How to Make Profit: Exploiting Fluctuating Electricity Prices with Albatross, A Runtime System for Heterogeneous HPC Clusters

Timo Hönig, Christopher Eibel, Adam Wagenhäuser, Maximilian Wagner, and Wolfgang Schröder-Preikschat

8th International Workshop on Runtime and Operating Systems for Supercomputers (ROSS 2018), Tempe, AZ, USA.

June 12, 2018

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

DFG

Gode Wind Farm: 582 MW

North Sea, Europe

Photo: DONG Energy

Solana Generating Station: 280 MW

Photo: Abengoa Energy

Solana Generating Station: 280 MW

Arizona, USA

Photo: Abengoa Energy, Google

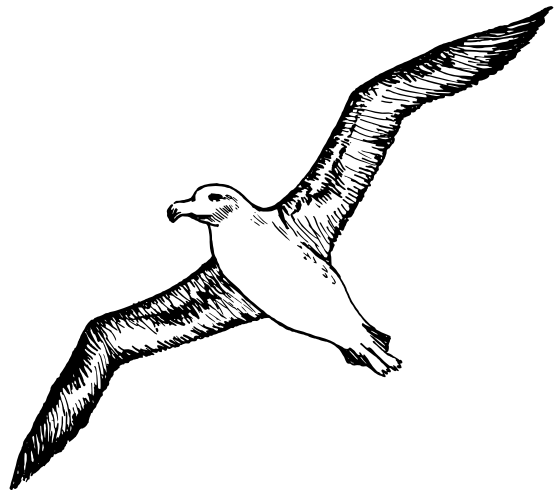
Arizona State University: 24 MW

Arizona, USA

Photo: Arizona State University

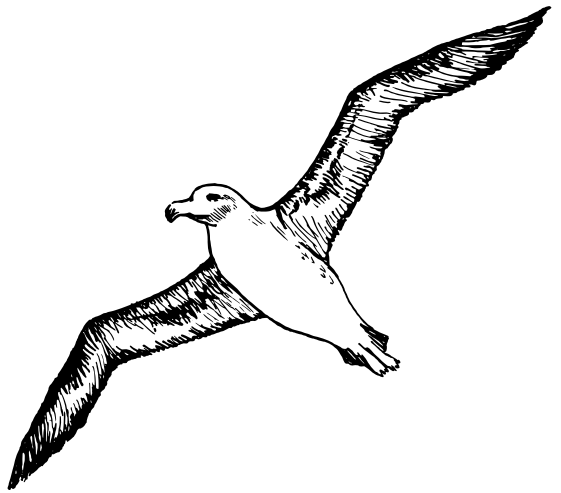
Building a Runtime System for Heterogeneous HPC Clusters

- Reenewable electricity sources have a great impact on the grid.



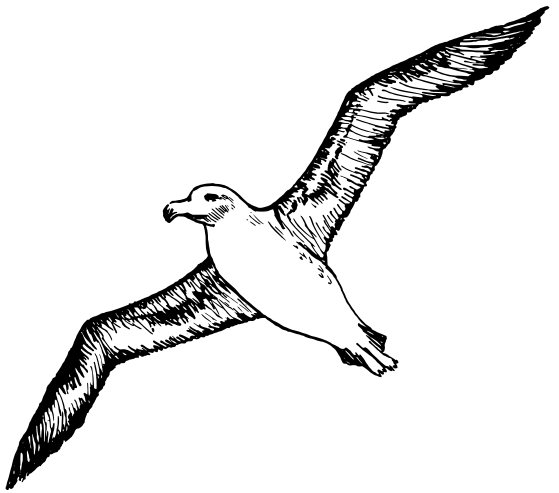
Building a Runtime System for Heterogeneous HPC Clusters

- Reenewable electricity sources have a great impact on the grid.
- Availability of renewables leads to fluctuating electricity prices.



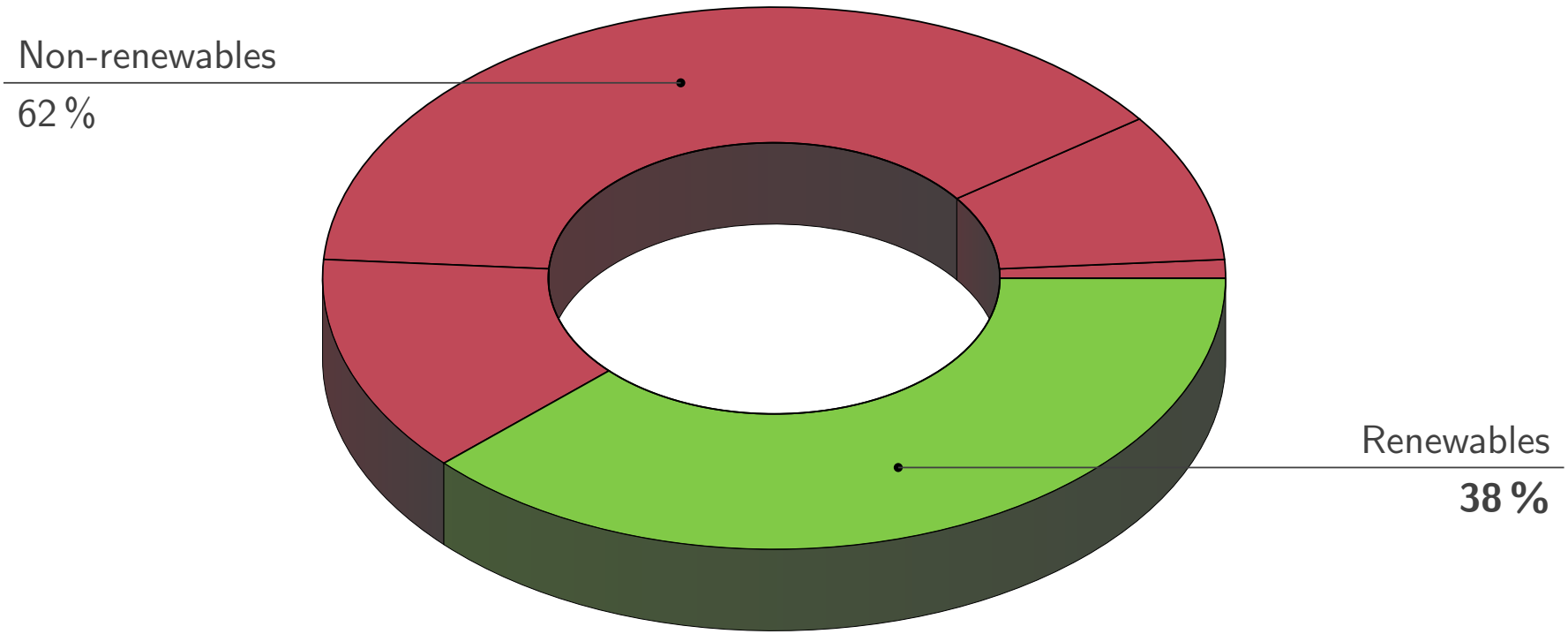
Building a Runtime System for Heterogeneous HPC Clusters

- Reenewable electricity sources have a great impact on the grid.
- Availability of renewables leads to fluctuating electricity prices.



