An advanced measurement meta-repository

Pedro A. Aranda Gutiérrez[1], Antal Bulanza[2], Marek Dabrowski[3], Baiba Kaskina[4], Juergen Quittek[5], Carsten Schmoll[6], Felix Strohmeier[7], Attila Vidacs[8], Kardos Sandor Zsolt[8]

[1]*Telefónica, Investigación y Desarrollo, Spain*
[2]*University of Brussels, Belgium*
[3]*Warsaw University of Technology, Institute of Telecommunications, Poland*
[4]*TERENA, The Netherlands*
[5]*NEC Europe, Germany*
[6]*Fraunhofer FOKUS, Germany*
[7]*Salzburg Research, Austria*
[8]*Budapest University of Technology, Hungary*
E-mail: [1]paag@tid.es, [2]bulanza@helios.iihe.ac.be, [3]mdabrow5@tele.pw.edu.pl, [4]baiba@terena.nl, [5]quittek@netlab.nec.de, [6]carsten.schmoll@fokus.fraunhofer.de, [7]fstrohmeier@salzburgresearch.at, [8]{vidacs, kardos}@tmit.bme.hu,

Abstract

The paper proposes meta-data models for information about public repositories of measurement and monitoring data and tools. The meta-data model consists of a set of attributes, which describe the features of the actual data of interest. The discussed meta-data models constitute a base for the advanced meta-repository of measurement tools and data, which was developed by the IST project MOME (Monitoring and Measurement Cluster). MOME is coordinating several European projects working on different issues related with traffic measurement and monitoring in IP-based networks. The MOME meta-repository is a platform for exchange of information between researchers working with the measurement data and tools. Exemplary usage scenarios described in this paper demonstrate the capabilities of the implemented MOME meta-repository.

1 Introduction

Traffic measurement and monitoring is considered as an important additional functionality in IP (*Internet Protocol*) based networks. The monitoring results are essential both for supporting current network operation tasks, and for enabling research on traffic characterisation and modelling. Several issues related with traffic measurements are currently intensively studied by many European research projects (e.g. [10][11][12][13][14][15][16]). The investigated topics include, among others:

1) Developing new measurement and monitoring methods. The goal is to design and deploy tools, which allow for more efficient and more accurate measurements of traffic characteristics, resource consumption, and QoS (*Quality od Service*) metrics.
2) Developing advanced architectures of measurement and monitoring systems for IP and IP QoS networks. Special focus is put on aspects related with performing measurements in the multi-domain environment.
3) Developing new algorithms for network and traffic control functions. The efficiency of some of those functions, like traffic engineering, admission control, or fault management, can be significantly increased thanks to the support by measurements.
4) Developing and validating realistic traffic models. Analytical models attempt to reproduce the characteristic behaviour of traffic observed on different levels (packets, flows, sessions). New models should be validated with various, sufficiently representative sets of raw measurement data, captured in real operational networks. Therefore, easy access to the repositories of raw measurement data is required.
5) Evaluating the packet transfer characteristics in operational and research networks. This topic covers both designing controlled trials for certification of new mechanisms, as well as methods for on-line verification of QoS delivered to the customers (SLA/SLS monitoring). The design of appropriate tests requires deep knowledge of the capabilities and limitations of the available measurement tools.

Effective work on the above topics requires access to up-to-date information about available measurement tools and data. Currently, several projects and organizations collect the measurement data (e.g. [6][7][8][9]). However, the tools and data formats used by them are different and often incompatible. Moreover, finding the appropriate data among many uncoordinated repositories is not easy due to lack of a standard way for describing the assumed measurement scenario and environment. As a consequence, the research using measurement data is quite a difficult task.

The main objective of the IST-MOME (*Monitoring and Measurement Cluster*) [1] is coordinating the work of European projects active in the fields of network monitoring and measurements. The coordination tasks include, among others, collecting and disseminating information about known **measurement tools** and available **measurement data**.

The information stored by MOME has a form of **meta-data**, i.e. "data about data". Therefore, unified meta-data models (description formats) were developed for annotating various measurement tools and measurement data. Based on the proposed meta-data models, the specialised MOME meta-data repository was implemented and deployed. The access to the repository is publicly available via the MOME project web page.

The general overview of the assumed approach is presented in Figure 1. The MOME meta-repository consists of the **Interoperability Database**, which stores descriptions of measurement ***tools***, and the **MOME Database**, which stores descriptions of available measurement ***data***.

