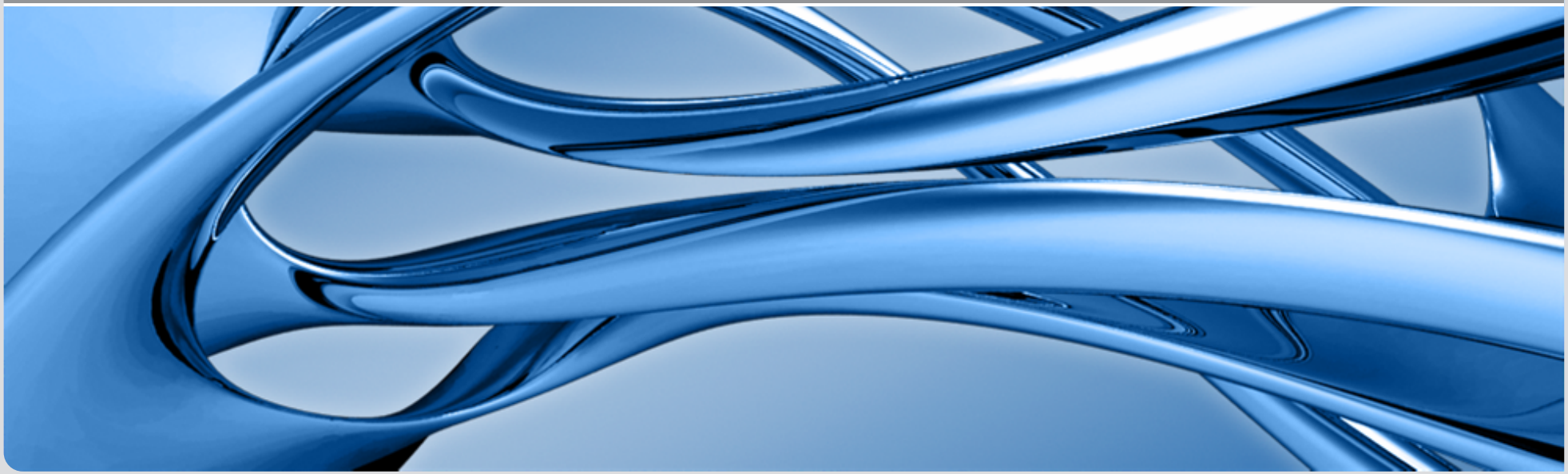
Karlsruhe Institute of Technology

Modeling and Simulation of Load Balancing Strategies for Computing in High Energy Physics

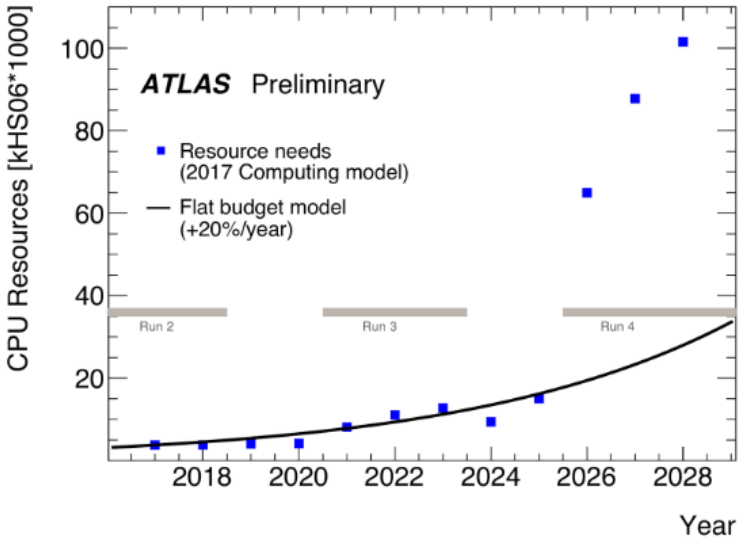
René Caspart[1], **Patrick Firnkes**[2], **Manuel Giffels**[1], **Anne Koziolek**[2], **Günter Quast**[1], **Ralf Reussner**[2], **Maximilian Stemmer-Grabow**[2]

