DAaaS: Behind the scenes

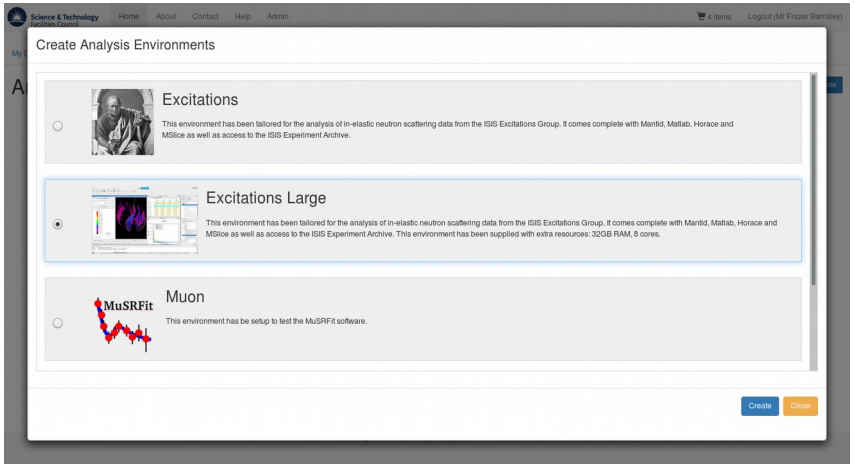
Frazer Barnsley, May 2019

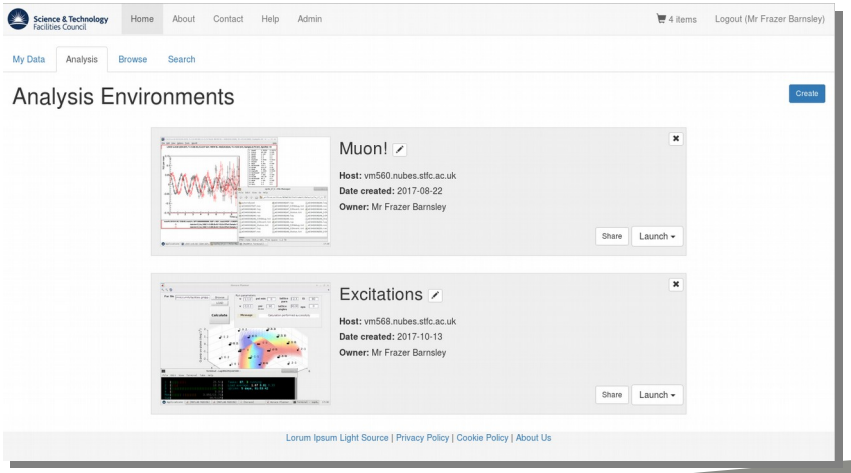
Overview

- User facing capabilities
- Virtual Machine Manager (VMM)
- VM Configuration
- ISIS Cloud
- External Clouds
- Data Movement Service (DMS) / Data Transfer Service (DTS)
- HPC integration

User Facing Capabilities

- Users can select a virtual research environment for a specific type of analysis
- Near instant – takes a couple of seconds
- Access via SSH, RDP, NoVNC