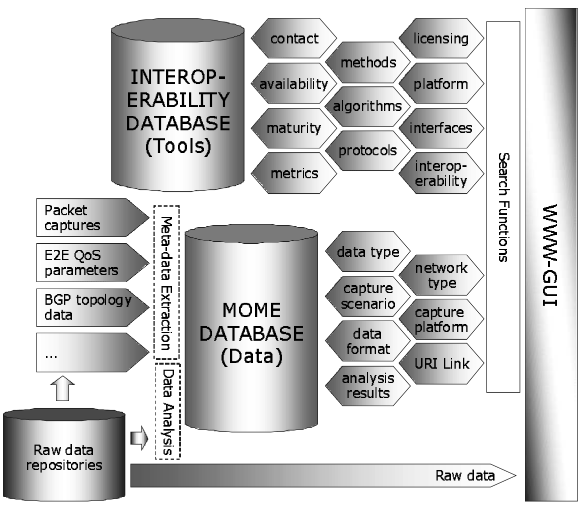


Figure 1. General overview of the MOME meta-repository

This remaining paper is structured as follows. Section 2 provides a short overview of previous works related with measurement meta-data models and meta-databases. Section 3 introduces us to the concept of MOME meta-repository and presents the proposed meta-data models for measurement tools and data. Section 4 documents the implementation of the database installed and maintained by the MOME project. Section 5 presents exemplary usage scenarios, which demonstrate the capabilities of the deployed system. Finally, section 6 concludes the paper.

2 Related works

The flexible meta-databases (i.e. a databases storing ***information about data***) have been recognised by several authors as an effective approach for creating unified access overlay to heterogeneous data.

For example, the measurement meta-database was proposed in [4]. The discussed system, called Scalable Internet Measurement Repository (SIMR), consists of three main components: (1) measurement repositories, which store the actual measurement results, (2) clients, which are the researchers who can download the measurement data with their web-browsers, and (3) a centralized database, which coordinates the access of the clients to the measurement repositories. SIMR is not designed as a public database and assumes accepting submissions only from known researchers.

Similar approach was recognized by the CAIDA project Correlating Heterogeneous Measurement Data to Achieve System-level Analysis of Internet Traffic Trends [3]. The main objective of this project is to foster the progress of measurement-based network research by creation of a meta-data repository that facilitates the access to the distributed raw data repositories. This database, named as the Internet Measurement Data catalogue (IMDC), will assist development of reports and new tools for data analysis and visualisation. Furthermore, a new language for labelling and annotating data sets will be created. At the present time, CAIDA is working at the creation of this central database.

The meta-database approach was also adopted by the Sprint IP Monitoring Project [5]. The goal of the project is to collect very large sets of detailed packet-level traffic traces from a backbone IP network and complement them with equipment configuration files and routing information. The proposed meta-data model is based on four key abstractions of the problem: (1) raw input data sets, (2) analysis programs, (3) result data sets, and (4) analysis operations. However, contrary to the MOME objectives, the project is aimed at the specific needs of the commercial network operator and it is not intended to provide the public exchange platform of captured measurement data.

Notice, that the meta-data approach has been used not only for network measurements and monitoring, but also in other research areas, where coordinated access to large sets of raw data is crucial. For example, the European Grid of Solar Observations (EGSO) [2] project has developed the meta-database as an access overlay to data collected by astronomical observations of the Sun. The suitable meta-data model has been defined and the implemented meta-repository is successfully used for exchanging information about data available from solar observations.

For storing meta-data about measurement tools (e.g. authors, licenses, capabilities, etc.) several platforms are available on the Internet. Therefore for developing the MOME Interoperability database, two options have been considered: implementation from scratch (based on a public domain database), or modifying already existing databases that serve similar purposes. Below we describe some of the systems, which were carefully investigated and evaluated.

Sourceforge [19] aims at supporting collaborative development within the open source community (developers and end-users). It provides a large repository of open source codes and applications, posted freely by projects and users. The advanced features available in Sourceforge are: user registration and authentication, forum system and user management for each project, advanced projects search tools (by category, description and name), information on where to download the code, and more. The Sourceforge source code is available for download and use.

SourceWell [20] is an announcement and retrieval system for software applications. SourceWell concept is based on the same principles (user authentication and authorization system, multilingual support, mailing lists, etc), as Sourceforge. The portal is maintained by a group of developers at BerliOS in Berlin. It aims at providing a neutral platform for the open source community.

The Freshmeat [21] is a free repository of open source software. It is designed to maintain application archives contributed by users. Thousands of Unix-based or other cross-platform applications are catalogued in the Freshmeat database. Links to the application source code are provided, as well as features such as: user management and authentication, category search, different user roles, web-based admin, and more.

The software repositories described above present functionality similar to the MOME Interoperability database [22]. However, due to their complex structure, directly adapting them to the basic requirements of the MOME database design could be quite difficult. Many services provided by those systems were neither mandatory nor useful for our implementation. Therefore, building a database system from scratch was preferred to using available web-based software repository.

3 The MOME meta-data models

In this section we discuss the details of the design of MOME databases. The implementation of databases can be approached either in an application dependent or an application independent way. Uniform data can be handled by an application specific database very well. However, such database is not sufficiently flexible for storing heterogeneous types of data. For example, the term “measurement data” covers collections of packet traces, sets of QoS measurement results, and much more. Thus, the measurement data can be characterized as highly heterogeneous. Databases designed for heterogeneous data will profit from an application independent implementation. The meta-databases are an example of application independent approach.

