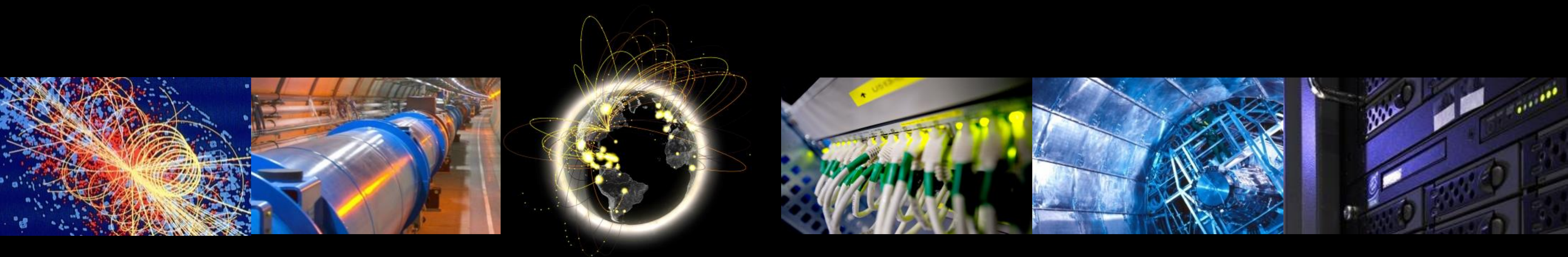
Cost and system performance modelling in WLCG and HSF: an update

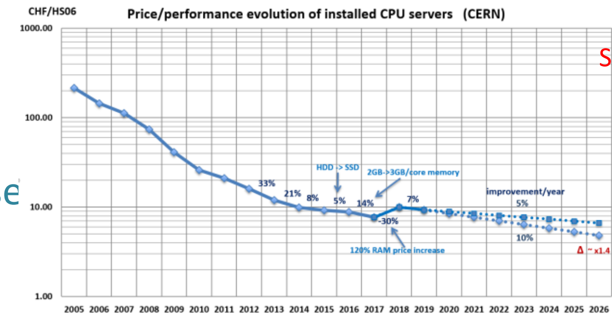
Andrea P. Sciabà
on behalf of the HSF/WLCG Systems performance and cost modelling WG

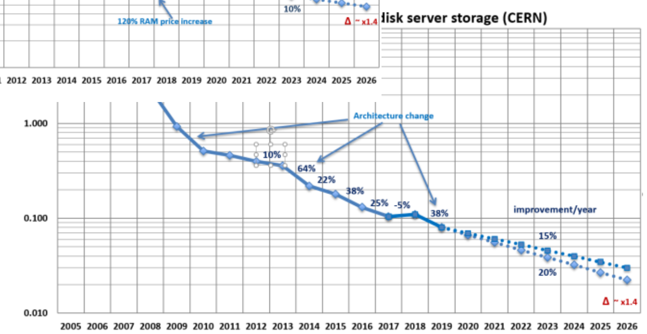
HEPiX Spring 2019 Workshop
La Jolla, 25-29 March 2019



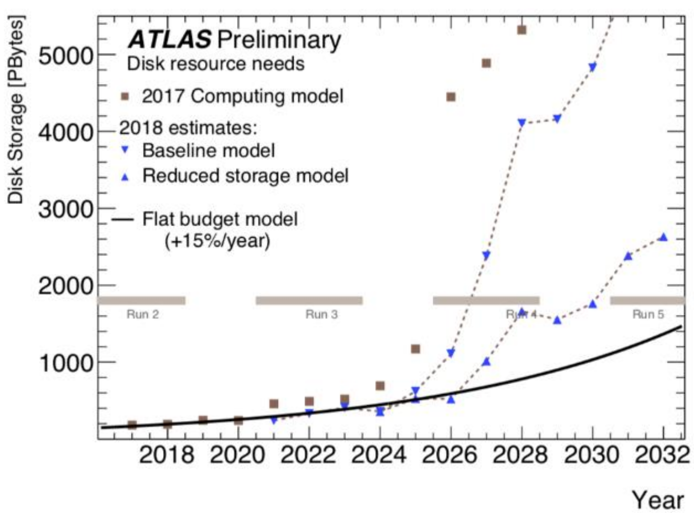
The High Luminosity challenge

- There is still a significant gap between the estimations of needed and available resources
  - 10x increase in trigger rates, 5x increase in pile-up, NLO & NNLO, …
  - Price/performance advances slowing down, 10-15%/year at best
- CPU and disk short by a factor ≈ 2
  - Even if the gap is reducing!
- Strong need to quantitatively understand our efficiency and how we can optimize performance

