

CREATING A CLUSTER - GOING IT ALONE

Simon Atack

University of Bristol

Simon.Attack@Bristol.Ac.Uk



OVERVIEW

In 2019 Bristol University went live with its homemade cluster. This is our experience of the trials and tribulations encountered during this endeavor from a sysadmin perspective. I will discuss the technical aspects and design considerations (including networking, deployment, scheduling [software stack] etc). And will finish with our reflections on this experience

BACKGROUND

Vision of the system

- A system that can be expanded (that could replace small standalone individual clusters)
- researchers can buy into the cluster
- researchers could buy at every new cluster eg 3 or 4 years
- Central irregular top ups
- Add in new equipment test things

INITIALLY

- Initially dev clusters created from scrap equipment
- Initial eval hardware purchased

CLUSTER STAGES

In order to satisfy end users

- stopped being a development
- became testing
- ultimately sneaking into production before continuing to develop & fine tune while a production system

MGMT SYSTEMS

- we went from classical HA systems
- to HV VMs doing HA

PROVISIONING/DEPLOYMENT

- foreman - as previously in system (non HPC)
- self developed wrapper round templating - as a quick lightweight solution

NETWORKING

- not aimed for low latency/high performance
- Initially as a 1G
- investigation of higher speeds
- settled on higher 100G/25G backbone
- 10G to nodes

- Previously Cumulus OS – Do we go Dell OS/Sonic?

QUEUING

- PBS was used to start (due to experience)
- switch to slurm to consolidate

STORAGE

- Initially basic NFS storage
- multiple namespaces
- moved to over to GPFS multicluster
- aimed as a storage for HPC rather than a storage for a cluster

SOFTWARE

- considering software management
- easybuild considered - due to time constraints couldnt go that route

JACK OF ALL TRADES

- We became on top hpc sysadmin
- Devops
- Developers
- Network Engineers
- continual racking & settingup

ISSUES OCCURED

- MGMT Stack wasnt stable reboot due to deployment shortcuts in building (chicken/egg issue)

OUTCOMES

- Upskilled team in understanding of clusters, how connected
- Gained working knowledge
- Single points of failure in experience/knowledge -- due to deadline
- Buddy development Work
- Documentation fell behind
- Minimum Viable product - missing features
- increased usage of VMs
- Investigating alternative finance structure for 'selling' resources of machines

LESSONS LEARNT

- We are not developers! (skills)
- Have a development system support structure
- This takes time (much more than you expect)
- Success has depended on goodwill of staff going above and beyond
- System is as you want it (you choose what you want)
- You've taken on the full support -- buck stops here.
- Scaling - lots of work for one institution to support

THANKS TO ACRC HPC TEAM

- Callum Wright
- Amaurys Avila Ibarra
- Isaac Prior
- Tom Batstone
- Dianaimh Green
- Ethan Williams

QUESTIONS

