Management of the Internet and Complex Services

European Sixth Framework Network of Excellence FP6-2004-IST-026854-NoE

Deliverable D2.1
Virtual Laboratory Integration Report

The EMANICS Consortium

Caisse des Dépôts et Consignations, CDC, France
Institut National de Recherche en Informatique et Automatique, INRIA, France
University of Twente, UT, The Netherlands
Imperial College, IC, UK
International University Bremen, IUB, Germany
KTH Royal Institute of Technology, KTH, Sweden
Oslo University College, HIO, Norway
Universitat Politecnica de Catalunya, UPC, Spain
University of Federal Armed Forces Munich, CETIM, Germany
Poznan Supercomputing and Networking Center, PSNC, Poland
University of Zurich, UniZH, Switzerland
Ludwig-Maximilian University Munich, LMU, Germany
University of Surrey, UNIS, UK
University of Pitesti, UPI, Romania

For more information on this document or the EMANICS Project, please contact:

Dr. Olivier Festor
Technopole de Nancy-Brabois — Campus scientifique
615, rue de Jardin Botanique — B.P. 101
F—54600 Villers Les Nancy Cedex
France

Phone: +33 383 59 30 66
Fax: +33 383 41 30 79
E-mail: <olivier.festor@loria.fr>

Document Control

Title: Virtual Laboratory Integration Report
Type: Public
Editor(s): Jürgen Schönwälder
E-mail: j.schoenwaelder@iu-bremen.de
Author(s): WP2 Partners
Doc ID: D2.1.doc

AMENDMENT HISTORY

Version	Date	Author	Description/Comments
V0.1	2006-06-30	Jürgen Schönwälder	First version, integrating the text from the wiki
V0.2	2006-07-02	Jürgen Schönwälder	Added missing sections
V0.3	2006-07-05	Jürgen Schönwälder	Integrated corrections submitted by the partners

Legal Notices

Table of Contents

(This page is left blank intentionally.)

1 Executive Summary

The EMANICS Network of Excellence (NoE) has committed itself to establish a virtual laboratory and a common testbed, which can be used by EMANICS partners for educational and research purposes. In addition, the joint creation and operation of such an infrastructure should foster the integration of the partners and enable joint project work.

This first “Virtual Laboratory Integration Report” starts by surveying the lab resources readily available within the NoE and which form the basis of building the virtual lab and common testbeds. Since a joint infrastructure will be easier to plan and more successful if there are concrete usage plans, a number of projects have been defined which will foster the creation of labs and testbeds. The projects are very different in size and nature.

The first larger project deals with the construction of a Voice over IP (VoIP) [1,2,3] testbed between EMANICS partners. This testbed will be used for educational purposes, the development of VoIP management tools, to gain experience with VoIP management issues, and to collect VoIP specific traces, which are highly interesting for research in this area. Some of the VoIP infrastructure might also be used internally to simplify the communication between participants.

The second larger project aims at collecting traces of network management traffic from real-world production networks and to provide the necessary tools and repositories for the analysis of such traces. This project explores the fact that several partners have access to different types of production networks and thereby it seems feasible to collect a large number of network management traffic traces, which will provide a more complete picture about network management operations in real networks. This project will also explore the question to what extent Grid [4] computing technologies can be used to support a distributed analysis of large traces.

Next to these two larger projects, there are two more focused projects to address some concrete needs of some partners. The first one aims at making existing traces available for network replay. Being able to replay the behaviour of a production network will allow researchers to evaluate distributed monitoring algorithms on real-world scenarios. The second focused project has the goal to collect basic resource usage statistics from a large number of different machines in order to evaluate load management algorithms.

This report concludes the first phase of this work package in which data was collected, exchanged, and some initial ideas were developed into concrete project plans. This work package is now entering the second phase where the focus shifts to the implementation of the virtual laboratory and common testbeds according to the projects defined in this report.

2 Introduction

The work package “Virtual Laboratory and Common Testbeds” is part of the integration activity of the EMANICS Network of Excellence (NoE). The objectives of this work package are stated in the work package description as follows:

- Integration of existing labs into virtual labs accessible to EMANICS members
- Establishment of a collaboration environment to support the organization / operation of the virtual lab (in coordination with work package 4)
- Creation and maintenance of repositories of measurement data traces

The discussions during and after the EMANICS kick-off meeting in January 2006 made it clear that it will be difficult to establish a virtual laboratory and a common testbed without having concrete usage scenarios in mind. As a consequence, it was decided that the work in this work package is organized around some concrete projects. The definition of these projects and the involvement of the various partners was a major work item during the first six months of the project. This report basically concludes this first phase of this work package in which information was collected, exchanged, and some initial ideas were developed into concrete project plans. The work package is now entering the second phase since the focus has shifted to the implementation of the virtual laboratory and common testbeds according to the projects defined in this report.

2.1 Purpose of the Document

This first virtual laboratory integration report focuses on two aspects. The first aspect concerns the existing lab infrastructure owned by the various partners that can be integrated. Two questionnaires were used to collect the relevant information about the existing labs, their current and future usage, as well as the availability of network traces and supporting repositories. The relevant data was collected in February/March 2006 and formed the basis for the formulation of projects that drive the development of the virtual laboratories and the common testbeds. The description of these projects, which were defined and refined in April/May 2006, is the second main aspect covered by this report.

2.2 Document Outline

Section 3 describes the lab infrastructure owned by the EMANICS partners in terms of hardware components, connectivity and current usage. Section 4 list the data repositories that are currently in place by the various partners and documents their specific interests in traffic traces. Section 5 describes the joint projects to be carried out in the virtual labs and testbeds and lists the contributions expected by the various partners. Section 6 outlines the integration and collaboration achieved between EMANICS partners through this work package before Section 7 concludes this report.

