US ATLAS Overview
Plans and Resource Gaps for HL-LHC

Jana Schaarschmidt (Washington), Torre Wenaus (BNL),

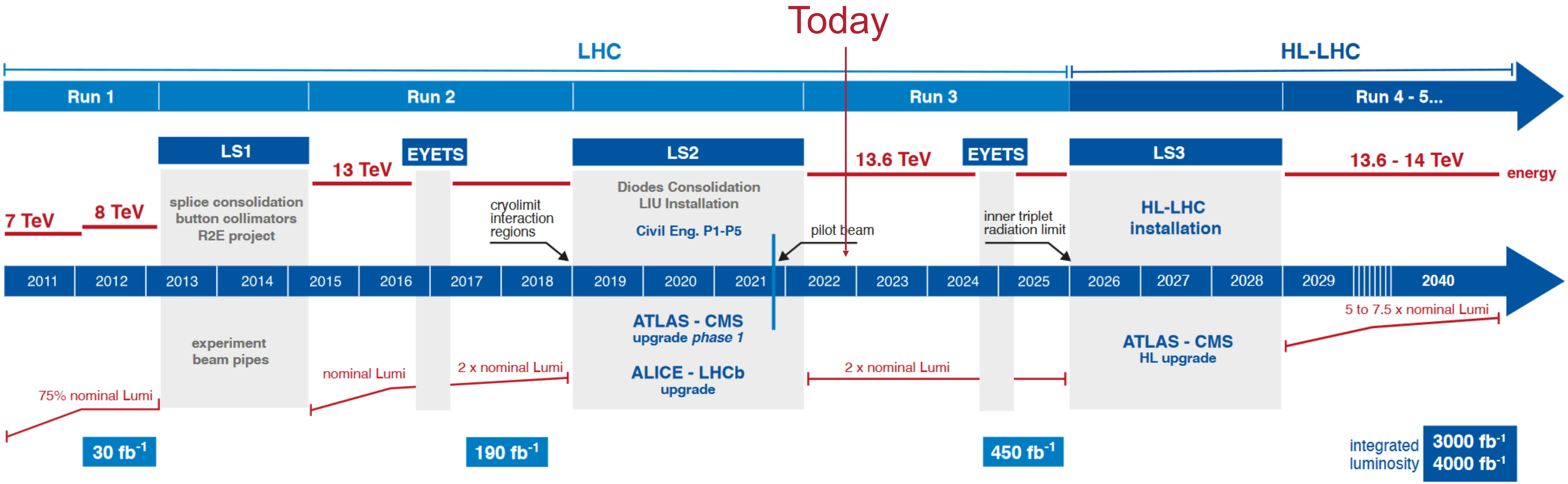
Verena Martinez Outschoorn (UMass Amherst), Paolo Calafiura (LBNL)

A Coordinated Ecosystem for HL-LHC Computing R&D

Washington DC

7-9 November 2022

LHC Schedule



LHC

Run1: 25 fb^{-1} usable data at 7-8 TeV

Run2: 140 fb^{-1} usable data at 13 TeV

Run3: Expect to take ~250 fb^{-1} of data at 13.6 TeV

→ ~10% of the total LHC dataset

HL-LHC brings unprecedented physics opportunities and computing challenges

HL-LHC

Run 4 & beyond: 3-4 ab^{-1} at 13.6-14 TeV

→ ~90% of the total LHC dataset

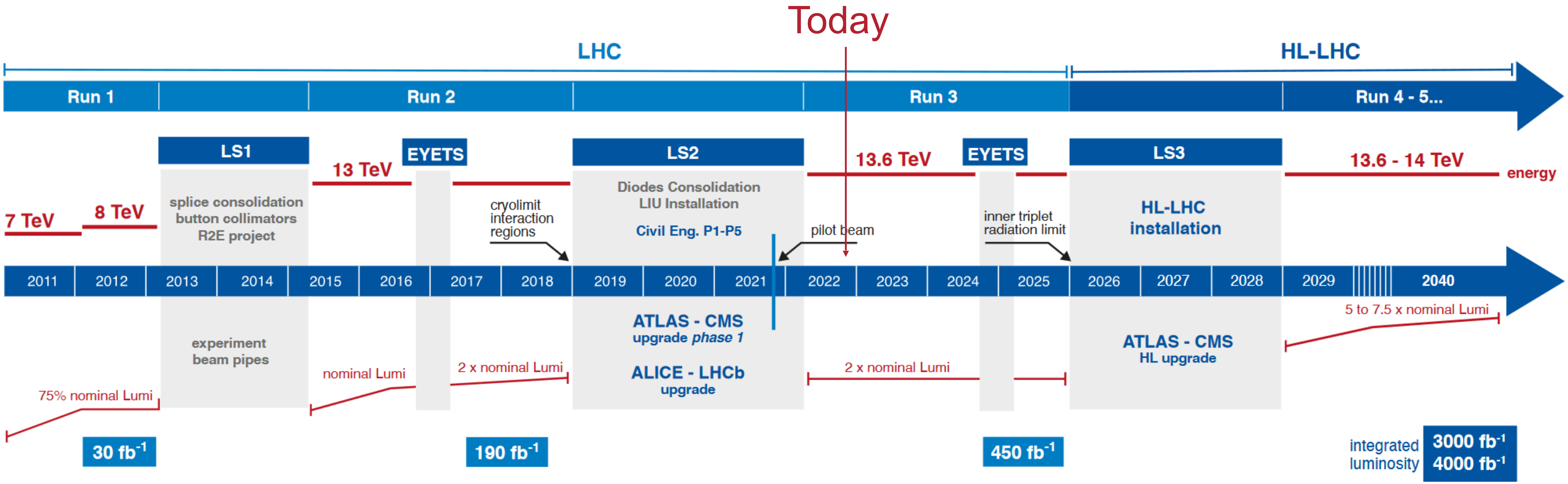
Data processing challenges

5-7x increase in luminosity (LHC upgrade)

4-5x increase in event size (new detectors)

10x increase in event rate (trigger upgrade)

ATLAS Software & Computing Goals



Critical for ATLAS physics program that by the start of Run 4 software and computing is ready to

- process incoming data
- store it with sufficient redundancy
- produce enough simulated MC events
- provide efficient analysis infrastructure

HL-LHC

Run 4 & beyond: 3-4 ab^{-1} at 13.6-14 TeV
→ ~90% of the total LHC dataset

Data processing challenges

5-7x increase in luminosity (LHC upgrade)
4-5x increase in event size (new detectors)
10x increase in event rate (trigger upgrade)

ATLAS Software & Computing HL-LHC Timeline

US ATLAS

Conceptual Design Report

HL-LHC Roadmap (Road to Run 4 or R2R4)

Technical Design Report

Run 4 Feature Freeze

Run 4 Performance Freeze

Run 4 Release Validated

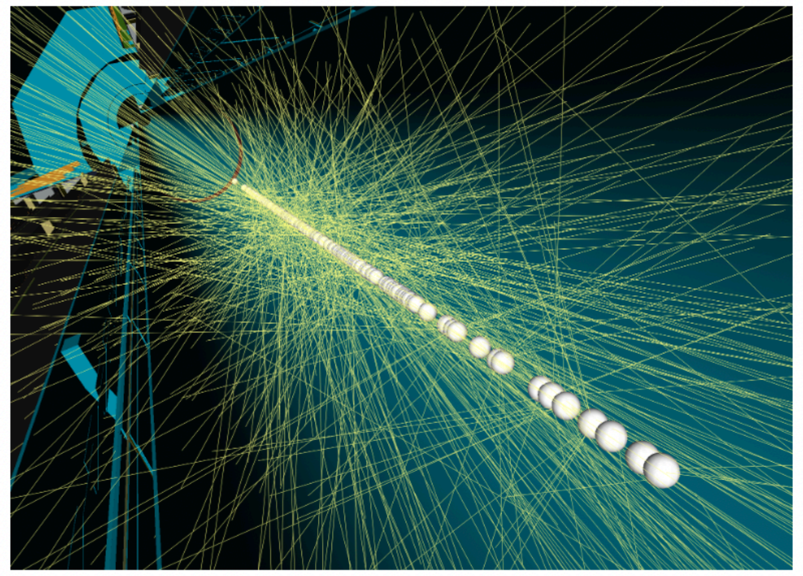
2020 2021 2022 2023 2024 2025 2026 2027 2028 2029

Ready for HL-LHC data-taking

ATLAS EXPERIMENT

HL-LHC CDR

ATLAS HL-LHC Computing Conceptual Design Report



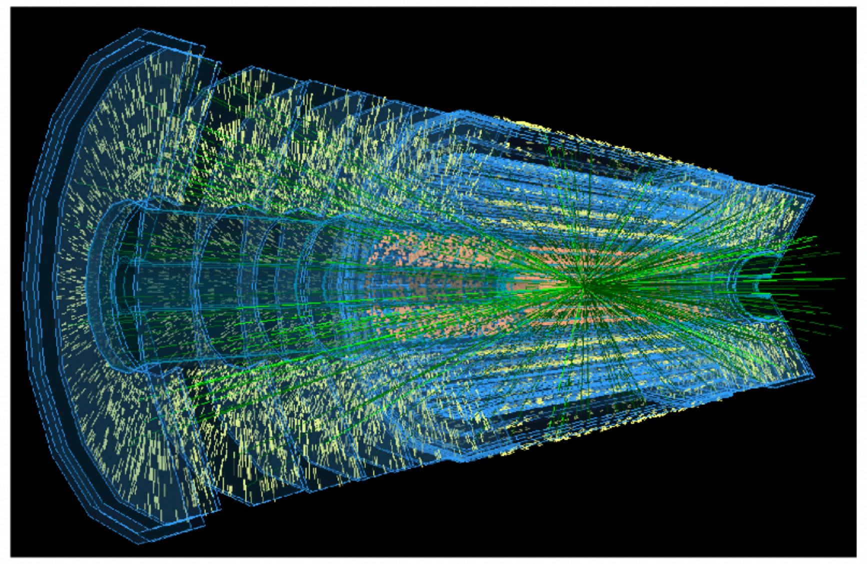
R&D including ML/accelerators and other new techniques & ideas

Low level reconstruction
(tracking, clustering, object acceptance)
High level reconstruction

ATLAS EXPERIMENT

HL-LHC Roadmap

ATLAS Software and Computing HL-LHC Roadmap

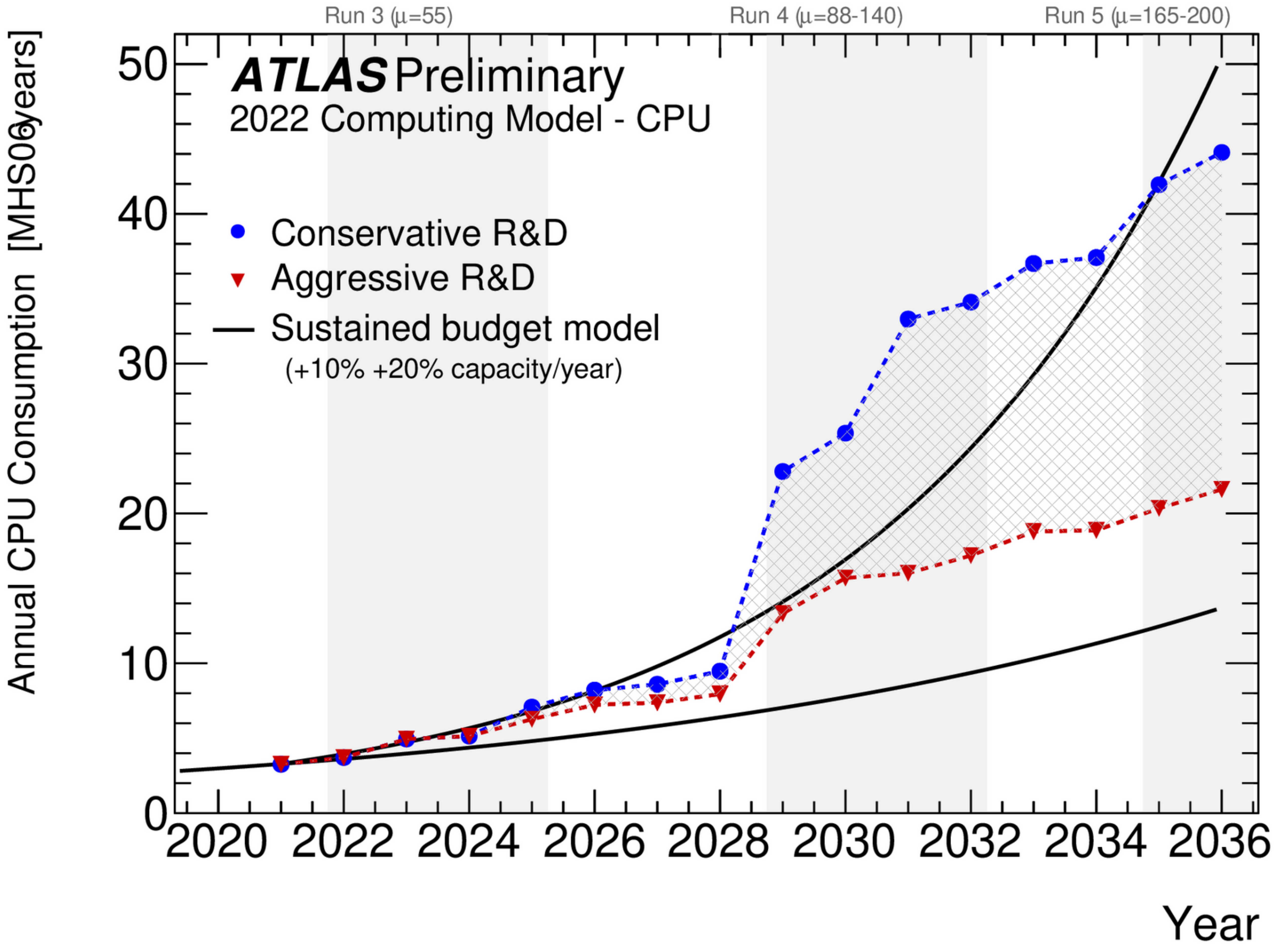


Performance tuning

Validation

Bug-fixes/ contingency

Resource modeling: CPU Projections



Run 4 requires aggressive R&D to stay within budget

Key elements

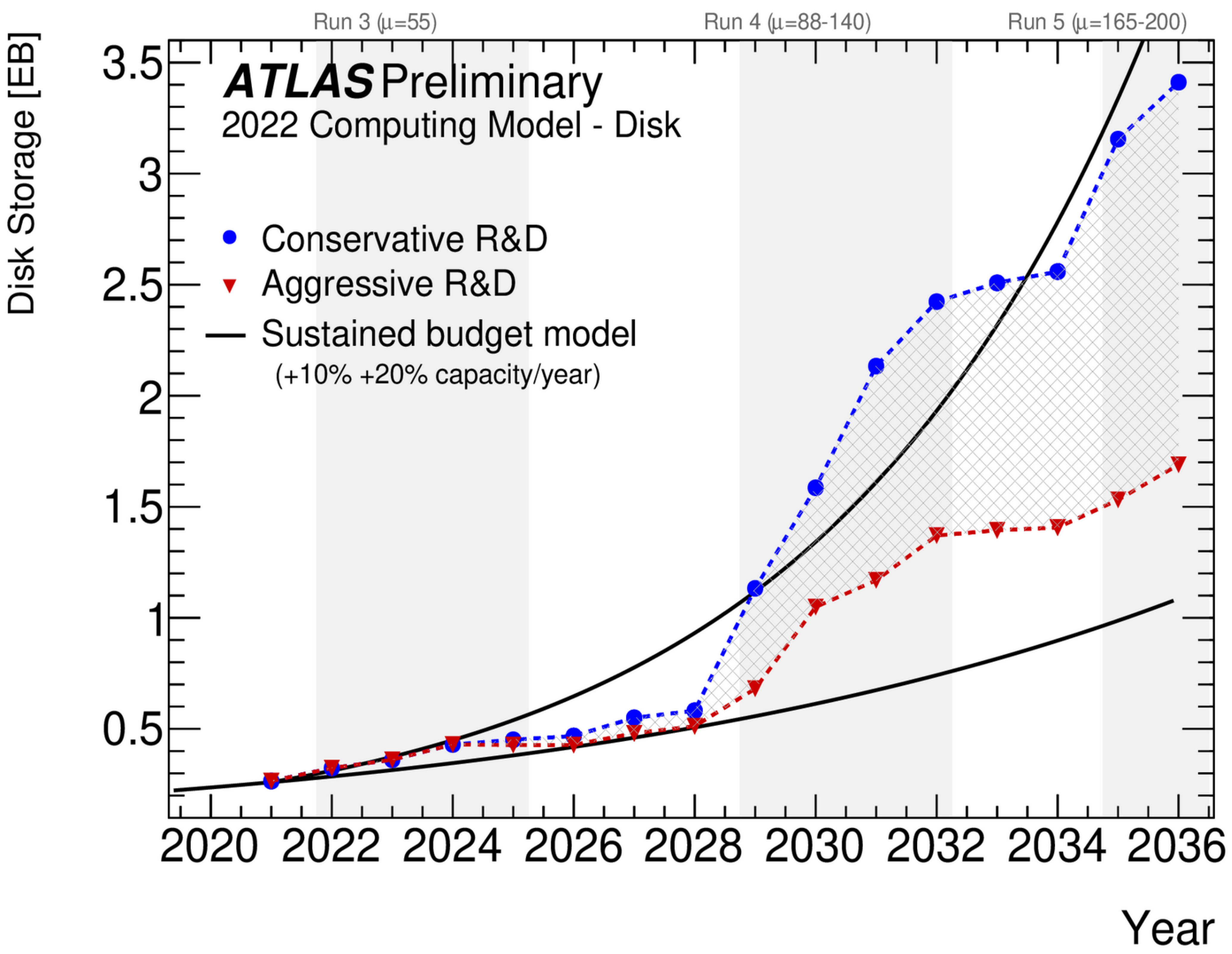
Fast simulation used for 70% (conservative) or 90% (aggressive) of MC simulation

