PIC
port d'informació
científica

IFAE
Institut de Física
d'Altes Energies

ATLAS Distributed Computing Operations towards HL-LHC

US ATLAS Meeting (Madison, USA)

10th July 2024

Mario Lassnig (CERN) and Andreu Pacheco (IFAE/PIC)
ATLAS ADC Coordination

CERN

ATLAS EXPERIMENT

Purpose of this presentation

- To discuss the **current status and future of ADC Operations**.
- To highlight **challenges faced by ADC Operations**, like **declining workforce** and **limitations of the tools** we use.
- To explore **opportunities for improvement**, such as **tool development** and **US ATLAS contributions**.
- To identify **critical central services and** discuss their **evolution**.
- To address **concerns about SAM, HammerCloud, and communication with sites and clouds**.
- To emphasize the **importance of US ATLAS Distributed Computing and US ATLAS site engagement** and **collaboration in shaping the future of ADC**.
- To discuss the **integration of US HPCs** as "Standard Sites".

ADC Operations (simplified)

- In charge of the daily smooth computing and data management processing in ATLAS
  - Following the evolution of the number of running slots, pending jobs, transfer states
  - Following and solving relevant incidents reported by shifters, sites, activity coordinations, production managers, users.
  - Addressing the relevant expert as a function of the issue.
  - Configuring relevant parameters in Rucio, CRIC, …
  - Reporting sites with issues relevant to them.
  - Integration of new developments or technologies (IAM,...)
- We rely on cloud and site support for issues related to sites
  - Too many sites and infrastructures to cover.
  - ADC Operations tries to help but sometimes it is not sufficient.
  - New OSG/WLCG sites, HPCs
  - Site upgrades
  - There are cases that site issues are global operational issues but they are hard to identify.

ADC Operations (simplified)

- Several teams involved:
  - Distributed Production and Analysis (DPA)
  - Distributed Data Management (DDM)
  - Workflow Management System (WFMS)
  - Tier-0
  - Central Services Operations (CSOPS)
  - Distributed Analysis Support Team (DAST)
- Related operation teams:
  - WLCG Operations
  - CERN IT
- Filter and report the issues to developers
  - IAM, Rucio, CRIC, Prodsys2, Panda, …
- Philosophy:
  - Solve the issues affecting production as fast as possible with contacting directly the appropriate experts.

Operations in the ADC Structure

ATLAS DISTRIBUTED COMPUTING
March 2024

ADC COORDINATION
Mario Lassnig, Andreu Pacheco Pages

PHYSICS
Production Coordination
M. Borodin
Analysis Coordination
A. Forti

Centralised Production
- Monte Carlo Production
- Group Production
- Data Reprocessing
- Physics Validation
- HLT Reprocessing

Physics Analysis
- User Analysis Tools
- Analysis Model Group
- DAST

FABRICS
Coordination
V. Garonne

Infrastructure
- Tier-0
- Grid
- HPC
- Cloud
- BOINC
- Analysis Facilities

Operations
- Computing Run Coordination
- DA Operations
- DPA Operations
- Central Services
- CRIC
- HammerCloud
- Monitoring
- ADCoS

DATA MANAGEMENT
Coordination
S. McKee, P. Vokac

System
- Rucio

Operations
- System Deployment
- DDM Central Operations
- Monitoring

Research
- Networks
- Caches
- Storage
- Cloud

WORKFLOW MANAGEMENT
Coordination
R. Walker, F. Barreiro Megino

System
- Workflow Definition
- Workload Management
- Workload Execution

Operations
- System Deployment
- Monitoring

Research
- Data Analytics
- Analysis Facilities
- Cloud
- HPC

Operations are spread across several ADC areas

ADC Operations meetings (see ADC twiki)

