CMS Data Transfer operations after the first years of LHC collisions

R Kaselis¹, S Piperov², N Magini³, J Flix⁴,⁵, O Gutsche⁶, P Kreuzer⁷, M Yang⁸, S Liu⁹, N Ratnikova⁹, A Sartirana¹⁰, D Bonacorsi¹¹, J Letts¹²

¹ Vilnius University (Lithuania), ² Institute for Nuclear research and Nuclear Energy, Bulgarian Academy of Sciences (Bulgaria), ³ CERN (Switzerland), ⁴ Port d’Informació Científica (Spain), ⁵ Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (Spain), ⁶ Fermi National Accelerator Laboratory (USA), ⁷ Rheinisch-Westfälische Technische Hochschule (Germany), ⁸ Massachusetts Inst. of Technology (USA), ⁹ Karlsruhe Institute of Technology (Germany), ¹⁰ École Polytechnique (France), ¹¹ Universita e INFN (Italy), ¹² University of California San Diego (USA)

E-mail: rapolas.kaselis@cern.ch

Abstract. CMS experiment utilizes distributed computing infrastructure and its performance heavily depends on the fast and smooth distribution of data between different CMS sites. Data must be transferred from the Tier-0 (CERN) to the Tier-1s for processing, storing and archiving, and time and good quality are vital to avoid overflowing CERN storage buffers. At the same time, processed data has to be distributed from Tier-1 sites to all Tier-2 sites for physics analysis while Monte Carlo simulations sent back to Tier-1 sites for further archival. At the core of all transferring machinery is PhEDEx (Physics Experiment Data Export) data transfer system. It is very important to ensure reliable operation of the system, and the operational tasks comprise monitoring and debugging all transfer issues. Based on transfer quality information Site Readiness tool is used to create plans for resources utilization in the future. We review the operational procedures created to enforce reliable data delivery to CMS distributed sites all over the world. Additionally, we need to keep data and meta-data consistent at all sites and both on disk and on tape. In this presentation, we describe the principles and actions taken to keep data consistent on sites storage systems and central CMS Data Replication Database (TMDB/DBS) while ensuring fast and reliable data samples delivery of hundreds of terabytes to the entire CMS physics community.

1. Introduction
The Compact Muon Solenoid (CMS) Experiment [1] is one of the four large particle physics experiments that have been recording proton-proton collisions at the center-of-mass energy of 7 TeV since March 2010. Recording lead ion-lead ion collisions at the center-of-mass energy of 2.76 TeV per nucleon pair since November 2010. In 2012 it started to record at 8 TeV at the LHC accelerator at CERN, Geneva, Switzerland [2]. CMS is located in an underground cavern at LHC Point 5 (P5), in Cessy, France. When LHC is running and CMS is recording data, firstly they are written to a large disk buffer at P5 [3]. Then data are transferred to CERN (Tier-0) [4] and processed for the first time. The data is stored on tape at CERN and also distributed to Tier-1 centers for a second archival copy on tape. At this stage, transfer operators start monitoring the data on its journey across the distributed
computing infrastructure. Also Monte Carlo (MC) simulations, which are produced at various Tier-2 centers, are sent back to Tier-1s for archiving. And in the end, data and MC is replicated back to Tier-2 centers for the collaboration to analyze.

2. CMS collaboration infrastructure [4]
The CMS computing model has several tiers of computing facilities:

- Tier-0 at CERN (T0), used for repacking, prompt reconstruction/processing/calibration, tape backup and exporting CMS data to Tier-1 centers.
- 7 Tier-1 (T1) centers, used for the tape backup, fast skimming and reprocessing CMS data at large-scale, furthermore to distribute data products to the Tier-2 centers. Also they are used for simulated data production and processing. The Tier-1 centers have large computing facilities and archival storage (based on tapes) systems.
- 54 Tier-2 (T2) facilities, where data analysis and Monte Carlo production are primarily carried out. These centers do not have tape backup systems, only disk storage. Each Tier-2 is “associated” to a Tier-1 for some services provided by T1 (i.e. FTS), usually on the basis of geographical proximity, but is not restricted to.
- There are also Tier-3 sites, which are not paid for through CMS, and not officially supported. Any individual or university can create a T3 site.