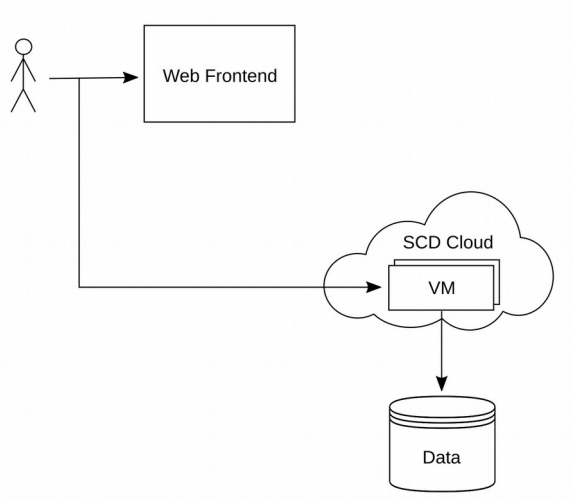




Science & Technology Facilities Council

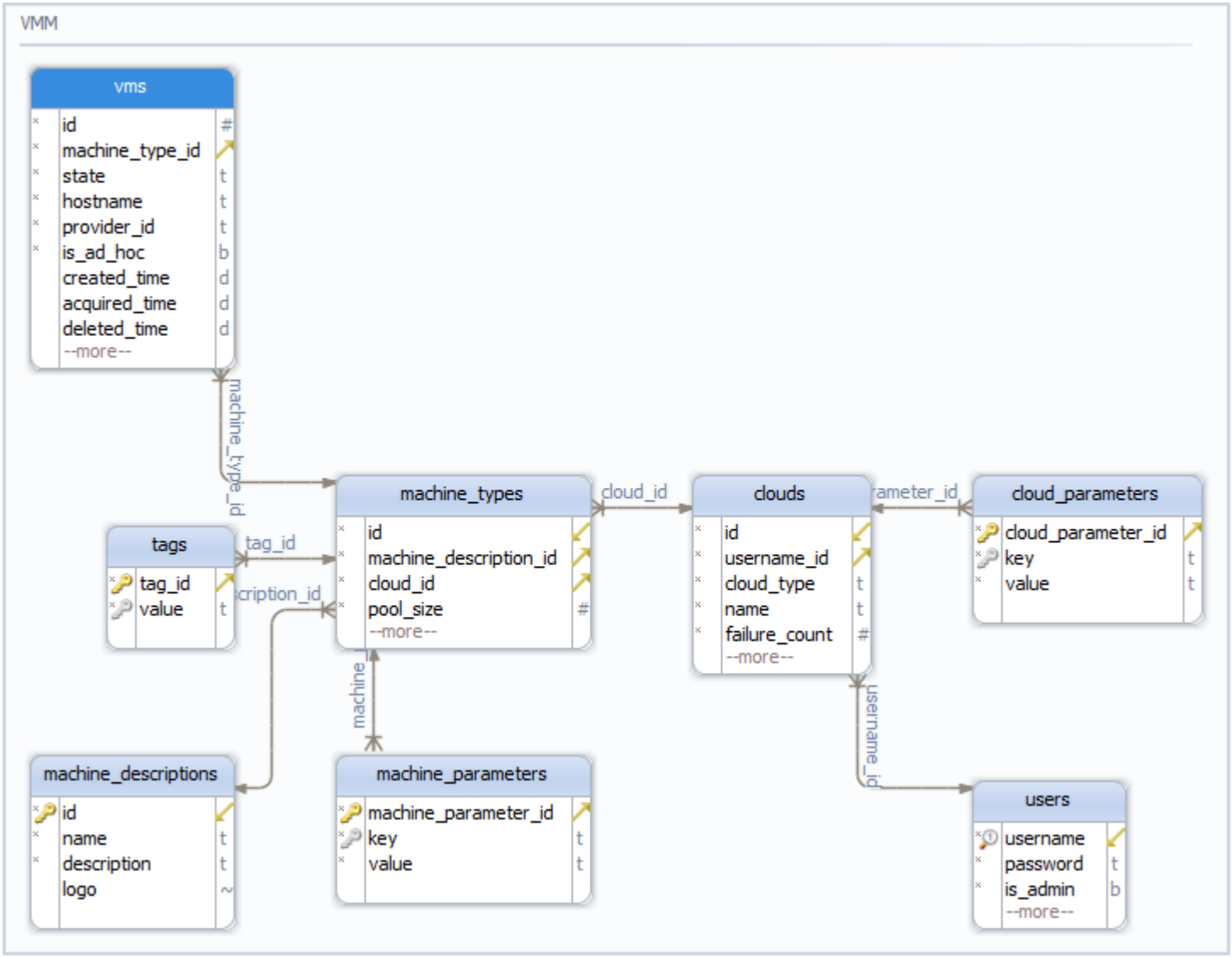
UK Research and Innovation

User Facing Capabilities

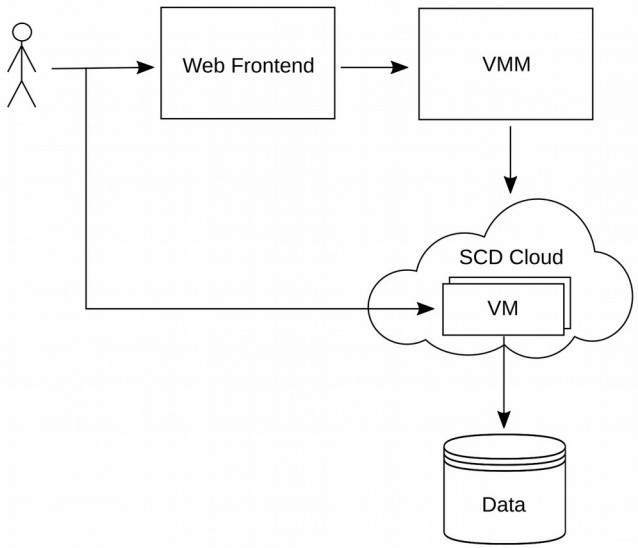


Virtual Machine Manager

- Manages the creation, deletion of VMs
- Provides a web API
- Uses python libcloud
- VM pools
- Machine types – links to environments
- Key Value pairs associated with machine types

