Un nouvel outil de monitoring du point de vue du site

By Julia Andreeva, Max Boehm, Adrian Casajus, Benjamin Gaidioz, Costin Grigoras, Lukasz Kokoszkiewicz, **Elisa Lanciotti**, Ricardo Rocha, Pablo Saiz, Irina Sidorova, Andrei Tsaregorotsev.
CERN, IT Dept.

Réunion des sites LCG-France
Grenoble 27-28 Novembre 2008

Plan

- Introduction
  - Motivation et objectif de l'activité
- Structure du nouvel outil de monitoring
  - Flux de l'information de la source jusqu'à' l'affichage final
- Objet du monitoring:
  - Les activités principales des VOs aux sites
- L'affichage: Gridmap
  - Structure du Gridmap
  - Apparence du premier prototype
- État courant de l'activité et questions ouvertes

Monitoring des activités pendant le CCRC08

- Après l'expérience du CCRC08, impressions des administrateurs des sites:
  - Les principaux outils de monitoring utilisés pendant cet expérience étaient des outils spécifiques pour les VOs. Dans la plupart des cas, ils ont bien fonctionné, mais ils ne sont pas très faciles á utiliser pour quelqu'un qui est externe á la VO
  - Les sites qui servent plusieurs Vos ont du apprendre á utiliser des outils très différents entre eux: temps d'apprentissage nécessaire et peu d'efficacité
  - Comme résultat, souvent les administrateurs des sites ne savaient pas bien si son site contribuait á l'activité de la VO comme on s'y attendait
- Les administrateurs des sites voudraient avoir la possibilité de comparer la performance du site de point de vue de la VO avec la vision qu'ils obtiennent de ses propres systèmes de monitoring

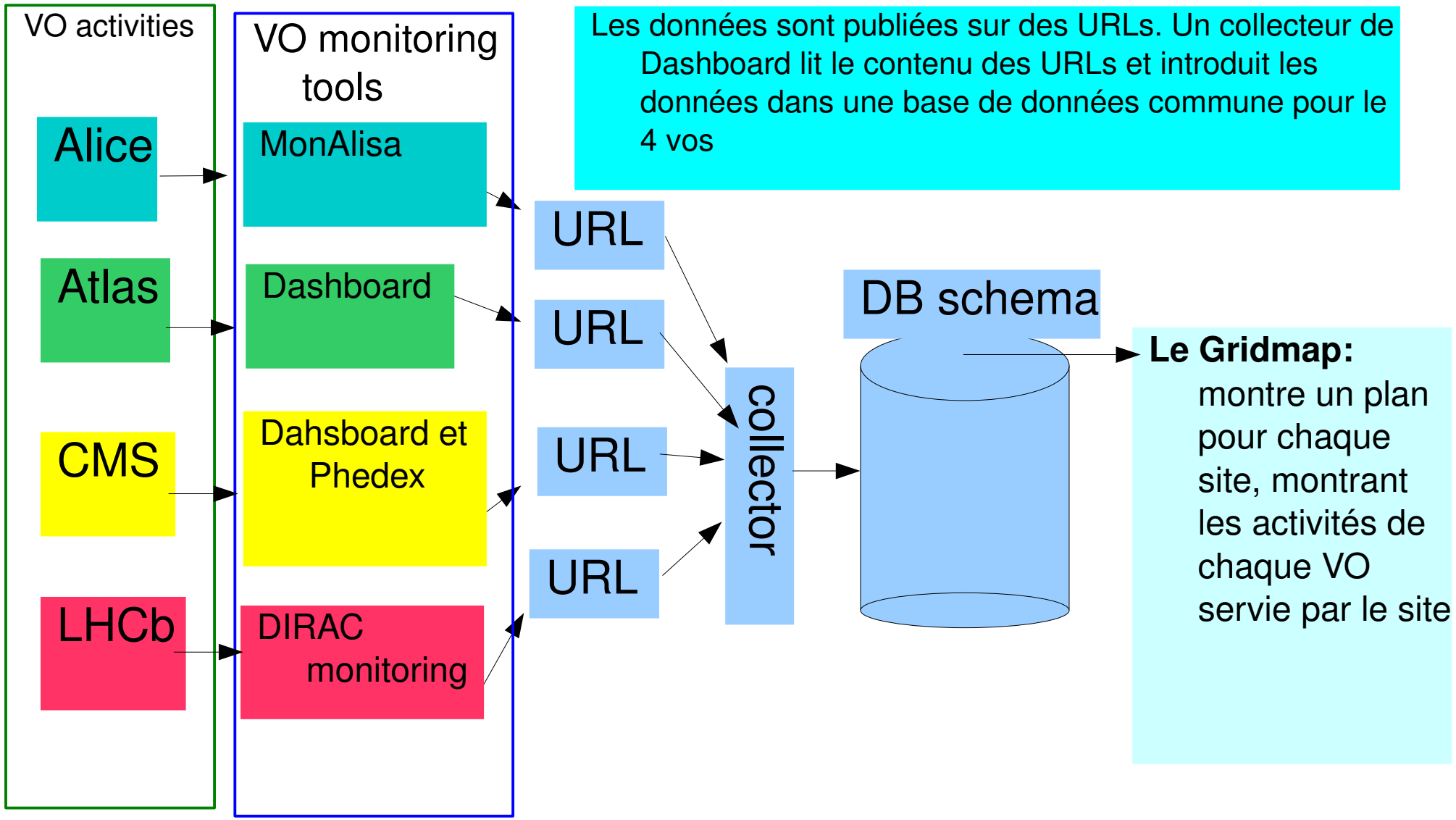
Objectif de cette activité

Fournir un nouvel outil de monitoring qui devrait:

- Á partir de la même vue, donner une vue générale de l'état des services dans le site. Il devrait être un outil simple á utiliser, même pour quelqu'un externe á la VO, et qui n'a pas de connaissances spécifiques de l'expérience --> **un aperçu général des activités du site á partir de une unique vue**
- Cet outil obtiendra l'information des outils de monitoring particuliers de chaque VO (Dashboard, MonALISA, Dirac, Phedex) et montrera le résultats essentiels de une façon simplifiée, en fournissant des liens vers les sources de l'information --> **un outil de haut niveau**
- Il montrera l'information en utilisant la technologie de Gridmap --> visualisation **immédiate des problèmes grâce aux plans avec couleurs**

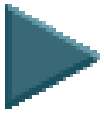
Une requête ultérieure des sites c'est aussi de obtenir une définition claire des objectifs des VOs relativement aux activités qui se déroulent au site: cette information devrait être montrée par cet outil á coté des valeurs observées pour permettre une évaluation des performances

Flux de l'information



Objets du monitoring: fournis par les outils de monitoring particuliers de chaque VO

- L'information sur les activités des VO est obtenue des différents outils de monitoring spécifiques de la VO :
  - Dashboard, MonALISA, DIRAC, Phedex
- Pour construire une base de données commune aux 4 VOs il est nécessaire:
  - Identifier des grandeurs qui sont communes á toutes les VOs
  - Définir clairement le signifié de chaque grandeur
  - Vérifier que ces grandeurs sont utiles pour le site et que elles sont disponibles dans les outils de monitorage



Effort de coordination: Définir un group de grandeurs qui ont le même signifié pour toutes le VOs

Construire un modèle de donnes et dessiner un schéma pour la base de donnés commune -> une seule et unique base de données pour le grandeurs des différentes expériences

Activités et grandeurs objet du monitoring

ACTIVITÉS

état général du site

job processing

transfert de données

...

GRANDEURS.

Nombre de jobs, nombre de jobs achevés (pendant la dernière heure)

wall time et temps CPU (si possible en KSI2K)

...

taux de réussite (moyenne de la dernière heure)

taux de transfert (moyen dans la dernière heure)

...

Il y a 2 activités principales: job processing et transfert de données Leurs grandeurs sont rassemblés pour les 4 expériences LHC

Pour chacune des activités principales il y a un groupe de activités secondaires: elles sont optionnelles et dépendent de la VO et de son modèle de calcul

Job Processing: MC production, data reconstruction, user analysis, SAM tests...

Data transfer: production transfer, transfer t0-t1, transfer t1-t1...

