GIP Agence Nationale de la Recherche

Programme de Recherche "Calcul Intensif et Grilles de Calcul"

Appel à projets 2005
Annexe : Formulaires de dépôt de dossier

Date limite de déclaration d'intention de déposer un dossier (obligatoire) :

Lundi 11 Juillet 2005 à 12h

Date limite de dépôt des dossiers :

Vendredi 9 Septembre 2005 à 12h

Contact (pour toute demande d'information) :

François ROBIN, CEA : anr-ci@cea.fr

Composition du dossier de réponse à l'appel à projets 2005
"Calcul Intensif et Grilles de Calcul"

Pour être complet, **le dossier soumis doit obligatoirement comporter les éléments suivants** :

Fiche de synthèse	Une seule	Rédigée par le coordonnateur du projet
Présentation du projet	Une seule	Rédigée par le coordonnateur du projet
Fiche de synthèse des demandes d'aide	Une seule	Rédigée par le coordonnateur du projet

Fiche partenaire privé	Une par partenaire privé	Rédigée par le partenaire
Fiche demande d'aide partenaire privé	Une par partenaire privé	Rédigée par le partenaire

Fiche partenaire public	Une par partenaire public	Rédigée par le partenaire
Fiche demande d'aide partenaire public	Une par partenaire public	Rédigée par le partenaire

Le dossier pourra être rédigé en langue anglaise.

Le coordonnateur du projet (qui est par convention le partenaire n.1 du projet) doit veiller à la cohérence, à la qualité et à la clarté du dossier. Celui-ci doit être fourni (voir également le texte de l'appel à projets) :

- par courrier électronique : en version électronique[1]
    - au format PDF[2]
    - au format XLS pour le tableau de synthèse des demandes d'aide,
    - au format RTF[3] (pour les autres tableaux et documents à fournir)
- par courrier postal : sous forme papier.

On trouvera à la fin de ce document les notes explicatives correspondant aux notes dans les tableau.

En cas de labellisation, le titre et le résumé du projet, le coordonnateur, la liste des partenaires impliqués (publics et privés), le montant total du projet et la durée du programme ont vocation à être publiés par le GIP ANR, notamment sur un site web. En déposant un dossier, les partenaires acceptent implicitement cette règle.

L'acronyme du projet doit être le même que celui indiqué lors de la déclaration d'intention sauf si le GIP ANR a imposé la modification de celui-ci. Dans ce cas, c'est l'acronyme modifié qui doit être utilisé.

---

1 Afin d'éviter toute confusion lors de la réception des propositions de projets par mail, les fichiers envoyés en pièce-jointe devront avoir un nom commençant par xxxx_ avec xxxx = l'abréviation en 4 caractères de l'acronyme du projet.

2 Un document PDF unique contenant dans l'ordre l'ensemble des éléments demandés (voir tableau ci-dessus) avec dans la mesure du possible un "signet" par élément du document.

3 Un ficher RTF par élément demandé.

Programme de Recherche
"Calcul Intensif et Grilles de Calcul"
Appel à projets 2005 – Fiche de Synthèse

Acronyme du projet : VIF

Titre du projet (2 lignes maximum)

Infrastructures virtuelles pour le calcul haute performance dynamique.

Virtual Infrastructures for high performance dynamic computing

Thématique de l'appel à propositions concernée par le projet proposé (cocher la ou les case(s)) :

- ❑ Grands défis applicatifs
- ❑ Méthodes et applications de la simulation numérique
- ❑ Maîtrise des architectures matérielles et logicielles avancées
- X Données et grilles de calcul

Résumé du projet (10 lignes maximum)

L'objectif de ce projet est d'explorer une approche en rupture avec les principes actuels (orientation «services») développés dans les grilles afin d'en améliorer conjointement la sécurité, le contrôle des performances et la tolérance aux pannes. Il s'agit d'étudier des grilles simples intégrant les apports des récents standards de communication. VIF mettra en oeuvre un nouveau paradigme de contrôle des ressources de calcul et du réseau d'une grille pour instancier le concept de clusters virtuels. L'infrastructure globale - aggrégation des resources matérielles distribuées- sera ainsi partitionnée dynamiquement en infrastructures virtuelles isolées, protégées et multiplexées dans le temps et dans l'espace pour répondre à des besoins de calcul spécifiques. Ce projet s'appuie sur les principes de virtualisation conjointe du réseau et du système d'exploitation, de contrôle du partage de bande passante, d'identifiants cryptographiques ainsi que sur les protocoles Ipv6 et HIP.
La proposition sera validée et évaluée sur l'instrument national Grid5000 avec des applications biomédicales, exigeantes en sécurité , en performance et tolérance aux pannes. Le modèle VIF sera comparé aux modèles Globus et PlaneteLab et il sera montré que les applications fonctionnent sur VIF de manière quasi transparente, comme sur une grappe étendue.

The goal of this project is to explore an approach in a break with current servicesoriented principles developed in grids to jointly enhance their security, performance controllability and fault tolerance. The aim is to build and study simple grids, based on efficient concepts by integrating recent communication standards. VIF will develop a new paradigm for resource control in grids to instantiate the concept of confined virtual clusters.
The global infrastructure (computers, disks, networks) in thus partitioned in virtual infrastructures (VIF) dynamically composed. These VIF are multiplexed in time and space and are isolated and protected.. The project is based on network and system virtualization, bandwidth sharing, cryptographic identifiers concepts and on the IPv6 and HIP protocols.
The proposal will be validated and evaluated on the Grid5000 testbed with biomedical applications, demanding in security , performance and reliability. Comparison with Globus and Planetlab approaches and demonstration of functional tranparence for the application will be the goals of the experiments.

Coordonnateur du projet

(nom, prénom, appartenance, coordonnées (messagerie électronique, adresse, téléphone, télécopie)

Pascale Vicat-Blanc Primet.

INRIA - Equipe RESO
Laboratoire d'Informatique du Parallélisme : LIP
UMR CNRS - INRIA - ENS - UCBL N° 5668
Ecole Normale Superieure
46, Allee d'Italie 69364 LYON Cedex 07 FRANCE
Tel: 33 (0) 4 7272 8802
Fax: 33 (0) 4 7272 8080
e-mail : Pascale.Primet@inria.fr,
web : http://perso.ens-lyon.fr/pascale.primet

Partenaires de la recherche publique et de la recherche privée			
1	*INRIA*	*7*	
2	*I3S*	*8*	
3	*HealthGrid*	*9*	
4		*10*	
5		*11*	
6		*12*	

Durée du projet (en mois) : 3 ans	**Montant total du projet** (HT) :	656k€
dont laboratoires publics EPIC (ADEME, BRGM, CEA, CNES, CSTB, IFREMER, ONERA, …) :		k€
dont laboratoires publics EPST (CNRS, INRA, INRETS, INRIA, INSERM, Universités, .. …) : et : hommes-mois[i]		512k€
dont **partenaires privés** :		144k€

Cadre réservé à l'administration

Résumé du contexte et de la motivation du projet (voir B2):

Moving one step further from computer clusters, grid computing is a promising technology that brings together geographically distributed resources to build a very high-performance computing environment for data-intensive or computing-intensive applications. Ten years ago, it has been recognized that current Internet technologies did not address the issues that Grid concept raises. The Grid community hardly tried -via solutions like Globus- to find an integrated approach, over current TCP/IP technology, to coordinated use of resources at multiple sites for computation. Grid technology seems today to break her promise. In this project, we aim at exploring a new virtualisation approach and at demonstrating at large scale how new Internet and system technologies, like Ipv6, HIP, Overlays and Virtual Private Network solutions, advanced security and resource sharing protocols, OS virtualization, may help in building trust, reliable and optimal environments for secure and performance-assured computing.

Résumé des retombées scientifiques et techniques attendues (voir B4) :

The VIF's approach may ease a large scale deployment of grid technology by the fact that it enables transparent Grid applications development. A comparison of VIF approach with Globus and PlaneteLab solutions will help in understanding their respective advantages and limits. The production of the project will be analyses, large scale demonstrations, quantitative evaluation and open source softwares development. The work will be published and demonstrated nationaly and internationaly in Grid and networking conferences and journals.
Reusable building blocks and a full Ipv6 integrated solution will be developed, disseminate in Grid community, and if meaningful, proposed for standardization in GGF. Moreover, stand-alone grid benchmarks for security, performance-control and dependability evaluation of this solution, systematic use cases and demonstration of biomedical grid applications on the Grid5000 testbed will increase grid evaluation knowledge, software and experience.
The project will also leaverage the knowledge in virtualization, grid security and grid networking and will increase the expertise in Grid5000 and GdX usage.

Résumé des retombées industrielles et économiques escomptées (le cas échéant) (voir B5)

HealthGrid partner will disseminate the scientifical and technical results in the large grid user community it represents. When the scientifical and technological results are convincing, the consortium will put a particular effort in developping advanced prototypes. Thanks to these demonstrators and the associated scientific outcome, the academic partners will contact societies potentially interested by the technology for a possible valorization. Because of the timeliness of the present project and high demand in integrated grid management solutions, we believe that the industrial and economical outcomes of the projet are high.
Indirectly, the results of this exploration will help the industrial community to better understand the potential and the limits of grid technology and Internet computing, (ie risk evaluation) and may help users in making appropriate choice for their computing resource investment.

Tableau récapitulatif des délivrables (voir B.6) :

Deliverable name	Deliverable type	Content	Responsible	Participant	Due date
D1.1	document	State of the art : Network , OS and Grid virtualization: approaches and technologies	INRIA (Gd large)	INRIA (Grand large, Reso, Planete)	T+6
D2.1	document	Case s and Requirements studies for VIF	I3S	All	T+6
D1.2	document	VIF general design and interfaces document	INRIA (Reso)	INRIA (Reso, Grand large, Planete)	T+12
D2.2	document	Application Use case design document	HealthGrid	(HealthGrid, I3S)	T+12
D1.3	document	Security design document	INRIA (Planete)	INRIA (Planete et Reso)	T+12
D1.4	document	Fault Tolerance design document	INRIA (Gd Large)	INRIA (Grand large, Reso)	T+12
D1.5	document	Performance control design document	INRIA (Reso)	INRIA (Grand large, Reso,, Planete)	T+12
D1.6	document	VIF general implementation document	INRIA (GLarge	INRIA (Grand Large, Reso, Planete	T+24
D2.3	software	Use cases software	Health Grid	HealthGrid, I3S	T+24
D1.7	software	Security software	INRIA (Planete)	INRIA (Planete et Reso	T+24
D1.8	software	Fault Tolerance software	INRIA (Grand Large)	INRIA (Grand large, Reso	T+24
D1.9	software	Performance control software	INRIA (Reso)	INRIA (Grand large, Reso,, Planete	T+24
D1.10	demonstration	VIF model demonstration	INRIA (Reso)	INRIA (Grand large, Reso,, Planete	T+32
D2.4	demonstration	Applications running on VIF demonstration	I3S	all	T+32
D1.11	document	Security experimentation and results document	INRIA (Planete)	INRIA (Planete et Reso	T+36
D1.12	document	Dependability experimentation and results document	INRIA (Gd large)	INRIA (Grand large, Reso	T+36
D1.13	document	Performance experimentation and results document	INRIA (Reso	INRIA (Grand large, Reso,, Planete)	T+36
D1.14	document	VIF Application and comparison	Health Grid	I3S, Health Grid	T+36

B.1. Project title

VIF
Infrastructures virtuelles pour le calcul haute performance dynamique.
Virtual InFrastructure for high performance dynamic computing

B.2. Context and motivation

Moving one step further from computer clusters, grid computing is a promising technology that brings together geographically distributed resources to build very high-performance computing environment for data-intensive or computing-intensive applications[4]. It has been demonstrated through large world wide programs (DataGRID, BioGrid, EGEE, LCG, GridWorld[5]..) that many future high end applications will require such high performance dynamic infrastructures.