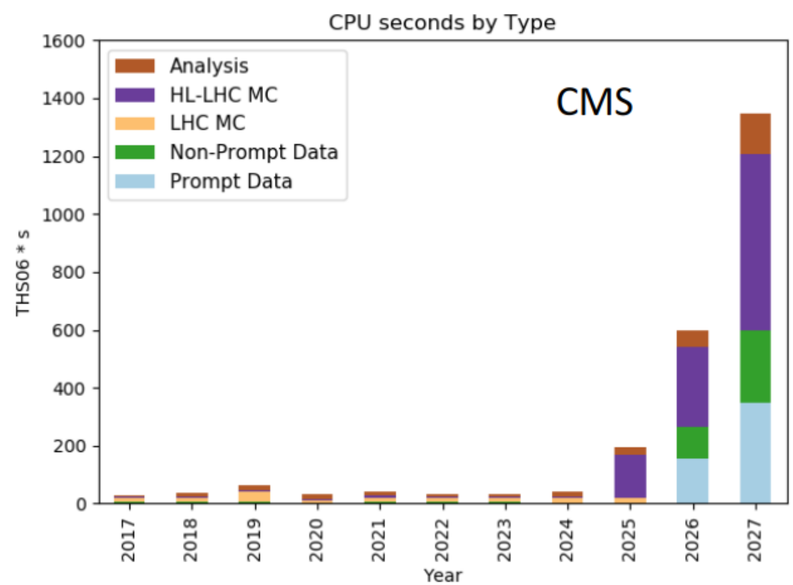




Source: B. Panzer-Steindel



Source: D. Costanzo



Source: D. Lange

The Working Group

- Main motivation is to help WLCG to fit into the available resources for Run3 and Run4
  - Develop a deep understanding of current workloads, resource utilisation, and site costs
  - Explore future scenarios, estimate possible improvements in efficiency
  - Develop tools and methods for the above
- The group was created in November 2017
  - Already presented status reports at HEPiX
  - Conveners: J. Flix, M. Schulz, A. Sciabà
  - ~35 active members from LHC experiments, sites and including software experts
  - By-weekly meetings
    - https://indico.cern.ch/category/9733/
- Dedicated session at the HOW workshop last week

WLCG Worldwide LHC Computing Grid

Current areas of work

- Some of the current areas of work and goals
  - Identify representative experiment workloads
  - Define which metrics best characterise such workloads
  - Understand how to estimate resource needs
  - Define a process to evaluate the TCO of an infrastructure
  - Measure the impact of new storage configurations
  - Identify potential areas for savings

WLCG
Worldwide LHC Computing Grid

Metrics and workload characterization

- Identify the metrics that best characterize the resource usages of HEP workloads
  - To quantify the impact of changes in the workload implementations → software experts
  - To quantify the impact of changes in the infrastructure → site administrators
  - Ultimately, to guide design decisions towards improved efficiencies
- Basic metrics
  - CPU (i.e. how much do I use?)
  - Memory (i.e. do I swap a lot?)
  - Disk I/O (i.e. how much do I read/write?
  - Network (i.e. what are the access patterns?)
- Can use procFS for the above → PrMon (GitHub)
- Hardware counters for deeper analysis → Trident

Metrics	Source	Comments
Process #	pstree /proc/<pid>/task/<pid>/children	Multi-process applications
Thread #	/proc/<pid>/stat	Multi-thread applications
CPU	/proc/<pid>/stat	User and System usage
Memory	/proc/<pid>/smaps	VMEM, RSS, PSS
IO	/proc/<pid>/io	Total bytes read/written
Network	/sys/class/net/<device>/statistics/*	<device>-level information

WLCG Worldwide LHC Computing Grid

Reference workloads

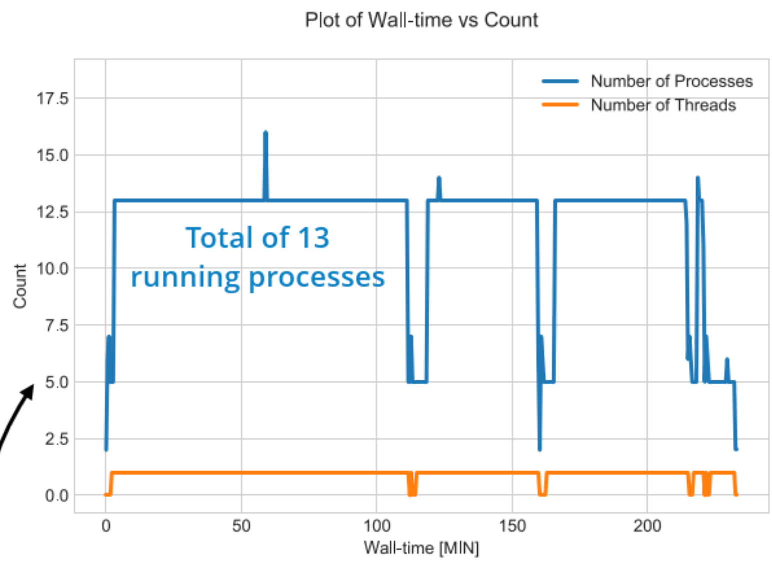
- ALICE (link) (GitLab)
  - A simple p-p simulation job using Geant3
- ATLAS (link)
  - A ttbar simulation job using Geant4 (GitLab)
  - A digitization + reconstruction job
  - A DxAOD derivation job
- CMS (link)
  - Generation + simulation of ttbar events (GitLab)
  - Digitization and trigger simulation with premixed pile-up
  - Reconstruction job producing AODSIM and MINIAODSIM
- LHCb (link)
  - Generation + simulation using Geant4 of $D^{*}(\rightarrow\pi(D^{0}\rightarrow K\ \pi))\ \pi\ \pi\ \pi$ events (GitLab)

Job	Events	Processes/ threads	Wallclock time (sec)	CPU efficiency	Time per event (sec)	Memory per core (GB)	Read rate per core (MB/sec)	Write rate per core (MB/sec)
ALICE sim	1000	1	10901	100%	10.90	0.96	0.0788	0.1660
ATLAS sim G4	1000	8	33627	100%	269.02	0.44	0.0152	0.0090
ATLAS digireco	2000	8	13981	87%	55.92	1.12	0.3174	0.2412
ATLAS deriv	95741	8	8401	98%	0.70	1.20	0.6849	0.0705
CMS gensim	1000	8	2651	99%	21.21	0.19	0.0473	0.0377
CMS digi	1000	8	737	78%	5.90	0.65	0.2854	0.3004
CMS reco	1000	8	1221	83%	9.77	0.45	0.3073	0.2153
LHCb gensim	10	1	1782	100%	178.20	0.89	0.3115	0.0117

WLCG Worldwide LHC Computing Grid

PrMon example: ATLAS digi-reco (1/2)

