Data intensive ATLAS workflows in the Cloud

Gerhard Rzehorz
Supervised by G. Kawamura, O. Keeble, A. Quadt
II. Institute of Physics, Georg-August-University Göttingen & CERN

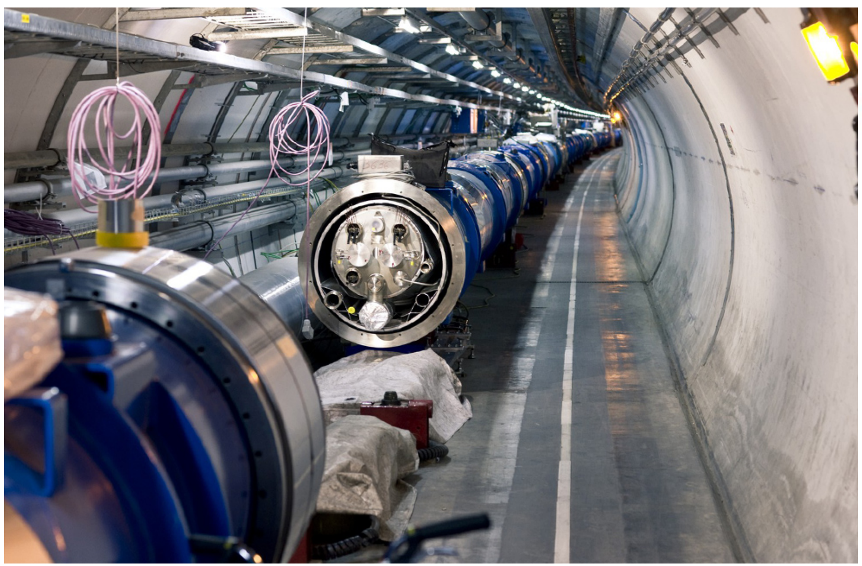
ISGC 2018, Taipei
23.03.2018

SPONSORED BY THE
Federal Ministry of Education and Research

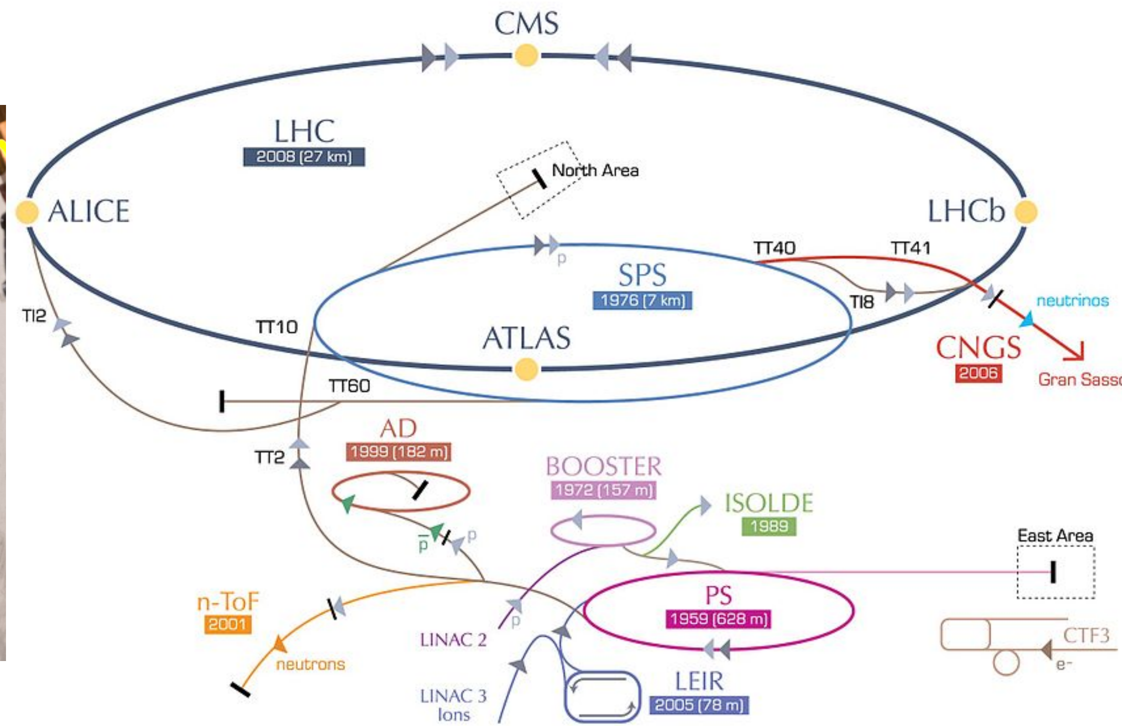
Large Hadron Collider (LHC)

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- CERN: international research organization, physics experiments
- LHC, accelerator ~ 27 km circumference, 50-175 m underground
- $>10^{11}$ protons per bunch, 40 MHz, 2017: 5 million billion collisions
- Events independent - parallelisable



From CERN webpages

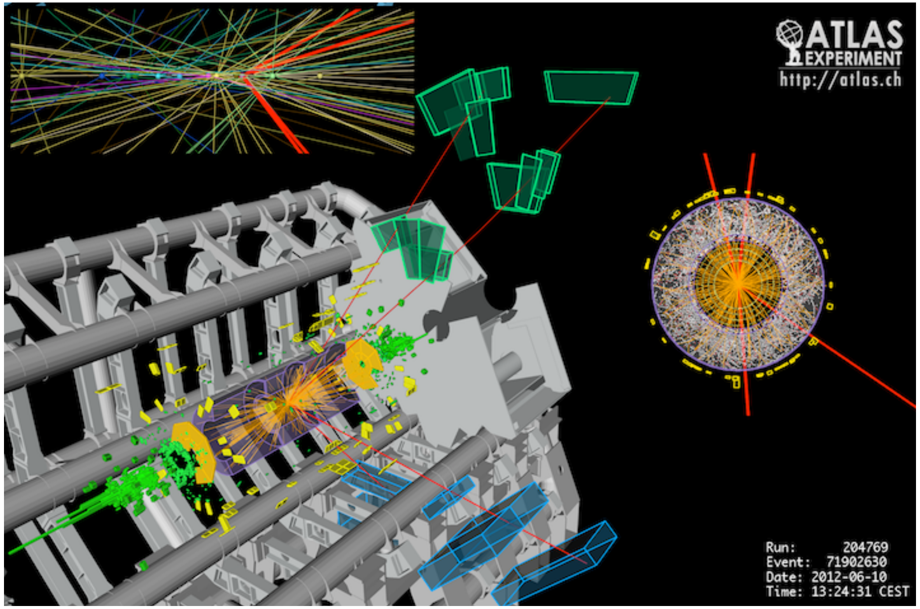


From CERN webpages

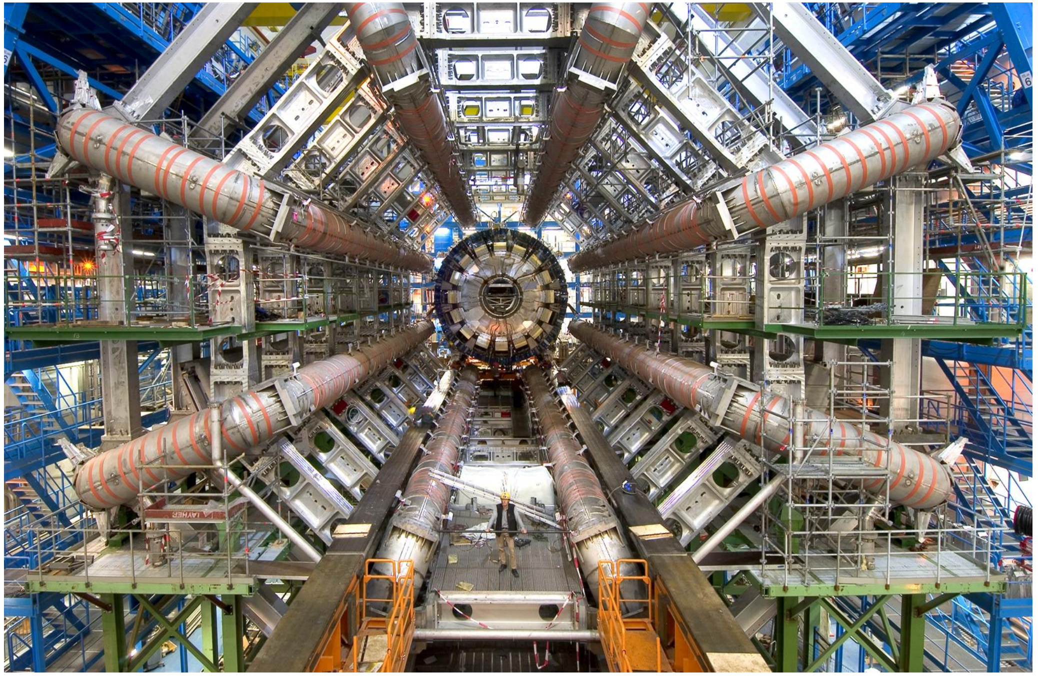
ATLAS experiment

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- A Toroidal LHC ApparatuS (ATLAS) particle detector
- 46m long, 25m diameter, 7000t heavy
- Collaboration: > 3000 scientists, ~182 institutions, Higgs Boson
- Trigger, processes + store events, data available everywhere
- October ‘17 ~ 12.3 PB data



