

CMS EXPERIMENT DATA PROCESSING AT RDMS CMS TIER 2 CENTERS¹

V. Gavrilov¹, I. Golutvin², V. Korenkov², E. Tikhonenko², S. Shmatov²,
V. Zhiltsov², V. Ilyin^{3,4}, O. Kodolova³, L. Levchuk⁵

¹*Institute of Theoretical and Experimental Physics, Moscow, Russia*

²*Joint Institute for Nuclear Research, Dubna*

³*Skobeltsyn Institute of Nuclear Physics, Moscow State University, Moscow, Russia*

⁴*National Research Centre "Kurchatov Institute"*

⁵*National Science Center "Kharkov Institute of Physics and Technology", Kharkov, Ukraine*

1. Introduction

Russia and Dubna Member States (RDMS) CMS collaboration was founded in the year 1994 [1]. The RDMS CMS takes an active part in the Compact Muon Solenoid (CMS) Collaboration [2] at the Large Hadron Collider (LHC) [3] at CERN [4]. RDMS CMS Collaboration joins more than twenty institutes from Russia and Joint Institute for Nuclear Research (JINR) member states. RDMS scientists, engineers and technicians were actively participating in design, construction and commissioning of all CMS sub-detectors in forward regions. RDMS CMS physics program has been developed taking into account the essential role of these sub-detectors for the corresponding physical channels. RDMS scientists made large contribution for preparation of study QCD, Electroweak, Exotics, Heavy Ion and other physics at CMS. The overview of RDMS CMS physics tasks and RDMS CMS computing activities are presented in [5-11]. RDMS CMS computing support should satisfy the LHC data processing and analysis requirements at the running phase of the CMS experiment [12].

2. Current RDMS CMS Activities

During the last decade, a proper grid-infrastructure for CMS tasks has been created at the RDMS CMS institutes, in particular, at Institute for High Energy Physics (IHEP) in Protvino, Joint Institute for Nuclear Research (JINR) in Dubna, Institute for Theoretical and Experimental Physics (ITEP) in Moscow, Institute for Nuclear Research (INR) of the Russian Academy of Sciences (RAS) in Moscow, Skobetsyn Institute for Nuclear Physics (SINP) in Moscow, Petersburg Nuclear Physics Institute (PNPI) of RAS in Gatchina, P.N.Lebedev Physical Institute (LPI) in Moscow and National Scientific Center "Kharkov Institute of Physics and Technology" (NSC KIPT) in Kharkov. In the CMS global grid-infrastructure these RDMS CMS sites operate as CMS centers of the Tier-2 level with the following names: T2_RU_IHEP, T2_RU_JINR, T2_RU_ITEP, T2_RU_INR, T2_RU_SINP, T2_RU_PNPI, T2_UA_KIPT.

Since 2012 year, the CMS basic requirements to the CMS "nominal" Tier2 grid-site are:

- persons responsible for site operation at each CMS Tier-2 site;
- site visibility in the WLCG global grid-infrastructure (BDII);
- availability of recent actual versions of CMS Collaboration software (CMSSW);
- high efficiency of regular file transfer tests;
- certified links with CMS Tier-1 and Tier-2 grid-sites;
- regular CMS Hammer Cloud (HC) tests;
- 9.8 kHS06 of Processing Resources;
- disk space of 620 TB for: 30 TB of stage-out space; 200 TB of group space (100 TB per

¹ These activities are partially supported by a grant of the Russian Foundation for Basic Research (RFBR) and the Ukrainian Academy of Science (12-07-90402-Ukr_a)

group), 150 TB of central space, 120 TB of regional space and 120 TB of user space (~40 users of 3 TB each).

The integrated RDMS Computing infrastructure as a whole completely satisfies these requirements.

The RDMS CMS computing model provides a valuable participation of RDMS physicists in processing and analysis of CMS data. Since 2008, the RDMS Tier-2 centers have been associated with CMS Exotics Physics Analysis Group and CMS Muon Physics Object Group (both groups hosted at the JINR site), CMS Heavy Ion Physics Analysis Group (hosted at the MSU site) and JetMet/HCAL Physics Object Group (hosted at the ITEP). Some later the KIPT site was associated with CMS Electroweak Analysis Group. The special tests shown that the RDMS Tier-2 sites are satisfied all requirements for such hosting including the additional requirements for certification of data transfer links between RDMS sites and other Tier-2 centers associated also with the same CMS Physics Groups. In general, RDMS CPU resources are sufficient for processing and analysis of experimental data provided by the LHC and for simulation.

