CHEP 2009

An extremely biased personal view

Dario Barberis

CERN & Genoa University/INFN

Outline

This is not a summary of the summaries

Anyway most of you were listening to many talks during the whole week

And the rest of the people were certainly looking at the slides on their laptops…

… unless they spent the whole time working out which session to attend next!

Computing in High Energy and Nuclear Physics
Prague | Czech Republic | 21 - 27 March 2009

Some CHEP'09 Statistics (1)

- ~600 participants for CHEP
  - ~240 people for the WLCG Workshop (almost complete overlap)
  - Personal bias: I counted O(100) ATLAS members in the participants list
    - about 50% of people contributing actively to Software & Computing
- Programme:
  - 7 tracks
  - 29 plenary and summary talks
  - ~100 parallel talks
  - ~300 posters
  - Lots of discussions
- Imagine if everyone writes 10 pages for each parallel talk or poster, and 20 pages for each plenary talk…
  - 4600 pages of proceedings!
    - Hopefully nobody will ever try to print the whole lot

Computing in High Energy and Nuclear Physics
Prague | Czech Republic | 21 - 27 March 2009

Some CHEP'09 Statistics (2)

The 100 most common words in abstracts (titles and contents):

access (191) alice (95) analysis (417) applications (251) architecture (108) atlas (441) available (109)
based (194) beam (84) build (83) calibration (100) centers (99) cern (317) cluster (93) cms (384) code (103) collaboration (105) commissioning (87)
community (88) components (105) computing (452) configuration (110) control (108) data (1291) database (150)
describe (143) design (90) detector (233) developed (323) different (207) distributed (273)
efficient (134) energy (111) environment (130) event (183) experiment (401) framework (205) generic (86)
grid (567) implementation (88) improvements (92) information (221) infrastructure (207) interface (166) job (153)
level (159) lhc (336) lhcb (92) local (81) management (267) middleware (85) model (198) monitoring (294)
network (167) number (98) offline (90) online (100) operations (166) order (115) packages (81) paper (129)
performance (332) physics (288) possible (89) present (256) problems (98) processing (274)
production (207) project (109) provides (236) quality (89) reconstruction (173) requirements (91) resources (244)
results (140) root (101) running (107) scalability (85) services (402) simulation (120) sites (146) slides (191)
software (384) solution (87) status (112) storage (213) support (136) system (791) tools (201) track (88)
transfer (88) trigger (133) usage (84) user (209) validation (85) virtual (86) web (133) wlcg (87) work (109) years (108)

Computing in High Energy and Nuclear Physics
Prague | Czech Republic | 21 - 27 March 2009

Starting from the top: "Data"

The 100 most common words in abstracts (titles and contents):

access (191) alice (95) analysis (417) applications (251) architecture (108) atlas (441) available (109)
based (194) beam (84) build (83) calibration (100) centers (99) cern (317) cluster (93) cms (384) code (103) collaboration (105) commissioning (87)
community (88) components (105) computing (452) configuration (110) control (108) data (1291) database (150)
describe (143) design (90) detector (233) developed (323) different (207) distributed (273)
efficient (134) energy (111) environment (130) event (183) experiment (401) framework (205) generic (86)
grid (567) implementation (88) improvements (92) information (221) infrastructure (207) interface (166) job (153)
level (159) lhc (336) lhcb (92) local (81) management (267) middleware (85) model (198) monitoring (294)
network (167) number (98) offline (90) online (100) operations (166) order (115) packages (81) paper (129)
performance (332) physics (288) possible (89) present (256) problems (98) processing (274)
production (207) project (109) provides (236) quality (89) reconstruction (173) requirements (91) resources (244)
results (140) root (101) running (107) scalability (85) services (402) simulation (120) sites (146) slides (191)
software (384) solution (87) status (112) storage (213) support (136) system (791) tools (201) track (88)
transfer (88) trigger (133) usage (84) user (209) validation (85) virtual (86) web (133) wlcg (87) work (109) years (108)

Computing in High Energy and Nuclear Physics
Prague | Czech Republic | 21 - 27 March 2009

