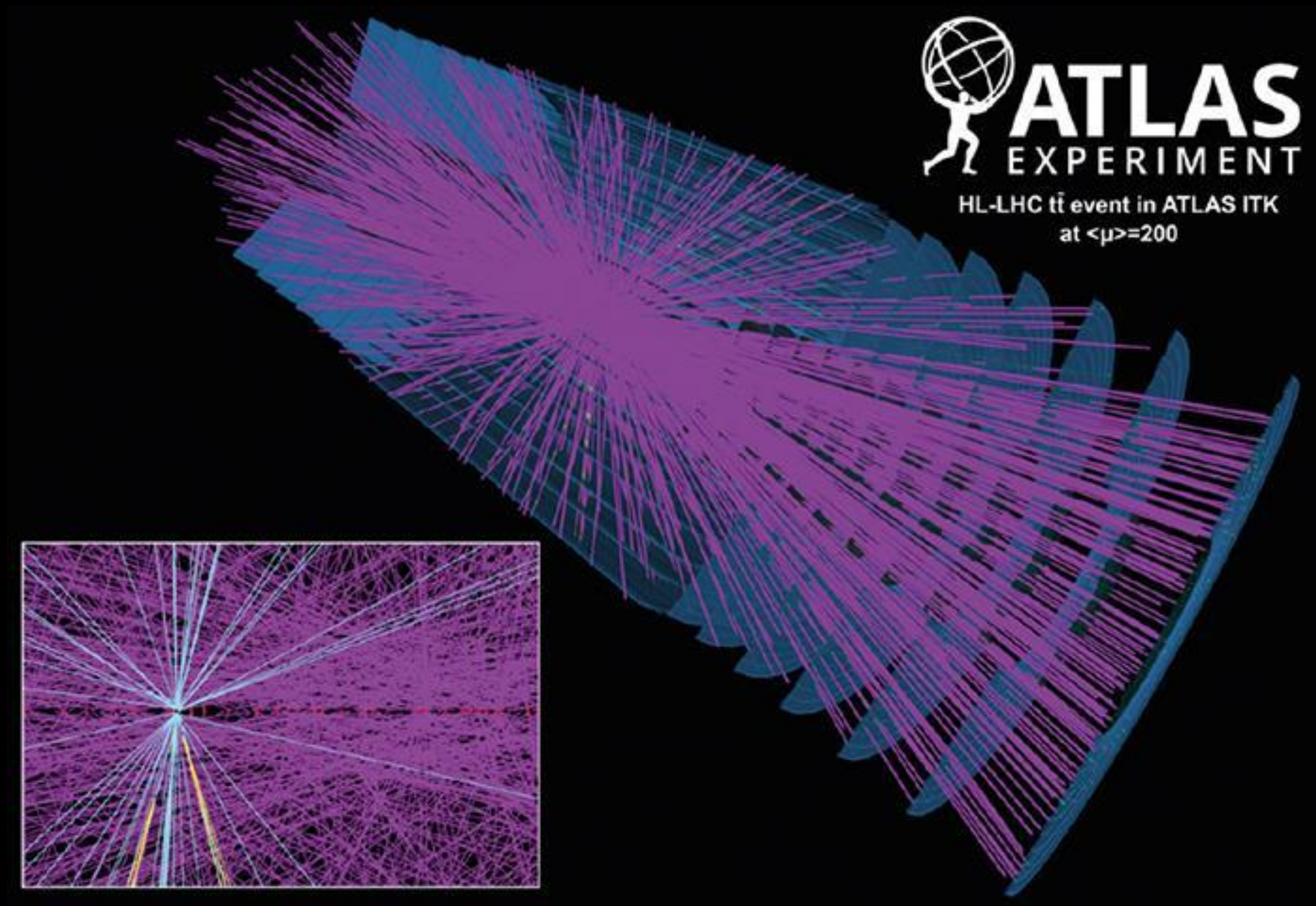
Software and Computing up to 2025

David Cameron (University of Oslo)
Slides mainly taken from James Catmore's talk @ NorduGrid 2018

