1.1 CMS (COMPACT MUON SOLENOID) APPLICATION

A new revolution is pending in high-energy physics. With the launch of the LHC accelerator at CERN in 2007, there will be a flood of new data. No single computing centre can aspire to process it all. It is important to a scientists and researchers of the Baltic States to partake in reaping the fruits of this cornucopia of data, in order to stay on the front-line. High-energy physics is fundamental of our understanding of the physical world.

There are several hints of new physics beyond the Standard Model. The discovery of the neutrino mass is a clear indication that the Standard Model is not a complete theory of particle physics. In the Standard Model, neutrinos are massless; we need new physics to provide the masses and the data from the LHC machine to help determine this physics experimentally.

Also, the LHC is needed to discover the scalar Higgs doublet, the particle whose vacuum expectation value obtained in electroweak symmetry breaking gives masses to weak vector bosons and quarks.

In addition, supersymmetric particles such as gluino or wino, for example, could be discovered in the LHC. Supersymmetry is an elegant symmetry relating fermions to bosons needed for gauge coupling unification and string theory.

1.1.1 Simulation of CMS events

As CMS is one of the four large experiments at the LHC, then it is quite well suited to the Grid as the Grid has been originally designed and deployed for exactly such applications. As the collisions happening in experiment are continuously happening and are independent of each other, then every single collision can be studied separately. The amount of collisions per year is estimated to be around 10^7, which in turn means that the tasks, which CMS physicists face, are highly parallelizable.

The requirement of a single job is to have access to 1 GB of memory and access to stored datasets. An average analyzable chunk size is ca 500-2000 events, which constitutes to approximately 2 GB data files depending on the contents. Minimum bias (collisions without producing new particles) simulated datasets are smaller and signal datasets as well as final detector response will be bigger. The total amount to be processed and analyzed on the Grid per year should be in the order of 5 PB.

The huge amount of computing power needed for the scientific task comes from the probabilistic nature of the quantum world. Using Monte Carlo techniques it is needed to simulate one physically interesting processes several million times to get information of the process. The whole cycle of Monte Carlo simulation consists of : generation of collision event using known laws of physics and presumable new physics; simulating the propagation of millions of particles, that come into being after the collision, through the complex body of detector taking account of all interactions between particles and material of the detector; finding the detector response and the digital signal that can be read out in the case of working process of the real CMS detector and reconstructing the information back to the particles which were traveling through it; finally, interpreting all reconstructed particles and their specific signatures through complex analysis. Of course in the case of real working LHC experiment the nature makes the first three steps of described tasks, but it is crucial first to simulate the detector behavior to know how the detector works and to be able to interpret the complex signatures of particle to find whether new physics exists or not.