The first step in designing the MOME meta-repository was to develop appropriate meta-data models for measurement tools and data. The meta-data model consists of a set of attributes, which together describe in detail the actual data of interest. Notice, that the description format should take into account the diversity of types of collected measurement data, as well as of types and capabilities of known measurement tools.

3.1 Measurement data

Currently, many organisations perform network measurements and publish results on their web pages. For assessing, if the considered data is representative and appropriate for using in particular research work, it should be accompanied by the detailed documentation of measurement environment and scenario. The description should include such information, like: the type of measured network, level of traffic aggregation, location and type of measurement equipment, etc. The MOME Database [24][25] aims at storing and providing public access to such information about available measurement data. In this way, MOME assists researchers in finding and retrieving the measurement data appropriate for their purposes, and by keeping the data available for further research.

The list of attributes, which constitute the MOME meta-data model, was defined based on the careful review of the publicly accessible data repositories (see e.g. [6][7][8][9]), and discussions about the specific requirements of the projects cooperating with MOME (e.g. [14]). First, the MOME meta-data model comprises the basic information about the data, like: identifier of the person who performed the measurement, type of the captured data, start- and end- time of the measurement process, submission date, type and size of the raw data file, as well as the link, from which this file can be retrieved. The information of this kind is stored as so-called *common attributes*, which have the same format for each entry annotated in the MOME database.

The selection of *detail attributes*, which describe the measurement environment and scenario, depends of the type of the measurement. In particular, the stored data can consist of: packet-level traces, flow-level traces, HTTP-level traces, measured values of QoS metrics, or collections of routing table snapshots. For each of those types of data, different information is relevant for describing the details of the measurement process. Therefore, different *detail attributes* have been defined for each type of measurement data. Notice, that the database design also allows adding not yet covered types of measurement data in future.

Additional information about the data can be obtained by employing statistical data analysis methods. The analysis provides aggregated information, e.g. in the form of model parameters, and allows us for thorough understanding of the characteristic features of measured traffic. The selected data analysis results are included in the MOME meta-data model as the *analysis attributes*.

Below, we provide the description of *detail attributes* and *analysis attributes*, defined for each of the following types of measurement data: packet traces, flow traces, QoS results, routing data, and HTTP traces.

The **packet traces** consist of time-stamped records of packets, observed in certain passive measurement point. They can be useful for several purposes. First of all, the analysis of packet traces can provide knowledge about the state and availability of network resources, which is essential for efficient traffic engineering and network dimensioning. Moreover, collecting packet traces collected in two measurement points constitutes a base for the passive method of QoS measurement. Finally, in-depth analysis of collected packet traces is necessary for developing and validating analytical models of traffic at the packet level.

The most popular measurement tool, which allows for capturing packet traces, is the *tcpdump* [17]. Using *libpcap*, it collects the packets observed on certain network interface and stores the results in the *pcap* data format. Another method for collecting packet traces assumes attaching the external passive monitor (e.g. the DAG capture card [23]) to the considered link.

The list of attributes, which describe packet traces in the MOME meta-repository, is presented in Table 1. Notice, that the values of *Analysis Attributes* give us basic information about the type, volume and characteristics of traffic, which was captured in the trace.

Table 1. The packet traces attributes

Detail attributes
- Type of network, where the trace was captured (e.g. LAN or WAN), - Location of the capture device, - Type of captured traffic (e.g. if the trace was taken from operational network or from controlled testbed), - Type and speed of the monitored link, - Capture method (e.g. if the trace was recorded directly by the router, or by external capture device), - Filter rules (i.e. if the trace contains all observed packets, or only a subset, defined according to certain rules), - Number of captured packets, - Contents of the record (i.e. if the trace contains entire TCP/IP headers with timestamps, or only selected header fields), - Trace anonymisation rules (i.e. if the IP addresses were removed from the trace or scrambled to protect privacy of users), - Capture platform (e.g. if it was Linux router or DAG card), - Data format (e.g. [17][18]: *Libpcap*, *DAG*, *tcpdump*, *PSAMP*, *sFlow*), - Additional information.
Analysis attributes
- Average bit rate, - Average packet inter-arrival time, packet size and packet rate, - Packet size distribution, - Average bit rate split between different transport protocols and applications, - Time-series of bit rates, calculated over non-overlapping measurement intervals of various length, - Variance of bit rate, - Required value of effective bandwidth, - Required values of traffic descriptor parameters, - Value of the Hurst parameter, which describes the level of self-similarity.

The **flow traces** contain the records of flows (identified by the pairs of IP addresses and port numbers), observed in certain passive measurement point. Flow-level traces are often collected in operational networks for the purpose of usage-based accounting. Moreover, they are important for the research on modelling the traffic on the flow, or connection level. The flow level traces can be collected from the routers using e.g. Cisco *NetFlow* [17], or IPFIX [26] capable tools. Another method assumes filtering the previously captured packet traces and classifying observed packets based on the values of appropriate header fields.