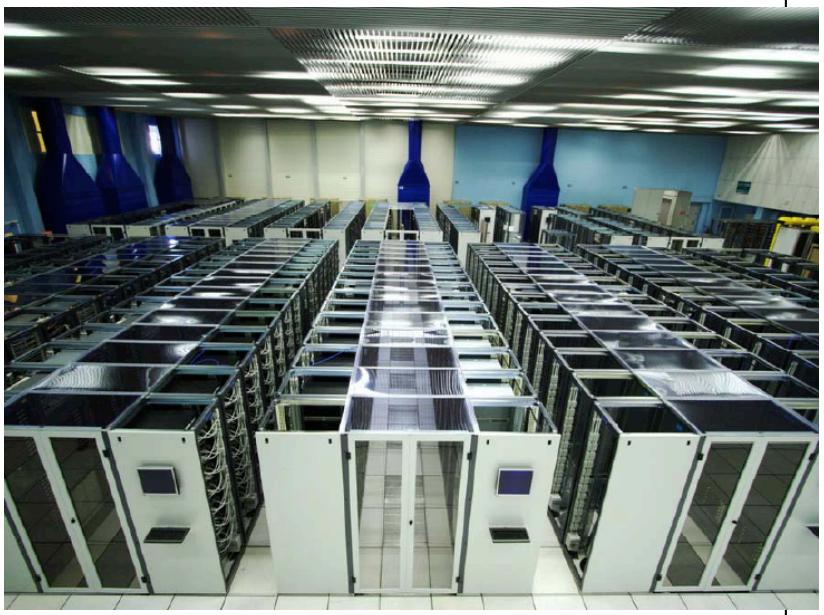
Starting from the top: "Data"

The 100 most common words in abstracts (titles and contents):

access (191) alice (95) analysis (417) applications (251) architecture (108) atlas (441) available (109)
based (194) beam (84) build (83) calibration (100) centers (99) cern (317) cluster (93) cms (384) code (103) collaboration (105) commissioning (87)
community (88) components (105) computing (452) configuration (110) control (108) data (1291) database (150)
describe (143) design (90) detector (233) developed (323) different (207) distributed (273)
efficient (134) energy (111) environment (130) event (183) experiment (401) framework (205) generic (86)
grid (567) implementation (88) improvements (92) information (221) infrastructure (207) interface (166) job (153)
level (159) lhc (336) lhcb (92) local (81) management (267) middleware (85) model (198) monitoring (294)
network (167) number (98) offline (90) online (100) operations (166) order (115) packages (81) paper (129)
performance (332) physics (288) possible (89) present (256) problems (98) processing (274)
production (207) project (109) provides (236) quality (89) reconstruction (173) requirements (91) resources (244)
results (140) root (101) running (107) scalability (85) services (402) simulation (120) sites (146) slides (191)
software (384) solution (87) status (112) storage (213) support (136) system (791) tools (201) track (88)
transfer (88) trigger (133) usage (84) user (209) validation (85) virtual (86) web (133) wlcg (87) work (109) years (108)

Data Management, Access, Analysis,…

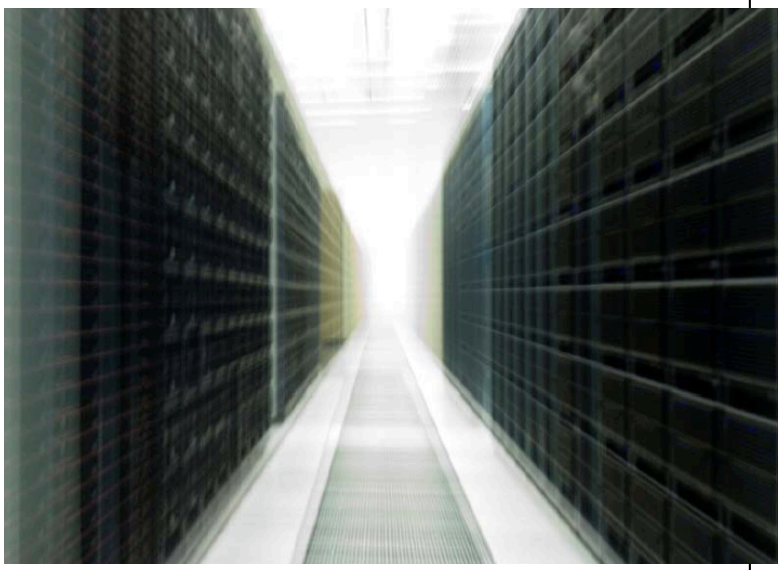
- The major problem of LHC (and RHIC and all other) experiments
- Some LHC experiments tried at first (2003-2004) to rely on Grid middleware developments of data management tools (storage elements, file catalogues)
- Then (2005) the WLCG project defined "baseline services" — mostly related to data management tools
  - These are effectively the tools we use now to access and move data between storage elements
    - SRM, FTS, LFC and various incarnations of monitoring tools



- We have shown we can use all these tools to run large-scale organized productions (simulation, real data processing and reprocessing)
  - Albeit still with too much human interventions when services temporarily fail or sites go down for scheduled or unscheduled interventions

Users want data!

