



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team reso

*Protocols and softwares for very
high-performance network*

Grenoble - Rhône-Alpes

Theme : Networks and Telecommunications

Activity
R *eport*

2010

Table of contents

1. Team	1
2. Overall Objectives	1
2.1. Project-team presentation overview	1
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 2010	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
2.8. Highlights	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	11
4. Application Domains	12
5. Software	12
5.1. Lyatiss Weaver suite	12
5.2. MPI5000	13
5.3. ShowWatts: Real time energy consumption grapher.	13
5.4. WattM : Monitoring framework for energy consumption of data centers	14
5.5. ANPI	14
5.6. OVNI5000	14
6. New Results	14
6.1. Optimized protocol implementation and networking equipments	14
6.1.1. Specifying and provisioning Virtual Infrastructures with HIPerNET	14
6.1.2. Joint elastic cloud and virtual network framework for application performance optimization and cost reduction	15
6.1.3. A Virtual Switch Architecture for Hosting Virtual Networks on the Internet	15
6.1.4. Size-Based Flow Scheduling in a CICQ Switch	15
6.1.5. Reliability Support in Virtual Infrastructures	15
6.1.6. HIPCAL: final report	16
6.1.7. Large Scale Autonomic Service Deployment in Next Generation Networks	16
6.1.8. Energy-efficiency in computing and networking for large-scale distributed systems	16
6.2. Quality of service and Transport Protocols for Future Networks	17
6.2.1. Auction-based Bandwidth Allocation Mechanisms for Wireless Future Internet	17
6.2.2. Traffic-aware Routing	17
6.2.3. Traffic classification techniques supporting semantic networks	18
6.3. High Speed Network's traffic metrology and statistical analysis	19
6.3.1. Modeling TCP Throughput: an Elaborated Large-Deviations-Based Model and its Empirical Validation	19

6.3.2.	A recurrent solution of Ph/M/c/N-like and Ph/M/c-like queues	19
6.3.3.	A note on simulating a M/G/1 queue for various service time distributions	19
6.3.4.	An approximate solution for Ph/Ph/1 and Ph/Ph/1/N queues	19
6.3.5.	Performance evaluation of a single node with general arrivals and service	20
6.3.6.	Methodology for Multifractal Analysis of Heart Rate Variability: From LF/HF Ratio to Wavelet Leaders	20
6.3.7.	Multifractal Analysis for In Partum Fetal-ECG Diagnosis	20
6.4.	Network services for high demanding applications	21
6.4.1.	Design and development of an MPI gateway	21
6.4.2.	High availability for clustered network equipments	21
6.4.3.	High availability for stateful network equipments	22
7.	Contracts and Grants with Industry	23
7.1.	Grants with Industry	23
7.2.	INRIA actions	23
7.2.1.	ARC MISSION	23
7.2.2.	GRID5000: ADT Aladdin	23
7.3.	National Initiatives	24
7.3.1.	ANR RESCUE	24
7.3.2.	FUI CompatibleOne Project	24
7.3.3.	ANR PETAFLOW	24
7.3.4.	ANR DMASC	25
7.3.5.	Action Interfaces Recherches en Grilles	25
7.4.	European Initiatives	25
7.4.1.	SAIL	25
7.4.2.	GEYSERS	26
7.4.3.	AUTONOMIC INTERNET	26
7.4.4.	OGF-EUROPE - 2008-2010	26
7.4.5.	PrimeEnergyIT	27
7.4.6.	COST Action IC804 on Energy Efficiency in Large Scale Distributed Systems	27
7.4.7.	Euro-NF Specific Joint Research Project : Security and Privacy Concerns in energy Efficient Computing	27
7.5.	Visitors	28
7.5.1.	Collaboration with University of California Santa Cruz (UCSC), US	28
7.5.2.	Collaboration with Universitat Politècnica de Catalunya (UPC)	28
7.5.3.	Collaboration with University of Thiés (Senegal)	28
7.5.4.	Collaboration with University of Tsukuba (Japan)	28
8.	Dissemination	28
8.1.	Conference organisation, editors for special issues	28
8.1.1.	Editorial Boards	28
8.1.2.	Chairing and Organisation of Conferences and Workshops	28
8.1.3.	Program committee members	28
8.1.4.	Participation in steering committees	29
8.1.5.	International expertise	29
8.1.6.	National expertise	29
8.1.7.	Public Dissemination	29
8.2.	Animation of the scientific community	30
8.3.	Teaching	31
8.3.1.	Graduate teaching	31
8.3.2.	Miscellaneous teaching	31
8.4.	Participation in boards of examiners and committees	32
8.5.	Seminars, invited talks	33

9. Bibliography**33**

1. Team

Research Scientists

Pascale Vicat-Blanc Primet [Team leader until June 1, 2010, Research Director (DR2) Inria, HdR]
Paulo Gonçalves [Team leader since June 1, 2010, Research Associate (CR1) Inria, HdR]
Laurent Lefèvre [Research Associate (CR1) Inria]

Faculty Members

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

Technical Staff

Abderhaman Cheniour [Expert Engineer INRIA - AUTONOMIC INTERNET (6 months)]
Romaric Guillier [Expert Engineer INRIA - HIPCAL (6 months), GEYSERS (1 month), SAIL (5 months)]
Matthieu Imbert [Research Engineer INRIA – 40%]
Philippe Martinez [Expert Engineer INRIA - DTI]
Jean-Christophe Mignot [Research Engineer CNRS – 40%]
Olivier Mornard [Expert Engineer INRIA - AUTONOMIC INTERNET (6 months), CNRS (6 months, 80%)]
Augustin Ragon [Expert Engineer INRIA - GEYSERS]
Armel Soro [Associated Engineer INRIA - ADT Aladdin]
Sébastien Soudan [Expert Engineer INRIA - EC-GIN (1 months), HIPCAL (5 months)]

PhD Students

Doreid Ammar [PhD student, INRIA , Bell Labs- 2009/2012]
Fabienne Anhalt [PhD student, INRIA - 2008/2011]
Mohammed Diouri [PhD student, Univ. Lyon 1, MENRT - 2010/2013]
Pierre-Solen Guichard [PhD student, INRIA-Bell Labs - 2008/2011, PhD interrupted in October 1, 2010]
Guilherme Koslovski [PhD student, INRIA - 2008/2011]
Dinil Mon Divakaran [PhD student, INRIA-Bell Labs - 2007/ Jul. 2010]
Van Dan Nguyen [PhD student (co-advisement with Univ. of Hanoi), AUF fellowship - 2010/2013]
Anne-Cécile Orgerie [PhD student - ENS, MENRT - 2008/2011]
Shubhabrata Roy [PhD student, INRIA SAIL, Dec. 2010/Nov. 2013]
Marina Sokol [PhD student, INRIA , Bell Labs- 2009/2012]

Post-Doctoral Fellow

Marcos Dias de Assunção [Post-doc INRIA - GREEN-NET (8 months), AUTONOMIC INTERNET (4 months)]

Visiting Scientist

Alexandre Brandwajn [Professor at University of California Santa Cruz (US) with RESO in April-June 2010]

Administrative Assistants

Sylvie Boyer [Secretary (SAR) INRIA – 10%]
Laetitia Lecot [Secretary INRIA since October 1, 2010 – 50%]
Caroline Suter [Secretary INRIA until October 1, 2010 – 50%]

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 analyzing 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 provides the context to devise, to model and to test innovative protocols, to design and to evaluate softwares. Moreover, the proposed solutions are implemented and experimented on real or emulated local or wide area testbeds, under real conditions and confronted 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 a standard and capacities of 100 Gbps are emerging. New technologies like 10 Gbps Ethernet or 10Gbps 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, a lot of researches and developments around the concept of grid and utility computing have underlined the strength 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 generalization of coarse Grain Web Servers like Google, Yahoo, Amazon providing the content, control, storage and computing resources for the users. All these trends 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 it is becoming a vast **cloud** constantly increasingly to embed the computational and the 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. Now, the network community strives not only at forwarding the packets taking into account the application semantics, but it also exploits new functionalities of the network to meet specific applications' requirements. Reciprocally, distributed systems and applications have traditionally been designed to run on top of the Internet whose architecture was taken as granted. Although the convergence of communication and computation at every level shows natural, it is still very difficult to efficiently explore the full range of possibilities it can bring in. So far, most of the approaches that were proposed to exploit this convergence, violate the initial design philosophy of the Internet protocol stack (end-to-end argument for example), or if implemented in the application layer, they have to face several issues regarding performance, resilience and scalability. Recently, ambitious research programs like FIND (NSF) or FIRE (EU) were launched to support the design of a new protocol architecture for the Future Internet and to circumvent the structural pitfalls of the current architecture. We believe that redesigning the Internet is an outstanding opportunity to better understand and to assess higher-level system requirements that in turn should steer 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 the main motivations of RESO's outcomes.

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 tackled not only to support the accelerating and dynamic growth of data traffic but also to address 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). Then, enabling ultra high performance machine to machine communications leads to new bandwidth and network resource sharing paradigms. RESO investigates different challenges such as quality of service, transport protocols, energy consumption 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 on key issues such as :

- where and how to integrate the autonomy required to manage and control high speed networks at large scale?
- how to deal 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)?
- To cope with applications' elasticity, scenarii for dynamic reconfiguration of virtual infrastructures rely on the central concepts of time-scale, regularity (burstiness) and *optimal* decision strategies. Can we address these aspects from the prospect of time series analysis?
- how to efficiently share the resources in a network that interleaves multimedia applications and computing applications (traffic awareness)?
- 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 processes 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 incontrovertible in large scale distributed system. In our research axis 1 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. A lot of the performance problems on end systems, but also on router's data plane, have to be studied and solved to make the virtualization approach viable. This is one of the research goals of this investigation axis.

On an other hand, the key enabling factor of new network services is programmability at every level; that is that of self-configuration ability of new softwares. We explore the concept of "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 mainly supported by the EU FP7 "Geyser" (2010-2012) and "SAIL" (2010-2012) projects, both concerning network virtualization.

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 following points:

- how does the traffic statistical properties really impact Quality of Service (QoS)?
- Characterize and model the resource usage variability inherent to specific applications, and use online measurement of these to control dynamic infrastructure adaptation.
- Real time identification and classification of transiting flows, according to a sensible typology?

This third axis is largely covered by “Semantic Networking” activities carried out in the common laboratory between INRIA and Alcatel-Lucent (2008-2012), as well as by the ANR project (*Programme Blanc*) “PetaFlow” (2009-2012) and the EU FP7 “SAIL” project.

