

Le réseau pour les grilles de calcul

L'exemple du projet EGEE

6^e Journées RÉSeaux (Marseille, déc. 2005)

Jean-Paul Gautier, Mathieu Goutelle (CNRS UREC)

- **Quelques mots sur EGEE ;**
- **Le réseau et la grille :**
 - Quels sont les besoins ?
 - Quels sont les services ?
- **Les activités réseau dans EGEE ;**
- **Utilisation d'un service réseau :**
 - Mise en place de "Service Level Agreement" ;
 - « Réservation » du service ;
 - « Activation » du service.
- **Monitoring ;**
- **Interface opérationnelle;**
- **Conclusion.**

- **Un projet européen sur 2 ans dans un programme de 4 ans (2004-2008), 70 partenaires (850 personnes) répartis dans 27 pays, environ 32 M€ pour les deux premières années ;**
- **Fédérer les grilles régionales et nationales en une architecture homogène :**
 - Plus de 170 sites dans le monde,
 - Plus de 17 000 CPU,
 - Environ 5 Po de capacité de stockage.
- **Un nouveau middleware capitalisant l'expérience accumulée ;**
- **Applications :**
 - Physique des Hautes Énergies (LHC),
 - Applications biomédicales,
 - Sciences de la Terre et Géophysique,
 - Astronomie,
 - Simulation financière,
 - ...
- **Et le réseau au milieu !**





- Pays fournissant des ressources
- Pays envisageant de participer

Grille EGEE / LCG :

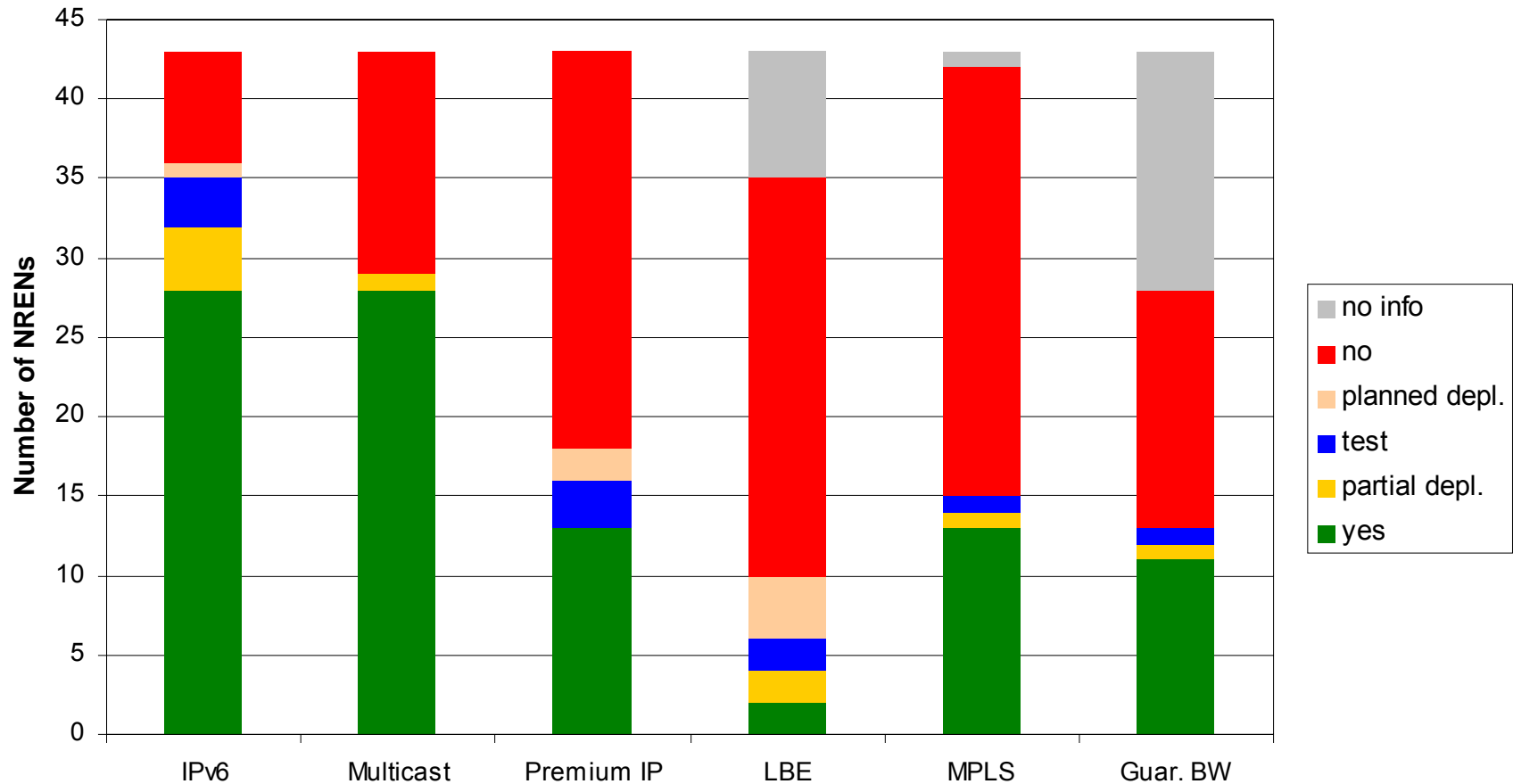
- 174 sites, 40 pays
- plus de 17 000 CPU
- environ 5 Po de stockage

