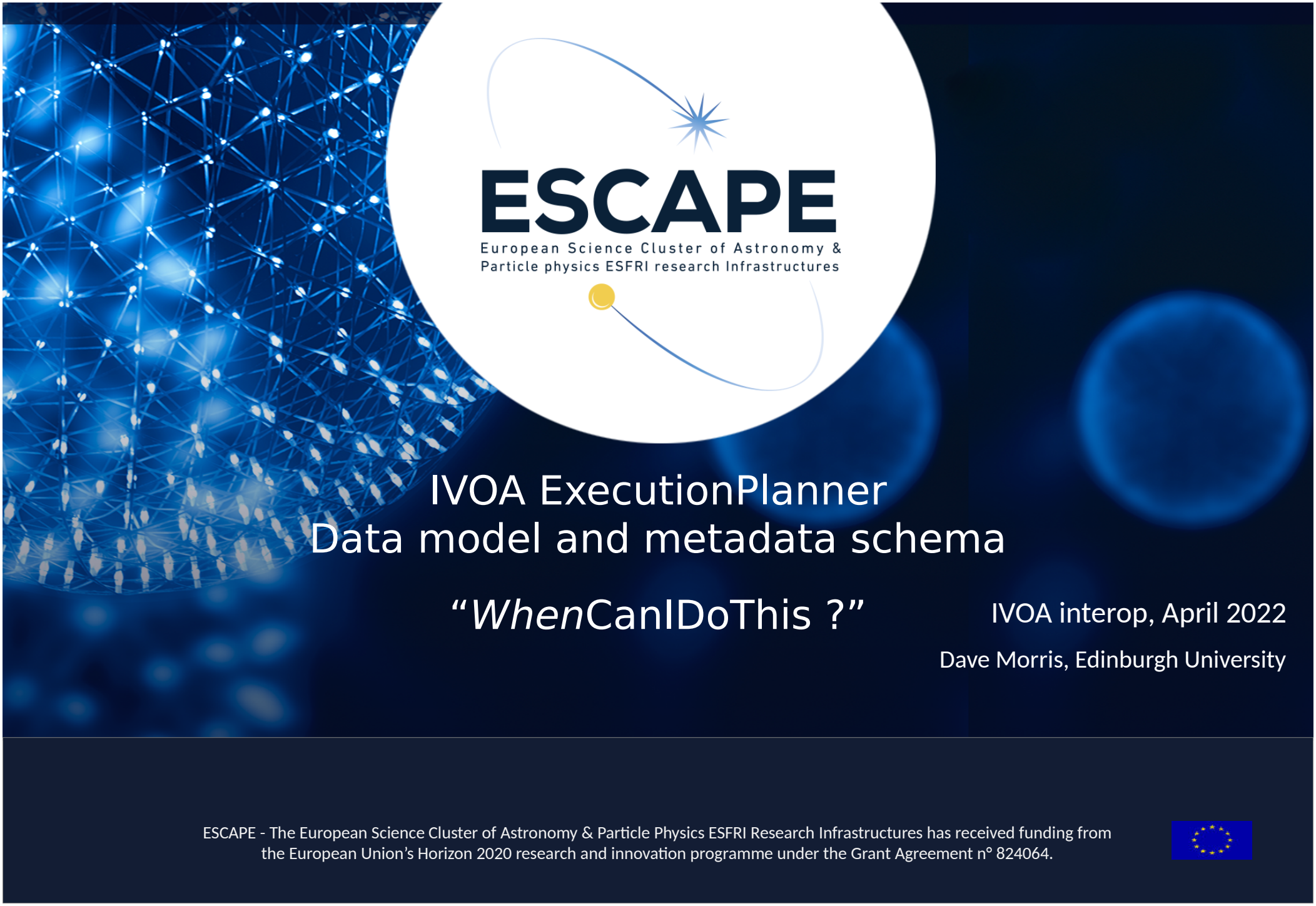


IVOA ExecutionPlanner
Data model and metadata schema

"*When*CanIDoThis ?"

