



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team reso

*Optimized protocols and software for high
performance networks*

Grenoble - Rhône-Alpes

THEME COM

Activity
R *eport*

2008

Table of contents

1. Team	1
2. Overall Objectives	2
2.1. Project-team presentation overview	2
2.2. Context	2
2.3. Research area	3
2.3.1. Axis 1 : Optimized protocol implementations and networking equipments	4
2.3.2. Axis 2 : Quality of Service and Transport layer for Future Networks	4
2.3.3. Axis 3 : High Speed Network's traffic metrology and statistical analysis	4
2.3.4. Axis 4: Network Services for high demanding applications	5
2.4. Application domains	5
2.5. Methodology	6
2.6. Goals	6
2.7. Summary of the main contributions of the team in 2008	6
2.7.1. Axis 1 : Protocol implementations and networking equipments	6
2.7.2. Axis 2 : End-to-end Quality of Service and Transport layer	6
2.7.3. Axis 3: High Speed Network's traffic metrology and statistical analysis	7
2.7.4. Axis 4: Network Services for high demanding applications	7
3. Scientific Foundations	7
3.1. Optimized Protocol implementations and networking equipments	7
3.2. Quality of Service and Transport layer for Future Networks	8
3.3. High Speed Network's traffic metrology and statistical analysis	10
3.4. Network services for high demanding applications	10
4. Application Domains	11
5. Software	12
5.1. BDTS: Bulk Data Transfer Scheduling Service	13
5.2. FLOC: Flow control	13
5.3. NXE: Network eXperiment Engine	13
5.4. HSTTS: High Speed Transport protocols Test Suite	13
5.5. HIPerNet: management of virtual private execution infrastructure	14
5.6. SNE (Stateful Network Equipment)	14
5.7. ShowWatts: Real time energy consumption grapher.	14
5.8. XCP-i (Interoperable eXplicit Control Protocol)	14
6. New Results	15
6.1. Optimized protocol implementation and networking equipments	15
6.1.1. Evaluation and optimization of network performance in virtual end systems and routers	15
6.1.2. High performance Autonomic Gateways for large scale distributed systems and Grids	15
6.1.3. High availability for clustered network equipments	16
6.1.4. High availability for stateful network equipments	17
6.1.5. Router assisted network transport protocol	17
6.1.6. XCP-i: a new interoperable XCP version for high speed heterogeneous networks	17
6.1.7. Autonomic Service Deployment in Next Generation Networks	18
6.1.8. Energy-efficiency in computing and networking for large-scale distributed systems	18
6.2. Quality of service and Transport Protocols for Future Networks	19
6.2.1. Flow Scheduling and endpoint rate control	19
6.2.2. Dynamic bandwidth provisioning	19
6.2.3. Flow-aware networks	20
6.2.4. Integrating very large packets in networks	20
6.2.5. Fair Rate Allocation for Multihop Wireless Networks	21
6.2.6. Fairness, Simplicity and Efficiency in Multihop Wireless Networks	21

6.2.7.	Available Bandwidth Estimation in Multihop Wireless Networks	22
6.2.8.	Congestion Collapse in Grid'5000	22
6.2.9.	Towards a User-Oriented Benchmark for Transport Protocols Comparison in very High Speed Networks	22
6.2.10.	A language for virtual resources and interconnection networks description	22
6.3.	High Speed Network's traffic metrology and statistical analysis	23
6.3.1.	Experimental investigation on the relationship between heavy tailed distributions and long range dependence	23
6.3.2.	Maximum likelihood estimate of heavy-tail exponents from sampled data	24
6.3.3.	A nonlinear model for satellite images time series: analysis and prediction of land cover dynamics	24
6.3.4.	EMD-based analysis of rat EEG data for sleep state classification	25
6.4.	Network services for high demanding applications	25
6.4.1.	Design and development of an MPI gateway	25
6.4.2.	Virtualizing and scheduling optical network resource for emerging IT services	26
6.4.3.	Automatic Network Services aligned with Grid Application requirements in CARRIOCAS	26
6.4.4.	Development of a metrology platform on Grid5000	27
6.4.5.	Programmable network services for context aware adaptation	27
7.	Contracts and Grants with Industry	28
7.1.	INRIA actions	28
7.1.1.	GRID5000: ADT Aladdin	28
7.1.2.	INRIA ARC GREEN-NET	28
7.2.	INRIA Bell Labs common laboratory: Semantic Networking	29
7.3.	ANAGRAN: flow-router evaluation in Grid5000	29
7.4.	France Telecom R&D	29
7.5.	CARRIOCAS	29
8.	Other Grants and Activities	30
8.1.	National actions	30
8.1.1.	ANR IGTMD	30
8.1.2.	ANR DSLLAB	30
8.1.3.	ANR HIPCAL	30
8.1.4.	ANR DMASC	31
8.2.	European actions	32
8.2.1.	AUTONOMIC INTERNET	32
8.2.2.	OGF-EUROPE	32
8.2.3.	AEOLUS	32
8.2.4.	EC-GIN	32
8.3.	International actions	33
8.3.1.	NEGST: JSPT-CNRS	33
8.3.2.	AIST Grid Technology Research Center: GridNet-FJ associated team	33
8.3.3.	Collaboration with University of Otago, New Zealand	34
8.3.4.	Collaboration with University of Lisbon, Portugal	34
8.3.5.	Collaboration with University of Istanbul, Turkey	34
8.4.	Visitors	34
8.4.1.	Collaboration with University of Sevilla, Spain	34
8.4.2.	Collaboration with Vrije University, Netherlands	35
8.4.3.	Collaboration with AIST GTRC, Japan	35
9.	Dissemination	35
9.1.	Conference organisation, editors for special issues	35
9.2.	Graduate teaching	36

9.3. Miscelleneous teaching	37
9.4. Animation of the scientific community	37
9.5. Participation in boards of examiners and committees	38
9.6. Seminars, invited talks	38
10. Bibliography	39

1. Team

Research Scientist

Pascale Vicat-Blanc Primet [Research Director (DR2) Inria, HdR]
Paulo Gonçalves [Research Associate (CR1) Inria]
Laurent Lefèvre [Research Associate (CR1) Inria]

Faculty Member

Jean-Patrick Gelas [Maître de conférences]
Olivier Glück [Maître de conférences]
Isabelle Guérin-Lassous [Professeur, HdR]

Technical Staff

Jean-Christophe Mignot [Research Engineer CNRS – 40%]
Aurélien Cedeyn [ENS Engineer – Grid5000]
Matthieu Imbert [Research Engineer INRIA – 40%]
Pierre Bozonnet [Expert Engineer INRIA –contrat CARRIOCAS until 31/08/08]
Marcelo Pasin [Expert Engineer INRIA-FP6 EC-GIN until 15/10/08]
Damien Ancelin [Expert Engineer INRIA -FP6 EC-GIN]
Olivier Mornard [Expert Engineer INRIA– ANR HIPCAL]
Philippe Martinez [Expert Engineer INRIA– ANR HIPCAL since 1/10/08]
Abderhaman Cheniour [Expert Engineer INRIA – FP7 Autonomic Internet]
Oana Goga [Associated Engineer INRIA– ADT Aladdin]
Adrien Kiren [Temporary Engineer INRIA– contrat CARRIOCAS]

PhD Student

Lucas Nussbaum [ATER - 2008/2009]
Narjess Ayari [PhD student with France Telecom R-D - CIFRE - 2005/2008 until]
Romaric Guillier [PhD student, INRIA - CARRIOCAS - 2006/2009]
Ludovic Hablot [PhD student, ENS, MENRT - 2006/2009]
Patrick Loiseau [PhD student, ENS , MENRT- 2006/2009]
Dino Lopez Pacheco [PhD student - Mexican Government Grant- 2004/2008]
Sébastien Soudan [PhD student, ENS, MENRT - 2006/2009]
Rémi Vanier [PhD student, INRIA - AEOLUS 2006/2009]
Dinil Mon Divakaran [PhD student, INRIA-Bell Labs - 2007/2010]
Pierre-Solen Guichard [Internship (Master degree) PhD student, INRIA-Bell Labs - 2008/2011]
Anne-Cécile Orgerie [Internship (Master degree) then PhD student - ENS, MENRT - 2008/2011]
Fabienne Anhalt [Internship (Master degree) then PhD student, INRIA - 2008/2011]
Guilherme Koslovski [Internship (Master degree) then PhD student, INRIA - 2008/2011]

Post-Doctoral Fellow

Manoj Dahal [Postdoc INRIA - CARRIOCAS contract since 5/11/08]
Olivier Grémillet [Postdoc INRIA Bell Labs since 1/12/08]
Emanouelli Draminitos [Postdoc INRIA AEOLUS contract since 15/11/08]

Visiting Scientist

Hugo Carrão [PhD student from University of Lisbon, with Reso from july to december 2008]
Süleyman Baykut [PhD student from University of Istanbul, with Reso from january to october 2008]
Alejandro Fernandez-Montes Gonzalez [Phd student from University of Sevilla, Spain with RESO in November 2008]
Pablo Neira Ayuso [Phd student from University of Sevilla, Spain with RESO from June to July 2008]

Administrative Assistant

Sylvie Boyer [Secretary (SAR) INRIA – 10%]
Christine Garcia [Secretary ENS – 50% until 31/08/08]
Jocelyne Brajon [Secretary ENS – 50% since 1/10/08]

2. Overall Objectives

2.1. Project-team presentation overview

The RESO team belongs to the “Laboratoire de l’Informatique du Parallélisme” (LIP) - Unité Mixte de Recherche (UMR) CNRS-INRIA-ENS with Université Claude Bernard of Lyon. It consists of twenty members in average, including six permanent researchers and teaching researchers. The research activities of the RESO project fits the "communicating" scientific priority of the INRIA's strategic plan 2008-2012. In this direction, RESO is focusing on communication software, services and protocols in the context of high speed networks and applying its results to the domain of high demanding applications and Future Internet.

The RESO approach relies on the theoretical and practical analysis of limitations encountered in existing systems or protocols and on the exploration of new approaches. This research framework at the interface of a specific network context and a challenging application domain, induces a close interaction, both with the underlying network level as well as the application level. Our methodology is based on a deep evaluation of the functionalities and performance of high speed infrastructures and on a study of the high end and original requirements before designing and analysing new solutions. RESO gathers expertise in the design and implementation of advanced high performance cluster networks protocols, long distance networks and Internet protocols architecture, distributed systems and algorithms but also scheduling theory, optimization, queuing theory and statistical analysis. This background work provides the context model for innovative protocols and software design and evaluation. Moreover, the proposals are implemented and experimented on real or emulated local or wide area testbeds, with real conditions and large scale applications.

2.2. Context

Wavelengths multiplexing and wavelengths switching techniques on optical fibers allow core network infrastructures to rapidly improve their throughput, reliability and flexibility. Links of 40 gigabits per second are now available and 100 gigabits per second are emerging. New technologies like 10 Gigabit/s Ethernet or 10Gigabit/s Myrinet is also driving the increase of bandwidth in local area networks. These improvements have given the opportunity to create high performance distributed systems that aggregate storage and computation resources into a virtual and integrated computing environment. During a decade lot of researches and developments around the concept of grid and utility computing have underlined the strenght of this approach. Today, the communication, computation and storage aspects of the Internet tend to converge. They are combined with the deployment of ultra high capacities interconnection networks with predictable performance and the emergence of coarse Grain Web Servers like Google, Yahoo, Amazon providing the content, control, storage and computing resources for the users. All these trends will strongly influence the development of the future Internet. They raise major research issues in networking and services, requiring a new vision of the network of networks and its protocol architecture. Indeed, the current Internet stack (TCP/IP) and its associated simple network management protocol is not consistent with the evolution of the network infrastructure components and its use by emerging services which aim to deliver supercomputing power available to the masses over the Internet. The coordination of networking, computing and storage requires the design, development and deployment of new resource management approaches to discover, reserve, co-allocate and reconfigure resources, schedule and control their usages. The network is not only a black box providing pipes between edge machines, but is becoming a vast cloud increasingly embedding the computational and storage resources to meet the requirement of emerging applications. These resources will be located at important crossroads and access points throughout the network. During the last few years, we have seen that the distinction between packet forwarding and application processing has become blurred. The network community now starts to worry not only about forwarding packets without regard to application semantics, but is increasingly trying to exploit new functionalities within the network to meet the requirements of application. Reciprocally, distributed systems and applications have traditionally been designed to run on top of the Internet, and to take the architecture of the Internet as given. Although the convergence of communication and computation at every level appears to be natural, it is still very difficult to efficiently explore the full range of possibilities it can bring. Most of

the proposals exploiting this convergence break the initial design philosophy of the Internet protocol stack (end-to-end argument for example), or if implemented in the application layer present lot of performance, resilience and scalability issues. Recently, ambitious research programs like FIND (NSF) or FIRE (EU) have been launched towards the design of a new protocol architecture for Future Internet and solving the critical issues facing the current architecture. We think that the Internet re-design raises the opportunity to better understand and assess higher-level system requirements, and use these as drivers of the lower layer architecture. In this process, mechanisms that are implemented today as part of applications, may conceivably migrate into the network itself, and this is one of main driver of the researches of RESO. One of the key challenge for large deployment of new high end applications in the Internet is the provisioning of a secure, flexible, transparent and high performance transport infrastructure for data access and processing. Consequently, future high-speed optical networks are addressed not only to support the accelerating and dynamic growth of data traffic but also the new emerging network requirements such as fast and flexible provisioning, QoS levels, and fast recovery procedures of such data intensive applications.

2.3. Research area

The use of networks for on-demand computing is now gaining in the large Internet, while the optical transport layer extends to the edge (fiber to the home). Enabling ultra high performance machine to machine communications lead then to new bandwidth and network resource sharing paradigms. RESO is investigating several issues such as quality of service, transport protocols, energy efficiency, traffic metrology , traffic modeling and network resource scheduling to deliver the emerging traffic in a timely, efficient, and reliable manner over long distance networks. In particular, RESO focuses of key issues such as :

- where and how to integrate the autonomy required to manage and control high speed networks at large scale?
- how dealing with the cost of resource and network virtualization on communication performance?
- which type of congestion control should be used in the context of a large scale deployment of high speed access networks (fiber to the home)?
- is the grid and high speed networks traffic self similar as the Internet traffic is? What are the critical scales? In which sense is self similarity harmful?
- how to efficiently share the bandwidth in a network that interleaves multimedia applications and computing applications?
- how to improve the interactions of message based communications or interactive traffic and transport layer in wide area networks?

To address some of these issues, our work follows four major research axes :

- Axis 1 : Optimized protocol implementations and networking equipments
- Axis 2 : Quality of Service and Transport layer for Future Networks
- Axis 3: High Speed Network's traffic metrology, analysis and modelling
- Axis 4: Network Services for high demanding applications

A large part of the axis 2 and axis 3 research topics is integrated in the ADR *Semantic Networking* of the common INRIA Bell Labs laboratory we are animating. The motivation of our research work in the common lab is to build and to exploit the knowledge that comes along with traffic. The goal is to act in a better way and to make better decisions at router and network level. The knowledge that comes part of traffic is what we refer to the 'semantics' of traffic. The main topics we are exploring in this research axis of the common laboratory are

- Traffic identification and classification
- Traffic sampling
- Flow analysis
- Flow scheduling
- Sampling-based scheduling
- Flow-based routing

2.3.1. Axis 1 : Optimized protocol implementations and networking equipments

In this research axis we focus on the implementation and on the optimization of the mechanisms and process within networking devices. Since several years, virtualization of the operating system is used in end system to improve security, isolation, reliability and flexibility of the environments. These mechanisms become a must in large scale distributed system. In our research axis1 we explore how these mechanisms can be also adapted and used in data transport networks and specifically in switching and routing equipments.

However, virtualization introduces an overhead which must be integrated to system performance models in order to forecast their behavior. Lot of performance problems on end systems but also on router's data plane have to be studied and solved to make the virtualization approach viable. Investigating these issues is one of the goals in this research axis.

On the other hand, the key enabling factor of new network services is programmability at every level; that is the ability for new software capabilities to self-configure themselves over the network. We explore the concept, "dynamic programming enablers" for dynamic service driven configuration of communication resources.

In this research axis we also explore the integration of context-awareness functionality to address two important issues : reliability of communications and energy consumption.

This direction is supported mainly by the EU FP7 "Autonomic Internet" project (2008-2010), with the INRIA "Action de Recherche Concertée" titled "Green-NET" (2008-2010). The ANR HIPCAL (2007-2009) grant helps our studies around network virtualisation

2.3.2. Axis 2 : Quality of Service and Transport layer for Future Networks

The goal of this axis is to guarantee quality of service in machine/user to machine/user communication while using efficiently the resources of the future networks. The two problems that are tackled here are: i) dynamic bandwidth sharing and congestion control in Future Internet and ii) control and flow management in semantic networks.

In this research axis, we focus on the three following questions:

- 1) which type of congestion control and transport protocol should be used in the context of large scale deployed high speed networks (fiber to the home, for example)?
- 2) how to efficiently share, but also dynamically provision the bandwidth of a network dedicated to computing tasks?
- 3) is the "flow-aware" approach a valuable solution to solve the end-to-end quality of service issue in the very high speed Future Internet?

2.3.3. Axis 3 : High Speed Network's traffic metrology and statistical analysis

Metrology of wide-area computer networks (i.e. the deployment of a series of tools allowing for collecting relevant information regarding the system status), is a discipline recently introduced in the context of networks, that undergoes constant developments. In a nutshell, this activity consists in measuring along time, the nature and the amount of exchanged information between the constituents of a system. It is then a matter of using the collected data to forecast the network load evolution, so as to anticipate congestion, and more widely, to guarantee a certain Quality of Service, optimizing resources usage and protocols design.

From a statistical signal processing viewpoint, collected traces correspond to (multivariate) time series principally characterized by non-properties: non-gaussianity, non-stationarity, non-linearities, absence of a characteristic time scale (scale invariance). Our research activity is undertaking the development of reliable signal analysis tools aimed at identifying these (non-)properties in the specific context of computer network traffic. In the course, we intend to clarify the importance of granularity of measurements.

Another challenge in network metrology is the effectiveness of packet sub-sampling. It means, to collect only a fraction of the overall traffic (supposedly redundant), and to study the possibility of inferring from that partial measurement, the most complete information about the system. Non trivial questions as, which fraction, which sub-sampling rule, adaptativity of this latter, smart sampling, statistical inference, open up a broad scope of investigation.

In this research axis, we focus on the two following questions:

- how does the traffic statistical properties really impact Quality of Service (QoS)?
- how to identify and to classify, in real time, transiting flows, according to a sensible typology?

Within the framework of the common laboratory between INRIA and Alcatel-Lucent, axis, "Semantic networking" brings in a new field of metrology research in RESO.

