

22nd International Conference on Computing in High Energy and Nuclear Physics, Hosted by SLAC and LBNL, Fall 2016

Un retour

Catherine Biscarat (LPSC), Michel Jouvin (LAL), Fabian Lambert (LPSC)

Avec l'aide de Luisa Arrabito (LUPM), Jérôme Fulachier (LPSC),
Sébastien Gadrat (CC-IN2P3), Ghita Rahal (CC-IN2P3)

Survue

- 570 contributions / 36 présentations plénières / 239 posters
- 512 personnes enregistrées.
- Participation française : 12 participants

Luisa Arrabito (LUPM), Catherine Biscarat (LPSC), Jérôme Fulachier (LPSC), Julien Houles (CPPM), Sébastien Gadrat (CC-IN2P3), Julius Hrivnac (LAL), Ivana Hrivnacova (IPNO), Michel Jouvin (LAL), Miguel Rubio-Roy (LLR), Fabian Lambert (LPSC), Ghita Rahal (CC-IN2P3), Grigory Rybkin (LAL)

- 4,5 journées de présentations, dont :
 - 2,75 jours de pléniers
 - 1,75 jours avec 7 sessions en parallèles → 12,25 jours de présentations.

Nous ne serons pas exhaustifs !

Choix de présenter les faits « marquants » - subjectifs – ouverts à discussions.

Organisation de cette présentation

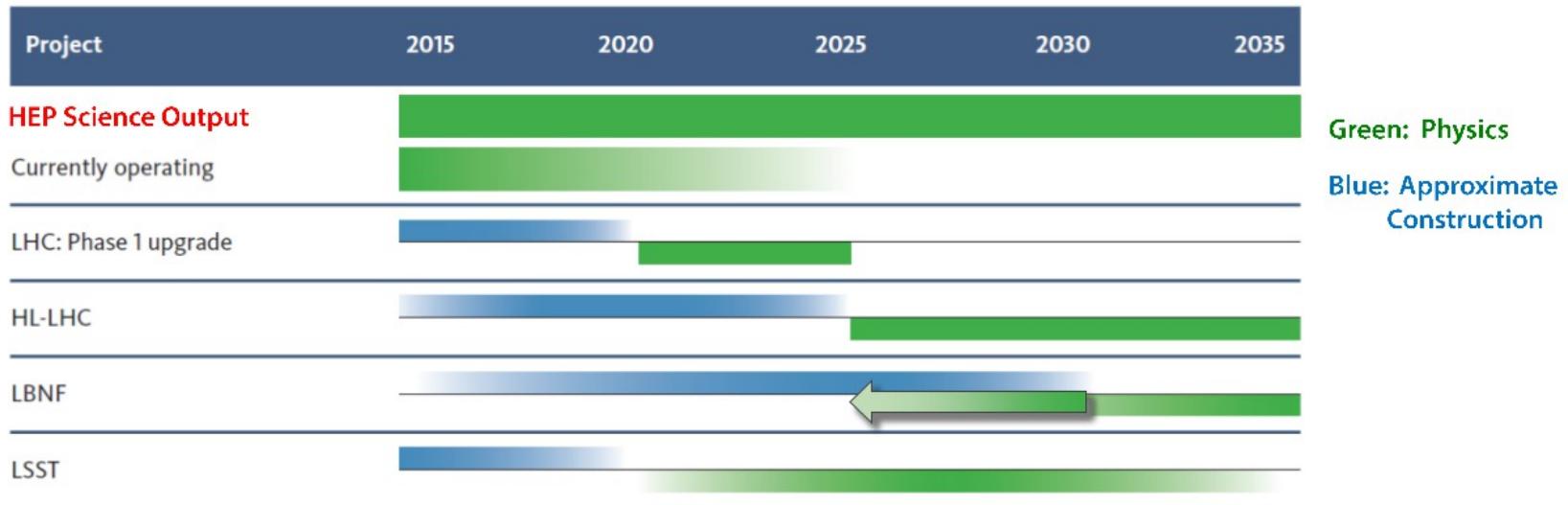
- Le contexte (Fabian)
- Le software (Fabian)
- Exploiter la performance des nouvelles architectures (Michel)
- Evolution des modèles de calcul et des sites (Catherine)

Le contexte

Science drivers of particle physics

- P5* (2011 → 2015) identified 5 science drivers
 - Use the Higgs boson as a new tool for discovery
 - Pursue the physics associated with neutrino mass
 - Identify the new physics of dark matter
 - Understand cosmic acceleration: dark energy and inflation
 - Explore the unknown: new particles, interactions, and physical principles
- Vision throughout 20-year timeframe (→ 2035)
 - Will produce results continuously
 - Realizing this vision will require a shift in approaching the networking and computing challenges

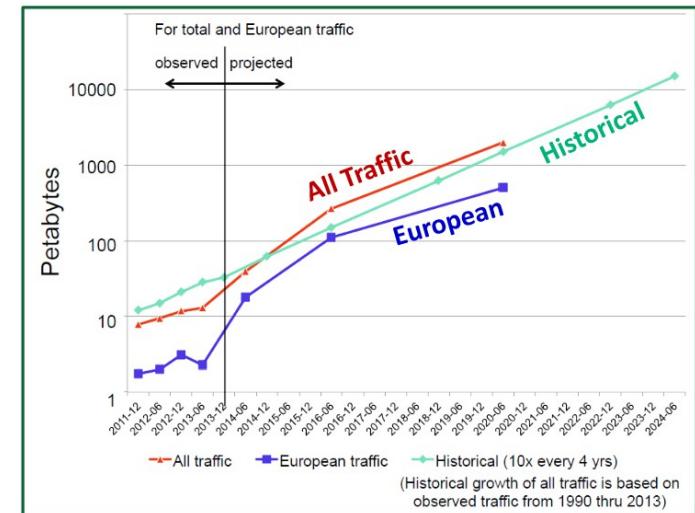
*P5 = Particle Physics Project Prioritization Panel



