Network Measurement Virtual Observatory: An Integrated Database Environment for Internet Research and Experimentation

Tamás Sebők, Zsófia Kallus, Sándor Laki, Péter Mátray, József Stéger(✉), János Szüle, László Dobos, István Csabai, and Gábor Vattay

Department of Physics of Complex Systems, Eötvös Loránd University, Budapest, Hungary
{sebok,kallus,laki,matray,steger,szule,dobos, csabai,vattay}@complex.elte.hu

Abstract. To understand the long-term dynamics of networks engineers and network scientists collect tremendous amount of data and distribute them across many different data warehouses. In EU FP7 OpenLab project we developed the nmVO, which helps handling distinct data sources together in a common way efficiently. It also supports data collecting systems with a permanent data storage, such as SONoMA, and provides a public front-end to run measurements and access data, called GrayWulf. Furthermore, the on-line analysis of data, yielding the behavior and the structure of the Internet is convenient by using server side scientific functionalities.

Keywords: Database federation · Network measurement data · Virtual observatory

1 Introduction

The data tsunami of the last decade forced scientists to find new ways of dealing with data. It turned out early that the same data management problems and solutions can be shared among very distant fields of science disciplines. *Virtual Observatories* (VO) appeared originally for astronomical data [1] and later in other fields of science [2] to make the multi-terabyte science archives manageable and, most importantly, to make them accessible for researchers. As large archives of data became available on the Internet an obvious step forward was to try and federate these distant databases and provide a unified, searchable view of them to extract aggregated information. VOs provided a simple way to share data among members of research groups world-wide. Processing the unprecedented amounts of data required new techniques and relational database management systems have become an every day tool of astronomers, geophysicists, network scientists, biologists, etc. Soon data analysis kits also became part of VOs driven by the realization that in many cases the computation is much easier taken to the data than the data be downloaded to operate on. While certain computations can be

V.C.M. Leung et al. (Eds.): TridentCom 2014, LNICST 137, pp. 65–74, 2014.
DOI: 10.1007/978-3-319-13326-3_7

formulated in SQL, the lingua franca of VOs, other problems require extensions, preferably accessible from the same SQL interfaces. Recently, network research has been facing with very similar issues.

In this paper, we present the Network Measurement Virtual Observatory[1] (nmVO) designed to facilitate Internet related network sciences. Since it's first variant [3], which was developed mostly on the basis of existing VO technologies, it has gone through radical changes, becoming an integrated database environment covering various new functionalities from permanent object storage and data access capabilities to the federation of heterogeneous remote SQL-based data sources. In OneLab [4] project[2], nmVO has become the primary data federation tool, through which all kind of network measurement data collected by different research groups and stored in heterogeneous databases like PostgreSQL, MySQL or SQLServer can be accessed, analyzed and visualized with ease.

The rest of the paper is organized as follows. In Section 1.1 we give a brief introduction to the field of Internet network research. Section 1.2 presents SONoMA, the web-service-based abstraction layer developed to help execute experiments with a heterogeneous measurement infrastructure. The concept of nmVO is explained in Section 2.

1.1 Internet Measurements and Data Archives

Since the 1960's, Internet has gone through an enormous evolution, becoming one of the most complex artificial systems in the world. Besides the growth in its size, the high number of users and various applications generate huge amount of network traffic to be handled everyday. As a consequence, traffic control, forecasting, performance analysis and monitoring are becoming fundamental issues for network operators and interesting targets for researchers as well.

Similarly to other scientific areas, different network research groups make significant efforts to examine one or another aspects of this global system, using high variety of measurement tools and analysis methodologies. However, in order to reveal real dependences between various mechanisms and to obtain a global view on how the system works the data of different aspects need to be handled together. For example, queuing delay tomography [5] or loss inference methods require both topology and one-way delay or packet loss measurements. Inspired by this idea, the nmVO was established based on existing VO solutions [3] in 2007.