2.3.4. Axis 4: Network Services for high demanding applications

In strong interaction with the three fundamentals axes, this axis focuses on the application of the solutions to the grid context and on their implementation in a real environment such as the national research instrument Grid5000. Indeed, we believe that the precise structure of future applications and services is difficult to design without building large scale instruments and systems for real use based on real and performant hardware. Therefore, in this research axis we develop prototypes and deploy them within the Grid5000 testbed. For example, we design special measurement and routing systems at the edge of each Grid5000 site to explore new approaches or difficult problems alike. Topics that are investigated in this axis are strongly focusing the usage and the evolution of the Grid5000 instrument:

- Studies on the interactions of MPI and transport layer in wide area networks,
- Design, development and evaluation of a dynamic bandwidth provisioning service,
- Studies on the virtual private execution infrastructure concept for grid and cloud computing environments.
- Large scale deployment and evaluation of a high speed network measurement infrastructure.

RESO pursues researches for improving communications in grid environments. Thanks to systematic experiments of the behavior MPI in large scale environment, we merge optimizations of current implementations and propose new optimizations in the communication layers in order to execute more efficiently MPI applications on the Grid. We also study the impact of using TCP protocol for WAN communications (inter-site communications in the grid) and its interactions with MPI applications.

This research direction is mainly supported by FP6 EC-GIN grant, Grid'5000-ALADDIN initiative, CARRIOCAS project, HIPCAL project, JSPS-NEGST project. The DSLAB and OGF-Europe project provide resources.

2.4. Application domains

RESO applies its research to the domains of high end applications, distributed computing and to Grid and Cloud communications in particular. Grid computing aims at bringing together large collection of geographically distributed resources (e.g., computing, storage, visualization, etc.) to build on demand very high performance computing environments for compute and data-intensive applications. These large scale cybernetic infrastructures gain increasing attention from a broad range of actors: from research communities to computer providers, large companies, and telecommunication operators (telcos). Whereas grids have been widely in use in the scientific community, they are now moving into the commercial environment through the concept of Cloud computing solutions. Cloud computing fits a recentralization scenario which offers suitable business and security model for large scale distributed resource sharing. Telcos are now moving toward infrastructure sharing and grid computing. Different scenarios for telcos are envisioned: telcos (1) deploy grids internally, e.g. for rapid dynamic service provisioning to new customers; (2) link different sites via VPNs; (3) act as a service broker. These scenarii are explored with industrial partners. Researches conducted

these last years reveal that grid technology raise new challenges in terms of network optimisation as well as of protocol architecture and of transport paradigms. We believe that a broad deployment of the grid and cloud technology can modify and influence the design of the future Internet as other emerging communicating applications. RESO design network services and network middleware, to simplify the programming and to optimize the execution of high end communicating applications while fully exploiting the capacities of the evolving networking infrastructure.

2.5. Methodology

The RESO approach relies on a methodology based on a three-steps cycle: 1) a fine analysis of limitations encountered in existing protocols (mainly TCP/IP), 2) the exploration of disruptive solutions, 3) the theoretical and experimental evaluation of these proposals. This research focuses an heavily ossified research object (the Internet protocol architecture) and lies between a challenging emerging application domain on a specific network context. These factors induce a close interaction with both the application level and the underlying network level as well as a deep technical and scientific knowledge of protocols and network equipments. The methodology is then based on a continuous study of the high end and original requirements and on experimental evaluation of the functionalities and performance of emerging dedicated high speed infrastructures. RESO gathers expertise in advanced high performance local and cluster area networks protocols, in distributed systems and algorithmics, in protocol and protocol architecture design, in long distance networking, in time series analysis and in statistical inference. This background work provides the basis for innovative protocols and software design. Moreover, we implement and experiment our proposals on real, emulated local or wide area testbeds with real conditions and large scale applications.

2.6. Goals

RESO aims at providing software solutions but also original processes for high performance and flexible communications on very high speed networking infrastructures and for an efficient exploitation of these infrastructures. The goal of our research is to provide analysis of the limitations of the current communication and network software and protocols designed for standard networks and traditional usages, and to propose optimization and control mechanisms for the end-to-end performance, quality of service, energy efficiency and resource optimization. RESO explores original and innovative end-to-end transport services and protocols that meet the needs of high end applications. These solutions must scale in increasing bandwidths, heterogeneity and number of flows.

RESO studies high speed networks and their traffic characteristics, high end applications requirements, creates open source code, distributes it to the research community for evaluation and usage and help in shortening the wizard gap between network experts and novices. The long term goal is also to contribute to the evolution of protocols, standards and networking equipments, prompting the introduction of metrology as an intrinsic component of high-speed networks. An important effort is naturally dedicated to the dissemination of these new approaches.

2.7. Summary of the main contributions of the team in 2008

During this year, RESO team had main contributions in the following fields:

2.7.1. Axis 1 : Protocol implementations and networking equipments

- Exploration of the data plane virtualization cost and opportunities in software routers (VXRouter);
- High performance Autonomic Gateways for large scale distributed systems and Grids;
- High availability for clustered network equipments
- High availability for stateful network equipments

2.7.2. Axis 2 : End-to-end Quality of Service and Transport layer

- Optimization algorithms for network resource sharing in very high speed networks (BDTS).
- Algorithms for dynamic bandwidth provisioning based on flows scheduling and aggregation
- Design of a language for specifying virtual infrastructures (VXDL)
- Study on network and system virtualisation for virtual private infrastructure creation and usage.
- Distributed algorithms for bandwidth sharing in mobile or very high speed environments.
- Analysis of opportunity of extending the size of transfer units (XLFrames).
- Analysis of flow scheduling and sampling-based scheduling
- Analysis of flow-based routing

2.7.3. Axis 3: *High Speed Network's traffic metrology and statistical analysis*

- Design and development of a fine grain traffic capture and traffic analysis system dedicated to 10Gb/s speed links.
- Comparison of sampling methods for characterizing heavy tailed distributions in high speed networks traffic.

2.7.4. Axis 4: *Network Services for high demanding applications*

- Design of a network service for dynamic bandwidth provisioning in very high speed environments;
- Specification of a network resource scheduling, virtualization and reconfiguration component in a service oriented approach (Carriocas)
- Definition and development of MPI5000, a gateway for MPI over wide area network improvment
- Traffic monitoring of LCG (CERN LHC Grid) 10Gb/s link at packet resolution to characterize Grid traffic;
- Design and development of the HIPerNet software for network-aware virtual cluster management tool.
- Pursue the collaborations for the development and usage of the GRID5000 international optical interconnections to Netherland (DAS3) and Japan (Naregi) in collaboration with RENATER;
- Design and development of a metrology infrastructure for fine grain traffic monitoring in Grid5000.

3. Scientific Foundations

3.1. Optimized Protocol implementations and networking equipments

Participants: Narjess Ayari, Jean-Patrick Gelas, Olivier Glück, Laurent Lefèvre, Pascale Vicat-Blanc Primet, Jean-Christophe Mignot, Ludovic Hablot, Sébastien Soudan, Fabienne Anhalt, Olivier Mornard.

The initial goal of the DARPA Internet Architecture was to develop an effective technique for multiplexed utilization of existing interconnected networks. Robustness was the first priority which strongly colored the design decisions within the Internet architecture. An architecture primarily for commercial deployment would have clearly placed the resource management at the beginning of the priority list. Some of the most significant problems with the Internet today relate to lack of sufficient tools for distributed management. For example, in the large Internet being currently operated, routing decisions need to be constrained by policies for resource usage. Today this can be done only in a very limited way, which requires manual setting of tables. This is error-prone and at the same time not sufficiently powerful. A key enabling factor of new services and protocols is then the ability for new software capabilities to self configure themselves over the network. Moreover, in the Future Internet, only trusting nodes should be able to communicate at will. Nodes should also be protected from nodes they do not want to communicate with. Virtualization, context-awareness and energy efficiency are then promizing concepts for Future networks. However their potential and limits in the context of dynamic and self-organized high speed networks have to be studied.

Since several years, virtualization of the operating system is used in end system to improve security, isolation, reliability and flexibility of the environments. These mechanisms become a must in large scale distributed system. Virtualized resources is a new way of sharing in which group of users or activities (or trusting nodes) are given static shares, and only within these groups there is dynamic sharing. Virtual networks present ideal vantage point to monitor and control the underlying physical network and the applications running on the virtual trusting nodes. How virtualization can be also adapted and used in data transport networks and specifically in switching and routing equipments is an open question. For example, virtualization introduces an overhead which must be integrated to system performance models in order to forecast their behavior. Lot of performance problems on end systems but also on router's data plane have to be studied and solved to make the virtualization approach viable. Investigating these issues is one of the goals in this research axis.

On an other hand, the key enabling factor of new network services is programmability at every level; that is the ability for new software capabilities to self-configure themselves over the network. We explore the concept, "dynamic programming enablers" for dynamic service driven configuration of communication resources. Dynamic programming enablers apply to an executable service that is injected and activated into the network system elements to create the new functionality at runtime. The basic idea is to enable trusted parties (users, operators, and service providers) to activate management-specific service and network components into a specific platform. We study mechanisms and infrastructures required to support these components. We aim at providing new functionality to services using Internet facilities, addressing the self-management operations in differentiated and integrated services. The goal is the enhancement of the creation and the management (customization, delivery, execution and stop) of Internet services.

In this research axis we also explore the integration of context-awareness functionality to address two important issues : reliability of communications and energy consumption.

Session awareness : Most of the NGN services involve a session model based on multiple flows required for the signalling and for the data exchange, all along the session lifespan. New service-aware dependable systems are more than ever required. Challenges to these models include the client and server transparency, the low cost during failure free periods and the sustained performance during failures. Based on our previous work with FT R&D ("Procédés de gestion de sessions multi-flux", N. Ayari, D. Barbaron, L. Lefèvre, France Telecom R&D Patent, June 2007), we continue to explore and propose session aware distributed network solutions which support the reliability mandatory to operators services (VOIP).

Energy awareness : Large scale distributed systems (and more generally next generation Internet) are facing infrastructures and energy limitations (use, cooling etc.). In the context of monitored and controlled energy usage, we plan to explore the proposal of energy aware equipments and frameworks, which allow users and middleware to efficiently use large scale distributed architectures.

We are developing solutions to dynamically monitor energy usage, inject this information as a resource in distributed systems and adapt existing jobs (OAR) and network (BDTS) schedulers to autonomically benefit from energy information in their scheduling decisions. This research is linked with experimental evaluation on Grid'5000 platform and inside the ALADDIN initiative.

3.2. Quality of Service and Transport layer for Future Networks

Participants: Pascale Vicat-Blanc Primet, Dino Lopez Pacheco, Laurent Lefèvre, Sébastien Soudan, Romaric Guillier, Dinil Mon Divakaran, Guilherme Koslovski, Isabelle Guerin-Lassous, Remi Vannier, Marcelo Pasin, Pierre Bozonnet.

Congestion control is the most important and complex part of a transport protocol in a packet switched shared network. The congestion control algorithm is then a key component which has to be considered to alleviate the performance problems in the future networks environments. TCP has shown a great scalability in number of users, but not in link capacity and link diversity. For example, TCP performance can be very low and unstable in data-center applications and interactive communications within high speed long distance networks infrastructures, like lambda grids environments. The conservative behavior of TCP with respect to congestion in IP networks is at the heart of the current performance issues faced when the traffic load is highly dynamic.

On the application side, one can observe that traditional applications were originally characterized by very basic communication requirements related to performance, reliability and order. The rapid deployment of new heterogeneous network technologies has pushed the development of an important number of new multimedia applications presenting complex requirements in terms of delay, bandwidth constraints and tolerance to losses. These applications need specific mechanisms to adapt to network congestion or changing medium conditions. To solve this problem, protocol enhancements and alternative congestion control mechanisms have been proposed for very high speed optical networks, wireless networks and for multimedia applications (see PFLDNET conference series). Most of them are now implemented in current operating systems, but these protocols are not equivalent, and not all of them are suitable for all environments and all applications, moreover they may not cohabit well. Since a couple of years, the evaluation and comparison of new transport protocols received an increasing amount of interest (see IRTF TMRG and ICCRG groups). However, TCP and other alternatives are complex protocols with many user-configurable parameters, and a range of different implementations. Several aspects can be studied, and various testing methods exist. The research community recognises that it is important to deploy measurement methods so that the transport services and protocols can evolve guided by scientific principles. Researchers and developers need agreed-upon metrics, a common language for communicating results, so that alternative implementations can be compared quantitatively. Users of these variants need performance parameters that describe protocol capabilities so that they can develop and tune their applications. Protocol designers need examples of how users will exercise their service to improve the design.

As the Internet has evolved from a research project into a popular consumer technology, it may not be reasonable to assume that all end hosts would fairly cooperate. Indeed, concerns were raised that the recently started deployment of non-IETF-approved high-speed TCP variants could lead to an "arms race" that would eventually have a detrimental effect on the overall performance of the Internet. As another example, commercial Internet accelerators can provide better performance for a single user at the expense of other users. In the future, expecting billions of Internet devices to fairly cooperate to prevent network congestion is overly optimistic. New bandwidth sharing approach have to be investigated.

Flow scheduling [3], based on the in-advance knowledge of resource requirement of an application or online estimation of these requirements can be applied. Signaling or real time flow analysis and also scalability issues have to be explored. Distributed and lagrangian relaxation-based solution for *bandwidth sharing* is also an interesting approach in Future Internet. This approach addresses well the dynamic feature, due to node mobility or traffic variation. However, some problems remain open. First, the sharing models are often very complex to compute, while still being inaccurate. Second, some parameters of allocation algorithms based on lagrangian relaxation are difficult to set, and are often obtained by trial ad error; and hence not optimized. Finally, the proposed solutions are often tested on home-made simulators that are far from being realistic.

We believe some network resource control has to be associated with the end to end flow control approach to offer better quality of experience in Future networks. Network resource control is classified into three timescales: data, control, and management. Each timescale corresponds with a level of aggregation : 'data' deals with packets; 'control' deals with aggregates of packets, i.e. flows; and 'management' deals with aggregates of flows. All three timescales must be addressed, since they all affect the service perceived by users, and the ease and efficiency with which the network can be operated. The current Internet protocols do not well address the control time scale, and do not consider packet aggregates. TCP deals with resource control at data timescales; while routing protocols, such as BGP and OSPF operate at timescales of the order of minutes or hours (management time-scale).

We propose to explore packet aggregates and address control timescale in the context of Future Internet not only for performance, but also for manageability and security purposes.

On an other hand, the optical fiber communication will be the predominant mechanism for data transmission in core network and may be also at the access. To address the anticipated terabit demands, dynamically reconfigurable optical networks are envisioned. This vision will be realized with the deployment of configurable optical components, which are now becoming economically viable. To meet the terabit challenge, network designers will enhance core functionality by migrating to devices equipped with tunable transceivers, optical

crossconnects and optical add/drop multiplexers. Optical Cross-Connects (OXC) becomes more and more, cheap, simple and controllable. The control-plane, traditionally in the hand of telco migrates progressively to the customers. Studying the interactions of components required to accomplish the tasks of bandwidth reservation, path computation and network signaling is an other goal.

3.3. High Speed Network's traffic metrology and statistical analysis

Participants: Pascale Vicat-Blanc Primet, Paulo Gonçalves, Patrick Loiseau, Matthieu Imbert, Damien Ancelin, Olivier Grémillet.

Tools for measuring the end-to-end performance of a path between two hosts are very important for transport protocol and distributed application performance optimization. Bandwidth evaluation methods aim to provide a realistic view of the raw capacity but also of the dynamic behavior of the interconnection that may be very useful to evaluate the time for bulk data transfer. Existing methods differ according to the measurements strategies and the evaluated metric. These methods can be active or passive, intrusive or non-intrusive. Non-intrusive active approaches, based on packet train or on packet pair provide available bandwidth measurements and/or the total capacity measurements. None of the proposed tools, based on these methods, enable the evaluation of both metrics, while giving an overview of the link topology and characteristics.

That is the reason why a metrology activity including data processing, statistical inference, time series and stochastic processes analysis, deemed important to embed in the main research realm of RESO. Our goal is for these analyses to become in the near future a plain component not only in the study and in the development of infrastructures and computing networks, but also in real-time resources identification and management.

Grids specificities, such as the cooperating equipments number and heterogeneity, the number of independent processes, the treatments, bandwidth and stock capacities, turn indispensable to revisit the algorithms, as well as the control and operating mechanisms, in order to reach appropriate and optimal performances.

To validate a priori hypothesis that sustain already investigated approaches (e.g. overlay, virtualizing network resources, distributing network treatments, middleware programming), we foresee to resort to metrology and to the statistical analysis of the collected data. Indeed, we believe that automatic identification of static and dynamic properties of network resources is a prerequisite for developing adequate algorithms.

To drive us in this task, we will rely on the impressive amount of studies devoted to the internet traffic analysis, and on the established results that have been obtained in the last years [87], [91], [74], [75]. For instance, we are interested in verifying if the conjecture relating long range dependance (LRD) in traffic flows with heavy tailed distributions of files sizes [91], still holds with grid networks. Difficulty dwells in a reliable estimation of density functions from loosely sampled data, and in the even more tricky one of accurate estimation of LRD parameters and tail exponents from incomplete data sets [5]. To tackle these issues, P. Gonçalves (former Mistis (ex IS2) project member) and P. Loiseau (PhD student, Ms in Physics) supply RESO with the necessary inter-disciplinary competence in signal processing and statistics.

Finally, the great investment that has been granted to Grid5000 (and to the interconnections Grid5000-NAREGI and Grid5000-DAS 3) will profitably be used providing us with a high-performance and quite novel experimental setup to confront the proposed theoretical models with real traffic measurements.

3.4. Network services for high demanding applications

Participants: Pascale Vicat-Blanc Primet, Olivier Glück, Laurent Lefèvre, Jean-Patrick Gelas, Paulo Gonçalves, Lucas Nussbaum, Patrick Loiseau, Olivier Mornard, Sébastien Soudan, Ludovic Hablot, Romaric Guillier, Manoj Dahal, Marcelo Pasin, Pierre Bozonnet, Aurélien Cedeyn, Oana Goga.

The purpose of Computational Grids was initially to aggregate a large collection of shared resources (computing, communication, storage, information) to build an efficient and very high performance computing environment for data-intensive or computing-intensive applications [83]. But generally, the underlying communication infrastructure of these large scale distributed environments is a complex interconnection of multi-IP domains with non controlled performance characteristics. Consequently *the Grid Network cloud* exhibits extreme heterogeneity in performance and reliability that considerably affect the global application performance.

The performance problem of the grid network cloud can be studied from different but complementary view points.

- Measuring and monitoring the end-to-end performance helps to characterize the links and the network behavior. Network cost functions and forecasts, based on such measurement information, allow the upper abstraction level to build optimization and adaptation algorithms.
- Optimally using network services provided by the network infrastructure for specific grid flows is of importance.
- Modeling, managing and controlling the grid network resource as a first class resource of the global environment: transfer scheduling, data movement balancing, bandwidth reservation and dynamic provisioning...
- Creating enhanced and programmable transport protocols adapted to heterogeneous data transfers within the grid may offer a scalable and flexible approach for performance control and optimization.

