JOB EFFICIENCY MONITORING
Getting your cluster and supercomputer under control

Jordi Blasco

18/06/2020

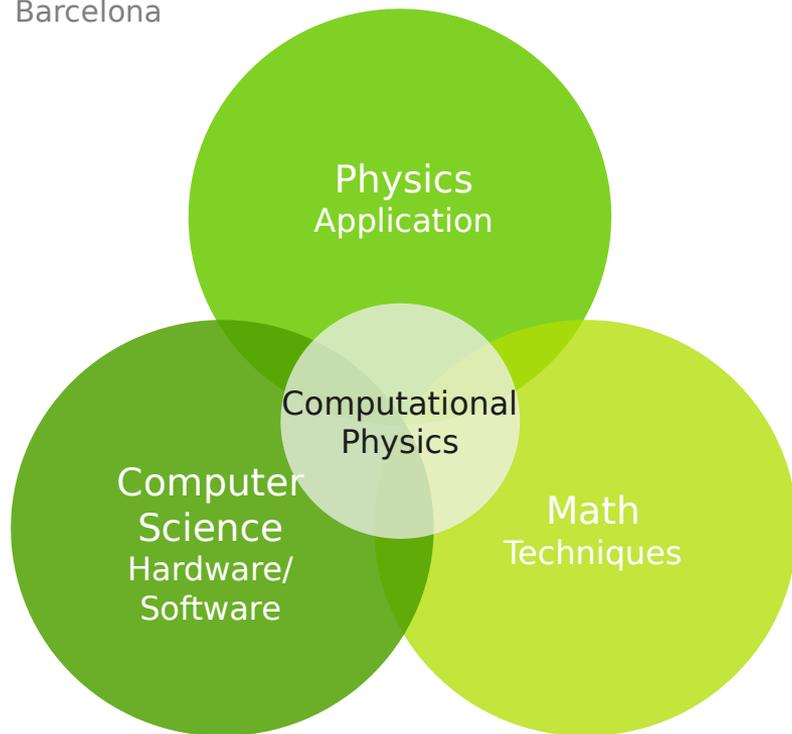www.hpcnow.com

HPC Knowledge
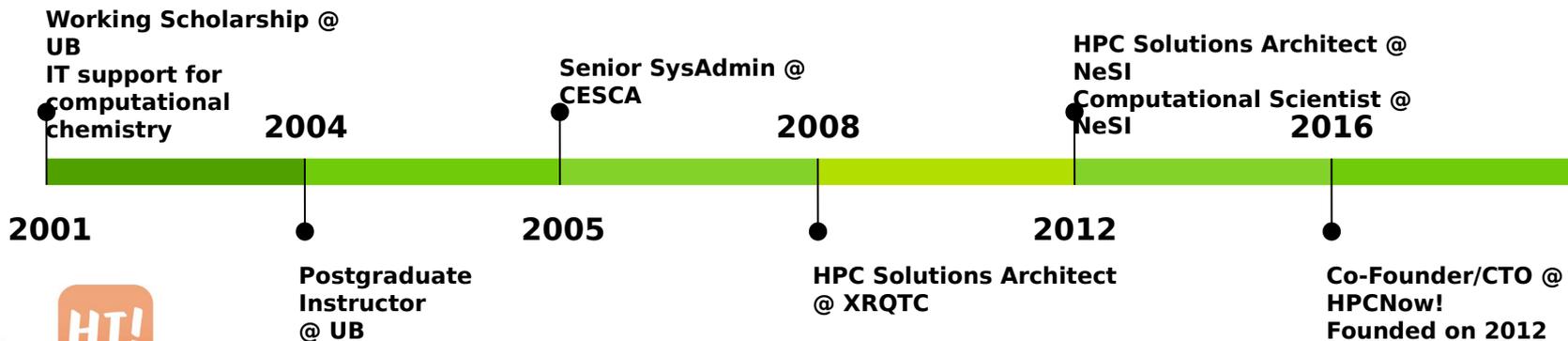Meeting'20

HPC Now!

Working Scholarship @ UB
IT support for computational chemistry

2004

Senior SysAdmin @ CESCA

2008

HPC Solutions Architect @ NeSI
Computational Scientist @ NeSI

2016

2001

Postgraduate Instructor @ UB

2005

HPC Solutions Architect @ XRQTC

2012

Co-Founder/CTO @ HPCNow!
Founded on 2012

HI!

HPC Now!