By the spring of the year 2011 CMS Tier-2 sites (computing centers) were considered in the context of the CMS computing requirements as “ready” for the data-taking phase of the experiment in the case of:

- site visibility and CMS virtual organization (VO) support;
- availability of disk and CPU resources;
- daily SAM tests availability > 80%;
- daily HC efficiency > 90%;
- commissioned links TO Tier-1 sites ≥ 2 ;
- commissioned links FROM Tier-1 sites ≥ 4 .

The status of readiness of RDMS CMS Tier-2 sites in October, 2012 is shown on Fig.1.

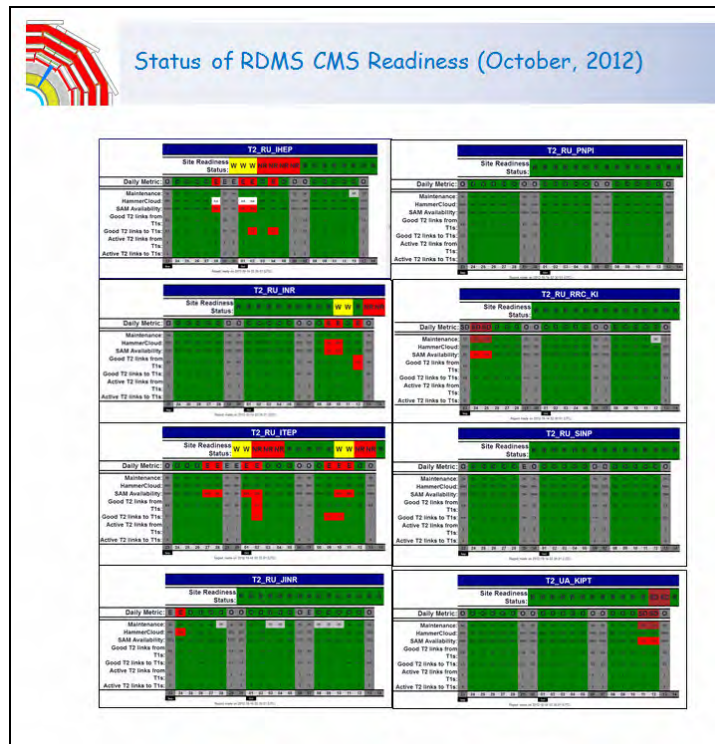


Fig.1: Readiness of RDMS CMS Tier-2 sites. October, 2012 (the permanent updated link see here <http://lhweb.pic.es/cms/SiteReadinessReports/SiteReadinessReport.htm>)

More than 560 TB were transferred to the RDMS Tier-2's from November 2011 to October 2012 (Fig.2). The maximum transfer rate to RDMS Tier-2 was more than 100 MB/s (Fig.3).

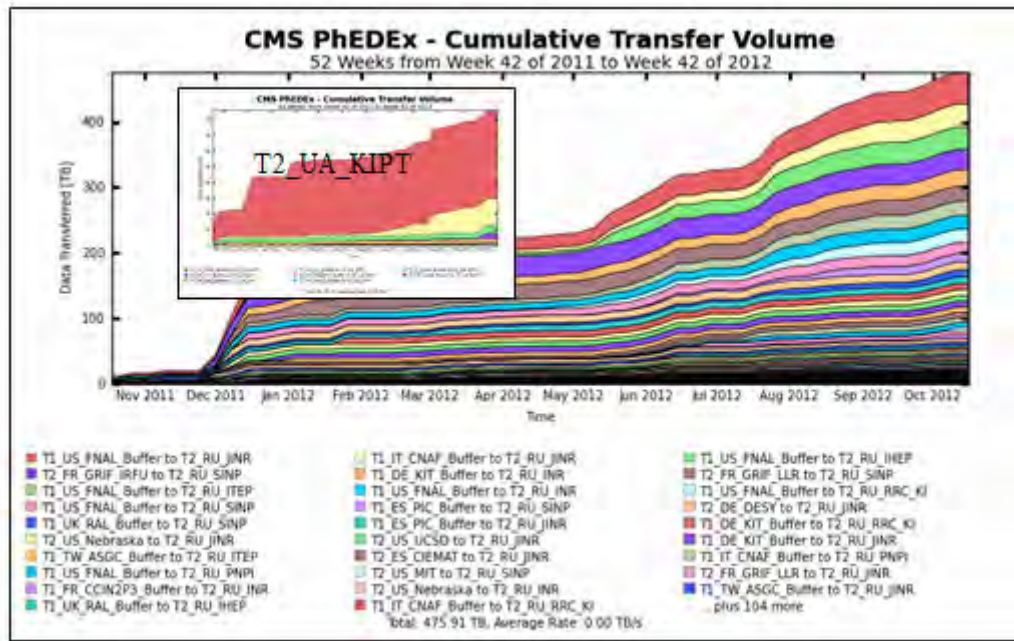


Fig.2: The cumulative transfer volume for the RDMS T2-sites from November, 2011 to October, 2012

