE LAN interfaces. All nodes acted as layer 2 Ethernet switches.
Internal performance monitoring counters have been used to measure data traffic on LAN and WAN ports, data collection took place in 1s periods. In addition integration over counting periods, NE internal communication delays and signalling delays to set up several paths within the network had to be considered, so that in this demonstrator reaction times of the network to bandwidth changes were constrained to an order of magnitude of several seconds. In addition there was a significant impact from the adaptation strategy and the available raw performance counters in HW devices, which will be further discussed in the summary section below.

Tests and Results

The first series of tests considered autonomous bandwidth adaptations to variable demands. Three cases are shown in the figures below:

- a point-to-point connection over several links (figures 2a/b)
- two flows in a meshed network which have to share the WAN bandwidth which may become oversubscribed (figure 3a/b)
- a single flow through the meshed network affected by a link failure and subsequent spanning tree reconfiguration (figure 4).

Figure 2a: Traffic pattern



Figures 2a/b show a logarithmically shaped traffic pattern of incoming traffic (2a) the traffic loss experienced within the network (2b). The figure shows how the total traffic loss results from superposition of the losses in subsequent nodes along the same data path. Since all nodes react independently of each other, i.e. collect their own measurement data and adjust bandwidth according to these local views, the adaptation delay of all nodes adds up to a total loss seen by the client. Reaction times are larger than a few seconds in this example, a consequence of using a slow adaptation algorithm in this particular case.

Figures 3a/b summarize the results from superimposing several data streams over the same transport switch capacity in the WAN. In figure 3a data measurements (LAN-RX, WAN-TX and Loss) and bandwidth provisioning (WAN-BW) at three nodes is shown. In this test case WAN transport capacity was large enough to accommodate both data flows. The expected and also observed effect is lossless transmission and transport capacity adaptation to the aggregate data stream. The small peaks for data loss, which can be seen nonetheless, result from latency in the adaptation at points in time with extreme traffic pattern changes. Figure 3b shows the equivalent result measured externally at the traffic generator, when the transport capacity is exceeded, i.e. not all traffic can be transported over the available WAN bandwidth. Here, traffic loss is expected, but the less obvious result is that traffic loss does not prefer one or the other data stream. This is a consequence of the fact that only two streams are merged and would not be valid if more streams had been added topologically after each other onto the same switching infrastructure. This is of course a generic problem known from Ethernet LANs and its removal would require other techniques such as RPR or VPLS. Beside the expected large traffic loss peaks there again are some small traffic losses observed which are caused by implementation inaccuracies of this demonstrator SW. The implication of these results is that – utilizing the existence of slow variations in data traffic, as discussed in the introduction, the WAN bandwidth can be time shared between

Figure 2b: Delayed adaptation in a chain of 3 nodes

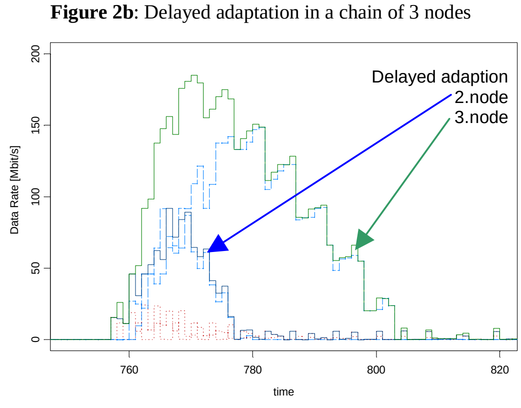
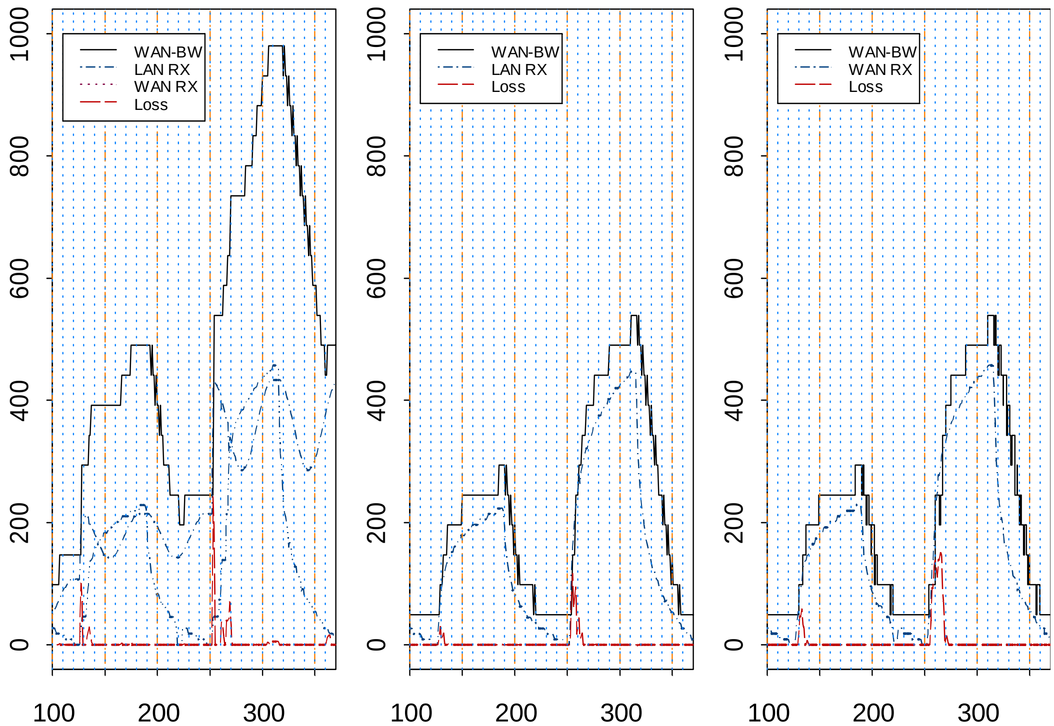