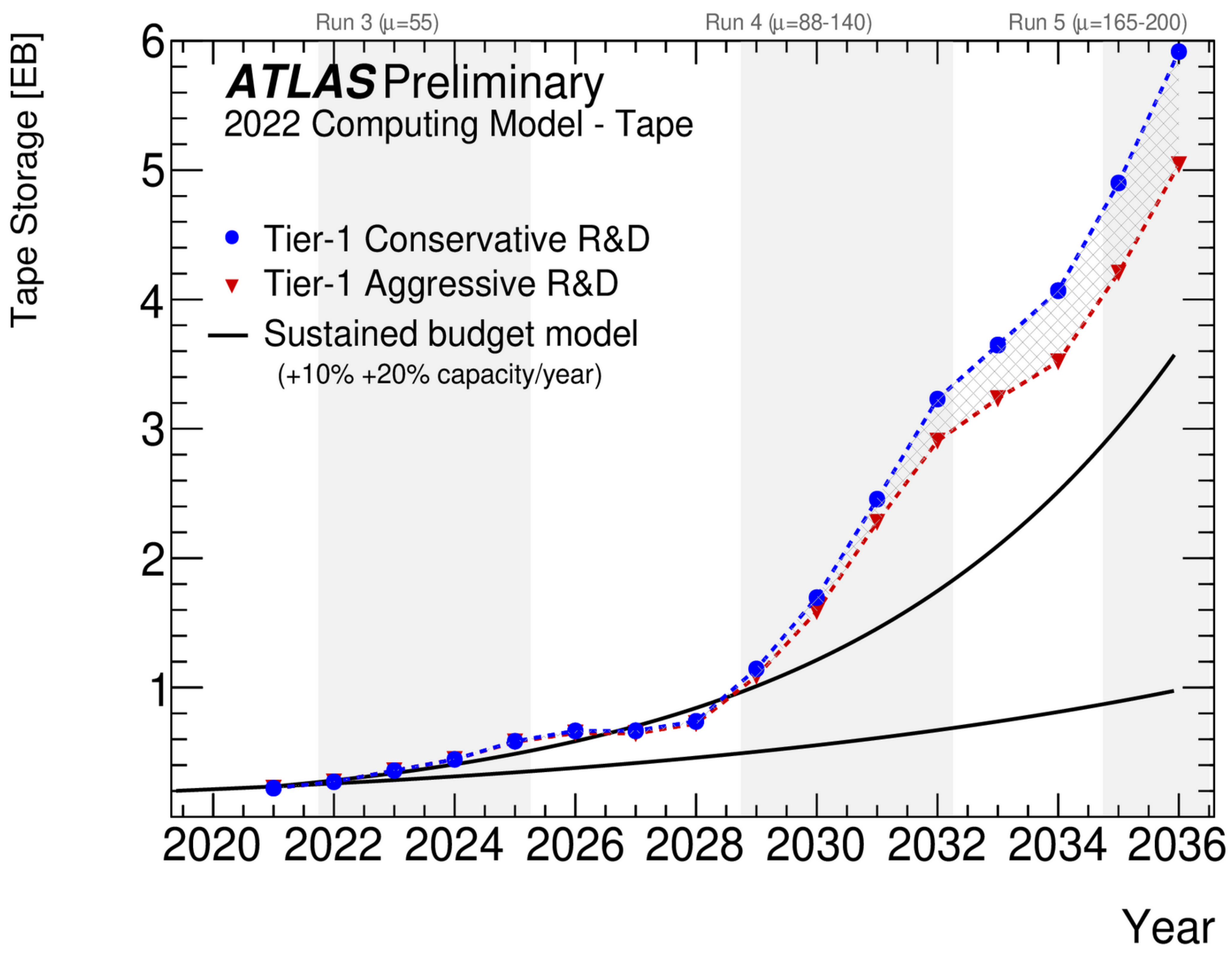
Faster event generation and reconstruction. Aggressive also means new person power is needed!

There is room for improvement!

Accelerators such as GPUs not considered in this projection, but are expected to play an important role too. Impact of GPUs will be quantified in the TDR.

Resource modeling: Disk and Tape Projections





Key elements

Disk needs mainly arise from storing formats for analysis. Aggressive R&D can reduce disk storage by almost factor 2: common analysis formats, aggressive data compression, increased use of tape, etc…

Tape storage budget has large uncertainties, but huge resources are needed.
Tape is required for long-term storage of data and simulation.

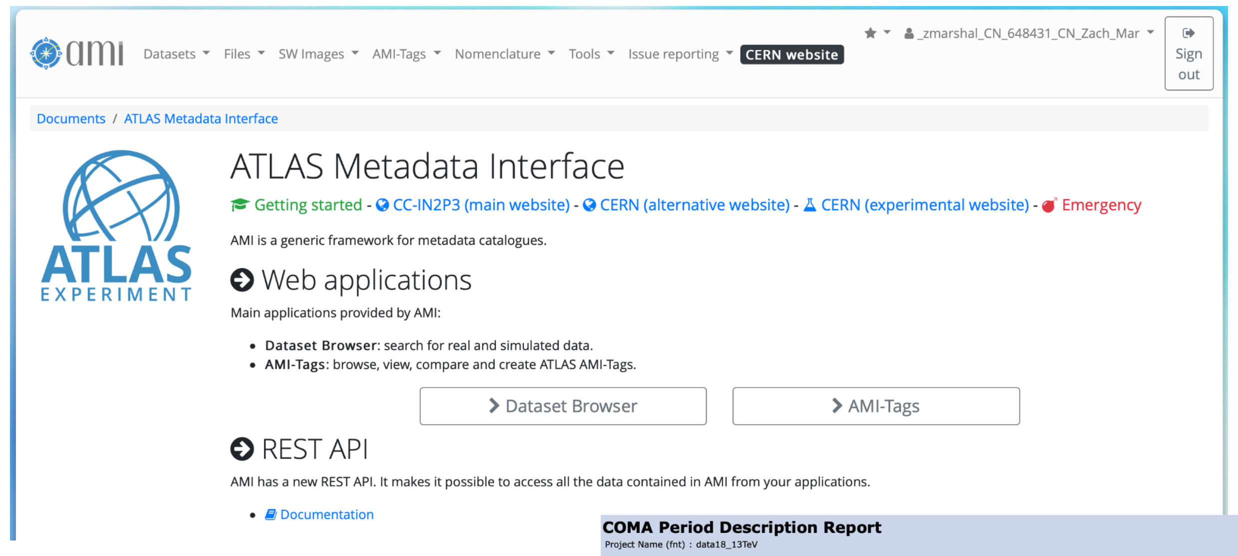
Areas of R&D in ATLAS Software & Computing

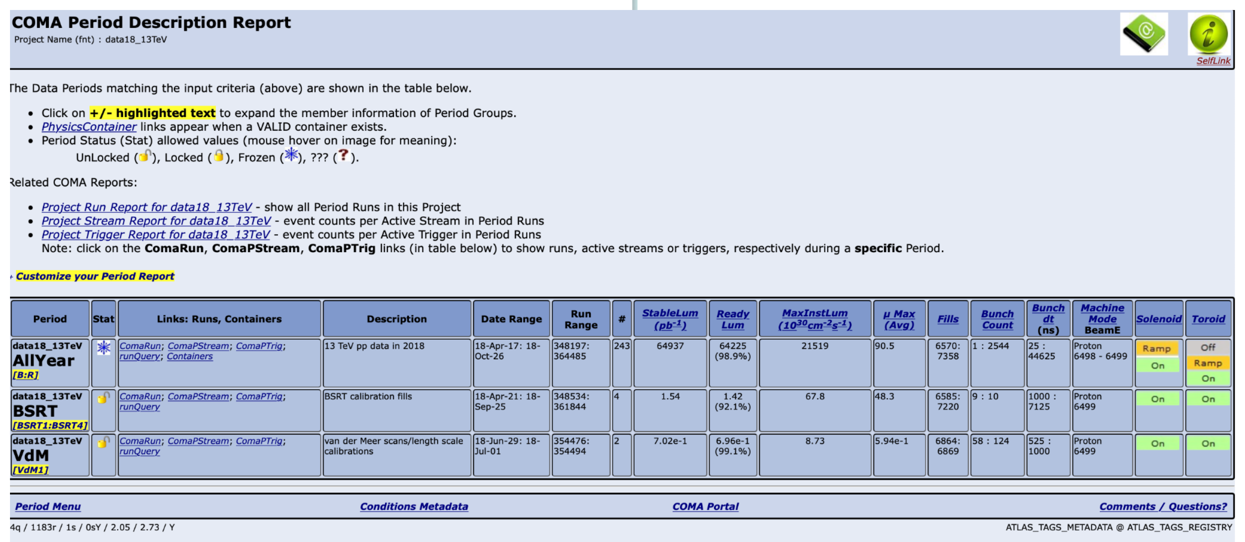
Database infrastructure, conditions evolution & metadata

ATLAS relies on **large amounts of auxiliary data** (run configuration, calibration & alignment data, data quality, TDAQ, geometries, ...)

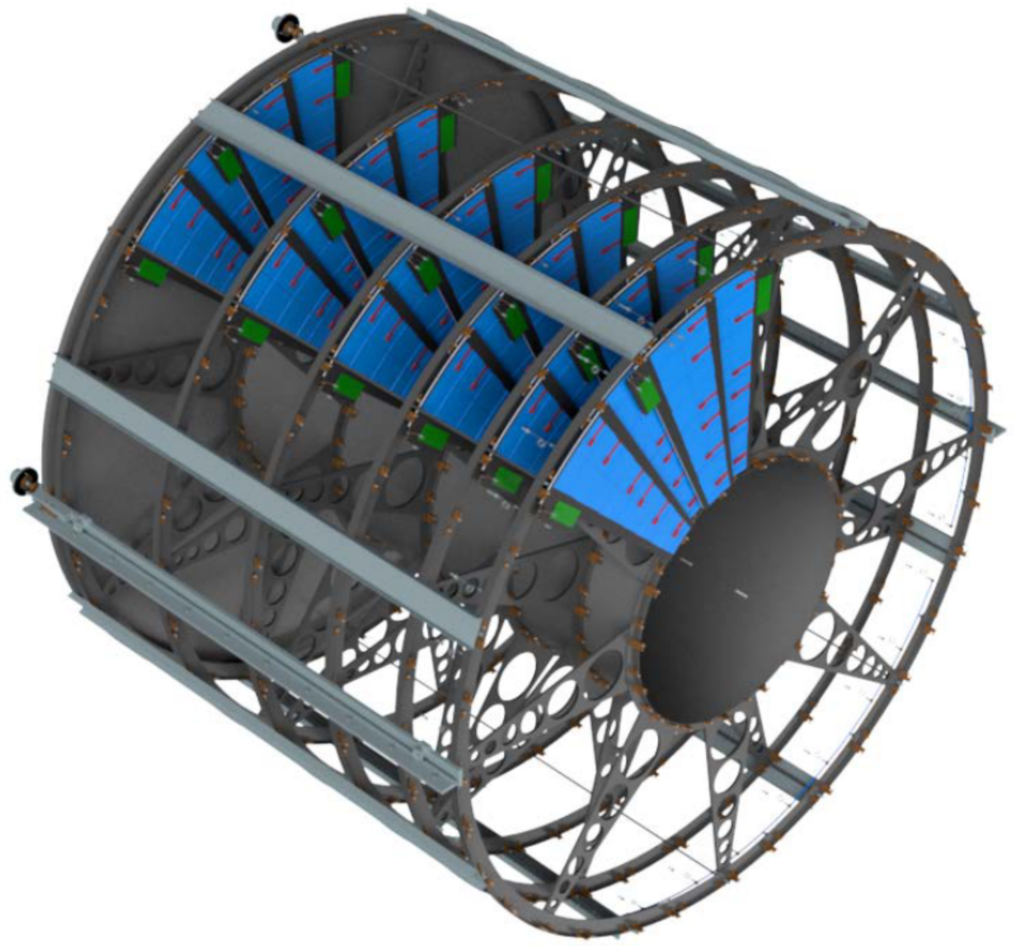
Migration of the **conditions database** from COOL to CREST
Relies more strongly on caches and requires less accesses from Athena grid jobs (1-2 TB per data taking period)

Other databases need to evolve to be ready for Run 4: COMA (run database), AMI (ATLAS metadata interface), EventIndex





Detector description



ITk strip endcap

Current GeoModel service is 20 years old. It needs updates for new detector changes and software improvements

Run 4 detector upgrades (talk) include new Inner Tracker (ITk), new high-granularity timing detector (HGTD), and new muon chambers

Plans are a **unified GeoModel of the whole detector** steered through XML-based database, improved (3D) **visualization** system, several **validation** tools

Areas of R&D in ATLAS Software & Computing

US ATLAS

Core software & heterogeneous computing/accelerators

AthenaMT allows the parallel processing of data, while concurrently analyzing multiple parts of an event at the same time (eg. tracking) — used in Run 3

Accelerators are a relatively new resource that is **not yet widely used, eg. in production workflows.** Ongoing R&D to integrate GPU in Athena.

Dedicated forum **HCAF** (Heterogeneous Computing and Accelerator Forum) to understand how to efficiently use accelerators — interest in **CPU/GPU hybrid workflows.**

Decide on technology for algorithm parallelization.

ATLAS Trigger/DAQ exploring FPGA accelerators.

Event generation

MC event generation uses a major fraction of CPU resources

Theory advances lead to more precise MC including higher order correction which require more resources.

Ongoing R&D

Optimization of settings and choices for the generators and efficient event filtering

Efficient description of theoretical systematic uncertainties

Use of HPCs for events with large number of final-state particles

New generators optimized for GPUs

New ML methods for event sampling

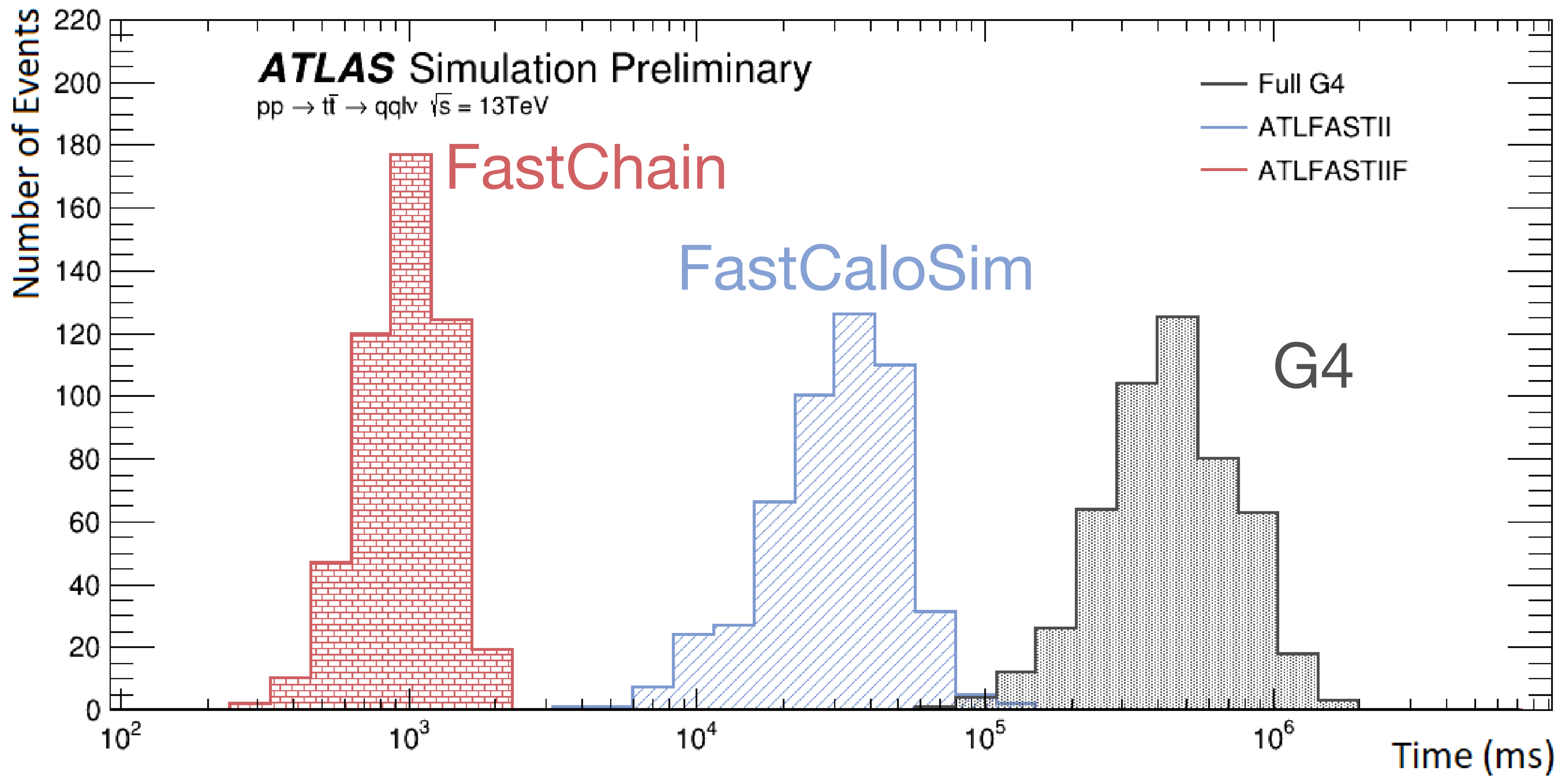
Areas of R&D in ATLAS Software & Computing

Simulation & Digitization

FastChain is a set of new components to make event simulation and digitization faster.