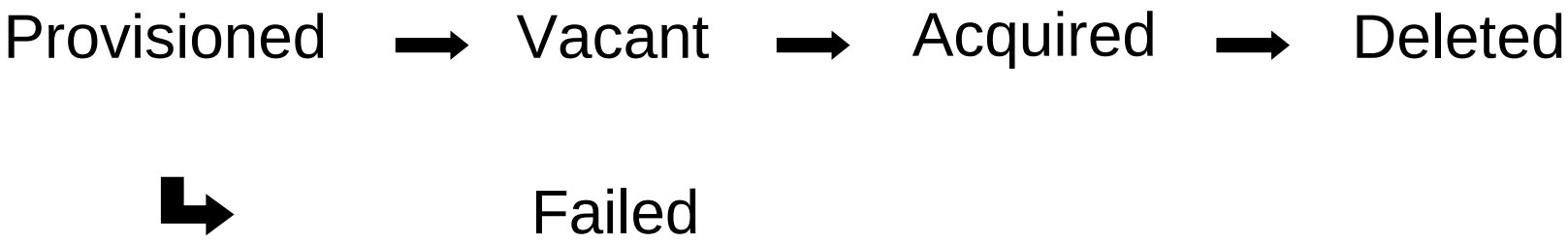

VMM
vms
id
machine_type_id
state
hostname
provider_id
is_ad_hoc
created_time
acquired_time
deleted_time
--more--
machine_type_id
machine_types
id
machine_description_id
cloud_id
pool_size
--more--
cloud_id
clouds
id
username_id
cloud_type
name
failure_count
--more--
rameter_id
cloud_parameters
cloud_parameter_id
key
value
tags
tag_id
value
tag_id
cription_id
machine
machine_descriptions
id
name
description
logo
machine_parameters
machine_parameter_id
key
value
username_id
users
username
password
is_admin
--more--

Virtual Machine Manager

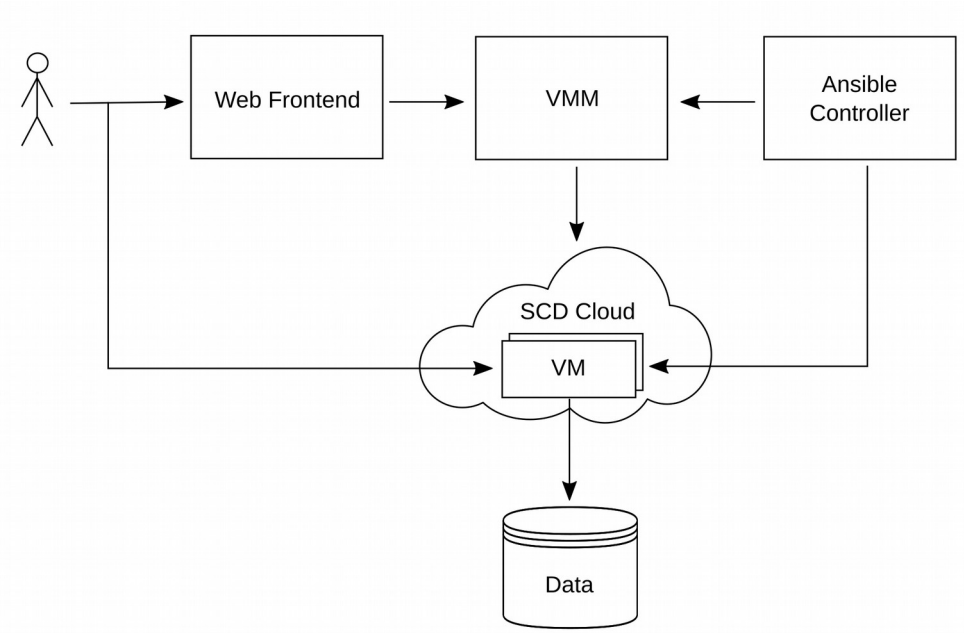


VM Configuration

- All environments based off Centos 7 Mininal
- Completely configured via Ansible
- Service called ansbilecontroller
- Pool management: ansible controller changes state from Provisioned to Vacant once configured