Le collecteur

- Un nouveau collecteur a été développé dans le cadre du Dashboard, en analogie avec le collecteur du module Site Status Board (développé par Pablo Saiz pour CMS)
- Le collecteur se exécute périodiquement (toutes les heures), lit les grandeurs publiées dans les URLs par les outils de monitoring des Vos, et les introduit dans la base de données

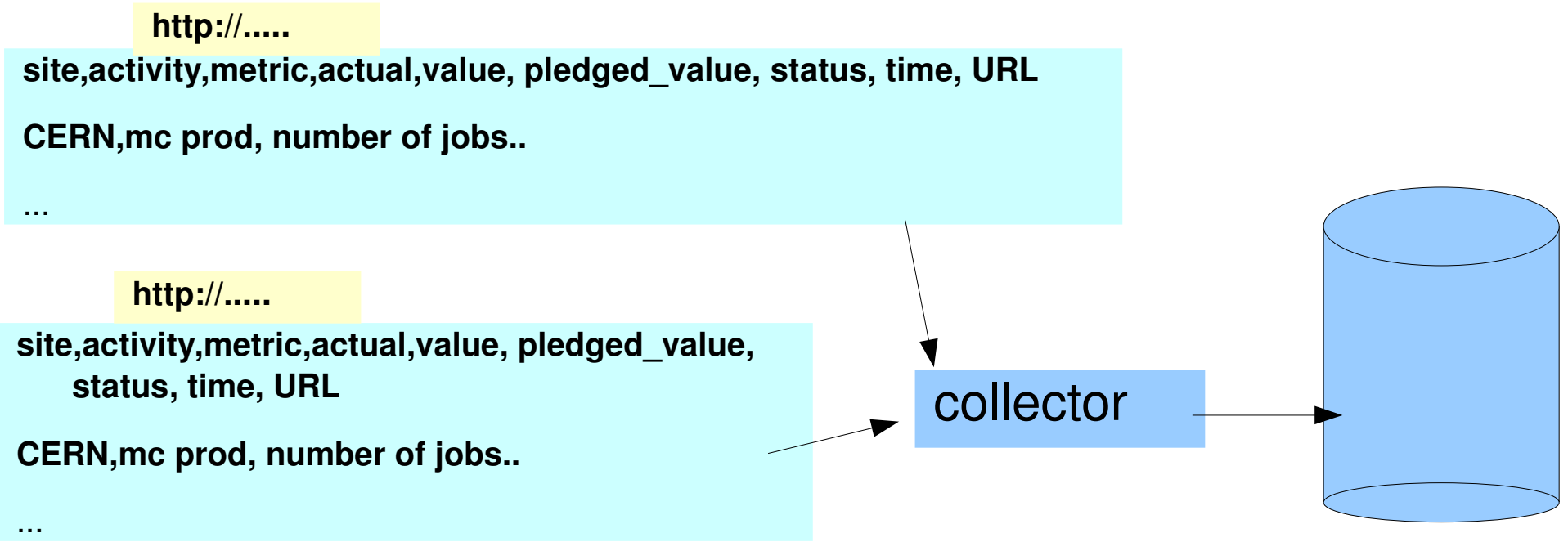
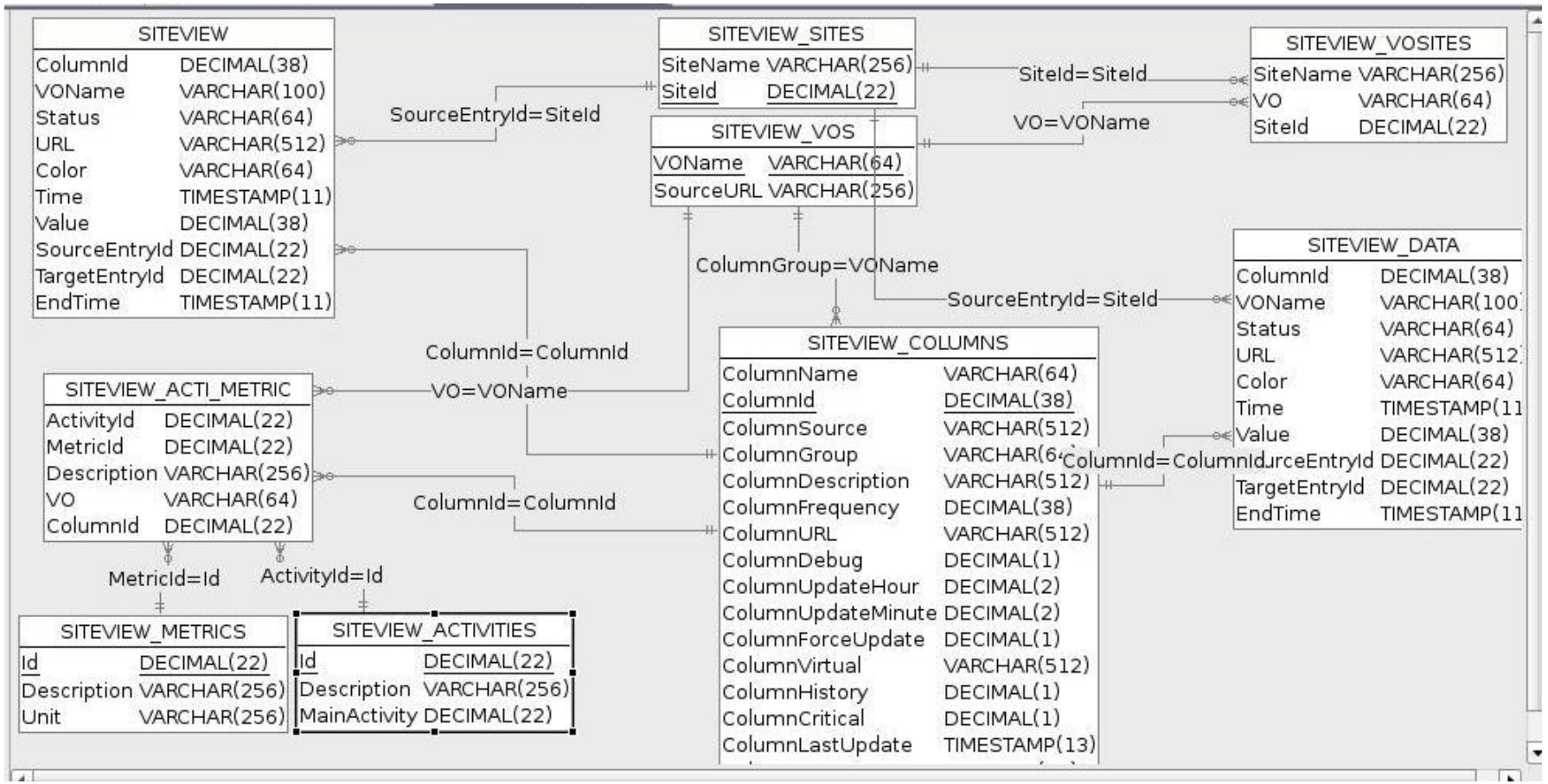


Schéma de la base des données

Implémentation sur un serveur de test de Oracle



Très flexible: d'autres activités et grandeurs peuvent être ajoutées ensuite, selon le besoin de la VO

