ETOMIC advanced network monitoring system for future Internet experimentation

ISTVÁN CSABAI, ATTILA FEKETE, PÉTER HÁGA, BÉLA HULLÁR, GÁBOR KURUCZ, SÁNDOR LAKI, PÉTER MÁTRAY, JÓZSEF STÉGER, GÁBOR VATTAY

Eötvös Loránd Univeristy, Budapest, Hungary
{csabai,fekete,haga,hullar,kurucz,laki,matray,steger,vattay}@etomic.org

Keywords: network monitoring, ETOMIC, traffic measurement, GPS

ETOMIC is a network traffic measurement platform with high precision GPS-synchronized monitoring nodes. The infrastructure is publicly available to the network research community, supporting advanced experimental techniques by providing high precision hardware equipements and a Central Management System. Researchers can deploy their own active measurement codes to perform experiments on the public Internet. Recently, the functionalities of the original system were significantly extended and new generation measurement nodes were deployed. The system now also includes well structured data repositories to archive and share raw and evaluated data. These features make ETOMIC as one of the experimental facitlities that support the design, development and validation of novel experimental techniques for the future Internet. In this paper we focus on the improved capabilities of the management system, the recent extensions of the node architecture and the accompanying database solutions.

1. Introduction

The European Traffic Observatory Measurement InfrastruCture (ETOMIC) was launched in 2004 [1,2]. It is targeted to provide the scientific community with an Internet measurement platform that is fully open and reconfigurable, extremely accurate and GPS-synchronized.

The ETOMIC system has been designed to allow researchers to perform any kind of active network measurement. The users are provided with a web-based graphical user interface for the definition of the experiments to run. Researchers may either choose from a number of built-in measurement scripts that cover the most popular measurement techniques, like traceroute or packet-pair experiments, or they can provide their own code for the experiments. To avoid conflicts in resource utilization each measurement has to be scheduled to exclusively reserve node resources for its execution. The node reservations are performed through the web-based user interface. The ETOMIC management kernel takes care of the software upload and experiment execution in a fully automated fashion.

After the successful duty of the measurement nodes since 2004 the renewal of the system components was necessary. In the Onelab project [3] we have extended the capabilities of the measurement hardware to match the current technological level and to incorporate the software evolution of the last years that are important from the perspective of network measuerments. The ETOMIC infrastructure now provides two ways of collecting experimental data. One possibility is when the researcher reserves and configures the measurement nodes and sets the parameters of the experiment through the Central Management System.

In this case, besides the original ETOMIC nodes, newly deployed enhanced measurement boxes can also be used for experimentation. To meet the requirements of high precision measurements the nodes are equipped with a DAG card (for the original nodes) or an ARGOS card (for the new generation nodes) to provide nanosecond-scale timestamping of network packets. Besides these nodes a third type of hardware component was also introduced, which is called Advance Probing Equipment (APE). APE is a low cost hardware solution developed to serve as a measurement agent for user applications: it provides a web service interface to conduct experiments.

This approach enables autonomic software components to automatically collect relevant network data from the ETOMIC system they rely on for their operation. As a consequence of a development in the system kernel the nodes of the PlanetLab platform [4] can also be used as measurement nodes by the ETOMIC system. The goal of this integration was to enable the federated usage of the high precision ETOMIC nodes and the numerous PlanetLab nodes.

To make it easier to handle and archive the huge amount of data collected by the ETOMIC platform we have created data repositories. There are two different interfaces for these data archives. The periodic measurements web interface can be used to poll automatically collected measurement data through pre-defined queries. As another approach, the Network Measurement Virtual Observatory (nmVO) [5] provides standard SQL database access to the user community. The nmVO provides a graphical user interface and a web service interface for the users to access raw and evaluated measurement.

