Dynamic Use of Resources for Affordable Exascale and Beyond

Organizers:

- Pierre-François Dutot (UGA)
- Antonio Peña (BSC)
- Jesus Carretero (UC3M)
- Martin Schreiber (UGA)
- Estela Suarez (FZJ)

Speakers:

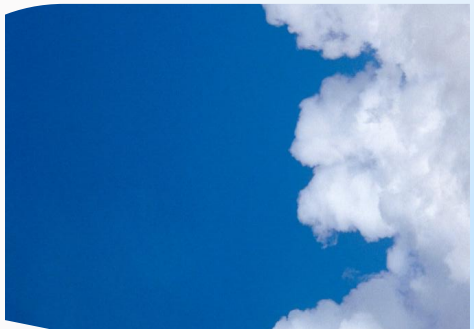
- Hans-Christian Hoppe (FZJ)
- Dominik Huber (TUM)
- Sergio Iserte (BSC)
- Daniel Milroy (LLNL)
- Martin Schulz (TUM)
- David E. Singh (UC3M)

Agenda

00:00 - 00:15
2 minute (strict!) appetizers

00:15 - 00:50
Discussion / questions from the audience

00:50 - 01:00
Final statements



Speakers

Sergio Iserte (BSC)

Senior Researcher

Dominik Huber (TUM)

PhD candidate

Daniel Milroy (LLNL)

Computer Scientist

+ ?

Martin Schulz (TUM)

Professor in CE

Hans-Christian Hoppe (FZJ)

Senior Project Lead

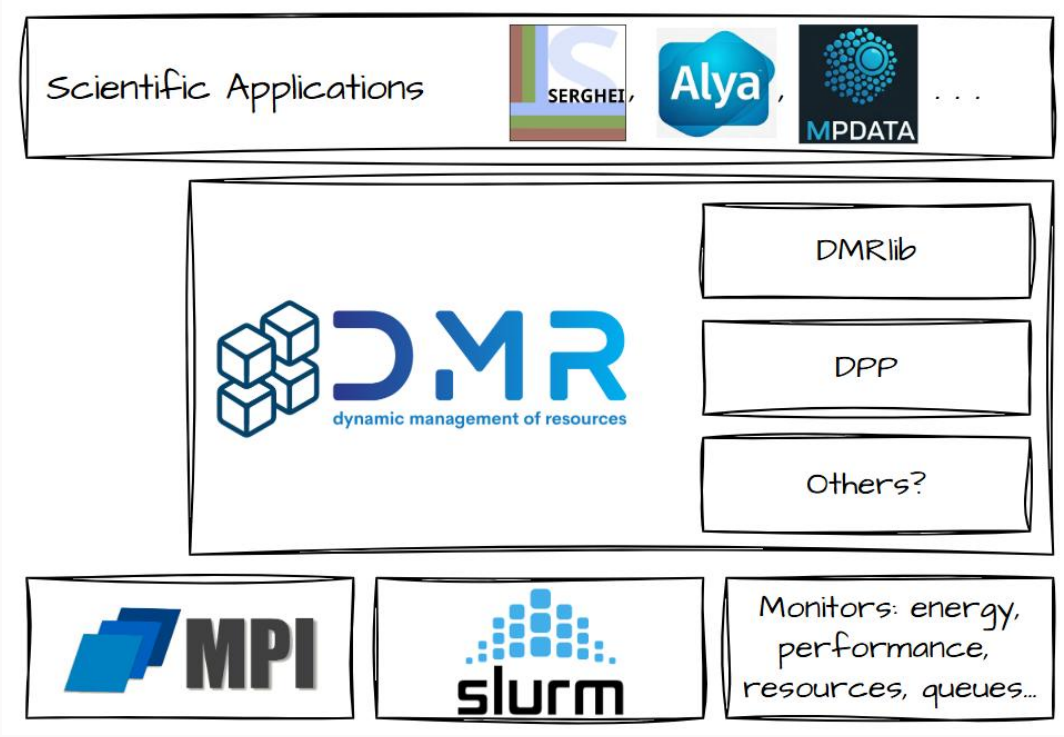
David E. Singh (UC3M)

Associate Professor

Sergio Iserte (BSC)

The Dynamic Management of Resources (DMR) Framework

An API that simplifies the development of dynamic applications.

