Service and Network Monitoring Support for Integrated End-to-End QoS Management

Mamadou Sidibé
CNRS-PRiSM Laboratory, University of Versailles
45 Avenue des Etats-Unis, 78035 Versailles, France
Mamadou.Sidibe@prism.uvsq.fr

Ahmed Mehaoua
CRIP5 Laboratory, University of Paris 5
45 Rue des Saints-Pères, 75270 Paris, France
Ahmed.mehaoua@math-info.univ-paris5.fr

Abstract—**Successful mass market provision of audio-visual services that would produce revenues for the content/service providers and network operators necessitates the use of an integrated management based on the end-to-end QoS over heterogeneous networks and terminals. This paper first briefly introduces the IST ENTHRONE II project[1] framework for QoS-enabled multimedia contents access, where a central concept is the integrated end-to-end management. Then, the focus of the paper is to present a novel service oriented monitoring system architecture as an essential part of this management. The proposed QoS monitoring system aims at providing monitoring information to service providers for providing quantified QoS-based services and their dynamic adaptation and to network operators for dynamic resource allocation allowing better usage of network resources. Performance evaluation conducted using a test-bed shows the responsiveness and the measurements accuracy of proposed end-to-end monitoring solution.**

Keywords: End-to-end monitoring; network measurements; NQoS; PQoS; SLA/SLS; cSLS/pSLS; service adaptation; integrated management

I. INTRODUCTION

Access to multimedia services over heterogeneous networks and terminals is of increasing market interest, while providing end-to-end Quality of Service (QoS) guarantees in multi-domain environment is still a challenge. This paper describes a service oriented monitoring system designed for use in multi-domain heterogeneous networking environment for the purpose of supporting cross-network audiovisual service offering. The presented monitoring system is that developed in the IST ENTHRONE II project for providing the means for service assurance and resource management. This QoS monitoring system is aimed at providing monitoring information to (1) service providers for providing quantified QoS-based services and its dynamic adaptation; (2) network operators for making provisioning decisions and allowing dynamic resource allocation for optimizing the usage of network resources; (3) service/network providers to verify whether the QoS performance guarantees committed in service agreements are in fact being met.

[1] This work is partially funded by European Commission ((ENTHRONE EU project IST 507637). See www.enthrone.org. The authors would like to thank ENTHRONE EU project partners for their inputs and valuable discussions.

In this work, we assume that the performance and traffic requirements of a requested service are described by a Service Level Agreement (SLA) and consequently its SLS (Service Level Specification) part [16]. Both the SLA and SLS are the basic elements in the operation of our proposed QoS-based monitoring system. SLAs are the means to formalize service level negotiations conducted between a customer (or Content Consumer) and a service provider for a specific class of service. The SLS is a subset of a SLA that denotes the technical characteristics of a service offered. These service technical characteristics refer to the provisioning aspects of the service, e.g. request, activation and delivery aspects from the network perspective. In this section, two types of SLSs (and consequently of SLAs) are distinguished: customer-to-provider SLSs (cSLSs), and provider-to-provider SLSs (pSLSs) [11], [14]. In ENTHRONE, the cSLA/cSLS is established between end-customers and service providers. The pSLS on the other hand is established between the service and network providers or between network providers. The pSLS is an agreement between providers for exchanging traffic in the Internet, with the purpose of expanding the geographical span of their offered services. Additionally, pSLSes are meant to support aggregate traffic (i.e. serving many customers), and it is assumed that they are already in place prior to any cSLS agreements with end-customers. On the other hand cSLSes can differ depending on the type of services offered because different cSLS types can have different QoS requirements.

Solving the problem of end-to-end QoS monitoring can not be simply reduced to the concatenation of single domain QoS measurements but some multi-dimensional aspects must also be taken into consideration. One important aspect is the co-operation of providers in the service delivery chain. Here, it is assumed for monitoring at inter-domain scale that it is essential for providers to co-operate based on an agreed framework formulating the configuration of monitoring elements and service, the execution of measurements, the composition of results in an appropriate way, and the exchange of measurement data between providers. Building on the above requirements, the functions required for QoS monitoring over heterogeneous networks include: (1) QoS-based service monitoring at both QoS performance and perceived quality levels, (2) QoS-based resource monitoring for performance monitoring at traffic class, node, path, and network levels, (3) a set of protocols for exchanging the monitoring results.