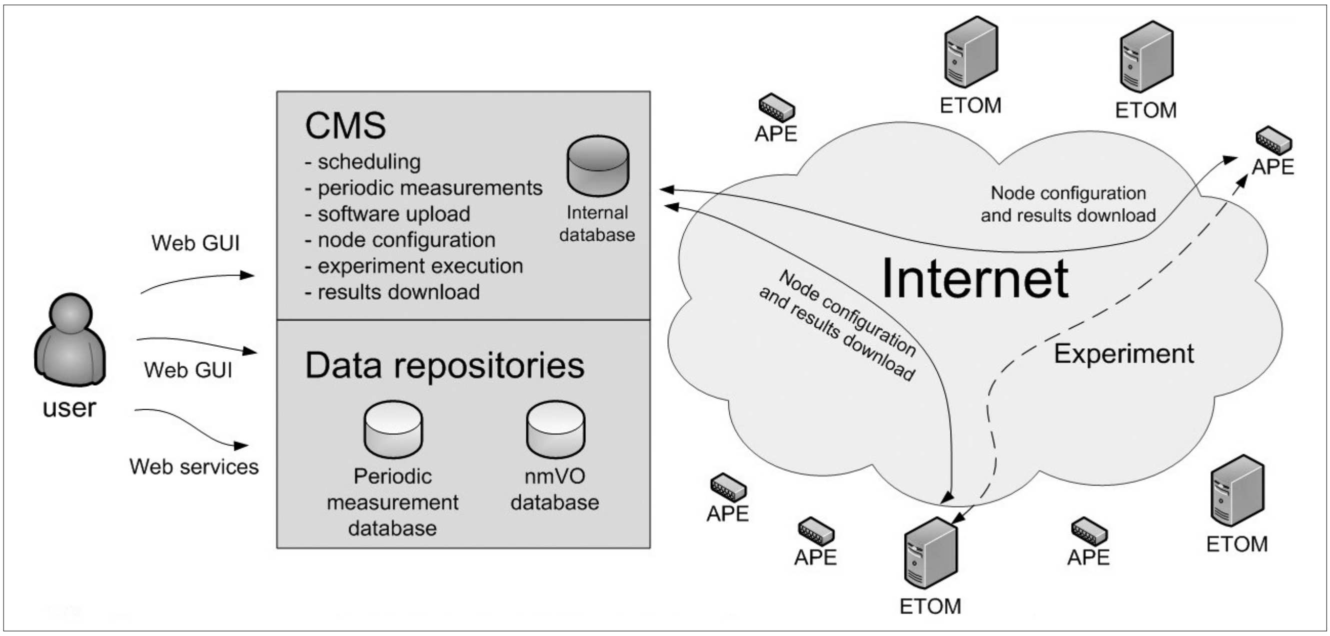


Figure 1. ETOMIC system architecture

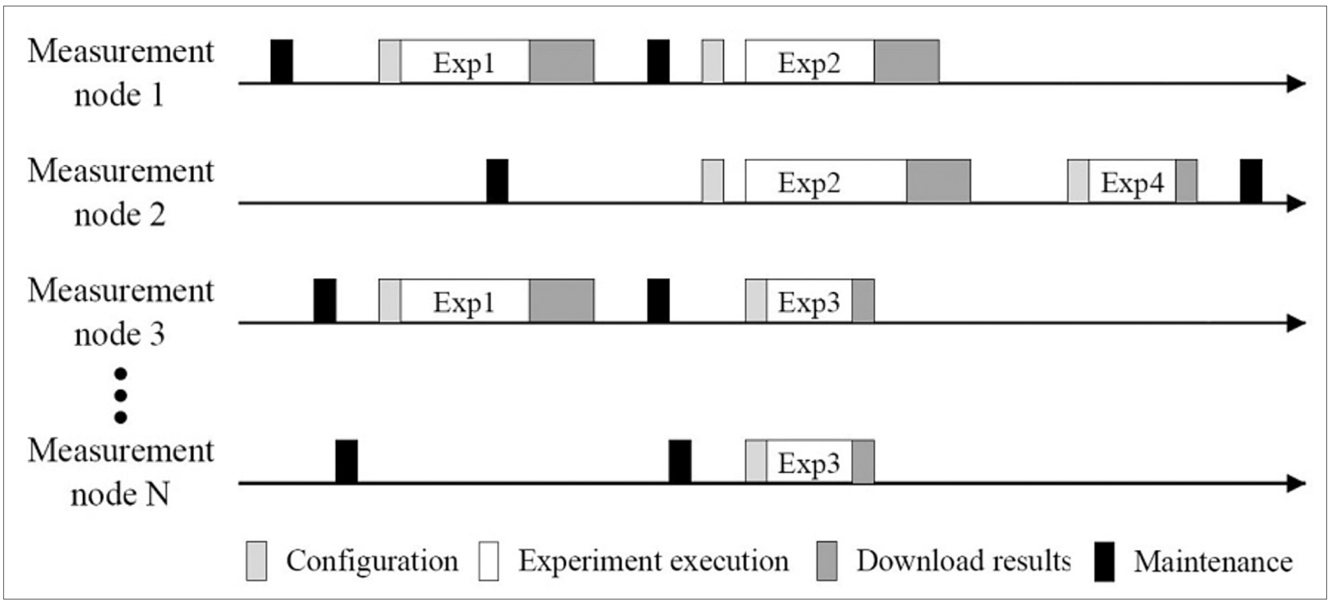
2. System architecture

The ETOMIC infrastructure is constituted of high precision measurement equipment modules placed hosted by European universities, research institutes and company laboratories. The clocks of the measurement nodes are synchronized via GPS signals, which allow not only packet round-trip time estimation, but also precise one-way delay measurements. The ETOMIC platform is very flexible, since researchers can develop and run any kind of active experiments.

A Central Management System (CMS) is in charge of system control, comprising not only the scheduling and execution of measurements experiments, but also system monitoring and configuration. The main software component of the CMS is called the *management kernel* which is running on a dedicated server computer. The kernel is responsible for scheduling tasks, deployment of user software to the measurement nodes, node configuration, experiment execution and the collection of measurement results from the nodes. The CMS provides a web-based graphical user interface where the researchers can configure the system and reserve system resources for their measurements. An internal database is attached to the kernel where the system and user level management information are stored. The results of a finished measurement are also collected and stored in this database until the user downloads them. The system components are depicted in *Fig. 1*.

As an important add-on for the original ETOMIC system the architecture has been extended with large capacity data repositories that are publicly available. The

Figure 2. Scheduling of experiments and maintenance tasks



system provides several different interfaces for these repositories through which the users can reach the collected datasets. The interfaces allow the users to run intelligent queries in order to filter and process the raw data on the server side.

3. Management kernel

3.1 Central Management System (CMS)

The ETOMIC management kernel constitutes the core of the Central Management System (CMS). It is in charge of user management, experiment scheduling and keeping the corresponding results in the temporary data storage.

In order to isolate the different measurements and to schedule experiments and maintenance tasks a calendar is used for each measurement node, as it is shown in *Fig. 2*. The management tasks can be divided into two branches: tasks that correspond to experiment definitions and tasks that correspond to the execution of the scheduled experiments.