High-speed networks of 1-10 Gbps interconnect these sites and the CMS computing model envisions commissioning all links between all of these sites for production data transfers [5] creating a full mesh of 3782 links (and accordingly more, if a new T1 or T2 center is added).

3. PhEDEx
To archive and analyze collected data, CMS and other experiments rely on the Worldwide LHC Computing Grid (WLCG) [6], a distributed data grid of many processing and storage clusters varying in size, from tens of TBs to several PBs.

At the core of all CMS transfers there is the PhEDEx project (Physics Experiment Data Export) [7]. PhEDEx is based on a high-availability Oracle database cluster hosted at CERN (Transfer Management Data Base, or TMDB) acting as a “blackboard” for the system state (data location and current tasks). PhEDEx software daemon processes or “agents” connect to the central database to retrieve their work queue, and write back the results of their actions to TMDB. A set of central agents is running at CERN; their responsibility lies in distributing download tasks for each site, harvest statistics and provide data transfers summaries. At each site a different set of agents is running, which interacts with TMDB and initiates download transfers using site technology-specific back-ends through WLCG middleware.

PhEDEx also provides two interfaces for data transfers operations management and monitoring: an interactive web site, and a web data service to retrieve information (through representational state transfer (REST) [8] interface) from TMDB in machine-readable formats.

4. Data transfer operations
When LHC is running and CMS is recording proton-proton collisions data, there are 25-30 TB exported daily from Tier-0. During heavy-ion runs, these numbers are even more impressive, around 80 TB/day, with a record approximately 120 TB in one day (and 350 TB in total were transferred during the same day among all sites) [9]. Raw CMS detector data are transferred to Tier-1s centers from Tier-0 for archival storage, reconstruction and skimming [10]. Skimmed reconstructed data are then transferred for analysis to all Tier-2 sites as well as synchronized between the Tier-1 centers. Simultaneously, Tier-2 centers are generating Monte Carlo samples and distribute to Tier-1 sites for further archival.
4.1. Stress-Tests
To be able to satisfy required CMS data export rate, transfer tests are performed all the time. There is a feature built into PhEDEx designed exactly for testing purposes – transfer traffic generator. Using this tool, links between sites can be tested and problems spotted beforehand. The same tool is used to enable new links between various sites. In order to satisfy CMS needs, Tier-1 centers should be able to export data at minimum 20 MB/s rate per link, and Tier-2 centers at 5 MB/s per link. Once a needed rate on the link between sites is achieved, transfers for real data are enabled.

The most recent export stress testing from Tier-0 to Tier-1 centers in preparation for increased data taking rate was performed in the beginning of 2012 and lasted for more than a month. That was much longer than originally planned and shows that even those connections, which are working fine at decent rates might fail under increased load. In this case, there were some problems with one of the Tier-1 centers with achieving desired rate. After extensive tuning of various parameters on PhEDEx agents at site the best configuration was found and Tier-0 was able to export data at 1.2 GB/s while...
Tier-1 centers reached the planned download rate according to CMS tape pledges and the rate was sustained for 24h (Figure 2).

During the same tests, after the planned rates were achieved, it was decided to check if there is export rate reserve. Therefore, rates were increased by 25%, reaching 1.5 GB/s. At this rate visible degradation on Tier-0 storage performance was visible, however it was still able to transfer files out. The maximum export rate of 1.8 GB/s was achieved during the same tests. Therefore, from the transfers’ rate point of view, operations are safe, but they need to be constantly monitored.

4.2. Monitoring
While transfers are running, it is very important to keep them going steady, therefore monitoring is very important. So far, monitoring is done mostly by central transfers operators and the process is almost all manual. It is taking a lot of effort to keep transfers running at good level, both in time and manpower.

