Modelling workflows in the Cloud

Gerhard Rzehorz
Supervised by G. Kawamura, O. Keeble, A. Quadt
II. Physikalisches Institut, Georg-August-Universität Göttingen & CERN

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

CERN

Physics at the Terascale 2017, DESY
28.11.2017

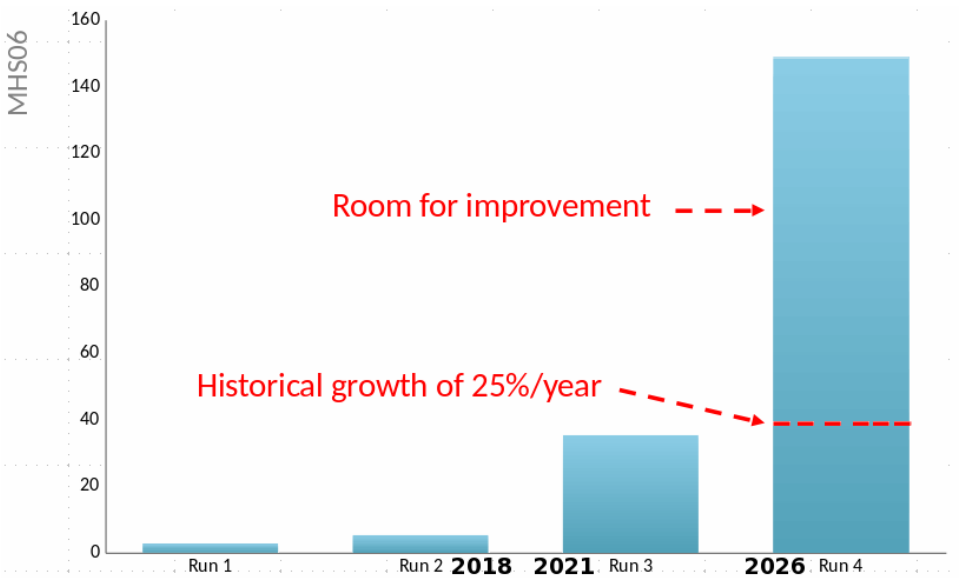
BMBF-Forschungsschwerpunkt
ATLAS-EXPERIMENT
Physik bei höchsten Energien mit dem ATLAS-Experiment am LHC

FSP 103
ATLAS

SPONSORED BY THE
Federal Ministry
of Education
and Research

Future Resource Needs

- Dependent on LHC performance (live-time), luminosity and pile-up
- 2016 data taking was already above expectations
- Run 3: manageable with technological evolution



[M. Schulz, Physics at the Terascale, Nov '16, DESY, slightly simplified]

- HL-LHC: CPU requirements ~ 60 times higher than '16
- Factor of ~10 considering steady technological growth of 20% per year
- Infrastructure improvement: Clouds

➢ Use Cloud resources in WLCG

Cloud Computing

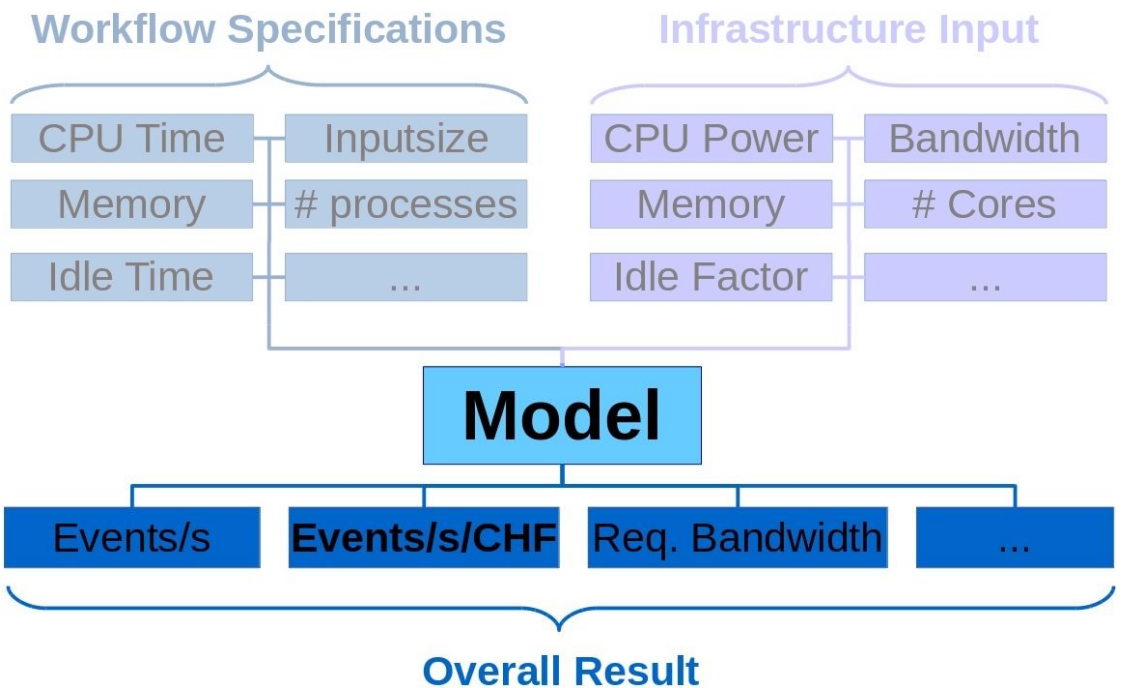
- IaaS from commercial provider, “renting” resources
- Workflows: Analysis too unpredictable → Evt Gen/MC Sim, Reconstruction (data intensive)
- Data intensive ≠ using storage
- Experience: Costly to set up storage (for short time scales)
- Cache-only site?
- “Trade” storage for network?