Possible to perform direct event generation → reconstruction without the intermediate formats. Several configurations of fast+standard tools are being considered for HL-LHC

Full simulation can also be improved, eg. by optimizations or by exploiting parallelism (Celeritas/ADePT, etc)



Fast Calorimeter Simulation (FastCaloSim) has a long tradition in ATLAS. The calorimeter simulation is most time-consuming part of full simulation (~80% with G4). New fast simulation (**AF3**) was released an year ago. Includes detailed modeling of the longitudinal and lateral shower evolution to model jet energy substructure. A GAN is used for hadronic shower modelling.

Several ongoing R&D projects are needed for Run 4 FastCaloSim on a GPU (HEP-CCE)

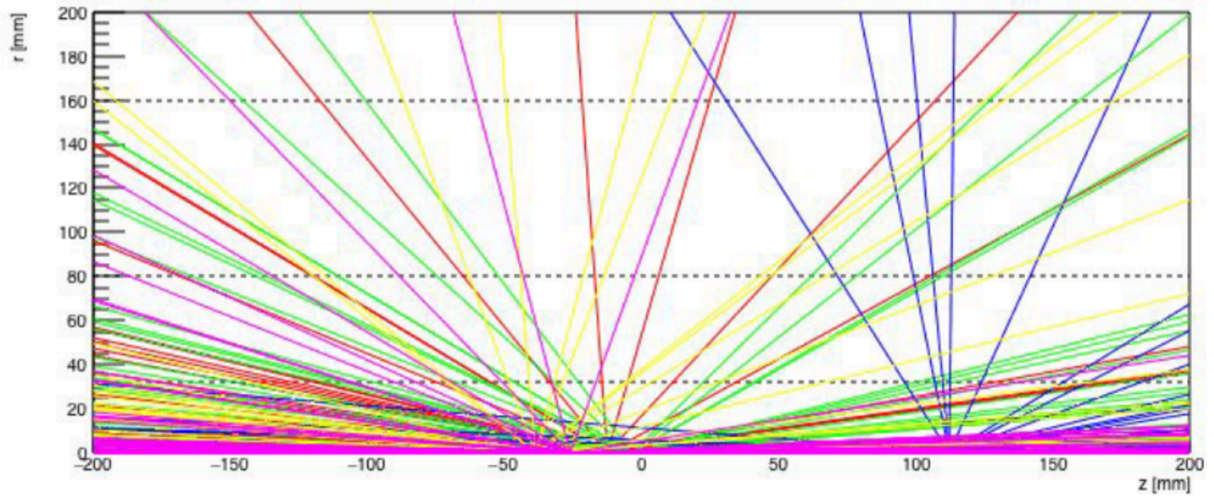
Pile-up effects will be a challenge during HL-LHC. In Run 2, pile-up and hard scatter events were digitized together. In the **new MC+MC Overlay** used in Run 3 pile-up samples are digitized separately and then overlayed onto the hard scatter. **Real data from zero-bias trigger** can be used for direct overlay to the digitized simulation.

For Run 4, another idea is **track overlay** where pile-up tracks are merged after reconstruction. Works well for topologies where hard scatter track reconstruction is not affected by PU.

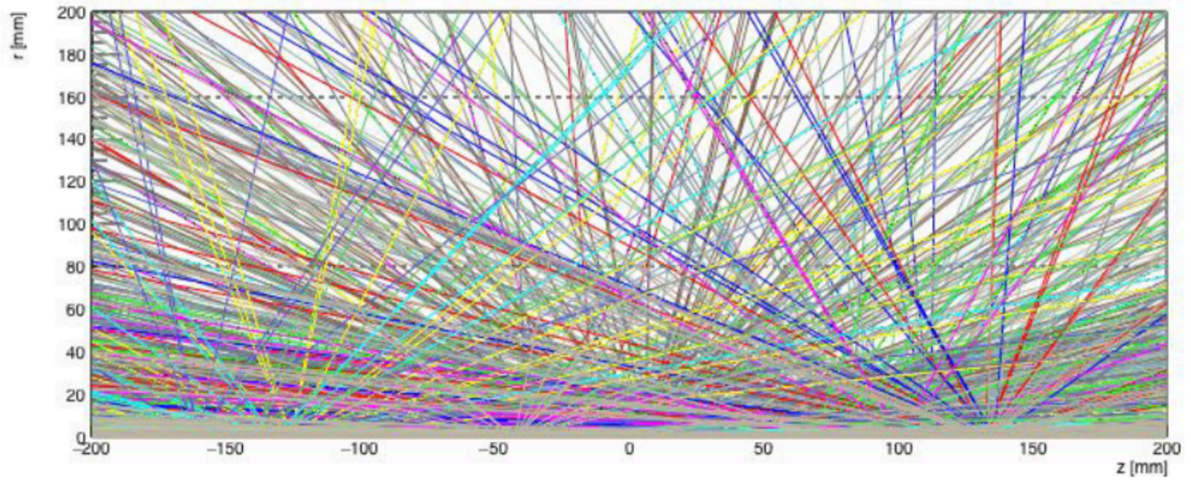
Areas of R&D in ATLAS Software & Computing

Reconstruction

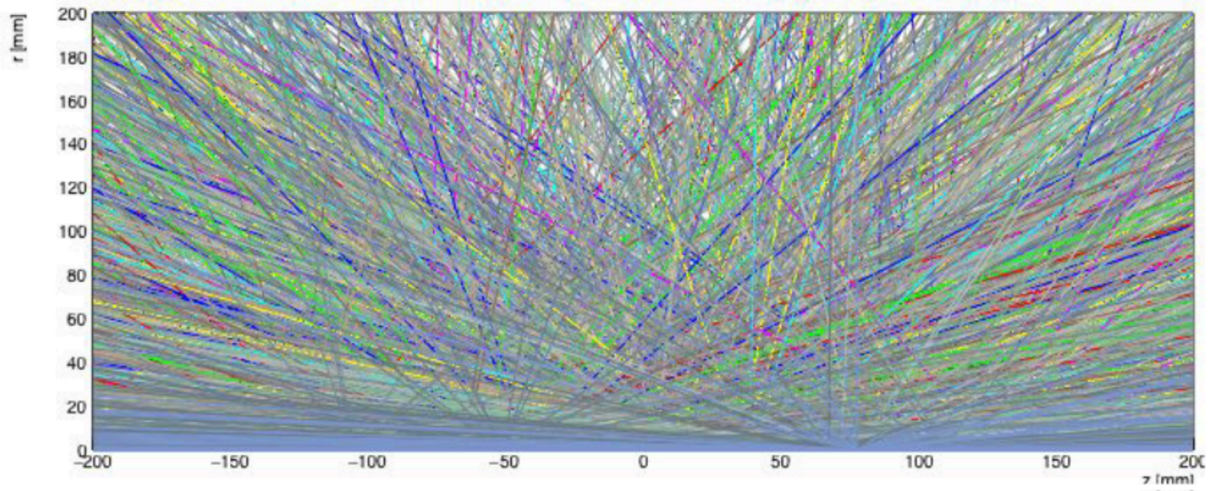
Object reconstruction **very challenging with 200 pile-up interactions on average** per event



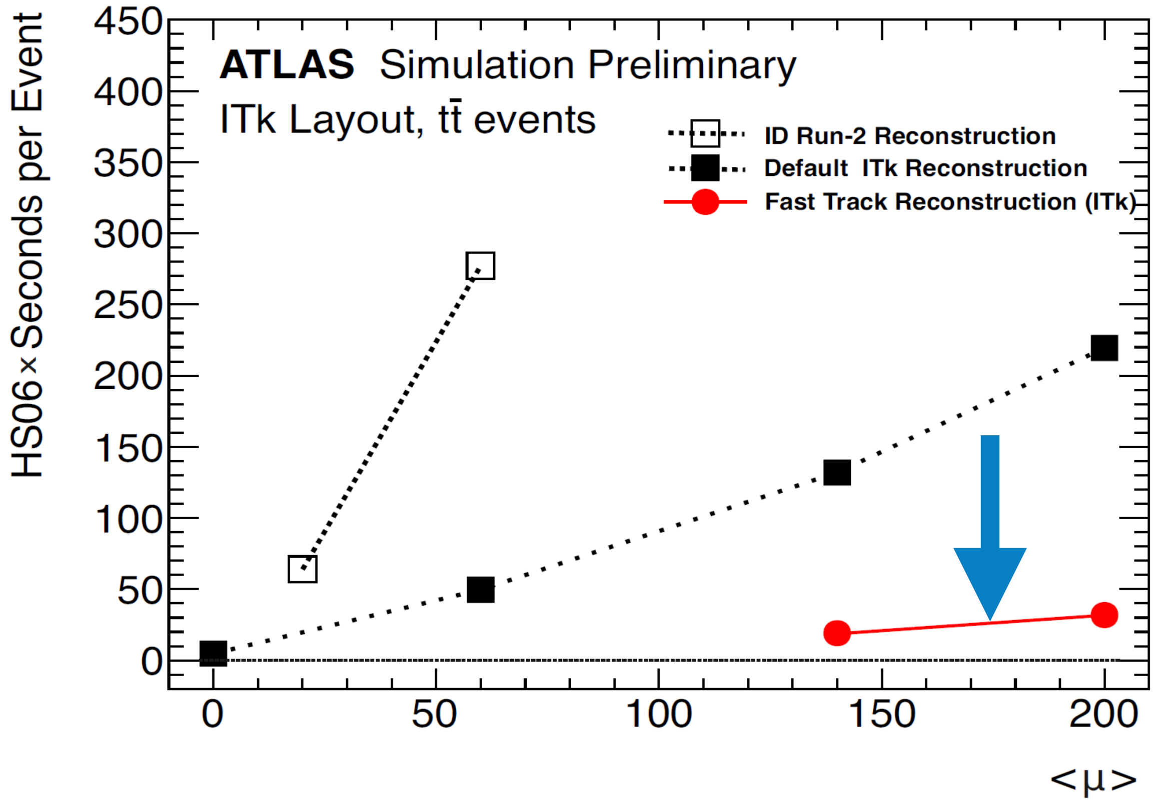
2010 <μ> ~ 5



2018 <μ> ~ 40



2029 <μ> ~ 200



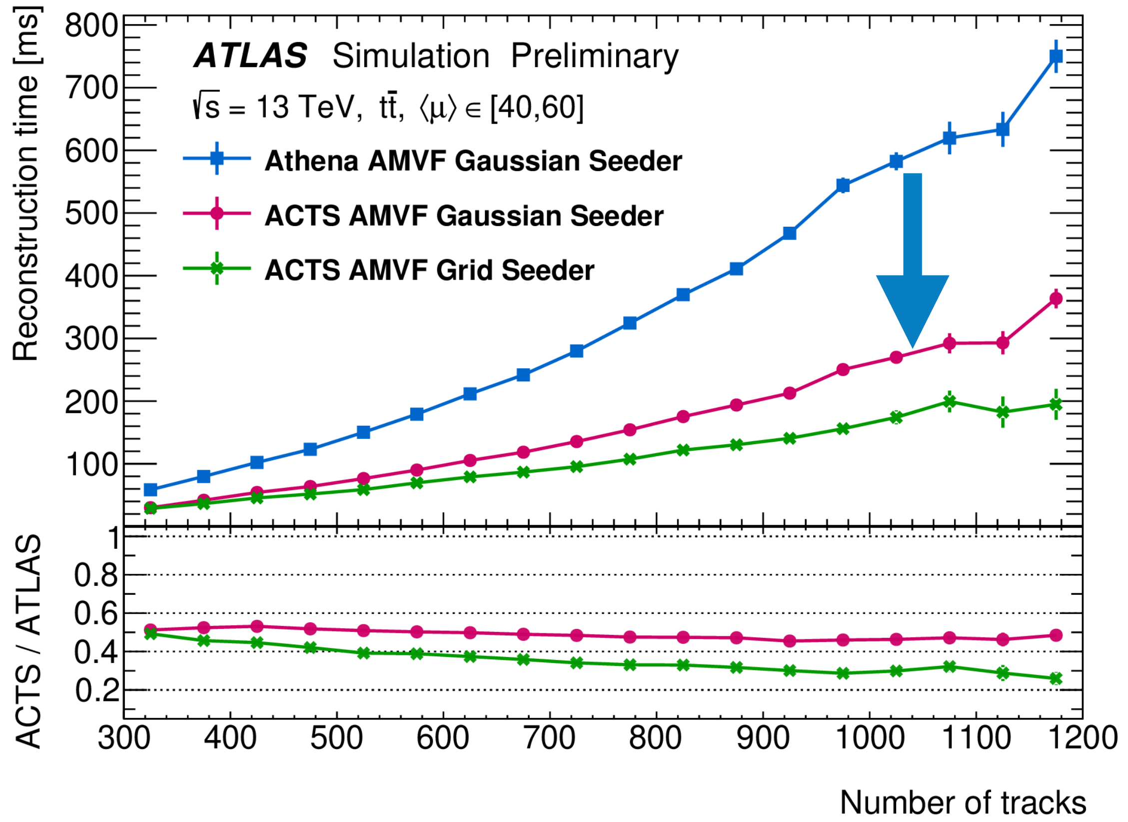
New & smarter algorithms to handle harsh conditions and benefit from new detectors

Fast track reconstruction (link) protoype showing large CPU gains possible

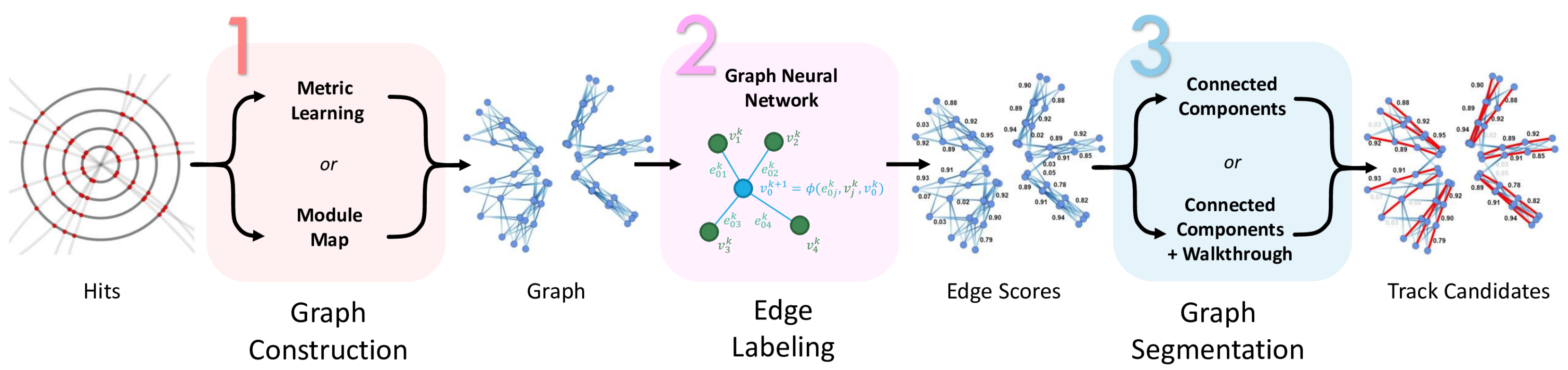
ACTS (A Common Track Reconstruction Software): experiment-independent project to achieve CPU reduction and great performance for tracking and also particle flow, supports multi-threading (link)

Integration in ATLAS progressing — already used for vertex finding

IRIS-HEP contributions have been crucial for the success of ACTS!