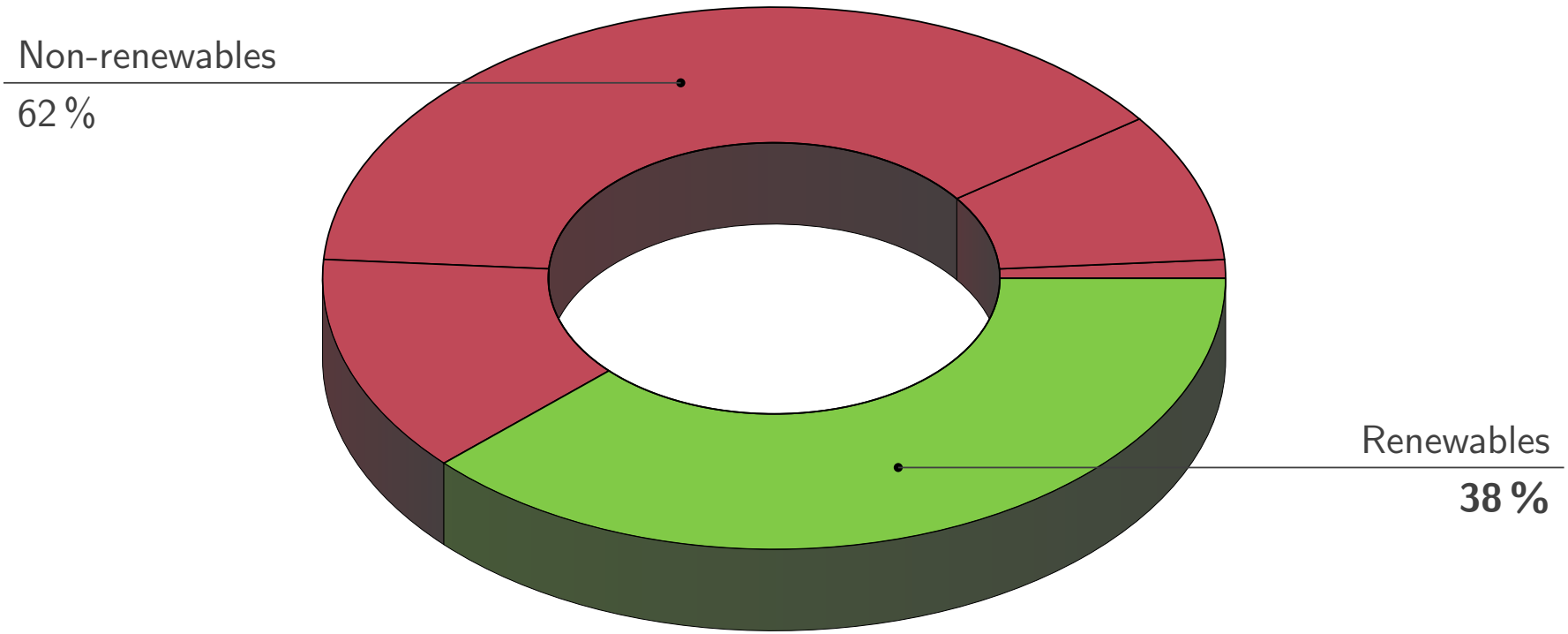
- **How can we design and build an HPC runtime system to exploit dynamic electricity pricing?**

Energy Mix (Germany, 2017)



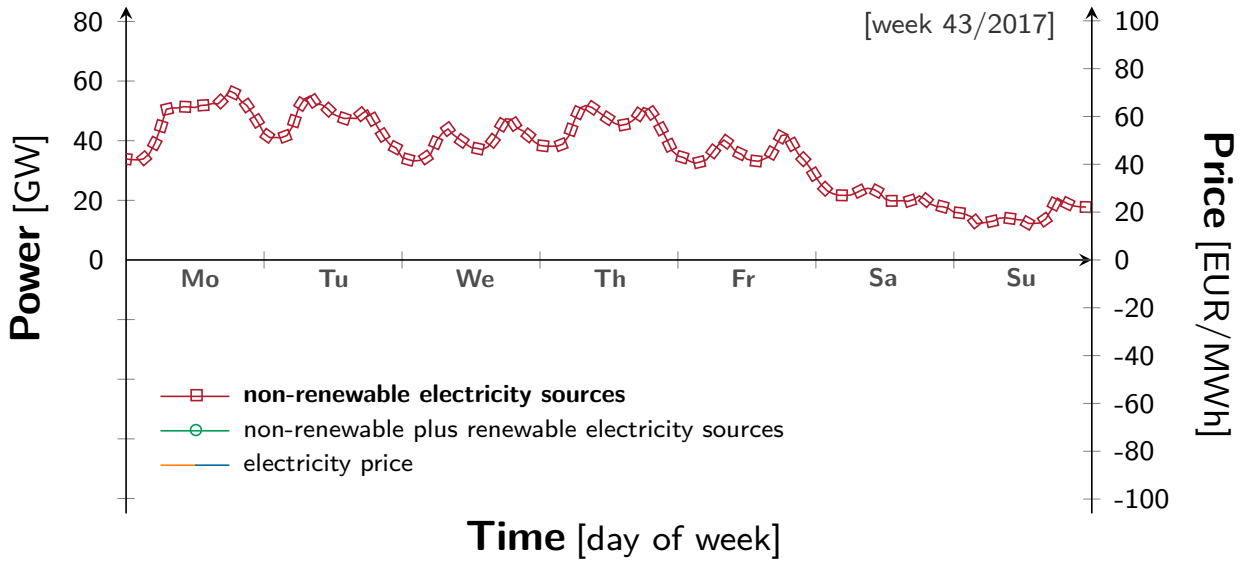
- non-renewables: 339.93 TWh, renewables: 209.97 TWh

Energy Mix (Germany, 2017)

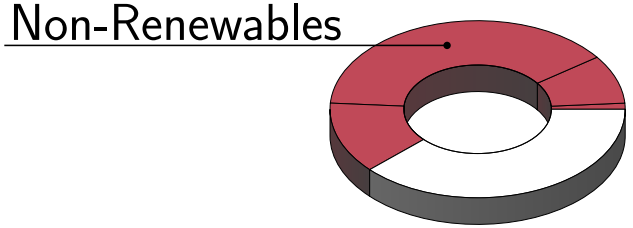


- non-renewables: 339.93 TWh, renewables: 209.97 TWh
- dynamic availability of renewables ➔ fluctuation electricity prices

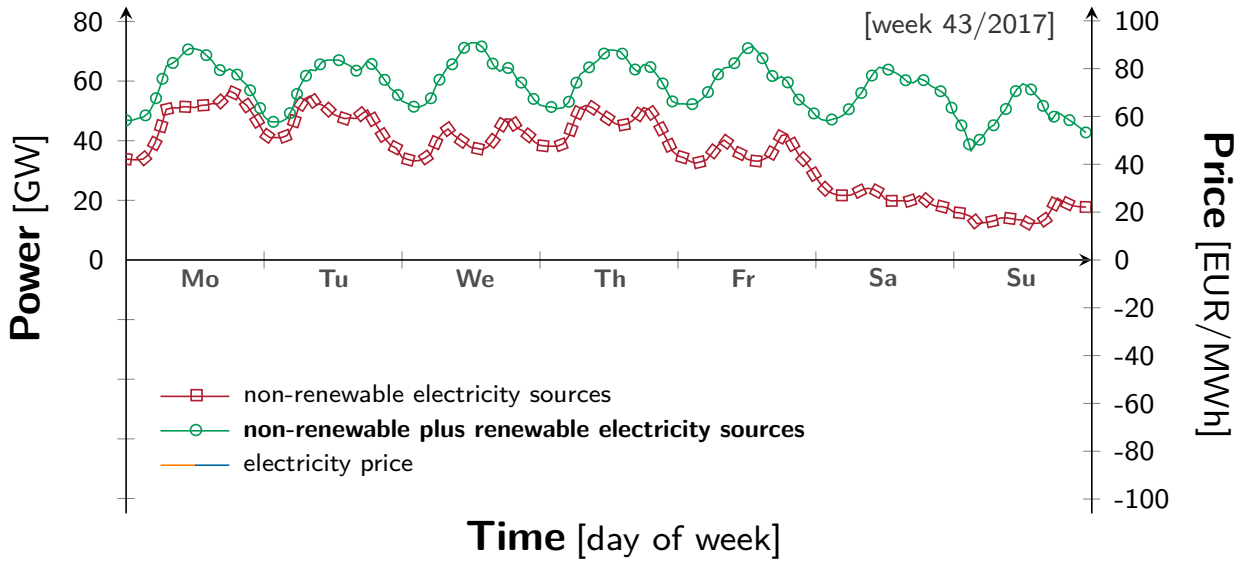
Fluctuation of Electricity Prices



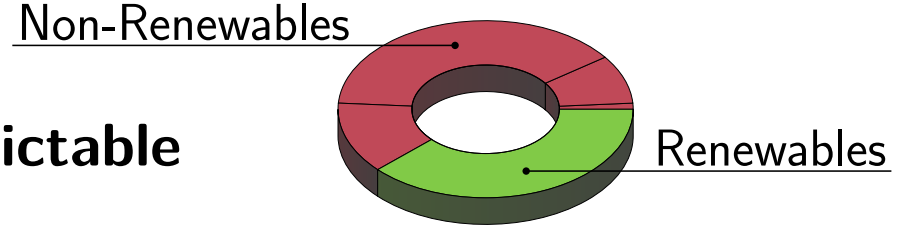
- non-renewables can be **planned**...



