Science & Technology Facilities Council

Virtualisation Cloud Computing at the RAL Tier 1

Ian Collier

STFC RAL Tier 1

HEPiX, Bologna, 18th April 2013

Virtualisation @ RAL

- Context at RAL
- Hyper-V Services Platform
- Scientific Computing Department Cloud
- GridPP Cloud Project

What Do We Mean By ‘Cloud’

For these purposes

- “does not require administrator intervention”
- Service owners don’t have to care about where things run

Context at RAL

- Historically requests for systems went to fabric team
  - Procure new HW – could take months
  - Scavenge old WNs – could take days/weeks
- Kickstarts & scripts took tailoring for each system
- Not very dynamic
- For development systems many users simply run VMs on their desktops – hard to track & risky

Evolution at RAL

- Many elements play their part
  - Configuration management system
    - Quattor (introduced in 2009) abstracts hardware from os from payload, automates most deployment
    - Makes migration & upgrades much easier (still not completely trivial)
  - Databases feeding and driving configuration management system
    - Provisioning new hardware much faster

Virtualisation & Cloud @ RAL

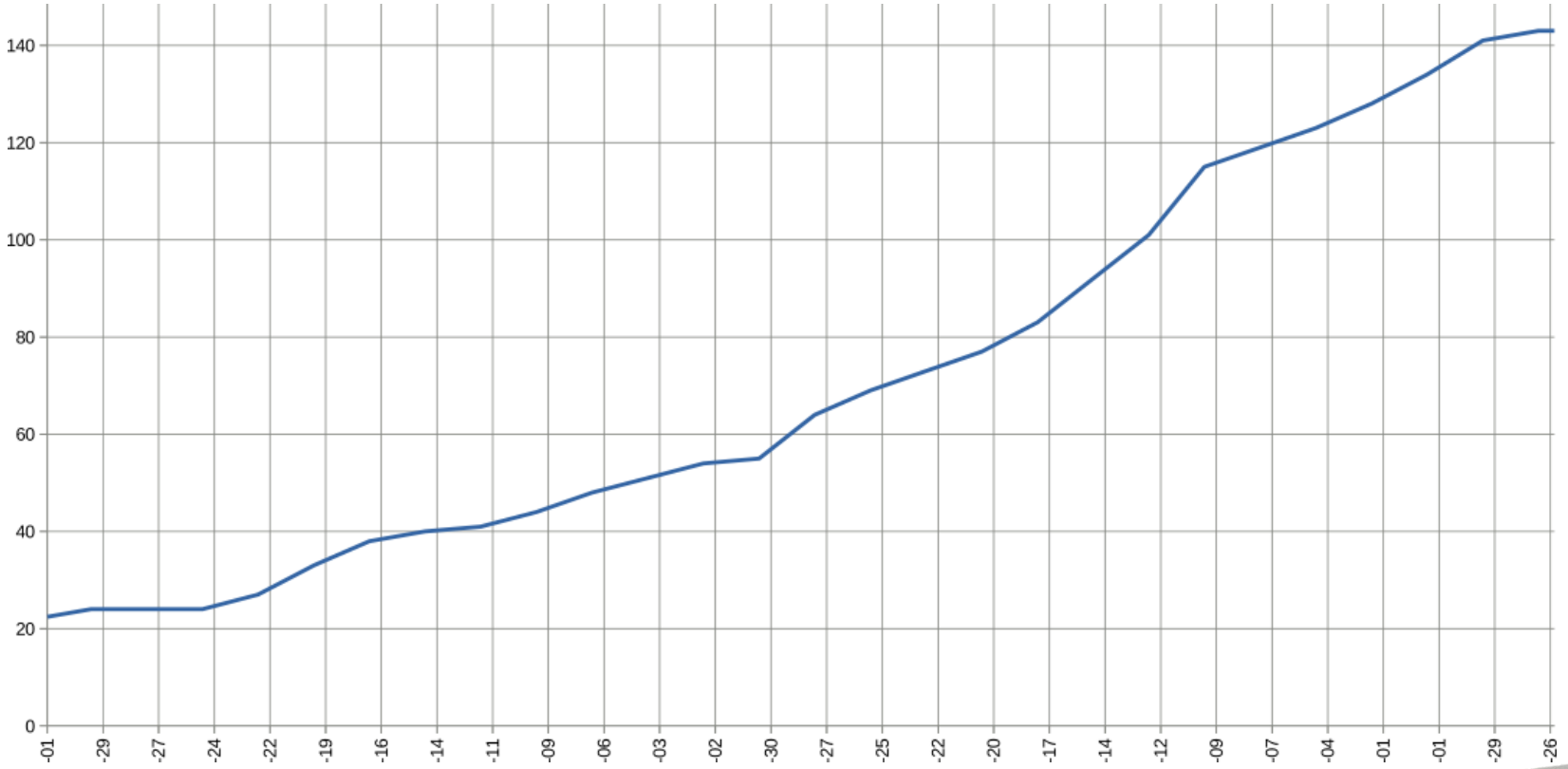
- Context at RAL
- Hyper-V Services Platform
- Scientific Computing Department Cloud
- WLCG Related Cloud