The reminder of the paper is organized as follows. Section II shows the related works. Section III introduces the

978-1-4244-2067-4/08/$25.00 ©2008 IEEE.

ENTHRONE end-to-end integrated service management approach. The overall inter-domain monitoring framework architecture is presented in Section IV. Section V discusses the service level monitoring aspects. Section VI then describes the experimental test-bed used to evaluate the performance of the end-to-end monitoring solution. Finally, our conclusion is provided in Section VII.

II. Related Works

There has been some work on monitoring and measurements at inter-domain level, by European research projects [7]. To mention some, the objective of the IST-INTERMON project has been to develop an integrated inter-domain QoS monitoring, analysis and modelling system to be used in multi-domain Internet infrastructure for the purpose of planning, operational control and optimization [7]. The proposed solution assumes that a centralized manager negotiates monitoring operations with each domain along the service delivery path. This results in a scalability problem for the INTERMON system as the inter-domain network expands. The focus of the IST-MoMe project has been the enhancement of inter-domain real-time QoS architectures with integrated monitoring and measurement capabilities. The objective of the IST-SCAMPI project was to develop an open and extensible network monitoring architecture for the Internet including a passive monitoring adapter at 10 Gbps speeds, and other measurement tools to be used for denial-of-service detection, SLS auditing, quality-of-service, traffic engineering, traffic analysis, billing and accounting [7]. IST-LOBSTER is its follow-up project aimed at deploying an advanced pilot European Internet Traffic Monitoring Infrastructure based on passive monitoring sensors at speeds starting from 2.5Gbps and possibly up to 10Gbps [7]. Finally, IST-AQUILA project is developing inter-domain QoS-metrics measurement mechanisms, based on the BGRP proposal, to enable measurement based admission control (MBAC) in large-scale IP environment [2], [7].

Our work differs from the previous IST projects in that: (1) its end-to-end scope and business model encompasses Content Providers (CPs), Service Providers (SPs), Network Providers (NPs) and Content Customers (CCs); (2) end-to-end service monitoring is tackled using an overlay network of service-level monitoring components communicating in a cascaded fashion; (3) network-specific measurements are collected and translated to network-independent format using XML-based data models. The overall aim is to provide the means to monitor the services, networks, and resources at both intra- and inter-domain levels. Additionally, as part of service level monitoring, this work utilizes (1) Quality Meters [3] at user-side to measure the perceived quality level (Delivered PQoS) of an audio-visual stream; (2) assessment of the perceived quality (Derived PQoS) from measured network performances (Measured NQoS) in access/core networks (NQoS to PQoS mapping).

III. Integrated End-to-End QoS Management

This section shortly introduces the ENTHRONE framework [9] dedicated to provides QoS-enabled media access services by creating, offering, transporting and delivering of content. These access services are based on cooperation of several business actors (SPs, NPs, CCs) over heterogeneous multi-domain networking environment.

ENTHRONE is a service-oriented project targeting to build a complete architectural solution for multimedia content offering concerned with end-to-end QoS management in terms of performance targets at the user, application, terminal, and network levels. A central concept is the ENTHRONE Integrated Management Supervisor (EIMS) subsystem [5] which has a number of functional facilities/components for each entity such as SP, CP, CC and NP for managing the end-to-end service delivery. ENTHRONE II [4] extends ENTHRONE I EIMS by adding a new set of functionality including among others improved dynamic service management (policy based), MPEG-21 standard cross-layer QoS adaptation, metadata management and enhanced monitoring system. In principle, a specific manager sub-system is added for each new functionality. The EIMS Dispatcher component and especially its Service Manager (EIMS-SM) located at the SP deals with the customer subscriptions (cSLAs), contracts with NPs through pSLSs, the services owned by SP and the access to the service, which has been chosen. In other hand, the EIMS-SM at the NPs deals with pSLSs [11]. On service disruption, the monitoring system provides input to cross-layer QoS Adaptation Manager (EIMS-AM) for content adaptation [8].