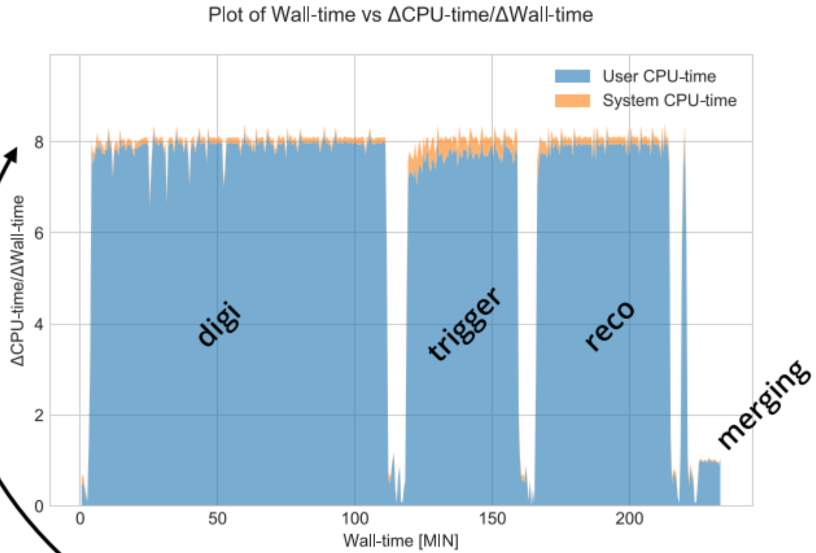
- Running digitization, trigger simulation, reconstruction+ in 8 parallel processes



Total # of processes/threads

"offset": 5 additional non-worker processes

6



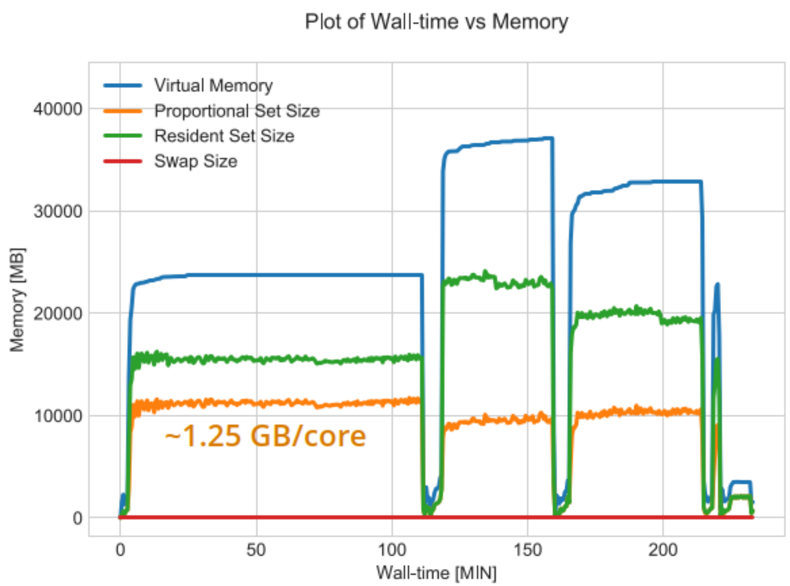
Effective CPU utilization

Source: Serhan Mete

Worldwide LHC Computing Grid

PrMon example: ATLAS digi-reco (2/2)

- Running digitization, trigger simulation, reconstruction+ in 8 parallel processes



Total memory usage

~220 MB/min
~3.5 MB/sec

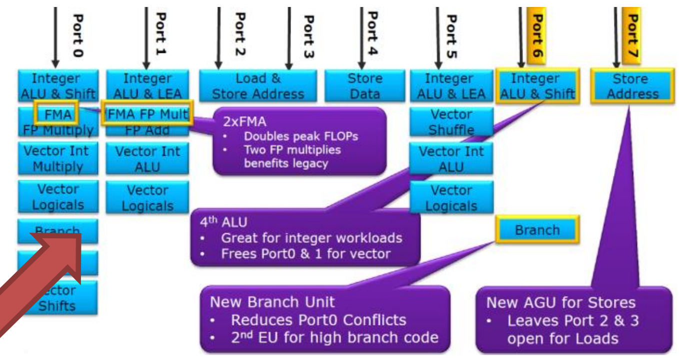
Disk I/O rate

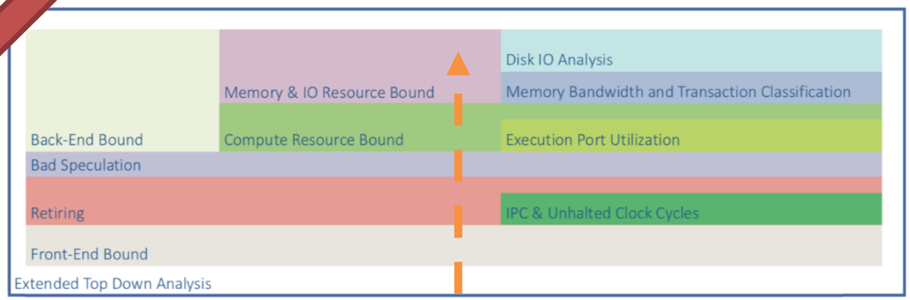
Source: Serhan Mete

WLCG Worldwide LHC Computing Grid

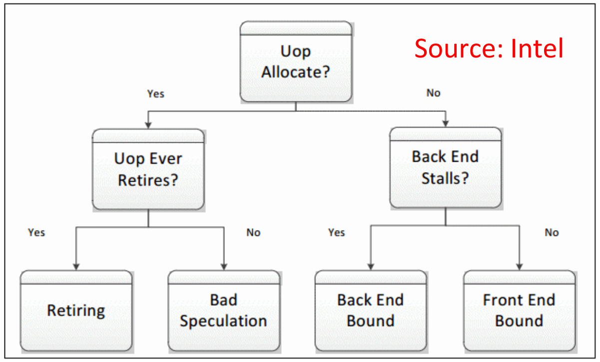
Measuring performance with Trident

- Measures CPU, IO and memory utilization based on hardware counters, memory and IO information, e.g. to identify bottlenecks
- Several metrics calculated
  - CPU: IPC, total cycles, top-down analysis (front-end bound, back-end bound, retiring, bad speculation)
  - Core backend utilization: compute (ports 0,1,5,6) vs memory (ports 2,3,4,7)
  - Memory: bandwidth usage, transaction classification (page-hit, page-empty, page-miss)
- Can be used to see how workloads differ (or resemble) the benchmarks we use (e.g. HS06)
- CPU counters are a powerful (but complex) tool and Trident makes them accessible



