CRAB3: 2012 Plan

- Overview

- WMAgent Changes

- Glidein Usage

- Transition to CRAB3

Eric Vaandering

OSG All Hands

UNL

20 March 2012

CRAB

- CRAB enables submission of CMSSW jobs to all CMS datasets within the data-location driven CMS computing structure

- The aim of CRAB is to hide as much of the complexity of the GRID as possible from the end user

- CRAB provides a user front-end to

  - Find data in and publish data to DBS

  - Split user jobs into manageable pieces

  - Transport user analysis code to the data location for execution (compiled on submitting node)

  - Execute user jobs, check status and retrieve output

Software Components

- CRAB interacts with many different pieces of CMS and Grid software on behalf of the user

  - DBS: What data exists? Publish the results.

  - Phedex: Where is the data?

  - BDII/SiteDB: What sites are available, who is the user?

  - Proxies/MyProxy: User authentication

  - Dashboard: Statistics and status of jobs

  - Grid middleware: Job submission, tracking

CRAB2 Development

- CRAB2 must still keep up with changes to CMSSW, CMS services, and Grid middleware. Unless you have a very good reason, try to stay current with releases.

  - Only CRAB 2.8.x currently working

- CRAB 2.x is now in maintenance mode. Only important bug fixes are being made

  - Freeze becomes deeper and deeper over time

CRAB3 is not CRAB2++

- What is CRAB3 and why are we doing it?

  - CRAB3 is a complete rewrite of CRAB. Only the name remains the same. No code in common.

  - Many preconceptions of "CRAB does X" should be discarded. It may not work that way anymore.

    - Uses WMAgents submission model

  - Same client/server model as CRABServer

  - Much thinner client, all work done on the server

    - CRAB2 did all job preparation on the client

    - No standalone mode, only server

CRAB3: Motivation

- More stable development model

  - Based on WMAgent, the current CMS workflow software

    - Consolidated development, WMAgent designed for data

  - Modular structure allows us to add features we've said "later" to for years

- Simplify by consolidating Stand Alone and Server modes

  - A number of functions implemented twice, slightly differently

- In CRAB2, two ways for users to get output

  - Return through Condor "output_files_to_transfer" or gLite sandbox

  - srmcp from worker node to home SE at end of job

  - Both methods are problematic

    - In fact, biggest source of unforced errors in CRAB2

- CRAB3 scraps all of this and uses centralized FTS transfers

  - Job on WN writes to /store/temp/user

  - FTS jobs submitted to move files to home SE

  - Should increase CPU efficiency on WN, make transfers more reliable

Changes to WMAgent

- We've had to modify/extend some portions of WMAgent for CRAB3

  - Mostly related to proxies. User's proxy must be delegated to the worker node for writing files to local SE, srmcp of logs, and glexec under Glideins

  - gLite submission added to WMAgent

    - BossAir instead of BossLite (again, all new code)

    - I expect this will never be used in production

  - AsyncStageout is specific for users too

- Glexec switching on the glidein submit node as well

  - Isolate users from each other, fair share when using glidein

- We really want glexec on the worker nodes

  - Ken has mentioned this in Tier2 computing mailings

- Input (user code) is handled differently than in CRAB2.

  - Stored on a central server, job (HTTP) requests from WN

  - Uses squids for caching. Simplistic now, may integrate use case into FRONTIER in the future.

- /store/temp/user (and maybe /store/temp/group) must be writable locally

  - Would be great to have SAM tests and/or pilots check for this

Comments

- Keep in mind that even though Glidein may do some things WMAgent does internally

    - We are not 100% sure of the status of gLite with WMAgent

    - We need to allow support LSF, Condor, PBS, ARC

    - So, we need to aim for the lowest common denominator

- Under active development. "Alpha" versions in the hands of integration

    - Underlying WMAgent is used for all Tier1 & MC work, not RelVals yet

- CRAB3 should be useful for some real workflows by expert users in June

- FTS based asynchronous stageout is working, still shaking out

- Very little effort yet on modifications for local submission mode

Current Development Priorities

- Monitoring to separate central cmsweb from submission/ job tracking

- Porting central components to cmsweb API

- MonteCarlo workflows

- Finishing up publication of output

- Support overflow (longer term)

Local Mode Plans

- One missing part of WMAgent is support for local schedulers

    - Scheduler plugins are easy, also need support for user switching with glexec because of server architechture

        - Of course, this is ~same as the switching used for Glidein submission

        - Delegated through myproxy

    - This is a requirement for the FNAL LPC Tier3

        - One of or the biggest analysis resources in CMS

        - Local Condor scheduling

- Bad news: Local mode will involve running a server

Physics Operations

- These plans are still being fleshed out

- Support model could change drastically with CRAB3

    - CRAB3 WMAgent is ~the same as production, so perhaps operators will be the same and PhysicsOps does not need expertise running a CRABServer

    - Some parts of CRABServer (ReqMgr, Stageout, etc) will be centralized at CERN.

    - Unclear if we need as many WMAgent instances as we have CRABServers now

- Perhaps PhysicsOps concentrates more on user support

Transition to CRAB3

- 2012 will be a year of transition

- Confident CRAB3 will be more reliable than CRAB2

- Expect that as feature set of CRAB3 grows, more people will move over

  - Can take over some non-CRAB2 workflows as well. (e.g. FWLite workflows)

- Still expect that CRAB2 will be supported through most of 2012

  - CRAB2 on the FNAL LPC will be supported until a replacement is available

- Grid submission for CMS should become much more reliable going forward

- Little that Tier2/Tier3 sites have to do. These are, after all, Grid jobs

- Be prepared for FTS transfers like MC production

- Squid proxy to cache user sandboxes

- Running client from your Tier3 will be easy

- Setting up a server to directly access local resources will be more challenging