IRIS Vcycle digital asset

Andrew McNab

University of Manchester

GridPP, DUNE, LHCb

IRIS Vcycle digital asset

- Vcycle/OpenStack context
- Digital Asset for IRIS
  - Library of VMs
  - Vcycle
  - ~~Vcycle Dashboard~~
    - Commercial Cloud
- Next steps

Vcycle in the OpenStack ecosystem

- IRIS has funded OpenStack and raised (some) expectations in new partners that OpenStack will be available

  - It is genuinely very flexible, and mainstream

- Sites increasingly becoming general HEP+Astro+???

  - "Manchester Tier-2 Centre" becomes "Blackett Computer Centre"

  - OpenStack helps support this generalisation

- Vcycle was developed for managing VMs that run jobs for HEP experiments

- So naturally fits into this as a way of running jobs from IRIS workload frameworks on OpenStack
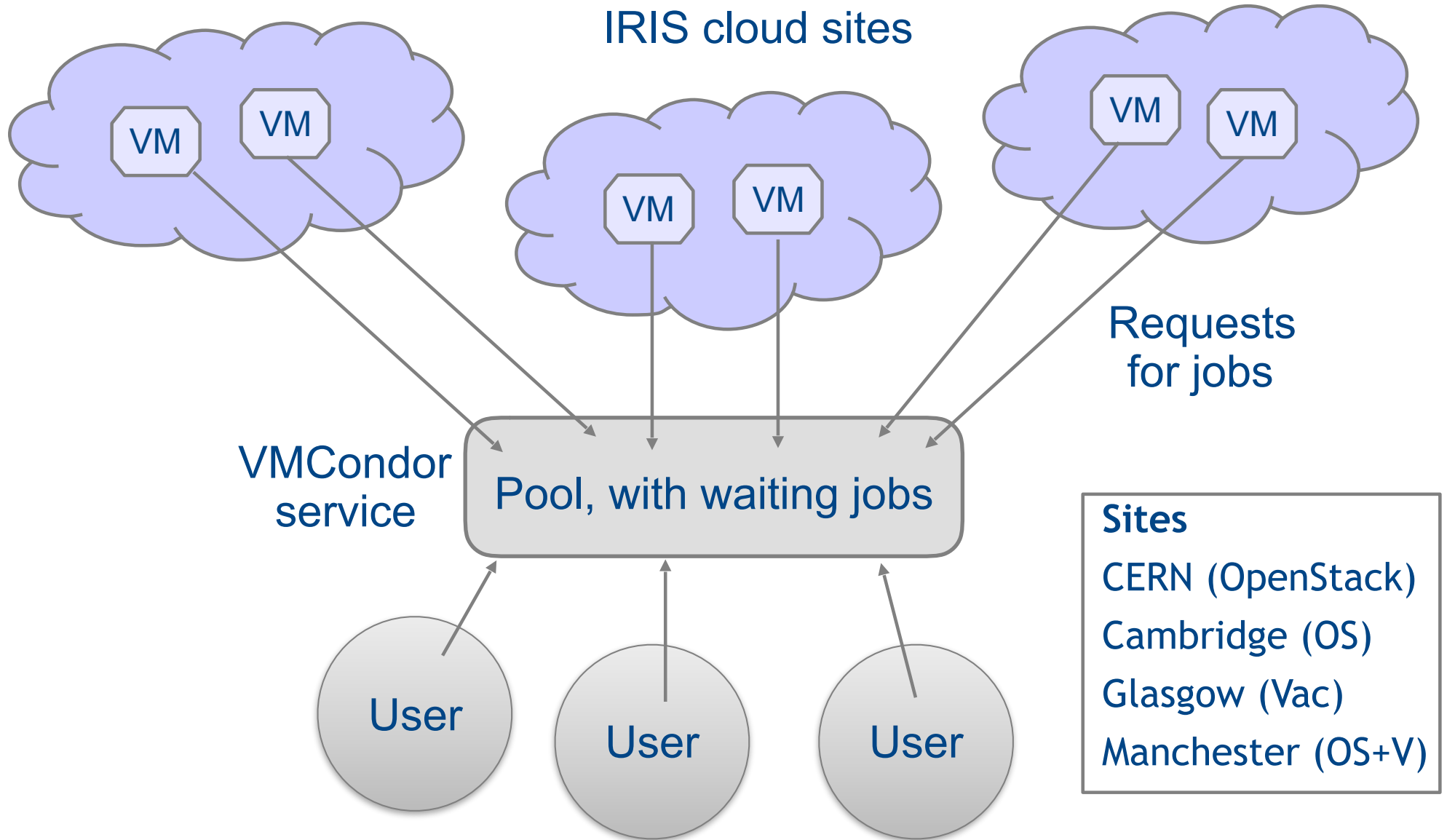
3