In a grid environment, two key points in the communication layers need to be taken in consideration in order to execute efficiently high performance applications: the heterogeneity of high-speed interconnects composing the grid and the Wide Area Network used to achieve inter-site communications. We explore new mechanisms to improve the application performance when it executes on the grid. We study, in particular, how a MPI application can benefit, during one execution, of several high-speed networks at the same time. In particular, it implies to find a way to communicate efficiently between these heterogenous interconnections. We also explore how to keep good performance execution when long-distance communications are necessary because the application is launched on multiple sites of the grid.

An efficient MPI implementation for the grid is one of our research topic in this axis with the aim of improving communications in grid environments. The MPI standard is often used in parallel applications for communication needs. Most of them are designed for homogeneous clusters, but MPI implementations for grids have to take into account the heterogeneity of high-speed interconnects composing the grid and the Wide Area Network used to achieve inter-site communications, in order to maintain a high performance level. These two constraints are not considered together in existing MPI implementations, and raise the question of MPI efficiency in grids. Our goal is to significantly improve the performance execution of MPI applications on the grid.

Finally, the resource mutualisation and sharing paradigm proposed by the Grid remains a very promising and powerful concept that we apply to network resource sharing at many levels. To explore new approaches or difficult problems alike, we design and deploy special shared network resource at the edge of Grid5000 sites [2]. The goal is develop "proof of concept" experiments for exploring, among others, traffic awareness, the buffer sizing problem, buffer and filtering "in route" approaches, router virtualization, multipath routing, and router assisted transport protocols and communication libraries (MPI5000).

4. Application Domains

4.1. Panorama

Keywords: *Autonomic Networks, Communication Software, End to End Transport, Grids, High Performance, Networks, Protocols, Quality of Service, Telecommunications.*

RESO applies its research to the domains of high performance Cluster and Grid communications. Existing GRID applications did already identify potential networking bottlenecks, either caused by conceptual or implementation specific problems, or missing service capabilities. We participated to the elaboration of the first GGF document on this subject [89] [88], [90]. Loss probability, important and incompressible latencies, dynamic behavior of network paths question profoundly models and technic used in parallel and distributed computing [82]. The particular challenge arises from a heavily distributed infrastructure with an ambitious end-to-end service demand. Provisioning end-to-end services with known and knowable characteristics in

a large scale networking infrastructure requires a consistent service in an environment that spans multiple administrative and technological domains. The first bottleneck is often located at the interface between the local area network (LAN) and the wide area network (WAN). RESO conducted several actions in the field of Grid High Performance Networking in the context of the OGF, the European or National projects. These activities have been done in close collaboration with other INRIA and CNRS French teams (Grand Large, Mescal, Graal) involved in the GRID5000 and the Grid Explorer projects and other European teams involved in pflidnet and Glif communities. RESO joined the CARRIOCAS project which studies and implements a very high bit rate (up to 40 Gb/s per wavelength) network interconnecting super computers, storage systems and high resolution visualization device to support data and computing intensive applications in industrial and scientific domains. Our activities cover networking intelligence for high performance distributed applications. Finally, the evolution of the Internet usage pushing the convergence of communication and computation at every level confirms our initial vision : the network should not be seen only as a black box providing pipes between edge machines, but as a vast cloud increasingly embedding the computational and storage resources to meet the requirement of emerging applications [6]. These resources are generally located at important crossroads and access points throughout the network. During the last few years we have seen that the distinction between packet forwarding and application processing has become blurred. The network community now starts to worry not only about forwarding packets without regard to application semantics, but is increasingly trying to exploit new functionalities within the network to meet the requirement of the application. Reciproquely, distributed systems and applications have traditionally been designed to run on top of the Internet, and to take the architecture of the Internet as given. The convergence of communication and computation at every level appears to be natural. It is however important to explore the full range of possibilities it can bring. Most of the proposals exploiting this convergence break the initial design philosophy of the Internet protocol stack (end to end argument for example), or if implemented in the application layer present lot of performance, resilience and scalability issues. We think that the Internet re-design raises the opportunity to better understand and assess higher-level system requirements, and use these as drivers of the lower layer architecture. In this process, mechanisms that are implemented today as part of applications may conceivably migrate into the network itself, and this is one of main driver of the researches of RESO and of our strong involment in the new INRIA-BellLabs "Semantic Networking" research axis.

- RESO is closely involved in the evolution of the Grid 5000 testbed, and responsible for the networking aspects. Grid5000 is a national initiative aiming at providing a huge experimental instrument to the grid software and computer science research community. RESO participate to the INRIA development action ALADDIN. Participating to the design, deployment and usage of such high performance experimental Network and Grid testbed allow us to gather a strong deep experience and unique expertize in high speed network and protocols exploration and tuning.
- RESO pursue the construction of an international community around Grid networks through the european EC-GIN project as well as with the OGF networking community.
- Through the ANR IGTMD project and is collaborating with the LCG and real physicists. A dedicated link deployed between IN2P3 (one of the largest computing center in France) and the FermiLab laboratory in Chicago, enable us to perform transport protocol experiments as well as traffic capture.
- RESO is bringing its expertize in Grids and Grid Networking to the CARRIOCAS project of the "pôle Ile de France System@tic. This collaboration enable us to explore the limits and the advantages of our previous results in the context of a 40Gb/s dynamically provisionable network.
- Through the ANR IGTMD project and is collaborating with the LCG and real physicists. A dedicated link deployed between IN2P3 (one of the largest computing center in France) and the FermiLab laboratory in Chicago, enable us to perform transport protocol experiments as well as traffic capture.
- Through the ANR HIPCAL project, RESO is collaborating with biology and medical imaging applications.

5. Software

5.1. BDTS: Bulk Data Transfer Scheduling Service

Keywords: *bandwidth provisioning, bulk data transfers, network bandwidth allocation.*

Participants: Marcelo Pasin, Sébastien Soudan, Dinil Mon Divakaran, Pascale Vicat-Blanc Primet.

The coordination of resource allocation among end points in controlled networks may require a service to transfer large data sets within time intervals. Such transfers commonly start some time after its request, use any variable bandwidth, and must complete before a deadline. BDTS is a software for Bulk Data Transfer Scheduling which gives users, applications or middleware the possibility to specify transfer requests as transfer jobs, and ensure a transparent control of them within controlled networks. BDTS manipulates and produces profiles which are step functions that express variable bandwidth assignement over time. BDTS incorporates a scheduler to divide the time windows of overlapping transfer jobs into multiple intervals. BDTS implements a multi-interval scheduling algorithm which minimizes the congestion factor of the network. It assigns independent bandwidth values to each transfer job at each interval, producing a bandwidth profile for each transfer job.

During this year, BDTS software has been enhanced to better fit the network model where users can provision their own infrastructure. This has been done first by defining a model of interaction between the different actors: users, service providers and network operator with the associated scheduling and network provisioning problem. And secondly by defining a solution based on a linear program to solve it. This part is now integrated in BDTS and will be reused to provision Carriocas's pilot network. BDTS has been demonstrated in SC'08. BDTS is distributed under LGPL licence and downloadable at : <http://www.ens-lyon.fr/LIP/RESO/Software>

5.2. FLOC: Flow control

Keywords: *end-host based traffic shaping, rate enforcement.*

Participants: Pascale Vicat-Blanc Primet, Sébastien Soudan.

This software solves the problem of enforcing a rate allocation profile made by an external bandwidth scheduler in a packet network. FLOC is the daemon present on end machine responsible to enforce the multi-interval bandwidth allocation profile received from a scheduler to a socket identified by a token and registered by user applications. FLOC changes GNU/Linux kernel's `qdisc` configuration according to current date and profile so that senders can only send at a given time at the rate they are allowed to. FLOC is distributed under LGPL licence and downable at : <http://www.ens-lyon.fr/LIP/RESO/Software>

5.3. NXE: Network eXperiment Engine

Keywords: *Network experiment, protocol evaluation, reproducible experiment, workflow.*

Participants: Pascale Vicat-Blanc Primet, Romaric Guillier.

NXE (for Network eXperiment Engine) is a tool developed to be able to execute any particular scenario over any given topology. A scenario is defined as a sequence of dates at which networking events such the start of a new bulk data transfer occurs. This software automate the selection, deployment , configuration and activation on distributed resources of pieces of software required to execute a large scale and reproducible networking experiment. This software has been demonstrated during the SuperComputing'2007 event on the INRIA booth and is adapted and used for deploying different types of networking experiment (controlled measurement, network device evaluation, virtualization overhead measurement, HIPerNet validation...) within RESO team. A graphical user interface has been developed to simplify the usage of the automation tool. NXE is distributed under LGPL licence and downable at : <http://www.ens-lyon.fr/LIP/RESO/Software>

5.4. HSTTS: High Speed Transport protocols Test Suite

Keywords: *Performance evaluation, TCP, transport protocol comparison.*

Participants: Pascale Vicat-Blanc Primet, Romaric Guillier.

HSTTS (for High Speed Transport protocol Test Suite) is software implementing a fixed set of data transfer scenarios. It is designed to help users evaluate the performance they ought to be able to get out of their networking infrastructure when they transfer data by using different types of transport protocols and services. This software has been presented during the SuperComputing'2007 event on the INRIA booth. BDTS is distributed under GPL licence and downloadable at : <http://www.ens-lyon.fr/LIP/RESO/Software>

5.5. HIPerNet: management of virtual private execution infrastructure

Keywords: *Grid5000, Resource discovery, Virtual private execution infrastructure, Virtualization, allocation, scheduling and management, selection.*

Participants: Pascale Vicat-Blanc Primet, Olivier Mornard, Jean-Patrick Gelas.

HIPerNet engine is software implementing discovery, selection, allocation, scheduling and management of virtual private execution infrastructures over the Internet. This software has been presented during the SuperComputing'2008 event on the INRIA booth. HIPerNet is distributed under GPL licence and downable at : <http://www.ens-lyon.fr/LIP/RESO/Software>

5.6. SNE (Stateful Network Equipment)

Keywords: *High Availability, fault tolerance.*

Participant: Laurent Lefèvre [contact].

Joint work with Pablo Neira Ayuso from University of Sevilla (spain).

SNE is a complete library for designing a stateful network equipment (contains Linux kernel patch + user space daemon). The aim of the SNE library is to support issues related to the implementation of high available network elements, with specially focus on Linux systems and firewalls. The SNE library (Stateful Network Equipment) is an add-on to current High Availability (HA) protocols. This library is based on the replication of the connection tracking table system for designing stateful network equipments. SNE is an open source project, available on the web (CECILL Licence) at <http://perso.ens-lyon.fr/laurent.lefevre/software/SNE>.

5.7. ShowWatts: Real time energy consumption grapher.

Keywords: *Energy, measure, user interface.*

Participants: Laurent Lefèvre, Anne-Cécile Orgerie, Jean-Patrick Gelas [contact].

Simple software used to display real time measures of energy consumed by processing nodes in a grid architecture. This software proposes a graphical interface connected to a set of powermeter devices. Graphical interface can display measures coming through a long distance secured network tunnel.

5.8. XCP-i (Interoperable eXplicit Control Protocol)

Keywords: *XCP, high performance transport protocol.*

Participants: Dino Martin Lopez-Pacheco, Anne-Cécile Orgerie, Laurent Lefèvre.

XCP (eXplicit Control Protocol) is a transport protocol that uses the assistance of specialized routers to very accurately determine the available bandwidth along the path from the source to the destination. We propose XCP-i [86] which is operable on an internetwork consisting of XCP routers and traditional IP routers without loosing the benefit of the XCP control laws

An ns-2 module simulating XCP-i has been developed and will be available on the web. Based on a Linux kernel, a software XCP-i router is currently under development.

6. New Results

6.1. Optimized protocol implementation and networking equipments

6.1.1. Evaluation and optimization of network performance in virtual end systems and routers

Keywords: *system virtualization, traffic control, virtual router.*

Participants: Fabienne Anhalt, Jean-Patrick Gelas, Pascale Vicat-Blanc Primet.

Virtualization techniques are applied to improve features like isolation, security, mobility and dynamic reconfiguration in distributed systems. To introduce these advantages into the network where they are highly required, an interesting approach is to virtualize the internet routers themselves. This technique could enable several virtual networks of different types, owners and protocols to coexist inside one physical network. Systematic analysis and experiments of the cost of network virtualization have been conducted. Optimisation of scheduler have been proposed.

The evaluation of Xen 3.1's network performance with TCP flows on end-hosts and routers show that the multiplexing level (Dom0) was the bottleneck . We have shown that the performance could be improved by manipulating the scheduler parameters. Giving more weight to dom0 improves throughput and fairness.

The evaluation of Xen 3.2's network performance with TCP and UDP flows on end-hosts and routers have shown that the performance improved significantly compared to previous 3.1 version of Xen. Better throughput was obtained and dom0's CPU overhead decreases. No more unfairness exists. A stated bottleneck is the forwarding of small sized packets.

We have proposed a model of a virtual router we have implemented with XEN and we have evaluated its properties. We show that the performance is close to the performance of non virtualized software routers, but causes an important processing overhead and unfairness in the share of the ressources. We study the impact of the virtual machine scheduler parameters on the network performance and we show that the module which is responsible of forwarding the packets between the virtual machines and the physical interfaces is the critical point of network communications. We analysed virtualization from the the data plane perspective. We explore the resulting network performance in terms of throughput, packet loss and latency between virtual machines, and also the correspondig CPU cost. The virtual machines act as senders or receivers, or as software routers forwarding traffic between two interfaces in the context of Xen. Our results show that the impact of virtualization on network performance is getting smaller with the successive Xen versions, making this approach a promising solution for data plane virtualization. The router migration with Xen 3.2 has been explored. The migration process is slowed down by the forwarding of flows by the virtual router, especially with TCP flows. It can take several minutes instead of several seconds in case of inactivity. Combining end-host, server and router virtualization could offer isolated and malleable virtual networks of different types, owners and protocols, all sharing one physical infrastructure.

6.1.2. High performance Autonomic Gateways for large scale distributed systems and Grids

Keywords: *execution environments, programmable and active networks.*

Participants: Jean-Patrick Gelas, Laurent Lefèvre.

In the framework of a cooperative industrial maintenance and monitoring project (TEMIC project), in which we are involved with different academic and industrial partners, we design devices to be easily and efficiently deployable in an industrial context. Once the hardware deployed and used, it must also be easily removable at the end of the maintenance or monitoring contract. In this project, we deploy our devices in secured industrial departments, restricted areas, or in an out-of-the-way locations. These devices must act as auto-configurable and re-programmable network nodes. Thus, the equipments must be *autonomic* and must not require direct human intervention.

The design of an autonomic network equipment must take into account specific requirements of active equipments in terms of dynamic service deployment, auto-settings, self-configuration, monitoring but also in terms of hardware specification (limited resources, limited mechanical parts constraints, dimension constraints), reliability and fault tolerance.

We proposed an adaptation of a generic high performance active network environment (Tamanhoir) in order to deploy on limited resources based network boxes and to increase reliability and scalability. The implementation process is based on a hardware solution provided by the Bearstech company. Through this approach we proposed the architecture of an **Industrial Autonomic Network Node** (called *IAN²*) able to be deployed in industrial platforms [81], [84]. We evaluated the capabilities of *IAN²* in terms of computing and networking resources and dynamic re-programmability.

6.1.3. High availability for clustered network equipments

Keywords: *fault tolerance, high availability, scalability.*

Participants: Narjess Ayari, Laurent Lefèvre, Pascale Vicat-Blanc Primet.

A key component for improving the scalability and the availability of network services is to deploy them within a cluster of servers. The main objective of this work is to design a network traffic load balancing architecture which meets fine grained scheduling while efficiently spreading the offered network traffic among the available cluster resources.

- **A scalable architecture for balancing the offered network traffic**

While a lot of researches have been conducted in the field of job and network load balancing, less interest has been granted to the impact of the granularity of the used mechanism on the reliable execution of the upper layer services. In fact, the currently used flow level network load balancing frameworks fail to achieve session awareness while efficiently spreading the offered network load among the available resources, typically, when the offered network session involves multiple and heterogeneous flows. Representative services range from familiar services like HTTP and FTP, to some recent services like multimedia streaming using RTSP/RTP/RTCP and Voice over IP using SIP. Our work aims to provide an architecture to efficiently balance the offered network sessions among the available processing resources within a cluster of servers.

- **A highly available architecture for balancing the offered network traffic**

High availability allows service architectures to meet growing demands and to ensure uninterrupted service. In our work, we are interested in providing the continuous execution of the offered network sessions in case of failure of the legitimate entry point to the cluster as well as in case of the failure of the processing server inside the cluster. We noticed that current fault tolerant frameworks need to support consistent transport and application level failover mechanisms, and that transport layer protocols do not provide high availability capabilities. Indeed, TCP does not distinguish between a packet loss due to congestion, or a packet loss due to a server overload or due to a server/link failure. Thus, it reacts the same way to packet losses and to delays, by retransmitting the same segment to the same remote end point of the connection. Moreover, TCP tolerates short periods of disconnection not longer than a few RTTs. It disconnects the communicating hosts once specific timers expire. On the other hand, transport protocols rely on an explicit association between a service and its physical location for the wired Internet. Thus, when a host fails, the end-to-end flow terminates.

In order to address this limitation, we proposed an active replication based system which enhances the reliability of the already established TCP flows. The proposed scheme is client transparent and does not incur any overhead to the end-to-end communication during failsafe periods, and performs well during failures. Parts of this work are protected by the Intellectual Property National Institute (INPI) patent disclosure N°FR0653546

[80], [79], [76], [78], [77]

Narjess Ayari has defended her PhD in October 2008.

6.1.4. High availability for stateful network equipments

Keywords: *fault tolerance, high availability.*

Participant: Laurent Lefèvre.

Joint work with Pablo Neira Ayuso from University of Sevilla (Spain).

In operational networks, the availability of some critical elements like gateways, firewalls and proxies must be guaranteed. Some important issues like the replication of these network elements, the reduce of unavailability time and the need of detecting failure of an element must be studied. We propose the SNE library (*Stateful Network Equipment*) which is an add-on to current High Availability (HA) protocols. This library is based on the replication of the connection tracking table system for designing stateful network equipments.

Proposing stateful network equipments on open source systems is a challenging task. We propose the basic blocks (SNE library) for building a stateful network equipment. This library can be combined with high-availability protocols (CARP, Linux HA...). We focus on Linux system in order to provide software solutions for designing high-available solutions for NAT, firewalls, proxies or gateways equipments...This library is based on components located in kernel and in user space of the network equipment. First micro-benchmark of communications mechanisms with Netlink sockets have shown the effectiveness of our approach

6.1.5. Router assisted network transport protocol

Keywords: *TCP, XCP, congestion control, estimations, variable bandwidth.*

Participants: Dino Martin Lopez-Pacheco, Laurent Lefèvre.

