INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team reso

Optimized protocols and software for high performance networks

Rhône-Alpes

THEME NUM

Table of contents

1. Team

Head of project-team
Pascale Vicat-Blanc Primet [Research Director (DR2) Inria, HdR]

Administrative Assistant
Danielle Bianchetti [Secretary ENS – 20%]
Sylvie Boyer [Secretary (SAR) INRIA – 10%]

Staff member INRIA
Paulo Gonçalves [Research Associate (CR1) Inria – with RESO since April 2006]
Laurent Lefèvre [Research Associate (CR1) Inria]

Staff member Université Claude Bernard Lyon1 (UCB)
Jean-Patrick Gelas [Maître de conférences – with RESO since September 2006]
Olivier Glück [Maître de conférences]
Isabelle Guérin-Lassous [Professeur – with RESO since September 2006, HdR]

Project technical staff
Pierre Bozonnet [Expert Engineer INRIA - RNRT Temic Project up to 31/7/2006 - Expert Engineer Alcatel-INRIA - 2006/2007]
Martine Chaudier [Expert Engineer INRIA - RNRT Temic Project up to 15/3/2006]
Jean-Christophe Mignot [Research Engineer CNRS – 40%]
Aurélien Cedeyn [ENS Engineer – Grid5000]

Ph. D. students
Narjess Ayari [PhD student with France Telecom R-D - CIFRE - 2005/2008]
Romaric Guillier [PhD student, ENS - 2006/2009]
Ludovic Hablot [PhD student, MENRT - 2006/2009]
Patrick Loiseau [PhD student, ENS - 2006/2009]
Dino Lopez Pacheco [PhD student - 2004/2007 - Mexican Goverment Grant]
Sébastien Soudan [PhD student, MENRT - 2006/2009]
Rémi Vanier [PhD student, INRIA - 2006/2009]

Student internship
Bin Bin Chen [Internship INRIA]
Damien Nicolet [Internship ENS]
Pablo Pazos Rey [Internship ENS]
Anne-Cécile Orgerie [Internship ENS]
Magi Sanchon [Internship ENS]
Rima Ben Mobesh [Internship ENS]

External collaborator
Loic Prylli [Myricom]

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. RESO is part of the " Numerical Systems " theme of the INRIA, part of the B subsection: Grids and high-performance computing. The research activities of the RESO project fits the first priority challenge of the INRIA’s strategic plan: "design and master the future network infrastructures and communication services platforms" . In this direction, RESO is focusing on communication software, services and protocols in the context of high performance short and long distance networking and applying its results to the domain of Grids.

2.2. Context

Wavelengths multiplexing and future wavelengths switching techniques on optical fibers allow core network infrastructures to rapidly improve their throughput and reliability. In a near future, access links of ten gigabits per second will be made available. New technologies like 10Gigabit/s Ethernet or 10Gigabit/s Myrinet is also driving the increase of bandwidth in local area networks. These improvements have given the opportunity to create high performance distributed systems called "computational and data grids" that aggregate storage and computation resources into a virtual and integrated computing environment. Grid computing is a promising technology harnessing distributed resources into virtual organizations for the future resource intensive scientific and business applications.

On the other hand, the volumes of heterogeneous data that are produced by various distributed sources (sensor networks but also sophisticated instruments) and managed in distributed data centers are rapidly increasing. Complex model computations performed on supercomputers produce petabytes of data, which have to be accessed and analysed by various user groups. Moving such enormous quantities of data among grid elements and ensuring efficient message passing between communicating processes raise specific challenges on the communication protocols and their related mechanisms. Grid technology can provide valuable help in this area if it provides a secure, flexible and transparent and high performance transport infrastructure for data access and processing. Consequently, future high-speed optical networks are addressed not only to support the accelerating and dynamic growth of data traffic but also the new emerging network requirements such as fast and flexible provisioning, QoS levels, and fast recovery procedures of such intensive computing applications. Enabling ultra high performance machine to machine communications lead then to new bandwidth sharing paradigms that have to be carefully examine.

Although grids theoretically offer solutions for resources aggregation, predictable and high performance for applications may be hard to obtain due to the improperness of communication protocols and software and to the fact that processors, memory, bus and disc speeds, involved into the protocol processing chain do no scale with network speeds. In order to deliver grid traffic in a timely, efficient, and reliable manner over long distance networks, several issues such as quality of service, security, traffic metrology and network resource scheduling have to be investigated.

2.3. Research area

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

- Optimized software architectures for efficient communications in end systems, cluster-based servers and programmable access equipments.
- Protocols and algorithms for efficient and customizable transport of heterogeneous traffic at high-speed.

Recently, a new research direction complementing the first ones has been introduced:

- Grid traffic metrology and statistical inference.

The first research axis explores how communication subsystems in end systems, in cluster networks and programmable access equipments can be enhanced and optimized. Our researches focus on high performance software solutions for clusters, new active network solutions for IP networks and interconnection of IP networks, networks of clusters or networks of data storage. We search at optimizing both data movements and I/O management that are closely inter-dependant, by using the intelligence of network interface cards (NICs).

The second research axis explores the problem of efficient transfer of heterogeneous flows in a high performance and high speed long distance networking infrastructure. The scientific directions we follow concern the study of flexible solutions exploiting innovating networking services in routers and the addition of packet processing software components at the edge of the core network for controlling the flows. Problems to be solved are modeling and quantifying the influence of the different performance parameters on a transport connection

and the end-to-end characterization of the network links with variable network services, the design of adaptive algorithms dedicated to the expressed flow needs, definition and introduction in the network of end-to-end protocol-oriented mechanisms, making the interaction between packet processing and forwarding smooth and efficient.

The third axis was recently brought into the RESO team activities (April 2006) and deals with the metrological aspects of grids' traffic. Intended as a diagnosis tool to serve our second research axis, we should be led to first define pertinent metrics to get an instantaneous snapshot of the network, and to assess the corresponding quality of service. In a second step, we foresee to develop methodological approaches, along with the necessary tools, to fit the measures with realiable (widesense) statistical models. Ultimately, based on these models, we plan to infer a short time forecast of the network capacities (instantaneous bandwidth, latency, losses,...), which will steer the current transport protocol to automatically adapt to the context.

2.4. Application domains

RESO applies its research to the domains of high performance computing and to Grid communications. Grid computing is a promising technology that brings together large collection of geographically distributed resources (e.g., computing, storage, visualization, etc.) to build on demand very high performance computing environments for compute and data-intensive applications. These large scale cybernetic infrastructures gain increasing attention from a broad range of actors: from research communities to computer providers, large companies, and telecommunication operators (telcos). Whereas grids have been widely in use in the scientific community, they are now on the verge of moving into the commercial environment. American, european and japanese telcos plan to move forward grid computing. Different scenarios for telcos can be envisioned: telcos may (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. Which senario will be developed remains an open question.

Researches conducted these last years reveal that grid technology raise new challenges in terms of network optimisation as well as of protocol architecture and of transport paradigms. A broad deployment of the grid technology can modify and influence the design of the future Internet as other emerging communicating applications.

The geographical topology of the Grid depends on the distribution of the community members. Though there might be a strong relation between the entities building a virtual organization, a Grid still consists of resources owned by different, typically independent organizations. Heterogeneity of resources and policies is a fundamental result of this. Web Service Resource Framework (WS-RF), the coupling of the notion of resource and web service, has recently been introduced. It has added the performance issue to the re-usability, interoperability, and openness advantages of web services. Grid services involve operations and strategies from application layer down to network layer, with service agreements defined at application layer and middleware developed for the communication between layers. In a typical implementation scenario, the grid middleware provisions the resource, and passes the delivery criteria to the network services. The network, accordingly, follows up to enforce the appropriate data transfer. In a Grid, the network performance requirements are very high and may strongly influence the performance of the whole distributed system. The construction of grid networks over the optical transport layer tackles the problem of communication performance from the transport medium perspective. However, our vision is that Grid applications, due to the heterogeneity and large scale factors, will continue to use traditional IP packet protocols, at least in the end systems and will rely on a complex interconnection of heterogeneous networks. In such context end-to-end flow performance is difficult to guarantee or predict. Thus, for achieving end-to-end QoS objectives, the remaining deficiencies of the network performance have to be masked by adaptation performed at the host level or somewhere in the datapath. RESO designs Grid network services and network middleware to keep applications network-transparent as much as possible, to simplify the programming and to optimize the execution of their communication parts while fully exploiting the capacities of the evolving networking infrastructure.

2.5. Methodology

The RESO approach relies on the theoretical and experimental analysis of limitations encountered in existing protocols and on the theoretical and experimental exploration of new approaches. This research framework between a challenging application domain and a specific network context, induces a close interaction with both the application level and the underlying network level. The methodology is based on a study of the high end and original requirements and on experimental evaluation of the functionalities and performance of high speed infrastructures. RESO gathers expertise in advanced high performance local and cluster area networks protocols, in distributed systems, in long distance networking, in time series analysis and in statistical inference. This background work provides the context model for innovative and adequate protocols and software design and evaluation. Moreover, the propositions are implemented and experimented on real or emulated local or wide area testbeds with real conditions and large scale applications.

2.6. Goals

RESO aims at providing software solutions for high performance and flexible communications fully exploiting the very high speed networking infrastructure of computational and data grids. The goal of our research is to provide analysis of the limitations of the current communication software and protocols designed for standard networks and traditional usages, and to propose optimization and control mechanisms for the end-to-end performance and quality of service. RESO explores original and innovative end-to-end transport services and protocols that meet the needs of grid applications. These solutions must scale in increasing bandwidths, heterogeneity and number of flows.

RESO studies high speed network characteristics, grid application 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 and networking equipments and to the dissemination of new approaches.

2.7. Summary of the main contributions of the team in 2006

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

2.7.1. Direction 1: Optimized communication software and equipments

- Study and optimizations of Message Passing Interface implementations for Grid platforms : MPICH-Madeleine, GridMPI, OpenMPI. The experiments were conducted on the national GRID'5000 testbed.
- Design and development of a high performance autonomic network node adapted to industrial context (IAN2). Proposition of adapted autonomic services;
- Validation of Tamanoir environment through internal and external projects (3DDL) ;
- Design of a fault tolerant and highly available architecture for clustered network equipments;
- Exploration of the Network Processor technology for designing high performance grid overlay gateways;
- Investigation of the behavior of Gigabit Ethernet switches in heavy congestion scenarios;

2.7.2. Direction 2: End-to-end transport and service differentiation

- Contribution to the analysis of the limitation of TCP protocol for very long distance high speed networks within the eponymous pfldnet community;
- Evaluation of new TCP variant in 10Gb/s real and emulated environment with a range of latencies;
- Evaluation and proposition around the XCP High Performance transport protocols;
- Optimization algorithms for network resource sharing and flow scheduling for delay constrainted bulk data transfers in high speed networks;

- Study and design of a high performance router assisted inter-operable transport protocol (XCP-i).