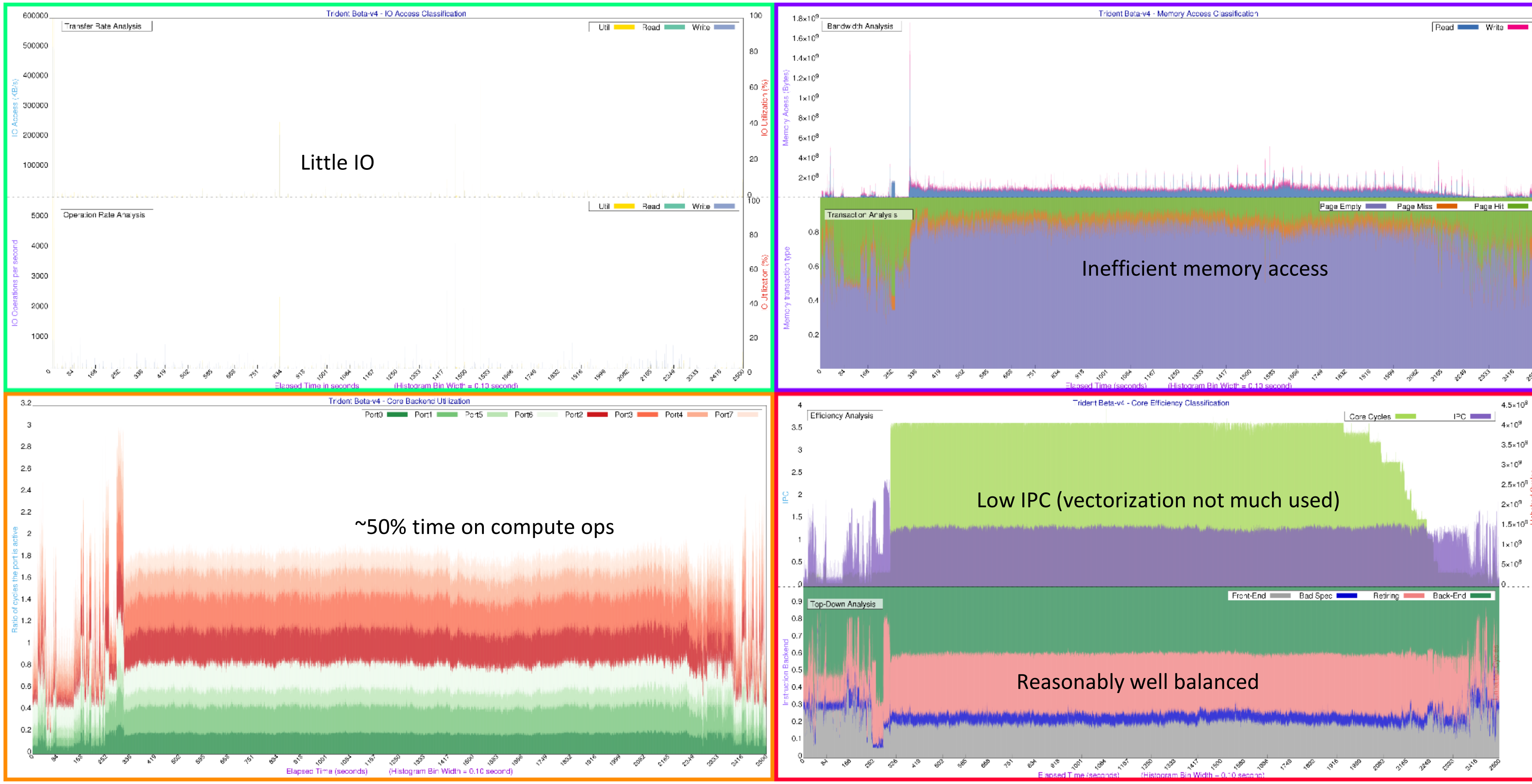


Full exploration of CPU utilization



Worldwide LHC Computing Grid

Trident plots: ATLAS Geant4



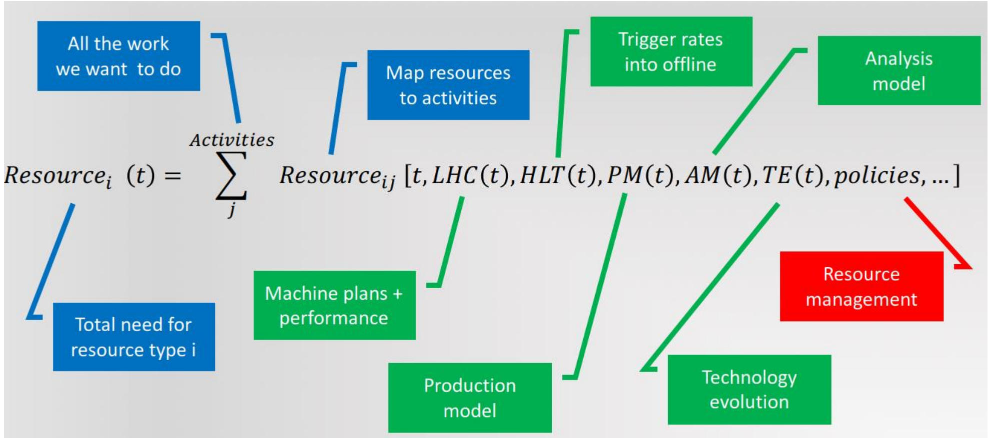
More on top-down analysis [here]