IV. End-to-End Monitoring System Architecture

In this section, we present the end-to-end QoS monitoring architecture and the signalling protocols adopted in ENTHRONE 1 [10], [1] and in the next section, we discuss some service level monitoring aspects addressed in ENTHRONE phase 2.

A. Monitoring System Components

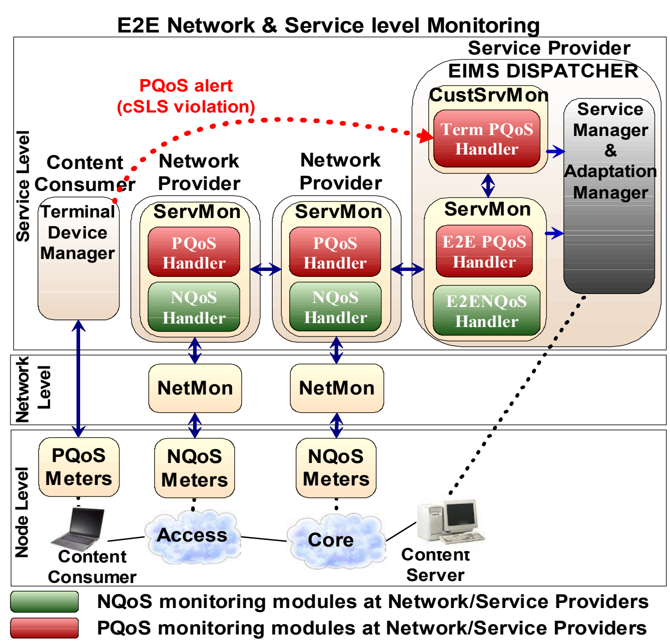

Figure 1. Overall monitoring system architecture

Three distinct monitoring components and two signalling protocols are defined in order to fulfil the aforementioned requirements. These monitoring components are called the Node, Network, and Service level monitors. EQoS-RM and EMon are the signalling protocols for monitoring exchanges at inter- and intra-domain levels. Fig. 1 shows the overall QoS monitoring system architecture. For efficiency and scalability reasons, the monitoring management architecture is structured in three levels: service-, network- and node-monitoring levels.

1) Node level Monitors (NodeMons): NodeMons are deployed only at network domain edges. They are used to perform active traffic measurements between any two edge nodes of an AS and to collect passive measurement information. NodeMons are configured with information about the variable to be monitored, the sampling and summarisation periods. Distinction must be made between network-level QoS measurements (NQoS) from application-level measurements (PQoS). PQoS probes, located at terminal, are used to examine application level perceived quality of audio-visual Digital Items [3]. PQoS Meters perform per-flow measurements, providing effective application-level QoS metrics and viewer-perceived quality. This helps to (1) detect cSLS violations (i.e., QoS degradations), to launch specific QoS failure location discovery (e.g., figure out the responsible domain/s), and (2) drive appropriate adaptation actions such as multimedia content adaptation, new load balancing initiatives, etc.

2) The Network level Monitor (NetMon): NetMon is responsible for intra-domain monitoring that utilizes network-wide performance and traffic measurements collected by all underlying NodeMons in order to build a physical and logical network view (i.e., the view of the established edge-to-edge QoS routes across the network). At NetMon level, the measurement information is further processed and aggregated so that only relevant QoS metrics are reported back to the monitoring component at the service level.

3) Service level Monitors: They are dedicated to perform customer/provider-related service level monitoring, auditing, reporting, and initiating some appropriate actions. Thus, they provide in-service verification of value-added services, verifying whether QoS performance guarantees committed in the SLSs are being met.

Two types of Service level Monitors are identified: the end-to-end provider-related (pSLS) Service Monitors (ServMons) and the Customer Service Monitor (CustSrvMonitor). These monitoring components are deployed as follows. Each NP will have its own ServMon (ServMon@NP) as a unique service monitoring subsystem while the SP will include both the ServMon (ServMon@SP) and the CustSrvMonitor subsystems. The monitoring operations performed by the ServMons and the CustSrvMonitor are presented in the next section.