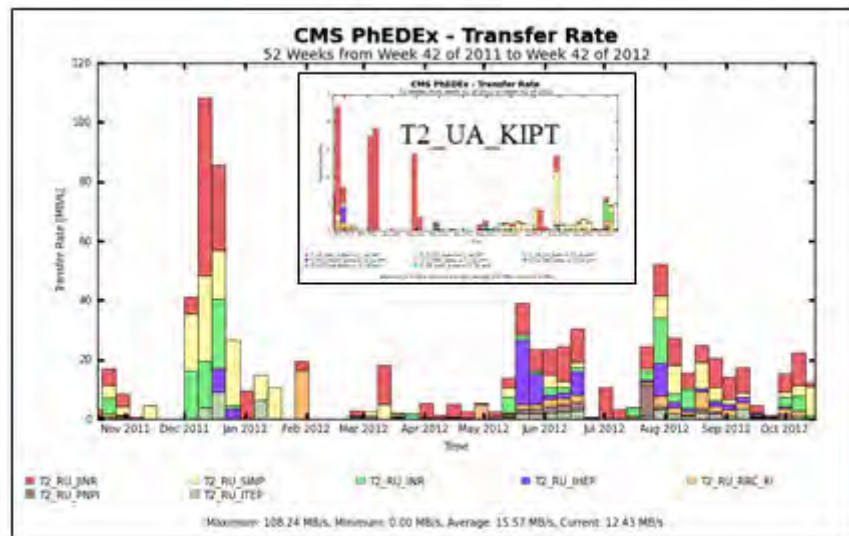


Fig.3: Transfer rates (more than 100 MB/s) at the RDMS Tier-2 sites during November, 2011 – October, 2012

The RDMS CMS Tier-2 sites are actively used by the CMS collaboration: 2,178,260 jobs (more than 7×10^9 events, see. Fig.4) of the CMS virtual organization were processed on the RDMS CMS Tier-2 sites in 2012.

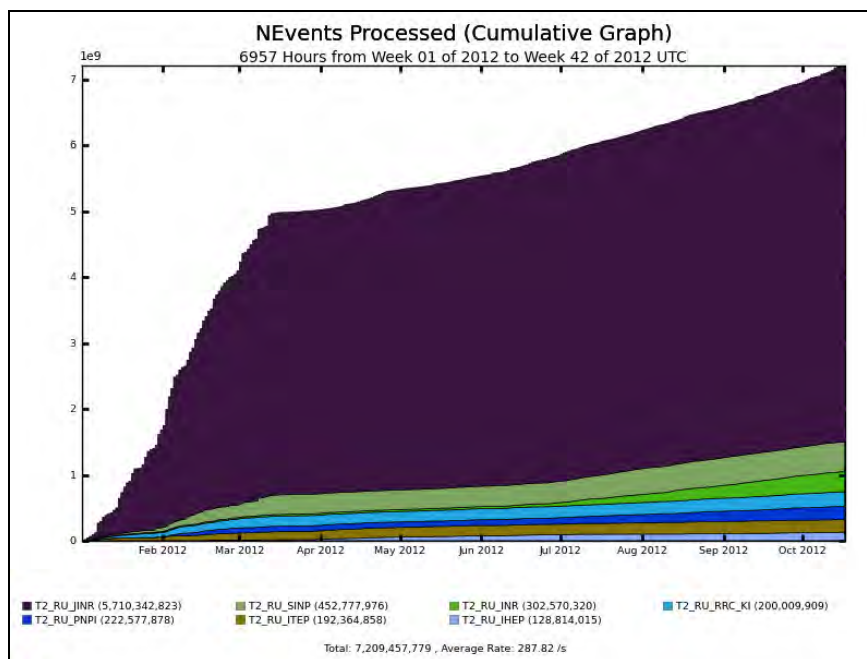


Fig.4: The number of processed events in 2012 with the RDMS CMS Tier-2 sites

In line with the CMS computing requirements for the data-taking phase of the experiment, now the RDMS CMS grid-sites provide:

- the computing and data storage resources in full;
- centralized deployment of actual versions of CMS specialized software (CMSSW);
- data transfers between the CMS grid-sites with the usage of the FTS grid-service on basis of VOBOX grid-services for CMS with the Phedex Server;
- SQUID proxy-servers for the CMS conditions DB access;
- certification of network links at the proper data transfer rates between JINR and CMS Tier1 and Tier2 centers;
- daily massive submission of CMS typical jobs by the CMS Hammer Cloud system;
- CMS data replication to the JINR data storage system in the accordance with RDMS CMS physicists' requests;
- participation in the CMS Monte-Carlo physical events mass production in the accordance with the RDMS CMS physicists' scientific program.

