2017

CMS Connect

J. Balcas

B. Bockelman

R. Gardner Jr.

K. Hurtado Anampa

B. Jayatilaka

See next page for additional authors

Authors
J. Balcas, B. Bockelman, R. Gardner Jr., K. Hurtado Anampa, B. Jayatilaka, F. Aftab Khan, K. Lannon, K. Larson, J. Letts, J. Marra Da Silva, M. Mascheroni, D. Mason, A. Perez-Calero Yzquierdo, and A. Tiradani

Journal of Physics: Conference Series

PAPER • OPEN ACCESS

CMS Connect

To cite this article: J Balcas *et al* 2017 *J. Phys.: Conf. Ser.* **898** 082032

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 082032 doi:10.1088/1742-6596/898/8/082032

CMS Connect

J Balcas[1], B Bockelman[2], R Gardner Jr.[3], K Hurtado Anampa[4], B Jayatilaka[5], F Aftab Khan[6], K Lannon[4], K Larson[5], J Letts[7], J Marra Da Silva[8], M Mascheroni[5], D Mason[5], A Perez-Calero Yzquierdo[9,10] and A Tiradani[5]

[1] California Institute of Technology, Pasadena, CA, USA
[2] University of Nebraska-Lincoln, Lincoln, NE, USA
[3] University of Chicago, Chicago, IL, USA
[4] University of Notre Dame, Notre Dame, IN, USA
[5] Fermi National Accelerator Laboratory, Batavia, IL, USA
[6] National Centre for Physics, Quaid-I-Azam University, Islamabad, Pakistan
[7] University of California, San Diego, La Jolla, CA, USA
[8] Universidade Estadual Paulista, São Paulo, Brazil
[9] Port d'Informació Científica, Barcelona, Spain
[10] Centro de Investigaciones Energéticas Medioambientales y Tecnológicas, Madrid, Spain

E-mail: `khurtado@nd.edu`

Abstract. The CMS experiment collects and analyzes large amounts of data coming from high energy particle collisions produced by the Large Hadron Collider (LHC) at CERN. This involves a huge amount of real and simulated data processing that needs to be handled in batch-oriented platforms. The CMS Global Pool of computing resources provide +100K dedicated CPU cores and another 50K to 100K CPU cores from opportunistic resources for these kind of tasks and even though production and event processing analysis workflows are already managed by existing tools, there is still a lack of support to submit final stage condor-like analysis jobs familiar to Tier-3 or local Computing Facilities users into these distributed resources in an integrated (with other CMS services) and friendly way. CMS Connect is a set of computing tools and services designed to augment existing services in the CMS Physics community focusing on these kind of condor analysis jobs. It is based on the CI-Connect platform developed by the Open Science Grid and uses the CMS GlideInWMS infrastructure to transparently plug CMS global grid resources into a virtual pool accessed via a single submission machine. This paper describes the specific developments and deployment of CMS Connect beyond the CI-Connect platform in order to integrate the service with CMS specific needs, including specific Site submission, accounting of jobs and automated reporting to standard CMS monitoring resources in an effortless way to their users.

1. Introduction

The Compact Muon Solenoid (CMS) is a high energy physics experiment analyzing proton-proton and heavy ion collisions produced at the Large Hadron Collider (LHC) at CERN. The study of the Higgs boson and the Standard model, the search for supersymmetric particles, extra dimensions and particles that could make up dark matter are some of the topics this detector has been designed to investigate [1].

However, the searches for new physics are not the only challenges facing the experiment. The LHC generated 75 Petabytes of data during its first three years of running and it is generating