In heterogeneous networks, where many flows, non-regulated and/or with a high QoS level, share the resources, the available best-effort bandwidth varies over time. This changes can be represented by an aggregation of UDP ON-OFF sources what produces a step-based variation model. In this type of environments, we have tested the performance of many transport control protocols (TCP New Reno, High Speed TCP, TCP Westwood+ and XCP) using the ns2 simulator. In our studies, XCP showed always the best performance, with a high stability and fairness level. But in heterogeneous networks, the lost of packets is very common, so we have tested XCP in a network where the lost in the reverse path cause some ACK losses. In the new results, we have found that the ACK losses produce many problems in the connections, caused by a wrong calculus of the congestion window size, specifically when the available bandwidth decreases. That is because the success of XCP is based on the network state information, provided by the routers to the sender in the ACK packets. Since, the problem is generated by the wrong calculus of the congestion window size in the sender side, we proposed to compute this value in the receiver side. We have called this new approach XCP-r

We repeated the simulations set using XCP-r and we found that XCP-r shows always more stability and better fairness level.

6.1.6. XCP-i: a new interoperable XCP version for high speed heterogeneous networks

Keywords: *TCP, XCP, XCP-i, available bandwidth, congestion control, virtual XCP-i router.*

Participants: Dino Martin Lopez-Pacheco, Laurent Lefèvre.

XCP (eXplicit Control Protocol) is a transport protocol that uses the assistance of specialized routers to very accurately determine the available bandwidth along the path from the source to the destination. In this way, XCP efficiently controls the sender's congestion window size thus avoiding the traditional slow-start and congestion avoidance phase. However, XCP requires the collaboration of all the routers on the data path which is almost impossible to achieve in an incremental deployment scenario of XCP. It has been shown that XCP behaves badly, worse than TCP, in the presence of non-XCP routers thus limiting dramatically the benefit of having XCP running in some parts of the network. In this work, we address this problem and propose XCP-i which is operable on an internetwork consisting of XCP routers and traditional IP routers without loosing the benefit of the XCP control laws.

XCP-i basically executes the next four steps to discover and compute a new feedback that reflects the state of the network where non-XCP routers are placed:

1. Discover where the non-XCP routers are in the data path.
2. Discover the upstream and downstream XCP-i routers of the non-XCP routers.
3. Estimate the available bandwidth where the non-XCP routers are placed.
4. Create a virtual XCP-i router that computes a new feedback using the estimated available bandwidth before.

The simulation results on a number of topologies that reflect the various scenario of incremental deployment on the Internet show that although XCP-i performances depend on available bandwidth estimation accuracy, XCP-i still outperforms TCP on high-speed links [86]. Dino Lopez Pacheco has defended his PhD in June 2008.

6.1.7. *Autonomic Service Deployment in Next Generation Networks*

Keywords: *autonomic network, programmability, service deployment.*

Participants: Abderhaman Cheniour, Jean-Patrick Gelas, Laurent Lefèvre.

RESO is involved in the FP7 Autonomic Internet project by focusing on autonomic service deployment solutions for large scale overlays.

Programmability in network and services encompasses the study of decentralised enablers for dynamic (de)activation and reconfiguration of new/existing services, including management services and network components. The challenge in Autonomic Internet FP7 project (AutoI) is to enable trusted parties (users, operators, and service providers) to activate management-specific service and network components into a specific platform. Dynamic programming enablers will be created that are applied to executable service code, which can be injected/activated into the system's elements to create the new functionality at runtime. Network and service enablers for programmability can therefore realise the capabilities for flexible management support required in AutoI.

RESO has proposed the ANPI : Autonomic Network Programming Interface which will support the service enablers plane of the AUTOI architecture. This interface is currently under development with the support of other AUTOI partners (Hitachi Europe, University College of London, UPC Barcelona, Univeristy of Passau).

6.1.8. *Energy-efficiency in computing and networking for large-scale distributed systems*

Keywords: *Energy-awareness, Energy-efficiency, Grid monitoring.*

Participants: Alejandro Fernandez, Jean-Patrick Gelas, Isabelle Guerin-Lassous, Laurent Lefèvre, Anne-Cécile Orgerie.

High performance computing aims to solve problems that require a lot of resources in terms of power and communication. While an extensive set of research project deals with the saving power problem of electronic devices powered by electric battery, few have interest in large scale distributed systems permanently plugged in the wall socket. The general common idea is indeed that, when they are not reserved, the grid resources should be always available, so that they should always remain fully powered on.

The large-scale distributed systems are sized to support reservation bursts. So they are not fully used all the time. Between the bursts, some resources remain free, so we can save energy during these gaps. This is our first approach taken in this work: to save energy by shutting down nodes when they are not used. We use the same approach for high performance data transport: the high-speed links are not always fully used and we can turn off the Ethernet cards and switch ports off to save energy.

Understanding the characteristic usage and workloads of the large-scale distributed systems is a crucial step towards the design of new energy-aware distributed system frameworks. Therefore we have studied the Grid5000 platform usage over long periods of time.

The analysis of these usage traces lead us to propose an energy-aware reservation infrastructure (EARI) which is able to shut down nodes when they are idle. This infrastructure proposes several energy efficient solutions for a reservation made by a user: several energy-efficient possibilities for his reservation. Thus the user is able to choose among these “green” solutions and this leads to an aggregation of the reservations. This infrastructure also includes a prediction algorithm to anticipate the next reservation in order to avoid shutting down nodes that we will need to be restarted quickly.

So, our infrastructure is based on three mechanisms:

- switching on and off the nodes;
- reservation aggregation with green policies and
- predictions of the next reservations.

This model has been validated over the Grid5000 traces by using a replay mechanism. The results are really encouraging and show that our infrastructure could make huge energy savings. This on/off model is a first step in our research on energy efficiency in computing and networking for large-scale distributed systems.

We are working on improving the prediction models with Alejandro Fernandez from University of Seville, Spain.

This work has been published in two international conferences: [35] and [36] and has been presented in three posters: [62], [64] and [63].

6.2. Quality of service and Transport Protocols for Future Networks

6.2.1. Flow Scheduling and endpoint rate control

Keywords: *bulk data transfers, flow scheduling, pacing, rate limitation.*

Participants: Pascale Vicat-Blanc Primet, Sébastien Soudan, Marcelo Pasin.

In this work we propose to manage explicitly the movements of massive data set between end point. We formulate the bulk data transfer scheduling problem and give an optimal solution to minimize the network congestion factor of a dedicated network or an isolated traffic class. The solution satisfying individual flows time and volume constraints can be found in polynomial time and expressed as a set of multi-interval bandwidth allocation profiles. To ensure a large scale deployment of this approach, we propose, for the data plane, a combination of a bandwidth profile enforcement mechanism with traditional transport protocols. We have examined several solutions for implementing such a mechanism in Linux kernel. The experimental evaluation shows that packet pacing performed at IP level offers a simple yet valuable and TCP-compatible solution for accurate bandwidth profile enforcement at very high speed.

6.2.2. Dynamic bandwidth provisioning

Keywords: *bandwidth provisioning, flow aggregation, network service, optical networks, optimization.*

Participants: Pascale Vicat-Blanc Primet, Sébastien Soudan, Pierre Bozonnet.

To address the anticipated sporadic terabyte demands generated by high-end time-constrained applications, dynamically reconfigurable optical networks services are envisioned. However, the time and rate granularities of a bandwidth reservation service and those of transfer tasks using the reserved capacity are not necessarily in the same order of magnitude. This may lead to poor resource utilisation and over-provisioning. We explore how request aggregation is able to limit this problem. Interactions between a bandwidth reservation service and a data mover service guaranteeing time-constrained data transfer are investigated. We formulate the underlying optimisation problem and propose an optimal strategy for bandwidth provisioning when malleable transfer requests are known in advance. Simulations show that the temporal parameters of requests (deadline and patience) are the dominant criteria and that a small malleability can improve performance a lot."

6.2.3. Flow-aware networks

Keywords: *QoS, cross-protect, flow analysis, flow scheduling, flow-aware, quality of service, sampling.*

Participants: Pascale Vicat-Blanc Primet, Dinil Mon Divakaran, Paulo Gonçalves, Pierre-Solen Guichard, Isabelle Guérin Lassous.

This work is conducted in the context of INRIA Bell Labs and in close collaboration with MAESTRO team (Eitan Altman). Flows crossing IP networks are not equally sensitive to loss or delay variations because they do not have the same utility functions and the same final usage. Since several years, research effort has been spent to solve the problem of the heterogeneous performance needs of the IP traffic. A class of solutions considers that the IP layer should provide more sophisticated services than the simple best-effort service to meet the application's quality of service requirements. Quality of service has been studied in IP networks in the context of multimedia applications. Re-thinking the fundamental paradigm of packet switching network in high speed networks is on the table. The idea is to go from a packet-level approach to a flow-oriented strategy. To cope with the scalability issues, we work on disruptive algorithms within equipments, and on fully distributed (or localized) solutions. Problems that need to be explored concern flow identification and classification (see also next research direction), flow admission control, flow routing, flow scheduling, interaction with transport protocols.

The problem of traffic identification and classification has received considerable attention from the research community. Our interest here is on how to build an efficient global knowledge plane that can be used for taking decisions on traffic identifications locally. Traffic can be classified at application level, trying to identify the specific application associated with the traffic. For better flexibility, the behaviour of traffic can be used to classification. In this way, the classification itself can be independent of any new application type.

Besides looking at traffic at a coarse level, it is also useful in analysing traffic at a finer level. Interesting decisions can be taken based on flow characteristics. Various important flow characteristics are size, age and rate. Decisions can be based on any one of these characteristics, or multiples of the same.

Since performing per-packet measurement for per-flow analysis is computationally challenging, there is growing interest in obtaining useful information on flow characteristics using sampling. Sampling reduces the processing required to obtain flow statistics.

One of the actions that can be taken based on flow characteristics is scheduling. Tremendous amount of work has been done in the area of scheduling of jobs, and of late, many researchers have applied this in the context of networking, to schedule flows.

We incorporate the idea of sampling to schedule flows so as to induce less processing overhead. We propose a simple and practical scheduling strategy, as well as analyse the mean response time of flows when the classification is accurate, and when the classification is performed based on sampled information.

An original, recent field of research in that domain is using flow-aware networking to achieve quality of service guarantees at flow level, which is the relevant granularity level for more and more users, applications and services like video and audio streaming or image guided surgery over long distances. By performing implicit differentiation between types of traffic and providing best quality of service for all admitted flows even in overload situations, Cross-protect is promising. In this work, we have partly evaluated this architecture, and then we have proposed a further evaluation and improvements concerning the implementation and failure tolerance, like adaptive routing for example.

6.2.4. Integrating very large packets in networks

Keywords: *jumbo frames, queueing delay analysis.*

Participants: Pascale Vicat-Blanc Primet, Dinil Mon Divakaran.

This work is conducted in the context of INRIA Bell Labs and in close collaboration with MAESTRO team (Eitan Altman). Looking into the future, this work (part of the INRIA Bell Labs research) addresses the need for larger packet size, called XLFrame (XLF), for an Internet which is soon to witness stupendous amounts of traffic that have to be processed and switched at amplifying line rates. Increasing the size of the basic transporting unit in the Internet has far-reaching incentives that otherwise appear hard to achieve. For a variety of reasons, we foresee a future Internet that has both packets (sand) and XLFs (rocks). As a first step, we analyse the effects of introducing XLFs in a network, and find the following: the amount of packet-header processing is greatly reduced, while the fair multiplexing of XLFs with standard packets can be achieved using a more careful queue management in routers.

We also look into how we can make improvements through incremental research. In this direction, studying the effect of having large packets (of size \gg current MTU) in the current network is important as well as useful. Such packets are called *XLFrames* (or XLFs in short). Some of the motivating reasons for having XLFs in a network are: (1) to reduce power consumption at equipments by reducing the processing required, (2) achieving maximum throughput with increasing line rates, and (3) reducing per-packet cost involved in protocol processing and interrupt handling at the end-hosts.

In this work, we find that, though XLFs greatly reduces per-packet cost, flows using XLFs throttle packet-switched flows. Besides, XLF-switched flows experience higher loss rates. A solution to the unfairness comes in the form of Deficit Round Robin (DRR) scheduling that can be deployed at an equipment. DRR combined with ECN is seen to reduce the loss rates considerably.

6.2.5. Fair Rate Allocation for Multihop Wireless Networks

Keywords: *distributed allocation, fairness.*

Participants: Rémi Vannier, Isabelle Guérin Lassous.

IEEE 802.11 is often considered as the underlying wireless technology of multihop wireless networks. But the use of 802.11 in such networks raises issues, like efficiency and/or fairness issues. Different kinds of solutions have been proposed to overcome these problems. One approach is to design new MAC protocols that provide alternatives to the IEEE 802.11 MAC protocol. Although these solutions are of some interest, it should probably take some time before new wireless network interface cards based on one of these solutions are developed and released. Another approach is to consider that 802.11 will remain the underlying wireless technology and to design solutions above it. Several solutions based on rate allocation have been proposed so far. The main drawback of the proposed solutions is that they rely on a radio medium sharing model that is difficult to compute in a wireless, distributed and mobile environment. Indeed, very few of these solutions have been derived into a network protocol.

In this work, we propose a distributed and dynamic rate allocation solution that is based on a simple sharing model. Due to its simplicity, we can derive a network protocol that can be practically used in multihop wireless networks. This protocol provides a fair bandwidth sharing between end-to-end flows while maintaining an efficient overall throughput in the network. This solution has been implemented in NS2 and evaluated by simulations.

6.2.6. Fairness, Simplicity and Efficiency in Multihop Wireless Networks

Keywords: *MAC protocols, complexity, efficiency, fairness.*

Participant: Isabelle Guérin Lassous.

The performance of ad hoc networks based on IEEE 802.11 DCF degrade when congestion increases. The issues concern efficiency and fairness. Many solutions can be found at the MAC layer in the literature, but very few solutions improve fairness and efficiency at the same time. In this work, we design a new backoff solution, called SBA. SBA uses only local information and two contention window sizes. By simulations, we compare SBA with IEEE 802.11 and several alternatives to 802.11 in ad hoc networks. We show that SBA achieves a good trade-off between fairness, simplicity and efficiency.

6.2.7. Available Bandwidth Estimation in Multihop Wireless Networks

Keywords: *bandwidth estimation, end-to-end guarantees.*

Participant: Isabelle Guérin Lassous.

Since 2005, IEEE 802.11-based networks are able to provide a certain level of Quality of Service by the means of service differentiation, thanks to the IEEE 802.11e amendment. However, no mechanism or method has been standardized to accurately evaluate the amount of resources remaining on a given channel. Such an evaluation would however be a good asset for bandwidth-constrained applications.

In multi-hop ad hoc networks, such evaluation becomes even more difficult. Consequently, despite the various contributions around this research topic, the estimation of the available bandwidth still represents one of the main issues in this field. In this work, we propose an improved mechanism to estimate the available bandwidth in IEEE 802.11-based ad hoc networks. Through simulations, we compare the accuracy of the estimation we propose to the estimation performed by others state of the art QoS protocols, BRuIT, AAC and QoS-AODV.

6.2.8. Congestion Collapse in Grid'5000

Keywords: *congestion collapse, high speed networks large scale experimentation.*

Participants: Pascale Vicat-Blanc Primet, Romaric Guillier.

Today, the capacity and the usage of the Internet is fundamentally changing. In the forthcoming year, millions homes will have access to the Internet with fiber lines. Network companies will offer speeds on fiber of up to 1 Gbps. Consequently ultra-high-speed applications will be enabled for low-cost, such as high-definition teleconferencing, telemedicine and advanced telecommuting for people working from home. We have designed, deployed and demonstrated a scenario which aims at showing that the traditional congestion control approach may present barrier to the large scale deployment of the envisioned high-end applications.

6.2.9. Towards a User-Oriented Benchmark for Transport Protocols Comparison in very High Speed Networks

Keywords: *High Speed networks, High Speed transport, Performance evaluation, Protocol Benchmark, TCP.*

Participants: Pascale Vicat-Blanc Primet, Romaric Guillier, Ludovic Hablot.

Standard TCP faces performance limitations in very high speed wide area networks, mainly due to a long end-to-end feedback loop and a conservative behaviour with respect to congestion. Many TCP variants have been proposed to overcome these limitations. However, TCP is a complex protocol with many user-configurable parameters and a range of different implementations. It is then important to define measurement methods so that the transport services and protocols can evolve guided by scientific principles and can be compared quantitatively. Users of these variants need performance parameters that describe protocol capabilities so that they can develop and tune their applications. The goal of this work is to make some steps towards a user-oriented test suite and a benchmark, called HSTTS, for high speed transport protocols comparison. We first identified useful metrics. We then isolated infrastructure parameters and traffic factors which influence the protocol behaviour. This enabled us to define classes of representative applications and scenarios capturing and synthesising comprehensive and useful properties. We finally evaluate this proposal on the Grid'5000 experimental environment, and present it to the IRTF TRMG working group.

6.2.10. A language for virtual resources and interconnection networks description

Keywords: *VXDL, optical networks, resource virtualization, virtual infrastructure.*

Participants: Pascale Vicat-Blanc Primet, Guilherme Koslovski, Marcelo Pazin.

VXDL was developed to help users, applications or middleware in the virtual components specification, model and representation. Basically, this language enables the description of virtual infrastructures which are composed by I) virtual resources, II) virtual network topology and III) virtual time. Using these three sets of features it is possible represents a virtual infrastructure composition (describing resources individually and in groups) detailing the network topology desirable (through links configuration and virtual routers) informing the execution timeline of each set of resources and links. Each component (resource or group) can have different parameters, allowing the configuration of size, software, hardware, location and functionality. In addition, VXDL is able to interact with some specific configurations for virtual infrastructures, as the definition of the virtual machines numbers that can be allocated in a physical resources; basically location (anchor) of a resource; and virtual routers usage. VXDL is defined using both BNF notation and XML standard, allowing its utilization in frameworks (or middleware) for management virtual environments. In this context, different systems can use VXDL for exchange information about the virtual infrastructures. The key concepts of VXDL were presented as a masters thesis at *Universidade Federal de Santa Maria (UFSM), Brazil*, and were published in *GridNets Conference*, both in 2008.

Benchmarks execution on virtual infrastructure: We investigate the benefit obtained with a detailed virtual infrastructure description, more precisely the network topology parameters provisioned by VXDL. For this, we selected some applications of the NAS benchmark for a execution in different virtual infrastructures allocated on the test environment Grid5000. The key point identified was the importance of a satisfactory network description mainly for applications sensitive to latency and bandwidth configurations. In addition, in this study we used virtual routers to help in the format definition of the virtual infrastructure. The results of this step were used in evaluation section of both publications cited above.