Autres grilles nationales et régionales : ~60 sites, ~6000 CPU

Country	sites	Country	sites	Country	sites
Austria	2	India	1	Russia	10
Belgium	1	Israel	2	Singapore	1
Bulgaria	4	Italy	25	Slovakia	3
Canada	6	Japan	1	Slovenia	1
China	1	Korea	1	Spain	13
Croatia	1	Netherlands	2	Sweden	2
Cyprus	1	Macedonia	1	Switzerland	2
Czech Republic	2	Pakistan	2	Taiwan	4
France	8	Poland	4	Turkey	1
Germany	8	Portugal	1	UK & Ireland	35
Greece	6	Puerto Rico	1	USA	3
Hungary	1	Romania	1	Yugoslavia	1

- **Le réseau est l'épine dorsale de la grille :**
 - Sans réseau haute performance, pas de grille !
- **Souvent oublié par la communauté grille :**
 - « Fonctionne depuis longtemps et de façon assez transparente » ;
 - « Ressource fiable et performante » ;
 - « Hors du domaine de responsabilité de la grille ».
- **Aller au-delà de Best Effort :**
 - Pour satisfaire les besoins des applications (transfert de données, accès aux données, trafic interactif) ;
 - Débit garanti, délai garanti ;
 - Problématique inter-domaine non-spécifique à « la grille ».
- **Faire reconnaître le réseau comme une ressource de la grille :**
 - Comme le calcul ou le stockage ;
 - Cela implique des fonctionnalités comme l'approvisionnement, la gestion, la surveillance voire la réservation de cette ressource.
- **Collaboration indispensable avec les fournisseurs de services (réseaux académiques ou commerciaux) pour y parvenir.**

- **Les applications sur la grille ont besoin (entre autres) :**
 - De transférer de grandes quantités de données (HEP) en continu ;
 - D'accéder à de grands volumes de données (HEP, Biomed) ;
 - D'échanger des messages interactivement (Biomed) ;
 - De réserver des ressources à l'avance (Biomed) ;
 - D'accéder à des ressources externes à la grille et inversement (avec une certaine qualité de service).
- **Ceci implique (entre autres...) :**
 - Des liaisons haut-débit garanti ;
 - Des liaisons haut-débit (mais avec des contraintes temporelles moins fortes) ;
 - Des liaisons à faible délai (ou à délai minimum) ;
 - Des mécanismes de réservations inter-domaines.