One major requirement for grid computing is then to control precisely how shared resources are used to provide users guarantees jointly in terms of **security, dependability and performance**. However, these requirements are not provided by current service oriented grid technologies based on vanilla Internet protocols. Over the past ten years, the Grid community has proposed security solutions that support management of credentials and policies when computations span multiple institutions, and resource management protocols and services that support secure remote access to computing and data resources. But **these solutions are heavy to develop and use, not fully secured and do not give performance insurance**. The problem of performance guarantees remains a big issue because grid applications are highly likely across multiple networks, and because multi-domain quality of services features involving IntServ and DiffServ QoS architectures are still under development and/or construction in IP networks.
Consequently some people start to think that grid technology break her promise. For us, Grid will become a very valuable approach for large scale high performance computing only if we **reconsider the initial technological and design choices in the light of Internet and operating system technology evolutions**. The best effort Internet philosophy was not defined for such intensive applications and such sensitive contexts and the one fit all solution could not correspond to such strong constrains ! Aiming to work with very high speed networks or to take the security constrains from the beginning in an efficient manner, the enhancement on Internet and communication protocols has been intensively pursued during the last years. For us, integrating these innovations in grid software will largely simplify their application development and improve performance.

B.3. Project description

The VIF project aims at exploring **new paradigms for gathering, exploting and controlling the sharing of distributed resources for high end applications**. It is an approach in a break with current service oriented principles developed in grids. The aim is to build and study **simple grids, based on efficient concepts and integrating advances in new**

---

4 All scientifical references are given in annexe 4.
5 more than 45 projects referenced at http://www.gridforum.org/ggf_grid_projects.htm

Internet and system technologies like Ipv6, HIP, Overlays and Virtual Private Network solutions, advanced crypto-based-security and resource sharing protocols, OS and machine virtualization. The goal is to reuse and combine existing mechanisms that have been proposed for new generation Internet and that will not require new APIs.

This project is thus in line with the ANR CIGC program as it focuses on the **optimization of software architecture** for high performance dynamic grid computing and as it aims at **developing and demonstrating at large scale innovative biomedical applications**. Indeed, many medical image analysis applications, such as epidemiology, statistical analysis, or validation of medical procedures, are facing challenging needs related to the amount of data to process and the constraints specific to medical data (privacy, widespreading...).
In this project, we demonstrate how advanced and innovative grid infrastructures enable new medical applications for which the data manipulation and computing cost are usually show-stoppers. The application deployed will be accessible in the medical community to ensure that the end-users get benefits of the developments performed. This project fits in themes *"Données et grilles de calcul "* of the program and aims at optimizing and simplifying the usage of shared resources.

Three types of locks - **security, performance-control and fault tolerance** - generally separatly examined in grids will be here **jointly studied** and, if possible, combinely solved.

Then the goal of the project is to define and to develop an integrated model for dynamic grid infrastructures design and control based on recent results from the networking and system communities. Briefly, the solution we propose to develop and to study will enable a user, who has access to an hardware grid infrastructure, to preempt a pool of resources for a given time window to execute his application by composing a dedicated virtual cluster. **The global infrastructure (computers, disks, networks) in thus partitioned in virtual infrastructures (VIF) dynamically composed**. These VIF are multiplexed in time and space and are isolated and protected. They constitute trusts domains. The project is based on **network and system virtualization, bandwidth sharing, cryptographic identifiers concepts and on the IPv6 and HIP protocols** *(HIPernet model)*.

-

The VIF project aims at developing, testing and demonstrating the virtualization approach at large scale. Virtualization is becoming a key feature of Grid where it essentiall provides the feature of abstracting some specific characteristics of the Grid infrastructure. The main motivations of machine virtualization in the context of Grid and large scale distributed system experimental platforms (PlanetLab, Grid eXplorer) are the following: a) each VM provides a confined environment where non-trusted applications can be run, b) VM allows to establish limits in hardware resource access and usages, through isolation techniques, c) VM allows adapting the runtime environment to the application instead of porting the application to the runtime environment, d) VM allows using dedicated or optimized OS mechanisms (scheduler, virtual memory management, network protocol) for each application, e) Applications and processes running within a VM can be managed as a whole.
A comparison with Globus and PlaneteLab approaches will help in understanding their respective actual limits and could ease a large scale deployment of grid technology via a simpler approach for resource and network control that enable a Grid application development similar to cluster application development. Development of stand-alone grid benchmarks for security, performance-control and dependability, systematic use case development, performance and security evaluation, **demonstration of data and computing intensive biomedical grid applications running on this middleware on the Grid5000 testbed** will be pursued. Finally, as Ipv6 will replace Ipv4 in the near future and as it has features for security and flow control allowing an elegant implementation of virtual networks, it becomes evident to us that all our concepts and approach have to be explored and developed in **a fully IPv6-based grid context**, what will be a « first» (at least in France).

Medical image analysis applications are a growing candidate for grid infrastructure usage and validation. For several years, the medical imaging community faces new challenges coming from

- the **amount of data** to process (the yearly production of images in a radiological department sums up to tens of TB of data);
- the **distribution of data** (widespreading of data sources over the territory);
- the **heterogeneity of data** (lack of well established standards);
- **confidentiality** of data to preserve patients privacy (medical data);
- and the complex data structure (medical files).

Data grid have been identified as a tool able to process large medical databases, enabling large scale applications such as epidemiology, statistical study over populations, medical simulation or research on rare diseases. The complex nature of medical images makes of the medical image processing area one of the most challenging and constraining application area for grid infrastructures. Two medical image analysis applications have been identified that will be used to validate our approach, measure and demonstrate the benefit of the middleware proposed.

Details of the backgroud and issues and ideas developed in the different tasks the VIF project proposes to achieve are given in annexe 1, 2, 3, all fundamental background references are in annexe 4 and on our website : http://www.ens-lyon.fr/LIP/RESO/VIF.

B.4. Retombée scientifiques et techniques

The VIF's approach may ease a large scale deployment of grid technology by the fact that it enables transparent Grid application development. A comparison with Globus and PlaneteLab solutions will help **in understanding their respective advantage and limits**.

The production of the project will be theoretical, qualitative and quantitative analysis, large scale demonstrations, security and performance evaluation and open source software development. The work will be **published and demonstrated nationaly and internationaly** in Grid and networking conferences and journals (Supercomputing, iGrid, Cluster, HPDC…). **Reusable building blocks and a full Ipv6 integrated solution will be developed, disseminated in Grid community in France through the Grid5000 projec**t and outside through the respective international collaborations of the partners (Europe, Japan,USA) , and if meaningful proposed for standardization in GGF.

Stand-alone grid benchmarks for security, performance-control and dependability, systematic use cases and **demonstration of biomedical grid applications on the Grid5000** testbed will provide grid evaluation knowledge, software and experience.

The project aims also at leaveraging the knowledge in virtualization, grid security and grid networking and will increase the expertise in Grid5000 and GdX usage.

B.5. Retombées industrielles et économiques escomptées (le cas échéant)

HealthGrid partner will disseminate the scientifical and technical results **in the large HealthCare user community** it represents. Our approach, aiming the best common denominator abstraction level for security and network resource control, is expected to be much lightweight than classical Globus and more scalable and computing oriented than Planetlab. Moreover, grid application development will be potentially straightforward or similar to cluster application development.

When the scientifical and technological results are convincing, the consortium will put a particular effort in developping advanced prototypes. Thanks to these demonstrators and the associated scientific outcome, the academic partners will contact societies (like Alcatel) potentially interested by the technology for a possible valorization. Because of the timeliness of the present project and high demand in integrated grid management solutions, we believe that the industrial and economical outcomes of the projet are high.

Indirectly, the results of this exploration will help the industrial community to better understand the potential and the limits of grid technology and Internet computing, (ie risk evaluation) and may help users in making appropriate choice for their computing resource investment.

B.6. Organisation et pilotage du projet

The project is divided in **two workpackages**. WP1 comprises the virtual infrastructures software design and development and WP2 concerns the virtual infrastructure software and concept evaluation.
The first workpackage is composed of three parts : a) state of the art analysis, b) solution design and c) solution development. Security aspects, performance aspects and dependability aspects will be examine independantly and jointly throughout this workpackage. Scientifical reports, software codes and demonstrations and documentations will be produced for each aspects and for the integrated model.
The second workpackage is divided in two parts : a) validation's scenario design and application adaptation, b) scenario deployment and evaluation in Grid5000 testbed. Security evaluation , performance evaluation and dependability evaluation scenarii based on two applications coming from the Biomedical field will be created and effectively executed and measured on the Grid5000 testbed.
These two workpackages are largely dependent and the teams will work in close interaction.
Globally, **the three years project is divided in three phases**: 1) problem and state of the art, evaluation techniques analyses ; 2) design and development of model and evaluation scenario 3) validation and experimentation.
A large part of the project man power and time resource will be allocated to experimentation and demonstrations development and perform on the Grid5000 as it is a real challenge of the project.

WP1 VIF Dynamic Infrastructure design- INRIA

WP1.1. Study of system and network virtualisation in Grids

The goal of this task will be to finely analyze the advantages, drawback and feasability of different kind of VM and VPN solutions for solving grid security, performance and dependtate of the art in terms of network, security and OS virtualisation, a technical implementation study, and a qualitative evaluation of their usage in grids will be produce. Then machine virtualization scalability, as is it an issue in grids will be studied. (see annexe1)

WP1.2. VIF model design and development

The results of the previous tasks will lead to the functional and architectural design of the VIF model. Network virtualization will permit multiple overlays to cohabit on a common shared infrastructure. All networks overlays and nodes will be kept isolated both at the network (Ipsec and bandwidth sharing control) and OS (filesystems, IPCs, etc.) level to insure security and performance. The next tasks concentrate on grid security and grid performance-control and dependability aspects of the model.

