User Analysis Support

Massimo Lamanna / CERN

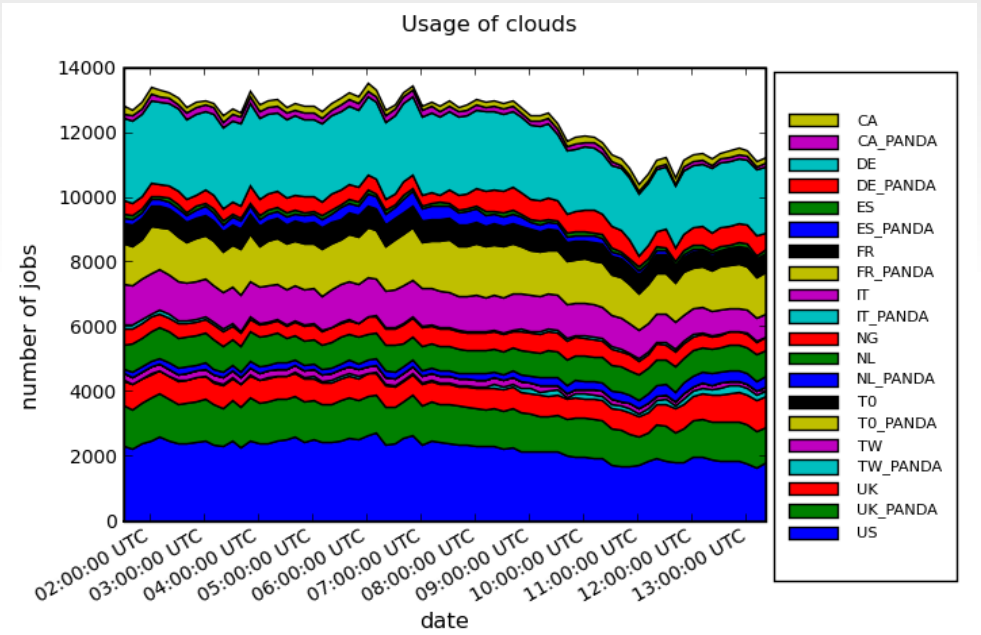
EGEE09 Conference – Barcelona
21-25 September 2009

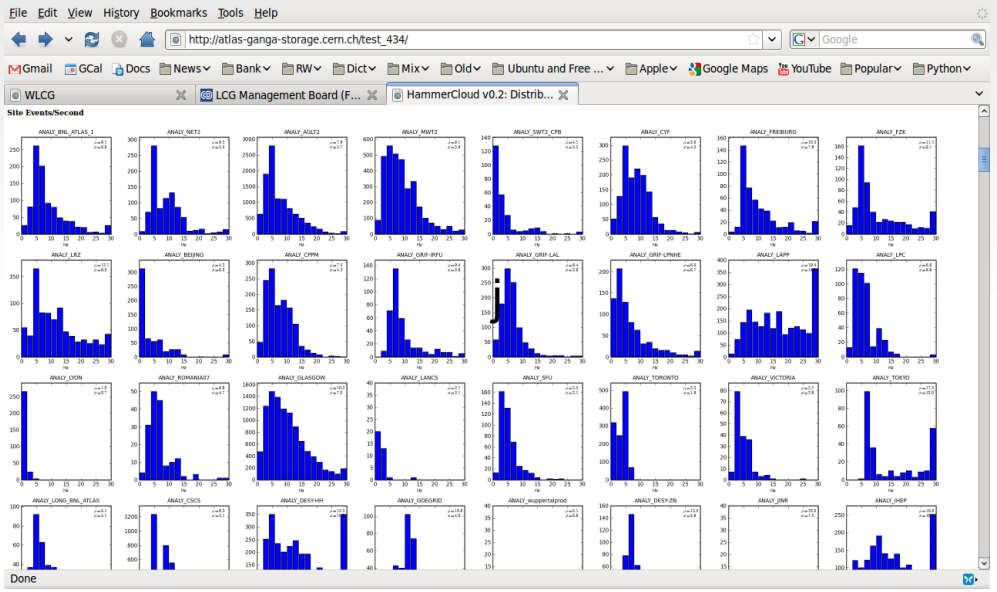
Background

- Specificity
  - Sites
  - Services
  - User
- Criticality
  - Direct impact on the applications
  - Direct impact on the grid world