Since then, only few other efforts can be found in the literature that aim at helping the network research community with shared data. Many research groups publish their measurement results in raw files with various formats. Furthermore, there are only few attempts towards the standardization of network measurement data formats. For this, some papers propose JSON and XML [6,7] since they are flexible and descriptive enough, while others prefer ontology-driven semantic representations [8,9]. Recognizing this need, CAIDA[3] has created an Internet

[1] http://nm.vo.elte.hu
[2] http://www.onelab.eu
[3] The Cooperative Association for Internet Data Analysis - http://www.caida.org

measurement data meta catalog, called DatCat[4], which is basically a searchable registry of dataset descriptions. It helps to find, annotate, cite and publish data contributed by others. DatCat also contains detailed descriptions of data sets including their location, reproducibility, formatting, etc. In some cases, the raw files are also available at DatCat servers, while at other cases only a link points to the real location. The key difference between DatCat and VO concepts is that VOs aim at offering a unified SQL-based interface to query and fetch data, while DatCat does not deal with this issue at all.

1.2 The Service Oriented Network Measurement Architecture

Distributed network measurements are essential means to characterize the structure, the dynamics and the operational state of the Internet. Although in the last decades several measurement and monitoring systems have been created, the easy access of these infrastructures and the orchestration of complex measurements were not solved. In 2010, we laid down the basis of a network measurement framework, called SONoMA [10], serving originally the natural needs of the network measurement community. The SONoMA provides easy-to-use web services, see Fig. 1, to carry out large-scale network measurements from heterogeneous networking elements including PlanetLab[5] nodes, BlackFin-based APE boxes and Etoms[6]. This approach has opened the door to perform atomic and complex network measurements in real time, furthermore, it automatically stores the measurement results in nmVO.

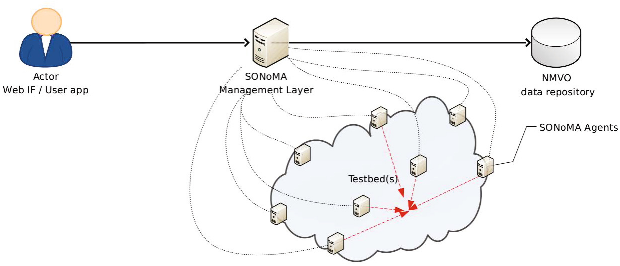

Fig. 1. The schematics of main SONoMA components and their control interactions

Recently, numerous new features have been added to SONoMA, especially focusing on making it more flexible, easier to extend and to support infrastructure monitoring. To this end the following enhancements have been applied:

4 http://www.datcat.org
5 http://www.planet-lab.org
6 http://www.etomic.org

1. *Flexible tool extension:* The new measurement agents decouple the control and the implementation of tools. Now standard tools are also supported via drivers speaking different protocols (like SOAP, REST, SSH). E.g., measuring available bandwidth via `iperf`, getting *RTT* using `fping` or reading cpu load and memory usage from `/proc` are made easy.
2. *Harmonization of network tools and network measurements:* Various testbeds and devices could provide the same metric by different tools or their outputs may differ. In the new framework the operator defines a mapping for each tool, indicating metric types and units properly. Two configuration schema have been investigated and tested: a semantic approach and a close-code implementation.
3. *Periodic measurements:* In the former model of operation, measurements were carried out on demand, whereas now the definition of periodic and continuous measurements are also possible. In this way, the platform can be instrumented to provide their users with accurate and up-to-date information on the available resources.
4. *Permanent storage of data:* One of the key advantage of the SONoMA system is that all the measurement data are automatically stored in a permanent repository, in nmVO, enabling researchers to query and analyze their data back in time via a SQL-based querying interface. To avoid data losses, the new SONoMA variant does it more efficiently in a two-stage fashion. Collected data is first stored in a temporal SQLite repository files, serving as a fast first stage database. The records from this database are then transferred to the permanent database of nmVO. The new schema is made extensible, the back-end schema also flexible, supporting the extension with new metrics and tools.
5. *IPv6 readyness:* The problem of IPv4 address space exhaustion, made us investigate the IPv6 capability of SONoMA. The new configuration mapping schema enables for inclusion tools operating in IPv6 world. Also the web service back end supports control calls over IPv6.