Karlsruhe Institute of Technology (KIT), [1] Institute for Experimental Particle Physics (ETP), [2] Institute for Program Structures and Data Organization (IPD)



Motivation

- Flat computing budget model does not cover future needs for computing resources
- Most efficient usage is mandatory

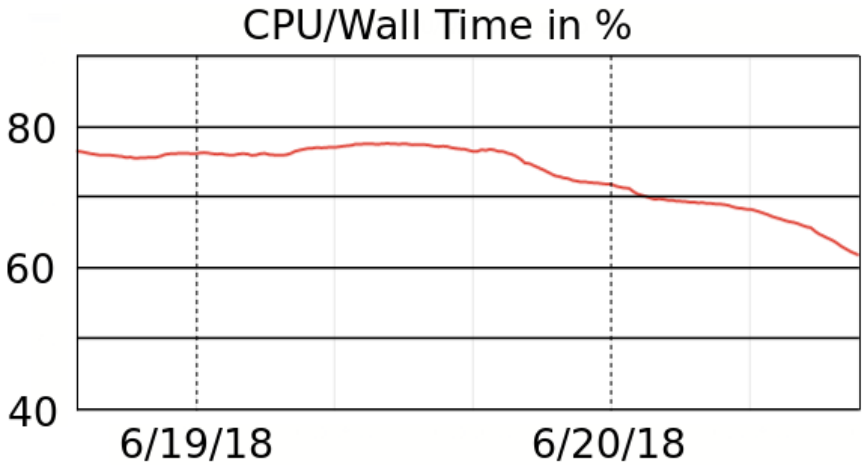


https://arxiv.org/abs/1712.06982

- Computing resources for user jobs in HEP will be increasingly distributed and heterogeneous
  - Addition of cloud resources like HNSciCloud and AWS
  - Institute clusters and Tier 3 centers
- Different job types and mixtures optimal for each resource
- Data placement will be more dynamic
  - Fewer centers hosting data
  - Leads to increase remote access

Scheduling as a key component

- Already for the current resources efficient scheduling for jobs is a challenge
- Even more important for distributed heterogeneous resources in the future
- Need to find ways to simulate the usage efficiency of these resources with different scheduling approaches



