Towards Connecting Base Stations over Metro Gigabit Ethernets

Rainer Baumann and Ulrich Fiedler
Computer Engineering and Networks Laboratory
ETH Zurich, Switzerland
baumann,fiedler@tik.ee.ethz.ch

Abstract

Deployment of Gigabit Ethernets in metropolitan areas creates new opportunities to save costs by converging data and telephone services. The primary question of our study is whether emerging metropolitan networks can meet the QoS requirements necessary to connect GSM and UMTS base stations. These requirements on delay, jitter, and loss are significantly more stringent than the ones for VoIP and have to be met at the presence of bursty cross traffic. Therefore, we have probed ETH's campus network which spans the metropolitan area of Zurich with traffic comparable to encapsulated E1 traffic from base stations and have measured perceived QoS. Our findings show that the campus network generally meets the QoS requirements. However, the perceived QoS degrades with increasing network utilization.

To further investigate the impact of the configuration of Ethernet switches on perceived QoS, we have conducted a simulation study. The results show that the perceived QoS is better when switch buffers are limited to small sizes of 1 MB compared to setups where the number of frames in the buffer is limited. From these results we infer that metropolitan Gigabit Ethernets are well suited for connecting GSM and UMTS base stations.

1 Introduction

The deployment of Gigabit Ethernets in metropolitan areas creates new opportunities to save costs by converging data and telephone services. These opportunities include connecting GSM and UMTS base stations over Gigabit Ethernets. However, to make this happen, it has to be carefully investigated whether these emerging metropolitan networks can meet the necessary QoS requirements at the presence of bursty data traffic. These QoS requirements on delay, jitter, and loss are significantly more stringent than the ones for VoIP. Therefore, in this paper, we address the following problems: Do metropolitan networks have a potential to meet necessary QoS requirements for connecting GSM/UMTS base stations? Is there a relation between perceived QoS and network utilization? Do switch buffer limitations impact perceived QoS and if so, how?

We report a probing experiment in ETH's campus network which spans the metropolitan area of Zurich. We have probed this network with traffic comparable to encapsulated traffic from base stations and have measured the perceived QoS. Moreover, we report a simulation study on switch buffer limitations.

To conduct the probing experiment, we have implemented generator, sink, and archiver applications which run on top of RTAI [1], a hard real-time Linux. The generator generates probing traffic that emulates encapsulated TDM E1 traffic from UMTS and GSM bases stations and injects it into the network. The sink receives the probing traffic and the archiver logs the traffic. After probing, we have analyzed the archived logs. To establish the relation between the average network utilization in this measurement period and QoS perceived by probes, we focused our investigation on six different routes. We chose these routes to include links with low, medium, and high average network utilization, i.e. smaller than 1%, around 1%, and significantly above 1% utilization[1]. Each route has been probed for at least 15 days.

Our results give strong indications that ETH's campus network can meet QoS requirements for connecting base stations as stated by the Metro Ethernet Forum (MEF) [8]. The only temporary violation of the QoS requirements were on routes that included a link with high average network utilization. Moreover, we have found a general relation between perceived QoS of probing traffic and average network utilization. Perceived QoS decreases as network utilization increases.

To further investigate the impact of the configuration of buffer limitations in Ethernet switches on perceived QoS,

---

[1] We note (i) that [9] studies utilization rates of data networks and reports average network utilizations of 1% - 5% in corporate and other networks and (ii) that an average utilization of 5% is sufficient to lead to significant temporary frame loss.
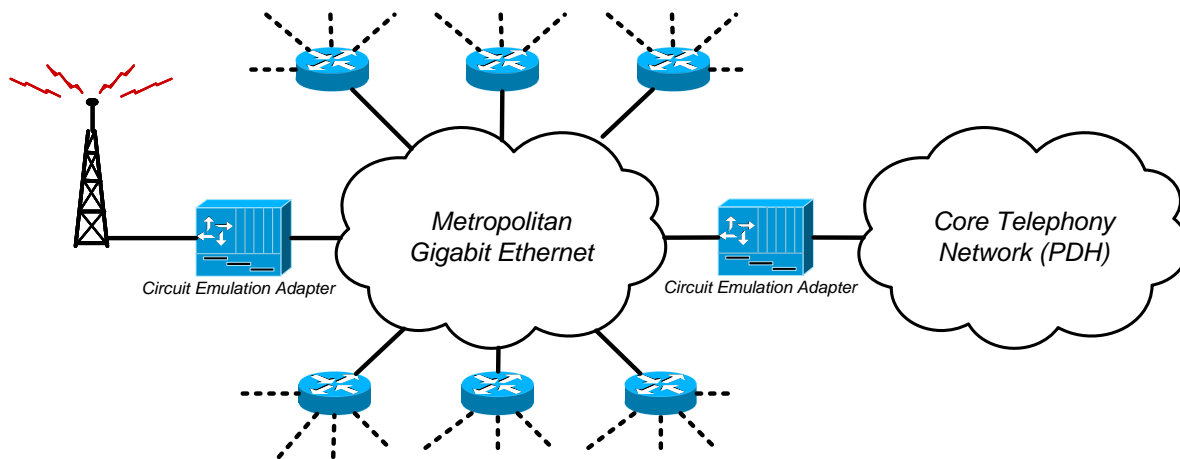
Figure 1. Scenario of our study