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 high-performing 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 Grid’5000-ALADDIN initiative. The 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 computing 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 used in the scientific community, they are now moving into the commercial environment through the concept of Cloud computing solutions. Cloud computing fits a re-centralization 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 scenarios are explored with industrial partners. Researches conducted these last years reveal that grid technology raises 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 on heavily ossified research objects (the Internet protocol architecture) and relies on a challenging emerging application domain on a specific network context. These factors induce a close interaction between 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 statistical analysis, in estimation theory and in performance evaluation. This background work provides the basis for innovative protocols and software design. Moreover, we implement and experiment our proposed prototypes 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 2010

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

2.7.1. Axis 1 : Protocol implementations and networking equipments

- Design of VxSwitch, a virtualized router architecture with dimensionable buffers and configurable schedulers.
- Design and prototype implementation of an OpenFlow controller to create and manage virtual networks.
- Design of an Autonomic Network Programming Interface
- High availability for clustered stateful network equipments
- Models and software frameworks for energy efficiency in large scale distributed systems

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.
- Traffic aware routing
- Resources allocation in hierarchical networks

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

- Adaptation and generalization of the semi-supervised machine learning scheme to the context of flows classification.
- Identification on a long-lived TCP flow, of a new type of scale invariance property with multifractal support. Demonstration of a (almost sure) large deviation principle guaranteeing the identifiability of the corresponding multifractal spectrum from a finite size real trace. Adaptation to TCP markovian models to characterize fairness and sources' synchronization.
- Development of a mathematically ground and numerically efficient estimator of the Large Deviations Multifractal Spectrum.

2.7.4. Axis 4: Network Services for high demanding applications

- Design and development of MPI5000, a communication layer for MPI over wide area network
- Design and development of a metrology infrastructure for fine grain traffic monitoring in Grid5000.

2.8. Highlights

In June 2010, RESO underwent a major change with the creation by Pascale Vicat-Blanc of LyaTiss, an INRIA spin-off. The aim of this start-up is to valorize and to transfer the technologies supporting a new software architecture for the Future Internet and the Cloud Networking. These technologies are aimed at promoting a sustainable and fair economical model. This vision is based on 20 years of fundamental and experimental research in the area of internet protocols, networks, collaborative systems, computing grids and the cloud services, partially developed within the team RESO.

LyaTiss was incubated at CREALYS between 2009 and 2010 and officially launched in June 2010. This enterprise project was granted the OSEO emergence prize in 2009 and the OSEO creation-development award in 2010. The start-up was among the 15 selected enterprises in the area of information and communication technologies to participate to the French Tech Tour 2010 organized by UbiFrance in the Silicon Valley in June 2010, where it was awarded the second prize. This award is the recognition of the strong potentialities of LyaTiss to economically develop in this region.

3. Scientific Foundations

3.1. Optimized Protocol implementations and networking equipments

Participants: Laurent Lefèvre, Pascale Vicat-Blanc Primet, Jean-Patrick Gelas, Olivier Glück, Fabienne Anhalt, Jean-Christophe Mignot, Sébastien Soudan, 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 promising 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 signaling 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.

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, Laurent Lefèvre, Isabelle Guérin-Lassous, Dinil Mon Divakaran, Sébastien Soudan, Romaric Guillier, Guilherme Koslovski.

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 advocated 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 recognizes 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 and 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 time-scales: data, control, and management. Each time-scale 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 time-scales 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

time-scales; while routing protocols, such as BGP and OSPF operate at time-scales 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 re-configurable 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 cross-connects 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: Paulo Gonçalves, Thomas Begin, Pascale Vicat-Blanc Primet, Shubhabraya Roy, Marina Sokol, Matthieu Imbert.

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 – and more recently machine learning schemes – 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 hypotheses that sustain already investigated approaches (e.g. overlay, virtualizing network resources, distributing network treatments, middleware programming), we 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, adaptive and self-reconfigurable solutions.

We ground our approach on our large scale, fully controllable and configurable experimental facility (Grid5000+MetroFlux [60]) to validate, to better understand and to extend anterior results that were either heuristically observed or theoretically derived. Conversely, we perform realistic experiments, under prescribed and reproducible conditions, to get new insights into the statistical specificities of internet traffic, and to precisely identify the role of the network parameters [58].

Difficulty dwells in reliable classifiers and estimators of statistical properties, from non-stationary and possibly incomplete traces. We address these issues by proposing signal processing techniques well-tailored to network traffic measurements [59]. To go beyond the statistically description of the Internet traffic, we aim at investigating their effects on networking equipments such as routers and switches. An empirical study,

implying the synthesis of realistic traffic traces, along with comprehensive sets of experimental measurements should allow us to achieve this goal. RESO also aims at carrying out complementary analytical studies, based on the definition of theoretical models, so as to gain new insights in the performance evaluation resulting from the application of " Internet-like " traffic into classical queueing systems.

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

3.4. Network services for high demanding applications

Participants: Olivier Glück, Laurent Lefèvre, Jean-Patrick Gelas, Paulo Gonçalves, Olivier Mornard, Sébastien Soudan, Romaric Guillier.

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 [56]. 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 heterogeneous 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

RESO applies its research to the domains of high performance Cluster, Grid communications and more recently that of cloud networking. 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 [62] [61], [63]. Loss probability, important and incompressible latencies, dynamic behavior of network paths question profoundly models and technic used in parallel and distributed computing [55]. 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 pfdnet and Glif communities.

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].

- 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 expertise in high speed network and protocols exploration and tuning.
- RESO pursue the construction of an international community around Grid and Cloud networks through the european "GEYSERS" and "SAIL" projects as well as with the OGF networking community.
- Through the ANR PetaFlow project, RESO is collaborating with biology and medical imaging applications.
- The ANR project RESCUE, following-up the INRIA ARC "MISSION", deals with QoS aspects in hybrid networks.

5. Software

5.1. Lyatiss Weaver suite

Participants: Pascale Vicat-Blanc Primet, Sébastien Soudan, Romaric Guillier, Sébastien Soudan, Dinil Mon Divakaran, Philippe Martinez, Guilherme Koslovski, Fabienne Anhalt, Olivier Mornard, Armel Soro, Jean-Patrick Gelas.

The following list of softwares, whose development was initiated at RESO, are in constant evolution as they represent a direct output of the current research carried out at RESO. They also form the core of the technological transfer to Lyatiss; embedded in the Lyatiss Weaver Suite, they implement the solutions for virtual resources orchestration and infrastructure services.

- BDTS Dynamic network bandwidth allocation, bulk data transfers scheduling (APP : JBDS version 1 du 15/12/2007: IDDN.FR.001.220025.000.S.P.2008.000.10700)
- FLOC Limitation and triggering of flow rate (APP: version 0.12 du 17 février 2009: IDDN.FR.001.290009.000.S.P.2009.000.10200)
- NXE Definition, configuration, deployment, run and analysis of a large scale experiment for protocol evaluation (APP: version 1.0 de novembre 2008: IDDN.FR.001.030005.000.S.P.2009.000.10800)
- VXcore Resource temporal database manager (APP: version 1.0 du 15 mars 2009 : IDDN.FR.001.290012.000.S.P.2009.000.10800)
- VXtopology Resource spacial database manager (APP: version 1.0 du 15 mars 2009 : IDDN.FR.001.290012.000.S.P.2009.000.10800)
- VXScheduler Adaptation of virtual infrastructure request and scheduling (patent version 1.0 du 15 mars 2009 : IDDN.FR.001.290010.000.S.P.2009.000.10800)
- VXDL parser interpretation and XML traduction of virtual infrastructures specifications (patent version 2.0 du 20 mars 2009: IDDN.FR.001.260009.000.S.P.2009.000.10800)
- SRVdemonstrator Scheduling, Reconfiguration and Virtualisation of Network resources for intensive computing environment.
- PATHNIF Systematically analysis and evaluate the capacity of potential bottlenecks of an end to end network path (APP: version 1.0 de mars 2009 : IDDN.FR.001.260002.000.S.P.2009.000.10800).
- HIPerNet v0.5: Cloud solution HIPerNet engine is a software implementing discovery, selection, allocation, scheduling and management of virtual private execution infrastructures over the Internet. HIPerNET v0.5 is focusing on virtual end-resource deployment and configuration (distributed under GPL license).

5.2. MPI5000

Participants: Olivier Glück, Jean-Christophe Mignot.

MPI5000 is a communication layer between the application (MPI for example) and the transport protocol (TCP) which improves communications of distributed applications over wide area network in grids. For instance, MPI5000 reduces the impact of retransmissions and the impact of congestion window in such a context. MPI5000 is automatically and transparently executed without modifying the application. The general principle is to introduce proxies at the interface between the local network and the long- distance network to differentiate communications. These proxies allows to put forward the split of TCP connections in order to avoid losses and retransmissions on the long-distance links. This mechanism also allows to keep the congestion window closer to available throughput on the long-distance network. This work is detailed and evaluated in [57], and shows which applications can benefit from these optimisations.

5.3. ShowWatts: Real time energy consumption grapher.

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

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.

5.4. WattM : Monitoring framework for energy consumption of data centers

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

- Functional description: Monitoring and exposing electrical usage of large scale number of resources.
- usage and deployment in GRID5000
- Status: Open Source software
- State: prototype

5.5. ANPI

Participants: Laurent Lefèvre, Cheniour Abderhaman.

In the context of the European STREP Autonomic Internet project, the ANPI (Autonomic Network Programming Interface) framework has been designed and developed. This framework allows the deployment and configuration of autonomic internet services. This software is part of the Autonomic Internet architecture and was used during Autonomic Internet reviews and demonstrations of large scale autonomic service deployment [42], [49], [34].

5.6. OVNI5000

Participants: Laurent Lefèvre, Mornard Olivier.

In the context of the European STREP Autonomic Internet project, the OVNI5000 (Orchestrated Virtual components framework In Grid5000) software framework has been designed and developed in order to use and deploy virtual infrastructure over Grid5000. It is designed as a distributed framework which doesn't require any centralized server or information. For optimization reason the root file system image is stored on the front end to speed up the deployment (root file-system image are big). The software is automatically installed in the different part of the infrastructure. The OVNI5000 framework was used during Autonomic Internet reviews and demonstrations of large scale autonomic service deployment [42], [49], [34].

6. New Results

6.1. Optimized protocol implementation and networking equipments

6.1.1. Specifying and provisioning Virtual Infrastructures with HIPerNET

Participants: Fabienne Anhalt, Guilherme Koslovski, Pascale Vicat-Blanc Primet.