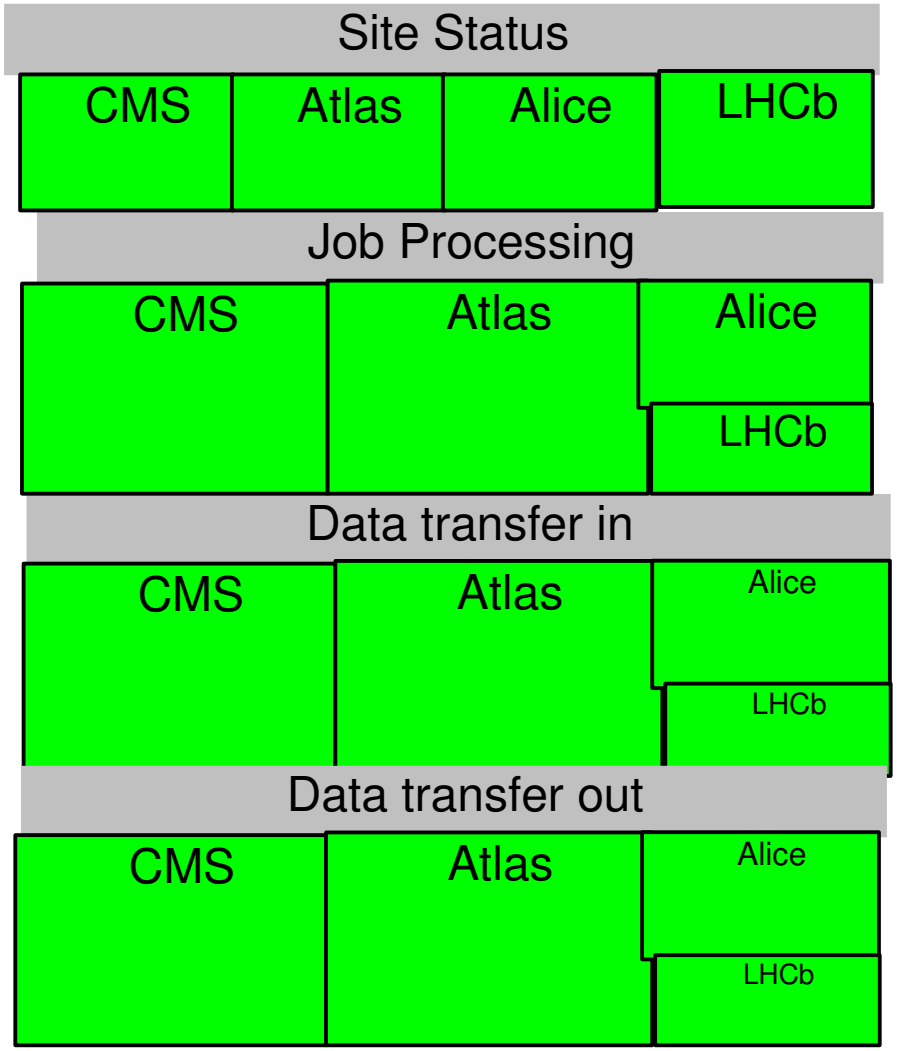
Le Gridmap

- Un Gridmap pour chaque site
- Dans le Gridmap du site, la hiérarchie est la suivante:
  - premier niveau: Activités
  - second niveau: VO

Le Gridmap principal ne montre que les activités principales, pour toute VO servie par le site

Les dimensions du rectangle qui représente la VO est proportionnel á une grandeur qui caractérise son activité (i.e. le nombre de jobs parallèles). Un offset constant est ajoute, á fin de pouvoir toujours visualiser toutes le Vos, même si elles ont une très faible activité en ce moment.

L'état (=COLOR) devrait être fourni para la VO. Chaque VO peut avoir une définition de état indépendante

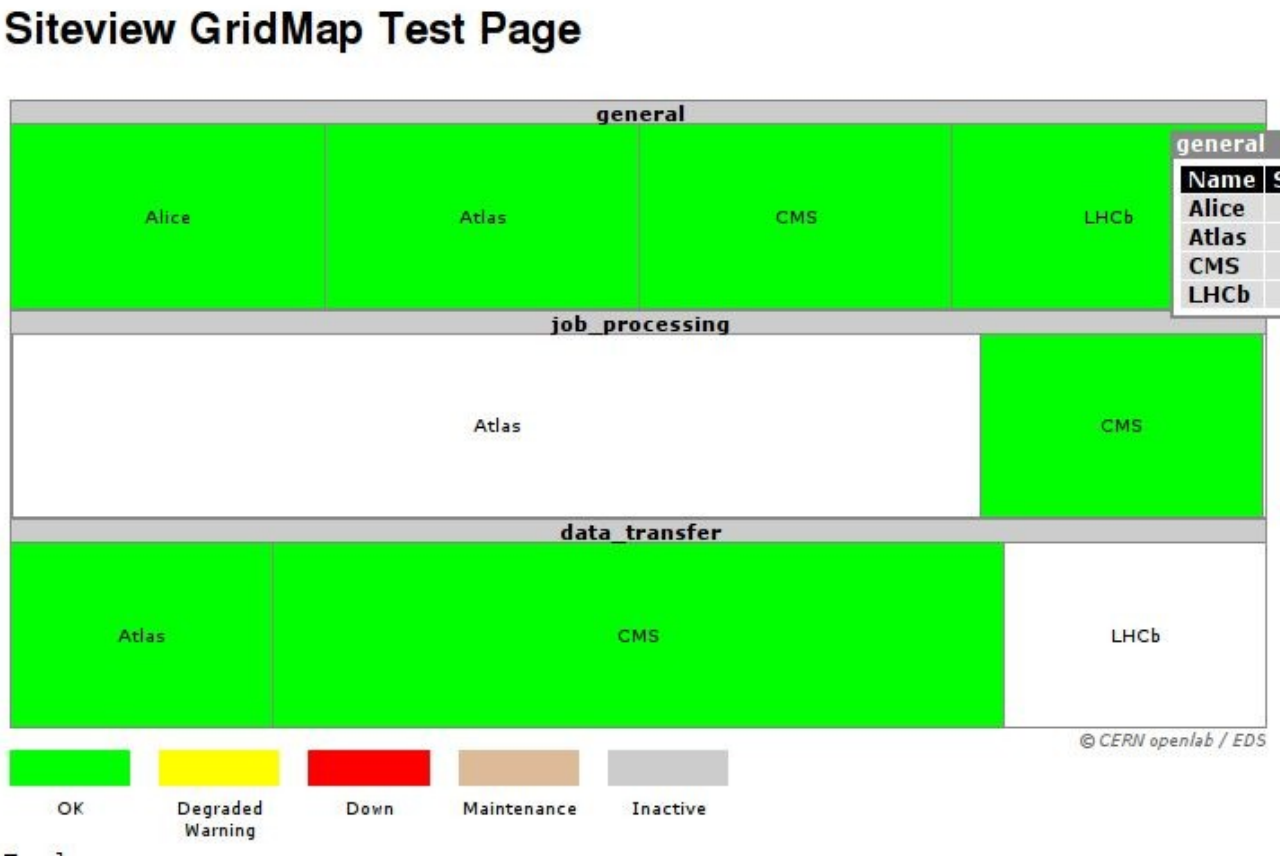


L'état du site

- Il s'agit d'une évaluation générale de l'état du site, du point de vue de la VO. Si l'état est 'ok' (vert), alors le site est en bon état pour la VO, même s'il n'y a aucun job en exécution
- Pour LHCb et Atlas l'état du site est calculé á partir des tests SAM spécifiques pour la VO
- Pour CMS non seulement le tests SAM, mais aussi le taux de réussite de l'activité de production, en ne prenant en compte que les échecs dus á un problème du site (vu le exit code). Et aussi si le site est visible sur le BDII
- Pour Alice il est calculé dans le cadre de Monalisa
- **Important**: si l'état du site apparaît rouge ça ne signifie pas nécessairement que le site est le responsable. Même si la définition de état est le plus possible relié au site, il peut arriver certaines fois que le site apparaît rouge pour des problèmes de la VO. Même en ce cas la, il est utile que le site soit au courant du problème

Un aperçu du prototype pour FZK

- Le Gridmap principal ne montre que les activités principales, pour toutes les Vos servies par le site