IVOA interop, April 2022

Dave Morris, Edinburgh University

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

The problem – different science platforms use different technologies

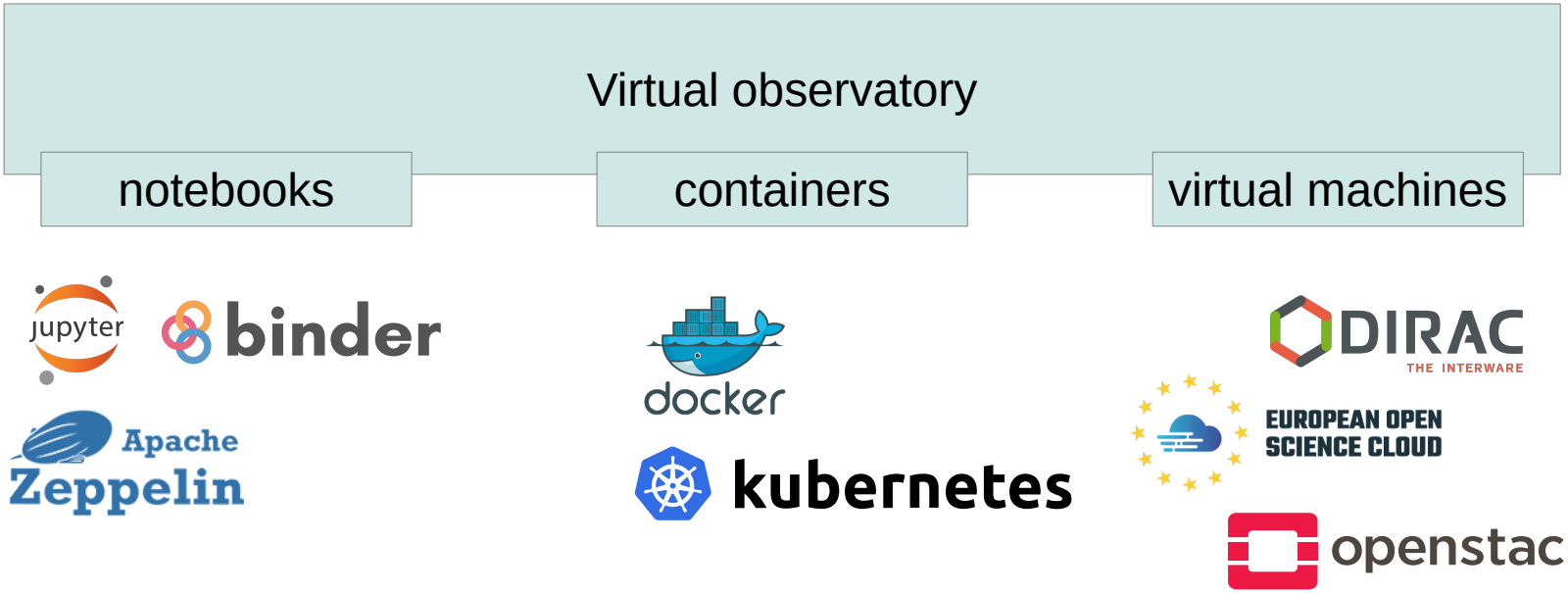
We end up having to understand all of them.

Which becomes more complex as the questions get more detailed.