User Analysis (→ Sites)

- User-intensive activities require specific site “testing”

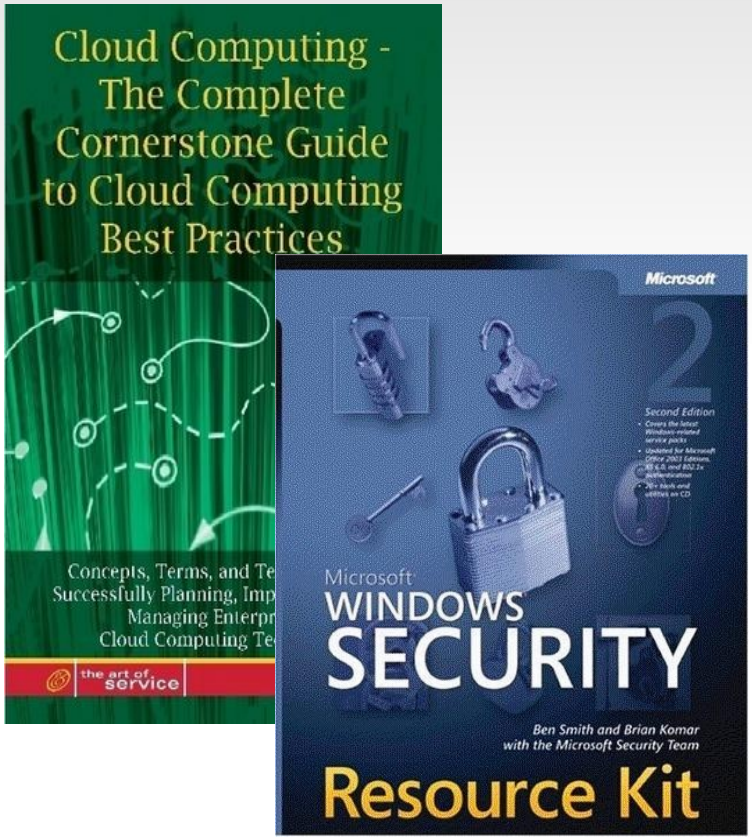




- The risk is sites won't be up to their job as good as the investment (compared to the hardware, middleware and people)

User Analysis (→ Services)

- Advanced application require application-specific services. They are critical to “get the job done”: e.g. pilot-job infrastructures. These services are **not** optional to get the job done



- The risk is to impact “production” activity
  - Stop/delay activities
  - Impact data custodial
  - Scare away resource centres
  - Our public image at stake

User Analysis (→ Users)

- Users generate load on the support teams as well. Often this is not a “clean” channel (misunderstanding in the tools, catch-all layer for **all** users programme: a site is down, the program has a (known) bug... it is not always “grid's fault”

“It's not just him. The whole system's down.”

- Untrained users generate **more** load than expert ones on the infrastructural
- Too much load on service **developers** and **operations** can ultimately put the system to halt

User Analysis (→ Applications)

- Large communities have larger resources (but many more potential problems)
  - Empower self-supporting communities

- E.g. ATLAS DAST (Distr. Analysis Shift Team):
  - 15h a day shifts (thanx to users in different time zone)
  - **Require a consistent effort to select/adapt tools**
  - Reuse then is simple

User Analysis (→ Grid world)

- We need happy and successful user communities
- Losing users is much easier than attracting them for the **first** time...
  - To get them back after they step out is basically impossible!

Recap

- Specificity
  - Sites
    - Load generators and appl. Specific testing
  - Services
    - Best practices (development, operations, audit)
  - User
    - Scalable model. Rely on self-sustained communities. Reuse existing components (make it easier to share as well)
- Criticality
  - Direct impact on the applications
    - Application success is the reason of grid computing
  - Direct impact on the grid world
    - Make our users happy (or die tryin')!