Cloud Computing

- IaaS from commercial provider, “renting” resources
- Workflows: Analysis too unpredictable → Evt Gen/MC Sim, Reconstruction (data intensive)
- Data intensive ≠ using storage
- Experience: Costly to set up storage (for short time scales)
- Cache-only site?
- “Trade” storage for network?
- Advantages: flexibility, (cost?)
- Unclear: Workflow performance, benefit in adapting infrastructure to workflows, procurement (what to ask for), less personpower intensive?

The Model - Concept

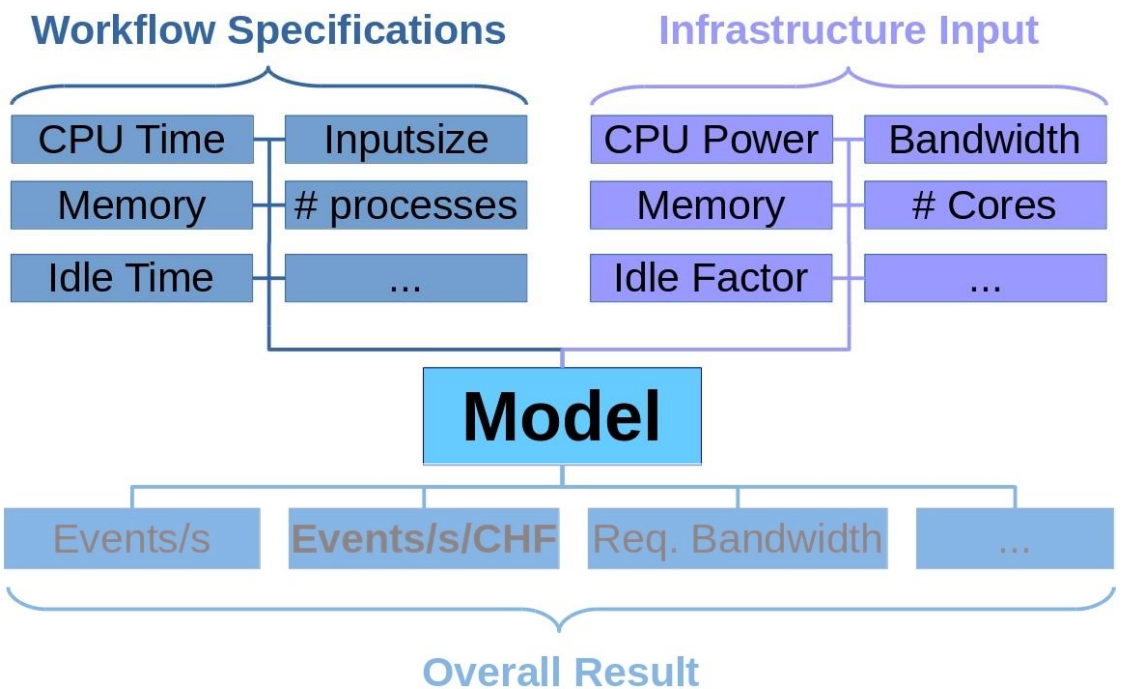
- Simple Model: linear combination
- Infrastructure inputs based on benchmarks