The MOME meta-data model attributes for flow traces are presented in Table 2. The *Analysis Attributes* store basic information about the flow-level characteristics of observed traffic, including its arrival pattern and typical flow sizes.

Table 2. The flow traces attributes

Detail attributes
- Type of network, where the trace was captured,
- Location of the capture device,
- Type of captured traffic (e.g. if the trace was taken from operational network or from controlled testbed),
- Type and speed of the monitored link,
- Capture scenario (e.g. if the trace was recorded directly by the router, or by external capture device attached to the link),
- Filter rules (i.e. if the trace contains all observed flows or only a subset, defined according to certain rules),
- Number of captured flows,
- Trace anonymisation rules (i.e. if the IP addresses were removed from the trace or scrambled to protect privacy of users),
- Capture platform (e.g. if it was Linux router or DAG card),
- Data format (e.g. [15][16]: *IPFIX*, *NetFlow*, *IPDR*)
- Additional information.
Analysis attributes
- Average flow inter-arrival time,
- Average flow duration,
- Average number of packets and bytes in a flow,
- Average flow arrival rate,
- Average bit rate.

The repositories of **QoS results** store the values of metrics of packet transfer quality (i.e. packet delay, packet delay variation, packet loss ratio) between certain two measurement points. The purpose of such measurement is e.g. to evaluate the QoS delivered to the customers of operational networks, or to perform trials for validation of new QoS mechanisms. The raw data sets of QoS results consist of collections of singleton values, obtained for the measurement packets transferred between considered measurement points.

Table 3 presents the attributes, which describe the QoS results in the MOME meta-database. The *Analysis Attributes* expose the aggregate information about the QoS offered to the measured traffic.

Table 3. The QoS results attributes

Detail attributes
- Type of network, where the measurement was performed,
- Measurement method (active or passive),
- Measured metrics (e.g. OWD, IPDV, RTT, packet losses…),
- Location of the measurement points,
- Measurement platform,
- Time synchronization method (e.g. NTP or GPS),
- Number of measured singleton values,
- Data format,
- Additional information.
Analysis attributes
- Minimum, average and maximum of observed packet delays,
- 10- and 90- percentiles of observed packet delays,
- Distribution of observed packet delays,
- Minimum, average and maximum of observed delay variation,
- Packet loss ratio.

The collections of **routing information** consist of the snapshots of routing tables, or records of update messages captured within certain time period. Usually, the routing repositories contain data related with BGP-4 inter-domain routing protocol. The collected information can be used for analyzing the logical topologies of multi-domain networks, detecting the faults and abnormal behaviours of inter-domain routing protocols. The snapshots of BGP-4 routing tables can be obtained with the help of routing collectors based e.g. on the *Zebra* software. The simplest way of collecting routing update messages assumes deploying *tcpdump* (or other packet capture tool), and filtering packets carrying the routing protocol messages.

The attributes defined for routing data in the MOME repository are presented in Table 4. The *Analysis Attributes* include the information about routing table size, which can be useful for studying the trends in growth of advertised network prefixes.

Table 4. The routing data attributes

Detail attributes
- Routing protocol (e.g. OSPF or BGP),
- Type of collected data (i.e. snapshots of routing tables, or collections of routing update messages),
- Location of collector device,
- Data format (i.e. *Zebra* or *Tcpdump*),
Analysis attributes
- Routing table size,
- Number of captured update messages.

The **HTTP traces** consist of the records of HTTP messages captured in passive measurement point, which is usually located in some operational WWW server. The HTTP traces can be collected for the purpose of assessing the load imposed on the web server by the incoming user requests. Modelling traffic at the session level (where session is related with the user activity, resulting in the exchange of HTTP requests and responses) is considered as a difficult task. The HTTP traces can be essential in the work on new session-level traffic models.

The MOME meta-data attributes corresponding to HTTP traces are presented in Table 5. The selected *Analysis Attributes* describe the basic parameters of pattern of session arrivals and server responses.

Table 5. The HTTP traces attributes

Detail attributes
- Location of the capture device,
- Filter rules,
- Number of captured entries,
- Trace anonymisation,
- Capture platform,
- Data format,
- Additional information.
Analysis attributes
- Average HTTP request inter-arrival time,
- Average response size in packets and bytes,
- Average request arrival rate.

The last type of data in MOME repository is covered by the general category of **web-based repositories**. It includes for example the collections of measurement data of different types, commonly stored in a single repository. The attributes describing such repositories include the link to the main web page, name of the providing institution, and the types of available measurement results.

3.2 Measurement tools

The MOME Interoperability database stores the collected meta-data about the measurement and monitoring tools, as well as the links to the original web sites of those tools. The database provides a searchable information base for people looking for a suitable tool for their specific measurement problems. Moreover, it allows projects developing new tools for working on the basis of the existing tools and technologies. Thus, they only need to add specific building blocks that are missing to fulfil the project needs.

The MOME platform will provide assistance for future projects by evaluating the interoperation of components developed by different projects with already existing tools and architectures. The long-term goal of the measurement tools database is to integrate and make available possibly complete information regarding tools interaction and interoperability.

