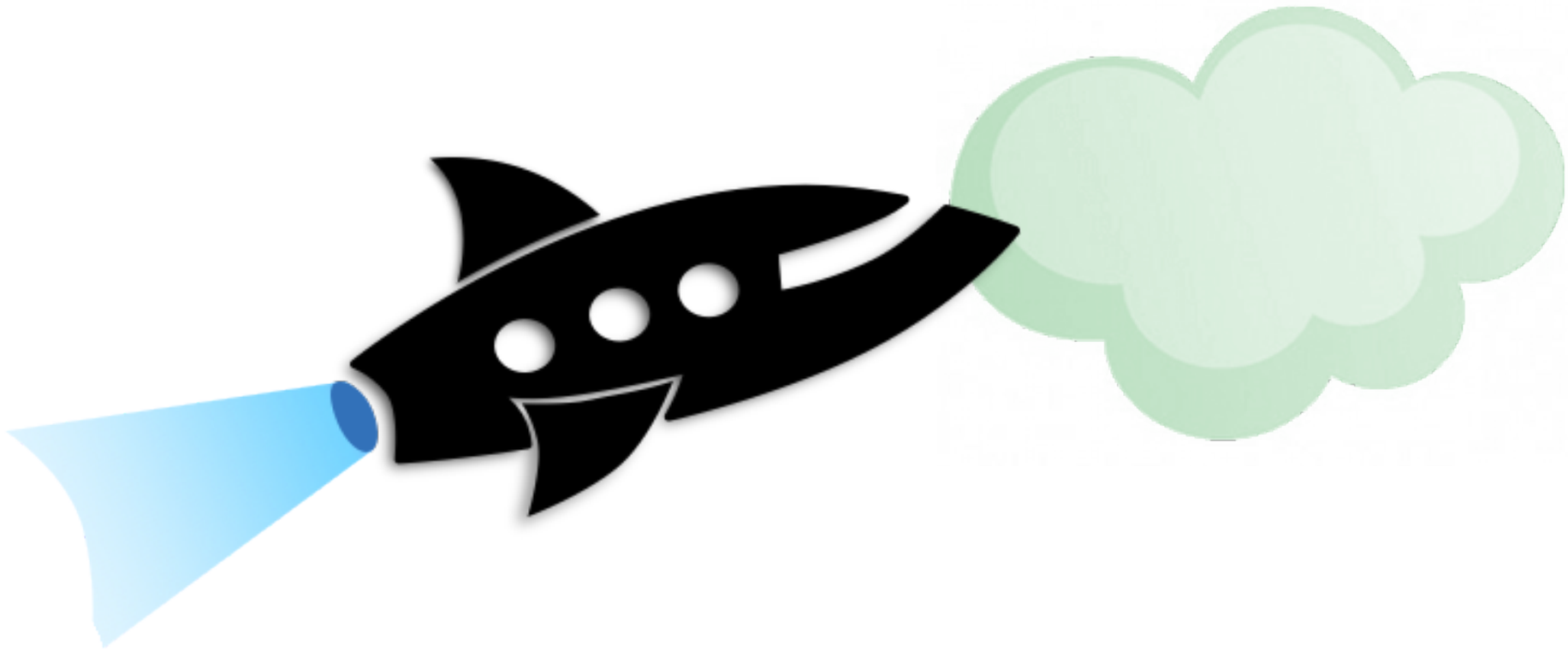
Green Discovery



Ehsan Ahvar, Adrien Lebre, Anne-Cécile Orgerie, and Anthony Simonet

Outline

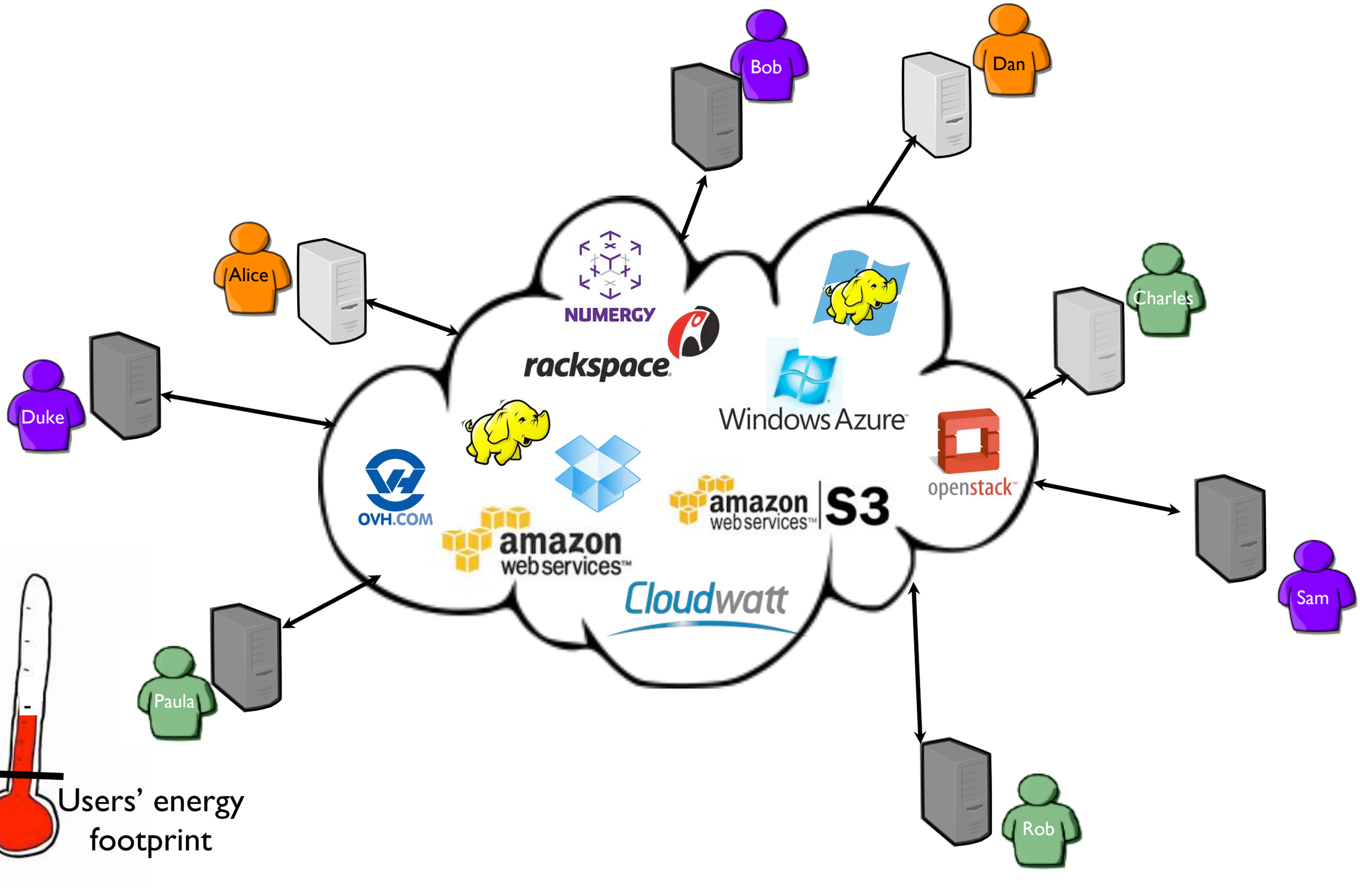
- Context
- Two post-docs
  - 1. OpenStack footprint
  - 2. Cost-benefit analysis centralized vs. distributed Clouds
  - 3. Exploiting renewable energy
- Conclusions

Localization is a key element to deliver ***efficient*** as well as ***sustainable*** **Utility Computing** solutions

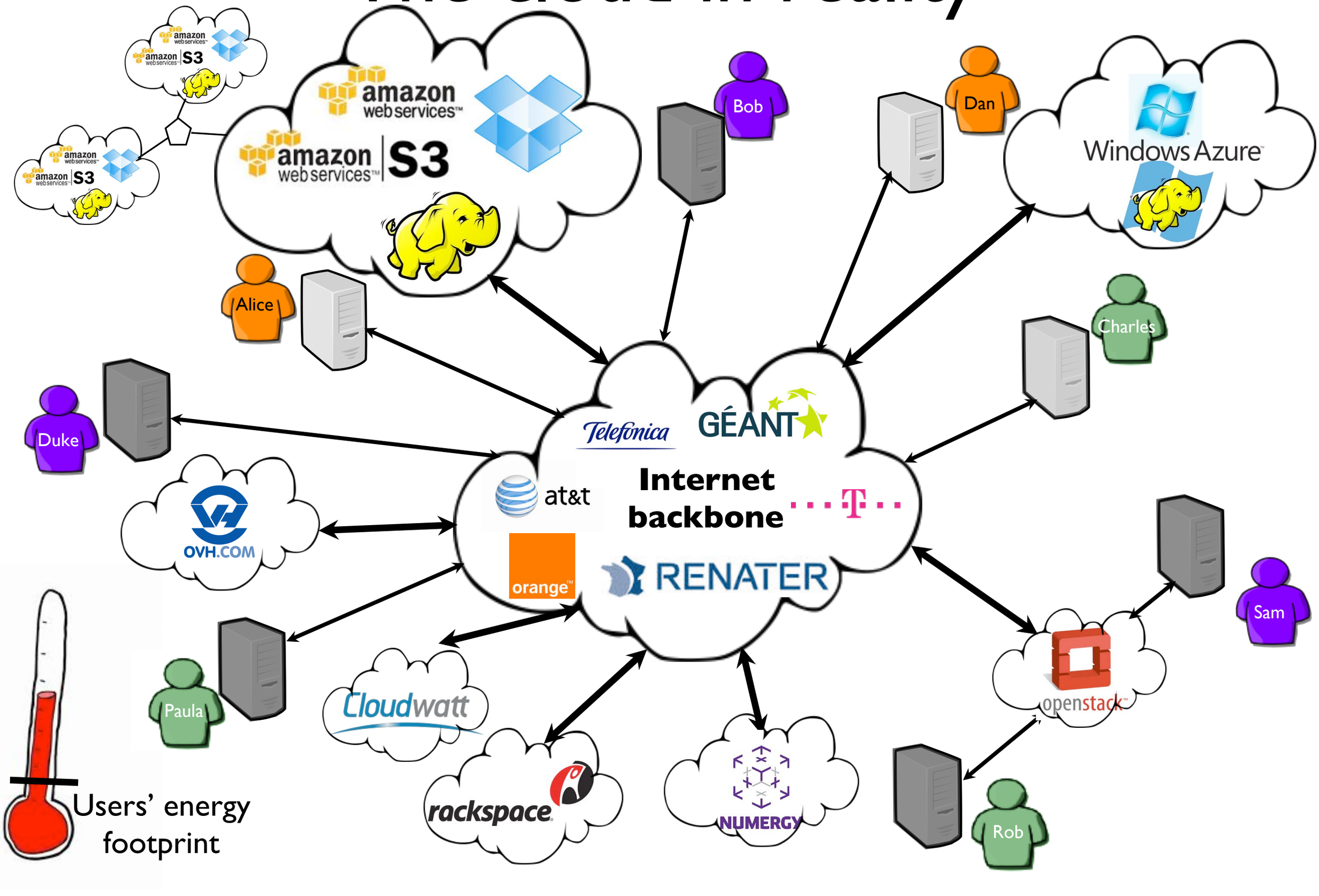
A simple Idea

Make Clouds greener!

The cloud from end-users



The cloud in reality



Google DC (Dalles)

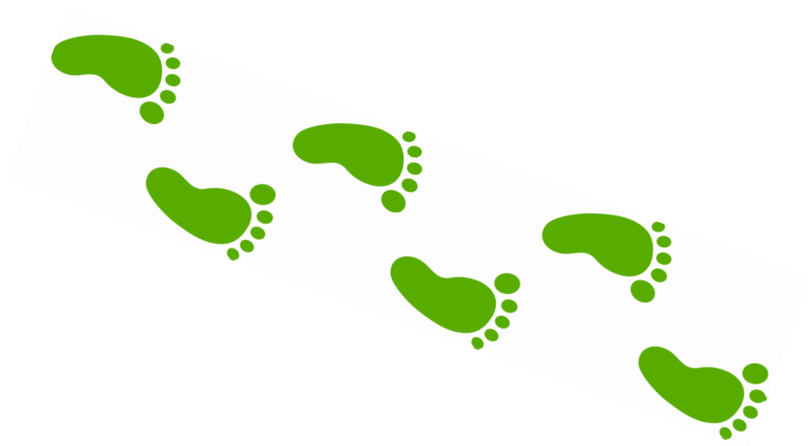
11.5 soccer fields
100 MWatts
100,000 servers

Average PUE: 1.12

https://www.google.com/about/datacenters/inside/locations/the-dalles/

Greener?