3.2 Scheduling and calendar maintenance

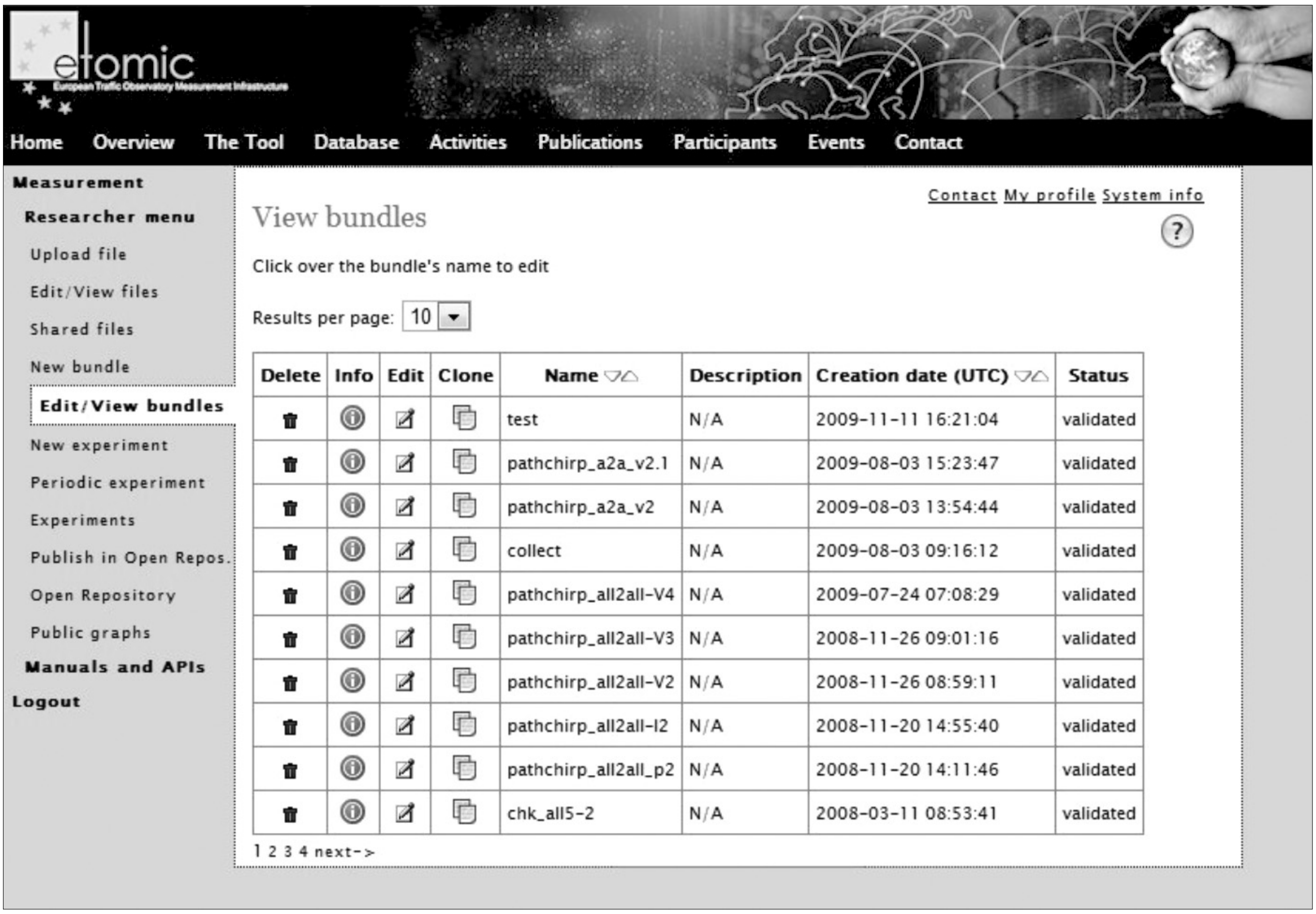
The researcher is expected to book measurement nodes for a certain time interval and to upload the applications neccessary for the measurement. The web interface is in charge of checking that the timeline for the experiment does not collide with any other previously registered measurement. In case of a successful resource reservation the CMS inserts the new experiment information into its internal database. The management kernel is continuously checking the internal database for new measurement requests. Once a new experiment has been defined and the deadline for execution approaches the management kernel uploads the measurement softwares and configures the nodes, executes the experiment and finally downloads the results to its temporary storage.

3.3 Web interface to CMS

Users are provided with a graphical interface for setting up the experiment beforehand. Then, the management kernel is in charge of experiment execution. An internal database is used to store all the necessary information to run the experiments. The stored information includes the applied softwares, external data files, the experiment results, the experiment status and the measurement node status.

Users can find manuals, the programming API and example codes here. The users can do the following operations: add a new program; upload the necessary data files; define the experiment bundle by scheduling the start and end times of the measurement; book ETOMIC time and reserve the measurement nodes; download the results; define periodic experiments with the repetition period; sharing files with other users.

Figure 3. List of experiment bundles on the web based graphical user interface



	ETOM w DAG	ETOM w ARGOS	APE	PlanetLab
platform	Intel server PC	HP server PC	Blackfin board	Variable
timestamping accuracy	60 ns	10 ns	100 ns	~10 ms
time synchronization	yes	yes	yes	no
GPS receiver	Garmin 35HVS	U-Blox LEA-4T	U-Blox LEA-4T	—
number of deployed nodes	18	20*	20*	~300**
user interface	web GUI	web GUI	web services	web GUI

*under deployment, **under integration

*Table 1. Available measurement nodes in the ETOMIC system *under deployment, **under integration*

3.4 Integration of Planetlab's nodes

PlanetLab is a global platform for supporting the development of new network services. This platform is also used for network experiments. The nodes of the PlanetLab platform are accessed interactively via remote shell. This access method enables the CMS to use the PlanetLab nodes as its own nodes. Although the main hardware capabilities of the PlanetLab and ETOMIC nodes significantly differ, the large number of Planet Lab nodes makes them very attractive to the user community.