A group of RDMS CMS specialists takes an active part in the CMS Dashboard development (grid monitoring system for the CMS experiments) (<http://dashboard.cern.ch/cms>).

The dedicated CMS remote worldwide-distributed centers (ROC) were built in different scientific organization [13]. The JINR CMS Remote Operation Center (ROC) was founded in 2009 to provide participation in CMS operations of a large number of RDMS CMS collaborating scientists and engineers. The JINR CMS ROC is designed as part of the JINR CMS Tier 2 center and provides the following functions:

- monitoring of CMS detector systems;
- data monitoring and express analysis;
- shift operations;
- communications of the JINR shifters with personal at the CMS Control Room (SX5) and CMS Meyrin centre;
- communications between JINR experts and CMS shifters;
- coordination of data processing and data management;
- training and information.

In 2010 the CMS ROC was founded and certified also at the SINP MSU to provide similar functions for CMS participants in Moscow.

RDMS CMS physicists work in the WLCG environment, and now we are having more than 30 members of the CMS Virtual Organization.

3. Summary

The RDMS CMS computing centers have been integrated into the WLCG global grid-infrastructure providing a proper functionality of grid services for CMS. During the last two years a significant modernization of the RDMS CMS grid-sites has been accomplished. As result, computing performance and reliability have been increased. In frames of the WLCG global infrastructure the resources of the both computing centers are successfully used in a practical work of the CMS virtual organization. Regular testing of the RDMS CMS computing centers functionality as grid-sites is provided.

All the necessary conditions for CMS data distributed processing and analysis have been provided at the RDMS CMS computing centers (grid-sites). It makes possible for RDMS CMS physicists to take a full-fledged part in the CMS experiment at its running phase.

References

- [1] V. Matveev, I. Golutvin, "Project: Russia and Dubna Member States CMS Collaboration / Study of Fundamental Properties of the Matter in Super High Energy Proton-Proton and Nucleus-Nucleus Interactions at CERN LHC" , 1996-085/CMS Document, 1996; <http://rdms-cms.jinr.ru>
- [2] CMS Collaboration, Technical Proposal, CERN/LHCC, 94-38, 1994; <http://cmsinfo.cern.ch>
- [3] <http://public.web.cern.ch/Public/Content/Chapters/AboutCERN/CERNFuture/WhatLHC/WhatLHC-en.html>
- [4] <http://www.cern.ch>
- [5] V. Gavrilov et al., "RDMS CMS Computing Model", in Proc. of the Int.Conference "Distributed Computing and Grid-Technologies in Science and Education", Dubna, 2004, pp.240-247.
- [6] V. Gavrilov et al., "RDMS CMS Computing", in Proc. of the 2nd Int.Conference "Distributed Computing and Grid-Technologies in Science and Education", Dubna, 2006, pp.61-66.
- [7] D.A Oleinik et al., RDMS - CMS Data Bases: Current Status, Development and Plans, in Proc.of the XXth Int. Symposium on Nuclear Electronics and Computing, JINR, Dubna, 2006, pp.216-221.
- [8] V. Gavrilov at al., Current Status of RDMS CMS Computing, in Proc. of the XXI st Int. Symposium on Nuclear Electronics and Computing, Dubna, 2008, pp.203-208.
- [9] D.A. Oleinik at al.,Development of the CMS Databases and interfaces for CMS experiment, in Proc. of XXI Int. Symp. on Nuclear Electronics & Computing (NEC`2007), Dubna, 2008, pp.376-381.
- [10] V. Gavrilov at al., RDMS CMS Computing activities before the LHC startup, in Proc.of 3rd Int.Conference "Distributed Computing and GRID-technologies in Science and Education, Dubna, 2008, pp.156-159.
- [11] V. Gavrilov et al., RDMS CMS data processing and analysis workflow, in Proc. of XXIII Int. Symp. on Nuclear Electronics & Computing (NEC`2011), Dubna, 2011, pp.148-153.
- [12] CMS Collaboration, The Computing Project, Technical Design Report, CERN/LHCC-2005-023, CMS TDR 7, 2005.
- [13] A.O. Golunov et al, "The JINR CMS Remote Operation Centre", Distributed Computing and Grid-technologies in Science and Education IV Int.Conference, Proceedings of the conference, Dubna, 2010, pp.109-113.