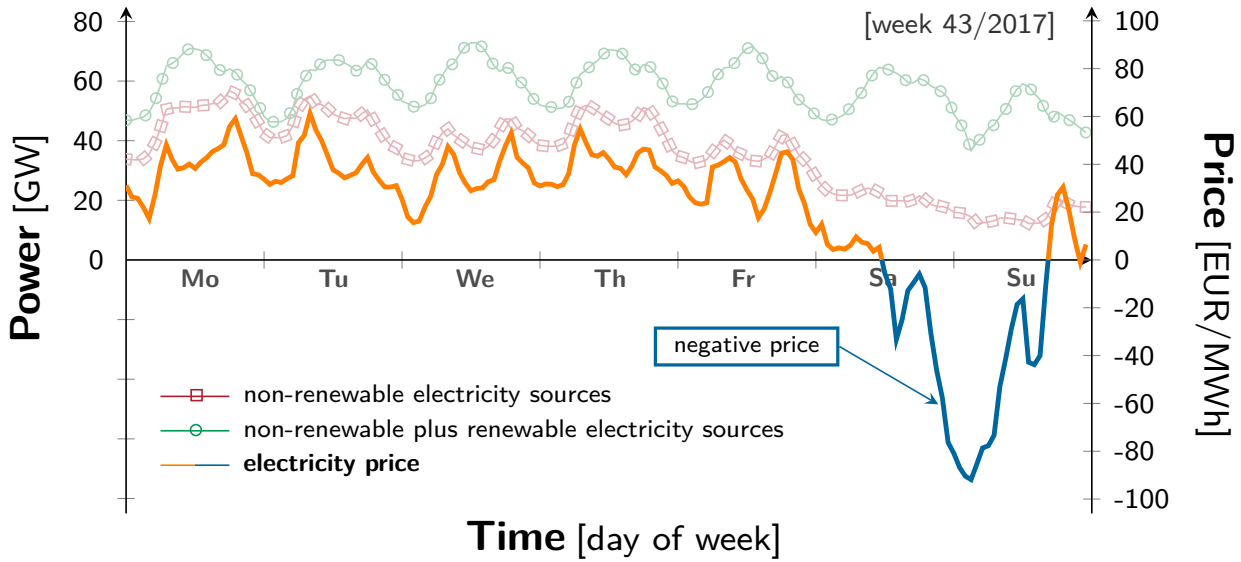
Fluctuation of Electricity Prices



- non-renewables can be **planned**...
- ...but renewables are often **unpredictable**

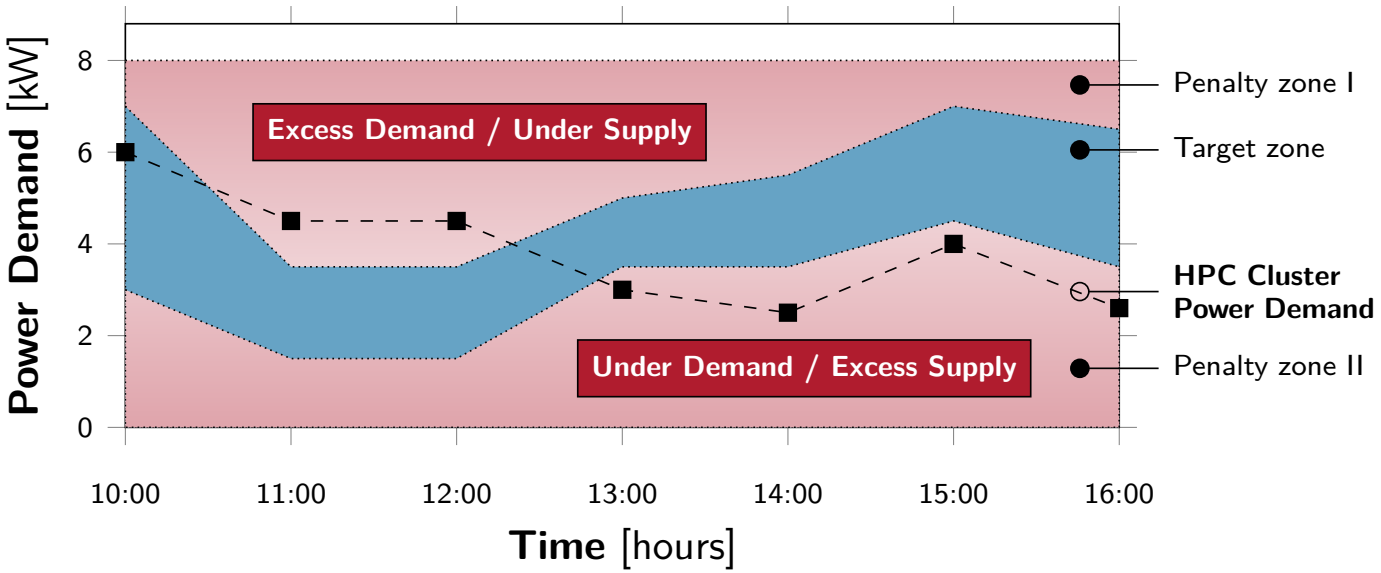


Fluctuation of Electricity Prices



- consequence: **strong fluctuation** of the electricity price
- extreme **imbalance of supply & demand** ➔ negative prices

Building a System that Benefits from Dynamic Electricity Prices



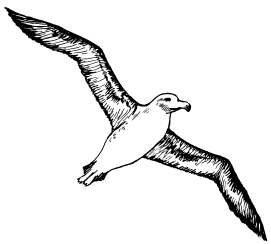
- challenge: take advantage of low and negative prices
- consider contracts (i.e., penalties enforced by grid operators)
- integrate with existing HPC cluster infrastructure

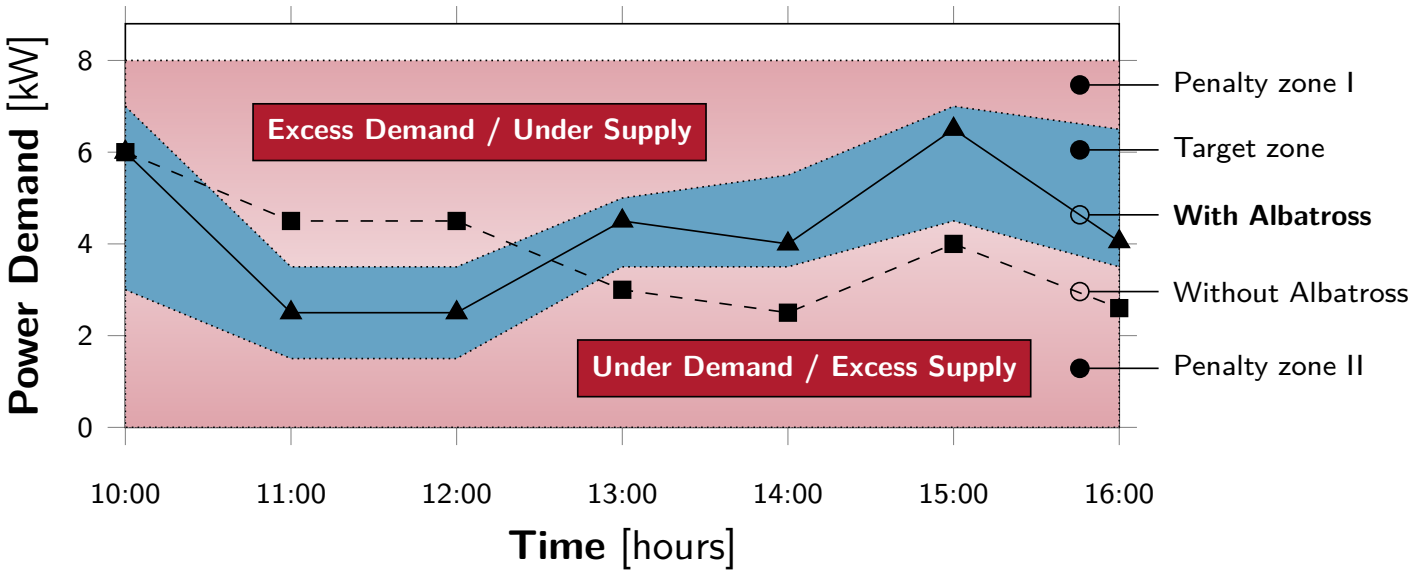
Building a System that Benefits from Dynamic Electricity Prices