3 Existing Lab Infrastructure

All EMANICS partners (except purely administrative partners) have existing network laboratories that can be potentially integrated or made available for specific experiments and teaching tasks. A survey was conducted in February 2006 in order to collect information and to make it accessible. The details are available on the wiki page of this work package where the data can also be easily maintained. Below is a summary based on the February 2006 survey.

3.1 INRIA

INRIA has the following hardware components:

- 1 CISCO 2621 with IPv6 basic support
- 5 Dell Workstations Precision 380 / Debian
- 3 CISCO IP Phones flashed to run SIP
- 2 Compaq Laptop Evo N800 / Debian
- 2 Linksys WRT54GS (OpenWRT Linux flashed) 802.11G APs with IPv6 support

The machines are generally IPv6 enabled and run open routing and signalling software (quagga, asterisk). The lab is connected to the French research network via a 100 Mbps uplink. The current usage is related to IPv6 experiments and mobility tests as well as local VoIP management experiments.

3.2 University of Twente (UT)

The University of Twente (UT) has the following hardware components:

- 1 Cisco 7206 router
- 1 Cisco AGS+ router
- 1 Cabletron SSR200 router
- 1 3Com 3C16610 SuperStack II Hub
- 1 IBM 8371 Hub
- 1 TByte file / web server for maintaining repositories
- Several PCs

The machines are connected to the Dutch research network via a 100 Mbps uplink. The current usage of the infrastructure is mainly for teaching and online education purposes. In particular, UT hosts online SNMP exercises, an online MIB browser and an online MIB validator. In addition, UT maintains a traffic trace repository with hundreds of hours of TCP/IP header data collected at various places in the Dutch research network SURFnet.

3.3 International University Bremen (IUB)

The International University Bremen (IUB) has the following hardware components:

- 2 Juniper J6300 router (Junos 7.6R1.9)
- 1 Riverstone RS 3000 switch/router (RS 9.1)

- 1 HP ProCurve 2824
- 2 Dual-Xeon Pentium 3 GHz, 2GB RAM / Debian
- 2 Pentium4 3 GHz, 1GB RAM / PlanetLab

The machines are connected to the German research network via a 100 Mbps uplink. Almost all machines support IPv4 and IPv6. The two Xeon server run XEN [5] virtualization software and almost all standard services (HTTP, FTP, SVN, XMPP, DNS, SMTP, RADIUS, TFTP, SIP) are provided using dedicated virtual machines. The equipment is used regularly for teaching purposes, software development and testing, and research projects. Usage of PlanetLab nodes is restricted to institutions that are members of the PlanetLab consortium [6].

3.4 Oslo University College (HIO)

The Oslo University College (HIO) has the following hardware components:

- 2 Juniper 2300 router
- 1 Juniper M7i router
- 1 Nortel BayStack gigabit Ethernet switch with optical module
- 1 Cisco 3660 router
- 2 Cisco AS 5300 router
- 1 Cisco Catalyst 5500 with supervisor module
- 1 Cisco Catalyst 5509 with router module
- 16 node Beowulf cluster
- 6 IBM Blades
- 1 Nortel/Alteon Load Balancer
- 2 virtual machine servers

The machines are connected to the Norwegian research network. Both, IPv4 and IPv6 connectivity is available. The server runs virtualization software such as XEN [5] and UML [7]. During Spring and Fall, the machines are used for teaching purposes. Administration is largely automated using tools such cfengine and MLN.

3.5 Universitat Politecnica de Catalunya (UPC)

The Universitat Politecnica de Catalunya (UPC) has the following hardware components:

- 3 Cisco 7204 router / IOS 12.2(4)T
- 2 Cisco 1750 router / IOS 12.1(4)T
- 1 Cisco 2612 router / IOS 12.0(7)T
- 6 PCs Pentium III, Linux Software.
- 3 WLAN Access Points 802.11 a/b

The testbed is logically separate from the university operational network but reachable through it. When necessary, the testbed can be connected to other networks via IP tunnels (IPv4 and IPv6). The infrastructure is used for ongoing research projects (active

node platforms, network and element management systems, policy-based management).

3.6 University of the Federal Armed Forces Munich (CETIM)

The University of the Federal Armed Forces Munich (CETIM) has the following hardware components:

- 2 Dell PowerEdge 2850 Intel Xeon 2,8 GHz, 2 GB Ram
- 4 Dell PowerEdge 2850 Intel Xeon 2,8 GHz, 2 GB Ram
- 2 Cisco 6503-E with Gigabit-Ethernet Ports + Supervisior Engine
- 3 Cisco 2522
- 2 Cisco 4000
- 2 HP Procurve 5304XL (J4850A) with modules: J4907A, J4852A, J4878B
- 2 HP Procurve 3400CL (J4905A)
- One System Management Host (Dell PowerEdge 750 Intel Pentium 4 2.8-GHz, 1GB Ram)
- Lab with 21 Pentium 4-PCs in a private LAN

The machines are connected to the university backbone via a firewall under control of the research group. IPv6 connectivity exists and will be expanded later this year. Standard services (SSH, HTTP, FTP, CVS, XMPP, DNS, RADIUS, LDAP, SIP) are available. Most servers run the XEN [5] virtualization software. HP OpenView and HP ProCurve Manager Plus are used for network management. The infrastructure is mainly used for teaching, the D-GRID project, and the generation and evaluation of network traces.

3.7 Poznan Supercomputing and Networking Center (PSNC)

The Poznan Supercomputing and Networking Center (PSNC) has access to its own networking lab as well as the monitoring infrastructure of the Polish national research network PIONIER.

