

# MapCenter : un modèle ouvert pour la découverte, la supervision et la visualisation des environnements distribués à large échelle

Franck Bonnassieux  
CNRS/UREC  
ENS LYON, 46 Allée d'Italie 69364 LYON Cedex 07  
franck.bonnassieux@ens-lyon.fr

Robert Harakaly  
CNRS/UREC  
ENS LYON, 46 Allée d'Italie 69364 LYON Cedex 07  
robert.harakaly@ens-lyon.fr

Pascale Vicat-Blanc Primet  
INRIA/RESO  
ENS LYON, 46 Allée d'Italie 69364 LYON Cedex 07  
Pascale.primet@ens-lyon.fr

Date de rédaction: 9 Octobre 2003

## Résumé

*Le concept de grille génère de nouveaux paradigmes et de nouvelles contraintes qui remettent en cause les solutions classiques de surveillance des réseaux et des systèmes. Une grille peut être définie comme une agrégation complexe de ressources de calcul et de stockage hétérogènes pour créer un environnement de calcul virtuel intégré. Dans ces environnements instables et en constante évolution, un système de supervision et de visualisation doit pouvoir s'adapter aux différentes politiques de sécurité, re-découvrir en permanence les ressources et services entrants et sortants, et générer en temps réels les vues abstraites et tableaux de bord nécessaires aux diverses communautés d'utilisateurs et d'administrateurs.*

*MapCenter a introduit un modèle de présentation souple et une architecture de supervision performante répondant à ces besoins. Son architecture modulaire et adaptable de découverte des ressources permet de générer dynamiquement et automatiquement l'ensemble des vues logiques et des cartes graphiques souhaitées.*

*Au delà des possibilités de supervision, MapCenter est également une interface d'accès aux divers systèmes d'information présents aujourd'hui dans les grilles de calculs et les environnements distribués à large échelle (Annuaire LDAP, SGBD, Relational Grid Monitoring Architecture, Web Services, ...).*

*Son extrême souplesse et sa grande facilité de mise en œuvre lui ont permis d'être déployé sur de nombreux projets majeurs de grilles à travers le monde. Son implémentation générique garantit une maintenance aisée et des évolutions rapides pour répondre aux nouveaux défis des grilles de demain.*

## Mots clefs

Grilles de calculs, Grid Monitoring, Environnements distribués, Supervision

# 1 Introduction

Les récents progrès dans la technologie des réseaux très haut débit courtes et longues distances ont rendu possible la construction de systèmes répartis haute performances, distribués à l'échelle planétaire dont certains des constituants sont des grappes de PC ou des calculateurs parallèles et que l'on nomme "grilles" de calcul.

Le concept de grille définit une agrégation complexe de ressources de calcul et de stockage hétérogènes pour créer un environnement de calcul virtuel intégré permettant à diverses communautés d'utilisateurs regroupés au sein d'organisations virtuelles d'exécuter de larges applications réparties dynamiquement sur des milliers d'ordinateurs ([3]). Dans un tel contexte, de nouveaux paradigmes et de nouvelles contraintes qui remettent en cause les solutions classiques de supervision des réseaux, des systèmes et des applications sont apparues.

Par exemple, une des caractéristiques essentielles de ces environnements est que les nœuds de calcul ou de stockage peuvent apparaître et disparaître rapidement, soit parce qu'ils tombent en panne, soit parce qu'ils sont déconnectés ou reconfigurés. Les ressources et les services qui les utilisent doivent être fréquemment et automatiquement redécouverts pour suivre l'évolution dynamique des environnements virtuels. De plus, chaque site de production interconnecté possède ses propres politiques et règles de sécurité, ainsi que ses propres procédures d'administration système et réseaux. Enfin, au dessus des nœuds et des sites, le grand nombre d'applications et d'organisations virtuelles activées requiert différentes vues abstraites de ces environnements. Un outil de visualisation doit être suffisamment ouvert et flexible pour représenter ces différentes vues logiques engendrées.

