

DOMA – What's New? What's Different?

Brian Bockelman






**Institute for Research & Innovation
in Software for High Energy Physics**

Emphasis for IRIS-HEP 2.0: The Road To Production

- The focus for the Institute in its current phase is having our R&D activities transition to production for the LHC experiments, making a sustainable impact.
- Recall – the 4 Computing Gaps outlined in the IRIS-HEP Strategic Plan:
 - G1: Raw Resource Requirements
 - G2: Scalability of the Distributed Cyberinfrastructure
 - G3: Analysis at the HL-LHC Scale
 - G4: Sustainability

Scaling the CI to HL-LHC Data Rates

- From first principles, we can estimate the minimum HL-LHC data rates – and they're imposing!
- The Data Challenge series of exercises are a community-wide plan to show scale milestones and maturity of technologies.
- IRIS-HEP Activities:
 - Rucio/SENSE integration (Added late in v1.0)
 - Maintain a reference platform (XRootD/HTTP-TPC).
 - Technology: Authorization Overhaul
 - Participation in the DC coordination.

<u>Year</u>	<u>Minimal</u> (Gbps, %)	<u>Flexible</u> (Gbps)
2021 	480, 10% 	960 
2023 2024	1,440, 30%	2,880
2025 2026	2,880, 60%	5,760
2027 2028	4,800, 100%	9,600

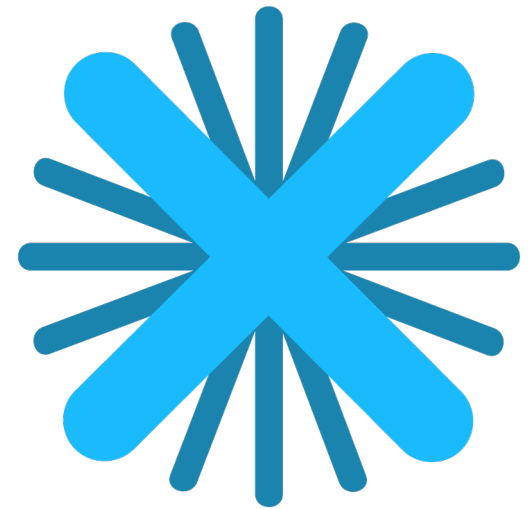
Original DC plan above – numbers are moving up at to the right!

Authorization technology overhaul

- The WLCG is in the middle of a large-scale authorization technology overhaul.
 - The token-based approach hews toward industry standards (G4. Sustainability) and is the path for improved security.
- The next 2-3 years are a critical juncture:
 - Initial set of services (CEs) upgraded. Switching to HTTP-TPC was first step in this direction for transfers.
 - Technology deployed widely (esp. at SEs); needs configuration and use.
 - DC24 is target for pure token-based transfers.
 - IAM & VOMS-Admin (obsolete) are living side-by-side; VOMS-Admin needs to be retired.

Delivering Columnar Data with ServiceX

- **New:** UT-Austin team, led by Peter Onyisi, is joining IRIS-HEP 2.0 with contributions to ServiceX.
- ServiceX has made significant strides this year in performance and integration with other parts of IRIS-HEP – much of this to be demo'd as part of the larger AGC Demonstration Event on Thursday.
- Participate in the ServiceX planning session at 1:30 today in the main room.



Core data streaming with XRootD and XCache

- **New** for IRIS-HEP 2.0! Done in conjunction with OSG-LHC area.
- During the strategic planning process, XRootD (and XCache) were identified as a foundational technologies for the LHC's distributed cyberinfrastructure. They are key to projects inside and outside of the Institute.
- For IRIS-HEP 2.0, we're supporting ~0.5FTE of a developer to:
 - **Evolve bulk data transfer:** refining HTTP-based data movement, scaling the data rates, implement new requirements for HTTP-TPC.
 - **Contribute to the core** development, testing, integration, and software delivery .
- Expected impact: Help sustain this core technology, ensure it's ready as a reference platform for DC24 and remains usable within Coffea-Casa.

Coffea-Casa

- Coffea-Casa – a joint project between AS, SSL, and DOMA – is an integration point for Analysis Facility R&D technologies.
 - It is delivered both as **software** (Helm chart built on top of JupyterHub, Coffea-Casa Docker containers, tutorials) and **services** (usable facilities at Nebraska and Chicago).
 - Used by the AGC team for the demonstration events that integrate Coffea, ServiceX, ML inference service, XCache, token-based authorization.
- Going into Phase 2, Coffea-Casa must build on the accomplishments above and:
 - Start pushing on the intersection with Analysis Preservation. How do we capture these analyses when they interact with so many new services?
 - Coordinating with the growing network of LHC analysis facilities. The pieces of the Coffea-Casa facilities serve as a reference point but aren't the global picture.

Continuing with the theme of “going into production”, how do we transition with the Ops programs?

Activities and Impact

Activity	G1. Resource Requirements	G2. Scaling the Distributed CI	G3. Analysis at the HL-LHC	G4. Sustainability
Scaling the CI to HL-LHC Data Rates		X		
Authorization Technology Overhaul				X
Columnar Data with ServiceX			X	
Coffea-Casa			X	
XRootd & XCache	?			X

Questions?

This project is supported by the National Science Foundation under Cooperative Agreements OAC-1836650 and PHY-2323298. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.