To demonstrate the potential of SONoMA we built and operate *Spotter*[7], a geolocation service, which uses measurement agents of known GPS coordinates to localize arbitrary IP addresses using round trip delay information. With this service it possible to tell the location of a computer with an accuracy of a few ten kilometers.

2 Network Measurement Virtual Observatory

Historical experimental data are essential to understand the long-term dynamics and structure of the Internet. There have been numerous projects by large collaborations and small research groups to measure and analyze a wide range of network parameters. The collected data are amassed in archives dispersed around the world in incompatible data formats, often not even accessible on-line. One of

[7] http://spotter.etomic.org

the main goal of nmVO is to federate and/or co-locate these datasets and build an on-line data warehouse for network scientists and other experimenters. This data warehouse, however, is much more dynamic than most science archives. Network experiments consists of large series of micro-measurements (thousands or millions of pings and traceroutes), even conducting the experiments requires a large amount of initial data. Raw numbers from micro-measurements are ingested into the central nmVO archive before analysis. This not only allows data analysis programs to leverage the functionalities provided by the database server, but also to re-analyze the data later, either to verify earlier results, or to apply more sophisticated algorithms and validate new models.

2.1 The nmVO Infrastructure

In the early stages, nmVO benefited a lot from open software developed for existing VO solutions in different fields of science. The SQL query batch framework and user interface of the database system, called CasJobs, was originally borrowed from SkyServer [11], a database containing astronomical data. However, during its operation nmVO's weaknesses and shortcomings were recognized inspiring us to rethink the whole architecture from scratch. During this process, the usage requirements were also identified and taken into account, leading to a more flexible and complex integrated database solution for accessing, federating, analyzing and visualizing data related to Internet research and experimentation. Our aim was to create an integrated database solution whose benefits can be exploited by both testbed users, network scientists, tool and infrastructure providers, as well.

Figure 2 depicts the logical configuration of nmVO. One can observe that the heart of this system consists of three main components: A cluster of local database servers storing measurement data, an object store component and a data federation, analysis and visualization tool called GrayWulf that has fully replaced the CasJobs-based solution detailed in [3]. One can also observe that nmVO can provide and access to other remote data sources and tools. The integration of remote databases require only a few configuration steps that can be done through a web interface easily. After that the remotely stored data can be accessed through GrayWulf's web interface. The figure also shows the system provides each user with a small database (MyDB) where the results of user queries can be stored for further analysis or visualization. To demonstrate how external tools can be integrated into the system, we extended the local nmVO database with the capability to call SONoMA[8] [10] if the requested data is not available in the local database. Specific stored procedures have been implemented, calling the web-service methods of SONoMA to carry out ping, traceroute and other measurements. Thus, besides accessing existing data stored in the local repository, nmVO can automatically collect data from network measurement tools in a transparent way, by submitting an ordinary SQL-query.