Advanced supercomputing services for science and engineering

We plan, install & support

You run

www.hpcnow.com

JOB EFFICIENCY MONITORING
Getting your cluster and supercomputer under control

Jordi Blasco

18/06/2020

2. Motivation

Do more with less
Identification of efficiency issues is very time-consuming

For a consulting company like HPCNow!, it means:

- We can offer more and better managed services with less resources.
- We can detect issues really quick.
- We can add more value to our services:
  - Better utilisation
  - Code/app tuning
  - Improve workflows
  - User education
- **We can quantify the impact of our work**

Developed as a tool to accelerate the way we deliver our services

Solve
Quick issue identification

Improve
Fast discovery of
opportunities

Evaluate
Increase return of
investment

Quantify
Accelerate research

What are the mechanisms for accelerating the research and discovery?

Buy a New Cluster
Large procurement usually involves long and complex RfP processes.
Not all the supercomputing facilities have a consolidated platform refresh cycle.

Use Cloud
Extend the current compute capacity by using cloud bursting to accommodate peaks of needs is definitely the best option. Unfortunately, ongoing and regular usage becomes expensive.

Improve Efficiency
By improving the performance and efficiency, you are somehow creating more allocation for new jobs with zero investment in hardware capacity.

Limitations of traditional technologies

Popular tools like Ganglia, Munin or Zabbix are not capable of representing the metrics required to identify inefficient jobs.

No link to user

No link to job

No link to other nodes allocated

No information about efficiency

Requirements / wishlist

The main goal is to identify opportunities to improve user workflows, user codes and applications, in addition to user

1 **Review resources requested vs used**

Key fundamental metrics to understand how well used are the resources requested.

2 **Real Time Monitoring**

Enables proactive job profiling and also the possibility to trigger real time actions.

3 **30 Seconds Resolution**

30 seconds resolution is quite reasonable for the majority of HPC workloads.

3. Architecture

![HPC Now! logo]

Key software components

Architecture

Large number of events to analyse requires the use of Big Data technologies. Data is gathered using custom codes and aggregated into ElasticSearch and InfluxDB, open source search and analytics engine which has high reliability and proven scalability. Finally, the data is represented through Grafana and Kibana, which are leading tools for querying and visualizing large datasets and metrics.
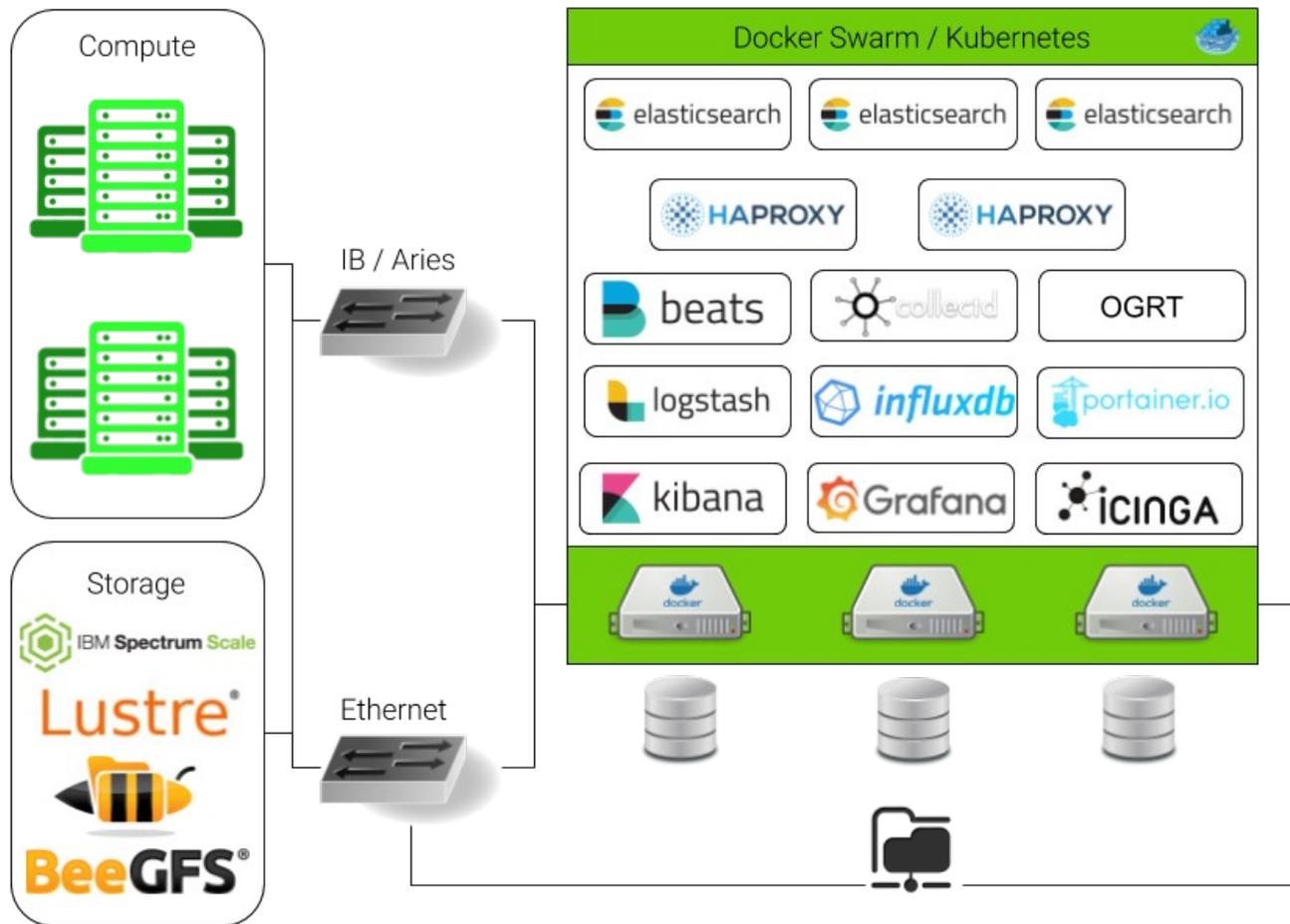
Magic Orchestration