- It was pointed out several times that most physicists don't want to take a 3-day tutorial to be able to access an ntuple and produce a histogram
  - Although we all provide reasonably up-to-date documentation and tutorials to our collaborators
- The problem is intrinsically complex
  - If we want to hide the complexity, our tools MUST NEVER FAIL
    - Similarly to components of an operating system
- Right now people have to learn about catalogues, storage elements, transfer tools, site downtimes and all that



- We are not at all certain that we'll be able to cope with massive user access to "hot" data
- Recent developments go towards having an extensive use of XROOT as data access protocol
  - It helps a lot when it comes to I/O bound (analysis) jobs
- Still, a directly monted file system is a more natural solution for a Unix user

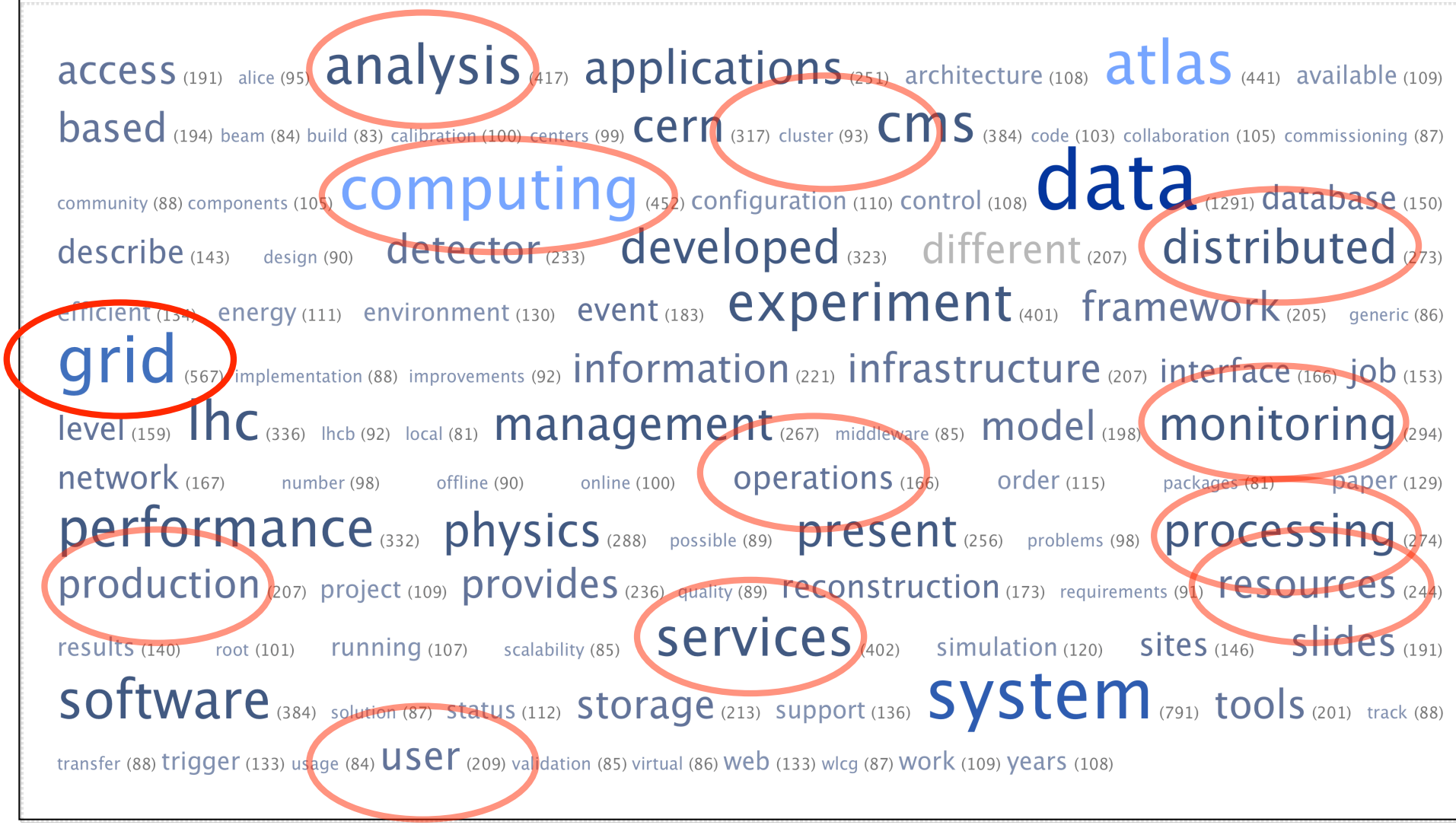
Next: "Grid"