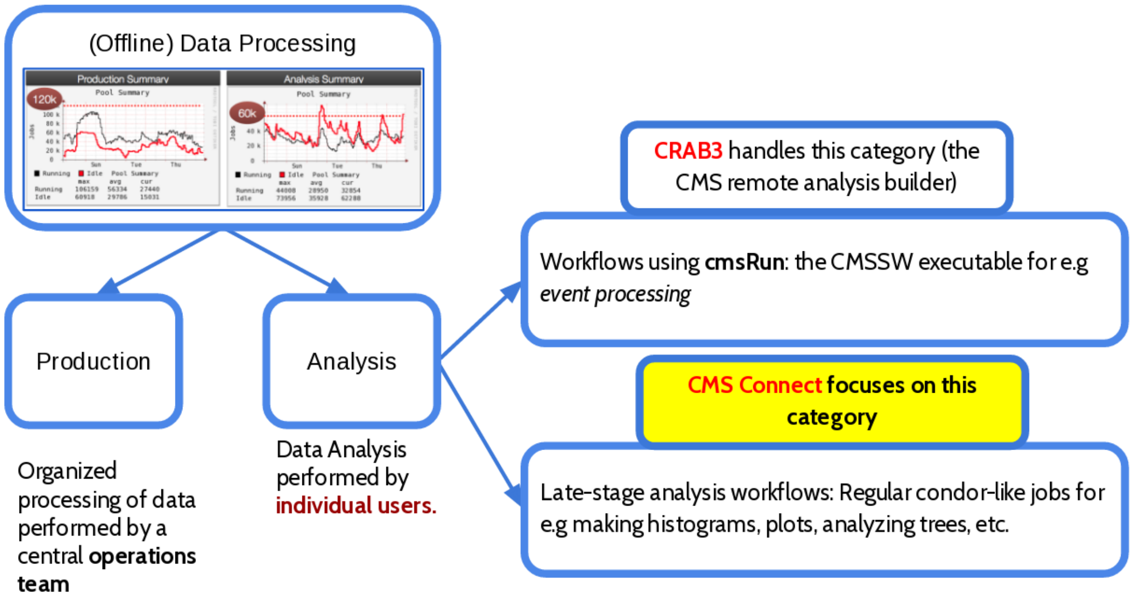

Figure 1. Offline data processing: Production and Analysis jobs are handled differently and both have their own submission tools or services to access the CMS Global Pool.

data at a rate of 40-50 Petabytes per year during its current second run [2]. Even after processing, selection and reduction of reconstructed data carried out centrally in CMS, a single physics analysis can still require from dozens to hundreds of terabytes of data and there is a particular high demand for computing resources for data processing in the experiment because of this. The CMS Global Pool [3] is the main source of computing resources to fulfill this need.

Different tools exist to submit jobs for simulation, reconstruction and analysis of events to the Global Pool. For instance, the production of simulated events is performed by a central operations team with the WMAgent framework [4], while analysis jobs based on the standard CMS executable (cmsRun) are submitted by individual users, mostly using the CMS Remote Analysis Builder (CRAB) [5], as seen in Figure 1. However, a variety of physics workflows going from generation of simulated events to the analysis of the detector data, background studies, creation of histograms and plotting do not depend on cmsRun and often either provide support for different batch systems or are easier to handle and submit via bare HTCondor jobs. Data Analysis Facilities like the LPC CAF at Fermilab [6] or Tier-3 sites are generally used for these workflows, but such alternatives are only valid to run locally.

CMS Connect was developed to provide users the ability to submit HTCondor jobs to CMS Global Pool resources on a submission interface that is similar to the analysis facilities that physicists are used to work with, yet fully integrated with CMS submission infrastructure components to guarantee proper fair share of resources per user, as well as reporting analysis activity to standard CMS monitoring services.

The sections below describe the platform CMS Connect was built on and the different existing components in the service that provide user account creation, group management, job submission, reporting of jobs and the service monitoring and notification system.

2. Base platform

CMS Connect is built on CI Connect [7], a platform developed for building multi-institutional virtual cluster environments. This platform relies on different well supported technologies to

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 082032 doi:10.1088/1742-6596/898/8/082032