2.7.3. Grid Network services and applications

- Contribution to the design and development of the GRID5000, national Grid testbed and its international interconnections to Netherland (DAS3) and Japan (Naregi) in collaboration with RENATER;
- Analysis of the properties and behavior of the Grid5000's 10Gb/s network infrastructure;
- Starting an activity on flow-level Grid Network performance and traffic monitoring in Grid5000;
- Contribution to the development, extension and usage of the high speed network in GridExplorer, national Grid emulator;
- Emulation instrument design and development for high performance protocols and Grid software evaluation. We designed, developed and evaluated the eWAN software for transforming a cluster in a virtual high performance grid network cloud;
- Integration of Web Services and programmable networks for improving flexibility of Active Grids.

3. Scientific Foundations

3.1. Optimized communication software and equipments

Participants: Narjess Ayari, Pierre Bozonnet, Martine Chaudier, Jean-Patrick Gelas, Olivier Glück, Laurent Lefèvre, Pascale Vicat-Blanc Primet, Jean-Christophe Mignot, Ludovic Hablot, Sébastien Soudan.

The emergence of high performance parallel applications has raised the need of low latency and high bandwidth communications. Massively parallel supercomputers provided integrated communication hardware to exchange data between the memory of different nodes. They are now often replaced by clusters of workstations based on high-speed interconnects such as MYRINET or INFINIBAND which are more generic, more extensive, less expensive and where communications are processed by dedicated network interfaces. A large amount of interesting work has been done to improve communications between cluster nodes at the application level through the use of the advanced features in the network interface card and *OS-bypass* techniques. Meanwhile, storage access needs to reach similar performance to read input data and store output data on a remote node without being the bottleneck. In a cluster environment, high performance applications running on high-speed interconnects require both efficient communication between computing nodes and fast access to the storage system. 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 how a MPI application can benefit, during one execution, of several high-speed networks at the same time. In particular, it implies to find a way to communicate efficiently between these heterogenous interconnections. We also explore how to keep good performance execution when long-distance communications are necessary because the application is launched on multiple sites of the grid.

In this research axis, we explore the design of autonomic network equipments able to dynamically deploy adapted services. These equipments have been used in industrial context (TEMIC project, 3DDL collaboration). In order to support network functions in the embedded equipments, we propose a high performance autonomic network environment execution architecture ($\text{Tamanoir}_{\text{embedded}}$ software suite). High availability, fault tolerance and scalability issues of cluster-based network equipments have been and are currently explored.

3.2. End-to-end High performance transport

Participants: Pascale Vicat-Blanc Primet, Dino Lopez Pacheco, Laurent Lefèvre, Sébastien Soudan, Romaric Guillier.

In TCP/IP networks, the end-to-end principle aims at simplifying the network level while pushing all the complexity on the end host level. This principle has been proved to be very valuable in the context of the traditional low capacity Internet. In packet networking, congestion events are the natural counterpart of the flexibility to interconnect mismatched elements and freely multiplex flows. Managing congestion in packet networks is a very complex issue. This is especially true in IP networks where, at best, congestion information is very limited (e.g., **ECN**) or, at worst, non-existent, forcing the transmitter to infer it instead (e.g., based on losses or delay) in TCP.

The conservative behavior of TCP with respect to congestion in IP networks (RFC 2581) is at the heart of the current performance issues faced by the high-performance networking community. Several theoretical and experimental analysis have shown that the dynamics of the traditional feedback based approach is too low in very high speed networks that may lose packets. Consequently network resource utilization is not optimal and the application performance is poor and disappointing. Many Grid-enabled computing applications wish to transfer large volumes of data over wide area networks and require high data rates in order to do so. However, Grid-enabled applications are rarely able to take full advantage of the high-capacity (2.5 Gbit/s, 10 Gbit/s and upwards) networks installed today. Recent data for Internet 2 show that 90% of the bulk TCP flows (defined as transfers of at least 10 Megabyte of data) use less than 5 Mbit/s, and that 99% use less than 20 Mbit/s out of the possible 622 Mbit/s provision. There are many reasons for such poor performances. Many of the problems are directly related to the end system, to the processor and bus speed, and to the NIC with its associated driver. TCP configuration (e.g., small buffer space or features such as SACK being improperly negotiated) will have a significant impact. TCP itself was designed first and foremost to be robust and when congestion is detected, TCP accommodates the problem but at the expense of reduced performance. There are also design problems with TCP itself. For example, for a standard TCP connection with 1500-byte packets and a 100 ms round-trip time, achieving a steady-state throughput of 10 Gbit/s would require an average congestion window of 83,333 segments, and a packet drop rate of at most one congestion event every 5,000,000,000 packet (or equivalently, at most one congestion event every 1 2/3 hours). HighSpeed TCP [58] and Scalable TCP [64] increase the aggressiveness in high-throughput situations while staying fair to standard TCP flows in legacy contexts. **FAST** [63] leverages the queueing information provided by round-trip time variations, in order to efficiently control buffering in routers and manage IP congestion optimally. These propositions are actively analyzed and experimented by the international community. Several issues have been already highlighted. Considering the traditional feedback loop will not scale with higher rate level under loss or congesting traffic conditions, it seems judicious to start examining alternative radical solution for end-to-end transport as well as for congestion control. These solutions can be based on pair to pair approaches, buffer in line or flow scheduling, fully exploiting not only the rate dimension of data transfer but also space, time and cross-layer parameters.

One of the direction recently investigated in Grids, is the capacity of dynamically establishing overprovisionned dedicated lambda path. The optical fiber communication will be the predominant mechanism for data transmission in the core. To address the anticipated terabit demands dynamically reconfigurable optical networks are envisionned. 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, equipped with tunable transceivers, optical crossconnects (OXC), and optical add/drop multiplexers. At the opposite side of the spectrum, dedicated high bandwidth channels are critical in large scale applications to ensure timely task completion, which in turn necessitates a high-performance control plane capable of scheduling such channels in advance. The control-plane, traditionally in the hand of telco will migrate to the users. Optical Cross-Connects (OXCs) becomes more and more, cheap, simple and controllable. Prototyping and studying the interactions of components required to accomplish the tasks of user-specified bandwidth reservation, path computation and network signaling is of importance. RESO starts to integrate this new technological perspective to understand how this optical component interacts with electronic component and how to configure, control and tune them with end computers.

Finally an other important issue is flow differenciation. Indeed, it is known that flows crossing IP networks are not equally sensitive to loss or delay variations. Since several years, research effort has been spent to solve the problem of the heterogeneous performance needs of the IP traffic. A class of solutions considers

that the IP layer should provide more sophisticated services than the simple best-effort service to meet the application's quality of service requirements. Quality of service has been studied in IP networks in the context of multimedia applications [57]. RESO explores various complementary or fundamentally different solutions to carry end-to-end quality of service to grid applications to ensure an efficient usage of the interconnected computing resources [60].

3.3. Metrology and Statistical inference on grids' traffic

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

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, time series analysis statistical inference, and stochastic processes, deemed important to embed in the main research contours of RESO. Our goal is for this analysis to become in the near future a plain component not only in the study and in the development of infrastructures and computing grids, but also in network resources modeling and in the static and dynamic properties automatic characterization.

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

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

To drive us in this task, we will rely on the impressive amount of studies devoted to the internet traffic analysis, and on the established results that have been obtained in the last years [68], [72], [50], [51]. For instance, we are interested in verifying if the conjecture relating long range dependance (LRD) in traffic flows with heavy tailed distributions of files sizes [72], still holds with grid networks. In the course, we will necessarily face the issue of reliable estimation of density functions from loosely sampled data, and even more crucially this of accurate estimation of LRD parameters and tail exponents from incomplete data sets [3]. No doubt that plural competences in time series analysis, statistics, and stochastic processes are required in order to complete this task successfully. That is why, RESO decided to broaden up the scope of its actual discipline with the venue of P. Gonçalves (former Mistis (ex IS2) project member) and the PhD work of P. Loiseau (Ms in Physics) initiated in september 2006.

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

3.4. Grid Network services and applications

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

The purpose of Computational Grids is 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 [61]. But generally, the underlying communication infrastructure of these large scale distributed environments is a complex interconnection of multi-IP domains with changing performance characteristics. Consequently *the Grid Network cloud* may exhibit extreme heterogeneity in performance and reliability that can considerably affect the global application performance. Performance and security are the major issues grids encountered from a technical point of view.

The performance problem of the grid network cloud can be studied from different but complementary view points. All these approaches are valuable and will fit the grid network services middleware framework under definition stage at GGF.

- 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.
- 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.
- Modeling, managing and controlling the grid network resource as a first class resource of the global environment: transfer scheduling, data movement balancing,...
- Advance reservations.

4. Application Domains

4.1. Panorama

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

RESO applies its research to the domains of high performance Cluster and Grid communications. Existing GRID applications did already identify potential networking bottlenecks, either caused by conceptual or implementation specific problems, or missing service capabilities. We participated to the elaboration of the first GGF document on this subject [70] [69], [71]. Loss probability, important and incompressible latencies, dynamic behavior of network paths question profoundly models and technic used in parallel and distributed computing [59]. 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. We argue that the first bottleneck is 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 GGF, the European or National projects. These activities have been done in close collaboration with other INRIA and CNRS French teams (Grand Large, Apache, Graal) involved in the GRID5000 and the Grid Explorer projects and other European teams involved in pfldnet and Glif communities.

- We continue the investigation of limits of the existing communication services or protocols and evaluate more efficient approaches within the Grid5000 national experimental infrastructure based on the RENATER network. Participating to the design, deployment and usage of such high performance experimental Grid testbed allows us to evaluate and measure the benefit that grid middleware and applications can get from enhanced networking technologies. The experience and expertize we get from this work are a tremendous gain for our research on performance bottlenecks.

- Grid 5000 is a national initiative aiming at providing a huge experimental instrument to the grid software research community. Lyon, with RESO and GRAAL projects, is part of this initiative. RESO is closely involved in the design and deployment of the testbed, and responsible for the networking aspects.
- We participate to the definition of the Grid Explorer physical architecture and to the design of the configuring, tuning and monitoring software. Grid Explorer, the largest cluster of Grid 5000 platform will be a very large scale instrument for grid software evaluation.

5. Software

5.1. EWAN: High Performance Network Emulation

Keywords: *eWAN*, *grid networking*, *network emulation*.

Participants: Magi Sanchon, Olivier Glück, Pascale Primet.