L'outil de surveillance et de visualisation de grille MapCenter a été conçu pour relever ces nouveaux défis. Une caractéristique essentielle d'un tel système de surveillance est qu'il doit être efficace mais non intrusif pour respecter les politiques hétérogènes et ne pas interférer avec les règles d'administration locales. L'architecture de MapCenter (décrite dans [1]) se décompose en trois couches :

- une couche de découverte et modélisation des ressources, basée sur des plug-ins modulaires
- une couche de surveillance, implémentée par un module efficace de polling multi-threads,
- une couche de visualisation, basée sur un modèle de présentation original

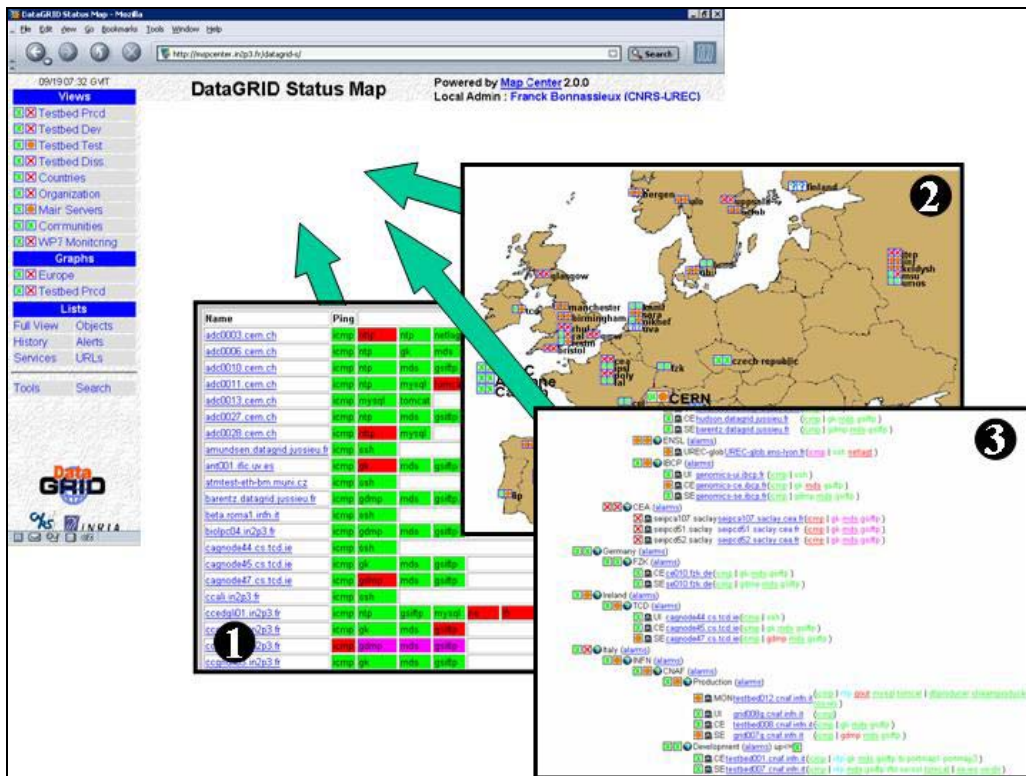


Figure 1 - Interface Homme-Machine de MapCenter

Les différentes possibilités de visualisation au travers de l'interface Homme-Machine de MapCenter sont illustrées en Figure 1 :

- la vue ❶ présente la liste des objets supervisés et de leurs états,
- la vue ❷ montre une carte graphique des sites d'une grille de calcul et de leurs états respectifs,
- la vue ❸ est une vue hiérarchique complète des pays, sites et machines dans ces sites.

Le modèle de données et de présentation de MapCenter (décrit également dans [1]) est composé d'objets, de symboles, de cartes et de liens entre cartes. Un objet est un élément de base représentant un ensemble de services sur une machine. Un symbole est la représentation graphique d'un objet dans une carte. Une carte est un ensemble de sous-cartes et de symboles. Les cartes et symboles permettent de créer une hiérarchie complète des objets supervisés par la couche de monitoring, et les liens servent à composer des vues graphiques et logiques au dessus de cette hiérarchie. Ce modèle ouvert est très flexible, et permet de structurer n'importe quelle approche logique d'un environnement distribué et de visualiser tous les tableaux de bords indispensables aux différents types d'utilisateurs finaux (chercheur, administrateur système, manager ...).