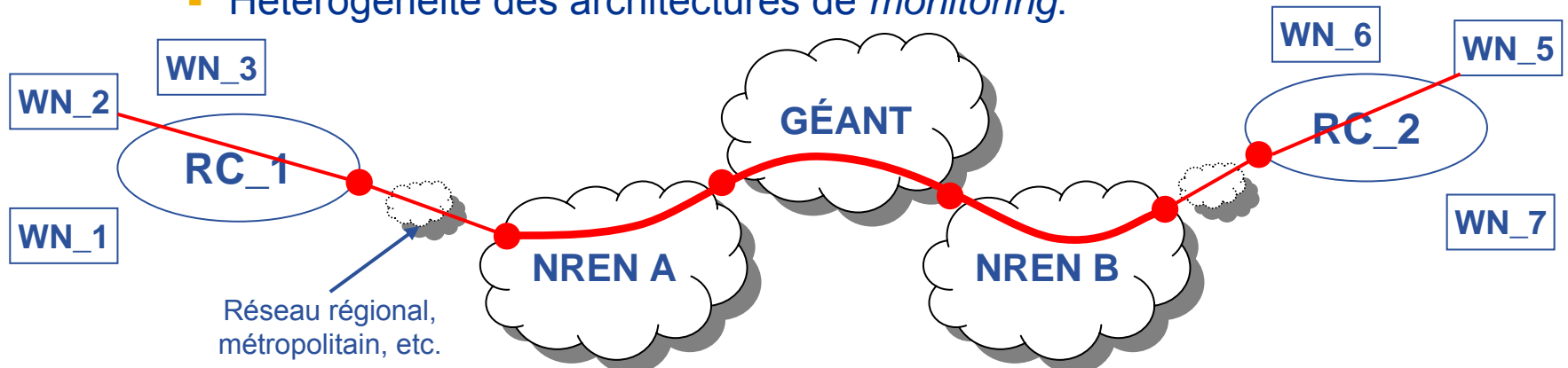
- “yes” signifie que le service est disponible à 100%.
- “no info” signifie qu’aucune information n’a pu être récupérée.

- **SA2 : “*Network Resource Provision*”**
 - “*Technical Network Liaison Committee*”,
 - Interface opérationnelle avec les réseaux,
 - Définition de SLA entre EGEE et les réseaux (Géant et réseaux nationaux).

- **JRA4 : “*Network Services Development*”**
 - “*Network Performance Monitoring*”,
 - “*Bandwidth Allocation and Reservation*”,
 - Évaluation des bénéfices d’IPv6 pour les grilles.

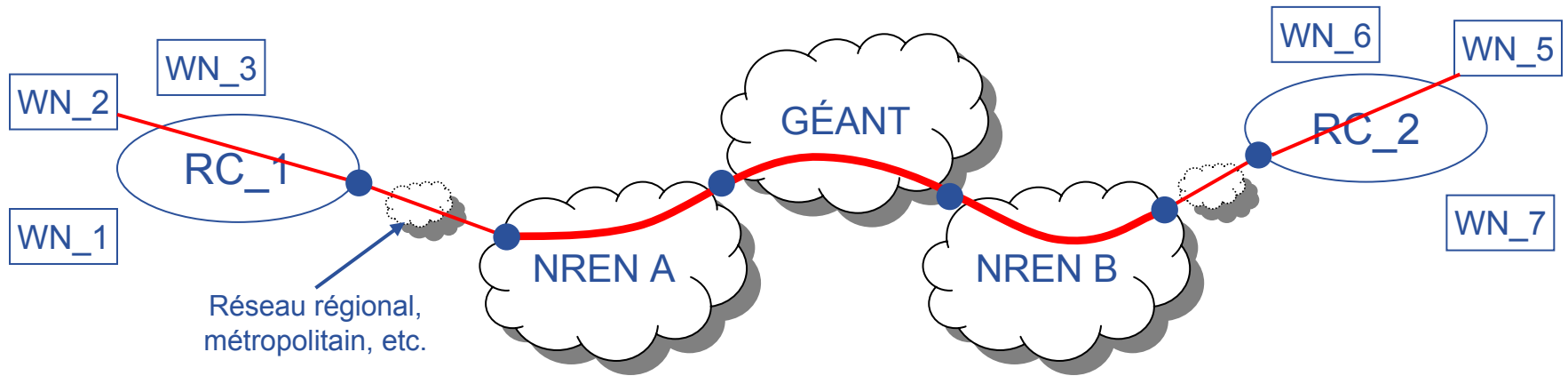
**Construire de fortes relations de travail
entre EGEE et les fournisseurs de réseau**