At present, software repositories store information about measurement and monitoring tools from various open source or commercial projects and companies. To ensure that the users are able to obtain and use the software catalogued in the database, detailed and appropriate information should be provided. This will enable users to reuse the tool and adjust it for their specific needs. The attributes such as: list of supported data input and output formats, information on interoperability and measured metrics, will permit the user to combine different tools by identifying possible interactions and interoperability.

Table 6 describes the attributes available for each tool entry. The attributes were defined based on the review of available software repositories, and on the interoperability parameters, developed taking into account specific features of known measurement tools and methods.

Table 6. The tools attributes.

- The tool name,
- The version of the tested tool,
- When the information on the tool was entered in the MOME database,
- When the last edit of the tool information occurred,
- A text describing the tool,
- A link to the tool homepage,
- A link to the author (or contact), if available,
- Related URLs,
- Category: packet capturing, traffic flow measurement, packet monitoring, connection monitoring, application-level monitoring, service monitoring, accounting, intrusion detection, sniffing, performance measurement, connectivity checking, route detection, topology detection, traffic visualization, traffic generation,
- Control input: CLI, GUI, Web interface,
- Data input: live interface, tcpdump packet file, other packet file, raw data files, csv text data files, text via stdin,
- Metrics measured by the tool: one-way-delay, round-trip-delay, jitter/delay variation, throughput, packet loss, etc.,
- Data output format: file, graphs, etc.,
- The time scope at which the tool works: real-time, seconds, minutes, etc.,
- Filter Attributes: what the tool can filter on,
- How data can be aggregated,
- Sampling schemes,
- Availability: freeware, open source, commercial, etc.,
- What type of license(s) the tool is available with,
- Which hardware the tool runs on,
- Which operating systems does the tool support,
- Which standardised metrics does the tool support,
- Which standardised protocols does the tool support,
- With which other tools this one can be used (information on interoperability),
- What is the schedule of data export,
- Additional features.

Additionally, the MOME tools database can store information on the performed tests of particular tool. The tested features can include for example the performance of the tool, or interoperability with another tool. The test attributes include: the identifier of the tester, the identifier of tool tested, the tool details (tool version), and the testing platform. The testing platform is described by such attributes, as: the operating system used, the system details, the hardware used, the network on which the tests have been performed, and some information regarding the test conditions. Finally, the results of the tests are also stored in the MOME Interoperability database.

4 Implementation of the MOME meta-repository

The meta-data models described in section 3 have been used as a base for the implementation of the MOME meta-repository [22][25]. The successful deployment of the MOME repository confirms, that the defined meta-data models adequately describe the available measurement tools and data. Figure 2 shows the implemented components of the meta-repository, which has been installed as an integral part of the MOME website hosted at the TERENA data centre in Amsterdam.

As web server, the original MOME server was used. It already hosts the MOME web site together with the survey tool, used for collecting information from associated projects. The web server implements the Graphical User Interface (GUI) providing access to all functionalities of the MOME Workstation. The user accesses the system through a standard browser by a link from the MOME web-site [1]. The GUI functionality is implemented in PHP language at the server side.

The objective of implementing the web-based user interface is to give the user a clear and complete view of the main features of the MOME platform. Thus, the interface helps the user understand at first sight the basic functionalities of the MOME database. Advanced functionalities are accessible to registered users. User registration is free of charge and can be done via the web interface. The location of the MOME database is: http://www.ist-mome.org/database/.

The data displayed in the GUI is taken from a database (*'mome'*), which is located on a separate machine (shell.ist-mome.org) and uses MYSQL as the database management system. This database stores meta-data information on the measurement data, as well as information on measurement tools. Additionally, raw data files can be locally stored on the same machine in a separate space (*'file server'*). The GUI also enables user access to raw data of third parties, like e.g. public repositories, and publishes measurement data from associated projects. Further, it is possible for associated measurement sites to upload measurement data and/or measurement meta-data to the MOME database automatically. Database access is implemented using the server-side script language PHP, which provides the integration between the database and the web-server.

Additionally, the MOME platform offers capabilities for performing selected analysis tasks on the measurement data annotated in the meta-repository. The user can request the analysis of measurement data by selecting appropriate option in the GUI. The MOME Workstation retrieves the raw data from its actual location, executes the external analysis tool, processes its output files and writes results into appropriate fields in the MOME Database. Since the analysis tasks are performed by external tools, adapters to the MOME system must be implemented by appropriate filters. As a first step, the analysis tool called *tcpdstat* was integrated in MOME, for performing basic statistical analysis of packet traces collected with *tcpdump*. The available *Analysis Attributes* are: average bit and packet rate, distribution of packet sizes, distribution of bit rate of traffic associated with different protocols and applications.