we have additionally conducted a simulation study. Our results indicate that limiting the switch buffer by size rather than by number of frames leads to better QoS for base station traffic. Specifically, we have observed that a switch buffer limitation of 1MB has led to a fulfillment of MEF's QoS requirements on delay and loss at an average data traffic network utilization of 1%. We explain this finding with the fact that large data frames are discarded when the buffer is close to full. At the same time the remaining buffer space is large enough to accommodate small frames as used to transport traffic from base stations.

The contributions of this paper can thus be summarized as follows:

- We show measurement results from a probing experiment that show that ETH's campus network generally meets the necessary QoS requirements for connecting GSM/UMTS base stations.

- The measurement results additionally indicate a general trend between perceived QoS and average network utilization.

- We show simulation results that indicate that limiting the switch buffer by size rather than by number of frames leads to better QoS for base station traffic.

Since data traffic burstiness is invariant [7] and average network utilization in metropolitan networks is typically low [9], we infer that these networks have indeed a potential to connect UMTS/GSM base stations.

The rest of the paper is structured as follows: Section 2 reviews the Metro Ethernet Forums QoS requirements for base station traffic in Metro Ethernets. Section 3 reports measurement results from the probing experiment in ETH's campus network. Section 4 reports the simulation study, before we conclude in Section 5.

2 QoS Requirements

The Metro Ethernet Forum (MEF) [8] is a consortium of ISPs and network equipment manufacturers that specify quality of service (QoS) requirements for circuit emulation over metropolitan Gigabit Ethernets. The main application for circuit emulation is connecting UMTS and GSM base stations. MEF's QoS requirements pertain (i) to a requirement for frame jitter, i.e. an upper bound to the time elapsed between receiving successive frames, (ii) to a requirement for the maximum delay and (iii) to a requirement for the frame loss rate as listed in table 1. These requirements relate to measurement periods of "30 hours or longer".

Metric	QoS requirement
Jitter	max. 10 ms
Delay	max. 25 ms
Loss rate	max. $8.75 \cdot 10^{-7}$

Table 1. Metro Forum's QoS requirements for base station traffic

Moreover, MEF specifies an additional composite requirement that the number of frames that violate any of the three above stated requirements must be bounded. This requirement is called the frame error rate, and is defined as

$$FER = JVR + DVR + LR. \qquad (1)$$

where
FER: frame error rate
JVR: jitter violation rate
DVR: delay violation rate
LR: loss rate
MEF specifies that this frame error rate is less than $8.75 \cdot$

10^{-7} which is equal to the loss rate. We conjecture that the reason behind this specification is that from an application perspective there is no difference between a late frame, a frame that violates the jitter requirement, and a lost frame.

3 Measurement study

For the measurement study we probe the operational campus network of ETH Zurich.
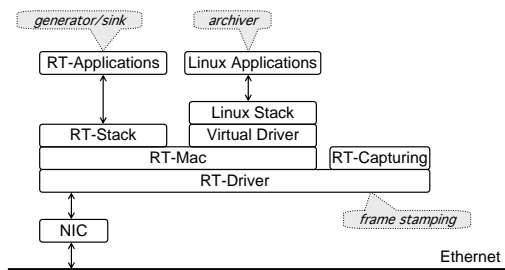
3.1 Measurement software

Figure 2. Protocol stack of measurement software

To perform the measurements we have implemented a probe generator, a probe sink, and an archiver which run on top of hard real-time Linux on commodity PCs. The protocol stack employed is depicted in figure 2. Our design conceptually follows Pasztor and Veitch [11] who show that such combinations of real-time software and commodity PCs can measure one-way network delays with an accuracy of $\pm 10~\mu$s. As a hard real-time Linux we have chosen RTAI [1]. RTAI is known for good real-time performance, and available under a GNU public license. For the real-time (RT) protocol stack, we have chosen RTnet [2], a Linux protocol stack modified for real-time networking. RTnet is known to offer atomic send and receive time stamps based on the time stamp counter processor register.