WP1.3 Security Issues and Propositions

The INRIA Reso team proposes to leverage on network (and sub-transport) layer security (IPsec, layer 3 and HIP, layer 3.5), self-certifying naming (i.e. CBID/HIT) and SPKI delegation certificates to implement the same security functionality the Globus GSI proposes. We believe that layer 3.5 approach allows for the maximum of genericity, because its "lowest layer" implementation it constitutes a least common denominator to deployed grid security infrastructure while keeping the interoperability properties the IP technology offer (comparing to layer2 VPN solutions). The Crypto-Based Identifiers (CBIDs) family is already used by networking community to help to solve several issues in the IPv6 world. One goal of this task will be to deeply analyze and design CBID and SPKI-based solutions for the Grid. INRIA Planete and Reso team will work on this aspect in concert.

One step further we want to explore solutions that combine high performance and security in communication. Indeed, high rate bulk data transfers to one or multiple destinations (through multicast/broadcast technologies in that case) raises specific challenges in terms of security. Packet source authentication and packet integrity are the basic services. Solutions based on shared group keys, if they are efficient from a computing perspective, are not sufficient in case of group communications since any group member may forge any packet. Solutions based on digital signatures with asymmetric key cryptography are not acceptable too because of prohibitive processing loads. In the VIF project, the INRIA Planète's standardization activity around TESLA (Timed Efficient Stream Loss-tolerant Authentication) a promising alternative will be continued, an open source implementation, with all the required features will be developed, integrated in the VIF middleware, and its performances evaluated in operational environments. In parallel, work will be carried out on the BiBa alternative.
The INRIA Planète team will also investigate on high performance cryptographic primitives, more precisely on joint FEC (Forward Error Correction) / cryptography techniques for Grid transmissions. If the performances are good, it may offer an alternative to VPN encryption in some target environments.

WP1.4 Performance- control Aspects

High performance controlled transfers are a key for high performance computing. In the in the context of the VIF approach, the INRIA Reso team proposes to develop an end to end hybrid network sharing model that combines classical on line IP QoS solutions with off-line advanced reservations. Here also we propose to leverage on network (and sub-transport) layer for performance control. The model we propose to evaluate maps the low volume, delay sensitive flows into existing classical prioritized QoS classes, the normal flows into standard best-effort class and isolates the large volume throughput sensitive ones with advance reservation techniques. Within each multi-point secured VPN, virtual sub-channels with deterministic performance for each type of grid traffic are defined. This network sharing approach take benefice from off-line scheduling for optimizing the transfer of huge data files and maximizing the allocated network resource utilisation while protecting other traffics from dangerous "elephants".
This task will look deeply at the following points :
- Solution to transparently differentiate the grid traffics from the source will be studied. For advance reservations the upper level reservations will be inegrated within workflow engine used by application. Such feature development involves both infrastructures and application teams (INRIA Reso and I3S).
- Tthe global network resource has to be divided from end to end into virtual sub-channels. As the interconnection comprises public and private networks, the provisionning and the flow differentiation will have to cross all the heterogeneous domains till end hosts application level. For the costly wide area access, dynamic bandwidth allocation may help in optimizing network resource provisionning for bulk transfers. This project concentrates on the private part of the network and explores the bandwidth sharing and flow differentiation problem from host's virtual OS to the WAN access equipment.
INRIA Reso develop the associated bandwidth allocation algorithms and the traffic control mechanisms in the Cariocas project of the IDF pôle de competitivite « Simulation Numerique » that has been recently labelled. All the valuable results and development will be reused, whenever is it convinient in the VIF project.

WP1.5 Dependability Aspects

Because of the virtualization principle, dependability covers at least two aspects: resource isolation and fault tolerance. First, users sharing the same resources should get a fair access to them. The notion behind this property is the virtual machine isolation.
Isolation may have a complex definition: 1) when several users request the same workload on the same resources, the isolation should ensure that no user is slowed down because other users are using the resources too much. 2) when a resource is requested by a subset of users, this subset should be able to share 100% of the resource capacities. In this task, INRIA Gd Large and Reso will design a benchmark suite and run experiments on a large variety of virtualization technologies in order to check the isolation of every computer resource independently. Currently, only results on global isolation are available and some results cannot be explained without a finer evaluation. If the isolation is found poor, we will investigate, design and develop mechanisms required to force a good resource isolation between user contexts. A second important aspect of virtualization, not yet addressed in the literature, is the fault tolerance of virtual machines. When deploying an application on thousands of processor running virtualization software, several issues raise: 1) the stability level is supposed lower than with standard OS, because virtualization tools have stronger limitations and some tools still need improvements 2) how to save the context of virtual OS? How to migrate virtual a OS context? these issues have already been discussed in the literature but not evaluated in the Grid context (there are new issues related to performance and security). 3) how to combine processor and network fault tolerance in the context of virtualization? Basically this should involve extending the context to virtual network and provide co-migration of network and OS virtual contexts. In summary, this task will investigate several key issues related to dependability in the context of network and OS virtualization in Grids.

<u>WP2 VIF Evaluation with BioMedical application and Grid5000 – I3S</u>

<u>WP2.1 : Use cases1 : Data intensive workflow for medical image databases analysis.</u>

The I3S team will develop an application use case for validating medical image registration algorithms. The *registration* of medical images is a key basic algorithm needed for many image processing algorithms (see annexe 1 for details)
It will be deployed through a workflow manager able to deal with the complexity of the processings (registration of different image pairs with several algorithms and intermediate transformation and comparison of results obtained) and to enact the computations efficiently on a grid infrastructure. In this project, the application will be deployed on the grid5000 hardware infrastructure and it will benefit from the secured and high performance transport services deployed. The role of this application will be to:

- Define precisely application requirements for the underlying infrastructure, especially considering the data transport and confidentiality needs.
- Develop realistic deployment scenarii on data sets distributed over different sites.
- Deploy the application on the infrastructure and adapt our workflow engine to control the computations.
- Optimize the performance of the workflow engine by taking into account the data transfer cost when scheduling computations.
- Demonstrate the usage of the infrastructure in real scale to validate the benefits in terms of performance and the interest in the medical image processin area.

<u>WP2.2 : Use cases2 : Radiotherapy treatments modelisation</u>
From a technological point of view, radiotherapy is a highly complex procedure, involving a variety of computational operations for data gathering, processing and control. Presently patients are treated with standardised radiation doses. Tissue density provided by medical imaging devices is used to calculate the dose delivered by photon and electron beams. Accuracy of Monte Carlo (MC) dose computation is awesome, provided that the computing power is sufficient to allow for enough runs to reduce the statistical noise. The Grid is a

natural alternative to costly parallel computers. In this way, MC dose computations could become standard for radiotherapy quality assurance (QA), planning, and plan optimisation years before individual departments could afford a local investment that is capable to support MC. Requirements needed for such deployment include the existence of a service level agreement between the departments and the Grid providers by which the Grid level of performances in terms of security, stability and response time is guaranteed.

HealthGrid has deployed of Monte-Carlo codes on several grid infrastructures but there are additional requirements in order to set-up a service for hospitals:

- Guaranteed response time to the hospital asking for the computation is needed. This response timeincludes time to move information on the grid, to execute the tasks to collect the outputs and concatenate it
- Medical data must be kept anonymous outside the hospital and medical images should be encrypted
- The infrastructure must be robust enough to always send back the computation in the guaranteed response time.

These three requirements not yet achieved on the existing grid infrastructures will be evaluated by HealthGrid in the VIF context.

WP2.3 : Demonstration and Experimentations on Grid5000

The proposed approach for resource security and control is a strong shift from the current situation in Grid where sites are basically running Globus or Grid middleware based on classical security and control systems. To test this approach at large scale, we need a platform that can be deeply reconfigured to run the IPV6 network protocol, virtualization technologies on the network and computer hardware. Currently, only Grid'5000 is providing such a possibility. So all tests will be conducted on this platform. Thanks to the reconfiguration capability of Grid'5000, we will be able to compare
the proposed approach to the existing ones (Globus, Unicore). From the networking point of view, Grid5000 is the ideal testbed, as it is based on a private network (MPLS L2VPN on RENATER3 and dedicated lambda on dark fiber in RENATER4). Such a well-provisionned and protected core network enable complete and transparent Ipv6 deployment and a fine control of achieved performances over the bottlenecks.

For comparing virtuel machine environments, for evaluating integrated network and OS virtualization performance, bandwidth sharing control algorithms or security approaches clear metrics and methodology will be defined and microbenchmarks will be developed. Metrics envisionned are for example the overhead of the virtualization technique (in terms of CPUs, memory, disk and network utilization), the evolution of the global performance when the number of virtual machines increases, the isolation between virtual machines (also in terms of CPUs, memory, disk and network utilization) and the virtual machine creation time. All these benchmarks will be executed in the Grid5000 testbed. In this task, a set of precise use case scenarii based on the biomedical applications will be defined and performed in Grid5000 with three environments: VIF, Globus and PlaneteLab. Applications and middleware runing in Grid5000 testbed will be demonstrate.

The table bellow sumarizes the tasks that are allocated to the different workpackages. The names deliverables are name as D1.X is related to WP1, D2.x concerns WP2.

Task	Deliverable name	Deliverable type	Content	Responsible	Participant	Due date
WP1.1	D1.1	document	Virtualization State of the art	Grand large	INRIA	T+6
WP2.1	D2.1	document	Application Requirements for VIF	I3S	All	T+6
WP1.2	D1.2	document	VIF general design and interfaces	RESO	INRIA	T+12

			document			
WP2.1	D2.2	document	Application Use case design document	HealthGrid	HealthGrid, I3S	T+12
WP1.3	D1.3	document	Security design document	Planete	INRIA	T+12
WP1.4	D1.4	document	Dependability design document	Grand Large	INRIA	T+12
WP1.5	D1.5	document	Performance control design document	RESO	INRIA I3S	T+12
WP1.2	D1.6	document	VIF general implementation document	Grand Large	INRIA	T+24
WP2.1	D2.3	software	Use case 1 software	I3S	I3S	T+24
WP2.2	D2.4	software	Use case 2 software	Health Grid	Health Grid	T+24
WP1.3	D1.7	software	Security software	Planete	INRIA	T+24
WP1.4	D1.8	software	Dependability software	Grand Large	INRIA	T+24
WP1.5	D1.9	software	Performance control software	RESO	INRIA	T+24
WP2.3	D1.10	demonstration	VIF model demonstration	RESO	INRIA	T+32
WP2.3	D2.4	demonstration	Applications running on VIF demonstration	I3S	All	T+32
WP2.3	D1.11	document	Security experimentation and results document	Planete	INRIA	T+36
WP2.3	D1.12	document	Dependability experimentation and results document	Grand Large	INRIA	T+36
WP2.3	D1.13	document	Performance experimentation and results document	Reso	INRIA	T+36
WP2.3	D1.14	document	Application and comparison results	Health Grid	I3S, Health Grid	T+36

The consortium gathers a small team of experts in next generation Internet technology (Vincent Roca, Claude Castellucia) , Grid instrastructures and core middleware (Franck Cappello, Pascale Vicat-Blanc Primet) and grid application development (Johan Montagnat, Yannick Legre). All the teams involved are also largely recognized for their past or actual strong involvment in large scale grid programs and international networking and grid communities.
The consortium is leaded by **Pascale Vicat-Blanc Primet**, graduated in Computer Science, (http://perso.ens-lyon.fr/Pascale.Primet) who is **senior researcher at INRIA,** (see annexe 4) and who has been involved in many large national and international grid projects and initiatives (DataGRID, DataTAG, VTHD++, eToile, Grid5000, GdX, GGF).
The project is structured in two groups of partners: infrastructure teams and application teams that will work in concert.
Infrastructure teams: the INRIA partner - RESO, GRAND LARGE and PLANETE projects - composes the networking and system scientific and technical expertise.
The involved teams of INRIA Grand-Large and INRIA RESO already work together since several years in two projects: ACI "masse de données" Grid eXplorer and GRID5000.