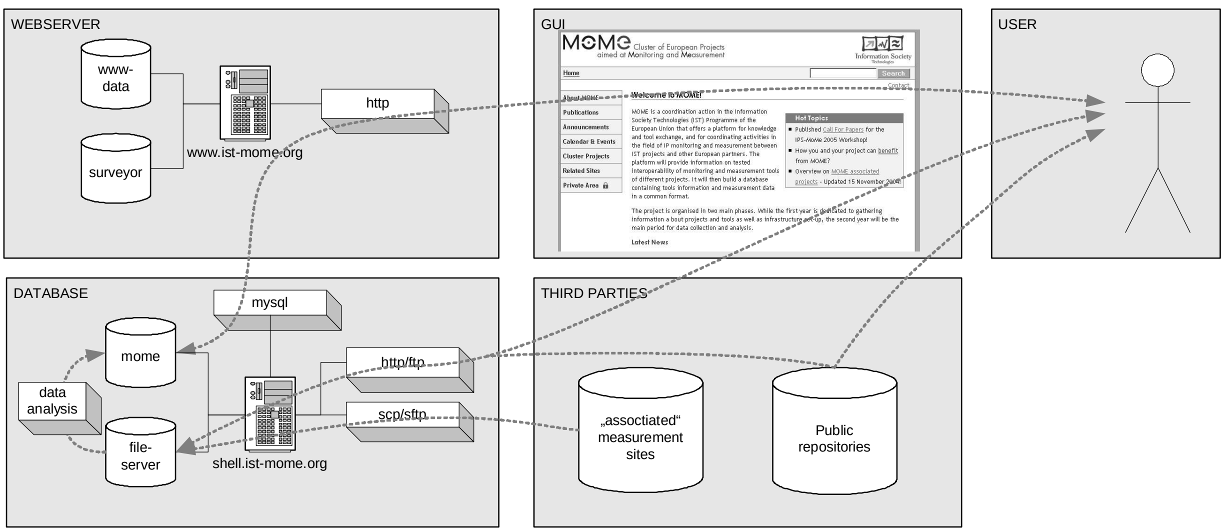


Figure 2. Deployment of the MOME meta-repository

5 Exemplary usage scenarios

In this section we demonstrate the functionalities available for users of MOME meta-repository in the following typical scenarios: browsing the meta-database, submitting new entries, analysing raw data, and performing system maintenance tasks. Access to particular functions depends on the privileges assigned for the user. The MOME database interacts with three kinds of users:

- *Not-registered users*: users with no relationship to MOME or its associated projects. Unregistered users can browse the database, but they are not authorised to submit new entries. Unregistered users are encouraged to register, to fully benefit from the database and share their data with other users.
- *Registered users*: users interested in the work of MOME and/or its associated projects, which contribute to the database and use it to retrieve measurements from other projects.
- *System administrator:* keeps the MOME database alive, by performing periodical checks and maintenance operations.

5.1 Browsing the MOME meta-repository

All users (registered and not-registered) are authorised to browse and search the MOME meta-repository. The access to the database is available via the public web page of the MOME project. The introductory screen describes the basic capabilities of the MOME server, and explains the offered service. A FAQ (Frequently Asked Questions) page is also available.

Browse the databases

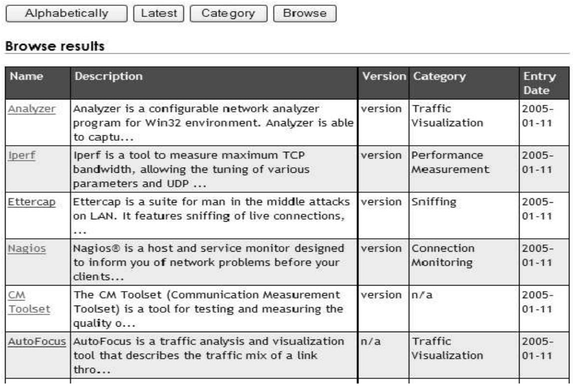
This function allows the user to get a quick overview of the contents of the MOME databases. The *"overview"* screen (see the exemplary view of the tools database on Figure 3) shows only the most important of the meta-data attributes: name of the tool or trace, its category, submission date, and a short description. When the user clicks on a selected entry, he is directed to the detailed view for this entry. Here, he can see all the meta-data attributes, which describe particular tool or data.

Search the databases

This function allows the user for searching the MOME repository for entries matching specific user-defined criteria, submitted in the form of text fields or drop-down lists. For example, the user can search for measurement data of particular type, or for a tool, which measures particular type of metric.

Download raw data

The MOME database allows the user for accessing the actual raw measurement data. He can retrieve the original data by clicking on the link, which is stored in the MOME database as one of the meta-data attributes. In the same way, the user can access the original web page associated with particular measurement tool. The download events are logged in the MOME system, for providing statistics about "most popular" measurement tools and traces.