The 100 most common words in abstracts (titles and contents):

access (191) alice (95) analysis (417) applications (251) architecture (108) atlas (441) available (109)
based (194) beam (84) build (83) calibration (100) centers (99) cern (317) cluster (93) cms (384) code (103) collaboration (105) commissioning (87)
community (88) components (105) computing (452) configuration (110) control (108) data (1291) database (150)
describe (143) design (90) detector (233) developed (323) different (207) distributed (273)
efficient (134) energy (111) environment (130) event (183) experiment (401) framework (205) generic (86)
grid (567) implementation (88) improvements (92) information (221) infrastructure (207) interface (166) job (153)
level (159) lhc (336) lhcb (92) local (81) management (267) middleware (85) model (198) monitoring (294)
network (167) number (98) offline (90) online (100) operations (166) order (115) packages (81) paper (129)
performance (332) physics (288) possible (89) present (256) problems (98) processing (274)
production (207) project (109) provides (236) quality (89) reconstruction (173) requirements (91) resources (244)
results (140) root (101) running (107) scalability (85) services (402) simulation (120) sites (146) slides (191)
software (384) solution (87) status (112) storage (213) support (136) system (791) tools (201) track (88)
transfer (88) trigger (133) usage (84) user (209) validation (85) virtual (86) web (133) wlcg (87) work (109) years (108)

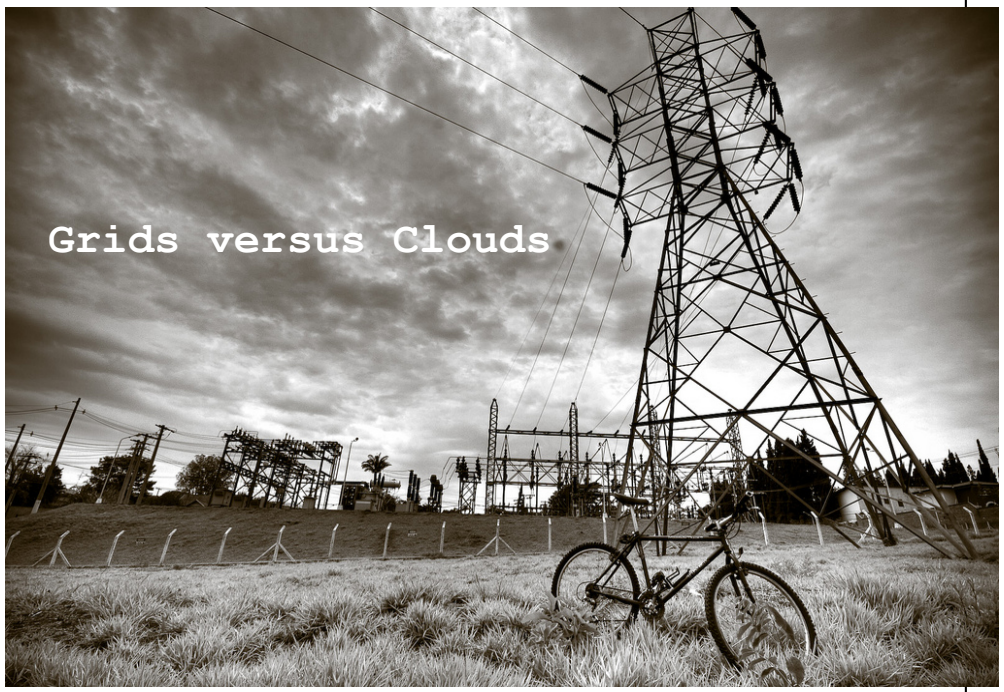
Next: "Grid"

The 100 most common words in abstracts (titles and contents):



Grids vs Clouds?