Monitoring [9] is done via web pages provided by PhEDEx, mostly checking rates table, where number of transferred files and errors are given for each link in real time and transfer quality plots, which provides an overall look on how transfers are proceeding. There are more than 2550 active links at the moment to watch with, a burst of links being added when new site is created.

Computing shifters do partial monitoring. They go through documented procedures what to check every two hours, which is very helpful in detecting stuck transfers. Transfers are stuck when data is requested to site, but there are no download attempts done by PhEDEx. Usually this happens because of no link between source and destination site exists. At the same time shifters also check if all needed central and site agents are running and reporting accordingly.

4.3. Transfer problems
With big number of sites and even bigger number of links there are many different kinds of issues. All transfer problems can be classified into 4 major groups (Table 1).

<table>
<thead>
<tr>
<th>Problem type</th>
<th>Description</th>
<th>Example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage issues</td>
<td>Issues and problems related to storage elements</td>
<td>Crashed disk server, corrupted tape, etc.</td>
</tr>
<tr>
<td>Network problems</td>
<td>Networking and physical level of transfers</td>
<td>Cable cut, router dropping packets, etc.</td>
</tr>
<tr>
<td>Authorization</td>
<td>Authorization is wrongly setup or not recognized</td>
<td>Not recognized or expired certificate, etc.</td>
</tr>
<tr>
<td>Misconfiguration</td>
<td>Configuration mistakes on PhEDEx, FTS, storage and so on.</td>
<td>Bad translation rule from logical file name to physical, PhEDEx not configured to export data to specific site, etc.</td>
</tr>
</tbody>
</table>

It is close to impossible to collect all error messages into one table, provide solution to it and leave monitoring to computing shifters. Error messages are changing with time, making automated checking for errors very complicated. At the same time, even errors for the same issue might be different, if the same purpose tools, made by different vendors, provide them. Also, quite often the error messages reported by the transfer system indicate a deeper underlying problem that requires expert’s attention. Therefore troubleshooting requires a lot of experience to be able to identify problematic side and suggest possible solution.

4.4. Troubleshooting
There are documented steps [11] in order to troubleshoot problematic transfers. Any site administrator
can follow them to debug transfer problems. Simply following those instructions, 80% of issues can be solved. Nevertheless, usually those steps are taken only when physicists miss their data or get an alarm from central operators about problematic transfers. Usually the same steps are being taken by central transfers operators as well. According to instructions, firstly transfer paths for the files should be checked whether PhEDEx knows source and destination sites. This easily can be checked in the PhEDEx web routing page. If no path exists, then transfer links should be checked. If link exists and is commissioned between the two sites, site agents should be checked if they are up and running. Next step is to check in which status files are. They can be stuck in staging from tape, or they can be transferring, if the latter, then transfers might be expiring in the download queue, or there are failures. Having failures with detailed error messages in front of the eyes helps to see where and what the problems are. After following these steps a Savannah ticket is opened to a suspected problematic site. Around 1300 tickets were opened in the last year and on average there are 15-20 open tickets pending to solve at any given time.

5. Improving transfers

A successful data transfers system, like PhEDEx for CMS, must rely on a stable, reliable and performing network infrastructure behind. The amount of data traffic, as well as the activated routes, in the distributed computing systems of LHC experiments comes from an original design based from the outcome of the MONARC project [13]. The MONARC project introduced the concept of a hierarchical multi-tiered computing infrastructure with defined functionalities for each Tier level. In the original LHC computing models based on these foundations, the major data traffic was expected to happen in the Tier-0 to Tier-1 routes, in the Tier-1 to Tier-1 routes and in the Tier-1 to Tier-2 routes. Once the experiments started to take proton-proton and heavy-ion data at 7 TeV, it soon became clear that the transfer routes were being used at a substantial scale in a previously unforeseen manner. For example, in the total production data volume transferred by the CMS experiment in 2010 the most used route was T1-T2 (48% of the total) and the second route was T2-T2. Since the network community had already design the LHCOPN network for T0-T1 and extended it also to T1-T1 routes, a work started to discuss how to "protect" the experiments traffic also in routes involving the large plethora of T2 sites. In this context, particularly relevant is the LHCONE project that is fully described elsewhere [14] and only introduced here.