- **Un besoin particulier des applications :**
 - Échange d'informations avec nécessité de garanties :
 - Capacité réservée à ces flux : 20 Mbit/s ;
 - Délai minimum, gigue « constante ».
 - Cette application est exécutée régulièrement entre deux sites :
 - Une application peut représenter plusieurs milliers de tâches à exécuter sur les fermes de PC ;
 - On ne connaît pas à l'avance les machines source et destination des flux ;
 - D'autres tâches s'exécutent sur la même ferme simultanément.
- **Problématique inter-domaine très forte :**
 - Différents services réseaux (pas de nécessairement compatibilité) ;
 - Différents politiques opérationnelles (attribution des services, gestion des incidents) ;
 - Hétérogénéité des architectures de *monitoring*.



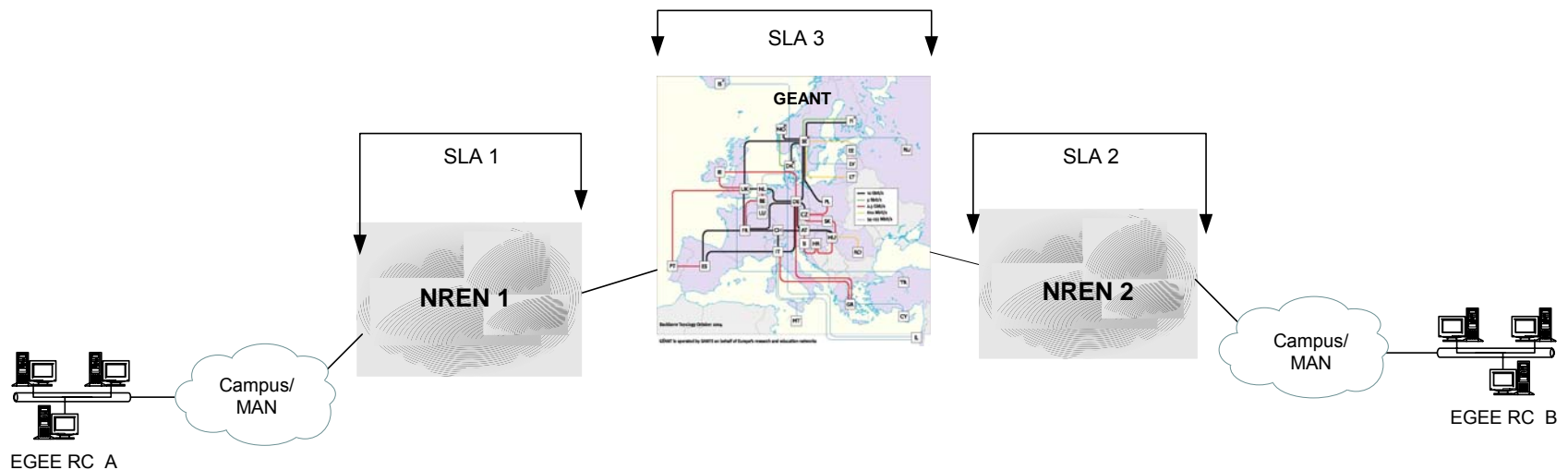
- **Pour utiliser un service réseau, plusieurs étapes :**
 - Établissement d'un modèle de SLA :
 - Accord entre EGEE et un domaine sur la fourniture d'un service ;
 - Validité à long terme (un an voire plus, la durée d'un projet),
 - Différents modèles selon le niveau de services.
 - Réservation d'un service :
 - Allocation de ressources à moyen terme (semaine, mois),
 - Agrégation des SLA de chaque domaine concerné,
 - Configuration du réseau pour autoriser et mettre en forme le trafic (peut induire un délai de configuration).
 - Activation du service :
 - Utilisation du service à court terme (durée d'un job, ~journée),
 - Configuration des LAN pour autoriser et mettre en forme le trafic,
 - Pas de délai de configuration (réponse immédiate !).
- **Les deux dernières étapes doivent être automatisées (interface logicielle).**

