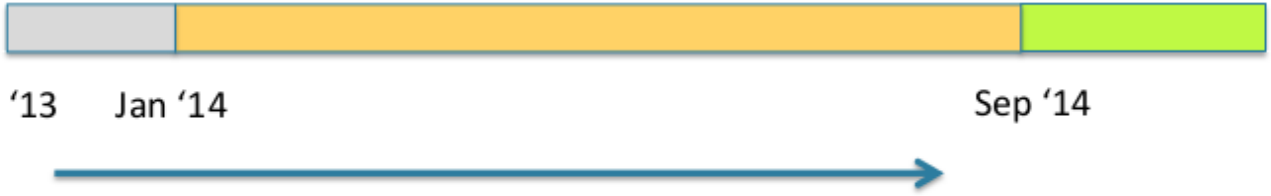
1.8.9

- DPM 1.8.9 has just been released
  - i/o monitoring “a la xrootd”
  - New logging capabilities
  - Performance improvements
  - End of YAIM support, full puppet support
  - Infosys improvements
  - MySQL/memcache improvements
  - ARGUS support re-enabled

DPM Configuration evolution (as presented during the previous Workshop)

YAIM

puppet labs



GridFTP updates

- Latest releases support a redirecting mode of operation
  - Previously, data was tunnelled within the system
  - Large potential performance benefits for gridftp-only transfers
  - Requires activation via configuration
    - And monitoring of client activity
    - Old clients can be inefficient – please move to gfal2!!
- Presentation focused on motivating some “early adopters” to validate the system

Dynamic Federations

- A project started a few years ago
- Goal: a frontend that presents what a certain number of endpoints would present together
  - Without indexing them beforehand
- These endpoints can be a very broad range of objects that act as data or metadata stores
  - We prefer to use HTTP/WebDAV things, yet that's not a constraint

HTTP Federation

starting PoC with **DPM** sites

Group by dst/clients that accessed *cern.ch*

Why HTTP/DAV?

- It's there, whatever platform we consider
  - A very widely adopted technology

- We (humans) like browsers, they give an experience of simplicity
- Mainstream and sophisticated clients: curl, wget, Davix, …
- ROOT works out of the box with HTTP access (LCG release >= 69)

- Goes towards convergence
  - Users can use their devices to access their data easily, out of the box
  - Web applications development can meet Grid computing
  - Jobs and users just access data directly, in the same way
  - Can more easily be connected to commercial systems and apps

Conclusions

- A r/o R&D prototype that exceeded expectations
  - 13 sites out of 19, the others are coming
  - Official site downtimes were always automatically detected so far

- Cleanness of LHCb repos helped
- Please evaluate it and help us improve
  - This is likely to be an actor of a next evolution in *large scale DM,* HEP meeting the Web through proper tools
- New features are coming. Smarter site detection, write support, logging, monitoring, …
  - Next RC of the Dynafeds expected in a couple of weeks
- High flexibility/scalability of the concept, able to deal with a broad range of endpoints
- Can be made to work with WebFTS to find the “right” sources
  - Also endpoint prioritization is pluggable

- Looking at exploiting the potential of mixing S3 storage with other techs
  - We are contributing to a r/w prototype for BOINC (See preGDB by Laurence Field)
- We are cooking an AGIS-based RUCIO-friendly prototype (>40 sites, >200 spacetokens)

Cloud collaboration

- EGI cloud federation in production → To be evolved into a use-case-driven cloud federation
  - Starting from the use cases to define the federation capabilities
  - Integration with the Operational tools
  - Security coordination of cloud service
  - Evolution of the AAI mechanisms
- Integration of cloud services with the EGI's operational fabric
  - Monitoring
  - Accounting
  - Support for X509 (with OCCI)
  - Integration in GOCDB and BDII
- This integration can be beneficial in particular for user communities who are already using the EGI infrastructure
- Discuss common use cases/requirements to leverage on the work done by both EGI and WLCG
  - E.g. data services (preservation, open data..), cloud broker..

Conclusions

- EGI services will continue to be provided and to evolve to fulfill the requirements of the users communities
- Collaboration opportunities for WLCG and EGI
  - Operations activities
  - Operations tools
  - Cloud
  - Other?
- Dedicated workshop with WLCG?