GNN4ITK (Exa.TrkX in the US) — Tracking using a GNN pipeline on ITk (link)



traccc: Demonstrator of tracking chain on accelerators

Developments ongoing in other areas of reconstruction beyond tracking

GPU-based implementations of calorimeter clustering are about to be integrated & validated

Overview of R&D Projects — Distributed Computing

Distributed Computing

Grid resources and software infrastructure for data management, workflow management, production system, monitoring, analytics, etc…

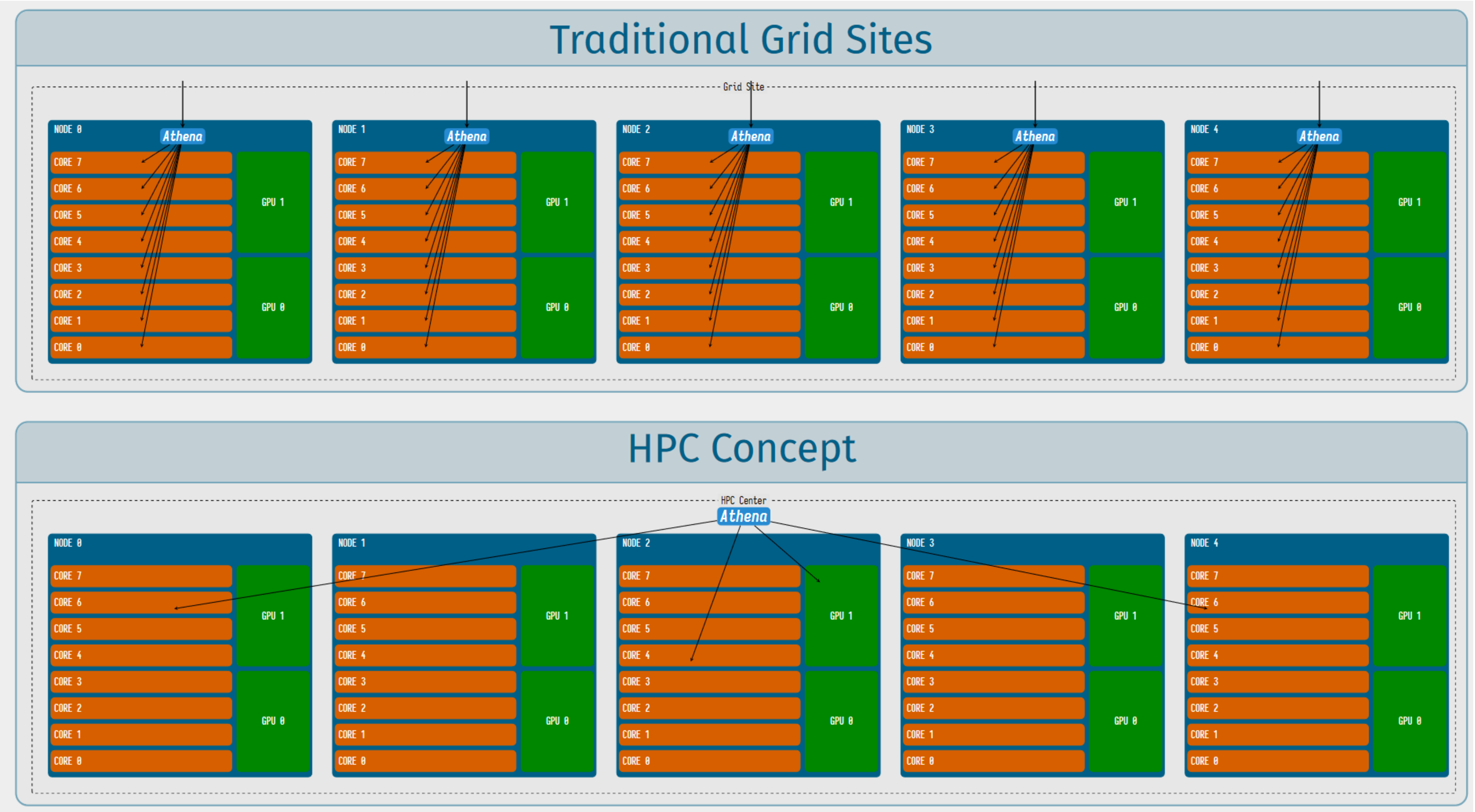
Data processing relies on distributed resources. ATLAS distributed computing infrastructure need to evolve for HL-LHC (scalability and efficiency)

Ongoing R&D
Token-based authentication
Increase network bandwidth
Integrate new resources, such as HPCs and cloud infrastructure

ATLAS workflows and data management system need to be adapted to serve a variety of HPC platforms

Refinement of fully **containerized workflows:** all the software in a single container deployed at the HPC site

New schedulers for heterogeneous architectures to use most appropriate resources and balance loads — Raythena, HPX



(Illustration: B. Stanislaus)

Overview of R&D Projects — Distributed Computing

Distributed Computing

iDDS and Data Carousel

Intelligent Data Delivery Service (iDDS) is a service to orchestrate the workload and data management systems to transform and deliver data needed to consumers while seeking be efficient in the use of storage, network and processing resources

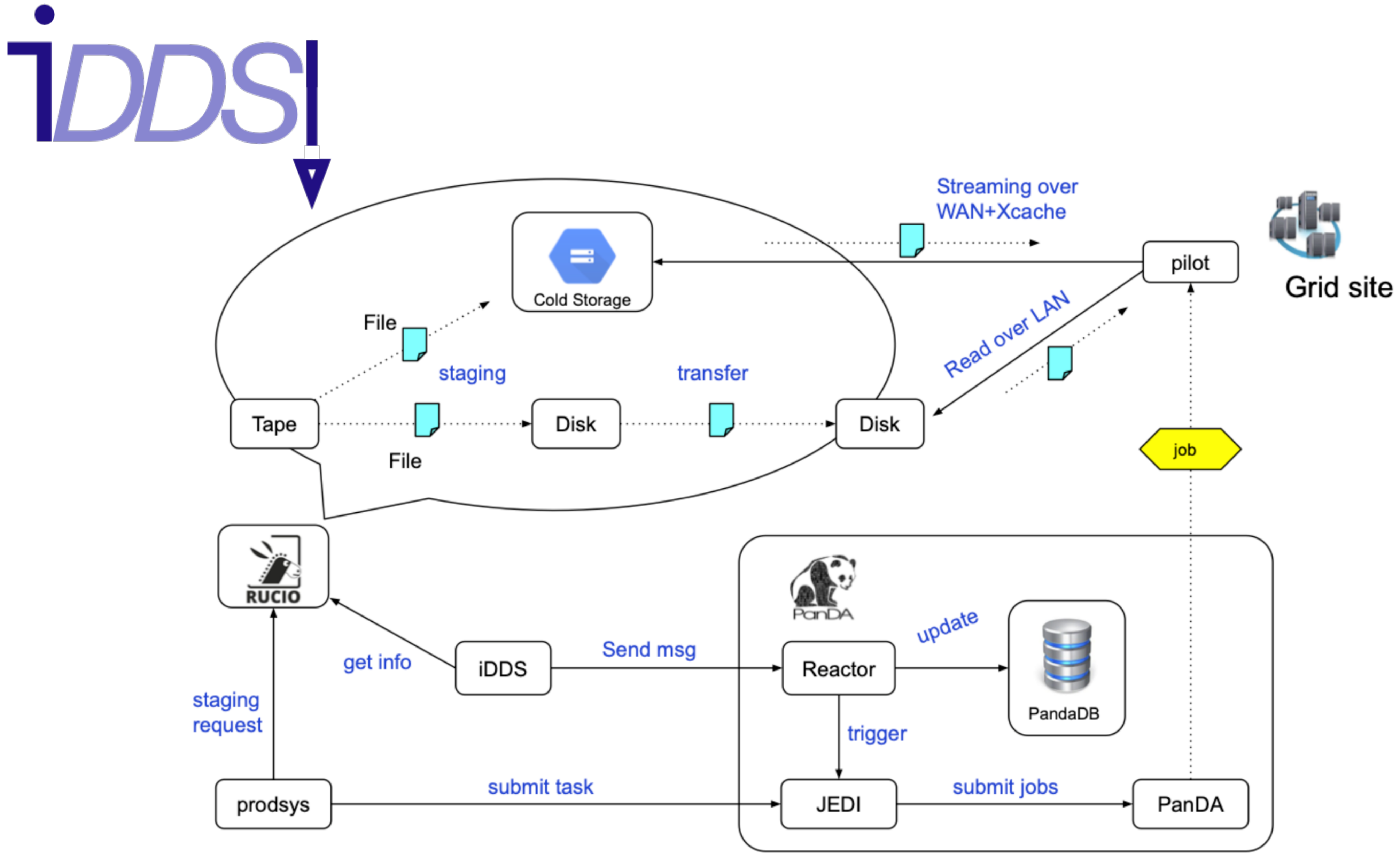
Data is stored in the lowest-cost media available (tape) and processing is executed by staging and promptly processing slices of inputs onto faster storage (disk), such that only the minimum required input data are available at any time

Data carousel is already used for AOD storage in run-3 (link)

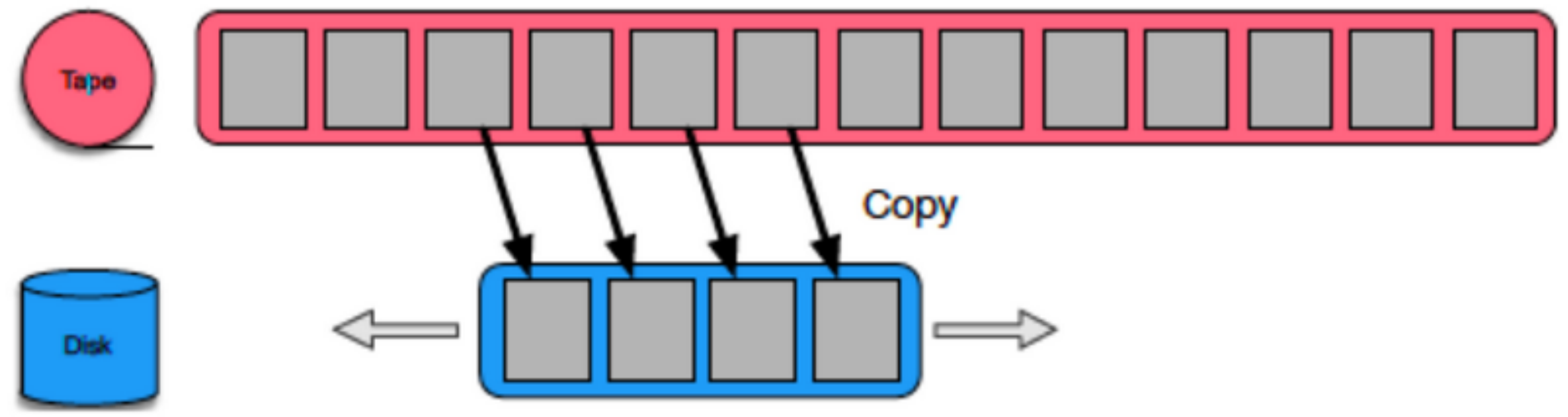
Plans for HL-LHC

Reconstructed data stored primarily on tape (with only very partial disk replicas), processed with the Data Carousel

iDDS provides streaming services for data delivery for a increasing range of complex workflows, such as AI/ML — service for is in production, used for HPO AF3 FastCaloGAN, available for analysis & software teams



ATLAS Data Carousel using iDDS



iris hep
Institute for Research & Innovation in Software for High Energy Physics

IRIS-HEP contributions have been crucial for the success of iDDS!

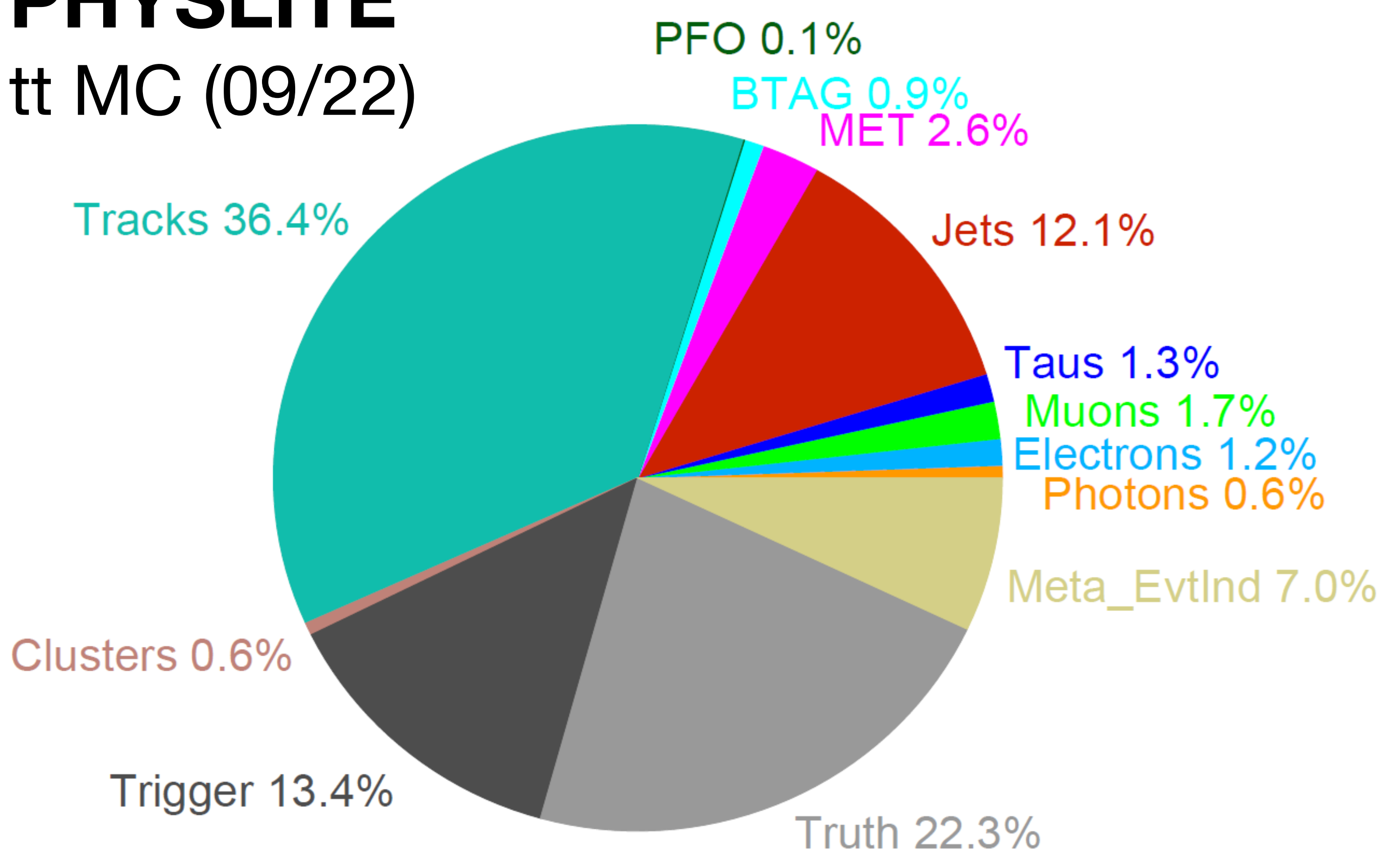
Areas of R&D in ATLAS Software & Computing

DAOD_PHYSLITE: future common **format for fast analysis**, unskimmed, contains calibrated objects. Small size: ~10 kB/event, total ~1PB/year, to be used by ~80% of all physics analysis in Run 4

Key concepts for DAOD_PHYSLITE
RDataFrame and RNtuple in ROOT 7
Columnar analysis and simplified systematics
Lossy compression

Analysis

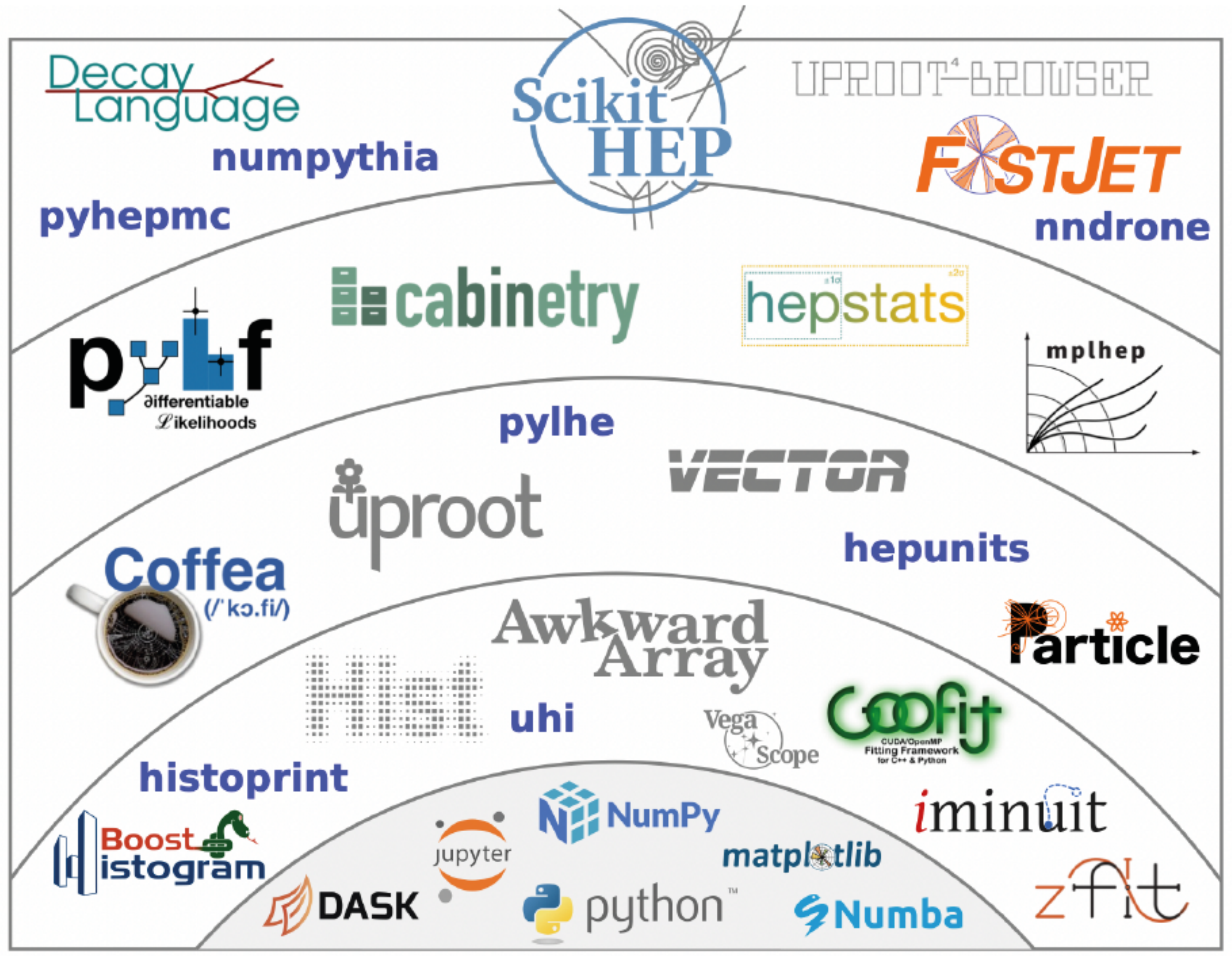
PHYSLITE
tt MC (09/22)