Note that ServMons have different functionalities and roles depending on their location. ServMon@NP is in charge of (1) partial NP related pSLS monitoring; (2) inter-domain QoS reporting on aggregated streams using XML-based measurement statistics; (3) processing partial NP related mapping feedback. In other hand, ServMon@SP is in charge of (1) end-to-end pSLS monitoring; (2) coordinating the pSLS related service level monitoring procedures and to proceed with the information provided by other ServMon entities of the networks involved in the end-to-end chain of QoS delivery; (3) providing end-to-end mapping feedback as input to the EIMS-AM, to calculate new adaptation.

B. ENTHRONE Monitoring Signalling Protocols

The presented monitoring framework introduced two signalling protocols namely an inter/intra-domain monitoring signalling protocol (EQoS-RM), and an intra-domain active measurement signalling protocol (EMon). These protocols ensure reliable and secure data exchanging and are presented in more detail in [1] and [10].

The inter-domain QoS monitoring signalling protocol (EQOS-RM) is part of the EQoS (ENTHRONE QoS) signalling protocol suite. This sub-protocol is dedicated to the end to end service monitoring task and operates at the EIMS inter/intra-domain level of both the NP, and SP. EQoS-RM uses a set of XML signalling messages based on the simple request/response model.

The EMon (Enthrone Monitoring) protocol has been defined to support fast-responsive intra-domain communication between NodeMon peers. The EMon protocol is based on a reliable transport protocol (TCP, secure TCP (STCP), SCTP) and undertakes the configuration, synchronization, and management of active measurement sessions between edge-domain NodeMon agents.

V. Service Level Monitoring Approaches

Service level monitoring aims at keeping track of the compliance of the level of end-to-end service provided to the customers. To achieve this, ENTHRONE phase 1 introduced two types of monitoring services: (1) the first one is performed inside the core and access networks on aggregated streams [10] by the ServMons; (2) The second one is performed at the EIMS Dispatcher level on a particular customer stream [5] by the CustSrvMonitor.

In ENTHRONE phase 2, the monitoring system is enhanced by the following modifications/improvements:

- ServMons are extended with new Derived PQoS from Measured NQoS handling modules (PQoS Handler) in order to support the new NQoS to PQoS mapping functions [13]. In fact, using the Measured NQoS, an approximation of PQoS (i.e., Derived PQoS) delivered to a number of application streams can be inferred, allowing the adoption of corrective actions to prevent cSLS violation.
- CustSrvMonitor retrieves from the ServMon end-to-end network conditions information in order to build dynamic context data to be forwarded to the EIMS-AM for adaptation decisions.

A. Monitoring in Core/Access Networks

We distinguish two monitoring tasks carried out by the ServMons:

1) *Continuous Monitoring*: The Continuous Monitoring is performed on a per-domain basis by the ServMon@NP and it involves periodic active and/or passive measurements of pre-established pSLSs. During this task, pSLS QoS performance is continuously monitored, and the retrieved results are made available to the ServMon@SP. The latter is driven by the EIMS-SM that usually initiates the Continuous Monitoring procedure after the pSLS invocation phase whereas the committed resources are indeed allocated at the NP level. In this procedure the EIMS-SM@SP requests for certain monitoring processes/jobs, each for a given measurement granularity. Effectively continuous service monitoring and its measurement reporting can either be NodeMon-driven using COPS_RM (push mode) or NetMon-driven using SNMP/CLI (pull mode). Fig. 2 depicts the sequence diagram of a continuous monitoring procedure using COPS_RM in push mode. Using the specified measurement frequency, the NodeMons regularly send back their measurement reports to the NetMon. The NetMon aggregates the different received measurement reports and forwards them to the upper layer ServMon. This last one evaluates the degree of satisfaction of the pSLS crossing their domain and sends an EQoS monitoring report (EQoS_RM Report) along the return path to the ServMon that initiated the current continuous monitoring procedure (i.e. the ServMon@SP). Finally, the EIMS-SM@SP can receive the monitoring reports from the ServMon@SP. In phase 2, ServMons, periodically or triggered by specific events occurring in the network, can provide network conditions information to the EIMS-SM/AM. This phase consists of a new adaptation calculation and dynamic service management allowing: (a) pSLS (several cSLS) violation avoidance by taking preventive actions such as pSLS modification/ negotiation (e.g. 90% of pSLS is consumed); (b) better dynamic behaviour of AC algorithm leading to a better utilisation of the network resources.

