



# IVOA - science platforms

Describing science software

Dave Morris

IVOA interop meeting  
Sydney, Australia  
May 2024

Dave Morris  
[dave.morris@manchester.ac.uk](mailto:dave.morris@manchester.ac.uk)

**\*\* name change \*\***

~~IVOA Execution Planner~~

**IVOA Execution Broker**



*International  
Virtual  
Observatory  
Alliance*



The functionality has evolved over time

It acts as a broker interface for  
compute platforms

It doesn't perform any planning  
functions itself

It provides the information needed  
to enable others to plan

**IVOA Execution Broker**

**Version 1.0**

**IVOA Working Draft 2024-04-25**

Working Group  
GWS

This version

<https://www.ivoa.net/documents/ExecutionBroker/20240425>

Latest version

<https://www.ivoa.net/documents/ExecutionBroker>

**<https://github.com/ivoa-std/ExecutionBroker>**

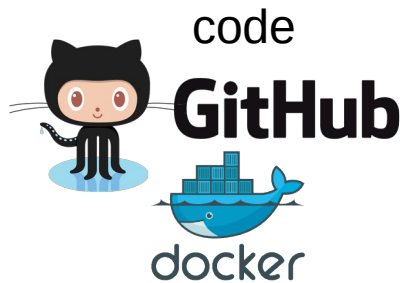
IVOA interop meeting  
Sydney, Australia  
May 2024

Dave Morris  
[dave.morris@manchester.ac.uk](mailto:dave.morris@manchester.ac.uk)

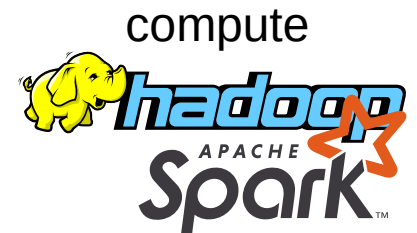
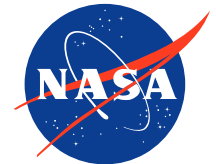
# The problem

Heterogeneous code, data  
and compute.

Everyone is slightly different.



data code



code data

IVOA interop meeting  
Sydney, Australia  
May 2024

Dave Morris  
dave.morris@manchester.ac.uk



## Centralized coordinator

Asking the wrong question

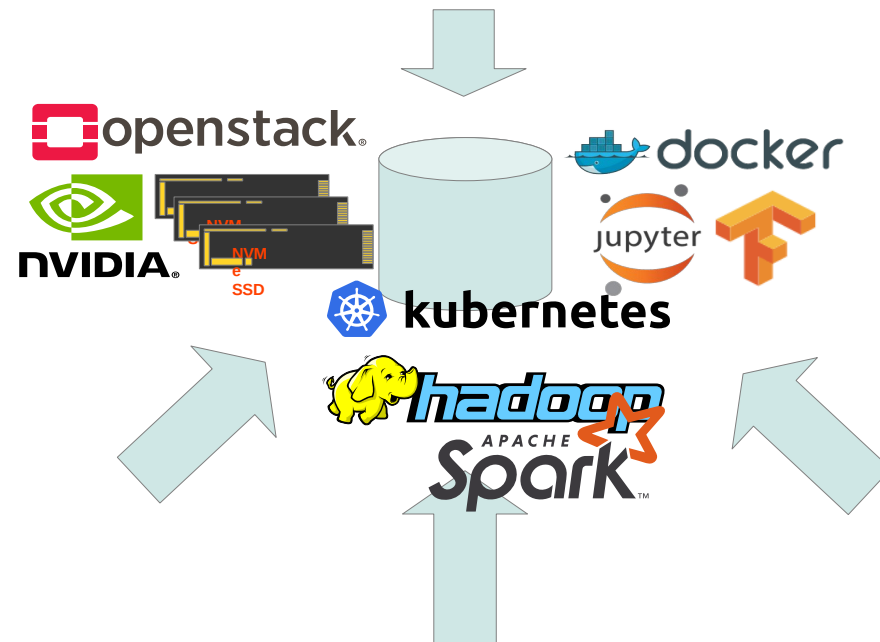
Harvest all the metadata into a central registry.

The central registry would need to understand the details of all the technologies.

It would also need to know the state of all the jobs running on all the compute platforms

**This solution does not scale.**

Find all the platforms that  
**<this> user** can run **<this>**  
task with **<this>** data.