Figure 3a: Superposition of traffic streams with sufficient WAN bandwidth



different data streams, permitting for real physical bandwidth reduction.

Figure 4 finally shows the adaptation of the transport capacity after a spanning tree re-configuration. For simplicity only one node is shown and furthermore the WAN bandwidth was kept at fixed size on one WAN port. This is the reason that WAN-bandwidth adaptation (WAN-BW) needs to be plotted only for the ASTN managed port. In figure 4, two spanning tree re-configurations are

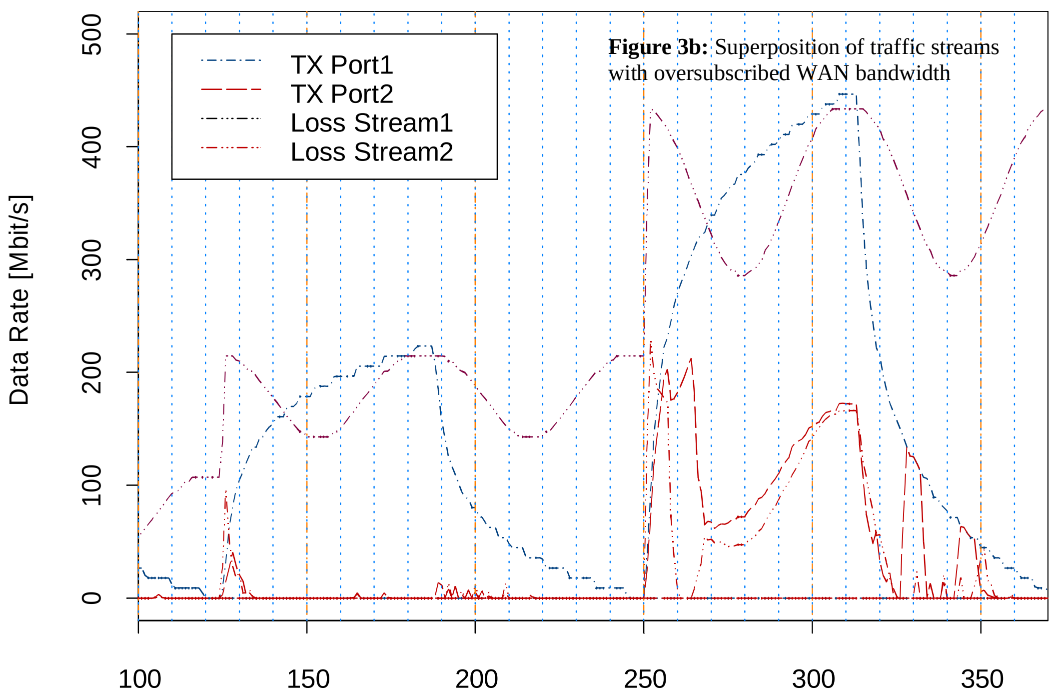


Figure 3b: Superposition of traffic streams with oversubscribed WAN bandwidth

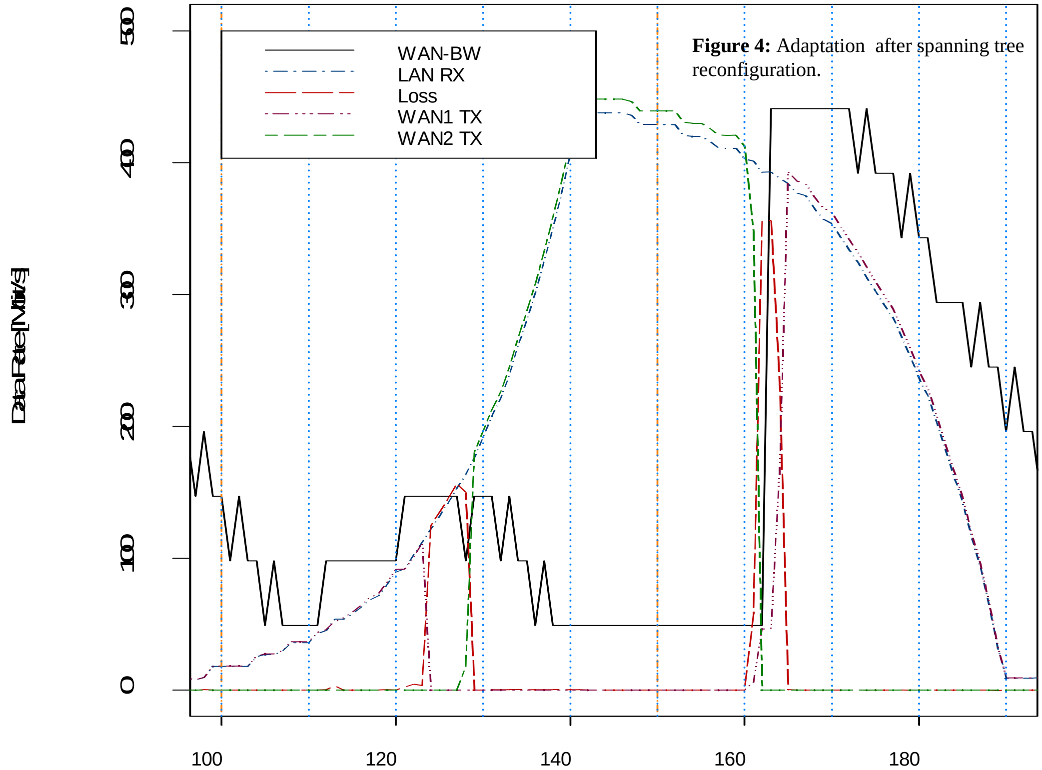


Figure 4: Adaptation after spanning tree reconfiguration.

shown, one from the ASTN managed port to the non-managed port at time 130, the opposite switch direction happens at time 160. The latter direction is the interesting one and it can clearly be seen that the WAN bandwidth is adjusted from its minimum value, which is 1 STS-1, to the value required to accommodate the entire client traffic. This shows the desired behavior of having minimized overcapacity on spanning tree disabled links and fast adaptation of low capacity links after change of the active topology. Traffic loss is dictated by the time it takes to establish all required SDH paths and adding them to the LCAS group (8 paths to be added in this example). We thus demonstrated that the traffic agnostic spanning tree can be adapted to actual traffic demands after topology changes as a result of e.g. network failures.

Summary

Using the fast autonomous path provisioning of the ASTN we observed lossless transport bandwidth adaptations for moderate traffic pattern changes suited to deal with typical traffic variations on wide area network connections [1-3]. We investigated among others scenarios like point-to-point private line transport, E-LAN situations with network failures, and competing traffic sources with under-dimensioned transport bandwidth. Recapitulating the results of this demonstrator-based study we can say that the presumption was confirmed that ASTN can be used to reduce overcapacity in switched Ethernet over SDH WANs. Even with variable traffic patterns and time variations which are short compared to the patterns we target at [1-3], no major obstacle like instabilities or spanning tree convergence issues were observed. From this we may not deduce that there are no such scenarios possible since the scope was too limited and the sophistication of the implementation did not surpass the demonstrator level.