Preliminary study of embed virtual infrastructures on distributed physical resources: This preliminary study investigate the representation of virtual and physical infrastructures as graphs, and the process of its co-allocation, sharing the physical capabilities. In addition our approach I) investigate the possibility of provisioning capacities (as resources and network), which can be requested on-demand for applications, and II) the virtual timeline specification which define an internal scheduler for each virtual infrastructure. The embedding problem is a NP-hard problem, which requires special heuristics that lead to an optimal result approximation. Currently, we are making a survey about related works and algorithms which was proposed, and studying a possible solution, considering the currently allocation of vertices and edges, in according with each capacities. Our first contribution in this context was the definition of a graph notation for represents both virtual and physical infrastructures, considering time attributes and subgraphs.

6.3. High Speed Network's traffic metrology and statistical analysis

6.3.1. *Experimental investigation on the relationship between heavy tailed distributions and long range dependence*

Keywords: *aggregated traffic, heavy-tailed flow size distributions, long range dependence.*

Participants: Patrick Loiseau, Paulo Gonçalves, Pascale Vicat-Blanc Primet.

After the seminal work by Taqqu et al. relating self-similarity to heavy-tailed distributions, a number of research articles verified that aggregated Internet traffic time series show self-similarity and that Internet attributes, like Web file sizes and flow lengths, were heavy-tailed. However, the validation of the theoretical prediction relating self-similarity and heavy tails remains unsatisfactorily addressed, being investigated either using numerical or network simulations, or from uncontrolled Web traffic data. Notably, this prediction has never been conclusively verified on real networks using controlled and stationary scenarii, prescribing specific heavy-tailed distributions, and estimating confidence intervals.

With this goal in mind, in [53], [66], we use the potential and facilities offered by the large-scale, deeply reconfigurable and fully controllable experimental Grid5000 instrument, to investigate the prediction observability on real networks. To this end we organize a large number of controlled traffic circulation sessions on a nationwide real network involving two hundred independent hosts. We use a FPGA-based MetroFlux measurement system [69], to collect the corresponding traffic at packet level. We then estimate both the self-similarity exponent of the aggregated time series and the heavy-tail index of flow size distributions, independently. On the one hand, our results complement and validate with a striking accuracy some conclusions drawn from a series of pioneer studies. On the other hand, they bring in new insights on the controversial role of certain components of real networks.

Pursuing our investigation on LRD and on its impact on the quality of service, we tackle the question of self-similarity of TCP traffic over lossy links. Relying on Taquu's model and its experimental evidences, it is unanimously admitted that TCP cannot be responsible for this self-similarity observed at large time scales. On the opposite side, we experimentally show in [70] that when the file size is heavy-tailed, the TCP congestion control mechanism under sufficiently high loss can annihilate the self-similarity that would be otherwise observed if no loss occurs.

6.3.2. *Maximum likelihood estimate of heavy-tail exponents from sampled data*

Keywords: *flow size, heavy-tail distributions, maximum likelihood estimation.*

Participants: Patrick Loiseau, Paulo Gonçalves, Pascale Vicat-Blanc Primet.

This work, submitted to ACM Sigmetrics 2009

, is a joint collaboration with the MISTIS team project.

We address the problem of estimating the flow size distribution corresponding to the traffic passing through an aggregated link. More precisely, we want to measure the tail exponent (hypothesizing that distributions are systematically heavy-tailed) from a sub-sampled measured series of packets stream. Considering a Pareto distribution as our theoretical a priori model, we formally derived the maximum likelihood estimate of the Pareto tail exponent α . Previously we proposed in [85] a heuristic estimate of α (assuming the same Pareto a priori), which turned out to be conceptually very close to the maximum likelihood estimate, providing us with an intuitive interpretation of this latter. Based on simulated data, we performed a systematic comparison of our estimator with different approaches proposed in the literature (e.g. stochastic counting, EM estimates, wavelet based estimates,...). Then, not only the proposed method significantly improves bias and variance estimates, but it also works with small sampling rates (reasonably up to 1/100), drastically contrasting with most rival estimators. To assess the performance of our method, we present a full performance evaluation based on numerical simulations, and also on a real traffic trace corresponding to internet traffic recently acquired.

6.3.3. *A nonlinear model for satellite images time series: analysis and prediction of land cover dynamics*

Keywords: *classification, land cover changes, land cover types, maximum likelihood estimate, non-linear time series model, prediction.*

Participants: Paulo Gonçalves, Hugo Carrão.

Albeit grounded on different physical origins, it is not rare that distinct real-world problems share common mechanisms and/or formulations. This similitude naturally fosters the development of unified frameworks which can then match a wide range of applications. Moreover, as statistical signal processing frequently stands at the junction of several scientific domains, it is not surprising that statistical studies go beyond the scope of the application areas they were initially addressing. That is why the RESO team (principally through one of his members) was led to participate to inter-disciplinary collaborations that do not straightforwardly relate to the main themes of the project activities, but which can capitalize with them.

That is notably the case for this joint work with the Portuguese Institute of Geography, University of Lisbon (Portugal).

Numerous efforts have been made to develop models to fit spectral bands and vegetation indices time series from satellite images for diverse land cover classes. The purposes of these models are to suppress short-frequency time series variations, usually related with atmospheric disturbances, and derive phenological attributes to monitor natural and human induced ecosystems dynamics. In [59], we propose an approach for predicting some phenological attributes of land cover classes. The prediction approach we introduce is performed with a parametric nonlinear harmonic model developed to fit intra-annual spectral bands and vegetation indices (VI) time series from satellite images. To the best of our knowledge, the model we use has the unique ability to fit intra-annual land cover classes time series with indiscriminate waveforms. The model is parsimonious, easily identifiable, robust to noise, versatile and can be estimated from few time observations. To evaluate the adequacy of our approach, we identified and predicted rice crops and shrubland phenological attributes from an intra-annual time series of weekly MODIS images acquired at 500 m nominal spatial resolution in Portugal. Results show that phenological attributes can be perfectly identified from model fits of complete intra-annual vegetation indices time series. Moreover, we show that phenological attributes are early predicted with high accuracy from few intra-annual time series observations. Within the same collaborative framework, in [59] we proposed a machine learning based classification of land cover types, featured on multi-temporal satellite images.

6.3.4. EMD-based analysis of rat EEG data for sleep state classification

Keywords: *empirical mode decomposition, heart beat rate variability, supervised classification.*

Participants: Paulo Gonçalves, Süleyman Baykut.

This is a collaborative work with University of Istanbul (Turkey), with the “Physiopathologie des Réseaux Neuronaux du Cycle Veille-Sommeil” lab of University Claude Bernard Lyon, and with the Physics lab of ENS Lyon.

Empirical Mode Decomposition (EMD)-based features from single-channel electroencephalographic (EEG) data are proposed for rats sleep state classification. The classification performances of the EMD-based features and some classical power spectrum density (PSD)-based features are compared. Supported by experiments on real EEG data, we demonstrate that classification performances can significantly improve, by simply substituting EMD to PSD in features extraction. This is in noticeably due to the natural adaptivity of EMD which show more robust to subjects variability.

6.4. Network services for high demanding applications

6.4.1. Design and development of an MPI gateway

Keywords: *Grid, Grid5000, MPI, heterogeneity, high-speed interconnects, relays.*

Participants: Ludovic Hablot, Olivier Glück, Jean-Christophe Mignot, Pascale Vicat-Blanc Primet.

The MPI standard is often used in parallel applications for communication needs. Most of them are designed for homogeneous clusters but MPI implementations for grids have to take into account heterogeneity and long distance network links in order to maintain a high performance level. These two constraints are not considered together in existing MPI implementations and raise the question of MPI efficiency in grids. Our goal is to significantly improve the performance execution of MPI applications on the grid.

We have done a state of the art, a performance evaluation, understanding and tuning of four recent MPI implementations for the Grid : MPICH-Madeleine, GridMPI, OpenMPI and MPICH2. The comparison is based on the executions of pingpong, NAS Parallel Benchmarks and a real application of geophysics. These experiments take place on the national GRID’5000 testbed. We show that a tuning of both TCP protocol and MPI implementation are necessary to obtain good performances on the grid. We study the impact on application time execution of a long-way latency between two groups of 8 MPI tasks for each NAS parallel benchmark. Our experiments and tunings presented in [42] lead to the conclusion that GridMPI performs better results than the others and that executing MPI applications on a grid can be beneficial if some specific parameters are well tuned.

Based on these results, we propose a new transparent layer called MPI5000 and placed between MPI and TCP allowing application composed of several tasks to be correctly distributed on available node regarding the grid topology and the application scheme. Thus, our layer needs two data files: a file describing the grid topology including available nodes, both latency and bandwidth between the nodes and between sites; another file describing the application communication patterns with the size and the amount of messages sent between MPI processes. Using these two data files, our layer should realise an efficient placement of tasks on grid nodes.

Our layer also propose to transparently split TCP connections between MPI processes in order to take into account the grid topology. This new architecture is based on a system of relays placed at the LAN/WAN interface. We replace each end-to-end TCP connection by three connections (two on the LAN between a node and a relay, one on the WAN between two relays). Thus, we expect a faster lost recovery on LAN as well as a reduction of memory used because the size of TCP buffers depends on RTT latency of the connection. On relays, we planned to use different TCP implementations or different protocols for local and distant communications. The relays could also implement a different scheduling strategy of MPI messages : for instance, we could give priority to small messages (usually MPI control messages). Finally, as MPI applications are mostly using small messages, they are more penalised if the network is congestionned by large flows. We planned to reserve bandwidth on the WAN in order to optimise MPI communications on the long distance shared links.

The implementation of our proposition is based both on a library between MPI and the operating system and on relays. Thus, the proposed architecture is independant of MPI implementations and is totally transparent for applications. For the moment, relays are implemented and the library is in a final phase. We are now finishing the developement of our layer. Then, we will deploy and test our toolkit in a grid context. Finally, we will implement the optimisations proposed previously.

6.4.2. Virtualizing and scheduling optical network resource for emerging IT services

Keywords: *bandwidth scheduling, network resource virtualization, optical networks.*

Participants: Pascale Vicat-Blanc Primet, Sébastien Soudan, Pierre Bozonnet, Manoj Dahal, Fabienne Anhalt.

Emerging IT services providers which aim to deliver supercomputing power available to the masses over the Internet, rely on high performance IT resources interconnected with ultra-high performance optical networks. To adjust their resource provisioning to demand variations, services for right-time access to capacities are required. These services are not supported by current telecom network infrastructures. We propose a flexible and open framework to implement the "virtual infrastructure as a service" concept to offer advanced IT service providers a dynamic access to extensible virtual private capacities, through on-demand and in-advance bandwidth and resource reservation services. We propose the virtual infrastructure service adapts to envisioned heterogeneous needs by multiplexing rigid and flexible requests as well as coarse or fine demands in order to optimize both resource provisioning and utility functions. This concept is implemented in the CARRIOCAS project within the SRV component which handles requests, aggregates them and provisions the resources accordingly. Considering the special case of the optical resource and dynamic bandwidth allocation, we formulate the underlying optimisation problem, study the impact of the malleability factor and evaluate, by simulation, the gain obtained.

6.4.3. Automatic Network Services aligned with Grid Application requirements in CARRIOCAS

Keywords: *VXDL, optical networks, resource virtualization, virtual infrastructure.*

Participants: Pascale Vicat-Blanc Primet, Sébastien Soudan, Pierre Bozonnet, Marcelo Pazin, Manoj Dahal, Guilherme Koslovski.

Automatic Network Services aligned with Grid Application requirements in CARRIOCAS Project, Automatic Service framework named Scheduling, Reconfiguration and Virtualization (SRV) is developed in CARRIOCAS project to enhance existing Telecom network infrastructures for supporting grid applications sharing IT resources interconnected with ultra-high performance optical networks. From the requirements of Grid applications a classification is proposed to specify the network services and their attributes. In large-scale collaborative environments, SRV solution is described to enable automatic network service operations according to high-performance computing service access. The resources hosted at datacenters are virtualized to be attached to transport network infrastructure offering uniform interfaces towards external customers. New level of service bindings is defined with network services during executions of Grid applications workflows. On-demand intensive computing and visualization services scenario is described in Telecom environment.

6.4.4. Development of a metrology platform on Grid5000

Keywords: *Gtrc-Net1, header extraction, metrology, monitoring, packet capture.*

Participants: Patrick Loiseau, Damien Ancelin, Aurélien Cedeyn, Matthieu Imbert, Romaric Guillier, Paulo Gonçalves, Pascale Vicat-Blanc Primet.

This activity is partially supported by the program GridNets-FJ (*Équipe associée*) between INRIA and AIST (Japan).

Researches in network traffic analysis embrace a large diversity of goals and are based on a variety of methodologies and tools. To have a better insight on the real nature and on the evolution of network traffic we argue that fine-grain analysis of real traffic traces have to complement simulations studies as well as coarse grain measurement performed by classical flow measurement systems. In particular, packet level measurements and analysis are needed. However, such methodologies are resource consuming and require very high performance devices to be operational in real high speed networks. In [69] we present the *Metroflux* system which aims at providing researchers and network operators with a very flexible and accurate packet-level traffic analysis toolkit configured for 1 Gbps and 10 Gbps speed links. This system is based on the GtrcNet FPGA-based device technology and on specific statistical analysis tools. We show the potential and the facilities offered by the *Metroflux* system coupled with the *Grid5000* large scale experimental platform and the Network eXperiment Engine (*NXE*) we have developed. In [69] we illustrate the application of *Metroflux* with the practical validation of the theoretical prediction relating self-similarity and heavy tails given by Taqqu theorem. We also illustrate several usages of this toolset, such as the investigation of conditions under which several traffic theories apply, as well as studies on traffic, protocols and systems interactions.

6.4.5. Programmable network services for context aware adaptation

Keywords: *execution environments, programmable networks.*

Participants: Laurent Lefèvre, Jean-Patrick Gelas.

Traditional industrial maintenance process (i.e. requiring regularly a human intervention on the exploitation area) are coming to their limits. Indeed, more and more industrial equipments are connected to communication networks. This allows us to consider optimised maintenance solutions. In addition to primary existing sensors (which only give some numeric values), we can now think about the use of multimedia sensors (video cameras, microphone, ...). Inside a cooperative industrial maintenance project (TEMIC project [92]) in which we are currently involved, our team designed equipments easily deployable in an industrial context, and also easily removable at the end of the maintenance contract.

The heterogeneity interms of networks, terminals and applications requires adaptive solutions for an efficient streams transmission on the platform networks. To respond to these various constraints, active services have to adapt and optimize the content of streams passing through the active network node. Multimedia data streams adaptation is performed dynamically in order to improve industrial maintenance solutions. The challenge is to provide an architecture running in a client/server environment, but involving no modification on the applications installed on the end-machines like web servers, video players,... For the Temic project, our team has worked on the design and adaptation of an industrial autonomic network node, which is derived from the Tamanoir environment. This Industrial Autonomic Network Node is designed to be deployed on

limited resources based network boxes, and so to be integrated into industrial platforms. We developed and tested active adaptation network services, specially written for the Tamanoir^{embedded}. Active services applying on multimedia streams crossing the network node may realize data compression, format transcoding, frame resizing... This kind of adaptation contributes to the saving of network bandwidth (by decreasing the output data rate) and to the reduction of the resources used on the client terminal playing the multimedia data (by reducing the framerate and the frame size). The adaptation is thereby transparent for the applications.

We base our developments and experimentations on mainly two industrial maintenance scenarios TEMIC project team to be used by a company through a maintenance contract on a restricted industrial area.

At this time, three active services have been developed for this project. They are designed to adapt multimedia data on the fly.

Our experiments show that our solution is efficient in reducing the amount of data transmitted on the network, and so the bandwidth consumed by the application, and also in reducing the CPU and resources needed on the client machine to decode the streams. However, our experiments clearly show some limitations in the performances of our industrial network node. These low performances impact directly the display quality on the user's device. We have now to improve our hardware equipment to obtain better performances.

7. Contracts and Grants with Industry

7.1. INRIA actions

7.1.1. GRID5000: ADT Aladdin

Participants: Olivier Glück, Magi Sanchon, Sébastien Soudan, Romaric Guillier, Ludovic Hablot, Laurent Lefèvre, Pascale Vicat-Blanc Primet, Paulo Gonçalves, Patrick Loiseau, Jean-Christophe Mignot, Aurélien Cedeyn.

ENS Lyon is involved in the GRID'5000 project, which is an experimental Grid platform gathering nine sites geographically distributed in France. ENS Lyon hardware contribution is done for now by two distinct set of computers. The Grid5000 of Lyon comprises now around 300 processors interconnected with a network of 500Mb/s Ethernet bisection and a 2Gb/s Myrinet interconnection for 64 nodes.

RESO is strongly involved in the choices of Grid5000's network components and architecture. Pascale Vicat-Blanc Primet is member of the national committee (comité de pilotage) of GRID'5000, of the Aladdin scientific committee, co-responsible of the Lyon site with Frederic Desprez, and coordinates networks aspects with Renater and RMU, Lyon's metropolitan network. Lyon site is nationally recognized to gather the "networking expertise" with skilled researchers and engineers and dedicated networking equipments (Metroflux, GNET10...). Working for the interconnection of the Grid5000 project at the international level, we are hosting the Japanese Naregi project remote hosts and are accessing to dedicated equipments within the Naregi testbed. We also participate to the ALADDIN ADT. Oana Gona, funded by the Aladdin ADT, is designing and developing an open traffic measurement and analysis infrastructure for the Grid5000 testbed. Aurélien Cedeyn is member of the national technical committee of GRID'5000. Year funding: 60 K euros

7.1.2. INRIA ARC GREEN-NET

Participants: Laurent Lefèvre, Jean-Patrick Gelas, Anne-Cécile Orgerie.

The GREEN-NET is a Cooperative Research Action (ARC : Action de Recherche Cooperative) supported by INRIA. This project explores the design of energy-aware software frameworks dedicated to large scale distributed systems. These frameworks will collect energy usage information and provide them to resources managers and schedulers. Large scale experimental validations on Grid5000 and DSSLAB platforms will be proposed. Laurent Lefèvre is leading the INRIA ARC GREEN-NET on "Power aware software frameworks for high performance data transport and computing in large scale distributed systems" which involved 4 partners : INRIA RESO, INRIA MESCAL (Grenoble), IRIT (Toulouse), Virginia Tech (USA). Official ARC GREEN-NET webpage : <http://www.ens-lyon.fr/LIP/RESO/Projects/GREEN-NET>

7.2. INRIA Bell Labs common laboratory: Semantic Networking

Keywords: *Flow analysis, Flow scheduling, Flow-based routing, Sampling, Traffic awareness.*

Participants: Pascale Vicat-Blanc Primet, Isabelle Guerin-Lassous, Paulo Gonçalves, Olivier Grémillet, Dinil Mon Divakaran, Pierre-Solen Guichard.

