

# Onboarding new VOs

Alastair Dewhurst



Alastair Dewhurst, 7<sup>th</sup> December 2018



# Introduction

- GridPP provides grid solutions to scientific computing problems.
  - It works well for those communities that both need it and understand why they need it.
- For those that don't need the Grid:
  - Provide something smaller and simpler.
  - Don't provide anything as it is not a use case we can deliver on.
- For those that need a Grid like solution but don't yet understand why:
  - It's difficult but we have to allow them to choose their destiny.
  - We are providing infrastructure as a service, for them to setup what they like.



# Which User?

- For the purposes of this discussion the new community is assumed to:
  - Probably only have a single person working on computing, who may have no experience of Grid middleware.
  - Need less than 500TB of archival tape storage.
  - Need less than 100TB of disk storage.
  - Need less than 2 million CPU hours in any one month (i.e. average usage of 3000 cores) and less than 6 million over a year.
- Any solution we propose should hopefully scale beyond this without forcing the user to start over again?



# UI as a service

- The Tier-I is going to setup a UI per VO (on OpenStack)
  - Access via ssh keys
- Allow direct submission to HTCondor batch farm.
  - The batch farm can automatically use cloud resources including high memory machines and GPU.
- Allow access to Echo S3 via DynaFed.
  - DynaFed can permit access based on hostname of UI.
  - HTCondor can be configured to upload job output directly to S3.
- Provide the ability to upload to CVMFS Stratum 0.
- Provide limited access to Castor tape.
  - Bulk archival will need to go via FTS into SRM.



# GridPP benefit?

- How do UIs submitting directly to RAL only resources help the rest of GridPP?
  - If we educate the user on Grid philosophies and some tools, it makes the transition to using the Grid fully much simpler.
  - I want to ensure there is a clear evolution from this to using the GridPP DIRAC instance.
- Other sites could choose to provide a similar setup.
  - The resources are silo but if the user rarely needs more than what a single site can provide that is fine.
  - If GridPP sites communicate and we have reasonably similar setups the user will be able to submit to multiple sites and learn why the Grid is the way it is.
- This is going back to basics, I expect it to evolve.



# First impressions matter

- I think it is very important that we get things right initially.
  - That buys plenty of good will.
- What are the first things we need to get right?
  - Understand their use case.
  - Software distribution.
  - What else?

