Simulation (Monte Carlo) Production for CMS

Stefano Belforte

WLCG-Tier2 workshop

Outline

- What MonteCarlo simulation is

- How is it done at CMS

- What Tier2 role is

- What services Tier2 need to provide

- Tutorials later this week to learn how exactly to setup and operate those services

Simulation Production

- Simulation (Monte Carlo) Production means
  - Generate, simulate, reconstruct data that look like data coming from detector, i.e. as realistic as possible, but all simulated
- Ideally:
  - Input is a configuration (few KB), output a dataset (TB)
  - Heavily CPU-bound activity
- In practice:
  - Simulation, reconstruction, addition of multiple interactions (pileup) look more like analysis: data in – executable – data out
  - Some step may be I/O bound
  - CMS will try to combine the flow and run CPU bound jobs

- **Expect a lot of CPU bound jobs, but do not build a Tier2 thinking only of CPU**
  - Tier2 also needs to run analysis anyhow

Simulation at Tier2

- In CMS Computing Model and Computing-TDR, Tier2's are the place where MC production is run
  - ➢ Link to C-TDR
- Tier2 comply with this requirement by offering a computing service that is exploited by CMS MC Operation team to produce simulation datasets according to CMS policy (defined by CMS physics groups)

- Tier2 do not define autonomously what data needs to be produced

- There will be "user level MC", small production that satisfy the need of single user and do not go through physics groups scheduling
  - ➢ Could still be managed centrally or run directly by that user
  - ➢ To a Tier2 it will look the same: a bunch of grid jobs
- Tier2 are not handed a list of jobs to run, they are used as service providers
- **Tier2 are "good grid sites that offer CMS-needed services"**

Simulation Workflow

- Simulation is a grid based activity
  - ➤ Jobs are submitted centrally
  - ➤ Expect 2-5 groups of operators to be able to keep the system busy and manage all production
  - ➤ Local submission is not forbidden, but un-economic
- Automated tools will be used
  - ➤ Central request queue (ProductionManager)
  - ➤ Chunks of (thousands of) jobs are handed over to a few ProductionAgents that submit, track, bookkeep, store output to final location and register in CMS databases
- Each ProductionAgent will need a human being (or two)
  - ➤ Ideally one ProductionAgent can fill the grid
  - ➤ In practice humans behind it will need to deal with operational failures on the grid, hence need to have more people and split competences
  - ➤ E.g. one group for OSG, one for EGEE
- **Your customers will be remote users**

- Simulation is an expensive task (many CPU hours): precious output
- Data will be stored at CMS Tier1's
  - ➤ Backup on tape
  - ➤ May need to re-process with new reconstruction
  - ➤ May need to distribute to more T2
  - ➤ Free disk while making it available even years later from tape

- But job output does not go directly from WN to T1:
  - ➤ Each WN output is stored locally
  - ➤ small files are merged locally (another set of central jobs)
    - ☞ Heavily I/O bound !
  - ➤ A validation step is run locally
  - ➤ Large files are sent to T1
  - ➤ Sometimes the output is short lived, may stay a week at T2, then get deleted
- **Full integration in CMS data management system needed**

Resource management

- In CMS Computing Mode Tier2 also support majority of user analysis
  - Each Tier2 more or less split: 50% analysis, 50% simulation
- Both activities can easily fill whatever resources are there
  - Simulation is planned, can fill queues with weeks-worth of jobs
  - Analysis by users is dynamic: I need this by tomorrow

- Need a way to partition resources

- Grid tools are not so developed here, especially on EGEE side
  - You will need to help, at least initially
  - EGEE PreProduction sites will be asked to setup an initial test configuration this summer, additional volunteers will be welcome

- **Simulation and analysis have difficult co-existance**

Simulation Operation

- Division of responsibility

- CMS Operation team focus on global issues:
  - CMS application: configuration, version, result
  - Validation of result
  - Bookeeping in CMS data catalogs
  - Ranks sites as yes/no based on works/does-not-work
  - No need to know details of dCache or DPM

- **Tier2 operation team focus on fabric issues:**
  - Racks, network switches, file servers
  - Cooling, heating, power, repairs
  - Operating system, middleware installation, SFT
  - Local failures tracking/fixing
  - No need to know details of CMS Dataset provenance tracking e.g.

- Pass all SiteFunctionalTests

- Working Computing Element
  - ➢ Good, solid, powerful batch farm
- Also provide good storage
  - ➢ Size
  - ➢ Performance
  - ➢ Services

- SRM access
- FTS channels (run on a server at some T1, but transfer is always a two end responsibility)
- Good I/O needed (those I/O bound cases)
- Good network (1Gbit/sec) capability to move data fast on WAN while data is being read/written fast locally
- **CMS nominal Tier2 (200 CPU, 200TB): not trivial to operate**

CMS services needed

- CMS Software distribution available
  - Simple, but critical
  - Central installation, but problems must be fixed locally
- Integration in CMS Data Management
  - Local (Trivial) File Catalog
  - PhEDEx (CMS's own high level dataset transfer tool, will work on top of FTS)

- Even simulation jobs will need to access CMS calibration and condition Database (how to make realistic simulation otherwise ?)
  - Need to deploy a Squid cache
  - CMS Database calls are mapped into the application to http calls
  - Squid is an open-source, widely-used, robust, high-performance, easy-to-install http cache server

- **Details on CMS services during tutorials**

Bringing it all toghether

- What are the really important things a Tier2 has to do in order to support CMS Simulation operation ?

- The answer is not:
  - ➢ Have lots of hardware
  - ➢ Even if that helps

- Next slide

What CMS Tier2 really need to do

- Sites need to be up and running and care for themselves
- They provide MC capacity as a service to remote operators and must be aware that those can not reach in and fix problems. Users can not login on your Worker nodes to run ps/dbg/kill !
  - ➢ You have to help
  - ➢ fixing SW area
  - ➢ spotting hung jobs, forwarding error logs, traces
  - ➢ Spotting/removing bad nodes
  - ➢ Monitor hardware and remove before it breaks
- **A good Tier2 is active, responsive, attentive, proactive**
  - ➢ Spot problems before the users
  - ➢ Fixe them before the users have problems
  - ➢ Study usage patters, talk to users, understand needs and plan for the future, reports problems to CMS Computing
- A Tier2 provides services to CMS globally, but also to local community, and reflects the obligations of local community versus CMS

Optional Service

- some site may host a "production operation unit" and so need to run ProductionAgent

- Or may decide to run a ProductionAgent for the convenience of local users

- Support will be available for running a ProductionAgent locally
  - **See tutorials**

Conclusion

● Tier2's are vital to CMS physics program

● Supporting CMS Simulation Production is easy:
  ➢ Just be a good grid site with a few easy additional services

● But it is also important, and difficult:
  ➢ Computers may stay up in unattended mode
  ➢ But quality of services usually not
  ➢ Continuous, proactive, intelligent care is needed

● You are the ones who will make your site productive and efficient
  ➢ No matter how fancy tools CMS develops centrally

● **A Tier2 will succeed or fail based on the dedication of the people running it**