The capabilities of PlanetLab are not described in this paper, here we only note that the PlanetLab nodes are usually up-to-date server PCs without any hardware components specialized for network measurements. The slice based management of PlanetLab nodes allows multiple users to run experiments simultaneously in the same remote node at the same time, while the CMS takes care of the unique resource allocation. In spite of the basic differences of PlanetLab and ETOMIC the federated usage of the high precision ETOMIC nodes and the numerous PlanetLab nodes could lead to new ways of experimentation.

The software installed on ETOMIC nodes has been adapted to make the joint usage possible, using a slice of PlanetLab that is automatically renewed by the CMS. This makes the whole range of ETOMIC and PlanetLab remote nodes available through the ETOMIC web interface. The most important challenge for the integration was the synchronization of the clocks in nodes from both platforms as they use different reference signals with highly different precisions.

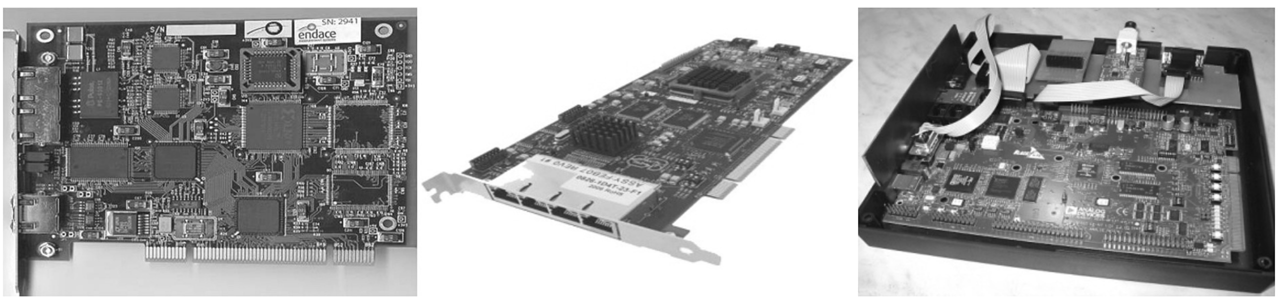
4. The measurement nodes

The nodes can be divided into two groups based on their hardware architecture. The ones that are built on server PC architecture are called *ETOM*. These nodes are accessible via the web-based graphical interface presented in Section 3.3. The ones that are based on a lightweight programmable board are called *APE*. The APE nodes are accessible via a web service interface. GPS receivers are connected to all types of measurement hardware to provide the precise time synchronization between the nodes and to provide the reference clock for the measurement cards.

Both of ETOM's and APE's hardware solutions is high-precision ones, due to the incorporated precision equipment (DAG and Argos cards) that are specifically designed to transmit packet trains with strict timing, in the range of nanoseconds.

APE is built on a development board with Blackfin processor. The board is manufactured by Analog Devices Inc. and has a number of different interfaces for hosting auxiliary hardware components that are responsible for specific network measurements tasks. Each measurement node is provided by two network interfaces:

Figure 4. Hardware components of ETOMIC responsible for the precision timestamping of the network packets. From left to right: the Endace DAG 3.6GE measurement card, the netFPGA based ARGOS measurement card and the APE node.