The hardware of the networking lab consists of the following components:

- 1 Dual Opteron AMD server: 2.2 GHz 64bit, 10/100/1000/10000 Ethernet
- 2 Pentium 4 class server
- 1 Cisco 7500 router with IPv6 & multicasts
- 1 IXIA traffic generator

The hardware of the PIONIER network monitoring infrastructure consists of the following components:

- 25 monitoring server: 2x Intel Xeon 3.00 GHz, 2GB RAM
- 1 central server: 2x Intel Xeon 3.00 GHz, 2GB RAM
- 25 NTP synchronization modules: GPS Elproma
- 25 3Com 3226 switches
- 25 Lantronix terminal server

The PIONIER network runs a 10 Gbps core. The networking lab is connected to the PIONIER network and supports IPv4 and IPv6. Management software consists of Cisco Works Campus Manager, HP OpenView Network Node Manager, MRTG, as well as rancid. The PIONIER monitoring infrastructure is currently being deployed and may be used for restricted experiments that do not impact the operation and monitoring of the PIONIER network. The networking lab is used in ad-hoc experiments by NOC staff or network researchers and dedicated reconfigurations are possible.

3.8 Ludwig-Maximilians University Munich (LMU)

The Ludwig-Maximilians University Munich (LMU) has the following hardware components:

- 4 Server Intel Pentium III 700MHz (grid testbed)
- 3 3Com switches (manageable)
- 11 PCs Intel Pentium III (computer network lab)
- 4 HP workstations (old, computer network lab)
- 4 ATM switches (computer network lab)
- 1 Intel Pentium 4, 2.8GHz (central server)
- 1 P4, 3.0 GHz, 1 GB RAM, 150 GB HDD (XEN)
- 1 Athlon64 3000+, 2 GB RAM, 80 GB HDD (XEN)

The machines are organized into a grid testbed, a computer network lab, and some other machines running XEN [5] virtualization software. The grid testbed is connected via a 100 Mbps uplink to the Munich science network while the other machines are connected via two firewalls to the Internet. The grid testbed runs the Globus toolkit and VOMS client and server software and is used for various projects. The network lab is mainly used for educational purposes related to intrusion detection, virtual circuits, and network management.

3.9 University of Pitesti (UPI)

The University of Pitesti (UPI) has the following hardware components:

- 1 Cisco 3662 router
- 1 Cisco 2950 switch
- 3 Allied Telesyn switches
- 3 Server / Linux Fedora 5
- 4 Grandstream GXV-3000 SIP video phones
- 1 2TB LaCie Ethernet Disk
- 5 Pentium4 3GHz, 1GB RAM / Linux Fedora

The machines are connected to a LAN (IPv4) protected by UPI firewalls. Standard services (WEB, FTP, SSH) are available. Monitoring is realized via intrusion detection systems such as snort and nessus. The equipment is used regularly for teaching purposes as well as software development and testing.

3.10 University Zurich (UniZH)

The University Zurich (UniZH) has the following hardware components:

- 22 Dell PowerEdge 850, Pentium 4, 3.6 GHz, 1 GB RAM
- 2 Dell PowerEdge 850, Pentium 4, 3.6 GHz, 2 GB RAM, 500 GB HDD
- 2 PlanetLab nodes, Pentium 4, 3.0 GHz, 1GB RAM
- 1 Cisco Catalyst 3750, 48xFE, 4x SFP
- 1 HP ProCurve Switch 2626, 24xFE, 2xGE
- 1 HP ProCurve Switch 2824, 24xGE
- 1 Linksys SRW2024, 24xGE

Most machines have public IPv4/IPv6 connectivity via a direct unfiltered 1 Gbps uplink to the Swiss research network SWITCH. The lab network is completely isolated from the University network. Standard services (SSH, WEB, DNS, DHCP, TFTP, SNMP, NIS, NFS, IPMI) are available. The nodes run Debian and use FAI for fully automated installation. The equipment is used for development, testing, and analysis of distributed Internet applications. Usage of PlanetLab nodes is restricted to institutions that are members of the PlanetLab consortium [6].

3.11 KTH Stockholm (KTH)

The KTH Stockholm (KTH) has the following hardware components:

- 16 Cisco 2600 routers
- 60 rack-mounted PCs
- 1 SmartBits 6000 traffic generator/analyzer
- 30 Intel IXP1200 network processors
- 2 Intel IXP2400 network processors

The machines are connected using three physically separated Ethernet LANs. One is a gateway to KTH's network and the Internet while the other two LANs are physically isolated for experimentation. The lab can be accessed form the outside via SSH. The equipment is currently used for research on distributed routers, distributed network management, self-organizing server clusters, and simulations.

3.12 Summary

As can be seen from the previous sections, there is a large collection of equipment that can be integrated (see Table 1). The hardware is covering a broad set of vendors such as Cisco, Juniper, HP, Riverstone, Cabletron, 3Com, Nortel, Linksys, or Dell.

Open Unix operating systems are frequently used on server systems. Some partners have rather advanced installations to automatically configure and re-configure machines. In addition, XEN virtualization software seems rather widely deployed on newer systems that are capable to host multiple virtual machines. This is an important observation since it implies that XEN virtual machines may be a viable approach to allow remote usage of resources in the virtual lab or to provide remote participants controlled access to machines in a controlled environment.

The description above makes it obvious that some partners have relatively good access to larger production networks, which is a valuable resource to explore within this NoE. Other partners have more interesting facilities for teaching purposes that might also be shared as part of efforts to exchange teaching materials between partners.