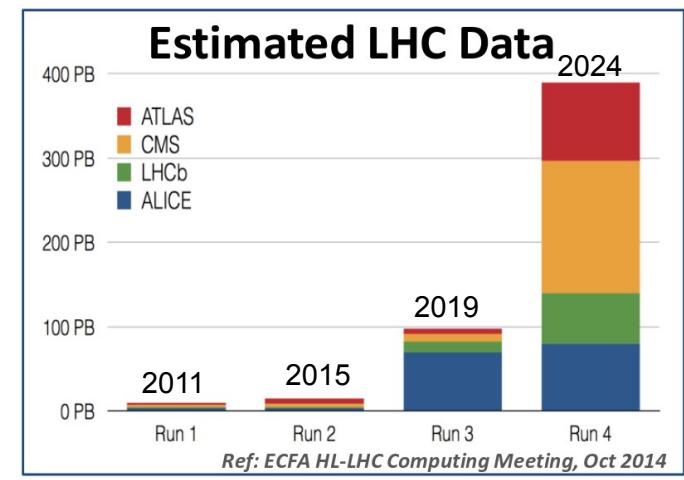
Source: DOE HEP viewpoint – Jim Siegrist – CHEP 2016

Exascale is coming !

- Future networking
 - ESNet Projected Traffic 43 PB/month in April 2016, 1 EB/month by ~2020, 10 EB/month by ~2024
 - Distributed environments where resources can be deployed flexibly to meet the demands
 - Software - Defined Networking (SDN)
- Future computing is Big Data
 - LHC Run 4 (2026 and beyond) will start the exabyte era
 - Own baseline resources that will be used at full capacity
 - Use service providers for peak cycles when needed
 - Strong emergence of data-centric computing



Evolution of ESNet traffic



Evolution of RAW data storage

Source: DOE HEP viewpoint – Jim Siegrist – CHEP 2016 & Future of Distributed Data and Workload Management – Ian Fisk – CHEP 2016

Future of HEP computing

- Adapt software
 - Not enough resources to rewrite software separately for everyone
 - Cooperate across experiments and laboratories
 - Review the existing softwares
 - Adopt big data tools
 - Spark, Hadoop for data access
 - Containers for interactivity and analysis preservation
- Optimize Hardware design
 - We can expect the technology improvements will give a factor of 6 - 10 improvements
 - But we still expect to be 10x 100x shortfall by 2025 (DOE point of view)
 - HEP community has opportunity to work with ASCR (Advanced Scientific Computing Research)
 - Optimized HW then becomes widely available and lower cost to HEP
- Explore new paradigms
 - Deep learning
 - Smart networking
 - Neuromorphic, Quantum computing

PARTIE I

CHEP 2016 - Software



Webinaire RI3
Retours CHEP 2016

Fabian Lambert – LPSC Grenoble

Plan

- Overview
- Software optimization
 - Memory Pattern
 - Economistics and ecologic methods
- Best practices
 - Mice Analysis User Software (MAUS)
 - Software Quality Control at Belle II
- Visualisation
 - A browser-based event display for the CMS Experiment at the LHC
 - Everware toolkit
- Big Data
 - Keywords

Overview

- Contributions about software are mainly in Track 5
 - But not only...
- Track 5 charges
 - Physics applications and machine learning
 - ROOT evolution
 - **Performance, analytics***, etc...
 - Processors and architectures
 - **Software development***
 - **Visualization***
 - **Big Data***



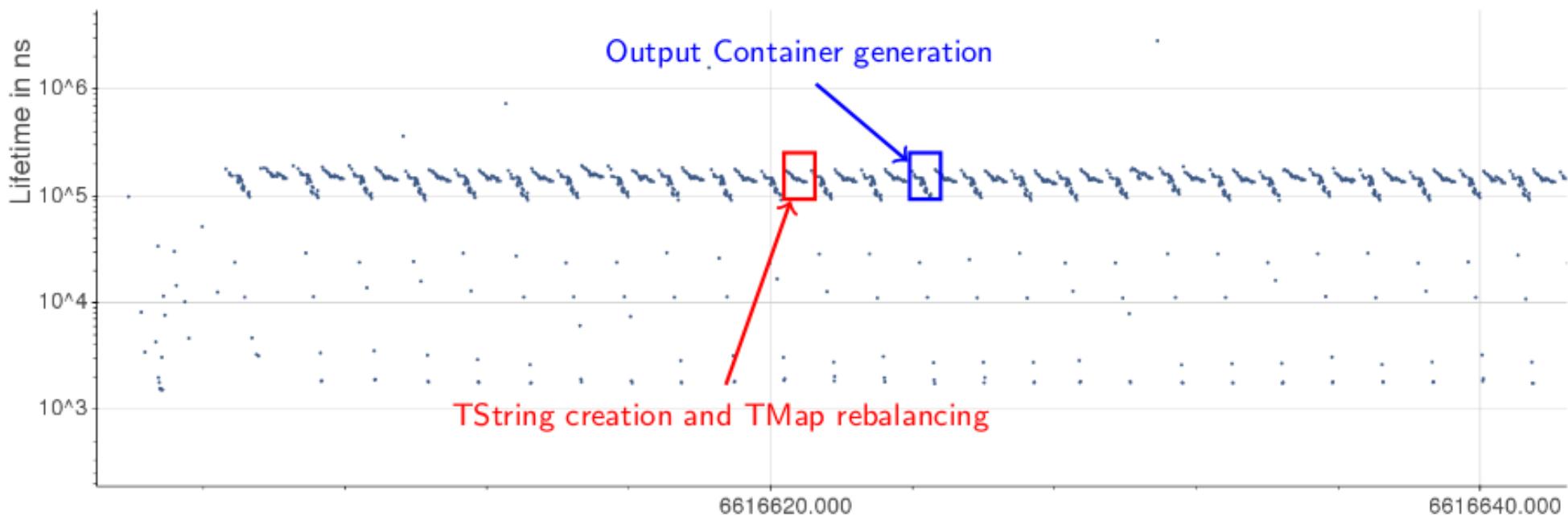
* my attended talk categories

Software Optimization

Software Optimization – Memory patterns

- Identifying Memory Allocation Patterns in HEP Software – Sami Kama et al.
- FOM-Tools provides a means to analyze memory allocation and utilization
 - <https://gitlab.cern.ch/fom/FOM-tools.git>
- Understand how and why we allocate/deallocate memory
 - HEP Applications usually consume large amounts of memory
 - Cost of memory per core is increasing
 - Significant amount of this memory is due to temporary allocations
 - Memory layout and access patterns are keys to performance