A particular topic in this context is limits to accuracy and capabilities of a control algorithm. It turned out that of

prime importance are adequate HW counters, which permit to deduce e.g. the exact required transport bandwidth. This is a non-trivial issue due to dependencies on the payload and the mapping schemes used.

Outlook

We conclude from our studies that the goal of Ethernet services provided over a resource efficient transport network need not be abandoned. By enriching mere Ethernet transport with the intelligence of an ASTN there is the possibility to maintain the simplicity of Ethernet while at the same time minimizing the required network overbuild.

Regarding the current attempts for Ethernet transport in metro networks, and the related strive for QoS, this work cannot be considered complete though. It had not been investigated yet what the benefit of such an approach in case of service differentiation into different quality classes would be. Open are questions regarding its extension into the metro space with its own traffic patterns and onto other emerging technologies like VPLS. Last not least, a quantitative study with regard to the network resource gains would be in place to prove the business value of this approach.

In spite of these current deficiencies, we see this work as a step towards a more converged network in the future.

Acknowledgement

This work has been funded in parts by the German Ministry for Research and Education (BMBF Grant SmartSyKon 01BP262)

References

[G.841] ITU-T standard G.841

[G.8080] ITU-T standard G.8080

[1] K. Thompson, G.J. Miller, R. Wilder, „Wide Area Internet Traffic Patterns and Characteristics“, IEEE Network (1997)

[2] J. Cao, W. S. Cleveland, D. Lin, D. X. Sun, “The Effect of Statistical Multiplexing on the Long-Range Dependence of Internet Packet Traffic”, Bell-Labs Technical Report (2004)

[3] Z.-L. Fhang et al., “Impact of Aggregation on Scaling Behavior of Internet Backbone Traffic”, Sprint Report TR02-ATL-020157 (2002)

[4] S.D.Maesschalck et al., “Intelligent Optical Networking for Multilayer Survivability”, IEEE Comm. Mag. (2002) 42