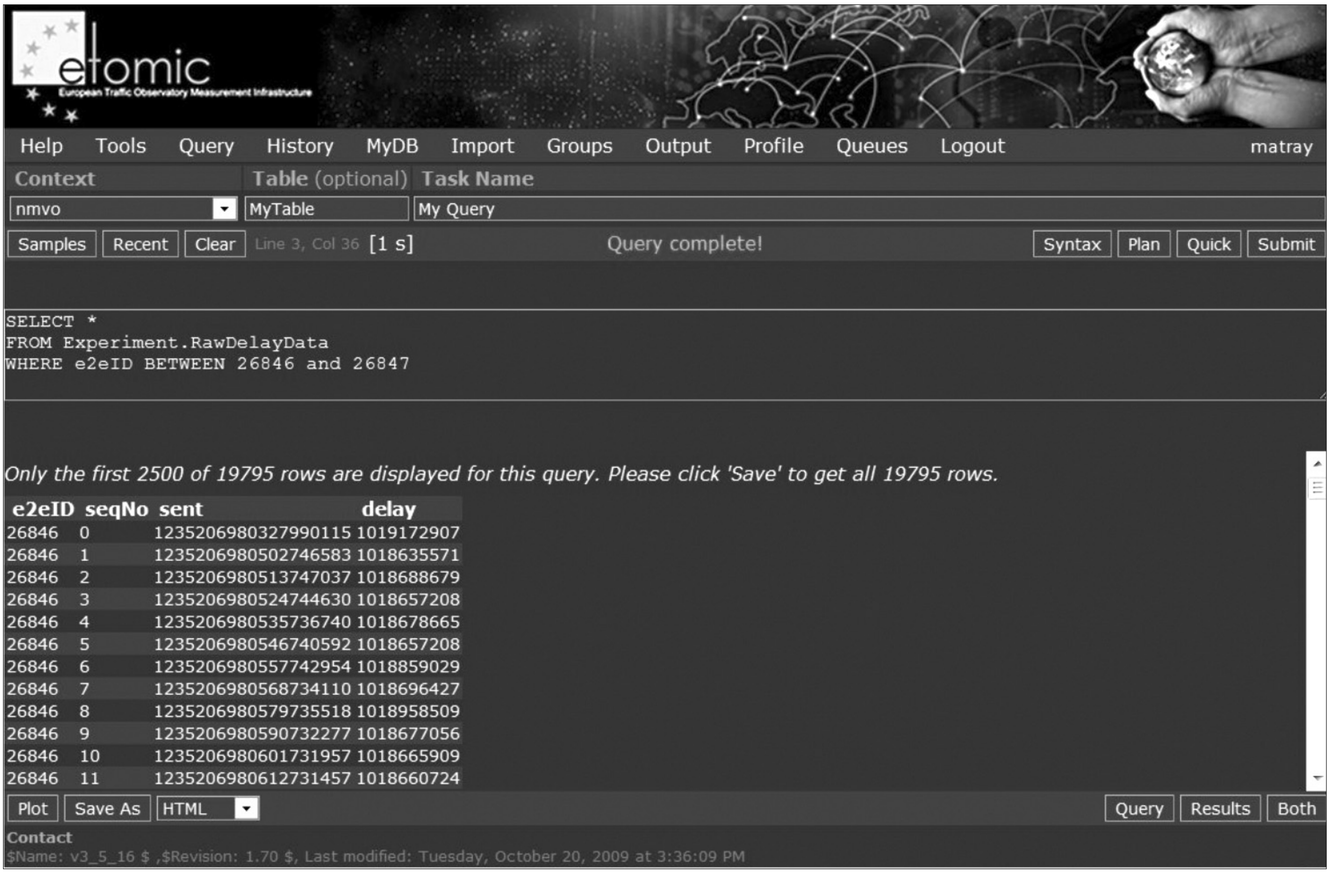


Figure 5. The web interface of the nmVO

a standard network interface card for management purposes (maintenance, software upload, data download) and an additional precision card for the probe traffic. The main feature of the APE nodes is that they provide measurement services which can be remotely called by user applications in an online fashion, without time slot reservation.

5. Accessing measurement data

5.1 Data repositories

For network measurements the collected raw data is traditionally stored in files in some standard (like traceroute dump, tcpdump) or custom formats. The files are then processed according to the research questions to be answered. Detailed analysis of complex networks requires large statistical samples. This requirement leads to substantial data size in case of measurements in high bandwidth networks, even if just a few parameters of the packets are recorded (like IP addresses, arrival time, protocol, size or delay).

Practically, measurements can produce dozens of megabytes at each monitoring node that sums up to hundreds of megabytes or even terabytes in multi-node experiment. Keeping only the results of the data analysis and discarding the raw data themselfes is not a good way to solve the data handling issues: measurement data gathered today cannot be reproduced in the future. Thus it is preferable to store the original datasets to allow further re-analysis and support the study of the long-term evolution of the network.

For these purposes we have created data repositories to store measurement data collected by means of the ETOMIC system. There are two different interfaces to reach raw and aggregated measurement data. The periodic measurements web interface can be used to poll data collected from automatic measurements in the system. The users can choose from pre-defined queries by selecting a given type of measurement, a set of nodes on which the measurement was conducted and a time frame. Long term data sets are available for one-way delay values, traceroute measurements and Paris traceroute measurements [6] data between all the ETOMIC node pairs. The results of these measurements are reachable through the periodic measurement repository of the ETOMIC website's Database / Open Repository menu [1].

