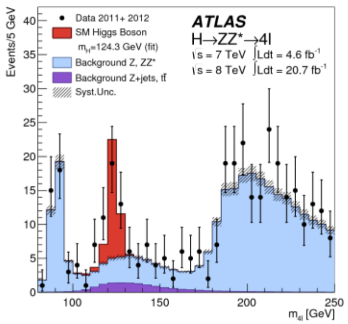
The ATLAS computing challenge for HL-LHC

Simone Campana, Torre Wenaus

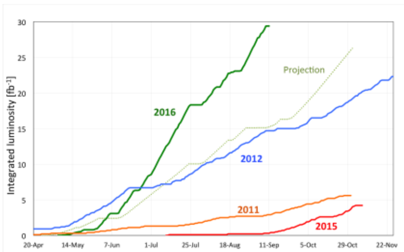
for the ATLAS collaboration

The Large Hadron Collider at CERN

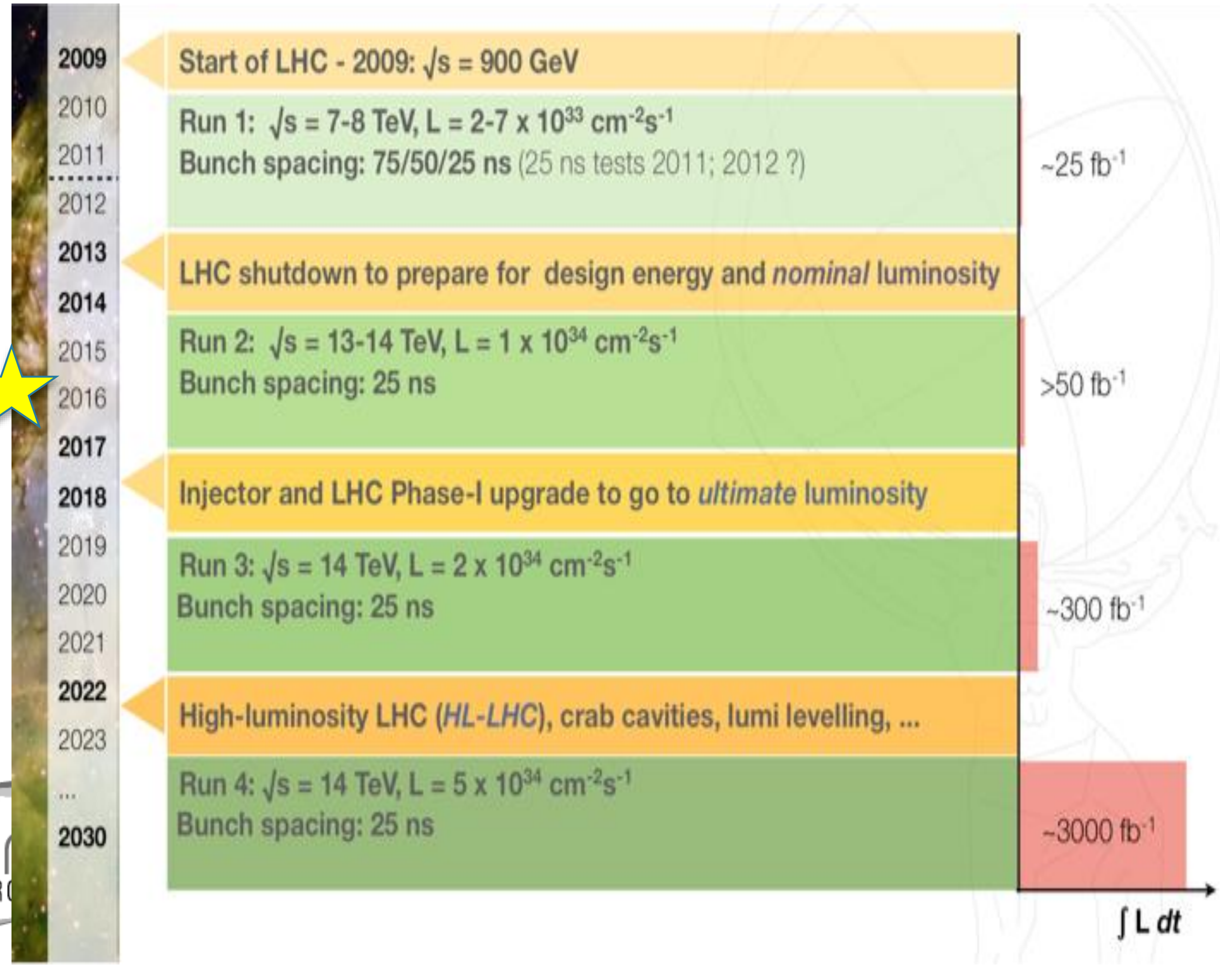
Higgs discovery in Run-1



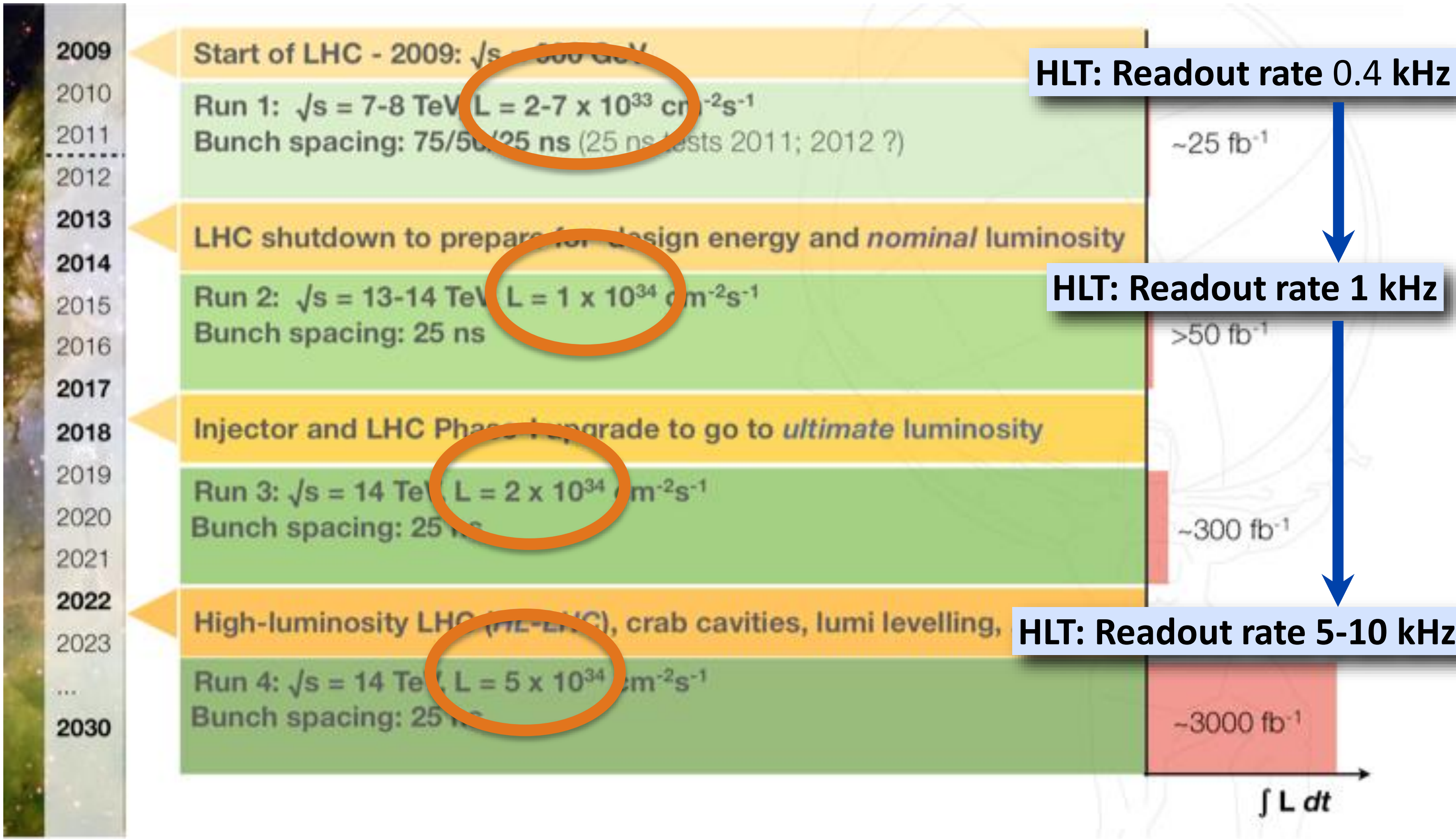
We are here: Run-2



This talk: computing challenge at HL-LHC in 2026 and beyond



HL-LHC rates: 10x more data per second



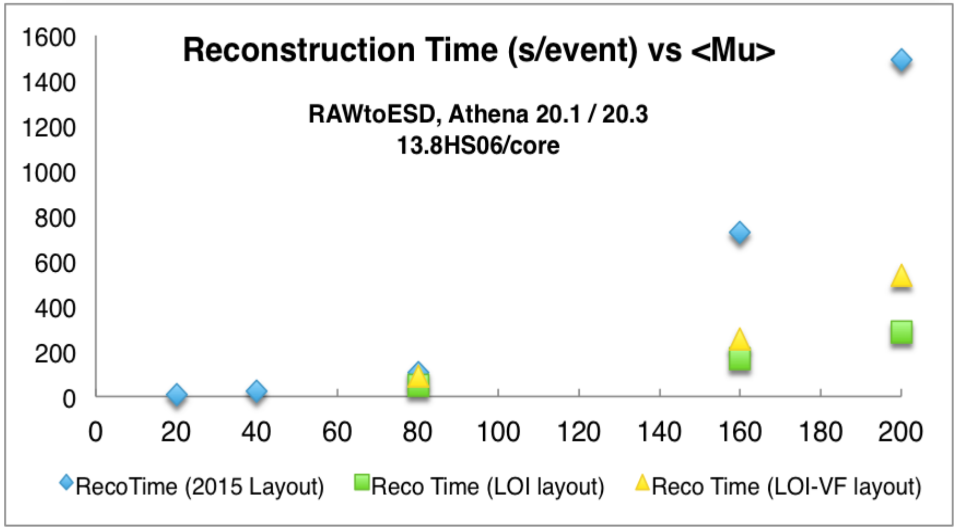
Effect of pile-up increase

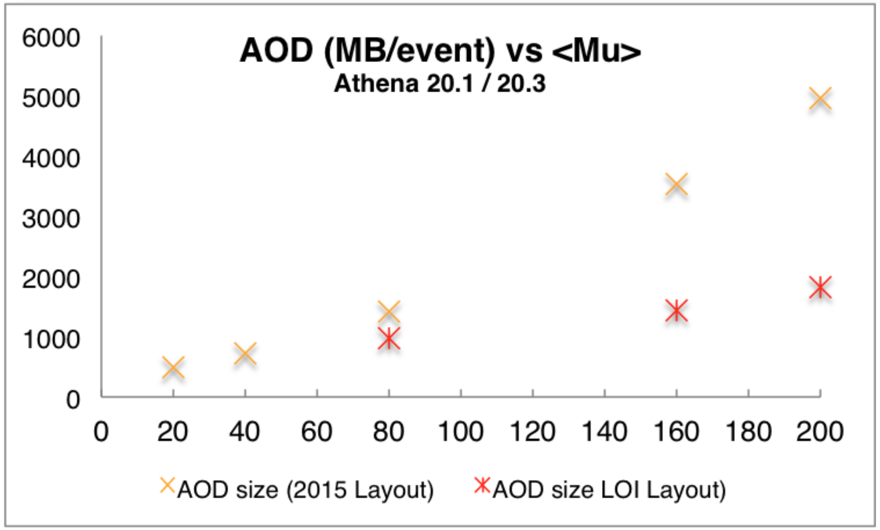
The average pile-up:

<mu>=14 in 2015
<mu>=23 in 2016
<mu> ≈ 35 in 2017
…
<mu> up to 200 in HL-LHC (10 years)

Higher pileup means:

Linear increase of digitization time
Factorial increase of Reco time
Larger events
Much more memory





Input parameters, assumptions, disclaimers

Input Parameters at HL-LHC
(LOI = the ATLAS Letter of Intent for Upgrade Phase-2)

Output HLT rate: 10kHz (5 to 10 kHZ in LOI)
Reco time: 288s/event, Simul Time: 454 s/event at mu=200
Nr Events MC / Nr Events Data = 2
Fast Simulation: 50% of MC events
LHC live seconds /year: 5.5M

Simplified Computing Model with respect to 2016/2017 resource requests:

Data from previous years not taken into account
=> Little difference at the beginning of the Run-4 but huge difference for Run-2 and Run-3

Projection of available resources in HL-LHC:

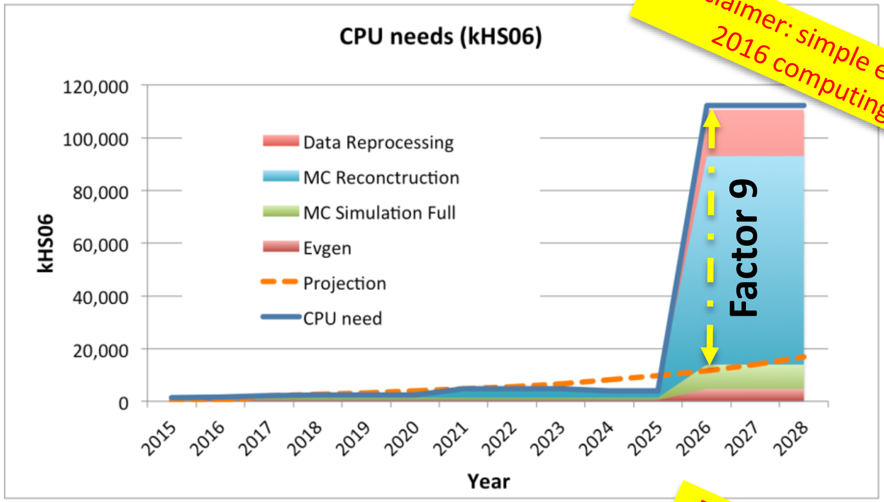
20% more CPU/year
15% more storage/year

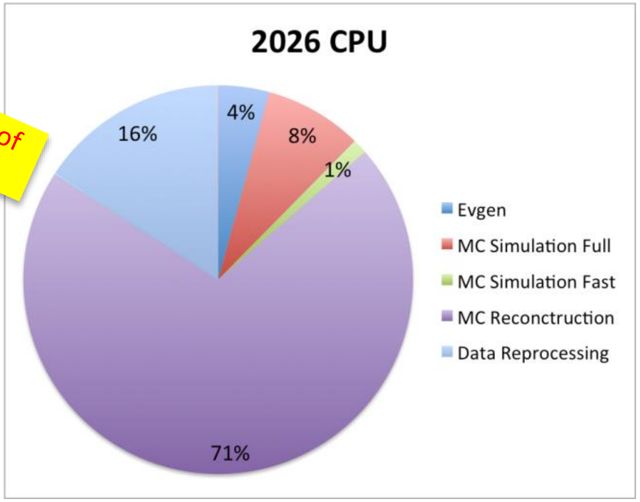
For the same cost

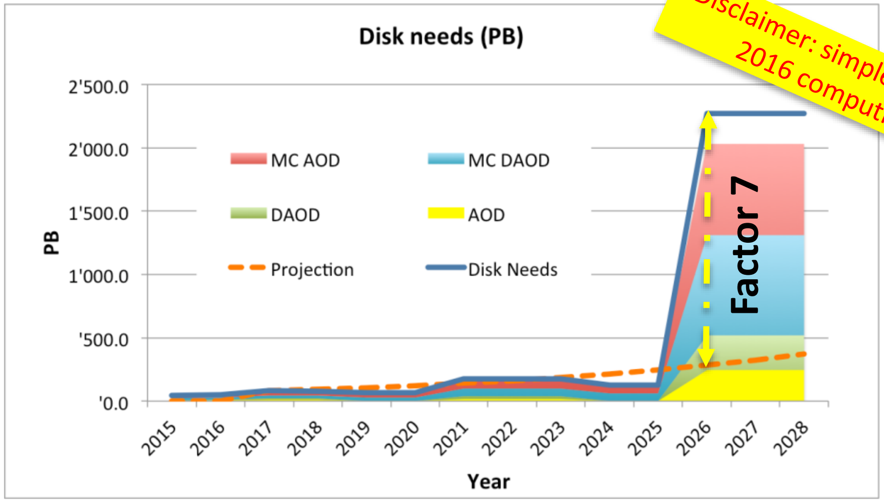
Projections evolve 2017 values
OF THIS SIMPLIFIED MODEL
(not the 2017 WLCG pledges)

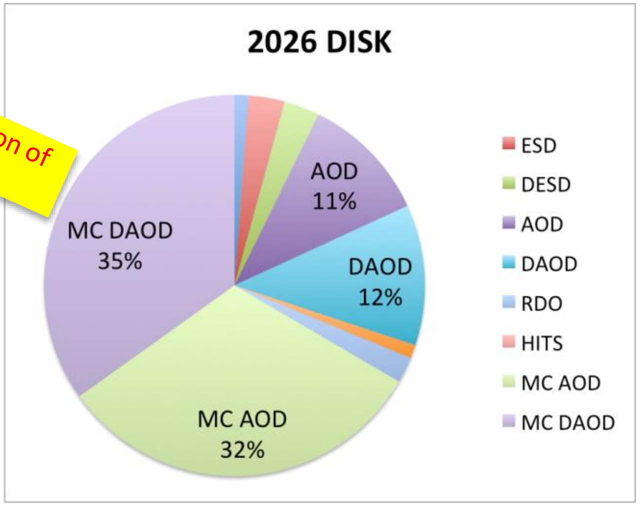
Conclusion: looking at absolute numbers makes little sense.
Relative differences between needs and projections at HL-LHC are meaningful. With caveats.

HL-LHC baseline resource needs









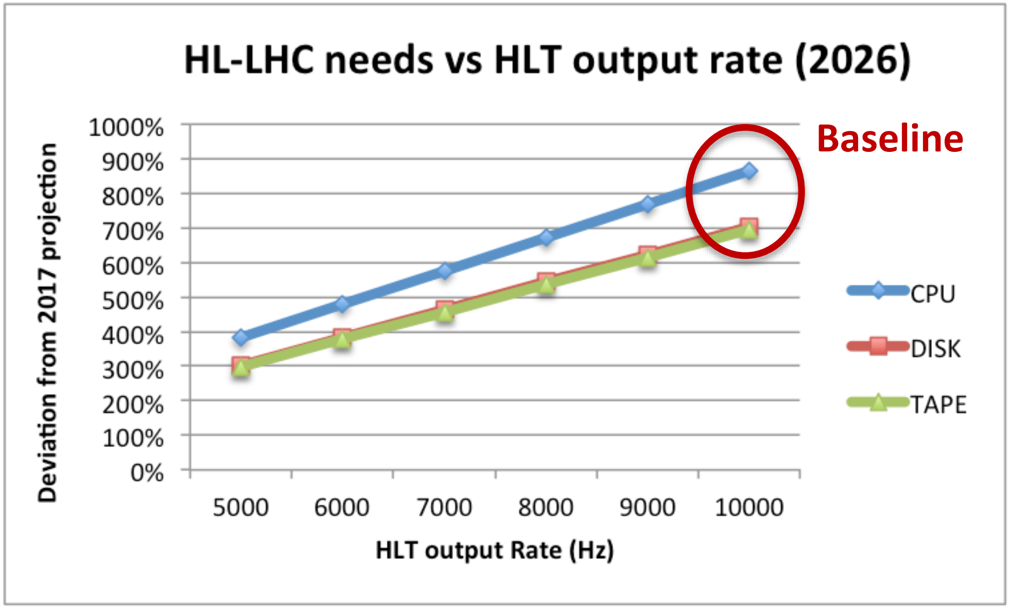
HLT output rate

The output trigger rate does not determine only the amount of data per year but also the amount of Monte Carlo to be produced.

The LOI foresees a value between 5 kHz and 10kHz. We use the latter as baseline in this study

The possibility to reduce the trigger rate to a lower value without impacting the ATLAS physics program will be analyzed in the years to come

If we consider the lower LOI limit (5kHz) the discrepancy with the projection of available resources reduces to x4 for CPU



Monte Carlo needs

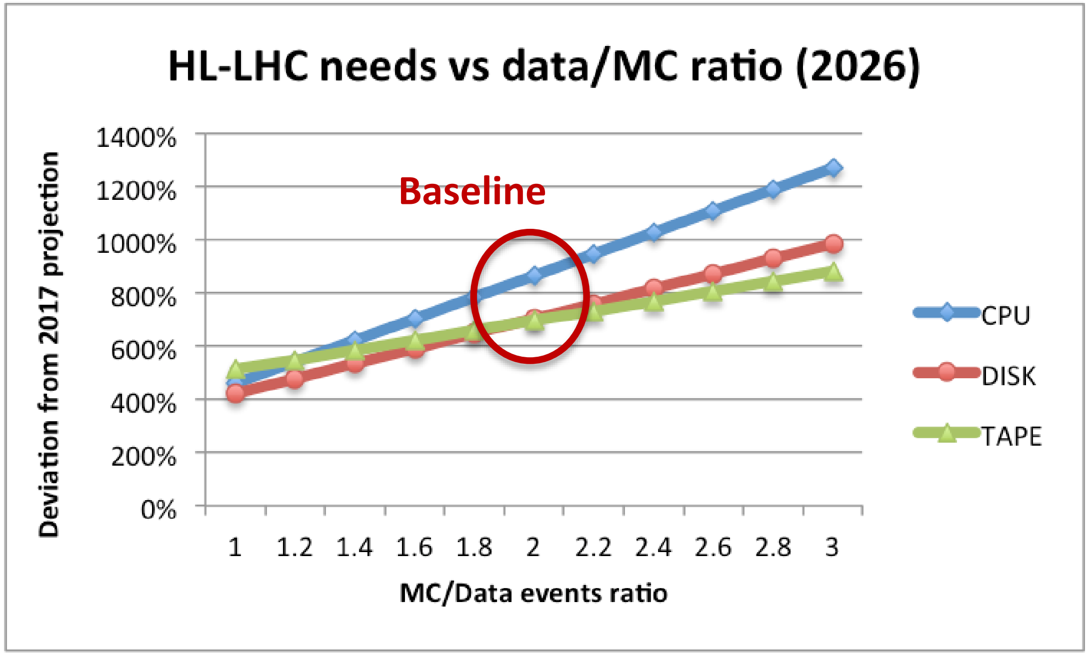
The physics case for HL-LHC will evolve in the next years. The high statistics of data collected in HL-LHC reduces the significance of statistical uncertainties. Therefore one might assume a lower need of MC with respect to data