- We had this week the first reports of serious usage of cloud computing
  - Both in plenary and parallel sessions (and posters)
- It is an interesting concept but I am not sure it makes financial sense for data-intensive applications
  - CPU power may be cheap but data storage and transfer is very expensive
  - It may be an option if/when we are short of simulation power
- On the other hand, commercial companies are evidently able to provide the software environment the user needs, without many of the existing restrictions of the Grid



Computing in High Energy and Nuclear Physics
Prague | Czech Republic | 21 - 27 March 2009

Grids AND Clouds!

- Virtualization is the keyword here!
- Can't we just do the same for our Grid sites?
  - Install "cloud middleware"...
  - And keep control of the resources and the data management tools
- All we need after all is a simple and reliable way to send jobs to where they can run fastest and most reliably
  - Virtual machines running user jobs can shield from local setup details
  - An existing implementatio of Grid middleware (ARC) is already very close to these needs
- I know some of this may sound like heresy to some in this room
- On the first day we had a long list of people who got into trouble after lecturing in Prague!

Dar

Computing in High Energy and Nuclear Physics
Prague | Czech Republic | 21 - 27 March 2009

Last (but not least): "Performance"

The 100 most common words in abstracts (titles and contents):

access (191) alice (95) analysis (417) applications (251) architecture (108) atlas (441) available (109)
based (194) beam (84) build (83) calibration (100) centers (99) cern (317) cluster (93) cms (384) code (103) collaboration (105) commissioning (87)
community (88) components (105) computing (452) configuration (110) control (108) data (1291) database (150)
describe (143) design (90) detector (233) developed (323) different (207) distributed (273)
efficient (134) energy (111) environment (130) event (183) experiment (401) framework (205) generic (86)
grid (567) implementation (88) improvements (92) information (221) infrastructure (207) interface (166) job (153)
level (159) lhc (336) lhcb (92) local (81) management (267) middleware (85) model (198) monitoring (294)
network (167) number (98) offline (90) online (100) operations (166) order (115) packages (81) paper (129)
performance (332) physics (288) possible (89) present (256) problems (98) processing (274)
production (207) project (109) provides (236) quality (89) reconstruction (173) requirements (91) resources (244)
results (140) root (101) running (107) scalability (85) services (402) simulation (120) sites (146) slides (191)
software (384) solution (87) status (112) storage (213) support (136) system (791) tools (201) track (88)
transfer (88) trigger (133) usage (84) user (209) validation (85) virtual (86) web (133) wlcg (87) work (109) years (108)

Computing in High Energy and Nuclear Physics
Prague | Czech Republic | 21 - 27 March 2009

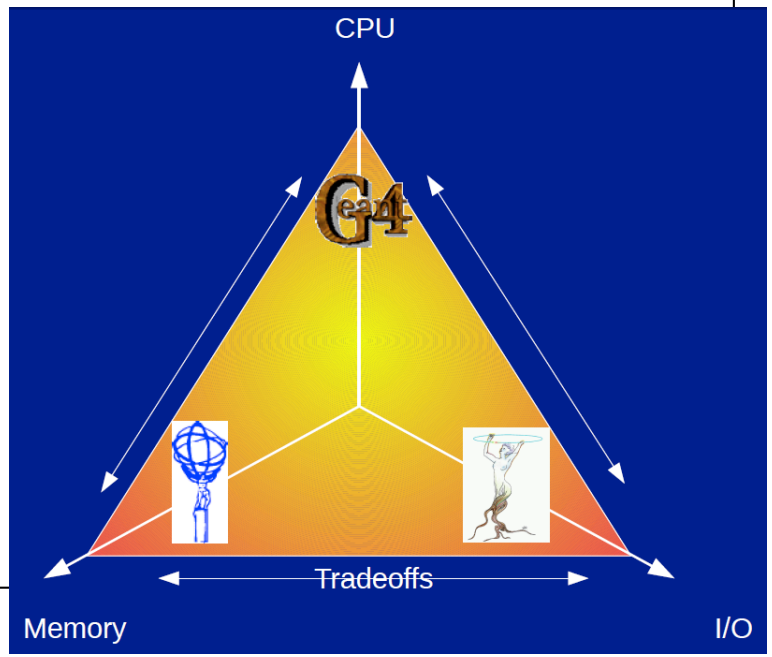
Last (but not least): "Performance"