Software Optimization - Memory patterns



An Interesting Pattern in Lifetime Plots

Software Optimization - Memory patterns

```
//In pseudo code
Root::TAccept selectionPassed;
for (itr in tracks) {
    if (some criteria){
        selectionPassed=m_trkFilter->accept(**itr,...);
    }
    if (selectionPassed)
    {
        //add track to output container
    }
}
```

Cause a copy of TAccept i.e
delete TString, remove TString
from ROOT internal maps,
construct a new TString,
insert it into ROOT maps
For all tracks!

```
// accept() signature
Root::TAccept& TrkFilter::accept(...){/*...*/}
```

Returns reference, no allocations. OK

```
namespace Root{
    class TAccept{
        // Class methods
        //....
        private:
            TString m_memberString;
    };
}
```

Has a member TString!

FOM identify a line with no apparent allocation but...

Software Optimization

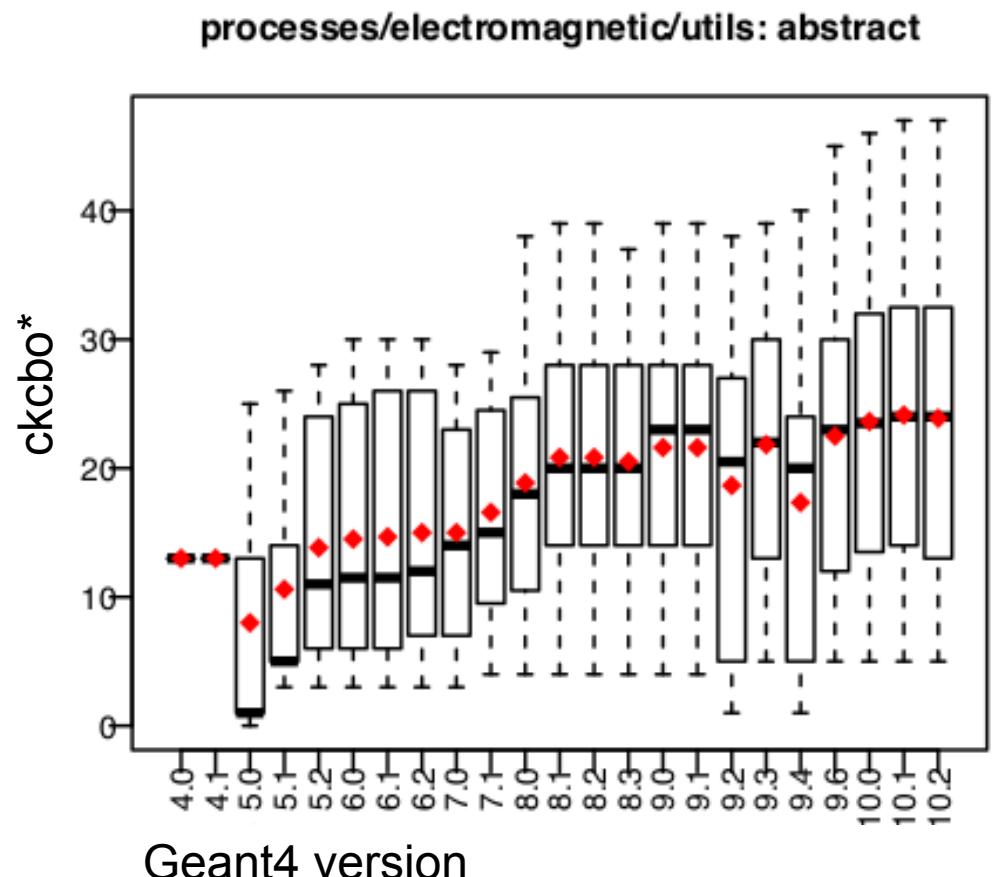
Econometric and ecologic methods

- Application of econometric and ecology analysis methods in physics software
 - Mincheol Han et al.
- Treat a software system as a sociosystem/ecosystem
- Apply data analysis concepts, methods and techniques developed in economy/ecology
 - Statistical techniques to identify patterns in a series of data
 - Used to forecast the future and also analysed passed event
 - Tests for statistical inference
 - Mann-Kendall test, Cox-Stuart test, Bartels test etc.
 - Related: change point detection

Software Optimization

Econometric and ecologic methods

- Coupling between classes
 - Excessive coupling between object classes is detrimental to modular design
- Trends in software functionnalities



* Chidamber and Kemerer Coupling Between Object

Best practices

Best Practices - MAUS

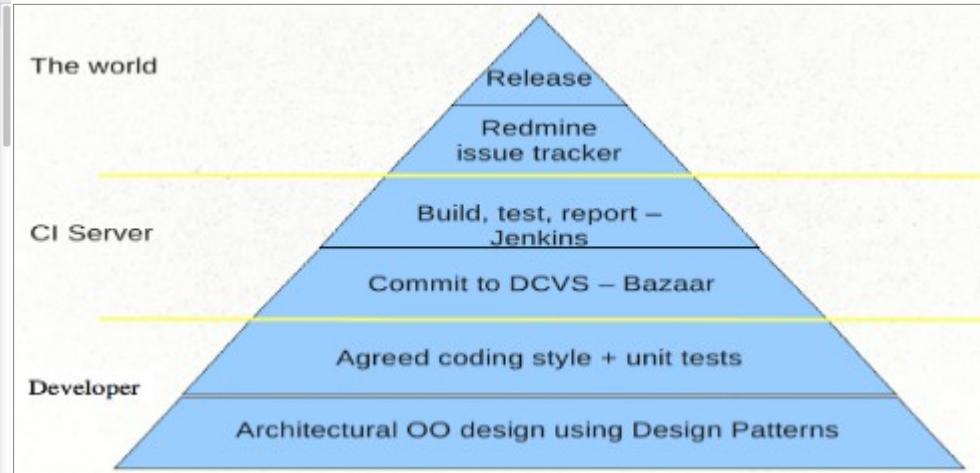
- Mice Analysis User Software (MAUS → track2) – Adam Dobbs et al.
 - Mice : Muon Ionization Cooling experiment
 - is both an accelerator physics & a particle physics experiment
 - Framework MAUS built on plugin modules (C++, Python) for analysis