HOWEVER

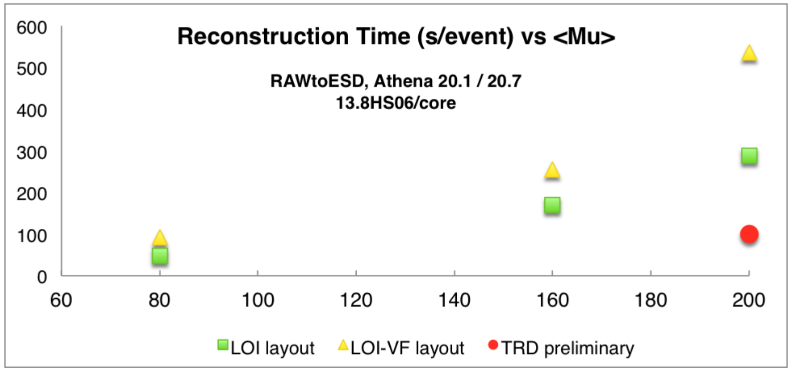
Things might change significantly once the physics case for HL-LHC evolves

Generators might become very expensive if we go to NNLO

In 2004 we expected a factor x0.3 MC with respect of data. We are at x2.0.



Layouts and Reconstruction



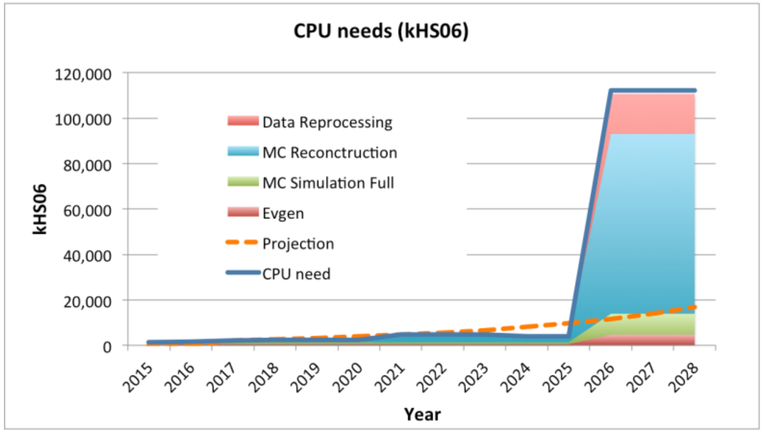
Reconstruction time dominates the CPU consumption in HL-LHC

Especially for MC, where trigger simulation utilizes the same offline algorithms (so it impacts twice as much)

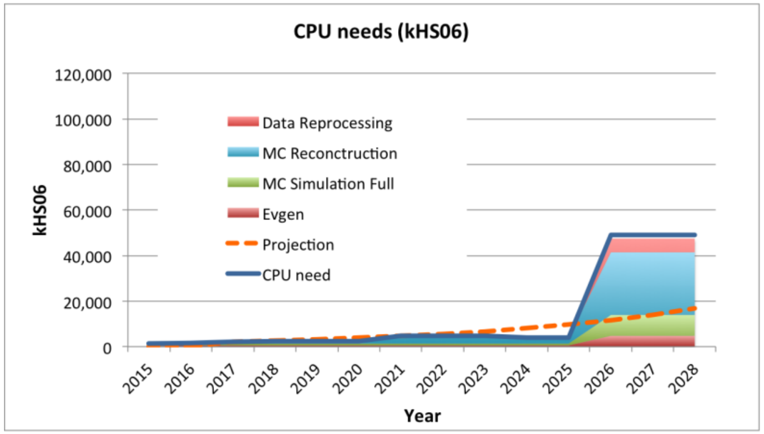
The detector layout will play an important role, together with the optimization/tuning of algorithms. Tracking will be the main consumer

Alternatives are also being investigated as R&D e.g. Machine Learning techniques

LOI Layout



Possible TDR Layout



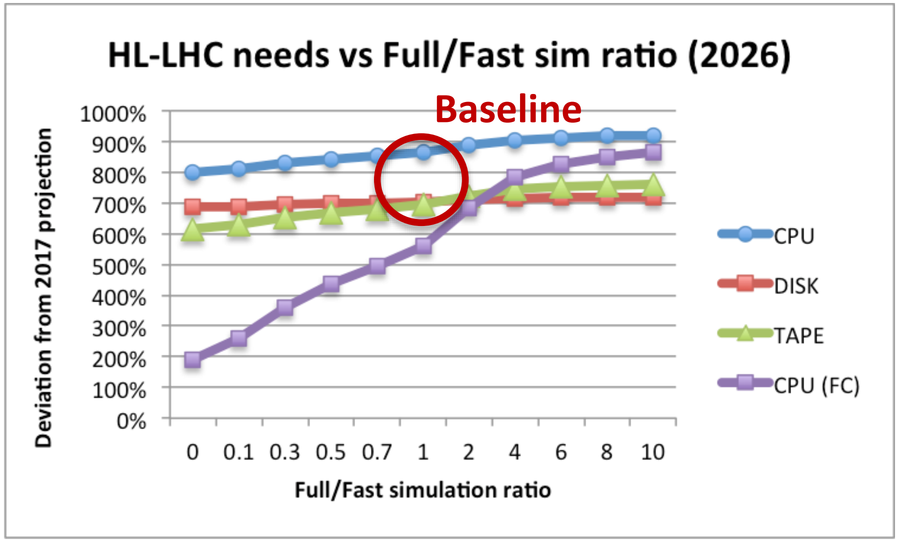
Fast Simulation and Fast Chain

Fast Simulation in Run-2 is x10 faster than Full Simulation (G4)

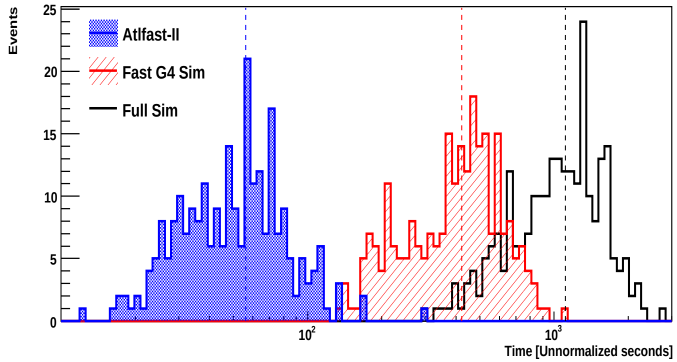
Fast Simulation can be used today only for a subset of analyses