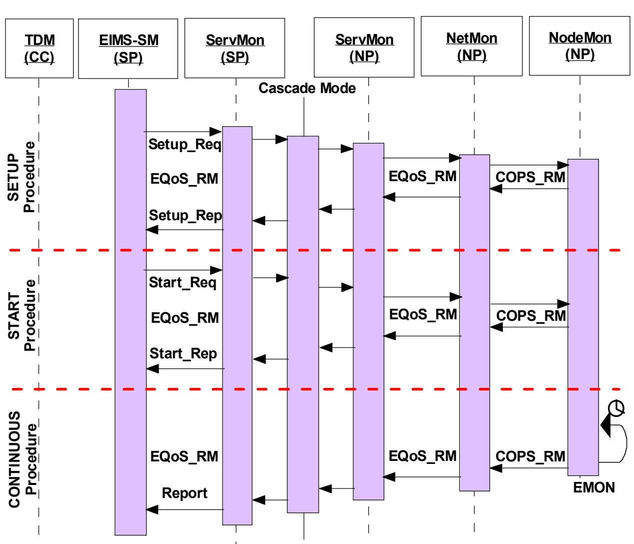

Figure 2. Continuous service monitoring procedure

2) *On-demand Monitoring:* The On-demand Monitoring operation is triggered by the CustSrvMonitor on PQoS alert and aims at locating the domain(s) that is/are the source(s) of end-to-end QoS degradation. This procedure allows verification of the conformance status of each pSLS (network conditions) by retrieving measurements of all pSLSs involved in the service provisioning process. Using the Terminal Device Manager (EIMS-TDM), service disturbance is reported to the CustSrvMonitor service. At this point, the ServMon@SP launches a "service disruption location" request by using the EQoS-RM protocol. Then, the service disruption location request spans all concerned ServMons@NP involved in the service provisioning chain in a cascaded fashion, asking them to retrieve the pSLSs measurements (if any) collected by the Continuous Monitoring procedure. If the pSLS measurements are not available at the NP level, the NetMon uses the COPS protocol to assess the pSLS fulfilment. Finally, each ServMon@NP aggregates the received reports and sends back an EQoS report in the service disruption location request reverse path. In phase 2, we also assume that the ServMon@SP receives a location response, and then provides input to CustSrvMonitor for remedial action taking (service adaptation/renegotiation). Fig. 3 shows the On-Demand service monitoring procedure.

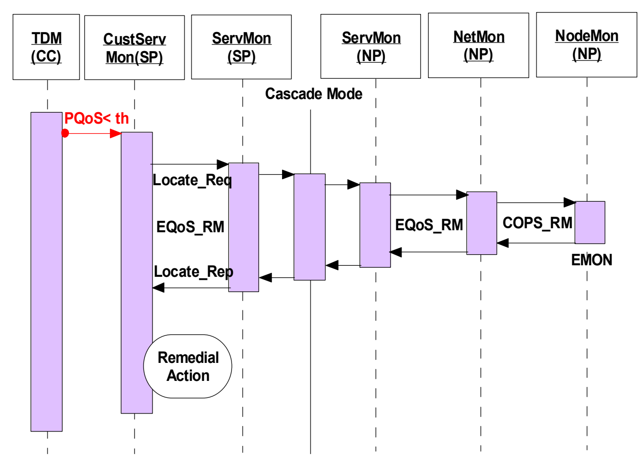

Figure 3. On-Demand service monitoring procedure

These two fundamental monitoring scenarios are preceded by a setup/configuration phase that often takes place after a successful pSLS negotiation on EIMS-SM@SP initiative. Thus, whenever a NP domain is crossed, a monitoring configuration from the ServMon to the NodeMon (monitoring job instantiation at specific network edges for a specific traffic class) takes place to basically register a new agreed pSLS, while the effective pSLS monitoring can be started later-on at the EIMS-SM@SP initiative.