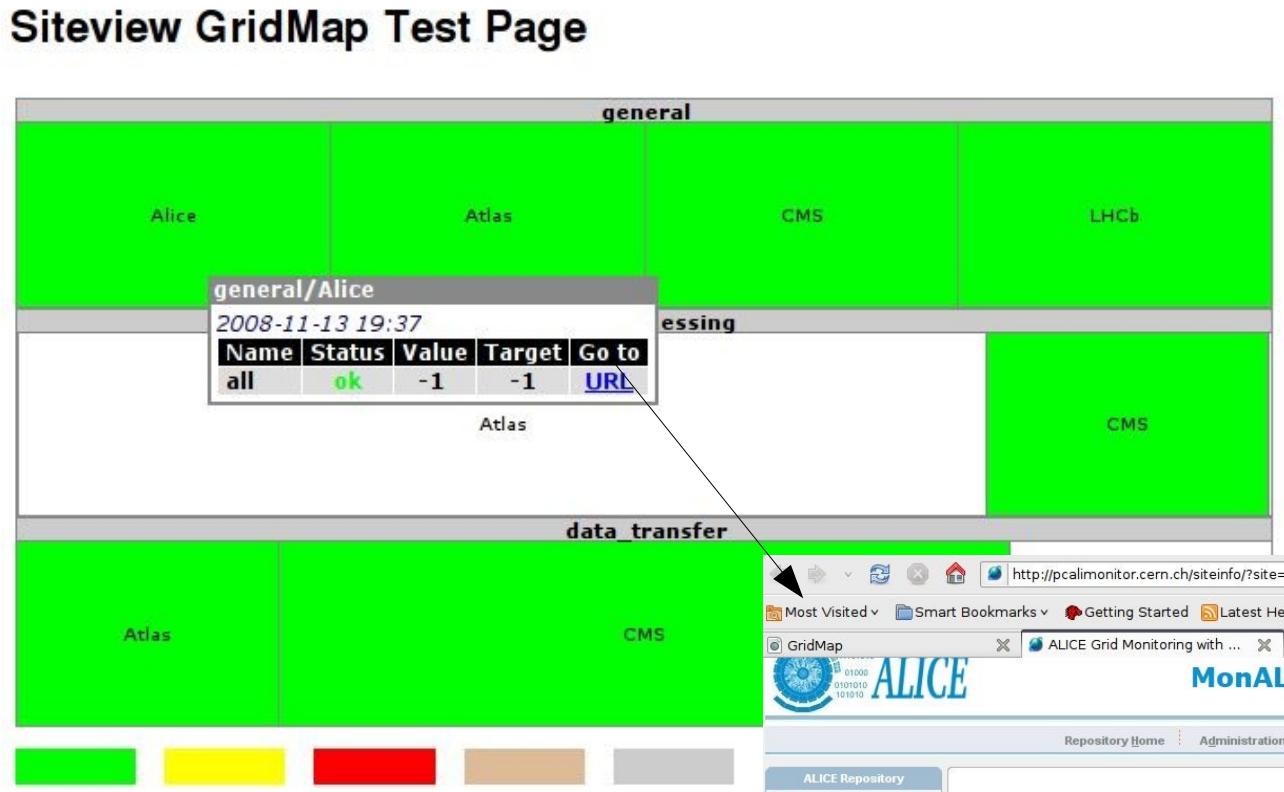


general

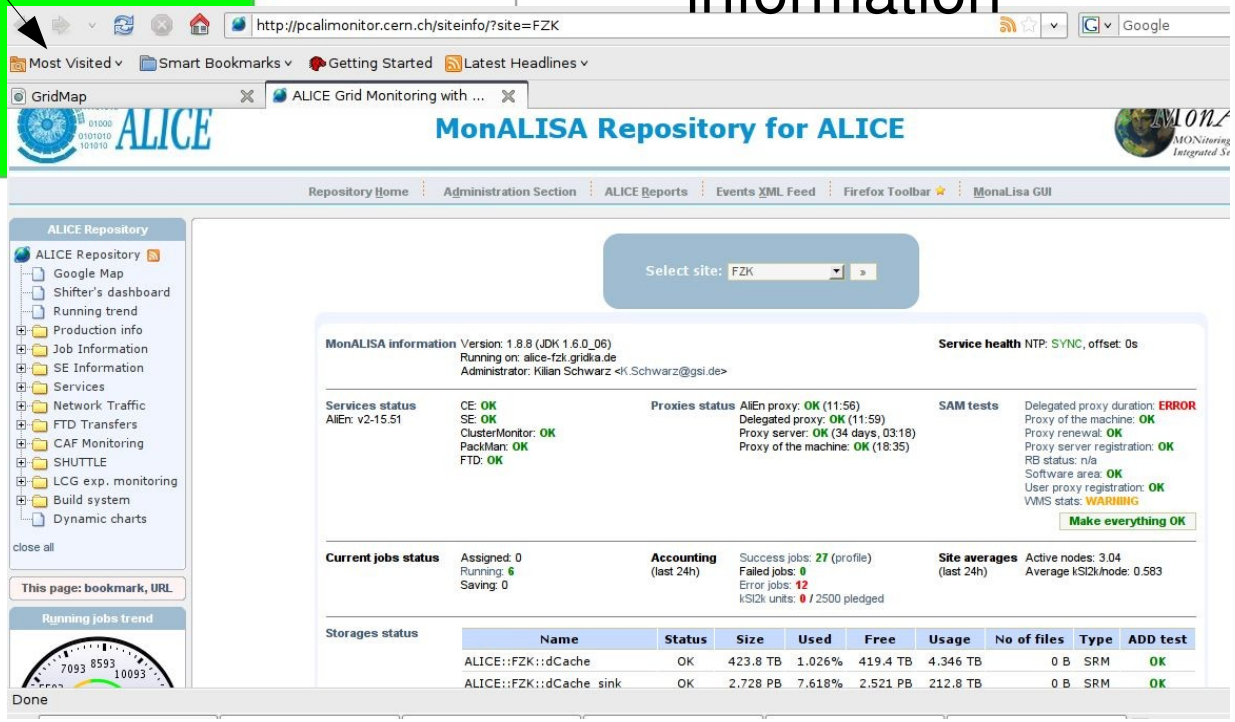
Name	Size	Status	Date
Alice	-1	ok	2008-11-13 19:37
Atlas	-1	ok	2008-11-13 19:37
CMS	-1	ok	2008-11-13 19:37
LHCb	-1	ok	2008-11-13 19:37

- En pointant la souris sur la barre d'en-tête il apparaît une fenêtre avec les noms de Vos, l'état et la date de la dernière mise a jour