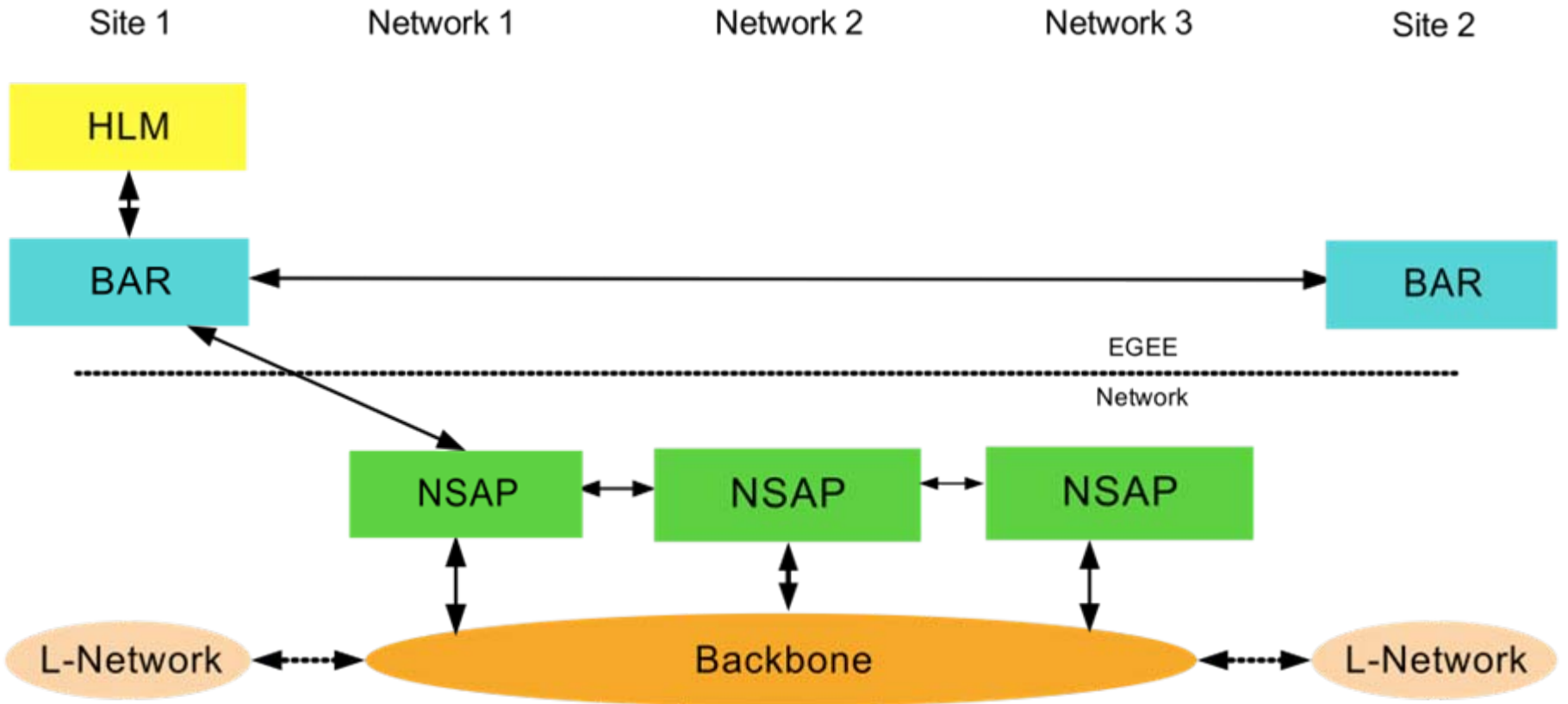
1. Modèles de SLA en place pour chaque domaine ;
2. “*Service Reservation*” : allocation de la ressource de bordure à bordure ;
3. “*Service Activation*” : allocation de la ressource de bout-en-bout.