- build a flexible runtime system that integrates data on electricity pricing into its operating decisions
- reduce energy demand when prices are high
- increase energy demand when prices are low (or negative)

- challenge: take advantage of low and negative prices
- consider contracts (i.e., penalties enforced by grid operators)
- integrate with existing HPC cluster infrastructure

Building a System that Benefits from Dynamic Electricity Prices



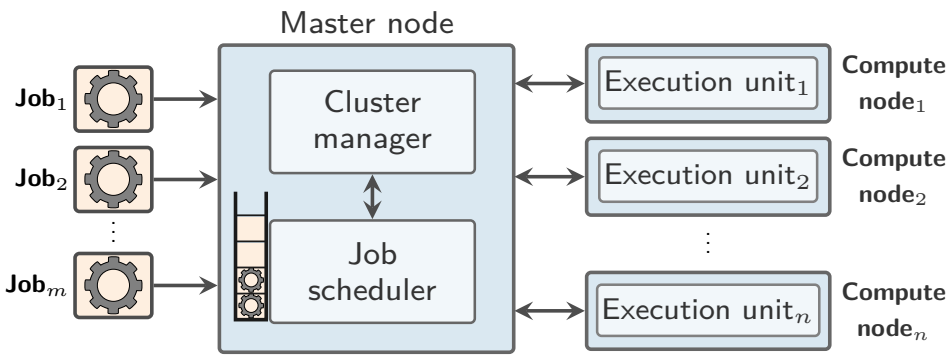


Our proposed runtime system Albatross:

- implements **low- and high-power** operation modes
- considers **heterogeneity** aspects of the cluster hardware
- respects **QoS and non-functional requirements** of workloads

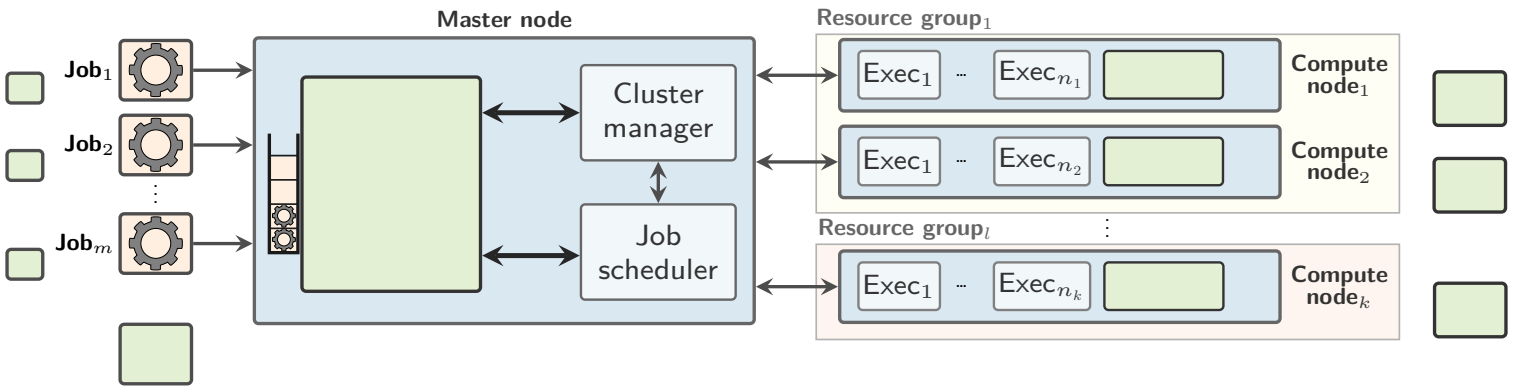
Albatross: Design and Implementation

- **generic** system model of an HPC cluster **without** Albatross

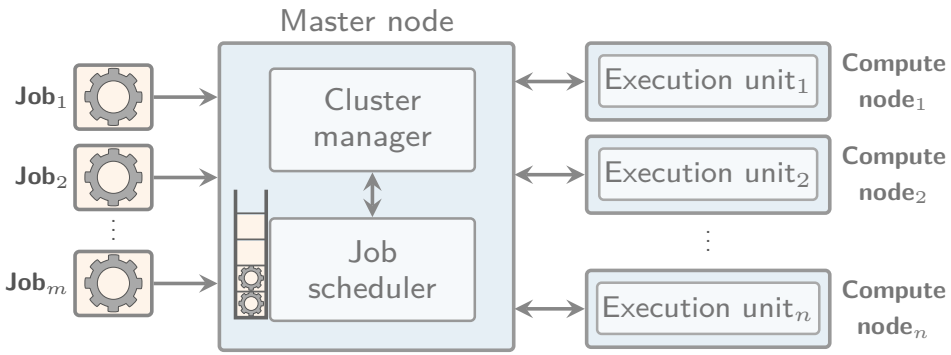


Albatross: Design and Implementation

- **extended** system model of an HPC cluster **with** Albatross

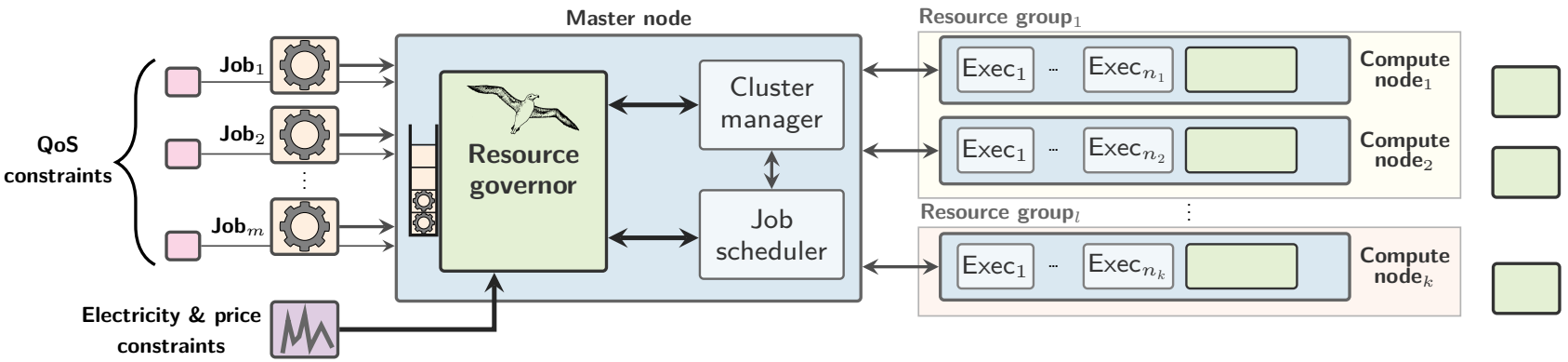


- **generic** system model of an HPC cluster without Albatross



Albatross: Design and Implementation

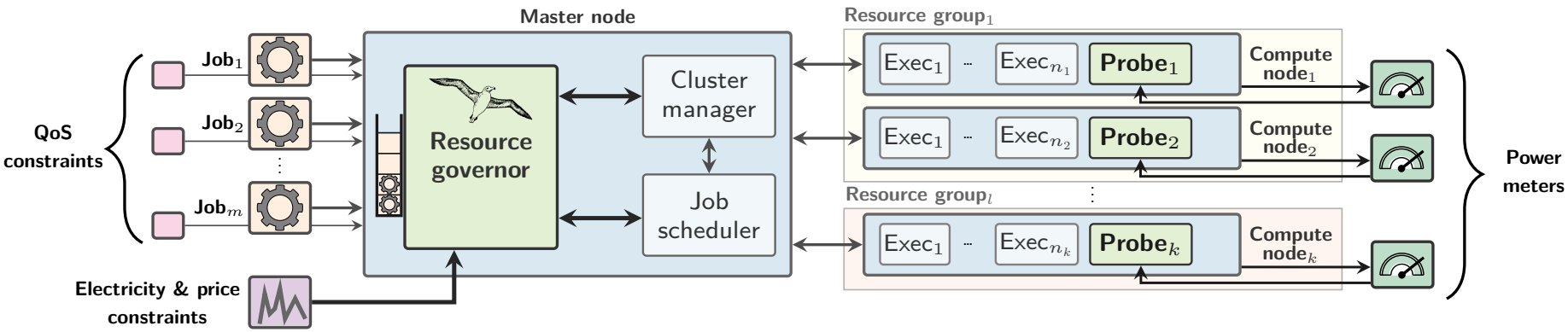
- **extended** system model of an HPC cluster **with** Albatross