Ongoing work



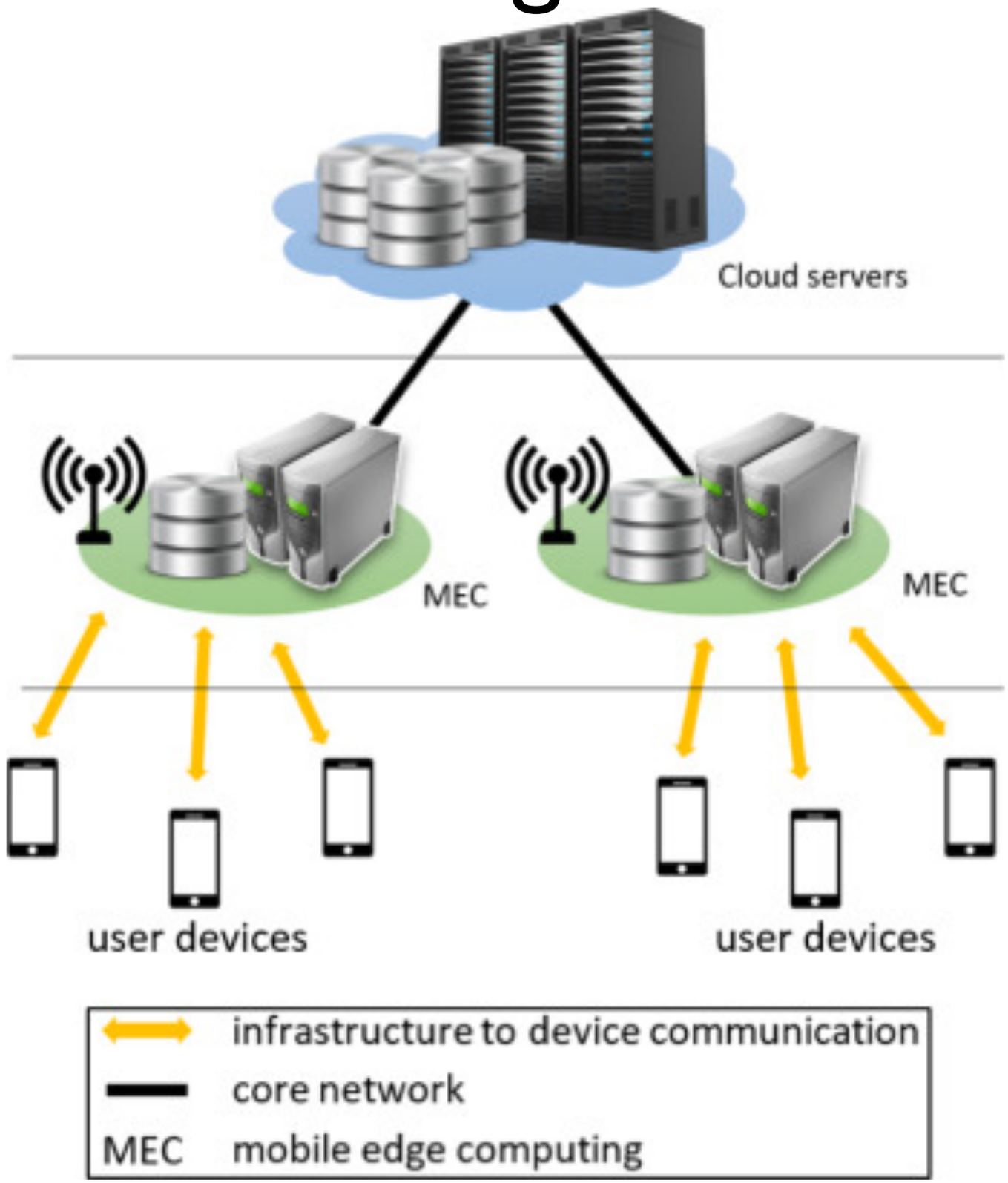
PoP level: hosting additional servers

- Cooling costs for these new servers
  - possibility of free cooling
  - load balancing over several sites to exploit renewable energy
- Energy costs
  - servers' + switches' energy conso ≤ available energy at PoP ?
  - higher kWh prices than in a DC (no wholesale price)
- More network traffic?

➔ Services closer to the users: global compensation? associated costs? More links to deploy?

- Managing costs ?

Edge Cloud



Stability
Availability
Latency

Low latency
Heterogeneity
Low capacity

"*Recovery for overloaded mobile edge computing*", D. Satria et al., FGCS 2017.

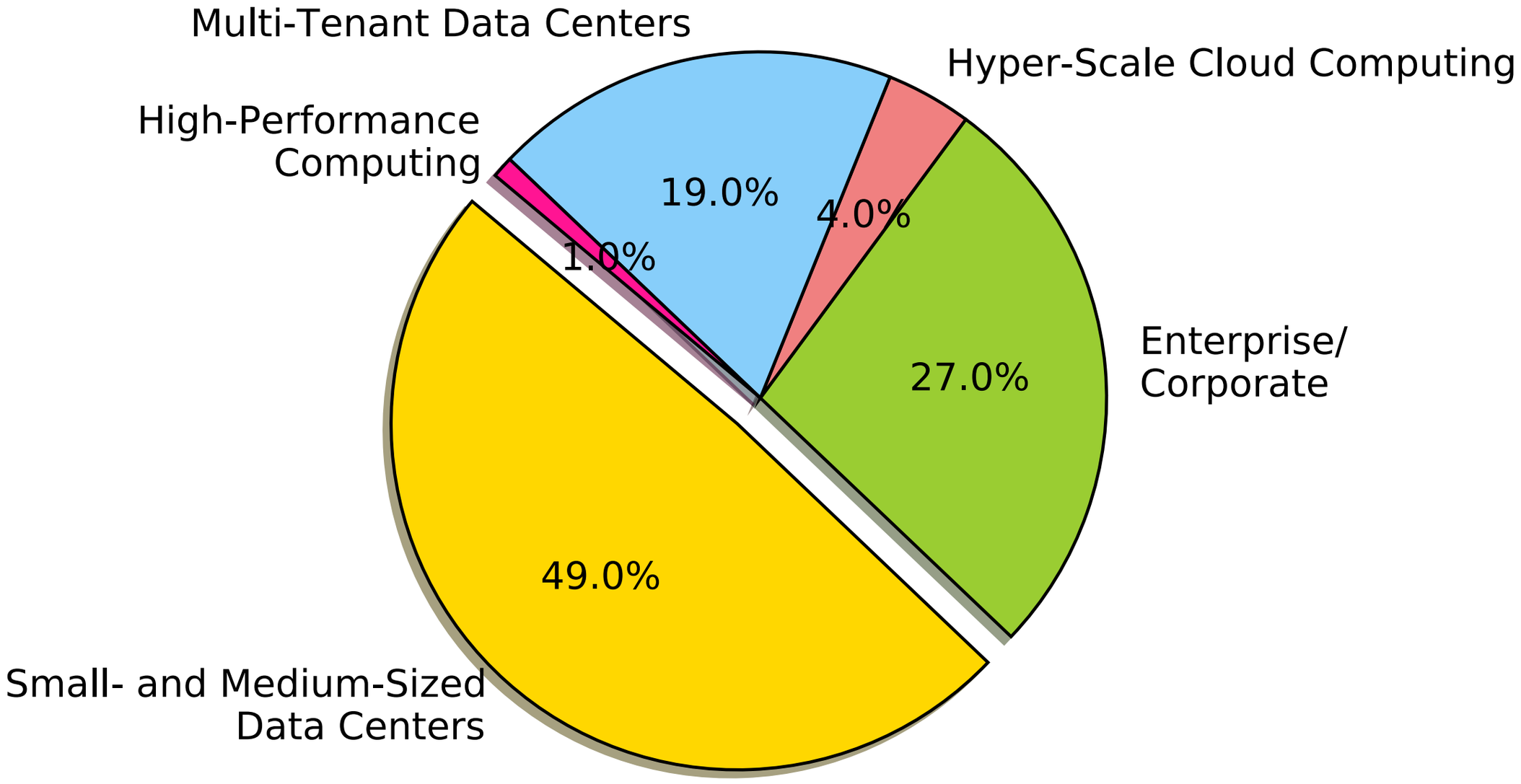
Frightening examples

- Google in 2010 : 900 000 servers, ~ 2 billion kWh
- Facebook in 2012: ~ 532 million kWh
- In 2012, about 509 147 datacenters worldwide, ~ electricity consumption of 30 nuclear power
- By 2017, more than 20 billion devices connected.
- *More than 50 billion in 2020?*

→ Electricity is a big problem in data centers

Source : **J. Koomey**, *Growth in Data Center Electricity Use 2005 to 2010*. Analytics Press, août 2011.

Data center's energy consumption



Estimated U.S. data center electricity consumption by market segment (2011)

Source: Data Center Efficiency Assessment, NRDC White paper, 2014.

ICT's impact

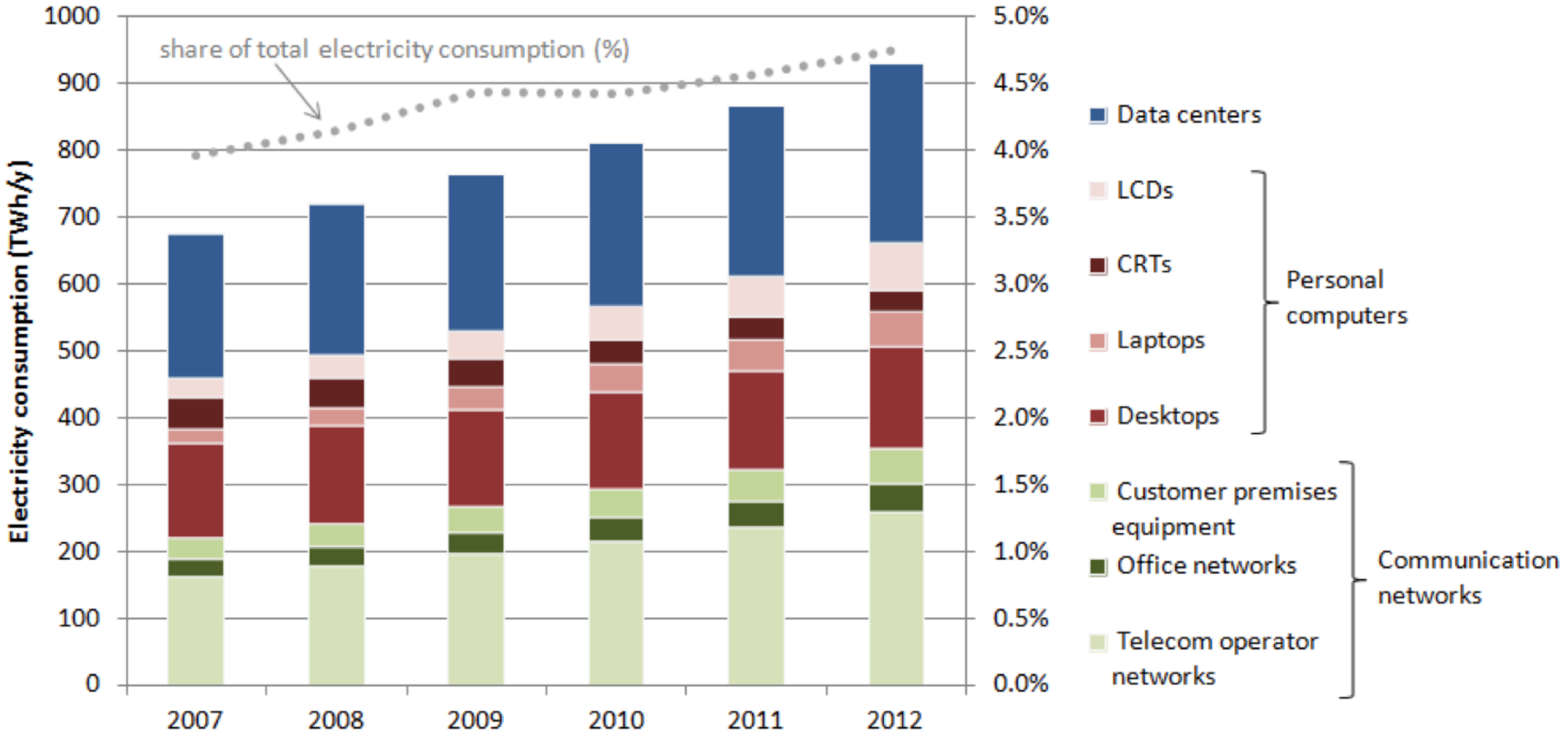


Figure 3-1: Worldwide use phase electricity consumption of communication networks, personal computers and data centers. Their combined share in the total worldwide electricity consumption has grown from about 4% in 2007 to 4.7% in 2012.

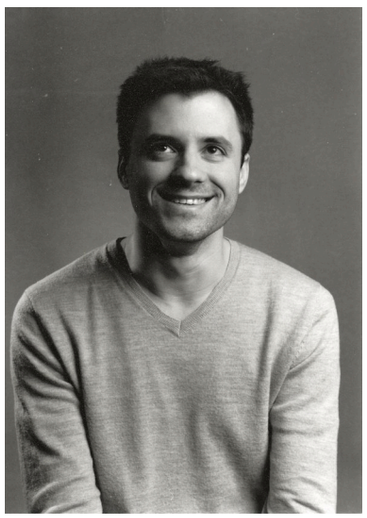
"*Overview of ICT energy consumption (D8.1)*" – Report FP7-2888021, European Network of Excellence in Internet Science, February 2013, The EINS Consortium.