B. *Monitoring at the EIMS Dispatcher level*

The CustSrvMonitor at EIMS Dispatcher performs customer service level (cSLS) monitoring using information collected at the terminal. Indeed, PQoS Meter sends a PQoS alert with some context data (e.g., measured PQoS value, date/time) to the CustSrvMonitor through the TDM indicating that the perceived QoS felt below a certain threshold, as defined within a given SLA (i.e., a degradation of QoS). This alert aims to initiate Digital Item (audiovisual content) adaptation remedial actions. Then, CustSrvMonitor triggers on-demand monitoring at ServMon level in order to collect related end-to-end network conditions information for building/updating full usage environment context information, which includes the perceived QoS, terminal capabilities, user characteristics, and collected network conditions. Finally, this information is provided to the MPEG-21 based EIMS-AM for calculating a new adaptation (MPEG21 Digital Item Adaptation) and/or service level agreement renegotiation.

VI. EXPERIMENTAL RESULTS

A. *Test-bed Configuration*

We set up a test-bed, comprising of two autonomous domains, AS1 and AS2 (Fig. 4, that represents the node level part of Fig. 1), representing two NPs configured to have an egde-to-edge domain RTT of 50 ms and 60 ms respectively. Each NP domain uses "NIST Net[2]" software to emulate the NP network-wide (WAN) behaviour. The test-bed is deployed to validate our implemented end-to-end monitoring solution and to evaluate its response time and accuracy. The one-way delay, loss and jitter are the IPPM QoS metrics [15], [6] of interest that have a short-term impact in the overall QoS delivery and degradation. The different NodeMons achieve edge-to-edge active measurements and are synchronized using the NTP protocol. We use Differentiated Services [6] traffic classes [6], termed Expedited Forwarding (EF) for Golden services, Assured Forwarding (AF) for Silver services, and Best-Effort (BE) for the Bronze services.

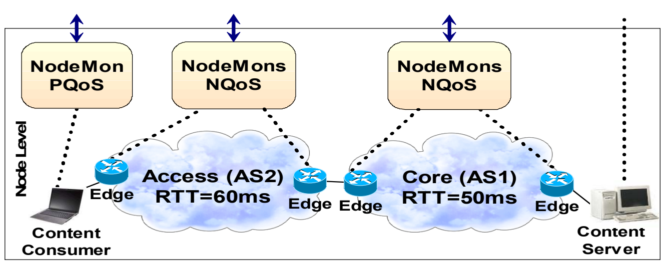

Figure 4. Node level part of monitoring system deployement test-bed

B. *Monitoring System Response time Analysis*

Fig. 5 shows the response time of the monitoring system when the network load is gradually increased in steps by 4% of the total capacity of links between two edge routers. Here, the response time stands for the time elapsed between the monitoring order issuance at EIMS Dispatcher and the time when the monitoring results are received. We assume that the EF traffic has a fixed bandwidth share that allows the traffic to be serviced even during the congestion periods. The signalling traffic is marked as EF traffic and then, it is not affected by network conditions. In Fig. 5, it's clearly revealed that the measured values of the response time for each service class are rather stable over the time. The oscillations are due to the fact of TCP/SCTP natural behaviour as explained above. Since all signalling traffic was marked as EF traffic, fairly good response time was maintained (oscillates around 500 ms for all services classes). Hence, the network load dynamics affect only the user traffic and not signalling traffic.