Task Force (TF) news (1/2)

- gLexec deployment
  - Integration in PanDA completed
    - Each site can be independently configured whether to use gLexec never/always/when it works
    - Ongoing deployment in ATLAS, then will study effects on success rates and based on it decide on the migration
- SHA-2 migration
  - OSG added the new VOMS servers to their vomses files yesterday
  - ALICE is already happily using them, the other experiments still need to conclude their tests

Working Group (WG) news

- Middleware Readiness
  - ATLAS requested the WG to include HTCondor
  - The Tier-0 will contribute for EOS and FTS3
  - DPM 1.8.9 being tested at the *Volunteer Sites*
  - Setting up for dCache at TRIUMF, NDGF and PIC
  - Setting up for StoRM at CNAF and QMUL
  - Developing a “MW readiness dashboard”
  - Next meeting November 19th at 4pm CET with Vidyo
- Network and transfer metrics
  - pS 3.4 patched to address the POODLE vulnerability
  - All sites must upgrade to 3.4+ according to broadcasted instructions
  - Validating the Datastore
  - Very soon the metrics will be usable for production and instructions made available to troubleshoot network problems

TF news (2/2)

- Machine/job features
  - Converged on a single implementation for Cloud infrastructures, to be proposed to WLCG
  - TF conclusion expected by the next GDB
- IPv6 deployment
  - Discussed at HEPix, agreed to strengthen the collaboration with OSG to avoid effort duplication
  - Recommendation to install perfSONAR in dual-stack from version 3.4
- Squid monitoring and HTTP proxy discovery
  - Campaign this month to register all squid instances in GOCDB and OIM (instructions)
    - It is needed to know which squid servers need to be monitored

Virtual Memory

- Many sites limit vmem because they want to limit RSS+swap
  - Kernels have changed years ago and vmem doesn't mean RSS+swap anymore it's the size of the address space
    - SCORE 32bit vmem-RSS+swap was still negligible in first approximation
    - 64bit address space much larger difference will increase
- Standard tools do not report the memory correctly anymore nor are able to limit RSS+swap
  - Processes may look like they are using 40GB of vmem but if one looks at RSS+swap with other tools the same processes don't go above 20GB
  - ulimit used to be able to distinguish for example it could limit RLIMIT_RSS now it limits only RLIMIT_AS which affects all memory allocation and mapping functions

Memory multicore case

- To the previous slide we need to add that multicore (v)memory is wrong by default because the shared memory is accounted multiple times.
  - Even without counting the experiments asking for more to cover the 5 minutes peaks
- Some sites limiting the (v)memory had to increase the limit
  - Problem when limit = allocation of resources
- Some sites are oversubscribing the memory by a factor
  - Useful particularly for multicore when most of the time the memory is not used.
  - Recipes for maui and HTcondor exist

Memory and cgroups

- Some sites are enabling cgroups.
  - Allows more accurate monitoring (see plots next slide)
  - Allows smart soft limit without allocating memory
    - If jobs exceed this the kernel pushes them back to a smaller value
  - Allows hard limit job gets killed

cgroups and BS

- Can it work everywhere?
  - Really easy to enable in Htcondor
  - Supported in SLURM
  - UGE has been patched
  - SoGE/OGE no support
    - Most GE sites use this I think
  - torque/maui no support
    - At last count still 100 sites
- Sites moving away from torque should look into it though
  - HTCondor recipe really easy
  - SLURM probably easy too

HEP Software Foundation: Progress so Far

Pere Mato/CERN on behalf of the HSF Startup Team
GDB Meeting, 12 November 2014

Motivation

- Much of our HEP software is now old (> 20 years)
  - it needs to be adapted to more modern standards
- Paradigm-shift resulting from the evolution of CPU architectures
  - our code has to be re-engineered to make use of the full capabilities
- Make use of all resources available to our community
  - HPC facilities, commercial clouds, volunteer resources
- Must attract people with the required advanced skills and experience
- Ensure interoperability with software developed by other scientific communities
- Opportunity for sharing software between different experimental programs

Summary

- Communication tools
  - Foundation Web: http://hepsoftwarefoundation.org
  - Forum: https://groups.google.com/forum/#!forum/hep-sf-forum
  - Mailing lists:
    - hep-sf-forum@googlegroups.com please sign up at to participate in the HSF
    - hep-sw-comp@googlegroups.com more general list for HEP SW & COMP
- Began consultations and community engagement
- White Paper synthesis document available
- Preparation of the SLAC workshop
  - Preliminary announcement at http://hepsoftwarefoundation.org/workshop-slac-jan-2015