Alphabetically | Latest | Category | Browse

Browse results

Name	Description	Version	Category	Entry Date
Analyzer	Analyzer is a configurable network analyzer program for Win32 environment. Analyzer is able to captu...	version	Traffic Visualization	2005-01-11
Iperf	Iperf is a tool to measure maximum TCP bandwidth, allowing the tuning of various parameters and UDP ...	version	Performance Measurement	2005-01-11
Ettercap	Ettercap is a suite for man in the middle attacks on LAN. It features sniffing of live connections, ...	version	Sniffing	2005-01-11
Nagios	Nagios® is a host and service monitor designed to inform you of network problems before your clients...	version	Connection Monitoring	2005-01-11
CM Toolset	The CM Toolset (Communication Measurement Toolset) is a tool for testing and measuring the quality o...	version	n/a	2005-01-11
AutoFocus	AutoFocus is a traffic analysis and visualization tool that describes the traffic mix of a link thro...	n/a	Traffic Visualization	2005-01-11

Figure 3. Browsing the meta-repository of measurement tools

Send feedback

The user can send a message to the originator of the raw data, or of the tool, that he has worked with. The messages are stored in the MOME repository as *user comments* and can be browsed by all users viewing particular entry in the database.

5.2 Submitting new meta-data entry

New entry can be added to the MOME meta-repository when new measurement tool is developed, or new measurement data is available. Only the registered users are authorised to add new entries to the MOME meta-repository. This restriction is needed as some kind of protection against inappropriate uses of the database. However, anyone can freely register in MOME. For setting up a new account, he just has to choose the unique username and password, and provide valid e-mail address.

Submit data

The registered user can submit new meta-data entry. First, the system verifies if the user is allowed to post the data to the MOME repository. Then, he can enter the values for all the attributes, which were defined in the meta-data model for the chosen type of measurement data. Figure 4 demonstrates the proposed concept, by showing the *common* and *detail* attributes submitted for the exemplary packet-level trace available from the well-known public repository [6]. One can observe, that the description of the trace is presented to the visitor of MOME database in a compact, easily readable form, and includes the essential information needed to assess the trace.

Modify data

The owner of the meta-data (i.e. the person who originally added the entry) is allowed to modify to the submitted attributes. The functions performed by the system when the user modifies the submitted data are similar, as in the case of previously discussed scenario for adding new database entry.

Dataset name	Bellcore BC-pAug89
Data type	PacketTrace
File size	5B
File compression	TL.Z
Start time	1989-08-29 11:25:00
End time	1989-08-29 11:25:52
Duration	0:00:52
Description	Trace of about 1 million Ethernet packets, captured at the Bellcore Morristown Research and Engineering facility
Dataset location	URL: ftp://ita.ee.lbl.gov/traces/BC-pAug89.TL.Z
MD5 Sum	n/a
Tool	n/a
Submitted by	mdabrowski @ 2005-01-26 21:11:38
Last Update	2005-01-26 21:11:38

a)

Network Type	LAN
Collector Location	Bellcore Morristown Research and Engineering facility, Morristown, US
Traffic Type	operational network, mostly local LAN traffic
Link Protocol	Ethernet
Capture Mode	n/a
Filter Rules	no filter
Number of Packets	1000000
Trace Anonymisation	n/a
Capture Platform	n/a
Data Format	The traces are in 2-column ASCII format, twenty bytes per line (including the newline). The first column gives the time in seconds since the start of the trace. The second column gives the Ethernet data length in bytes, not including the Ethernet preamble
Additional Info	Thetimestamps are reported to 6 decimal places, they have 4-microsecond precision, and further analysis indicates that the actual accuracy is about 10 microseconds (primarily due to bus contention).

b)

Figure 4. Meta-data entry for exemplary packet trace: a) common attributes, b) detail attributes

5.3 Analysing raw data

The owner of the entry in the MOME Database can initiate statistical analysis of a raw measurement data described by this entry. Selected analysis tools are integrated in the MOME Data Analysis Workstation and allow for performing typical analysis tasks on data stored using the most popular measurement data formats. Currently, the basic analysis of packet traces collected with *tcpdump* is supported. The user can choose the analysis procedures, which should be applied to the raw data. Then, the MOME Data Analysis Workstation executes the external analysis tool. After finishing the analysis task, the results are inserted into appropriate fields (*analysis attributes*) in the MOME measurement database.

5.4 System administration and maintenance

The administrator has access to all the functionalities available for not-registered and registered users. In particular, he is allowed to view, modify and delete any meta-data entry in both databases. Additionally, the administrator has privileges for performing maintenance functions, like checking the status of the database. The status information includes e.g. the statistics regarding the number of registered users, stored meta-data entries, and usage of resources (disk space). Additionally, the statistics corresponding to the execution of the analysis tasks, like e.g. the number of completed tasks and average analysis time, are also available.

6 Conclusions

Nowadays, cooperation between Internet Providers in order to share network measurement information is very low. Some factors are important to understand this issue. On the one hand, competitive providers invest huge amounts of money to increase networks capacity instead of investing to know how that capacity is utilised and trying to improve that use. This more engineering-oriented solution involves more resources and time. This lack of cooperation matches in time with the growth in the number of applications using distributed data and computing resources throughout the Internet.