Table 1: Summary of the equipment available for integration (partners listed in greyscale are not formally members of this work package)

Partner	Router / Switches	Hosts	Remarks
CETIM	11	28	
HIO	9	24	1 Notel / Alteon load balancer
IC	0	0	
INRIA	1	7	3 Cisco phones, 2 Linksys WRT54GS
IUB	4	2	2 PlanetLab nodes
KTH	16	60	
LMU	7	18	
PSNC	26	29	25 Lantronix terminal server, 1 IXIA
UNIS	0	0	
UPC	6	6	3 access points
UPI	5	8	4 GXV-3000 video phones
UT	5	0	
UniZH	4	24	2 PlanetLab nodes
Total	74 + 20	122 + 84	

4 Existing Repositories and Traces

Network research is to a large extent evaluated using simulation models. Analytical models usually are only able to capture specific aspects of a complex communication system and larger scale experiments are usually very difficult to conduct and difficult to repeat. Since networking research largely relies on simulations, it is crucial to validate the assumptions underlying these simulation models. This explains the special value of traces from production networks for the research community. Unfortunately, it is relatively difficult to get access to such traces due to the sensitive nature of the data and the general reluctance of operators to make traces openly available.

As a consequence, it has proven to be relatively difficult to establish openly accessible repositories of network traces. The approach taken within EMANICS to deal with this issue is to (a) explore cooperation and trust relationships between partners to make traces accessible within the NoE and (b) to explore the creation of distributed repositories where traces may be centrally indexed while the trace data remains under the control of the participating sites and an infrastructure is created to allow researchers to process traces at remote sides without compromising the privacy requirements of the traces.

4.1 Partner's Interest in Traces

Several EMANICS partners have expressed their specific interest in network traces:

- The University of Twente (UT) is currently very interested in SNMP traces but in addition, the UT has interest in VoIP and Skype traces, IDS traces (to replace the DARPA99 set), and general TCP/IP header traces (provided they are long).
- The University of Pitesti (UPI) is mainly interested in general TCP/IP header traces, SNMP traces and SYSLOG traces. Since UPI is developing a VoIP infrastructure, they are also interested in VoIP and Skype traces.
- The University of Zurich (UniZH) is interested in IDS traces (including payload information where possible). UniZH is also interested in Skype traces to inspect encrypted traffic as well as any flow level information (e.g., netflow data) for statistics, traffic management, load balancing mechanisms, multi-domain/distributed accounting, and pricing.
- The Ludwig-Maximilian University (LMU) is primarily interested in SNMP and SYSLOG traces in order to assess the "level of autonomy" of network components by analyzing management traffic to and from these devices, such as frequency of SNMP operations, distribution of OIDs, distribution of values set for different OIDs etc.
- The International University Bremen (IUB) is generally interested in network management traces. Initially, the focus is on SNMP but traces of other protocols used to carry management traffic such as SYSLOG, CLIs, NETCONF are also of high interest.
- The KTH Stockholm (KTH) is interested in network traces that can be replayed in order to evaluate distributed algorithms to aggregate monitoring data.

4.2 Existing Traces and Repositories

Several EMANICS partners already have some traces or they have interesting contacts that can be exploited to obtain new traces.

4.2.1 University of Twente (UT)

The University of Twente (UT) will provide SNMP traces from different sources. UT in addition collects traces for IDS purpose and potentially Skype traces (both signalling and data transfers). UT already has a large collection of TCP/IP header traces and is able to collect more such traces. UT has good contacts with the Dutch research network provider (SURFnet) and UT may collect additional data (for example, UT currently gets netflow data from the Dutch research network). Access to data maintained by UT must be discussed on a case-by-case basis. Next to SURFnet, UT is able to obtain traces from various measurement points within their campus network and UT maintains reasonable contacts with some ADSL providers and server hosting companies.

4.2.2 University of Pitesti (UPI)

The University of Pitesti (UPI) can provide general TCP/IP header traces as well as network management traces (SNMP, SYSLOG) and application traffic traces (voice streams, web traces). These traces originate mainly from the campus network, the national educational and research network (RoEduNet) and the faculty network. Since UPI has several contacts with some ADSL providers and server companies, UPI could get extra data from them.

4.2.3 University of Zurich (UniZH)

The University of Zurich (UniZH) currently collects flow-level data (about 4MB of data per day) but it is possible to also collect TCP/IP header traces or even full packet dumps including payload for specific purposes. The data currently originates from the link which connects UniZH's experimental Linux cluster with the Swiss research network SWITCH. The cluster hosts two PlanetLab [6] nodes, several test nodes used by various research projects, and eventually the hosts connected to the EMANICS VoIP testbed. Discussions are currently underway to determine whether additional traces can be obtained from the UniZH network and the SWITCH network.

4.2.4 Poznan Supercomputing and Networking Center (PSNC)

The Poznan Supercomputing and Networking Center (PSNC) is the operator of a metropolitan area network as well as the Polish national research network PIONIER. PSNC is also a high performance computing center. Thus, it is quite easy for PSNC to provide different types of traces (under specific conditions). PSNC is ready to provide TCP/IP header traces, SNMP traces, and application specific traces from PSNC's computational environment. Partners should specify detailed requirements before PSNC collects traces.

4.2.5 Ludwig-Maximilian University Munich (LMU)

The Ludwig-Maximilian University Munich (LMU) collects netflow data from ten backbone routers of the Leibniz Supercomputing Center (LRZ). The LRZ is the network service provider for all major educational institutions in Munich (more than 50000 hosts). It links several campus networks in a metropolitan area style network. Traces might be obtained from this network on request.

4.2.6 International University Bremen (IUB)

The International University Bremen (IUB) can provide general TCP/IP header traces of the link connecting the PlanetLab [6] nodes to the Internet. In addition, IUB has network management traces from their networking lab (which also hosts services for other computer science research groups). IUB maintains good contacts to the operator of a local metropolitan area network, which provides connectivity between the major research institutions as well as governmental and industrial organization. This network makes intensive use of wireless links.