From CERN webpages

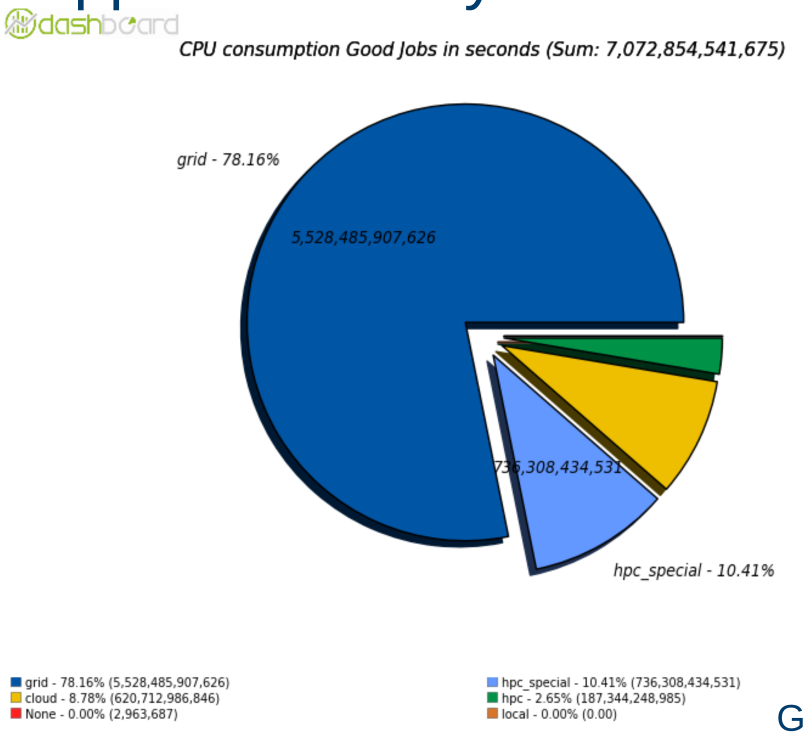


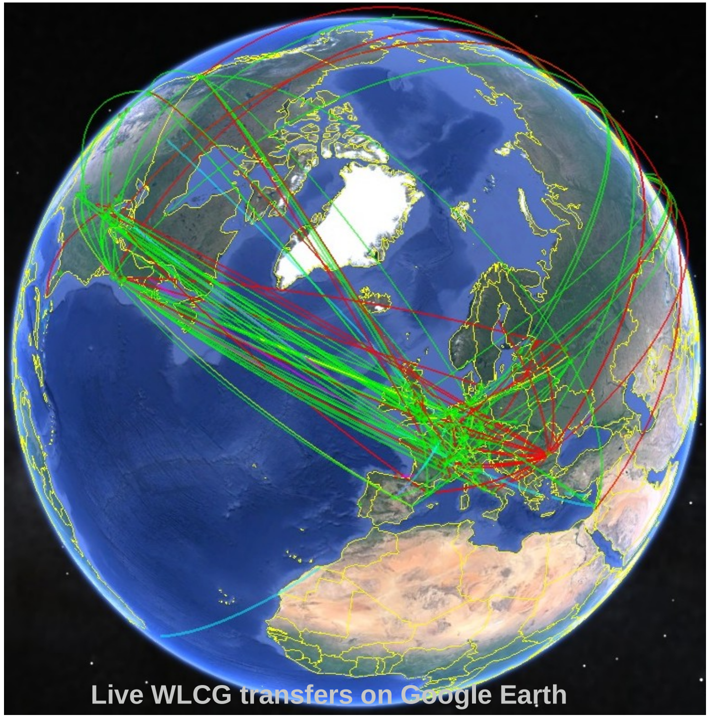
From CERN webpages

WLCG

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

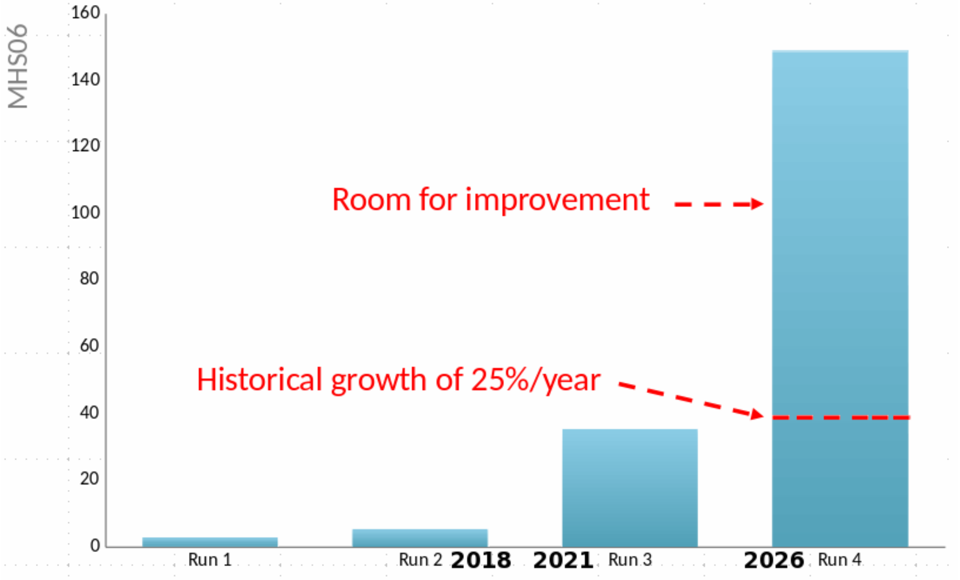
- Worldwide LHC Computing Grid (WLCG), tiered structure
- 42 countries, 170 computing centres, 2 million tasks per day, 750k computer cores, static resources, pledged, high availability
- 400 PB physics data on disk and 400 PB on tape
- Opportunistically used resources: HPC, voluntary computing, Cloud





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]

- Run 4/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
- Data intensive workflows ≠ 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
- Data intensive workflows ≠ 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?

HNSciCloud

- Joint pre-commercial procurement
- Procurers: CERN, CNRS, DESY, EMBL-EBI,
- ESRF, IFAE, INFN, KIT, STFC, SURFSara
- Procurers committed: funds, manpower, use-cases
- Total budget > 5 M€
- Prototype Phase: 3 consortia