Outline

- Context
- **Two post-docs**
- 1. OpenStack footprint
- 2. Cost-benefit analysis centralized vs. distributed Clouds
- 3. Exploiting renewable energy
- Conclusions

Two Discovery post-docs

Anthony Simonet (Oct. 2015 – Apr. 2017) in Nantes



- Energy footprint of distributing OpenStack
- Cost-benefit analysis of a massively distributed Cloud

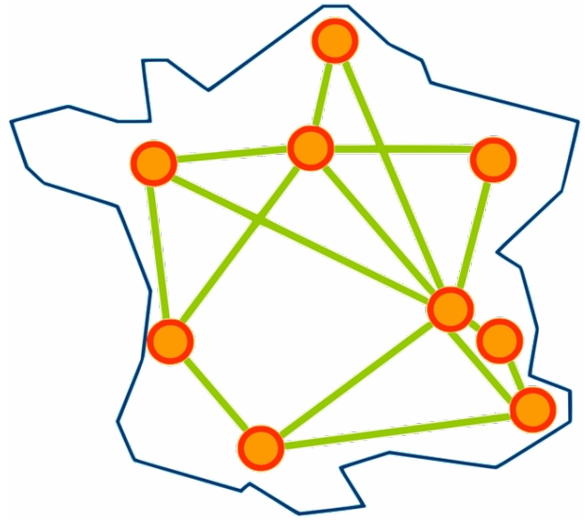
Ehsan Ahvar (June 2017 - …) in Rennes

- Energy cost analysis of an edge cloud
- Leveraging renewable energy to power edge clouds

Outline

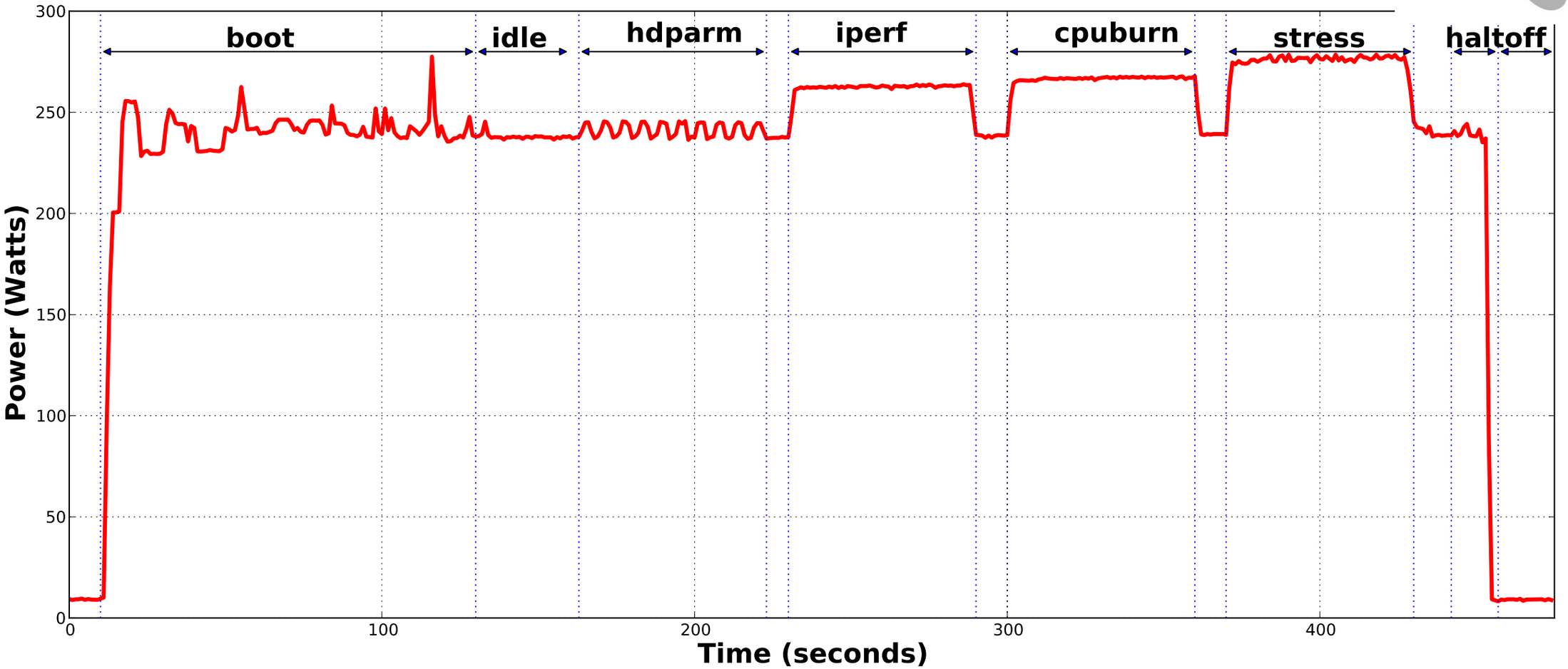
- Context
- Two post-docs
- **1. OpenStack footprint**
- 2. Cost-benefit analysis centralized vs. distributed Clouds
- 3. Exploiting renewable energy
- Conclusions

Wattmeters

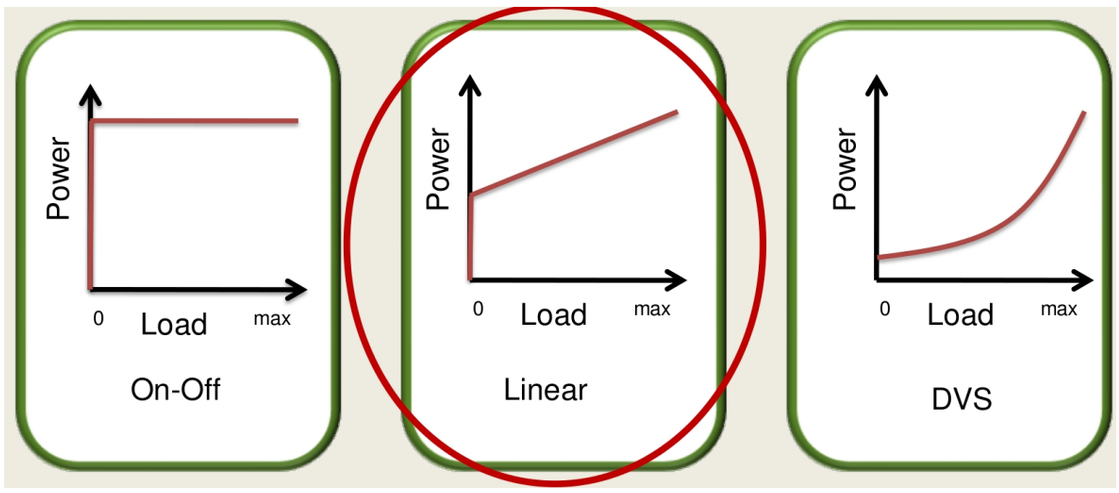


Grid'5000

Server power consumption

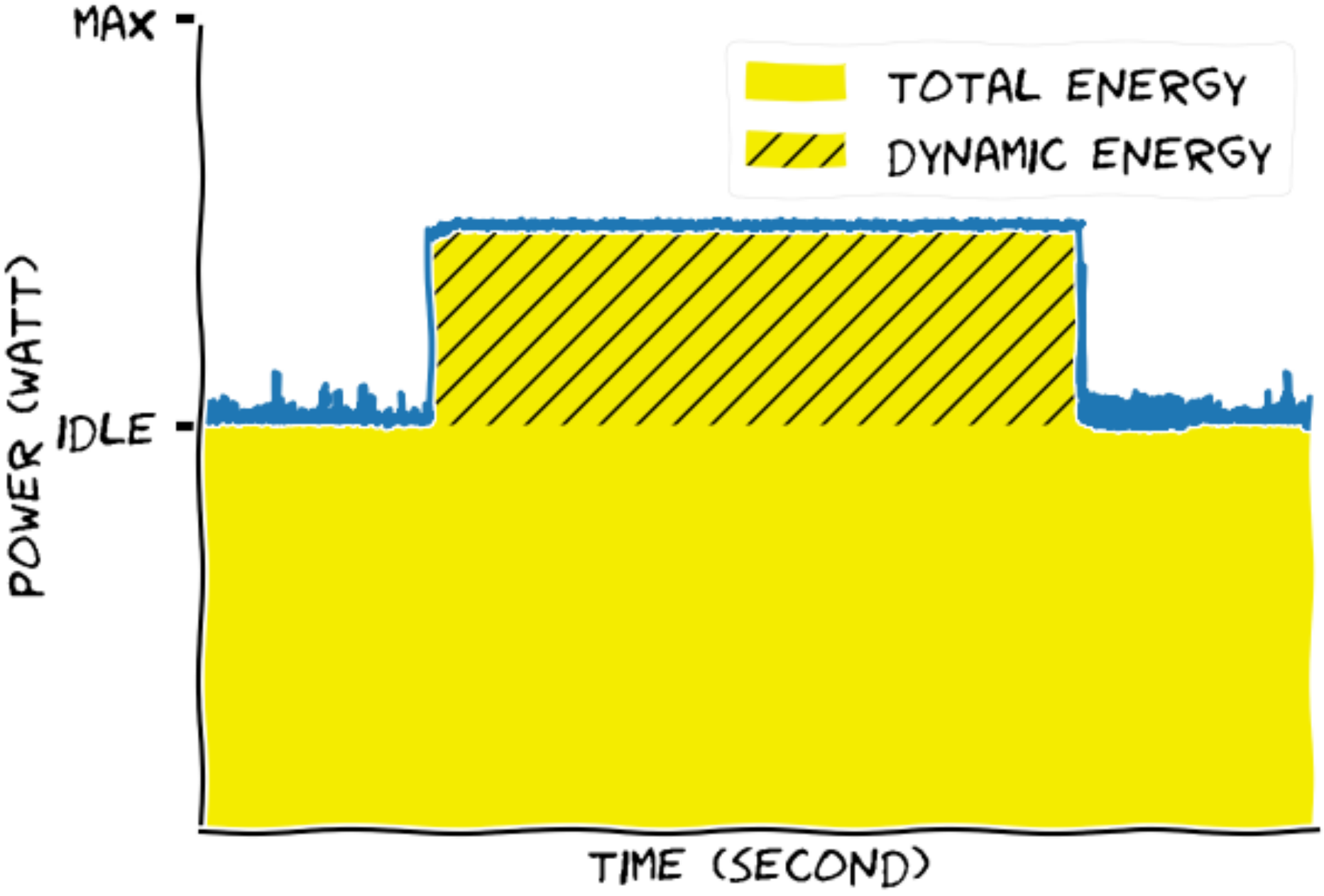


- 'Useless' consumption when powered on: 50 to 90%
- Processors: main consumers when in use

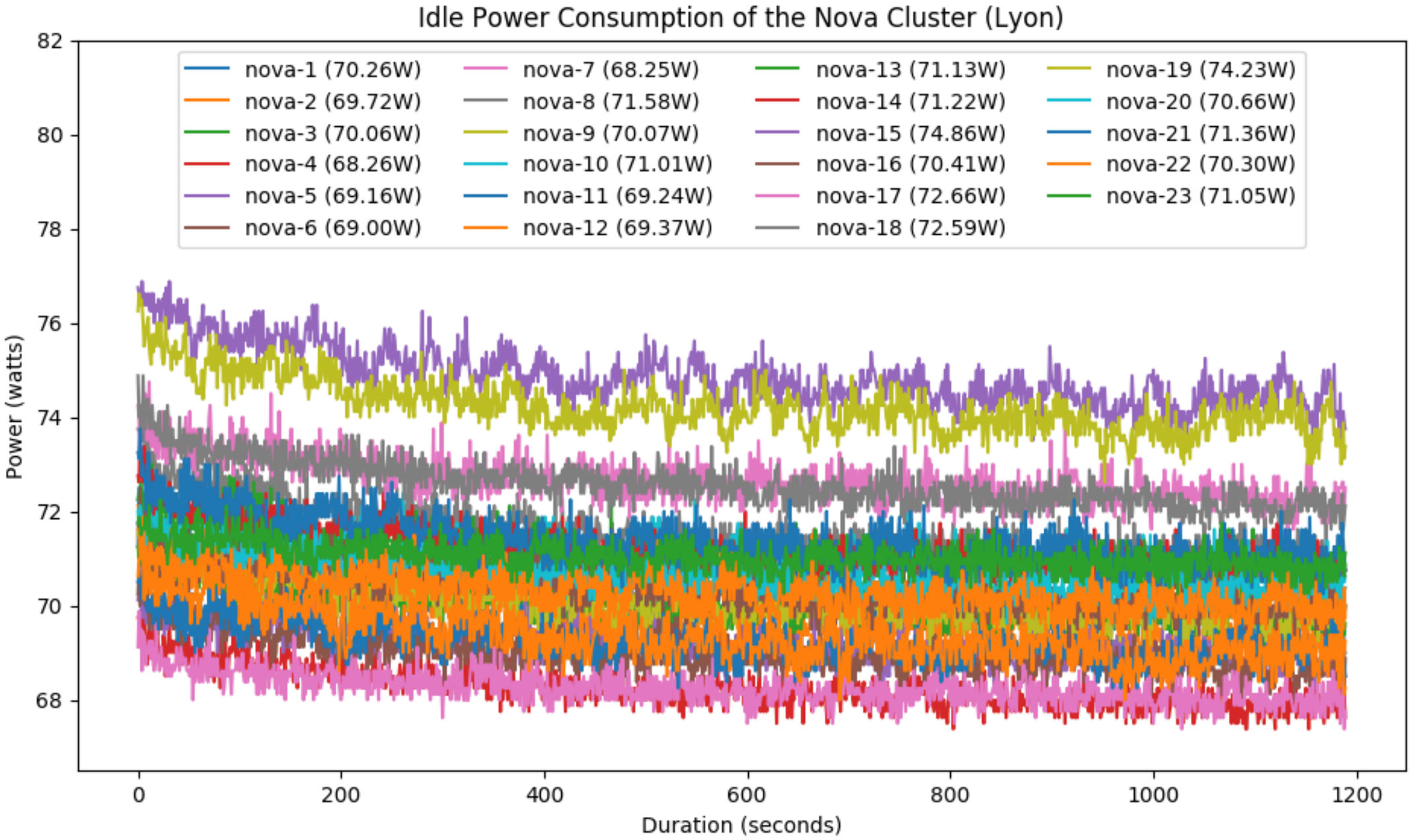


"*The Green Grid'5000: Instrumenting a Grid with Energy Sensors*", M. Dias de Assunção, J.-P. Gelas, L. Lefèvre and A.-C. Orgerie, INGRID 2010.