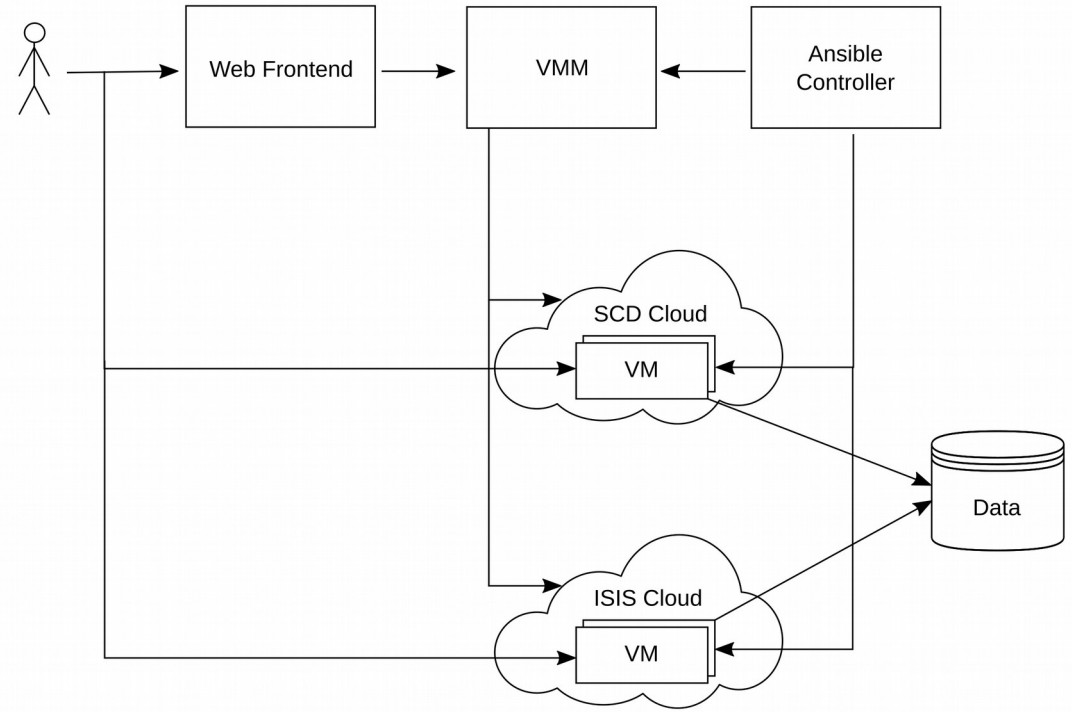
VM Configuration



ISIS Cloud

- Another Openstack cloud
- Provides redundancy and extra capacity
- Add a new cloud entry into the VMM
- Up to web frontend on how to present multiple clouds to user