The probe generator is a real-time application that uses the RT-driver of the Ethernet card to stamp frames with the sending time before injecting the frames into the network. In addition to the send-time stamp, the generator also provides frame sequence numbers to detect frame duplicates, drops and reordering. The probe sink is a real-time application that stamps incoming frames with a receiver time stamp, and forwards the frames to the archiver. The archiver is a non real-time application that writes the sinked frame into a log file which resides on the workstation's hard disk. Once a day, this log file is transferred to a tape archive. Finally, an off-line task analyzes this log file for the quality of service associated with the archived probing traffic.

For the measurements reported in this paper, we have used Dell precision 340 workstations. These workstations are equipped with an Intel Pentium III processor with 1 GHz, 512 MB RAM and a 3COM 3C595 Fast EtherLink network interface card. The factor that limits the accuracy of one-way network delay measurements is the workstation's interrupt latency. This latency depends on the frequency of the PCI 8254 bus, which runs at 1193180 Hz and is as short as $0.8\mu s$. To calibrate the clock skew between sender and receiver side, we have directly connected the workstations with a crossover cable for 12 hours. The accuracy after calibration is listed in table 2. This accuracy is at least one order of magnitude smaller than the smallest minimal one-way network delay in our measurements.

Avg. delay	$1.30\mu s$
Std. dev. of delay	$0.51\mu s$
Max. delay	$13.55\mu s$
Sample size	$3.3 \cdot 10^8$

Table 2. Accuracy of one-way delay measurements

3.2 Probing measurement

We have probed the ETH Zurich network at different buildings throughout the city with traffic that emulates encapsulated E1 TDM traffic from base stations (8000 frames per second, 32 bytes payload per frame plus Ethernet, VLAN, IP, and UDP header). The rate of this traffic has been $5.248Mbit/s$. Measurement periods were between 15 and 29 days, i.e. significantly longer than the 30 hours minimum specified in the MEF's QoS requirements.

3.3 Topology

The topology of the probed network is depicted in figure 3. This topology follows the CISCO recommendations on using a three level layered hierarchy with a dual core [5]. The core routers are connected to an ISP which connects ETH to the outer world.

3.4 Routes of probing traffic

Probing traffic has been injected at the access layer. All routes probed went through a router in the core layer. Both forward and backward direction of the routes have been probed. Forward directions are listed in table 3. The different routes have been probed during disjunct observation intervals.
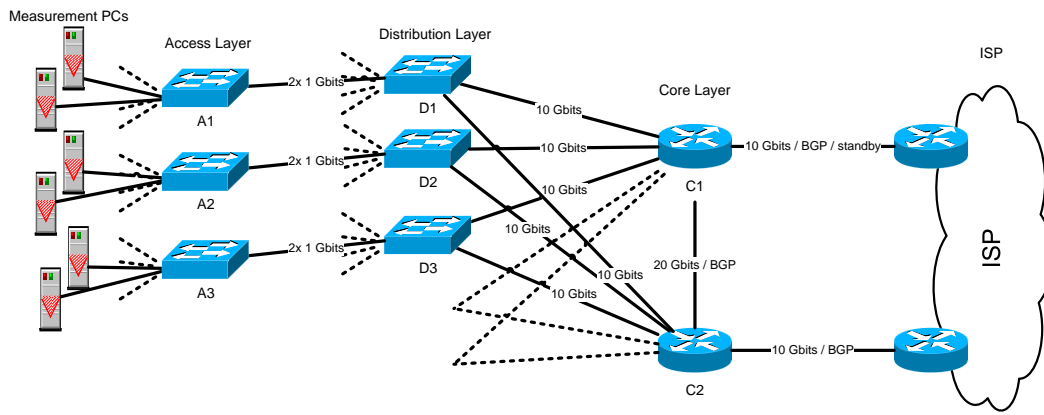
Figure 3. Topology of the ETH Zurich network (abstraction)

Probing	Route
A1 → A3	A1 → D1 → C2 → D3 → A3
A1 → A2	A1 → D1 → C2 → D2 → A2
A2 → A3	A2 → D2 → C2 → D3 → A3

Table 3. Routes of probing traffic (forward direction)

3.5 Measurement periods

Table 4 lists the paths of the probing measurements, the sample size in terms of number of frames and the duration.