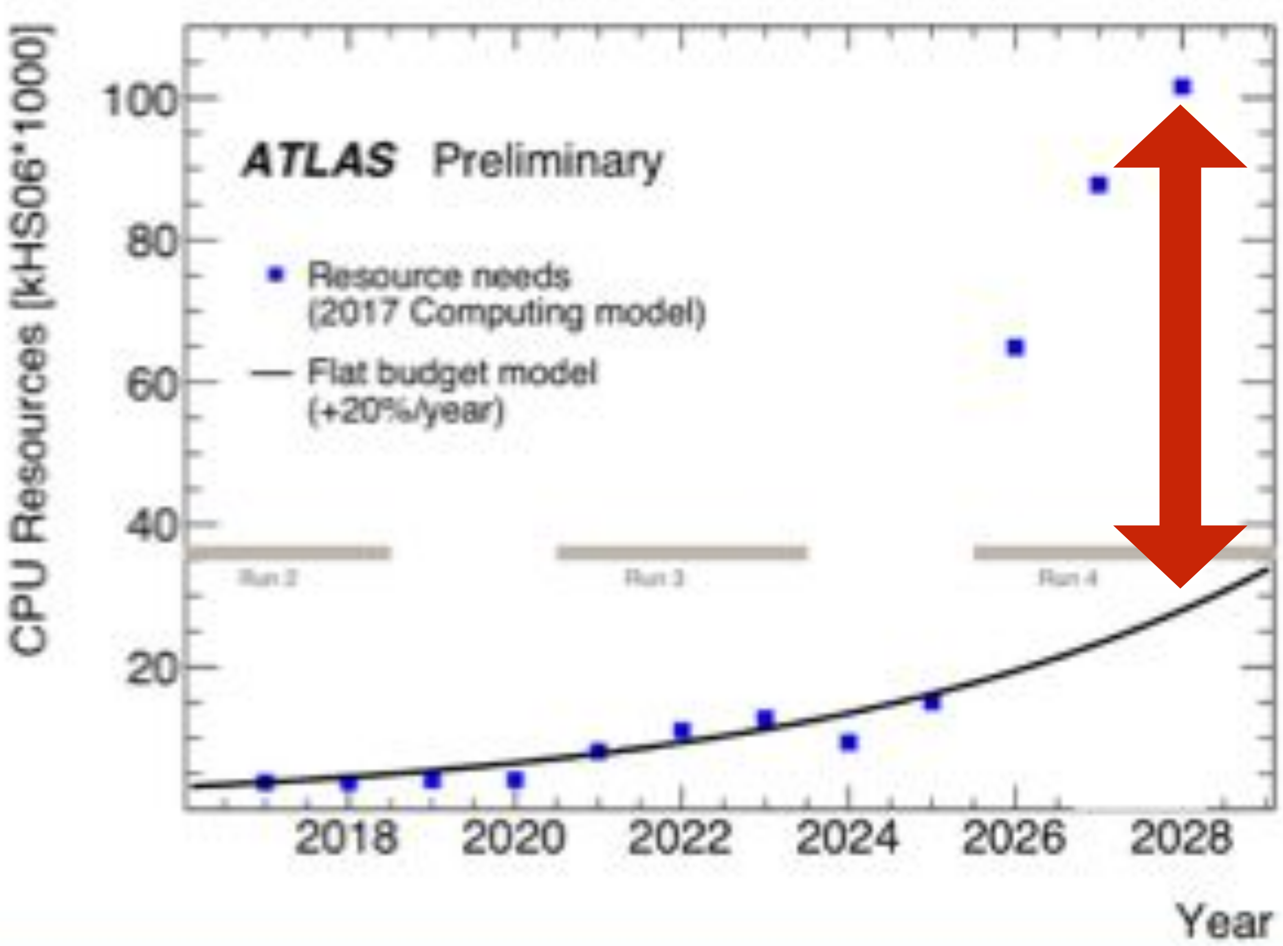


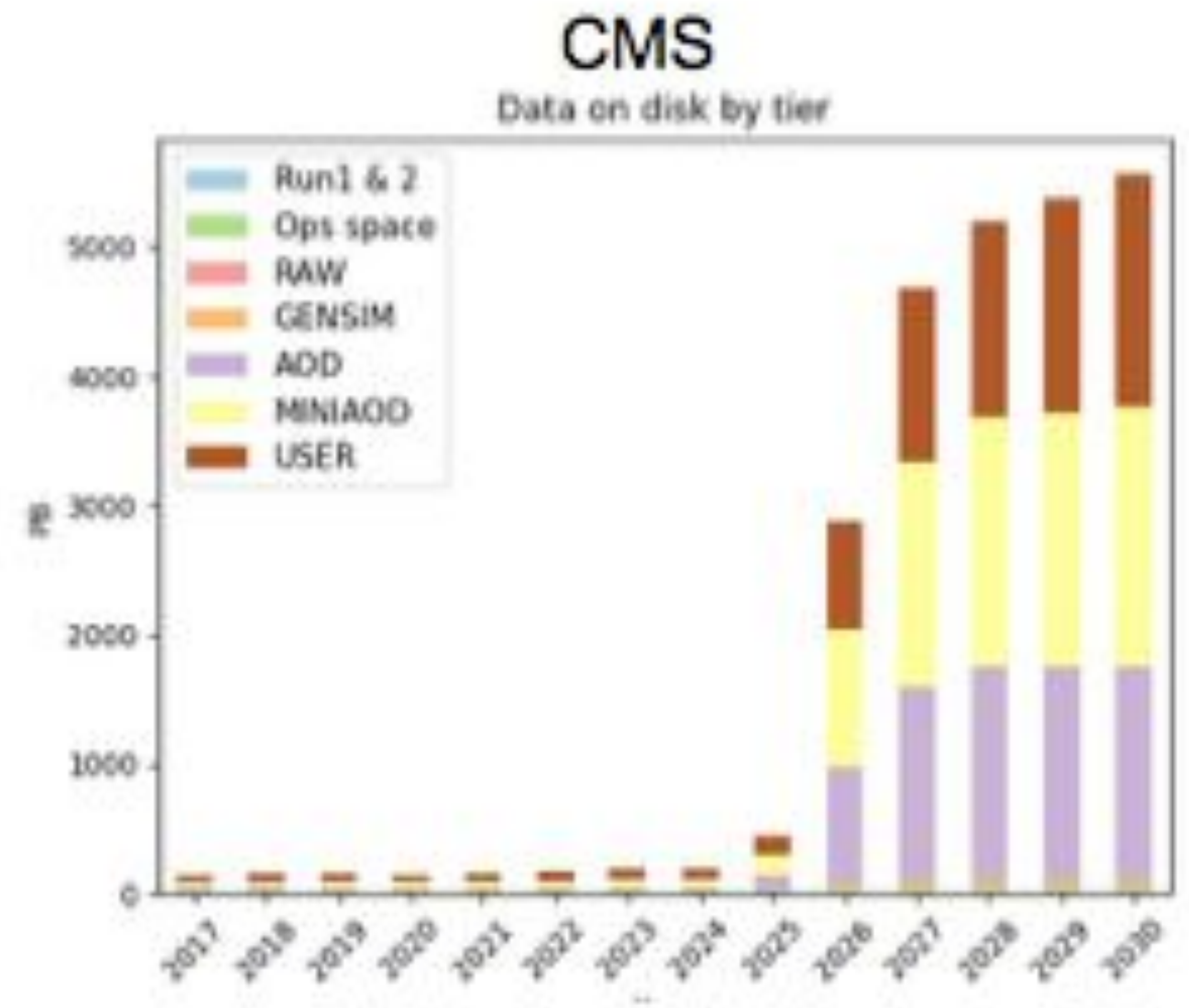
LHC Runs 3 and 4

	Run 2	Run 3	Run 4-5 (HL-LHC)
Year	2015-2018	2021-2023	2026++
CoM energy (TeV)	13	14	14
Lumi/Nominal	1.0 (2015-2016) 2.5 (2017-2018)	2.5	5-7
$\int L$ at end of run (fb^{-1})	150	300	3000