The screenshot shows the Jenkins dashboard with the following details:

- Build Queue:** No builds in the queue.
- Build Executor Status:** 2 Idle, 1 Idle (Building MAUS_full_install #608), 4 Idle (hepinm071, hepinv157, Building MAUS_si64 #115, rogers_si64 (offline)).
- Build History:** A table showing build history for various MAUS components, including build times, last success, last failure, and duration.

W	Name	Last Success	Last Failure	Last Duration
MAUS_build_and_unit_test	5 hr 21 min - #357	1 day 5 hr - #356	46 min	
MAUS_per_commit_third_party_hepinm071	16 hr - #171	1 day 7 hr - #170	2 hr 54 min	
MAUS_geometry_download	19 hr - #173	N/A	3 hr 22 min	
MAUS_full_install	1 day 5 hr - #607	19 days - #596	2 hr 54 min	
MAUS_per_commit_third_party_hepinv157	1 day 7 hr - #58	22 days - #51	4 hr 9 min	
MAUS_per_commit_third_party_brunel_test_box	1 day 13 hr - #45	7 mo 5 days - #22	2 hr 31 min	
MAUS_per_commit_third_party	1 day 13 hr - #224	19 days - #218	2 hr 44 min	
MAUS Rogers load tests	4 days 21 hr - #125	N/A	1 day 0 hr	
MAUS_online	6 days 6 hr - #201	13 hr - #203	5 hr 41 min	
MAUS_integration_tests	6 days 6 hr - #324	1 day 2 hr - #326	1 day 9 hr	
MAUS_si64	6 days 11 hr - #114	N/A	6 hr 11 min	
adobbs_merge_candidate	7 days 17 hr - #1	N/A	4 hr 9 min	
MAUS_karadzkov	12 days - #106	2 days 17 hr - #111	2 hr 36 min	
MAUS_release	12 days - #71	8 mo 19 days - #63	5 hr 45 min	

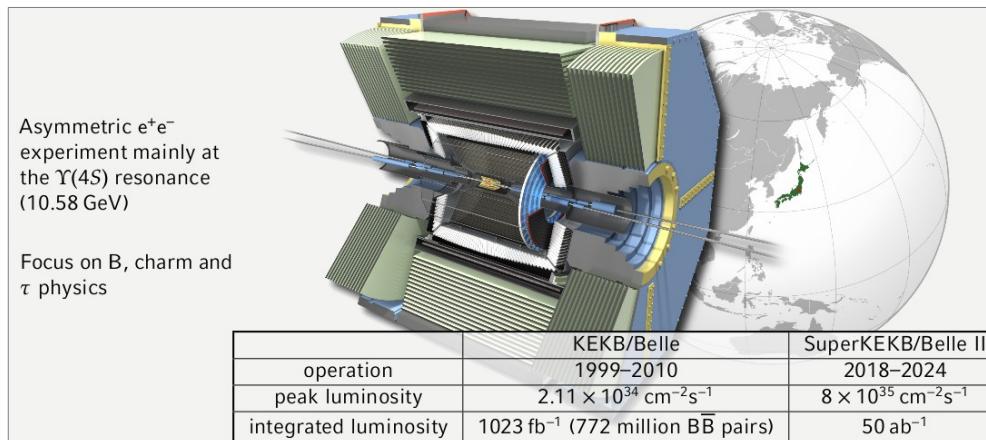
Continuous Integration (Unit + Integration tests)



Code management

Best Practices - Belle II

- Software Quality Control at Belle II - Timothy Gebhard et al.
- Software
 - Modular approach
 - Utilize new technologies: C++11 (GCC 5.2), ROOT 6, Geant 4.10, Python 3.5
 - Parallel processing support using fork
- Code organisation
 - approximately one million lines of code organized in 30 packages
 - one librarian responsible for each package



Belle II experiment

Best Practices - Belle II

- Code Management
 - switched from SVN to Git this year
 - Issues tracking with Atlassian JIRA
 - sophisticated build service and monthly integration builds
- Code Quality
 - enforced unified code style
 - cppcheck, doxygen, memcheck, clang static-analyzer
- Validation and Testing
 - 3 levels of testing: Unit tests, framework tests and physics validation
 - all runnable by the user
- Others
 - short weekly meetings, detailed monthly meetings
 - software quality shift

Results of development build						
Monday, October 03, 2016 Revision: 541441f						
failure						
Package details						
Package	Librarian	Build Result	Intel Build Result	Clang Build Result	Cppcheck	Test Result
alignment	Tadeas Šíká	✓ OK	✓ OK	✓ OK	Remarks: 4	0/4, 0/0
analysis	Anze Zupanc	✓ OK	✓ OK	✓ OK	Remarks: 30	✓ 0/14, 0/15
arith	Luka Santoj	⚠ Warnings: 2	✓ OK	✓ OK	⚠ Warnings: 2 Remarks: 71	✓ 0/0, 0/4
b2bill	Anze Zupanc	✓ OK	✓ OK	✓ OK	✗ Errors: 1 Remarks: 1	None
background	Mirko Staric	✓ OK	✓ OK	✓ OK	✓ OK	None
beast	Igor Jaćođo	⚠ Warnings: 4	✓ OK	✓ OK	Remarks: 13	None
bklm	Lea Pilonen	✓ OK	✓ OK	✓ OK	✓ OK	✓ 0/0, 0/0
calibration	Tadeas Šíká	✓ OK	✓ OK	✓ OK	Remarks: 5	✗ 0/0, 1/2
cde	Nakano Etsushi	✓ OK	✓ OK	⚠ Warnings: 1	✓ OK	✓ 0/0, 0/0
daq	Ryosuke Itoh	✓ OK	✓ OK	✓ OK	✓ OK	None
decflles	Philip Urquiza	✓ OK	✓ OK	✓ OK	✓ OK	None
display	Christian Pultwmacher	✓ OK	✓ OK	✓ OK	✓ OK	None
ewl	Tobian Ferber, Miyabayashi Kenkichi	✓ OK	✓ OK	⚠ Warnings: 1	⚠ Remarks: 55	✓ 0/12, 0/0
eklm	Krill Chikin	✓ OK	✓ OK	✓ OK	Remarks: 57	None