[8] http://sonoma.etomic.org

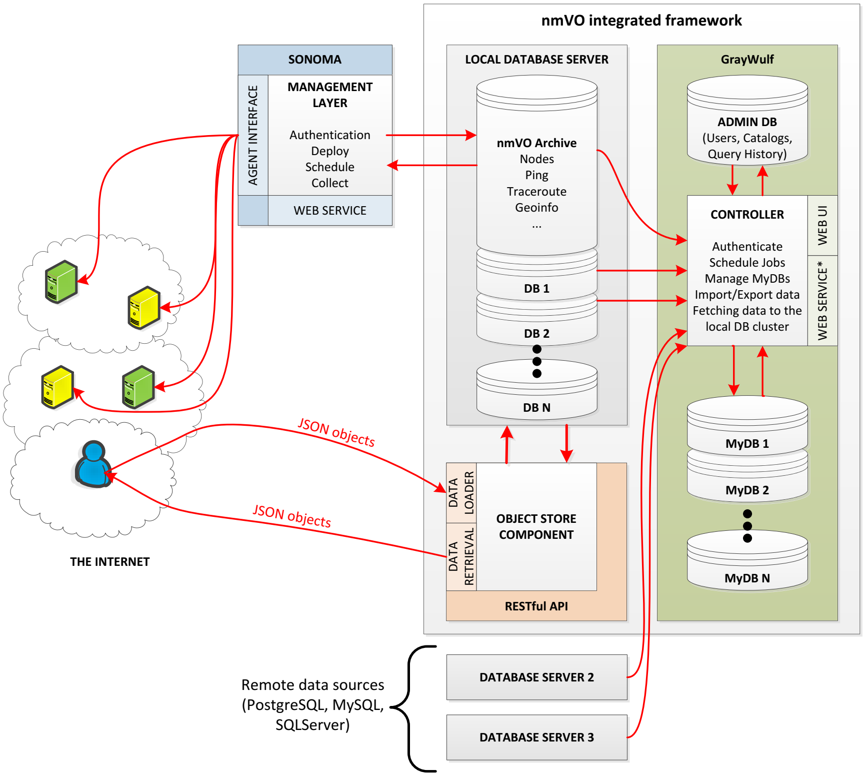

Fig. 2. The complex architecture of the nmVO environment

Finally, the object store component offers a RESTful interface to submit any kind of JSON objects for storing permanently in the local database of nmVO. Currently, Packet Tracking [12] and periodic `iperf` measurements are stored like this, but other types can easily be covered.

2.2 The Local nmVO Database

We have designed and built a multi-terabyte relational database to store archival and recent network measurement data including raw measurements and results from data analysis. The nmVO database [3] is organized into *Collections*. Collections group together measurements, sub-measurements and analyzed results belonging to the same experiment. A single experiment usually consists of thousands of micro-measurements. For example, in case of a geolocation experiment, in which we want to determine the most likely geographic coordinates of a given host, the network topology around the host has to be mapped with traceroute

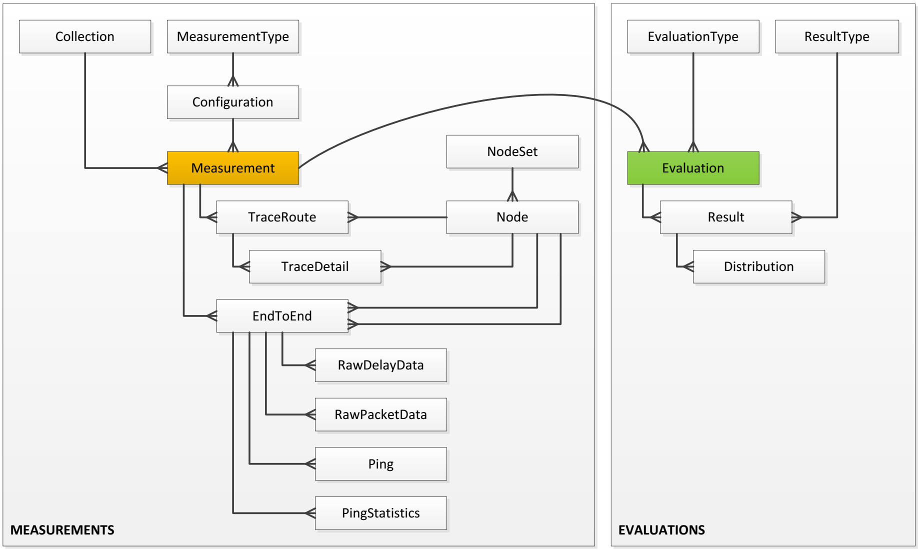

Fig. 3. A high-level view of the nmVO database schema. Measurements are organized into collections. Both raw measurements (Ping table) and results of analysis (PingStatistics table) are stored.

measurements, then delay roundtrip times along various routes have to be measured. Collection may contain multiple evaluations of the same raw data, usually based on different models, methods or initial parameters.

The two basic types of measurements the nmVO database stores are end-to-end measurements consisting of ping delay times and traceroutes consisting of router chains. For ping delays, raw measurements and aggregated delay times with statistics are stored.