HELIX NEBULA THESCIENCECLOUD

Consortium 1: T-Systems, Huawei, Cyfronet, Divia

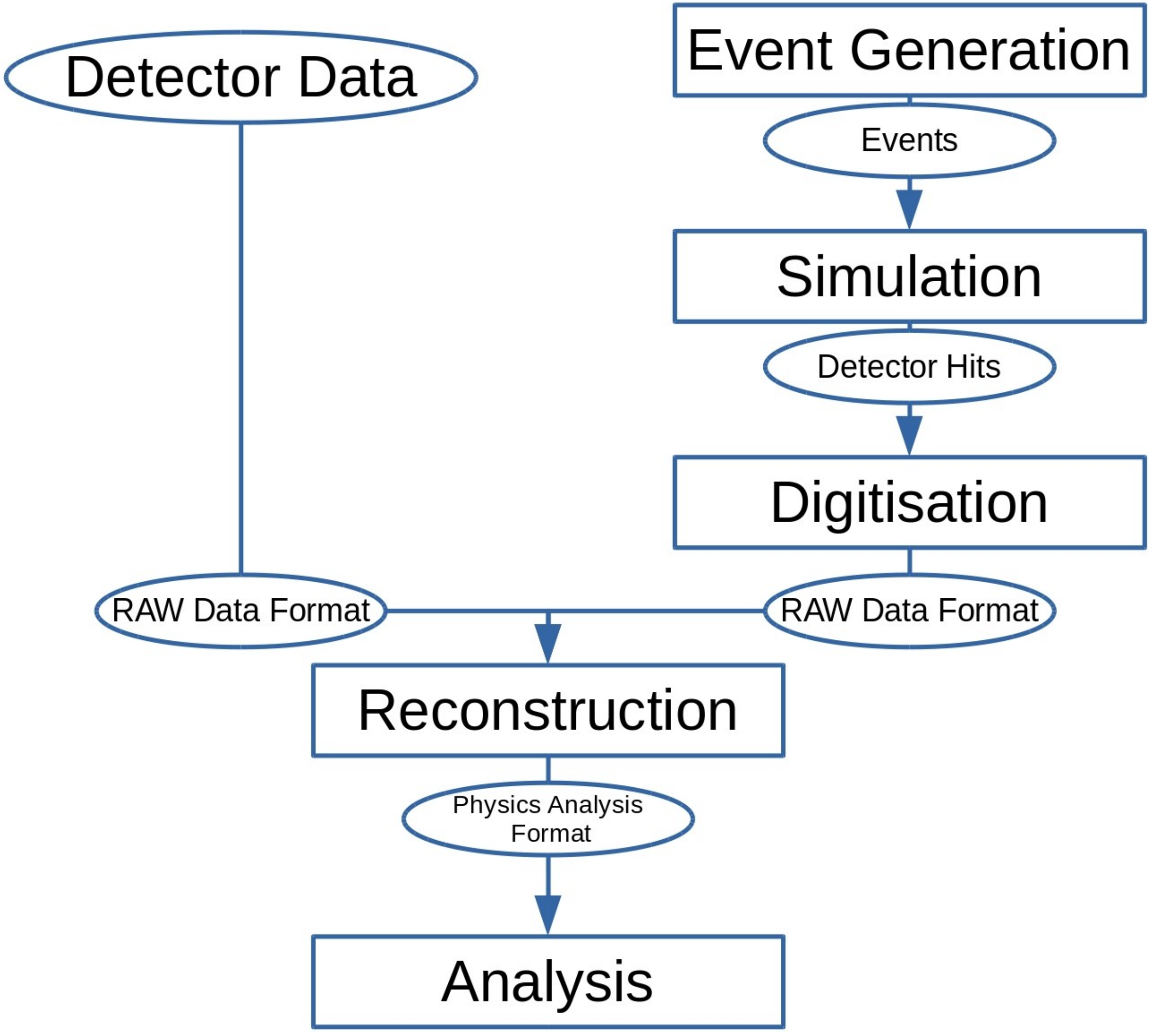
Consortium 2: IBM

Consortium 3: RHEA Group, SixSq, Exoscale, T-Systems

- Tests on Exoscale, IBM and T-systems infrastructure

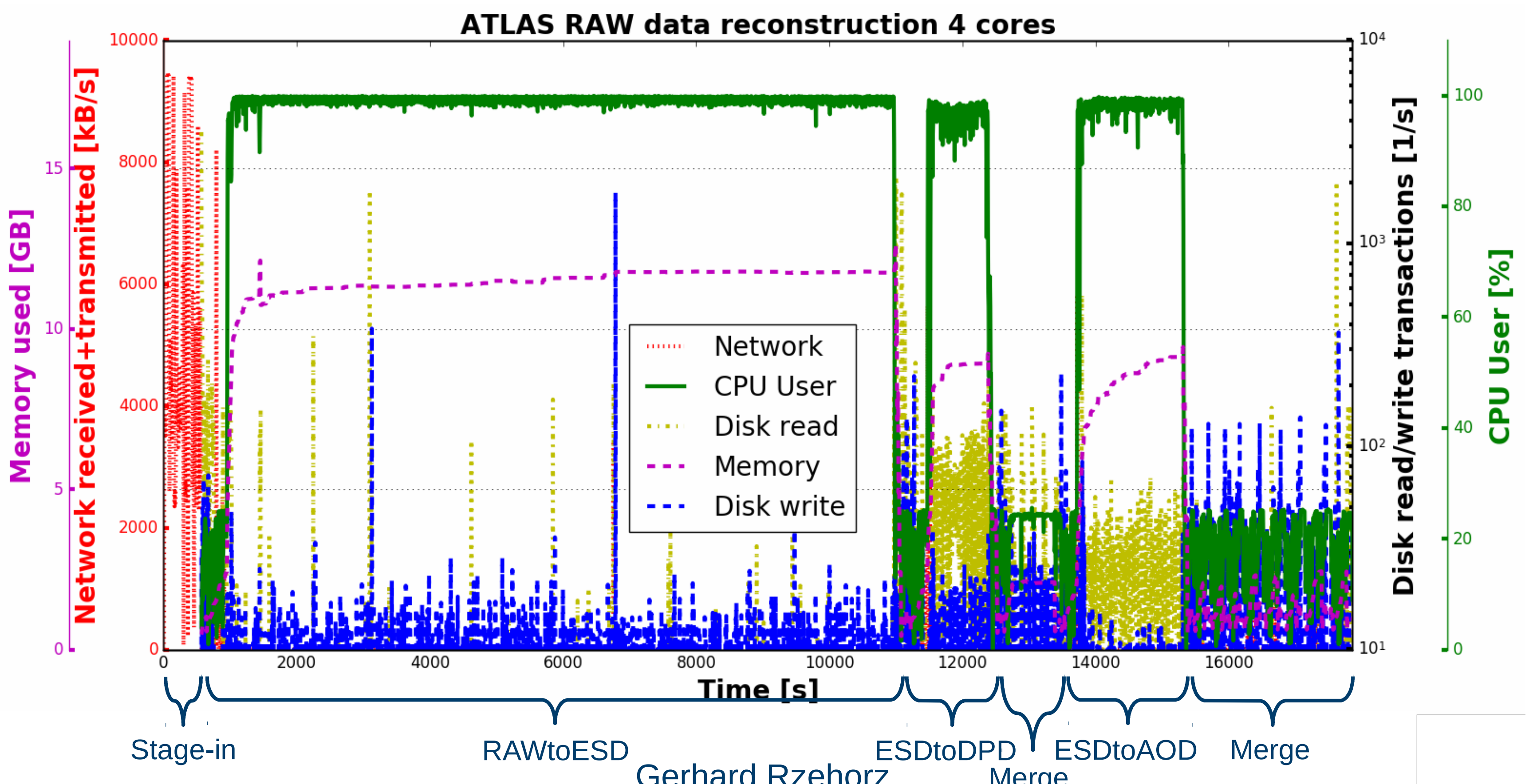
Workflows

- Event Generation (single-core, CPU intensive), Monte-Carlo simulation (CPU intensive), Reconstruction (data intensive), Analysis (data intensive)
- Analysis runs user code, unpredictable