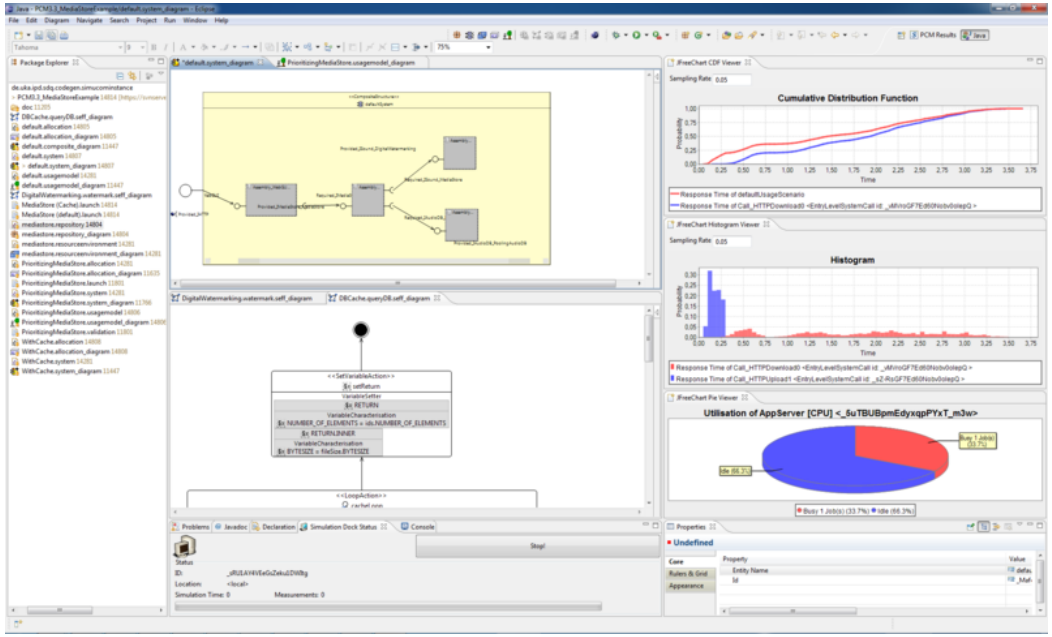
Approach

- Simulate these resources using different scheduling approaches
  - Need to model workflows and resources
  - Need to implement different scheduling approaches
  - Check the effect of these approaches on the usage efficiency
- Make use of the Palladio Simulator[1]
  - Established tool in the computer science community
  - Actively developed and used by computer scientists at KIT
  - Extended to model jobs and available resources at computing centers
- As a proof of concept simulate CMS workflows at the Tier 1 center GridKa

[1] The Palladio component model for model-driven performance prediction
https://doi.org/10.1016/j.jss.2008.03.066

Palladio Simulator

- Performance predictions for design decisions
- Software architecture simulator
- Successfully applied for
  - Solving industrial problems
  - Optimizing cloud infrastructure (chemical computing)[1]



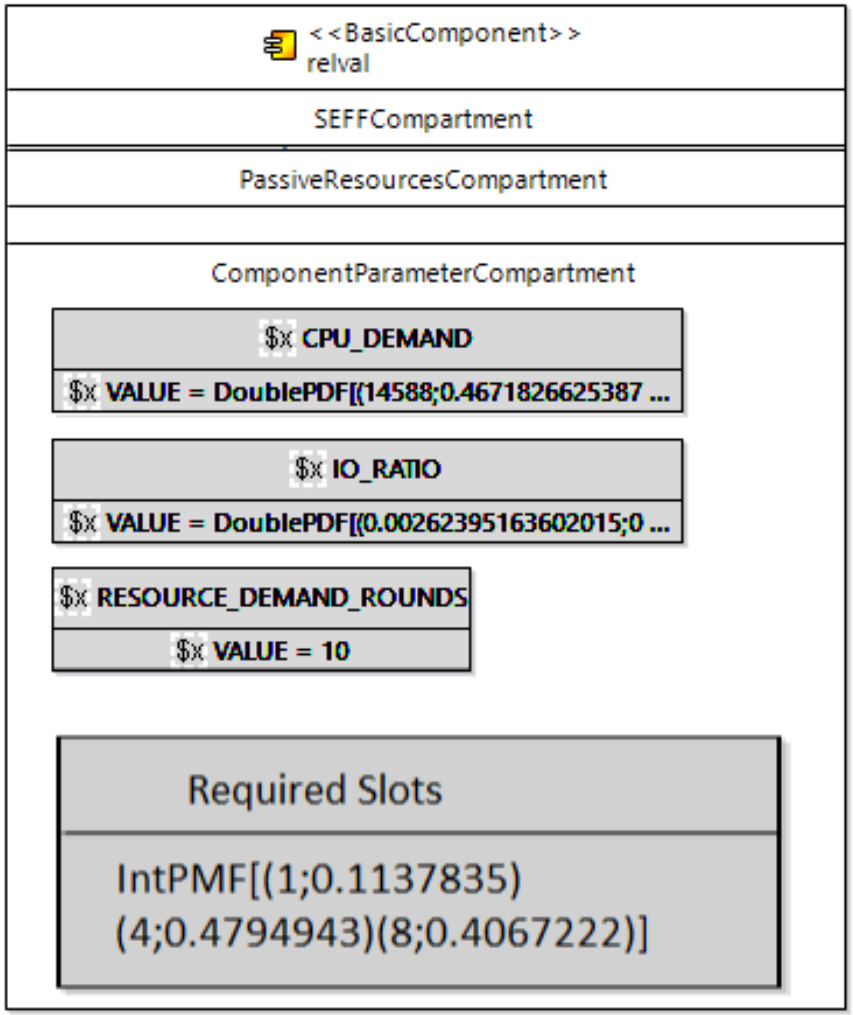
Graphical Interface of the Palladio Eclipse Plugin