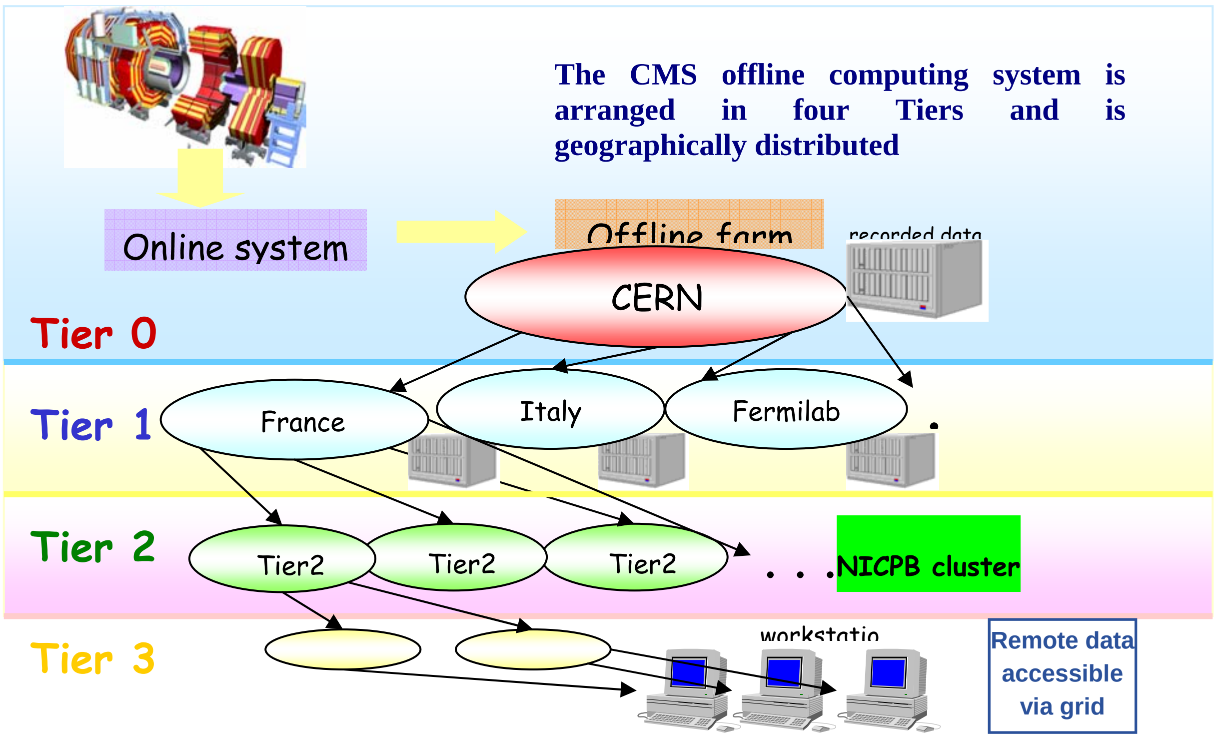

Fig. 1. The architecture of data distribution within CMS experiment simlation

1.1.2 The structure of CMS software

In addition to the requirement that CMS software (CMSSW) should be preinstalled in a computing element, the respective Tier 2 centers which contribute resources to CMS has also to run additional services like PhEDEx (CMS data management and transfer tool), Frontier (calibration data cache) etc. These tools are only needed at participating sites and ordinary simpler jobs like monte carlo production can be run on any Grid resource which has CMS software installed.

The application itself (CMSSW) is distributed as a versioned package. CMS uses remote Grid installations to unify the installed software base across all sites. The first installation initiates a separate RPM (RedHat Package Manager) package repository and uses the apt package management software to install the required software and all its dependencies. The multitude of dependencies will not be described here, but their separate installation is not required as CMS keeps track on what versions work together and always distributed to the sites versions, which are interoperable. Starting from CMSSW, which was released during summer 2006, CMS uses a single software release and unified execution methodology to perform all of the possible tasks (MC production, detector simulation, reconstruction as well as analysis).

To enable a cluster to support CMS VO requires from the cluster only that CMS VO is granted access with enough disk space in the VO specific software area. A single release is ca 2-4GB and usually at least 4-5 releases are maintained at sites although with overlapping software. The actual requirement is in the order of 10 GB per computing element. To support purely the Monte Carlo production jobs the worker nodes must comply with minimum requirement of 1GB of memory per job and outbound access to allow CMS jobs to send monitoring information during job execution as well as stage out produced files to a central location. SRM client software is also needed to be installed at the worker nodes. For software installations rpm build operating system utility is required as some packages are rebuilt on the worker node to best match to the running environment.

1.1.3 Implementation of CMS jobs

For Grid users the implementation of CMS jobs is easy through a special application called CRAB. Users specify their own analysis code and configuration files as well as which datasets they want to run on and how they would like the dataset to be split (number of events per job etc). CRAB then performs the resource location through available datasets as well as clusters supporting all of the requirements, splits the jobs according to user specifications and submits them. In the first year of BalticGrid project there was a steady transformation of the architecture of CRAB software, from the old form to a very recent one, as it is shown in fig. 2 below.