Route	Sample size	Duration [days]
A1→A3	$1.04 \cdot 10^{10}$	15.05
A3→A1	$1.04 \cdot 10^{10}$	15.05
A1→A2	$2.03 \cdot 10^{10}$	29.37
A2→A1	$2.03 \cdot 10^{10}$	29.37
A2→A3	$1.19 \cdot 10^{10}$	17.22
A3→A2	$1.19 \cdot 10^{10}$	17.22

Table 4. Summary of probes conducted

3.6 Perceived QoS

Figure 4 depicts the quality of service which the probing traffic received over the full measurement duration. Specifically, this figure lists the loss rate, jitter violation rate, delay violation rate and frame error rate for each of the paths. The frame error rate limitation (FER limitation) bar depicts the Metro Ethernet Forum's overall quality of service requirements.

3.7 Utilization

To be able to make the relation between QoS perceived by probing traffic and network utilization, we have additionally measured link utilizations. The average utilization of the links on the routes that have been probed is listed in table 5 and table 6.

Link	Utilization Up	Utilization Down
A1↔D1	0.3%	0.5%
A2↔D2	0.2%	0.2%
A3↔D3	0.4%	1.0%

Table 5. Avg. util. for access-distribution layer links

Link	Utilization Up	Utilization Down
D1↔C2	0.2%	0.5%
D2↔C2	0.5%	1.5%
D3↔C2	3.1%	5.1%

Table 6. Avg. util. for distribution-core layer links

From these measurements, we infer that the high average utilization on the link from C2 to D3 is the underlying cause for the frame loss rates that violate MEF's requirements. To further investigate this violation, we have looked at the temporal distribution of frame losses. Figure 8 shows the temporal distribution of frame losses on the path that experienced a violation of MEF's requirements, i.e. $A2 \rightarrow A1$ and $A1 \rightarrow A3$, respectively. From this figure, we infer that the loss rate may be acceptable for several days before heavy data bursts cause a violation of MEF's frame loss rate
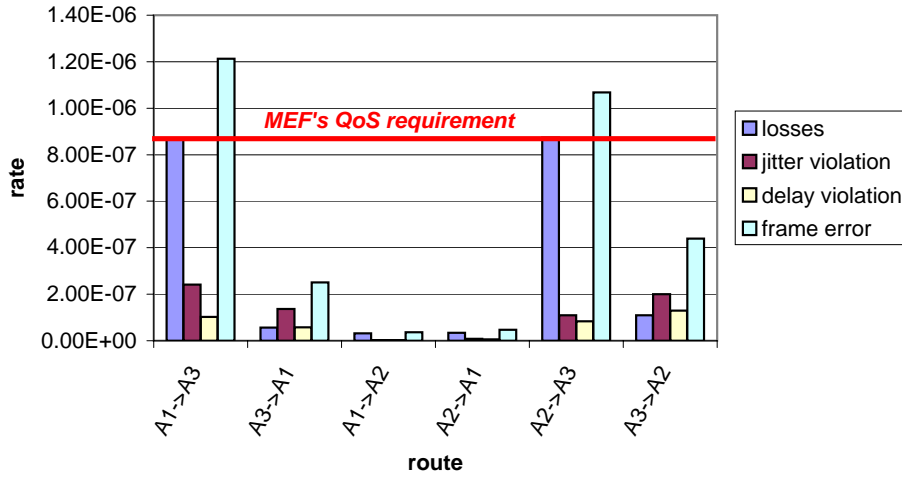
Figure 4. Quality of service perceived by probes over the full measurement period

requirement. This finding can be explained with the Noah effect [6] in long-range dependent data traffic. In addition to that, one can rank routes according to the frame error rate and compare this ranking to the ranking according to the network utilization on the bottleneck link. This comparison reveals a general correlating trend (see table 7).