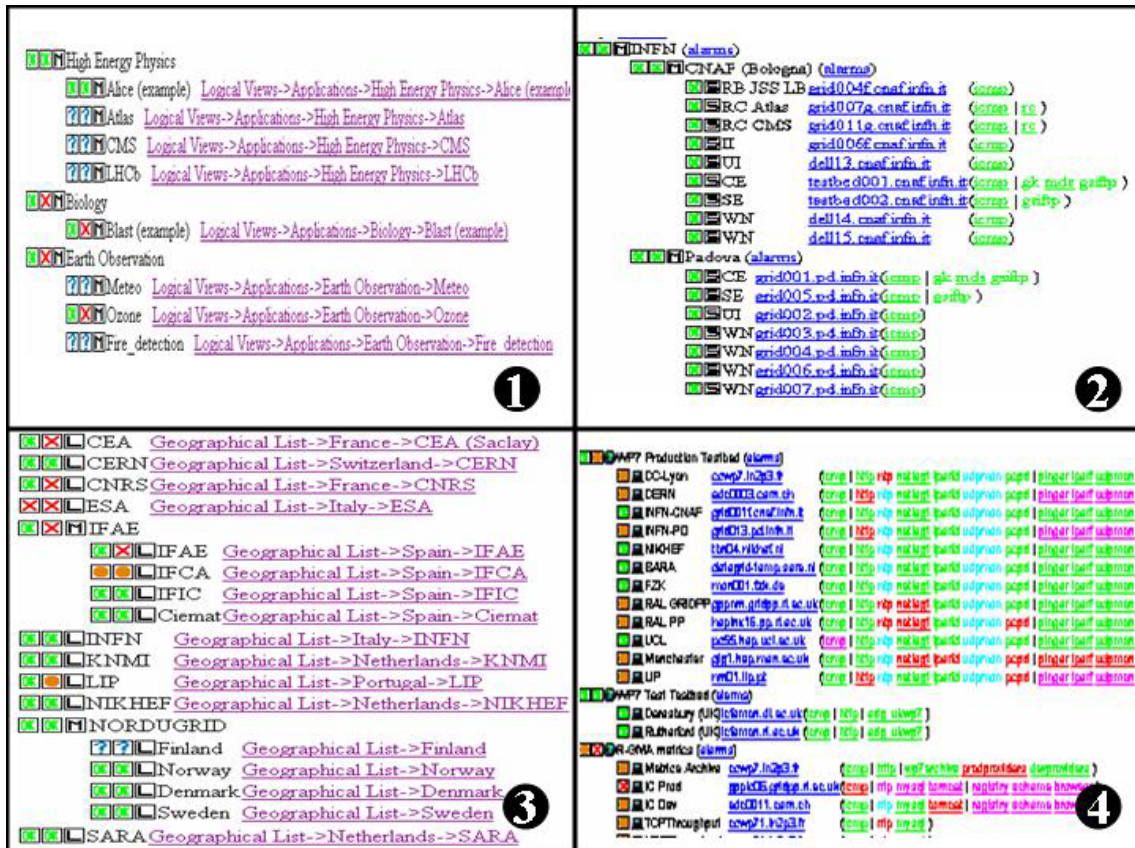


Figure 2 - Différentes vues logiques d'une grille

La Figure 2 fournit quelques exemples de vues logiques extraites de la grille du projet DataGrid ([12]):

- la vue ❶ est une vue de l'état des applications en production sur la grille,
- la vue ❷ est une vue hiérarchique des sites et des différents éléments sur chaque site (gestionnaire de grappe, nœud de calcul, nœud de stockage, interface utilisateur, ressource broker, gestionnaire de répliquas...),
- la vue ❸ est une vue de la grille par organismes et sous-organismes participants,
- la vue ❹ est une vue regroupant tous les éléments et les fonctions de monitoring réseaux.

Toutes ces vues logiques ont été générées très simplement et rapidement à l'aide de cartes et de liens, et pointent in fine sur les mêmes objets supervisés, sans redondance de supervision, ni de déclaration.

MapCenter réalise la supervision des objets grâce à plusieurs capteurs (ICMP, TCP, UDP, HTTP) depuis un serveur central ([2]), et un système complet peut-être déployé en quelques heures. Un point clef pour un système de monitoring est de rester transparent aux ressources supervisées, sans interférences ni effets de bords pour les machines ou les administrateurs systèmes. Pour cela, la vérification des services TCP est effectuée en mode furtif (demi connexion TCP, [6]), évitant ainsi le réveil des démons correspondants et la génération d'entrées parasites dans les fichiers de log.

Pour obtenir une solution performante, l'architecture de monitoring est entièrement multi-threadées ([2]), tous les timeouts et les mécanismes de reprises sur erreur ont été également optimisés pour augmenter les performances. MapCenter est ainsi capable de superviser plusieurs dizaines de milliers de machines en quelques minutes.

Dans la suite de cet article, nous nous attacherons à montrer les nouvelles fonctionnalités de MapCenter conçues, développées et déployées récemment qui sont une nouvelle conception générique de la découverte des ressources, ainsi que la généralisation des interfaces d'accès aux différents systèmes d'information qui composent les grilles de calcul actuelles.