# IVOA ExecutionBroker

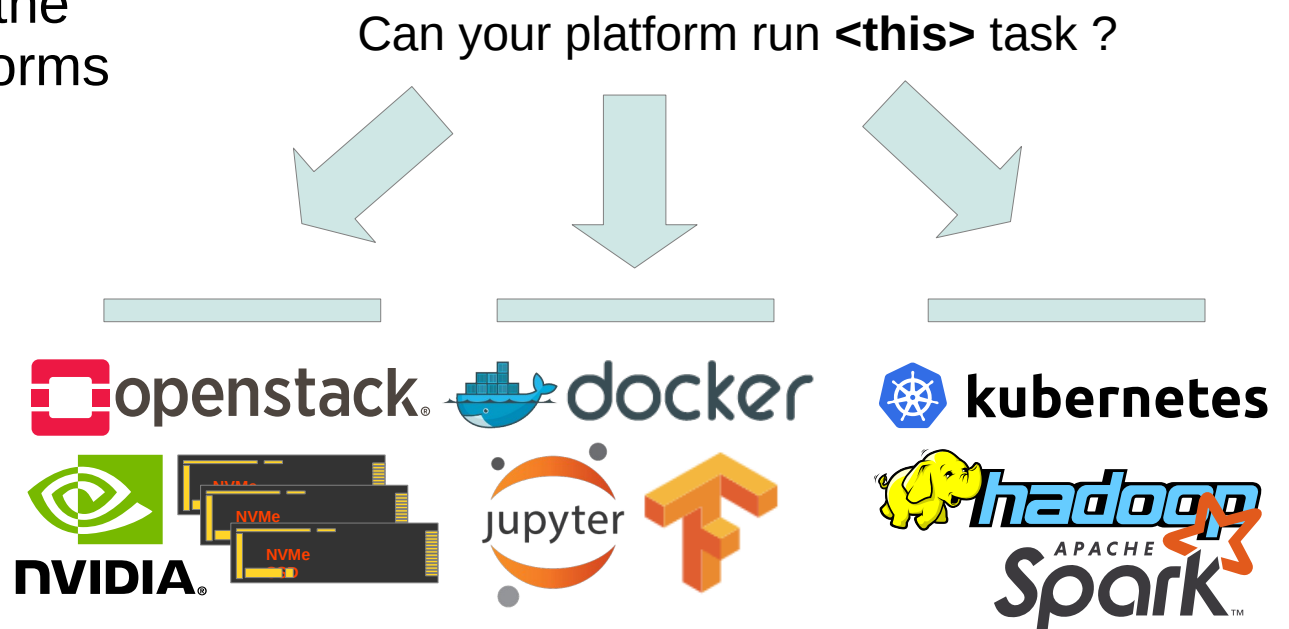


Asking the right question

Alternatively we can delegate the question to the individual platforms

Each platform only needs to understand the technologies it provides.

If a platform doesn't understand the question, it can just say no.



This architecture is much better at scaling to meet changes in capacity and complexity, and at adapting to new technologies.



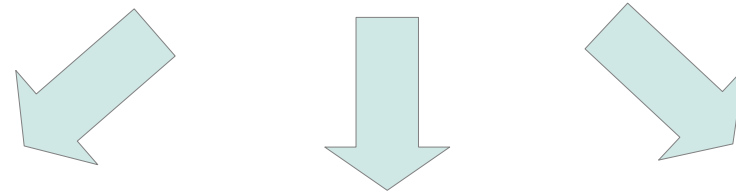
## The science use case

User looks for some data

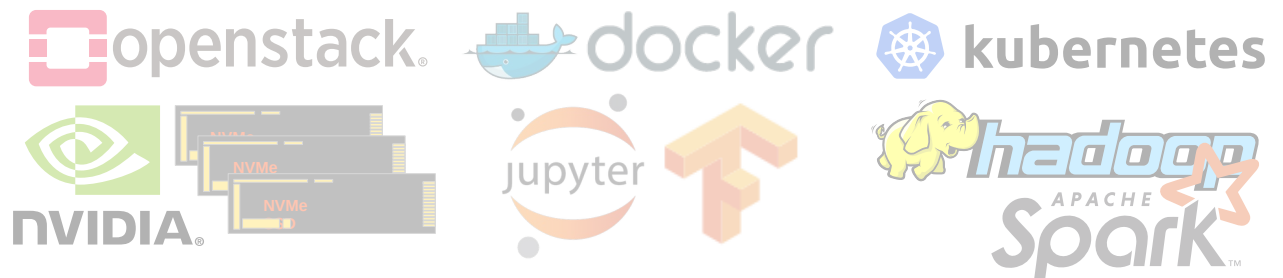
User looks for some software

User looks for a compute platform  
to run the software on their data

Can your platform run **<this>** code  
on **<this>** data ?



ExecutionBroker interface





## The science use case

User looks for some data

User looks for some software

User looks for a compute platform  
to run the software on their data

FAIR data discovery

Common vocabulary

ObsCore DataLink

Can your platform run **<this>** code  
on **<this>** data ?