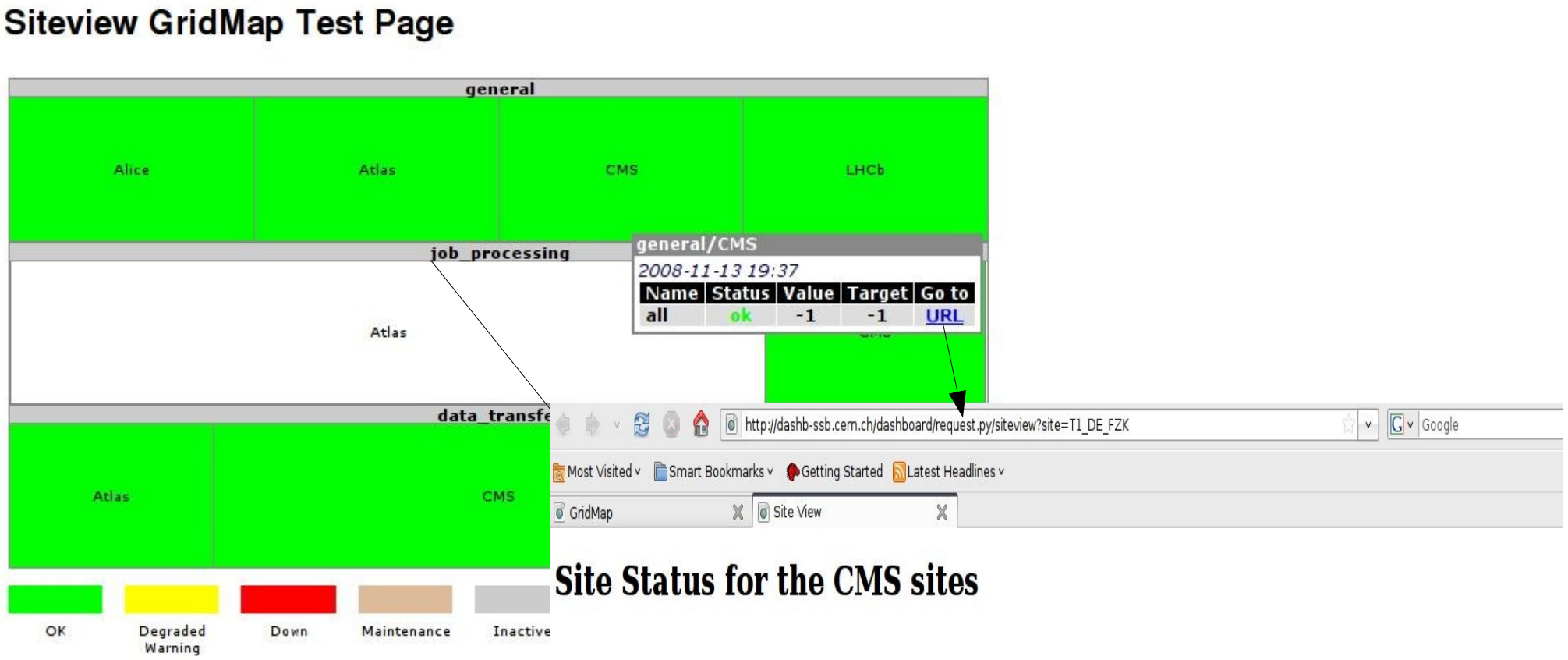
Lien a l'état du site pour Alice



Les infobulles fournissent un lien vers la source de l' information

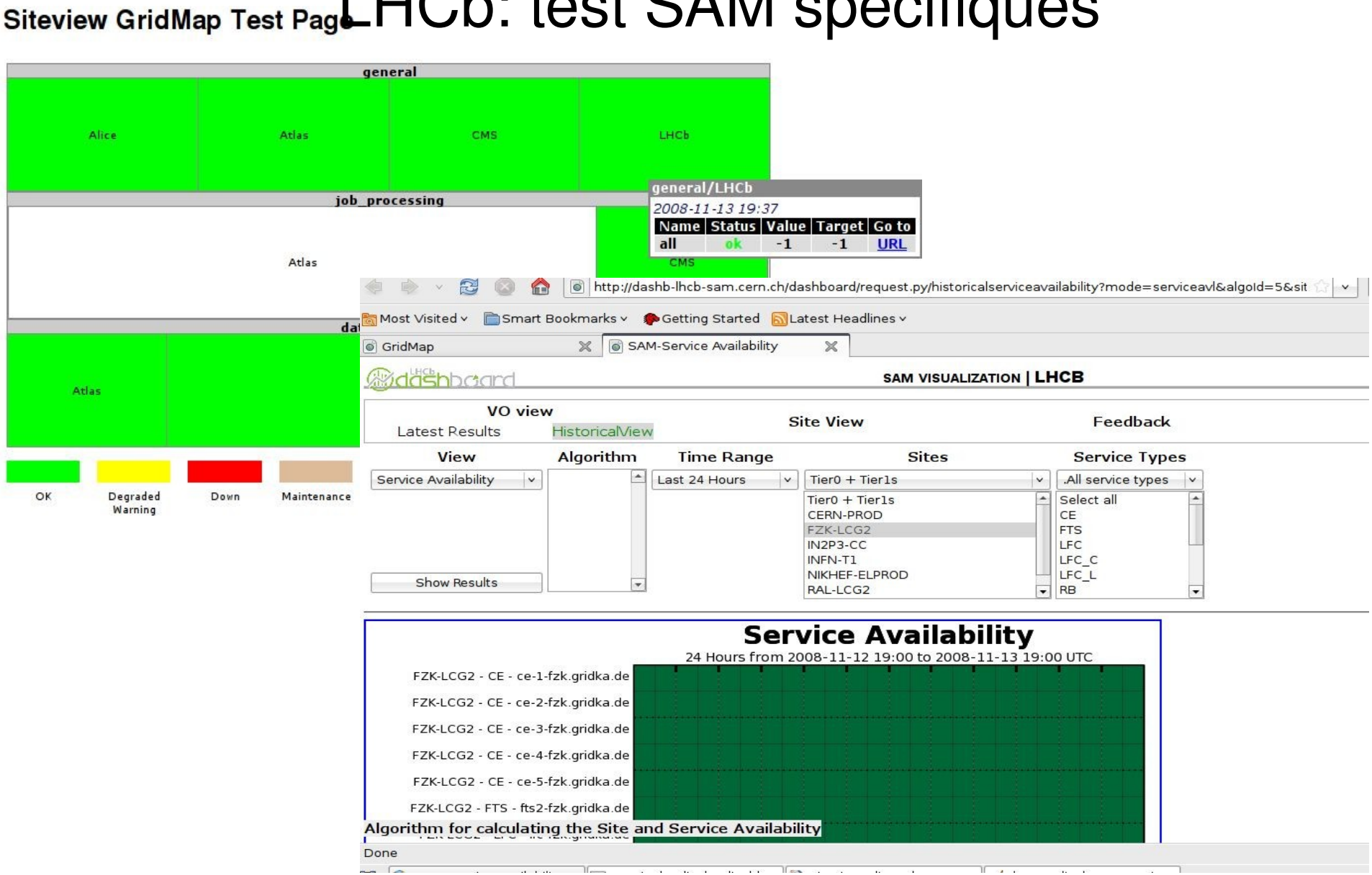


Lien pour l'état du site pour CMS: Site Status Board



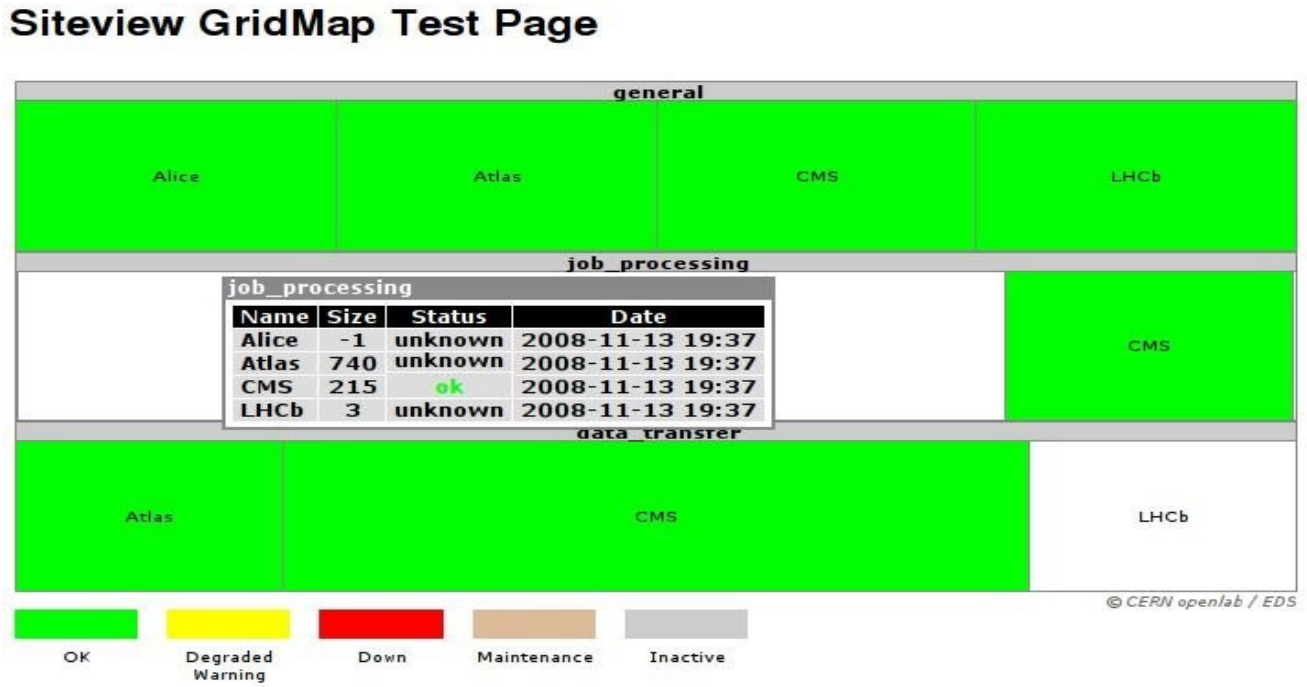
A partir d'ici on a tous les liens nécessaires pour continuer la recherche