The application partners Healthgrid and I3S gather a strong experience in developing and evaluating grid applications in current grid environment like Globus, DataGrid or EGEE.
These application partners use to work with INRIA Reso through previous EU projects (DataTAG, DataGRID), national project (eToile) and they have published collaborative papers.

From a practicall point of view, the VIF steering committee is composed by the scientific leaders of each partner. The technical team, gathering any member charge of a development and experimentation task (permanent or CDD researchers, permanent or CDD engineers, PhD students) will be leaded by a permanent researcher or engineer in order to insure the consistency and the perenniality of the achievements and of the technical knowledge.

B.7. Propriété intellectuelle :

As all partners are french and in order to keep things clear we write this section in french :
Principes de l'accord de propriété intellectuelle
Les règles de propriété intellectuelle seront fixées au sein d'un document contractuel de partenariat indépendant de la présente proposition (protocole d'accord), une fois que la présente proposition sera acceptée par l'organisme auprès duquel le projet fait l'objet de la présent proposition. En tout état de cause, les principes de propriété intellectuelle qui s'appliqueront sont ceux donnés en annexe 5.

Apports spécifiques :
Nous donnons ci-dessous les éléments concernant les apports spécifiques des différents partenaires à ce projet.
Apports logiciels :
Les software qui seront utilisés (et produits) sont (et seront) sous licence de type logiciel libre, comme c'est la pratique chez l'ensemble des partenaires (voir http://www.inrialpes.fr/planete/people/roca/mcl/). Il s'agit de logiciels développés par les partenaires ou du domaine public comme des implémentations HIP de Boeing sur LINUX et Ericsson sur FreeBSD, de l'interface HIP/OpenDHT pour la resolution HIT vers IP, des environnements de virtualisation XEN, VMware ou UML, FLUTE/ALC.

Seuls les certains algorithmes de traitement d'image propriétaires seront utilisés pour les expérimentations de I3S sur Grid5000 après demande d'accord explicite auprès de leur propriétaires.

Normes: VIF utilisera les standards de l'Internet et en particulier les documents de l'IETF et du GGF auxquels ont participé les différents partenaires: http://www.ietf.org/html.charters/hip-charter.html,
http://www.inrialpes.fr/planete/people/roca/doc/draft-faurite-rmt-tesla-for-alc-norm-00.txt

Les modèles, algorithmes et protocoles qui seront apportées dans VIF par les différents partenaires ont été antérieurement publiés dans des conférences internationales ou des rapports de recherche (voir le site web de VIF sur www.ens-lyon.fr/LIP/RESO/VIF). We will base our work particularly on

- Julien Laganier, Pascale Vicat-Blanc Primet : HIPernet: Decentralized Network Security in Grid Environments. GRID2005 workshop of IEEE Supercomputing Conference, Seattle nov 2005.

	Programme de Recherche "Calcul Intensif et Grilles de Calcul" Appel à projets 2005 – <u>Fiche de synthèse des demande d'aide</u>

Acronyme du projet : VIF

	Nom du partenaire	*Type de partenaire (public ou privé)*	*Aide demandée Année 1*	*Aide demandée Année 2*	*Aide demandée Année 3*	*Aide demandée Total*
1	*INRIA*	*public*	*155 917*	*156 447*	*79 591*	*391 955*
2	*I3S*	*public*	*22 000*	*50 500*	*47 500*	*120 000*
3	*HEALTH GRID*	*Privé (association)*	*48 000*	*48 000*	*48 000*	*144 000*
	Total		225 917	254 947	175 091	655 955

Note : le tableau ci-dessus doit être fourni également sous forme d'une feuille de calcul Excel.

	Programme de Recherche "Calcul Intensif et Grilles de Calcul" Appel à projets 2005 – <u>Fiche partenaire public</u>

Acronyme du projet : VIF	**Numéro du partenaire : 1**

Nom du laboratoire (sigle développé) :
INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE
Unité de Recherche pilote : Unité de Recherche INRIA Rhône-Alpes (Projet de Recherche INRIA-RESO)
Projets de Recherche INRIA participants :
RESO et PLANETE (UR INRIA Rhône-Alpes)
GRAND LARGE (UR INRIA Futurs)

Adresse :

NB : Le Projet de Recherche INRIA-RESO est un Projet Commun, hébergé dans les locaux du Laboratoire Informatique et Parallélisme (LIP-UMR 5668) de l'Ecole Nationale Supérieure de Lyon, à l'adresse suivante :
Ecole Normake Supérieure de Lyon
46 Allée d'Italie
69364 Lyon Cedex 07 - FRANCE

Adresse de l'établissement gestionnaire du projet soumis :
Unité de Recherche INRIA Rhône-Alpes
Zirst - 655 avenue de l'Europe - Montbonnot - 38334 Saint Ismier Cedex - France
Téléphone : +33 (0)4 76 61 52 00 - Fax : +33 (0)4 76 61 52 52

Organisme(s) de rattachement (université, CNRS, INRIA, …) :
Projet de Recherche INRIA-RESO (LIP - UMR5668) :
Institut National de Recherche en Informatique et en Automatique (Etablissement gestionnaire du projet soumis)
Ecole nationale Supérieure de Lyon
Centre national de la Recherche Scientifique
Université Claude Bernard de Lyon

Projet de Recherche INRIA PLANETE
Institut National de Recherche en Informatique et en Automatique

Projet de Recherche INRIA GRAND LARGE
Institut National de Recherche en Informatique et en Automatique
Centre national de la Recherche Scientifique
Université Paris Sud
Université des Sciences et Technologies de Lille

Code du laboratoire : UMR5668

Responsable scientifique ou technique du projet (contact pour l'instruction du dossier)

Nom : VICAT-BLANC PRIMET Prénom : Pascale Fonction : Responsable d'équipe
Adresse postale : Ecole Normale Supérieure de Lyon – 46 Allée d'Italie – 69364 LYON Cedex

Tél : 04.72.72.88.02 Fax : 04.72.72.88.06 Mel : Pascale.Primet@inria.fr

Nombre de personnes affectées au projet : 15

Nombre et catégorie des salariés (Pr/PH/DR/MCU/MCUPH/IR/TR/Post doctorants, ..) :
3 Ingénieurs Experts
2 Post-Doctorants
4 Chargés de Recherche/Maîtres de Conférences
3 Directeurs de Recherche/Professeurs
3 Doctorants

Correspondant administratif et financier du projet (contact pour l'instruction du dossier)

N° SIRET : 180 089 047 000 70 — Code APE : 731Z

Adresse postale :
Unité de Recherche INRIA Rhône-Alpes
Service Administratif et Financier
Zirst - 655 avenue de l'Europe - Montbonnot Saint Martin
38334 Saint Ismier Cedex - France

Responsable administratif et financier :

Nom : DESLORIEUX — Prénom : Colette — Fonction : Responsable Administratif et Financier

Adresse postale :
Unité de Recherche INRIA Rhône-Alpes
Service Administratif et Financier
Zirst - 655 avenue de l'Europe - Montbonnot Saint Martin
38334 Saint Ismier Cedex - France
Tél : 33 (0)4 76 61 53 06 — Fax : +33 (0)4 76 61 54 55 — Mel : Colette.Deslorieux@inrialpes.fr

Visa d'un responsable ayant pouvoir de contracter et d'engager juridiquement l'organisme

Nom : KAHN — Prénom : Gilles — Fonction : Président Directeur Général

Adresse postale :
Institut National de Recherche en Informatique et en Automatique
Domaine de Voluceau – Rocquencourt BP 105
78153 Le Chesnay Cedex - France
Tél : +331 3963 5122 — Fax : +331 3963 5888 — Mel : Gilles.Kahn@inria.fr

Le signataire s'engage à mettre en oeuvre tous les moyens nécessaires à la réalisation de l'opération aidée dans les conditions prévues par le document d'Appel à Projets et par le document "Règlement relatif aux modalités d'attribution des aides du GIP Agence Nationale de la Recherche" documents dont il reconnaît avoir pris connaissance et souscrire aux obligations qui en découlent en ce qui le concerne.

