



LCG Service Challenges – Tier 2 Planning

Introduction

Whereas LCG Service Challenges 1 and 2 (SC1/SC2) involve only the Tier0 (CERN) site and selected Tier1s, SC3 – due to start in July 2005 – includes a small number of Tier2 sites. SC4 – due to complete at least 6 months prior to LHC data taking – should include all Tier2 sites – possibly as many as 100. Ramping up the required services at such a large number of sites will need significant coordination, which can best be done using existing structures, such as national or regional projects or bodies – e.g. INFN in Italy, GridPP in the UK, or meetings such as HEPiX.

The role of the Tier2 sites differs across the four main LHC experiments, but typically these sites concentrate on the production of simulated data, producing calibration constants and playing an active role in ‘chaotic’ end-user analysis¹. They may also cache small subsets of other data types – including RAW and ESD – typically for the purposes of code development and verification. Although the Tier2s are expected to contribute significant amounts of CPU capacity, they do not provide Mass Storage Systems: indeed, their entire data sample must be considered non-permanent.

A discussion of Tier2 sites cannot be complete without considering the services offered to them by Tier1s, which include the safe-keeping of simulated data produced at the Tier2s as well as the distribution of data required for analysis or other purposes.

In contrast to the distribution of data from Tier0 to Tier1 sites, where there is the requirement to keep up with data taking, the data flows to and from Tier2 sites can be somewhat more relaxed. For example, provided that sufficient storage is available at the Tier2 site, uploading of data to a Tier1 site can afford to be delayed for a period of possibly days – in the worst case (other than massive disk crashes!), Monte Carlo production would have to stall until connectivity to a suitable Tier1 can be re-established.

Although there is no fixed dependency between a given Tier2 site and a Tier1, and indeed there is the need for any Tier2 site to be able to request and / or access data residing at any Tier1 or even the Tier0, practical considerations, including network topology and associated cost, would indicate some degree of ‘soft coupling’. That is, whilst a given Tier2 may indeed request data from any Tier1, the data may flow via a third-party, rather than directly from a ‘remote’ – in terms of network connectivity – partner. In other words, a ‘logical mesh’ does not imply a physical one.

In addition to disk pool management software – the subject of a future document – a Tier2 site would also have to provide reliable file-transfer services, presumably based on the same tools as those envisaged to be used between CERN and the Tier1 sites.

For the remainder of this document, we focus on the steps required to establish such services, with a view to participation in SC3 and beyond.

¹ LHCb is different in this respect as its Tier2 sites focus on Monte Carlo production and do not take part in analysis.



Relevant Tier2 Services

Tier2 sites are expected to run some form of disk pool manager that offers a conformant² SRM interface. They should also run a reliable file transfer service conformant with the then-current LCG standard. The management of the local disk pool(s) and local end of the file transfer services are the responsibility of the Tier2 site. Clearly, appropriate network connections must also be provisioned, typically of the order of 1Gbit/s per supported experiment³. We note that certain analysis models can generate much higher network traffic.

These responsibilities will be described in more detail in the LCG Memorandum of Understanding (MoU) and in the Computing Models of the experiments that a given Tier2 will support.

Joining the Service Challenges

In order to join the LCG Service Challenges, Tier2 sites need first to install and configure the disk pool and file transfer management software and appropriate services. Basic file transfers can then be exercised between sites to verify the installation, performance and reliability. Further tests need then to be scheduled as part of the verification of the corresponding experiments' Computing Model, as instantiated by their offline software.

² The exact SRM version will be defined as part of the milestones of the corresponding service challenge.

³ The network bandwidth is typically dominated by the incoming traffic, e.g. replenishment of the local copy of analysis and/or other data. The reverse is true of ALICE, due to the large simulated RAW event size (300MB).