D.Morris
Institute for Astronomy,
Edinburgh University

The problem – different science platforms use different technologies

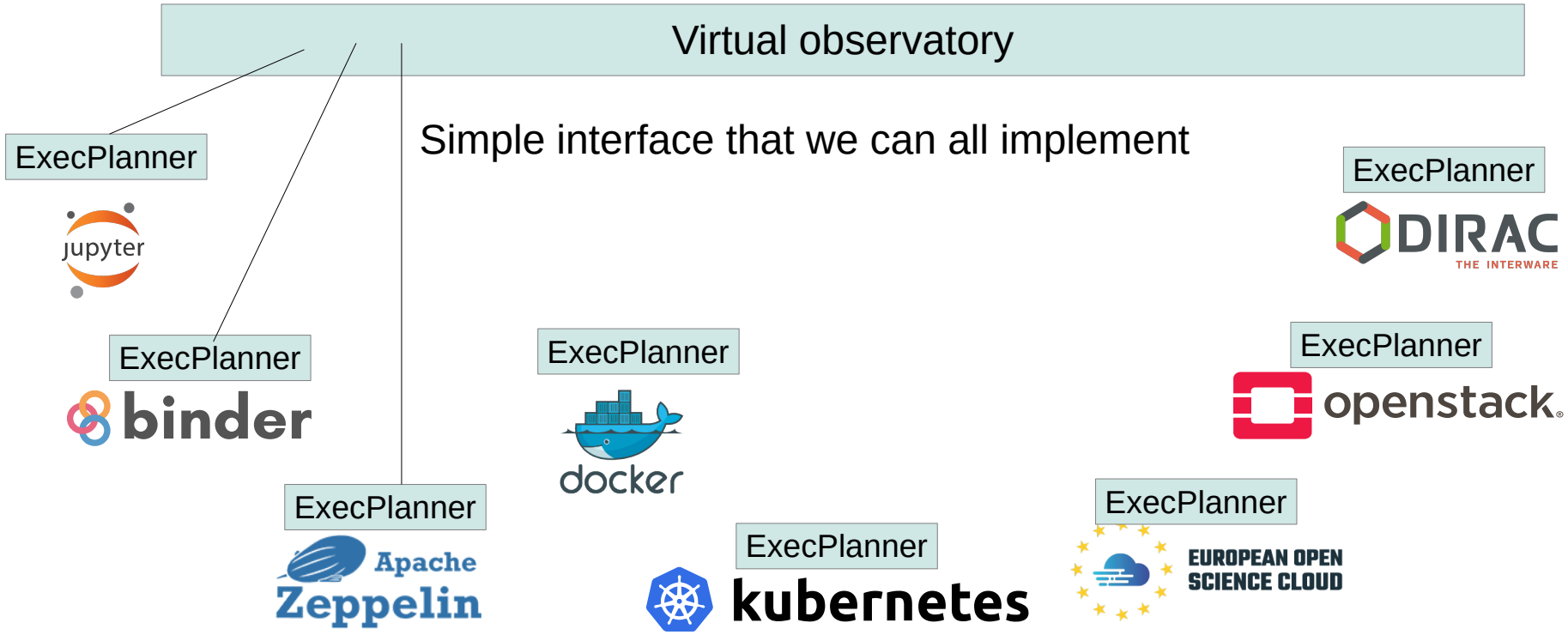
Plugin architecture helps



D.Morris
Institute for Astronomy,
Edinburgh University

The solution

Delegation is the logical extension



D.Morris
Institute for Astronomy,
Edinburgh University

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

Current customers

ESCAPE Science Analysis Platform (ESAP)

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

binder jupyter docker DIRAC THE INTERWARE EUROPEAN OPEN SCIENCE CLOUD openstack

Gaia Data Mining platform (Gaia DMp)

gaia

openstack kubernetes Apache Zeppelin APACHE Spark

Expressions of interest

CANFAR Science Platform

SciServer, JHU

D.Morris
Institute for Astronomy,
Edinburgh University

IVOA interop
April 2022

Simple interface that we can all implement

Common language for describing things

Meeting all the use cases adds complexity

```
{
  "type": "uri://docker-container",
  "task": {
     "image": "docker.io/example:1.0"
    }
  "data-resources": [
       ....
      ]
  "compute-resources": [
       ....
      ]
  "storage-resources": [
       ....
      ]
}
```

New sections added to the schema

Based on ESAP and GaiaDMp use cases

D.Morris
Institute for Astronomy,
Edinburgh University

Simple interface that we can all implement

Common language for describing things

Meeting all the use cases adds complexity

Lower the barrier to entry by making details optional

simple

```
{
"type": "uri://docker-container",
"task": {
   "image": "docker.io/example:1.0"
   }
}
```

complex

```
{
  "type": "uri://docker-container",
  "task": {
     "image": "docker.io/example:1.0"
     }
  "data-resources": [ .... ]
  "compute-resources": [ .... ]
  "storage-resources": [ .... ]
 }
```

D.Morris
Institute for Astronomy,
Edinburgh University

Simple interface that we can all implement

Common language for describing things

Meeting all the use cases adds complexity

Lower the barrier to entry by making details optional

Flexible vocabulary should not be a problem

Customer : Can I buy an <apple>

Shop : No, we are a <bakery>, we only sell <bread>.