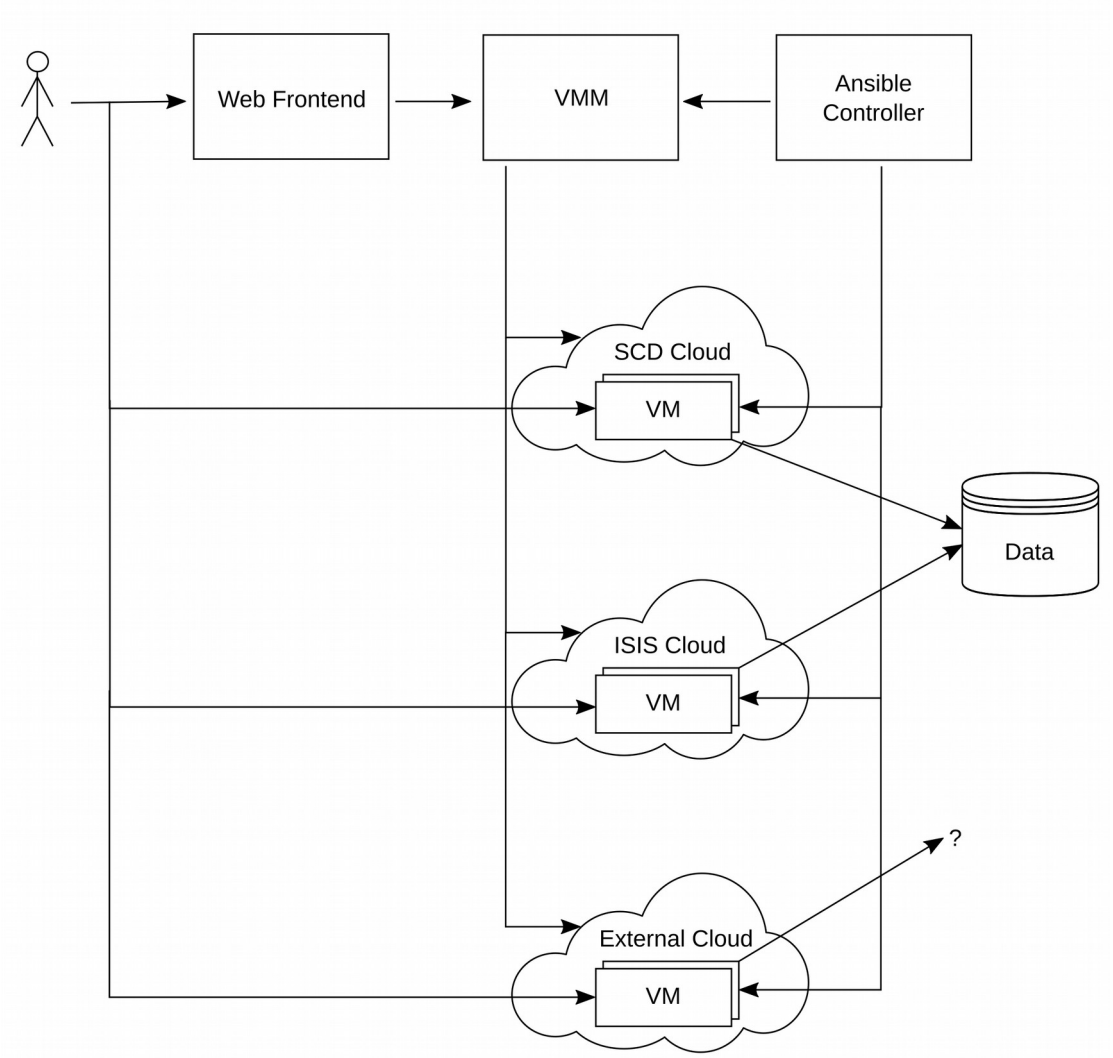
ISIS Cloud


Web Frontend
VMM
Ansible Controller
SCD Cloud
VM
Data
ISIS Cloud
VM

External Clouds

- IRIS Clouds, AWS etc …
- Same as adding ISIS cloud