5.2 Network Measurement Virtual Observatory (nmVO)

The basic concept presented in [5] is an approach to efficiently store and share research data. Beyond the simple data collecting and archiving functions it aims at providing easy-to-use analysis tools via both human and machine readable interfaces.

One of the main features of the nmVO is that it provides SQL access to the databases that are integrated under its framework, thus the users can edit and run their customized queries through either the web-based SQL interface or the web services interface. The main advantage of this solution is that the researchers can filter out the relevant information from the huge archives using server side processing. Hence, only the necessary datasets and results have to be downloaded from the server.

To sketch the nmVO principle through a possible application, consider a scenario in peer-to-peer overlay networks where management information is needed to optimize the routing between the peers. It would be unthinkable to use gzipped files for such real-time evaluation. On the contrary, the scenario is feasible if one turns to the nmVO to get the typical loss rate, the average delay on certain routes or the shortest path between the peers. This means that beyond the data itself, analysis tools are also needed to perform such data filtering and transformation queries efficiently. Using these stored procedures we can move the typical filtering and preprocessing tasks to server side.

The majority of the experimental data collected in the ETOMIC system is inserted into the data repositories and can be reached through both the nmVO web graphical user interface and the web service interface. The developed nmVO API is integrated into the APE nodes' software, so that all measurement data from the APE boxes are automatically copied into the nmVO data repository.

In addition to the historical raw data collections and the evaluated results of periodic measurements, also non-ETOMIC traceroute logs and topology data [7], one-way delay values [8], queueing delay tomography data [9], available bandwidth results [10], router interface clustering and IP geolocalization data [11] can be found in the archive.

The nmVO can be accessed through the ETOMIC website's Database / CasJobs Query Interface menu [1] and also via Web services for client applications.

6. Conclusions

In this paper we presented the enhanced ETOMIC network measurement infrastructure. We described the key components of the architecture and the new features of the Central Management System. The improved system kernel includes support for periodic measurements and the federated usage of the high precision ETOMIC nodes and the numerous PlanetLab nodes. Besides the kernel development novel hardware components have been developed and deployed. New lightweight measurement equipments have been installed that provide measurement services which can be remotely called by user applications via web services.

The system now also includes well structured data repositories to archive and share the experimental data. Periodic measurement data can be polled with customizable pre-defined queries, while the nmVO framework gives full SQL access to its archive. The recent developments make ETOMIC an easy to use experimental facitlity with versatile feature for network research.

Acknowledgements

The authors thank the partial support of the EU ICT OneLab2 Integrated Project (grant agreement No.224263), the EU ICT MOMENT Collaborative Project (grant agreement No.215225) and the National Office for Research and Technology (NAP 2005/ KCKHA005).

Authors

ISTVÁN CSABAI is a professor of physics at the Department of Complex Systems, Eötvös Loránd University. His research interest is quite wide, ranging from artificial intelligence through cancer research to cosmology. Although these disciplines are diverse, the way of handling of their inherent complexity and the necessary computational tools are often common. He has written his first paper about the dynamics of communication networks 15 years ago, recently he is working on the development of the ETOMIC network measurement system and the Network Measurement Virtual Observatory concept.

ATTILA FEKETE was born in 1975, Budapest. He received his MSc degree in physics from Eötvös Loránd University (ELTE) in 1999. After graduation he studied TCP dynamics in collaboration with ELTE Communication Networks Laboratory and Ericsson Research. Between 2003 and 2008 he worked in Collegium Budapest Institute for Advanced Studies on the modeling of complex networks. He received his PhD degree in physics from ELTE, where he is currently working as a teaching assistant, in 2009.

PÉTER HÁGA is an assistant professor at the Eötvös Loránd University. He received his PhD degree in physics in 2008. From 1999 he is a permanent member of the Communication Networks Laboratory and he is one of the founders of the ETOMIC measurement infrastructure. His research interest includes network measurement techniques, traffic modeling and adaptive protocols.

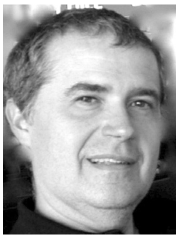
BÉLA HULLÁR received his MSc diploma in Computer Science in 2007 from the Eötvös Loránd University. He has been working in ELTE Communications Network Laboratory since then, as a PhD student. His research interests are in monitoring and measuring computer networks, developing the appropriate measurement infrastructure and its management system.