Route	Frame error rate	Rank	Bottleneck utilization	Rank
A1→A3	$1.21 \cdot 10^{-6}$	1	5.1%	1
A3→A1	$2.50 \cdot 10^{-7}$	4	3.1%	3
A1→A2	$3.59 \cdot 10^{-8}$	6	1.5%	5
A2→A1	$4.68 \cdot 10^{-8}$	5	0.5%	6
A2→A3	$1.07 \cdot 10^{-6}$	2	5.1%	1
A3→A2	$4.39 \cdot 10^{-7}$	3	3.1%	3

Table 7. Ranking routes according to perceived frame error rate and according to bottleneck util.

4 Simulation study

In addition to the probing measurement, we have conducted a simulation study to investigate the relation between switch buffer limitations and QoS perceived by traffic from the base stations.

4.1 Simulation software

We use the OpNet modeler [10] as our simulation environment. OpNet modeler is a discrete event simulator that offers hierarchical network models with a focus on layer 2.

For our study, we have enhanced OpNet modeler for collecting detailed frame delay and loss statistics to measure the QoS that base station traffic has perceived.

4.2 Switch buffer limitations

We make two considerations for the limitation of switch buffer capacity.

- We refer to a widely used rule-of-thumb and

- we consider the impact of the MEF's QoS requirements

The rule-of-thumb states that the number of frames N the switch buffer should be able accommodate can be calculated with

$$N = RTT * C/MTU \qquad (2)$$

where RTT denotes the average round-trip time of flows through the link, C the data rate, and MTU the size of the maximum transfer unit. For the data traffic in metropolitan networks we estimate RTT = 10ms and C = 1GBits/s, hence

$$RTT * C = 1.25MB \qquad (3)$$

This estimation and a MTU of 1500 bytes lead to a switch buffer capacity of 874 frames.

Alternatively, we consider that the MEF's QoS requirement limits the frame delay to 25ms. This upper limit leads to maximum buffer size of 3.2MB.

We thus consider the following configurations for limiting buffer capacity.

Buffer space	Number of frames
1 MB	500
2 MB	1000

Table 8. Switch buffer limitations considered in our simulation study

4.3 Topology

Since each switch in a metropolitan network that transports base station traffic has to be configured in a way that the QoS requirements for delay and loss can be met, we focus on representing the most congested switch on this path. This leads to the topology depicted in figure 5 where all links are 1000BaseZX and hence have link bandwidth of 1Gbit/s.
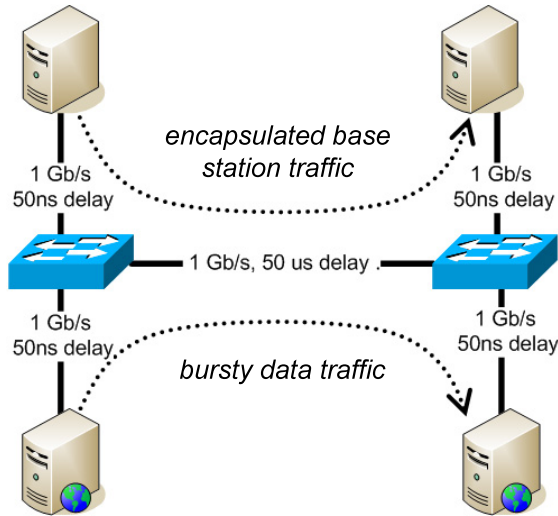
Figure 5. Simulation topology

4.4 Workload generation

Encapsulated base station traffic is generated with a constant bit rate of 5.248 Mbit/s. The bursty data traffic is generated with a set of superposed on/off sources. Sources either subsequently send frames (are in on state) or wait (are in off state). The number of frames send in on state is determined by the size of the associated transfer. Frames are sent at full line speed. The size of transfers is Pareto distributed. The parameters of this distribution are as listed in table 9. We refer to Feldmann et al. [3] for this choice. Moreover, we set a bound on transfer sizes at 4.1GB. This limit of 2^{32} bytes accounts for the fact that files stored in popular file systems are not arbitrarily large. The duration of the wait period after the last frame of the transfer is sent

is also Pareto distributed. The shape parameters of the wait time distribution is set equal to the shape parameter of the transfer size distribution. The average of this distribution is chosen so that the data traffic produced consumes 1% of the capacity of the 1 GBit/s network.