Immediate plans

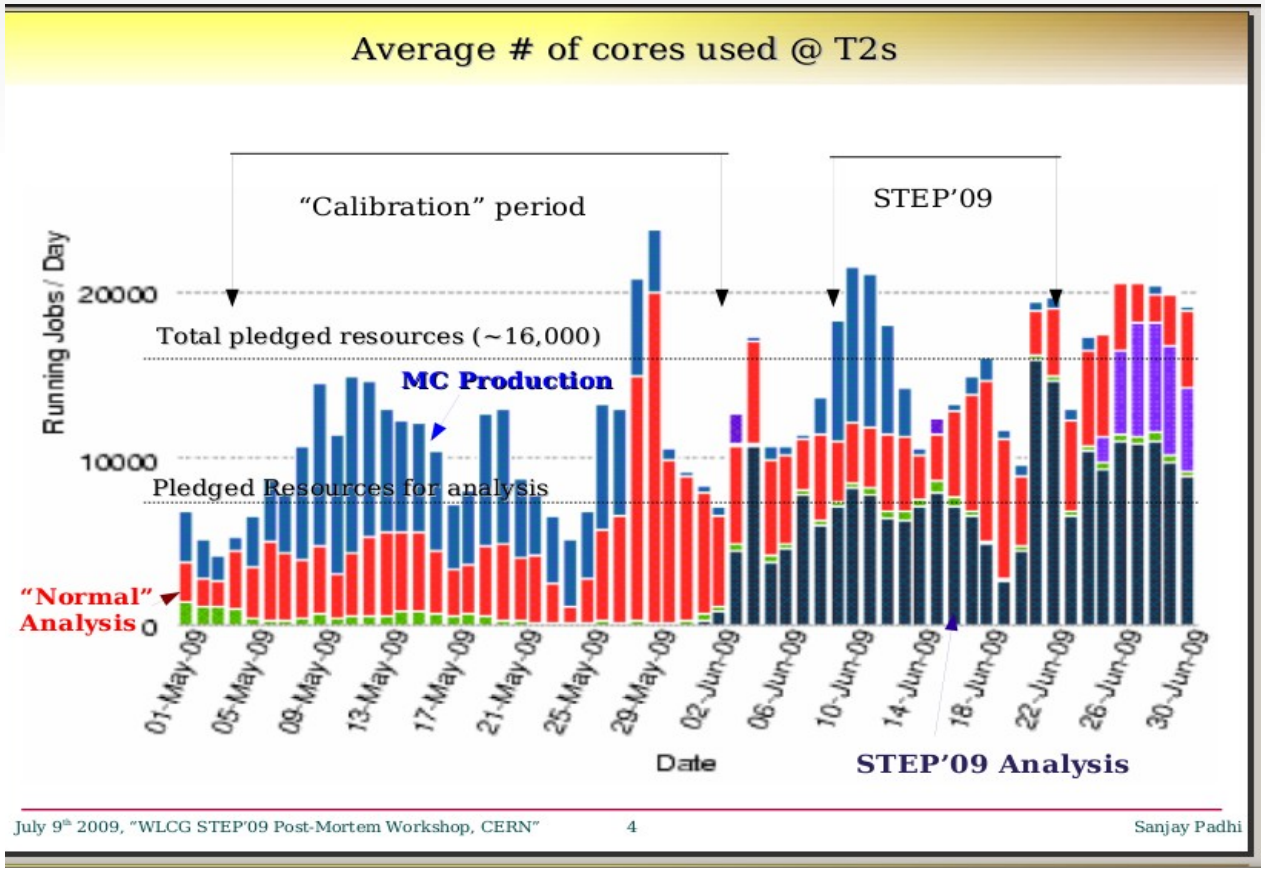
- All experiments are getting ready
  - CMS has a 2-week analysis jamboree starting October the 5th
  - ATLAS plans an analysis in October (second half)
- STEP09 was very encouraging
  - Robot activities to top-up user analysis
    - Very realistics
    - But we want more humans :)

Real Users on the Grid

- Why we do not want **only** robot activities
  - We do not test the full **diversity** of analysis jobs
    - Although good “portfolios” of users applications is available in the robot system
  - We do not reach the **all the users** who will be agressively using the system at turn on
  - We do not stress the **user support**

Example of October jamborees (CMS)