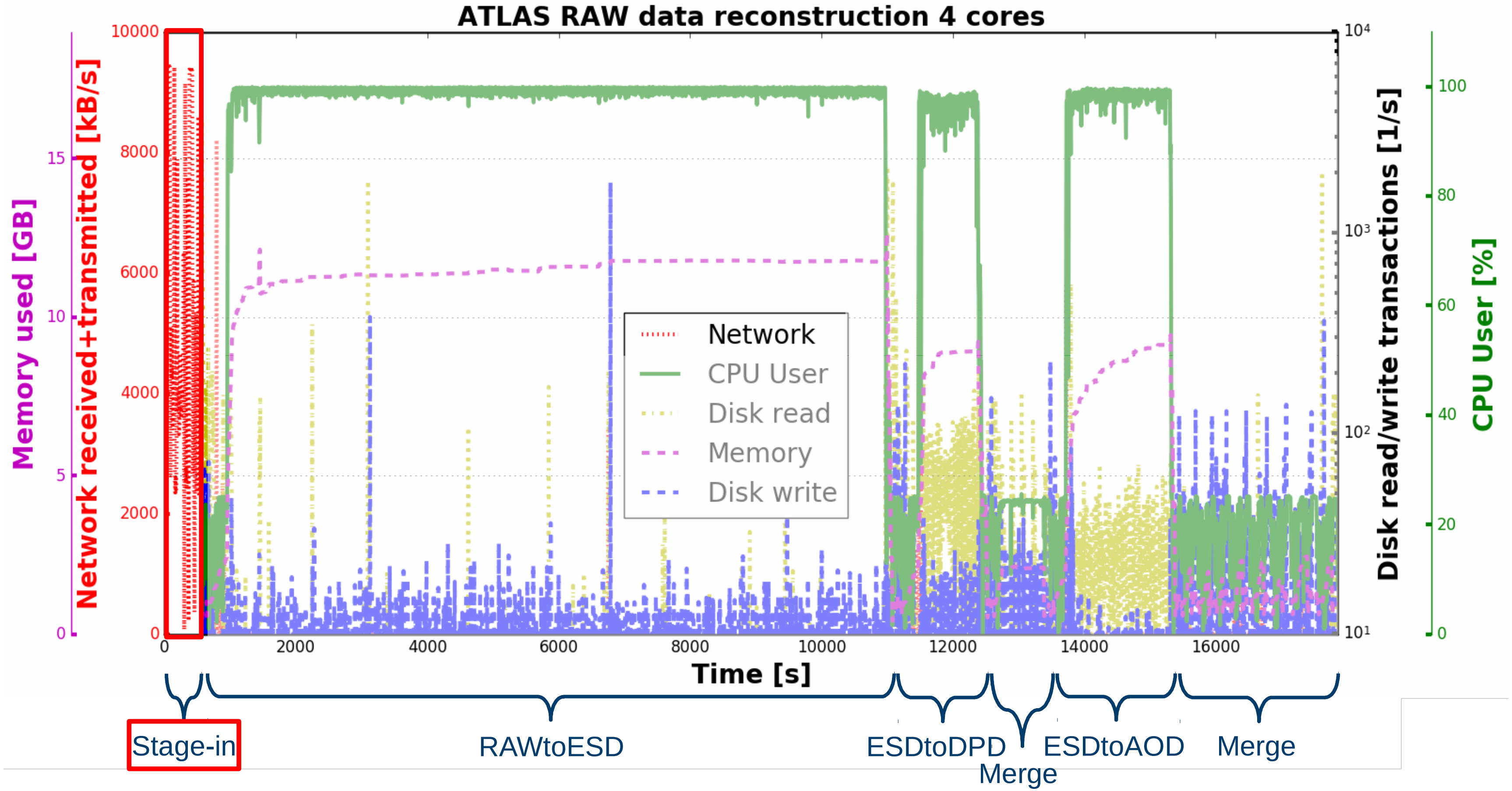
Workflows - Reconstruction

- Reco on private VM, profiling with 'sar' (sysstat), xrdcp from EOS
- ATLAS RAW data reco: combination of transformations



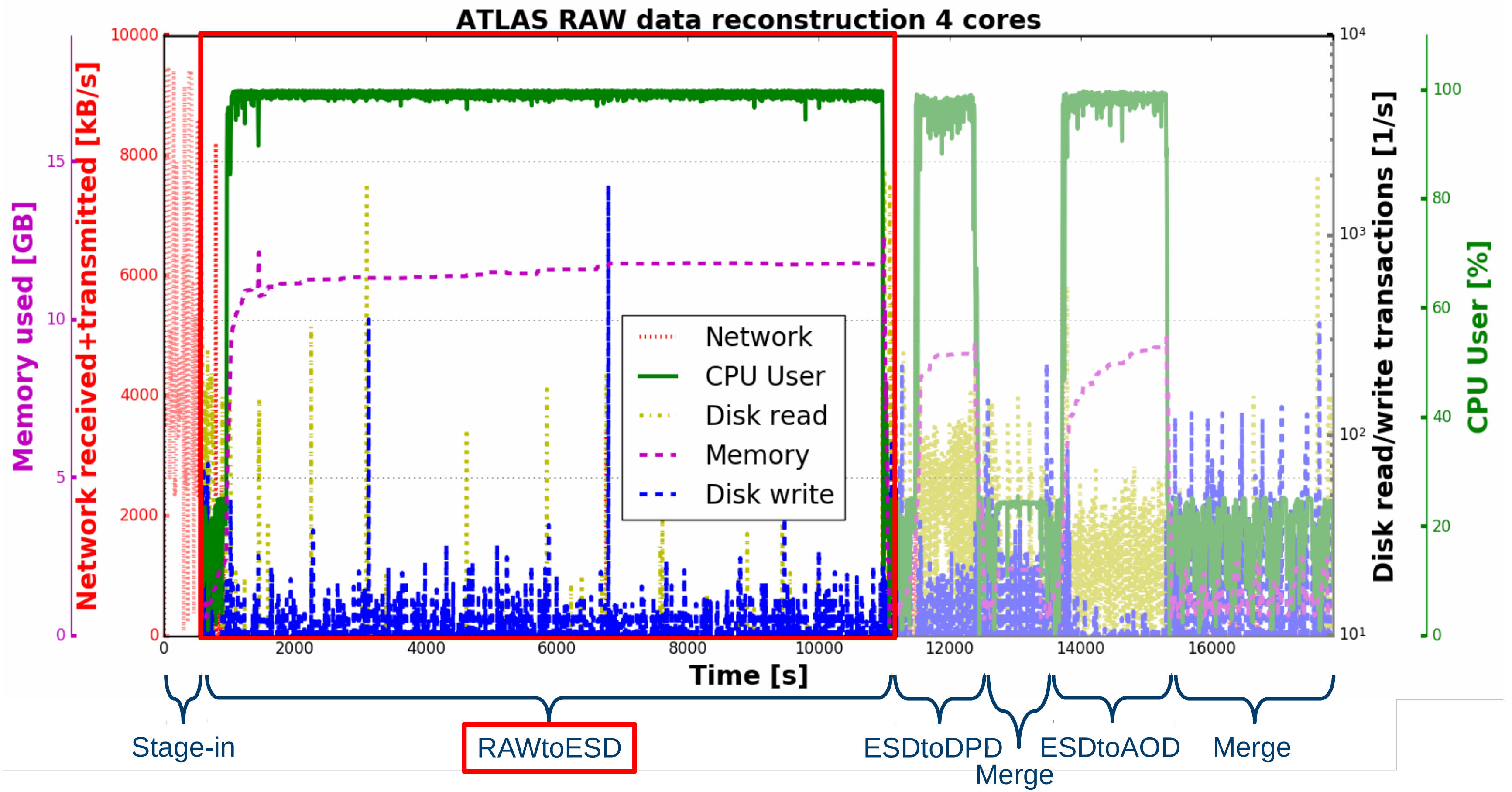
Workflows - Reconstruction

- Reco on private VM, profiling with 'sar' (sysstat), xrdcp from EOS
- ATLAS RAW data reco: combination of transformations