Parameter	Value
Number of On/Off sources	100'000
On-time: Transfer size Pareto-distributed	Avg. 12KB, $\alpha = 1.2$ Max. 4.1GB
Off-time Pareto-distributed	Avg. 1025 sec $\alpha = 1.2$

Table 9. Parameters for the On/Off model to generate data traffic

4.5 Impact of data traffic burstiness

Before presenting simulation results, we give a theoretical argument showing that an average data traffic network utilization of 1% is sufficient to cause switch buffer overflows at frequencies that can be expected to degrade the QoS of traffic from base stations beyond MEF's QoS requirements. In the next subsection, we then present the results of a simulation that accounts for effects that ameliorate the negative impact of frequent buffer overflows on base station traffic.

For the theoretical argument, we assume that we have only data traffic in our simulation and that access links have unlimited capacity. The frequency how often data frames overflow the switch buffer can then be estimated from the fraction of transfers that are larger than the switch buffer. This fraction p is given with

$$p = F^{-1}(x) = k(1-x)^{-\frac{1}{\alpha}} \qquad (4)$$

where F is the cumulative distribution function of transfer sizes and x is the size of the switch buffer[2]. The cumulative distribution of the transfer sizes in our simulation is given by

$$F(x) = 1 - \left(\frac{k}{x}\right)^\alpha \qquad for \quad x \epsilon [k, \infty) \qquad (5)$$

where the location parameter k can be inferred from the shape parameter $\alpha = 1.2$ and the average transfer size a=12kB employing the relation

$$k = \frac{\alpha - 1}{\alpha} a \qquad (6)$$

We now evaluate Equation 4 with setting x to the amount of available switch buffer. The expected frequency of buffer

---

[2] see [4] why quantiles converge fast and lead to meaningful expected values

overflows can than be inferred by assuming that the data traffic generation model can be expected to induce 106.6 transfers per second. This number of transfers is calculated from the average utilization of the link and the average transfer size (see table 9).

From this argument it follows that for a buffer size of 1 MB, 5.6% of the transfers cause buffer overflows since

$$1 - F^{-1}(1 \; MB) = 0.056 \tag{7}$$

Hence data traffic frames overflow the buffer every 16.7 seconds. Setting a bound on the size of transfers e.g. at 4.1GB has no major impact on these results since

$$1 - F^{-1}(4.1 \; GB) = 2.5 \cdot 10^{-8} \tag{8}$$

which is much smaller than 0.056. For a buffer size of 2 MB the corresponding values are 2.4% and 39.0 seconds.

Queue size	Corresponding transfer size percentile	Expected overflow frequency
1MB	99.944	every 16.7sec
2MB	99.976	every 39.0sec

Table 10. Time to overflow switch buffer with data traffic (at unbounded access link capacity)

These theoretical results indicate that data traffic is so bursty that an average data traffic network utilization of 1% is sufficient to cause significant buffer overflows. Moreover, we have checked traffic burstiness with variance time plots and have verified that the Hurst parameter is 0.9 for at least 5 magnitudes of aggregation on the time axis.

4.6 Simulation results

For each buffer limitation listed in table 8 we have seeded the random number generator with 16 different seeds and have run the simulation for 2 hours and 47 minutes simulation time. This leads to a total simulation time of more than 44 hours for each buffer limitation. The resulting QoS that the base station traffic perceived during this time is depicted in figure 6.

Detailed results for the different switch buffer limitations are as follows: Limiting the number of frames in the switch buffer as currently done in most operational networks leads to a violation of the QoS requirements. When limiting the switch buffer at 500 frames, only one out of 16 simulation runs meets MEF's QoS requirements for delay and loss. When limiting at 1000 frames, only three out of 16 simulation runs meet the QoS requirements for delay and loss.
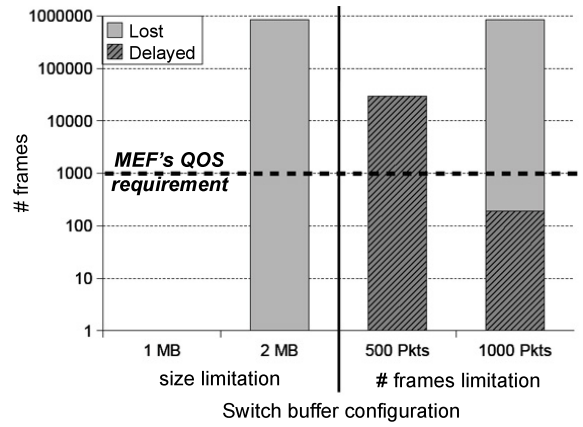

Figure 6. QoS of base station traffic for the switch buffer limitations listed in table 8

We explain this finding with the high burstiness of the data traffic in our simulation.