- Similar envelope of STEP09
  - User++, applications++

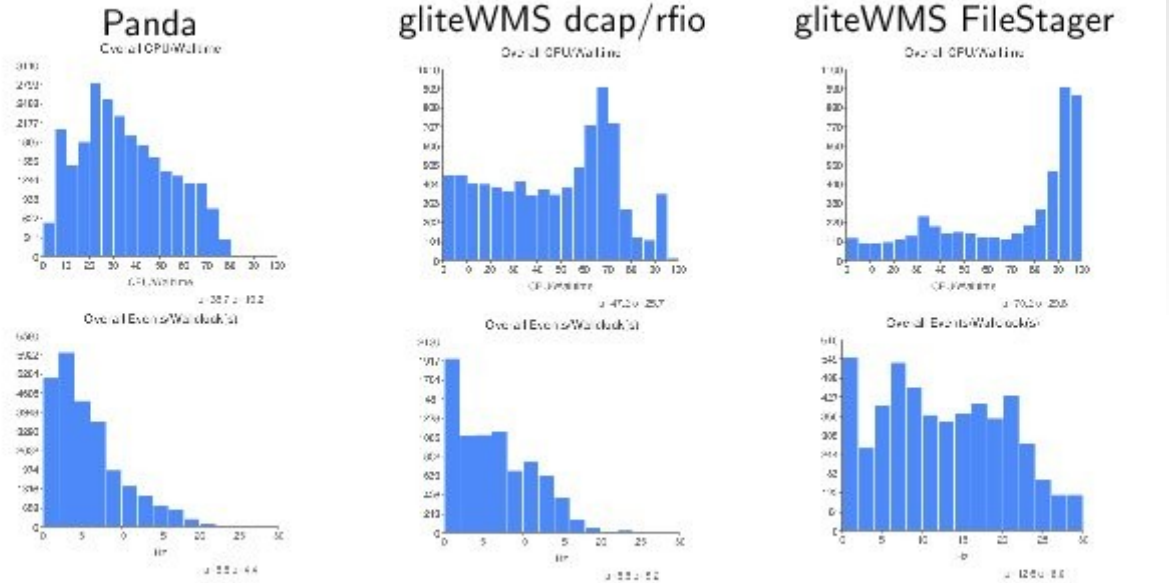


Example of October jamborees (ATLAS)

- In each cloud (10 clouds)
  - 200 M events distributed on the cloud T2
  - o(10) expert "job submitters"
  - Top up with HammerCloud
  - Very large number (maybe ~100?) of "data downloaders" to use dq2-get (Grid → Laptop)

User support

Recent DE cloud HC tests: 573, 574, 575



	Jobs compl.	Files	Events	Comment
Panda	25127	78611	717m	24h test
dcap	8054	39018	354m	Problems with HC balancing
FileStager	5167	30039	277m	only 13h test (crash)

More in the Users on the Grid (Thursday) – A. Maier

Recruitment + training + credits

- EU crew is relatively well staffed after recent additions
  - However more people are always welcome
- US has a number of new and also potential recruits
  - EVO-based training day(s) will be held in October, organized by Nurcan Ozturk
- DAST/support credit for analysis site admins?
  - Admins of popular analysis sites feel they are on permanent support shifts, esp. because of DDM-related questions. Discuss with ADC if this could qualify as ATLAS service work?
- Credit for DAST is being clarified this week
  - Many potential recruits ask about this – knowing it will help attract people
  - Latest from Kaushik: DAST is service work, recieves full credit for this (8 hours/day, 5 days, per week on shift). However DAST is not a Cat.1 activity (critical activities at Point1).

E-Learning

- E-learning is important for areas where users have to learn all details of the tasks (such as physics analysis toolkit).
- The motivations:
  - Learn to do good quality analysis code using PAT (Users point of view)
  - Get feedback and train expert (Developers point of view)

Crab Feedback Forum

- HN forum, present since years
- User are generally satisfied to use it (from “ASTF crab survey”)
- CRAB users are encouraged to subscribe
- Often other users jump in to help
- Sometimes it is used to broadcast, in an efficient way, messages to users

Outlook (1/2)

- More tests (and associated training and coordination)
  - "Hit the road running"™ (at LHC turn on)
  - Essential for an effective detector commissioning
    - Burn in, preliminary calibrations, detector "understanding"
  - **Empower the users**
    - **Support our tools (CRAB, Ganga, …)**
    - **Support the users (Dashboard, ...)**

Outlook (2/2)

- More tools
  - “Acid test” still missing
  - Collaborations opportunities
    - “Shifter tools”
    - No development, integrate existing (in use or from established sources e.g. Google tools)
      - Now: GoogleCalendar, varios ticketing system, etc...
      - In perspective: GoogleWave etc...
  - **Empower the support team (and protect the developers)**