4.2.7 University of Federal Armed Forces Munich (CETIM)

The University of Federal Armed Forces Munich (CETIM) has TCP/IP header traces over approximately one week, measured at the uplink of the building, where our computer science students live, to the campus. CETIM can collect additional traces as required.

5 Projects

During the initial discussions among the partners involved in this work package, it was decided that testbeds must serve a concrete purpose since building up an infrastructure without concrete usage scenarios in mind bears the risk to achieve only a small degree of efficiency, both in terms of actual usage and achieved collaboration. During an open discussion process, the testbed activities described in the following sections were identified.

5.1 VoIP Management Testbed (VOIP)

The objective of this activity is to provide a testbed for VoIP communication based on open source components.

5.1.1 Motivation

The VoIP management testbed is important for several reasons. First, the management of VoIP services based on technologies such as SIP [1] has not been addressed yet in the network management research community. Knowing what and how to perform practical real world VoIP management on larger scale networks is a major help for the research community.

The second reason is VoIP behaviour identification: One of the main issues in the management of VoIP services is related to traffic characterization of VoIP traffic (both signalling and data). Traditional models, based on the Erlang model and fixed reserved bandwidth used over PSTN, are not viable anymore in next generation converged networks. Preliminary studies on VoIP showed that users tend to have very long lasting calls and bandwidth is highly dependent on the codec/user behaviour. Having accurate traffic models for VoIP networks is highly important for several management tasks:

1. **Configuration management**: Knowing how the traffic looks like and what to expect in terms of network behaviour from a VoIP network is essential for configuration purposes.
2. **Security management**. VoIP will be the major issue in network security over the future years. Network intrusion detection and detecting fraudulent usage can be possible only if we know how real signalling traffic differs from abnormal traffic.

The work to be done for this testbed is multi-fold:

1. **Educational resource for VoIP courses**: Several partners have teaching activities centered on VoIP and one current major issue is related to allowing to the students the end-to-end management of VoIP services over a shared network. End system management done locally (PBX, SER router) are possible with the existing resources at each partner site, but no end to end management is possible due to the lack of a common testbed spawned over the Internet.
2. **Development of new open source based VoIP tools**: There is a lack of VoIP tools in the open source community and this lack is amplified when we consider only the management and security issues. In order to manage the testbed, several security and management tools must be developed from scratch. We will develop a looking glass and security monitoring and assessment tools in order to manage the testbed.
3. **Experience in VoIP management**. The testbed can not only serve as an excellent learning resource (as pointed out in the first point) for undergraduate

and graduate students, but also research activities and experimentations can benefit from it. We intend to provide a global environment allowing to experiment, test and validate management scenarios.

4. **Trace data collection**: As of today, the sensitive nature of VoIP data makes it impossible to obtain realistic traces. Such traces are however invaluable to the research community working on VoIP network dimensioning, fault management and more specific event correlation and root cause analysis. This testbed should provide this type of data to the partners having research activities on these issues.

5.1.2 Infrastructure and Operation

The VoIP testbed will be constructed using only open source software (e.g., Asterisk and SIP router software) as a guarantee for transparent and trusted software. Each partner will install the required platform elements and a global VoIP testbed will be jointly developed. Each partner will perform local configuration management on its systems in order to comply with local security policies and privacy regulations. A limited guest account for simple management tasks will be made available to other testbed partners.

5.1.3 Partners and their Roles

The VoIP testbed activity is lead by INRIA. The EMANICS partners involved in this activity are INRIA, UPI, CETIM, IUB, and UniZH.

5.1.3.1 INRIA

INRIA is in charge to define the general VoIP testbed and will animate the following activities:

- Develop a Web based centralized looking glass to the managed devices over the network
- Provide global activities synchronisation and animation for the partners
- Provide support for VoIP software installation and operation
- Develop VoIP assessment and monitoring tools
- Monitor and audit the security of the overall platform

The following work needs to be done:

1. Development of a Web-based management interface for all SIP devices deployed in the testbed
2. Development of a security monitoring tool for the VoIP testbed
3. Development of a comprehensive assessment tool suite for the VoIP testbed

5.1.3.2 University of Pitesti (UPI)

The University of Pitesti (UPI) will provide the following support:

- Manage locally a SIP based router plus PBX interface towards the PSTN
- Provide call detail records and SIP traces to all partners

The following work needs to be done:

1. Install and configure a SIP router and Asterisk PBX

2. Collect SIP level packet traces and Asterisk call detail records
3. Install and maintain a data repository for these traces

5.1.3.3 University of Federal Armed Forces Munich (CETIM)

The University of Federal Armed Forces Munich (CETIM) will provide the following support:

- Manage local SIP-based router (Asterisk + SER)
- Establish connectivity to the campus PSTN
- Installation of DUNDi in order to easy the routing of phone calls between EMANICS partners
- Integration of Asterisk with LDAP
- Evaluation of VoIP clients (HW, SW) and their interaction with Asterisk
- Provide SIP traces to partners

The following work needs to be done:

1. Install and configure a SIP router and Asterisk PBX
2. Install and configure DUNDi
3. LDAP integration within the testbed
4. Perform a comprehensive test of VoIP SIP clients and report the results

5.1.3.4 International University Bremen (IUB)

The International University Bremen (IUB) will provide the following support:

- Manage local SIP based router (Asterisk)
- Participate as a user fo the EMANICS VoIP testbed

The following work needs to be done:

1. Install and configure a SIP router (Asterisk)
2. Integrate VoIP clients

5.1.3.5 University of Zurich (UniZH)

The University of Zurich (UniZH) (not formally involved in this work package) will provide the following support:

- Participate as a user of the EMANICS VoIP testbed
- UniZH will integrate its own VoIP gateway with a campus PSTN connection into the testbed and get access to the infrastructure, software, traces, and support provided by the other partners

The following work needs to be done:

1. Install and configure a SIP router and Asterisk PBX

5.1.4 Reporting

A more detailed description of the VoIP testbed will be provided in a subsequent deliverable of this work package. The research results enabled by the existence of the testbed will be published as papers and presented at conferences and journals, like IM/NOMS, DSOM, or eTNSM.