Signature :

ANR – Agence Nationale de la Recherche GIP	Programme de Recherche "Calcul Intensif et Grilles de Calcul" Appel à projets 2005 – <u>Fiche de demande d'aide partenaire public</u>

Acronyme du projet : VIF	**Numéro du partenaire : 1**
Nom du laboratoire (sigle développé) : **INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE** Unité de Recherche INRIA Rhône-Alpes- Projets de Recherche RESO et PLANETE Unité de Recherche INRIA Futurs- Projet de Recherche GRAND LARGE	

Estimation du coût marginal du projet pour le laboratoire :

	Année 1			**Année 2**			**Année 3**			**To (Eu**
	Nbre h/m	**Coût h/m**	**Coût total**	**Nbre h/m**	**Coût h/m**	**Coût total**	**Nbre h/m**	**Coût h/m**	**Coût total**	
Dépenses de personnel										
Ingénieur Expert	12	3.831,12€	45.973,44€	26	3.869,43€	100.605,20€	16	3.908,13€	62.530,00€	
Post-Doctorant	24	3.831,12€	91.946,88€	9	3.869,43€	34.824,88€				
Total			137.920,32€			135.430,08€			62.530,00€	335.8
Equipements	0€			0€			0€			(
Achat de petits matériels, de consommables etc	0€			0€			0€			0
Prestation de service	0€			3.000,00€			2.000,00€			5.00
Frais de missions	12.000,00€			12.000,00€			12.000,00€			36.0(
Frais généraux (4% des dépenses)	5.996,81€€			6.017,20€€			3.061,20€€			15.07
Total (Euros)	155.917,13€			156.447,28€			79.591,20€			391.9
Aide demandée (Euros)	155.917,13€			156.447,28€			79.591,20€			391.9

Evaluation (pour information) du coût complet du projet pour le laboratoire

EQUIPEMENT	FONCTIONNEMENT				Total Equipement + fonctionnement
	Dépense de personnel	Prestation de service	Autres dépenses de fonctionnement	Total fonctionnement	
a	b	c	d	e = b + c + d	f = a + e
€	2.038.781,84€	5.000,00 €	51.075,22€	2.094.857,05€	**2.094.857,05€**

	Programme de Recherche "Calcul Intensif et Grilles de Calcul" Appel à projets 2005 – <u>Fiche partenaire public</u>

Acronyme du projet : VIF	**Numéro du partenaire : 2**

Nom du laboratoire (sigle développé) : I3S (Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis)

Adresse : 2000 route des lucioles
BP 121
06903 Sophia Antipolis cedex
Organisme(s) de rattachement (université, CNRS, INRIA, …) : CNRS
Code du laboratoire : UMR6070

Responsable scientifique ou technique du projet (contact pour l'instruction du dossier)

Nom :MONTAGNAT Prénom : Johan Fonction : Chargéde Recherhce
Adresse postale : ESSI
930 route des Colles
BP 145
06903 Sophia Antipolis
Tél : 04 92 96 51 03 Fax : 04 92 96 50 55 Mel : johan@i3s.unice.fr

Nombre de personnes affectées au projet : 3 (+1 sur financement projet)
Nombre et catégorie des salariés (Pr/PH/DR/MCU/MCUPH/IR/TR/Post doctorants, ..) : 1CR, 1 MCF, 1 doctorant (+1 IR sur financement projet)

Correspondant administratif et financier du projet[1] (contact pour l'instruction du dossier)

N° SIRET : 180 089 013 016 82 Code APE : 732Z
Adresse postale : CNRS DR20 (délégation côte d'azur)
250 rue Albert Einstein
06560 Valbonne - Sophia Antipolis
<u>Responsable administratif et financier</u> :

Nom : Boisson Prénom : Jean-Paul Fonction : Délégué régional
Adresse postale :
250 rue Albert Einstein
06560 Valbonne - Sophia Antipolis
Tél : 04 93 95 41 90 Fax : 04 93 95 41 15 Mel : Helene.Faradji@dr20.cnrs.fr

Visa d'un responsable ayant pouvoir de contracter et d'engager juridiquement l'organisme

Nom : Boisson Prénom : Jean-Paul Fonction : Délégué régional
Adresse postale : 250 rue Albert Einstein
06560 Valbonne - Sophia Antipolis

Tél : 04 93 95 41 90 Fax : 04 93 95 41 15 Mel : Helene.Faradji@dr20.cnrs.fr

Le signataire s'engage à mettre en oeuvre tous les moyens nécessaires à la réalisation de l'opération aidée dans les conditions prévues par le document d'Appel à Projets et par le document "Règlement relatif aux modalités d'attribution des aides du GIP Agence Nationale de la Recherche"[1], documents dont il reconnaît avoir pris connaissance et souscrire aux obligations qui en découlent en ce qui le concerne.

Signature :

	Programme de Recherche "Calcul Intensif et Grilles de Calcul" Appel à projets 2005 – Fiche de demande d'aide partenaire public

Acronyme du projet : VIF	**Numéro du partenaire : 2**
Nom du laboratoire (sigle développé) : : I3S (Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis)	

Estimation du coût marginal du projet pour le laboratoire :

	Année 1			Année 2			Année 3			Total (Euros)
	Nbre d'unités	*Coût unitaire*	*Coût total*	*Nbre d'unités*	*Coût unitaire*	*Coût total*	*Nbre d'unités*	*Coût unitaire*	*Coût total*	
Dépenses de personnel (1) (catégorie 1) (catégorie 2) …				1	36K	36K	1	36K	36K	72K
Equipements de R&D (3)										
Achats de petits matériels, de consommables etc. (3)	4,5K			3K						7,5K
Prestations (2) (3) **:** Information scientifique et technique Propriété industrielle Faisabilité technique Conception, analyse de la valeur Essais, tests, caractérisation Prototypage Etudes économiques Autres prestations	6K (prototypage)									6K
Frais de missions (3)	10K			10K			10K			30K
Frais généraux (4 % des dépenses)	1,5K			1,5K			1,5K			4,5K
Total (Euros)	22K			50,5K			47,5K			120K
Aide demandée (Euros)	22K			50,5K			47,5K			120K

(1) Personnel non statutaire directement affecté au projet exprimé en hommes mois. Les dépenses éligibles se limitent aux salaires et aux charges sociales.
(2) Chiffré par type de prestation
(3) Y compris TVA non récupérable.

Evaluation (pour information) du coût complet du projet pour le laboratoire

EQUIPEMENT (1) (2)	FONCTIONNEMENT				TOTAL
	Dépenses de personnel (3)	Prestations de service (1)	Autres dépenses de fonctionnement (1)	Total fonctionnement	
(a)	212,4K (b)	6K (c)	42K (d)	380,4 K(e) = (b) + (c) + (d)	380,4 K (f) = (a) + (e)

(1) Coût HT majoré le cas échéant de la TVA non récupérable
(2) Équipement : matériel dont la valeur unitaire est supérieure à 4 000 euros HT
(3) Dépenses de personnel y compris les charges sociales
.

ANR — Agence Nationale de la Recherche GIP	Programme de Recherche "Calcul Intensif et Grilles de Calcul" Appel à projets 2005 – <u>Fiche partenaire privé</u>

Acronyme du projet : VIF		**Numéro du partenaire : 3**
Nom ou Raison sociale de l'entreprise : HEALTHGRID		
N° SIRET : 449 708 841 00019 Adresse du siège social : 14 bis Avenue de Cournon – 63430 Pont-du-Chateau Forme juridique actuelle : Association loi 1901		*Code APE :* 913E
Date de création de l'entreprise : 13 avril 2003		Effectif de l'entreprise : 1
En cas d'appartenance à un groupe : Nom du groupe :		Effectif du groupe :
Etablissement où seront effectués les travaux		
Nom de l'établissement : HEALTHGRID Adresse de l'établissement : ISIMA – Campus des Cezeaux 24 avenue des Landais – 63177 Aubière		
Responsable technique du projet (contact pour l'instruction du dossier)		
Nom : LEGRE Adresse postale : 14 bis Avenue de Cournon – 63430 Pont-du-Chateau Tél : 06.86.32.57.44	Prénom : Yannick Fax : 04.73.40.50.01	Fonction : Président Mel : yannick.legre@healthgrid.org
Responsable administratif et financier du projet (contact pour l'instruction du dossier)		
Nom : LEGRE Adresse postale : 14 bis Avenue de Cournon – 63430 Pont-du-Chateau Tél : 06.86.32.57.44	Prénom : Yannick Fax : 04.73.40.50.01	Fonction : Président Mel : yannick.legre@healthgrid.org
Visa d'un responsable ayant pouvoir de contracter et d'engager juridiquement l'entreprise		
Nom : LEGRE Adresse postale : 14 bis Avenue de Cournon – 63430 Pont-du-Chateau Tél : 06.86.32.57.44	Prénom : Yannick Fax : 04.73.40.50.01	Fonction : Président Mel : yannick.legre@healthgrid.org
Le signataire s'engage à mettre en oeuvre tous les moyens nécessaires à la réalisation de l'opération aidée dans les conditions prévues par le document d'Appel à Projets et par le document "Règlement relatif aux modalités d'attribution des aides du GIP Agence Nationale de la Recherche"[1], documents dont il reconnaît avoir pris connaissance et souscrire aux obligations qui en découlent en ce qui le concerne. Signature :		

	Programme de Recherche "Calcul Intensif et Grilles de Calcul" Appel à projets 2005 – Fiche de demande d'aide partenaire privé

Acronyme du projet : VIF	**Numéro du partenaire : 3**
Nom ou Raison sociale de l'entreprise : HEALTHGRID	

Estimation (en coûts complets H.T.) du projet pour l'entreprise :

	Année 1			Année 2			Année 3			Total (Euros)
	Nbre d'unités	*Coût unitaire*	*Coût total*	*Nbre d'unités*	*Coût unitaire*	*Coût total*	*Nbre d'unités*	*Coût unitaire*	*Coût total*	
Amortissements d'équipements de R&D										
Dépenses de personnel (1) *(catégorie 1)* *(catégorie 2)* …	0,5	76K	38K	0,5	76K	38K	0,5	76K	38K	114K
Prestations de service (2) **:** Information scientifique et technique Propriété industrielle Faisabilité technique Conception, analyse de la valeur Essais, tests, caractérisation Prototypage Etudes économiques Autres prestations										
Frais de mission	2K			2K			2K			6K
Autres dépenses de fonctionnement										
Dépenses liées à l'utilisation d'autres équipements de R&D que ci-dessus (3)										
Autres dépenses (3)										
Frais (assistance, encadrement, coût de structure) (4)	8K			8K			8K			24K
Total H.T. (Euros)	48K			48K			48K			144K
Aide demandée (Euros)	48K			48K			48K			144K

(1) *Personnel directement affecté au projet, chiffré en hommes mois par catégorie de personnel.*
(2) *Chiffré par type de prestation*
(3) *Justifiées selon une procédure de facturation interne*
(4) *Ces frais seront remboursés jusqu'à un plafond défini par les règles propres au GIP ANR. Ce plafond est calculé en fonction des éléments donnés dans le tableau (frais de personnel, équipement, …)*

Note : *En cas de décision de financement de ce projet, le porteur de projet devra alors fournir un dossier complémentaire comportant un planning de déroulement du projet et des documents administratifs et financiers (entreprises, laboratoires et/ou associations) dont la liste lui sera précisée.*

Evaluation (pour information) du coût complet du projet pour le laboratoire

EQUIPEMENT [1] [2]	FONCTIONNEMENT				TOTAL
	Dépenses de personnel [3]	Prestations de service [1]	Autres dépenses de fonctionnement [1]	Total fonctionnement	
(a)	(b)	(c)	(d)	(e) = (b) + (c) + (d)	(f) = (a) + (e)

(1) Coût HT majoré le cas échéant de la TVA non récupérable
(2) Équipement : matériel dont la valeur unitaire est supérieure à 4 000 euros HT
(3) Dépenses de personnel y compris les charges sociales.