Customer : Can I buy an <apple>

Shop : Yes, we can offer <5> different varieties of <apple>.

D.Morris
Institute for Astronomy,
Edinburgh University

Simple interface that we can all implement

Common language for describing things

Meeting all the use cases adds complexity

Lower the barrier to entry by making details optional

Adding more detail should not be a problem

Customer : Can I buy an <iPad> with <64G memory> and <5G network>

Shop : No, we are a <bakery>, we only sell <bread>.

Customer : Can I buy an <iPad> with <64G memory> and <5G network>

Shop : Yes, we can offer 2 different payment contracts.

D.Morris
Institute for Astronomy,
Edinburgh University

ROYAL OBSERVATORY EDINBURGH

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

Working with limited resources

Small task, large cloud

Simple answer

YES

docker

amazon web services

D.Morris
Institute for Astronomy,
Edinburgh University

THE UNIVERSITY OF EDINBURGH

IVOA interop
April 2022

ROYAL OBSERVATORY EDINBURGH

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

Working with limited resources

Large task, limited cloud

Not so simple

Depends ……

Apache Zeppelin

APACHE Spark

openstack.

D.Morris
Institute for Astronomy,
Edinburgh University

ROYAL OBSERVATORY EDINBURGH

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

Working with limited resources

Big data, complex analysis

gaia

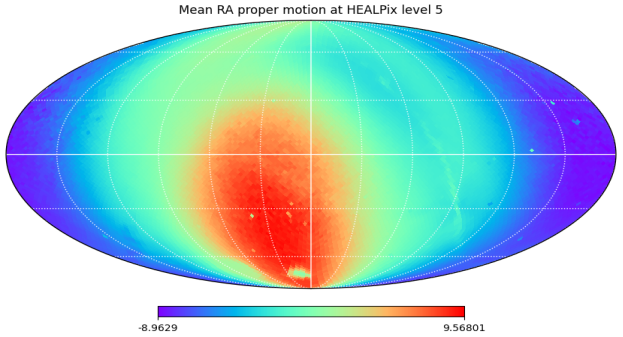
Gaia Data Mining Platform (Gaia DMp)