Lien pour l'état du site pour les VOs Atlas et LHCb: test SAM spécifiques



Activité de job processing

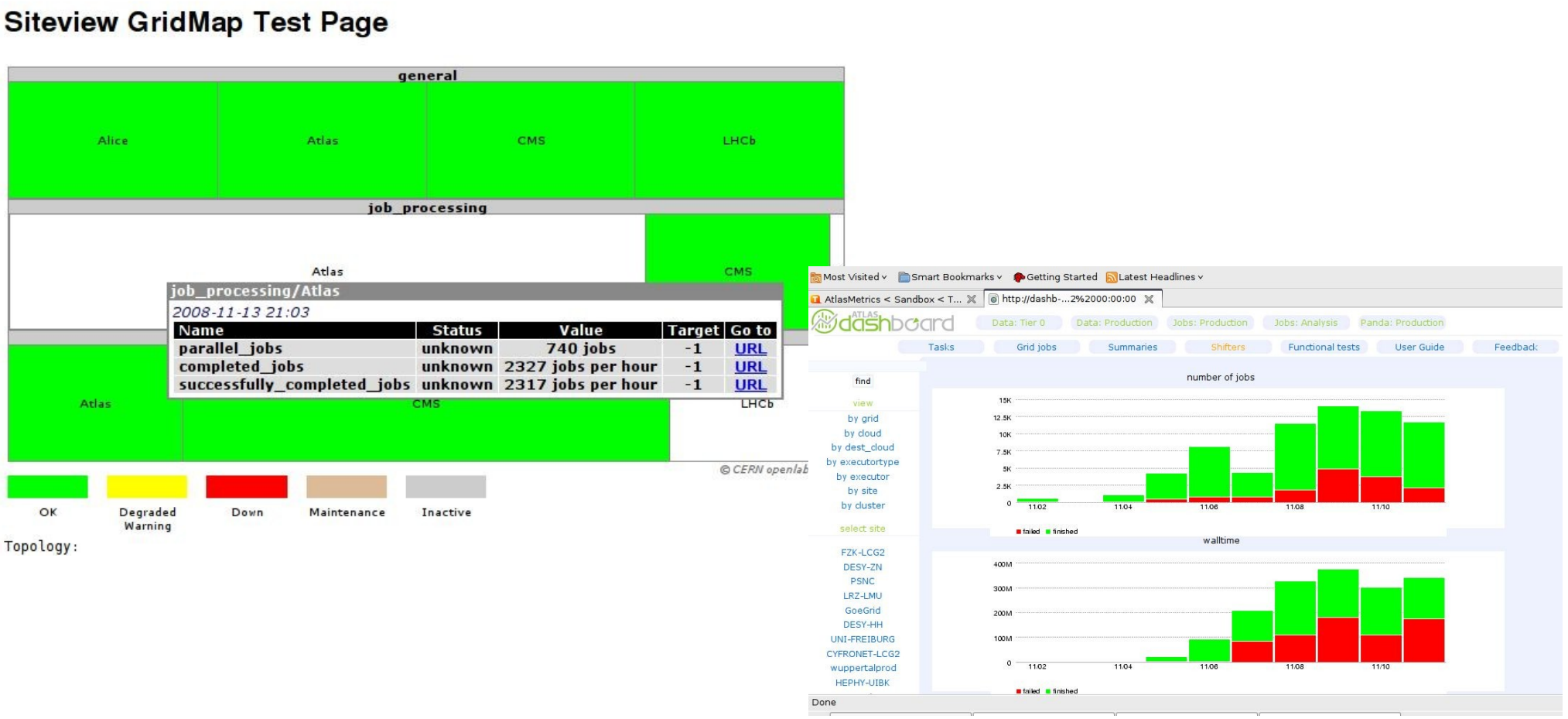
- En pointant la souris sur l'en-tête: information sur les Vos qui sont servies par le site.
- On montre le nombre de jobs en exécution (c'est le paramètre qui détermine les dimensions du rectangle), l'état (qui détermine la couleur), et la date de la dernière mise a jour



- **Important**: si l'état est rouge, il n'indique pas un problème du site. Par contre, il indique un problème relié a l'activité de la VO sur ce site

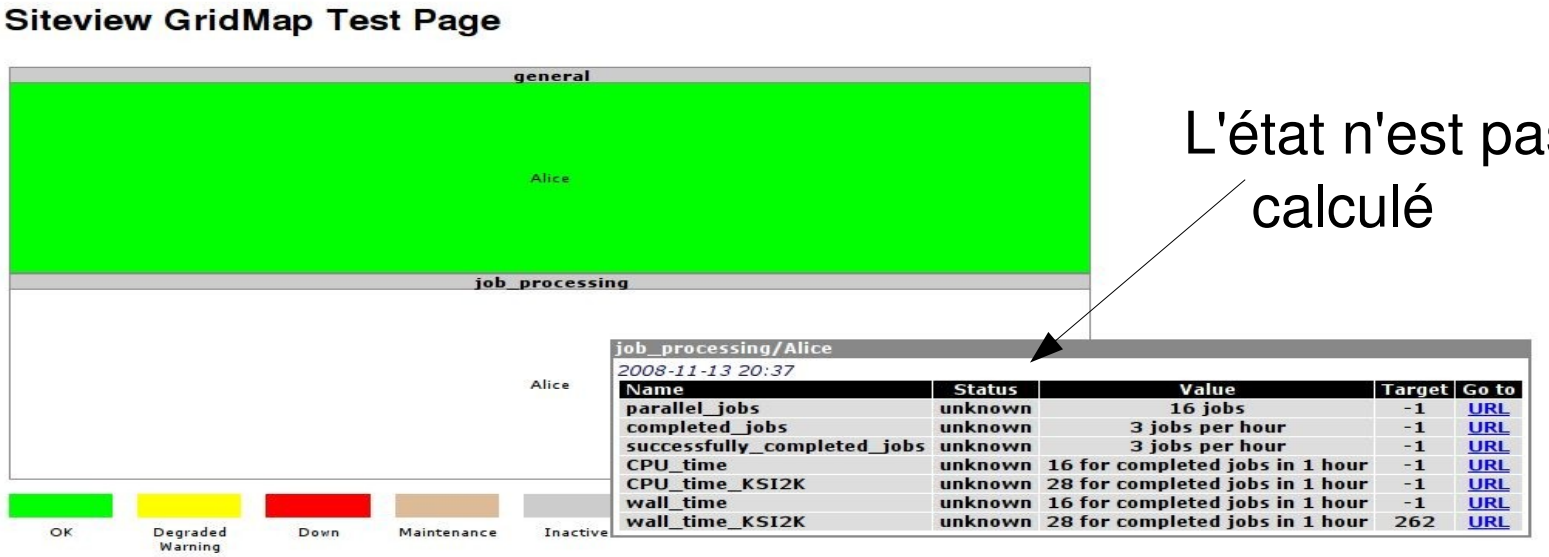
Job processing: Atlas

- Toutes les grandeurs relatives á l'activité de job processing (jusqu'à maintenant que MC production, mais bientôt aussi user analysis et autres types de taches) sont montrées
- Le lien affiché renvoie vers la page de Dashboard relative á la production MC au site

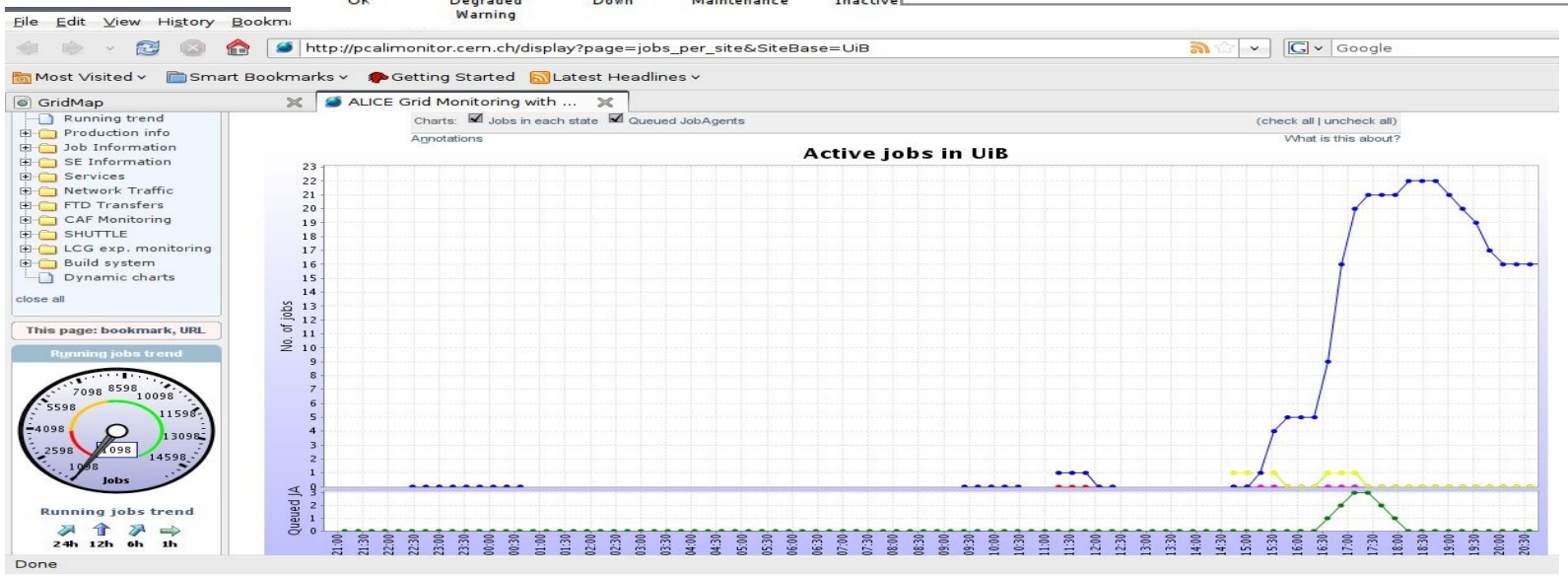


Job Processing: Alice

- Toutes les grandeurs relatives au job processing sont affichées
- Le wall time est affiché, **normalisé en ksi2k** avec aussi la valeur attendue