In order to accelerate the time-to-solution, HPCNow! has invested a lot of efforts and resources to automate the deployment by taking advantage of container-centric orchestrater technologies.

Noun: automagic (uncountable)

/ɔːtəˈmadʒɪk/

1. A process carried out automatically in such a clever way that the result appears to be magic.

2. Blend of automatic + magic; from the principle (often called Clarke's third law) that any sufficiently advanced technology is indistinguishable from magic.
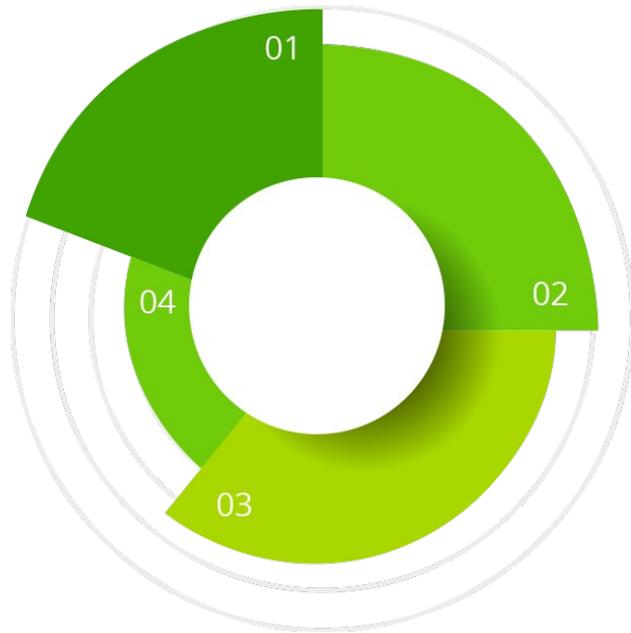
Key Features

Short turnaround from issue identification to solution
Quick discovery, fast analysis

Users can customise the dashboards to meet site needs
Flexible visualization

Easy and affordable way to scale out
Scalable solution

Orchestration based on Docker Swarm and Kubernetes
Quick orchestration

It has 20MB memory footprint. No penalty on HPL
No overhead

Towards better understanding of real needs

Demo

Taking as starting point predefined dashboards with the most demanded and useful metrics, these dashboards can also be tailored according to the customer needs.

Computationa l Scientists

Platforms Teams

End Users

End User Support Teams

4.Case Studies

More than just a working solution

Nowadays, we install this solution in every single client we provide managed services.

We have installed this solution as a request for some clients.

Pfizer

CSIRO

Bureau of

Australian Government
Bureau of Meteorology

5.Conclusions

HPC Now **!**

Scalability

Performance

Efficiency

Detect user mistakes early

It has been key to improve reliability, and avoid massive waste of computational resources.

Increase HPC user maturity

It helped to accelerate research and discovery; helped to improve the user workflows. End users acquired more knowledge and awareness about the their needs.

Quantify the impact of our services

In terms of return of investment, research outputs, utilisation, user satisfaction, ...

THANK YOU!

HPCNow!

Parc Tecnològic, Marie Curie, 8 08042 Barcelona (Spain)

34 Fernly Rise, 2019 Auckland (New Zealand)

info@hpcnow.com

HPC Now!