With the expansion and convergence of communication and computing, dynamic provisioning of customized networking and processing infrastructures, as well as resource virtualization, are appealing concepts and technologies. Therefore, new models and tools are needed to allow users to create, trust and enjoy such on-demand virtual infrastructures within a wide area context. This paper presents the HIPerNET framework we are designing and developing for creating, managing and controlling virtual infrastructures in the context of high-speed Internet. The key idea of this proposal is the combination of network- and system-virtualization associated with controlled resource reservation to provide fully isolated environments. HIPerNET's motivations and design principles are presented. Then we examine specifically how this framework handles the virtual infrastructures, called Virtual Private eXecution Infrastructures (VPXI). To help specifying customized isolated infrastructures, HIPerNET relies on VXML, a language for VPXI description and modeling which considers end-host resource as well as the virtual network topology interconnecting them, including virtual routers. After the specification, allocation and scheduling phases, HIPerNET helps in provisioning, deploying and configuring virtual private execution infrastructures. This means, it triggers the dynamic configuration of all the equipments involved. In this paper we concentrate on network configuration, particularly to achieve network performance isolation. We also study and evaluate mechanisms to implement and configure virtual-link control. Experimental results obtained within the Grid'5000 testbed are presented and analyzed.

6.1.2. Joint elastic cloud and virtual network framework for application performance optimization and cost reduction

Participants: Fabienne Anhalt, Guilherme Koslovski, Pascale Vicat-Blanc Primet.

Cloud computing infrastructures are providing resources on demand for tackling the needs of large-scale distributed applications. To adapt to the diversity of cloud infrastructures and usage, new operation tools and models are needed. Estimating the amount of resources consumed by each application in particular is a difficult problem, both for end users who aim at minimizing their costs and infrastructure providers who aim at controlling their resources allocation. Furthermore, network provision is generally not controlled on clouds. This paper describes a framework automating cloud resources allocation, deployment and application execution control. It is based on a cost estimation model taking into account both virtual network and nodes managed by the cloud. The flexible provisioning of network resources permits the optimization of applications performance and infrastructure cost reduction. Four resource allocation strategies relying on the expertise that can be captured in workflow-based applications are considered. Results of these strategies are confined virtual infrastructure descriptions that are interpreted by the HIPerNet engine responsible for allocating, reserving and configuring physical resources. The evaluation of this framework was carried out on the Aladdin/Grid'5000 testbed using a real application from the area of medical image analysis.

6.1.3. A Virtual Switch Architecture for Hosting Virtual Networks on the Internet

Participants: Fabienne Anhalt, Dinil Divakaran, Pascale Vicat-Blanc Primet.

The future Internet is envisioned to host a large number of virtual networks managed by different operators sharing the same physical infrastructure. In such a scenario, an operator may not even own physical resources as such, but lease virtual resources to have their own virtual networks. While control-plane virtualization with several routing instances becomes common in equipments, the data plane is generally shared relying on logical isolation or resource segmentation, lacking in efficient sharing with performance guarantees. This makes it necessary to look into layer 2, and virtualize the switching fabric to have control over the sharing of the most critical resources for packet switching. In this article, we come up with a flexible architectural design for virtualizing a switching fabric. This new architecture enables a multitude of choices to customize virtual switches as needed. Our simulations show the relative loss in performance brought by virtualization, and demonstrate how the offered flexibility can help in exploiting the resources in a new way, satisfying independent virtual switch requirements.

6.1.4. Size-Based Flow Scheduling in a CICQ Switch

Participants: Dinil Divakaran, Fabienne Anhalt, Pascale Vicat-Blanc Primet.

Size-based (SB) scheduling policies have been shown to improve response times of small flows, without degrading the performance of large flows. But these differentiating policies are designed for Output-queued switch architecture, which is known to have scalability issues. On the other hand, the buffered-crossbar (BX) switch architecture is currently being pursued as a potential next-generation scalable switch architecture. This work looks into the problem of performing SB scheduling in BX switches. In particular, the design goals, w.r.t each output port, are (i) to transmit high-priority packet(s) as long as there is at least one present, and (ii) to respect the FIFO order among high-priority packets. In this direction, we propose to use PIFO queue at each crosspoint of a CICQ switch. The initial design presented as pCICQ-1 switch is simple and guarantees that packet-priorities are respected once they are in the crosspoint queues. But it does not maintain the FIFO order of high-priority packets, besides letting a bounded number low-priority packets to depart through an output, when there are one or more high-priority packets for the same output. To solve this, we propose an enhancement, as pCICQ-2 switch, that achieves both the design goals.

6.1.5. Reliability Support in Virtual Infrastructures

Participants: Guilherme Koslovski, Pascale Vicat-Blanc Primet.

Through the recent emergence of joint resource and network virtualization, dynamic composition and provisioning of time-limited and isolated virtual infrastructures is now possible. One other benefit of infrastructure virtualization is the capability of transparent reliability provisioning (reliability becomes a service provided by the infrastructure). In this context, we discuss the motivations and gains of introducing customizable reliability of virtual infrastructures when executing large-scale distributed applications, and present a framework to specify, allocate and deploy virtualized infrastructure with reliability capabilities. An approach to efficiently specify and control the reliability at runtime is proposed. We illustrate these ideas by analyzing the introduction of reliability at the virtual-infrastructure level on a real application. Experimental results, obtained with an actual medical-imaging application running in virtual infrastructures provisioned in the experimental large-scale Grid5000 platform, show the benefits of the virtualization of reliability.

6.1.6. HIPCAL: final report

Participants: Fabienne Anhalt, Romaric Gillier, Guilherme Koslovski, Pascale Vicat-Blanc Primet.

HIPCAL studies a new paradigm (grid substrate) based on confined virtual cluster concept 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 clusters are multiplexed in time and space, isolated and protected. The goal of this project is to explore an approach in a break with principles developed in grids to jointly enhance the application portability, the communication's performance control and their security. The project is providing a IaaS (Infrastructure as a service) solution based on end to end bandwidth reservation, control overlay, network and system virtualization, cryptographic identification principles. The proposal has been demonstrated, 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 100 Mbps to 10 Gbps, few microseconds to 100ms will be involved in these experimentations. Comparison with Globus, Planetlab and Cluster on Demand approaches will be one of the specific goals of the experiments. We aim at demonstrating the functional transparency, enhanced predictability and efficiency for applications offered by the HIPCAL approach. The LYaTiss startup, has been launched on the 1st June 2010 to pursue the development and the commercial diffusion of the solution and results of the project.

6.1.7. Large Scale Autonomic Service Deployment in Next Generation Networks

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

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) and the OVNI (Orchestrated Virtual components framework In Grid5000) which support the service enablers plane of the AUTOI architecture. This interface is currently has been demonstrated on the Grid5000 platform, used and validated by the AUTOI partners (Hitachi Europe, University College of London, UPC Barcelona, University of Passau) and disseminated in deliverables and publications [42], [49], [34].

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

Participants: Marcos Dias de Assuncao, Diouri Mohammed, 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. The energy monitoring infrastructure designed by RESO under the name "Green Grid5000" is used for different exploration of energy awareness [27], [26], [48], [47]. RESO has launched the ICT Energy Logs 4all repository (<http://www.ens-lyon.fr/LIP/RESO/ict-energy-logs/>) in order to share and disseminate energy logs collection among research communities.

We propose an Energy-efficient Reservation Infrastructure for large-scale Distributed Systems (ERIDIS) which is able to collect and expose energy information of large scale distributed systems infrastructures, to aggregate some resources reservations in time and space and to express Green Policies. The ERIDIS approach can be applied in the context of Grids, Clouds and Energy Efficient Networks.

This model has been validated over the Grid5000 traces by using a replay mechanism and through some simulations frameworks [39], [25], [33]. The ERIDIS framework has been evaluated and validated in the context of the Green Cloud [16], [46].

6.2. Quality of service and Transport Protocols for Future Networks

6.2.1. Auction-based Bandwidth Allocation Mechanisms for Wireless Future Internet

Participant: Isabelle Guérin Lassous.

An important aspect of the Future Internet is the efficient utilization of (wireless) network resources. In order for the - demanding in terms of QoS - Future Internet services to be provided, the current trend is evolving towards an "integrated" wireless network access model that enables users to enjoy mobility, seamless access and high quality of service in an all-IP network on an "Anytime, Anywhere" basis. The term "integrated" is used to denote that the Future Internet wireless "last mile" is expected to comprise multiple heterogeneous geographically coexisting wireless networks, each having different capacity and coverage radius. The efficient management of the wireless access network resources is crucial due to their scarcity that renders wireless access a potential bottleneck for the provision of high quality services.

In this work, we propose an auction mechanism for allocating the bandwidth of such a network so that efficiency is attained, i.e. social welfare is maximized. In particular, we propose an incentive-compatible, efficient auction-based mechanism of low computational complexity. We define a repeated game to address user utilities and incentives issues. Subsequently, we extend this mechanism so that it can also accommodate multicast sessions. We also analyze the computational complexity and message overhead of the proposed mechanism. We then show how user bids can be replaced from weights generated by the network and transform the auction to a cooperative mechanism capable of prioritizing certain classes of services and emulating DiffServ and time-of-day pricing schemes. The theoretical analysis is complemented by simulations that assess the proposed mechanisms properties and performance. We finally provide some concluding remarks and directions for future research.

6.2.2. Traffic-aware Routing

Participants: Thomas Begin, Isabelle Guérin Lassous.

Internet is undergoing a strong increase in the number of constrained applications it serves, and this leads to a ever greater and challenging issue for QoS solutions. Many QoS solutions are based on a best-effort routing that does not take into account parameters such as delay or loss rate. For instance, DiffServ, often considered as the most implemented QoS solution by operators, is actually based on QoS-unaware routes, and therefore attempts to provide QoS mechanisms above the routes provided the classical routing protocols! QoS routing is often viewed as the missing piece in a full QoS architecture for the Internet.

In a first work, we study the multi-class routing issue under a very simple classification that is based on the used transport protocol. We assume that sensitive traffic is sent with UDP while elastic traffic is sent with TCP. We investigate the choice of different routing protocols according to the used transport protocol. As a first step, we only consider routing based on static metrics that are the number of hops, the propagation delay, the propagation delay summed to queue length within buffers and the link capacity. To compute the routes, Dijkstra algorithm is applied for the first three metrics while the widest shortest path is used for the fourth metric.

Our preliminary results tend to show that the best achievable configuration is the use of a routing on the number of hops coupled with a routing on propagation delay summed to queue size independently of the transport protocol (UDP with the first routing and TCP with the second one or the opposite). The worst one is the two routings both on capacity. However, there is no outstanding configuration and several combinations are very close in terms of performance. This may be due to the use of static metrics which do not well reflect the real use of the network and/or the used classification UDP/TCP that may be inappropriate for multi-class routing.