Source: Servesh Muralidharan

WLCG Worldwide LHC Computing Grid

Resource estimation

- The initial goal was to define a common framework for modelling the computing requirements of the LHC experiments
  - Models as collection of parameters and generic calculations
  - Using as input the characteristics of the workflows
  - Allow to play with different scenarios



Source: D. Lange

- However, now it seems that difference among experiments are large enough to make such tool either very complex or not realistic enough
  - Still, lot of progress came from sharing ideas and code

WLCG Worldwide LHC Computing Grid

General observations from resource estimation

- Models address CPU, disk and tape
  - But not network (not OK), and no GPUs (still OK)
- Yearly granularity in estimates is sufficient
  - Easy to fill the resources at all times
- Effect of R&D or big changes in computing models difficult to anticipate
  - Better to use the model to set targets given the constraints
- Experiments are talking to each other more than ever (also thanks to the working group)
- Discussion if still important to separate T0/1/2s in the context of HL-LHC
- Tape I/O as an extremely important metric at HL-LHC
  - RAW data will increase by 50x (CMS)
  - Expect bursts in tape I/O, need to have good estimates

WLCG Worldwide LHC Computing Grid

Site cost estimation models

- The goal is to develop a method to calculate the cost of providing the needed resources
  - At first, several site people in the WG went through a cost estimation exercise starting from an “example” workload
    - Significant differences due to different metrics, methodologies, understanding of metrics
- A model should rely on
  - Common metrics
  - Common measurement methods
  - Common framework for computing costs
- First implementation as a spreadsheet using as input prices and trends to calculate required budgets

WLCG Worldwide LHC Computing Grid

Cost evolution model

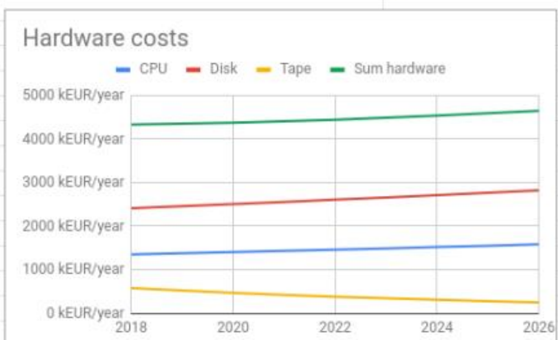
Example with dummy data

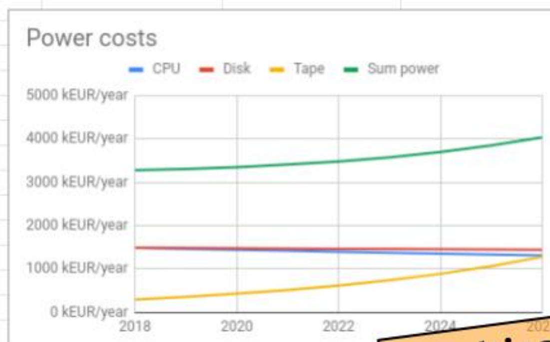
	Category	Metric	current value	yearly evolution type	yearly evolution rate
SITE DATA (input params)	Local power situation	Datacenter PUE	1,7	none	none
		Power price	0,10 EUR/kWh	linear	0,030 EUR/kWh.year
	Power consumption	CPU	2,00 W/HS06	exponential	-18,0 %/year
		Disk	10,00 W/TB	exponential	-17,0 %/year
		Tape	1,00 W/TB	exponential	0,0 %/year
	Procurement	CPU price	10,00 EUR/HS06	exponential	-15,0 %/year
		Disk price	100,00 EUR/TB	exponential	-15,0 %/year
		Tape cartridge price	10,00 EUR/TB	exponential	-25,0 %/year
	Tape budget	cartridge budget over total tape budget	40%		
	Hardware lifetime	CPU lifetime	5 years		
		Disk lifetime	6 years		
		Tape lifetime	7 years		

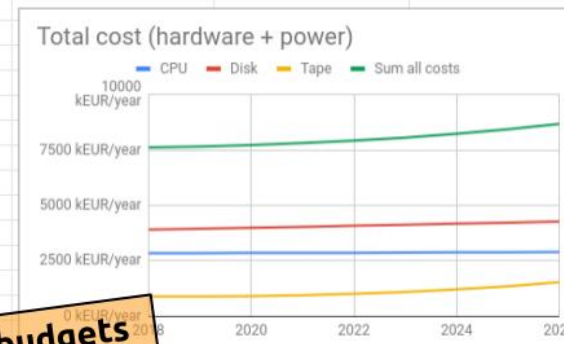
Site input

		current capacity	capacity growth rate
Capacity planning (input params)	CPU	500 kHS06	20,0 %/year
	Disk	100 PB	20,0 %/year
	Tape	200 PB	20,0 %/year

Experiment input







Resulting budgets

renaud.vernet@cc.in2p3.fr 2019-03-20 CCIN2P3 6

WLCG Worldwide LHC Computing Grid

Site cost Tier-1 survey

- Launched in September a survey among Tier-1 sites (and open to Tier-2s) to understand their costs for CPU, disk and tape
  - Questionnaire available here
  - Eight Tier-1s and one Tier-2 answered
- Average costs
  - CPU: €10.3/HS06, -12%/y
  - Disk: €126/TB, -15%/y
  - Tape: €22/TB, -14%/y
- 20-50% spread among sites

Source: R. Vernet

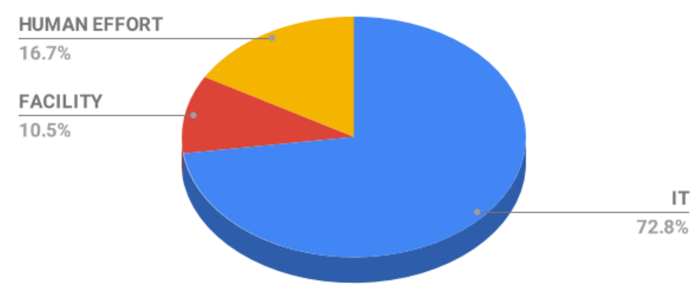
WLCG Worldwide LHC Computing Grid

Total cost of ownership

- Two approaches being considered
- "Atomic" approach
  - Cost of hardware, power, network equipment, building, electricity, PUE, FTEs included
  - Implemented as a self-documented [spreadsheet]
- "Holistic" approach
  - Take the complete budget of the datacenter
  - Remove tertiary expenses
  - Categorize per sector
  - Deduce TCO per unit of resource capacity
- Networking expenses not yet modelled
- The goal is to come up with a unified approach usable by all WLCG sites
  - Discussions are still ongoing