In this paper, we showed how the study of the state of the art has revealed that the main problematic of coordinating Internet measurements is heterogeneity. This is due to the myriad of tools to collect and process network measurements, which have been proven to be relevant by their day to day use. This study has also revealed that there are projects trying to establish an umbrella to correlate Internet measurements. These projects are in a very early stage and haven't produced any database definition, which MOME could reuse. The study of the state of the art in multidisciplinary measurement repositories has shown a common denominator: the use of the meta-data paradigm. Traces are no longer stored locally by the projects. Instead, they provide value added data, which includes pointers to publicly available traces. This is the approach followed by the MOME project. We also presented in this paper the resulting structure for the MOME metadata database, the intertwining of the tools and measurement databases, which finally led to a unique database design for all workpackages in the project. To give a more practical view of the MOME database, we also have shown the different functions provided to a MOME user.

In our opinion, the meta-database architecture is the best way to achieve the goal of creating an access overlay to the measurements gathered. In addition to the measurements proper, the meta-database model provides links to tools, which can be used to analyse the data. The added value of MOME is to provide this coordinating layer, not to provide the actual storage space.

7 Acknowledgments

The work was supported by the IST Coordination Action MOME (*Monitoring and Measurement Cluster*), IST-001990.

8 References

[1] IST-MOME (Monitoring and Measurement Cluster) project web site, *http://www.ist-mome.org*

[2] B. Bentley et. al., "EGSO - The European Grid of Solar Observations", in *Proceedings of the 10th European Solar Physics Meeting*, Prague, September 2002

[3] CAIDA "Correlating Heterogeneous Measurement Data to Achieve System-Level Analysis of Internet Traffic Trends" project, *http://www.caida.org/projects/trends/*

[4] M. Allman, E. Blanton, W. Eddy, "A Scalable System for Sharing Internet Measurements", in P*roceedings of Passive and Active Measurements*, PAM 2002, Fort Collins, March 2002

[5] S. B. Moon, T. Roscoe, "Metadata Management of Terabyte Datasets from an IP Backbone Network: Experience and Challenges", in *Proceedings of ACM SIGMOD Workshop on Network-Related Data Management*, Santa Barbara, May 2001.

[6] The "Internet Traffic Archive" web site, *http://ita.ee.lbl.gov/index.html*

[7] The "NLANR Network Traffic Packet Header Traces" web site, *http://pma.nlanr.net/Traces/*

[8] The "NLANR Measurement and Network Analysis" web site h*ttp://watt.nlanr.net/active/maps/ampmap_active.php*

[9] The "RIPE NCC Routing Information Service Raw Data" web site, *http://data.ris.ripe.net/*

[10] I. Miloucheva et al., "Intermon architecture for complex QoS analysis in inter-domain environment based on discovery of topology and traffic impact", in *Proceedings of the International Performance and Simulation Workshop*, IPS2004, Budapest, March 2004

[11] T. Engel et al., "AQUILA Adaptive Resource Control for QoS Using an IP-based Layered Architecture", *IEEE Communications Magazine*, Vol. 41 No. 1, January 2003

[12] A. Bąk, W. Burakowski, F. Ricciato, S. Salsano, H. Tara-siuk, "A framework for providing differentiated QoS guar-antees in IP-based network", *Computer Communications* 26 (2003), Elsevier, pp. 327-337.

[13] The IST-EuQoS (Endt-to-end Quality of Service support over heterogeneous networks) project, *www.euqos.org*

[14] The IST-EuroNGI (Design and Engineering of the Next Generation Internet) project web site, *http://eurongi.enst.fr/en_accueil.html*

[15] J. Coppens, E.P. Markatos, J. Novotny, M. Polychronakis, V. Smotlacha, S. Ubik, "SCAMPI - A Scaleable Monitoring Platform for the Internet", in *Proceedings of the International Performance and Simulation Workshop*, IPS2004, Budapest, March 2004

[16] The IST LOBSTER project web page, *http://www.ist-lobster.org/*

[17] C. Schmoll et al., "State of Interoperability", *Deliverable D11 of IST MOME*, June 2004

[18] J. Quittek et al, "MOME Standardisation Plan and Recommendations", *Deliverable D31 of IST MOME*, October 2004

[19] The Sourceforge web site, *http://sourceforge.net*

[20] The Sourcewell web site, *http://sourcewell.berlios.de*

[21] The Freshmeat web site, *http://freshmeat.net*

[22] C. Schmoll et al., "MOME Interoperability Database", *Deliverable D12 of IST MOME*, June 2004

[23] The ENDACE web site, *www.endace.com*

[24] P. Aranda Gutierrez et al., "MOME Database", *Deliverable D21 of IST MOME*, September 2004

[25] P. Aranda Gutierrez et al., "MOME Data Analysis Workstation", *Deliverable D22 of IST MOME*, September 2004

[26] J. Quittek, T. Zseby, B. Claise, S. Zander, "Requirements for IP Flow Information Export (IPFIX)", *RFC 3917*, October 2004