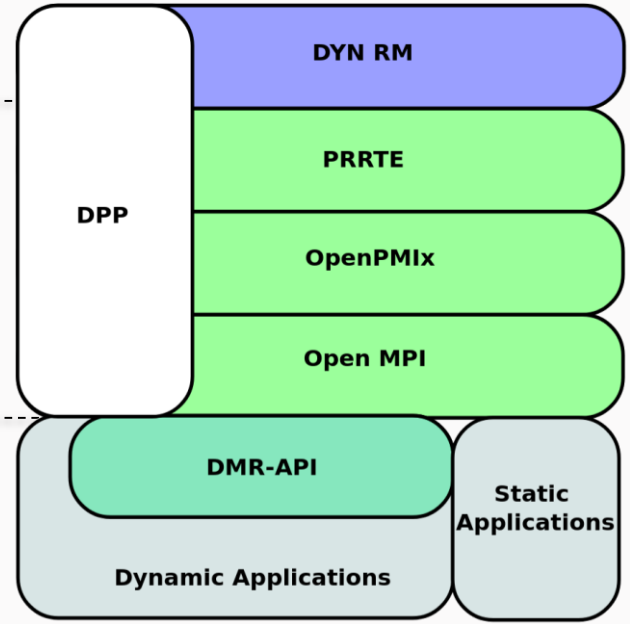


Dominik Huber (TUM)

Dynamic Resource Manager

Dynamic
Middleware Layer

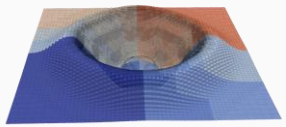
Application Layer

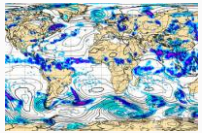


PMIx 10^18

MPI

https://gitlab.inria.fr/dynres





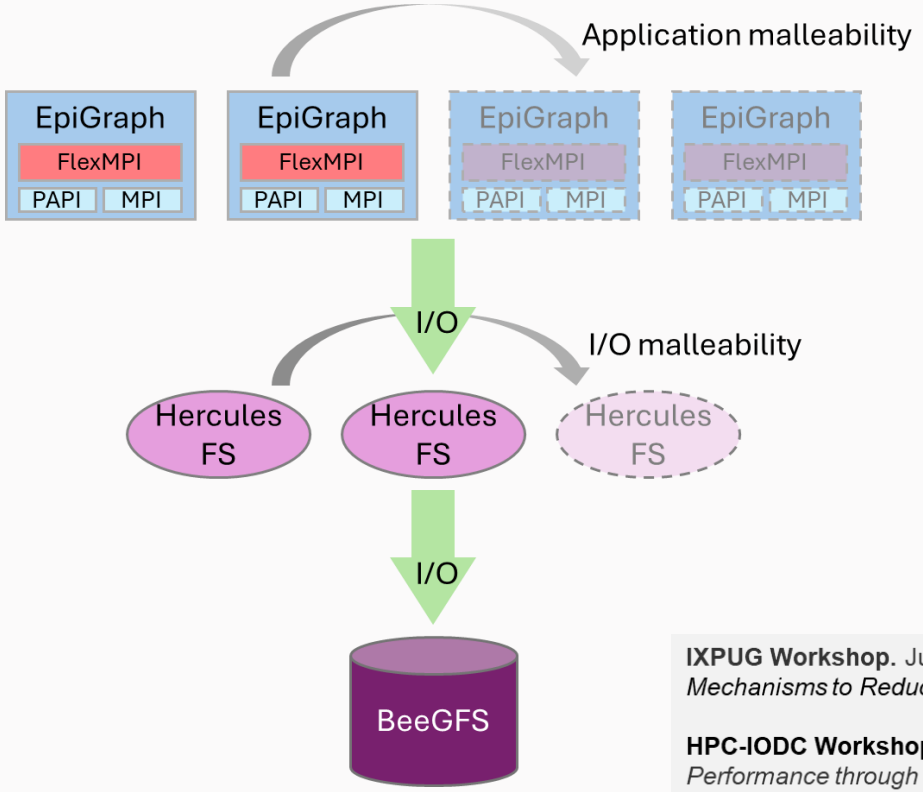
Bridging the Gap Between Genericity and Programmability of Dynamic Resources in HPC

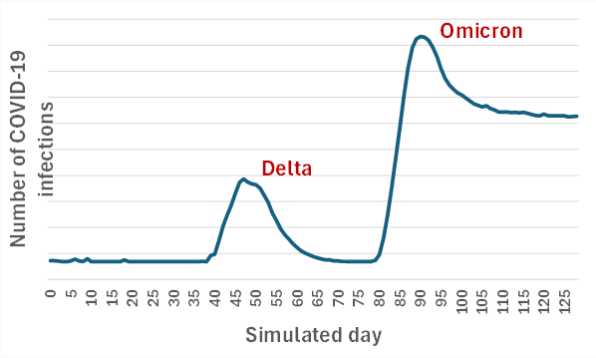
Wednesday, June 11, 2025 11:10 AM to 11:35 AM · 25 min. (Europe/Berlin)