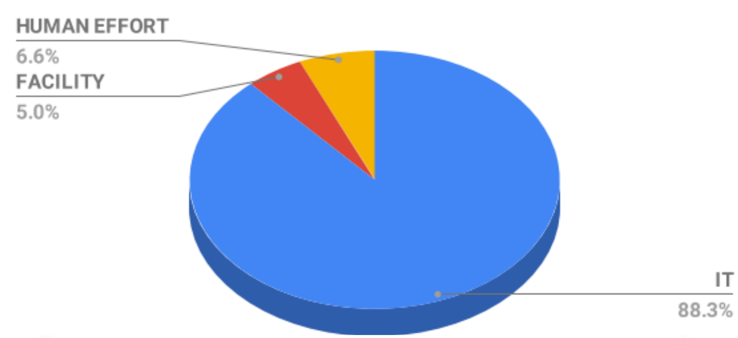
CPU server expenses

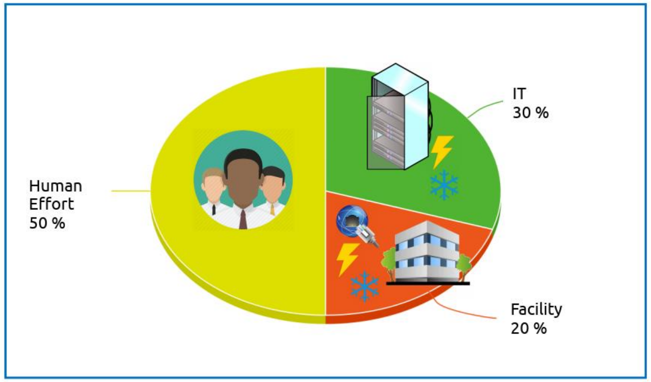
CC-IN2P3



Storage server expenses

https://goo.gl/ctiMkE

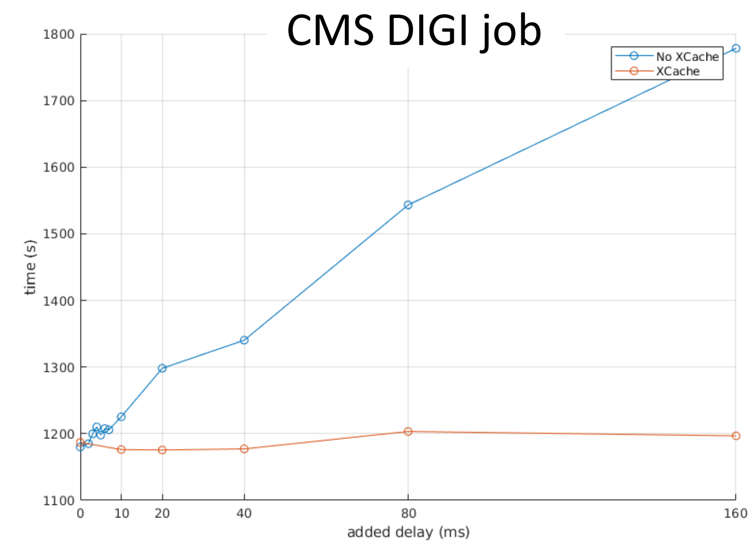


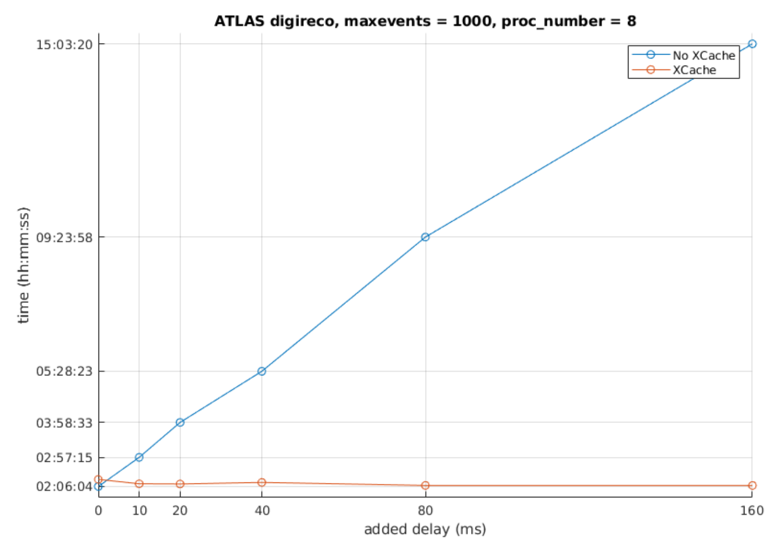


Source : R. Vernet

Storage modeling and popularity studies

- Investigated the feasibility of storage caches and the impact of network latency
  - Work is now converging into the DOMA (Data Organization Management Access) working group in WLCG
  - Studies with a focus on cost will stay in this WG
- Examples
  - Investigating the effect of data losses on non-redundant storage (spreadsheet)
  - Effect on reference workloads of an artificial latency, with and without a cache