In a second work, we investigate the multi-constrained routing under a more realistic framework. Among the several existing multi-constrained routing algorithms, we choose PIRA algorithm since it was recently listed as one of the most efficient. This algorithm works in two steps. First, all hops k-shortest paths are iteratively discovered. Then, another iteration is used to find a feasible path that meets the expressed constraints. As PIRA imposes no specific types of metrics, we are able to implement any instance of it, fitting it to the constraints required by the application. Thus, we transform the PIRA algorithm into a QoS routing protocol and implement it in NS2.

First, our results tend to show that, with the configuration of today's networks, setting up routes for inelastic traffic with two link metrics, i.e. available bandwidth and loss rate, yields results as good as those with three metrics (including delay). Secondly, it appears from our experiments that the fresh space's criticality widely differs according to the type of QoS routing. For instance, relying only on links available bandwidth to decide how to route inelastic traffic can be effective, yet it requires a refresh pace much higher than this needed with two (or three) link metrics. Therefore, a network operator may face a trade-off between a multi-constrained routing with lower pace of refresh and a single-constrained routing but with higher pace of refresh. Investigating more around appropriate values for refresh pace is a piece of our future works.

6.2.3. Traffic classification techniques supporting semantic networks

Participants: Olivier Grémillet, Paulo Gonçalves, Pascale Vicat-Blanc.

The Semantic Networking concept has been introduced to solve the QoS, scalability and complexity challenges for the Future of Internet. Based on traffic awareness and considering flow entities, it contributes to an adaptive management of the network and provides better knowledge of the transported traffic. Studying the processing time of the classification compatible with real-time operation of such networks is a key question for implementation purposes. In this paper, we present interesting techniques for classification of traffic in semantic networks. The Sample & Hold and multi-stage filter schemes are studied to detect the biggest flows. Their performance is evaluated on real traffic traces. In addition the classification of traffic according to the originating application is investigated. In particular, we analyze the influence of many parameters derived from a traffic flow on the performance of application identification and classify them according to their accuracy. By doing this, a light scheme is proposed able to classify accurately the traffic. We finally discuss the architecture of an hardware implementation to validate the concept of semantic networking.

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

6.3.1. Modeling TCP Throughput: an Elaborated Large-Deviations-Based Model and its Empirical Validation

Participants: Paulo Gonçalves, Pascale Vicat-Blanc.

Joint work with Patrick Loiseau (Post-Doctoral fellow at Sysiphe, INRIA Rocquencourt).

In today's Internet, a large part of the traffic is carried using the TCP transport protocol. Characterization of the variations of TCP traffic is thus an important issue, both for resource provisioning and Quality of Service purposes. However, most existing models are limited to the prediction of the (almost-sure) mean TCP throughput and are unable to characterize deviations from this value. In this paper, we propose a method to describe the deviations of a long TCP flow throughput from its almost-sure mean value. This method relies on an ergodic large deviations result, which was recently proved to hold on almost every single realization for a large class of stochastic processes. Applying this result to a Markov chain modeling the congestion window evolution of a long-lived TCP flow, we show that it is practically possible to quantify and to statistically bound the throughput variations at different scales of interest for applications. Our Markov-chain model can take into account various network conditions and we demonstrate the accuracy of our method prediction in different situations using simulations, experiments and real-world Internet traffic. In particular, in the classical case of Bernoulli losses, we demonstrate: (i) the consistency of our method with the widely-used square-root formula predicting the almost-sure mean throughput, and (ii) its ability to additionally predict finer properties reflecting the traffic variability at different scales.

6.3.2. A recurrent solution of Ph/M/c/N-like and Ph/M/c-like queues

Participants: Thomas Begin, Alexandre Brandwajn.

This work, submitted to Journal of Applied Probability [41], was performed during the visit of Pr. Brandwajn in the RESO team project.

We propose an efficient semi-numerical approach to compute the steady-state probability distribution for the number of requests at arbitrary and at arrival time instants in Ph/M/c-like systems in which in the inter-arrival time distribution is represented by an acyclic set of memoryless phases. Our method is based on conditional probabilities and results in a simple computationally stable recurrence. It avoids the explicit manipulation of potentially large matrices and involves no iteration. Due to the use of conditional probabilities, it delays the onset of numerical issues related to floating-point underflow as the number of servers and/or phases increases. For generalized Coxian distributions, the computational complexity of the proposed approach grows linearly with the number of phases in the distribution.

6.3.3. A note on simulating a M/G/1 queue for various service time distributions

Participants: Thomas Begin, Alexandre Brandwajn.

In this note [36], we highlight a potential pitfall that can occur when certain types of distributions are handled within a model. Our study relies on a simple queueing model that has a nice analytical resolution, and therefore can be used as a fair comparison basis. Our results suggest that using certain types of distributions (including Pareto distribution) within models tends to increase the complexity of simulations as a means to solve these models. We show that this effect grows with increasing values of the service time coefficient of variation. Thus, if a distribution within a model is only defined by its first two (or n) moments, then one should choose the distribution that is the simplest to simulate.

6.3.4. An approximate solution for Ph/Ph/1 and Ph/Ph/1/N queues

Participants: Thomas Begin, Alexandre Brandwajn.

In order to adequately represent heavy-tailed distributions (used to characterize workloads in computer networks and I/O subsystems) as phase-type distributions, it is necessary to use a large number of phases. This may cause problems for classical numerical solution methods for the resulting queueing system. We propose a simple approximation to assess the steady-state probabilities of the number of customers in Ph/Ph/1 and Ph/Ph/1/N queues, as well as probabilities found by an arriving customer. The latter include the probability of buffer overflow in the case of the Ph/Ph/1/N queue. The phase-type distributions considered are assumed to be acyclic. Our method involves iteration between solutions of an M/Ph/1 queue with state-dependent arrival rate and a Ph/M/1 queue with state-dependent service rate. We solve these queues using a simple and efficient recurrence. By iterating between these two simpler models our approximation divides the state space, and is thus able to easily handle phase-type distributions with large numbers of stages (in excess of a hundred). The proposed method converges typically within a few tens of iterations. Our approximation is asymptotically exact, and its overall accuracy is good: generally within a few percent of the exact values, except when both the inter-arrival and the service time distributions exhibit low variability. In the latter case, especially under moderate loads, the use of our method is not recommended.

6.3.5. Performance evaluation of a single node with general arrivals and service

Participants: Thomas Begin, Alexandre Brandwajn.

Queueing delays experienced by packets buffered at a node are among the most difficult to predict when considering the performance of a flow in a network. The arrivals of packets at a node tend to be highly variable so that a finite-buffer single-server queue with general arrivals and service emerges as a natural model of a network link. In this paper we propose an approach to the solution of such a queue when the times between arrivals and service times are represented as acyclic phase-type distributions. The proposed solution approach, based on the use of conditional probabilities, is conceptually simple, easy to implement in a standard computer language, numerically robust and reasonably fast.

In addition to standard steady-state probabilities and queue size averages, the proposed approach produces the probabilities of the state of the queue found by an arriving packet, in particular, the packet loss probability, directly linked to the QoS perceived by the user.

6.3.6. Methodology for Multifractal Analysis of Heart Rate Variability: From LF/HF Ratio to Wavelet Leaders

Participant: Paulo Gonçalves.

Joint work with Patrice Abry (CNRS, ENS Lyon), Stéphane Jaffard (Univ. de CrÃ©teil) and Muriel Doret (Hospices civils de Lyon).

The present contribution aims at proposing a comprehensive and tutorial introduction to the practical use of wavelet Leader based multifractal analysis to study heart rate variability. First, the theoretical background is recalled. Second, practical issues and pitfalls related to the selection of the scaling range or statistical orders, minimal regularity, parabolic approximation of spectrum and parameter estimation, are discussed. Third, multifractal analysis is connected explicitly to other standard characterizations of heart rate variability: (mono)fractal analysis, Hurst exponent, spectral analysis and the HF/LF ratio. This review is illustrated on real per partum fetal ECG data, collected at an academic French public hospital, for both healthy fetuses and fetuses suffering from acidosis.

6.3.7. Multifractal Analysis for In Partum Fetal-ECG Diagnosis

Participant: Paulo Gonçalves.

Joint work with Patrice Abry (CNRS, ENS Lyon) and Muriel Doret (Hospices civils de Lyon).

In partum fetal suffering surveillance is a key task to prevent fetal and neonatal mortality due to asphyxia. This is partially conducted by monitoring and analyzing fetal Electrocardiogram recorded during the delivery phase: A strong variability measured on the corresponding heart beat time series indicates a normal process. Though satisfactory in practice, the currently used analysis/decision criteria lead to a high number of false positives. Multifractal analysis can be envisaged as a new tool to revisit time series variability analysis. Applied to data collected at Hospices Civils de Lyon, France, wavelet Leader based multifractal analysis is shown here to achieve a significant discrimination between the True Negative, True positive and False Positive classes of patients. This hence open promising tracks to decrease the number of False Positives achieved with current tools.

6.4. Network services for high demanding applications

6.4.1. Design and development of an MPI gateway

Participants: Olivier Glück, Jean-Christophe Mignot.

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 propose a new transparent layer called MPI5000 and placed between MPI and TCP to transparently split TCP connections between MPI processes in order to take into account the grid topology. The implementation of MPI5000 is based both on a library between MPI and the operating system and on relays. Thus, the proposed architecture is independent of MPI implementations and is totally transparent for applications.

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, it allows 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. As MPI applications are mostly using small messages, they are more penalized if the network is congestionned by large flows. Thanks to the communication aggregation between relays, we have showed that our architecture allows to keep the congestion window closer to available throughput on the long-distance network.

We have analyzed for many points the overhead and the benefits of using our proxies. The theoretical analysis is supported by experiments. We have concluded that for MPI applications that are using collective operations, the benefit on losses and retransmissions generally do not hide the overhead added by the splitting of the connections. Other applications benefit from this mechanism if they communicate sufficiently.

Then, thanks to our architecture, we have proposed to use different TCP implementations for local and distant communications. We have tested the following TCP variants: Cubic, Bic, Reno, Htcp, HighSpeed, Scalable and Illinois. The results obtained show that Htcp, Bic and Cubic have the best results for WAN communications.

Finally, on each relay, we have implemented different scheduling strategies for MPI messages: the first one gives the priority to MPI control messages, the second one to MPI collective operations and the last one to MPI small messages. We have done several experiments on the national GRID'5000 testbed to evaluate the impact of our scheduling strategies on application execution time. We have launched 16 MPI tasks on two different sites for each NAS parallel benchmark and measured the application execution time for each strategy. The results show that the better strategy is the one that gives priority to MPI control messages. Using this strategy, some applications can decrease their execution time by almost 10%.