- input interfaces for job submission and control
  - submission interface for electricity-pricing data
  - real-time data as provisioned by grid operator
  - detailed job description, added QoS constraints

Albatross: Design and Implementation

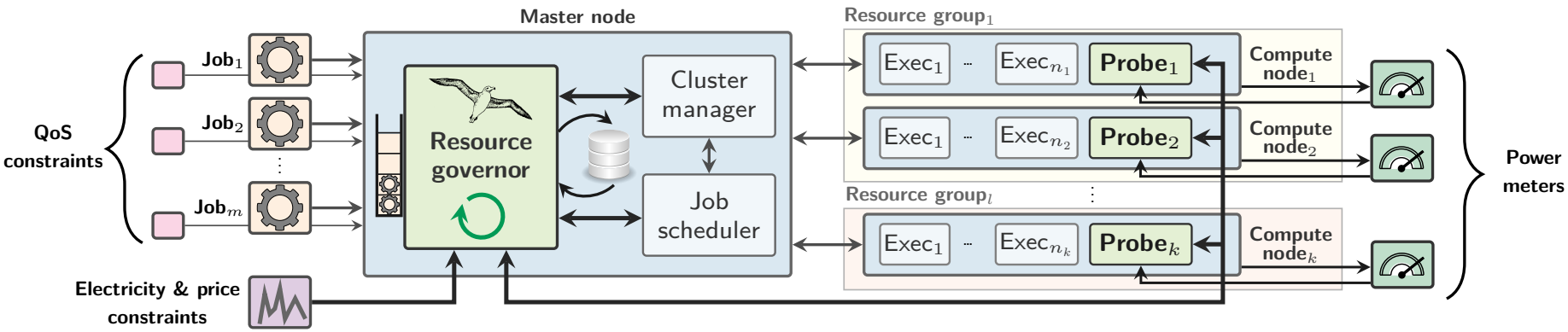
- **extended** system model of an HPC cluster **with** Albatross



- power monitoring and control interfaces
  - probes operate power-management features (e.g., power capping)
  - hardware-specific metering (e.g., RAPL, measuring devices)
  - delivery of base data for feedback loop to the resource governor

Albatross: Design and Implementation

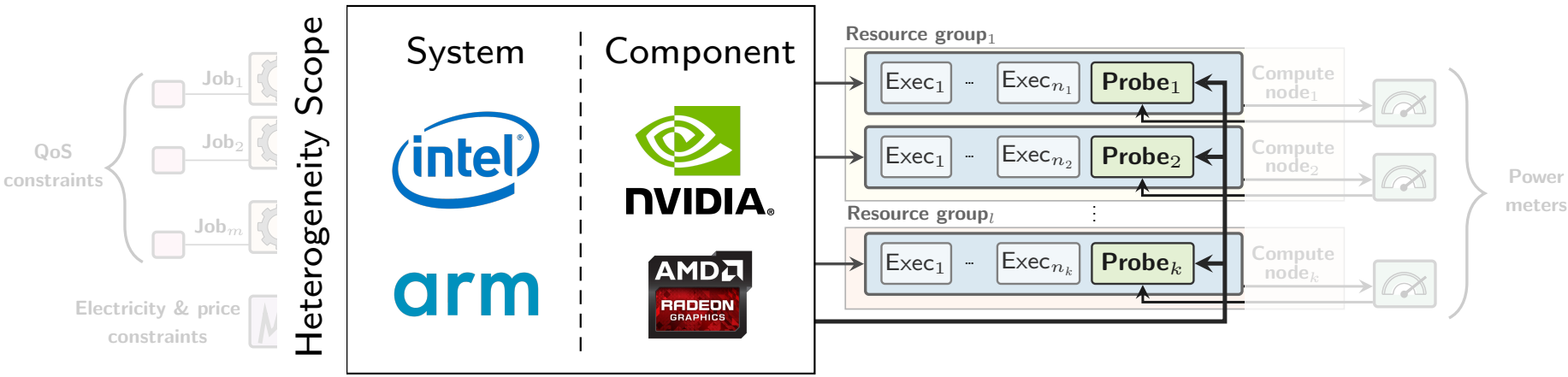
- **extended** system model of an HPC cluster **with** Albatross



- runtime control of the HPC system by the resource governor
  - job submission based on QoS data and current electricity prices
  - control and measurement of power demand for individual jobs
  - consider heterogeneity aspects (system, component, and configuration)

Albatross: Design and Implementation

- **extended** system model of an HPC cluster **with** Albatross



- runtime control of the HPC system by the resource governor
  - job submission based on QoS data and current electricity prices
  - control and measurement of power demand for individual jobs
  - consider heterogeneity aspects (**system, component, and configuration**)

Albatross: Prototype Implementation

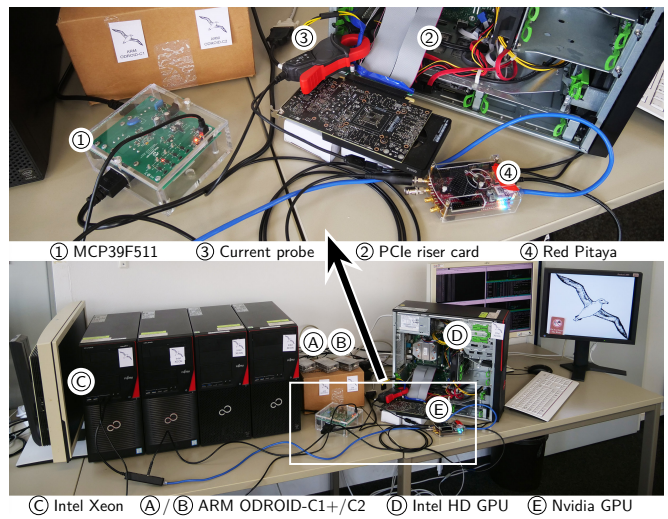
- core of Albatross implementation is based on Slurm
- constraint-aware job-to-node–assignment strategy

Feature	Moab	Tivoli	Univa	Slurm	
Job pinning	✗	✗	✓	✓	✓
CPU allocation	✓	✓	✗	✓	✓
Quality of service	✓	✗	✓	✓	✓
Generic resources	✓	✗	✓	✓	✓
Cluster status	✓	✓	✓	✓	✓
Heterogeneity aware	✓	✗	✓	✗	✓
Power/price aware	✗	✗	✗	✗	✓

Table: Feature comparison of Albatross with other workload managers.