The objective of LHCONE is to provide a collection of access locations that are effectively entry points into a network that is private to the LHC T1/2/3 sites. LHCONE is not intended to replace the LHCOPN but rather to complement it. It addresses Tier-2/3 level, on GPN [15] infrastructures in different nations. It is intended to grow as a robust and scalable solution for a global system serving LHC Tiers’ needs and to fit the new less-hierarchical computing models.

5.1. CMS tests to prepare for LHCONE

In this context, the CMS experiment started well in advance a T2-T2 "commissioning" work, which was finalized in 2010 already. In view of the inclusion of more T2/3 sites into the LHCONE prototype, the CMS experiment launched a program to get prepared. The basic idea is to test a link before, for benchmarking, and after it joins the prototype, in order to see the effects of the change at the network fabric level, and eventually give feedback to the network experts. Besides LHCONE, these tests were seen by CMS as a good review of the status of the connectivity among CMS sites at higher Tier levels, and so the effort was designed as much as much possible as integrated into the usual PhEDEx Operations practices.

CMS decided to adopt a strategy consisting of two complementary approaches. The first is based on the LoadTest (LT) infrastructure, i.e. use the Debug instance of PhEDEx to test links with Debugging Data Transfers (DDT) procedures, adapted from a "commissioning" scope to a "performance monitoring" scope. This approach, being based on PhEDEx, is fully CMS-specific. The effort needed to perform such tests is too much if the whole matrix of 2.5k permutations among more than 50 T2 sites is taken intro consideration, so only a subset of sites were selected, according to a list
of prioritized T2 sites which was written by the ATLAS and CMS computing coordination team. For CMS, the initial list consisted of a dozen of sites. The second approach is based on the gLite FTS and the FTM monitoring. With respect to the first one, this gives complementary, customizable, detailed info for each link under test (rate per stream, throughput), so once data is collected, the problem is addressed and scales for N links. Additionally, this is not CMS-specific, and it could be adopted by other experiments. In the rest, we mainly focus on the first approach, based on PhEDEx.

From the operational point of view, the basic idea of the LHCONE tests with the PhEDEx LoadTest is described in the following. A test on a single link (e.g. the link from the T2 named "A" to the T2 named "B") consists of:

- a careful check of the running traffic in the Prod and Debug instance, in terms of PhEDEx subscriptions, completion, etc. in order to decide whether to give green-light for a test on this link with no interference with existing activities, or instead to pass on to another link;
- if a good condition is found, the file injections from A and B to any other site are stopped and the subscriptions in Debug from any site to A and B are suspended (of course, the link from A to B is kept active, and also the T1-T2 links are kept active, otherwise the site readiness evaluation for that T2 would be affected);
- a one-time injection in the A-B unidirectional route is made, usually as big as 1k files, i.e. a data volume of roughly 2-3 TB; after some time of wait for completion, ad-hoc scripts that query the central PhEDEx transfer management database are able to extract information about the outcome of the A-B test;
- a clean-up is made and the cycle restarts with another route (e.g. C-D). Given how the test has been designed, some parallelism is also possible (e.g. A-B and C-D can be tested simultaneously with no interference).

For each test, a timestamp, the number of files transferred (and the corresponding size), maximum hourly transfer rate, the overall latency for the transfer of the full sample, and the quality of the transfer exercise (defined as the number of successes per the number of attempts) are extracted and noted down.

As of now, all links in the pilot matrix have been tested, and full matrixes of results, for each of the observables listed above, have been filled, yielding to a full set of information on all the links. These figures represent the benchmark, which CMS will be comparing the situation to, every time a site plugs in to the LHCONE prototype.

6. Data consistency

Increasing amount of data and frequent access to it by physicists greatly increases chances that one or another file gets corrupted or lost. Bigger load on the systems increases the possibility to crash, corrupting and/or destroying files. Therefore active monitoring of storage consistency is crucial. CMS data management has two central databases to keep meta-data:

- DBS (Dataset Bookkeeping Service) [16] catalogues all CMS event data from Monte Carlo and Detector sources.
- TMDB (Transfer Management Data Base) contains information about active data, transfers, where data have been produced and at which sites copy exists.