Dedicated computing resources for more complex end-user analysis workflows such as training complex ML algorithms, exploiting python-based ecosystems for vectorized computations, visualizations etc

Several R&D projects ongoing to identify new tools: scalable Jupyter notebooks, new data delivery services, etc…

Analysis facilities, focus on **interactivity, usability, user-support,** with federated access, store the data and MC samples needed, can provide GPUs and specialized software (e.g. ML), deployed in various forms, for example integrated in large computing centers, as HPCs, or on the cloud



Several analysis tools developed with IRIS-HEP participation!

iris hep
Institute for Research & Innovation in Software for High Energy Physics

Education, Training & Workforce Development

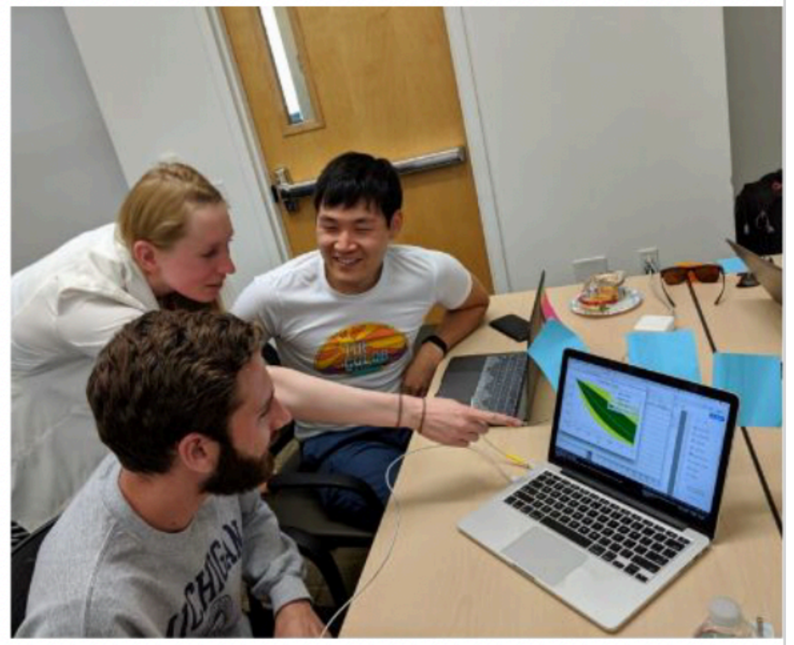
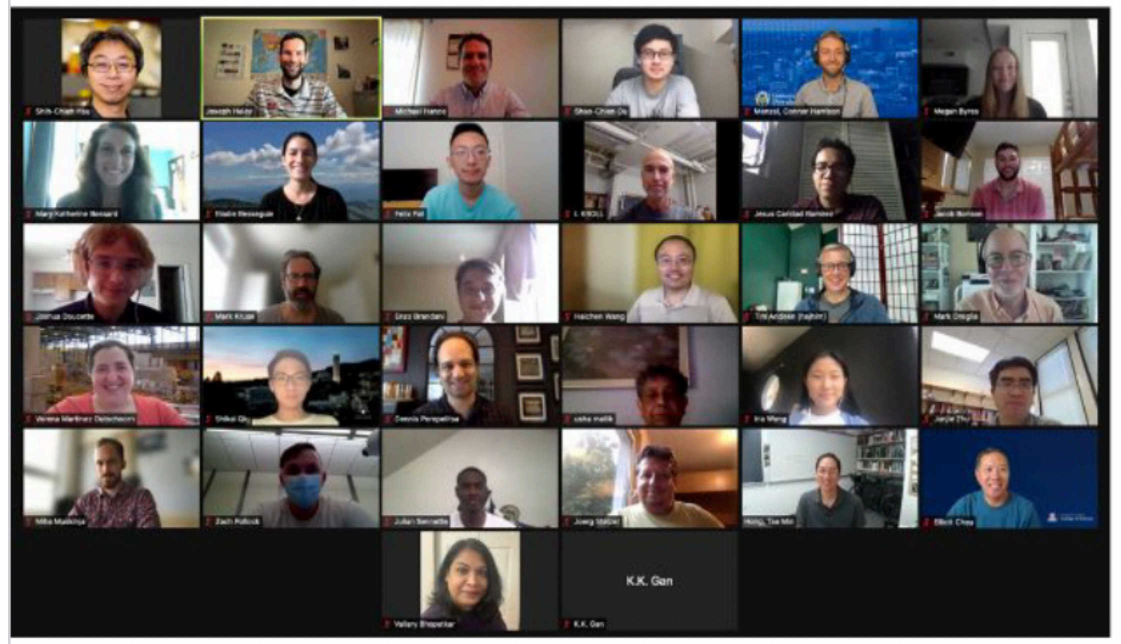
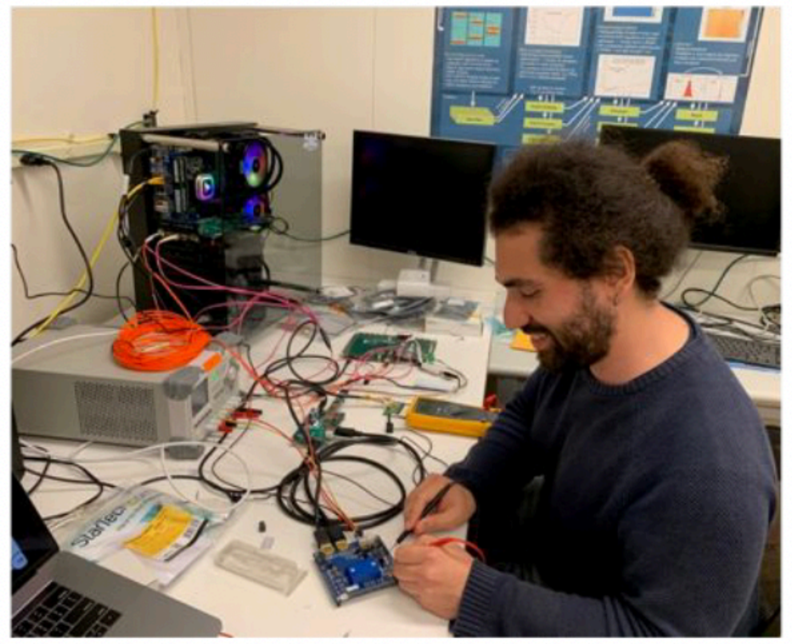
US ATLAS

Education & training in software & computing is **critical** to prepare the next generation of scientists
— analysts at most levels will need **new skills** to handle increasing datasets
R&D projects are an exciting opportunity to **engage young scientists and develop careers**
— also creates a pipeline of software and computing professionals for operations etc

iris hep
Institute for Research & Innovation in Software for High Energy Physics

Several training initiatives in US ATLAS and ATLAS — **some events organized with IRIS-HEP participation!**



A priority for the future is training current users on new tools — important for widespread use
Communities such as IRIS-HEP can make increasing contributions to training infrastructure, curriculum development and interfacing with experiment-specific tools

ATLAS R&D Demonstrators

Conceptual Design Report | HL-LHC Roadmap (Road to Run 4 or R2R4) | Technical Design Report | Run 4 Feature Freeze | Run 4 Performance Freeze | Run 4 Release Validated

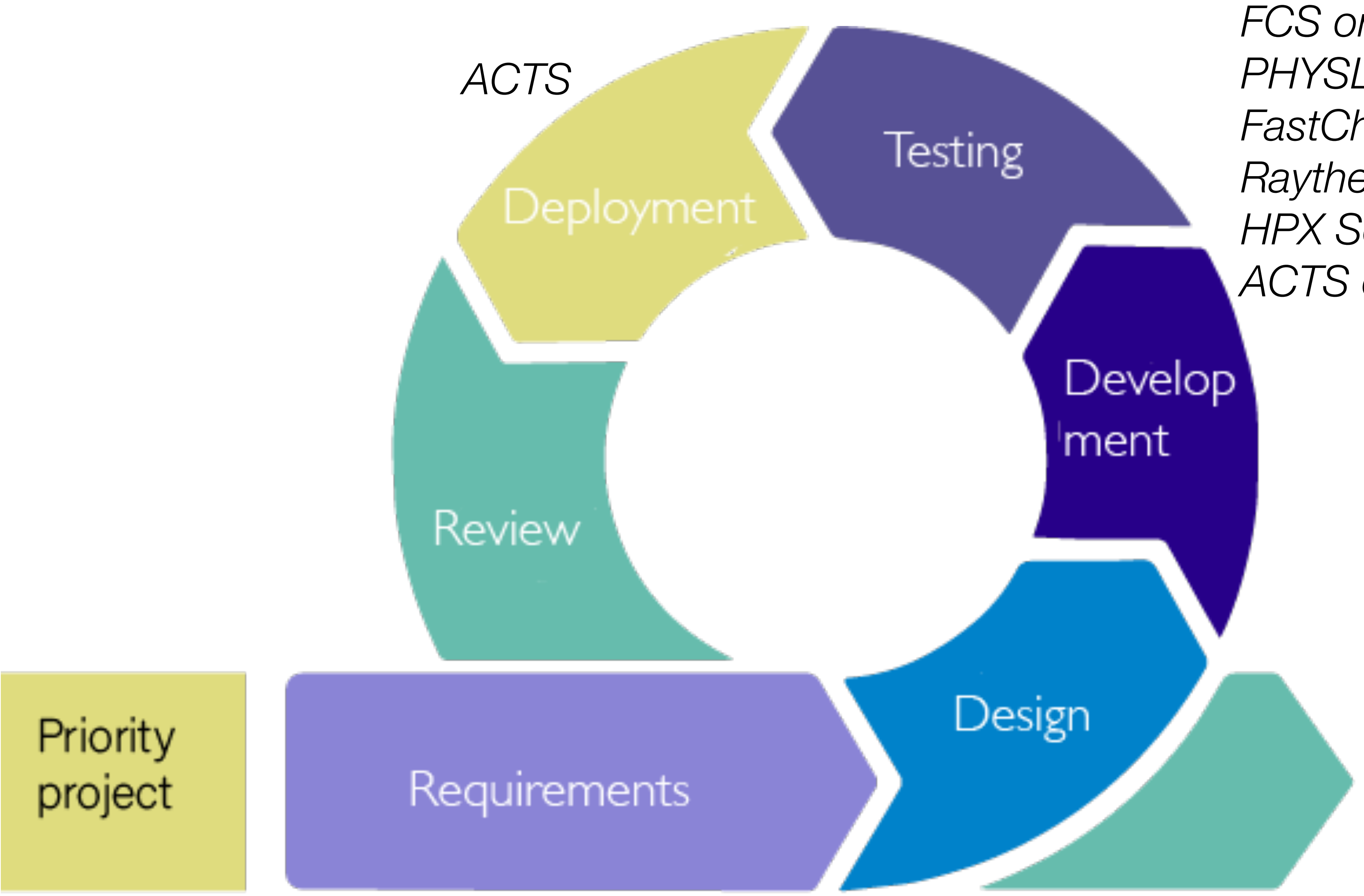
2020 · 2021 · 2022 · 2023 · 2024 · 2025 · 2026 · 2027 · 2028 · 2029

Ready for HL-LHC data-taking

R&D including ML/accelerators and other new techniques & ideas

Project Life Cycle

ACTS



FCS on GPUs
PHYSLITE
FastChain
Raythena
HPX Scheduler
ACTS on GPUs

Deployment, commissioning, production

ML services
iDDS
Data carousel
ONNX Runtime
AF3 FastCaloGAN
Containerization
Enhanced XRootD
gridftp replacement
Raythena (end 22)

Or not

Our calls have been good so far!

Demonstrators planned in next year

proof-of-concept demonstrations to estimate the impact of R&D projects on hardware resources
e.g. CPU time, GPU time, RAM, disk, tape & network

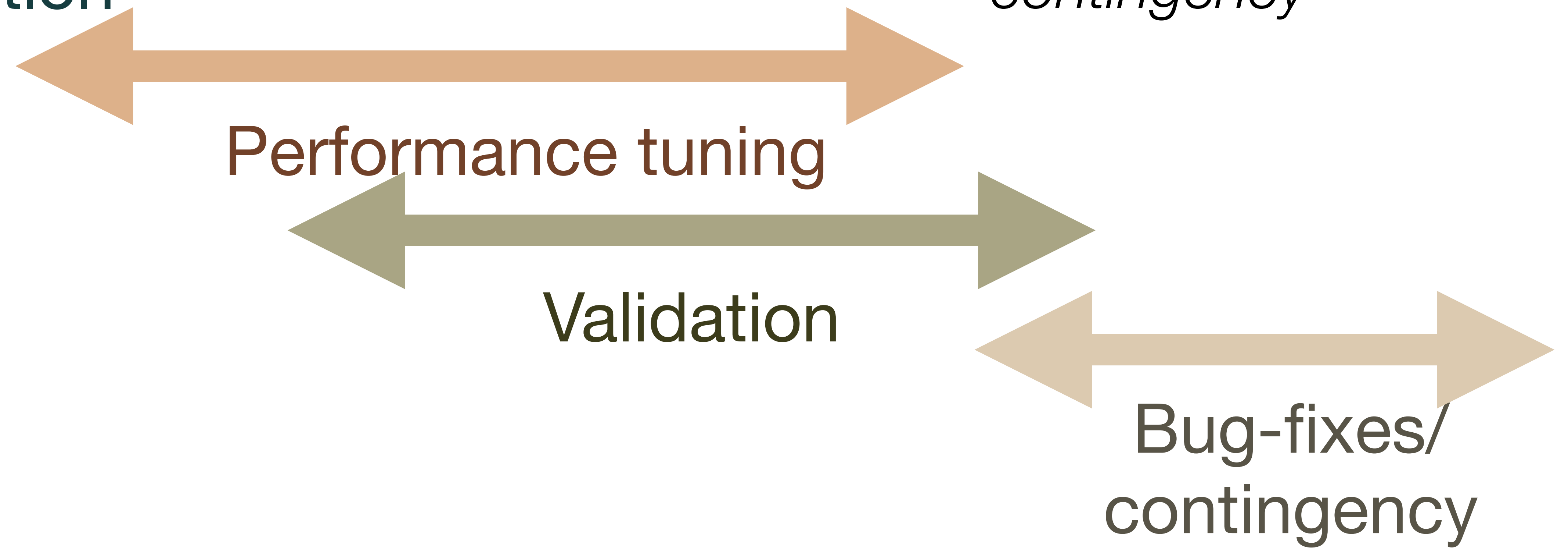
work ongoing to define configurations, datasets & platforms for these impact studies

Key process for all R&D developments to participate in!