Detector Simulation in general is not the driving cost in HL-LHC

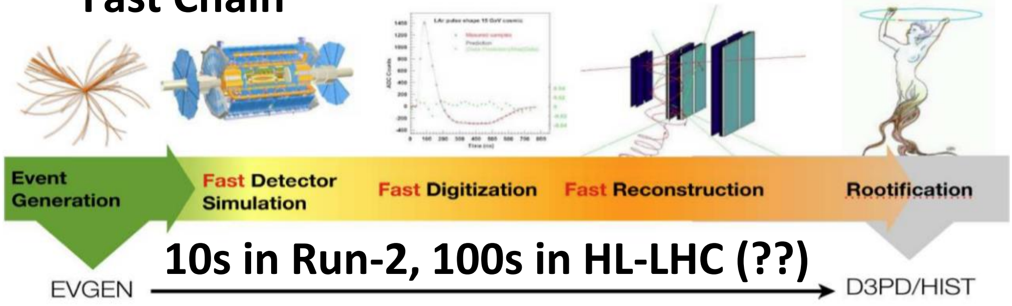
The gain will come with Fast Chain



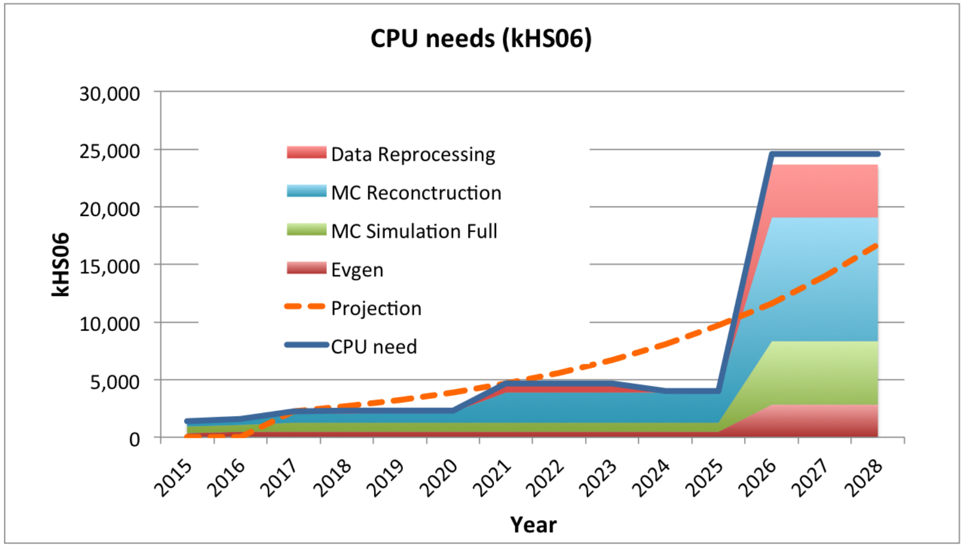
Fast Simulation



Fast Chain



If we want a very optimistic scenario …



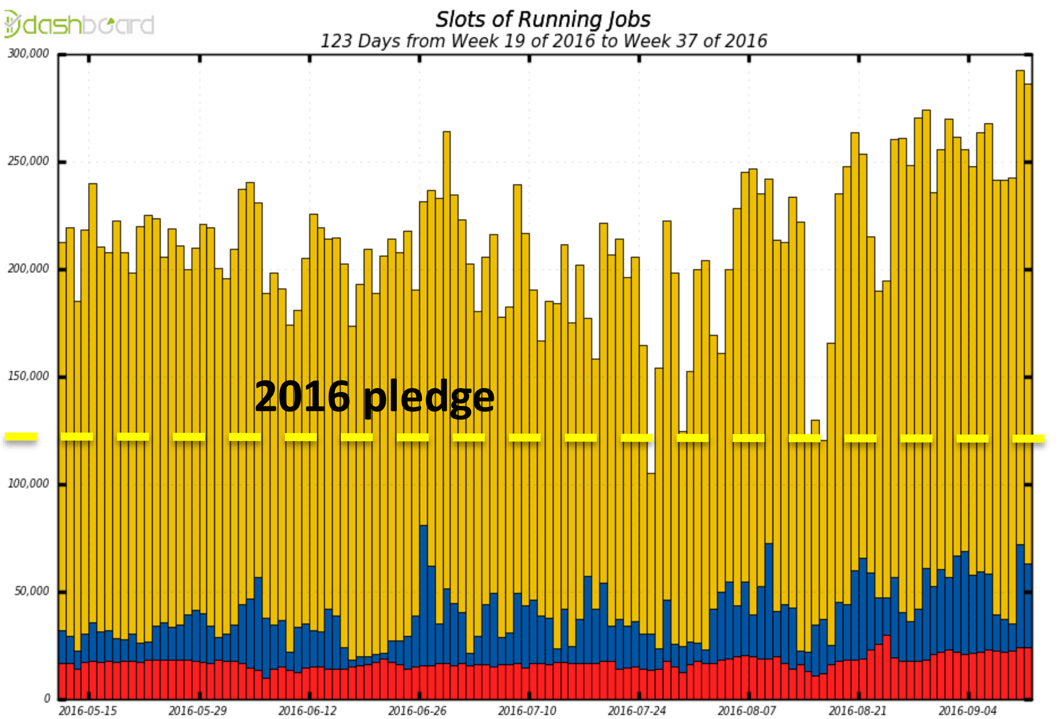
In a very optimistic scenario, the discrepancy for CPUs reduces to 200% (from almost 900%).

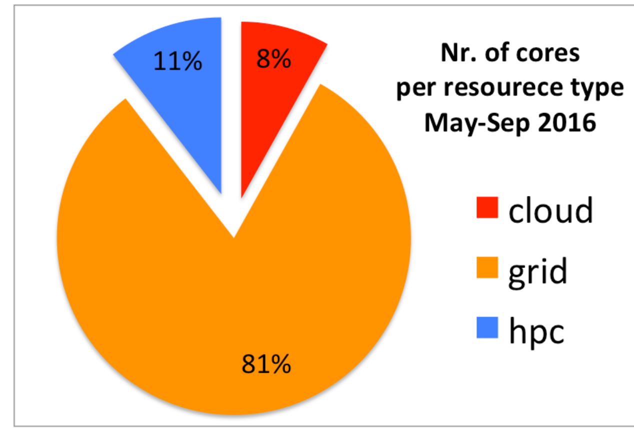
Which, given all the uncertainties, means problem solved

DO NOT GET TOO EXCITED AND LISTEN TO THE REST OF THE TALK

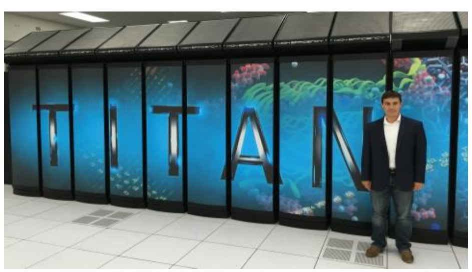
	Baseline Scenario	Optimistic Scenario
HLT output rate	10kHz	7.5kHz
Reco and Simul Time/Evt	from LOI	From preliminary TDR studies
Nr. Events MC / Nr. Events Data	2.0	1.5
Fast Simulation	50% of MC events	50% of MC events
Fast Chain	None	50% of MC events
LHC live seconds/year	5.5M	5.5M

Heterogeneous Resources



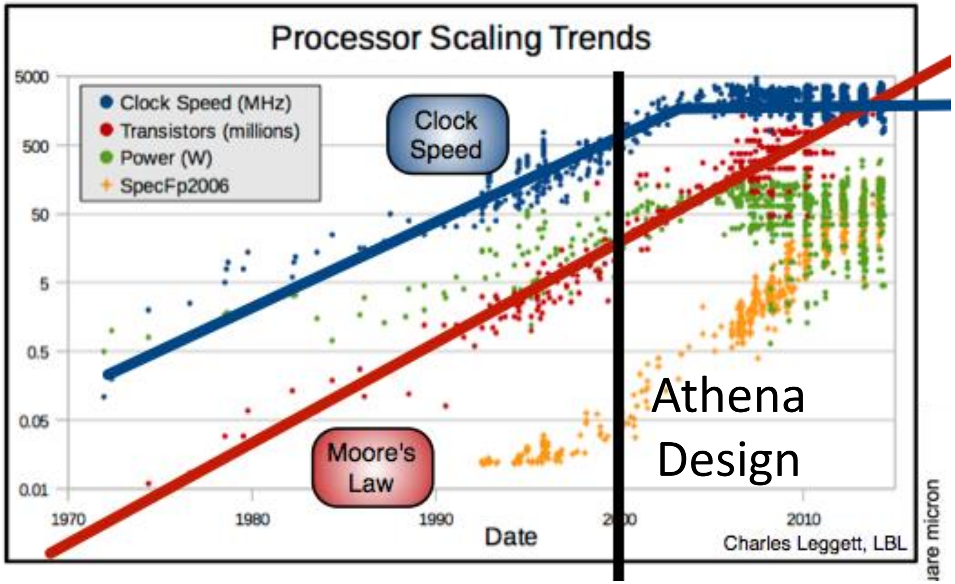


Integration of non Grid resources in ATLAS is a big investment with the potential of a big return