6.4.2. High availability for clustered network equipments

Participant: Laurent Lefèvre.

Joint work with Narjess Ayari (Orange).

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 No. FR0653546

[54], [53], [50], [52], [51]

Narjess Ayari has defended her PhD in October 2008. In 2010, this research activity has resulted to the following publication : [12].

6.4.3. High availability for stateful network equipments

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

In 2010, this research activity has resulted to the following publication : [18].

7. Contracts and Grants with Industry

7.1. Grants with Industry

7.1.1. INRIA Bell Labs common laboratory: *Semantic Networking*

Participants: Pascale Vicat-Blanc Primet, Isabelle Guerin-Lassous, Paulo Gonçalves, Thomas Begin, Dinil Mon Divakaran, Pierre-Solen Guichard, Doreid Ammar, Marina Sokol.

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.

7.2. INRIA actions

7.2.1. *ARC MISSION*

Participants: Thomas Begin, Paulo Gonçalves, Isabelle Guérin Lassous.

The project Mobile SubStitutiOn Networks (MISSION) is focused on the performance study, the possibilities and the feasibility to deploy a fleet of mobile wireless routers to help a wired network that can not offered its services anymore. This project deals with the theoretical aspects as the practical aspects of such a deployment. From a theoretical point of view, one problem is to minimize the number of used routers while rebuilding the network to replace. The main difficulty lies in the possibility to offer the services provided by the wired network in a transparent way. The controlled mobility allows a redeployment or an adaptation of the built network according to the needs or to the on-going traffic on the network. This controlled mobility should improve the network performance.

7.2.2. *GRID5000: ADT Aladdin*

Participants: Laurent Lefèvre, Gelas Jean-Patrick, Olivier Glück, Paulo Gonçalves, Matthieu Imbert, Armel Soro, Olivier Mornard, Jean-Christophe Mignot, Diouri Mohammed, Orgerie Anne-Cécile.

ENS Lyon is involved in the GRID'5000 project, which is an experimental Grid platform gathering ten 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 the 10 Gbit per second network. Lyon site is nationally recognized to gather the "networking expertise" with skilled researchers and engineers and dedicated networking equipments (Metroflux, GNET10...). Lyon site also hosts an important part of the Green Grid5000 infrastructure by hosting a set of 1500 wattmeters and exposing energy measurements to the Grid5000 community.

RESO is strongly involved in the choices of Grid5000's network components and architecture. Laurent Lefèvre is member of the national committee (comité de direction) of GRID'5000, of the Aladdin scientific committee and responsible of the Lyon site.

7.3. National Initiatives

7.3.1. ANR RESCUE

Participants: Thomas Begin, Paulo Gonçalves, Isabelle Guérin Lassous.

Access and metropolitan networks are much more limited in capacity than core networks. While the latter operate in over-provisioning mode, access and metropolitan networks may experience high overload due to evolution of the traffic or failures. In wired networks, some failures (but not all) are handled by rerouting the traffic through a backup network already in place. In developed countries, backup networks are adopted wherever possible (note that this is generally not the case for the links between end users and their local DSLAM). Such a redundant strategy may not be possible in emerging countries because of cost issues. When dedicated backup networks are not available, some operators use their 3G infrastructure to recover some specific failures; although such an alternative helps avoid full network outage, it is a costly solution. Furthermore, availability of 3G coverage is still mainly concentrated in metropolitan zones. When no backup networks are available, it would be interesting to deploy, for a limited time corresponding to the period of the problem (i.e., failure or traffic overload), a substitution network to help the base network keep providing services to users.

In the RESCUE project (2010-2013), we will investigate both the underlying mechanisms and the deployment of a substitution network composed of a fleet of dirigible wireless mobile routers. Unlike many projects and other scientific works that consider mobility as a drawback, in RESCUE we use the controlled mobility of the substitution network to help the base network reduce contention or to create an alternative network in case of failure. The advantages of an on-the-fly substitution network are manifold: Reusability and cost reduction; Deployability; Adaptability.

The RESCUE project addresses both the theoretical and the practical aspects of the deployment of a substitution network. From a theoretical point of view, we will propose a two-tiered architecture including the base network and the substitution network. This architecture will describe the deployment procedures of the mobile routing devices, the communication stack, the protocols, and the services. The design of this architecture will take into account some constraints such as quality of service and energy consumption (since mobile devices are autonomous), as we want the substitution network to provide more than a best effort service. From a practical point of view, we will provide a proof of concept, the architecture linked to this concept, and the necessary tools (e.g., traffic monitoring, protocols) to validate the concept and mechanisms of on-the-fly substitution networks. At last but not least, we will validate the proposed system both in laboratory testbeds and in a real-usage scenario.

7.3.2. FUI CompatibleOne Project

Participants: Laurent Lefèvre, Jean-Patrick Gelas, Olivier Mornard.

The project CompatibleOne (Nov 2010-Nov 2012) funded by the Fonds Unique Interministériel² (FUI) is dealing with the building of a Cloud architecture open software stack.

In this project, RESO is focused on the design and provisioning of energy aware and energy efficient components in order to include energy aspects in QoS, SLAs and billing in clouds architectures. RESO is leading the task T3.4 on energy management and will participate in activities on virtual machines design and migration.

7.3.3. ANR PETAFLOW

Participants: Paulo Gonçalves, Pascale Vicat-Blanc Primet, Matthieu Imbert.

This ANR (Appel Blanc International) started in october 2009 and will end in september 2012. It is a collaborative project between the GIPSA Lab (Grenoble), MOAIS (INRIA Grenoble), RESO (INRIA Grenoble), the University of Osaka (the Cybermedia Center and the Department of Information Networking) and the University of Kyoto (Visualization Laboratory).

We aim at proposing network solutions to guarantee the Quality of Service (in terms of reliability level and of transfer delay properties) of a high speed, long-distance connection used in an interactive, high performance computing application. Another specificity of this application is the peta-scale volume of the treated data corresponding to the upper airway flow modeling.

7.3.4. ANR DMASC

Participant: Paulo Gonçalves.

Started in october 2008, this ANR project, led by J. Barral (Sisyphé, INRIA Roquencourt), 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).

Its main objective is to develop advanced multifractal analysis tools, from mathematically ground results to efficient estimators. We apply these methods to the analysis, to the modeling and to the classification (for non invasive diagnoses) of cardio-vascular systems.

7.3.5. Action Interfaces Recherches en Grilles

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

RESO is leading the Action Interfaces Recherches en Grilles : "Energy efficiency in Grids from experimental to operational platforms" (supported by Institut des Grilles and Aladdin INRIA : 2010-2011). The goal of this action is to propose and evaluate energy aware software components able to be deployed in production data centers of Grids. Some energy monitoring infrastructure will be deployed and validated in operational centers (like the IN2P3 Data center).

7.4. European Initiatives

7.4.1. SAIL

Participants: Pascale Vicat-Blanc, Paulo Gonçalves, Thomas Begin, Guilherme Koslovski, Shubhabrata Roy, Romaric Guillier.

SAIL? objective is the research and development of novel networking technologies using proof-of-concept prototypes to lead the way from current networks to the Network of the Future. SAIL leverages state of the art architectures and technologies, extends them as needed, and integrates them using experimentally-driven research, producing interoperable prototypes to demonstrate utility for a set of concrete use-cases. SAIL reduces costs for setting up, running, and combining networks, applications and services, increasing the efficiency of deployed resources (e.g., personnel, equipment and energy). SAIL improves application support via an information-centric paradigm, replacing the old hostcentric one, and develops concrete mechanisms and protocols to realise the benefits of a Network of Information (NetInf). SAIL enables the co-existence of legacy and new networks via virtualisation of resources and self-management, fully integrating networking with cloud computing to produce Cloud Networking (CloNe). SAIL embraces heterogeneous media from fibre backbones to wireless access networks, developing new signalling and control interfaces, able to control multiple technologies across multiple aggregation stages, implementing Open Connectivity Services (OConS). SAIL also specifically addresses cross-cutting themes and non-technical issues, such as socioeconomics, inclusion, broad dissemination, standardisation and network migration, driving new markets, business roles and models, and increasing opportunities for both competition and cooperation. SAIL gathers a strong industry-led consortium of leading operators, vendors, SME, universities and research centres, with a valuable experience acquired in previous FP7 projects, notably 4WARD. The impact will be a consensus among major European operators and vendors on a well-defined path to the Network of the Future together with the technologies required to follow that path.

7.4.2. GEYSERS

Participants: Pascale Vicat-Blanc, Paulo Gonçalves, Fabienne Anhalt, Sébastien Soudan, Romaric Guillier.

GEYSERS vision is to qualify optical infrastructure providers and network operators with a new architecture, to enhance their traditional business operations. Optical network infrastructure providers will compose logical infrastructures and rent them out to network operators; network operators will run cost-efficient, dynamic and mission-specific networks by means of integrated control and management techniques. In the GEYSERS concept, high-end IT resources at users' premises are fully integrated with the network services procedures, both at the infrastructure planning and connection provisioning phases.

Following this vision, GEYSERS will specify and implement a novel optical network architecture able to support "Optical Network + Any-IT" resource provisioning seamlessly and efficiently. Energy consumption metrics for the end-to-end service routing are part of this efficiency.

GEYSERS is proposing to:

- Specify and develop mechanisms that allow infrastructure providers to partition their resources (optical network and/or IT), compose specific logical infrastructures and offer them as a service to network operators. This will be done overcoming the current limitations of networks/domain segmentation, and will support dynamic and on-demand changes in the logical infrastructures.
- Specify and develop a Network Control Plane for the optical infrastructure, by extending standard solutions (ASON/GMPLS and PCE), able to couple optical network connectivity and IT services automatically and efficiently, and provide them in 1 step, dynamically and on-demand, including infrastructure re-planning mechanisms.

These achievements will enable infrastructure providers, network operators and application providers to participate in new business scenarios where complex services with complex attributes and strict bandwidth requirements can be offered economically and efficiently to users and applications. The GEYSERS outcomes will be validated in an EU-wide optical network test-bed.

7.4.3. AUTONOMIC INTERNET

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

Autonomic Internet (AutoI - FP7.ICT.2007.Call1-216404 - 2008-2010) 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> In 2010, this project has resulted to the following publications : [35], [34] and dissemination event [49]

7.4.4. OGF-EUROPE - 2008-2010

Participants: Laurent Lefèvre, Augustin Ragon.

RESO participates in the OGF-Europe to reinforce the french participation to OGF standardization activities. We mainly concentrate our contribution on Telco interaction and Energy-efficiency in Grid context.