Annexe 1 : Virtualisation concept

1. Introduction

The concepts of Virtual Machine (VM) and Virtual network is not recent and was proposed very early in the history of computers and networks. For Virtual Machine, there are several approaches at the hardware (virtual machine) and the OS levels. An example of pioneer works was the VM of the IBM 370. VM were identical copies of the hardware where every copy runs its own OS. Currently, there is a strong trend toward VM concept adoption. Microprocessor vendors (AMD, Intel, IBM, etc.) are proposing new hardware mechanisms to improve virtualization performance in their last generation of products.

A first objective in virtualization research is to provide transparence, portability, confinement and performance isolation. The different virtualization techniques provide different security degrees between virtual machines or networks. Overhead minimisation and performance isolation are challenges that any virtualisation environment have to deal with. For example, workload imbalance between virtual machines running on the same machine should not compromise the fair share of the physical machine resources between virtual machines. In networking field, the fairness and performance isolation problem is solved by various bandwidth sharing control models (TCP, IP QoS, traffic engineering…) according the level of abstraction considered. Virtualization approaches present specific trade-off in the management of these issues .

2. Virtualization and Grids

Virtualization is becoming a key feature of Grid where it essentiall provides the feature of abstracting some specific characteristics of the Grid infrastructure. This abstraction can be done through services and reconfiguration. In the former case the application has to be written making calls to services (services are supposed to play the role of abstract interfaces to the local middleware, OS and hardware). In the later case the infrastructure is reconfigured (the level of the reconfiguration depends on the infrastructure) to match the application requirements. OGSA and the work on WSRF (Web Services Resource Framework) typically belong to the service oriented approach, while work on PlanetLab and Lambda network typically correspond to the reconfiguration approach. In these two cases, Virtual Machines can play a central role, by presenting interfaces to users and applications identical to the ones of a real host while confining users and applications actions into a restricted scope within the hosting resource.

The main motivations of machine virtualization in the context of Grid and large scale distributed system experimental platforms (PlanetLab, Grid eXplorer) are the following: a) each VM provides a confined environment where non-trusted applications can be run, b) VM allows to establish limits in hardware resource access and usages, through isolation techniques, c) VM allows adapting the runtime environment to the application instead of porting the application to the runtime environment, d) VM allows using dedicated or optimized OS mechanisms (scheduler, virtual memory management, network protocol) for each application, e) Applications and processes running within a VM can be managed as whole. For example, time allocation can be different for every VM and under the control of a VM manager (also called scheduler, hypervisor or orchestrator).

As a consequence of its strong potential, many Grid projects are studying the integration of machine virtualization in their research and development.
In some projects (like GridIreland), virtual machine is envisioned to reduce to number of Globus Gateways upfront each Grid site. In NAREGI, a lightweight Grid virtual machine is

developed to support Metacomputing while providing strong securit and fault tolerance. Some Grid services are proposed to provide machine virtualization in Grid, allowing high security and flexibility by decoupling the service execution from the actual resources (VMPlants). Recently, several projects have extended the concept of machine virtualization to the
one of Virtual Clusters. Virtuoso extend VNET, a virtual private network tool implementing a virtual local area network over a wide area using tunneling, for virtual machines in Grids. The In-VIGO grid computing system uses virtualization technologies to create dynamic pools of virtual resources. Machine virtualization already plays a significant role in Grid and P2P research platforms. In PlanetLab, machine virtualization is used to confine the experiment of every user in a dedicated domain, named slice. In MicroGrid, virtualization allows Globus applications to run without modification. Machine
virtualization may also be used to execute a large number of execution contexts on the same machine. For example, the Grid eXplorer project (in large scale distributed system
emulators) is relying on VM technologies to build a distributed system emulator scaling to 100†000 virtual nodes running over 1000 physical nodes.

If the concept of machine virtualization become popular for the Grid, then it is likely to find situations where many virtual machines or virtual OS will run over a single Grid node. This immediately raises the issue of machine virtualization scalability. Although, there are several previous works on QoS for Virtual Machines (including research on fair-share
(eclipse), to our knowledge, there is surprisingly only few published results on VM scalability.

As the network virtualization permit multiple overlays to cohabit on a common shared infrastructure, VPN technology appears to be a viable solution for building trust domains in grids. But the choice of the VPN conceptual layer (in terms of OSI model) has a very important impact on the level of dynamism, performance guarantees and scalability one can achieve. For example, layer2 VPN as implemented in the Grid5000 testbed, offer very good communication performance but is not much scalable. It is very difficult to add a new site. On an other hand, end to end VPN layer 3 overlays are more scalable than edge to edge VPN overlay, but they do no offer the possibility to easily manage the shared bandwidth.

Annexe 2 : Backgroud and Details of the VIF propositions

1. Security Aspects

Functionally, the classical GSI solution allows for the dynamic deployment of secure virtual organization (or trust domain) overlays, while handling entity identification and rights delegation. But when it comes to implementation, GSI is based on transport layer security (TLS, layer 4) and PKIX/X509 certificates. Such an approach is not general enough, requires lot of application modifications and presents performance drawbacks that will be studied in this task.
The INRIA Reso team proposes to leverage on network (and sub-transport) layer security (IPsec, layer 3 and HIP, layer 3.5), self-certifying naming (i.e. CBID/HIT) and SPKI delegation certificates to implement the same security functionality GSI proposes. We believe that layer 3.5 approach allows for the maximum of generality, because its "lowest layer" implementation it constitutes a least common denominator to deployed grid security infrastructure while keeping the interoperability properties the IP technology offer (comparing to layer2 VPN solutions). Nevertheless, we will demonstrate that this architecture can also act as an additional security barrier complementing existing security mechanisms (e.g.GSI).

The Crypto-Based Identifiers (CBIDs) family is already used by networking community to help to solve several issues in the IPv6 world: identifier ownership for mobility, neighbor discovery, and multicast group membership as well as infrastructure-less opportunistic encryption. One goal of this task will be to deeply analyze and design CBID and SPKI-based solutions for the Grid. Planete and Reso team will work on this aspect in concert.

One step further we want to explore solutions that combine high performance and security in communication. Indeed, high rate bulk data transfers to one or multiple destinations (through multicast/broadcast technologies in that case) raises specific challenges in terms of security. Packet source authentication and packet integrity are the basic services, since a single forged packet in a stream can corrupt a whole transfer (this is amplified by FEC encoding when this building block is used to provide more reliability and/or scalability to the transport protocol, since bogus data can be propagated to the whole object) and even introduce security threats (e.g. an attacker can send a forged packet that changes the meta-data associated to an object in order to point to a system file) . Solutions based on shared group keys, if they are efficient from a computing perspective, are not sufficient in case of group communications since any group member may forge any packet. Solutions based on digital signatures with asymmetric key cryptography are not acceptable too because of prohibitive processing loads. TESLA (Timed Efficient Stream Loss-tolerant Authentication) is one possible solution, and BiBa (Bins and Balls signature) from the same author, is another one.
The INRIA Planète team is already working actively on the use of TESLA for file multicast/broadcast applications, and has submitted an IETF Internet Draft on the subject. In the VIF project, this standardization activity around TESLA will be continued, an open source implementation, with all the required features will be developed, integrated in the existing transmission tools for the Grid, and its performances evaluated in operational environments. In parallel, work will be carried out on the BiBa alternative, which supposedly solves some of the limitations of TESLA (in particular the high buffering requirements at a receiver). If the first theoretical evaluation of BiBa shows there's an interest, we will implement it, integrate it in the same application as would have been done with TESLA, evaluate its performances compared to TESLA, and standardize it at IETF if it is worth.
The INRIA Planète team is also working on high performance cryptographic primitives, more precisely on joint FEC (Forward Error Correction) / cryptography techniques. The goal is to have both operations done at once, which makes sense when sending large confidential contents on public networks. In that case encryption is achieved almost for free (i.e. without any major performance impact, neither during encoding nor decoding) which is a major

asset. If the work, from some aspects, begins to be mature (patent deposit under progress), it has also opened some new perspectives that will be studied, and we are interested in studying its applicability to the Grid transmission tools. If the performances are good, it may offer an alternative to VPN encryption in some target environments.

2. Performance- control Aspects

High performance controlled transfers are a key for high performance computing. The INRIA Reso team proposes to develop in the context of the VIF approach an end to end hybrid network sharing model that combines classical on line IP QoS solutions with off-line advanced reservations. Here also we propose to leverage on network (and sub-transport) layer for performance control. The model we propose to evaluate maps the low volume, delay sensitive flows into existing classical prioritized QoS classes, the normal flows into standard best-effort class and isolates the large volume throughput sensitive ones with advance reservation techniques. The main idea is to define, within each multi-point secured VPN, virtual sub-channels with deterministic performance for each type of grid traffic. This network sharing approach take benefice from off-line scheduling for optimizing the transfer of huge data files and maximizing the allocated network resource utilisation while protecting other traffics from dangerous "elephants".

There are however differents issues to solve to make this approach a reality: This task will look deeply at the following points :
- As classical communication programming interfaces do not provide users the ability to specify QoS requirements when sending their data, we will propose a way to transparently differentiate the grid traffics. For advance reservations we plan to integrate the upper level reservations within workflow engine used by application. Such facility is still a technical issue and will be part of this task, involving both infrastructures and application teams.
- End to end resource provisionning: the global network resource has to be divided from into virtual sub-channels. Each channel has to be well provisionned from host to host to meet the various requirement of the grid. As the interconnection comprises public and private networks, the provisionning and the flow differentiation will have to cross all the heterogeneous domains till end hosts applicaiton level. For the costly wide area access, dynamic bandwidth allocation may help in optimizing network resource provisionning for bulk transfers. In this project we will concentrate at the private part of the network and exploring the bandwidth sharing and flow differentiation problem from host's virtual OS to the WAN access equipment.

There two other important points, complementing the task we propose here : the bandwidth allocation algorithm and the traffic control mechanisms. The Reso has already explored solution and will develop them in the Cariocas project of the IDF pôle de competitivite « Simulation Numerique » that has been recently labelled. All these results and development will be reuse, whenever is it convinient in the VIF project.

3. Dependability Aspects

Because of the virtualization principle, dependability covers at least two aspects: resource isolation and fault tolerance. First, users sharing the same resources should get a fair access to them. The notion behind this property is the virtual machine isolation.
Isolation may have a complex definition: 1) when several users request the same workload on the same resources, the isolation should ensure that no user is slowed down because other users are using the resources too much. 2) when a resource is requested by a subset of users, this subset should be able to share 100% of the resource capacities. In this task we will design a benchmark suite and run experiments on a large variety of virtualization technologies in order to check the isolation of every computer resource independently. Currently, only results on global isolation are available and some results cannot be explained without a finer evaluation. If the isolation is found poor, we will investigate, design and develop mechanisms required to force a good resource isolation between user contexts. A second important aspect of virtualization, not yet addressed in the literature, is the fault

tolerance of virtual machines. When deploying an application on thousands of processor running virtualization software, several issues raise: 1) the stability level is supposed lower than with standard OS, because virtualization tools have stronger limitations and some tools still need improvements 2) how to save the context of virtual OS? How to migrate virtual a OS context? these issues have already been discussed in the literature but not evaluated in the Grid context (there are new issues related to performance and security). 3) how to combine processor and network fault tolerance in the context of virtualization? Basically this should involve extending the context to virtual network and provide co-migration of network and OS virtual contexts. In summary, this task will investigate several key issues related to dependability in the context of network and OS virtualization.

