Contribution ID: **68** Type: **Talk (15min + 5min)**

Reconciling Life Sciences and HPC

Wednesday, March 6, 2024 4:10 PM (20 minutes)

Despite the immense computational power offered by HPC clusters(and the resources governments pour into obtaining these resources), it remains uncharted territory for many researchers. The primary deterrents include the perceived bureaucratic hurdles associated with accessing HPC resources, the opacity of usage procedures, and a notable lack of accessible support systems.
The intricate process of gaining access to HPC systems, coupled with the convoluted nature of usage procedures, has led to a situation where life science researchers seek alternatives. Institutions, responding to this demand, resort to establishing additional localized infrastructures, resulting in a fragmented and non-standardized compute ecosystem with lots of redundancy. The consequence is a landscape marked by redundancy, inefficiency, and a lack of standardized practices, all of which hinder the realization of the true potential of HPC in life science research.
To bridge this gap, a pivotal step involves the adoption of contemporary workflow systems that seamlessly integrate with HPC batch systems and support remote file management for research data management support. This talk tells the story of incorporating native batch system support into the snakemake workflow management system, advancements by which life science researchers can overcome traditional obstacles and tap into the full capabilities of distributed cluster computing. Highlighted are success stories from massive pharmaceutical ligands screens and genome oriented research.
As the talk explores these advancements, we extend an invitation to engage with the Snakemake developer community. Collaboratively, we can contribute to making bioinformatic workflow solutions more accessible and aligned with Open Science goals. Join us in this practical journey towards a more efficient and collaborative future in life science research.

Slot length

Primary author: MEESTERS, Christian (Johannes Gutenberg Universität Mainz)

Co-authors: Dr KÖSTER, Johannes; HELLMANN, Lukas (Johannes Gutenberg Universität Mainz)

Presenter: MEESTERS, Christian (Johannes Gutenberg Universität Mainz)

Session Classification: Parallelization and HPC Infrastructure

Track Classification: HPC and Massive Data: Parallelization and HPC Infrastructure