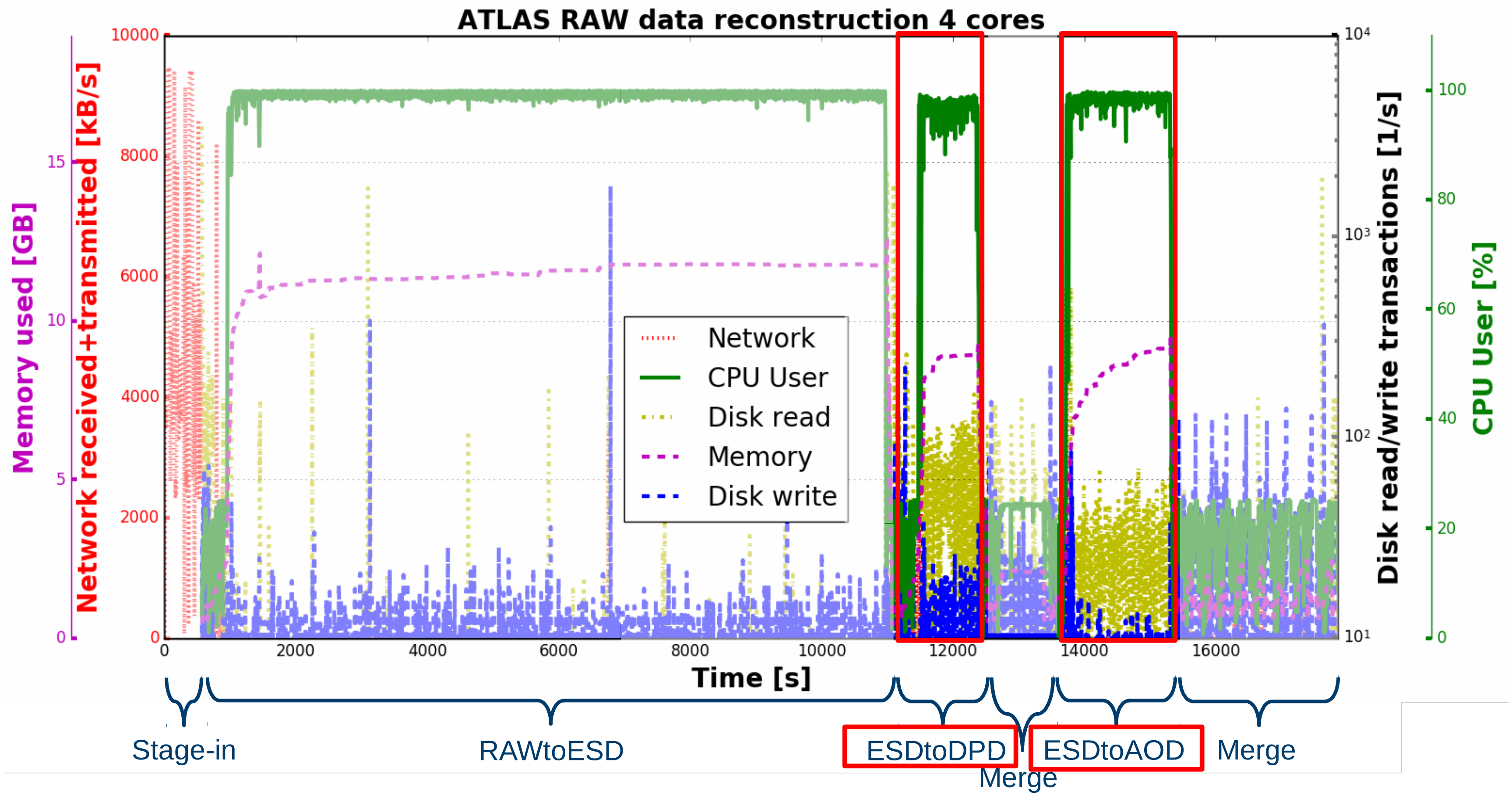
Workflows - Reconstruction

- Reco on private VM, profiling with 'sar' (sysstat), xrdcp from EOS
- ATLAS RAW data reco: combination of transformations



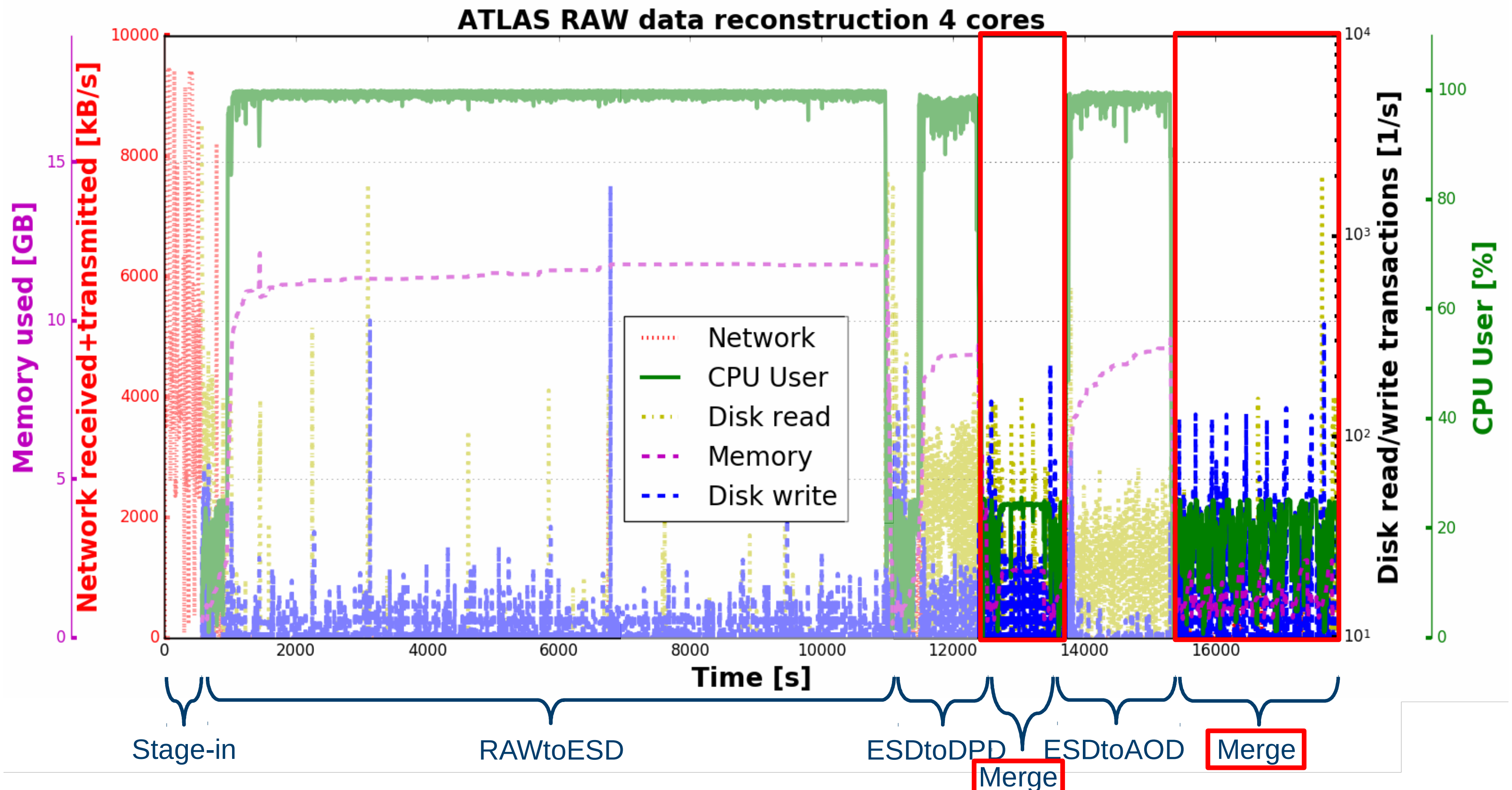
Workflows - Reconstruction

- Reco on private VM, profiling with 'sar' (sysstat), xrdcp from EOS
- ATLAS RAW data reco: combination of transformations



Workflows - Reconstruction

- Reco on private VM, profiling with 'sar' (sysstat), xrdcp from EOS
- ATLAS RAW data reco: combination of transformations



Workflows – Fluctuations

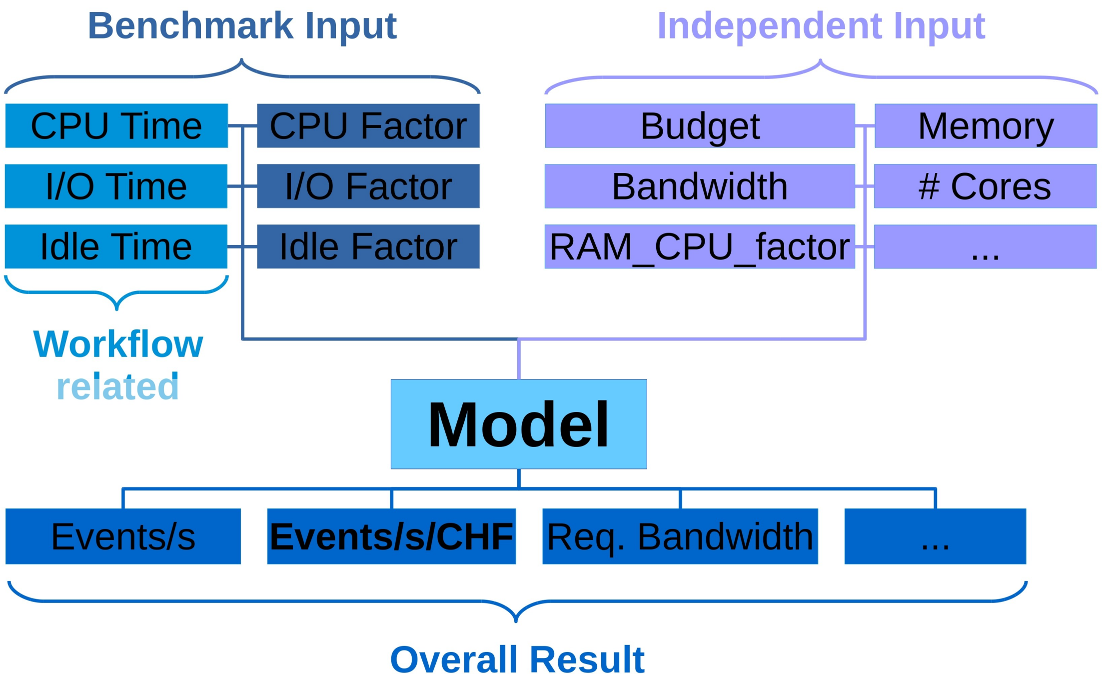
GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- Running same workflow twice ⟶ two different wall times
- Fluctuations:
  - Measure under laboratory conditions (controlled environment)
  - Use different input data ⟶ additional variation
- Fluctuations are low, average converges
- Few benchmarks represent entire workload

- Plot: y-Axis does not start
- at zero
- Error = StdDev / sqrt(n)