Challenges: resource provisioning, non standard architecture, GPU processing capacity, memory

Hardware trend and implications



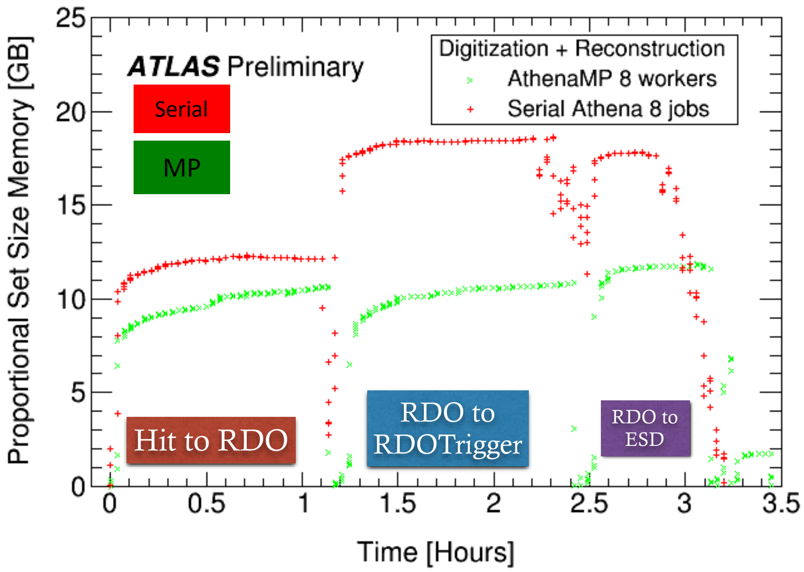
Clock Speed stalled but transistor density keeps increasing. Exploiting hardware becomes more complicated (vectors, memory…)

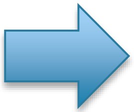
Example: Cori@NERSC (Intel Knights Landing)
1PB of Memory, 9304 nodes
68 cores/node, 4 HW threads/core
=> Approx 300 MB/thread

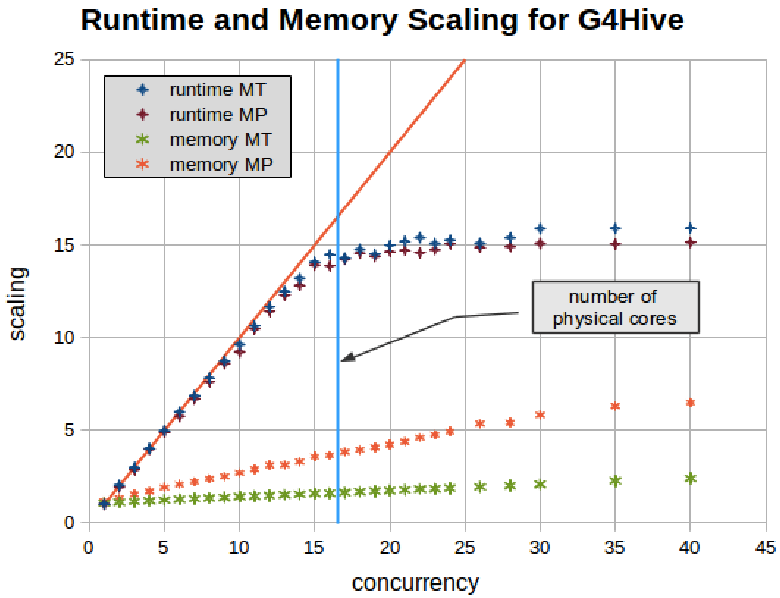
From Multi Processing to Multi Threading

AthenaMP (multiprocessing) will not be sufficient anymore. We will need (and we are developing) AthenaMT (multithreading). Will be in production for Run-3 (2020) already.

Parallel processing in a multithreaded environment will come with its challenges both for developers, operations and infrastructures

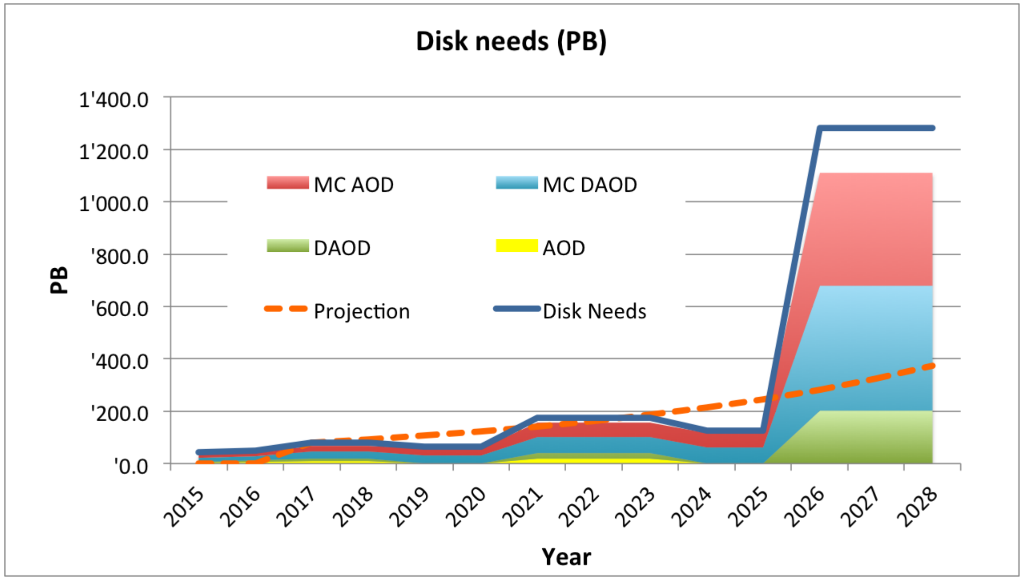






What about Storage ?

Optimistic Scenario + No AOD on disk



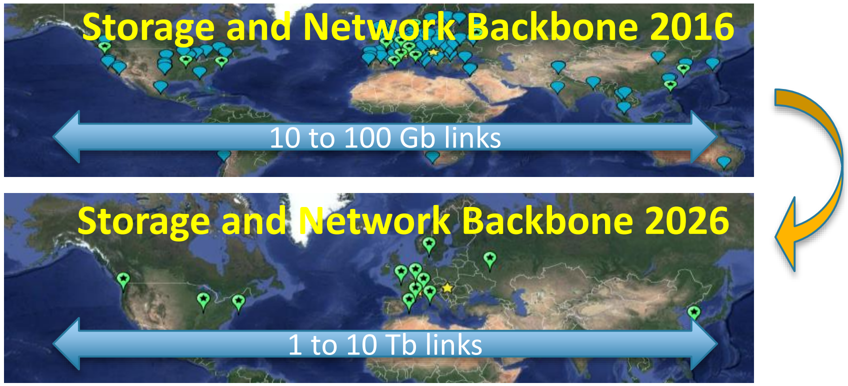
Even in the optimistic scenario, we are still far from solving the problem

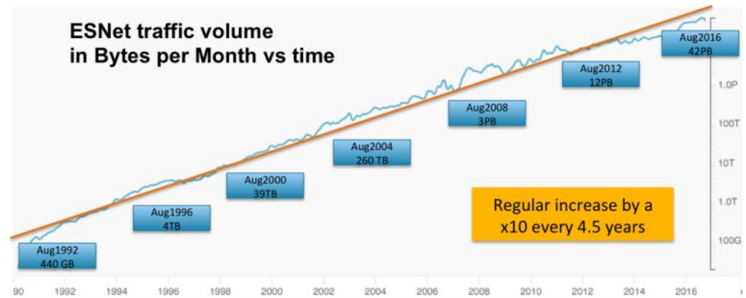
AODs and DAODs are the main consumers.