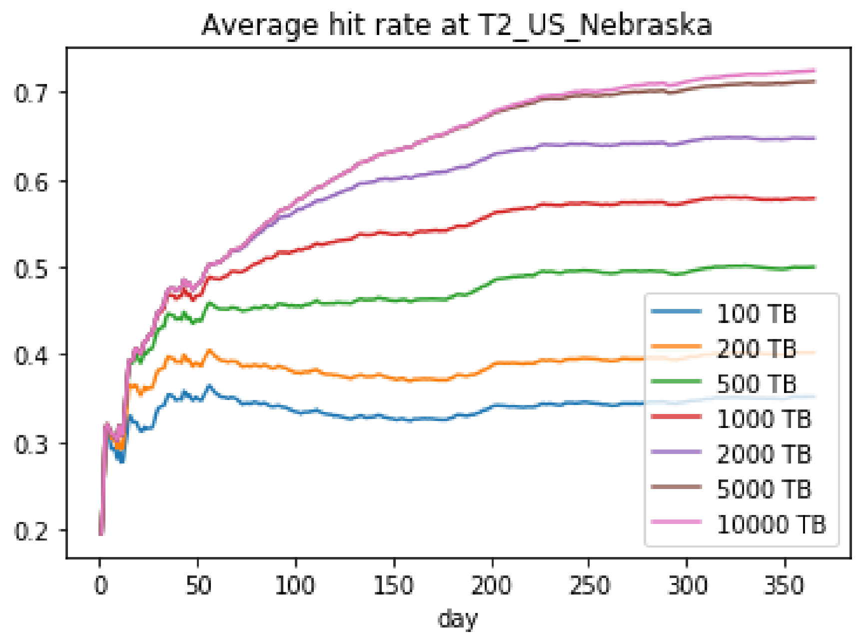


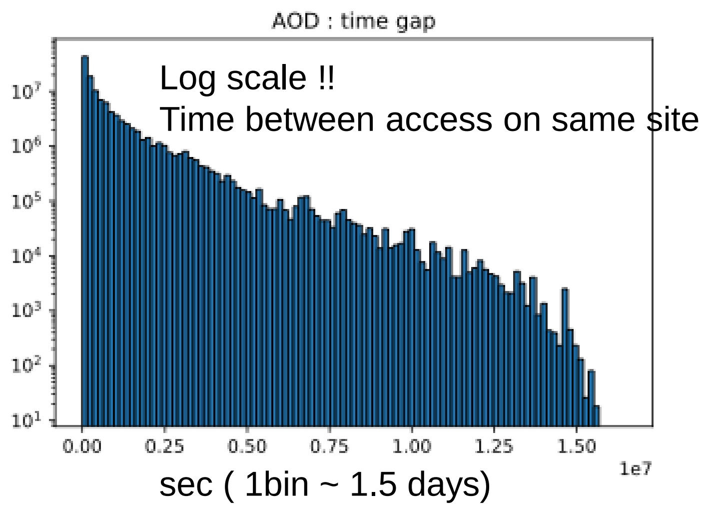


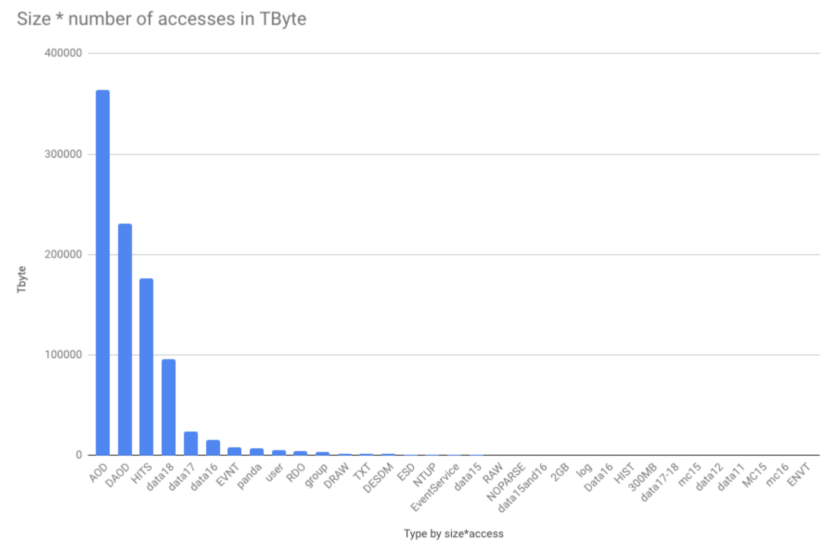
WLCG Worldwide LHC Computing Grid

Storage modeling and popularity studies

- Examples (cont’d)
  - Simulating a cache using CMS popularity data
    - Hit rates vs. cache size
  - Studying data access patterns in ATLAS
    - Using Rucio logs
    - Looked at “impact”(file size * no. of accesses) vs. file type, time between accesses, etc.
- Studies are all preliminary
  - The point here is to give a taste of what is going on







Worldwide LHC Computing Grid

Other areas of potential savings

- Many “small” improvements can stack to provide significant gains
  - Some of these estimates got more precise as a result of recent studies (data popularity, TCO estimates, effect of data loss, …)

Summary of how this evolved:

Change	Effort Sites	Effort Users	Gain
managed storage on 15 sites + caches elsewhere	Some on large sites/gain on small sites	little	40% decrease in operations effort for storage
Caches at most sites (dataLake strawman)	Some everywhere	Frameworks some	15% of storage
Reduced data redundancy	Some large sites	Frameworks some	30-50% disk costs
Reduced data replication and cold data	little	Frameworks some	30% disk costs
Compact data formats for analysis	none	Some	>15% disk costs
Scheduling and site inefficiencies	Some	Some	10-20% gain CPU
Reduced job failure rates	Little	Some-Massive	5-10% CPU
Compiler and build improvements	None	Little-Some	15-20% CPU
Improved memory access/management	None	Realistic	10%-15% CPU
Exploiting modern CPU architectures	None	Massive	100% CPU
Paradigm shift algorithms	Some	Massive-Infinite	Factor 2-100 CPU
Paradigm shift online/offline data	Little	Massive-Infinite	2-10 CPU 10-20 Storage

