SiteDB: Marshalling people and resources available to CMS

Simon Metson, H.H. Wills Physics Laboratory, Bristol, UK**; Daniele Bonacorsi**, *University of Bologna / INFN, Italy;*
Ricky Egeland University of Minnesota, Twin Cities, USA; Marco Dias, *SPRACE, Brazil*

What is SiteDB?

SiteDB is a catalogue of all sites used by CMS. It records the CMS resources at the site, as well as pledged resources for the future. It is also used to keep track of CMS personnel at each site, and the roles and duties they fulfil to the collaboration. Collaborators are automatically added to the database on registering with the CMS Hypernews system and can be assigned roles for a site by any other site contact.

As well as providing resource pledge management SiteDB acts as a simple monitoring portal, bringing in and providing links to content from PhEDEx, DBS, SAM tests, CMS Dashboard, GOC and GSTAT. The database also maintains a mapping between the different names a site can be known by, for instance the institute name (University of Bristol), it's SAM/WLCG name (UKI-SOUTHGRID-BRIS-HEP), and the sites CMS name (T2_UK_SGrid_Bristol), and associations between sites (e.g. the parent T1 site for a T2).

Why build SiteDB?

CMS relies on close contact to the sites it uses. The distributed nature of CMS computing makes it very difficult to track who is responsible for what over time. Some site roles are technical (for instance running PhEDEx, administering FTS) others are related to CMS computing policy (e.g. the site's Data Manager) or site managerial roles (Site Executive).

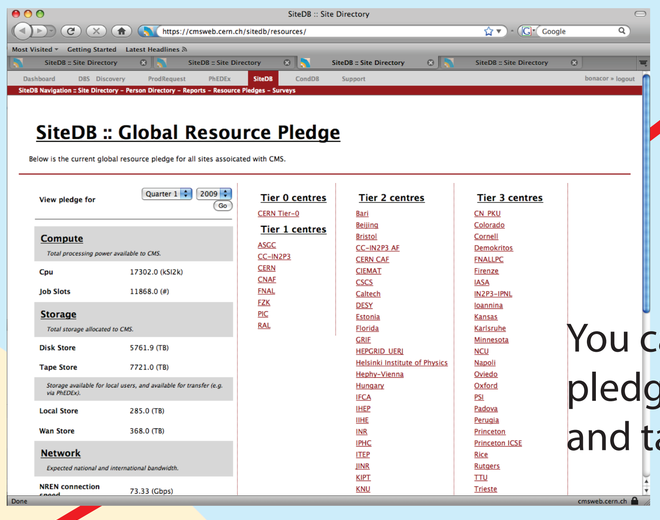
SiteDB allows us to track who has what role at a site and enables us to contact them easily. For users who may not know what the various roles and responsibilities are at a site, SiteDB provides a simple link into the CMS Savannah portal. This has a group per site, dynamically maintained from the information in SiteDB and allows us to easily track problem resolution across our distributed infrastructure.

SiteDB architecture

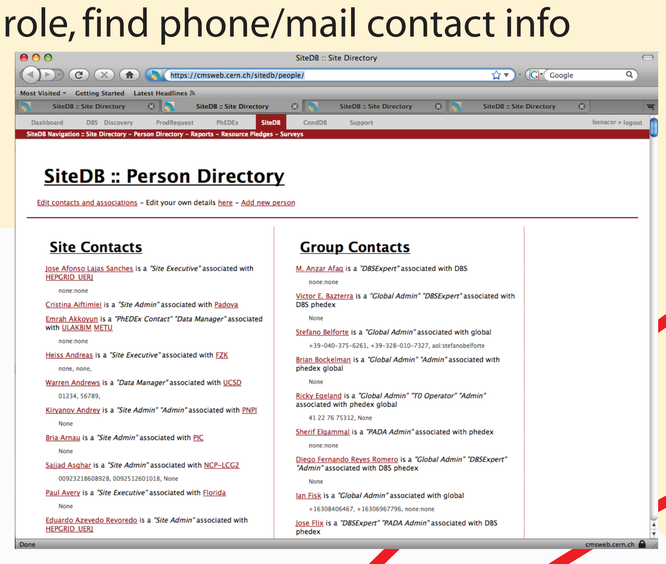
SiteDB uses an Oracle database as its backend, managed by CERN/IT. This is accessed through a web interface that provides both a GUI and an API. This is built using the CMS DMWM web tools framework and run on voboxes by the CMS Facilities Operations group at CERN. The API returns information in both xml and json format, and restricts access to "sensitive" information for CMS collaborators only. The API's are used by other CMS applications (both on the web and off) to retrieve information about CMS sites (for instance their pledged CPU) and collaborators (for instance the user name associated to a given DN).