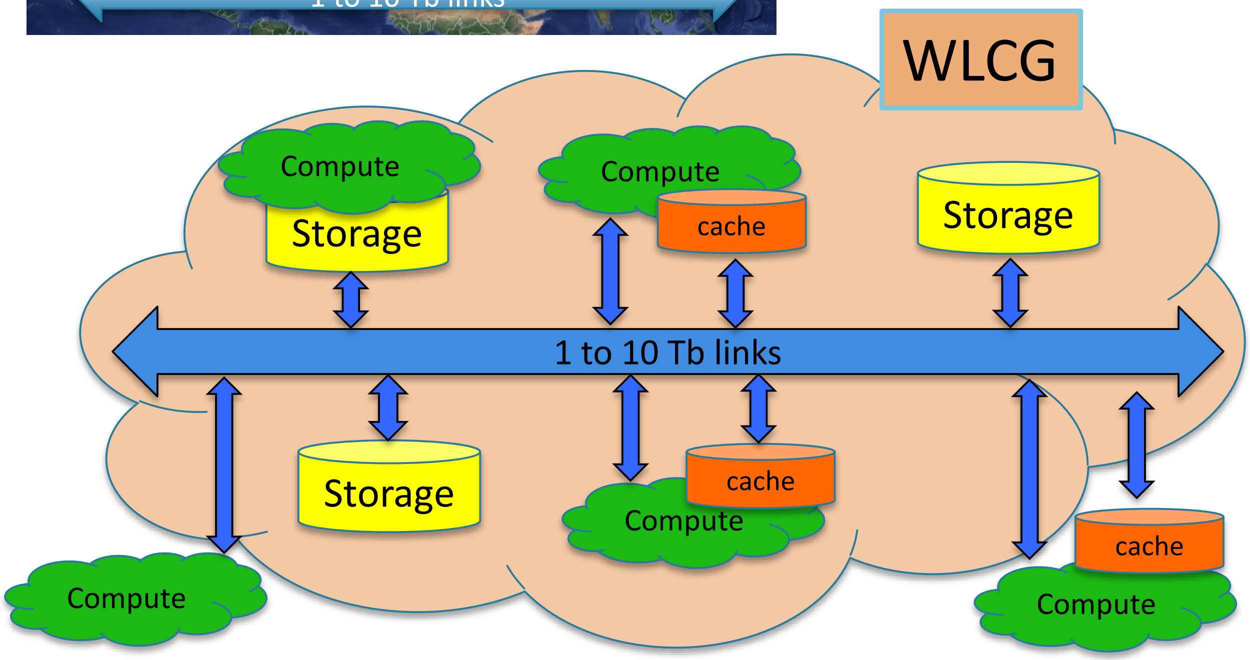
With no AOD on disk (run Train Analysis from AODs on TAPE) you get x4 above the resource projection

The remaining gain must come from re-thinking of distributed data management, distributed storage and data access. A network driven data model allows to reduce the amount of storage, particularly for disk. Tape today costs at least 4 times less than disk.

Computing infrastructure in HL-LHC







A data cloud for science

Storage and Compute loosely coupled but connected through a fast network

Heterogeneous Computing facilities (Grid/Cloud/HPC/ …) both in and outside the cloud

Different centers with different capabilities, fo different use cases

Data Management: Challenges and Opportunities

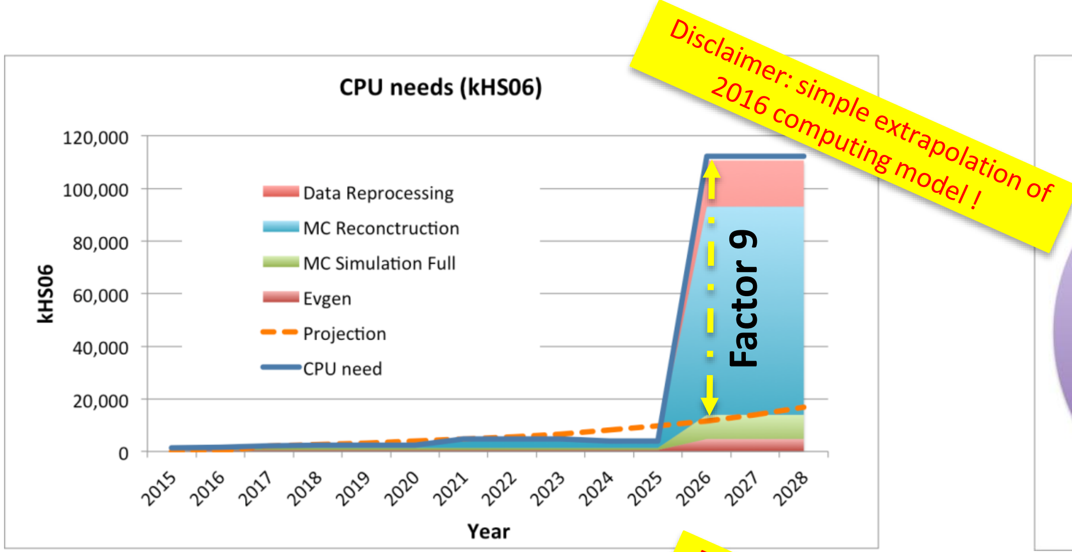
- "Funny how tape never seems like the cheap option when you have to pay for it". One could say the same about network
- A fast WAN does not imply fast data access. The infrastructure and the I/O layers need to be optimized from end to end
- Multilevel caching should be built **IN** the infrastructure rather than **ON** top of it
- A unique opportunity to define and implement a common data management and data access layer
- Today WLCG is a data Grid. Tomorrow we will have a data cloud
  The challenge is always the data