5.2 Network Management Trace Collection and Analysis (TRACE)

The objective of this activity is to collect traces of network management traffic from production networks and to provide the necessary tools and repositories for the analysis of such traces.

5.2.1 Motivation

Management protocols like the Simple Network Management Protocol (SNMP) [8] are widely deployed to monitor, control, and configure network elements. Even though management protocols like SNMP are well documented and understood, it remains relatively unclear how they are used in practice and what the typical usage patterns are [9]. As a consequence, researchers and protocol designers often base their work on assumptions that lack a sound justification. The purpose of the work done in this activity is to collect management traces from many different production networks and to provide tools to efficiently process the traces and extract meaningful data.

The work carried out in this activity is organized into three different tasks:

1. **Trace Collection**: It is crucial for this activity to obtain traces from production networks. As an initial step, several EMANICS partners will collect network management traces in their respective institution or related organizations. In a second step, partners may help other organizations located in their area with the collection of additional traces. The initial focus will be on SNMP.
2. **Tool Development**: The goal is to develop generic tools for network management trace analysis. Initially, the focus will be on tools supporting the analysis of SNMP traces. Tool development includes the creation of a generic anonymization library that can be used for traffic trace anonymization.
3. **Distributed Trace Analysis**: The nature and complexity of collecting and analyzing management traces among distributed partners depends on the type and size of the scenario. As long as few trace providers share a small amount of traces, based on peer-to-peer legal agreements or implied trust, efforts to run these processes and its associated bureaucracy can be kept to a minimum. At the same time, however, growth and scalability in terms of the number of partners and resources, the amount of trace data, and the efficiency in terms of number of analyses per time are severely limited.

   In the proposed activity, we want to model a scenario that is still feasible for small groups of partners, yet offers more scalability and automation. In this scenario, multiple trace providers offer traces to multiple trace analysts, who run jobs on the provided traces. Two cases of collaboration will be modelled:

   1. The traces are copied from the trace provider to the trace analyst, who then processes these traces locally.
   2. The trace analyst submits an analysis job to the trace provider, who runs the job on the local traces and returns the results.

   As distributed collection and analysis of traces is close to other tasks, which grid computing has successfully been applied to, we want to explore how existing grid middleware can be used to support features like job management, account and permission management, and access permissions. It is the overall goal to reach a high level of automation while providing checkpoints for trace providers, so that they remain in control which trace analyst runs what kind of jobs on their data and which results are sent back to the analyst.

As a proof of concept, a demonstrator should be set up, running on selected resources of the partners, that contribute to this activity, to show the applicability of grid middleware to this task. In a pragmatic 80/20 approach, existing trace analysis tools should be wrapped as grid services and trace files in a common data formats should be wrapped as grid resources.

Note that the actual interpretation of network management traces is not part of this activity. The work done in this work package is purely oriented towards enabling research based on real-world traces. Some of the research enabled through the availability of the infrastructure establish through this work package will be carried out in work package 7 (Scalable Management).

5.2.2 Infrastructure and Operation

Traces will be collected by using hardware components already maintained by the partners. The partners who are collecting the traces in cooperation with the organizations providing access to their network will handle the setup of appropriate measurement points. The distributed trace analysis prototype will be initiated on existing grid computing infrastructures. LMU encourages other TRACE or EMANICS partners to provide additional grid resources. The precise conditions for handling access to the distributed trace analysis infrastructure will be specified as part of the activity.

5.2.3 Partners and their Roles

The TRACE activity is lead by IUB. The EMANICS partners involved are IUB, UT, LMU, PSNC, CETIM, and INRIA.

5.2.3.1 International University Bremen (IUB)

The International University Bremen (IUB) will contribute with the following work items:

- Development of SNMP trace analysis tools. This includes the steering necessary to ensure that a coherent set of tools is realized and not just a collection of point solutions that are difficult to maintain
- Collect network management traces from a regional network operator
- Contribute to the requirements specification for the distributed trace analysis prototype

5.2.3.2 University of Twente (UT)

The University of Twente (UT) will contribute with the following work items:

- Collect traces from at least four different locations
- Contribute to the development of analysis tools
- Contribute to the requirements specification for the distributed trace analysis prototype

5.2.3.3 Ludwig-Maximilian University Munich

The Ludwig-Maximilian University Munich (LMU) will contribute with the following work items:

- Collect traces from the Munich backbone network
- Model the distributed trace analysis scenario by
    - defininig terminology

- developing an organizational model with roles, activities, processes and responsibilities
- specifying an information model with documents, jobs, and permissions

- Installation of a demonstrator based on grid middleware

5.2.3.4 Poznan Supercomputing and Networking Center (PSNC)

The Poznan Supercomputing and Networking Center (PSNC) will support this activity in the following way:

- Collect traces from their local organization, a metropolitan area network, and a national research network

5.2.3.5 University of Federal Armed Forces (CETIM)

The University of Federal Armed Forces (CETIM) will contribute with the following work items:

- Collect traces from the University network
- Provide grid resources for the distributed trace analysis prototype

5.2.3.6 INRIA

INRIA is interested to access the SNMP traces to establish monitoring patterns which are useful for feeding benchmarking models developed at INRIA.