The nmVO database is tightly integrated with a series of easy-to-use on-line network data analysis tools, many of them accessible directly from the database using SQL, or through intuitive web-based user interfaces. Figure 3 shows a high-level overview of the nmVO database schema.

Using SQL queries answering topology related questions is vastly simplified compared to any file-based approach. For example, one can easily determine the number of discovered paths between two given nodes, or the nature of route changes in a given network segment and time interval.

2.3 Federating Network Measurement Archives

Network experiments have been conducted by various groups around the world for years. Long-term dynamical analysis of the Internet is impossible without these data. The main issue of federation of data is their format. Historically, network scientists collected raw data in huge file, often storing redundant information using lengthy data models. To build the foundation of a federated system for network measurement, we designed a clear and comprehensive database schema to store raw data from a variety of typical network experiments.

Our system currently covers different databases with different database management systems, like Tophat which uses PostgreSQL or ETOMICDB which uses MySQL. The federated data set is reachable via a web interface, where the users easily can handle all of the joined databases and create complex queries and cross-joins.

2.4 Data Access and Query Processing

Public access to the nmVO database is provided via GrayWulf. The data reduction and analysis tasks are formulated as SQL queries. This way scientist can easily delegate all processing to the nmVO servers where data are co-located. The GrayWulf infrastructure is hosted by a dedicated server containing a database for the SQL query batch service, the batch service itself, and sandbox databases of registered users called MyDBs.

Besides local, co-located databases, nmVO can also connect to other, remote databases. Query results from remote databases get stored in the MyDB and can be joined with data from the local data sets. Users can also upload their own data to MyDB, and download data in various data formats.

An example query shows how to join tables from databases (`nmVO:` and `TOPHAT:`) with distinct locations. So location information is fetched from the local nmVO whereas path length is retrieved from the remote data source.

```
SELECT t.agent_id as src, t.destination_id as dst, t.hop as ip,
     t.hopcnt as hopcnt, r.lat as lat, r.lng as lon
 FROM TOPHAT:traceroute_hops t
 INNER JOIN nmVO:geo.routercity r ON t.hop = r.ip_string
 WHERE t.first between '2012-06-07 0:0' and '2012-06-08 0:0'
```

2.5 Data Visualization

The best way to understand our data is the visualization. We built in a gnuplot based visualization tool, so the user can plot directly from database. It uses the servers resources, the plot displayed in the web interface either in HTML 5 canvas or one of many commonly used graphic format, such as postscript, PNG or JPEG. The visualization tool has 2 basic functions, plotting curves and histograms. Figure 4 shows a simple histogram made by the visualization tool.

2.6 Interactive Experiments

Since in nmVO databases are closely bound with the measurement infrastructure (SONoMA, Spotter) there is not anything to prevent us from initiating measurements directly from the database server, using SQL! Being able to conduct experiments straight from SQL scripts has a great potential. For example, missing data can be gathered on fly. To demonstrate how to integrate web services with databases, we wrapped the SONoMA ping web service interface into a SQL user-defined function.

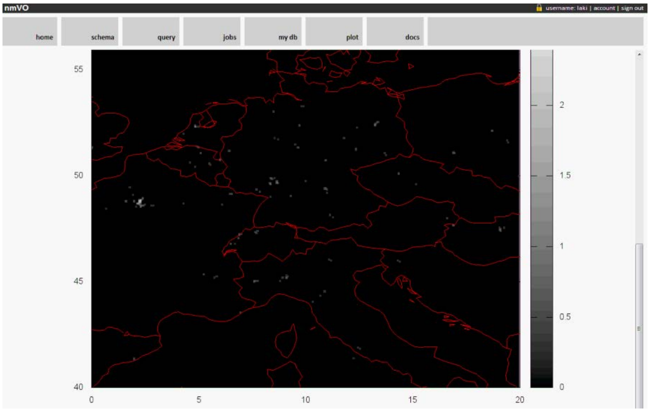

Fig. 4. Histogram of geographical coordinates made by the gnuplot based built in visualization tool