Idle power consumption

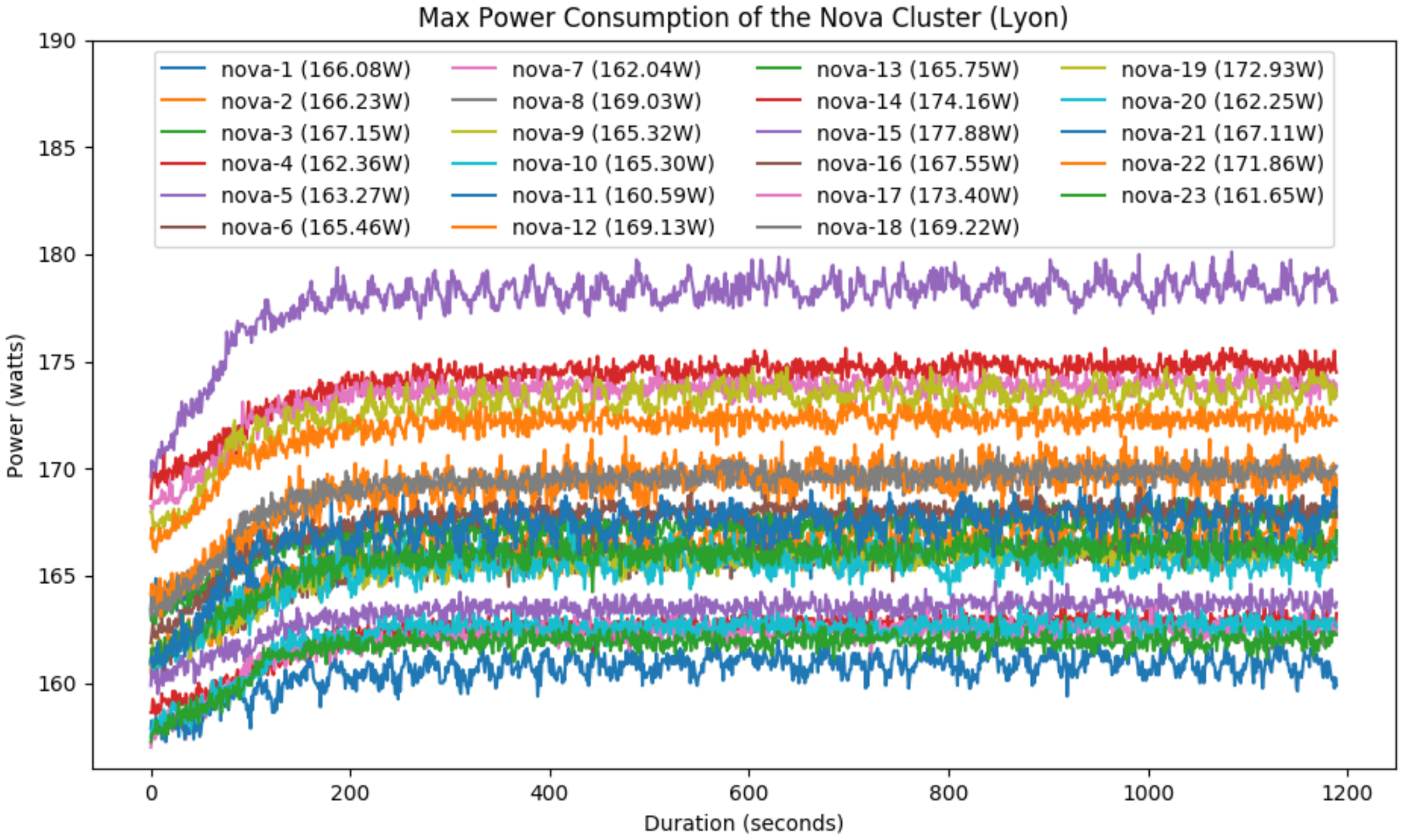


Homogeneous servers



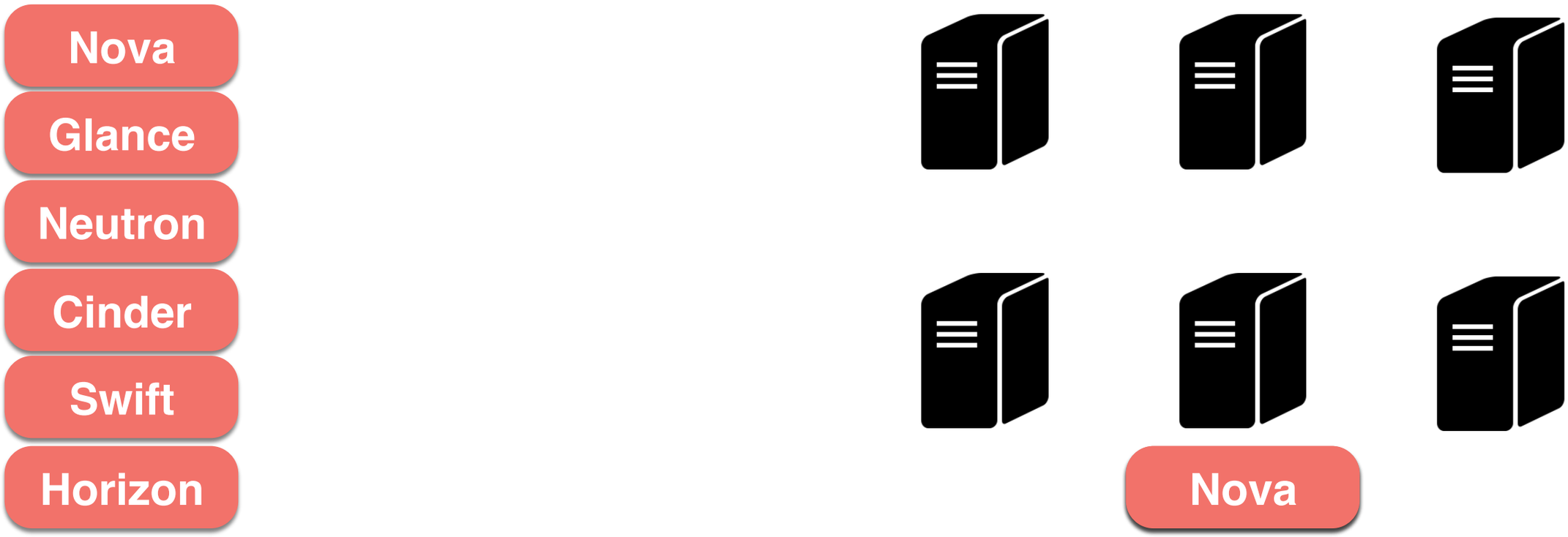
Courtesy of David Guyon

Homogeneous servers



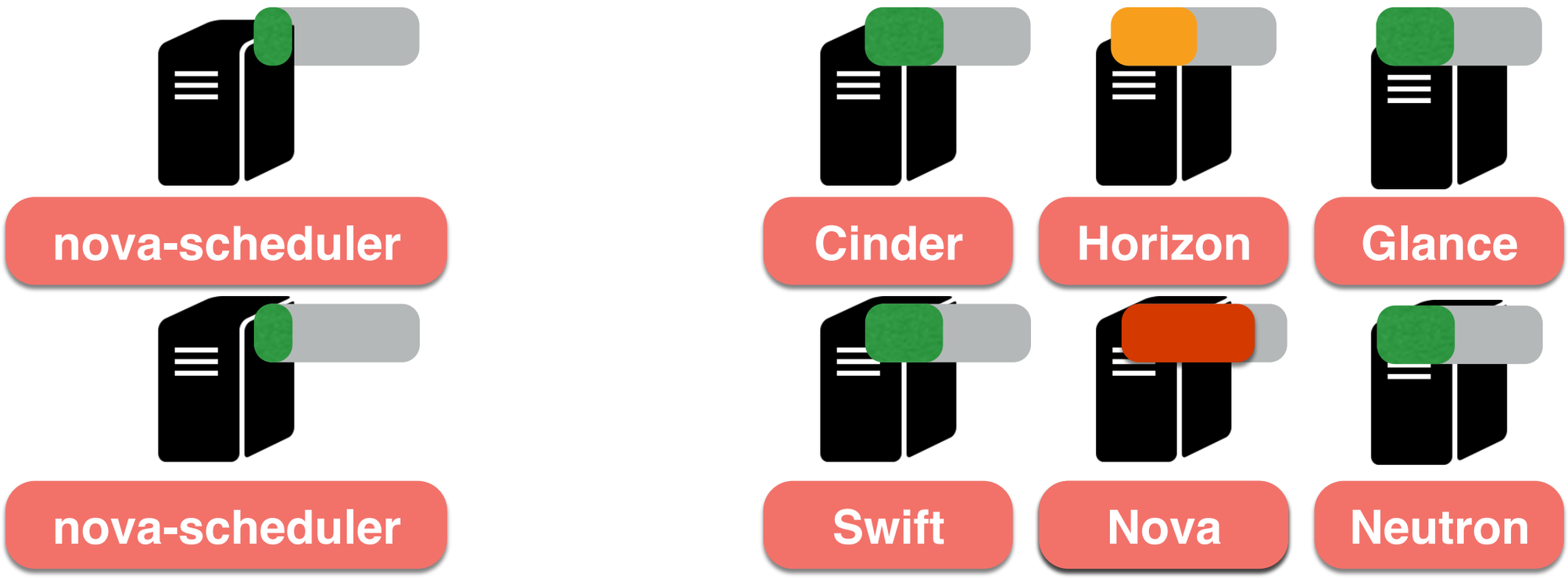
Courtesy of David Guyon

Many Ways to Deploy OpenStack



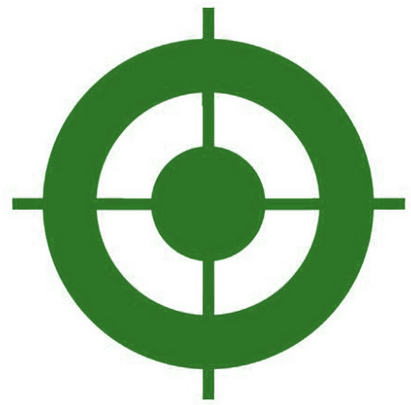
with consequences on resources utilisation and energy consumption

Example: nova-scheduler



Pinned to one core, leaves the other cores unused

Goals

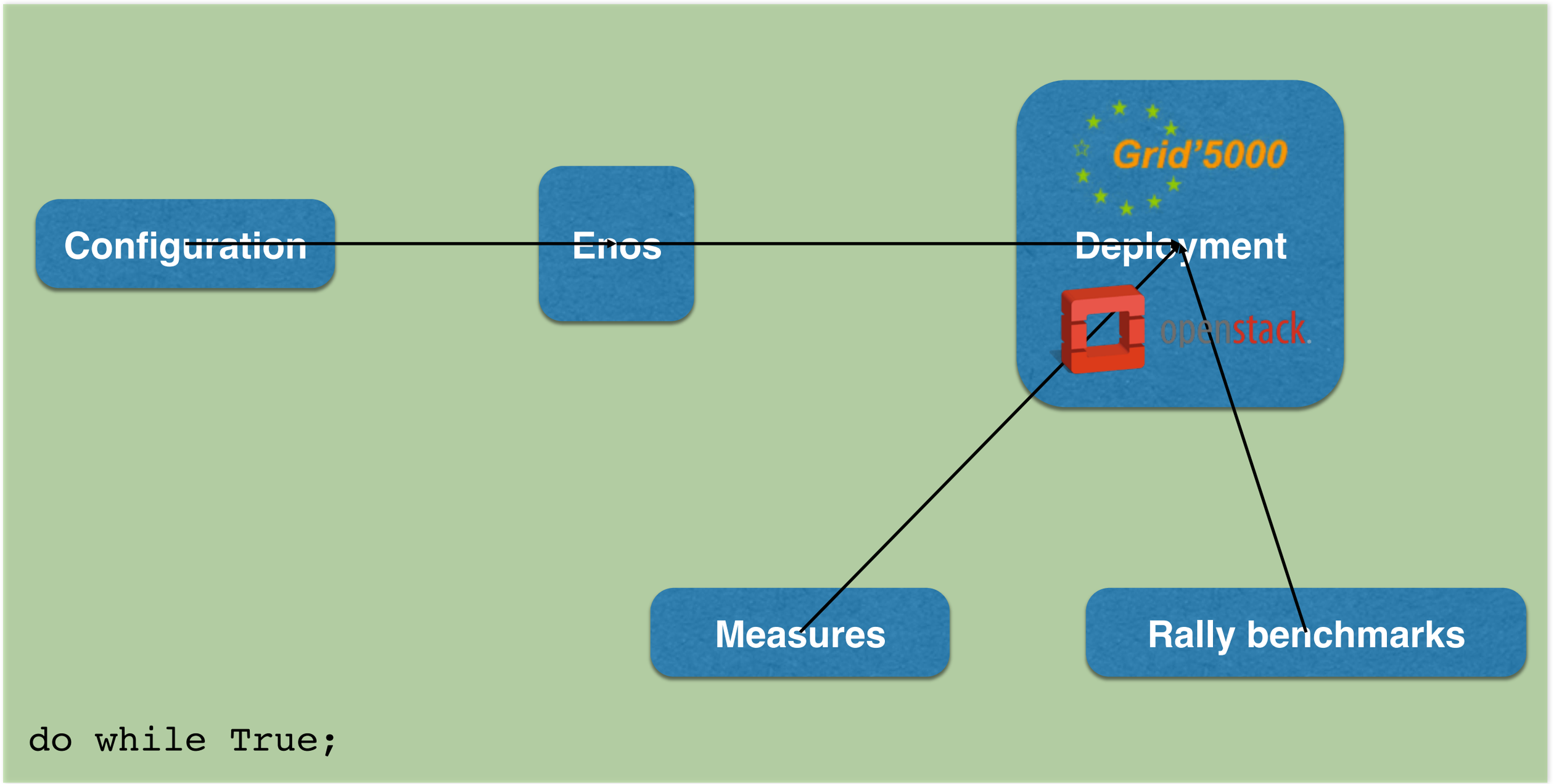


- Measuring the energy consumption of **support services** during common operations (*booting a VM, creating a VNET, allocating a VIP, migrating a VM, etc.*) when varying the number of compute nodes ([1, 800])