3.7 Tbytes of numerical data

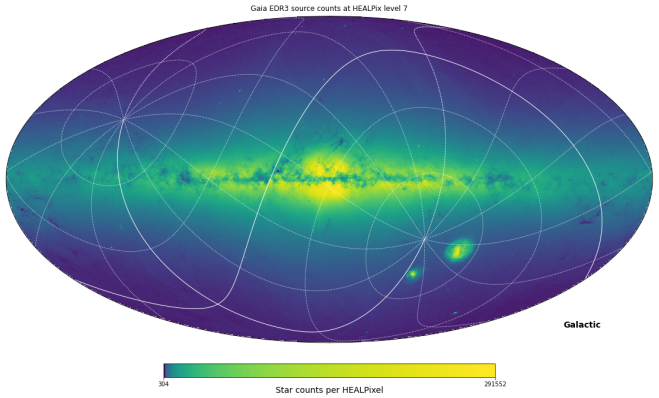
Zeppelin - 54 cores, 86G memory

Spark - 6x 26 cores, 43G memory

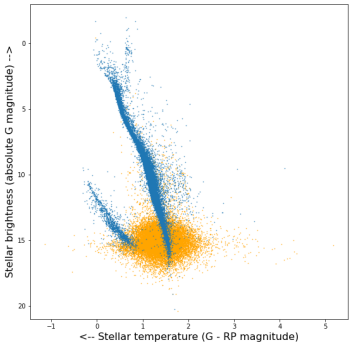
> 9hrs for a complex analysis



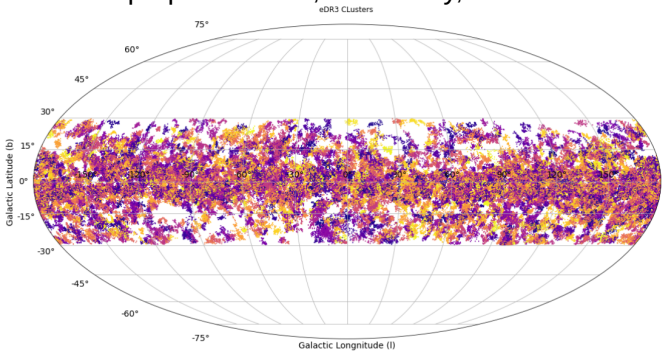
Mean proper motions, N. Hambly, 2022



Mean proper motions, N. Hambly, 2022



Stellar classifier N. Hambly, D. Crake 2022



HDBSCAN Clustering, D. Crake, 2022

D.Morris
Institute for Astronomy,
Edinburgh University

THE UNIVERSITY OF EDINBURGH

IVOA interop
April 2022

ROYAL OBSERVATORY EDINBURGH

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

Working with limited resources

Gaia DMp

gaia

3.7 Tbytes of numerical data

Zeppelin - 54 cores, 86G memory

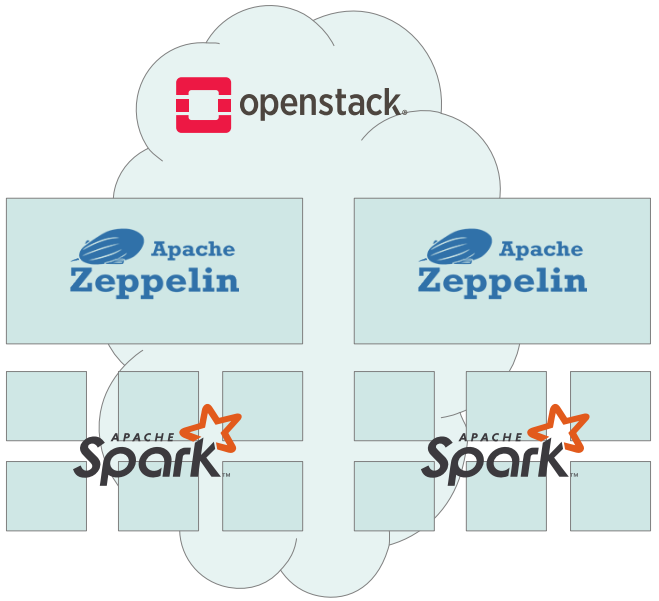
Spark - 6x 26 cores, 43G memory

> 9hrs for a complex analysis

Limited cloud

> 50% of the available resources

We can’t run more than one at a time



D.Morris
Institute for Astronomy,
Edinburgh University

THE UNIVERSITY OF EDINBURGH

IVOA interop
April 2022

When can I do this ?