Build service using buildbot and Atlassian Bamboo

Directory	Line Coverage	Functions
conditions/src	49.1 %	111 / 226
core/src	76.7 %	1072 / 1397
database/src	75.7 %	465 / 614
dataobjects/src	84.5 %	490 / 580
datastore/src	92.8 %	532 / 573
dbobjects/src	98.1 %	52 / 53

Unit tests with Google test

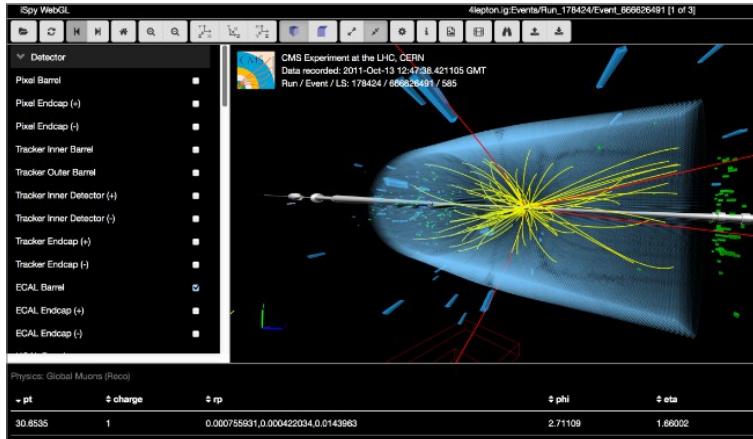
Visualisation

Visualisation - CMS

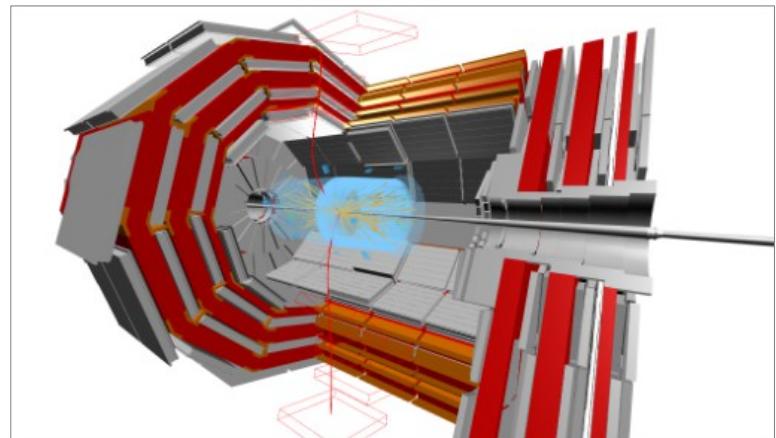
- A browser-based event display for the CMS Experiment at the LHC using WebGL
 - <http://cern.ch/ispy-webgl>
 - Production of high-quality public event display images for CMS Experiment
 - Use in CMS masterclasses for high-school students
 - Application in the CERN Open Data Portal (<http://opendata.cern.ch>)
- Using open-source technologies
 - WebGL ([OpenGL](#) based API) : <https://www.khronos.org/webgl/>
 - [three.js](#) ([JavaScript](#) API for webGL) : <http://threejs.org/>
- Input Format
 - ig format: a zip archive containing one or more JSON file
 - Extract information from CMS event format with CMSSW C++ code and convert to ig format

Visualisation - CMS

- Features
 - Standard event display control
 - Can be run completely client-side and offline
 - Tree, 3D, and sortable table view
 - Touch events enabled so works on mobiles, tablets
 - Import/export of 3D files
 - Can be read into SketchUp (<http://www.sketchup.com>)