Annexe 3 : Details of the Biomedical applications

1. Use cases1 : Data intensive workflow for medical image databases analysis

Over the past years, many processing algorithms were developed to process digital medical images such as algorithms for signal filtering, anatomical and functional structures modeling, structures segmentation, or quantification of medical parameters from medical images. Applying these algorithms to full image databases is increasingly needed for medical studies or the statistical validation of medical image analysis procedures. Although computing grids are well adapted to tackle intensive computing emerging when processing large image databases, the effective control of the data flow involved when applying multiple processings on data distributed over different sites remains an opened research area. Moreover, data confidentiality needs to be preserved when data is transported from acquisition sources to processing sites and for storage of intermediate computing results. We are interested in the deployment of data intensive workflows on grid infrastructures to face the need of image-based medical studies.

The *registration* of medical images is a key basic algorithm needed for many image processing algorithms. Several medical images acquired with different imaging devices or at different instant in time are alined in a same spatial frame. This processing is compensating for patient position or acquisition geometry differences in order to compare different images. Images to register can be acquired with the same modality (*mono-modality* registration) or different imaging modality (*multi-modality* registration). The registration may apply to data coming from a same patient, for which only position and orientation needs to be compensated for (*rigid* registration). It may also apply to images from different patients for which anatomical differences need to be compensated for (*non-rigid* registration). Many registration algorithms have been developed to deal with the different cases encountered [MICCAI'98].

A medical image registration computation produces a transformation (translation plus rotation in the case of rigid registration or dense deformation field in the case of non-rigid registration). As no reference, or *gold standard*, does exist to validate the result of the computation (as often in the medical image processing field), it is very difficult to assess the quality of the result in an objective maner. A visual inspection can only reject obvious outlier results but an objective, reproducible, and precise evaluation if not possible, especially when considering 3D data sets that are difficult to visualize. We use the so called *bronze standards* technique to statisticaly quantify the maximal error resulting from a given registration algorithm. The larger the sample image database is, and the larger the number of registration algorithms to compare with is, the most accurate is the method. These two criterions make the *bronze standards* technique a very compute intensive application. Moreover, the need for processing a large number of input images implies the usage of distributed image databases (coming from several acquisition sites) that should be efficiently transported to the computing resources while preserving the confidentiality of these data. Given the large computations involved, the underlying infrastructure also needs to provide massive parallelism and fault tolerance to face temporary hardware failures (network interuptions, reboots...) by automatic recovery procedures.

This application for validating medical image registration algorithms is deployed through a workflow manager able to deal with the complexity of the processings (registration of different image pairs with several algorithms and intermediate transformation and comparison of results obtained) and to enact the computations efficiently on a grid infrastructure. In this project, the application will be deployed on the grid5000 hardware infrastructure and it will benefit from the secured and high performance transport services deployed. The role of this application will be to:

- Define precisely application requirements for the underlying infrastructure, especially considering the data transport and confidentiality needs.
- Develop realistic deployment scenarii on data sets distributed over different sites.

- Deploy the application on the infrastructure and adapt our workflow engine to control the computations.
- Optimize the performance of the workflow engine by taking into account the data transfer cost when scheduling computations.
- Demonstrate the usage of the infrastructure in real scale to validate the benefits in terms of performance and the interest in the medical image processin area.

[MICCAI'98]
"The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration"
A. Roche, G. Malandain, X. Pennec and N. Ayache
Medical Image Computing and Computer-Assisted Intervention (MICCAI'98),
pp 1115-1124, Oct. 1998, Cambridge, USA.

[Biogrid'03]
"Using grid technologies to face medical image analysis challenges"
J. Montagnat, V. Breton, I. E. Magnin
Biogrid'03, proceedings of the IEEE CCGrid03, pp 588-593, May 2003, Tokyo, Japan.

[JGC'04]
"Medical images simulation, storage, and processing on the European DataGrid testbed"
J. Montagnat, F. Bellet, H. Benoit-Cattin, V. Breton, L. Brunie, H. Duque, Y. Legré, I.E. Magnin, L. Maigne, S. Miguet, J.-M. Pierson, . Seitz, and T. Tweed
Journal of Grid Computing, vol 2(4), pp 387-400, Kluwer, dec. 2004

2. Use Case 2: Radiotherapy treatments modelisation

From a technology point of view, radiotherapy is a highly complex procedure, involving a variety of computational operations for data gathering, processing and control. Presently patients are treated with standardised radiation doses. Tissue density provided by medical imaging devices is used to calculate the dose delivered by photon and electron beams. Accuracy of Monte Carlo (MC) dose computation is awesome, provided that the computing power is sufficient to allow for enough runs to reduce the statistical noise. The Grid is a natural alternative to costly parallel computers. In this way, MC dose computations could become standard for radiotherapy quality assurance (QA), planning, and plan optimisation years before individual departments could afford a local investment that is capable to support MC. Requirements needed for such deployment include the existence of a service level agreement between the departments and the Grid providers by which the Grid level of performances in terms of security, stability and response time is guaranteed.

The patient comes to hospital in order to image the presumed cancer disease located in his body. Images are typically IRM and/or CT scans. Those images are produced in a DICOM format. After the examination of the patient and the analysis of medical images, the physician is able to make a diagnosis of the tumor. The physician elaborates a prescription to kill the tumor. In our case, the medical cure is based on radiotherapy and brachytherapy solutions.

Brachytherapy is the use of encapsulated radioactive sources to treat cancer. Radioactive sources are used to deposit therapeutic doses near tumours while preserving surrounding healthy tissues. In this case, Monte Carlo technique is used to optimize the dose calculations around the brachytherapy source.
Radiotherapy involves directing a beam of megavoltage x rays or electrons (occasionally protons) at a very complex object, the human body and especially a tumour. Currently, Monte Carlo simulation techniques are the most accurate method for dose calculation in radiotherapy, in particular when radiation is transported from one medium to another.

The principle of a Monte Carlo simulation in physics, is to simulate the radiation transport knowing the probability distributions governing each interaction of particles in materials.

Different possible trajectories or histories of a particle could be produced. Then, simulations store physical quantities of interest for a large number of histories to provide information on required quantities. That process involves the use of random numbers responsible for the name Monte Carlo given to this technique which was introduced during the development of atomic energy in the post World War II era.

In the case of radiotherapy-brachytherapy applications, the goal is to calculate accurately the dose distribution in a located tumor. Most of the commercial systems, named TPS (Treatment Planning Systems), use an analytic calculation to determine these dose distributions and so, errors near inhomogeneities in the patient can reach 10 to 20%. Such codes are very fast (execution time below one minute to give the dose distribution of a treatment), thus allowing their usage in medical centres. Accurate calculations of Monte Carlo simulations involve a very high computing time compared to analytic calculations used in routine cancer treatment planning. Especially for radiotherapy treatments, Treatment Planning System (TPS) use approximations in the beam model and the dose calculation(e.g, the exclusion of electron transport) to speed up the computation. This may introduce significant uncertainties in the dose distributions in a patient, especially in the presence of heterogeneities such as the air-tissue, lung-tissue and tissue-bone interfaces.
To follow medical requirements, errors in the dose calculation should be kept below 2%. That's why, for specific applications the use of Monte Carlo simulations seems to be the best way to compute complex cancer treatment.

Even though the cost of computing resources is continually decreasing thereby facilitating consequent calculations, Monte Carlo simulations can't currently be used for clinical treatments planning for which the computing time remains too high on a single machine.
The computing time of a Monte Carlo simulation depends on different parameters: the number of particles generated during a simulation, the medium where particles interactions occur. Depending on the type of material filling the medium and the type of particles generated, the number of physical interactions can vary.
So, there is a real interest for parallel and distributed Monte Carlo simulations in order to provide very accurate studies in medical physics.

Deployment of Monte-Carlo codes has been achieved on several grid infrastructures but there are additional requirements in order to set-up a service for hospitals:

- Guaranteed response time to the hospital asking for the computation is needed. This response time includes time to move information on the grid, to execute the tasks to collect the outputs and concatenate it
- Medical data must be kept anonymous outside the hospital and medical images should be encrypted
- The infrastructure must be robust enough to always send back the computation in the guaranteed response time.

These three requirements are not yet achieved on the existing grid infrastructures.

References:
L.Maigne, D. Hill, P. Calvat, V. Breton, D. Lazaro, R. Reuillon, Y. Legré, D. Donnarieix, "Parallelization of Monte Carlo Simulations and Submission to a Grid Environment", *Parallel Processing Letters*, Vol. 14, No. 2 (June 2004), p 177-196.

Annexe 4: Main references for the project

I. Foster, C. Kesselman, and S. Tuecke, “The Anatomy of the Grid: Enabling Scalable Virtual Organizations,” International J. Supercomputer Applications, no. 15(3), 2001.

V. Welch, F. Siebenlist, I. Foster, J. Bresnahan, K. Cajkowski, J. Gawor, C. Kesselman, S. Meder, L. Pearlman, and S. Tuecke, “Security for grid services,” in Proc. of 12th IEEE International Symposium on High Performance Distributed Computing (HPDC 2003), June 2003.

Carl A. Waldspurger, Memory resource management in VMware ESX server, SIGOPS Oper. Syst. Rev.36, 2002, ACM Press

Paul Barham and Boris Dragovic and Keir Fraser and Steven Hand and Tim Harris and Alex Ho and Rolf Neugebauer and Ian Pratt and Andrew Warfield, Xen and the art of virtualization, SOSP '03: Proceedings of the nineteenth ACM symposium on Operating systems principles, 2003},

Larry Peterson and Tom Anderson and David Culler and Timothy Roscoe, A Blueprint for Introducing Disruptive Technology into the Internet, Proceedings of HotNets, 2002

Hans-Jorg Hoxer and Kerstin Buchacker and Volkmar Sieh, "Implementing a User Mode Linux with Minimal Changes from Original Kernel", Proceedings of 9th International Linux System Technology Conference, 2002

Van Krsul and Arijit Ganguly and Jian Zhang and Jos A. B. Fortes and Renato J. Figueiredo, « VMPlants: Providing and Managing Virtual Machine Execution Environments for Grid Computing », Supercomputing'04: Proceedings of the 2004 ACM/IEEE conference on Supercomputing (CDROM).

J. Ioannidis, A. Keromytis, and al., “Implementing a Distributed Firewall,” in ACM Conference on Computer and Communications Security, 2000.