- Providing a model for the energy footprint of OpenStack

- Providing guidelines for energy-aware OpenStack operations

Methodology




```
do while True;
```

OpenStack Idle Energy Consumption

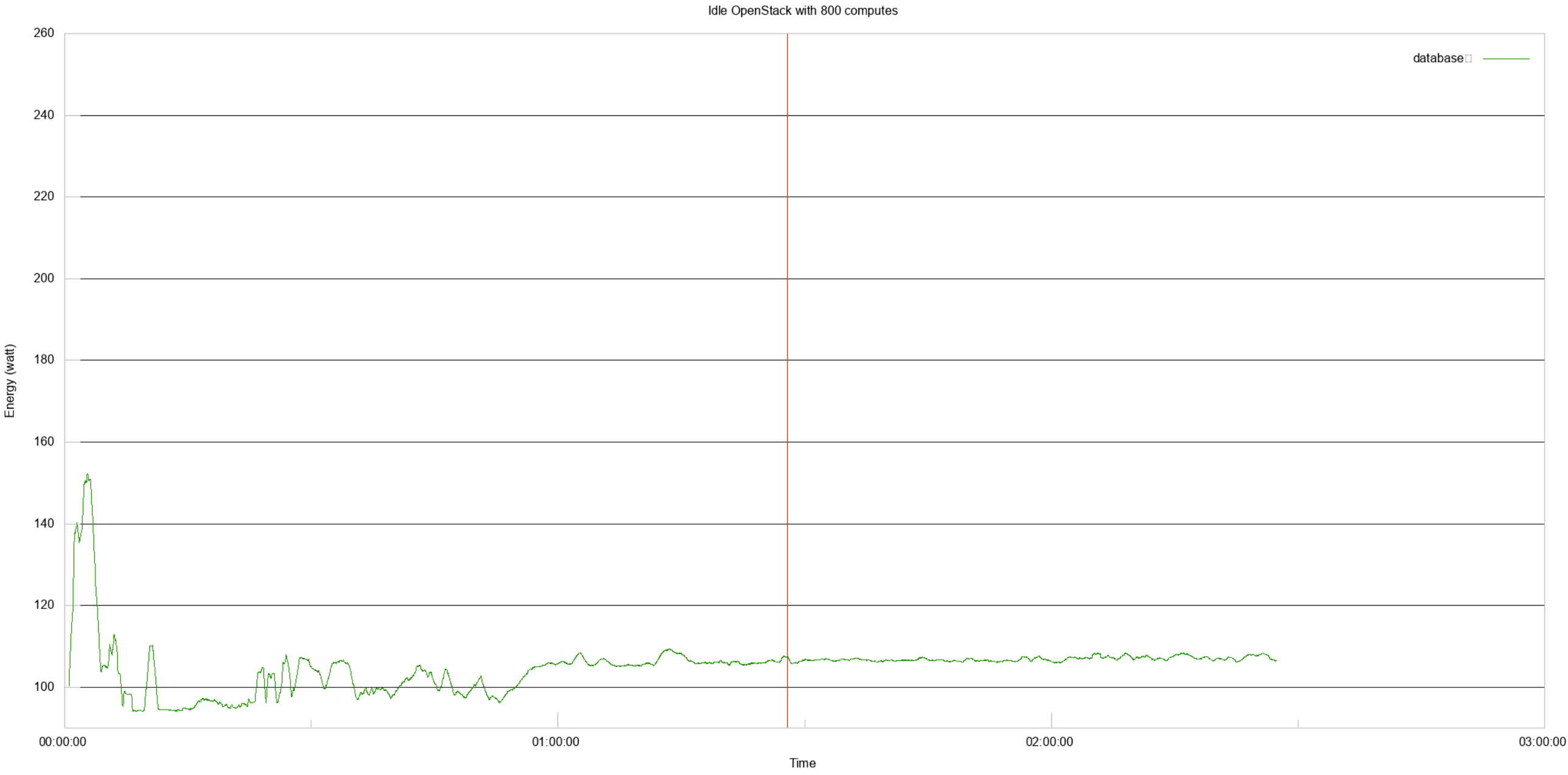
- Measure of all services **except compute**
- On the **Taurus** cluster of Grid'5000
- Some services isolated on a specific node
  - Glance, Cinder, Neutron, Horizon, MariaDB, nova-scheduler
  - (RabbitMQ, HAProxy)
- Remaining services on a *controller* node
- 1 to **800** (fake) compute nodes

Methodology

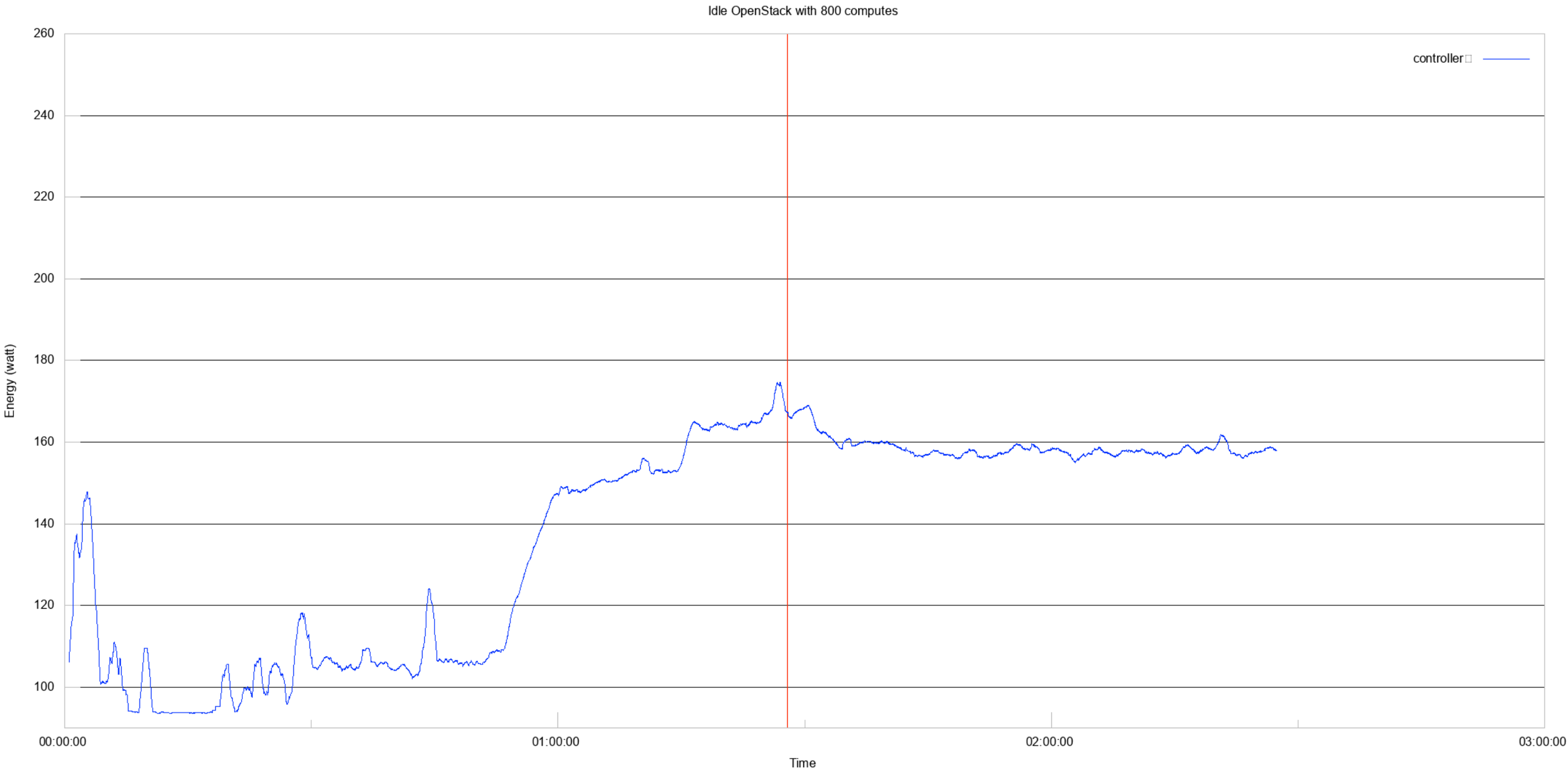
For n in {1, 10, 100, …, 800} compute nodes

1. Deploy OpenStack
2. **Do nothing for one hour**

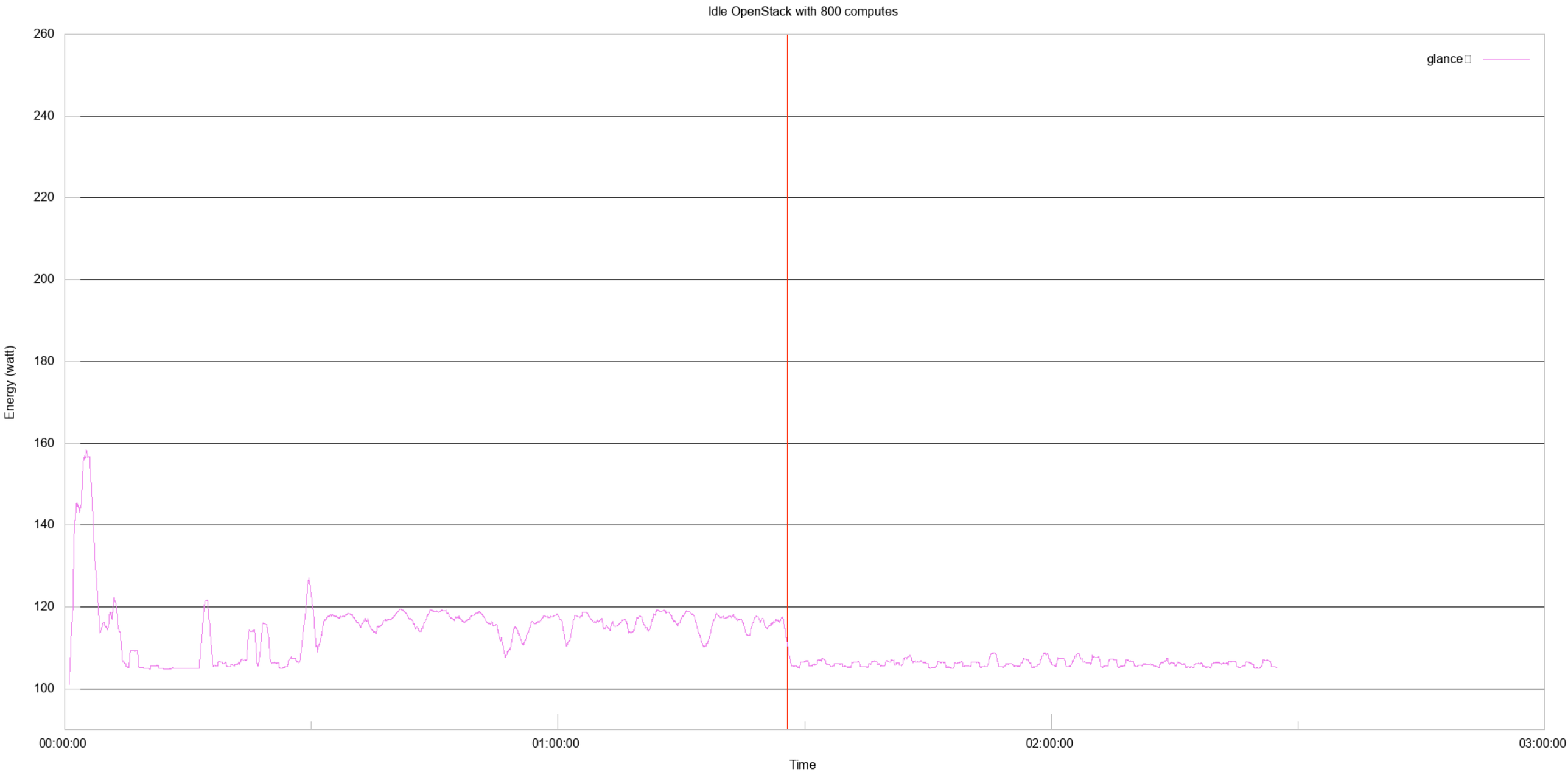
Idle Energy Consumption



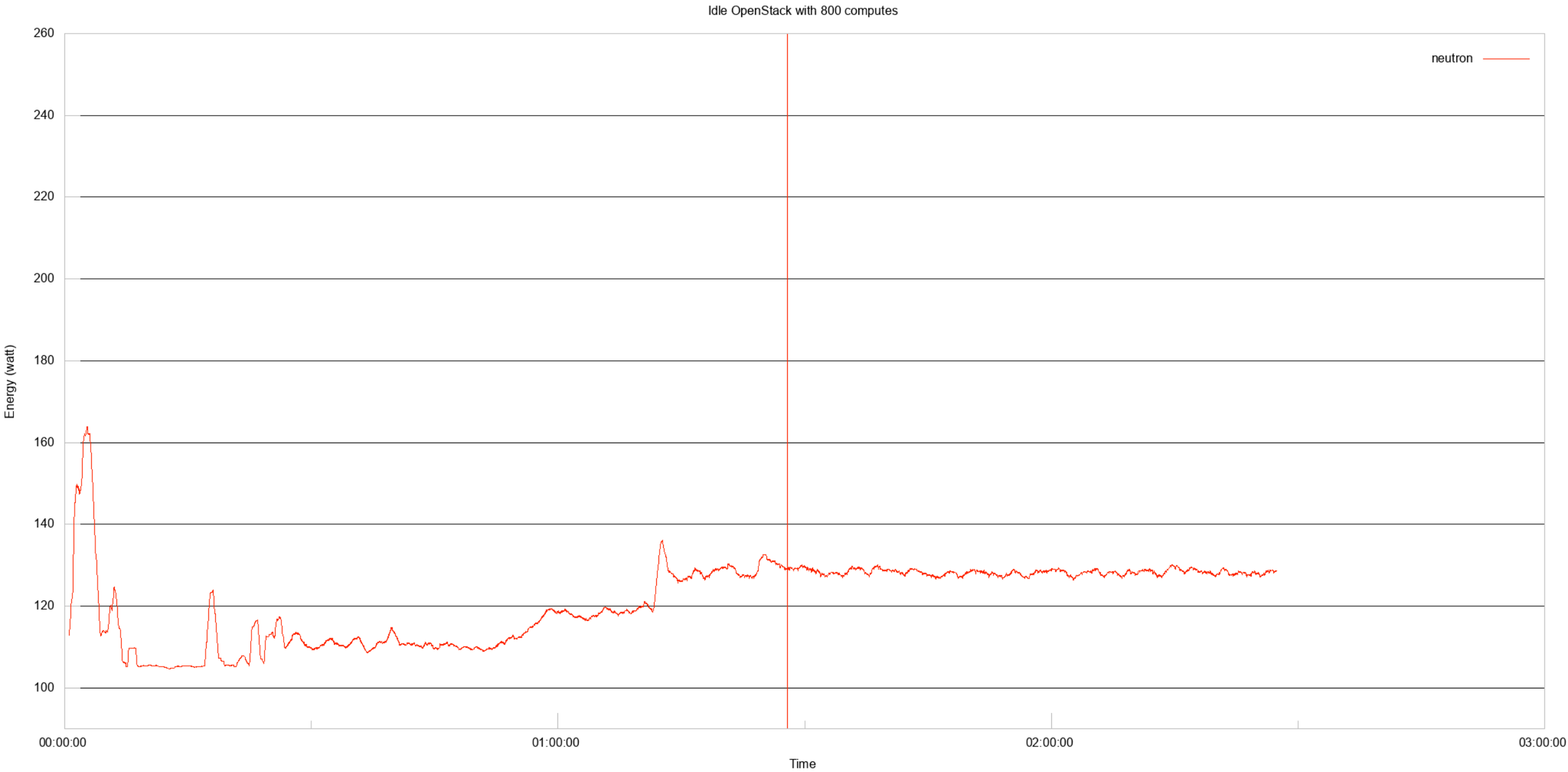
Idle Energy Consumption



Idle Energy Consumption



Idle Energy Consumption



Idle Energy Consumption

Difficulties

- Complex deployments, frequent failures for various reasons
- Variability idle measurements (sudden change of consumption with no charge)