Limiting the switch buffer by size leads to a fulfillment of QoS requirements when the buffer size is not too large. In our simulation, we have found no frame loss and no delayed frames when the switch buffer has been limited to 1MB. However, limiting the buffer at a larger size such as 2MB leads to a significant number of delayed frames which violates the delay requirement. Only three out of 16 simulation runs meet the delay requirement when the switch buffer has been limited to 2MB although no frames have been lost. We explain the finding that no base station traffic frames are lost when the switch buffer is limited by size with a lockout phenomenon (see figure 7 for an illustration). Large data traffic frames are locked out at the end of switch buffer while there is remaining buffer space for small base station traffic frames.

We are aware that our simulations tend to over-estimate the effect of this lock-out phenomenon since (i) in real networks the fraction of data traffic frames that have a size close or equal to the MTU of 1500 bytes is smaller than in our simulation and (ii) depending on encapsulation and sampling, base station traffic frames can be larger than the 78 bytes in our simulation. However, limiting the switch buffer by size rather by number of frames remains to be the better choice when employing metropolitan networks to transport base station traffic.

5 Conclusion

In this paper, we have investigated whether metropolitan networks can potentially meet necessary QoS requirements for connecting GSM/UMTS base stations. We have conducted a measurement study and have probed the operational ETH campus network with traffic comparable to en-
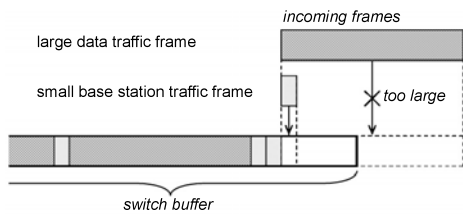
Figure 7. The lock-out phenomenon in the switch buffer.

capsulated E1 traffic from base stations for several weeks. We have then measured the QoS our probes perceived. From our results we infer that this network, which spans the metropolitan area of Zurich, can generally meet QoS requirements as specified by the Metro Ethernet Forum. The only temporary violations of the QoS requirements we have found were on routes that included links with high utilization. The results show a general correlating trend between QoS perceived by probing traffic and average network utilization.

We have additionally conducted a simulation study to investigate the effect of configurations in buffer limitations of Ethernet switches on perceived QoS. From our results we conclude that limiting the switch buffer by size rather than by number of frames leads to better QoS for base station traffic. Specifically, we have found that a switch buffer limitation to 1MB leads to a fulfillment of MEF's QoS requirements on delay and loss at an average data traffic network utilization of 1%.

Assuming that data traffic burstiness and network utilization are the major factors that determine the QoS perceived by base station traffic, we infer that metropolitan Gigabit Ethernets have indeed a potential to connect UMTS/GSM base stations.

Finally, we would like to stress that the network probing software used in our measurement study is available on request from the authors. This software will enable the research community, network service providers, and other service providers to assess the QoS of their networks.

6   Acknowledgment

References

[1] RTAI, Real-Time Application Interface. www.rtai.org.

[2] RTnet, Hard Real-Time Networking for Linux/RTAI. www.rts.uni-hannover.de/rtnet.

[3] A. Feldmann, A. C. Gilbert, P. Huang, and W. Willinger. Dynamics of IP traffic: A study of the role of variability and the impact of control. In *Proc. of SIGCOMM*, pages 301–313, 1999.

[4] U. Fiedler and B. Plattner. Using Latency Quantiles to Engineer QoS Guarantees for Web Services. In *Proc. of IWQoS*, 2003.

[5] Cisco Systems Inc. Hierarchical campus design-at-a-glance, 2004. available via www.cisco.com.

[6] J. Beran. *Statistics for Long-Memory Processes. Monographs on Statistics and Applied Probability*. Chapman and Hall, NY, 1994.

[7] K. Park and W. Willinger. *Self-Similar Network Traffic and Performance Evaluation*. Wiley-Interscience, NY, 2000.

[8] Metro Ethernet Forum. Technical Specification MEF3, Circuit Emulation Service Definitions, Framework and Requirements in Metro Ethernet Networks, April 2004.

[9] A. Odlyzko. Data networks are lightly utilized, and will stay that way. *Review of Network Economics*, 2(3), September 2003.

[10] OPNET Technologies Inc. OpNet Modeler 10.5A (PL3), 2004.