- Abstract models: Express resource needs as statistical distributions

[1] Rapid Testing of IaaS Resource Management Algorithms via Cloud Middleware Simulation
https://arxiv.org/pdf/1801.09484.pdf

Model Computing Jobs

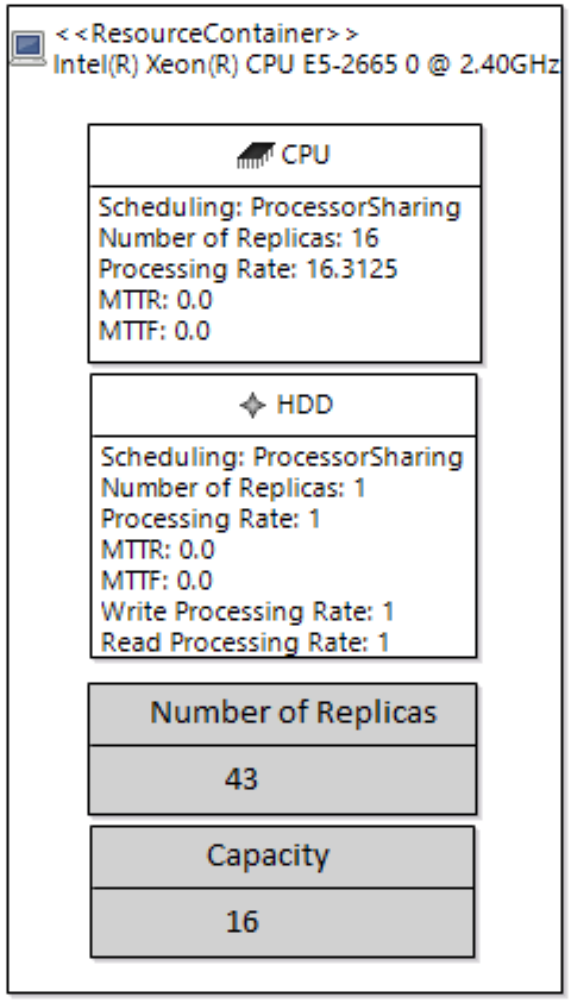
- Model each kind of computing job with its resource usage
  - CPU & I/O
  - Required job slots
  - Number of events
- Model high load on system
  - Closed workload
  - Enough jobs to guarantee that system never idles
  - Each job type has configurable share of load



Screenshot: Computing job component in Palladio

Model Resource Container

- Model each type of computing node
  - Number and processing speed of cores
  - I/O capabilities
  - Number of instances of node
- Model load balancing strategy
  - First fit search based on available job slots
  - Easily modifiable to evaluate new strategies