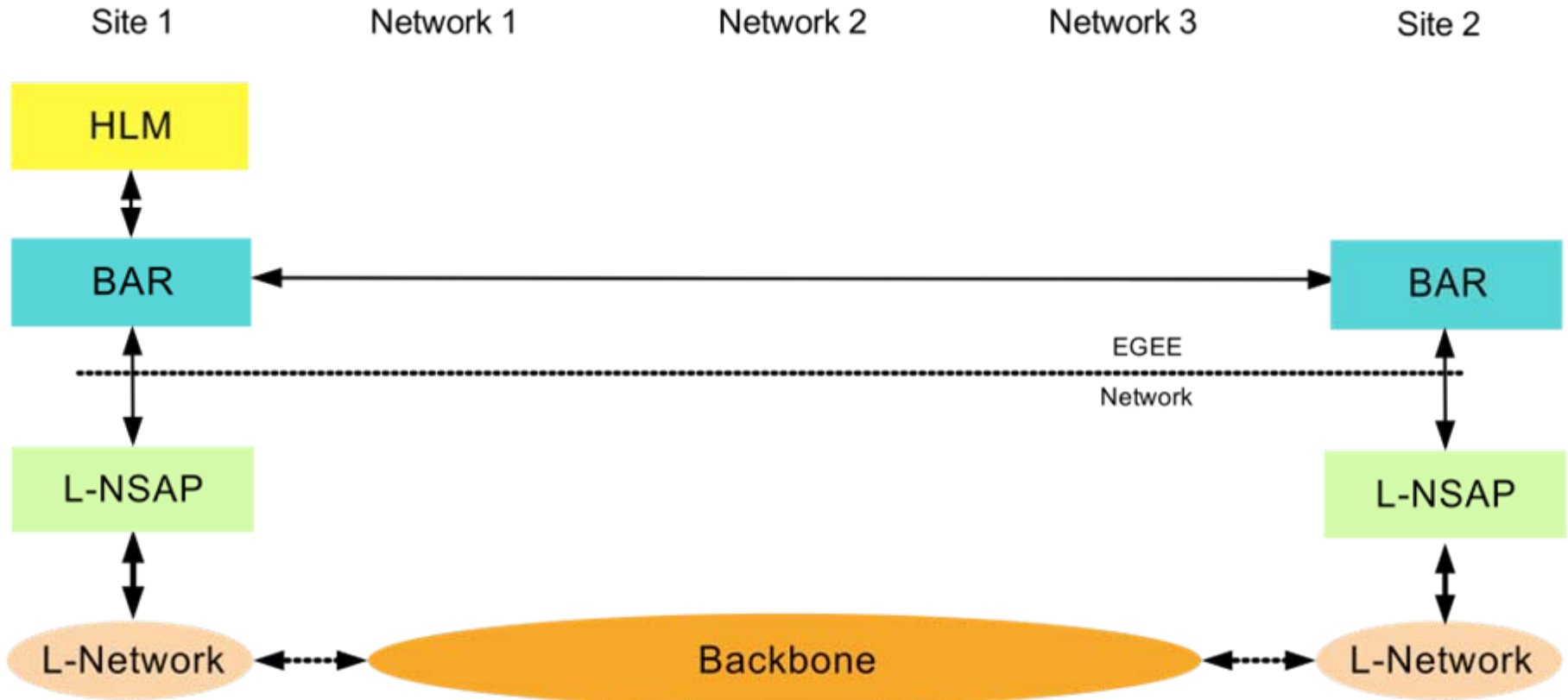


- Configuration du réseau de cœur ;
- Choix des extrémités du chemin (émetteur et récepteur) ;
- Configuration des équipements des sites pour autoriser uniquement les extrémités précédemment choisies et mettre en conformité le trafic (gestion de la capacité à l'intérieur de la réservation).