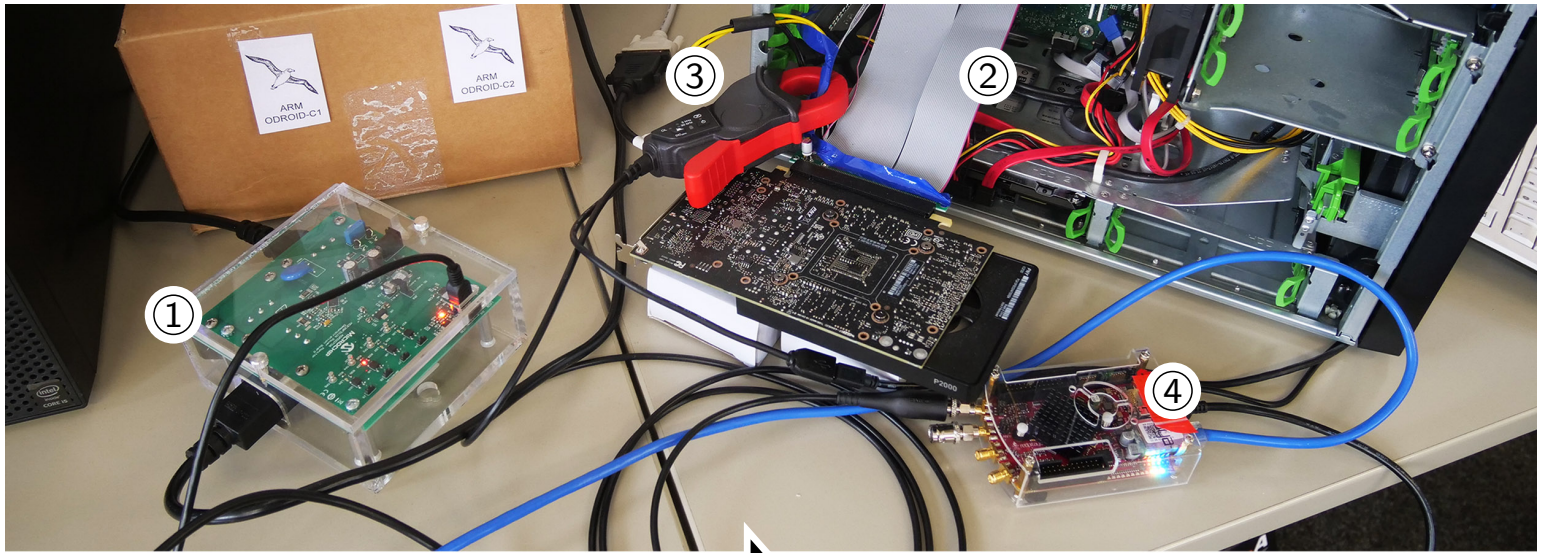
Evaluation Overview

- evaluation scenario
  - </> NAS Parallel Benchmarks (NPB) suite, Albatross runtime system
  - heterogeneous cluster with Intel & ARM CPUs, Nvidia & Intel GPUs, and various power meters

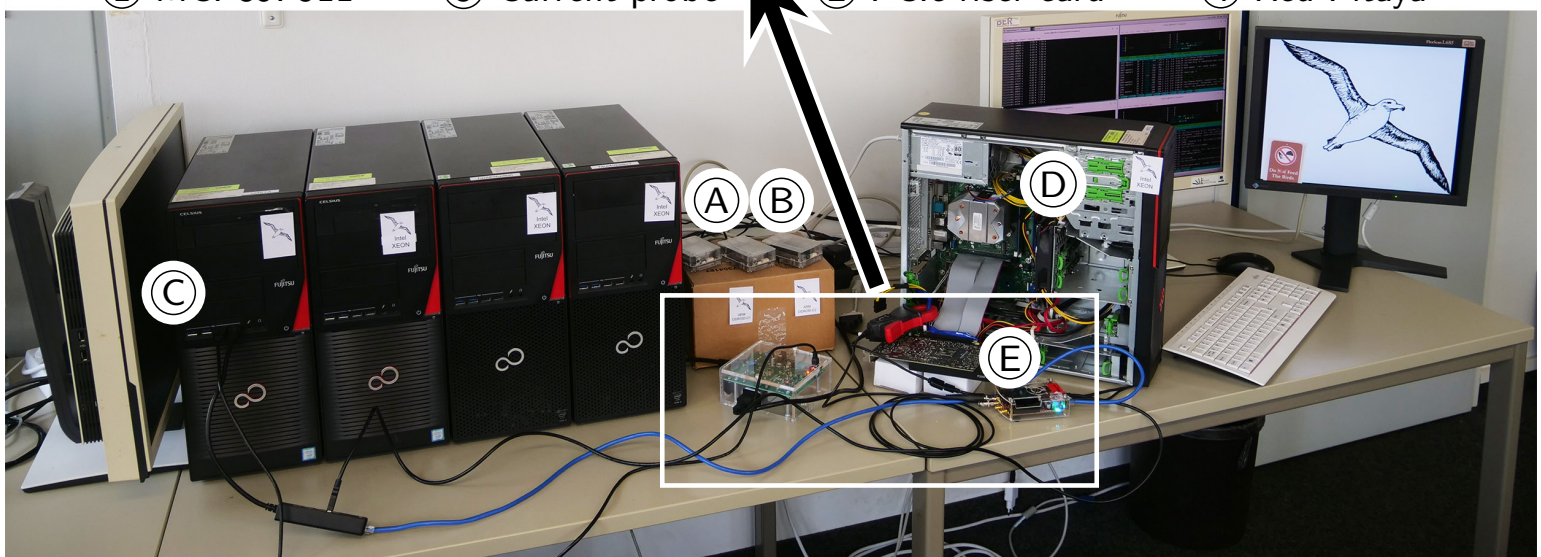


- evaluation experiments and goals
  1. impact of system configuration on **energy demand** and **execution times**
  2. influence of heterogeneity on the **energy–delay product** and power demand
  3. combination of **power** and **price awareness** to exploit dynamic electricity pricing

Evaluation Overview



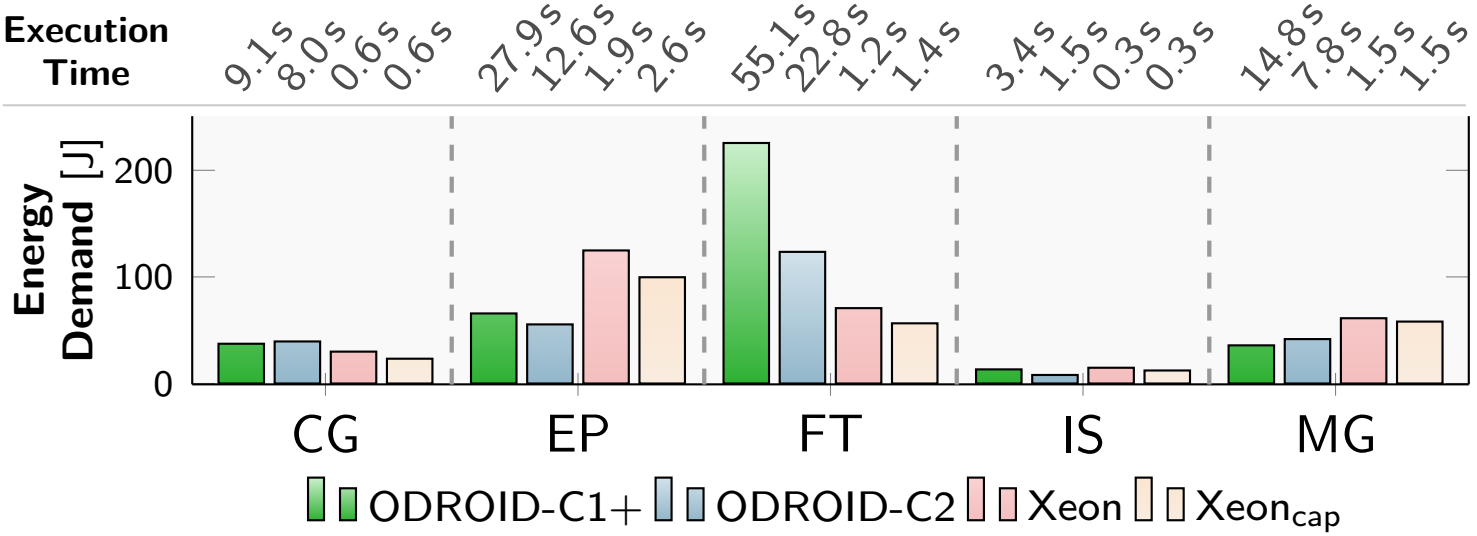
① MCP39F511 ③ Current probe ② PCIe riser card ④ Red Pitaya



Ⓒ Intel Xeon Ⓐ / Ⓑ ARM ODROID-C1+/C2 Ⓓ Intel HD GPU Ⓔ Nvidia GPU

Experiment 1: Energy Demand and Execution Times

- impact of system-level and configuration-level heterogeneity



- Xeon and Xeon$_{cap}$ dominate with performance
- energy demand in some cases lower for ARM platforms
- occasional relationship between process energy demand and execution time