Integration in ATLAS Operations

Integration in operations

Mature R&D developments shown to have significant impact need to be integrated with the rest of the software & computing operations
Focus on experiment specific needs and optimizations for ATLAS
Concern about long term support for several community projects



Long term integration and support through operations is key to sustainable projects

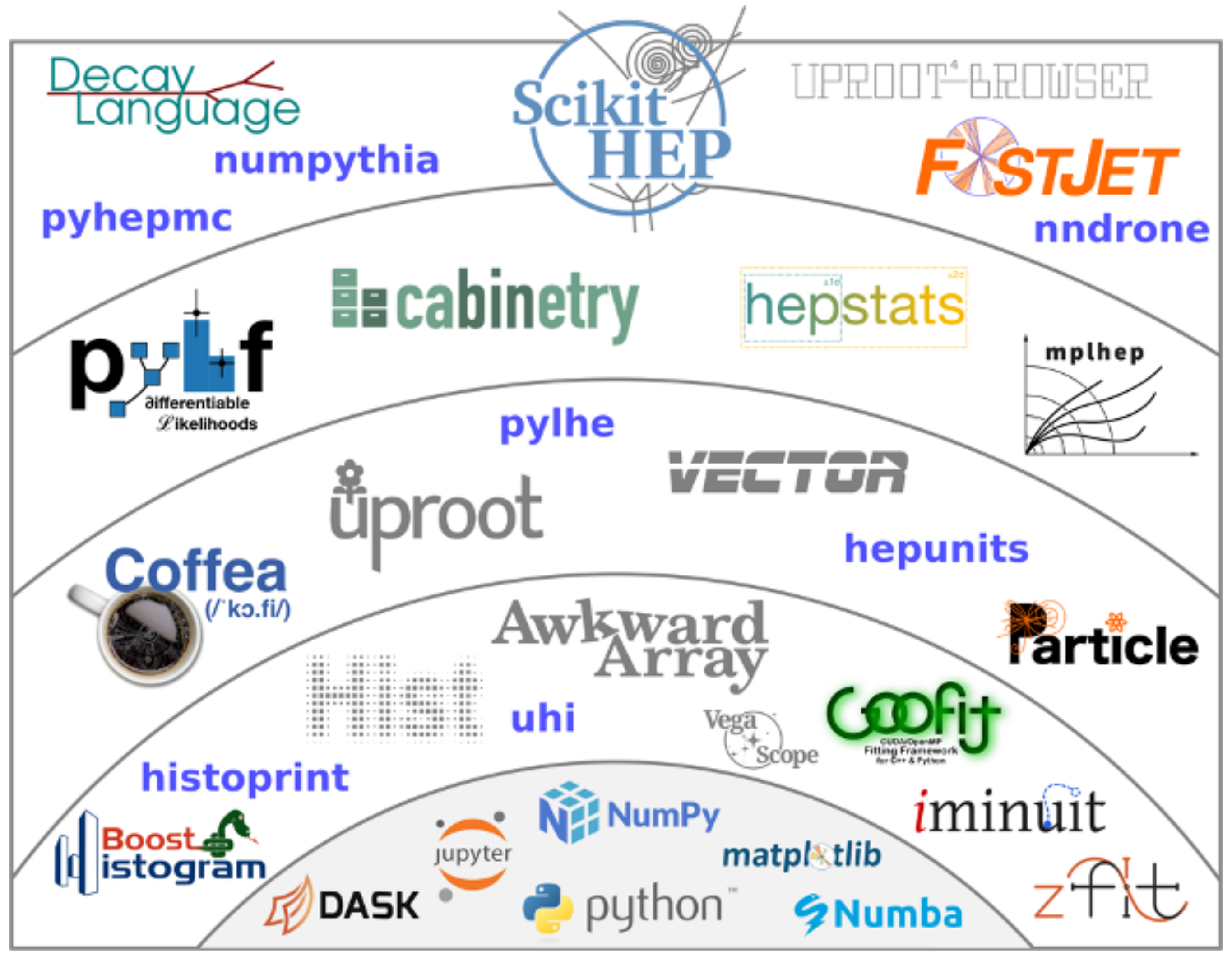
Improving Synergy with Operations

Ready for HL-LHC data-taking

- Several US-funded **R&D** projects **key to bridging the resource gap** for the HL-LHC
  - Projects like OSG-LHC, ACTS, iDDS have already made valuable contributions to ATLAS
- Key questions to the **successful transition of R&D to operations**
  - Does the R&D provide **significant gain** & is crucial?
    - Key to participate in demonstrator process — Prototype systems deployed by TDR (2024)
  - What is needed for R&D developments to be **integrated** in the ATLAS **infrastructure**?
    - Early integration to understand scope of work — play active role in organizing development and integration
  - Will there be **adequate maintenance** throughout the lifetime of ATLAS?
    - Embed new projects in experiments soon to increase participation — need beginning-to-end systems (full featured prototypes) for experiments in Run 3

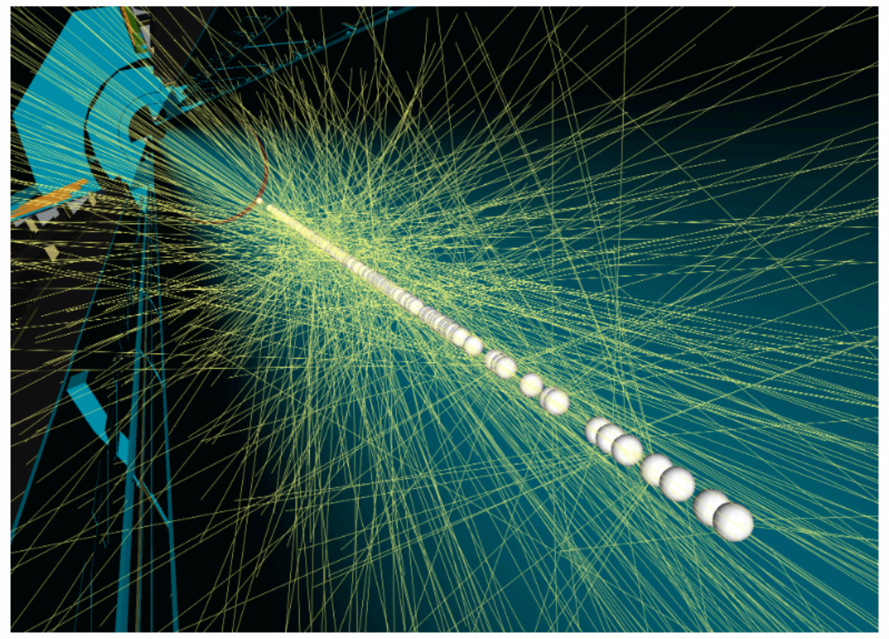
Improving Synergy with Operations — Analysis

- Key goal is to **integrate new ideas for analysis in** the **current** ATLAS analysis **model**
  - Several of the ideas developed in US-funded R&D projects could be integrated in the ATLAS ROOT-based infrastructure — some already are!
  - Topics such as data formats (PHYSLITE, …) are experiment specific, but topics such as transformations to columnar formats & tools to handle systematic variations are possible joint projects
- Development of **Analysis Facilities considering** the **ATLAS event and analysis model**
  - Workflows consider the particular way ATLAS deals with data access, handling, calibration & description of experimental uncertainties
  - Integration of new data analysis and management tools (coffea-casa, serviceX, etc) may require significant dedicated effort to enable widespread use in ATLAS



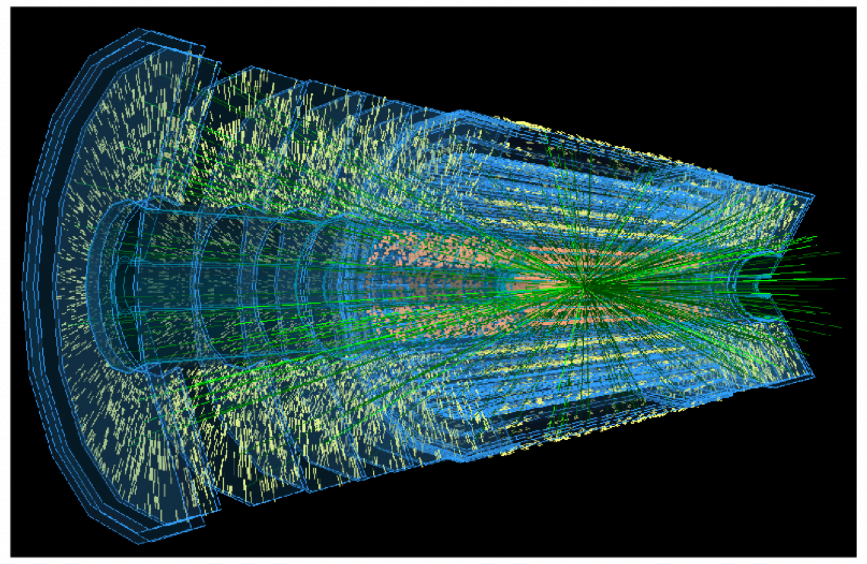
HL-LHC CDR

ATLAS HL-LHC Computing Conceptual Design Report



HL-LHC Roadmap

ATLAS Software and Computing HL-LHC Roadmap



Summary & Outlook

Ready for HL-LHC data-taking

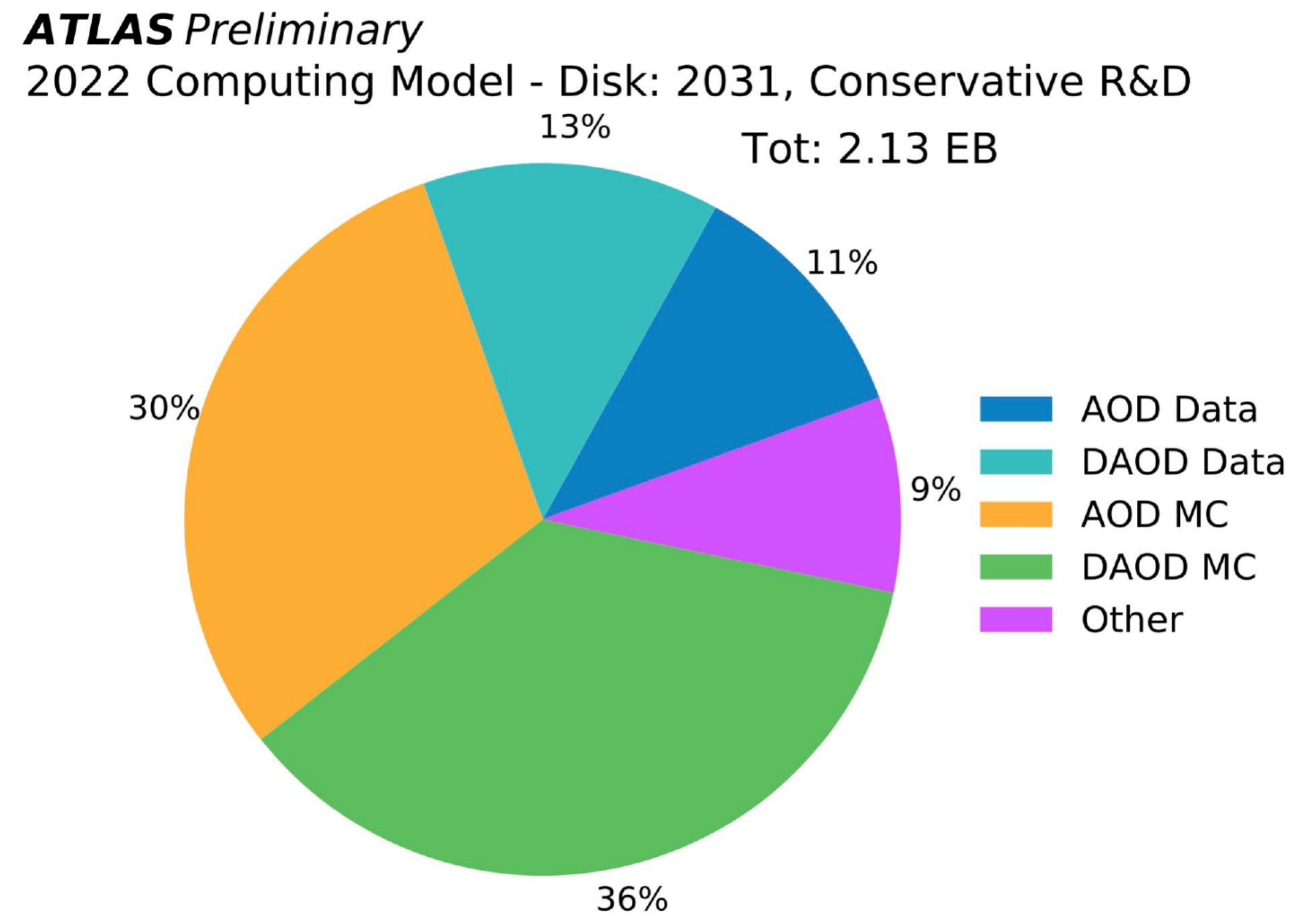
- HL-LHC brings unprecedented physics opportunities and computing challenges
- ATLAS is greatly benefitting from collaboration in common software and computing projects
  - Enables sharing ideas, effort and expertise across HEP experiments and with colleagues from other science communities
  - **Several projects with participation from programs such as IRIS-HEP have been key to success!**
- Many exciting R&D projects ongoing, impact will become clear in the next few years
  - The focus in 2023 will be to demonstrate R&D projects, evaluate their impact on resources and effort
  - Down-selection of R&D projects will lead to the TDR in 2024, which serves as a blueprint for the future

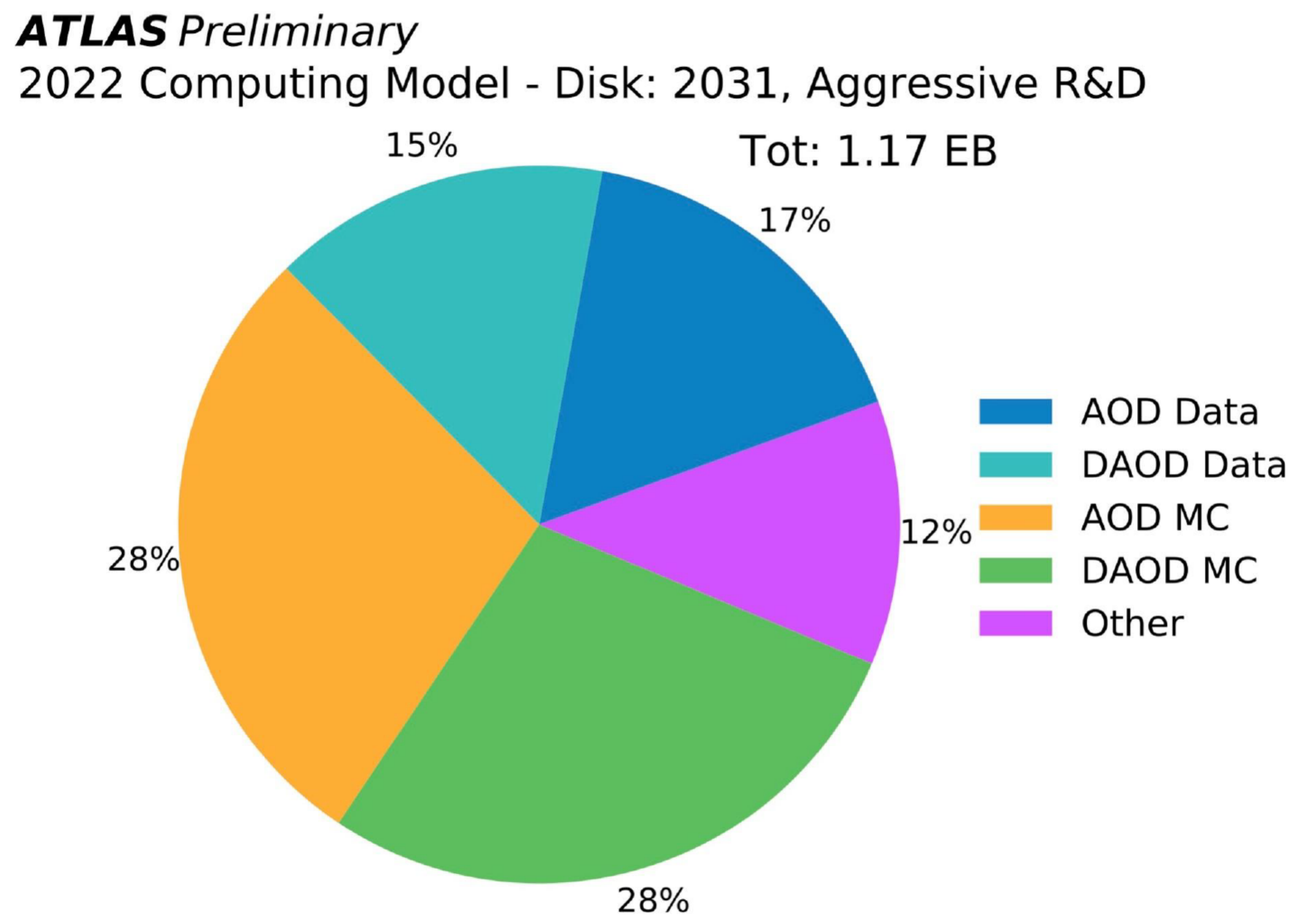
Thank you!