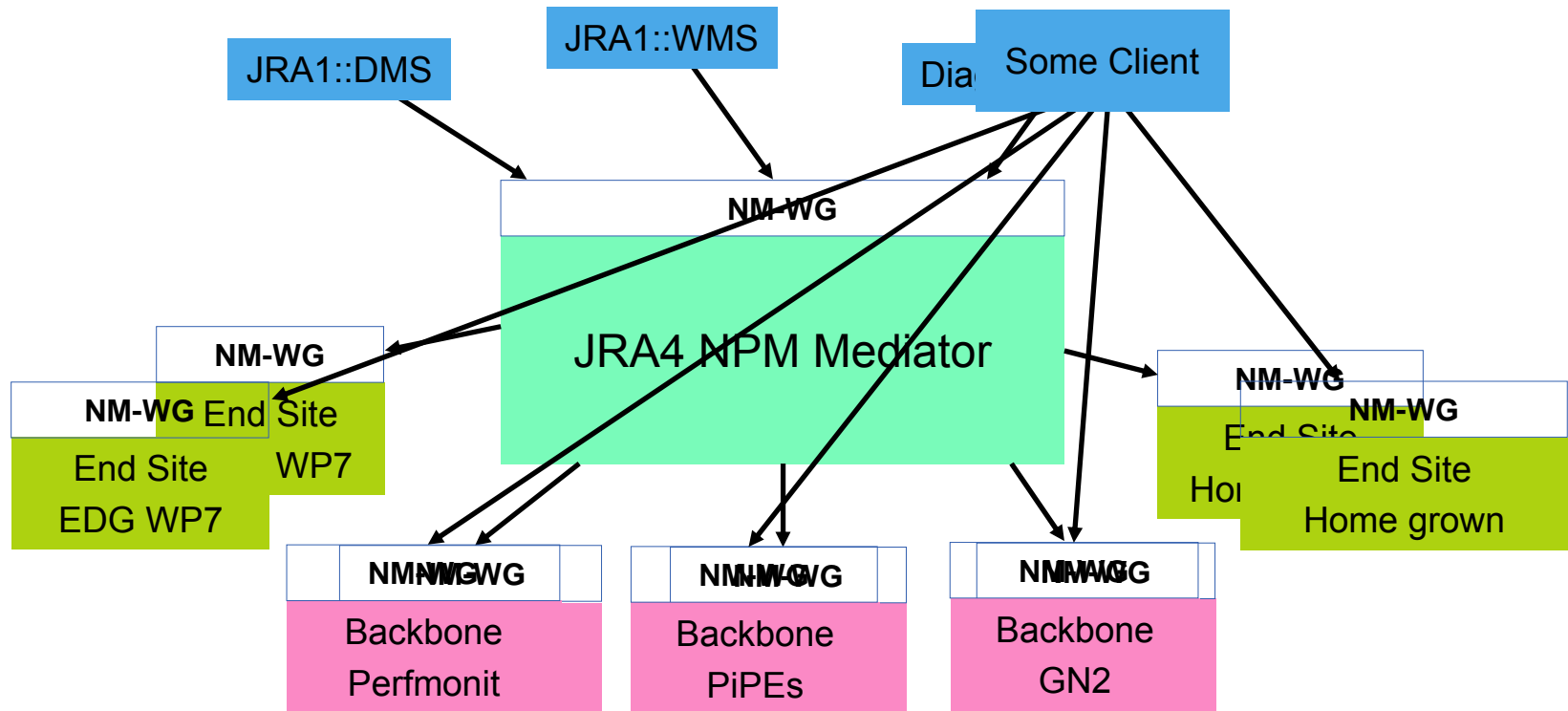
- **Définition de SLA :**
 - S’appuie sur des travaux et études antérieurs ;
 - Définition en collaboration avec GN2.
- **Modèle de SLA de bout-en-bout pour EGEE :**
 - « Agrégation » des SLA de chaque domaine entre les « frontières » du nuage réseau ;
 - Problématique du *monitoring* des paramètres du SLA ;
 - Difficulté pour s’adapter et prendre en compte le “*last mile*”.