Experiment 1: Energy Demand and Execution Times

- impact of system-level and configuration-level heterogeneity

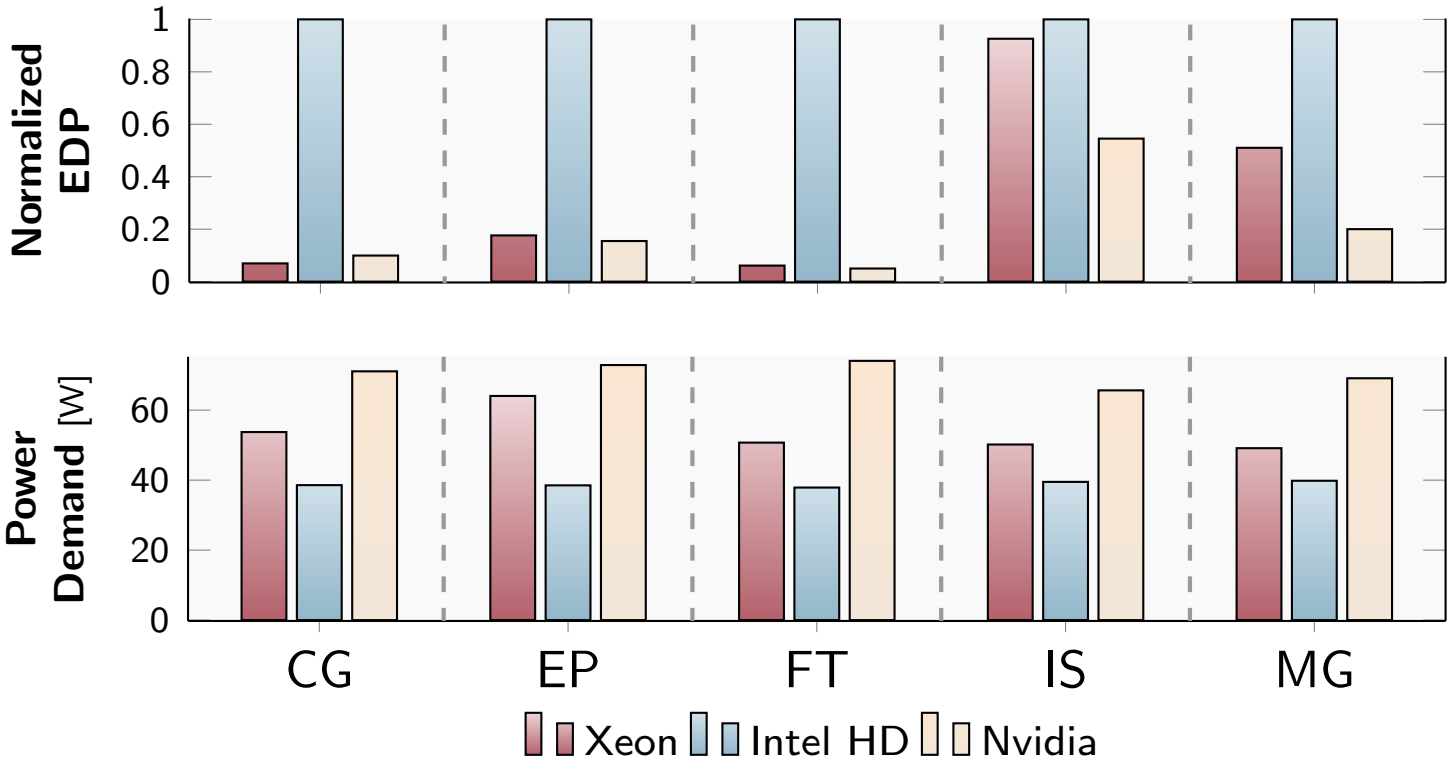
Key take-away

- race to sleep does not always reduce energy demand
- knowledge on systems' impact is absolutely necessary for operation, requires measurements

- energy demand in some cases lower for ARM platforms
- occasional relationship between process energy demand and execution time

Experiment 2: Energy–Delay Product and Power Demand

- impact of component-level heterogeneity



- EDP is best for Nvidia for all but one benchmark
- Intel HD has the worst EDP results, but lowest average power

Experiment 2: Energy–Delay Product and Power Demand

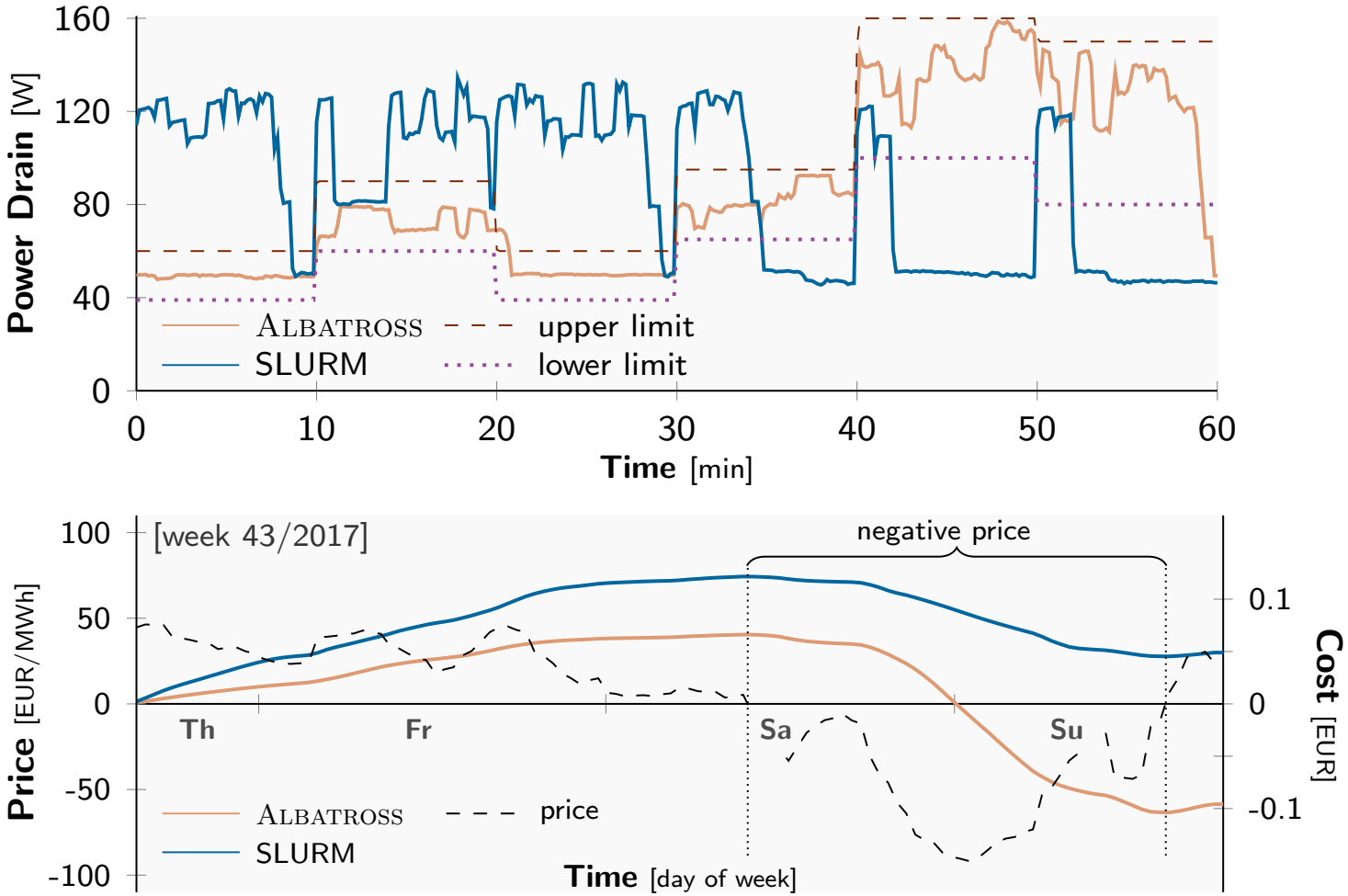
- impact of component-level heterogeneity

Key take-away

- both, EDP and average power must be considered
- to balance the cluster's power drain Albatross trades average power for EDP (and vice versa)

- EDP is best for Nvidia for all but one benchmark
- Intel HD has the worst EDP results, but lowest average power