Where next?

Future developments of SiteDB will initially consolidate the code used onto the new CMS WMCore code base (of which the web tools is now a part). Next will be improved resource monitoring and tracking of resources, to further aid in both short and long term planning of computing operations. Ideally this should hook into an improved monitoring portal, such that we can easily see if sites are delivering to CMS what they have pledged.

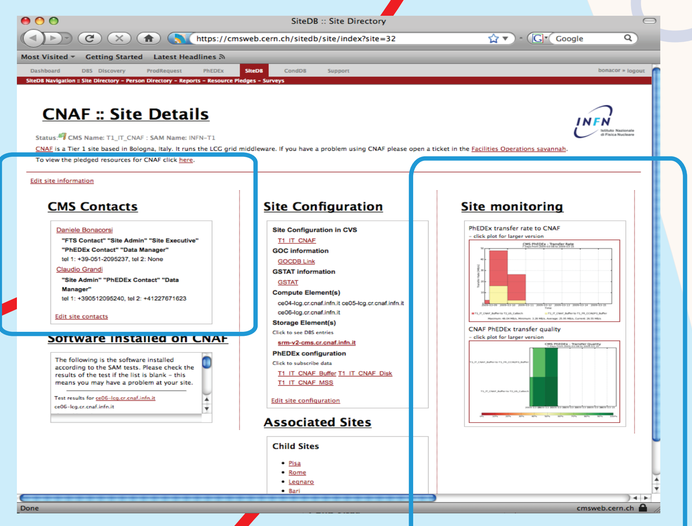


You can check the pledges for CPU, disk and tapes

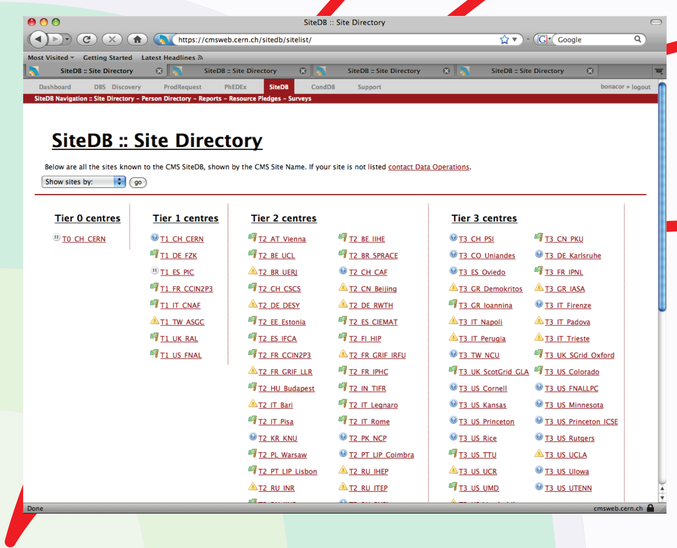


You can list all contacts, query for a given role, find phone/mail contact info

Selecting a site, you can see the contacts and their "roles"...



...and access monitoring information for the site



When you log-in the main SiteDB page displays the list of sites supporting the CMS VO, and their "Tier" responsibility