Science & Technology Facilities Council

Hyper-V Platform

- Over last three years
  - Local storage only in production
  - ~200 VMs
- Provisioning transformed
  - Much more responsive to changing requirements
  - Self service basis – requires training all admins in using management tools – but this
- Progress of high availability shared storage platform (much) slower than we'd have liked
  - Planning move to production now

Hyper-V Platform



Science & Technology Facilities Council

Hyper-V Platform

- Mostgrid services virtualised now
  - argus, apel, bdii, cream-ce, fts, myproxy, ui, wms, etc.
- Internal databases & monitoring systems
- Also test beds (batch system, CEs, bdiis etc)
- Move to production very smooth
  - Team had good period to become familiar with environment & tools

Hyper-V Platform

- When a Tier 1 admin needs to set up a new machine all they have to request is a DNS entry
- Everything else they do themselves
- Maintenance of underlying hardware platform can be done with (almost) no service interruption.
- This is already much, much better – especially more responsive – than what went before.

- Has many characteristics of private cloud
  - But we wouldn't usually call it 'cloud'

Hyper-V Platform

- However, Windows administration is not friction or effort free (we are mostly Linux admins….)
  - Share management server with STFC corporate IT – but they do not have resources to support our use
  - Troubleshooting means even more learning
  - Some just ‘don’t like it’
- Hyper-V continues to throw up problems supporting Linux
  - None show stoppers, but they drain effort and limit things
  - Ease of management otherwise compensates for now
  - **Much** better with latest SL (5.9 & 6.4)
- Since we began open source tools have moved on
  - We are not wedded to Hyper-V

Virtualisation & Cloud @ RAL

- Context at RAL
- Hyper-V Services Platform
- Scientific Computing Department Cloud
- WLCG Related Cloud

Science & Technology Facilities Council

SCD Cloud

- Prototype E-Science Department cloud platform
- Began as small experiment 18 months ago
- Using StratusLab
  - Share Quattor configuration templates with other sites
  - Very quick and easy to get working
  - But has been a moving target as it develops
- Deployment done by graduates on 6 month rotation
  - Disruptive & variable progress

SCD Cloud

- Initially treat systems much like any Tier 1 system
- We allow users in whom we have high levels of trust
  - Monitor that central logging is active, sw updates are happening
- Cautiously introducing new user groups
- Plan to implement further network separation
  - Waiting for reorganisation of Tier 1 Network Martin spoke about

SCD Cloud

- Resources
  - Began with 20 (very) old worker nodes
    - ~80 cores
    - Filled up very quickly
    - 1 year ago added 120 cores in new Dell R410s – and also a few more old WNs
    - This month adding 160 cores in more R410s
- Current
  - ~300 cores – enough to
    - continue development to cover further use cases
    - Run a meaningful test bed

SCD Cloud Usage

- 30 or so regular users (dept of ~200)
- ~100 VMs at any one time
  - Typically running at 90-95% full
- Exploratory users from other departments
- Also adding very selective external (GridPP) users
- Proof of concept more than successful
  - Full time ‘permanent’ staff in plan
  - It is busy – lots of testing & development
    - People notice when it is not available

SCD Cloud Future

- Develop to full resilient service to users across STFC
- Participation in cloud federations
- Have been evaluating storage solutions
  - For image store/sharing and S3 storage service
  - Ceph looks very promising for both
    - Have new hardware delivered for 80TB ceph cluster
    - Will be deploying in coming weeks
- Integrating cloud resources in to Tier 1 grid work
- Reexamine platform itself.
  - Things have moved on since we started with StratusLab

Virtualisation & Cloud @ RAL

- Context at RAL
- Hyper-V Services Platform
- Scientific Computing Department Cloud
- WLCG Related Cloud

Science & Technology Facilities Council

WLCG Related Cloud

Dynamically-provisioned worker nodes

- Allow a traditional batch system to make opportunistic use of cloud resources by dynamically creating worker nodes
- Testing two implementations:
  - HTCondor
    - A service monitors the state of the pool which creates & destroys VMs as necessary. Condor startd daemons on each VM then advertise themselves to the Condor collector.
  - SLURM
    - Makes use of existing power save logic: instead of powering up & down nodes, the SLURM controller creates & destroys VMs as necessary.