The Grid aims at expanding the cluster based parallel computing paradigm towards large scale distributed systems based on IP networks. EWAN [75] [34] is a high performance network environment emulator. It takes place in the research effort on computer grids, aggregations of computer resources inter-connected by a wide area network. EWAN offers an emulation framework needed by experiments in this field, bringing a great flexibility, a high level of performance and a precise control. EWAN provides features to control key characteristics of grid or transport protocol evaluation scenarios. To achieve correct performance and enable test at gigabit speed with minimum noise and overhead, the different functional entities are deployed on separate, non shared and reserved machines and local networks. Compared to Emulab, the particularity of EWAN is to exploit within a limited time (one to few ours depending on the experiment needs) any cluster composed of several tens or more common PCs. The main functionnalities that have been identified are:

- link emulation with key characteristics control: like latency (from 1ms to 500ms) loss rate (with different distributions), capacities (from (10Mb/s to 10Gb/s);
- topology emulation (chain, star, ring, mesh, dumpbell, fish bone...);
- IP version (v4 and v6) and jumbo frame support;
- traffic generation;
- process application running;
- traffic and performance monitoring and logging.

The eWAN software can be divided into two main parts: an interface for creation of simple topologies and an engine for deploying every sort of topologies. We have evaluated several network emulation solutions and have configured a 12 nodes cluster to test EWAN software on this cluster. As emulation solutions, we have compared Nistnet, netem and GtrcNET. Nistnet is a well known software to emulate network link, netem is an equivalent recently included in the Linux kernel and GtrcNET is an hardware network emulator developed by the AIST. EWAN manages all the three solutions and allows the user to choose one of them. All material (source code, documentation, results, ...) about EWAN can be found at http://www.ens-lyon.fr/LIP/RESO/Software/EWAN/. EWAN is also distributed in the DataGridExplorer (ACI MD) project. A paper written in collaboration with Tomohiro Kudoh and Yuestu Kodama (AIST GTRC) has been published in the proceedings of the International Pfldnet2006 conference. [34]

5.2. SNE (Stateful Network Equipment)

Keywords: *High Availability*, *fault tolerance*.

Participant: Laurent Lefèvre (contact).

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

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

5.3. Tamanoir^{embedded} (Active execution environment for embedded autonomic network equipments)

Keywords: *autonomic networking*, *programmable network equipments*.

Participants: Martine Chaudier, Jean-Patrick Gelas (contact), Laurent Lefèvre.

We designed an Execution Environment called $Tamanoir^{embedded}$ based on the Tamanoir software suite. The original Tamanoir version is a prototype software with features too complex for an industrial purpose (cluster-based approach, Linux modules, multi-level services...).

Due to some typical industrial constraints (e.g code maintenance), we reduced the code complexity and removed all unused classes and methods or actually useless for this project. It allows us to reduce the overall size of the software suite and make the maintenance and improvement of the code easier for service developers.

$Tamanoir^{embedded}$ is a dedicated software platform fully written in Java and suitable for heterogeneous services. Tamanoir provides various methods for dynamic service deployment. $Tamanoir^{embedded}$ also supports autonomic deployment and services updating through mobile equipments. Inside automatic maintenance projects, we deploy wireless based IAN^2 (Industrial Autonomic Network Node) nodes in remote industrial environments (no wire connections available) [65]. In order to download maintenance information, human agents can come near IAN^2 nodes to request informations. During this step, mobile equipments (PDA, Tablets, cellulars) are also used as mobile repositories to push new services and software inside autonomic nodes.

Tamanoir is an open source software suite, available on the web and protected by APP (Agence Francaise de Protection des Programmes).

5.4. LSCAN (Large Scale Deployment of Autonomic Networks)

Keywords: *Grid5000*, *Nagios*, *autonomic networks*.

Participants: Pablo Pazos Rey, Laurent Lefèvre (contact).

In the context of the Grid5000 project, we have developped the LSCAN software suite which is dedicated to the deployment (graphical), management (through web services) and experiment of large scale deployment of autonomic network nodes. LSCAN is an open source project, adapted from the Nagios software suite and is available on the web.

5.5. XCP-i (Interoperable eXplicit Control Protocol)

Keywords: *XCP*, *high performance transport protocol*.

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

XCP (eXplicit Control Protocol) is a transport protocol that uses the assistance of specialized routers to very accurately determine the available bandwidth along the path from the source to the destination. We propose XCP-i which is operable on an internetwork consisting of XCP routers and traditional IP routers without loosing the benefit of the XCP control laws [24], [25], [42], [41]. An ns-2 module simulating XCP-i has been developed and will be available on the web. Based on a Linux kernel, a software XCP-i router is currently under development.

6. New Results

6.1. Optimized communication software and equipements

6.1.1. Optimisation of MPI applications

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

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

State of the art of MPI implementations is followed by analysis of one of them, MPICH-Madeleine, which manages heterogeneity with efficiency. Our optimizations lead to obtain a bandwith of 600 Mbps instead of 95 Mbps for sending MPI messages over a Wide Area Network. The experiments have been performed on the Grid'5000 french national plateform.

6.1.2. Study of Layer 2 equipements behavior in presence of contending flows

Keywords: *cross-layering*, *ethernet switches*, *queue management*, *transport protocol*.

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

The goal of this study is to examine the interactions between layer 2 (Ethernet) switches and TCP in high bandwidth delay product networks. First, the behavior of a range of Ethernet switches when two long lived connections compete for the same output port is investigated. Then, the report explores the impact of these behaviors on TCP protocol in long and fast networks (LFNs). Several conditions in which scheduling mechanisms introduce heavy unfair bandwidth sharing and loss burst which impact TCP performance are shown [43].

This work shows several conditions in which the standard packet scheduling mechanisms in switches introduce heavy unfairness (or starvation) on large intervals (300 ms) and loss burst which impact TCP performance. These conditions correspond to situations where huge data movements occur simultaneously. It also shows that behaviors are different from switch to switch and not easily predictable. These observations offer some tracks to better understand layer interactions. They may explain some congestion collapse situations and why and how parallel transfers mixing packets of different connections take advantages over single stream transfers.

We plan to pursue this investigation of layer two - layer four interaction and explore how to model it and better adapt control algorithms to fit the new applications requirements.

6.1.3. Design of Grid gateway with network processor technology

Keywords: *Grid gateway*, *IXP2400*, *flow control*, *network processor*.

Participants: Sébastien Soudan, Pascale Vicat-Blanc Primet.

Due to network bandwidth growth, grid computing has been envisionned. As grids are aggregation of computer clusters based on high performance local networks and interconnected by very high speed core networks, access links to core networks remain bottleneck locations. In order to optimize grid computing overall performance, we propose to control data transfers from one cluster to an other. This control must allow ressource schedulers to schedule data transferts in addition of tasks. Based on a grid network resource reservation model, we design a grid gateway, participating to the flow scheduling and control processes, located at the LAN/WAN interface. To have a better insights of the constraints and on the implementation issues, we have developped a prototype based on *Network Processors* INTEL *IXP2400* which is able to control flows at 1 gigabits/s speed without impacting the central processor performance. We discuss the main design choices and experimental results [49], [45].

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

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

Participants: Jean-Patrick Gelas, Laurent Lefèvre.

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

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

We proposed an adaptation of a generic high performance active network environment (Tamanoir) in order to deploy on limited resources based network boxes and to increase reliability and scalability. The implementation process is based on a hardware solution provided by the Bearstech company. Through this approach we proposed the architecture of an **I**ndustrial **A**utonomic **N**etwork **N**ode (called IAN^2) able to be deployed in industrial platforms [55], [65]. We evaluated the capabilities of IAN^2 in terms of computing and networking resources and dynamic re-programmability. [47]

6.1.5. High availability for clustered network equipments

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

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

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

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

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

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

  High availability allows service architectures to meet growing demands and to ensure uninterrupted service. In our work, we are interested in providing the continuous execution of the offered network sessions in case of failure of the legitimate entry point to the cluster as well as in case of the failure of the processing server inside the cluster. We noticed that current fault tolerant frameworks need to support consistent transport and application level failover mechanisms, and that transport layer protocols do not provide high availability capabilities. Indeed, TCP does not distinguish between a packet loss due to congestion, or a packet loss due to a server overload or due to a server/link failure. Thus, it reacts the same way to packet losses and to delays, by retransmitting the same segment to

the same remote end point of the connection. Moreover, TCP tolerates short periods of disconnection not longer than a few RTTs. It disconnects the communicating hosts once specific timers expire. On the other hand, transport protocols rely on an explicit association between a service and its physical location for the wired Internet. Thus, when a host fails, the end-to-end flow terminates.

In order to address this limitation, we proposed an active replication based system which enhances the reliability of the already established TCP flows. The proposed scheme is client transparent and does not incur any overhead to the end-to-end communication during failsafe periods, and performs well during failures. Parts of this work are protected by the Intellectual Property National Institute (INPI) patent disclosure N°FR0653546 [36].

6.1.6. High availability for stateful network equipments

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

Participant: Laurent Lefèvre.

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

Proposing stateful network equipments on open source systems is a challenging task. We propose the basic blocks (SNE library) for building a stateful network equipment. This library can be combined with high-availability protocols (CARP, Linux HA...). We focus on Linux system in order to provide software solutions for designing high-available solutions for NAT, firewalls, proxies or gateways equipments...This library is based on components located in kernel and in user space of the network equipment. First micro-benchmark of communications mechanisms with Netlink sockets have shown the effectiveness of our approach [14].

6.2. End-to-end High performance transport

6.2.1. Evaluation of High Speed TCP variants and study of large flow interactions in high-speed shared networks

Keywords: *bulk data transfers*, *congestion control*, *high speed transport protocol*, *transfer delay predictability*, *transport protocol experimentation*.

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

We consider the problem of huge data transfers and congestion control in contexts where transfer delay bounds are required. We investigate high-speed TCP variants and contribute to their evaluation by providing accurate measurements. This work gives an insight on the behaviour of alternative protocols under different realistic congestion and long latency conditions in the 10 Gbps experimental environments provided by the Grid5000 testbed and by the GtrcNET10 latency emulation device. This work also gives experimental results on performance of a large number of parallel flows (up to 110 parallel streams) and on large flow interactions in a real very high-speed networks [39], [43]. This work complements the general studies on transport protocol benchmarking which we explored within the international cpfldneet community [74].

6.2.2. Router assisted network transport protocol

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

Participants: Dino Martin Lopez-Pacheco, Laurent Lefèvre.

In heterogeneous networks, where many flows, non-regulated and/or with a high QoS level, share the resources, the available best-effort bandwidth varies over time. This changes can be represented by an aggregation of UDP ON-OFF sources what produces a step-based variation model. In this type of environments, we have tested the performance of many transport control protocols (TCP New Reno, High Speed TCP, TCP Westwood+ and XCP) using the ns2 simulator. In our studies, XCP showed always the best performance, with a high stability and fairness level. But in heterogeneous networks, the lost of packets is very common, so we have tested XCP in a network where the lost in the reverse path cause some ACK losses. In the new results, we have found that the ACK losses produce many problems in the connections, caused by a wrong calculus of the congestion window size, specifically when the available bandwidth decreases. That is because the success of XCP is based on the network state information, provided by the routers to the sender in the ACK packets. Since, the problem is generated by the wrong calculus of the congestion window size in the sender side, we proposed to compute this value in the receiver side. We have called this new approach XCP-r [26]. We repeated the simulations set using XCP-r and we found that XCP-r shows always more stability and better fairness level.