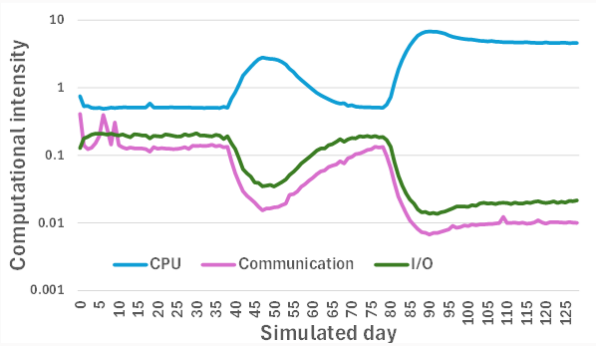
Hall F - 2nd floor

Research Paper

David E. Singh (UC3M)







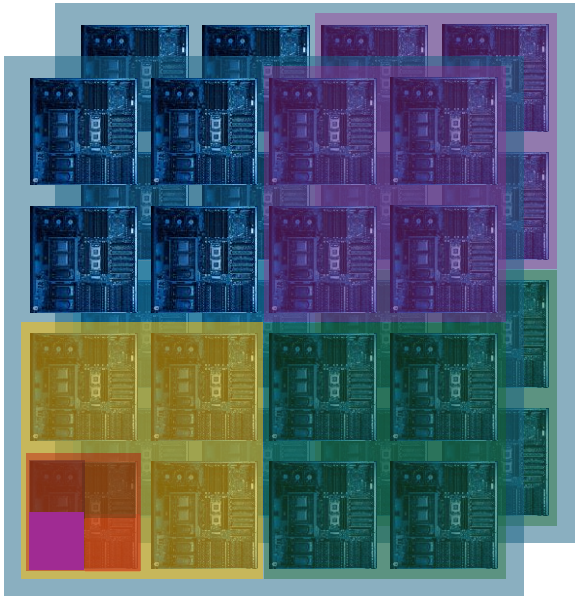
IXPUG Workshop. June 13, 3pm. *Combining Malleability and Distributed Control Mechanisms to Reduce I/O Contention.* Research paper

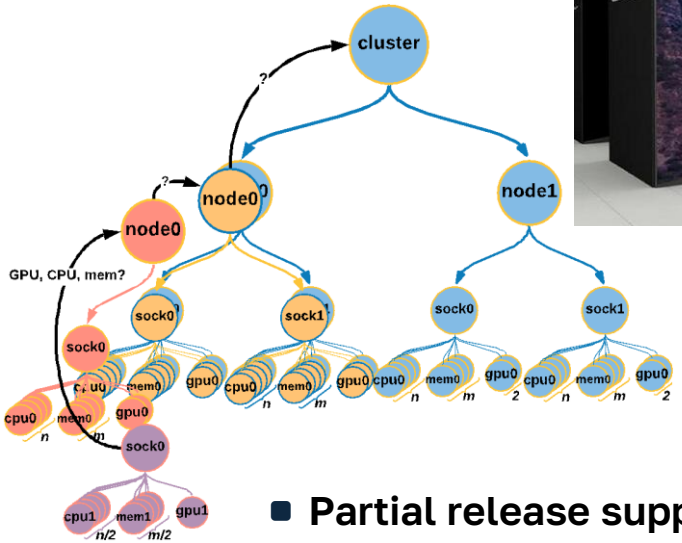
HPC-IODC Workshop. June 13, 5:30pm. *Enhancing Parallel Computing CPU and I/O Performance through Malleable Resource Management.* Invited talk

flux support for elasticity is maturing.

Daniel Milroy (LLNL)

Flux-core (resource management, core services) and **flux-sched** (queuing and scheduling) have different levels of support for elasticity.





- **Partial release supported on El Capitan**
- **Support for growing, shrinking resource graph**
- **Supports growing pre-configured instance size**
- **More in progress on GitHub**

Prepared by LLNL under Contract DE-AC52-07NA27344.

Hans-Christian Hoppe (FZJ)

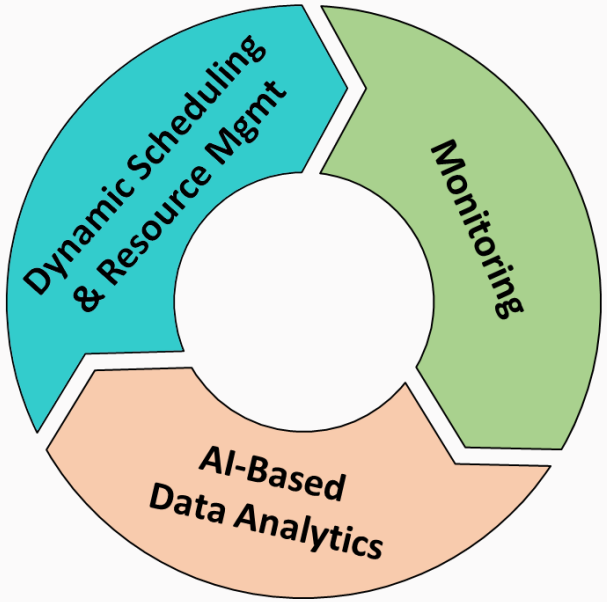
SEANERGYS
ENERGY EFFICIENT EXASCALE

It's time to combine existing approaches to provide a SW suite for optimising HPC/AI center operational energy efficiency