## 2 Découverte automatique des ressources

Les environnements fortement distribués (grilles de calcul, grilles de stockage, système pair à pair) sont en constante évolution. Les nœuds de calculs ou de stockages peuvent apparaître et disparaître rapidement, les organisation virtuelles (communautés de chercheurs) se créent et se modifient continuellement, mêmes les organes centraux ("ressources brokers", gestionnaires de réplication des données, bases de registres...) ne sont pas figés et se déplacent fréquemment. L'arrivée (voir l'avènement) de OGSA ([9]) développe ce paradigme en définissant tout élément de la grille comme un service basé sur le concept de « soft state » : tout service à une durée de vie limitée dans le temps, et donc rien n'est immuable.

Pour suivre cette évolution dynamique et rapide des environnements fortement distribués, la découverte des ressources et la génération des cartes et des vues logiques et graphiques sont réalisées automatiquement par MapCenter pour de nombreux environnements de grilles ou de pair à pair. En particulier des modules de découverte ont été développés pour les annuaires LDAP, les bases MySQL, RGMA ([7]), ou encore via des clients d'interrogation spécifiques. Ce chapitre décrit cette architecture de découverte automatique.

### 2.1 Architecture générale

Une première architecture de découverte automatique des ressources à été ébauchée dans [2]. Sa conception a depuis été entièrement revue pour satisfaire les besoins de souplesse, et s'adapter très rapidement aux différents systèmes d'information.

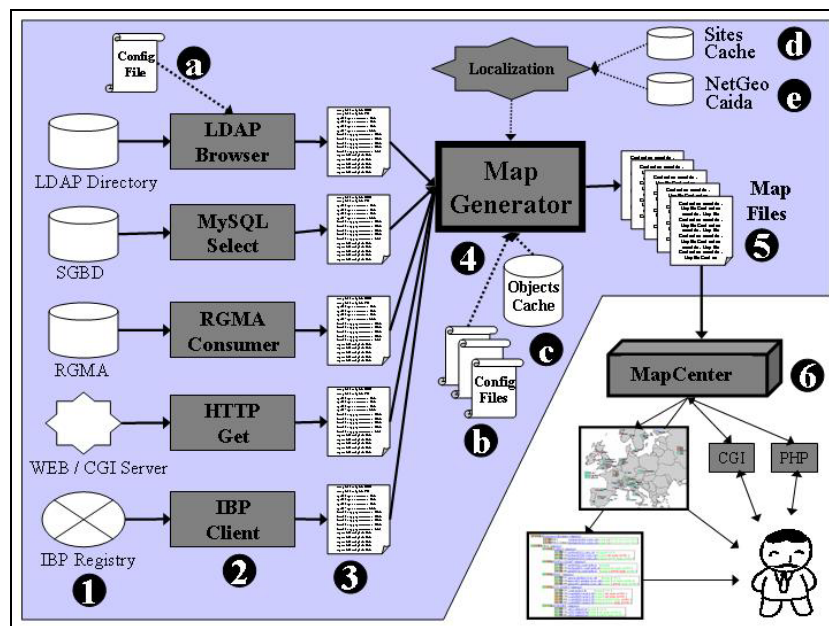


Figure 3 - Architecture du processus de découverte automatique

La ligne directrice de l'architecture proposée (Figure 3) est la généricité, basée sur une grande modularité afin de pouvoir s'interfacer très rapidement aux divers systèmes distribués existants et futurs avec le minimum de développements.

Le processus de découverte se décompose en trois phases principales :

- En fonction du type de la source d'information contenant les objets du système à superviser ❶, un **module de découverte** ❷ adapté est exécuté pour générer un fichier ❸ contenant la liste de ces objets, ainsi que leurs types.

- A partir de ce fichier, un **générateur de carte** (4) est ensuite exécuté pour générer le fichier de données (Map File (5) ) reconnu par MapCenter, contenant tous les objets, les services, les symboles, les cartes, les liens et les vue logiques et graphiques.
- MapCenter lui-même est ensuite exécuté (6) à partir de ce fichier de donnée pour réaliser les fonctions de supervision et de visualisation souhaitées

L'ensemble de ce processus de découverte est relancé régulièrement, avec une fréquence dépendant de la volatilité de l'environnement supervisé. A titre d'exemple, dans le cadre des grilles de calcul existante (DataGRID, LCG...), ce processus est actuellement exécuté une fois par jour.