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

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

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

XCP (eXplicit Control Protocol) is a transport protocol that uses the assistance of specialized routers to very accurately determine the available bandwidth along the path from the source to the destination. In this way, XCP efficiently controls the sender's congestion window size thus avoiding the traditional slow-start and congestion avoidance phase. However, XCP requires the collaboration of all the routers on the data path which is almost impossible to achieve in an incremental deployment scenario of XCP. It has been shown that XCP behaves badly, worse than TCP, in the presence of non-XCP routers thus limiting dramatically the benefit of having XCP running in some parts of the network. In this work, we address this problem and propose XCP-i which is operable on an internetwork consisting of XCP routers and traditional IP routers without loosing the benefit of the XCP control laws [24], [25], [42], [41]. XCP-i basically executes the next four steps to discover and compute a new feedback that reflects the state of the network where non-XCP routers are placed:

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

The simulation results on a number of topologies that reflect the various scenario of incremental deployment on the Internet show that although XCP-i performances depend on available bandwidth estimation accuracy, XCP-i still outperforms TCP on high-speed links.

6.2.4. Scheduling bulk data transfers in grid networks

Keywords: *bandwidth reservation*, *deadline*, *flexible start time*, *flow scheduling*, *grid networks*, *multi-rate*.

Participants: Chen Binbin, Sébastien Soudan, Pascale Vicat-Blanc Primet.

In this long term research area, we consider the problem of bulk data transfers and bandwidth sharing in the context of grid infrastructures and propose to explore a disruptive approach for congestion control in high pseed networks. Indeed, in grid computing which empowers high-performance computing in a large-scale distributed environment, network bandwidth, which makes the expensive computational and storage resources work in concert, plays an active role on carrying grid applications traffic. Due to specific traffic patterns and application scenarios, grid network resource management encounters new challenges. From the bandwidth sharing perspective, we look at network bandwidth shared among computing and storage elements and explore a session level network resource control approach.

- In our first investigations of this field, we introduced a specific network model: a hierarchical bipartite graph with two sets of bottlenecks called ingress and egress points and defined bulk data transfer job request. Referred to as short-lived, grid data requests with transmission window and volume are scheduled in the network. By manipulating the transmission window, the request accept rate and network resource utilization are to be optimized. The formulated optimization problem, considering this network model, is proven NP-complete. Associated with proposed heuristics, simulations are carried out to illustrate the pros and cons of each bandwidth sharing strategy and its application scenarios. A tuning factor, that allows for adapting performance objective, is introduced to adjust network infrastructure and workload [27].
- We then continue this study of bandwidth reservation problem for bulk data transfers in grid networks. We generalize our grid networks model as a set of distributed sites interconnected by any network with potential bottlenecks, and transfer requests arrive online with specifed volumes and deadlines. Current reservation schemes such as RSVP are designed for requests with fixed transmission start time and single rate. In comparison, our definition of request in terms of volume and deadline allows more flexibility in the design of reservation schemes. We define the extended design space by formalizing three schemes families, namely, NOW (or immediate), Single Rate (SR) and Multirate (MR), with increasing generality, complexity and potential performance. Maximal packing (MaxPack) and minimize delay (MinDelay) are set as criteria to select candidate scheme from each family. The proposed reservation schemes is shown to achieve a much better performance then RSVP-type schemes, and can be implemented in both centralized and distributed architectures [56].
- The following study explored the same problem of bandwidth scheduling for transfers with specified volume, active time window (arrival time and deadline) and route, but consider both periodic and sporadic bulk data transfers. For periodic transfers, their request definitions are available off-line and network capacity is dimensioned to accept all of them. An important objective of scheduling, thus, is to minimize the (weighted) maximum required capacity in network along time axis. If request must be served with non-zero bandwidth in a continuous interval, the optimal scheduling is NP-complete even for a single link. In comparison, if the active window of request can be divided into multiple sub-intervals, each with different data rates (possibly zero), the optimal scheduling problem can be modeled as a multicommodity network flow problem which employs polynomial solution. Remained network capacity from periodic transfers is then used to serve sporadic transfers which arrive dynamically. The performance metric for sporadic transfers includes both accept probability and flow time, both of which can be potentially improved if bandwidth is scheduled flexibility [37].
- Finally, we start examining how the advance reservation and off-line data transfers jobs scheduling will interfere with an unified control plane allowing the creation of bandwidth guaranteed tunnels across optical core network and Ethernet local network. We propose a model for such networks and study the problem of bandwidth sharing with bulk data transfers in this GMPLS context. Several allocation algorithms based on QoS routing works have been proposed and compared [44], [33]. This work will be pursued by algorithm development, simulations and experiments in Grid5000 and DAS3 testbeds.

6.3. Grid Network services and applications

6.3.1. Evaluation of 10 GbE links in Grid'5000

Keywords: *10 GbE*, *Grid'5000*, *TCP*, *TCP tuning*, *wizard gap*.
Participants: Ludovic Hablot, Sébastien Soudan, Romaric Guillier, Pascale Vicat-Blanc Primet.

This activity [40] concerns an end-to-end evaluation of the 10Gb/s infrastructure of Grid'5000, the French research grid built with 3000 processors on 9 sites [13]. These sites are connected with a dedicated DWDM provate network, which provides 1 GbE or 10 GbE links to end sites. This work shows how TCP parameters tuning impacts the performance of individual connection as well as the utilization of the infrastructure. We evaluate TCP problems in Grid'5000 links and highlight several software and hardware configuration issues detected in the Grid5000 infrastructure.

6.3.2. Large Scale Gigabit Emulated Testbed for Grid Transport Evaluation

Keywords: *emulation*, *grid computing*, *performance analysis*, *transport protocols evaluation*.

Participants: Magi Sanchon, Pascale Vicat-Blanc Primet.

Evaluating the performance of Grid applications running on high performance platforms interconnected by high speed and long distance networks with new transport services and protocols is highly required. This work examines the performance of the eWAN integrated environment enabling large scale grid emulation at gigabit speed. It discusses features provided to control key characteristics (topology, round trip time, packet size, drop rate, link capacity) of an evaluation scenario. A method to increase the accuracy of rate control under various delay configuration is proposed and some experimental results are detailled: grid computing, performance analysis, transport protocols evaluation, emulation. [35]

6.3.3. Context aware network services: video adaptation for mobile platforms

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

Participant: Laurent Lefèvre (contact).

Joint work with René Fuentes Riquelme from 3DDL and Jean-Marc Pierson from LIRIS Lab, INSA Lyon.
Active Networks allow user or applications to inject customized programs into the network nodes. The creation of new services is an original way to think about development and deployment of customized modules to perform computation within the network. This can lead to massive improvement of network functionalities.

New mobile phone generations integrate more and more a Java Virtual Machine. This JVM allows providers to propose applications and games working on heterogeneous phones (without having to redo some specific development and to adapt them individually for specific features).

We propose to benefit from active and programmable networks by deploying active nodes on data path to efficiently adapt streams on the fly. This research follows three main goals:

- to reduce development costs and the complexity for managing a version of a game for each mobile class. The active node will adapt the files on the fly;
- to reduce the usage of bandwidth and interactions between clients and data server;
- to efficiently support deployment of mobile applications without adding too much latency on real networks.

We design the architecture of active transcoding services (ActiveWapS) deployed inside the Tamanoir Execution Environment. These services transform on the fly, parts of the mobile applications and video streams in order to adapt them to target mobile phones. These services are deployed in network equipments and applied on the data streams before reaching mobile phones. We also validate this approach by emulation on a local platform [22].

6.3.4. Programmable network services for context aware adaptation

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

Participants: Martine Chaudier, Laurent Lefèvre, Jean-Patrick Gelas.

Traditional industrial maintenance process (i.e. requiring regularly a human intervention on the exploitation area) are coming to their limits. Indeed, more and more industrial equipments are connected to communication networks. This allows us to consider optimised maintenance solutions. In addition to primary existing sensors (which only give some numeric values), we can now think about the use of multimedia sensors (video cameras, microphone, ...). Inside a cooperative industrial maintenance project (TEMIC project [73]) in which we are currently involved, our team designed equipments easily deployable in an industrial context, and also easily removable at the end of the maintenance contract.

The heterogeneity interms of networks, terminals and applications requires adaptive solutions for an efficient streams transmission on the platform networks. To respond to these various constraints, active services have to adapt and optimize the content of streams passing through the active network node. Multimedia data streams adaptation is performed dynamically in order to improve industrial maintenance solutions. The challenge is to provide an architecture running in a client/server environment, but involving no modification on the applications installed on the end-machines like web servers, video players,... For the Temic project, our team has worked on the design and adaptation of an industrial autonomic network node, which is derived from the Tamanoir environment. This Industrial Autonomic Network Node is designed to be deployed on limited resources based network boxes, and so to be integrated into industrial platforms. We developed and tested active adaptation network services, specially written for the Tamanoir^{embedded}. Active services applying on multimedia streams crossing the network node may realize data compression, format transcoding, frame resizing... This kind of adaptation contributes to the saving of network bandwidth (by decreasing the output data rate) and to the reduction of the resources used on the client terminal playing the multimedia data (by reducing the framerate and the frame size). The adaptation is thereby transparent for the applications.

We base our developments and experimentations on mainly two industrial maintenance scenarios [15]. They were planned by the TEMIC project team to be used by a company through a maintenance contract on a restricted industrial area.

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

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

6.3.5. Inter-Planetary Grid Networking

Keywords: *Delay Tolerant Networking*, *Grid*, *autonomic networks*.

Participants: Jean-Patrick Gelas, Laurent Lefèvre.

The idea to extend the computer network protocols in order to tremendously extend the range of Internet through space was born and supported by the same persons who design TCP/IP 30 years ago, like Vint Cerf. Due to some constraints, transport protocols, among other (ex: routing, name space) must be radically changed to fit the requirement of this unusual environment, namely space! In the same time, the Delay Tolerant Networking (DTN) community works on networks which must deal with high latencies, frequent disconnections, no end-to-end path and power saving constraints. The new proposed protocols are designed to support high latencies and long disconnection (i.e. more than few milliseconds). They also should resist to planned or unplanned disconnection. We consider that the concept of Interplanetary Networks based on Disruption Tolerant Network solutions can be applied to Grid infrastructures.

Programmable and active networks allow specified classes of users to deploy dynamic network services adapted to data streams requirements. We have proposed the Active Grid Architecture (A-Grid) which focuses on active network adaptation for supporting Grid environments and applications. This Active Grid architecture proposes solutions to support multi-cluster infrastructures. This architecture is based on programmable network nodes distributed on network path used as gateways of clusters. In this architecture the network will take part in the Grid computing session by providing efficient and intelligent services dedicated to Grid data streams transport.