During this year we conducted the following researches:

- State-of-the-art on the different aspects covered within the Semantic Networking Aspects. Particular focus on the X-protect approach of France Telecom R&D.
- Study of the impact of large and small flows in current networks and analyse on how to handle both in Semantic Networks
- Proposals of new ideas through INRIA/Alcatel-Lucent discussions that will lead to patents, in elephant flow monitoring and scheduling/control.
- Global Semantic Networking architecture and high-level view of Semantic node defined.
- Development of the 10Gb/s packet capture system. Trace of 10Gb/s traffic on a real production network captured. The fine-grain analysis of these data is ongoing.

Year funding: 120Keuros

7.3. ANAGRAN: flow-router evaluation in Grid5000

Keywords: *Flow-based router evaluation.*

Participants: Pascale Vicat-Blanc Primet, Adrien Kiren, Romaric Guillier, Matthieu Imbert.

We have designed and ran experimentations of the ANAGRAN FR router within Grid5000.

7.4. France Telecom R&D

Participants: Laurent Lefèvre, Pascale Primet.

Since 2005 RESO has a collaboration with France Telecom R&D (Lannion) on “Network load balancing on layer 7 switching for high performance and high available Linux based platforms”. A CIFRE grant has supported this collaboration. Ayari Narjess has done her PhD on this topic from June 2005 to october 2008 [78], [79], [80], [77]. A patent has been deposed on "Multi-flow sessions management solutions" in 2007 [76]

7.5. CARRIOCAS

Participants: Pascale Vicat-Blanc Primet, Pierre Bozonnet, Marcelo Pasin, Manoj Dahal, Romaric Guillier, Guilherme Koslovski.

Carriocas project studies and implements an ultra high bit rate (up to 40 Gbps per wavelength) network interconnecting super computers, storage servers and high resolution visualization devices to support data and computing intensive applications in industrial and scientific domains. The R&D activities cover high bit rate transmission systems, advanced networking intelligence, and high performance distributed applications. CARRIOCAS is a three year project started in October 2006 which aims to be an experimental step of the transition from local to external storage and computing systems. This transition is valuable to share the cost of powerful systems among several users, to provide scalable and resilient architecture through distributed resource and to enable virtual collaborative working environments between different actors working on a same project. The following points are especially investigated:

- Supporting the high bandwidth requirements through the migration of networks from 10 gbp/s to 40 Gbps/s per wavelength in a cost effective way.
- Building architectural, protocol and algorithmic solutions able to provide to the network the agility to dynamically adapt to the application needs with a high level of automation and optimisation, while taking into account the administrative and business constraints.
- Developing and demonstrating on a network testbed distributed applications bringing performance enhancements for concrete scientific and industrial needs.
- Investigating the definition and the associated business models of high added value services integrating computing, visualization, storage and network resources.

In this project, RESO is in charge of the design and prototyping of the "Resource Scheduling Reconfiguration and Virtualisation - SRV" component. Year funding : 100Keuros

8. Other Grants and Activities

8.1. National actions

8.1.1. ANR IGTMD

Participants: Pascale Vicat-Blanc Primet, Chen Cheng, Romaric Guillier.

The aim of this project (2006-2008) is to design, develop and validate mechanisms that concretely make the interoperability of heterogeneous grids a reality. The project concentrates on the following topics: a) Bulk data transfers, b) replication and referring mechanisms, c) information system and job management interoperability, d) grid control and monitoring, e) usage of statistics and accounting data. A particular emphasis will be put on disk to disk bulk data transfers over very long distance with optimal performance. The key idea is to fully exploit the specificity of LCG applications (Computing Grid Project to find the grid middleware developed for the future Large Hydron Collider in CERN) and their real infrastructures to analyse and experiment new communication and replication models, alternative transport protocols emerging within the international scientific community. The participation to a standardization process for a generic grid transport service for bulk exchanges between heterogeneous grids will be a strong goal of the project. Despite the fact that the interoperability and the unification of a generic data transport in Grids are very often perceived as a necessity, they are in fact very little studied. The present project would allow France to get a leading position in this computing area that will be absolutely crucial to insure the Large Hydron Collider (LHC) data exploitation. The very strong experience of the partners in deployment and exploitation of international research and production computing instruments gives a promising perspective to this project and its ambitious experimental approach. In this project, RESO is responsible for all research activities concerning high speed transport protocols and services. The key idea is to fully exploit the specificity of LCG applications and their real infrastructures to analyse and experiment new communication and replication models, alternative transport protocols emerging within the international scientific community. We are also exploiting the LCG traffic circulating on the IGTMD link for packet capture and grid flow analysis Year funding: 70Keuros

8.1.2. ANR DSLLAB

Participants: Laurent Lefèvre, Pascale Vicat-Blanc Primet, Jean-Patrick Gelas.

RESO is partner of the DSLlab research project (2006-2008) which aims at building and using an experimental platform about distributed systems running on DSL Internet. The objective is twofold:

- to provide with accurate and customized measures of availability, activity and performances in order to characterize and tune the models of the ADSL resources;
- to provide with a validation and experimental tool for new protocols, services and simulators and emulators for these systems.

DSLlab consists of a set of low power, low noise computers spread over the ADSL. These computers are used simultaneously as active probes to capture the behavior traces, and as operational nodes to launch experiments. We expect from this experiment a better knowledge of the behavior of the ADSL and the design of accurate models for emulation and simulation of these systems which represents now a significant capability in terms of storage and computing power. The DSLLAB platform will be deployed in 2007.

In this project, RESO is responsible for the definition, design and development of flow control algorithms and mechanisms, enabling distributed computing applications to fully exploit the DLS links.

8.1.3. ANR HIPCAL

Participants: Pascale Vicat-Blanc Primet, Jean-Patrick Gelas, Olivier Mornard, Fabienne Anhalt, Guilherme Koslovski, Philippe Martinez, Lucas Nussbaum.

HIPerCAL studies a new paradigm (grid substrate) based on confined virtual private execution infrastructure for resource control in grids. In particular, we propose to study and implement new approaches for bandwidth sharing and end to end network quality of service guarantees. The global infrastructure (computers, disks, networks) is partitioned in virtual infrastructures (aggregation of virtual machines coupled with virtual channels) dynamically composed. These virtual infrastructures are multiplexed in time and space, isolated and protected. The goal of this project is to explore an approach in a break with current services-oriented principles developed in grids to jointly enhance the application portability, the communications performance control and their security. The project aims at providing a grid substrate based on end to end bandwidth reservation, control overlay, network and system virtualization, cryptographic identification principles. The proposal is to be validated and evaluated at different scales on the Grid5000 testbed with biomedical applications, demanding in security, performance and reliability. 10 to 1000 processors, links with 100Mb/s to 10Gb/s, few microseconds to 100ms will be involved in these experimentations. We aim at demonstrating the functional transparency, enhanced predictability and efficiency for applications offered by the HIPerNet approach. RESO has developed, deployed and tested the first version of the HIPerNet software on the Grid5000 testbed.

Year funding: 100Keuros

8.1.4. ANR DMASC

Participants: Paulo Gonçalves, Patrick Loiseau.

Started in October 2008, this ANR project, led by J. Barral (Sisyphé, INRIA Rocquencourt), is a partnership between INRIA (Sisyphé and Reso), university Paris 12 and university Paris Sud (équipe d'accueil EA 4046 Service de Réanimation Médicale CHU de Bicêtre).

Numerical studies using ideas from statistical physics, large deviations theory and functions analysis have exhibited striking scaling invariance properties for human long-term R-R interval signals. These signals are extracted from electrocardiograms and represent the time intervals between two consecutive heartbeats. The scaling invariance measured on these empirical data are reminiscent of geometric fractal properties verified theoretically by certain mathematical objects (measures or functions), which are called (self-similar) multifractals. These numerical studies also reveal that the scaling invariance may have different forms, according to the fact that the patients have a good health or suffer from certain cardiac diseases. These observations suggest that a good understanding of multifractal properties of cardiac signals might lead to new pertinent tools for diagnosis and surveillance. However, until now, neither satisfactory physiological origin has been associated with these properties nor mathematical objects have been proposed as good models for these signals. It is fundamental for possible medical applications in the future to go beyond the previously mentioned works and achieve a deepened study of the scaling invariance structure of cardiac signals. This requires new robust algorithms for the multifractal signals processing; specifically, it seems relevant to complete the usual statistical approach with a geometric study of the scaling invariance. In addition, it is necessary to apply these tools to a number of data arising from distinct pathologies, in order to start a classification of the different features of the observed scaling invariance, and to relate them to physiological concepts. This should contribute to develop an accurate new flexible multifractal mathematical model whose parameters could be adjusted according to the observed pathology. It is also important to strengthen the information by performing the multifractal analysis of another fundamental signal in cardiology, namely the blood pressure, as well as the simultaneous multifractal analysis/modeling of the couple (R-R, Blood Pressure). This project aims at achieving such a program. It also proposes to contribute to explain the origin of the scaling invariance properties by developing a reduced order dynamical system, which shall describe the heart's electromechanical activity and simultaneously shall generate multifractal outputs in accordance with the R-R signals models. A 1-D model of cardiac fiber would be already very satisfactory. This aspect of the project is closely related to the delicate issue of understanding the link between multifractal phenomena and PDEs, another topic that will be investigated. The project team consists in six members representing four partners: two specialists of multifractal analysis, one specialist of cardio-vascular system modeling and PDEs control, one specialist of statistical signal processing and two physiologists (among which one cardiologist) specialists of cardio-vascular signals processing. The project will benefit of a wide data's bank of long term (24h) R-R interval signals already recorded in various clinical settings including diabetes, acromegaly and sleep apnea, and

a prospective data bank will be established in the field of medical intensive care unit, namely in patients presenting cardiovascular pathologies like heart failure, arterial hypertension and chock states. The data bank will include both R-R interval signals and arterial blood pressure signals.

Year funding: 2,5Keuros

8.2. European actions

8.2.1. AUTONOMIC INTERNET

Participants: Laurent Lefèvre, Jean-Patrick Gelas, Abderhaman Cheniour.

Autonomic Internet (AutoI - FP7.ICT.2007.Call1-216404) project suggests a transition from a service agnostic Internet to service-aware network, managing resources by applying autonomic principles. In order to achieve the objective of service-aware resources and to overcome the ossification of the current Internet AutoI will develop a self-managing virtual resource overlay that can span across heterogeneous networks that can support service mobility, security, quality of service and reliability. In this overlay network, multiple virtual networks co-exist on top of a shared substrate with uniform control. The overlay will be self-managed based on the system's business goals, which drive the service specifications, the subsequent changes in these goals (service context) and changes in the resource environment (resource context). This will be realised by the successful co-operation of the following activities: autonomic control principles, resource virtualisation, enhanced control algorithms, information modelling, policy based management and programmability. RESO is mainly involved in the programmability of the AUTOI overlay by proposing an Autonomic Network Programming Interface which will support large scale service deployment. Laurent Lefèvre is leading the workpackage 5 on "Service Deployment". Official webpage : <http://ist-autoi.eu>

8.2.2. OGF-EUROPE

Participants: Laurent Lefèvre, Pascale Vicat-Blanc Primet.

RESO is mainly involved in the interactions with the OGF Grid network community.

8.2.3. AEOLUS

Participants: Isabelle Guérin-Lassous, Rémi Vanier.

AEOLUS (Algorithmic Principles for Building Efficient Overlay Computers) is an IP project that has been started since September, 1st, 2005. The university of Patras (Greece) is the prime contractor. The goal of this project is to investigate the principles and develop the algorithmic methods for building an overlay computer that enables an efficient and transparent access to the resources of an Internet-based global computer. In particular, the main objectives of this project are:

- To identify and study the important fundamental problems and investigate the corresponding algorithmic principles related to overlay computers running on global computers.
- To identify the important functionalities such an overlay computer should provide as tools to the programmer, and to develop, rigorously analyze and experimentally validate algorithmic methods that can make these functionalities efficient, scalable, fault-tolerant, and transparent to heterogeneity.
- To provide improved methods for communication and computing among wireless and possibly mobile nodes so that they can transparently become part of larger Internet-based overlay computers.
- To implement a set of functionalities, integrate them under a common software platform in order to provide the basic primitives of an overlay computer, as well as build sample services on this overlay computer, thus providing a proof-of-concept for our theoretical results.

8.2.4. EC-GIN

Participants: Pascale Vicat-Blanc Primet, Paulo Gonçalves, Patrick Loiseau, Damien Ancelin, Sébastien Soudan, Romaric Guiller, Ludovic Hablot, Marcelo Pasin.

EC-GIN (Europe-China Grid InterNetworking) is an European STREP project started in November 1st 2006. The university of Innsbruck (Austria) is the prime contractor.

The Internet communication infrastructure (the TCP/IP protocol stack) is designed for broad use; as such, it does not take the specific characteristics of Grid applications into account. This one-size-fits-all approach works for a number of application domains, however, it is far from being optimal - general network mechanisms, while useful for the Grid, cannot be as efficient as customised solutions. While the Grid is slowly emerging, its network infrastructure is still in its infancy. Thus, based on a number of properties that make Grids unique from the network perspective, the project EC-GIN will develop tailored network technology in dedicated support of Grid applications. These technical solutions will be supplemented with a secure and incentive-based Grid Services network traffic management system, which will balance the conflicting performance demand and the economic use of resources in the network and within the Grid.

By collaboration between European and Chinese partners, EC-GIN parallels previous efforts for real-time multimedia transmission across the Internet: much like the Grid, these applications have special network requirements and show a special behaviour from the network perspective. However, while research into network support for multimedia applications has flourished, leading to a large number of standard protocols and mechanisms, the research community has neglected network support for Grid computing up to now. By filling this gap and appropriately exploiting / disseminating the project results, EC-GIN will, therefore, cause a "snowball effect" in the European and Chinese networking and Grid computing research communities.

Technically, EC-GIN will make the Grid work, operate, and communicate better. By appropriately utilising the underlying network, Grid resources in general will be used more efficiently and amplify the impact of Grid computing on the society and economy of Europe and China. Year funding: 100Keuros

8.3. International actions

8.3.1. *NEGST: JSPT-CNRS*

Participants: Olivier Glück, Magi Sanchon, Sébastien Soudan, Romaric Guillier, Ludovic Hablot, Laurent Lefèvre, Pascale Vicat-Blanc Primet, Paulo Gonçalves, Patrick Loiseau, Jean-Christophe Mignot.

The objective of this project is to promote the collaborations of Japan and France on grid computing technology. In order to promote the collaborative researches, we consider that this project is organized for the following three parts:

1. Grid interoperability and applications
2. Grid Metrics
3. Instant Grid and virtualization of grid computing resources.

RESO mainly participates to the Grid Metrics topic.

Despite the development of strong technologies in all these domains, many issues are still open about the measurement methodology itself, the emulation or simulation of Grid platforms and the understanding of Grid software stack, application performance, and fault tolerance. The Grid Metrics topics, basically gathers all researches about applications, programming models, libraries, runtimes, operating systems and network evaluation, either in synthetic environment (emulators and simulators) or real environment (real network and Grids).

8.3.2. *AIST Grid Technology Research Center: GridNet-FJ associated team*

Participants: Pascale Vicat-Blanc Primet, Olivier Gluck, Ludovic Hablot, Sébastien Soudan, Romaric Guillier, Olivier Gluck, Paulo Gonçalves, Patrick Loiseau.

Since 2007, RESO is pursuing its collaboration with AIST through the Gridnet-FJ associated team. We followed and even increased our working program on four parts: 1) High speed transport protocol over very high speed links, 2) Bandwidth allocation and control in Grids, 3) Optimisation of MPI communications in Grids, 4) Co-design of GtrcNET-packet capture functionality.

On point 1) with the high speed testbed for protocol evaluation we have deployed within Grid5000 and which integrates GtrcNET1 and GtrcNET10, we pursued our work on TCP variants comparison. We highlight the problem of congestion level which makes TCP behave very strangely (long TCP stops) and the problem of congesting reverse traffic. During our visit to AIST, we had long discussions on the TCP stop problem. This issue has been now solved and a patch to LINUX TCP stack has been posted; We also work together on the INRIA HSTTS (High Speed Transport Test Suite) and gathered very interesting and constructive remarks from AIST colleagues.

The collaboration between AIST GTRC team and INRIA RESO team on the point 2) aims at studying how BDTS, a scheduled data transfer service could benefit this flexibility offered by advance provisioning of some network path and to develop a service which use the interface.

AIST GTRC is collaborating with Pr Ishikawa team at University of Tokyo on GridMPI implementation (point 3). The AIST develop GridMPI, an MPI implementation designed for grids, that uses a similar system of relays as we do in our MPI5000 layer. This architecture has two goals: transmission of messages from a private cluster to a public cluster and also allowing the transfer from a Myrinet cluster to TCP on the long distance link. Their implementation is able to manage many relays to forward datas on long distance links. Their objective is different but the involved mechanisms are similar to the MPI5000 layer. The AIST planned to test their relays on Grid5000 to have a real network platform instead of emulating latency as they are doing for the moment. These experiments are realised jointly with the RESO team.

During the stay of Yuestu Kodama and Tomohiro Kudoh at ENS, the INRIA RESO team and the AIST GTRC team design and develop the Metroflux system for 10Gb/s speed and deploy 10 of such equipment within Grid5000/ALADDIN.

8.3.3. Collaboration with University of Otago, New Zealand

Participant: Laurent Lefèvre.

Laurent Lefèvre has been hosted as invited researcher in Otago University (Dunedin, New Zealand) from July to September 2007. He has worked in the team of Prof. Zhiyi Huang on advanced network solutions for media streaming[28], [27].

8.3.4. Collaboration with University of Lisbon, Portugal

Participant: Paulo Gonçalves.

P. Gonçalves is co-advising the PhD program of Hugo Carrão from ISEGI, University of Lisbon. A grant from the "Programme Actions Universitaires Intégrées Luso-Françaises" supports our collaboration (ends in january 2009). H. Carrão spent a 6 months visit at RESO (july - december 2008), and his PhD defense is expected in the spring 2009.

8.3.5. Collaboration with University of Istanbul, Turkey

Participant: Paulo Gonçalves.

P. Gonçalves oriented the 1st year program of Süleyman Baykut PhD work, from the Department of Electronics & Communications Engineering, University of Istanbul. S. Baykut was granted a TUBITAK fellowship (Scientific and Technological Research Council of Turkey) which funded his long-term stay of 10 months with the RESO team.

8.4. Visitors

8.4.1. Collaboration with University of Sevilla, Spain

Participants: Laurent Lefèvre, Anne-Cécile Orgerie.

RESO has hosted two PhD students from University of Sevilla for short term periods during the year 2008 : Pablo Neira Ayuso (June-July 2008) to work on fault tolerant issues in stateful firewalls and Alejandro Fernandez (November 2008) to work on the prediction models for energy efficiency in large scale distributed systems.

8.4.2. Collaboration with Vrije University, Netherlands

Participant: Pascale Vicat-Blanc.

RESO has hosted one PhD student from Vrije University for short term period during the year 2008 : Andrei Agapi (June-July 2008) to work on flow scheduling within the Internet.