- Access to SNMP traces to establish monitoring patterns

5.2.4 Reporting

The output produced by this activity includes concrete tool implementations, which will be available under an open source license. Tools will be documented by providing manual pages, papers describing the tools, and presentations. Concrete data exchange formats supported by the tools may be submitted to the IETF/IRTF for publication.

The availability of network management traces will enable joint research work, which is expected to lead to a number of publications at conferences and journals, like IM/NOMS, DSOM, or eTNSM.

The work on distributed trace analysis will lead to a demonstrator. All software developed for the prototype will be shared within EMANICS and if possible made openly available.

5.3 Trace Collection for Network Replay (REPLAY)

The objective of this activity is to collect traffic traces, which allow researchers to replay network behaviour such that distributed data aggregation algorithms can be evaluated.

5.3.1 Motivation

KTH is engineering a novel protocol (A-GAP) [10] for adaptive real-time monitoring of large-scale networks. Specifically, KTH wants to monitor network-wide metrics computed from device counters using aggregation functions, such as SUM, AVERAGE and MAX. Examples of such metrics include the total number of VoIP flows and the maximum link utilization in a network domain. A-GAP is a decentralized and asynchronous protocol that minimizes the generated overhead for a configurable accuracy of the estimation.

Real traffic traces are needed to evaluate A-GAP in realistic scenarios and compare the results to predictions based on the stochastic model that underlies A-GAP. The idea is to use previously recorded traces to replay IP packets in a test environment. Note that, for replaying, only TCP-IP header data is needed; packet payload is not needed. For privacy reasons the payload is therefore not included on the traces; for the same reason IP addresses need to be anonymized.

Within the REPLAY project the UT will make available its previously collected and anonymized traffic traces. These traces were collected at four locations:

1. **Location 1:** On location #1 the 300 Mbit/s Ethernet link (a trunk of 3 x 100 Mbit/s) has been measured, which connects a residential network of a university to the core network of this university. On the residential network, about 2000 students are connected, each having a 100 Mbit/s ethernet access link. The residential network itself consists of 100 and 300 Mbit/s links to the various switches, depending on the aggregation level. The measured link has an average load of about 60%. Measurements have taken place in July 2002.
2. **Location 2**: On location #2, the 1 Gbit/s Ethernet link connecting a research institute to the Dutch academic and research network has been measured. There are about 200 researchers and support staff working at this institute. They all have a 100 Mbit/s access link, and the core network of the institute consists of 1 Gbit/s links. The measured link is only mildly loaded, usually around 1%. The measurements are from May - August 2003.
3. **Location 3**: Location #3 is a large college. Their 1 Gbit/s link (i.e., the link that has been measured) to the Dutch academic and research network carries traffic for over 1000 students and staff concurrently, during busy hours. The access link speed on this network is, in general, 100 Mbit/s. The average load on the 1 Gbit/s link usually is around 10-15%. These measurements have been done from September - December 2003.
4. **Location 4**: On location #4, the 1 Gbit/s aggregated uplink of an ADSL access network has been monitored. A couple of hundred ADSL customers, mostly student dorms, are connected to this access network. Access link speeds vary from 256 kbit/s (down and up) to 8 Mbit/s (down) and 1 Mbit/s (up). The average load on the aggregated uplink is around 150 Mbit/s. These measurements are from February - July 2004.

In addition to these existing traces, the UT will also collect some new traces from the core of a large national research network.

5.3.2 Infrastructure and Operation

The funding received by the UT will primarily be used for buying two systems: one system that can be used to collect new traces and one system two anonymize and distribute traces. Both systems will be equipped with high-speed network connections, and disks capable of storing between 0,5 and 1 TB of data. The machines will be administrated by the UT.

5.3.3 Partners and their Roles

Two partners will participate in this project: UT and KTH. The role of the UT is to lead this project, to collect traces (like TCP/IP headers), perform some pre-processing (like anonymization) and host a KTH student for a short period. The role of KTH is to use these traces to evaluate the A-GAP protocol. Note that KTH is not formally involved in this work package.

5.3.4 Reporting

Progress will be reported on a quarterly basis and included in the EMANICS quarterly management report. In addition, results of the research enabled by the network replay data collection activity will be reported at the EMANICS review and be published as papers and presented at conferences and journals, like IM/NOMS, DSOM, or eTNSM.

5.4 Resource Usage Data Collection (ABLOMERS)

The objective of this activity is to collect monitoring data of computational resources in distributed systems for optimal load balancing and service scheduling in heterogeneous platforms.

5.4.1 Motivation

ABLOMERS is a Java software agent that contacts local SNMP agents in every monitored node for generating resource status and statistical load reports in XML format. The communication between the ABLOMERS software agent and the physical resources is via SNMP, requiring only 'read' access rights. The communication between ABLOMERS agents and resource managers that will collect and format the information will be designed shortly. About the later we can already say that it will not be based on SNMP.

The initial application of ABLOMERS is to test a generic resource manager system that is fed with the above mentioned statistical computational resources performance. The final goal is to design a near-optimal resource manager based on heuristic algorithms. This is done in the context of a Ph.D. thesis at UPC. The resource status information should be obtained from as many nodes as possible where all of them are used for different computational activities (file servers, web servers, pc desktops, laptops, etc.) into their respective networks.

ABLOMERS can be seen as a general-purpose tool easily adaptable to any kind of network resource monitoring. In addition, the specific experiment proposed in the field of resource monitoring can also be beneficial for the results themselves; in fact, ABLOMERS will offer statistical resource performance in every enabled node. Other partners interested in this type of data could use the monitored information for any purpose. UPC plans to collaborate with interested partners to design specific experiments based on ABLOMERS for their particular use.

