ATLAS software and computing road-map

James Catmore, UiO
HL-LHC and ATLAS computing

**Annual CPU Consumption [MHS/years]**

**ATLAS Preliminary**
2022 Computing Model - CPU

- Conservative R&D
- Aggressive R&D
- Sustained budget model (+10% +20% capacity/year)

**Disk Storage [EB]**

**ATLAS Preliminary**
2022 Computing Model - Disk

- Conservative R&D
- Aggressive R&D
- Sustained budget model (+10% +20% capacity/year)
ATLAS road-map for HL-LHC computing

• Prepared as input to the LHCC review of computing for HL-LHC
  • Lays out a concrete plan to deliver the vision set out in the Conceptual Design Report prepared in 2020
  • Consists of a set of milestones to be met in the coming years across all areas of software and computing in ATLAS
  • Milestones are categorised as being as
    • “Maintenance and operation”: needed just to stand still
    • “Conservative R&D”: development work that can be carried out with existing person power within that domain
    • “Aggressive R&D”: development work that will require new person-power or existing personnel committing to new activities, not assumed to come from the relevant domain
  • More aggressive milestones → potentially greater impact on resources
  • Progress will be followed up internally every six months
Roadmap topics

- Project management
- Databases, metadata, conditions
- Core software & heterogeneous computing
- Detector description, simulation, digitisation
- Reconstruction
- Analysis
- Distributed computing
- HPC and analysis facilities
- Collaboration with common projects
- Current or planned Norwegian involvement
## Milestones

### Project management
- Milestone 1: [Start date] - [End date]
- Milestone 2: [Start date] - [End date]
- Milestone 3: [Start date] - [End date]

### Core software & heterogeneous computing

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milestone 1</td>
<td>[Start date] - [End date]</td>
</tr>
<tr>
<td>Milestone 2</td>
<td>[Start date] - [End date]</td>
</tr>
<tr>
<td>Milestone 3</td>
<td>[Start date] - [End date]</td>
</tr>
</tbody>
</table>

### Detector description, simulation, digitization

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milestone 1</td>
<td>[Start date] - [End date]</td>
</tr>
<tr>
<td>Milestone 2</td>
<td>[Start date] - [End date]</td>
</tr>
<tr>
<td>Milestone 3</td>
<td>[Start date] - [End date]</td>
</tr>
</tbody>
</table>

### Reconstruction

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milestone 1</td>
<td>[Start date] - [End date]</td>
</tr>
<tr>
<td>Milestone 2</td>
<td>[Start date] - [End date]</td>
</tr>
<tr>
<td>Milestone 3</td>
<td>[Start date] - [End date]</td>
</tr>
</tbody>
</table>

### Analysis

<table>
<thead>
<tr>
<th>Milestone</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milestone 1</td>
<td>[Start date] - [End date]</td>
</tr>
<tr>
<td>Milestone 2</td>
<td>[Start date] - [End date]</td>
</tr>
<tr>
<td>Milestone 3</td>
<td>[Start date] - [End date]</td>
</tr>
</tbody>
</table>

### Databases, conditions, metadata

- Milestone 1: [Start date] - [End date]
- Milestone 2: [Start date] - [End date]
- Milestone 3: [Start date] - [End date]

### Distributed computing

- Milestone 1: [Start date] - [End date]
- Milestone 2: [Start date] - [End date]
- Milestone 3: [Start date] - [End date]
Current involvement

- ATLAS distributed computing management, NorduGrid/ARC middleware development, ATLAS@Home, HPC integration
- Tau reconstruction
- Derivation framework, data formats, distributed analysis
- Development and follow-up of the milestones
Contributions to the milestones
(Non-exhaustive; based on current expertise and interests)

- GPU programming in core software
- Machine learning in reconstruction
- High performance analysis
- HPC activities, new workflows

NFR application
(Catmore & Gramstad)

NFR application
(Catmore & Gramstad)

NFR application
(Catmore & Gramstad)

ATLAS QT (Langrekken)

DC-6
Integrating next generation of HPCs
5.1 Integration of at least 2 EuroHPC sites
5.2 Integration of next generation US HPCs for production

DC-6
Exploratory R&D on GPU-based workflows for next generation HPC

DC-7
HL-LHC datasets replicas and versions management
7.1 Replicas and versions detailed accounting
7.2 DAOD replicas reduction
7.3 DAOD versions reduction

DC-8
Data Carousel for storage optimization
6.1 Investigate with sites the cost of Tape infrastructure and the estimated cost in case of sensible increase of read/write throughput
6.2 Reduce the AOD on disk to 50% of the total AOD volume, using Data Carousel to orchestrate the stage from tape for DAOD production.

DC-9
Disk management: secondary(cached) dataset
9.1 Evaluate the impact on job brokering and task duration if disk space for secondary data is reduced

AN-3
Prototyping & review of columnar data operations for end-to-end analysis
3.1 Tests of basic reading performance using TTree version of PHYSLITE
3.2 Prototyping of tools for columnar CP operations and other systematics
3.3 Adoption of ROOT7 data structures
3.4 Performance and ease-of-use assessment leading to decision on adoption

AN-4
Development of columnar analysis infrastructure
4.1 Prototyping of framework for orchestrating columnar CP operations
4.2 Development of columnar skimming/augmentation
4.3 Demonstrate end-to-end analysis using columnar tools/flw + dist. comp.
4.4 Development & roll-out of documentation/training for run 4 analysis

CS-12
Make ATLAS Data Model classes accelerator-friendly
12.1 Prototype GPU-friendly xAOD classes
12.2 Support for reduced/mixed precision in ATLAS EDM
12.3 Decision on xAOD API evolution
12.4 Event-batching and EDM
Summary

• **ATLAS people**: please read the document and consider whether new long-term commitments could contribute to delivering these milestones whilst fitting into the strategy of NorCC

• **Non-ATLAS people**: are there areas where we could meaningfully collaborate?