→ experimental work

Rally Benchmarks

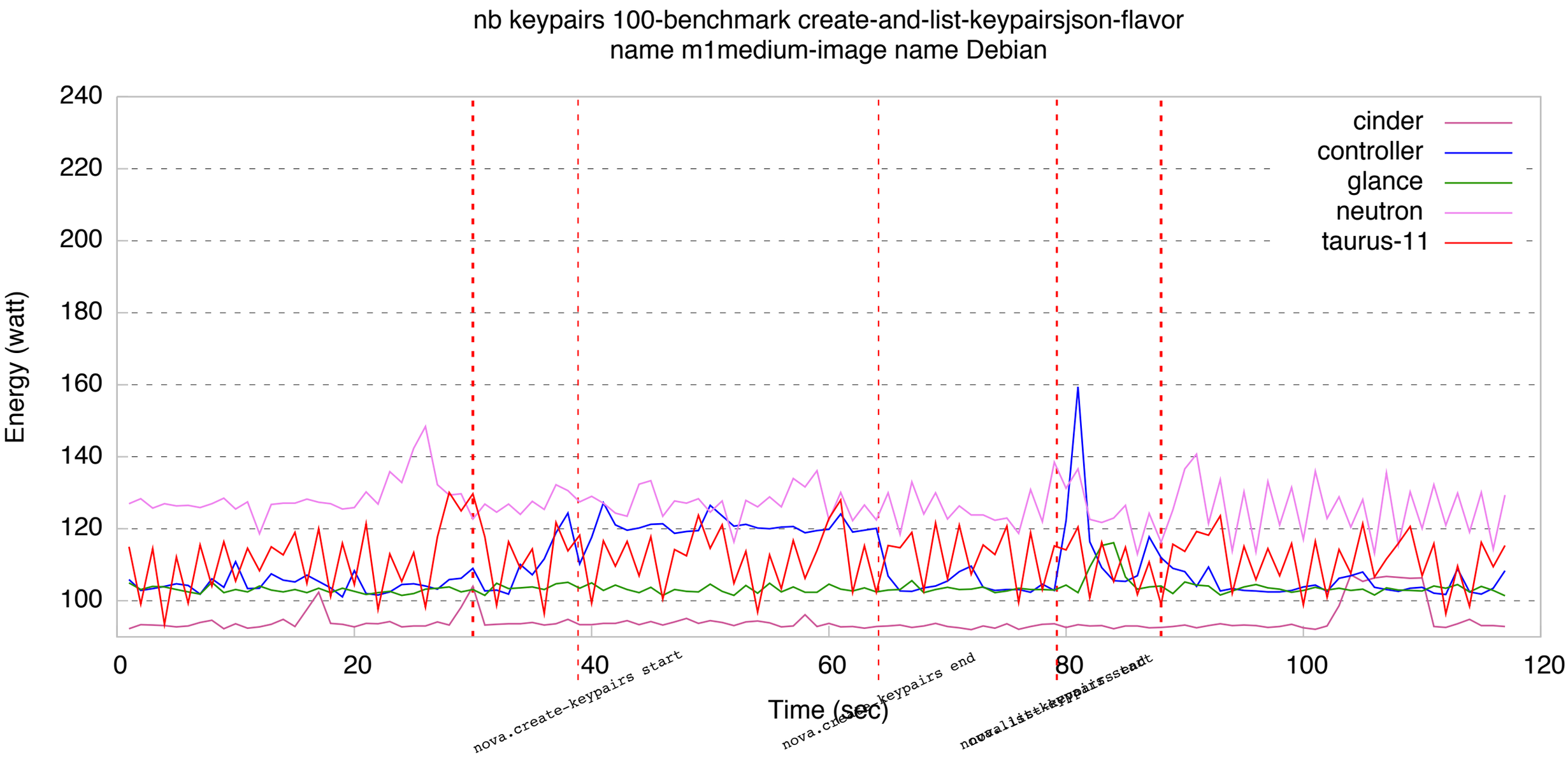
- The official benchmark suite for OpenStack
- Still on Taurus
- Isolated services: glance, cinder, neutron
- 1 controller
- 1 (actual) compute

Methodology

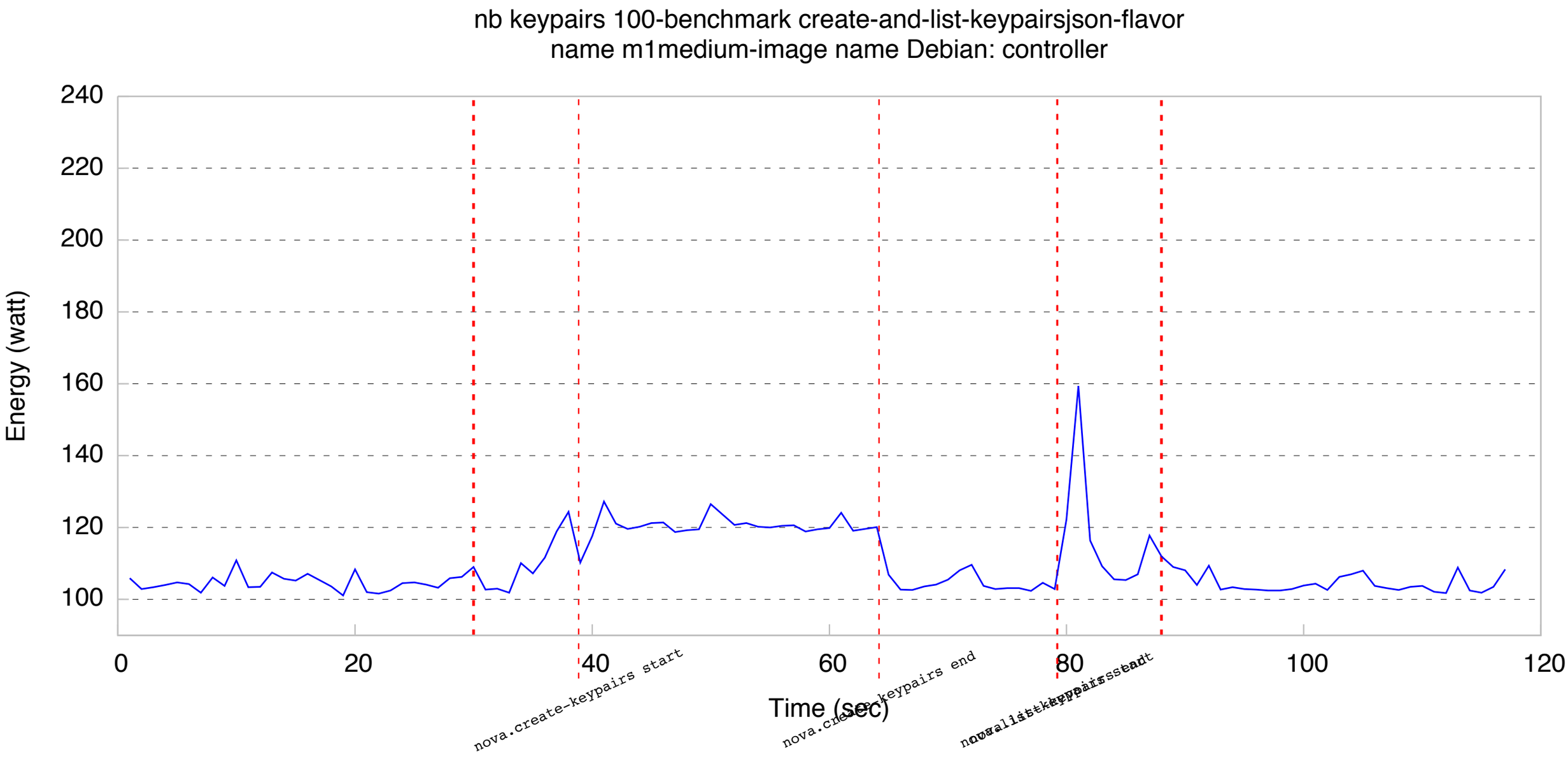
1. Deploy OpenStack
2. For each rally benchmark