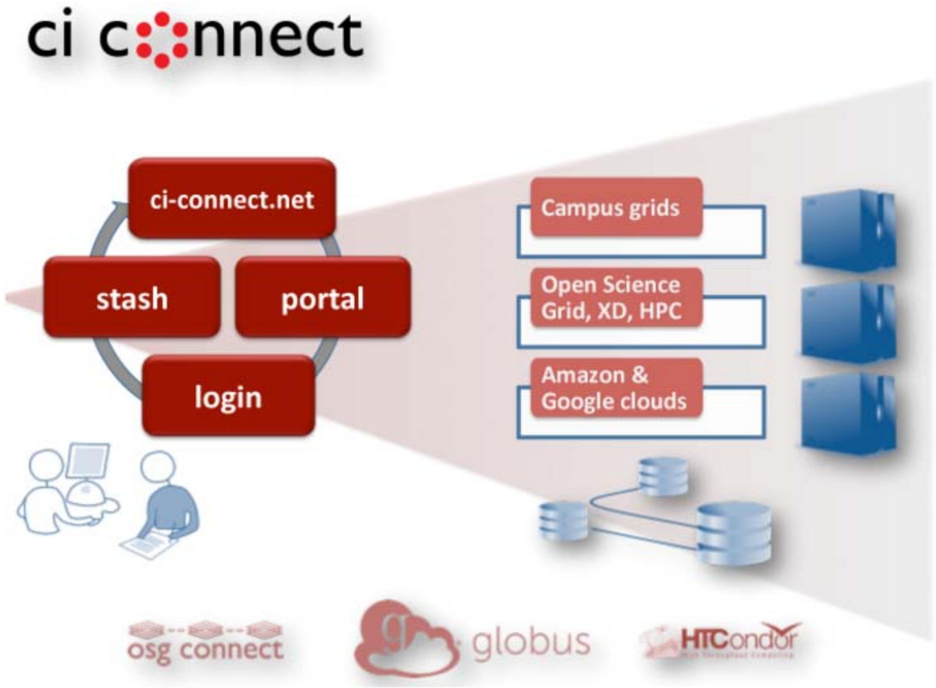

Figure 2. Diagram of the CI Connect platform.

provide a set of computing services focused on batch oriented analysis. CI Connect has been adopted by different organizations and institutions, such as the Open Science Grid [8], ATLAS [9] and Duke University [10].

The different elements of this platform can be seen in Figure 2. A web portal integrated with Nexus [11], a Globus service to provide identity and group management, is used to register and organize users in different group projects. A user account to the login node is auto provisioned after approval, giving the user interactive login access to a submission point with HTCondor, access to distributed software through CVMFS [12] and access to data through the wide-area network with XRootD [13].

CI Connect services are hosted in a 100 Gbps Science DMZ, a scalable network architecture explicitely designed for high-performance scientific applications [14].

3. CMS Connect Components

As illustrated in Figure 3, CMS Connect starts with the CI Connect elements that manage the user account creation and storage configuration, followed by the components that provide integration with the CMS Global Pool and perform job reporting. Each one of these components is described below.

3.1. Account creation and group management

User identities are managed by Globus Nexus, as previously mentioned in Section 2. Nexus requires only minimal information about the user (username, Full Name, Institution, email address and password), verifies password strength, username uniqueness and provides a group interface that allows organization of users per institution, as shown in Figure 4. Managers can create analysis subgroups allowing for better organization and monitoring of activity in the service.

3.2. Stash Area

Even though all CMS users have access to storage for analysis at either Tier-2 sites or Central Analysis Facilities, it is sometimes convenient to have storage on the submission machine for certain activities.

CMS Connect users have access to a temporary data storage provided by the Open Science Grid to assist with pre-stage job input files, sandboxes, etc. This area can be accessed via fuse in the login node, or XRootD and HTTP remotely. The latter can be used as an alternative to HTCondor transfer mechanisms to avoid adding more work and network activity to the submission machine job scheduler.

3.3. Job submission

HTCondor is used at the login node to allow users to submit jobs. Additional components are added in a wrapper that runs on top of HTCondor in order to integrate the job scheduler with the CMS Global Pool.

Users are required to send their CMS proxy certificates with their jobs. This allows the service to preserve fair share accounting in the pool by mapping the user's Distinguished Name (DN) with its corresponding CERN username (which is used for such accounting). This requirement also increases the security and traceability in the submission system and avoids issues with workflows that need to transfer data from and to other CMS Tier sites (CMS proxy certificates are mandatory for this).

CMS Connect submits to all computing resources available at the Global Pool unless the user specifies a list of desired sites, which is done by adding a single HTCondor ClassAd (a mapping from attribute names to expressions) in the submit file [15]. Such ClassAd is later recognized by the GlideinWMS infrastructure in the Global Pool and handled properly for resource provisioning.