Vcycle Digital Asset

- November 2018 to March 2020
- Three original work packages:
  - Library of VM definitions
  - Vcycle improvements
  - Vcycle dashboard

WP1 Library of VM definitions

- Existing VMs (GridPP, LHCb, ATLAS; ALICE next) updated to CernVM 4 (~CentOS7)

- SKA and LSST requirements supported in GridPP DIRAC VMs (mainly multiprocessor job support)

- DUNE VM in production, contacts GlideinWMS pool at FNAL for jobs

- EUCLID VM ongoing: requires that VMs join a slurm batch system in Edinburgh over a VPN

- VMCondor now an end-to-end "simple batch" system

  - Users do not need X.509 certificates

  - VMCondor wrappers of HTCondor commands

  - VMs at multiple sites join an HTCondor pool

  - https://github.com/iris-ac-uk/iris-vmcondor/wiki

VMCondor architecture

IRIS cloud sites

VM   VM

VM   VM

VM   VM

Requests for jobs

VMCondor service

Pool, with waiting jobs

Sites
CERN (OpenStack)
Cambridge (OS)
Glasgow (Vac)
Manchester (OS+V)

User   User   User

WP2 Vcycle hardening and scalability

- Hardening and documentation
  - LHCb Vcycle instance at CERN used to identify some race conditions and improve robustness
- Support added for multiple co-operating Vcycle instances
  - Scalability: going beyond ~4000 VMs.
  - Live upgrades: take one Vcycle instance down at a time for upgrades etc.
  - Failover: one Vcycle instance dies, the VMs it created become managed by another instance
  - Ongoing work to fix bugs, race conditions, as all this scales up.
- Underpinned by shared filesystem between instances rather than a database