This tolerant design expects that applications remain efficient even if networks generate high latencies for communications. This approach focused on latency can be generalized to disrupted infrastructures. But, we want to propose global solutions as transparent as possible for users, applications and Grid middleware. Our approach allows us to modify only the system used as Programmable Network Gateway (PNG) located between clusters and the external network (i.e. Internet) [21].

The proposed architecture of an Interplanetary Grid can also be applied to Grid infrastructures dealing with unreliable long distance network connections. We are currently implementing the model exposed in the previous section and we plan to emulate first experimentations and evaluations of this approach [62].

6.3.6. Integrating web services and programmable networks for improving flexibility of active Grids

Keywords: *Web services*, *programmable networks*.

Participants: Laurent Lefèvre, Pablo Pazos Rey.

Joint work with Chien-Jon Soon and Paul Roe from Queensland University of Technology, Brisbane, (Australia).
Active Grids [53], [54] are a form of grid infrastructure where the grid network is active and programmable. These grids directly support applications with value added services [66] such as data migration, compression, adaptation and monitoring. Services such as these are particularly important for eResearch applications which by their very nature are performance critical and data intensive.

We propose an architecture for improving the flexibility of Active Grids through web services. These enable Active Grid services to be easily and flexibly configured, monitored and deployed from practically any platform or application. The architecture is called WeSPNI ("Web Services based on Programmable Networks Infrastructure") [23].

7. Contracts and Grants with Industry

7.1. France Telecom R&D

Participants: Laurent Lefèvre, Pascale Primet.

In 2005, RESO has launched a collaboration with France Telecom R&D (Lannion) on "Network load balancing on layer 7 switching for high performance and high available Linux based platforms". A CIFRE grant has been accepted for supporting this collaboration. Ayari Narjess has begun her PhD on this topic in June 2005 [52].

7.2. INTEL

Participants: Pascale Vicat-Blanc Primet, Sébastien Soudan, Laurent Lefèvre.

This collaboration aims at studying the potential of the network processor technology for building High performance (several Gigabits links) network emulators and dynamically programmable routers. The goal is to show that network processors improve performance and enhance capacities of Software network emulators and programmable routers based on Linux platforms. Network interface cards with network processors have been integrated within the GRID5000 testbed. A Grid Gateway prototype for flow control has been developed.

7.3. 3DDL

Keywords: *java*, *programmable networks*.

Participant: Laurent Lefèvre.

RESO has established a long term collaboration with 3DDL SME. This collaboration concerns the design and deployment of software components inside the network in order to support the deployment of mobile applications on heterogeneous terminals [67] [22]. Funded by Région Rhone-Alpes with collaboration of LIRIS, INSA Lyon.

7.4. Bearstech

Keywords: *embedded PC*, *network services*.

Participants: Jean-Patrick Gelas, Laurent Lefèvre (contact).

Since 2004, RESO is launching a collaboration with this young company targeted on embedded computers and network equipments. This collaboration has allowed an improved design of "Tamanoir embedded" software suite.

7.5. Alcatel

Participant: Pascale Vicat-Blanc Primet.

This collaboration concerns the extension of grid technology to Internet scale. It consists of an analysis of concrete solutions for a very large scale and distributed grid information management functionality (resource discovery, localization, selection, and maintenance).

8. Other Grants and Activities

8.1. Regional actions

8.1.1. Fédération Lyonnaise de Calcul Scientifique Haute Performance

Participant: Laurent Lefèvre.

RESO is a member of the "Fédération Lyonnaise de Calcul Scientifique Haute Performance", that is building a regional grid infrastructure with several high-performance clusters and parallel machines. Supported by the Rhone-Alpes region (2004-2005).

8.2. National actions

8.2.1. ACI GRID: GRID5000

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

ENS Lyon is involved in the GRID'5000 project, which aims at building an experimental Grid platform gathering eight sites geographically distributed in France.

The operational status of Lyon's part of Grid'5000, is as follow. For the hardware, Foundry switches and IBM computers have been upgraded to the latest firmware. For the security point of view a firewall has been configured to run on the gateway to provide an isolation from the computers which are not part of Grid'5000, and a set of proxy to render basic services (DNS, NTP) have been activated. Finally for the users point of view, they have a set of computers whose clocks are synchronized (by NTP), where they have a uniform login/account handled by an LDAP server (previously done using NIS), and a common home directory delivered by NFS, the frontend provide them with the necessary compilation tools for the AMD64 architecture (optimized PathScale compiler is available), the currently available distribution on the node are Debian or Gentoo, which run using the native 64bits mode.

With the help of funding of INRIA Rhone-Alpes, the platform has been upgraded with 150 processors and a 10Gb/s core lan. The Grid5000 of Lyon comprises now around 300 processors interconnected with a network of 250Mb/s Ethernet bisection and a 2Gb/s Myrinet interconnection for 64 nodes.

RESO has been strongly involved during this year in the design of the national prototype platform of GRID'5000 and in the choices of network components and architecture. Pascale Vicat-Blanc Primet is member of the national committee (comité de pilotage) of GRID'5000, co-responsible of the Lyon site with Frederic Desprez, and coordinates networks aspects with Renater and RMU, Lyon's metropolitan network. She defines with Renater the new dark fiber core infrastructure that will enable to interconnect each sites with 10G/s access links. She is also working for the interconnection of the Grid5000 project and the japanese Naregi project. Aurélien Cedeyn is member of the national technical committee of GRID'5000. Actual funding: 530K euros

8.2.2. RNRT Temic

Participants: Laurent Lefèvre, Jean-Patrick Gelas, Martine Chaudier, Pierre Bozonnet.

(2003-2006) The RNRT Temic project is focused on providing solutions for collaborative management of large and complex industrial process. In this project, RESO provides dynamic and adaptative networking solutions for efficiently supporting heterogeneous data streams and equipments. Experiments and platforms based on active and programmable network technology will be designed. RESO also proposes multimedia adaptive network services for industrial sensors. Funding : 2 Engineers for 1 year

8.2.3. ACI Grandes Masses de Données GridExplorer

Participants: Jean-Patrick Gelas, Olivier Glück, Laurent Lefèvre, Dino Lopez Pacheco, Pascale Vicat-Blanc Primet.

(2003-2006) : The aim of this project was to create a large scale grid and network emulator. RESO is involved in the design of the platform and is interested in designing a high performance transport protocol test methodology in this environment. EWAN [75] [34], our high performance network emulator, is one of the main RESO contributions to this project. Pascale Vicat-Blanc is responsible of the network theme. RESO has participated to the definition of the architecture and technical choices of cluster hardware.

The second part of our contribution to this project was the evaluation of high speed transport protocol. This activity has started within the GridExplorer project and is now continuing within the ANR IGTMD project.

8.2.4. ANR IGTMD

Participants: Pascale Vicat-Blanc Primet, Romaric Guillier.

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

8.2.5. ANR DSLLAB

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

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

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

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

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

8.2.6. ANR HIPerCAL

Participants: Pascale Vicat-Blanc Primet, Olivier Glück, Paulo Gonçalves, Isabelle Guérin Lassous, Laurent Lefèvre.

HIPerCAL (2007-2009) 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) in 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 current services-oriented principles developed in grids to jointly enhance the application portability, the communications performance control and their security. The project aims at providing a grid substrate based on end-to-end bandwidth reservation, control overlay, network and system virtualization, cryptographic identification principles. The proposal will be validated and evaluated at different scales on the Grid5000 testbed with biomedical applications, demanding in security, performance and reliability. 10 to 1000 processors, links with 100Mb/s to 10Gb/s, few microseconds to 100ms will be involved in these experimentations. 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 HIPerCAL approach.

In this project, RESO is responsible for the HIPerCAL software integration and for the design and development of advance network control solutions.

8.3. European actions

8.3.1. AEOLUS

Participants: Isabelle Guérin-Lassous, Rémi Vanier.

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

- To identify and study the important fundamental problems and investigate the corresponding algorithmic principles related to overlay computers running on global computers.
- To identify the important functionalities such an overlay computer should provide as tools to the programmer, and to develop, rigorously analyze and experimentally validate algorithmic methods that can make these functionalities efficient, scalable, fault-tolerant, and transparent to heterogeneity.

- To provide improved methods for communication and computing among wireless and possibly mobile nodes so that they can transparently become part of larger Internet-based overlay computers.
- To implement a set of functionalities, integrate them under a common software platform in order to provide the basic primitives of an overlay computer, as well as build sample services on this overlay computer, thus providing a proof-of-concept for our theoretical results.

8.3.2. EC-GIN

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

EC-GIN (Europe-China Grid InterNetworking) is an European STREP project started in November 1st 2006. The university of Innsbrück (Austria) is the prime contractor.
The Internet communication infrastructure (the TCP/IP protocol stack) is designed for broad use; as such, it does not take the specific characteristics of Grid applications into account. This one-size-fits-all approach works for a number of application domains, however, it is far from being optimal - general network mechanisms, while useful for the Grid, cannot be as efficient as customised solutions. While the Grid is slowly emerging, its network infrastructure is still in its infancy. Thus, based on a number of properties that make Grids unique from the network perspective, the project EC-GIN will develop tailored network technology in dedicated support of Grid applications. These technical solutions will be supplemented with a secure and incentive-based Grid Services network traffic management system, which will balance the conflicting performance demand and the economic use of resources in the network and within the Grid.
By collaboration between European and Chinese partners, EC-GIN parallels previous efforts for real-time multimedia transmission across the Internet: much like the Grid, these applications have special network requirements and show a special behaviour from the network perspective. However, while research into network support for multimedia applications has flourished, leading to a large number of standard protocols and mechanisms, the research community has neglected network support for Grid computing up to now. By filling this gap and appropriately exploiting / disseminating the project results, EC-GIN will, therefore, cause a "snowball effect" in the European and Chinese networking and Grid computing research communities.
Technically, EC-GIN will make the Grid work, operate, and communicate better. By appropriately utilizing the underlying network, Grid resources in general will be used more efficiently and amplify the impact of Grid computing on the society and economy of Europe and China.

8.4. International actions

8.4.1. Programme d'Actions Intégrées "Sakura" with AIST GTRC Japan + Equipe associée

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

After the first France-Japan Grid workshop in Paris (Mars 2004), INRIA RESO team and AIST GTRC group decided to work together. Both team focus their activities in High Performance GridNetworking area. AIST GTRC networking group is studying approaches that activate and use some intelligence in a dedicated equipment in the path, named GtrcNet1. Two GtrcNet1 equipments have been installed within the GRID5000 node in Lyon. Both our teams adopt the same type of solutions based on IP technology, and exploiting some "intelligence" within the network (i.e. in network interface cards or in programmable equipments located in edge networks) to tackle the same kind of problems: high end-to-end throughput, performance control and measurement. A Memorandum of Understanding has been signed between INRIA and AIST GTRC in july 2004. A "Programme d'Actions Integrees" SAKURA project has been accepted for the 2005-2007 period wich enable a fruitful collaborative research on high performance evaluation, network processing and protocol benchmarking.