5.4.2 Infrastructure and Operation

Partners supporting the resource usage data collection activity have to provide standard PC hardware to run the ABLOMERS software. Monitored nodes must run an SNMP agent. A more detailed specification of the hardware requirements and the set of preferred MIB modules for collecting resource usage statistics is available from UPC.

The hosting partner will be involved in the initial installation and configuration of the ABLOMERS software. During the monitoring phase, support from hosting partners will not be necessary except in case of any failure or misbehaviour where support from a local administrator might be necessary.

5.4.3 Partners and their Role

This project is lead by UPC. UPC is responsible to provide the data collection software. In addition, UPC will be in charge of tracking the appropriate installation of the agents in all the remote sites involved, the data collection in the experiment mentioned above,

and the reporting process. In addition, UPC will carry out the adaptation of ABLOMERS to any related specific experiment if required from a partner.

The EMANICS partners UT, UNIS and KTH already agreed to support the data collection activity by hosting the ABLOMERS software in their environments. They will help with the initial ABLOMERS installation and configuration. Note that UNIS and KTH are not formally involved in this work package.

5.4.4 Reporting

Results from the initial experiment described above are expected to be available along the last quarter of this year. Therefore UPC expects to distribute a report by the end of 2006.

6 Collaboration

The work package "Virtual Labs and Common Testbeds" is part of the EMANICS integration activities. This section briefly discusses the collaboration achieved within this work package. Table 2 summarizes which partner is involved in the activities carried out in this work package.

Table 2: Summary of the partners involved in the various activities (partners listed in greyscale are not formally members of this work package)

Partner	VOIP	TRACE	REPLAY	ABLOMERS	Total
CETIM	X	X			2
HIO					0
IC					0
INRIA	X	X			2
IUB	X	X			2
KTH			X	X	2
LMU		X			1
PSNC		X			1
UNIS				X	1
UPC				X	1
UPI	X				1
UT		X	X	X	3
UniZH	X				1
Total	5	6	2	4	

As can be seen from Table 2, the two larger activities VOIP and TRACE integrate 5 or 6 partners respectively while the smaller activities REPLAY and ABLOMERS integrate 2 or 4 partners respectively. In addition, 1 partner participates in three activities, 4 partners participate in two activities, 6 partners in one activity and only one partner (HIO) is not actively involved in this work package. It should be noted that HIO is leading efforts in other work packages such as the organization of the AIMS conference

and therefore decided to not participate actively in this work package during the first EMANICS year.

Overall, the level of integration achieved looks very promising. The planned activities require the collaboration of partners in order to produce meaningful output and therefore it can be expected that by carrying out these activities, a closer integration of the involved partners will be achieved.

7 Summary and Conclusions

The first "Virtual Laboratory Integration Report" documents the existing labs and trace repositories of the EMANICS partners and defines several projects that will be carried out in order to build the virtual laboratories and common testbeds. The two larger projects are centered around VoIP technology and its management, which is becoming increasingly important, and the gathering and analysis of network management traces from production networks. The two smaller projects focus on traces for network replay and resource usage statistics for load heuristic balancing algorithms.

All these projects have in common that they require the collaboration of partners within the EMANICS consortium in order to be successful. As a consequence, a decent level of real integration can be expected once the joined labs and testbeds become operational.

This report concludes the first phase of this work package in which data was collected, exchanged, and some initial ideas were developed into concrete project plans. This work package is now entering the second phase where the focus shifts to the implementation of the virtual laboratory and common testbeds according to the projects defined in this report.

8 References

[1] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, *SIP: Session Initiation Protocol;* RFC 3261, June 2002.

[2] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, *RTP: A Transport Protocol for Real-Time Applications;* RFC 3550, July 2003.

[3] S. A. Baset, H. Schulzrinne, *An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol;* Columbia University Technical Report CUCS-039-04, September 2004.

[4] I. Foster, C. Kesselman, S. Tuecke, *The Anatomy of the Grid: Enabling Scalable Virtual Organizations;* International Journal of Supercomputer Applications (15)3, 2001.

[5] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, A. Warfield, *Xen and the Art of Virtualization;* Proceedings of the 19th ACM Symposium on Operating Systems Principles, October 2003.

[6] L. Peterson, T. Anderson, D. Culler, T. Roscoe, *A Blueprint for Introducing Disruptive Technology into the Internet;* First Workshop on Hot Topics in Networking (HotNets-I), October 2002.

[7] J. Dike, *User Mode Linux;* Prentice Hall, 2006.

[8] J. Case, R. Mundy, D. Partain, B. Stewart, *Introduction and Applicability Statements for Internet Standard Management Framework;* RFC 3410, December 2002.

[9] J. Schoenwaelder, *SNMP Traffic Measurements;* Internet-Draft, May 2006.

[10] A.G. Prieto, R. Stadler, Distributed Real-time Monitoring with Accuracy Objectives; IFIP Networking 2006, IFIP May 2006.

9 Abbreviations

CLI	Command Line Interface
DNS	Domain Name System
DUNDi	Distributed Universal Number Discovery
FTP	File Transfer Protocol
LDAP	Lightweight Directory Access Protocol
PBX	Private Branch Exchange
PSTN	Public Switched Telephone Network
RADIUS	Remote Authentication Dial-In User Service
SIP	Session Initiation Protocol
SMTP	Simple Mail Transfer Protocol
SNMP	Simple Network Management Protocol
SSH	Secure Shell
SVN	Subversion
TFTP	Trivial File Transfer Protocol
VoIP	Voice over Internet Protocol
XMPP	Extensible Messaging and Presence Protocol

10 Acknowledgements

This deliverable was made possible due to the large and open help of the WP2 team of the EMANICS NoE. Many thanks to all of them.