All mistakes are mine
Thanks to contributions from
Alessandro Di Girolamo & Zach Marshall
Rob Gardner, Gordon Watts, Heather Gray, Mike Hance
and others I forgot to write down

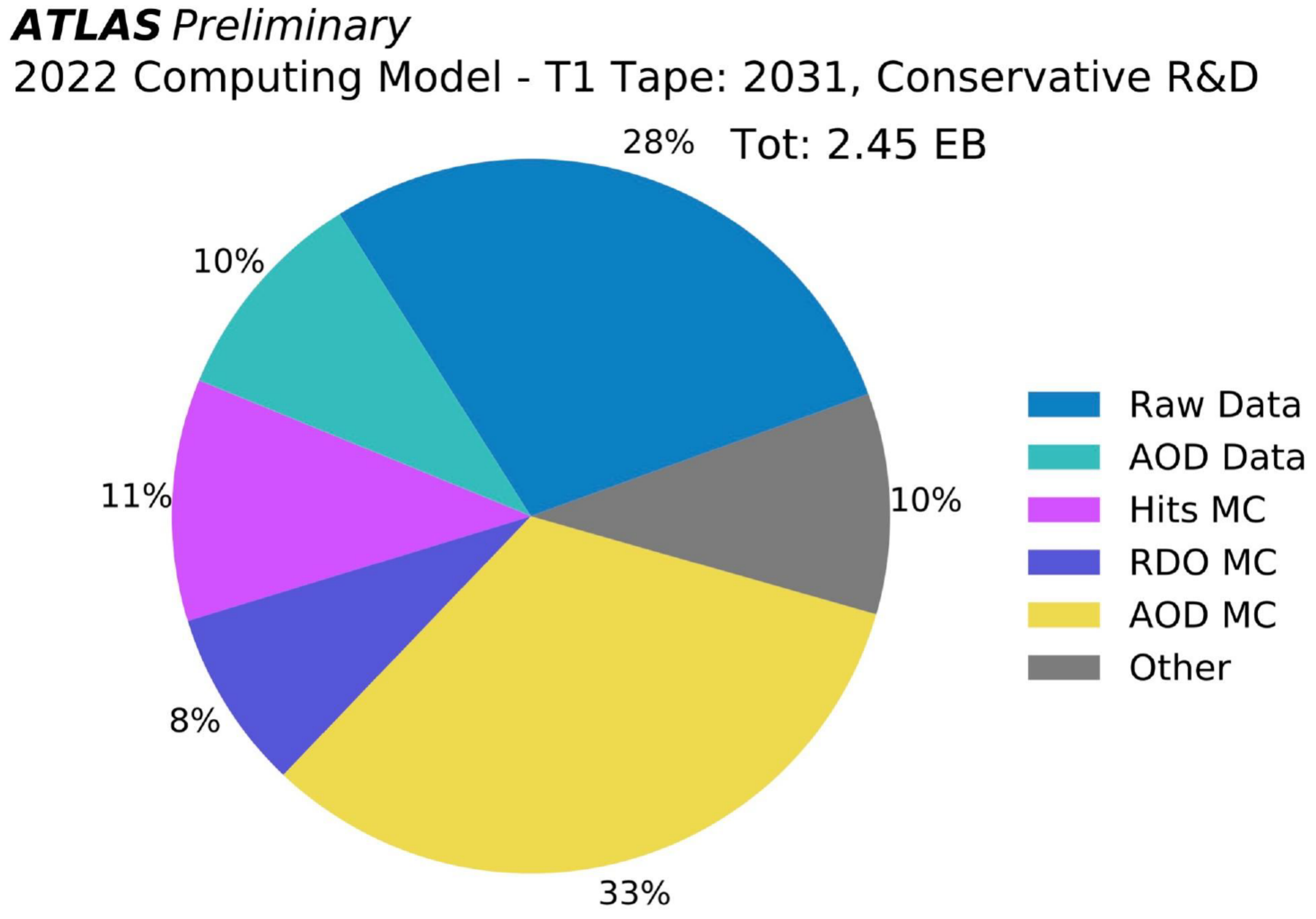
Resource modeling: Disk and Tape Projections

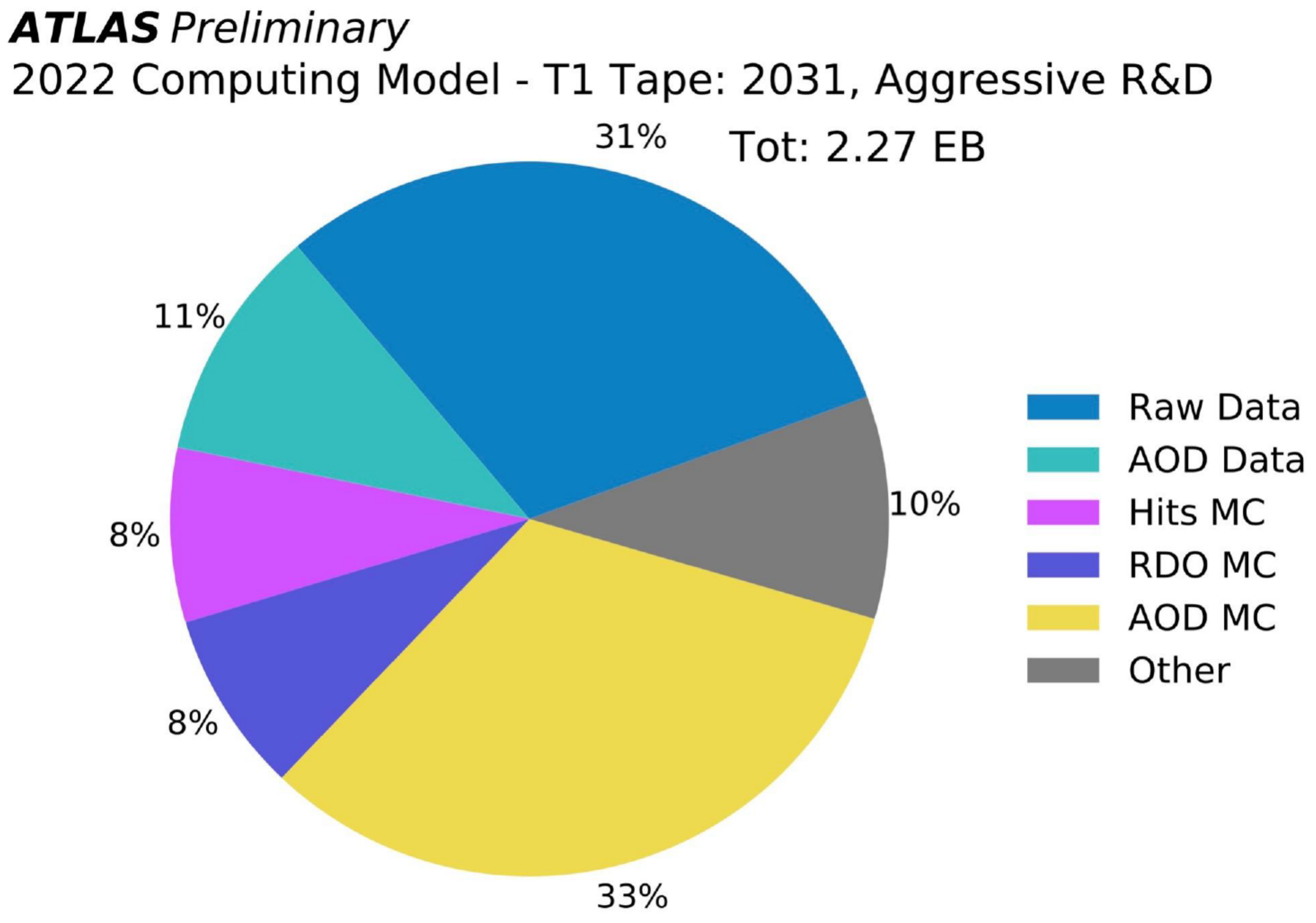
Disk





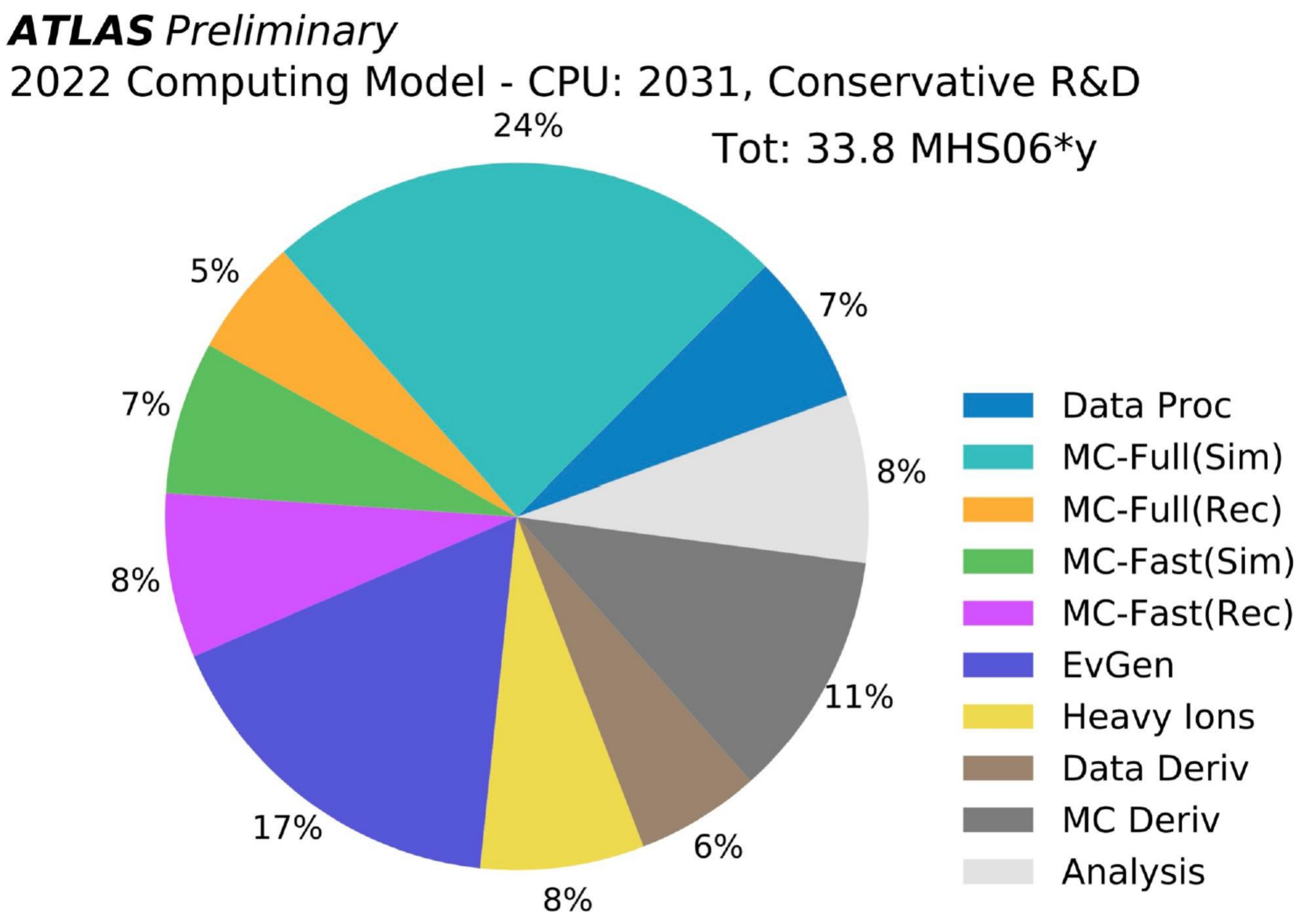
Tape

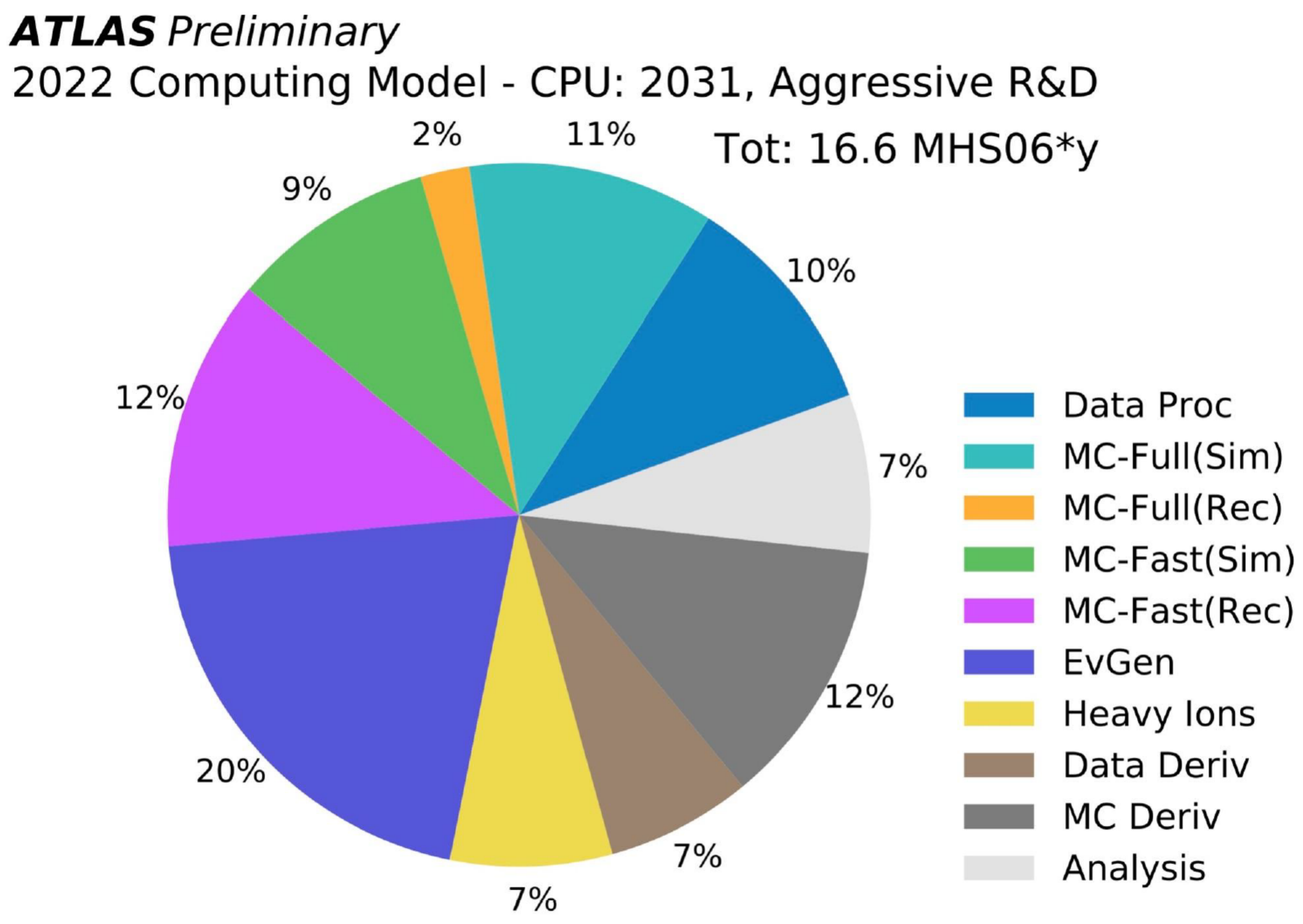




Fractions similar for both scenarios, but aggressive R&D can reduce disk storage by almost factor 2

Resource modeling: CPU Projections





Good mix, but still dominated by FullSim fraction (24%) and EvGen production (17%)