Monitoring
- Scalable system/facility/environment sensor data capture
- Inclusion of non-structured data
- Establishment of a system data plane

AI-based data analytics (secret sauce?)
- Automatic workload characterisation and prediction
- Directing scheduling & resource mgmt. decisions
- Actionable feedback to application developers and users

Dynamic scheduling & resource management
- Dynamic control of system & facility operating points
- Exploit dynamicity/malleability for energy- /power- /carbon-aware scheduling
- Rightsize moldable workload resource requirements



Martin Schulz (TUM)

Implementing all these ideas means ***breaking*** long standing ***dogmas*** in HPC

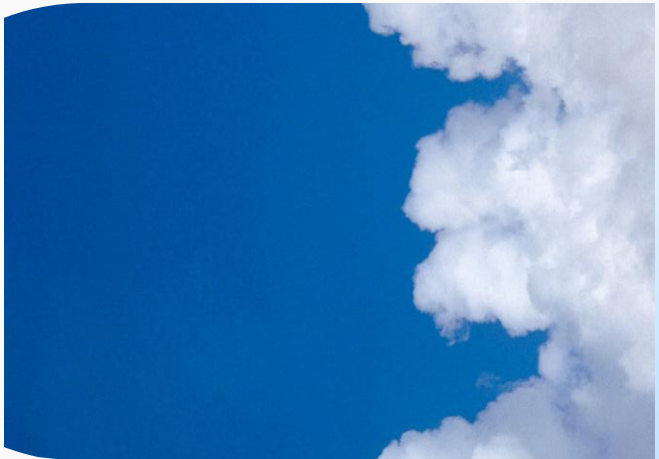
One Node = On Job	**On-Node Co-Scheduling** Requires deep understanding and dynamic tracking of topology	Effective use of complementary accelerators
Jobs have static resource usage from start to end	**Dynamic Process Allocation** Requires new resource management interfaces and management approaches	Adapt to external conditions (e.g., contention)
Worst case and fixed power distribution	**PowerStack / SEANERGYS** Requires continuous monitoring and live feedback & interactions with the grid	Adaptive power steering on over-provisioned nodes

Agenda

00:00 - 00:15
2 minute (strict!) appetizers

00:15 - 00:50
Discussion / questions from the audience

00:50 - 01:00
Final statements



Questions

How feasible is a widespread adoption of these dynamic strategies given existing system constraints?

Questions

How do you get
end-users / application developers
/ owners / administrators
on-board?

Questions

How many savings do you expect with respect to “regular” operational costs?

Questions

What do you see as the
main challenge for research?

Questions

Where do you foresee the next steps?
(Applications, System, Vendors, …)

Questions

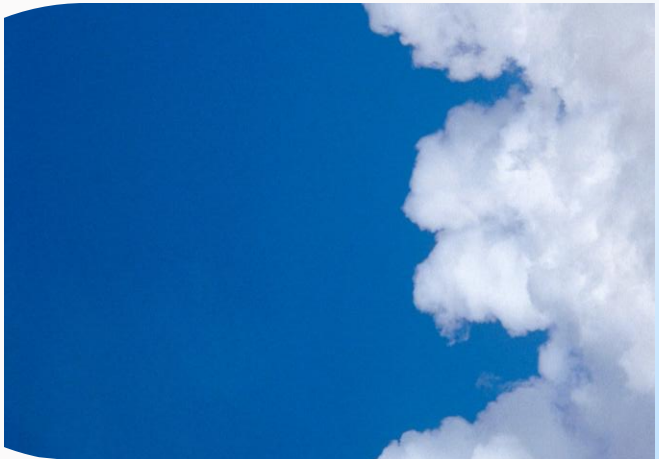
Which time span do you foresee to become operational?

Agenda

00:00 - 00:15
2 minute (strict!) appetizers

00:15 - 00:50
Discussion / questions from the audience

00:50 - 01:00
Final statements



Final statement

2 min for each speaker

Final statement

Thank you for attending - how to get involved?

- MPI Forum:
  MPI Session working group working on dynamic resources since over 5 years

- OpenMP ARB:
  Dynamic resources currently under discussion

- Dynamic Resource Consortium (in general):
  Email to sergio.iserte@bsc.es to join