Event display in browser

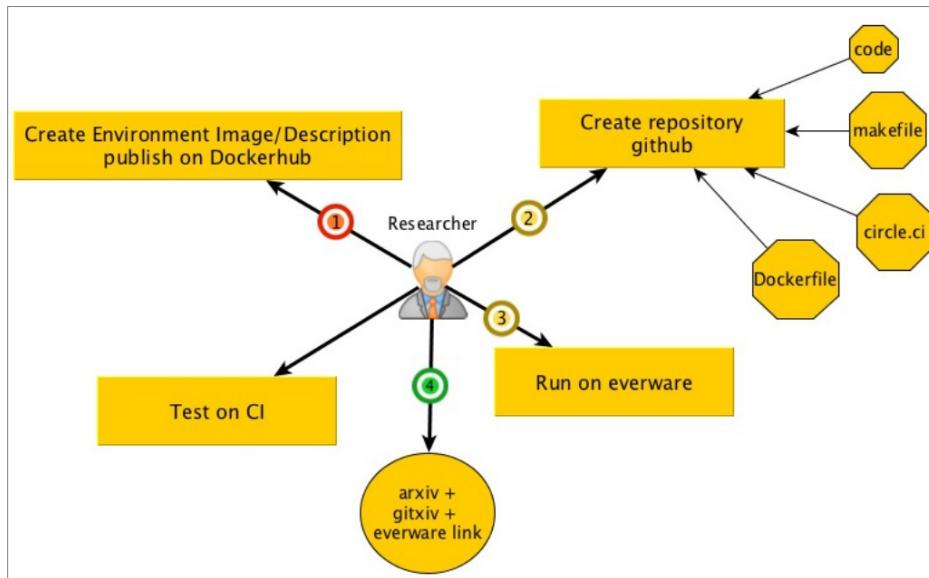


Detailed CMS geometry in Sketchup

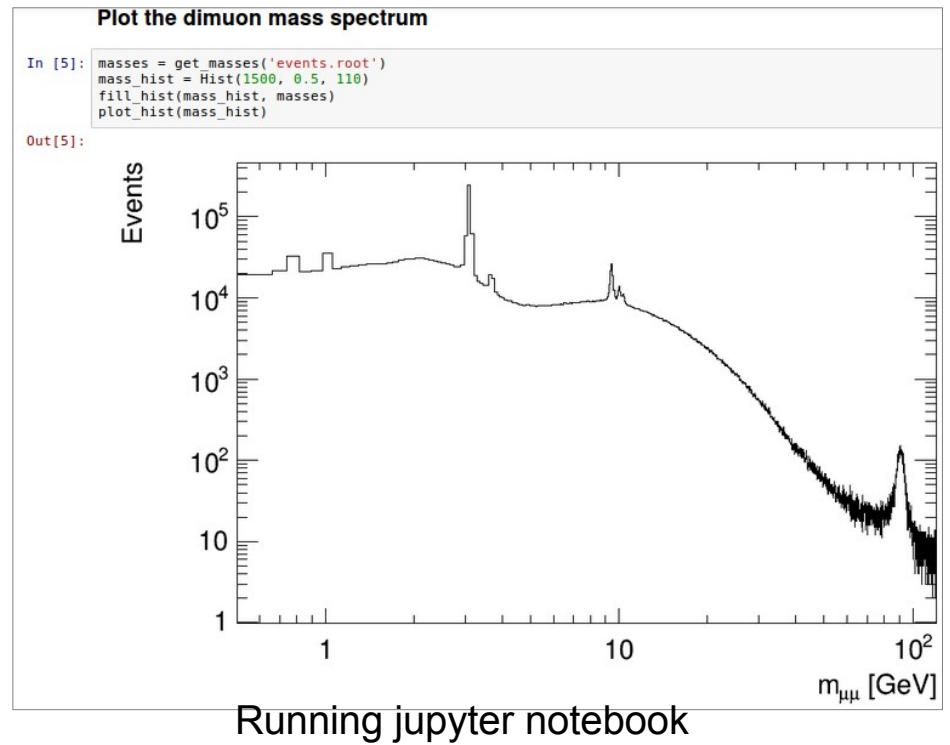
Visualisation - Everware

- Everware toolkit - Alexander Tiunov et al.
 - <https://github.com/everware/everware>
 - Demo
 - video available [here](#)
 - <https://github.com/everware/everware-dimuon-example>
 - Everware is about re-usable science
 - it allows people to jump right into your research code
 - lets you launch Jupyter notebooks from a git repository with one click
 - Everware toolkit
 - extension for JupyterHub
 - spawner for building and running custom docker images
 - integrated with
 - dockerhub and github (for authentication and repository interaction)

Visualisation - Everware



Basic research workflow



Big Data

Big Data - keywords

- Track 5 Afternoon Session
- Flume, Spark, Hadoop
 - Appears to be standard technos for Big Data



PARTIE II

Exploiter la Performance des Nouvelles Architectures

Webinaire Retour CHEP 2016

Michel Jouvin

25/11/2016

Présentation séparée

PARTIE III

Evolution des modèles de calcul et des sites

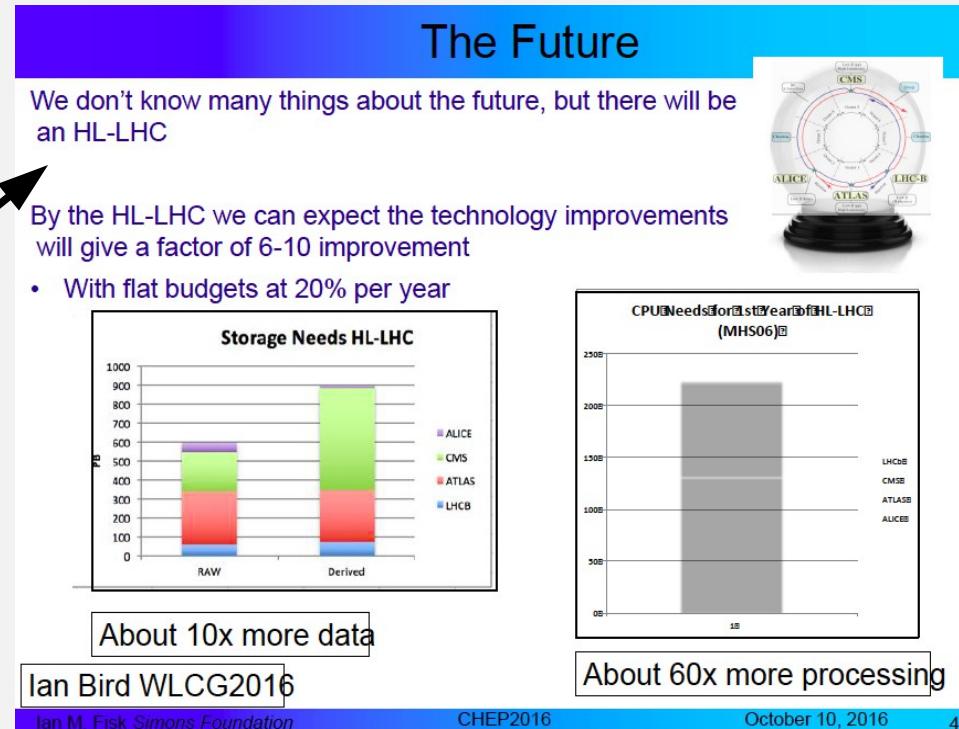
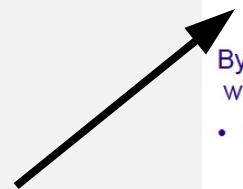
- Run 4 (HL-LHC) : augmentation des besoins du LHC (sans compter la venue d'autres expériences dans le paysage)
→ les besoins ne peuvent pas être servis à coût constant (notre hypothèse)

Optimisation des coûts ?

Agrégation de nouvelles ressources ?