## 2.2 Modules de découvertes

Cinq modules de découverte ont été développés, correspondant à différents systèmes d'information existant actuellement dans le domaine des grilles et des environnements distribués

- Un module d'interrogation des annuaires LDAP, entièrement configurable (via le fichier de configuration (7)) qui permet entre autres de retrouver tous les éléments stockés dans le MDS [8] d'une grille basée sur Globus Toolkit 2.
- Un module d'interrogation des bases de données MySQL, basés sur les requêtes SQL *select* standard.
- Un module d'interrogation de RGMA ([7]), basé sur un consommateur RGMA accédant à toutes les informations produites par l'ensemble des producteurs RGMA de la grille
- Un module HTTP, utilisant wget, lynx ou tout autre butineur en ligne de commande, permettant d'accéder à toutes listes de machines stockées en format brut dans des pages HTML accessibles par Internet.
- Un module spécifique pour le L-Bone, utilisant le client C de IBP.

Tous ces modules permettent d'ores et déjà de lancer des requêtes de découverte automatique sur le majorité des grilles existantes. De part notre architecture modulaire, de nouveaux modules pourront facilement être développés et intégrés pour accompagner les nouveaux systèmes d'information émergents (OGSA, Web Services, WSDL/XML...).

## 2.3 Générateur de cartes

Une fois l'ensemble des objets découverts, il est ensuite nécessaire d'organiser et de formater ces données pour les rendre conformes au format du fichier de données de MapCenter, et pour représenter l'ensemble des vues logiques et des cartes graphiques souhaités. C'est le rôle du module « générateur de cartes ». Ce module est générique et unique pour l'ensemble des modules de découvertes décrit précédemment.

Différents fichiers de configuration (8), adaptables pour chaque grille ou environnement distribué, permettent de définir l'ensemble des services implémentés dans cet environnement, les services associés à chaque type d'objet, la liste des vues logiques souhaités, et enfin l'ensemble des cartes graphiques générées avec leurs informations de localisation géographique associées.

La génération des cartes graphiques nécessite d'obtenir les informations de localisation géographique des objets et des sites. Plusieurs mécanismes ont été implémentés pour réaliser cette localisation automatique :

- L'identification des sites s'effectue par extraction du nom de domaine de chaque objet. Par exemple, nous considérons que l'objet « *grid1f.cnaf.infn.it* » fait parti du site « *cnaf.infn.it* ». un fichier spécial (cache des sites (9)) contient les informations de localisation géographique des sites principaux (longitude et latitude). Un objet d'un site répertorié est ainsi immédiatement localisé.
- Ensuite, l'outil Netgeo [13] (10) est utilisé pour les objets dont le site est inconnu dans notre cache des sites. Cet outil fournit de bon résultat pour les noms de domaine pas trop récent.
- Enfin, chaque objet découvert est stocké dans un autre fichier spécial (cache des objets (11)). Les informations géographiques sont conservées dans ce fichier de cache des objets pour une localisation immédiate lors de la découverte suivante.

Une des difficultés majeures de la découverte automatique des ressources à intervalle régulier réside dans le fait qu'elle ne permet pas de refléter les arrivées et départs d'objets et de services en temps réel. Le fichier de cache des objets nous permet également d'associer une durée de vie à chaque objet ; ainsi un objet est réellement supprimé de notre système après plusieurs recherches successives ne référant pas cet objet.