- Pause for 2 minutes
- **Execute Rally benchmark**
- Pause for 2 minutes

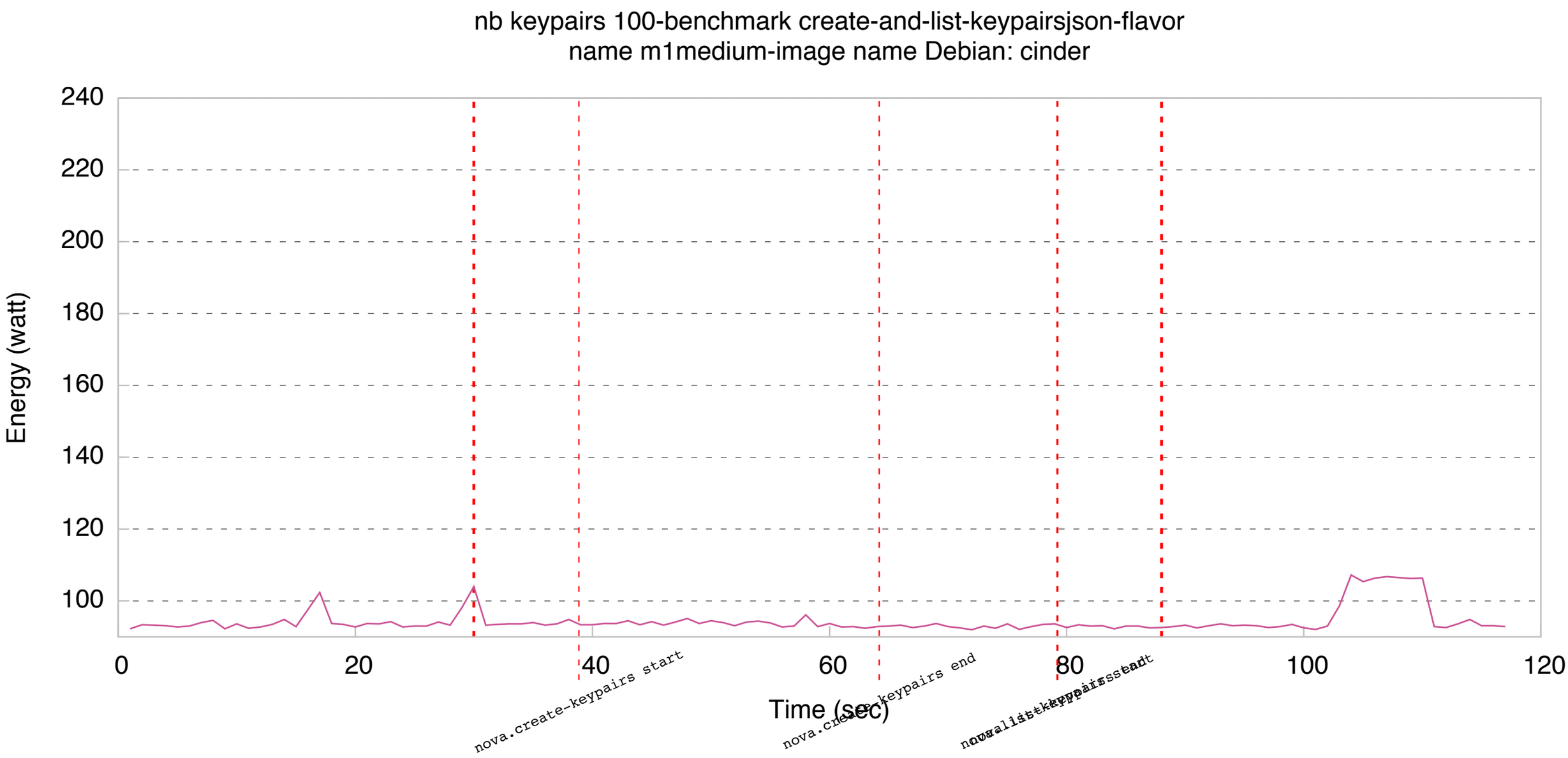
Rally benchmark



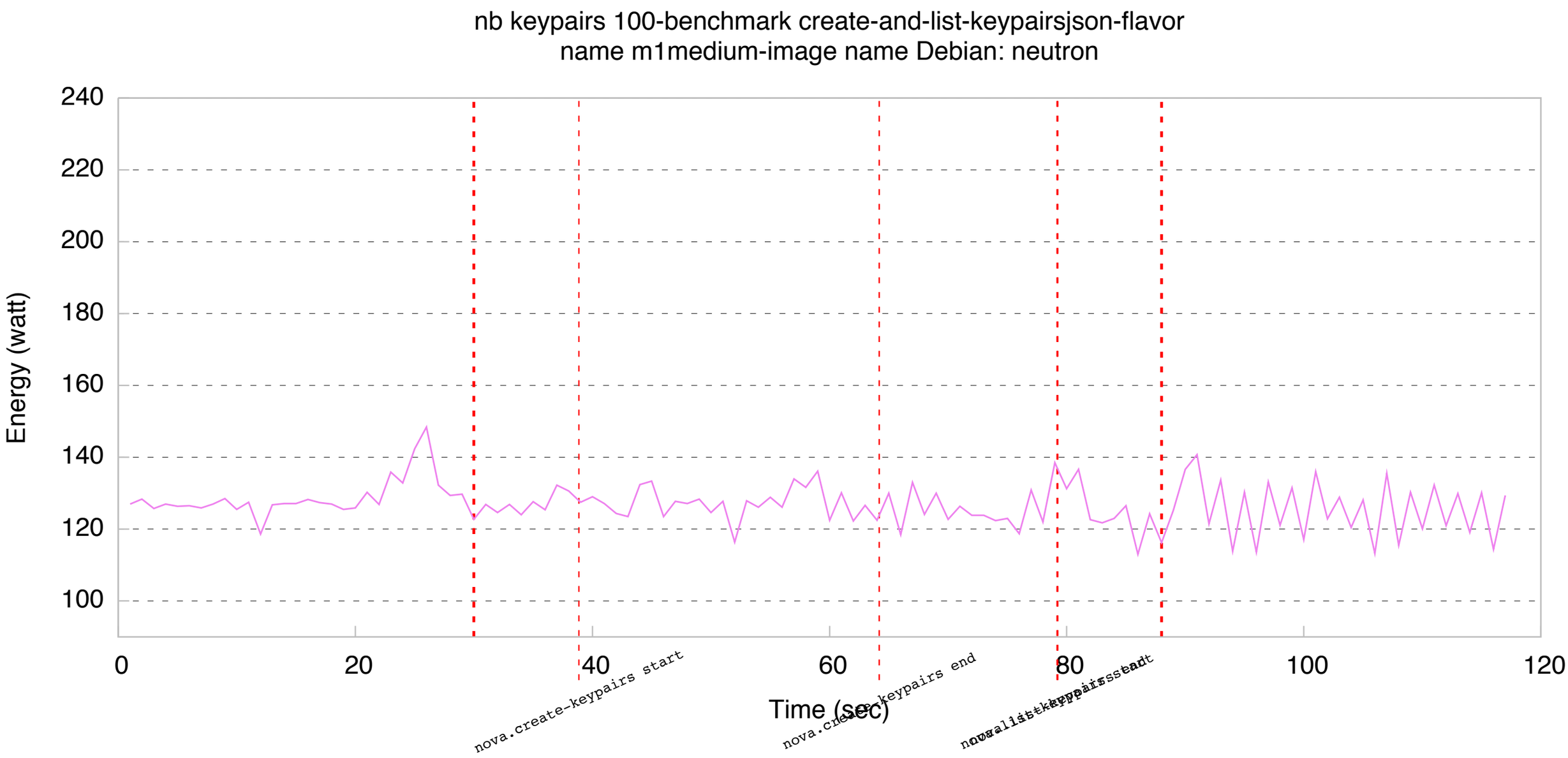
Rally benchmark



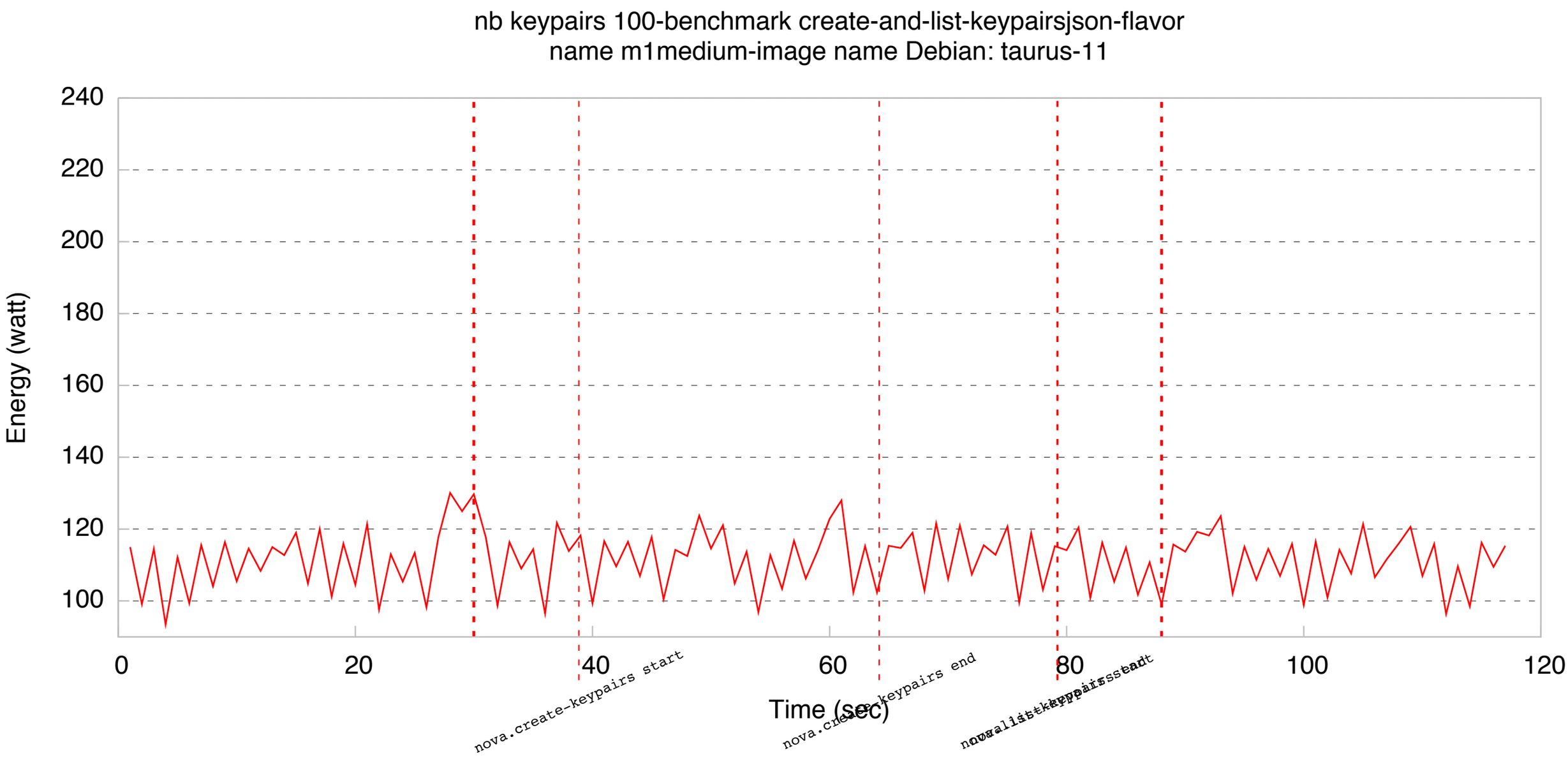
Rally benchmark



Rally benchmark



Rally benchmark



Rally benchmarks

Difficulties

- Precision of measures (precision of wattmeters and Kwapi)
- Very short time for measuring most operations (only a couple seconds)
- Difficulty to get the exact timing of operations

Goals

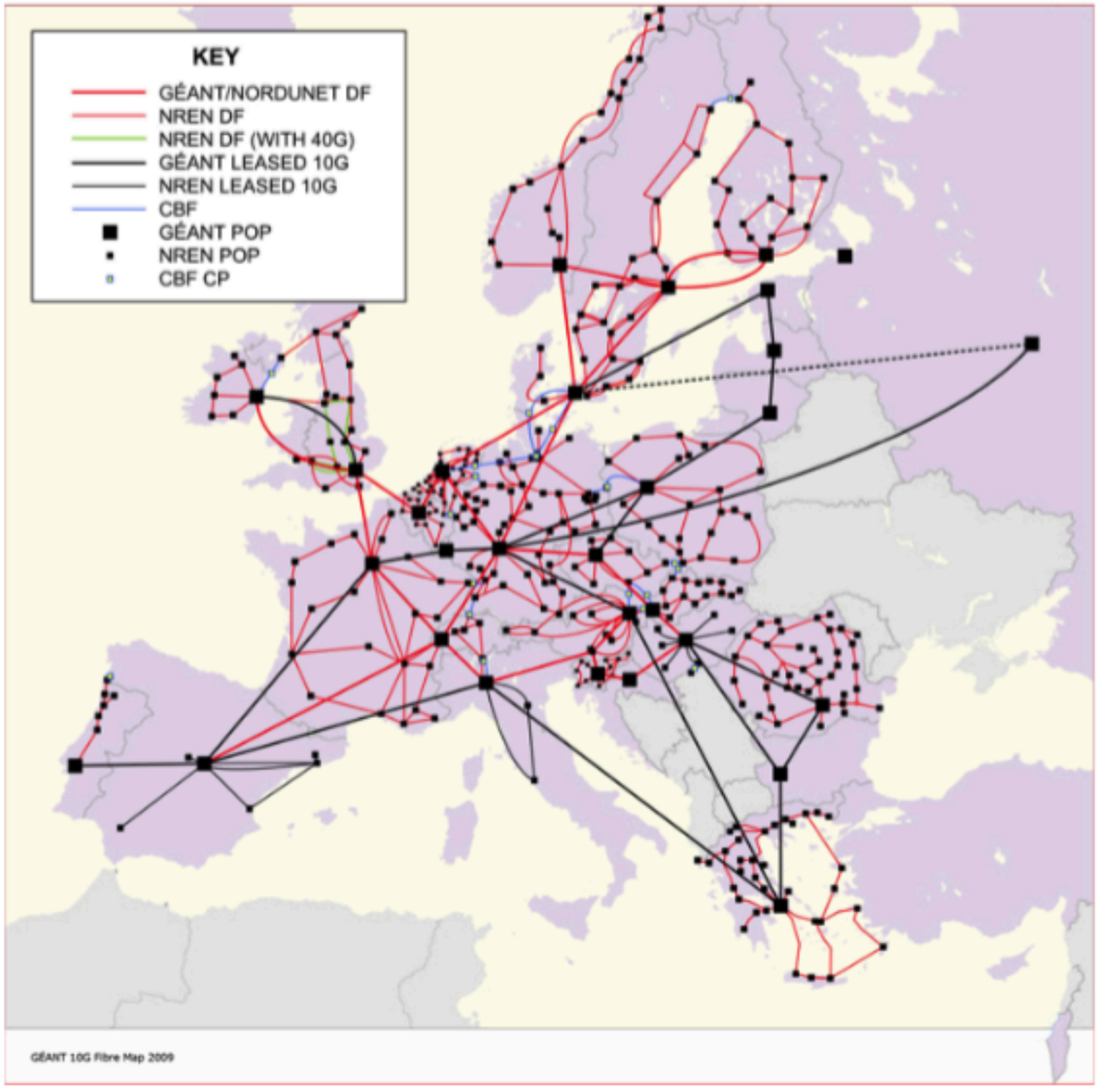
- Study how energy consumption scales with the number of compute nodes and operations
- Isolate more services (nova-schedulers, RabbitMQ & HAProxy)

And so?

Outline

- Context
- Two post-docs
- 1. OpenStack footprint
- **2. Cost-benefit analysis centralized vs. distributed Clouds**
- 3. Exploiting renewable energy
- Conclusions

Distributed infrastructures



Pros:

- Closer to users (legislation and latency friendly)
- DCs easier to deploy
- Reduces usage of backbone network
- Favor use of local renewable energy

Cons:

- More difficult to operate

Figure: The Géant network

Operating distributed clouds?

- Will it work?
- **Who would/could do that?**
  - Cloud computing providers
  - Internet service providers
  - Telecom Operators
  - New players
- What would the infrastructure look like?
- **Would it be economically viable?**

Classification of actors

Level of control	Outsourced → Controlled				
	1	2	3	4	5
Inter-DC network	✗	✗	✓	✗	✓
Compute/storage/ intra-DC network	✗	✓	✓	✓	✓
Building	✗	✗	✗	✓	✓

Table: Level of control on elements of the infrastructure, from fully outsourced to fully controlled.

"Deploying Distributed Cloud Infrastructures: Who and at What Cost?"
Adrien Lebre, Anthony Simonet and Anne-Cécile Orgerie, Workshop InterCloud, April 2016.

Cost model

Cost model with 7 categories of cost:

1. Servers
2. Storage
3. Network
4. Power
5. Cooling
6. Building
7. Maintenance

Instantiation

	Facilities	Servers	Storage	Network	Power	Cooling	Maintenance
Outsourced	0€	1,400€	320€	0€	0€	0€	1,000€
Outsourced+BW	0€	1,400€	320€	3,750€	0€	0€	1,000€
NREN	44€	525€	45€	315€	405€	363€	1,500€
Telecom Operator	44€	525€	45€	1354€	405€	363€	1,500€

Table: Monthly estimated cost for operating a DC in a DCC infrastructure.