The Model - Concept

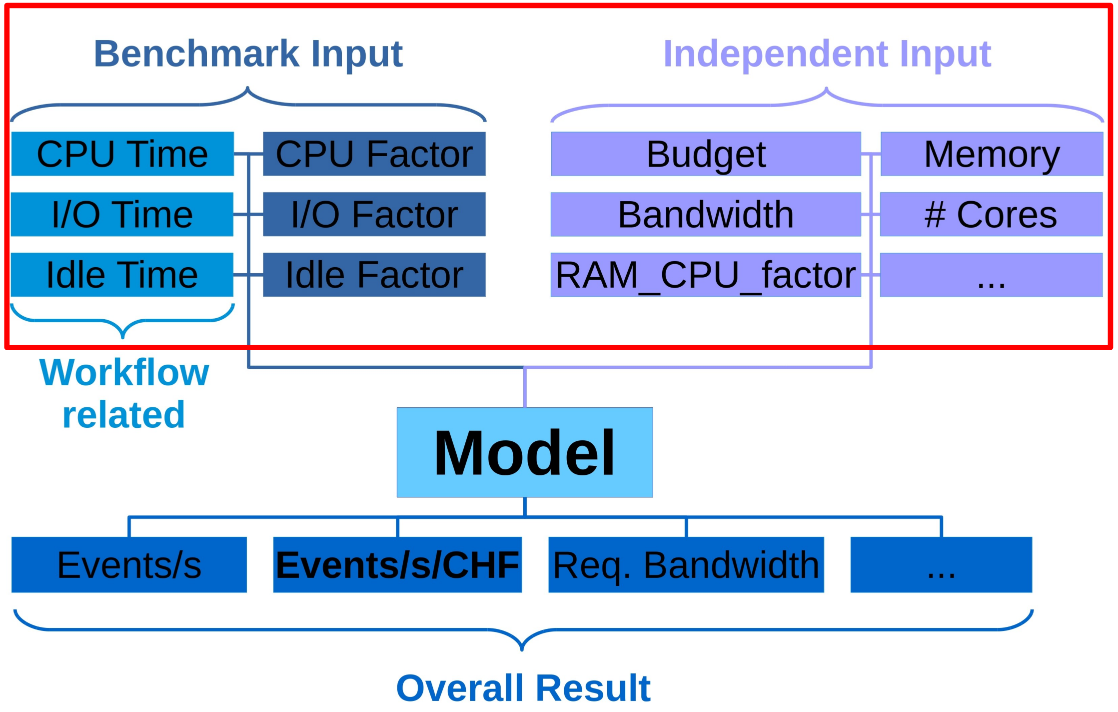
- Simple Model: linear combination
- Infrastructure inputs based on benchmarks



- 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} $Dollar^{-1}$, “physics“ per time and money) or infrastructure metric (e.g. bandwidth)
- Assessment of Clouds

The Model - Input

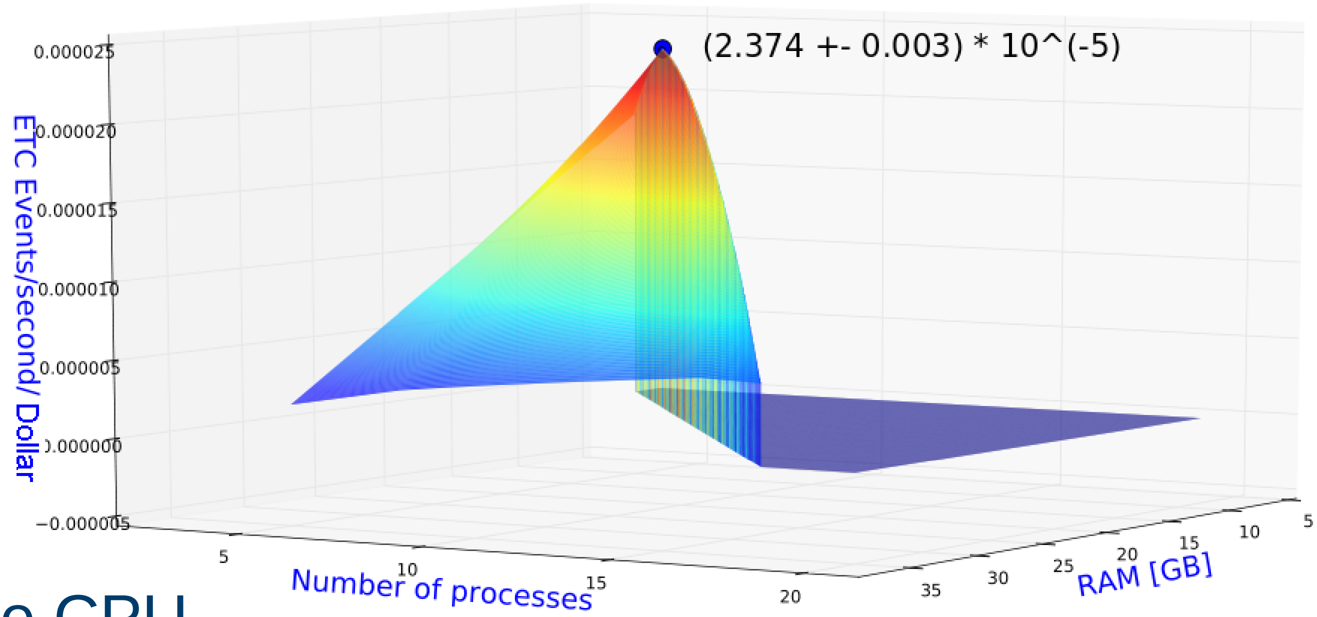
- Infrastructure as well as workflow parameters needed
- Workflow specifics obtained from anywhere (Grid)



- 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

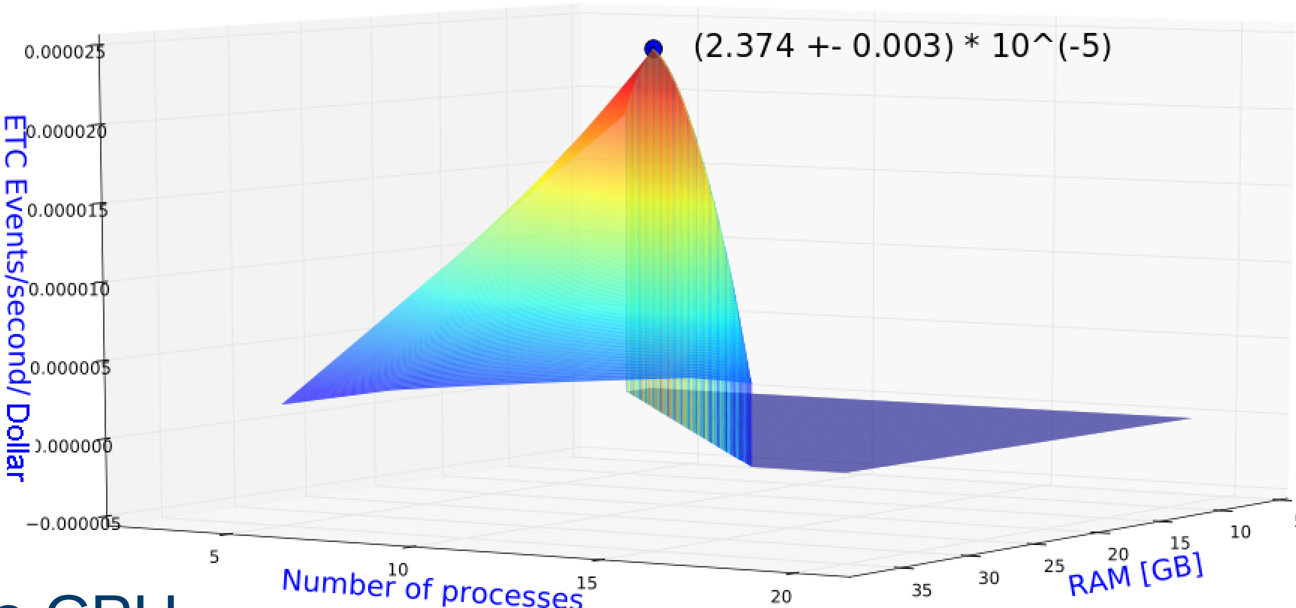
The Model - Example

- Investigate <u>overcommitting</u>
- RAW reconstruction
- Fixed budget
- Example: few VMs (cost CloudSigma)
- Vary inputs
- “Trade” RAM for more CPU



The Model - Example

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN



- Investigate <u>overcommitting</u>
- RAW reconstruction
- Fixed budget
- Example: few VMs (cost CloudSigma)
- Vary inputs
- “Trade” RAM for more CPU
- Result: 1000 Dollars 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)

Workflows - Validation

- Model Validation: cover workflow and infrastructure aspects
- Two separate dedicated infrastructures, major ATLAS workflows
- Reference + target VM: Model target, compare to measurement
- Difference Model prediction to measurement with respect to the average measured duration