- Generic: outside physics
- Correlations: e.g. CPU-power impact required bandwidth
- Evaluation: find inefficiencies
- Configuration: SSD? Faster CPU? 4- or 8-core?
- Result: combined (e.g. Events s^{-1} CHF^{-1}, “physics“ per time and money) or infrastructure metric (e.g. bandwidth)
- Assessment of Clouds

The Model - Input

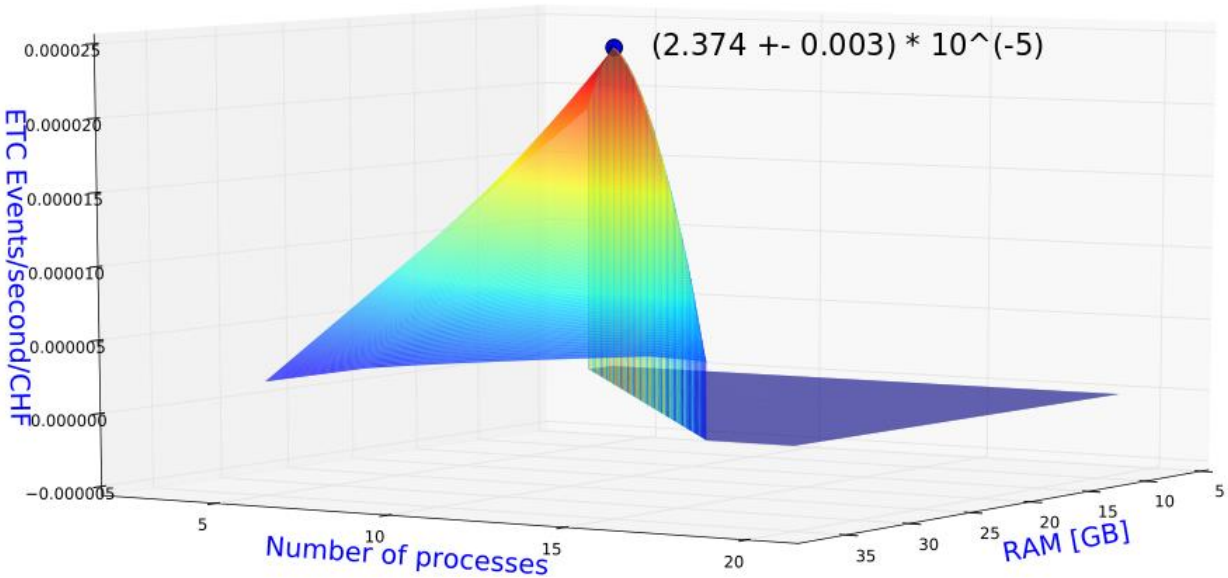
- Infrastructure as well as workflow parameters needed
- Workflow specifics obtained from anywhere (Grid)



- Infrastructure inputs during Cloud procurement phase
- With access to Cloud: Run (ATLAS) benchmark job
- Without access to Cloud: Benchmark suite (tendering phase) provides input
- Classify jobs