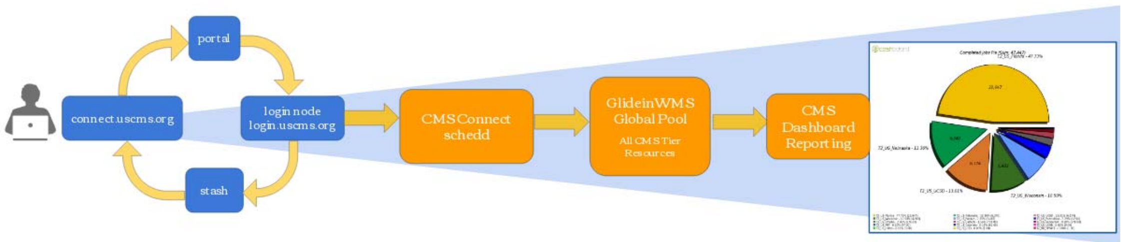

Figure 3. CMS Connect main components.

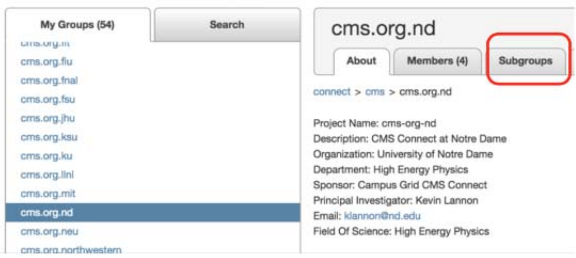

Figure 4. Group Management in CMS Connect.

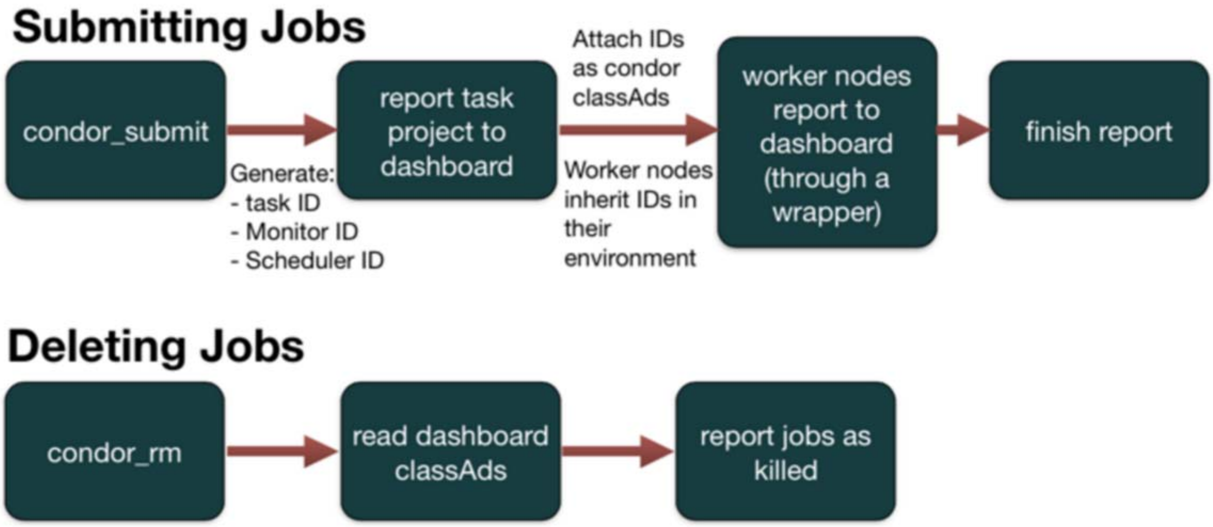

Figure 5. Diagram of the CMS Dashboard integration.

3.4. CMS Dashboard

The CMS Dashboard [16] is the service where the experiment archives production and analysis activity on the computing resources. CMS Connect integrates dashboard reporting for HTCondor jobs submitted from this service.

The dashboard reporting procedure is pictured in Figure 5. When users submit jobs, three parameter IDs required by the dashboard are generated to report the task project. Such parameters are added to the job ClassAds and worker nodes inherit such Ids in their environment. Each job status is then reported from their respective worker node through a wrapper. When a job is deleted, CMS Connect reads the dashboard ClassAds in the job to report it as killed.