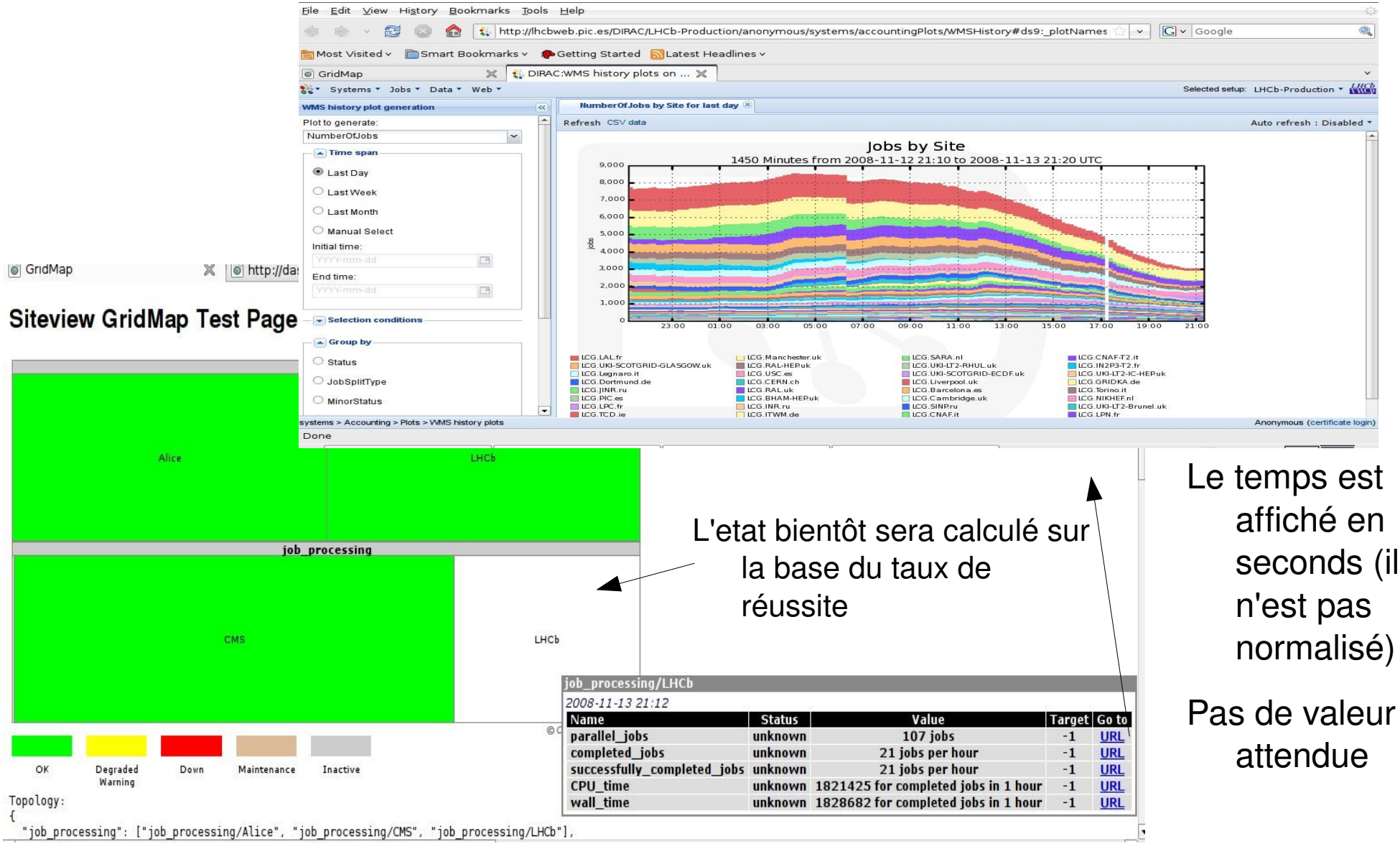
L'état n'est pas calculé



Le lien dans le infobulle renvoie a la URL de Monalisa pour ce site et cet activité

Job processing LHCb

- Job processing LHCb. Lien vers le history plot du Dirac WMS



L'etat bientôt sera calculé sur la base du taux de réussite

Le temps est affiché en seconds (il n'est pas normalisé)