- Deux pistes d'optimisation :
 - Modèles de calcul
 - Infrastructure
 - Et le réseau ?
 - Sans parler du software

I. Bird : 5 ans de computing au LHC
== coût du détecteur ATLAS ou CMS



Ref : plénier I.Fisk

Les *clouds* – l'ère commerciale

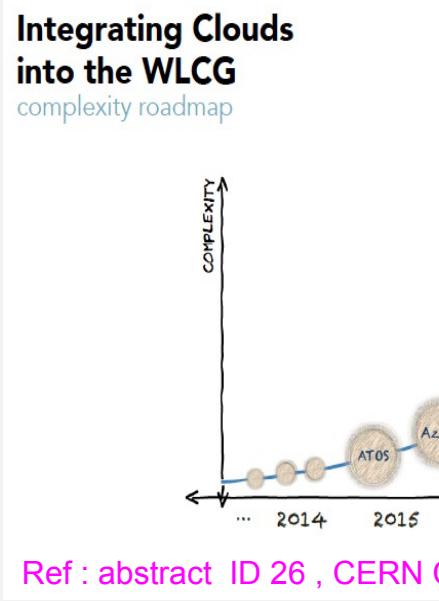
- On a beaucoup parlé de \$\$ et de clouds à CHEP
 - Privés (flexibilité, fédération) ref : plénier « Clouds and competitors », D. Sobie
 - Commerciaux (coûts)
- !! pas de remise en cause de la grille, où grille = connexion de sites, fédération
Mais plutôt une discussion sur ressources « privée » vs « commerciale »
- HEP est devenu modeste comparé aux géants commerciaux :
 - Amazon : 40 millions de cœurs / WLCG : 350 milliers
 - Google : 10-15 exaB / WLCG : 700 PB
- Les cloud commerciaux grossissent très vite
 - Une opportunité pour le monde académique ?
 - Absorption des pics (notre activité n'est pas plate) – « cloud bursting »
 - Meilleurs prix (achat de grosses volumétries)

Que faire de nos données

- Offre dans les clouds commerciaux : compute & data
 - CPU : volatile
 - Data : ne le sont pas !
 - Elles sont précieuses (coût des détecteurs + accélérateurs + opérations)
 - Elles nous coutent chers : WLCG : 2/3 du coût du matériel est passé dans le disk
- A ce jour, stocker les données dans le cloud coûte cher (les sortir)
 - Ne pas devenir prisonnier d'un cloud commercial
 - Les données restent dans notre giron
 - Délivrer les données aux clouds providers
- Cette idée peut être (est) étendue à toutes les ressources (sites dédiés, centres HPC)
 - Les technologies existent
- Projets américains AWS BNL/ATLAS ou bien AWS Fermilab/CMS
 - BNL : bucket S3 à l'extérieur du cloud + peering ESNET
 - Ces projets sont sur des *grants*

En parlant de coûts

- Le coût des clouds commerciaux est toujours plus cher que dans nos centres
- CERN et Europe - le projet HNSciCloud
 - Evaluation des couts dans un cloud hybride
 - Des workflow avec data (IO)
 - Encore un challenge à ce jour



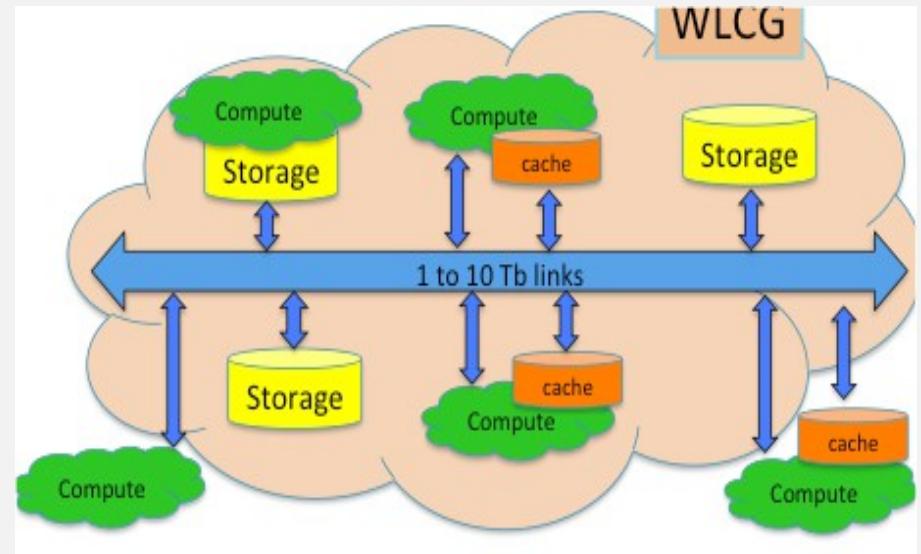
Cloud vs. In-House

- In-House is more cost-effective for both computing and disk storage in a 3-yr scenario
 - \$3.3M less for disk storage
 - \$0.6M less for computing
- Staff costs not significantly different with either solution
- Cost of data movement not estimated for either solution
- Cost of data replication not included for in-house or cloud solution
- So where does cloud computing fit at BNL?

[Ref : The role of dedicated computing centers in the age of cloud computing]

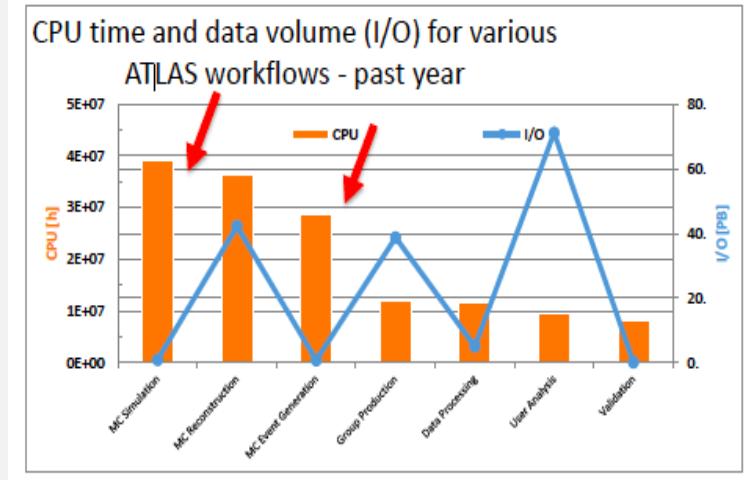
Découpler le calcul des données

- Un data center virtuel
 - Agréger les ressources dans un cloud scientifique (HTC/HPC /Grille/Cloud)
 - Très bonne connexion entre les sites (Tb)
- Un petit nombre de mega centre de réduction des données
 - Simplification des opérations
 - Les données réduites sont facilement distribuées
 - Mais : comment ca fit avec les Agences de Financements ?
 - Notre système distribué permet de trouver des ressources hors pledges



Aller plus loin...