7.4.5. PrimeEnergyIT

Participants: Laurent Lefèvre, Dias de Assuncao Marcos.

The increasing use of powerful IT services in all public and private service sectors as for example administration, health services and entertainment has lead to a growing energy demand for centralized IT equipment in data centers and central IT units of companies. According to EU and US studies this trend will continue unless energy efficient technology and efficient operation of equipment is broadly implemented. Business-as-usual would lead to a doubling of energy consumption within a few years thereby also significantly increasing energy costs in data centers. The implementation of energy efficient technologies and optimized hardware operation however allows energy and cost savings of up to 60

PrimeEnergyIT european project (Intelligent Energy in Europe : 2010 - 2012)supports the market development and demand for energy efficient central IT hardware and infrastructure providing tools and services for IT and infrastructure managers, consultants and other relevant experts. This project is led by Austrian Energy Agency. The PrimeEnergyIT initiative is operated by an international consortium of national agencies and research institutions in cooperation with a number of associate partners from industry. RESO participates in this project as energy efficiency expert and is involved in metrics design and service based energy efficient solutions.

7.4.6. COST Action IC804 on Energy Efficiency in Large Scale Distributed Systems

Participants: Laurent Lefèvre, Marcos Dias de Assuncao, Anne-Cécile Orgerie, Jean-Patrick Gelas.

The main objective of the Action is to foster original research initiatives addressing energy awareness/saving and to increase the overall impact of European research in the field of energy efficiency in distributed systems. The goal of the Action is to give coherence to the European research agenda in the field, by promoting coordination and encouraging discussions among the individual research groups, sharing of operational know-how (lessons-learned, problems found during practical energy measurements and estimates, ideas for real-world exploitation of energy aware techniques, etc.).The Action objectives can be summarized on scientific and societal points of view: sharing and merging existing practices will lead the Action to propose and disseminate innovative approaches, techniques and algorithms for saving energy while enforcing given Quality of Service (QoS) requirements. Laurent Lefèvre is Management Committee member and French representative in this COST action.

7.4.7. Euro-NF Specific Joint Research Project : Security and Privacy Concerns in energy Efficient Computing

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

SPEC is a Euro-NF specific joint research project (contract #216366) led by university of Passau (Germany) in collaboration with university of Wien (Austria), the CERTH-ITI (Greece) and INRIA (France). This newly formed consortium has agreed to initiate a specific joint research project which will focus on the impact of security and privacy on emerging and evolving ICT energy efficiency mechanisms.

The focused problems are of high importance and span multiple domains of technology, including energy efficiency, virtualization and IT security/privacy, and consequently have a deep impact on the evolution of next generation networks and future Internet.

Project is divided in four work package (WP). While INRIA is involved in each WP, we are leading in particular WP3 entitled Deploying smart meters and threats to user privacy. We rely on our expertise on large scale energy system monitoring (acquired on the Lyon Grid5000 platform).

The results of the project will have a wide impact and it is expected that they will provide first, a joint publication in a high profile conference and a journal paper, and second, provide the basis for one or two of the next EU FP7 calls. The expertise gained in this highly specific and focused project will provide excellent grounds for further, more extended research in this area.

7.5. Visitors

7.5.1. Collaboration with University of California Santa Cruz (UCSC), US

RESO has hosted Pr. Alexandre Brandwajn for 3 months (April-June 2010) as invited researcher to work on the design a new semi-numerical approaches to solve queueing systems. His action also covered some aspects of semantic routing.

7.5.2. Collaboration with Universitat Politècnica de Catalunya (UPC)

RESO has hosted Dr. Monica Aguilar Iguarta and Carolina Tripp, PhD student, for 2 weeks in June 2010 as invited researchers to work on the design of QoS routing solutions in vehicular networks. Isabelle Guérin Lassous has also visited their research group in UPC in June 2010.

7.5.3. Collaboration with University of Thiés (Senegal)

RESO has hosted Dr. Cheikh Sarr for 3 months in (May - July) 2010 as invited researcher to work on the design of available resources evaluation techniques in multihop wireless networks.

7.5.4. Collaboration with University of Tsukuba (Japan)

RESO has hosted Takayuki Imada for 6 weeks (January-February 2010) as invited resarcher to work on the desing of energy efficient cluster based architectures.

8. Dissemination

8.1. Conference organisation, editors for special issues

8.1.1. Editorial Boards

- *Computer Communications*, Elsevier, I. Guérin Lassous
- *Ad Hoc Networks*, Elsevier, I. Guérin Lassous
- *Discrete Mathematics and Theoretical Computer Science*, I. Guérin Lassous
- *EURASIP*, officer of the Local Liaison Board, P. Gonçalves.

8.1.2. Chairing and Organisation of Conferences and Workshops

- Isabelle Guérin Lassous is Program chair of ACM PE-WASUN 2010 (Seventh ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks), October 17-21 2010, in Bodrum, Turkey.
- Laurent Lefevre is :
 - Program vice-chair of GreenCom 2010 : The 2010 IEEE ACM International Conference on Green Computing and Communications, Hangzhou, China, December 18-20, 2010
 - Program chair of Grid 2010 : The 11th IEEE ACM International Conference on Grid Computing, Bruxelles, October 25-29, 2010
 - Publicity chair of The 4th International Conference on Network and System Security (NSS 2010), Melbourne, Australia, September 1-3, 2010.
 - Organizer of the Green Wired Networks Days - COST Action IC804, Lyon June 10-11, 2010
- Jean-Patrick Gelas is co-chair of E2GC2 workshop : Energy Efficient Grids, Clouds and Clusters , during the IEEE Grid2010 conference, Brussels, Belgium, 25-29 Oct 2010

8.1.3. Program committee members

- International Conference on Networking and Computing (ICNC 2010), Laurent Lefevre, 2010
- E2GC2 workshop : Energy Efficient Grids, Clouds and Clusters , during the IEEE Grid2010 conference, Laurent Lefevre, 2010
- 1st International Conference on Energy-Aware High Performance Computing, Laurent Lefevre, 2010
- 4th International Conference on Network and System Security (NSS 2010), Laurent Lefevre, 2010
- VTDC10 - The 4th International Workshop on Virtualization Technologies in Distributed Computing, Laurent Lefèvre, 2010
- ICCS 2010 : International Conference on Computational Science 2010 , Laurent Lefevre, 2010
- CCGrid 2010 : 10th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, Laurent Lefèvre, 2010, 2009
- IEEE Wireless Communications and Networking Conference (WCNC 2010) (Network Track), Laurent Lefevre, 2010
- HPPAC2010: The Sixth Workshop on High-Performance, Power-Aware Computing, Laurent Lefèvre, 2010
- eEnergy 2010 : 1st International Conference on Energy-Efficient Computing and Networking, Laurent Lefevre, 2010
- HotP2P'10 : 7th International Workshop on Hot Topics in Peer to Peer Systems, Laurent Lefèvre, 2010
- Track on Power-Aware Design and Optimization (PADO) , during the 25th ACM Symposium On Applied Computing (SAC), Laurent Lefevre, 2010
- ACOMP 2010 : International Conference on Advanced Computing and Applications, Laurent Lefèvre, 2010
- 8th Australasian Symposium on Parallel and Distributed Computing , Laurent Lefevre, 2010, 2009, 2008
- MedHocNet 2010 (June 23-30, Juan-les-Pins, France), Isabelle Guérin Lassous
- HotMesh 2010 (June 14-17, Montreal, Canada), Isabelle Guérin Lassous
- Algotel 2010 (May 31 - June 3, Belle Dune, France), Isabelle Guérin Lassous
- IEEE PerSens 2010 (March 29-April 2, Mannheim, Germany), Isabelle Guérin Lassous
- IEEE PerCom 2010 (March 29-April 2, Mannheim, Germany), Isabelle Guérin Lassous
- NumBig 2010 (march 31-April 2, Grenoble, France), Paulo Gonçalves

8.1.4. Participation in steering committees

- ACM GridNets, Pascale Vicat-Blanc Primet, since 2006
- PFLDNET workshop serie, P. Vicat-Blanc Primet, since 2005
- IEEE/ACM CCGrid conference, Laurent Lefèvre, since 2004
- ICPS, Laurent Lefèvre since 2006
- Grid, Laurent Lefèvre since 2010

8.1.5. International expertise

8.1.6. National expertise

- Paulo Gonçalves, member of the visiting committee for the AERES evaluation of LTSI (Rennes I).

8.1.7. Public Dissemination

- Laurent Lefèvre - Participant for event : "1000 chercheurs parlent d'avenir - 1000 researchers speak of the future" , book and show on Pantheon walls, Paris, October 18-24, 2010
- Laurent Lefèvre - Interview for paper : "La programmation eco responsable en chantier, Dossier développement durable - Ecodevelopment, Sustainable Development" (in french), 01 Informatique Business et Technologies, Number 2054, pages 44-45, September 30, 2010
- Laurent Lefèvre - Interview for paper : "Le défi de l'informatique verte - Challenges of Green IT" (in french), Magazine La Recherche, Number 444, pages 64-66, September 2010
- INRIA Booth at Supercomputing Conference(SC), Laurent Lefèvre, Marcos Dias de Assuncao, Anne-Cécile Orgerie, Landry Ghislain, 2010.
- Laurent Lefèvre - Interview radio : "L'empreinte écologique de l'Internet" (in french), Magazine EJDG (Ecole de Journalisme) - RCF Isère , March 25, 2010
- Laurent Lefèvre - Interview radio : "Le développement numérique durable" (in french), "Magazine du Développement Durable", Radio Méditerranée Internationale, February 25, 2010
- Laurent Lefèvre - Interview : "Le cout écologique d'Internet - Ecological cost of Internet" (in french), Dossier Pour la Science, Number 66, pages 40-41, January-March 2010

8.2. Animation of the scientific community

- *E2GC2 2010 (Energy Efficient Grids, Clouds and Clusters Workshop)*
Co-Organized by Hiroshi Nakamura, Jean-Marc Pierson and Jean-Patrick Gelas during the IEEE Grid 2010 conference, 25-29 October 2010 in Brussels, Belgium.