In 2005 Pascale Vicat-Blanc Primet spent two weeks in Japan for deploying and experimenting the eWAN software within the AIST SuperCluster [34]. In 2005 Pascale Vicat-Blanc Primet, Sébastien Soudan and Romaric Guillier spent two weeks at AIST for studying transport protocol issues as well as 10Gb/s networks issues. Several publications have been collaboratively written and a Associated Team application prepared. RESO and GTRC AIST will pursue and amplify their collaborations. The accepted INRIA Associated-Team program will run during 2007-2009 period.

8.4.2. NEGST: JSPT-CNRS

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

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

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

RESO mainly participates to the Grid Metrics topic.
Despite the development of strong technologies in all these domains, many issues are still open about the measurement methodology itself, the emulation or simulation of Grid platforms and the understanding of Grid software stack, application performance, and fault tolerance. The Grid Metrics topics, basically gathers all researches about applications, programming models, libraries, runtimes, operating systems and network evaluation, either in synthetic environment (emulators and simulators) or real environment (real network and Grids). Pascale Vicat-Blanc has participated to the periodical NEGST workshops (Paris in june 06; Tampa in november 06).

8.4.3. Programme d'Actions Intégrées "Fast" with Queensland University of Technology

Participants: Laurent Lefèvre, Jean-Patrick Gelas, Pablo Pazos Rey.

Joint work with Paul Roe (from Queensland University of Technology, Brisbane (Australia).
This project focuses on the design of Web Services based on Programmable Networks Infrastructure (WeSPNI). This collaboration between RESO team and Programming Language and System group (PLAS) in Queensland University of Technology (Brisbane, Australia) aims to bring together researchers able to design next generation of overlay networks. We observe a real convergence between Grid infrastructure and Web Service solutions. Based on the Open Grid Service Infrastructure, Grid researchers have proposed the WSRF (Web Service Resource Framework) where Web Services naturally fit in Grid requirements. This collaboration exploits this convergence by providing network solutions adapted to Grid requirements. This Fast project is supported by French Ministry of Foreign affairs (2005-2006) [23]. In 2006, L. Lefèvre has spent one month in QUT, Australia (October 2006) and P. Roe has visited RESO (December 2006).

This project focuses on the design of Web Services based on Programmable Networks Infrastructure (WeSPNI). This collaboration between RESO team and Programming Language and System group (PLAS) in Queensland University of Technology (Brisbane, Australia) aims to bring together researchers able to design next generation of overlay networks. We observe a real convergence between Grid infrastructure and Web Service solutions. Based on the Open Grid Service Infrastructure, Grid researchers have proposed the WSRF (Web Service Resource Framework) where Web Services naturally fit in Grid requirements. This collaboration exploits this convergence by providing network solutions adapted to Grid requirements. This Fast project is supported by French Ministry of Foreign affairs (2005-2006) [23]. In 2006, L. Lefèvre has spent one month in QUT, Australia (October 2006) and P. Roe has visited RESO (December 2006).

8.4.4. NSF-INRIA with Aerospace Organization

Participant: Laurent Lefèvre.

A NSF-INRIA project is running with Aerospace Organization-USA (C. Lee team) on support of programmable networks for Grid middleware and overlays. (2004-2006). A call for papers for a journal special issue is currently under preparation.

8.5. Visitors

8.5.1. Collaboration with AIST GTRC, Japan

Participant: Pascale Vicat-Blanc.

Collaboration with Tomohiro Kudoh and Kodama Yuetsu from AIST (Japan).
RESO has hosted Dr Tomohiro Kudoh and N. Yuetsu for 1 week as invited researchers in April and October 2006 to work on experimentation with GtrcNet1 equipment integrated in the Grid5000 cluster.

9. Dissemination

9.1. Conference organisation, editors for special issues

- Pascale Vicat-Blanc is
  - Workshop chair of the IEEE International Conference on High Performance Distributed Computing (HPDC2006) in Paris;
  - co-chairing the International Workshop on Grid networks (GridNets2006) of the IEEE Broadnet Conference in San Jose (California- USA);
  - general chair of first International Conference on Grid networks (GridNets2007);
  - co-chairing the high Performance networks track of the International conference Europar2007;
  - member of program committees : VECPAR2006, CFIP2006, GRIDNETS2006, PFLDNET2006, EUROPAR2007, GRIDNETS2007. She has been reviewer for international journal and conferences: Communication Network Journal, Parallel letter, JPDC, Calculateurs Parallèles, TSI.

    She is co-editor of a special issue of the Elsevier Communication Network Journal on Hot Topic in Transport Protocols for very high speed long distance networks.
- Laurent Lefèvre is:
  - organizer and *program chairman* of workshops series "Distributed Shared MemOry on Clusters" DSM2006 (Singapore) within IEEE International Symposium on Cluster Computing and the Grid (CCGrid);
  - program Vice-chair of HPCC06: International Conference on High Performance Computing and Communications, Munich, Germany, September 13-15, 2006;
  - program Co-Chair of ICPS 2006: International Conference on Pervasive Services, Lyon, France, June 26-30, 2006;
  - co-organizer of the INRIA Booth during the Supercomputing conference (SC06) in Tampa Bay, USA, November 2006;
  - *Steering Committee* member of:
    * IEEE International Symposium on Cluster Computing and the Grid (CCGrid conference) series since 2004;
    * ICPS2007: IEEE International Conference on Pervasive Services, Istanbul, Turkey, July 2007;

    * IWAN2006: Eight International Workshop on Programmable and Active Networks, Paris, France, September 25-29, 2006 during Autonomic Networking 2006 conference;
  - member of the following Program Comittees: e-Science 2006 : 2nd IEEE International Conference on e-Science and Grid Computing ,SuperComputing 2006, Grid2006 : The 7th IEEE/ACM International Conference on Grid Computing, Euro PVMMPI 2006 : 13th European PVM/MPI Conference, ICPP 2006: 35th International Conference on Parallel Processing, Session on Network Management, Modelling and Performance in the 5th International Conference on Communication Systems, Networks and Digital Signal Processing, CSNDSP 2006, DFMA'06 : Second International Conference on Distributed Frameworks for Multimedia Applications.

- Isabelle Guérin Lassous is:
  - co-organizing the École d'Été ResCom 2007;
  - a member of the following program committees: 1st International Conference on Late Advances in Networks (ICLAN 2006) Paris, France - 2nd International Workshop on Localized Communication and Topology Protocols for Ad hoc Networks (LOCAN 2006), Vancouver, Canada - ACM/IEEE MSWiM 2006, Torremolinos, Malaga, Spain - IFIP 1st International Conference on Ad-Hoc Networking, Santiago Chile - Med-Hoc-Net 2006, Lipari, Italy - Seventh ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 2006), Firenze, Italy, Second International Workshop on Sensor Networks and Systems for Pervasive Computing (PerSeNS 2006), Pisa, Italy, IEEE International Workshop on Foundations and Algorithms for Wireless Networking (FAWN 2006), Pisa, Italy.
- Paulo Gonçalves is local co-chair of GRIDNETS2007 and member of the technical program committee for 2007 IEEE Statistical Signal Processing (SSP) Workshop.
- Jean-Patrick Gelas is:
  - PC member of HPCC 2006 : The 2006 International Conference on High Performance Computing and Communications, Munich, Germany, 13-15 September 2006.
  - Local Organisation comitee member and chairman at ICPS 2006 : International Conference on Pervasive Services, Lyon, France, June 26-30, 2006

9.2. Graduate teaching

- **since 2004**: P. Vicat-Blanc Primet
  Advanced protocols for high speed networks. *Réseaux avancés et leurs protocoles*.
  Master Research (Ecole Normale Supérieure de Lyon, University Claude Bernard Lyon 1), lecture: 28h/year.
- **since 2004**: O.Glück
  Client/Server Model, Internet Applications, Network and System Administration.
  Master 2 SIR (University Claude Bernard Lyon 1), lecture 30h, others 30h.
- **since 2004**: JP.Gelas
  High-speed networks, QoS and Multimedia ; Initiation to Java ; Local Area Networks .
  Master 2 SIR and CCI (University Claude Bernard Lyon 1), lecture 30h, others 40h.
- **since 2005**: JP.Gelas
  Long distance networks ; Networks and Transport Protocols ; Routing ; Advanced Java and Web services Master 2 SIR (University Claude Bernard Lyon 1), lecture 45h, others 45h.

9.3. Miscelleneous teaching

- **since 2004**: O. Glück
  Computer Networks.
  Licence Informatique, (University Claude Bernard Lyon 1), lecture 30h, others 30h.
- **since 2004**: O. Glück
  Programming on the Web.
  Master 1 Informatique, formely Maîtrise d'Informatique, (University Claude Bernard Lyon 1), lecture 15h, others 15h.
- **since 2002** : L. Lefèvre
  High peformance Networks, Internet and tools.
  Maitrise Informatique (Université Antilles Guyane, Pointe à Pitre), 40h eq TD/an.
- L. Lefèvre
  Networks.
  Licence Informatique L3if (Ecole Normale Supérieure de Lyon), 30h eq TD/an.

9.4. Animation of the scientific community

- Pascale Vicat-Blanc
  - is member of the "Networks" expert committee of the CNRS;
  - is within the Global Grid Forum, standardization entity for grid middleware, is co-chair of the Data-Transport Research Group. RESO is also active in the Network Monitoring Working Group as in the Grid High Performance Networking;
  - is within the Grid5000 project, member of the steering committee;
  - is within the DataGrid explorer project (ACi Masses de Données), member of the steering committee and responsible for the topic "networks".
- Isabelle Guérin Lassous is:
  - member of the CNRS TAROT action (Techniques Algorithmiques, Réseaux et d'Optimisation pour les Télécommunications);
  - the INRIA scientific leader of the european project AEOLUS (Algorithmic Principles for Building Efficient Overlay Computers);
  - the INRIA scientific leader of a contract with FT R&D, "Bandwidth problems in multihop wireless networks";
  - member of the ARC INRIA Iramus (Radio Interface for Multihop Networks);

9.5. Participation in boards of examiners and committees

- Pascale Vicat-Blanc
  - is Chair of the board of examiners for recruitments of *Chargés de Recherche CR2* of the Rhône-Alpes INRIA research unit in 2006.
  - is co-chair of Technical Committee 3 (CT3 (Protocol Architecture)) and member of Evaluation Commission of the ANR Telecom. Participation to the extended Executive Board for 2007 Call for Project definition.
  - is member of two PhD examining boards: Paul Staerzt (LSR IMAG - reviewer) and Mathieu Gineste(LAAS - reviewer).
- Isabelle Guérin Lassous is member of:
  - the specialists committee (section 27) of the ENS Lyon;
  - the hearing committee of INRIA Rhône-Alpes;

  - the SPECIF committee that allocates PhD awards;
  - four PhD examining boards: Nathalie Mitton (INSA de Lyon - Co-supervisor), Luigi Iannone (Paris 6), Dang Quan Nguyen (Paris 6 - reviewer) and Fanilo Harivelo (La Réunion - reviewer).
- Olivier Glück is a member of
  - the “commissions de spécialistes 27ème section” of University Claude Bernard Lyon 1 and University Pierre et Marie Curie Paris 6.
  - the “conseil de l’UFR d’Informatique” of University Claude Bernard Lyon 1.
  - the “Conseil des Etudes et de la Vie Universitaire” of University Claude Bernard Lyon 1.
- Laurent Lefèvre
  - is member of the “commissions de spécialistes de 27ème section” of University Jean Monnet (Saint-Etienne), University Antilles Guyane (Pointe à Pitre) and University Lumière (Lyon2);
  - has been reviewer of the PhD thesis jury of Bruno Volckaert : "Architectures and algorithms for network and service aware Grid resource management", University of Gent, Belgium, May 2006;
  - has been reviewer for the Agence Nationale de la Recherche (ANR), Telecommunications programs, Call for Proposals, 2006.