- Daily meetings:
  - ADC Operation meetings daily 9:00 CERN time (zoom)(codi)
    - Any issues related to DPA, DDM, DA, Tier0,...
  - (New) US ATLAS Operation meetings daily at 9:30 CDT (16:30 CEST)
    - Chaired by the same person as the morning ADC Ops meeting
- Weekly dedicated meetings after CERN morning meetings:
  - DDM devops meeting weekly at 9:30 CEST on Tuesdays (codi)
  - Tier0 ops meeting weekly at 9:30 CEST on Wednesdays (codi)
  - Central services ops meeting weekly at 9:30 CEST on Thursdays (codi)
- Weekly meetings with devops teams
  - DPA meeting weekly at 10:30 CEST on Thursdays (indico)
  - WFMS meeting weekly at 15:30 CEST on Thursdays (indico)
  - Rucio meeting weekly at 15:00 CEST on Thursdays (indico)
  - Monitoring meeting fortnightly at 16:00 CEST on Fridays (indico)
  - ATLAS - IT STORAGE ops meeting weekly at 10:00 CEST on Mondays (indico)
  - WLCG Ops meeting weekly at 15:00 CEST on Mondays (indico)
- Monthly meetings
  - ATLAS Fabrics: ADC, Sites & Services Providers (indico)

List of questions on ADC Operations

- How do we see the evolution of ADC Operations?
- What are critical central services?
- What about SAM (Some people think that they are not sure that site availability is 100% correct based on SAM testing)?
- What about HammerCloud?
- How do you see communication between sites and clouds?
- How site's concerns could be addressed?
- What ADC Coordinators want to hear from US ATLAS Facilities and Distributed Computing?
- How ADC sees future of HPC and HPC effort in ADC?

How do you see the evolution of ADC Operations?

- Challenges:
  - Declining Workforce: Fewer people are contributing to central operations, raising concerns about long-term sustainability.
  - Tool Limitations: Existing tools may not be optimized, potentially hindering efficiency.
- Opportunities:
  - Developments: Upgrading central operations tools could reduce overall workload, including cloud operations.
  - US Contribution: The US has a strong track record of central operations contributions, offering a potential source of support.
- Considerations:
  - Focus on Practical Solutions: Prioritize practical solutions over complex automation with large language models (LLMs) to avoid the "operational intelligence trap."
- Explore opportunities to improve central operations staffing and tools for long-term efficiency.

What are critical central services?

- Challenges:
  - Dynamic Landscape: Critical central services evolve over time. (e.g. VOMS phasing out)
  - Incomplete Lists: Listing all critical services is difficult, potentially overlooking important ones.
- Current Services (Examples):
  - Identity & Access Management (IAM)
  - Rucio (Data Management)
  - File Transfer Service (FTS)
  - HammerCloud (HC)
  - Workflow Management Ecosystem (Panda/DEfT/JEDI/ProdSys2/Harvester)
  - Monitoring Tools (Monit/Grafana, BigPanda, Kibana)
  - ARC Control Tower (ACT)
  - Job Submission (Pilot)
  - Computing Resource Information Catalogue (CRIC)
  - Frontier servers

What are critical central services?

- Uncertainties:
  - Unclear future direction for central services which keep changing with technology evolution.
    - Supporting new infrastructures
    - Change of operating systems
    - Change of technology
    - Change of the service implementation.
- Proposal:
  - Centralized Review: Regularly evaluate core functionalities and identify trends to optimize service usage and recommend best practices.

What about SAM?

- Current State:
  - Agreement that existing SAM/ETF tests are sub-optimal and don't provide accurate reliability metrics.
- Proposed Improvement:
  - Update SAM/ETF tests to reflect a more accurate picture of availability and reliability.
- Concerns:
  - Potential for confusion (and instability over time) if metrics are updated too frequently.
  - Reliance on assumptions and potentially biased opinions from a small number of sites.
- Call to Action:
  - Carefully consider the risks and benefits before updating SAM/ETF tests.
  - Gather broader input to ensure revised metrics accurately represent system performance.
- SAM/ETF tests affect reliability and availability numbers, not production

What about HammerCloud?