+ Paulo Gonçalves is :

- INRIA scientific representative of the EU project SAIL
- INRIA scientific representative of the EU project GEYSERS
- INRIA-RESO scientific representative of the ANR project DMASC
- INRIA-RESO scientific representative of the ANR project PETAFLOW

+ Isabelle Guérin Lassous is :

- leading the Action de Recherche Semantic Networking within the common laboratory INRIA - BellLabs
- the laboratory leader of the INRIA ARC MISSION
- the laboratory leader of the ANR Rescue

+ Laurent Lefèvre is :

- leading the WorkPackage 5 on "Service Deployment" of the FP7 STREP Project "Autonomic Internet"
- the INRIA representative of the OGF-Europe project
- the INRIA scientific representative for the GreenTouch initiative
- Management Committee Member and French representative of the COST Action IC0804 on Energy efficiency in large scale distributed systems.
- Organizer of the Green Wired Networks Days - COST Action IC804, Lyon June 10-11, 2010
- Co-Organizer of the INTECH seminary day on "Green IT, Green by IT and sustainable development", INRIA, Grenoble, June 8, 2010

8.3. Teaching

8.3.1. Graduate teaching

- **since 2008:** P. Gonçalves
Models for Network Traffic
Master 2 (ENS Lyon, Dept. of Computer Science), lecture 24h.
- **since 2009:** P. Gonçalves
Responsible for the axis “Models and Optimization for Emergent Infrastructure” of the ENS Lyon Computer Science Master (Informatique fondamentale)
- **since 2009:** T. Begin
Computer Networks.
Master 1 (Université Claude Bernard Lyon 1), lecture 12h, others 48h.
- **since 2009:** T. Begin
High-Speed and Long-Distance Networks.
Master 2 SIR (Institut de la Fracophonie pour l’Informatique, in Hanoi, Vietnam - Université Claude Bernard Lyon 1), lecture 30h, others 30h.
- **since 2004** JP. Gelas
Long Distance networks ; Networks and Transport Protocols ; 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 SIR, TI, Image and App (Université Claude Bernard Lyon 1), lecture 30h, others 30h.
- **since 2008** JP. Gelas
Introduction to System, Computer Networks and Client/Server architecture.
Master 2 CCI (Université Claude Bernard Lyon 1), lecture 20h, others 30h.
- **since 2010** JP. Gelas
Software development on mobile Apple platforms (iPhone, iPodTouch and iPad).
Master 2 SIR, TI, Image and App (Université Claude Bernard Lyon 1), lecture 6h, others 18h.
- **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.
- I. Guérin Lassous
Multimédia and Quality of Service Master 2 SIR (Professional) / RTS (Research) (University Claude Bernard Lyon I), lecture 18h, others 12h.
- I. Guérin Lassous
Networking
Master 2 CCI (Professional)(University Claude Bernard Lyon I), lecture 18h, others 12h.
- I. Guérin Lassous
Autonomic Computing
Master 2 RTS (Research) University Claude Bernard Lyon I), lecture 15h.

8.3.2. Miscellaneous teaching

- **since 2009:** JP. Gelas
co-manager of the professional master CCI (SIRR) (University Claude Bernard Lyon 1)
- **since 2009:** I. Guérin Lassous is the header of the Networking speciality of the computer science master of University Claude Bernard Lyon 1.
- **since 2009:** T. Begin
Computer Networks.
Licence Informatique, (University Claude Bernard Lyon 1), tutorials and practical works 34h.
- **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.
- **since 2004:** O. Glück
Computer Networks.
Licence Informatique, (University Claude Bernard Lyon 1), lecture 30h, others 30h.
- **since 2006:** I. Guérin Lassous
Ad Hoc Networks Master 1 (University Claude Bernard Lyon I), lecture 6h, others 6h.

8.4. Participation in boards of examiners and committees

- + Olivier Glück is a member of
 - the 27ème section selection committee of University Claude Bernard Lyon 1;
 - the “conseil du département Informatique” of University Claude Bernard Lyon 1;
 - the “conseil de l’UFR Faculté des Sciences et Technologies” 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 CNU Section 27.
- + Isabelle Guérin Lassous is:
 - a member of CNU, section 27;
 - the chair of a selection committee (section 27) of the University Claude Bernard Lyon 1;
 - a member of a 27e section selection committee of University Claude Bernard Lyon 1;
 - a member of a 27e section selection committee of ENSIMAG (Grenoble);
 - a member of the scientific committee of the ResCom group (CNRS ASR);
 - a member of one HdR examining board: Ernesto Exposito (INPT - reviewer);
 - a member of seven PhD examining boards: Benoit Miscopein (INSA Lyon - examiner), Joseph Ramé (Paris-Sud 11 - reviewer), Eryk Schiller (INPG - reviewer), Cédric Gueguen (UPMC Paris 6 - reviewer), Bashir Yaya (UVSQ - examiner), Mohamed Oussama Cherif (UTC - reviewer), Julien Champ (Montpellier II - examiner);
 - a member of the “conseil du département Informatique” of University Claude Bernard Lyon 1.

8.5. Seminars, invited talks

- Isabelle Guérin Lassous gave:
 - a seminar of one week on "QoS issues in heterogeneous networks" at UPC, Barcelona, Spain, June 2010.
- Laurent Lefèvre has been invited to give the following talks :
 - "From Green Grids to Green Clouds : towards energy efficient systems", Laurent Lefèvre, Grids to Clouds day, Lyon, France, December 13, 2010
 - "Energy efficient approach for grids, clouds and networks", Laurent Lefevre, CCGSC 2010 : Clusters and Computational Grids for Scientific Computing Workshop, Flat Rock, North Carolina, USA, September 7-10, 2010
 - Participant in the Green Panel of HPCS2010 conference : "Energy-aware, Power-aware, and Green Computing for Large Distributed Systems and Applications", Caen, June 30, 2010
 - Invited speaker in Green-IT conference : "Technologies vertes de l'information : une réponse aux défis environnementaux" - "Green IT Technologies : an answer to environmental challenges", Lyon, June 9, 2010 evening organized by Enviscope and ITADVICE

9. 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", February 2005, vol. 21, n^o 2, 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 "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, 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", December 2005, vol. 100, n^o 472, 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.

Publications of the year

Doctoral Dissertations and Habilitation Theses

- [9] D. M. DIVAKARAN. *Dealing with Elephants in the Internet: Towards a Flow-Based Internet Architecture*, ENS-Lyon, Université de Lyon, 46, allée d'Italie, 69364 Lyon cedex 07, France, July 2010.
- [10] P. GONÇALVES. *Analyse temps-fréquence et modèles d'invariance d'Échelle : Contribution à l'Étude de systèmes complexes*, École doctorale de mathématiques et d'informatique fondamentale de Lyon, 2010, Habilitation à Diriger des Recherches.

Articles in International Peer-Reviewed Journal

- [11] F. ANHALT, G. KOSLOVSKI, P. VICAT-BLANC PRIMET. *Specifying and provisioning Virtual Infrastructures with HIPerNET*, in "ACM International Journal of Network Management (IJNM) - special issue on Network Virtualization and its Management", 2010.
- [12] N. AYARI, L. LEFÈVRE, D. BARBARON. *Design and Evaluation of a Session-Aware Admission-Control Framework for Improving Service Providers Profitability*, in "Journal of Internet Engineering - Special issue on Service Oriented Infrastructures", December 2010, vol. 4, n^o 1.
- [13] T. BEGIN, A. BRANDWAJN, B. BAYNAT, B. WOLFINGER, S. FDIDA. *High-level Approach to Modeling of Observed System Behavior*, in "Performance Evaluation", May 2010, vol. 67, n^o 5, p. 386 – 405.
- [14] H. CARRÃO, P. GONÇALVES, M. CAETANO. *A nonlinear model for satellite images time series: analysis and prediction of land cover dynamics*, in "IEEE Trans. in Geosciences and Remote Sensing", 2010, vol. 48, n^o 2, p. 1919-1930.
- [15] M. DORET, H. HELGASON, P. ABRY, P. GONÇALVES, C. GHARIB, P. GAUCHERAND. *Multifractal analysis of fetal heart rate variability in fetuses with and without severe acidosis during labor*, in "American Journal of Perinatology", 2010, To appear.
- [16] L. LEFÈVRE, A.-C. ORGERIE. *Designing and Evaluating an Energy Efficient Cloud*, in "The Journal of SuperComputing", March 2010, vol. 51, n^o 3, p. 352-373.
- [17] 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*, in "IEEE/ACM Transactions on Networking", August 2010, vol. 18, n^o 4, p. 1261-1274.
- [18] P. NEIRA AYUSO, R. GASCA, L. LEFÈVRE. *Communicating between the kernel and user-space in Linux using Netlink sockets*, in "Journal Software: Practice and Experience", August 2010, vol. 40, n^o 9, p. 797-810.

- [19] T. TRUONG HUU, G. KOSLOVSKI, F. ANHALT, P. VICAT-BLANC PRIMET, J. MONTAGNAT. *Joint elastic cloud and virtual network framework for application performance optimization and cost reduction*, in "Journal of Grid Computing (JoGC)", 2010.
- [20] M. D. DE ASSUNÇÃO, A. DI COSTANZO, R. BUYYA. *A Cost-Benefit Analysis of Using Cloud Computing to Extend the Capacity of Clusters*, in "Cluster Computing", September 2010, vol. 13, n^o 3, p. 335–347, <http://dx.doi.org/10.1007/s10586-010-0131-x>.

Articles in National Peer-Reviewed Journal

- [21] L. LEFÈVRE. *Le cout écologique d'Internet - Ecological cost of Internet (in french)*, in "Pour la Science", January 2010, vol. 66, p. 40-41.