ExecutionBroker interface

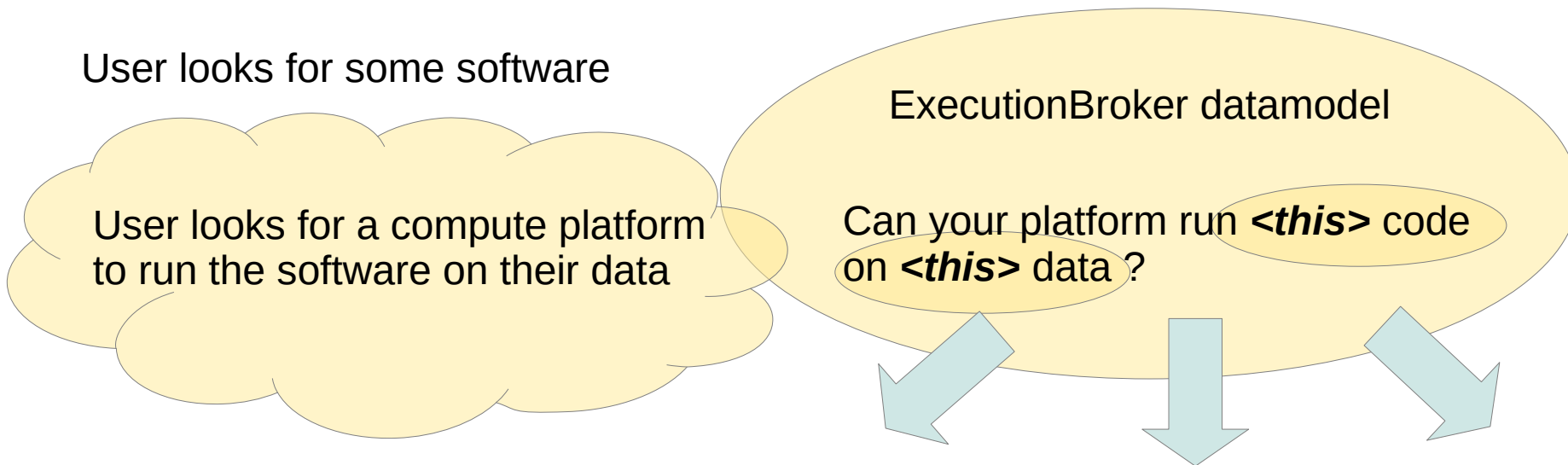




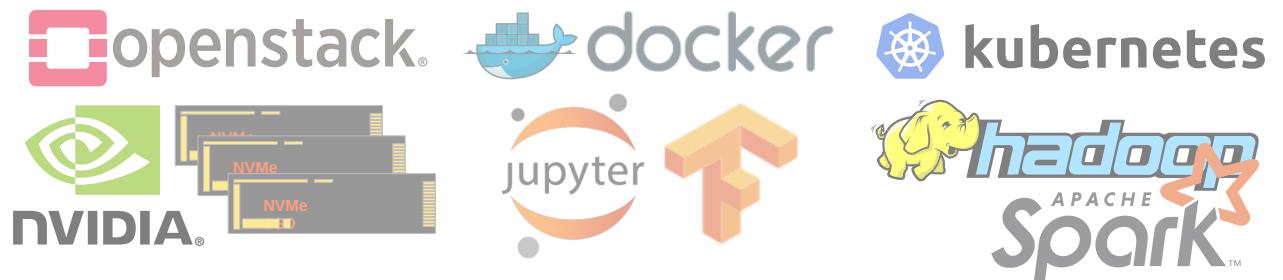
# The science use case

User looks for some data

User looks for some software



ExecutionBroker interface







# The science use case

User looks for some data

User looks for some software

User looks for a compute platform  
to run the software on their data

ExecutionBroker datamodel

Can your platform run **<this>** code  
on **<this>** data ?

Work in progress SKA & CADC (SP-4241)

```
executable:  
  type: "docker-container"  
  spec:  
    repo: "....."  
    image: "....."
```

```
compute-resource:  
  type: "generic-compute"  
  spec:  
    cores:  
      min: 8  
    memory:  
      min: 8M
```

```
data-resource:  
  type: "S3-object"  
  spec:  
    endpoint: "....."  
    bucket: "....."  
    object: "....."
```

IVOA interop meeting  
Sydney, Australia  
May 2024

Dave Morris  
dave.morris@manchester.ac.uk



## The science use case

User looks for some data

User looks for some software

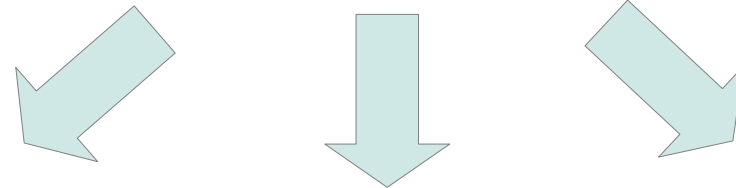
User looks for a compute platform  
to run the software on their data

Code discovery ?