8.4.3. Collaboration with AIST GTRC, Japan

Participants: Tomohiro Kudoh, Yuetsu Kodama, Pascale Vicat-Blanc.

RESO has hosted Dr Tomohiro Kudoh for 2 weeks and Dr Yuetsu Kodama for 1 week as invited researchers to work on the design of a packet capture feature within GtrcNet10 equipment integrated in the Grid5000 cluster. We also worked together on the OGF Network Service Interface use case document.

9. Dissemination

9.1. Conference organisation, editors for special issues

- Pascale Vicat-Blanc is
 - member of steering committee of PFLDNET international conference serie, ACM Grid-Nets international conference,
 - Program co-Chair of the ACM International Conference on High Speed Networks for Grid Applications (GridNets2008) in Beijing.
 - Program co-Chair of the International Conference Broadnets2009, Future Internet track..
 - General co-chair of the international workshop on Protocols for Very Long Distance Networks PFLDNET2008 in Manchester.
 - member of program committees : CCGrid2008, CFIP2008, BROADNET2008, CC-Grid2009, ITC21 (2009), ITCSS 2009,
 - Guest editor of the special issue of Future Generation Computer System, the journal of the Grid on "High Speed Networks for Grid Applications", to be edited beginning of 2009.
- Paulo Gonçalves
 - is member of the program committee of ASSESS workshop in conjunction with CCGrid 2008;
 - is co-editor of "Scaling, Fractals and Wavelets", John Wiley Ed.
- Laurent Lefèvre is:
 - Tutorial and workshop chair of PDCAT'08 : The Ninth International Conference on Parallel and Distributed Computing, Applications and Technologies, Dunedin, New Zealand, December 2008
 - General chair of CCGrid 2008 : 8th IEEE/ACM International Symposium on Cluster Computing and the Grid, Lyon, France, May 2008
 - Steering Committee member of IEEE International Symposium on Cluster Computing and the Grid (CCGrid conference) series since 2004

- member of the following Program Committees: International Workshop on Advances in High-Performance E-Science Middleware and Applications (AHEMA 2008), held in conjunction with the IEEE eScience Conference 2008, 6th International Workshop on Middleware for Grid Computing - MGC 2008, in conjunction with ACM/IFIP/USENIX 9th International Middleware Conference 2008, Euro PVMMPI 2008 : 15th European PVM/MPI Conference, International Symposium on Pervasive Grid (PGrid08), HPDC 2008 : IEEE International Symposium on High Performance Distributed Computing (HPDC-17), ICCS 2008 : International Conference on Computational Science 2008, ICA3PP-2008 : The 8th International Conference on Algorithms and Architectures for Parallel Processing, HPPAC2008: The Fourth Workshop on High-Performance, Power-Aware Computing, HotP2P'08 : Fifth International Workshop on Hot Topics in Peer-to-Peer Systems, 6th Australasian Symposium on Grid Computing and e-Research
- Isabelle Guérin Lassous is:
 - in the Editorial Team of the international electronic journal Discrete Mathematics & Theoretical Computer Science;
 - in the Editorial Board of the international journal Computer Communications, Elsevier;
 - Guest co-Editor in Elsevier's Performance Evaluation journal for the special issue on 'Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks';
 - Program co-Chair of ACM PE-WASUN 2008 (Fifth ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, October 2008, Vancouver, Canada);
 - a member of the following program committees: OPODIS 2008 (Luxor, Egypt), VTC-Fall 2008 (Calgary, Canada), Med-Hoc-Net 2008 (Palma de Mallorca, Spain), IFIP Networking 2008 (Singapore), PerSeNS 2008 (Hong Kong), ICDCN 2008 (Kolkata, India).
- Olivier Glück is:
 - a member of the following Program Committees : RenPar'18, CCGrid 2008, GridNets 2008, HPDC 2008.
 - the local arrangement co-chair of the 8th IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2008), May 19-22, 2008, Lyon, France.

9.2. Graduate teaching

- **since 2006** I. Guérin Lassous
Multimédia and Quality of Service Master 2 SIR (Professional) / RTS (Research) (University Claude Bernard Lyon I), lecture 18h, others 12h.
- **since 2006** I. Guérin Lassous
Networking
Master 2 CCI (Professional)(University Claude Bernard Lyon I), lecture 18h, others 12h.
- **since 2006** I. Guérin Lassous
Autonomic Computing
Master 2 RTS (Research) University Claude Bernard Lyon I), lecture 15h.
- **since 2006** I. Guérin Lassous
Wireless Networks
Master 2 Professional (IFI, Hanoi, Vietnam), lecture 40h, others 20h.
- **since 2004** O.Glück
Client/Server Model, Internet Applications, Network and System Administration.
Master 2 SIR (University Claude Bernard Lyon 1), lecture 30h, others 30h.

- **since 2004** JP.Gelas
High-speed networks, QoS and Multimedia ; Initiation to Java ; Local Area Networks .
Master 2 SIR and CCI (University Claude Bernard Lyon 1), lecture 30h, others 40h.
- **since 2005** JP.Gelas
Long distance networks ; Networks and Transport Protocols ; Routing ; Advanced Java and Web services.
Master 2 SIR (Université Claude Bernard Lyon 1), lecture 45h, others 45h.
- **since 2007** JP.Gelas
Embedded System and Software.
Master 2 (Université Claude Bernard Lyon 1), lecture 15h, others 15h.
- **since 2008** JP.Gelas
Introduction to System, Computer Networks and Client/Server architecture.
Master 2 (Université Claude Bernard Lyon 1), lecture 20h, others 30h.

9.3. Miscellaneous teaching

- **since 2006:** I. Guérin Lassous
Ad Hoc Networks Master 1 (University Claude Bernard Lyon I), lecture 6h, others 6h.
- **since 2004:** O. Glück
Computer Networks.
Licence Informatique, (University Claude Bernard Lyon 1), lecture 30h, others 30h.
- L. Lefèvre
High performance Networks.
Maitrise Informatique (Ho Chi Minh Ville University, Vietnam), 30h eq TD.
- L. Lefèvre
is responsible of training periods for Research Master in ENS-Lyon
- **since 2007:** JP. Gelas
Long Distance and High Performance Network.
Graduate students of the "Institut de la Francophonie pour l'Informatique" in Hanoi, Vietnam, 60h lectures/year.

9.4. Animation of the scientific community

- Pascale Vicat-Blanc
 - the INRIA scientific leader of ADR Semantic Networking of the INRIA Bell Labs common laboratory
 - is member of the "Networks" expert committee of the CNRS.
 - is within the Grid5000 project and ADT ALADDIN, member of the steering committee and co-leader of the Grid5000@Lyon site.
 - is leading the ANR CIS HIPCAL project.
 - is leading the INRIA team within the CARRIOCAS System@tic project.
 - is leading the INRIA team within the european EC-GIN project.
 - is leading the LIP team of the ANR (blanc) IGTMD project.
- Isabelle Guérin Lassous is:
 - member of the CNRS TAROT action (Techniques Algorithmiques, Réseaux et d'Optimisation pour les Télécommunications);

- the INRIA scientific leader of the european project AEOLUS (Algorithmic Principles for Building Efficient Overlay Computers);
- member of the new INRIA - Bell Labs common research laboratory (INRIA scientific leader of the WP4 - Mechanisms for QoS control and management of flows in the Semantic Networking research activity).
- Laurent Lefèvre is :
 - leading the INRIA ARC GREEN-NET on “Power aware software frameworks for high performance data transport and computing in large scale distributed systems” which involved 4 partners : INRIA RESO, INRIA MESCAL (Grenoble), IRIT (Toulouse), Virginia Tech (USA).
 - leading the WorkPackage 5 on “Service Deployment” of the FP7 STREP Project “Autonomic Internet”
 - the INRIA representative of the OGF-Europe project.

9.5. Participation in boards of examiners and committees

- Pascale Vicat-Blanc : president of the hearing committee of INRIA Rhône-Alpes;
- Isabelle Guérin Lassous is member of:
 - CNU, section 27;
 - the specialists committee (section 27) of the UJF (Université Joseph Fourier) - Grenoble;
 - two HdR examining board: Filali (Eurecom - reviewer), Peyrard (Toulouse - reviewer);
 - seven PhD examining boards: Karbaschi (Paris 6 - reviewer), Samper (INPG - reviewer), Watteyne (INSA Lyon - examiner), Ben Abdesslem (Paris 6 - reviewer), Busnel (Rennes 1 - examiner), Toham (INPG - reviewer), Pigné (Havre - president);
 - selection committee of the Ecole Nationale d’Administration (ENA).
- Olivier Glück is a member of
 - the “commissions de spécialistes 27ème section” of University Claude Bernard Lyon 1 and University Pierre et Marie Curie Paris 6.
 - the “conseil de l’UFR d’Informatique” of University Claude Bernard Lyon 1.
 - the “Conseil des Etudes et de la Vie Universitaire” of University Claude Bernard Lyon 1.
- Laurent Lefèvre
 - is member of the “commissions de spécialistes de 27ème section” of Ecole Normale Supérieure (Lyon), University Antilles Guyane (Pointe à Pitre) and University Lumière (Lyon2);
 - has been reviewer of the PhD thesis of :
 - * Benjamin Quetier : “EMUGRID : etude des mecanismes de virtualisation pour l’emulation conforme des grilles a grande echelle”, University Paris XI, France, September 2008 - Examinateur
 - * Frank Chiang : “Self-Adaptability, Resilience and Vulnerability in Autonomic Communications with Biology-inspired Strategies”, University of Technology, Sydney, Australia, July 2008 - Reviewer

9.6. Seminars, invited talks

- Pascale Vicat-Blanc was invited to give talks at:

- at CCGrid2008 *Presentation of the CARRIOCAS project* - July 2008
- at CARRIOCAS-PHOSPHORUS workshop meeting - July 2008
- at AIST - *Dynamic bandwidth provisioning in the CARRIOCAS project* july 2008
- at Tokyo Technical University (TiTech) -*Flow scheduling and network virtualization* july 2008
- at Osaka University - *Presentation of the Naregi-Grid5000 project* july 2008
- at the ITU-T System and Network Operation meeting at Beijing -*Network virtualization* October 2008
- at Optical Networks workshop of GridNets2008- *Presentation of the CARRIOCAS project* Beijing - October 2008.
- at GridNets2008 - *Network virtualization,:: perspectives in Grids* Beijing - October 2008.
- Invited talk at INRIA Visiting committee - "A new way of sharing Network" - december 2008.
- Isabelle Guérin Lassous gave:
 - an invited talk on "resources allocation for multihop wireless networks" at IFI, Hanoi, Vietnam, october 2008.
- Laurent Lefèvre has been invited to give the following talks :
 - “Energy Efficiency Challenges for Large Scale Distributed Systems”, Laurent Lefèvre and Anne Cécile Orgerie, Deakin University, Australia, December 15, 2008
 - “Towards Energy Aware Resource Infrastructure for Large Scale Distributed Systems”, Laurent Lefèvre, CCGSC 2008 : Clusters and Computational Grids for Scientific Computing Workshop, Flat Rock, North Carolina, USA, September 14-17, 2008
 - “GREEN-* : Towards Energy Efficient Solutions for Next Generation Large Scale Distributed Systems”, Laurent Lefèvre, ACOMP 2008 : International Workshop on Advanced Computing and Applications, Ho Chi Minh City, Vietnam, March 12-14, 2008
 - “Proposing inter-operable router-assisted transport protocol for transferring large volume of data in high performance Grids”, Laurent Lefèvre, Seminar, University of Sevilla, Spain, March 6, 2008
 - “Towards new services and capabilities for next generation Grids”, Laurent Lefèvre, Seminar, University of Sevilla, Spain, March 5, 2008
- Jean-Patrick Gelas has been invited to :
 - give a talk titled "Towards energy aware resource infrastructure for large scale distributed systems (GREEN-*)" to the members and students of the "Institut de la Francophonie pour l'Informatique" in Hanoi, Vietnam, November 20th, 2008.

10. Bibliography

Major publications by the team in recent years

- [1] F. BOUHAFS, J.-P. GELAS, L. LEFÈVRE, M. MAIMOUR, C. PHAM, P. VICAT-BLANC PRIMET, B. TOURANCHEAU. *Designing and Evaluating An Active Grid Architecture*, in "The International Journal of Future Generation Computer Systems (FGCS) - Grid Computing: Theory, Methods and Applications", vol. 21, n^o 2, February 2005, p. 315-330.

- [2] F. CAPPELLO, F. DESPREZ, M. DAYDE, E. JEANNOT, Y. JEGOU, S. LANTERI, N. MELAB, R. NAMYST, P. VICAT-BLANC PRIMET, O. RICHARD, E. CARON, J. LEDUC, G. MORNET. *Grid5000: a nation wide experimental grid testbed*, in "International Journal on High Performance Computing Applications", 2006.
- [3] B. B. CHEN, P. VICAT-BLANC PRIMET. *Supporting bulk data transfers of high-end applications with guaranteed completion time*, in "IEEE ICC2007 International Conference on Computer Communication", IEEE (editor), 2007.
- [4] B. GOGLIN, O. GLÜCK, P. VICAT-BLANC PRIMET. *An Efficient Network API for in-Kernel Applications in Clusters*, in "Proceedings of the IEEE International Conference on Cluster Computing, Boston, Massachusetts", IEEE Computer Society Press, September 2005.
- [5] P. GONÇALVES, R. RIEDI. *Diverging moments and parameter estimation*, in "Journal of American Statistical Association", vol. 100, n^o 472, December 2005, p. 1382–1393.
- [6] J. LAGANIER, P. VICAT-BLANC PRIMET. *HIPernet: a decentralized security infrastructure for large scale grid environments*, in "6th IEEE/ACM International Conference on Grid Computing (GRID 2005), November 13-14, 2005, Seattle, Washington, USA, Proceedings", IEEE, 2005, p. 140-147.
- [7] L. LEFÈVRE, J.-P. GELAS. *Chapter 14 on "High Performance Execution Environments"*, in "Programmable Networks for IP Service Deployment", A. GALIS, S. DENAZIS, C. BROU, C. KLEIN (editors), Artech House Books, UK, may 2004, p. 291-321.
- [8] D. LOPEZ PACHECO, C. PHAM, L. LEFÈVRE. *XCP-i : eXplicit Control Protocol for heterogeneous inter-networking of high-speed networks*, in "Globecom 2006, San Francisco, California, USA", November 2006.

Year Publications

Doctoral Dissertations and Habilitation Theses

- [9] N. AYARI. *Contributions to Session Aware Frameworks for Next Generation Internet Services*, 153 pages, Ph. D. Thesis, École normale supérieure de Lyon, 46, allée d'Italie, 69364 Lyon cedex 07, France, October 2008.
- [10] D. LOPEZ PACHECO. *Propositions for a robust and inter-operable eXplicit Control Protocol on heterogeneous high speed networks*, 137 pages, Ph. D. Thesis, École normale supérieure de Lyon, 46, allée d'Italie, 69364 Lyon cedex 07, France, June 2008.

Articles in International Peer-Reviewed Journal

- [11] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *Fault Tolerance for Highly Available Internet Services: Concepts, Approaches, and Issues*, in "IEEE Communications Surveys and Tutorials", vol. 10, n^o 2, July 2008, p. 34-46.
- [12] H. CARRÃO, A. ARAÚJO, P. GONÇALVES, M. CAETANO. *Multitemporal MERIS images for land cover mapping at national scale: the case study of Portugal*, in "International Journal of Remote Sensing", To appear, 2008.
- [13] H. CARRÃO, P. GONÇALVES, M. CAETANO. *Contribution of multispectral and multitemporal information from MODIS images to land cover classification*, in "Elsevier, Remote Sensing of Environment", vol. 112, n^o 3, 2008, p. 986-997.

- [14] B. GOGLIN, O. GLÜCK, P. VICAT-BLANC PRIMET. *Interaction efficace entre les réseaux rapides et le stockage distribué dans les grappes de calcul*, in "Technique et Science Informatiques", vol. 27, n^o 7/2008, September 2008, p. 911–940, <http://www.ens-lyon.fr/LIP/Pub/Rapports/RR/RR2006/RR2006-04.pdf>.
- [15] L. LEFÈVRE, J.-P. GELAS. *IAN2 : Industrial Autonomic Network Node architecture for supporting personalized network services in the industrial context*, in "The International Journal of Future Generation Computer Systems (FGCS) - Grid Computing: Theory, Methods and Applications", vol. 24, n^o 1, January 2008, p. 58-65.
- [16] R. RAZAFINDRALAMBO, I. GUÉRIN LASSOUS, L. IANNONE, S. FDIDA. *Dynamic Packet Aggregation to Solve Performance Anomaly in 802.11 Wireless Networks*, in "Computer Networks", vol. 52, n^o 1, January 2008.
- [17] R. RAZAFINDRALAMBO, I. GUÉRIN LASSOUS. *Increasing Fairness and Efficiency using the MadMac Protocol in Ad Hoc Networks*, in "Ad Hoc Networks", vol. 6, n^o 3, May 2008.
- [18] S. SARR, C. CHAUDET, G. CHELIUS, I. GUÉRIN LASSOUS. *Bandwidth Estimation for IEEE 802.11-based Ad Hoc Networks*, in "IEEE Transactions on Mobile Computing", vol. 7, n^o 10, 2008.
- [19] S. SOUDAN, B. B. CHEN, P. VICAT-BLANC PRIMET. *Flow Scheduling and endpoint rate control in GridNetworks*, in "Future Generation Computer Systems", To appear., 2008.

International Peer-Reviewed Conference/Proceedings

- [20] L. ANDREW, C. MARCONDES, S. FLOYD, L. DUNN, R. GUILLIER, W. GANG, L. EGGERT, S. HA, I. RHEE. *Towards a Common TCP Evaluation Suite*, in "PFLDnet 2008", March 2008.
- [21] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *A Session Aware Admission Control Scheme for next Generation IP Services*, in "Fifth annual IEEE Consumer Communications and Networking Conference (IEEE CCNC 2008), Las Vegas, USA", January 2008.
- [22] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *On improving the Reliability of Cluster based Voice over IP Services*, in "FastAbstract : DSN 2008 : The 38th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, Anchorage, Alaska, USA", June 2008.
- [23] N. AYARI, L. LEFÈVRE, D. BARBARON. *On improving the Reliability of Internet Services through Active Replication*, in "PDCAT 2008 : The Ninth International Conference on Parallel and Distributed Computing, Applications and Technologies, Dunedin, New Zealand", December 2008, p. 259-262.
- [24] N. AYARI, P. NEIRA AYUSO, L. LEFÈVRE, D. BARBARON. *Towards a Dependable Architecture for Highly Available Internet Services*, in "ARES'08 : The Third International Conference on Availability, Reliability and Security, Barcelona, Spain", March 2008, p. 422-427.
- [25] L. CHANG, A. GALIS, B. MATHIEU, K. JEAN, R. OCAMPO, L. MAMATAS, J. RUBIO LOYOLA, J. SERRAT, A. BERL, H. DE MEER, S. DAVY, Z. MOVAHEDI, L. LEFÈVRE. *Self-organising Management Overlays for Future Internet Services*, in "MACE 2008 : 3rd IEEE International Workshop on Modelling Autonomic Communications Environments during the ManWeek 2008 : 4th International Week on Management of Networks and Services, Samos Island, Greece", September 2008, p. 74-89.