International Peer-Reviewed Conference/Proceedings

- [22] P. ABRY, H. HELGASON, P. GONÇALVES, E. PEREIRA DE SOUZA NETO, M. DORET, C. GHARIB. *Multifractal analysis for in partum fetal-ECG diagnosis*, in "IEEE Int. Conf. on Acoust. Speech and Sig. Proc.", Dallas, TX (US), March 2010.
- [23] P. ABRY, H. WENDT, S. JAFFARD, H. HELGASON, P. GONÇALVES, E. PEREIRA, C. GHARIB, P. GAUCHERAND, M. DORET. *Methodology for Multifractal Analysis of Heart Rate Variability: From LF/HF Ratio to Wavelet Leaders*, in "32nd Int. Conf. of the IEEE Engineering in Medicine and Biology Society", Buenos Aires (Argentina), Sept. 2010.
- [24] F. ANHALT, D. M. DIVAKARAN, P. VICAT-BLANC PRIMET. *A Virtual Switch Architecture for Hosting Virtual Networks on the Internet*, in "11th International Conference on High Performance Switching and Routing", Dallas, Texas, USA, 6 2010, <http://perso.ens-lyon.fr/fabienne.anhalt/papers/VXB-switch.pdf>.
- [25] G. DA-COSTA, M. DIAS DE ASSUNCAO, J.-P. GELAS, Y. GEORGIU, L. LEFÈVRE, A.-C. ORGERIE, J.-M. PIERSON, O. RICHARD, A. SAYAH. *Multi-facet approach to reduce energy consumption in clouds and grids: The GREEN-NET Framework*, in "e-Energy 2010 : First International Conference on Energy-Efficient Computing and Networking", Passau, Germany, April 2010.
- [26] M. DIAS DE ASSUNCAO, J.-P. GELAS, L. LEFÈVRE, A.-C. ORGERIE. *The Green Grid5000: Instrumenting a Grid with Energy Sensors*, in "5th International Workshop on Distributed Cooperative Laboratories: Instrumenting the Grid (INGRID 2010)", Poznan, Poland, May 2010.
- [27] M. DIAS DE ASSUNCAO, A.-C. ORGERIE, L. LEFÈVRE. *An Analysis of Power Consumption Logs from a Monitored Grid Site*, in "IEEE/ACM International Conference on Green Computing and Communications (GreenCom-2010)", Hangzhou, China, December 2010.
- [28] D. M. DIVAKARAN, F. ANHALT, E. ALTMAN, P. VICAT-BLANC PRIMET. *Size-Based Flow Scheduling in a CICQ Switch*, in "11th International Conference on High Performance Switching and Routing (HPSR 2010)", Dallas, Texas, USA, June 2010, p. 140–145.
- [29] D. M. DIVAKARAN, G. CAROFIGLIO, E. ALTMAN, P. VICAT-BLANC PRIMET. *A Flow Scheduler Architecture*, in "IFIP/TC6 NETWORKING 2010", Chennai, India, May 2010, p. 122–134.
- [30] G. FEDAK, J.-P. GELAS, T. HÉRAULT, V. INIESTA, D. KONDO, L. LEFÈVRE, P. MALECOT, L. NUSSBAUM, A. REZMERITA, O. RICHARD. *DSL-Lab: a Low-Power Lightweight Platform to Experiment on Domestic*

Broadband Internet, in "International Symposium on Parallel and Distributed Computing (ISPDC 2010)", Istanbul, Turkey, July 2010.

- [31] O. GRÉMILLET, P. GONÇALVES, P. VICAT-BLANC PRIMET, A. DUPAS. *Traffic classification techniques supporting semantic networks*, in "1st Int. Wireless Comm. and Mobile Comp. Conf., TRaffic Analysis and Classification W.", Caen (France), Jun-jul 2010.
- [32] P. LOISEAU, P. GONÇALVES, J. BARRAL, P. VICAT-BLANC PRIMET. *Modeling TCP Throughput: an Elaborated Large-Deviations-Based Model and its Empirical Validation*, in "IFIP, Performance", Namur (Belgium), Nov. 2010.
- [33] A.-C. ORGERIE, L. LEFÈVRE, J.-P. GELAS. *Demystifying Energy Consumption in Grids and Clouds*, in "The Work in Progress in Green Computing (WIPGC) Workshop, in conjunction with the first IEEE sponsored International Green Computing Conference", Chicago, USA, August 2010.
- [34] J. RUBIO-LOYOLA, A. ASTORGA, J. SERRAT, W. K. CHAI, L. MAMATAS, A. GALIS, S. CLAYMAN, A. CHENIOUR, L. LEFÈVRE, O. MORNARD, A. FISCHER, A. PALER, H. DE MEER. *Platforms and Software Systems for an Autonomic Internet*, in "IEEE Globecom 2010 - Next Generation Networking Symposium", Miami, USA, December 2010.
- [35] J. RUBIO-LOYOLA, A. ASTORGA, J. SERRAT, L. LEFÈVRE, A. CHENIOUR, D. MULDOWNNEY, S. DAVY, A. GALIS, L. MAMATAS, S. CLAYMAN, D. MACEDO, Z. MOVAHEDI, G. PUJOLLE, A. FISCHER, H. DE MEER. *Manageability of Future Internet Virtual Networks from a Practical Viewpoint*, in "FIA Book : "Towards the Future Internet - Emerging Trends from European Research"", Valencia, April 2010.

National Peer-Reviewed Conference/Proceedings

- [36] T. BEGIN, A. BRANDWAJN. *Note sur la simulation d'une file M/G/1 selon la distribution du temps de service*, in "Proceedings of the 12th Rencontres Francophones sur les Aspects Algorithmiques de Télécommunications, ALGOTEL10", Belle Dune, France, June 2010, 4.

Workshops without Proceedings

- [37] G. KOSLOVSKI, W.-L. YEOW, C. WESTPHAL, T. TRUONG HUU, J. MONTAGNAT, P. VICAT-BLANC PRIMET. *Reliability Support in Virtual Infrastructures*, in "2nd IEEE Int. Conf. on Cloud Computing Technology and Science (CloudCom 2010)", IEEE, 2010.

Scientific Books (or Scientific Book chapters)

- [38] Z. HUANG, J. HINE, L. LEFÈVRE, T. MCGREGOR, Y. PAN, H. SHEN, Z. XU. *Special Issue on "Emerging Research in Parallel and Distributed Computing" - Journal of Supercomputing*, Journal of Supercomputing - Springer Netherlands eds., March 2010, vol. 51, n^o 3.
- [39] A.-C. ORGERIE, M. DIAS DE ASSUNCAO, L. LEFÈVRE. *Chapter on book "Grids, Clouds and Virtualization" - Massimo Cafaro and Giovanni Aloisio editors*, in "Energy Aware Clouds", Springer Book, October 2010, p. 145-170, ISBN : 978-0-85729-048-9.

Research Reports

- [40] F. ANHALT, C. BLANCHET, R. GUILLIER, T. TRUONG HUU, G. KOSLOVSKI, J. MONTAGNAT, P. VICAT-BLANC PRIMET, V. ROCA. *HIPCAL: final report*, INRIA, 2010.

- [41] A. BRANDWAJN, T. BEGIN. *A Recurrent Solution of Ph/M/c/N-like and Ph/M/c-like Queues*, INRIA, 2010, n^o RR-7321, <http://hal.archives-ouvertes.fr/inria-00492748/en/>.
- [42] A. CHENIOUR, J.-P. GELAS, L. LEFÈVRE, L. MAMATAS, O. MORNARD, J. RUBIO-LOYOLA. *Autonomic Internet AutoI Service Plane and Interfaces - Deliverable D5.2*, AUTOI European Project, December 2010.
- [43] D. M. DIVAKARAN, F. ANHALT, E. ALTMAN, P. VICAT-BLANC PRIMET. *Size-Based Flow Scheduling in a CICQ Switch*, INRIA, 2010, n^o RR-7183, <http://hal.archives-ouvertes.fr/inria-00450054/en/>.
- [44] A.-C. ORGERIE, L. LEFÈVRE. *A year in the life of a large-scale experimental distributed system: usage of the Grid'5000 platform in 2008*, INRIA, December 2010, n^o 7481, <http://hal.inria.fr/inria-00544992/en/>.
- [45] P. VICAT-BLANC PRIMET, G. KOSLOVSKI, F. ANHALT, T. TRUONG HUU, J. MONTAGNAT. *Exploring the Virtual Infrastructure as a Service concept with HIPerNet*, INRIA, 2010, n^o RR-7185.

Other Publications

- [46] M. DIAS DE ASSUNCAO, L. LEFÈVRE, A.-C. ORGERIE. *Investigating the Energy Consumption of Virtual Machines*, April 2010, Poster in E-Energy 2010 conference : The First International Conference on Energy-Efficient Computing and Networking, Passau, Germany.
- [47] M. DIAS DE ASSUNCAO, L. LEFÈVRE, A.-C. ORGERIE. *Monitoring, Collecting and Exposing Energy Usage of Large Scale Distributed Infrastructures*, April 2010, Poster during Grid5000 Spring School, Lille, France.
- [48] M. DIAS DE ASSUNCAO, L. LEFÈVRE, A.-C. ORGERIE. *Monitoring, Collecting and Exposing Energy Usage of Large Scale Distributed Infrastructures*, April 2010, Poster in E-Energy 2010 conference : The First International Conference on Energy-Efficient Computing and Networking, Passau, Germany.
- [49] A. GALIS, S. CLAYMAN, A. FISCHER, A. PALER, Y. AL-HAZMI, H. DE MEER, J.-P. GELAS, L. LEFÈVRE, O. MORNARD, A. CHENIOUR, A. ASTORGA, J. SERRAT, J. LOYOLA, S. DAVY. *Future Internet Management Platforms for Network Virtualisation and Service Clouds*, December 2010, Service Wave 2010 Demonstration.

References in notes

- [50] N. AYARI, D. BARBARON, L. LEFÈVRE. *Procédés de gestion de sessions multi-flux*. France Telecom R&D Patent, June 2007.
- [51] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *Implementation of an Active Replication based Framework for Highly Available Services*, September 2007, NetFilter Workshop 2007, Karlsruhe, Germany.
- [52] 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.
- [53] 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.

-
- [54] 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.
- [55] S. FLOYD, V. JACOBSON. *Link-sharing and Resource Management Models for Packet Networks*, in "IEEE/ACM Transaction on Networking", 4, August 1995, vol. 3.
- [56] I. FOSTER, C. KESSELMAN. *The Grid : Blueprint for a new Computing Infrastructure*, in "Morgan Kaufmann Publishers Inc.", 1998.
- [57] L. HABLOT, O. GLÜCK, J.-C. MIGNOT, R. GUILLIER, S. SOUDAN, P. VICAT-BLANC PRIMET. *Interaction between MPI and TCP in grids*, INRIA, 06 2009, n^o 6945, <http://hal.inria.fr/inria-00389836/>.
- [58] 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*, in "IEEE/ACM Transactions on Networking", December 2009, to appear, http://perso.ens-lyon.fr/patrick.loiseau/articles/Invest_SS_and_HT.pdf.
- [59] P. LOISEAU, P. GONÇALVES, S. GIRARD, F. FORBES, P. VICAT-BLANC PRIMET. *Maximum Likelihood Estimation of the Flow Size Distribution Tail Index from Sampled Packet Data*, in "ACM Sigmetrics", June 2009, <http://portal.acm.org/citation.cfm?id=1555349.1555380>.
- [60] 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*, in "TridentCom", April 2009.
- [61] V. SANDER. *Networking issues of GRID Infrastructures*, in "GRID Working Draft of the GRID High-Performance Networking Research Group, Global GRID Forum", 2003.
- [62] V. SANDER, F. TRAVOSTINO, J. CROWCROFT, P. VICAT-BLANC PRIMET, C. PHAM. *Networking Issues of Grid Infrastructures*, Open Grid Forum, october 2004, <http://forge.gridforum.org/projects/ghpn-rg/>.
- [63] D. SIMEONIDOU. *Optical Network Infrastructure for Grid*, in "Grid Working Draft of the Grid High-Performance Networking Research Group, Global GRID Forum", 2003.