Model difference	Wall Time [%]
EvGen	0,49
MC Sim	2,68
Reconstruction	-0,28

- Results from 25 measurements
- Good agreement
- Move to the Cloud

The Model - Results

- Which of the HNSciCloud providers is best?
- Example: same price for VMs between the providers
- Model estimates uncertainty on prediction from the benchmarks

Events/second/Dollar	**Exoscale**	**IBM**	**T-Systems**
EvGen	$0{,}37 \pm 0{,}00$	$0{,}25 \pm 0{,}00$	$0{,}24 \pm 0{,}01$
MC Sim	$0{,}72 \pm 0{,}03$	$0{,}39 \pm 0{,}03$	
Reco 1	$12{,}88 \pm 0{,}08$	$8{,}19 \pm 0{,}36$	$9{,}38 \pm 0{,}03$
Reco 2	$4{,}44 \pm 0{,}05$	$3{,}29 \pm 0{,}07$	$2{,}51 \pm 0{,}07$
Reco 3	$8{,}71 \pm 0{,}03$	$5{,}70 \pm 0{,}08$	$4{,}37 \pm 0{,}15$
Digi Reco	$2{,}38 \pm 0{,}00$	$1{,}19 \pm 0{,}01$	$1{,}04 \pm 0{,}01$

- Reco 1: with merging 2015 data, Reco 2: no merging 2015 data, Reco 3: no merging 2017 data

The Model - Results

- Which of the HNSciCloud providers is best?
- Example: same price for VMs between the providers
- Model estimates uncertainty on prediction from the benchmarks

Events/second/Dollar	Exoscale	IBM	T-Systems
EvGen	1. $0{,}37 \pm 0{,}00$	2. $0{,}25 \pm 0{,}00$	3. $0{,}24 \pm 0{,}01$
MC Sim	$0{,}72 \pm 0{,}03$	$0{,}39 \pm 0{,}03$	
Reco 1	$12{,}88 \pm 0{,}08$	$8{,}19 \pm 0{,}36$	$9{,}38 \pm 0{,}03$
Reco 2	$4{,}44 \pm 0{,}05$	$3{,}29 \pm 0{,}07$	$2{,}51 \pm 0{,}07$
Reco 3	$8{,}71 \pm 0{,}03$	$5{,}70 \pm 0{,}08$	$4{,}37 \pm 0{,}15$
Digi Reco	$2{,}38 \pm 0{,}00$	$1{,}19 \pm 0{,}01$	$1{,}04 \pm 0{,}01$

- Exoscale would be the preferred provider in this example
- Model easily adapted to include correct price schemata

Large scale - Measurement

- Larger scale: 10 VMs per provider, 10+ Jobs per VM
- Standard deviation (StdDev) much larger than for single-VMs

	Exoscale Wall Time [s]	StdDev [%]	IBM Wall Time [s]	StdDev [%]	T-Systems Wall Time [s]	StdDev [%]
EvGen	2915	4,70	3927	3,02	4089	3,07
MC Sim	1279	4,77	2321	15,19	4808	27,00
Reco 1	5737	7,68	8193	6,97		
Reco 2	5700	3,34	8165	7,32	15430	16,93
Reco 3	4547	10,50	7061	8,90	8681	21,05
Digi Reco	8381	8,12	15116	15,85	25821	23,82

Large scale - Measurement

- Larger scale: 10 VMs per provider, 10+ Jobs per VM
- Standard deviation (StdDev) much larger than for single VMs

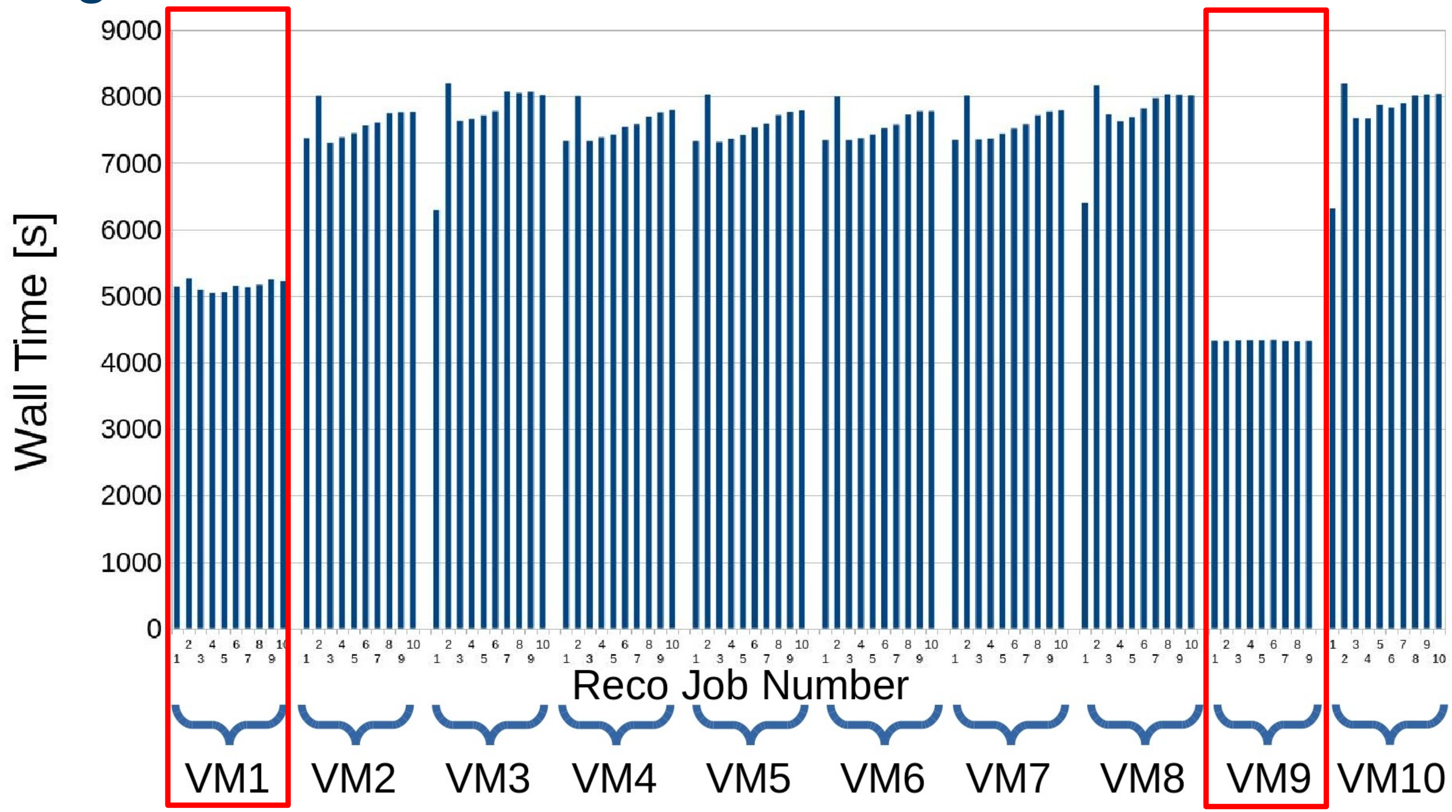
	Exoscale Wall Time [s]	StdDev [%]	IBM Wall Time [s]	StdDev [%]	T-Systems Wall Time [s]	StdDev [%]
EvGen	2915	4,70	3927	3,02	4089	3,07
MC Sim	1279	4,77	2321	15,19	4808	27,00
Reco 1	5737	7,68	8193	6,97		
Reco 2	5700	3,34	8165	7,32	15430	16,93
Reco 3	4547	10,50	7061	8,90	8681	21,05
Digi Reco	8381	8,12	15116	15,85	25821	23,82

- Large StdDev not necessarily bad, depends on homogeneity
- Workflow impact on fluctuations small