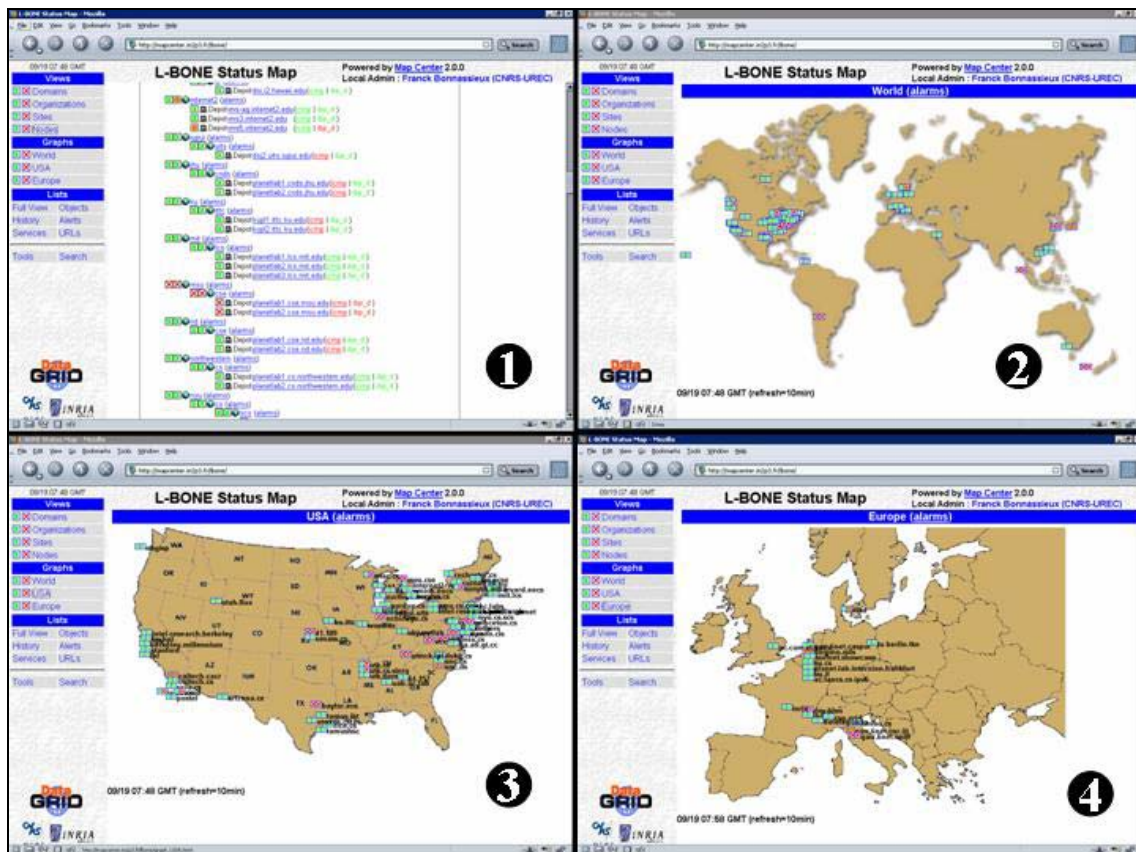


Figure 4 - *Vues et Cartes automatiques*

La Figure 4 fournit quelques exemples de vues logiques et de cartes graphiques découvertes automatiquement sur le L-BONE : ❶ est une vue logique des sites et des dépôts IBP sur ces sites, ❷ ❸ et ❹ sont des cartes géographiques des dépôts IBP dans le monde, et des zooms sur les USA ainsi qu'en Europe.

### 3 Accès aux systèmes d'information

Les systèmes d'information des grilles actuelles sont extrêmement divers et basés sur des technologies hétérogènes et incompatibles (Annuaire LDAP, SGBD, RGMA [7], ...) L'émergence des Web Services, utilisant SOAP, WSDL et le XML, ainsi que OGSA [9] permettent d'espérer de voir apparaître des standards plus fort structurant ces systèmes. Cependant, l'état actuel des implémentations des systèmes de grille contraint l'utilisateur (chercheur, administrateur système, responsable d'organisation virtuelle...) à utiliser et à maîtriser une multitude d'outils et d'interfaces clientes différentes pour accéder à l'ensemble des informations.

Au delà des possibilités de supervision et de visualisation, MapCenter a été également conçu pour répondre à ces besoins en permettant l'accès centralisé à l'ensemble des systèmes d'informations, de configuration et d'état de ces environnements. Pour cela, plusieurs plug-ins ont été développés en PHP ou sous forme de script CGI pour accéder directement à l'ensemble de ces informations depuis les différentes vues logiques, et ce de façon complètement transparente pour l'utilisateur final.

#### 3.1 Annuaire LDAP

Les annuaires LDAP sont largement utilisés pour organiser et stocker de nombreuses informations. Ils peuvent référencer l'ensemble des utilisateurs finaux, l'appartenance de ces utilisateurs aux diverses organisations virtuelles, l'authentification des personnes et des ressources aux travers de certificats, ainsi que les ressources de calcul et de stockage elle-même au travers du MDS [8] par exemple.

Une interface PHP implémentant un browser LDAP est intégrée à MapCenter. Cette interface permet aux différents utilisateurs d'accéder et de parcourir hiérarchiquement n'importe quel annuaire LDAP aussi facilement qu'un explorateur de fichier graphique.

### 3.2 SGBD

Les SGBD occupe également une place de choix dans les environnements distribués et les grilles. On les retrouve par exemple pour stocker les ressources de calculs, les droits et les polices d'autorisation des utilisateurs et des processus, ou encore les données de monitoring réseaux entre les différents sites.

Un module de visualisation de table MySQL a été développé en PHP. Il permet de lancer n'importe quelle requête SQL directement à partir des vues logiques de MapCenter, sur l'ensemble des bases supervisées et autorisant de telle requêtes.