3 Summary

We have presented a brief introduction to the world of network measurements. To get a better understanding of the data network scientists use, first we introduced the methods and the basic infrastructure network measurements are done with. SONoMA, our open measurement management service was explained to show how time and effort can be reduced with a Virtual Observatory infrastructure to carry out otherwise complicated complex experiments. SONoMA, while remains easy to use, contains all the tools and has access to all the important testbeds that network scientists need. Building application on the basis of SONoMA is simple thanks to its web service interfaces. SONoMA is deeply integrated with the nmVO database to automatically store and publish measurement data.

We have introduced the concept of nmVO, a Virtual Observatory for Internet measurement data and demonstrated how technology developed for other fields of science can be reused to build scientific data repositories with minimal effort. The most important feature of nmVO is, however, that the database is tightly integrated with the measurement infrastructure. Tools work from data stored in the database and save results there. The nmVO database can be used as a cache to look up whether certain measurements have been done earlier, and reuse them if possible, saving on network traffic and measurement time. Archived measurements also allow for time-domain investigation of the network.

Another interesting feature of nmVO is that certain measurement tools are readily available from the SQL interface: experiments can be executed on the fly by calling stored procedures. This makes nmVO more like a real observatory than a virtual one.

Acknowledgments. The authors thank the partial support of the EU FP7 OpenLab project (Grant No.287581), the European Union and the European Social Fund through project FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013), the OTKA 7779 and 103244, and the NAP 2005/KCKHA005 grants. EITKIC_12-1-2012-0001 project was partially supported by the Hungarian Government, managed by the National Development Agency, and financed by the Research and Technology Innovation Fund and the MAKOG Foundation.

References

1. George Djorgovski, S., Williams, R.: Virtual observatory: From concept to implementation. CoRR, abs/astro-ph/0504006 (2005)
2. Spaniol, M., Benczúr, A.A., Viharos, Z., Weikum, G.: Big web analytics: Toward a virtual web observatory. ERCIM News **89**, 2012 (2012)
3. Mátray, P., Csabai, I., Hága, P., Stéger, J., Dobos, L., Vattay, G.: Building a prototype for network measurement virtual observatory. In: MineNet, pp. 23–28. ACM (2007)
4. On federations..., (2009)
5. Rizzo, T., Steger, J., Péter, P., Csabai, I., Vattay, G.: High quality queueing information from accelerated active network tomography. In: TRIDENTCOM (2008)
6. Marcelo, B., et al.: Standardizing large-scale measurement platforms. ACM SIGCOMM Computer Communication Review **43**(1), 58–63 (2013)
7. Dhawan, M., et al.: Fathom: a browser-based network measurement platform. In: Proceedings of the 2012 ACM Conference on Internet Measurement Conference, ACM (2012)
8. van der Ham, J., Stéger, J., Laki, S., Kryftis, Y., Maglaris, V., de Laat, C.: The novi information models. Future Generation Computer Systems (2013)
9. Mátray, P., Csabai, I., Hága, P., Stéger, J., Vattay, G.: A semantic extension of the network measurement virtual observatory. Advances in Databases and Information Systems (2009)
10. Hullár, B., Laki, S., Stéger, J., Csabai, I., Vattay, G.: SONoMA: A service oriented network measurement architecture. In: Korakis, T., Li, H., Tran-Gia, P., Park, H.-S. (eds.) TridentCom 2011. LNICST, vol. 90, pp. 27–42. Springer, Heidelberg (2012)
11. Szalay, A.S., Gray, J., Thakar, A.R., Kunszt, P.Z., Malik, T., Raddick, J., Stoughton, C., vandenBerg, J.: The SDSS skyserver: Public access to the sloan digital sky server data (2002)
12. Santos, T., Henke, C., Schmoll, C., Zseby, T.: Multi-hop packet tracking for experimental facilities. SIGCOMM Comput. Commun. Rev. **40**(4), 447–448 (2010)