Increasing data volumes, rate and event complexity

Increasing complexity and rate





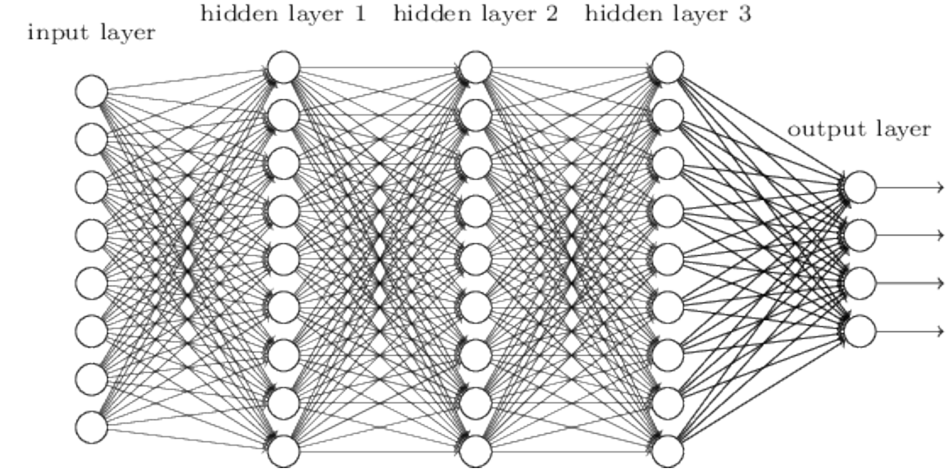
Rest of run 2 and run 3: smart & efficient use of existing model will see us through

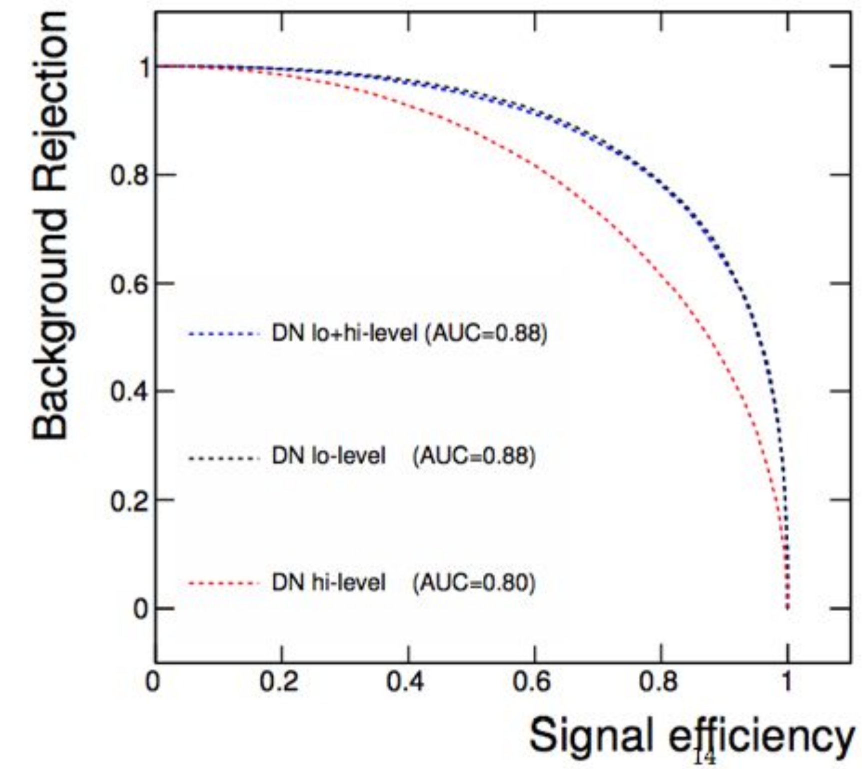
Run 4: <μ> ~ 200 and much higher rate - need to do things very differently → we must be able to make full use of evolving technology to have a hope of keeping up with the HL-LHC

This means that our software has to change radically

How to address the challenge

- Parallelism: CMS is already multi-threaded; ATLAS has a multi-process framework in use (AthenaMP) and plans to be multi-threaded in time for Run-3 (AthenaMT)
- Some elements of the frameworks are shared between experiments: further sharing could reduce the workload
- Machine learning
  - Long history in HEP: BDTs and shallow neural networks used since the 1980s
  - More powerful computers and bigger training datasets have led to the growth of "deep learning" and an accompanying cluster of high-performance open source tools from outside HEP
  - Possibly this has revolutionary potential at all levels of the field
    - Simulation, reconstruction, physics analysis, automation of shift work, optimisation of computing resources, anomaly detection