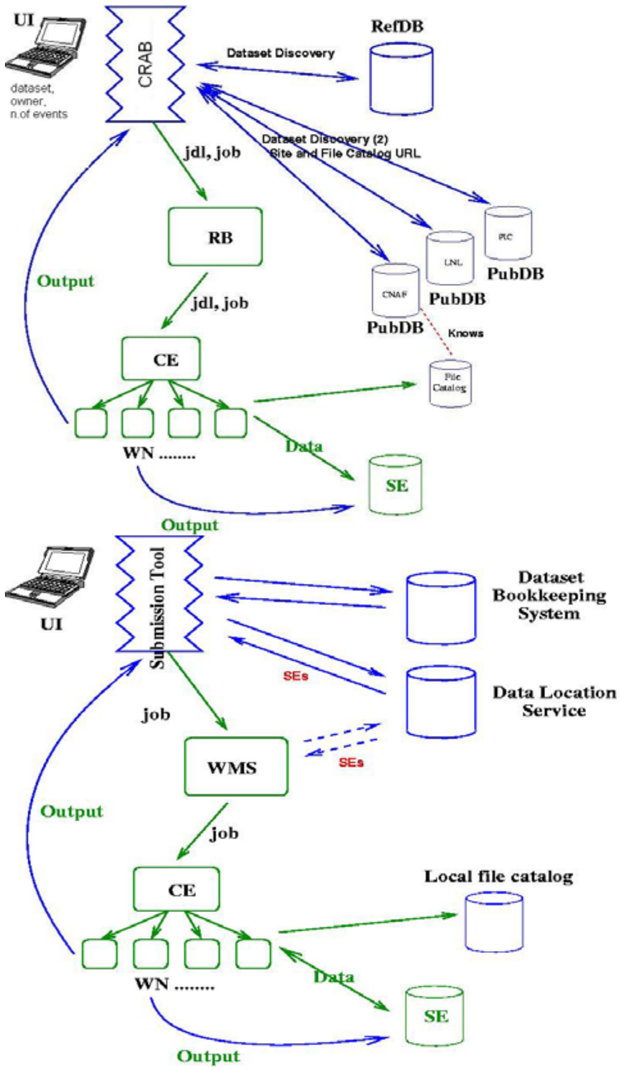

Fig. 2. The architecture of CRAB at the start of BG (left graph) and in recent days (right graph).

CRAB is a user-friendly tool whose aim is to simplify the work of users with no knowledge of grid infrastructure to create, submit and manage job analysis into grid environments:

- written in python and installed on UI (grid user access point)

Users have to develop their analysis code in a interactive environment and decide which data to analyse. They have to provide to CRAB:

- Dataset name, number of events
- Analysis code and parameter card
- Output files and handling policy

CRAB handles data discovery, resources availability, job creation and submission, status monitoring and output retrieval.

The actual results are also managed by the same application which brings the resulting output files to a central location allowing the user to just specify the configuration and then forget about the whole Grid integration part. For bigger Monte Carlo production activities as well as data reconstruction or re-reconstruction, they are performed centrally by a limited number of people who are competent and have special tools for such tasks.

During the first year of BalticGrid project CMS has been running a lot of Monte Carlo production as well as some analysis jobs on BalticGrid. During this period there was also bigger service maintenance when migration from old software to new software was performed and in which time CMS didn't run any significant number of jobs on Grid anywhere. Starting from June/July CMS started the analysis and data production challenge "Service Challenge 4 (SC4)" in which one of the BalticGrid partners (NICPB) took part. The challenge lasted until end of September. During that challenge almost 17 000 analysis jobs were run in T2_Estonia (NICPB) as can be seen in fig. 3:

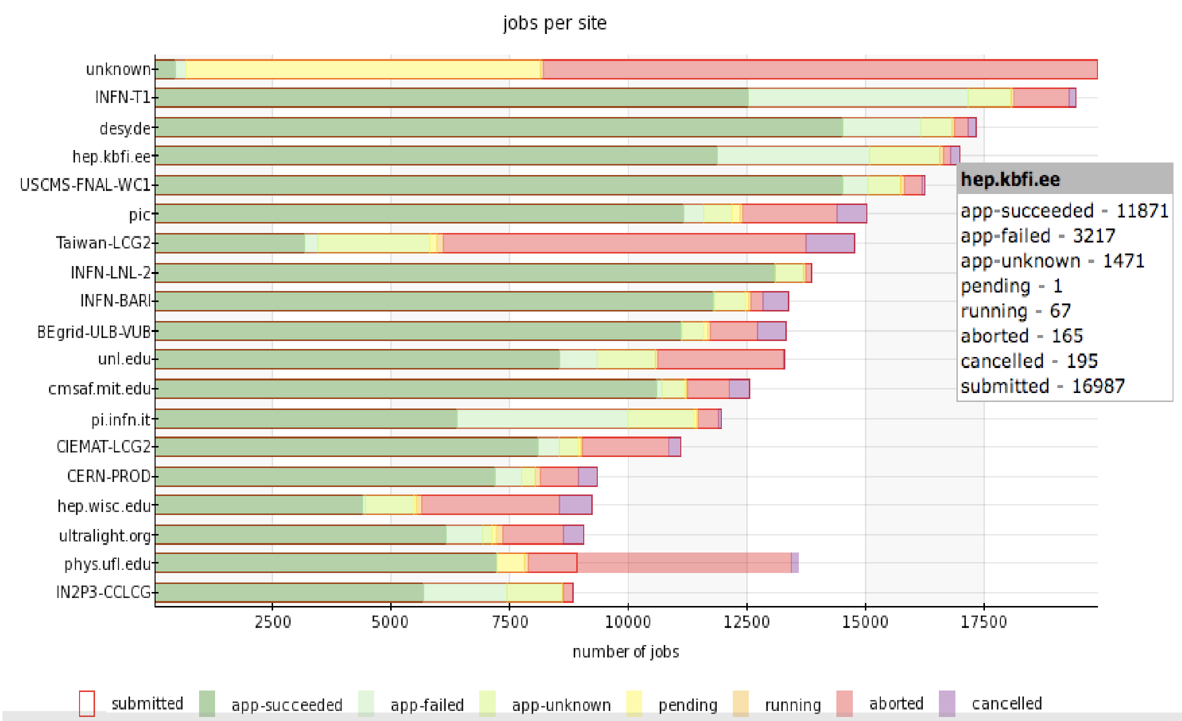

Fig. 3. Analysis' jobs of CMS on T2_Estonia cluster

In addition CMS produced 66Mio events out of which a fraction was produced in Estonia. Also some users of CMS VO were running their jobs on other clusters, which are tested in accordance to EGEE rules (clusters in Lithuania and Latvia). During SC4 also the transfer quality and possibility was tested to verify CMS computing model as well as individual computing centers and their connectivity. In that time T2_Estonia transferred to it’s location 160TB of data and exported a bit more than 90 TB of data as can be seen on the following figures:

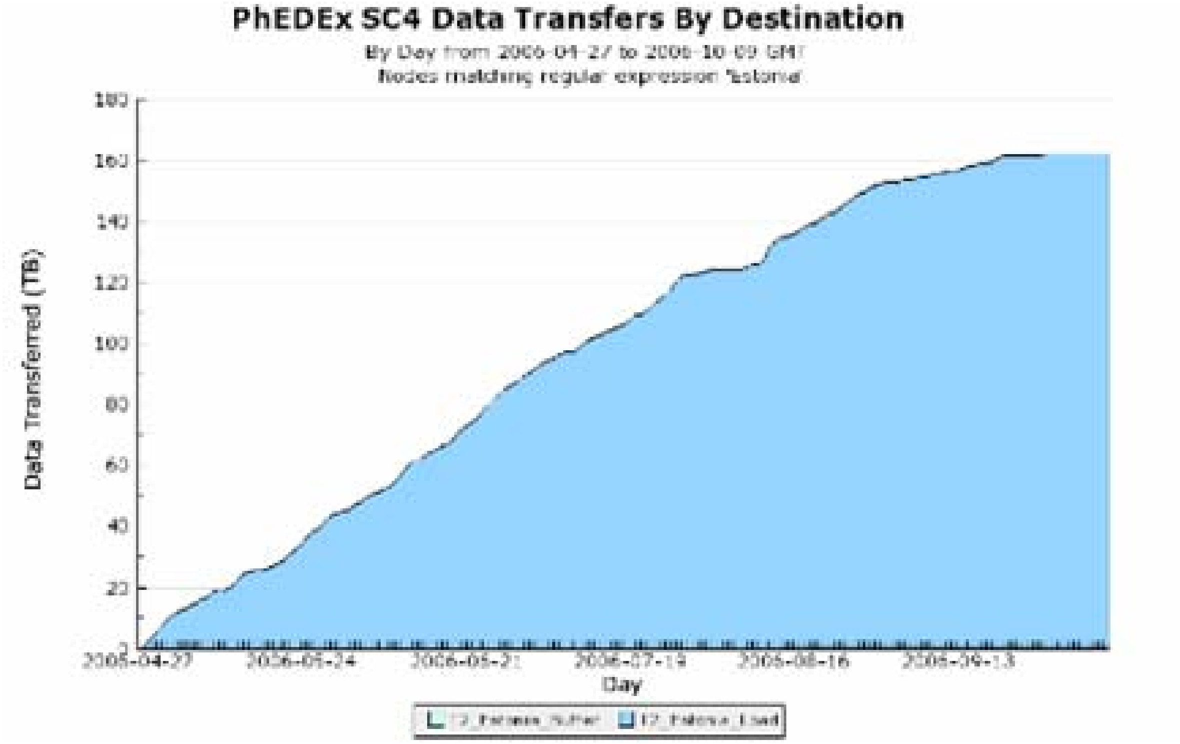

Fig. 4. PhEDEx data transfers by destination

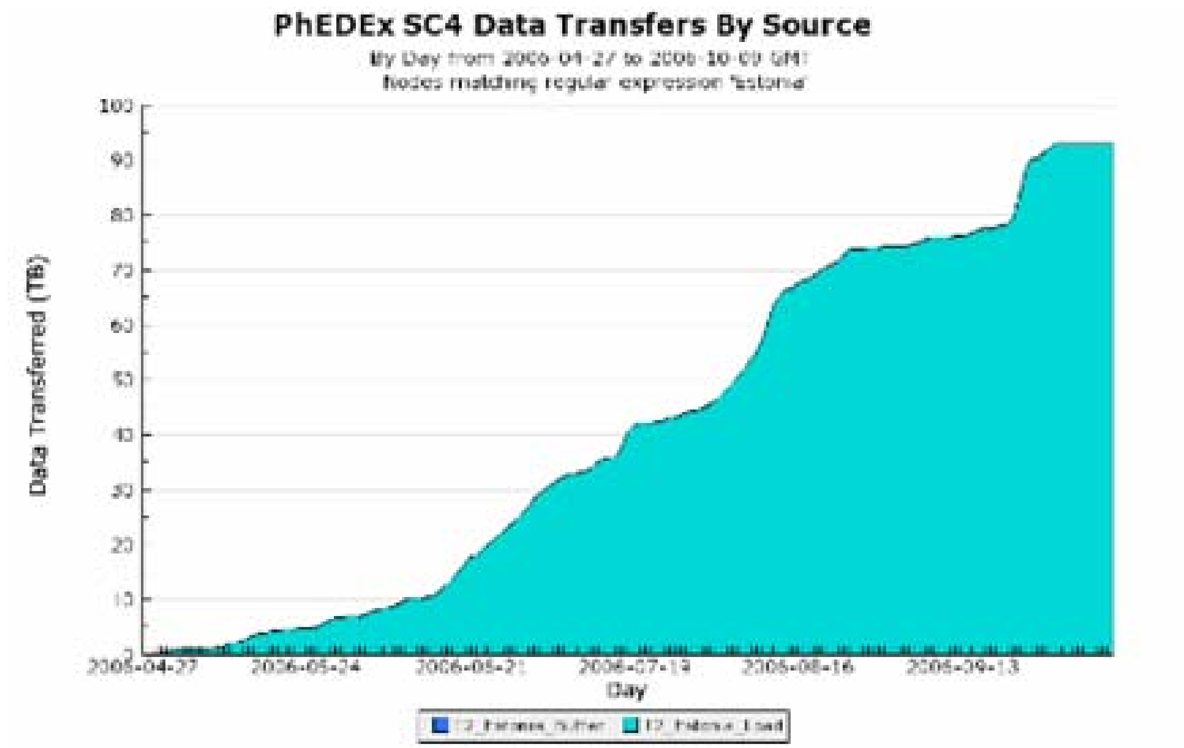

Fig. 5. PhEDEx data transfers by source

The average transfer rate of 20MB/s which was the target for SC4 was kept for majority of the challenge with some slower periods and with peaks to almost 50MB/s