Single instance Vcycle   Multi instance Vcycle

Vcycle DA - Andrew.McNab@cern.ch - IRIS F2F, London, 17 Dec 2019
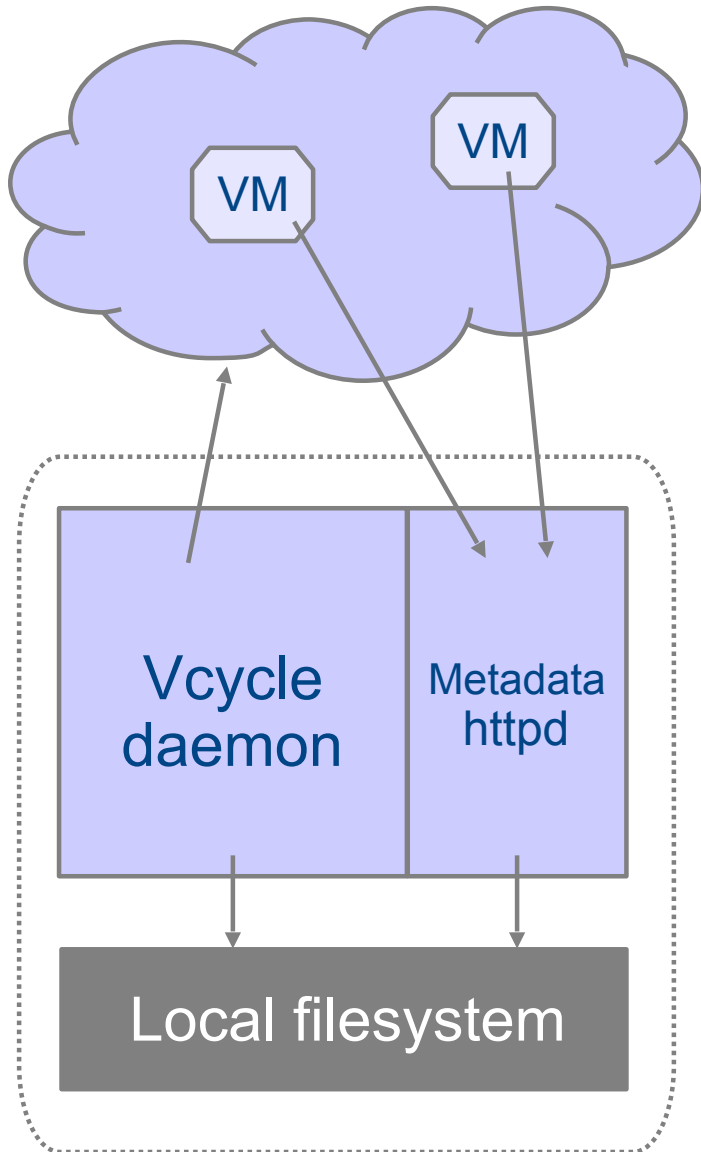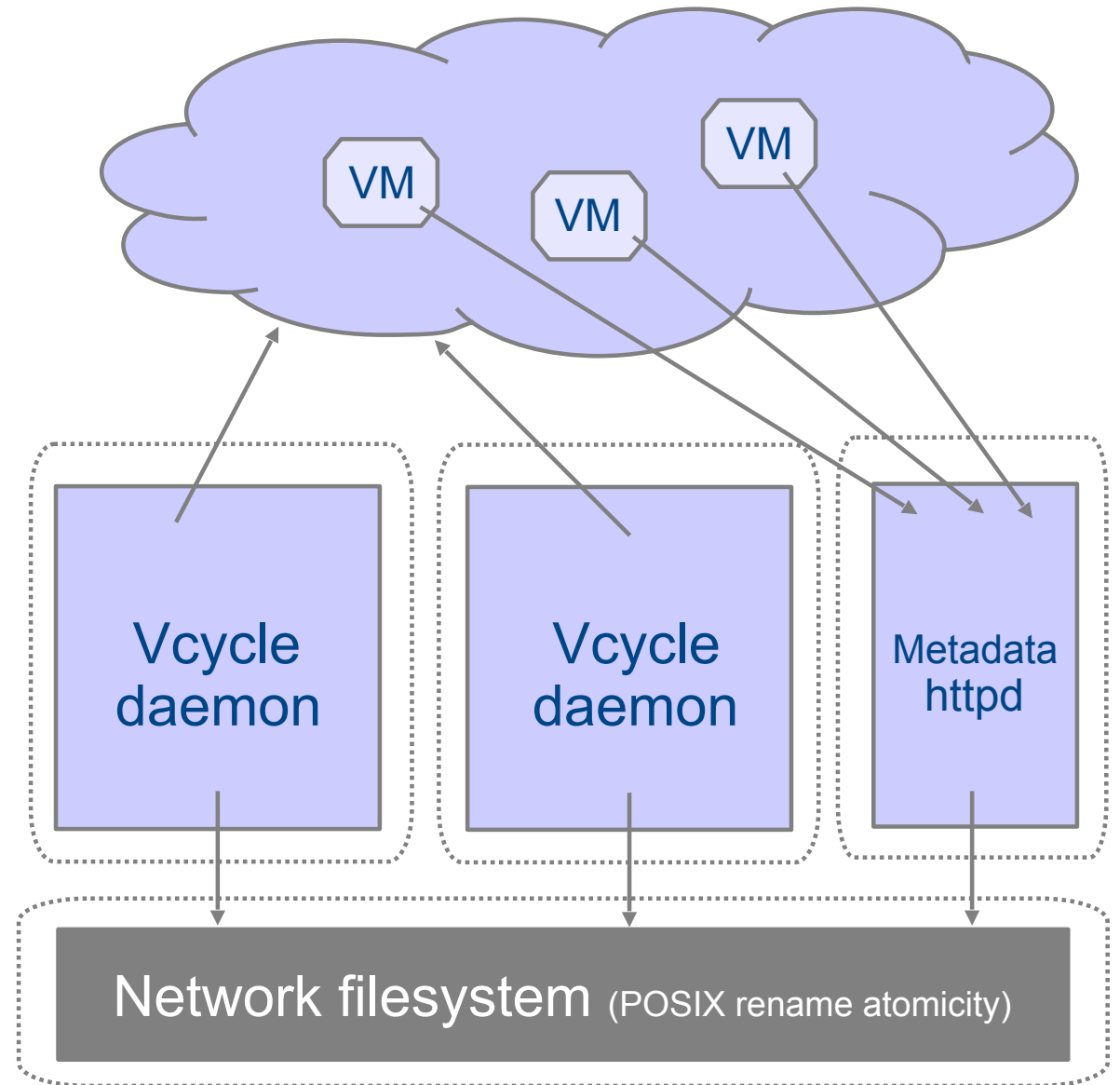
OpenStack sites used for testing