The 100 most common words in abstracts (titles and contents):

access (191) alice (95) analysis (417) applications (251) architecture (108) atlas (441) available (109)
based (194) beam (84) build (83) calibration (100) centers (99) cern (317) cluster (93) cms (384) code (103) collaboration (105) commissioning (87)
community (88) components (105) computing (452) configuration (110) control (108) data (1291) database (150)
describe (143) design (90) detector (233) developed (323) different (207) distributed (273)
efficient (134) energy (111) environment (130) event (183) experiment (401) framework (205) generic (86)
grid (567) implementation (88) improvements (92) information (221) infrastructure (207) interface (166) job (153)
level (159) lhc (336) lhcb (92) local (81) management (267) middleware (85) model (198) monitoring (294)
network (167) number (98) offline (90) online (100) operations (166) order (115) packages (81) paper (129)
performance (332) physics (288) possible (89) present (256) problems (98) processing (274)
production (207) project (109) provides (236) quality (89) reconstruction (173) requirements (91) resources (244)
results (140) root (101) running (107) scalability (85) services (402) simulation (120) sites (146) slides (191)
software (384) solution (87) status (112) storage (213) support (136) system (791) tools (201) track (88)
transfer (88) trigger (133) usage (84) user (209) validation (85) virtual (86) web (133) wlcg (87) work (109) years (108)

Software performance

- A large fraction of experiment software and common tools (Geant4 etc.) were designed at the end of last century
  - Development continued adiabatically for a decade
  - It is high time to go for a global clean-up and some re-thinking
- Performance was addressed extensively in plenary and parallel sessions
  - Becoming more and more important when
    - The attention shifts from simulation (CPU bound) tasks to reconstruction and analysis (I/O bound) tasks
    - Computing resources become saturated as LHC data arrive
    - Other limitations to large-scale data access are removed
- All LHC experiments invested considerable effort over the last couple of years to optimize CPU consumption, memory usage and event sizes



Hardware trends

- One stumbling point until now: full use of 64-bit architectures saves some 10% of CPU power but almost doubles the memory footprint
  - What is the trade-off in terms of €/$/£/¥/CHF?
  - Or is the solution in the use of a multi-threaded application?
    - How would that fit on the Grid?
    - Or in a virtual machine that runs on the Grid?
  - Good news at this CHEP (from CMS): VMEM increase can be reduced with custom linker script
- Multi-core 64-bit processors are nevertheless yesterday's technology.
  - We must think seriously of tomorrow's technology: many-core processors
- One possibility is to design
  - Parallel software that can occupy any number of cores in a machine
  - Virtual machines that can run in many-core processors
  - Grid middleware that allows the submission of this new kind of jobs
- These components must work simply and reliably together

Neglected but Important Topics

Networks

- They generally work!
  - Long-distance network capability is far ahead of our needs
  - End-to-end bandwidth is instead still an issue
    - Digital divide is broadening
    - "Last mile" problems persist in large European and North American labs
  - What the user sees is the total transfer time of her dataset, not the performance of high-speed transoceanic optical links

Collaborative Tools

- Many of them work reasonably now
  - Reports in plenary and parallel sessions
  - The habit of using more than just the telephone has not spread yet throughout the collaborations
    - Apart perhaps from the computing communities

Conclusions

- Keywords:
  - Data management and data access
  - Performance
  - Clouds
  - Virtualization
- Our "community" tasks till CHEP 2010:
  - Consolidate the current software and computing infrastructure and make sure it will work for the first LHC run
    - This may need more manpower than we would have liked, but I am confident we have learned to use what we have
  - Improve continuously the performance
  - Look ahead and re-think some critical parts of our models and of the components we use
    - So as to have the possibility to phase them in as soon as they perform better then what we have now

Computing in High Energy and Nuclear Physics
Prague | Czech Republic | 21 - 27 March 2009

Thanks!

- Thanks to all of you who contributed ideas (or more) for this talk
- Thanks to Milos Lokajiček & IAC for inviting me

Thanks to Milos and all Prague people and institutes behind the success of CHEP'09!

See you next year