A CMS dashboard pie chart of jobs completed via CMS Connect and the different sites such jobs were completed at can be seen Figure 6.

4. Service monitoring and notification system

Maintaining the high availability of a complex system is often a challenge, as the failure of a single component can affect the proper functioning of the entire service. Therefore, monitoring these components is extremely important.

CMS Connect uses Jenkins [17], an open source automation server to periodically test all service components, including access to the login node, Stash and CVMFS availability and submission of jobs to the CMS Global Pool, as shown in Figure 7. When one of these components fails, an email notification is immediately sent to the CMS Connect system administrator. Additionally, failures related to the CI Connect platform can be notified to the Open Science Grid help desk for the right support team to take action, allowing us to maximize the availability of the service to the users.

5. Conclusions

We have developed a service designed to complement existing submission tools in the CMS experiment for HTCondor-based analysis workflows. CMS Connect provides a virtual environment that resembles local cluster submission interfaces, but has all advantages the CMS Global Pool has to offer, besides providing integration with the CMS dashboard for reporting.

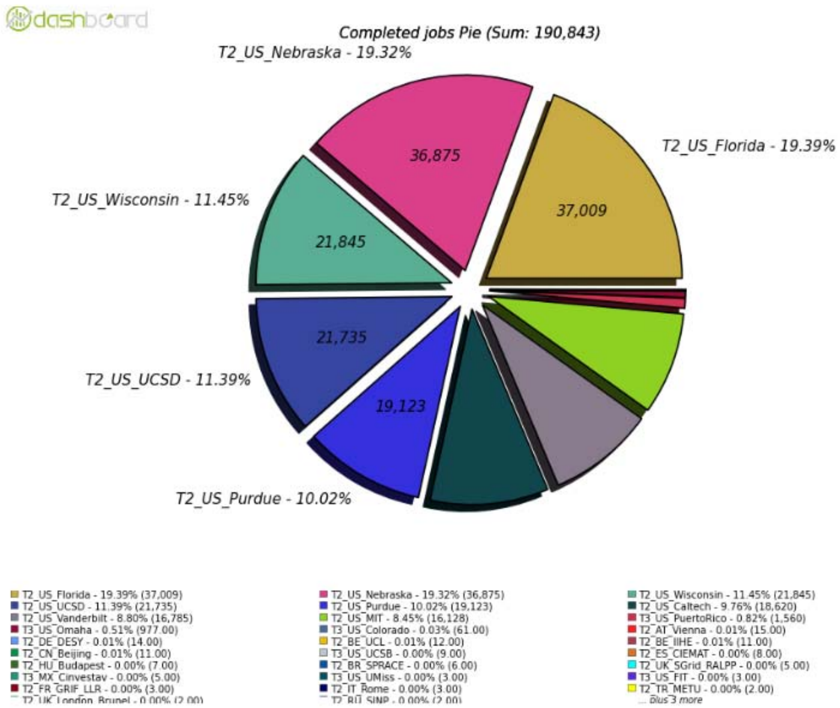

Figure 6. CMS Dashboard.

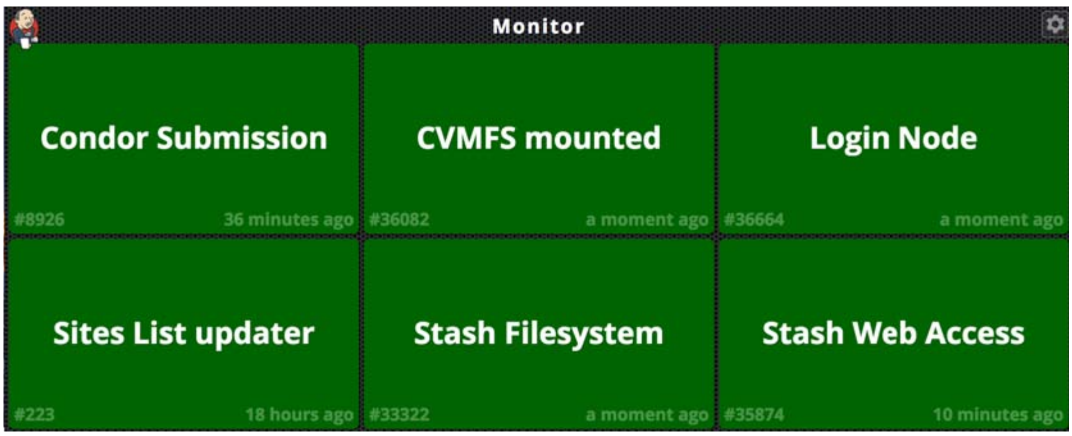

Figure 7. CMS Connect service monitoring via Jenkins.

