Optimizing CMS Physics Workflow Construction Across Heterogeneous Worldwide Resources

UNIVERSITY OF NOTRE DAME

Alan Malta Rodrigues
amaltar2@nd.edu

1. Introduction

- High Energy Physics (HEP) scientific workflows often consist of multiple tasksets, each representing a distinct processing step.
- Tasksets may have diverse resource requirements, including CPU architectures, GPU support, memory, and I/O constraints.

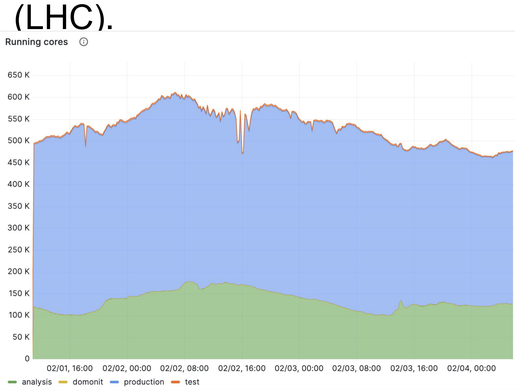
Transforming core-years into days of processing.

2. Research Questions

1. How can we **efficiently group tasksets** to maximize performance in a **heterogeneous computing environment?**
2. What are the **trade-offs between different workflow compositions**, and how do they impact throughput and resource utilization?

3. Background

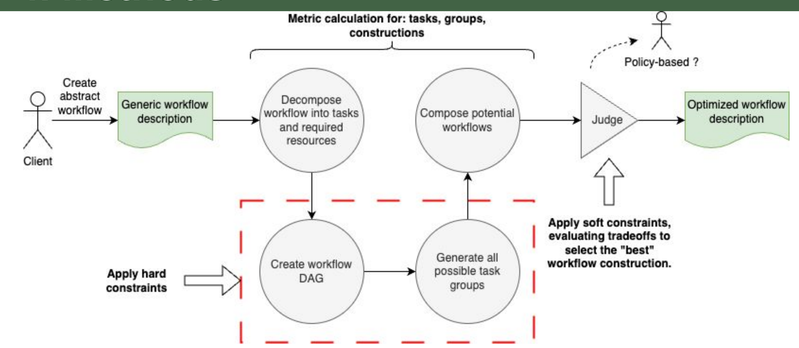
- The Worldwide LHC Computing Grid (WLCG) combines about 1.4 million CPU cores and 1.5 exabytes of storage, supporting global-scale data processing.
- The CMS (Compact Muon Solenoid) experiment at CERN collects data from particle collisions at the Large Hadron Collider (LHC).



CMS peaked at about 0.5 million CPU-cores in parallel.

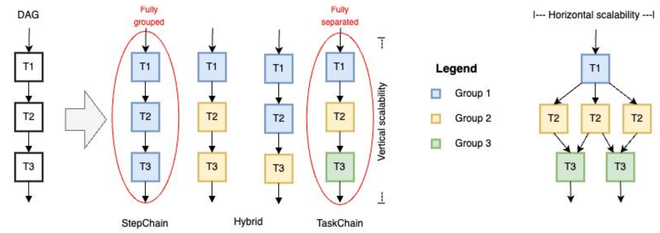
Producing about 1PB of data/week.

4. Methods



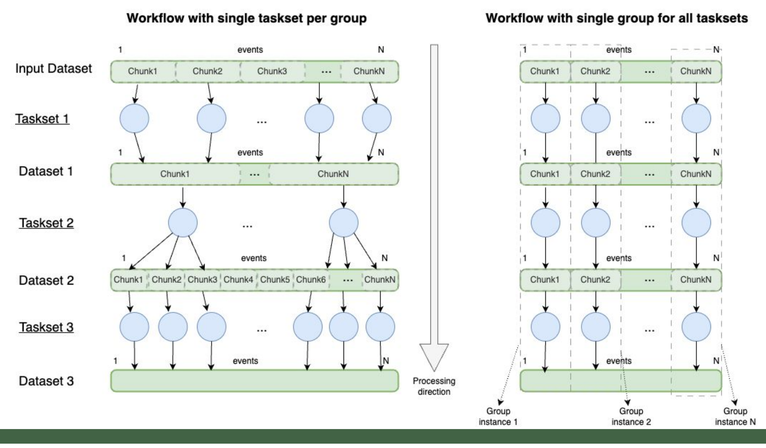
There currently exist two baseline workflow constructions - StepChain and TaskChain - however, between these extremes lies a **spectrum of hybrid workflow constructions**, which enable fine-grained control to prioritize:

- **event throughput;**
- **resource utilization efficiency;**
- **I/O patterns and data locality.**



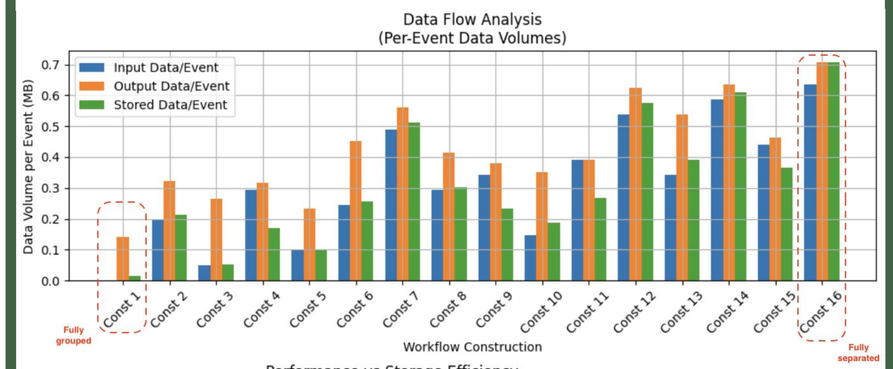
Taskset grouping and workflow composition can **scale** both:

- In **depth**: chaining tasksets into a single job to reduce overhead.
- In **width**: by increasing the chunk of data and events to be processed with a single job



5. Results & Discussion

Sequential DAG with 5 tasksets fully compatible.



6. Conclusions

1. **Flexible workflow composition** has a measurable impact on overall resource utilization and processing efficiency.
2. Enables **effective exploitation of heterogeneous resources**, allowing policy-driven decisions based on taskset characteristics and system constraints
3. Even **small composition improvements yield substantial benefits** at the scale of CMS computing, where millions of jobs are continuously processed.

7. Future Work

- Consider overhead caused by merge jobs (I/O and turnaround time)
- Execute a synthetic benchmark.

8. Acknowledgements

Dr. Douglas Thain
Dr. Kevin Lannon
Cooperative Computing Lab