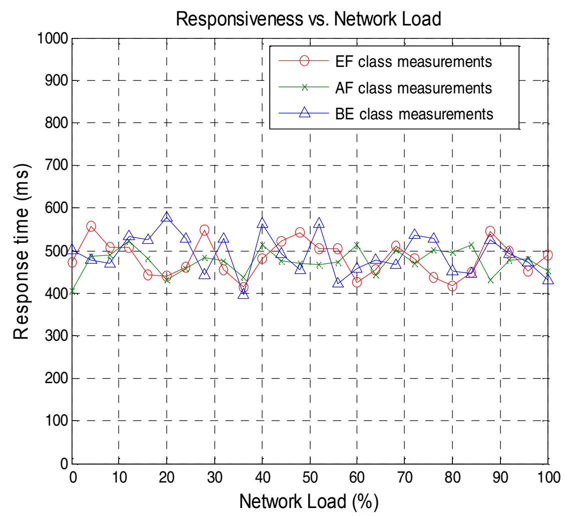

Figure 5. Monitoring system response time for different level of network load

C. *Monitoring System Accuracy Analysis*

In order to characterize our monitoring system accuracy, we explicitly introduced in the network specific delay (30ms for EF, see Fig. 6), jitter (10ms for AF, see Fig. 7), or loss rate (30%s for EF, see Fig. 8) for each service class, and then measured the QoS metrics related to these service classes.

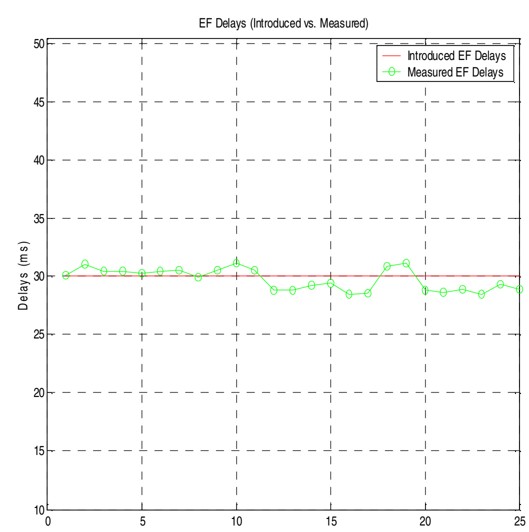

Figure 6. One-way delay measurements for EF traffic class

[2] NIST Net is a network emulation package. http://www-x.antd.nist.gov/nistnet/

The measurements were repeated 25 times to get more information about the "long-term" accuracy of our monitoring system and its ability to continuously perform measurements and produce measurement results. Given the potential NTP clock lag (1 millisecond) the QoS metrics measured by our monitoring system were very close to the ones introduced. Especially, both delay and jitter accuracy falls below 1 millisecond most of the time.

