

hpc-ch forum on Business Continuity / Crises Management & Energy Savings and Efficiency - Hybrid Event

Contribution ID: 9

Type: **not specified**

Community Development

Thursday, 23 May 2024 14:00 (1 hour)

- New community members
- Updates from the community

Efficient HPC Cluster Deployment in the Cloud: Strategies for Scalability, Speed, and Self-Service

This talk outlines strategies for deploying High-Performance Computing (HPC) clusters in the cloud to enhance operational efficiency and flexibility. We discuss our method for rapidly provisioning dedicated cloud clusters, providing a reliable overflow capacity during peak demands and a robust backup during on-premise outages or planned interventions. Our experience with seamless cloud bursting, alongside our future goals for user-driven resource management through AI-enhanced tools, highlights our commitment to advancing cloud-based HPC solutions. The presentation will detail the technical achievements and the roadmap for empowering users to autonomously manage computing and storage resources.

Speaker: Pablo Llopis, EPFL