9.6. Seminars, invited talks

- Pascale Vicat-Blanc was:
  - Invited talk at the national ARRU workshop organised by RENATER : "High Speed Transport Protocol evaluation in Grid5000". April 2006
  - Invited Session "QoS and Security" co-chair and speaker of the first International workshop ITU/OGF (International Telecommunication Union/Open Grid Forum) on Grids and New Generation Networks. Invited talk title: "QoS and Security issues in Grids". October 2006.
  - Invited talk at AIST booth - SC06 : "Research on High Speed Networks & Transport protocols for Grid applications",; November 2006
  - Invited talk at PARISTIC2006 "High Speed Transport Protocol evaluation in Grid5000". November 2006
  - Invited talk at GRID5000-DAS3 workshop "GridNetworking researches in Grid5000" : "High Speed Transport Protocol evaluation in Grid5000". December 2006
  - Invited talk at INRIA-ALCATEL steering comittee meeting "On Grids, Networks & Protocols" December 2006
  - Invited talk at 4th International TERENA workshop "NRENs and Grids" : "High Speed Transport Protocol evaluation in Grid5000". December 2006
- Laurent Lefèvre has been invited to give a talk on "Designing High Performance Autonomic Gateways for Large Scale Grids and Distributed Environments", during Clusters and Computational Grids for Scientific Computing - CCGSC2006, Flat Rock, USA, September 12, 2006
- Paulo Gonçalves was invited speaker at the “Nonlinear dynamical methods and time series analysis workshop", held in Udine (Italy), August 29 - September 1, 2006.

10. Bibliography

Major publications by the team in recent years

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

[2] 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, Massachussets", IEEE Computer Society Press, September 2005.

[3] P. GONÇALVES, R. RIEDI. *Diverging moments and parameter estimation*, in "Journal of American Statistical Association", vol. 100, n^{o} 472, December 2005, p. 1382–1393.

[4] M. GOUTELLE, P. VICAT-BLANC PRIMET. *Study of a non-intrusive method for measuring the end-to-end capacity and useful bandwidth of a path*, in "Proceedings of the 2004 International Conference on Communications, Paris, France", IEEE Communication Society, June 2004.

[5] J. LAGANIER, P. VICAT-BLANC PRIMET. *HIPernet: fully distributed security for grid environments*, GRID2005, USA, apr 2005.

[6] L. LEFÈVRE, J.-P. GELAS. *Programmable Networks for IP Service Deployment*, A. GALIS, S. DENAZIS, C. BROU, C. KLEIN (editors). , chap. Chapter 14 on "High Performance Execution Environments", Artech House Books, UK, may 2004, p. 291-321.

[7] D. LOPEZ PACHECO, C.-D. 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.

[8] L. MARCHAL, P. VICAT-BLANC PRIMET, Y. ROBERT, J. ZENG. *Optimal Bandwidth Sharing in Grid environment.*, in "IEEE HPDC 15th IEEE International Conference on High Performance Distributed Computing", 2006.

[9] P. VICAT-BLANC PRIMET, F. BONNASSIEUX, R. HARAKALY. *Network monitoring in the European DataGRID project*, in "International Journal of High Performance Computing Applications", vol. 18, n^{o} 3, January 2004, p. 293-304.

[10] P. VICAT-BLANC PRIMET, F. ECHANTILLAC, M. GOUTELLE. *Experiments of the equivalent differentiated service model in grids*, in "in International Journal Future Generation Computer Systems FGCS, special issue on "High Performance Networking and Services in Grids", vol. 21, n^{o} Issue 4, april 2005, p. 512-524 (.

Year Publications

Books and Monographs

[11] L. BRUNIE, S. HARIRI, L. LEFÈVRE, J.-M. PIERSON. *Proceedings of ICPS2006 : International Conference on Pervasive Services*, IEEE, Lyon, France, jun 2006.

Articles in refereed journals and book chapters

[12] A. ANTONIADIS, A. FEUERVERGER, P. GONÇALVES. *Wavelet Based Estimation for Univariate Stable Laws*, in "Annals of the Inst. of Stat. Math., Tokyo (JP)", To appear, 2006.

[13] 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.

Publications in Conferences and Workshops

[14] P. N. AYUSO, L. LEFÈVRE, R. M. GASCA. *High Availability support for the design of stateful networking equipments*, in "ARES'06 : The First International Conference on Availability, Reliability and Security, Vienna, Austria", April 2006.

[15] M. BRAHMA, M. CHAUDIER, E. GARCIA, J. GELAS, H. GUYENNET, F. HANTZ, L. LEFÈVRE, P. LORENZ, H. TOBIET. *TEMIC: a New Cooperative Platform for Industrial Tele-Maintenance*, in "DFMA06 : International Conference on Distributed Framework for Multimedia Applications, Penang, Malaysia", May 2006.

[16] H. CARRÃO, M. CAETANO, P. GONÇALVES. *Meris based land cover characterization: a comparative study*, in "Proceedings of the ASPRS international conference, Reno (NV, US)", May 2006.

[17] H. CARRÃO, P. GONÇALVES, M. CAETANO. *Use of intra-annual satellite imagery time-series for land cover characterization purpose*, in "Proceedings of the 26th EARSeL Symposium, Warsaw (Poland)", 2006.

[18] C. CHAUDET, I. GUÉRIN LASSOUS. *Etat des lieux sur la qualité de service dans les réseaux ad hoc*, in "Colloque Francophone sur l'Ingénierie des Procoles (CFIP), Tozeur, Tunis", November 2006.

[19] B. GONCALVES, N. MITTON, I. GUÉRIN LASSOUS. *Comparison of two Self-Organization and Hierarchical Routing Protocols for Ad Hoc Networks*, in "Second International Conference on Mobile Ad Hoc and Sensor Networks (MSN), Hong-Kong, China", December 2006.

[20] P. GONÇALVES, H. CARRÃO, M. CAETANO. *Parametric model for intra-annual reflectance time series*, in "Proceedings of the 26th EARSeL Symposium, Warsaw (Poland)", 2006.

[21] L. LEFÈVRE, J.-P. GELAS. *Towards interplanetary Grids*, in "Workshop on "Next Generation Communication Infrastructure for Deep-Space Communications" held in conjunction with the Second International Conference on Space Mission Challenges for Information Technology (SMC-IT), Pasadena, California", July 2006.

[22] L. LEFÈVRE, J.-M. PIERSON. *Just in time Entertainment deployment on mobile platforms*, in "ICIW'06 : International Conference on Internet and Web Applications and Services, Guadeloupe, French Caribbean", February 2006.

[23] L. LEFÈVRE, P. ROE. *Improving the flexibility of Active Grids through Web Services*, in "Fourth Australasian Symposium on Grid Computing and e-Research (AusGrid2006), Hobart, Australia", ISBN 1-920-68236-8, vol. 28, n^{o} 7, Australian Computer Society, January 2006, p. 3-8.

[24] D. LOPEZ PACHECO, C.-D. 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.

[25] D. LOPEZ PACHECO, C.-D. PHAM, L. LEFÈVRE. *XCP-i: eXplicit Control Protocol pour l'interconnexion de réseaux haut-débit hétérogènes*, in "CFIP 2006: Colloque Francophone sur l'Ingénierie des Protocoles, Tozeur, Tunisia", October 2006.

[26] D. LOPEZ PACHECO, C. PHAM. *Enabling Large Data Transfers on Dynamic, Very High-Speed Network Infrastructures*, in "Proceedings of ICN/ICONS/MCL 2006, Mauritius", april 2006.

[27] L. MARCHAL, P. VICAT-BLANC PRIMET, Y. ROBERT, J. ZENG. *Optimal Bandwidth Sharing in Grid environment.*, in "IEEE HPDC 15th IEEE International Conference on High Performance Distributed Computing", 2006.

[28] A. PINHEIRO, P. GONÇALVES, H. CARRÃO, M. CAETANO. *A toolbox for multi-temporal analysis of satellite imagery*, in "Proceedings of the 26th EARSeL Symposium, Warsaw (Poland)", 2006.

[29] R. RAZAFINDRALAMBO, I. GUÉRIN LASSOUS, L. IANNONE, S. FDIDA. *Aggrégation Dynamique de Paquets pour résoudre l'Anomalie de Performance des Réseaux Sans fil IEEE 802.11*, in "Colloque Francophone sur l'Ingénierie des Procoles (CFIP), Tozeur, Tunis", November 2006.

[30] R. RAZAFINDRALAMBO, I. GUÉRIN LASSOUS, L. IANNONE, S. FDIDA. *Dynamic Packet Aggregation to Solve Performance Anomaly in 802.11 Wireless Networks*, in "9th ACM/IEEE International Symposium on Modeling, Ananlysis and Simulation of Wireless and Mobile Systems (MSWiM), Terremolinos, Spain", October 2006.

[31] C. SARR, C. CHAUDET, G. CHELIUS, I. GUÉRIN LASSOUS. *Amélioration de la précision pour l'estimation de la bande passante résiduelle dans les réseaux ad hoc basés sur IEEE 802.11*, in "8es Journées Doctorales Informatique et Réseau (JDIR), Marne-la-Vallée, France", January 2006.

[32] C. SARR, C. CHAUDET, G. CHELIUS, I. GUÉRIN LASSOUS. *Improving Accuracy in Available Bandwidth Estimation for 802.11-based Ad Hoc Networks*, in "Third IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS), Vancouver, Canada", Short paper, October 2006.

[33] S. SOUDAN, P. VICAT-BLANC PRIMET. *Passerelle à base de processeurs réseaux pour le contrôle des flux de grille*, poster paper, Conference Francophone de l'Ingenierie des Protocoles, CFIP 2006, October 2006.