-
- [26] R. GUILLIER, P. VICAT-BLANC PRIMET. *Methodologies and Tools for Exploring Transport Protocols in the Context of High-Speed Networks*, in "IEEE TCSC Doctoral Symposium", May 2008.
- [27] S. HASAN, L. LEFÈVRE, Z. HUANG, P. WERSTEIN. *Cross Layer Protocol Support for Live Streaming Media*, in "AINA-08 : The IEEE 22nd International Conference on Advanced Information Networking and Applications, Okinawa, Japan", March 2008, p. 319-326.
- [28] S. HASAN, L. LEFÈVRE, Z. HUANG, P. WERSTEIN. *Supporting Large Scale eResearch Infrastructures with Adapted Live Streaming Capabilities*, in "6th Australasian Symposium on Grid Computing and e-Research, Wollongong, Australia", vol. 82, Conferences in Research and Practice in Information Technology - Australian Computer Society, January 2008, p. 55-63.
- [29] G. KOSLOVSKI, P. VICAT-BLANC PRIMET, A. S. CHARÃO. *VXDL: Virtual Resources and Interconnection Networks Description Language*, in "GridNets 2008", Oct. 2008.
- [30] D. LOPEZ PACHECO, L. LEFÈVRE, C. PHAM. *Lightweight Fairness Solutions for XCP and TCP Cohabitation*, in "IFIP/TC6 Networking 2008, Singapore", May 2008, p. 715-726.
- [31] P. NEIRA AYUSO, L. LEFÈVRE, R. M. GASCA. *hFT-FW : Hybrid Fault-Tolerance for Cluste-based Stateful Firewalls*, in "ICPADS 2008 : The 14th IEEE International Conference on Parallel and Distributed Systems, Melbourne, Australia", December 2008.
- [32] P. NEIRA AYUSO, R. M. GASCA, L. LEFÈVRE. *FT-FW: Efficient Connection Failover in Cluster-based Stateful Firewall*, in "PDP2008 : 16th Euromicro International Conference on Parallel, Distributed and network-based Processing, Toulouse, France", February 2008, p. 573-580.
- [33] P. NEIRA AYUSO, R. M. GASCA, L. LEFÈVRE. *Multiprimary support for the Availability of Cluster-based Stateful Firewalls using FT-FW*, in "ESORICS 2008 : 13th European Symposium on Research in Computer Security, Malaga, Spain", October 2008, p. 1-17.
- [34] P. NEIRA AYUSO, L. MACCARI, L. LEFÈVRE, R. M. GASCA. *Stateful Firewalling for Wireless Mesh Networks*, in "NTMS 2008 : The second IFIP International Conference on New Technologies, Mobility and Security, Tangier, Morocco", November 2008.
- [35] A.-C. ORGERIE, L. LEFÈVRE, J.-P. GELAS. *Chasing Gaps between Bursts : Towards Energy Efficient Large Scale Experimental Grids*, in "PDCAT 2008 : The Ninth International Conference on Parallel and Distributed Computing, Applications and Technologies, Dunedin, New Zealand", December 2008, p. 381-389.
- [36] A.-C. ORGERIE, L. LEFÈVRE, J.-P. GELAS. *Save Watts in your Grid: Green Strategies for Energy-Aware Framework in Large Scale Distributed Systems*, in "ICPADS 2008 : The 14th IEEE International Conference on Parallel and Distributed Systems, Melbourne, Australia", December 2008.
- [37] R. RAZAFINDRALAMBO, I. GUÉRIN LASSOUS. *SBA: un algorithme simple de backoff pour les réseaux ad hoc*, in "Algotel, St Malo, France", May 2008.
- [38] R. VANNIER, I. GUÉRIN LASSOUS. *Partage équitable de la bande passante dans les réseaux ad hoc*, in "Conférence Francophone sur l'Ingénierie des Protocoles, Les Arcs, France", Mar 2008, <http://hal.archives-ouvertes.fr/hal-00250058/fr/>.

- [39] R. VANNIER, I. GUÉRIN LASSOUS. *Towards a Practical and Fair Rate Allocation for Multihop Wireless Networks based on a Simple Node Model*, in "11th ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM), Vancouver, Canada", October 2008.
- [40] D. VERCHERE, O. AUDOUIN, B. BERDE, A. CHIOSI, R. DOUVILLE, H. POUYLLAU, P. VICAT-BLANC PRIMET, M. PASIN, S. SOUDAN, T. MARCOT, V. PIPERAUD, R. THEILLAUD, D. HONG, D. BARTH, C. CADÉRE, V. REINHART, J. TOMASIK. *Automatic Network Services aligned with Grid Application requirements in CARRIOCAS Project*, in "GridNets 2008", Oct. 2008.

National Peer-Reviewed Conference/Proceedings

- [41] J.-P. GELAS, O. MORNARD, P. VICAT-BLANC PRIMET. *Évaluation des performances réseau dans le contexte de la virtualisation XEN*, in "CFIP 2008 : Colloque Francophone sur l'Ingénierie des Protocoles, Les Arcs, France", March 2008.
- [42] L. HABLOT, O. GLÜCK, J.-C. MIGNOT, P. VICAT-BLANC PRIMET. *Etude d'implémentations MPI dans une grille de calcul*, in "Actes de Renpar'08", Février 2008.

Scientific Books (or Scientific Book chapters)

- [43] L. BONONI, I. GUÉRIN LASSOUS (editors). *Fifth ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks (PE-WASUN)*, ACM, ACM, Vancouver, Canada, October 2008.
- [44] P. VICAT-BLANC PRIMET, J. MAMBRETTI, T. KUDOH. *Proceedings of the 2nd International ACM/ICST GRIDNETS conference on Networks for Grid Applications*, ACM, oct 2008.

Books or Proceedings Editing

- [45] Z. HUANG, Z. XU, N. ROUNTREE, L. LEFÈVRE, H. SHEN, J. HINE, Y. PAN (editors). *Proceedings of PDCAT 2008 : The Ninth International Conference on Parallel and Distributed Computing, Applications and Technologies*, IEEE Computer Society, Dunedin, New Zealand, December 2008.
- [46] T. PRIOL, L. LEFÈVRE, R. BUYYA (editors). *Proceedings of CCGrid2008 : Eight IEEE International Symposium on Cluster Computing and Grid*, IEEE, Lyon, France, May 2008.

Research Reports

- [47] M. ABID, A. BERL, Z. BOUDJEMIL, A. CHENIOUR, S. DAVY, S. DENAZIS, A. GALIS, J.-P. GELAS, C. FAHY, A. FISCHER, J. RUBIO LOYOLA, L. LEFÈVRE, D. MACEDO, L. MAMATAS, H. DE MEER, Z. MOVAHEDI, J. STRASSNER, H. ZIMMERMANN. *Autonomic Internet Initial Framework - Deliverable D6.1*, Technical report, INRIA, August 2008.
- [48] A. AGAPI, T. KIELMANN, M. PASIN, S. SOUDAN, P. VICAT-BLANC PRIMET. *Globally Optimal Bulk Data Transfers in Overlay Routing Networks*, Technical report, n^o TR-0157, Institute on Grid Systems, Tools and Environments, CoreGRID - Network of Excellence, August 2008, <http://www.coregrid.net/mambo/images/stories/TechnicalReports/tr-0157.pdf>.
- [49] F. ANHALT, P. VICAT-BLANC PRIMET. *Analysis and evaluation of a XEN based virtual router*, Technical report, n^o 6658, INRIA Rhône Alpes, Sep 2008, <http://hal.inria.fr/inria-00320620/fr/>.

- [50] A. CHENIOUR, J.-P. GELAS, C. FAHY, A. FISCHER, J. RUBIO LOYOLA, L. LEFÈVRE, D. MACEDO, L. MAMATAS. *Autonomic Internet Initial Service Infrastructure - Deliverable D5.1*, Technical report, INRIA, December 2008.
- [51] G. FEDAK, J.-P. GELAS, T. HERAULT, V. INIESTA, D. KONDO, L. LEFÈVRE, P. MALECOT, L. NUSSBAUM, A. REZMERITA, O. RICHARD. *DSL-Lab: a Platform to Experiment on Domestic Broadband Internet*, Deliverable Projet ANR DSLLAB, Technical report, ANR, December 2008.
- [52] J.-P. GELAS, O. MORNARD, P. VICAT-BLANC PRIMET. *Evaluation des performances reseau dans le contexte de la virtualisation XEN*, "Also available : LIP report RR2008-01", Technical Report, n^o 6409, INRIA, January 2008, <https://hal.inria.fr/inria-00203461>.
- [53] P. LOISEAU, P. GONÇALVES, G. DEWAELE, P. BORGNAT, P. ABRY, P. VICAT-BLANC PRIMET. *Investigating self-similarity and heavy-tailed distributions on a large scale experimental facility*, Technical report, n^o 6472, INRIA, March 2008, <http://hal.inria.fr/inria-00263634/fr/>.
- [54] R. VANNIER, I. GUÉRIN LASSOUS. *Towards a Practical and Fair Rate Allocation for Multihop Wireless Networks based on a Simple Node Model*, Research Report, n^o 6538, INRIA, 05 2008, <https://hal.inria.fr/inria-00280407>.
- [55] P. VICAT-BLANC PRIMET, J.-P. GELAS, O. MORNARD, G. KOSLOVSKI, V. ROCA, L. GIRAUD, R. PODESTA, V. INIESTA. *HIPerNet design document*, Deliverable #3: HIPCAL ANR-06-CIS-005, Deliverable, INRIA, July 2008.
- [56] P. VICAT-BLANC PRIMET, J.-P. GELAS, O. MORNARD, G. KOSLOVSKI, V. ROCA, L. GIRAUD, R. PODESTA, V. INIESTA. *HIPerNet design document*, "Deliverable #3 : HIPCAL ANR-06-CIS-005", Technical report, INRIA, July 2008.

Other Publications

- [57] F. ANHALT. *Analyse et évaluation d'un routeur virtuel basé sur XEN*, Masters thesis, INSA de Lyon, Jun 2008.
- [58] F. ANHALT. *Intégration de la ressource routeur virtuel au logiciel HIPerNET de composition et d'exploitation d'infrastructures virtuelles à la demande*, Masters thesis, INSA de Lyon, Sep 2008.
- [59] H. CARRÃO, P. GONÇALVES, M. CAETANO. *A nonlinear model for satellite images time series: analysis and prediction of land cover dynamics*, preprint, December 2008.
- [60] R. GUILLIER, P. VICAT-BLANC PRIMET. *Congestion Collapse in Grid5000*, demo, Stanford Congestion Collapse workshop, April 2008.
- [61] G. KOSLOVSKI, P. VICAT-BLANC PRIMET, A. S. CHARÃO. *VXDL: Uma Linguagem para Descrição de Interconexões e Recursos em Grades Virtuais*, Masters thesis, Programa de Pós-Graduação em Informática, Brasil, August 2008.
- [62] L. LEFÈVRE, J.-P. GELAS, A.-C. ORGERIE. *How an Experimental Grid is Used : the Grid5000 Case and its Impact on Energy Usage*, Poster CCGrid2008 : 8th IEEE International Symposium on Cluster Computing and the Grid, Lyon, France, May 2008.

- [63] L. LEFÈVRE, A.-C. ORGERIE. *EARI : Energy Aware Resource Infrastructure for Large Scale Distributed Systems*, Poster INRIA Booth, Supercomputing 2008, Austin, USA, November 2008.
- [64] L. LEFÈVRE, A.-C. ORGERIE. *GREEN-NET : Power aware software frameworks for high performance data transport and computing in large scale distributed systems*, Poster National ARC Days - INRIA Sophia Antipolis, France, October 2008.
- [65] P. LOISEAU, P. GONÇALVES, G. DEWAELE, P. BORGNAT, P. ABRY, P. VICAT-BLANC PRIMET. *Investigating self-similarity and heavy-tailed distributions on a large scale experimental facility*, preprint, submitted for publication in IEEE/ACM Transactions on Networking, May 2008.
- [66] P. LOISEAU, P. GONÇALVES, G. DEWAELE, P. BORGNAT, P. ABRY, P. VICAT-BLANC PRIMET. *Vérification du lien entre auto-similarité et distributions à queues lourdes sur un dispositif grande échelle*, 9 ième Atelier en Evaluation de Performances, Aussois, France, June 2008.
- [67] P. LOISEAU, P. GONÇALVES, S. GIRARD, F. FORBES, P. VICAT-BLANC PRIMET. *Maximum likelihood estimation of the flow size distribution tail index from sampled data*, preprint, submitted to ACM Sigmetrics 2009, Seattle, WA, USA, June 15-19, 2008.
- [68] P. LOISEAU, P. GONÇALVES, R. GUILLIER, M. IMBERT, Y. KODAMA, P. VICAT-BLANC PRIMET. *Metroflux: A high performance system for analyzing flow at very fine-grain*, preprint, accepted at TridentCom 2009, Washington DC, USA, April 6-8, 2008.
- [69] P. LOISEAU, P. GONÇALVES, Y. KODAMA, P. VICAT-BLANC PRIMET. *Metroflux: A fully operational high speed metrology platform*, Euro-NF workshop: New trends in modeling, quantitative methods and measurements, in cooperation with Net-coop, THOMSON Paris Research Labs, France, September 2008.
- [70] P. LOISEAU, P. GONÇALVES, P. VICAT-BLANC PRIMET. *How TCP can kill Self-Similarity*, Euro-NF workshop: Traffic Engineering and Dependability in the Network of the Future, VTT, Finland, September 2008.
- [71] A.-C. ORGERIE. *Energy-Aware Frameworks for High-Performance Data Transport and Computations in Large-Scale Distributed Systems*, Masters thesis, Ecole Normale Supérieure de Lyon, Jun 2008.
- [72] P. VICAT-BLANC PRIMET, J.-P. GELAS, O. MORNARD, G. KOSLOVSKI, V. ROCA, L. GIRAUD, J. MONTAGNAT, T. TRUONG HUU. *A scalable security model for enabling Dynamic Virtual Private Execution Infrastructures on the Internet*, Submitted for publication, 2008.
- [73] P. VICAT-BLANC PRIMET, J.-P. GELAS, O. MORNARD, G. KOSLOVSKI, V. ROCA, L. GIRAUD, J. MONTAGNAT, T. TRUONG HUU. *A scalable security model for enabling dynamic virtual private execution infrastructures on the Internet*, "submitted", December 2008.

References in notes

- [74] P. ABRY, R. BARANIUK, P. FLANDRIN, R. RIEDI, D. VEITCH. *Multiscale nature of network traffic*, in "IEEE Signal Processing Magazine", vol. 19, 2002, p. 28–46.
- [75] P. ABRY, P. FLANDRIN, D. VEITCH. *Internet : comment réguler le trafic ?*, in "La Recherche", n° 384, March 2005, p. 50–53.

- [76] N. AYARI, D. BARBARON, L. LEFÈVRE. *Procédés de gestion de sessions multi-flux. France Telecom R&D Patent*, June 2007.
- [77] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *Implementation of an Active Replication based Framework for Highly Available Services*, NetFilter Workshop 2007, Karlsruhe, Germany, September 2007.
- [78] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *SARA: A Session Aware Infrastructure for High Performance Next Generation Cluster-based Servers*, in "ATNAC 2007 : Australasian Telecommunication Networks and Applications Conference, Christchurch, New Zealand", December 2007.
- [79] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *Session Awareness issues for next-generation cluster-based network load balancing frameworks*, in "AICCSA07 : ACS/IEEE International Conference on Computer Systems and Applications, Amman, Jordan", May 2007, p. 180-186.
- [80] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *T2CP-AR: A system for Transparent TCP Active Replication*, in "AINA-07 : The IEEE 21st International Conference on Advanced Information Networking and Applications, Niagara Falls, Canada", May 2007, p. 648-655.
- [81] M. CHAUDIER, J.-P. GELAS, L. LEFÈVRE. *Towards the design of an autonomic network node*, in "IWAN2005 : Seventh Annual International Working Conference on Active and Programmable Networks, Nice, France", November 2005.
- [82] S. FLOYD, V. JACOBSON. *Link-sharing and Resource Management Models for Packet Networks*, in "IEEE/ACM Transaction on Networking", 4, vol. 3, August 1995.
- [83] I. FOSTER, C. KESSELMAN. *The Grid : Blueprint for a new Computing Infrastructure*, in "Morgan Kaufmann Publishers Inc.", 1998.
- [84] L. LEFÈVRE, J.-P. GELAS. *IAN2 : Industrial Autonomic Network Node architecture for supporting personalized network services in industrial context*, in "Submitted to The International Journal of Future Generation Computer Systems (FGCS) - Grid Computing: Theory, Methods and Applications", 2007.
- [85] P. LOISEAU, P. GONÇALVES, P. VICAT-BLANC PRIMET. *A comparative study of different heavy tail index estimators of the flow size from sampled data*, in "MetroGrid Workshop, GridNets, New York, USA", ACM Press, October 2007.
- [86] D. LOPEZ PACHECO, L. LEFÈVRE, C. PHAM. *XCP-i : eXplicit Control Protocol pour l'interconnexion de réseaux haut-débit hétérogènes*, Also available as LIP Research Report RR2007-47, Research Report, n° 6385, INRIA, December 2007, <https://hal.inria.fr/inria-00195634>.
- [87] K. PARK, W. WILLINGER. *Self similar traffic analysis and performance evaluation*, Wiley, 2000.
- [88] V. SANDER. *Networking issues of GRID Infrastructures*, in "GRID Working Draft of the GRID High-Performance Networking Research Group, Global GRID Forum", 2003.

-
- [89] V. SANDER, F. TRAVOSTINO, J. CROWCROFT, P. VICAT-BLANC PRIMET, C. PHAM. *Networking Issues of Grid Infrastructures*, Technical report, Open Grid Forum, october 2004, <http://forge.gridforum.org/projects/ghpn-rg/>.
- [90] D. SIMEONIDOU. *Optical Network Infrastructure for Grid*, in "Grid Working Draft of the Grid High-Performance Networking Research Group, Global GRID Forum", 2003.
- [91] M. TAQQU, W. WILLINGER, R. SHERMAN. *Proof of a fundamental result in self similar traffic modeling*, in "Computer Communication Review", vol. 27, 1997, p. 5–23.
- [92] H. TOBIET. *Panorama des réseaux utilisés et services à valeur ajoutée TEMIC*, Deliverable D2.1 Projet RNRT Temic, Technical report, INRIA/ENS-LIP, February 2005.