- Issues around data access

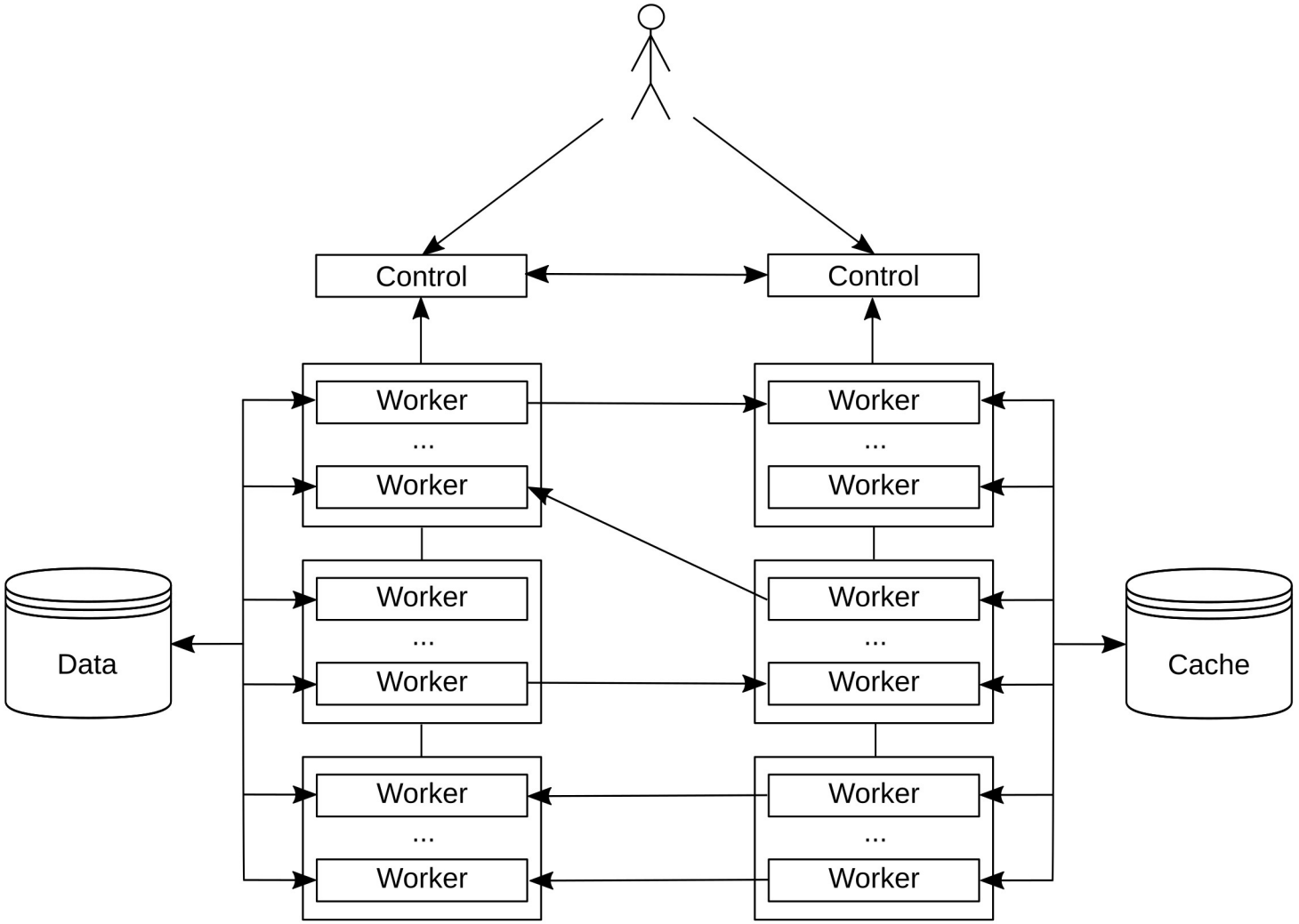
External Clouds



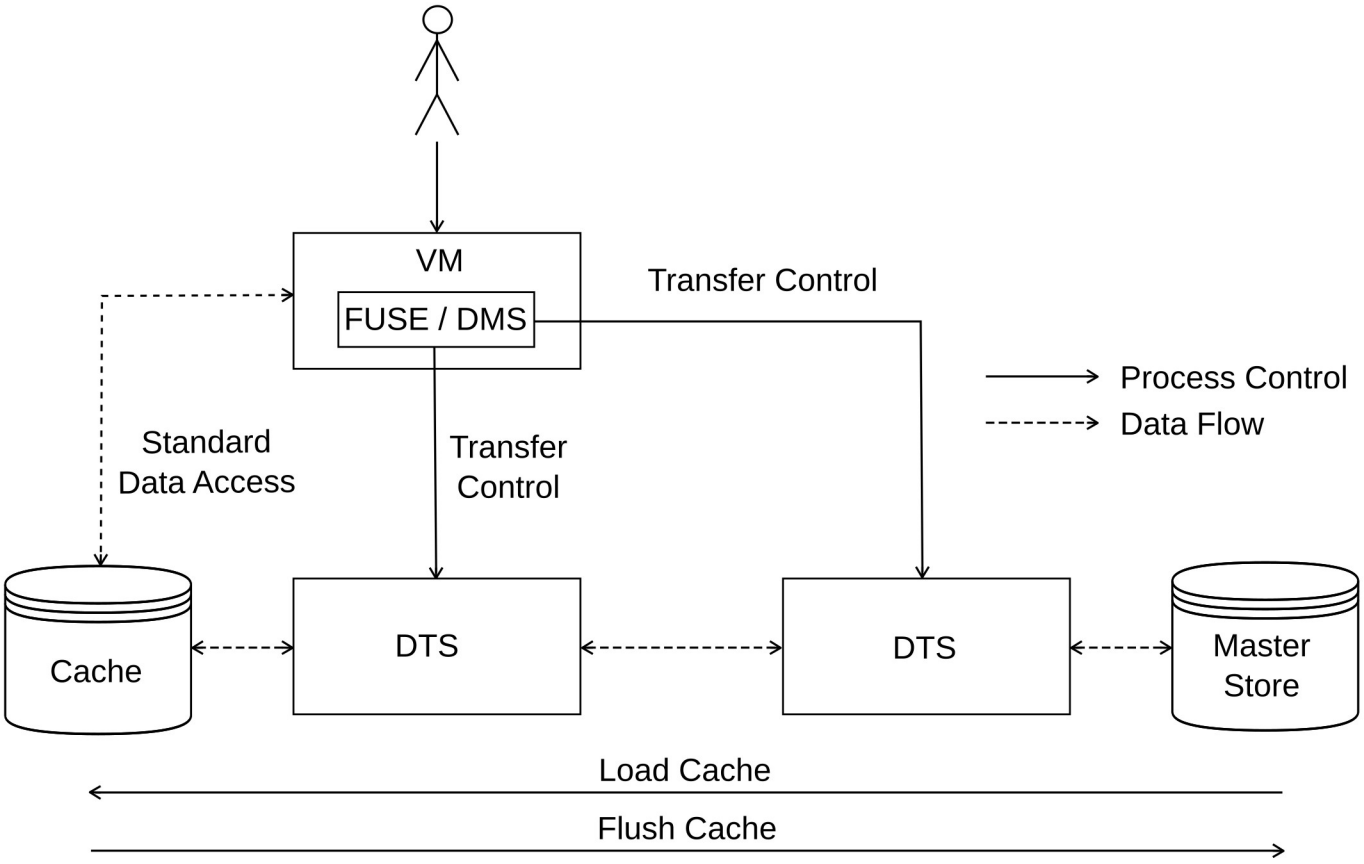
Data Movement Service (DMS) / Data Transfer Service (DTS)