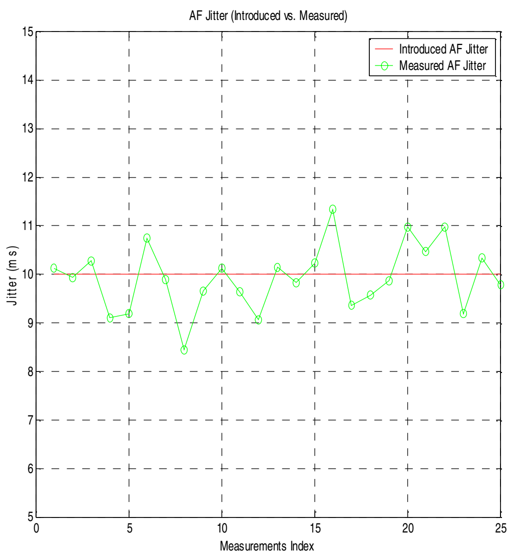

Figure 7. Jitter measurements for AF traffic class

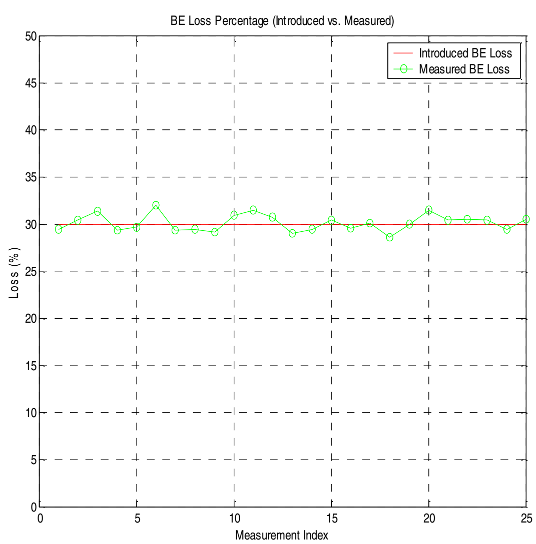

Figure 8. Packet loss ratio measurements for BE traffic class

The above results show that traffic classes and consequently pSLSs can be accurately and individually monitored. Thus, our monitoring system can provide real-time and accurate inputs to service and adaptation management entities for end users service assurance and network resources optimization.

VII. CONCLUSION

This paper describes a service-oriented end-to-end QoS monitoring system aimed at keeping track of the compliance of the level of end-to-end service provided to the customers by the service providers. The monitoring system provides the means for remedial actions to be taken (MPEG21 cross-layer adaptation) in case of service degradation or failure, e.g., on non-conformance to SLSs. It also assists network providers in making provisioning decisions for optimizing the usage of network resources. Thus, the service level monitoring supports the cross layer adaptation and the dynamic service management functions.

REFERENCES

[1] T. Ahmed, ed., ENTHRONE Deliverable D23f "End-to-end QoS Signalling & Policy-based Management Architectures," August 2005.

[2] A. Asgari, P. Trimintzios, G. Pavlou, R. Egan, "Scalable Monitoring Support for Resource Management and Service Assurance", IEEE Network Magazine, Dec./Nov. 2004, Vol. 18, No. 6, pp. 6-18.

[3] P. Brétillon, ed., ENTHRONE I Deliverable 25.2 "Perceived Quality Meters and Agents Prototypes", August 2005.

[4] P. Brétillon, ed., ENTHRONE II Deliverable D01, "Overall system architecture – version 2", February 2007.

[5] INESC, ed., D7 ENTHRONE I Deliverable 07 "IMS Implementation", 20 December 2005.

[6] IETF, For information on the various IETF working groups including IPPM and DiffServ, visit http://www.ietf.org

[7] IST, European IST research projects, for more information visit: www.cordis.lu/ist/. Specifically for IST-INTERMON visit: www.ist-intermon.org/, for IST-MoMe visit: www.ist-mome.org/, for IST-LOBOSTER visit: www.ist-lobster.org/, for IST-AQUILA visit: www-st.inf.tu-dresden.de/Aquila/, and for IST-SCAMPI visit: www.ist-scampi.org/.

[8] I. Kofler, ed., ENTHRONE II Deliverable D05i, "MPEG-21 based Cross-Layer Adaptation Decision-Taking Engine", April 2007.

[9] E. Le Doeuff, ed., ENTHRONE I Deliverable 01 "Overall system requirements and functional architecture specification", 31 March 2004.

[10] A. Mehaoua, T. Ahmed, H. Asgari, M. Sidibé, A. NAFAA, G. Cormentzas, and T. Kourtis, "Service-driven Inter-domain QoS Monitoring System for Large-scale IP and DVB Networks" in ELSEVIER's Computer Communication Journal, Special Issue on Monitoring and Measurements of IP Networks, Volume 29, Number 10, page 1687-1695 - june 2006

[11] P. Flegkas, ed., D1.1: "Specification of Business Models and a Functional Architecture for Inter-domain QoS Delivery", May 2003. http://www.mescal.org/ Public Deliverables.

[12] P. Souto, ed., ENTHRONE II Deliverable D03i, "EIMS for ENTHRONE 2", March 2007.

[13] M. Sidibé, ed., ENTHRONE II Deliverable D06i, "Service Management and Monitoring", April 2007.

[14] TRT, ed., ENTHRONE I Deliverable D24F, "Specification of protocols, algorithm, and components, the architecture, and design of SLS Management", July 2005.

[15] V. Paxson et al., "Framework for IP Performance Metrics," IETF RFC-2330, May 1998.

[16] Goderis D.et al., Internet Draft <drafttequila-sls-03.txt>, "Service Level Specification Semantics, Parameters and negotiation requirements", October 2003. Expired.