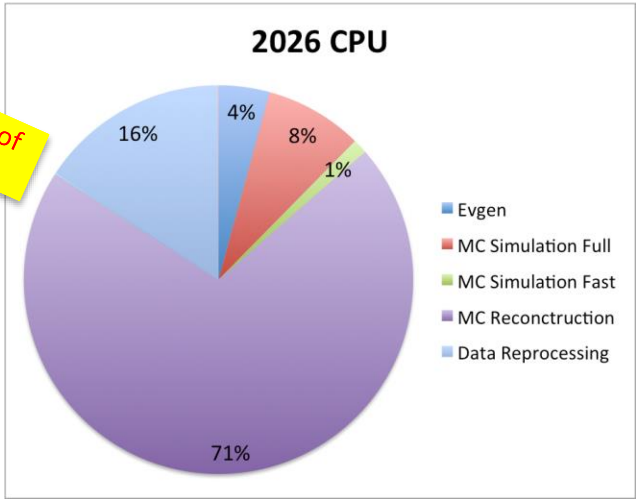
Conclusions

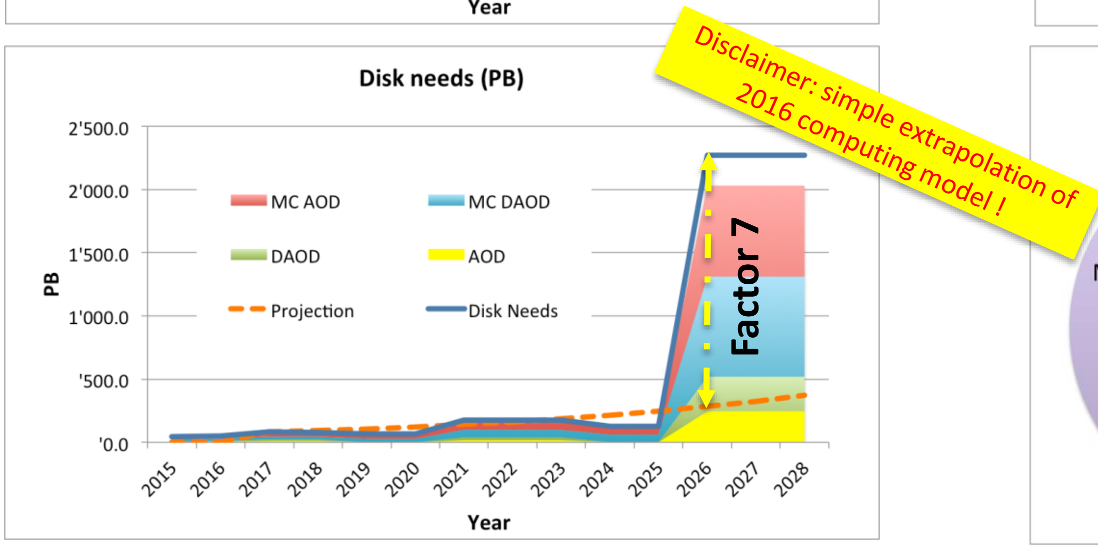
- We identified a concrete set of steps in preparation for computing at HL-LHC
- To keep cost of computing under control in 2026 we need to invest effort from now
- The effort spans many areas: online, offline software, distributed computing, physics, infrastructure and facilities. The detector layout will play a crucial role
- It is important to consider cost of computing when choices are made
- We are on schedule to define a computing model for HL-LHC in the next two years

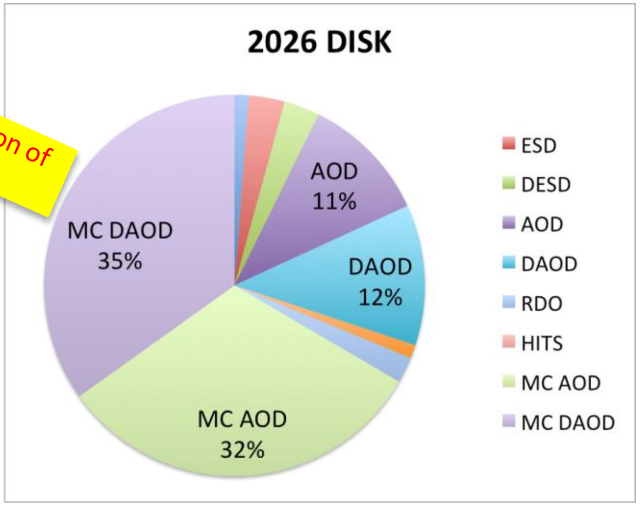
HIGHLIGHTS

HL-LHC baseline resource needs

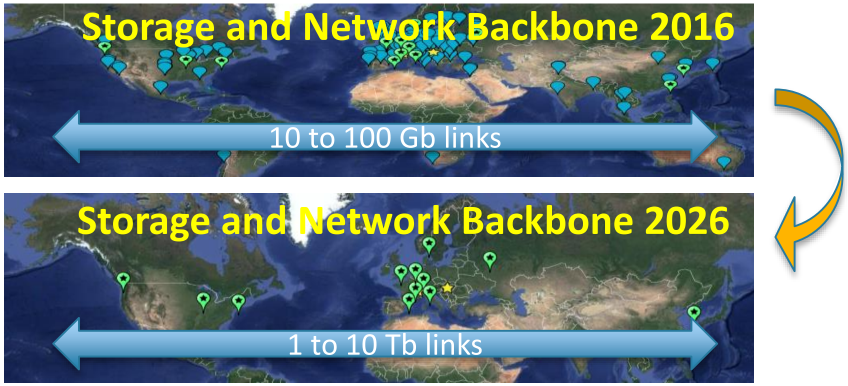


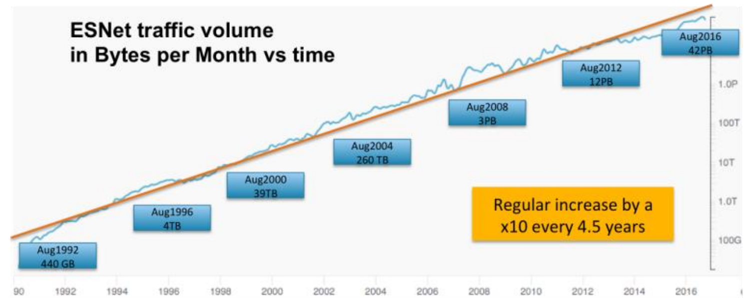


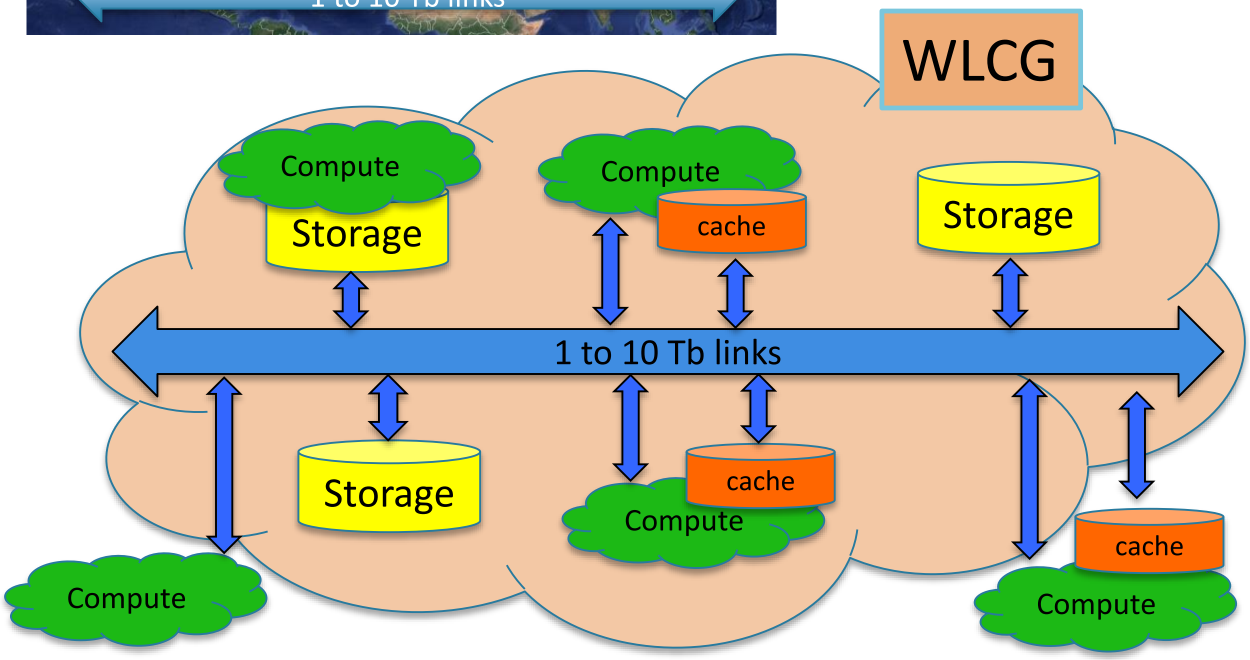




Computing infrastructure in HL-LHC







A data cloud for science

Storage and Compute loosely coupled but connected through a fast network

Heterogeneous Computing facilities (Grid/Cloud/HPC/ …) both in and outside the cloud

Different centers with different capabilities, fo different use cases