- DTS
  - Python service with web API
  - Inspiration from GridFTP
  - Allows 3rd party transfer
- DMS
  - FUSE client running on VMs
  - Provides a mirror image of remote data store
  - Intelligently caches data in background

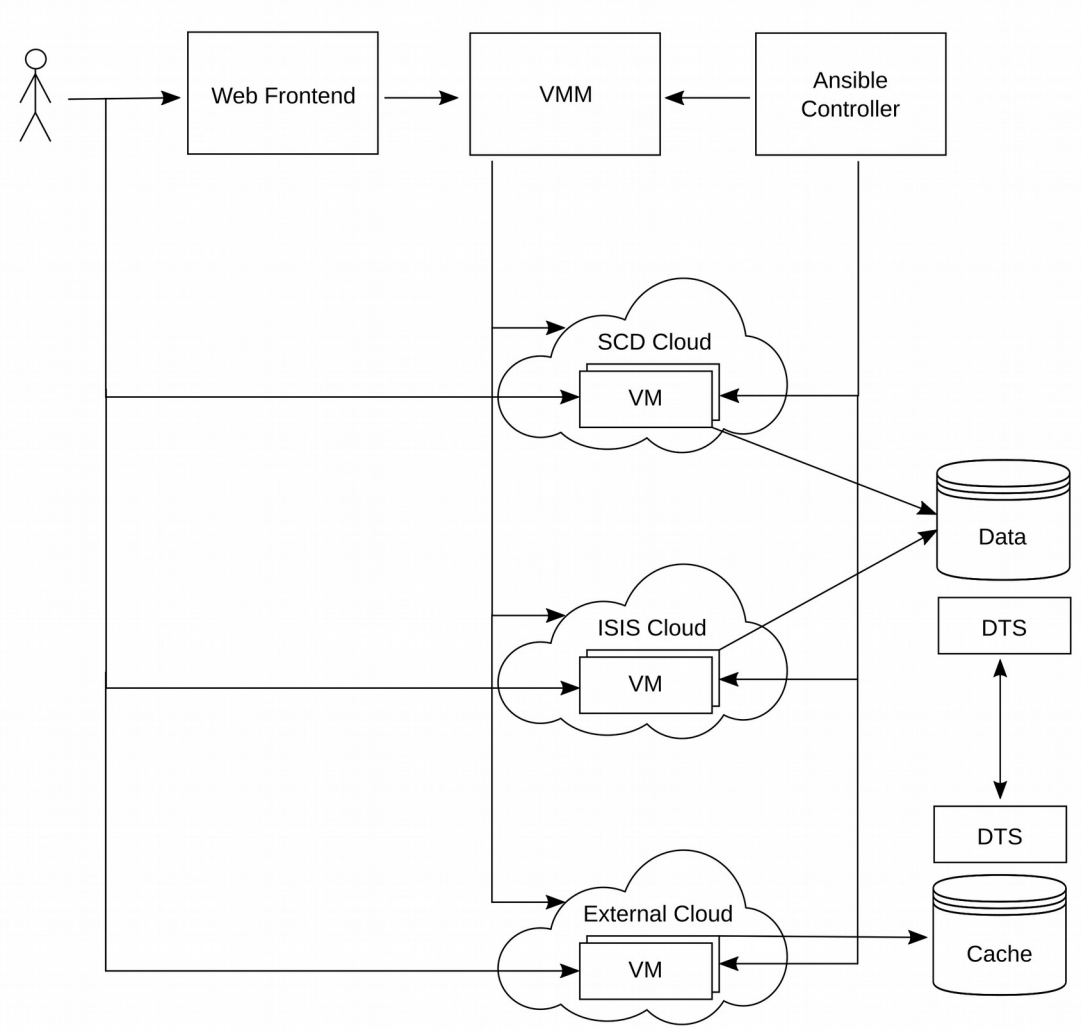
DTS


Control
Control
Worker
...
Worker
Worker
...
Worker
Worker
...
Worker
Worker
...
Worker
Worker
...
Worker
Worker
...
Worker
Data
Cache