T-Systems - Fluctuations

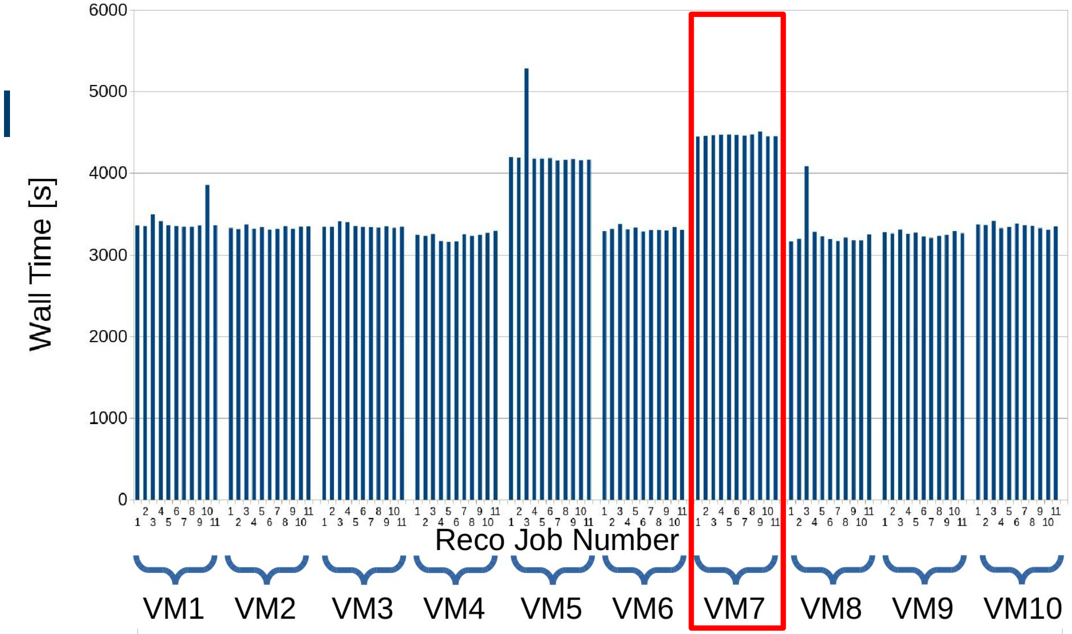
GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

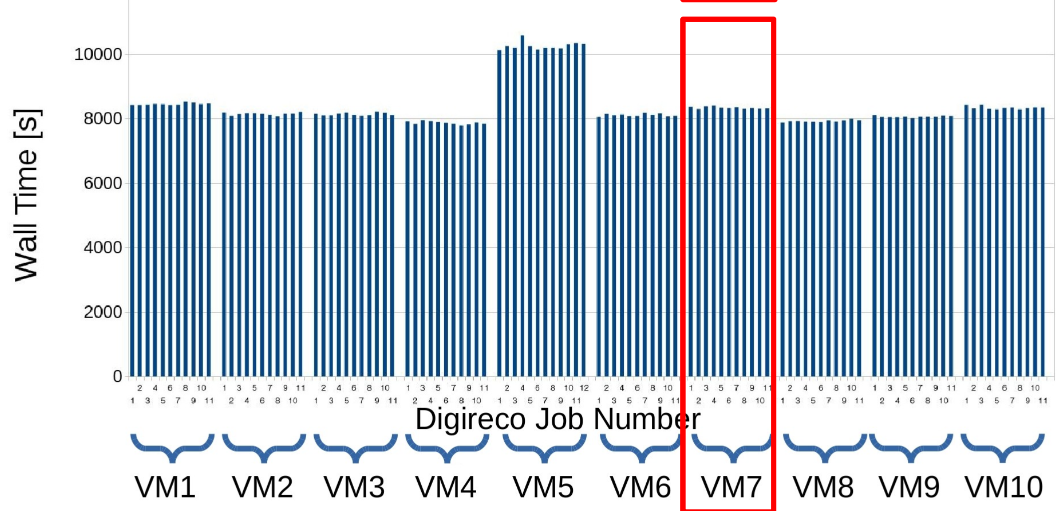
- Infrastructure fluctuations on all three providers
- Example: T-Systems
- Reco workflow, two generations of VMs
- Good provider: Better performing VMs as bonus
- More accurate model results by splitting the infrastructure



Exoscale - Fluctuations

- Infrastructure fluctuations on all three providers
- Example: Exoscale
- Top: Reco workflow
- Bottom: Digi+Reco workflow
- Two generations of VMs
- VM 7 appears faster for Digireco
- Migrated to better hardware? Fewer influences from neighbouring VMs?
- Include in uncertainties





Model discrepancy

- Model prediction error for the large scale measurement
- Large standard deviations in the measurements result in large discrepancies
- T-Systems: more than two generations?
- Biggest prediction discrepancies, due to benchmark-VM differences (heterogeneous infrastructure)

Model Error:	Exoscale [%]	IBM [%]	IBM fast [%]	T-Systems [%]	T-Systems fast [%]
EvGen	1,92	0,62		2,15	
MC Sim	4,47	6,39	3,07	-36,01	21,80
Reco 1	-4,46	-0,87	-1,38		
Reco 2	-3,97	-2,09	0,83	-18,80	7,71
Reco 3	-6,01	1,73	-2,26	-3,05	-5,70
Digi Reco 2	0,80	3,43	11,92	-4,77	14,23

Model discrepancy

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- Model prediction error for the large scale measurement
- Large standard deviations in the measurements result in large discrepancies
- T-Systems: more than two generations?
- Biggest prediction discrepancies, due to benchmark-VM differences (heterogeneous infrastructure)

Model Error:	**Exoscale [%]**	**IBM [%]**	**IBM fast [%]**	**T-Systems [%]**	**T-Systems fast [%]**
EvGen	1,92	0,62		2,15	
MC Sim	4,47	6,39	3,07	-36,01	21,80
Reco 1	-4,46	-0,87	-1,38		
Reco 2	-3,97	-2,09	0,83	-18,80	7,71
Reco 3	-6,01	1,73	-2,26	-3,05	-5,70
Digi Reco 2	0,80	3,43	11,92	-4,77	14,23

Cross-workflow estimations

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- Different workflow types: different input data, software stack and e.g. merging/no merging
- Single-VMs, compare low fluctuations top rows

Model Error [%]	IBM VM1	T-Systems VM1	Exoscale VM1	Exoscale VM2
Reco 1	1,11	-0,80	-0,91	-0,77
Reco 2	-3,74	1,64	0,97	0,77
Reco 1 with 2	-23,22	6,60	-13,59	-10,88

- Cross-workflow modelling introduces large error
- Not comparing “apples with oranges”: categorise jobs
- Careful also with: Number of events (overheads), Number of cores

Conclusion

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- 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 and compares providers
- Prerequisite: Carefully classify workflows
- Fluctuations and differences between VMs have to be considered when benchmarking

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

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

	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	

	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	

Single VM - Model

- Compare Model prediction to measurement
- Single VMs within different Cloud providers
- Error → discrepancy between Model and measurement
- Uncertainty → Model estimation from standard deviation of benchmark results

[%]	Event Generation		Simulation		Reconstruction	
	Error	Uncertainty	Error	Uncertainty	Error	Uncertainty
IBM	0,15	1,56	2,49	4,23	1.85	4.34
TSY	0,84	0,74	1,98	4,46	-0.40	0.35
EXO 1	-0,34	1,12	2,63	4,36	-0.17	0.62
EXO 2	0,02	1,78	1,83	4,33	-0.51	0.93
EXO 3	3,00	1,03	2,68	4,43	-0.05	0.94
GOE	0,68	1,11	2,49	4,41	1.02	0.73

Large scale – Model error

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- Model prediction error for the large scale measurement
- Large standard deviations in the measurements result in large discrepancies
- T-Systems: worst case scenario
- prediction far off, due to benchmark-VM differences

Model Error:	**Exoscale [%]**	**IBM [%]**	**T-Systems [%]**
EvGen	-6,48	0,62	1,88
MC Sim	-0,41	-22,93	-60,30
Reco 1	-3,19	-7,40	-33,89
Reco 2	0,88	-1,84	-38,00
Reco 3	-5,17	-14,19	-28,87

Large scale - Measurement

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- Larger scale: 10 VMs per provider, 10+ Jobs per VM
- Standard deviation (StdDev) much larger, large StdDev not bad

	Exoscale Wall Time [s]	StdDev [%]	IBM Wall Time [s]	StdDev [%]	T-Systems Wall Time [s]	StdDev [%]
EvGen	2915	4,70	3927	3,02	4089	3,07
MC Sim	1279	4,77	2321	15,19	4808	27,00
Reco 1	5737	7,68	8193	6,97		
Reco 2	5700	3,34	8165	7,32	15430	16,93
Reco 3	4547	10,50	7061	8,90	8681	21,05
Reco 4	4528	8,78	7009	8,62	10785	14,22
Reco 5	3147	2,79	4756	14,31	7099	16,95
Reco 6	3529	12,77	4919	9,52	7986	22,66
Reco 7	5550	16,99	7630	9,95	15348	18,80
Digi Reco 1	1210	10,21	2019	8,52	2789	18,77
Digi Reco 2	8381	8,12	15116	15,85	25821	23,82

Workflows

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

- Event Generation (single-core, CPU intensive), Monte-Carlo simulation (CPU intensive), Reconstruction (data intensive), Analysis (data intensive)
- CPU intensive workflows understood
- Analysis runs user code, unpredictable