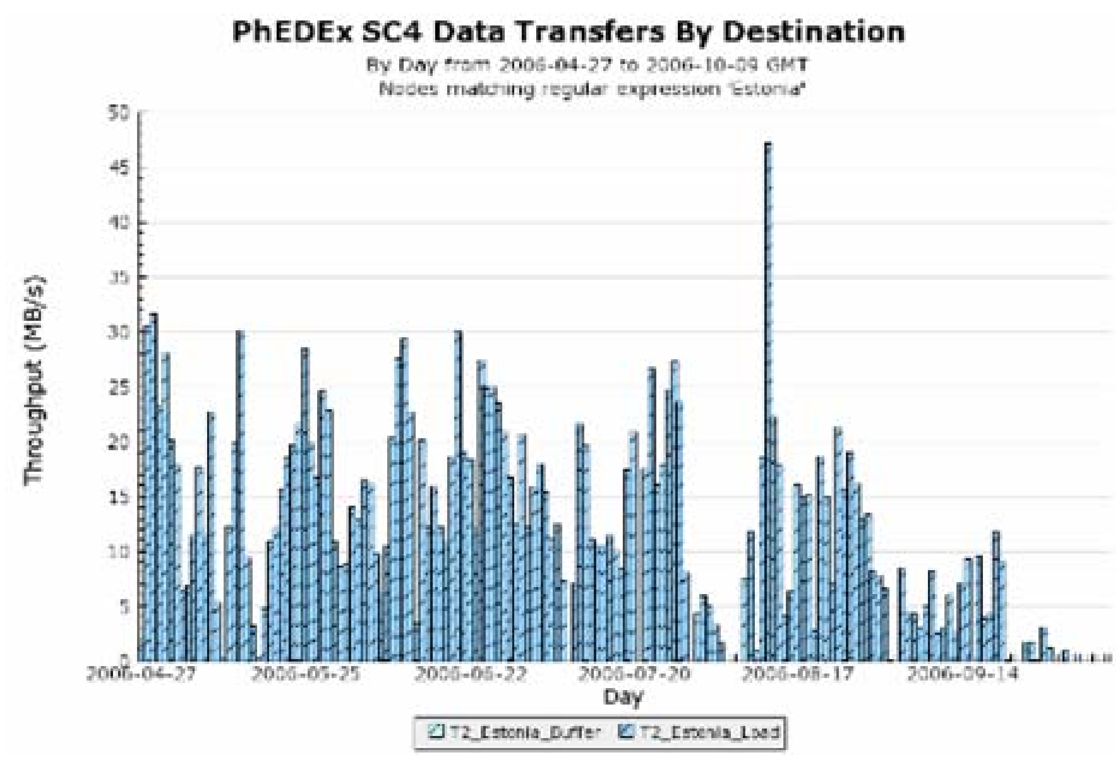

Fig. 6. PhEDEx SC4 data transfers by destination

In October 2006 CMS started the next challenge CSA06 (Computing Software and Analysis 2006), which is to demonstrate 25% capacity running of the whole computing model. This includes data reconstruction and processing at a constant pace at CERN, exporting of data to Tier 1 centers and from there to Tier 2 centers. Also analysis of re-construction results is to be performed. To the date of writing of this document only the data export and transfer part of the test has been in running