- De façon générale : spécialisation des sites
 - choix des transformations en fonction du matériel
 - Ne plus tout faire partout
- LSST annonce des « Data Access Centers » : en charge de servir les données aux utilisateurs et de fournir du support et des ressources pour les analyseurs finaux
 - Ref : plénier « Non-HEP: Data Management for LSST »
- Tous les événements ne sont pas « du Higgs »
 - Simulation rapide
- Déjà au Run 3 du LHC : calibration et reconstruction « en ligne »
 - (Upgrade des détecteurs ALICE et LHCb)
 - ALICE : construction d'une mega ferme « O2 » (online/offline)
 - LHCb : mise en place d'un « turbo stream » (on ne garde plus les données brutes) ref : abstract ID 447



Les sites s'organisent

- Agrégation de ressources
 - Extension de sites
 - Interconnexion aux clouds
 - Mise en commun de plateforme
- Utilisation des gros centres HPC voisins
 - Sites américains, Suisse, Chine

Motivation

- Flat budgets, increasing costs for energy, personnel, ...
- Increasing competition by commercial cloud providers
- HEP computing requirements (# PBs etc.) not unique any more. Other fields of research catching up.
- Funding agencies hesitant to fund HEP specific computing infrastructures and push for more cooperation between different research fields.

=> National and EC funded projects and infrastructures

Ref : abstract ID 161, Developing the WLCG Tier-1 center GridKa as topical center in a multidisciplinary research environment

KIT
Karlsruhe Institute of Technology

The diagram illustrates the shift from a single discipline to a multidisciplinary environment. It starts with a yellow box labeled 'HEP'. This is followed by a row of three boxes: 'Biology' (green), 'Earth Sciences' (blue), and 'Very different methods, workflows, ...' (light blue). Below this is a row of two boxes: 'HEP specific services, protocols' (yellow) and 'Common and specific Services, protocols' (light blue). The final stage shows two large boxes: 'Platform 1' (yellow) and 'Platform 2' (teal).

3 CHEP 2016 J. Heuser, P. Petrucci

THE LHCONCRAY PROJECT: OVERVIEW

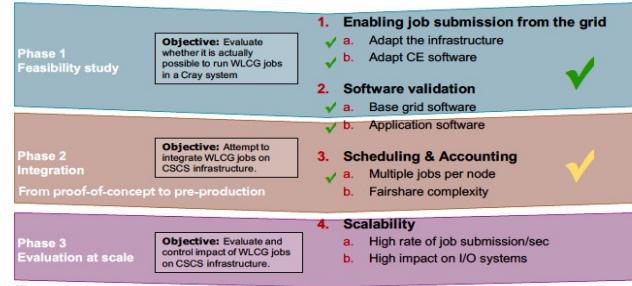
ATLAS and LHC COMPUTING ON CRAY

8

New approach to a Cray at CSCS, Lugano (2016)

- ▶ CRAY XC40 (Piz Dora) - we have access to the Test Development System (Brisi)

The original Compute plan



CSCS

LHConCray Project | 3

ETH Zürich

u^b

REF : abstract ID 88, ATLAS AND LHC COMPUTING ON CRAY



Toward a (semi-)elastic Data Center?

- Planning to upgrade the Data Center to host resources at least until the end of LHC Run 3 (2023)
- A complementary solution could be (dynamic) extension on remote farms
- Cloud bursting on commercial provider
 - Tests of opportunistic computing on Cloud providers
- Static allocation of remote resources
 - First production use case: part of 2016 pledged resources for WLCG experiments at CNAF are in Bari-ReCaS
- Also participating to HNSCloud EU PCP project

Ref : abstract #421, Extending the farm on external sites: the INFN Tier-1 experience

7

CHEP 2016

Oct. 10 2016

Istituto Nazionale di Fisica Nucleare

A l'autre bout du spectre

- Pour que les « petits » sites puissent aussi contribuer
 - US : US-ATLAS et US-CMS travaillent en collaboration rapprochée avec les sites
 - EU : une plus grande diversité de sites, de cultures de financements

Hardware Specifications

(aka the “brick”)



Hardware:

- 40 cores
- 12 x 4TB data disks (RAID6, XFS)
- 128 GB ram
- 2 x 10 gbit network interface

Software:

- Full HTCondor pool
- XRootD server, redirector, and proxy cache
- cvmfs w/ optional Squid

The brick is effectively a site in a box for \$10k

Ref : abstract ID 269, Site in a Box, Improving the Tier 3 Experience

CHEP 2016

4

10/13/2016

T2 simplifications

- **Reduce** the catalog of required services, where possible
- **Replace** classic, complex services with new, simpler portfolio
- **Simplify** deployment, maintenance and operation of services
- Some sites could offer a partial portfolio, e.g. just “worker nodes”
 - Need to get their contributions properly recognized

Ref : abstract ID 165, Lightweight sites:

computing resources

7

Pour finir

Le prochain CHEP

CHEP 2018 : 23-27 avril, 2018

WLCG Workshop : 21-22 avril 2018

Sofia

- Bulgaria's capital city with a population of 1.3 M
- It's in the [top 10](#) best places for business start-ups in the world
- An old city having many historical places and thousands of years of history
- International IT companies in Sofia: HP, IBM, VMware, Telerik, Microsoft