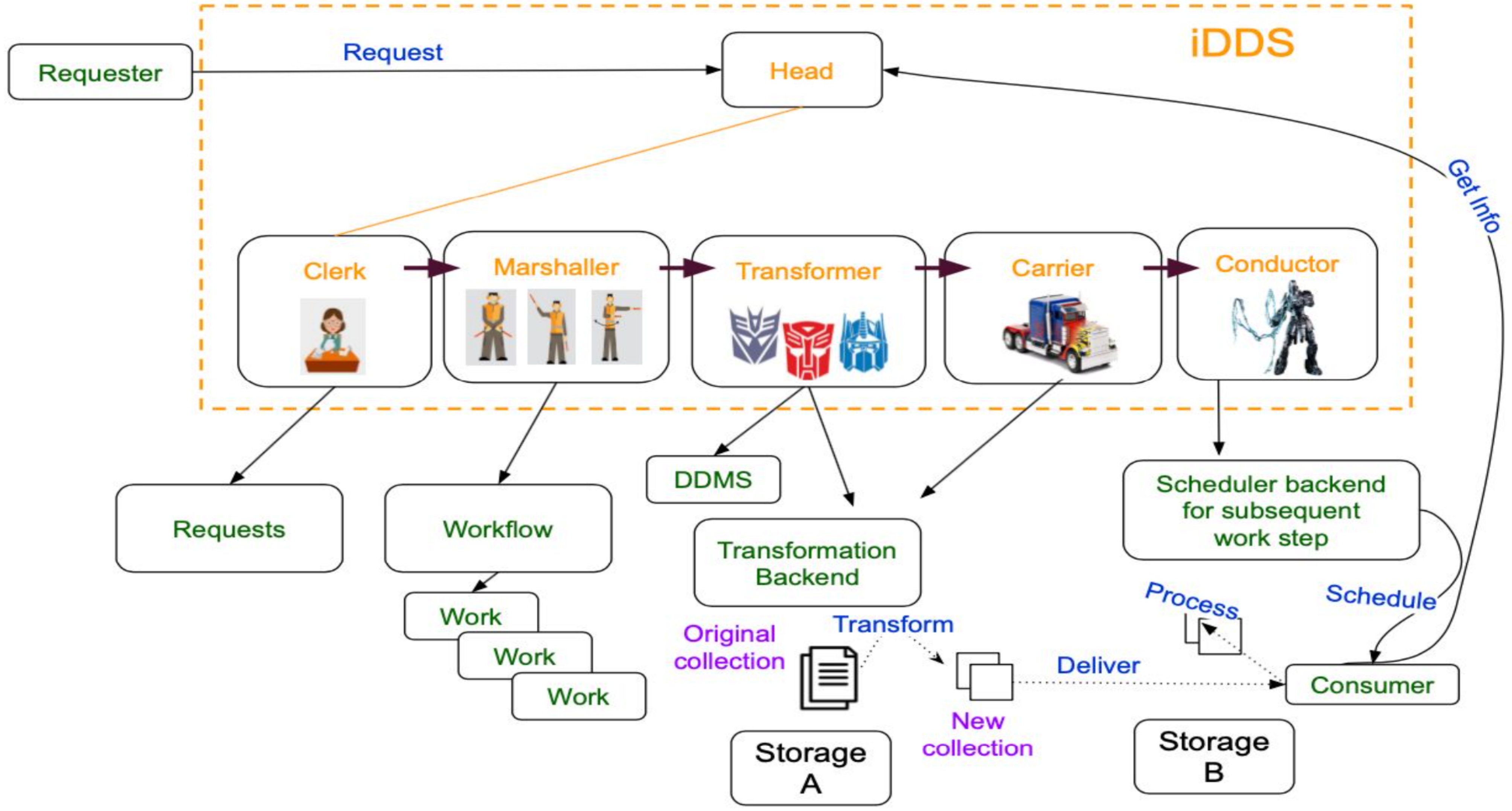
About a factor 2 overall speed-up

FullSim fraction significantly reduced

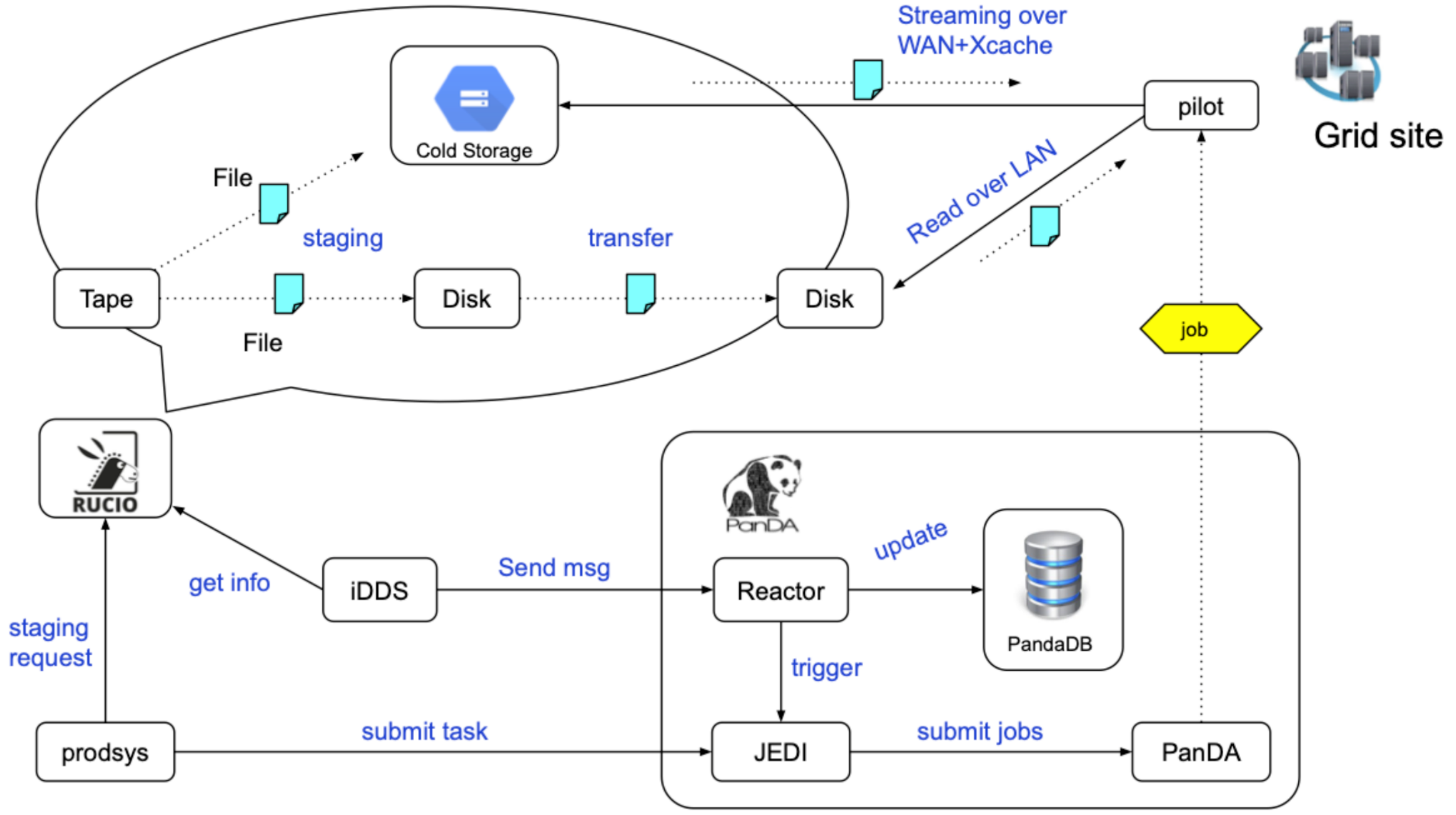
EvGen production now dominates the CPU

Intelligent Data Delivery Service (iDDS)

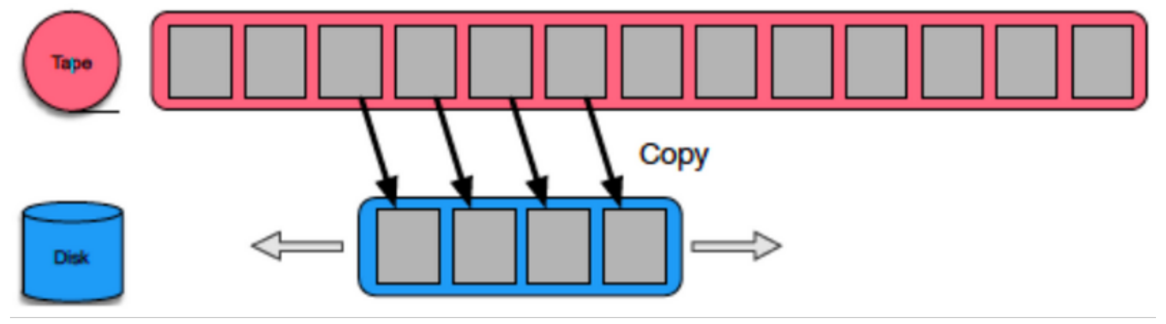
- iDDS is an experiment-agnostic add-on to PanDA or other workload manager supporting granular data delivery and **orchestration of complex workflows** that are efficient in their use of storage, network and processing resources
  - A joint project with IRIS-HEP, project hosted by HSF
  - Used by ATLAS, Rubin, sPHENIX
- Used in a growing list of applications important for HL-LHC readiness and serving/scaling analysis
  - **ATLAS Data Carousel** processes tape-resident data using a small disk storage footprint via a sliding window orchestrated by PanDA, iDDS and Rucio
    - In production for almost 2 years, reducing the storage needs of analysis object data, the dominant storage load for HL-LHC
    - Ongoing R&D to reduce the footprint and improve performance
  - **Highly scalable ML services**
    - Enable analysts to run processing-intensive AI/ML applications on large scale, geographically distributed, heterogeneous resources
    - Shorten optimization and training latencies by orders of magnitude
  - **Active learning services** drawing on the ML work



Intelligent Data Delivery Service (iDDS)

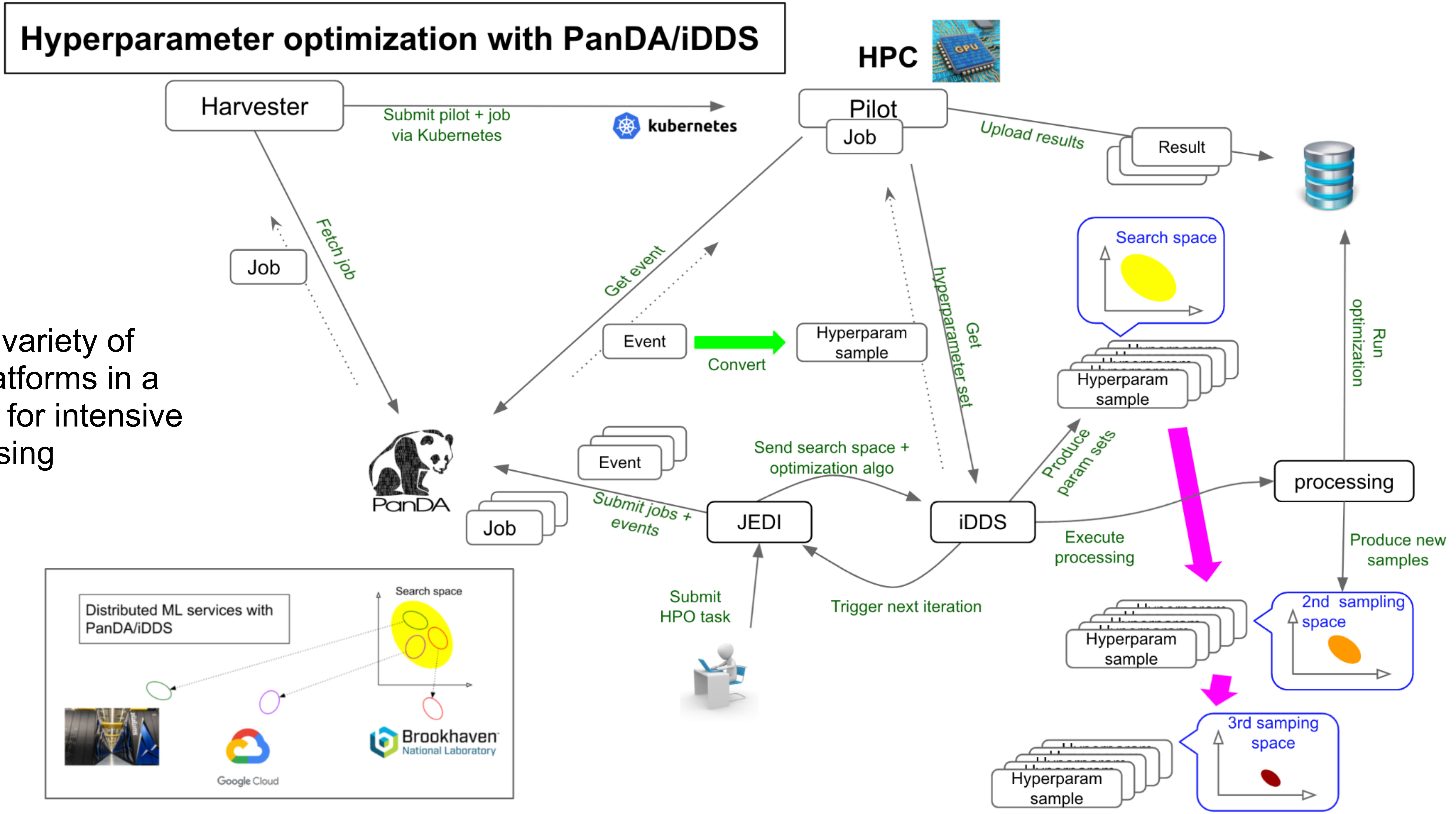


ATLAS Data Carousel using iDDS



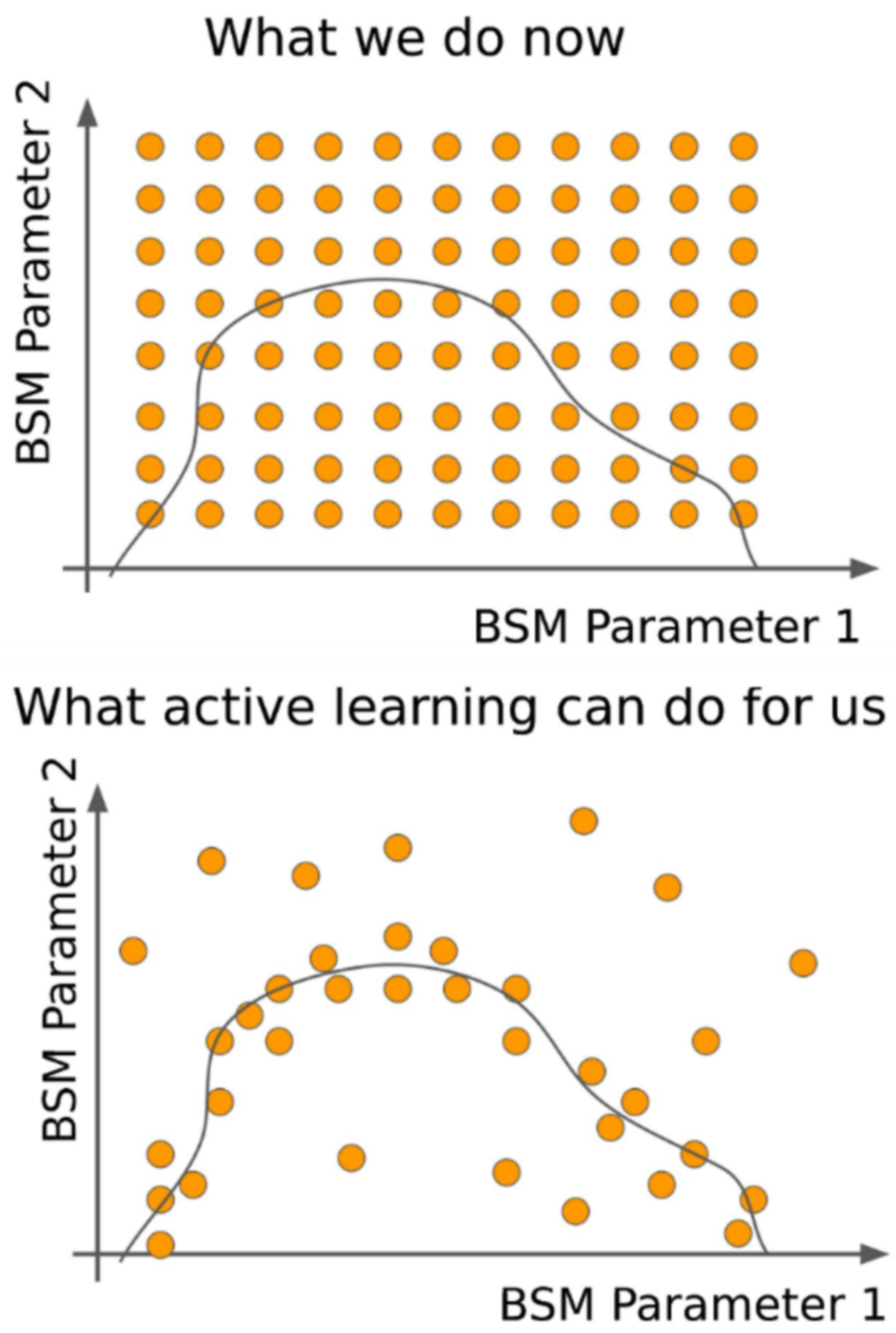
Distributed, scalable ML services

Leveraging a variety of distributed platforms in a coherent way for intensive AI/ML processing



Applying ML services via active learning in analysis

- The hyperparameter optimization (HPO) service we developed is in production use for FastCaloGAN, part of the production ATLAS fast simulation AtlFast3
- In working with the analysis community to find the next large scale application of the HPO service, what emerged was active learning, algorithmically a close relative
- The active learning technique we're using was developed by our ATLAS NYU colleagues
  - "Excursion Set Estimation using Sequential Entropy Reduction for Efficient Searches for New Physics at the LHC", Kyle Cranmer et al, ACAT 2019
  - in calculating an iso-contour surface f(x), conventional approach uses a grid with a sampling density not informed by the unknown f(x)
  - instead, use an iterative approach, using information about f(x) from previous evaluation cycles to sample parameter space more efficiently
  - find an iterative algorithm that suggests points to evaluate that help the most in finding the contour
  - Kyle and colleagues found a computationally efficient one
- In response to interest from analyzers (in particular that team), we adapted our ML hyperparameter optimization service to serve this similar iterative refinement algorithm
- The entire workflow from event generation -> simulation -> reconstruction -> derivation -> limit setting analysis and its iterative refinement loop is implemented and automated using PanDA and IDDS
  - It employs grid and REANA (Reproducible Research Data Analysis Platform) processing resources
- Modular and containerized
  - Analysts provide components specific to their analysis



Active Learning via iterative regression on a limit surface

Opportunistic Resources

- Projects to investigate the use of cloud resources for HEP
- Full ATLAS sites using the Rucio Storage Element and PanDA queue demonstrated on Amazon and Google
- Various payloads including analysis, simulation, athenaMT reconstruction have been demonstrate
- Several setups (Spot and Preemptible instances, virtual machines with different disk configurations) have been evaluated
- Next steps: Demonstrate features and capabilities not available (or limited) on WLCG and University resources. For example, use of ML tools, analytics and other features of GCP for physics analysis.

Machine Learning

- Artificial intelligence and machine learning are used extensively by ATLAS in all areas of event processing and physics analysis
- Prominent USATLAS examples
  - ONNX Runtime integration in athena to enable more applications of machine learning
  - Distributed training enabled the use of GANs to describe the fluctuations in hadronic showers in the calorimeter in Atlfast3
- Many opportunities to profit from advances in artificial intelligence machine learning
  - ML for simulation in detectors beyond calorimeters
  - ML to reduce event generation time
  - ML-based track reconstruction algorithms