### 3.3 Appel aux Web Services et Parseur XML

Apparus depuis quelques années, les Web Services occupent de plus en plus le devant de la scène dans le domaine des grilles, et de nouvelles possibilités basées sur ces technologies apparaissent chaque mois. En attendant l'avènement annoncé d'OGSA [9], plusieurs services ont déjà été implémentés par plusieurs équipes ou projet de grilles, en général bâtis au dessus du moteur Axis d'Apache.

On peut ainsi citer au sein du projet DataGrid les systèmes de gestion des données (Replica Location Service, Replica Metadata Catalog et Replica Optimisation Service), ainsi que l'interface d'accès aux éléments de stockage, tout deux basés sur les Web Services.

Des modules PHP spécifiques ont été développés pour accéder à ces services, en réalisant des appels aux services via SOAP, et en analysant et en formatant les résultats XML fournis. De tels modules peuvent être très rapidement conçues et ajoutés à MapCenter pour suivre l'arrivée et l'évolution rapide des Web Services pour les grilles.

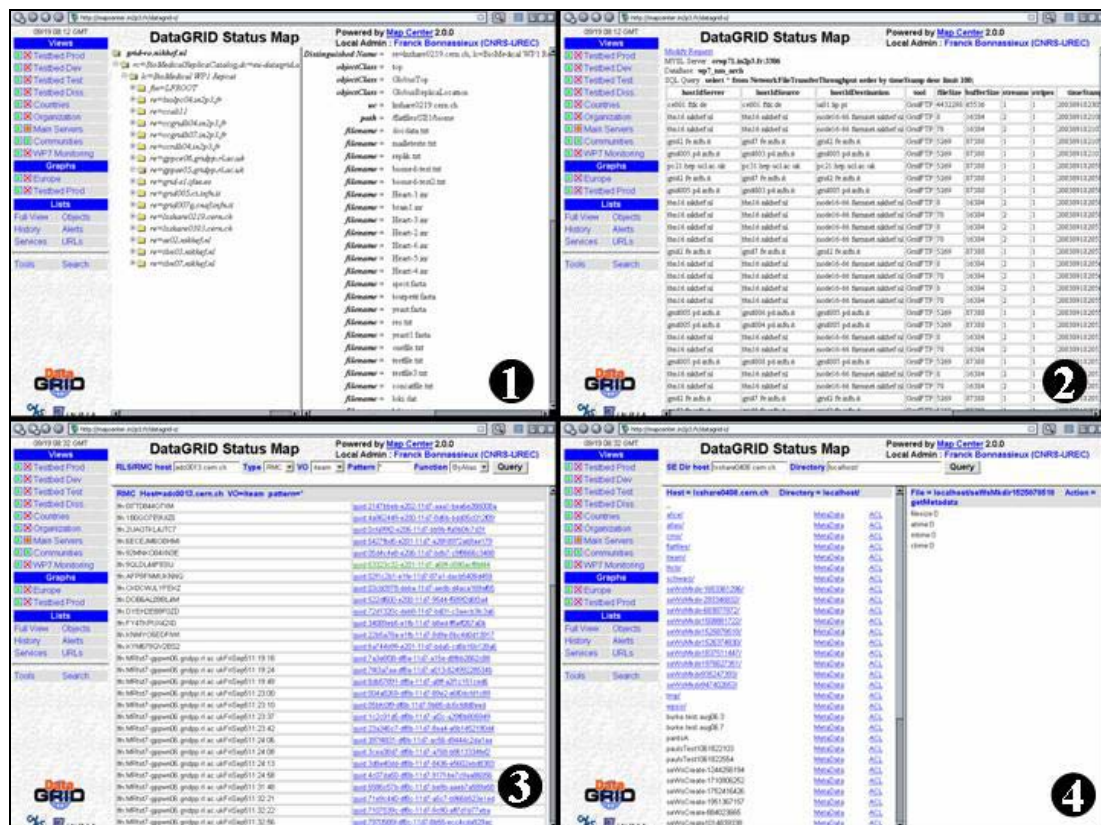


Figure 5 - Accès aux systèmes d'information

La Figure 5 fournit quelques exemples des interfaces d'accès à divers systèmes d'information :

- ❶ est l'interface du browser d'annuaire hiérarchique LDAP.
- ❷ est l'interface du module de visualisation de tables MySQL.
- ❸ est l'interface d'un module d'accès aux systèmes de gestion des répliquas (Replica MetaData Catalog, et Replica Location Service) implémentés sous forme de Web Services (Appel SOAP et analyse XML)
- ❹ est l'interface d'un module d'accès aux « Storage Elements » de DataGrid, implémentés également sous forme de Web Services.