- Vcycle instance at Manchester has access to and manages VMs at
  - Cambridge (Cumulus)
  - RAL (Tier-1)
  - Manchester
  - Edinburgh (ECDF)
  - CERN (donated by LHCb)
- All these are at the level of 1 - 10 VMs but it has been very useful for testing since they are all slightly different

Vcycle and IAM

- During the project support for OpenStack API credentials were added
  - The workflow is that you log into the OpenStack dashboard with username/password and retrieve a credential pair
  - Pair is then put in the Vcycle configuration
  - This is a natural and secure way to work, and a one off step when the site is configured in Vcycle
- One consequence is that however you authenticate to the dashboard (IAM?), you can retrieve and use API credentials
- So to first order, no special IAM support is needed

IRIS Commercial Cloud

- £50,000 to be spent on commercial cloud before end of March 2020

- Target Google and Amazon cloud services
  - Instantiate VMs as on OpenStack

- Approved and waiting for the money to be released
  - Google and Amazon accounts created and plugins for Vcycle being tested with "free" money

- Plan is to prioritise SKA, LSST, DUNE and backfill with LHCb to maintain the rate of spend
  - Have done something similar with LHCb in the past using "free" money on Google Cloud

WP3 ????

- Was "Vcycle dashboard" to provide a web interface for managing Vcycle instances

- Instead will propose to use the remaining three months for Amazon/Google cloud work

  - More than just running as many VMs as we can up to a limit and the demand?

  - eg add support for Amazon spot pricing

  - Use pre-emption support in some VM flavours

- Start to include elements of Compute Provisioning

  - (See next slide)

Compute Provisioning

- Several things are coming together under this umbrella

- In cloud terms, it means something like Vcycle which has access to a project or tenancy

  - And decides when to create VMs which workload management systems can then fill with work

- "when" to do this is the big issue when we have different priorities from different user communities

- Things like pre-emption are also part of the picture

- Longer term, want to be able to publish a plan for this year so data processing campaigns in the future (for example) are prioritised over nice-to-have but non essential simulation now

Summary and next steps

- VM library has produced useful results
  - Multiprocessor DIRAC VM (GridPP for SKA/LSST and LHCb)
  - New DUNE VM
  - New VMCondor simple batch VM
- Multi instance Vcycle work is in production
  - Scalability, live upgrades, failover
- API credentials support has indirectly led to IAM support
- Next steps:
  - More VM types
  - Provide these VM-based workloads to more IRIS sites
  - Deploy on capacity from IRIS Commercial Cloud project