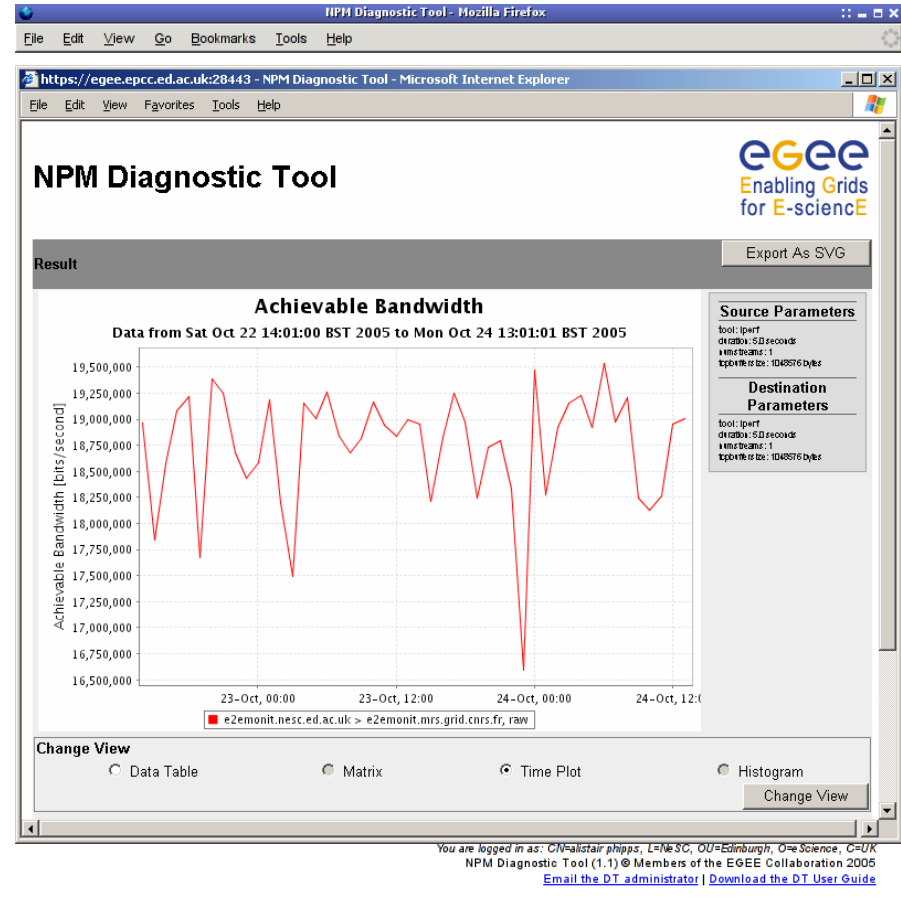




- Pas de travail dans EGEE pour produire encore une autre infrastructure de mesure :
 - Collecte des données dans les infrastructures existantes ;
 - Publication à travers une interface commune ;
 - Extensibilité intrinsèque du modèle.



- Développement d'un outil de diagnostic :
 - Présentation des données fournies par différents outils de monitoring (bout-en-bout, cœur du réseau) ;
 - Pour les équipes opérationnelles (diagnostic a priori et a posteriori).
- Utilisation des données par le middleware :
 - Choix des sites ;
 - Optimisation des transferts de données ;



- **Indispensable à la construction d'une infrastructure de grille opérationnelle ;**
- **Dialogue entre le support utilisateur de la grille et les centres opérationnels des réseaux :**
 - Information des utilisateurs sur les problèmes (incident, maintenance) dans les réseaux ;
 - Résolution des problèmes signalés par les utilisateurs (coupure, performances) en interaction avec les réseaux concernés.
- **EGEE Network Operations Centre (ENOC) :**
 - Point d'entrée unique pour les réseaux ;
 - Point de contact unique (vu des utilisateurs) pour les problèmes réseaux ;
 - Sera capable de surveiller l'état du « réseau EGEE » en utilisant les informations de *monitoring* fournies par les réseaux.

- **Problématique inter-domaine :**
 - Grande hétérogénéité du réseau dans le monde de la grille ;
 - Propriété inhérente à l'infrastructure ;
 - Les mécanismes élaborés sont transposables.
- **Mais besoin de services de bout-en-bout :**
 - Services réseaux mais aussi services opérationnels ;
 - Implémentation de la réservation inter-domaine dans la communauté des réseaux nationaux (activités GN2-SA3 et GN2-JRA3 de Géant2) ;
 - Extension possible à Internet en général — voir la présentation 68 (J.M. Uze) cet AM ou <http://www.ipsphere.org/>.
- **Touche aussi les réseaux de campus (si un cluster s'y installe) !**
 - Interface opérationnelle avec la grille ;
 - Interface avec les services de réservation ;
 - Problème identique par cascade pour les réseaux « intermédiaires » (métropolitains, régionaux...).
- **Tendance : donner plus de « contrôle » du réseau à l'utilisateur ?**
 - La grille n'est pas un utilisateur particulier des services réseaux ;
 - Cela pourrait profiter à toutes les communautés d'utilisateurs.

Merci à vous !

Des questions ?