## 4 Déploiement et résultats

MapCenter est aujourd'hui déployé et quotidiennement utilisé sur plus d'une douzaine de projets de grilles de calcul, ou d'environnements fortement distribués (Figure 6).



Figure 6 - Références MapCenter

Plusieurs facteurs permettent d'expliquer ce succès, qui relève principalement de sa facilité de mise en œuvre et de déploiement :

- Architecture centralisée, sans agents à déployer sur les sites.
- Respect des politiques de sécurité locale (aucun besoin de droits ou d'accès particulier, monitoring furtif des services TCP...)
- Découverte automatique des ressources et des services.
- Création immédiate de vues logiques et de cartes graphiques manuelles ou automatiques
- Accès instantané aux divers systèmes d'information de l'environnement supervisé.

Malgré son architecture centralisée, l'implémentation multi-threads du monitoring optimisé grâce à des timeouts adaptés, permet des performances élevées. A titre d'exemple, le site de DataGrid peut effectuer un cycle complet de monitoring de plus de 500 services sur plus de 200 machines réparties à travers l'Europe en moins de 2 minutes.

Les divers sites MapCenter supervisent aujourd'hui plusieurs milliers de machines à travers le monde. Des campagnes intensives de tests ont été effectuées et ont démontrées par la pratique la possibilité pour MapCenter de superviser depuis un seul serveur bas de gamme plusieurs milliers de nœuds en quelques minutes[2].

De tels résultats nous permettent d'affirmer que MapCenter est dès à présent prêt à relever les défis de demain : superviser en temps réel des grilles représentant des puissances de calcul énormes sur des peta-octets de données distribuées sur la planète.

## 5 Conclusion

MapCenter est maintenant utilisé par plusieurs grilles de calculs internationales majeures ([11]), et également par d'autres environnements fortement distribués. Son modèle flexible de présentation permet de fournir aux utilisateurs finaux une très large variété de tableaux de bords. Ses mécanismes de découverte automatique et son monitoring performant et transparent permettent de le déployer très rapidement sur de nombreux types d'environnements distribués. Le développement de nouvelles fonctionnalités et son déploiement continuera dans les années à venir.

## Références

- [1] Bonnassieux, F., Chanussot, F., Harakaly, R., Primet, P., "MapCenter : An Open Grid Status Visualization Tool", ISCA Parallel and Distributed Computing System Conference 2002, Louisville, Kentucky.
- [2] Bonnassieux F., Harakaly R., Primet P., "Automatic services discovery, monitoring and visualization of grid environments: the MapCenter approach". Across Grid conference, 13-14 Feb 2003, Santiago de Compostela, Spain
- [3] I. Foster, C. Kesselman, "the Grid: Blueprint for a New Computing Infrastructure", Morgan Kaufmann, Pub.ISBN 1-55860-475-8, August 1998.
- [4] Tierney, B., Aydt, R., Gunter, D., Smith, W., Taylor, V., Wolski, R., Swany, M, "A Grid Monitoring Service Architecture". Global Grid Forum Performance Working Group, 2001.
- [5] Tierney, B., Crowley, B., Gunter, D., Holding, M., Lee, J., Thompson, M, "A Monitoring Sensor Management System for Grid Environments". Proceedings of the IEEE High Performance Distributed Computing conference (HPDC-9), August 2000.
- [6] Fyodor, "The Art of Port Scanning". Phrack Magazine Volume 7, article 11 of 17, Issue 51 September 01, 1997.
- [7] Fisher, W. S., "Relational Model for Information and Monitoring" Grid Forum Informational Draft GWD-GP-7-1.
- [8] Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C. "Grid Information Services for Distributed Resource Sharing". Proceedings of the Tenth IEEE International Symposium on High-Performance Distributed Computing (HPDC-10), IEEE Press, August 2001.
- [9] Foster, I., Kesselman, C., Nick, J., Tuecke, S. "The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration". Open Grid Service Infrastructure WG, Global Grid Forum, June 22, 2002.
- [10] A. Iamnitchi, I. Foster, "On Fully Decentralized Resource Discovery in Grid Environments", International Workshop on Grid Computing, 2001
- [11] MapCenter Web Site for DataGrid: <http://ccwp7.in2p3.fr/mapcenter>
- [12] European DataGrid projet: <http://web.datagrid.cnr.it>
- [13] NetGeo - The Internet Geographic Database <http://www.caida.org/tools/utilities/netgeo/>