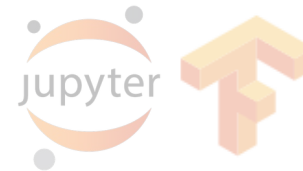
Vocabulary of science functionality

Input and output data types ?

Can your platform run **<this>** code  
on **<this>** data ?



ExecutionBroker interface





## The science use case

User looks for some data

User looks for some software

**a tool**

Vocabulary to describe functionality

What does it **do** ?

**This is what the user is looking for**

Source detection  
Edge detection  
Noise filter  
3D visualization

Code **application** discovery ?

Vocabulary of science functionality

Input and output data types ?

Vocabulary to describe inputs and outputs

What it contains

images  
spectra  
cubes  
visibilities  
measurement sets

What format it is in

jpeg, png, tiff  
FITS  
targz of {n}FITS

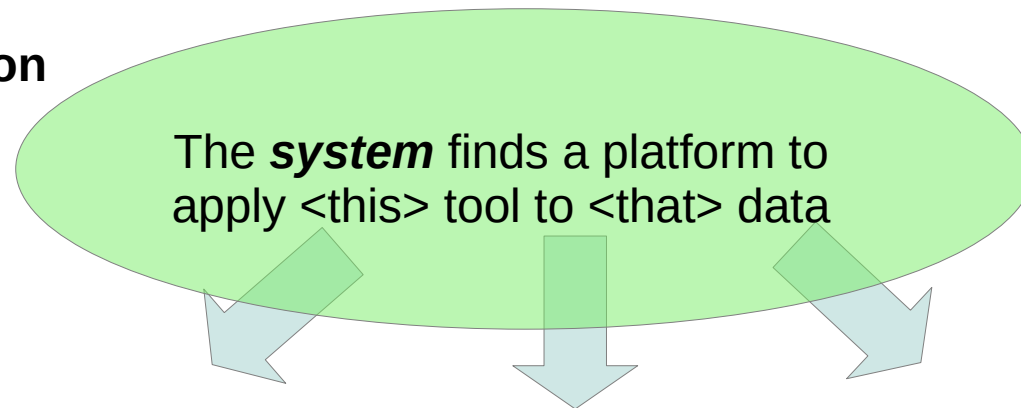


# The science use case

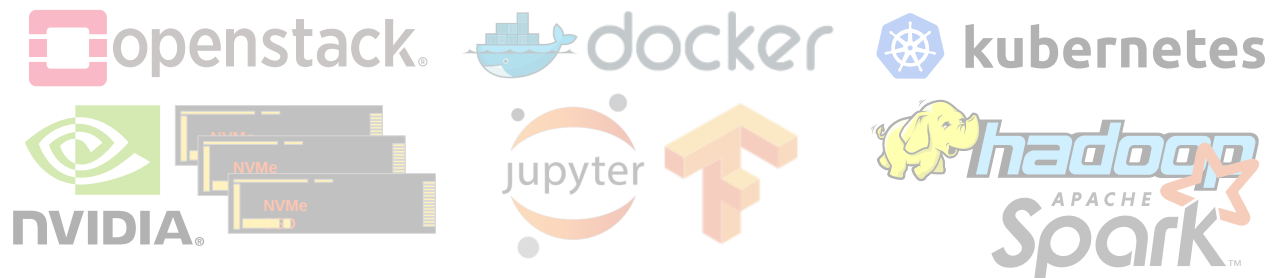
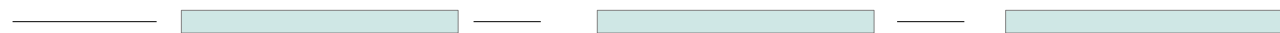
User looks for some data

User looks for ~~some software~~ **a tool**

User presses the big green button



ExecutionBroker interface





## New project – app discovery

Look for prior art.

Look for existing vocabularies.

Start to fill in the missing bits.

All help welcome

Contact Dave Morris

dave.morris@manchester.ac.uk

### Vocabulary to describe functionality

What does it **do** ?

#### ***This is what the user is looking for***

Source detection

Edge detection

Noise filter

3D visualization

### Vocabulary to describe inputs and outputs

What it contains

images

spectra

cubes

visibilities

measurement sets

What format it is in

jpeg, png, tiff

FITS

targz of {n}FITS