Screenshot: Resource container in Palladio

Simulating CMS Jobs at GridKa

Karlsruhe Institute of Technology

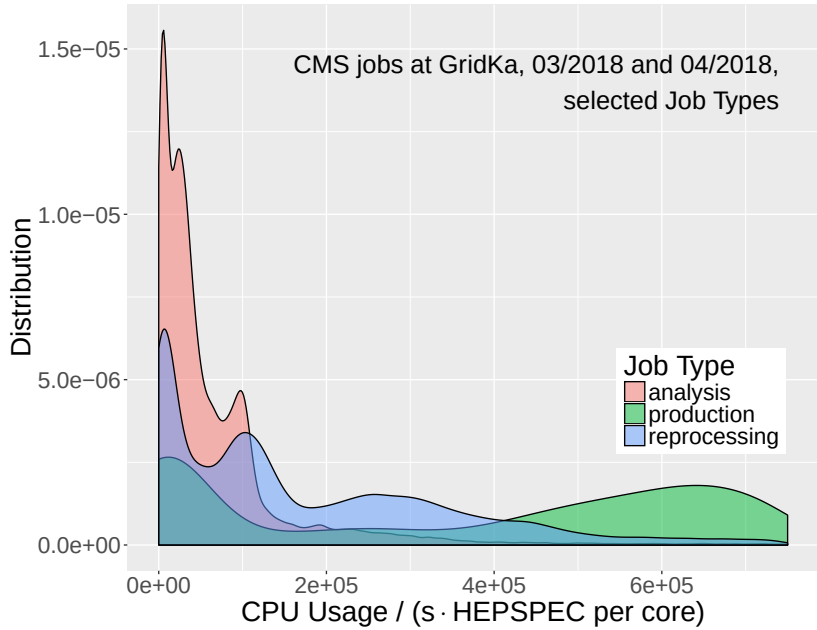
Data Sources:

- Global job monitoring data
  - JobMonitoring, WMArchive job reports

    from Hadoop analytix cluster with CMSSpark framework (Kuznetsov)
  - Currently extracting CMS jobs at GridKa
- Site-specific performance data
  - VO resource share
  - Node benchmarks

Realistic job CPU usage profiles

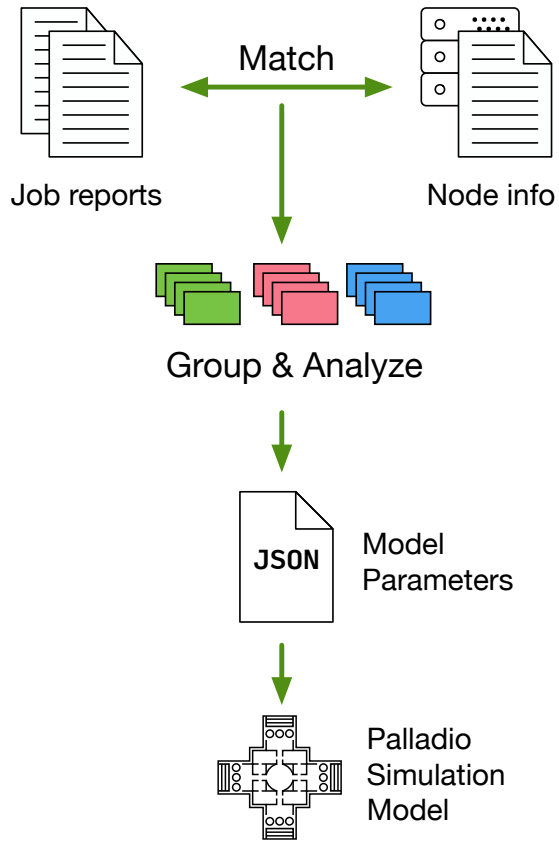
(including failing jobs)



Model Calibration Process

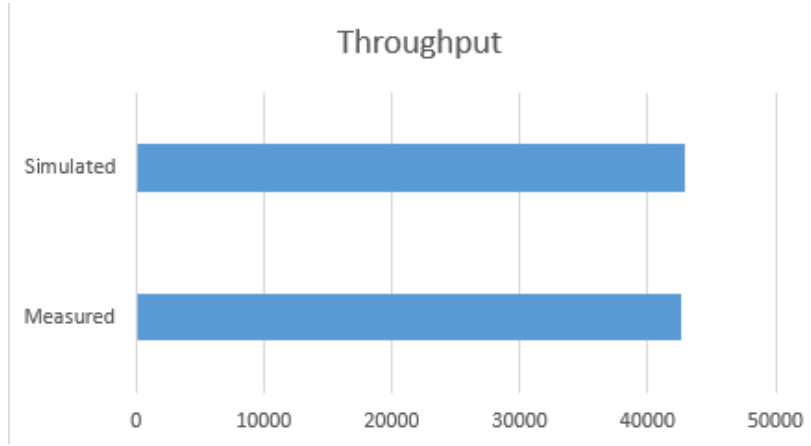
Automated parameter extraction:

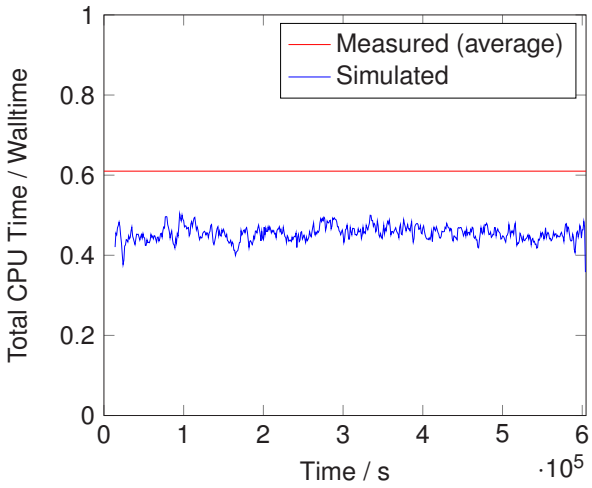
1. Match jobs and node performance information
2. Group computing jobs by type and requirements
3. Extract resource demand distributions and load composition

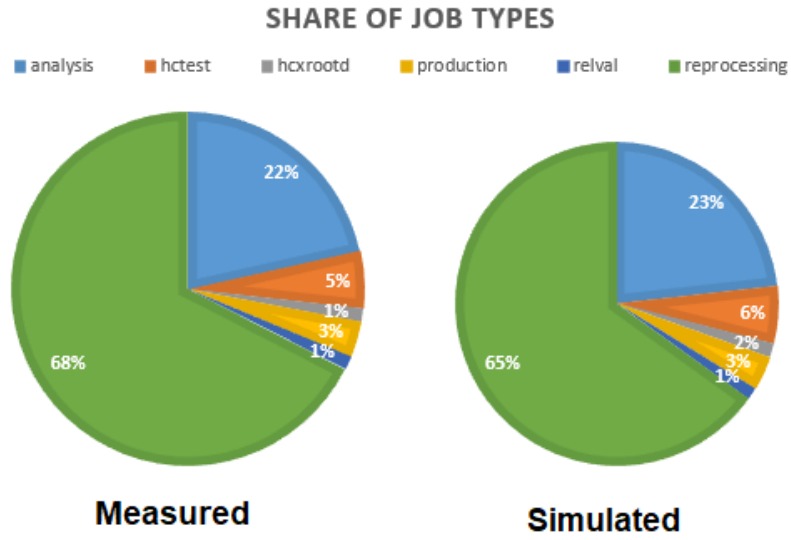


Validation

- Simulation of CMS computing jobs at GridKa for one week
  - Ran in 18 minutes on a laptop
- Metrics
  - Throughput
  - Share of job types
  - CPU efficiency







Conclusion

- Computing resources for user analyses will become increasingly distributed and heterogeneous in the future
- Scheduling will be a key component for the efficient usage
- In cooperation with computer scientists at KIT use the tool Palladio to model and evaluate different scheduling strategies
- A proof of concept for the usage of Palladio was performed simulating CMS workflows at the Tier 1 center GridKa
  - Fully automated model creation based on data from CMS workflow monitoring and GridKa
  - We were able to successfully model the situation at GridKa for key metrics
  - Investigate modeling for further metrics

Outlook

- Extend the current model to be able to simulate heterogeneous and distributed computing resources
- Include modeling of challenges due to remote data access and caching mechanism
- Use the result of these models as an input for the design of future scheduling systems
- Optimize scheduling decisions using modeling results at run-time