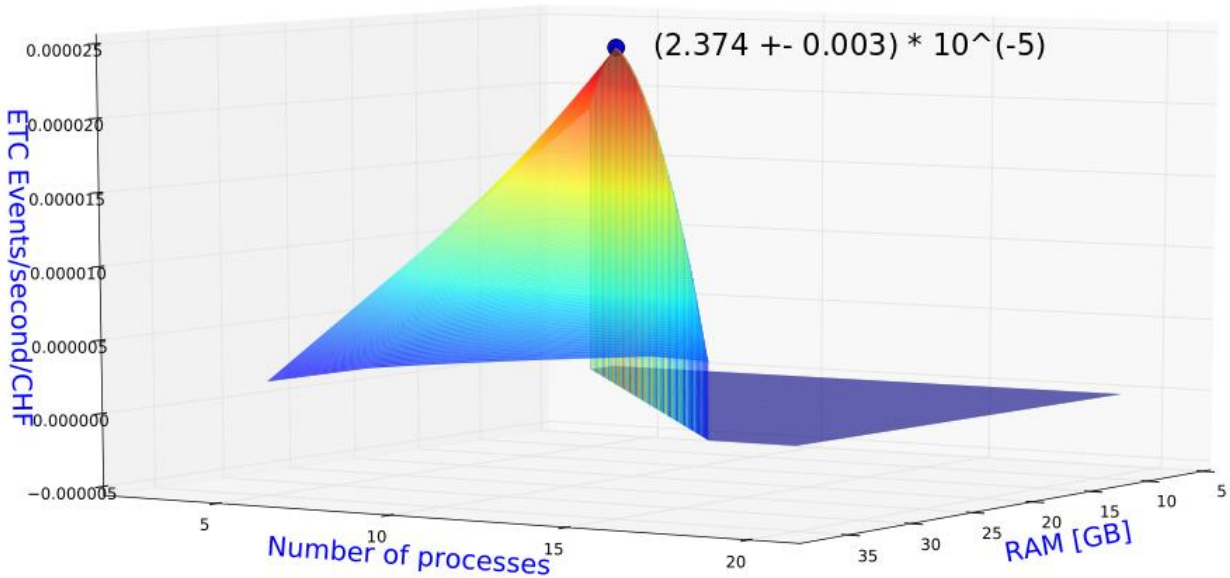
The Model - Example

- Investigate overcommitting
- RAW reconstruction
- Fixed budget
- Example: few VMs (cost known)
- Vary inputs
- “Trade” RAM for more CPU



The Model - Example

- Investigate overcommitting
- RAW reconstruction
- Fixed budget
- Example: few VMs (cost known)
- Vary inputs
- “Trade” RAM for more CPU
- Result: 1000 chf in 10000 s → (23740 ± 30) events reconstructed
- Process/RAM position of maximum - best configuration
- Maximum ETC value to compare different providers
- Result applicable to Grid (even with fixed RAM)



The Workflows - Behaviour

- Validation: cover all possible workflow and infrastructure aspects:
  - Event generation, MC Sim, Reconstruction, Digitisation
  - Different CPU/Disk/Network/Memory types/speeds
- Reference + target VM: Model target, compare to measurement
- VMs different dedicated hosts (no influencing neighbours)
- Variation of the workflow

EvGen	target
difference Wall Time %	**0,49**
difference CPU Time %	-0,02

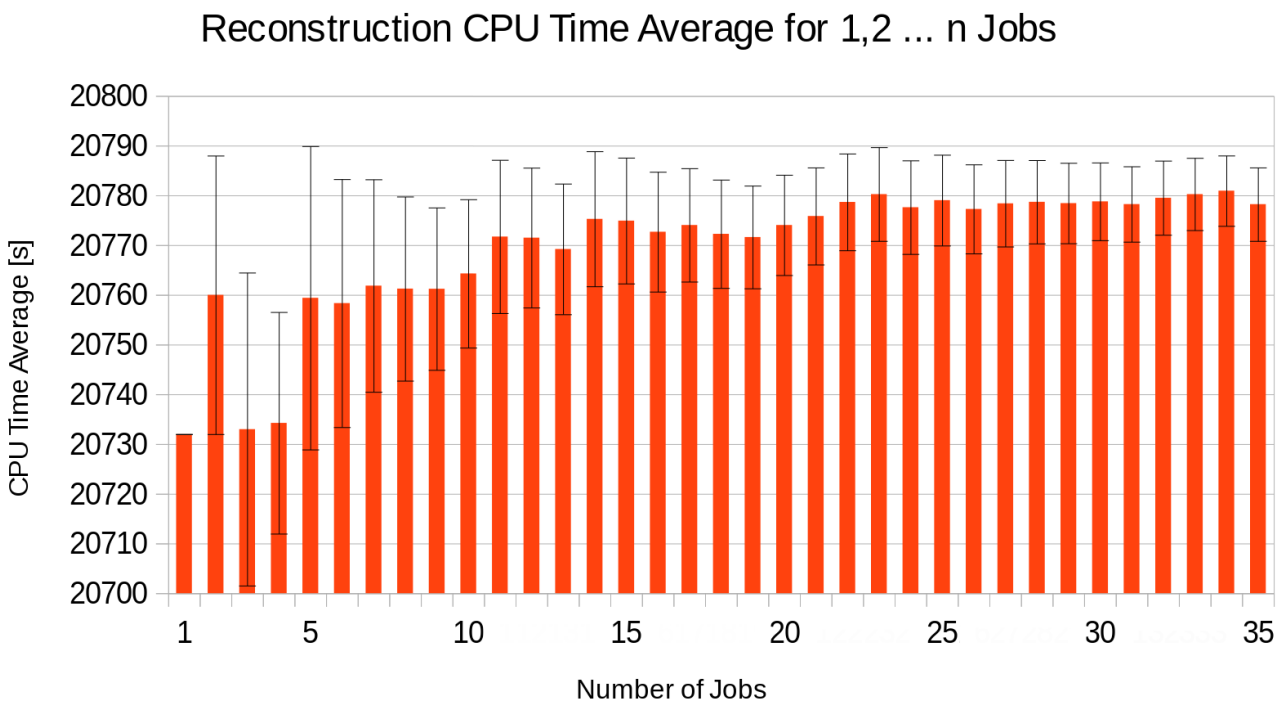
Reconstruction	target
difference Wall Time %	**-0,28**
difference CPU Time %	0,91