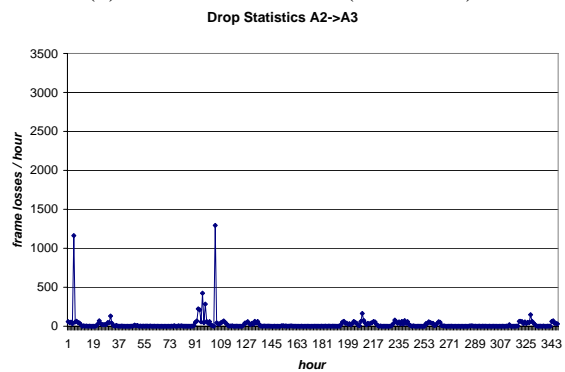[11] A. Pasztor and D. Veitch. A precision infrastructure for active probing. In *Proc. of PAM*, 2001.
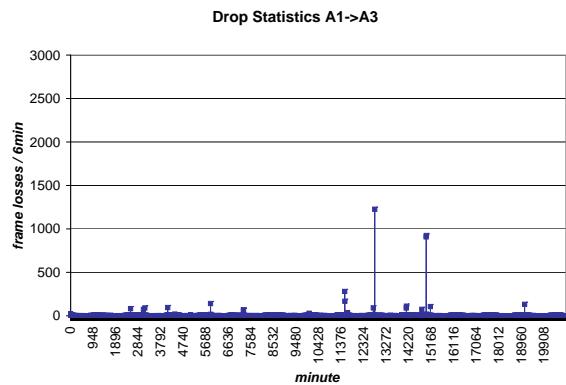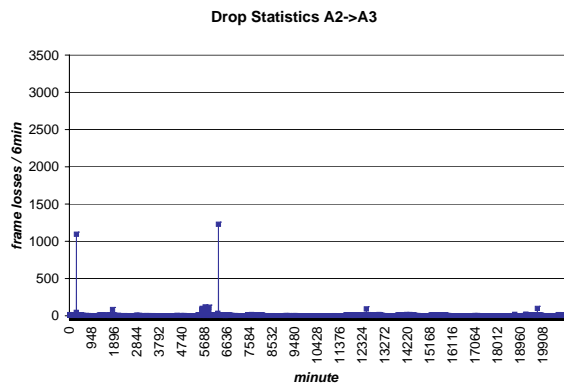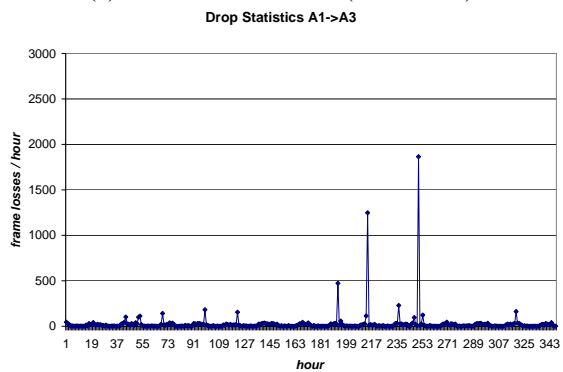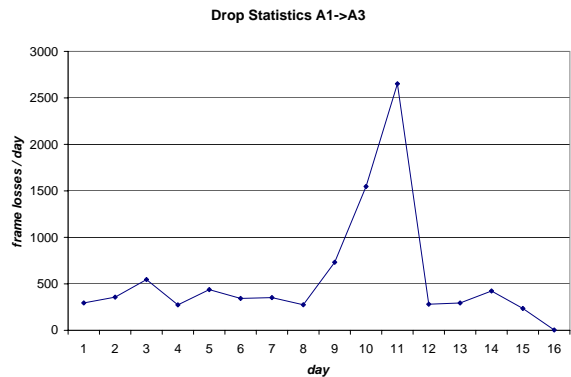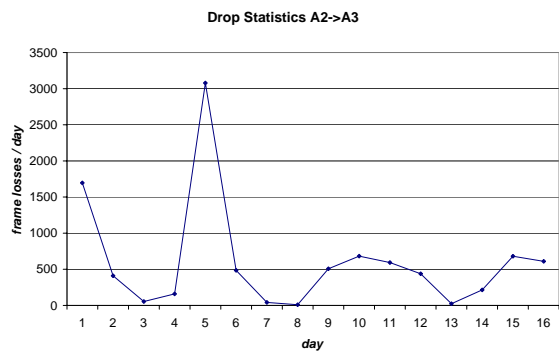
Figure 8. Temporal distribution of frame losses on path with high loss.