In addition, every site usually maintains local catalogues provided by their storage systems and accessible by various protocols depending on the technology.

There are several types of inconsistency with different consequences:

- Missing file. File location is recorded to be at some site, but it was lost, thus don’t exist anymore and any workflows requiring that file going to fail.
- Central data catalogues do not know about existing file at site (orphans). This type of inconsistency doesn’t create any issues to physicists, but makes inefficient storage utilization.
- File is damaged or corrupted. Either size or checksum, or both differ from what it should be.
Information again is stored in TMDB/DBS. This type of inconsistency might fail the workflow or even worse, produce the wrong results.

- Inconsistency between two central databases. Missing a file in one while it is present in another. That creates confusion and questions between users and physicists.

Even though whole CMS data transfer system is designed to do changes to the file transparently, in a highly distributed and complex environment, such as CMS experiment, data inconsistencies are still appearing due to different reasons. Failure in a file production system, human error factor when changes are made manually, heavy load on storage and processing systems might affect the regular site workflow, leading to inconsistency in the data catalogues. Storage systems migration to newer, bigger, better systems might create additional inconsistencies. Additionally, failed official deletions at sites create “orphans”. Maintaining clean and healthy CMS data established procedures for storage consistency checking are in use.

At the moment, storage consistency checks are being done centrally on monthly basis at all Tier-1s, except CERN. CERN contains data from the times before both central data catalogues existed, and is really complicated to know, whether data is really orphan, or just not catalogued, however steps are taken to catalog and clean all data and then include in monthly checks. Currently, the same storage consistency procedures are being extended to all Tier-2 centers.

6.1. Procedure
The procedure in its nature requires human judgment, so it cannot be fully automated. New tools are being designed and tested in order to make these tests semi automatic. It cannot be fully automated, because files are deleted after checks, and if some steps fail for one or another reason, like not mounted storage system, the result of automatic actions might be catastrophic with lots of data loss.

Communication with sites about consistency checks is held with the help of Savannah tickets. For each site, one ticket is opened asking for a full list of files stored under CMS used directories. Once a list is obtained, it is used to compare files with central data catalogues to find files, which are “orphan” – not known to central databases.

Checking for “dark” data is one half of consistency; another half is to check the storage element at site that all files are present, which should be present. For the latter task, BlockDownloadVerify (BDV) agent is used. But it can be used only if site is running recent PhEDEx version. This agent uses site-local plugins, which are optimized for concrete storage element and technology used at site. Big advantage in using this approach is that once site-local plugin is correctly installed and optimized, nothing for site needs to be done, tests can be injected centrally. At the same time, if one or several storage servers crash, it is easy to find out which files got corrupted or lost. Results of tests can be found on PhEDEx web page.

6.2. Challenges
Rapidly increasing amount of data is one of the several challenges. There are more than 20 millions physical files in the system, for some of which more than one copy is kept. And this number should match with information stored in central data booking catalogues. Distribution across many sites makes whole process a bit easier to handle as data can be checked at relatively small chunks, but at the same time big amount of sites makes the process much longer and almost never-ending. Once a one monthly round is finished, it is time to start another round of consistency checks in order to keep healthy data.

7. Summary
CMS experiment has a data transfer infrastructure, which have met raised requirements in the first years of LHC collisions. However, operational challenges are rising every day, which are tackled by transfer team. This paper presented the basic ideas how transfer operations are running and how the problems are addressed. Although the procedures are almost completely manual, there are steps taken
to automatize these processes and help in daily transfer tasks.

References
[7] T. Wildish 2012 From toolkit to framework - the past and future evolution of PhEDEx, these proceedings
[11] https://twiki.cern.ch/twiki/bin/viewauth/CMS/PhedexSiteAdminTransferTroubleshooting (requires authorisation)