GÁBOR KURUCZ received his first MSc degree in Telecommunications and Media Informatics at 2003 from the Budapest University of Technology and Economics (BME). He earned his second diploma in Security of Information and Communication Systems at 2008 from BME. In 2006, he joined the Department of Physics of Complex Systems at ELTE. His research interests are in developing anetwork equipment with a precise timestamping module under the OneLab project.

SÁNDOR LAKI received his MSc degree in computer science from the Eötvös Loránd University, Budapest, in 2007. Currently he is working towards his PhD in computer science at the Department of Physics of Complex Systems. His primary research interest includes Internet measurement techniques, adaptive protocols and network modeling, especially IP geolocation and traffic classification.

PÉTER MÁTRAY received his M.Sc. degree in mathematics and computer science from Eötvös Loránd University, Budapest in 2005. Currently he is doing his PhD studies at the Department of Physics of Complex Systems. His attention is mainly focused on the database aspects of network measurements. Besides that, he is partly involved in active probing Internet measurements (especially network tomography) and visualization.

JÓZSEF STÉGER graduated as a biophysicist at the Faculty of Science at Eötvös Loránd University in 2001. Now works as an assistant lecturer at the Dept. of Physics of Complex Systems. He is carrying out research tasks within the study of communication networks, like the Internet. By using the measurement architecture introduced in the article he made queuing delay tomography experiments possible, which enables the group to notice congested hotspots within the network. He is an active member of international projects Moment and Onelab2.



GÁBOR VATTAY is a professor of physics of complex systems at the Faculty of Sciences of Eötvös University. He received his PhD in 1994 and his DSc title in 2003 in mathematical physics of chaotic systems. His interest is in the complex dynamics arising in natural and man-made systems. Since 1998 his interest shifted toward the dynamics of networked systems. He is the founder and director of the Communication Networks Laboratory established at Eötvös Loránd University by Ericsson Research since 2000 and the leader of the EU supported Internet measurement effort ETOMIC.

References

[1] The ETOMIC website, *http://www.etomic.org*

[2] D. Morato, E. Magana, M. Izal, J. Aracil, F. Naranjo, F. Astiz, U. Alonso, I. Csabai, P. Haga, G. Simon, J. Steger, G. Vattay, ETOMIC: A testbed for universal active and passive measurements. IEEE TRIDENTCOM 2005, Best Testbed Award, pp.283–289, 23-25 Febr. 2005.

[3] OneLab – Future Internet Test Beds, *http://www.onelab.eu*

[4] L. Peterson, T. Anderson, D. Culler, T. Roscoe, A Blueprint for Introducing Disruptive Technology into the Internet. Workshop on Hot Topics in Networks, October 2002.

[5] P. Mátray, I. Csabai, P. Hága, J. Stéger, L. Dobos, G. Vattay, Building a Prototype for Network Measurement Virtual Observatory. ACM SIGMETRICS – MineNet 2007, 12 June 2007, San Diego, USA.

[6] B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, R. Teixeira, Avoiding traceroute anomalies with Paris traceroute. Internet Measurement Conference, October 2006.

[7] S. Garcia-Jimenez, E. Magana, D. Morato, M. Izal, Techniques for better alias resolution in Internet topology discovery. IFIP / IEEE IM 2009, 1-5 June 2009.

[8] A. Hernandez, E. Magana, One-way Delay Measurement and Characterization. International Conference on Networking and Services (ICNS '07) , 2007.

[9] T. Rizzo, J. Stéger, I. Csabai, G. Vattay, P. Pollner, High Quality Queueing Information from Accelerated Active Network Tomography. Tridentcom 2008, Innsbruck, Austria, March 17-20, 2008.

[10] P. Hága, K. Diriczi, G. Vattay, I. Csabai, Granular model of packet pair separation in Poissonian traffic. Computer Networks, Vol. 51, Issue 3 , pp.683–698, 21 Februar 2007.

[11] S. Laki, P. Mátray, P. Hága, I. Csabai, G. Vattay, A Detailed Path-latency Model for Router Geolocation. IEEE TridentCom 2009 Conference, 6-8 April 2009, Washington D.C., USA, 2009.