[34] P. VICAT-BLANC PRIMET, R. TAKANO, Y. KODAMA, T. KUDOH, O. GLÜCK, C. OTAL. *Large Scale Gigabit Emulated Testbed for Grid Transport Evaluation*, in "Proceedings of The Fourth International Workshop on Protocols for Fast Long-Distance Networks, PFLDnet'2006, Nara, Japan", February 2006, http://www.hpcc.jp/pfldnet2006/paper/s1_02.pdf.

[35] P. VICAT-BLANC PRIMET, R. TAKANO, Y. KODAMA, T. KUDOH, O. GLÜCK, C. OTAL. *Large Scale Gigabit Emulated Testbed for Grid Transport Evaluation*, in "PFLDNET06 International workshop on Protocols for Fat Long Distance Networks. Nara,", 2006.

Internal Reports

[36] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *Etude de de la problématique de la répartition de la charge réseau au niveau applicatif*, Research Report, France Telecom R&D Research report, March 2006.

[37] B. B. CHEN, P. VICAT-BLANC PRIMET. *A Flexible Bandwidth Reservation Framework for Bulk Data Transfers in Grid Networks*, Technical report, LIP ENS Lyon, INRIA RESO - inria-00078069, june 2006, https://hal.inria.fr/inria-00078069.

[38] B. GOGLIN, O. GLÜCK, P. VICAT-BLANC PRIMET. *Interaction efficace entre les réseaux rapides et le stockage distribué dans les grappes de calcul*, Also available as Research Report RR-5806, INRIA Rhône-

Alpes, Research Report, n^{o} RR2006-04, LIP, ENS Lyon, Lyon, France, January 2006, http://hal.inria.fr/inria-00070218.

[39] R. GUILLIER, L. HABLOT, Y. KODAMA, T. KUDOH, F. OKAZAKI, P. VICAT-BLANC PRIMET, S. SOUDAN, R. TAKANO. *A study of large flow interactions in high-speed shared networks with Grid5000 and GtrcNET-10 instruments*, Research Report, n^{o} 6034, INRIA, 11 2006, https://hal.inria.fr/inria-00116249.

[40] R. GUILLIER, L. HABLOT, S. SOUDAN, P. VICAT-BLANC PRIMET. *Evaluation des liens 10 GbE de Grid'5000*, Technical report, LIP ENS Lyon, INRIA RESO - RR-6047, November 2006, https://hal.inria.fr/inria-00118236.

[41] D. M. LOPEZ PACHECO, C. PHAM, L. LEFÈVRE. *XCP-i : eXplicit Control Protocol for heterogeneous inter-networking of high-speed networks*, Research Report, n^{o} 5946, INRIA, July 2006, http://hal.inria.fr/inria-00083955.

[42] D. LOPEZ PACHECO, C.-D. PHAM, L. LEFÈVRE. *XCP-i : eXplicit Control Protocol for heterogeneous inter-networking of high-speed networks*, Research Report, n^{o} RR2006-24, LIP, June 2006.

[43] S. SOUDAN, R. GUILLIER, L. HABLOT, Y. KODAMA, T. KUDOH, F. OKAZAKI, P. VICAT-BLANC PRIMET, R. TAKANO. *Investigation of Ethernet switches behavior in presence of contending flows at very high-speed*, Research Report, n^{o} 6031, INRIA, 11 2006, https://hal.inria.fr/inria-00115893.

[44] S. SOUDAN, P. VICAT-BLANC PRIMET. *Grid flow control: Prototype of a grid gateway based on Network Processors*, Technical report, LIP ENS Lyon, INRIA RESO - RR-5909, mai 2006, https://hal.inria.fr/inria-00071356.

[45] S. SOUDAN, P. VICAT-BLANC PRIMET. *Contrôle de flux de grille : Prototypage d'une passerelle de grille à base de Network Processors*, Technical report, n^{o} RR2006-18, RESO, LIP (UMR CNRS - ENS Lyon - UCB Lyon - INRIA 5668), LIP, Ecole Normale Supérieure de Lyon, FRANCE, May 2006.

Miscellaneous

[46] L. HABLOT, O. GLÜCK. *Evaluation et optimisation d'une implémentation de MPI*, Master Thesis also available as Research Report RR2006-26, INRIA Rhône-Alpes, Technical report, LIP ENS Lyon INRIA Reso, Juillet 2006, http://hal.inria.fr/inria-00090666.

[47] L. LEFÈVRE, J.-P. GELAS. *High Performance Autonomic Gateways for Large Scale Grids and Distributed Environments*, Poster INRIA Booth, Supercomputing 2006, Tampa Bay USA, November 2006.

[48] S. SOUDAN. *Partage de bande passante et plan de contrôle optique dans les grilles*, DEA2006-01, Technical report, ED MathIF, ENS-Lyon, June 2006.

[49] S. SOUDAN, P. VICAT-BLANC PRIMET. *Passerelle à base de processeurs réseaux pour le contrôle des flux de grille*, poster, CFIP 2006, October 2006.

References in notes

[50] P. ABRY, R. BARANIUK, P. FLANDRIN, R. RIEDI, D. VEITCH. *Multiscale nature of network traffic*, in "IEEE Signal Processing Magazine", vol. 19, 2002, p. 28–46.

[51] P. ABRY, P. FLANDRIN, D. VEITCH. *Internet : comment réguler le trafic ?*, in "La Recherche", n^{o} 384, March 2005, p. 50–53.

[52] N. AYARI, D. BARBARON, L. LEFÈVRE, P. VICAT-BLANC PRIMET. *A Survey on High Availability Mechanisms for IP Services*, in "HAPCW2005 : High Availability and Performance Computing Workshop, Santa Fe, New Mexico, USA", October 2005.

[53] A. BASSI, M. BECK, F. CHANUSSOT, J.-P. GELAS, R. HARAKALY, L. LEFÈVRE, T. MOORE, J. PLANK, P. VICAT-BLANC PRIMET. *Active and Logistical Networking for Grid Computing: the e-Toile Architecture*, in "The International Journal of Future Generation Computer Systems (FGCS) - Grid Computing: Theory, Methods and Applications", Elsevier B.V (ed),ISSN 0167-739X, vol. 21, n^{o} 1, January 2005, p. 199-208.

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

[55] M. CHAUDIER, J.-P. GELAS, L. LEFÈVRE. *Towards the design of an autonomic network node*, in "IWAN2005 : Seventh Annual International Working Conference on Active and Programmable Networks, Nice, France", November 2005.

[56] B. B. CHEN, P. VICAT-BLANC PRIMET. *Supporting bulk data transfers of high-end applications with guaranteed completion time*, 2007.

[57] V. FIRIOUS, J. LE BOUDEC, D. TOWSLEY, Z.-L. ZHANG. *Theories and Models for Internet Quality of Service*, in "IEEE", May 2002.

[58] S. FLOYD. *HighSpeed TCP for Large Congestion Windows*, in "Internet draft, work in progress", 2002, work in progress, http://www.icir.org/floyd/talks/floyd-tsvwg-Jul02.pdf.

[59] S. FLOYD, V. JACOBSON. *Link-sharing and Resource Management Models for Packet Networks*, in "IEEE/ACM Transaction on Networking", 4, vol. 3, August 1995.

[60] I. FOSTER, M. FIDLER, A. ROY, V. SANDER, L. WINKLER. *End to end Quality of Service for High End applications*, in "Computer Communications, special Issue on Network Support for Grid Computing", 2002.

[61] I. FOSTER, C. KESSELMAN. *The Grid : Blueprint for a new Computing Infrastructure*, in "Morgan Kaufmann Publishers Inc.", 1998.

[62] J.-P. GELAS, L. LEFÈVRE. *MoonGrid: Bring Processing Power to the Moon*, ISU Annual International Symposium: "Why the Moon?", Strasbourg, France, February 2007.

[63] C. JIN, D. X. WEI, S. H. LOW. *FAST TCP: motivation, architecture, algorithms, performance*, in "IEEE Infocom", March 2004.

[64] T. KELLY. *Scalable TCP: Improving Performance in Highspeed Wide Area Networks*, in "Protocol for Long Distance Networks Conference", n^{o} Pfldnet-1, February 2003.

[65] L. LEFÈVRE, J.-P. GELAS. *IAN2 : Industrial Autonomic Network Node architecture for supporting personnalized network services in industrial context*, in "Submitted to The International Journal of Future Generation Computer Systems (FGCS) - Grid Computing: Theory, Methods and Applications", 2007.

[66] L. LEFÈVRE. *Heavy and lightweight dynamic network services : challenges and experiments for designing intelligent solutions in evolvable next generation networks*, in "Workshop on Autonomic Communication for Evolvable Next Generation Networks - The 7th International Symposium on Autonomous Decentralized Systems, Chengdu, Jiuzhaigou, China", ISBN : 0-7803-8963-8, IEEE Society, April 2005, p. 738-743.

[67] L. LEFÈVRE, A. SAROUKOU. *Active network support for deployment of Java-based games on mobile platforms*, in "The First International Conference on Distributed Frameworks for Multimedia Applications (DFMA'2005), Besancon, France", IEEE Computer Society, February 2005, p. 88-95.

[68] K. PARK, W. WILLINGER. *Self similar traffic analysis and performance evaluation*, Wiley, 2000.

[69] V. SANDER. *Networking issues of GRID Infrastructures*, in "GRID Working Draft of the GRID High-Performance Networking Research Group, Global GRID Forum", 2003.

[70] V. SANDER, F. TRAVOSTINO, J. CROWCROFT, P. VICAT-BLANC PRIMET, C. PHAM. *Networking Issues of Grid Infrastructures*, Technical report, october 2004, http://forge.gridforum.org/projects/ghpn-rg/.

[71] D. SIMEONIDOU. *Optical Network Infrastructure for Grid*, in "Grid Working Draft of the Grid High-Performance Networking Research Group, Global GRID Forum", 2003.

[72] M. TAQQU, W. WILLINGER, R. SHERMAN. *Proof of a fundamental result in self similar traffic modeling*, in "Computer Communication Review", vol. 27, 1997, p. 5–23.

[73] H. TOBIET. *Panorama des réseaux utilisés et services à valeur ajoutée TEMIC*, Delivrable D2.1 Projet RNRT Temic, Technical report, February 2005.

[74] J. TOUCH, K. KOBAYASHI, P. VICAT-BLANC PRIMET. *Special issue "Hot topics in Transport Protocols for Very Long distance networks"*, in "International Journal of Computer Networks (COMNET)", january 2007.

[75] P. VICAT-BLANC PRIMET, O. GLÜCK, C. OTAL, F. ECHANTILLAC. *Emulation d'un nuage réseau de grilles de calcul: eWAN*, Research Report, n^{o} RR2004-59, LIP, ENS Lyon, Lyon, France, December 2004, http://www.ens-lyon.fr/LIP/Pub/Rapports/RR/RR2004/RR2004-59.pdf.