Source: M. Schulz

WLCG Worldwide LHC Computing Grid

Next steps

- After 1.5 years of activity it's time to re-examine the roadmap and the goals
- Some preliminary ideas
  - Archive the results of all current and future studies
  - Set clear priorities for studies
    - So far much of the work was done in a "spontaneous" way, leading to interesting results but without a long-term goal in mind
  - Harmonize studies performed on data from different experiments
    - Typically popularity data
  - Make sure that such studies are well coordinated with DOMA
  - Consolidate the tools to calculate TCO and cost trends
    - Very important to allow experiments to plan for the future
  - Improve the understanding of tape costs relative to disk
    - Including complexities like Oracle dropping out of tape business
  - Consider running PrMon for all jobs
    - Negligible overhead, may lead to a much better understanding of resource usage (currently experiments record averages, or max of metrics, not time series)

WLCG Worldwide LHC Computing Grid

Conclusions

- This working group was established to improve our understanding of the performance and the cost of computing for LHC (and HEP) and its evolution
  - HL-LHC requires us to squeeze all the performance we can get at all levels
- The WG is active in many areas, is already achieving important results and these activities are much more mature now
- Independent work also started (e.g. at KIT to model a datacenter), we should try to benefit from it
- Right time to re-evaluate priorities
- Better coordination with DOMA
- Work is still in progress but the time scale is long…

WLCG Worldwide LHC Computing Grid

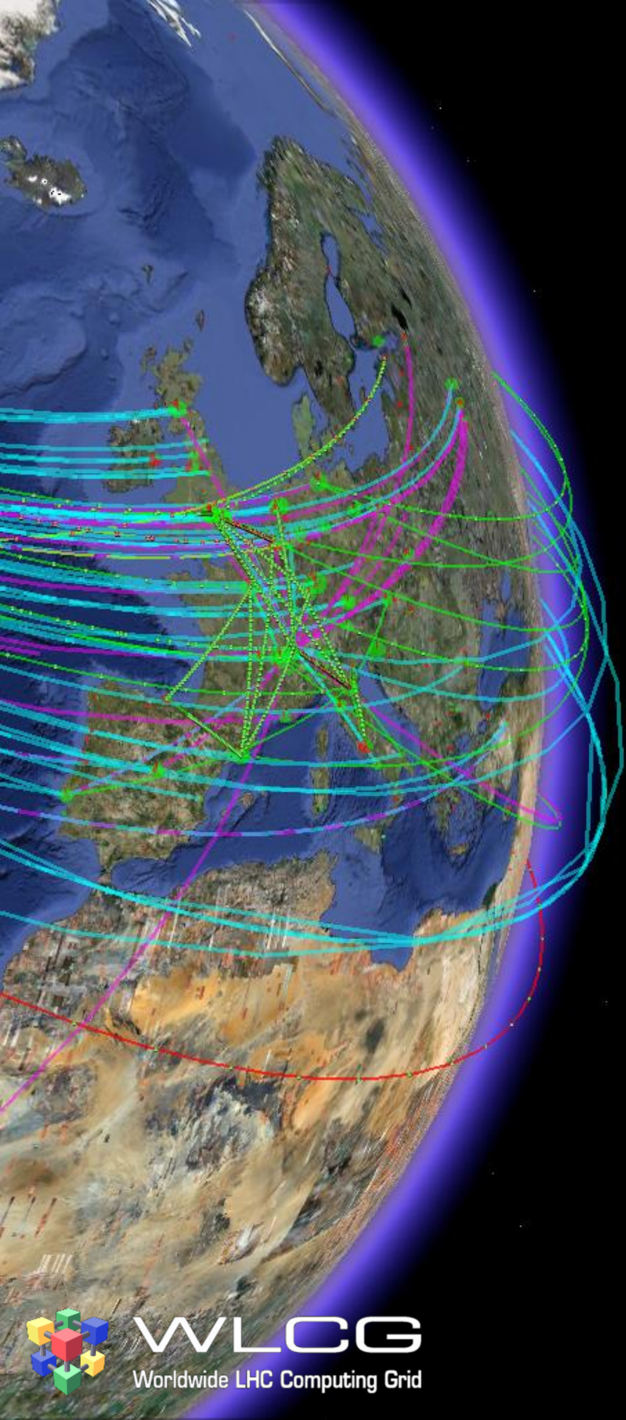
Membership

- Alaettin Serhan Mete, Alessandra Forti, Alessandro Di Girolamo, Andrea Sartirana, Andrea Sciabà, Andrea Valassi, Andrew Sansum, Andrey Kirianov, Antonio Pérez-Calero, Bernd Panzer-Steindel, Carlos Perez Dengra, Catherine Biscarat, Concezio Bozzi, Costin Grigoras, Daniele Bonacorsi, David Lange, David Smith, Davide Costanzo, Dirk Duellmann, Domenico Giordano, Duncan Rand, Eric Fede, Erik Mattias Wadenstein, Frank Wuerthwein, Gareth Roy, Graeme A Stewart, Helge Meinhard, Jan Iven, Johannes Elmsheuser, José Flix, Markus Schulz, Martin Gasthuber, Michel Jouvin, Michele Michelotto, Oxana Smirnova, Paul Millar, Raul Cardoso Lopes, Renaud Vernet, Servesh Muralidharan, Tommaso Boccali, Torre Wenaus, Xavier Espinal Curull, Xiaomei Zhang, Yves Kemp

WLCG Worldwide LHC Computing Grid

Further reading

- Indico category
  - https://indico.cern.ch/category/9733/



Backup slides

Differences in TCO

Holistic TCO

Atomic TCO

Facility
- Building & equipment
- PUE

IT
- Filled racks
- Network
- Power consumption

Human effort
- System admin
- Service operator
- Network expert

Facility
- Upgrades
- Maintenance

Other human effort
- Facility people
- Dev ops
- Administration
- Project managers
- User Support