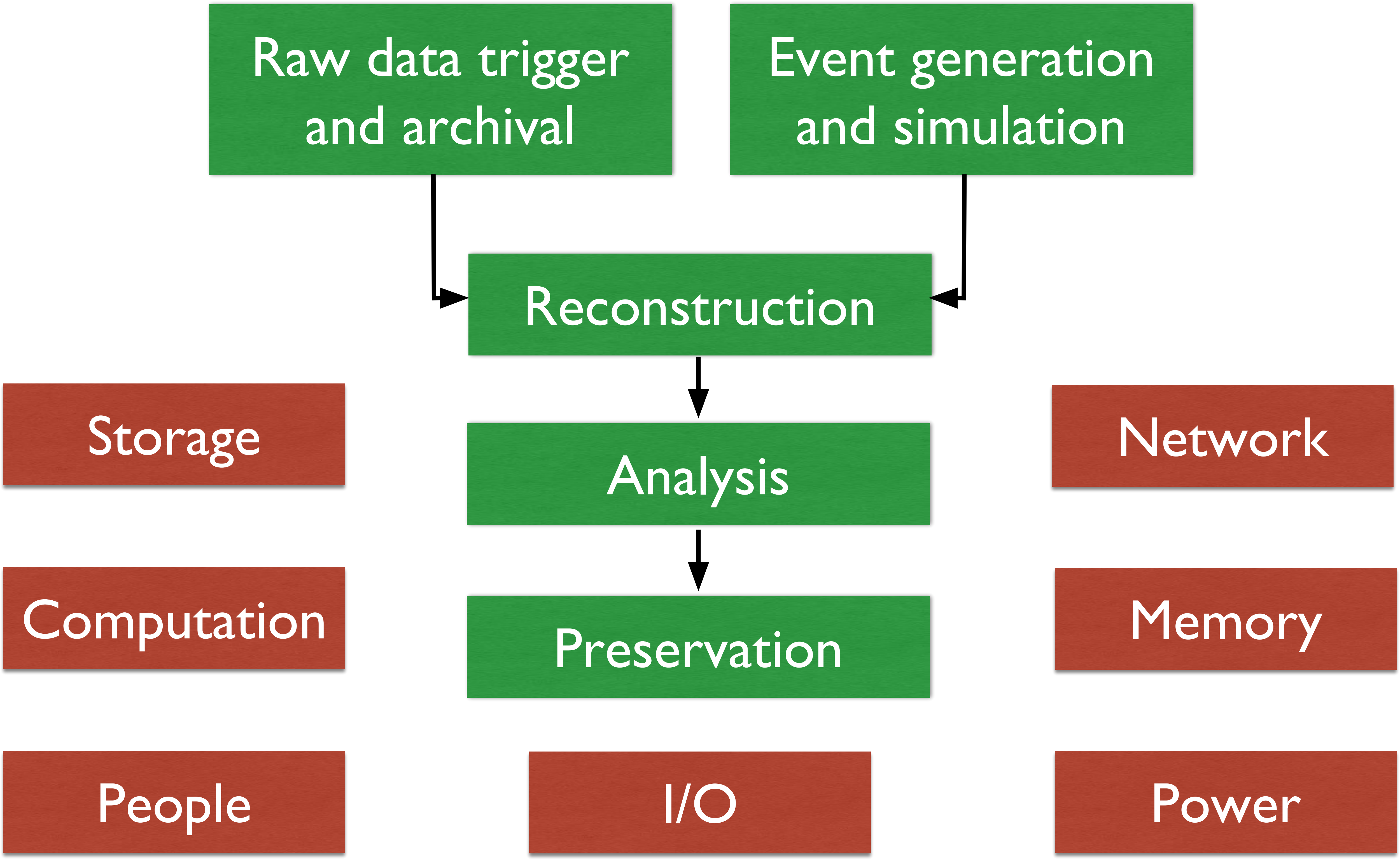




Facilities and data management

- Storage and computing are overwhelmingly from WLCG resources
  - This is expected to continue into Run 4, but with other resources in the mix
- Volunteer computing, commercial clouds, HPCs
  - Need to ensure our software can work seamlessly in these environments
- Strengthen links with other big data sciences especially with regards to sharing network resources
- Storage is the major challenge for HL-LHC
  - Sheer volume
  - How to support fast access for analysis and machine learning?
  - Different storage models: data lakes, quality of service
  - Technologies: more SSD, less disk? Relative costs of tape?