	Total	**Cost/VM**
Outsourced	**2,720€**	**26.15€**
Outsourced+BW	**6,470€**	**62.21€**
NREN	**3,197€**	**29.60€**
Telecom Operator	**4,236€**	**39.22€**

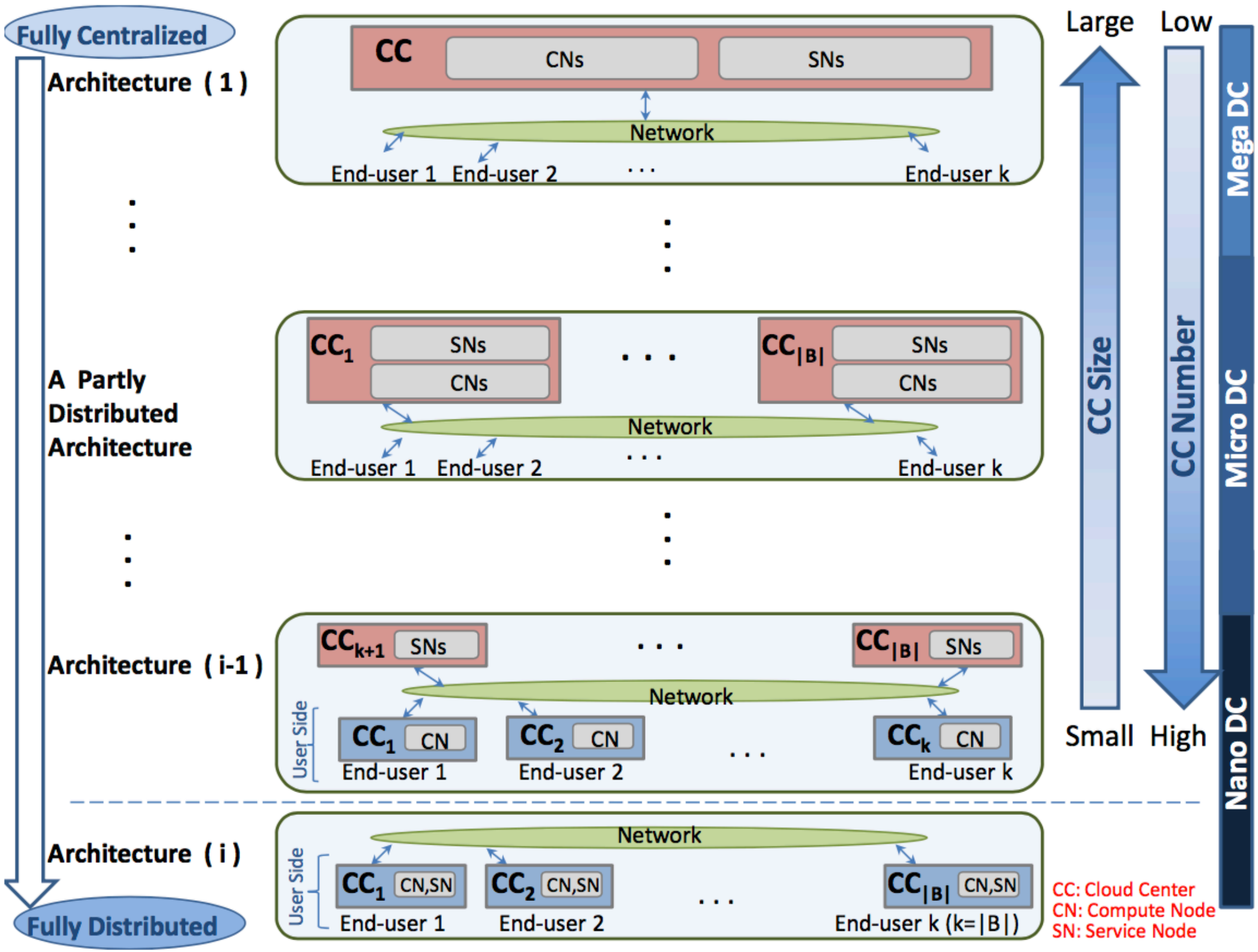
Table: Total cost and comparison of the cost of a VM with Amazon's 54€.

Conclusions

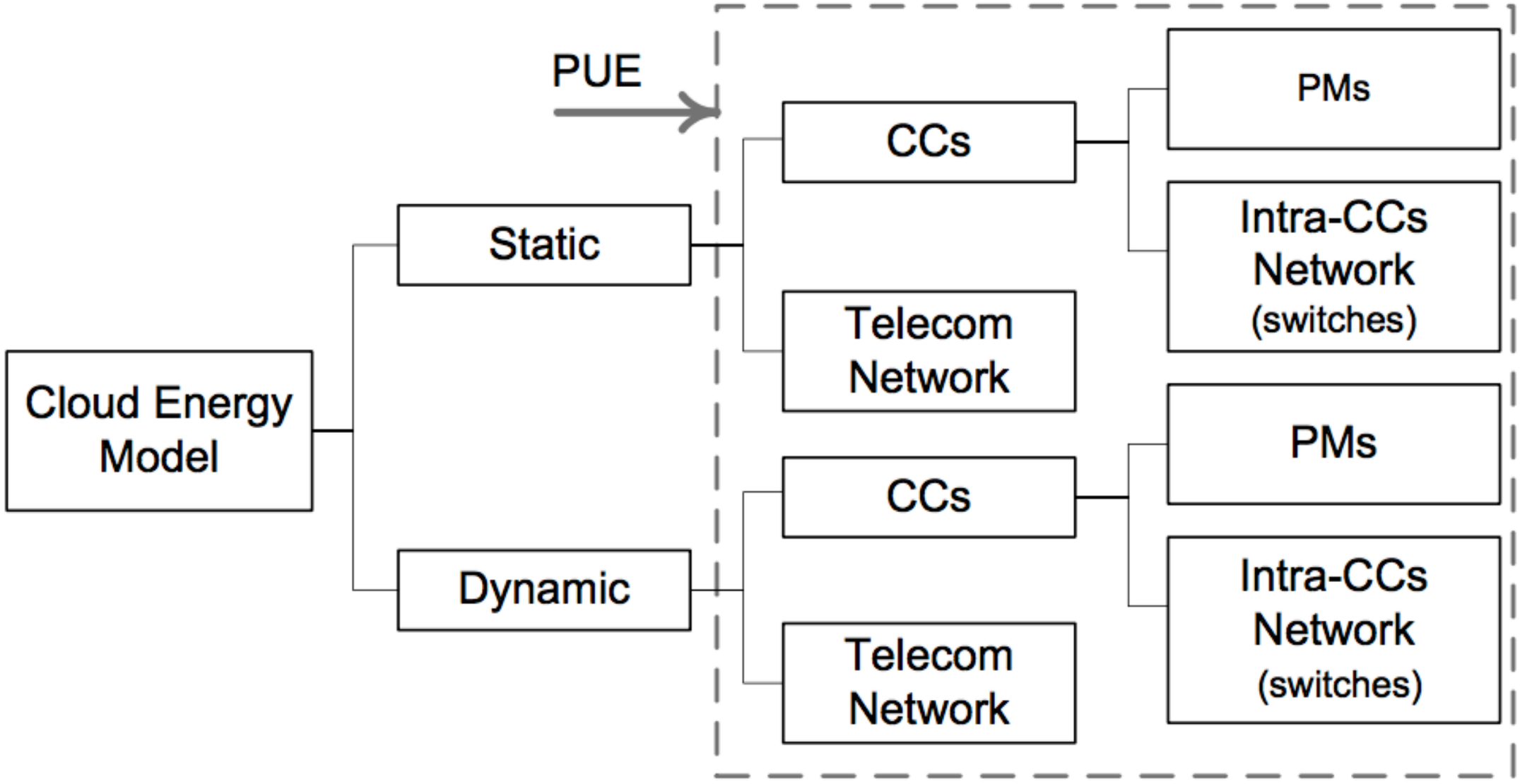
- Two main costs: maintenance & network
- Outsourcing viable when network & storage are monetized
- Building and backbone expenses should be shared by other activities
- Carefully chose the number of VMs per DC
- Many ways to deploy distributed clouds…

Energy cost of an edge cloud?

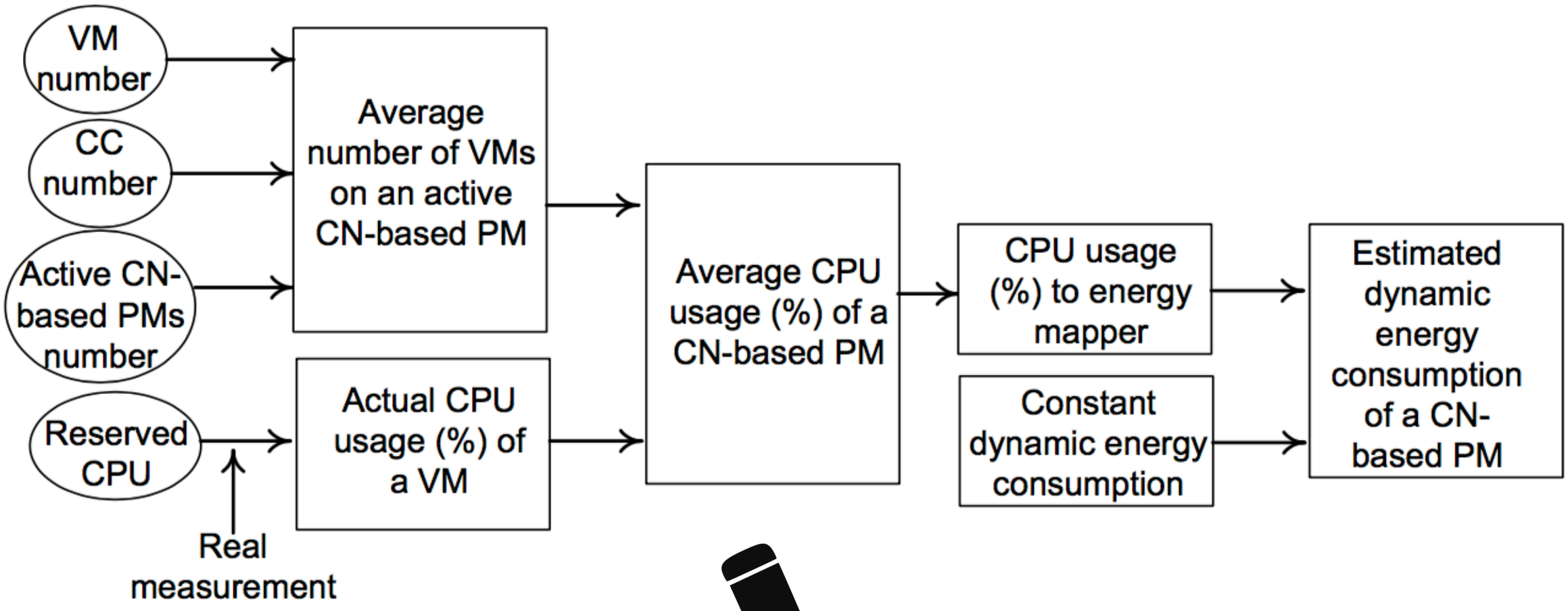
Architectures



Energy consumption splitting



Example for servers

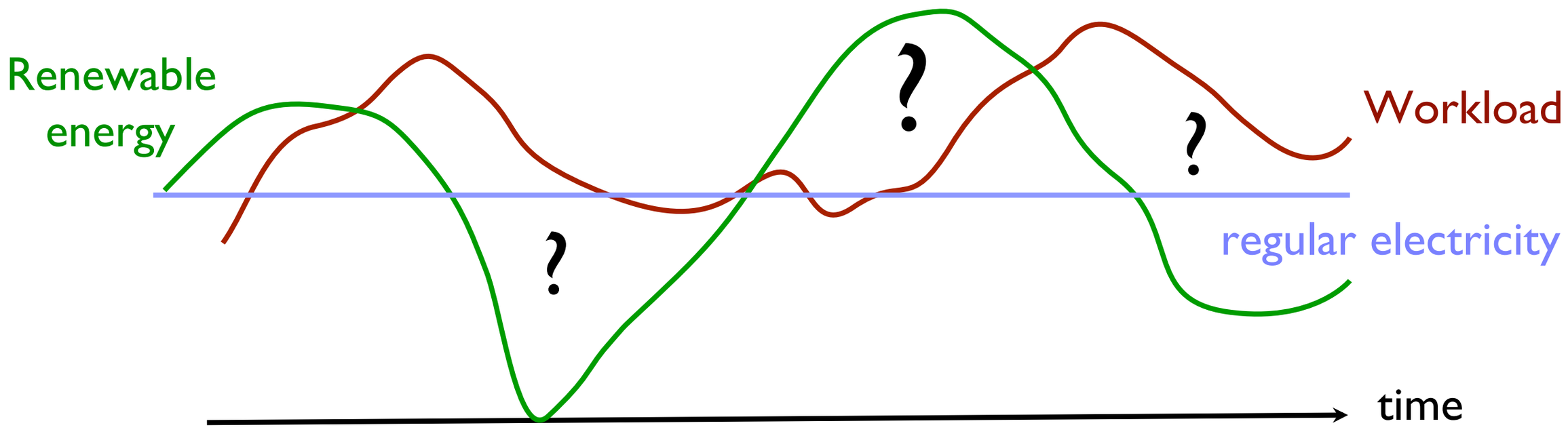


Outline

- Context
- Two post-docs
- 1. OpenStack footprint
- 2. Cost-benefit analysis centralized vs. distributed Clouds
- **3. Exploiting renewable energy**
- Conclusions

Goals

- Exploiting local production of renewable energy
- Designing green policies at local scale (on-off, consolidation, etc.) and infrastructure scale (resource allocation, migration, etc.)

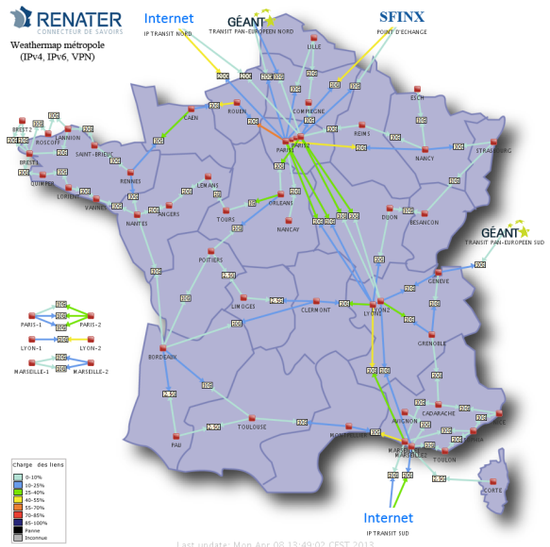


Outline

- Context
- Two post-docs
- 1. OpenStack footprint
- 2. Cost-benefit analysis centralized vs. distributed Clouds
- 3. Exploiting renewable energy
- **Conclusions**

Beyond Discovery !

- From sustainable data centers to a new source of energy



- Leverage "green" energy (solar, wind turbines...)

http://parasol.cs.rutgers.edu

Exploiting renewable energy

http://parasol.cs.rutgers.edu

Thank you for your attention

http://people.irisa.fr/Anne-Cecile.Orgerie