- Importance:
  - HammerCloud is a critical tool for current operations with unique functionalities.
- Problem:
  - HammerCloud lacks of sufficient support.
  - HammerCloud blacklisting may cause losing a lot of resources
    - Probes
    - Production payload causing overload
      - We have cases where 40% failures is acceptable by requesters.
    - Policy how to apply blacklisting
    - All or nothing policy can be rethought.
      - IO limits
      - reduction of load?
- Potential benefit:
  - More development effort (including maintaining core functionality during migrations) could improve HammerCloud as a tool for everyone.

How do you see communication with sites and clouds?

- The introduction of **monthly Fabrics meetings is a positive step** towards increased dialogue (Jul 3rd meeting)
  - Discussions are more and more interesting.
- We acknowledge a **perceived decrease in engagement from clouds in central operations discussions**.
- We emphasize the **value of user and site feedback**, even constructive complaints, in identifying areas for improvement.
- We acknowledge the **positive improvements in communication with the US cloud** in the past year.
  - Daily US-ATLAS Operations meeting in one example
- We aim to maintain **constructive and solution-focused discussions** moving forward.

How sites' concerns could be addressed?

- Existing Communication Channels
  - Established communication paths (meetings, email, Mattermost, Discourse, JIRA, GGUS)
  - Annual site jamboree (limited to once a year)
- Openness to Improvement
  - Acknowledges these methods might be insufficient.
  - Welcomes suggestions for alternative communication channels for sites to express concerns.
- Proactive Communication from Sites
  - Emphasizes the importance of sites proactively reaching out with suggestions.
  - Encourages sites not to wait for central operations to initiate communication.
- Possibility to import internal site information into the ATLAS monitoring.
  - Many issues cannot be understood centrally but if monitoring is there diagnosing would be easier.

What ADC Coordinators want to hear from US ATLAS Facilities and Distributed Computing?