Activities and Limitations



We should aim to minimise the limitations that computing and software impose on our ability to do physics research

HEP Software Foundation: facilitates cooperation and common efforts in High Energy Physics software and computing internationally

A Roadmap for HEP Software and Computing R&D for the 2020s: https://arxiv.org/abs/1712.06982

Advancing from here

- Community White Paper process has been a success
  - Engaged more than 250 people and produced more than 300 pages of detailed description in many areas
- Summary roadmap lays out a path forward and identifies the main areas we need to invest in for the future for our software upgrade
  - Supporting the HL-LHC Computing TDRs and NSF S2I2 strategic plan
  - You can still sign :-)
- HEP Software Foundation has proved its worth in delivering this CWP Roadmap
  - Achieving a *useful* community consensus is not an easy process
  - Sign up to our forum to keep in touch and get involved (hep-sf-forum@googlegroups.com)
- We now need to marshal the R&D efforts in the community, refocusing our current effort and helping to attract new investment in critical areas
  - The challenges are formidable, working together will be the most efficacious way to succeed
  - HSF will play a vital role in spreading knowledge of new initiatives, encouraging collaboration and monitoring progress
  - Next HSF workshop in March, shared with WLCG, should start to put our ideas into practice:
    - C++ Concurrency, Workload Management and Frameworks, Facilities Evolution, Analysis Facilities, Training, ...

Norwegian computing infrastructure for HEPP

Work package number	**3**	**Lead partner**	**UiO**
Work package title	ATLAS and ALICE Computing upgrade		
Partners	UiO, UiB, HiB		
Start month	2017.07.01	**End month**	2022.06.30

Objectives:

- Upgrade and scaling up of the Norwegian ALICE and ATLAS computing and integration into the Nordic Tier-1
- Integrate the Norwegian part of the WLCG into the national computing infrastructure, with the long term goal to match the Exascale nature and complexity of the extreme conditions prevailing at the high luminosity LHC

Description of work (tasks, lead partner and role of participants):

Task 3.1.1: ALICE computing, new investments, integration. Lead partner: UiB. (M2-M13)

Task 3.1.2: ALICE computing Stage 2, same work, partner, etc as 3.1.1 (M26-M37)

Task 3.2.1: ATLAS computing, new investments, integration. Lead partner: UiO. (M2-M13)

Task 3.2.2: ATLAS computing, Stage 2, same work partner, etc as 3.2.1 (M26-M37)

Deliverables (brief description and month of delivery measured from the project start):

D3.1.1: Installation of 50% (Stage 1) of computing hardware for ALICE. M13

D3.1.2: Full installation (Stage 2) of computing hardware for ALICE. M37

D3.2.1: Installation (Stage 1) of 50% of computing hardware for ATLAS. M13

D3.2.2: Full installation (Stage 2) of computing hardware for ATLAS. M37

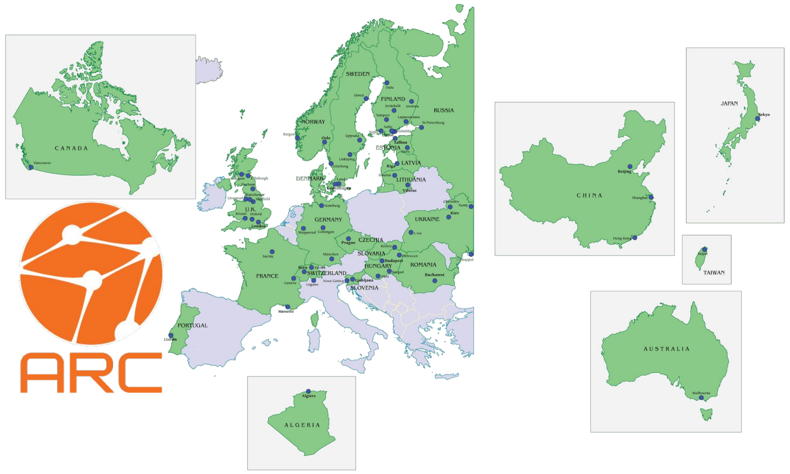
Future resources

	2017	2017	2018-2022	2018-2022
	ATLAS	ALICE	ATLAS	ALICE
CPU (kSpec)	11.2	10.69	22.4	20.6
Disk (pBytes)	1.20	1.21	3.0	1.93
Tape (pBytes)	2.84	1.37	3.8	2.55