MC Sim	target
difference Wall Time %	**2,68**
difference CPU Time %	1,91

- Result from 25 measurements
- Good agreement

The Workflow – Behaviour

- How big fluctuations?
- Convergence?
- How many benchmarks for “good” input?
- Note: y-Axis does not start at zero; Error = StdDev / sqrt(n)



The Model - Uncertainties

- Compare Model prediction to measurement
- 5 VMs on 3 different Cloud providers (HNSciCloud prototypes)

HELIX NEBULA
THESCIENCECLOUD

Helix Nebula Science Cloud Joint Pre-Commercial Procurement

Procurers: CERN, CNRS, DESY, EMBL-EBI, ESRF, IFAE, INFN, KIT, STFC, SURFSara
Experts: Trust-IT & EGI.eu

The group of procurers have committed
- Procurement funds
- Manpower for testing/evaluation
- Use-cases with applications & data
- In-house IT resources

Resulting services will be made available to end-users from many research communities

Co-funded via H2020 Grant Agreement 687614

Total procurement budget >5.3M€

The Model - Uncertainties

- Compare Model prediction to measurement
- 5 VMs on 3 different Cloud providers (HNSciCloud prototypes)
- Model: provide error estimation for every result
- Use standard deviation of benchmark results, error propagation to final result

- Model error prediction

	Reco 1 Wall diff %	Reco 2 Wall diff %	Reco 3 Wall diff %
IBM	2,06	4,41	0,38
TSY	2,60	0,32	1,95
Exoscale 1	1,22	0,60	1,57
Exoscale 2	0,97	0,93	10,76
Exoscale 3	0,80	0,92	

Measurement

	Reco 1 Wall diff %	**Reco 2 Wall diff %**	**Reco 3 Wall diff %**
IBM	1,11	-3,74	0,48
TSY	-0,80	1,64	-1,53
Exoscale 1	-0,91	0,97	-0,52
Exoscale 2	-0,77	0,77	2,40
Exoscale 3	-0,42	0,78	

The Model - Uncertainties

- Similar jobs (same task, SW stack, merging)

vs

Different jobs (SW stack, merging)

	Reco 1 Wall diff %	Reco 2 Wall diff %
IBM	1,11	-3,74
TSY	-0,80	1,64
Exoscale 1	-0,91	0,97
Exoscale 2	-0,77	0,77
Exoscale 3	-0,42	0,78

Mix factors	1 with 2 Wall diff [%]
IBM	-23,22
TSY	6,60
Exoscale 1	-13,59
Exoscale 2	-10,88
Exoscale 3	-8,13

- Not comparing “apples with oranges”: Categorise jobs
- Careful also with: Number of events (overheads), Number of Cores

Conclusion

- Future resource deficit → Cloud possible relief
- Infrastructure adaptations to workflows (e.g. bandwidth vs storage, overcommitting plus RAM, reco/evgen VMs/sites)
- Model compares sites, finds bottlenecks and optimal configurations
- Model indicates correlations and impact between parameters, e.g. CPU speed on required bandwidth
- Model quantifies Cloud benefits
- Prerequisite: Carefully classify workflows

Work sponsored by the Wolfgang Gentner Programme of the Federal Ministry of Education and Research.

SPONSORED BY THE
Federal Ministry of Education and Research

Backup

The Model

- The Workflow and Infrastructure Model solves the previous and following questions:
  - Evaluation of workflow behaviour on infrastructure: inefficiencies? bottlenecks?
  - Comparison of different configurations: SSDs? faster CPUs? 4- vs 8-core? only Simul?
  - Discovery of adaptations and optimisations: overcommitting with additional RAM?
  - Assessment of workflow requirements: bandwidth? storage?
  - (Cloud) site comparison

The Model

- Plethora of input parameters ⟶ graspable output for different scenarios
- Vary metrics against each other
- Find min/max of desired output value
- Highest level: site (Cloud) comparison
- Simple: less accurate, but not all Cloud aspects known

The Workflow

- Processing in controlled environment
- ATLAS RAW data reco: combination of transformations
- Split transformations - too complex