Experiment 3: Combining Power and Price Awareness



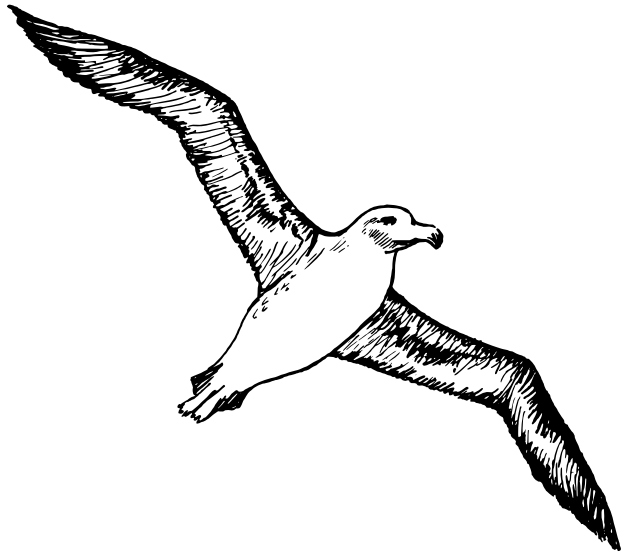
Conclusion

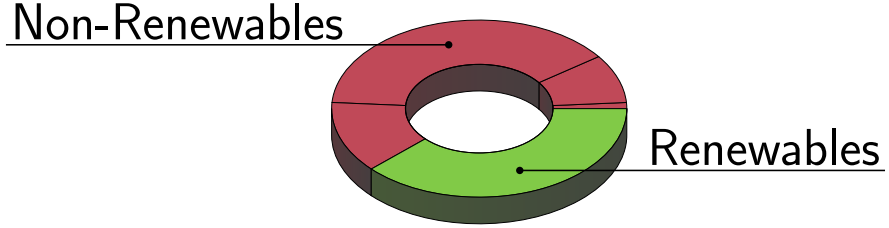
Albatross, an HPC runtime system that successfully exploits dynamic electricity pricing.

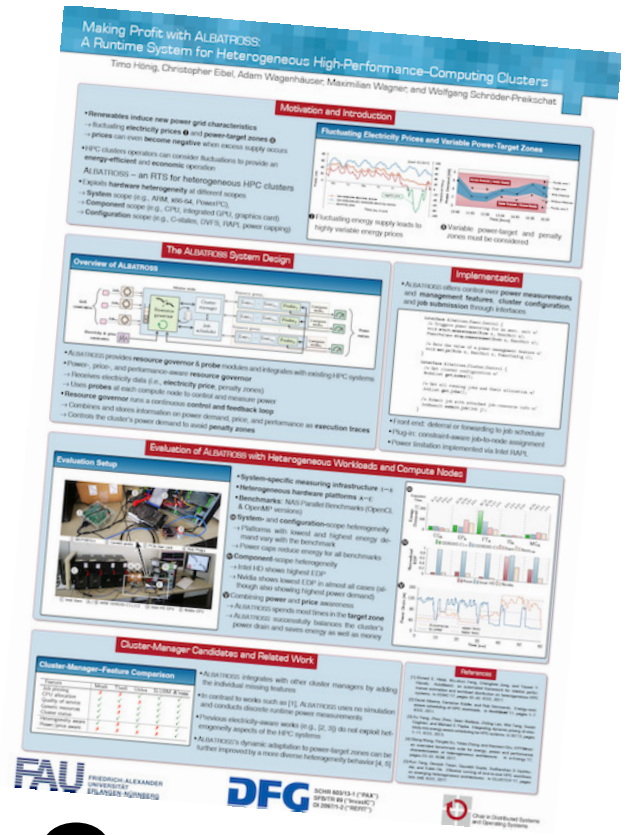
Albatross...

- considers QoS of jobs and electricity prices for deploying workloads
- exploits heterogeneity at system-level, component-level, and configuration-level
- prepares HPC clusters for future settings where energy is *just another operating resource*

Discussion

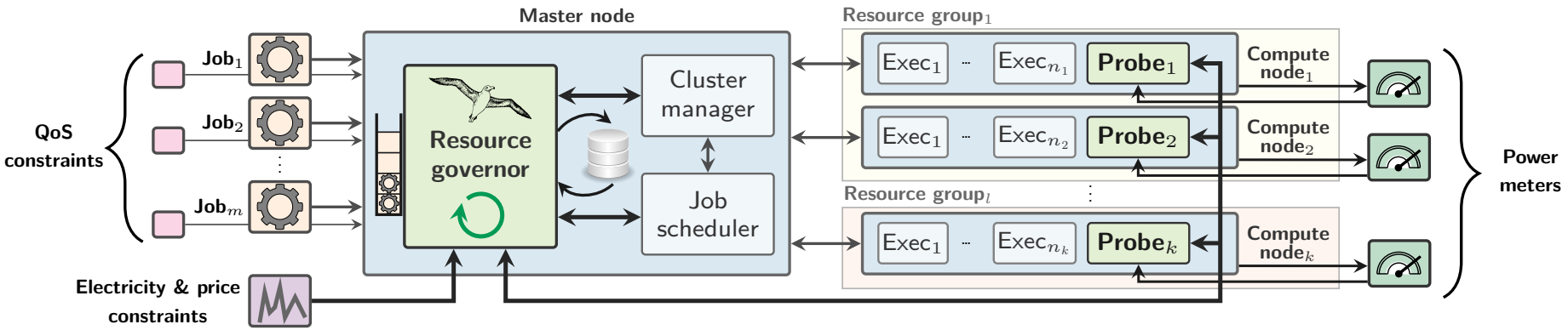






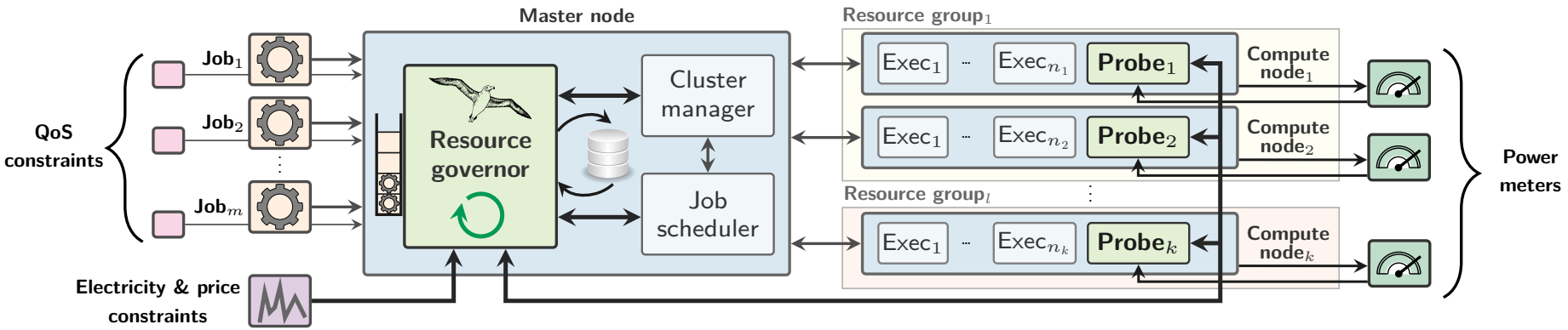
HPDC Poster Session
(Wednesday Evening)

Albatross: Design and Implementation




```
interface Albatross_Cluster_Control {
  /* Get cluster configuration */
  NodeList get_nodes();
  /* Get all running jobs and their allocation */
  JobList get_jobs();
  /* Submit job with attached job-resource info */
  JobResult submit_job(Job j);
}
```

Albatross: Design and Implementation




```
interface Albatross_Power_Control {
  /* Triggers power measuring for an exec. unit */
  void start_measurement(Node n, ExecUnit u);
  PowerValues stop_measurement(Node n, ExecUnit u);

  /* Sets the value of a power-management feature */
  void set_pm(Node n, ExecUnit u, PowerConfig c);
}
```

Timo Hönig, Christopher Eibel, Adam Wagenhäuser, Maximilian Wagner, and Wolfgang Schröder-Preikschat:

How to Make Profit: Exploiting Fluctuating Electricity Prices with Albatross, A Runtime System for Heterogeneous HPC Clusters

Proceedings of the 8th International Workshop on Runtime and Operating Systems for Supercomputers (ROSS'18), 12 June 2018, Tempe, AZ, USA

ACM, New York, NY, USA, ISBN: 978-1-4503-5864-4
DOI: 10.1145/3217189.3217193

Timo Hönig, Christopher Eibel, Adam Wagenhäuser, Maximilian Wagner, and Wolfgang Schröder-Preikschat:

Making Profit with ALBATROSS: A Runtime System for Heterogeneous High-Performance–Computing Clusters

Proceedings of the 27th International Symposium on High-Performance Parallel and Distributed Computing (HPDC'18), 11 - 15 June 2018, Tempe, AZ, USA

ACM, New York, NY, USA, ISBN: 978-1-4503-5899-6
DOI: 10.1145/3220192.3220457