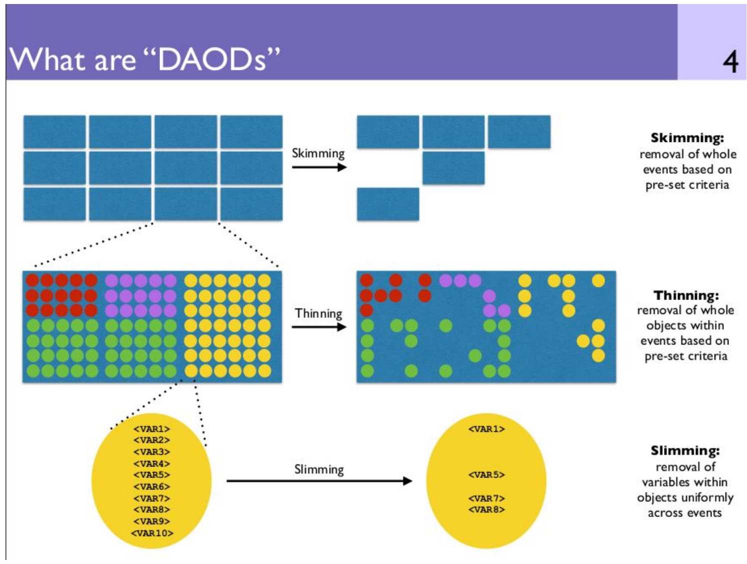
- Also exploring how to exploit different kinds of resources
  - Clouds, HPC (PRACE)

Norwegian S&C activities

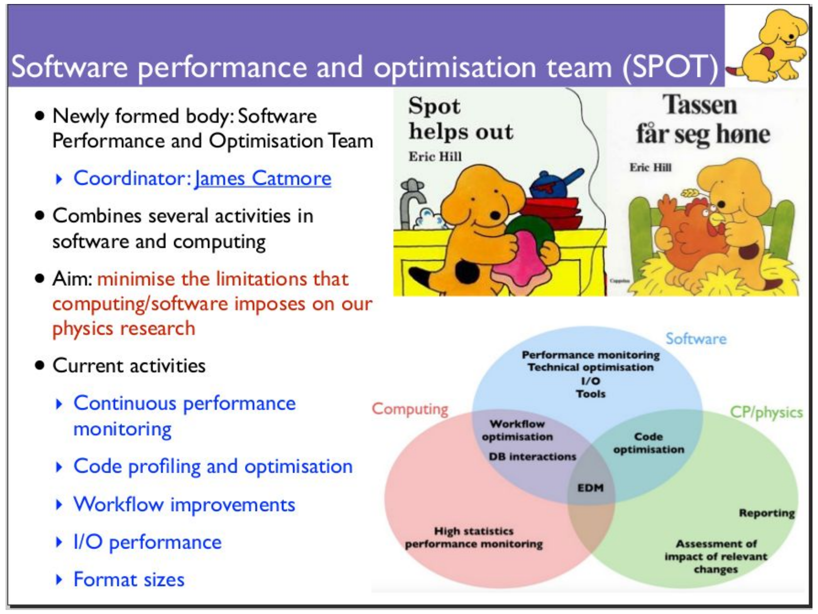
Rucio: Data management for ATLAS ++



ARC middleware



Derivation framework



Key roles in S&C activities in ATLAS

All of these must continue to evolve and expand to meet future challenges

People and training

- Not controversial to say that most experiments have a desperate shortage of people willing and able to work on software and computing
  - This comes at a time when we can't rely on technology improvements to keep us afloat
- We need to support
  - People who eventually plan to go into industry → need to stay relevant and provide training in modern technologies that are transferable to the commercial or public sectors
  - People who want long-term careers in HEP → need to recognise SW&C work as equal to detector and physics analysis work
    - As things become more complicated we need physics leaders with strong interests in SW&C
- We need to improve our citation and publication record
- We must invest in training our community at different levels from basic analysis to advanced software engineering
- Collaboration with those from other academic fields and industry is important in this regard