and has been running as expected. Tier 0 has been performing over the estimates (100% uptime) and Tier 1 and Tier 2 centers are capable of keeping the pace of incoming data. There have been tests (both intentional and unintentional) of site downtimes and elimination of backlog of transfers. The average rate NICPB has been able to keep over the whole 3 weeks of CSA06 has been in the excess of 20MB/s including downtimes and has been able to keep over 40MB/s the past few days:

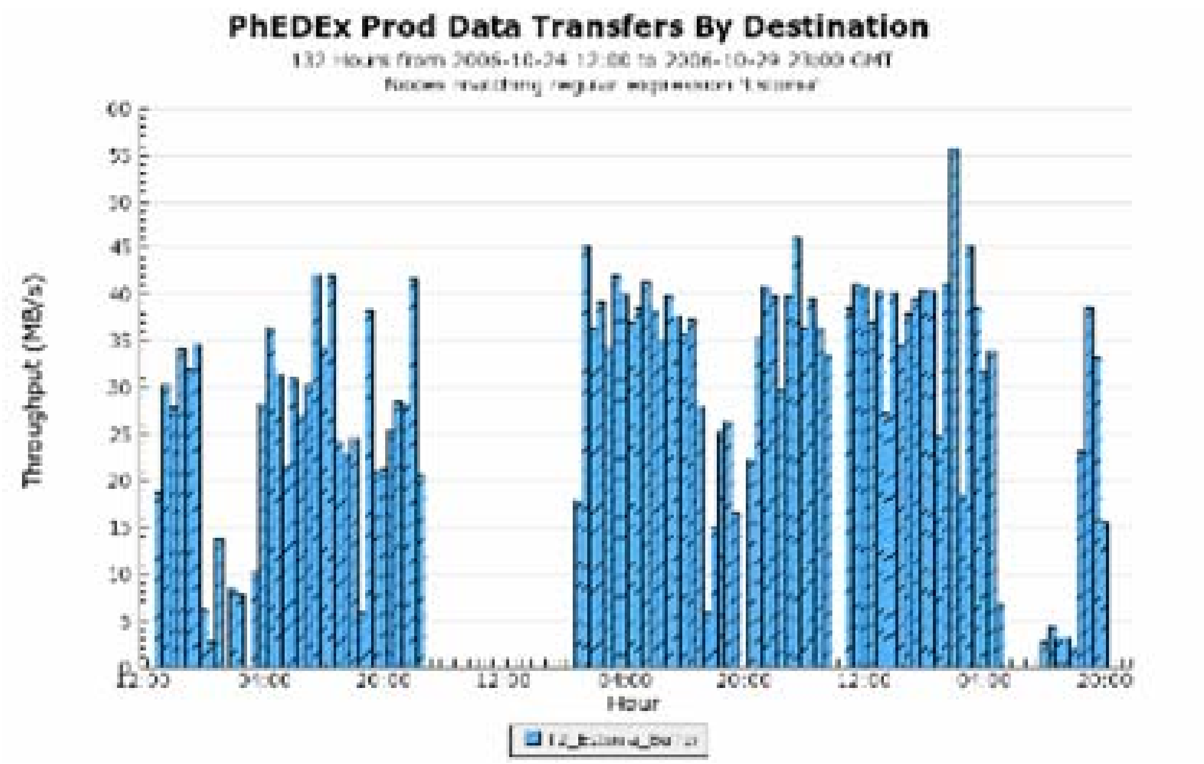

Fig. 7. PhEDEx production data transfers by destination