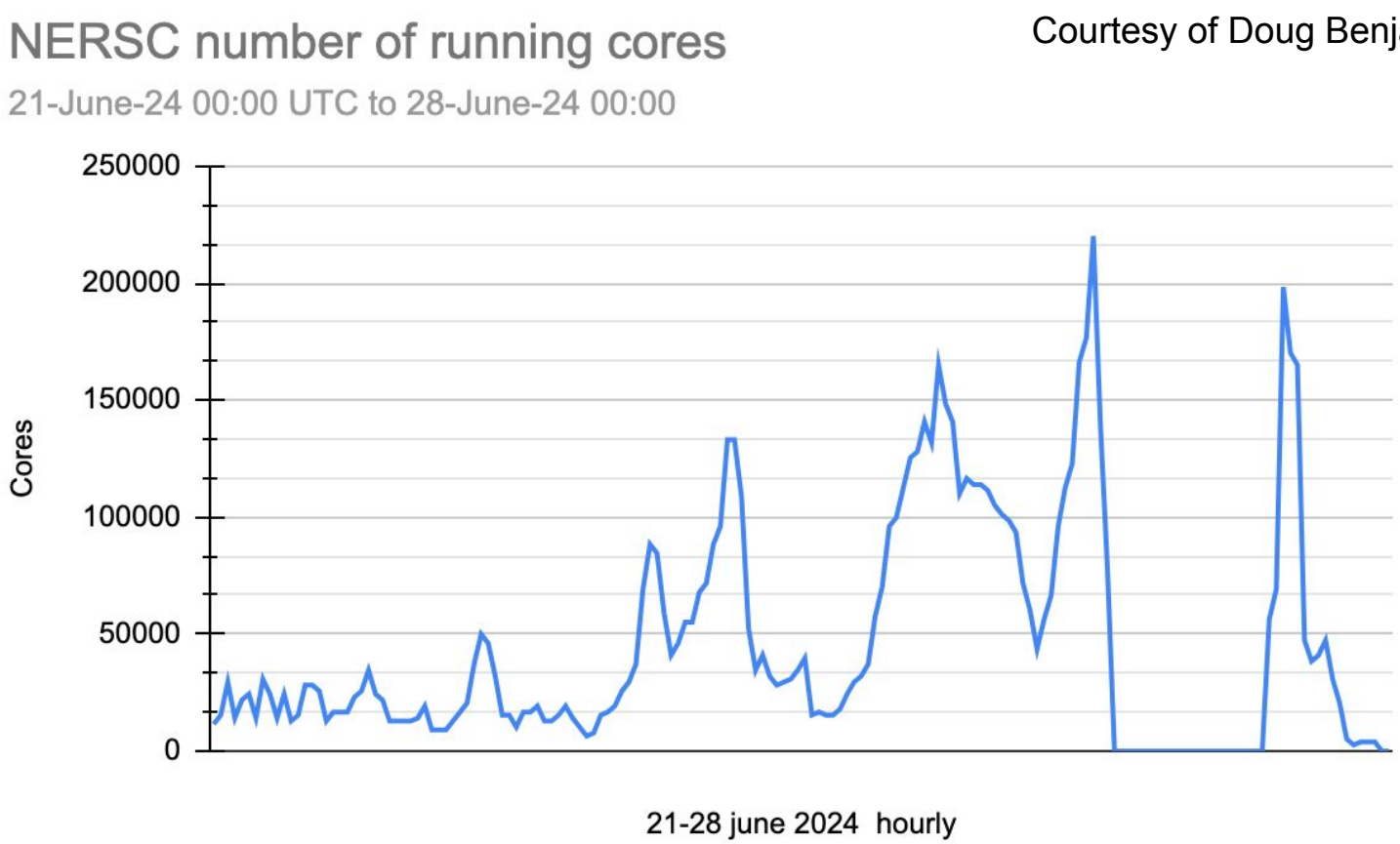
- Importance of US-ATLAS Engagement:
  - Experienced ADC personnel and site operators in the US hold valuable expertise.
  - Their feedback is crucial for improving central operations and shaping the future of ADC.
- Collaboration for Mutual Benefit:
  - Central operations relies heavily on the support of sites.
  - Increased site engagement directly translates to greater influence within the ADC community.
- Areas of joint collaboration:
  - Improve site monitoring accessible to ATLAS central services.
  - Collaboration to develop the next version of SAM/ETF and HammerCloud tools.
- Call to Action:
  - We actively encourage US sites to share their insights and concerns.
  - By working together, we can ensure a stronger, more effective ADC for everyone.

How ADC sees future of HPC and HPC effort in ADC?

- HPC Task Force:
  - This is one of the ADC Official Milestones (DC5)
- Importance of HPCs:
  - Currently 30% of the CPU contribution and going up! NERSC Perlmutter
  - HPCs are seen as a crucial part of a balanced resource plan due to their immense computing power that can relieve pressure on WLCG sites.
  - We have to distinguish between a real HPC and a cluster with the HPC name on it.
- Challenges of HPCs:
  - The large size and eventual retirement of HPCs, like Vega, pose a risk to resource stability.
  - HPCs work on allocation of resources which have to be used in a given time.
  - People supporting HPC are working on their own with specific solutions.
- Desired Outcome:
  - Creating an ADC Group dedicated to HPC operations could be a way to go.
- Focus Area:
  - Integrating US HPCs as "Standard Sites" is considered the most valuable effort to maximize the benefits of these resources and streamline scheduling.
  - Contribution to the HPC global effort.

NERSC_Perlmutter

Courtesy of Doug Benjamin



Conclusions

- ADC Operations faces challenges due to declining workforce and tooling limitations.
- Investment in tooling and US contributions are potential solutions for improvement.
- Critical central services need to be regularly evaluated and optimized.
- SAM/ETF tests and HammerCloud require attention and potential updates.
- Communication and collaboration with sites, especially in the US, are crucial.
- Integrating US HPCs as "Standard Sites" is a key focus area for the future.

Backup slides