Acknowledgments
The Open Science Grid and University of Chicago Team not only provide and maintain the CI-Connect based submission hardware and networking infrastructure, but also support the service on different levels, including hosting, registration, software updates and maintenance notifications, within other tasks. We are especially grateful to Robert Gardner, Lincoln Bryant, Suchandra Thapa and Balamurugan Desinghu who assisted closely in the integration of the CI-Connect platform with the CMS Global Pool and continue supporting the service operability. The present work is partially funded under grants from the U.S Department of Energy, the National Science Foundation and Spain Ministry of Economy and Competitiveness

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 082032 doi:10.1088/1742-6596/898/8/082032

grant FPA2013-48082-C2-1/2-R. The Port d Informació Científica (PIC) is maintained through a collaboration between the Generalitat de Catalunya, CIEMAT, IFAE and the Universitat Autonoma de Barcelona.

References

[1] Chatrchyan S *et al.* (The CMS Collaboration) 2008 The CMS experiment at the CERN LHC JINST **3** S08004
[2] Cancio G, Bahyl V, Kruse D, Leduc J, Cano E and Murray S 2015 Experiences and challenges running CERN's high capacity tape archive *J. Phys. Conf. Ser.* **664** 042006
[3] Balcas J *et al.* 2016 Stability and scalability of the CMS Global Pool: Pushing HTCondor and glideinWMS to new limits, to be published in these proceedings
[4] Boudoul G, Franzoni G, Norkus A, PolA , Srimanobhas P and Vlimant J R 2015 Monte Carlo Production Management at CMS *J. Phys. Conf. Ser.* **664** 072018
[5] Cinquilli M, Spiga D, Grandi C, Hernandez J M, Konstantinov P, Mascheroni M, Riahi H and Vaandering E 2012 CRAB3: Establishing a new generation of services for distributed analysis at CMS *J. Phys. Conf. Ser.* **396(3)** 032026
[6] Riahi H *et al.* 2011 Large scale and low latency analysis facilities for the CMS experiment: development and operational aspects *J. Phys. Conf. Ser.* **331** 072030
[7] The CI-Connect platform `http://ci-connect.net`
[8] Jayatilaka B, Levshina T, Rynge M, Sehgal C and Slyz M 2015 The OSG open facility: A sharing ecosystem using harvested opportunistic resources *J. Phys. Conf. Ser.* **664** 032016
[9] ATLAS Connect `https://connect.usatlas.org`
[10] Duke Connect `https://duke.ci-connect.net`
[11] Chard K, Lidman M, McCollam B, Bryan J, Ananthakrishnan R, Tuecke S and Foster I 2016 Globus Nexus: A Platform-as-a-Service provider of research identity, profile, and group management *Future Gener. Comput. Syst.* **56** 57183
[12] Blomer J, Aguado-Sánchez C, Buncic P and Harutyunyan A 2011 Distributing LHC application software and conditions databases using the CernVM file system *J. Phys. Conf. Ser.* **331** 042003
[13] Dorigo A, Elmer P, Furano F, and Hanushevsky A 2005 XROOTD - a highly scalable architecture for data access. *WSEAS Trans. Comp.* **4(4)** 34853
[14] Dart E, Rotman L, Tierney B, Hester M and Zurawski J 2013 The Science DMZ: A network design pattern for data-intensive science *SC: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis* Denver, CO USA ACM 85:1-85:10
[15] CMS Connect documentation `http://docs.uscms.org/Selecting+Sites`
[16] Karavakis E, Andreeva J, Maier G and Khan A 2012 CMS dashboard for monitoring of the user analysis activities, *AIP Conf. Proc.* **1504(1)** 983-86
[17] Jenkins `https://jenkins.io`