WLCG Related Cloud

CMS UK cloud activities

- Enabling CMS analysis jobs to be run on cloud resources in the UK
  - Users run the standard CMS tool (CMS Remote Analysis Builder) to create & submit jobs
  - GlideinWMS system at RAL instantiates VMs as needed & creates an on-demand overlay HTCondor batch system for running the user jobs

Virtualisation & Cloud @ RAL

- Context at RAL
- Hyper-V Services Platform
- Scientific Computing Department Cloud
- WLCG Related Cloud

Science & Technology Facilities Council

Summary

- Using range of technologies
- Many ways our provisioning & workflows have become more responsive, ‘agile’
- Private cloud has developed from a small experiment to beginning to provide real services
  - With constrained effort
  - Slower than we would have liked
  - The experimental platform is proving well used
- We look forward to being able to replace Hyper-V for resilient services

Science & Technology Facilities Council

Backup Slides

Science & Technology
Facilities Council

JASMIN/CEMS

The JASMIN super-data-cluster

- UK and European climate and earth system modelling community.
- Climate and Environmental Monitoring from Space (CEMS)
- Facilitating further comparison and evaluation of models with data.

6.6 PB Storage Panasas at STFC

- Fast Parallel IO to Compute servers (370 Cores)

Gnodal 10GB networking

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Science & Technology Facilities Council

JASMIN/CEMS

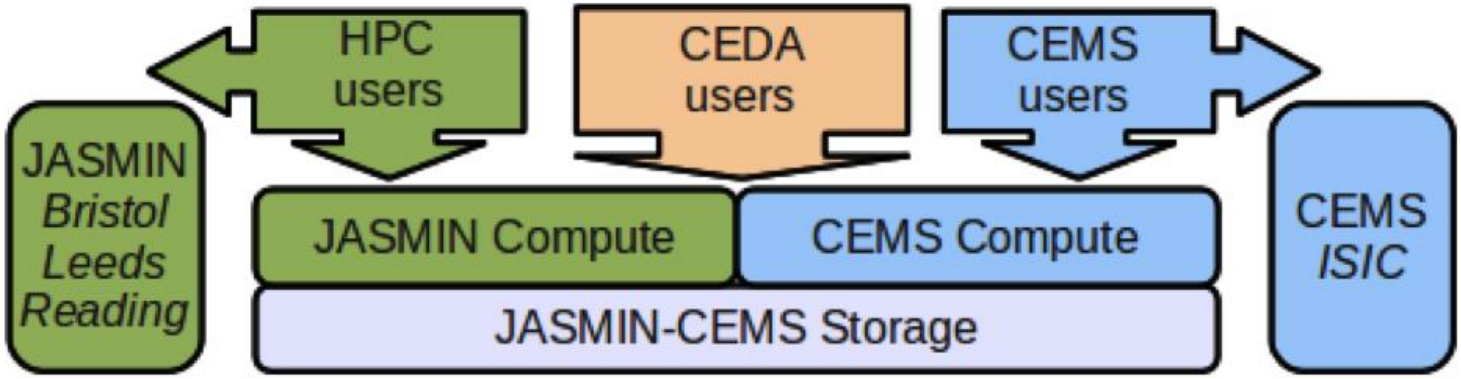

panasas
panasas
panasas
panasas
panasas
panasas
panasas
panasas
panasas

JASMIN Super Data Cluster

JASMIN	3.5 PetaBytes Panasas Storage 20 x Dell R610 (12 core, 3.0GHz, 96G RAM) 1 x Dell R815 (48 core, 2.2GHz, 128G RAM) 1 x Dell Equallogic R6510E (48 TB iSCSI VM image store) VMWare vSphere Center 1 x Force10 S4810P 10GbE Storage Aggregation Switch
CEMS	1.1 PetaBytes Panasas Storage 7 x Dell R610 (12 core 96G RAM)Servers 1 x Dell Equallogic R6510E (48 TB iSCSI VMware VM image store) VMWare vSphere Center + vCloud Director

Science & Technology Facilities Council

JASMIN Super Data Cluster



JASMIN provides three classes of service:

•Virtualised compute environment (not strictly a "private cloud").

•Physical compute environment.

- No private data connection

HPC service ("Lotus").

- Not easily reconfigurable to JASMIN cloud.
- Separate data connection.

Science & Technology Facilities Council

JASMIN Super Data Cluster

- Two distinct clouds
- One supports manual VM provisioning by CEDA and the climate HPC community
  - Configuration controlled at site
  - Therefore greater trst and greater network access
- One supports more dynamic provisioning by the academic users in the CEMS community.
  - Users provision own VMs
  - Access to Panasas
  - Otherwise less trusted
- So, they have different vCentre server installations.

Science & Technology Facilities Council