DMS


VM
FUSE / DMS
Transfer Control
Process Control
Data Flow
Standard
Data Access
Transfer
Control
Cache
DTS
DTS
Master
Store
Load Cache
Flush Cache

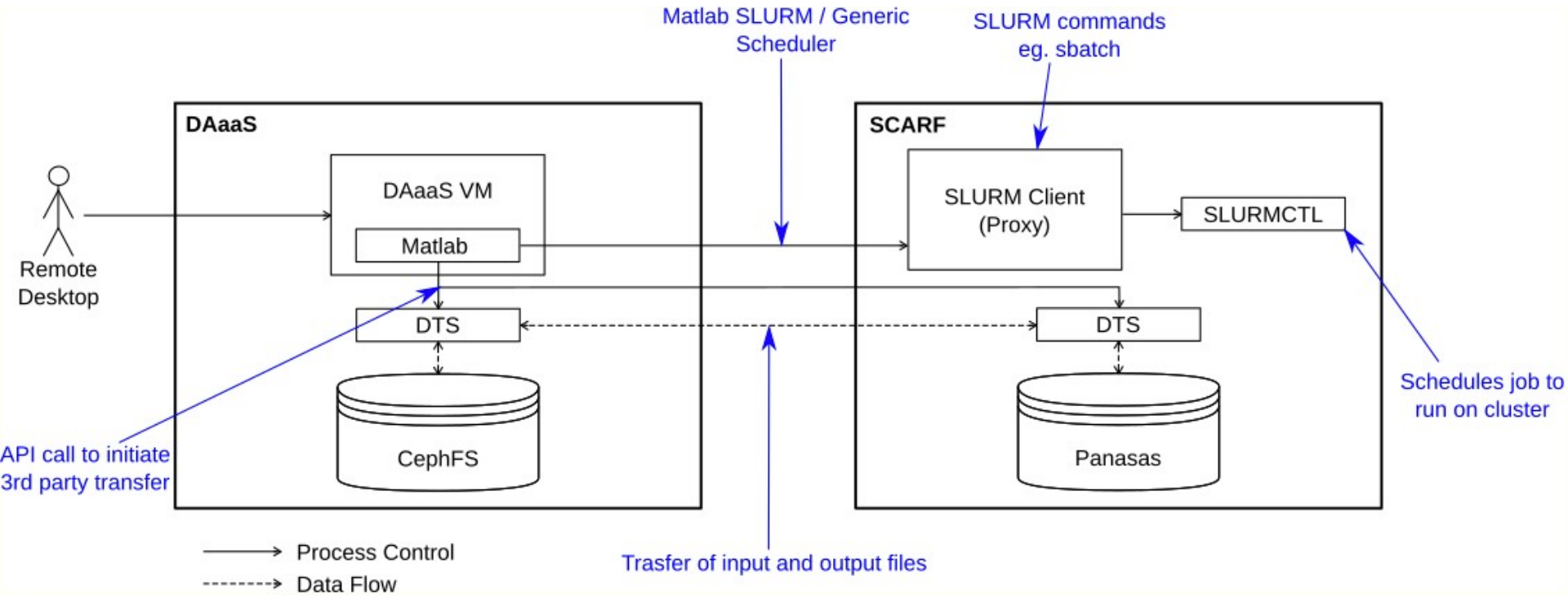
DMS / DTS


Web Frontend
VMM
Ansible Controller
SCD Cloud
VM
ISIS Cloud
VM
External Cloud
VM
Data
DTS
DTS
Cache

HPC

- Looking to link DAaaS VMs to HPC
- Aspiration to make this invisible to user
- Looking to create prototype for ISIS Pace project
- Pace project is linked the ISIS Excitations group (ie. one of the production DAaaS environments)

HPC


Matlab SLURM / Generic Scheduler
SLURM commands eg. sbatch
DAaaS
SCARF
DAaaS VM
Matlab
SLURM Client (Proxy)
SLURMCTL
Remote Desktop
DTS
DTS
CephFS
Panasas
Schedules job to run on cluster
API call to initiate 3rd party transfer
Process Control
Data Flow
Trasfer of input and output files

Next Steps

- VMM going live in production this Thursday
- Integration of ISIS Cloud set for 10th September
- HPC integration – aiming for 6 months

frazer.barnsley@stfc.ac.uk