C. Ellison and etc., SPKI Certificate Theory, IETF, RFC 2693, September 1999.

R. Moskowitz and P. Nikander, Host Identity Protocol Architecture, IETF, draft-ietf-hip-arch-02.txt, January 2005.

C. Castellucia, G. Montenegro, J. Laganier, and C. Neumann, “IPv6 Opportunistic Encryption,”
in Proc. of 9th European Symposium on Research in Computer Security. LNCS, Springer, September 2004.

C. G., K. S., S. Ch., and S. G., “Supernetworking: The Next Generation of Secure Enterprise Networking,” in ACSAC 2000, March 2000.

J. Saltzer and M. Schroeder, “The Protection of Information in Computer Systems,” 1975.

D. Price and A. Tucker, “Solaris Zones: Operating System Support for Consolidating Commercial Workloads,” in USENIX Large Installation System Administration Conference (LISA’04), November 2004.

P.-H. Kamp and R. Watson, “Jails: Confining the omnipotent root,” in Proceedings of the 2nd Intl. SANE Conference, May 2000.

Bassi and J. Laganier, “Towards an IPv6-based Security Framework for Distributed Shared Storage,” in Proc. of 7th IFIP Conference on Communications and Multimedia Security. LNCS,

I. Foster, *The Grid 2: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann, 2004.

M. Goutelle, Y. Gu, E. He, S. Hegde, R. Kettimuthu, J. Leigh, P. Vicat-Blanc Primet, M. Welzl, and C. Xiong. (2004, Feb.) A survey of transport protocols other than standard tcp. Grid working draft, Data
transport research group, Global GRID Forum. [Online]. Available: http://forge.gridforum.org/projects/data-rg/document/

I. T. Foster, M. Fidler, A. Roy, V. Sander, and L. Winkler, End-to-end quality of service for high-end applications,î *Computer Communications*, vol. 27, no. 14, pp. 1375-1388, 2004.

K. Czajowski, I. Foster, and C. Kesselman, Resource co-allocation in computational grids, in *Proc. IEEE the eighth International Symposium on High Performance Distributed Computing*, Aug. 1999, pp. 219-228.

L. Marchal, P. Vicat-Blanc Primet, Y. Robert, and J. Zeng, Network resource scheduling in grids, IEEE GlobeCom., 2005.

V. Firoiu, J. L. Boudec, D. Towsley, and Z. Zhang, Theories and models for internet quality of service, *Proceedings of the IEEE*, vol. 90, pp. 1565-1591, Sept. 2002.

Annexe 5: Project leader biography and partners description

Pascale Vicat-Blanc Primet

The consortium is leaded by **Pascale Vicat-Blanc Primet**, graduated in Computer Science, (http://perso.ens-lyon.fr/Pascale.Primet) who is **senior researcher at INRIA** . Her research interests include Distributed and Real-Time Systems, High Performance Grid and Cluster Networking, Active Networks, Internet protocols (TCP/IP), Network Quality of Service, Collaborative Work and toolkits,. Member of the LIP Laboratory board (Laboratoire de l'Informatique du Parallelisme.) of the Ecole Normale Supérieure in Lyon, she is actually leading the RESO team, labelled RESO project of the Institut National de la Recherche en Informatique et Automatique (INRIA). Since 2000, she has been very active in the international and national Grid community. In 2001-02, she has been managing the Network Workpackage of the European IST DataGRID project and was member of the DataGRID technical committee. She has been the scientifical coordinator of the French RNTL Grid Platform project, titled E-Toile and funded by the French Research Ministry, and was member of the RNRT VTHD++ project. She was the INRIA coordinator for the IST DataTAG project. Co-chair of the DataTransport Research Group in the Global Grid Forum, she has co-edited several Gridnetworking and Transport protocol GGF documents. She is member of the steering committee of the French ACI MD DataGRID Explorer (GdX) project and of the ACI Grid Grid5000 project. She is (has been) member (/or chair) of international conferences steering or program committees (IEEE Grid workshop/SC, Pfldnet, IEEE GridNets, IEEE HPDC conference) and reviewer for international conference and journal in Grids and Networking. She has published her work in more than 60 papers in Grid and Networking journals or conference like "Future Generation Computer Systems – Grid Computing", International Journal of High Performance Computing Applications, IEEE ICC, IEEE GlobeCom, IEEE HPSR (High Speed Switching and Routing) IEEE CCGrid, IEEEClusters, IEEE and many specialized workshops. She also coordinated several large scale demonstrations: DataGrid WP7 review, DataTAG, GridDemo, journées RNTL.

INRIA Reso studies and develops new solutions in terms of protocols, services and software for next generation Internet networks and infrastructure (grid computing). The team works on optimized software architectures for efficient communications in end systems, cluster-based servers and programmable access equipments, and develop protocols and computations for efficient and customizable transport of heterogeneous streams in very high speed networks. We provide solution that optimize both data movements and I/O management that are closely inter dependant, by using the intelligence of network interface cards (NICs). RESO designs Grid network services and network middleware to avoid the applications to be network-aware, to simplify the programming and to optimize the execution of their communication parts while fully exploiting the capacities of the evolving network infrastructure.

INRIA Grand-Large has a long experience in parallel and distributed computing, GRID computing, the associated languages and middleware. Grand-Large has introduced the XtremWeb and MPICH-V environments for P2P computing and fault tolerant parallel computing. Members of Grand-Large have also designed, implemented and evaluated many parallel and distributed codes for supercomputers, massively parallel machines, clusters, P2P multi-site platforms and GRID. Grand-Large is leading the Grid eXplorer and Grid'5000 project in France. Members of Grand-Large are involved in the organization of top Conferences and Journals in the Grid domaine. Grand-Large develops a large number of international collaborations (In 2005, more than 20 researchers visits were organized, about 10 in each direction).

INRIA Planete-RA is specialized in security for small devices like sensor networks and group communication applications. The team studies and develops new security building blocks for these target environments, from cryptographic primitives like encryption schemes that enable secure data aggregation, or lightweight cryptography for small devices, to privacy mechanisms, secure location mechanisms, security in ad'hoc networks, high performance source authentication building blocks.

The application partners Healthgrid and I3S gather a strong experience in developing and evaluating grid applications in current grid environment like Globus, DataGrid or EGEE.

The HealthGrid association (http://www.healthgrid.org) was founded to bring the necessary long term continuity between the projects addressing the deployment of Grid technology for healthcare. Its goal is to collaborate on the following activities:
. Identification of potential business models for biomedical Grid applications.
. Feedback to the Grid development community on the requirements of the pilot applications deployed by the European projects.
. Development of a systematic picture of the broad and specific requirements of physicians and other health workers when interacting with Grid applications.
. Dialogue with clinicians and those involved in medical research and Grid development to determine potential pilots.
. Interaction with clinicians and researchers to gain feedback from the pilots.
. Interaction with all relevant parties concerning legal and ethical issues identified by the pilots.
. Dissemination of pilot outcomes to the wider biomedical community.
. Interaction and exchange of results with similar groups worldwide.
. Formulation and specification of potential new applications in conjunction with end user communities.
The HealthGrid association is already identified at a world level as a key actor of the Grid deployment in healthcare. This international dimension has been strengthened by the edition of the HealthGrid White Paper in 2004 with contributions from more than 40 experts from Europe, Asia and North America.

I3S CNRS laboratory, RAINBOW team is specialized in distributed systems and component programing. In this team, Johan Montagnat (CR CNRS) is developing an activity in grid computing, especially applied to medical image analysis, with the help of Diane Lingrand (MCF, specialized in image processing) and Tristan Glatard (PhD student, working on workflow managers). Johan Montagnat has a background in computer science and a PhD in medical image analysis. He is the leader of the biomedical applications activity in the EGEE Eureopean (http://www.eu-egee.org/) project and the manager of the MEDIGRID ACI-GRID (http://www.creatis.insa-lyon.fr/MEDIGRID/) project that will come to end in november 2005.

Annexe 6 : Principe Accords de Propriété Intellectuelle

1) Définitions

Le terme **Connaissances Antérieures** désigne toute connaissance brevetée ou non, savoir-faire, secret de fabrique ou tout autre type d'information sous quelque forme que ce soit, appartenant à ou détenue par l'un des partenaires antérieurement au projet ou acquise par lui indépendamment de l'exécution du projet, mais nécessaires à la réalisation du projet ou à l'exploitation des résultats du projet.

Le terme **Résultats Propres** désigne toute connaissance brevetée ou non, savoir faire, secret de fabrique ou tout autre type d'information, sous quelque forme que ce soit, résultant directement des travaux du projet et obtenus par le personnel d'un seul partenaire, sans le concours du personnel de l'autre partenaire.

Le terme **Résultats Communs** désigne toute connaissance brevetée ou non, savoir faire, secret de fabrique ou tout autre type d'information, sous quelque forme que ce soit, résultant directement des travaux du projet et obtenus conjointement par le personnel des deux partenaires

2) Propriété intellectuelle

2-1 Connaissances antérieures

Chaque partenaire conserve la propriété totale et exclusive de ses connaissances antérieures.
Cependant, pour les seuls besoins du projet et pour la seule durée du projet, chaque partenaire pourra, sur demande expresse et sous obligation de confidentialité, utiliser sans contrepartie financière les Connaissances Antérieures de l'autre partenaire.

2-2 Résultats Propres

Chaque partenaire sera propriétaire de ses Résultats Propres, qu'il aura générés seul, sans le concours du personnel de l'autre partenaire, et décidera seul de l'opportunité et de la nature des mesures de protection à prendre, à ses seuls nom et frais.

2-3 Résultats Communs

Les Résultats Communs seront la propriété conjointe des partenaires, qui décideront ensemble de l'opportunité et de la nature des mesures de protection à prendre, en leurs noms conjoints et à leur frais partagés. Le partenaires élaboreront un règlement de copropriété avant toute exploitation.

3) Principe d'exploitation

3-1 Exploitation des Connaissances Antérieures

Chaque partenaire dispose librement de ses Connaissances Antérieures.
Chaque partenaire s'engage à concéder à l'autre partenaire, sur demande expresse de celui-ci, des licences sur ses Connaissances Antérieures nécessaires à la valorisation des Résultats Propres ou Communs du projet, à des conditions financières préférentielles.

3-2 Exploitation des Résultats Propres

Chaque partenaire pourra librement exploiter et utiliser les Résultats Propres dont il est propriétaire, directement ou par voie de concession de licence.
Les Résultats propres d'un partenaire nécessaires à l'exploitation des Résultats Propres de l'autre partenaire feront l'objet d'une licence à négocier à des conditions financières préférentielles.

3.3 Exploitation des Résultats Communs

Les partenaires copropriétaires des Résultats Communs, régleront les modalités d'exploitation dans le cadre du règlement de copropriété sus mentionné.