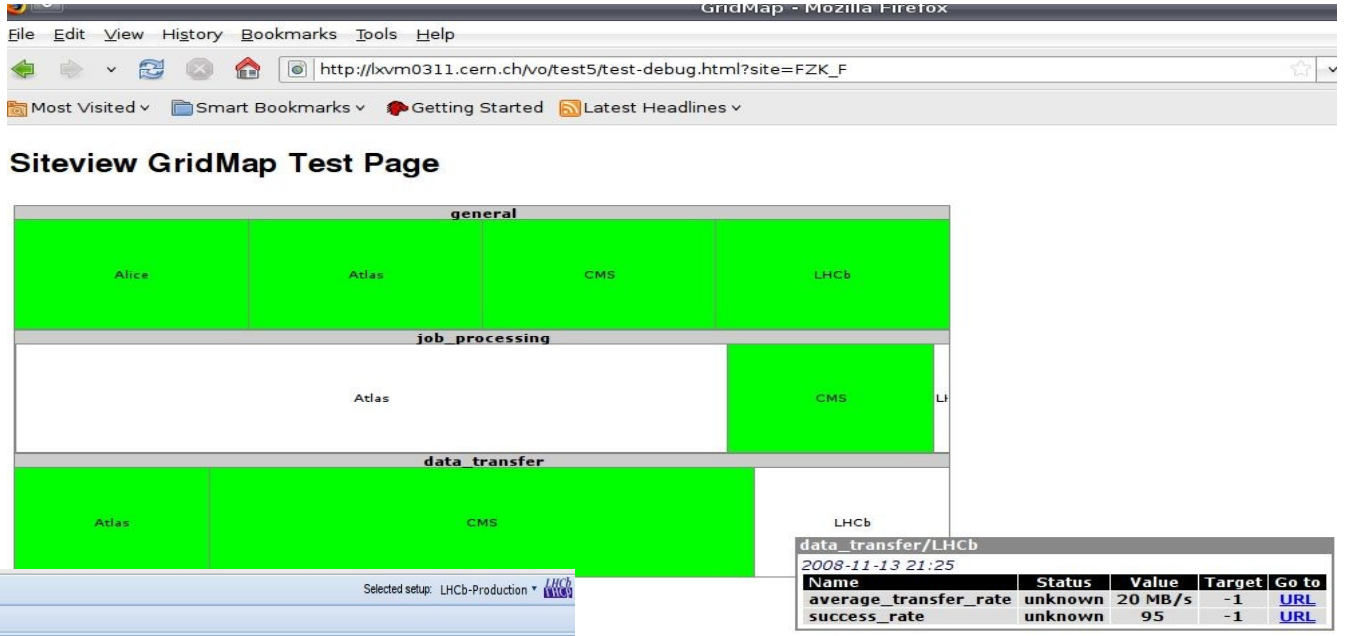
Pas de valeur attendue

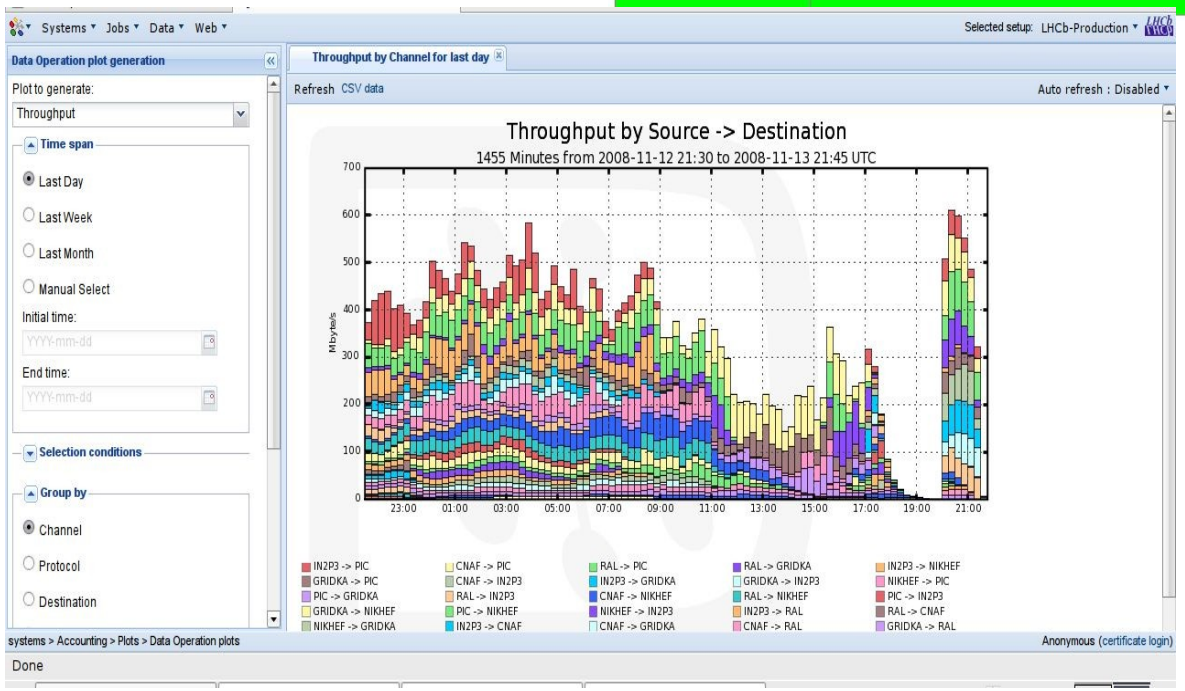
Le second niveau du Gridmap

- Par ex. activité de job processing de la VO CMS.
- En cliquant sur le plan, un second plan apparaît avec les activités secondaires. Par ex. pour CMS: reprocessing, MC production, user analysis, ...
- L'infobulle montre le nombre de jobs en exécution pour chaque activité secondaire

Transfert des données

- Dans la derniere version on montre le flux de données entrantes et sortantes separamment
- Le Gridmap de second niveau montre les differents types de transfert



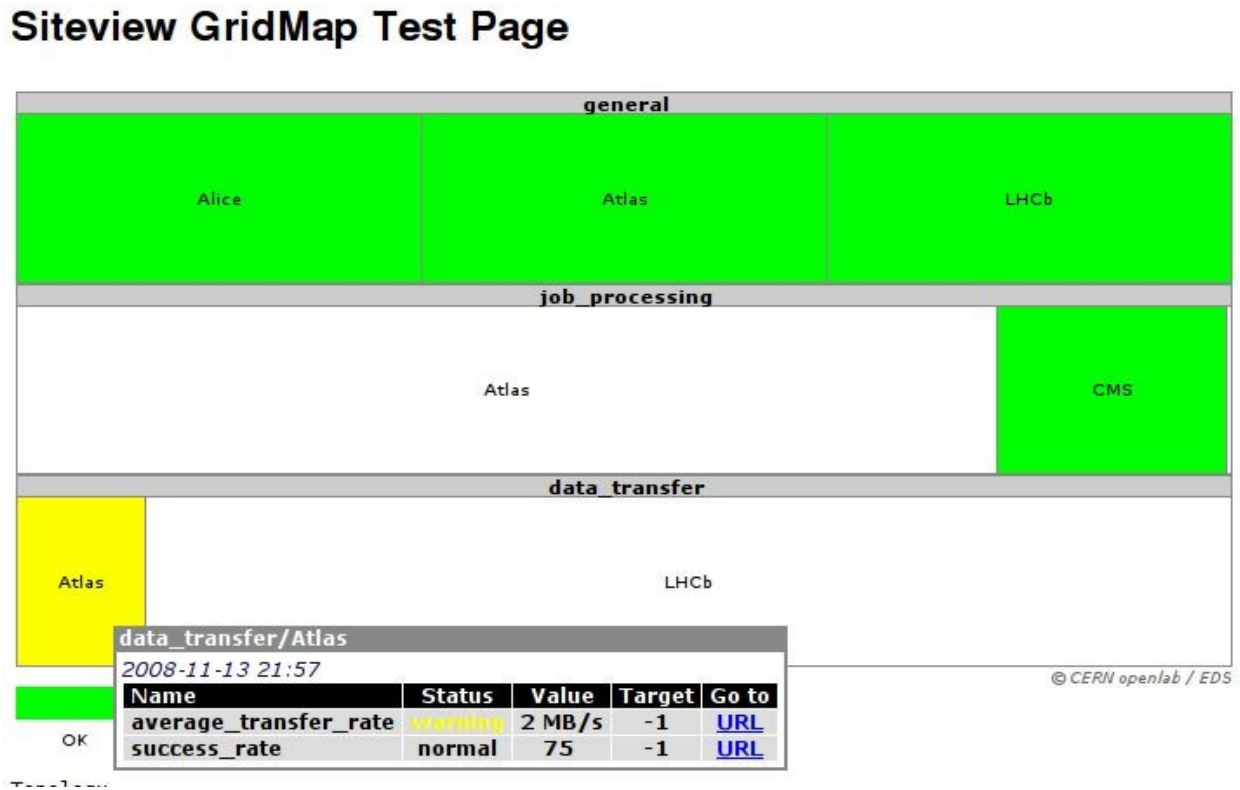


- en développement pendant cette semaine. Une vue plus recente sera montrée pendant la demo

Transfert des données pour Atlas

- **L'état et la valeur attendue sont affichés**
- Pour le transferts de T0 aux T1, l'état est calculé pour le taux de transfert moyen en comparant avec la valeur attendue: il peut être ok ou warning
- Pour le taux de réussite l'état est calculé suivant cette règle:

- SR<20% : critical
- 20<=SR<50%: warning
- 50<=SR<80%: normal
- SR>80% good



État actuel et plans pour l'avenir

- On a défini un groupe de activités et de grandeurs
- On a dessiné un schéma pour la base de données commune et on a vérifie sa validité pour nos objectifs
- On a développé un collecteur qui lit l'information publiée sur des URL et introduit les données dans le schéma commun
- Le Vos ont dejá fourni la plupart des grandeurs, mais quelques grandeurs sont encore manquantes
- Souvent les valeurs attendues et l'état ne sont pas fournis par les VOs

Partie facile du travail: l'outil a une structure très simple

pas très facile...

Difficile, mais réalisable

Plans: une version de test du Gridmap complet avant la fin de l'année

Questions ouvertes

- Souvent c'est difficile pour les VO de afficher les valeurs attendues pour les grandeurs et l'état pour les activités
- Souvent la valeur attendue pour une grandeur n'est pas disponible, ou bien elle est disponible mais elle se trouve dans un autre système ou base de données qui n'a aucune connexion avec l'outil de monitoring, donc la valeur observée ne peut pas être comparée avec la valeur attendue
- En général, le Vos sont très réticentes á déclarer un état pour son activité au site

Quelques observations á ce propos:

Si l'activité d'une VO sur un site apparaît rouge, cela ne signifie pas que le site est responsable, mais c'est plutôt la VO responsable du problème

L'état du site est montré dans le rectangle supérieure 'Site Status' : si cela est vert, alors le site est en bon état, même s'il n'y a aucun job en exécution

Cet outil n'est que un moyen informatif, qui a comme but de simplifier l'accès a l'information donnée par les outil de monitoring des Vos et d'aider les administrateurs des sites á dépister les problèmes de une façon simple et rapide

Remerciements

- Ce projet ne peut avancer que grâce a la collaboration des différentes Vos qui publient les données nécessaires pour peupler la base de données commune
- Merci a: P.Saiz (CMS), C.Grigoras (Alice), R.Rocha and B. Gaidioz (Atlas), A. Casajus and A.Tsaregorotsev (LHCb), W.Ollivier (LHCb and Atlas) pour sa contribution fondamentale

Feedback to: elisa.lanciotti@cern.ch