Gaia DMp

3.7 Tbytes of numerical data

Zeppelin - 54 cores, 86G memory

Spark - 6x 26 cores, 43G memory

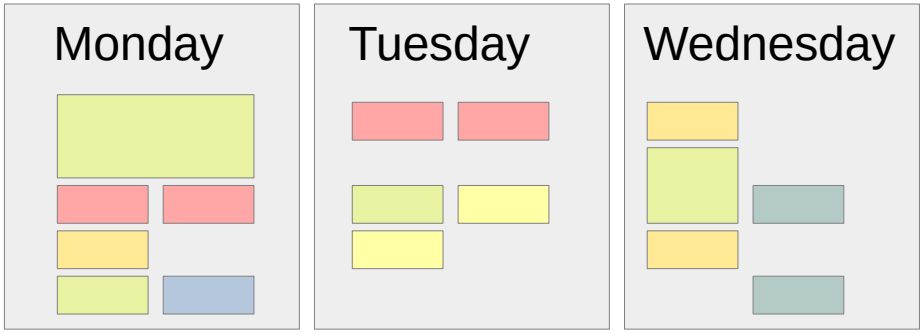
> 9hrs for a complex analysis

Booking system – resource leases

Offer : today at 14:00

Offer : today at 21:00

Offer : tomorrow at 8:00

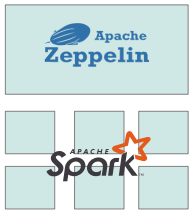


D.Morris
Institute for Astronomy,
Edinburgh University

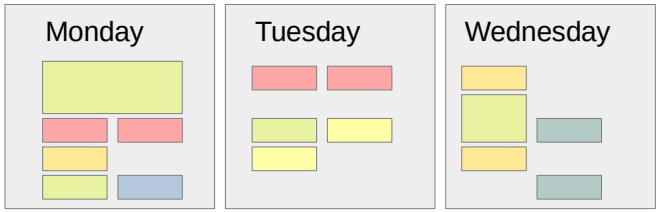
3 levels of detail

Can I do this ***simple thing*** ?

Can I do this ***complex thing***?



When can I do this complex thing ?



D.Morris
Institute for Astronomy,
Edinburgh University

ROYAL OBSERVATORY EDINBURGH

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

3 levels of detail

Can I do this ***simple thing*** ?

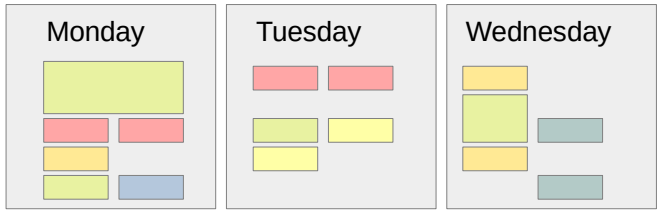
jupyter

Can I do this ***complex thing***?

Apache Zeppelin

Apache Spark

When can I do this complex thing ?



ESCAPE

ESAP is here

binder jupyter DIRAC THE INTERWARE docker EUROPEAN OPEN SCIENCE CLOUD openstack

gaia

Gaia DMp needs all three levels

Implemented in stages

2022 - 2023

Apache Zeppelin openstack

Apache Spark kubernetes

D.Morris
Institute for Astronomy,
Edinburgh University

THE UNIVERSITY OF EDINBURGH

IVOA interop
April 2022

ROYAL OBSERVATORY EDINBURGH

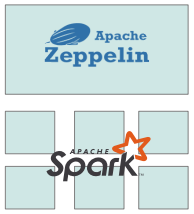
ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

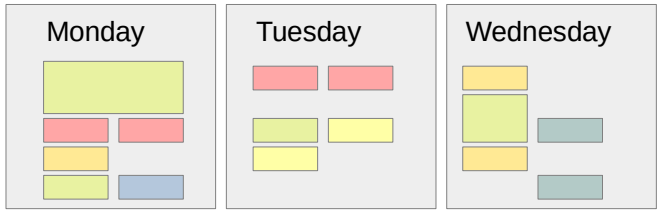
3 levels of detail

Can I do this ***simple thing*** ?

Can I do this ***complex thing***?



When can I do this complex thing ?



Are you interested in joining us ?

What level matches your use case ?

Hackathon session

dmr@roe.ac.uk

Dave Morris

D.Morris
Institute for Astronomy,
Edinburgh University

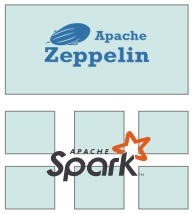
Extra slides added after the GWS session ...

D.Morris
Institute for Astronomy,
Edinburgh University

Separate the interfaces

Can I do this ***simple thing*** ?

Can I do this ***complex thing***?



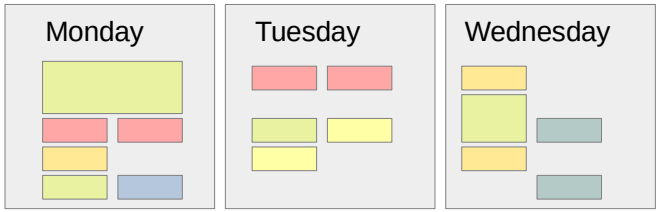
Original stateless ExecPlanner API

CanIDoThis?

Useful for unattended batch processing

- Jobs scheduled on a queue

When can I do this complex thing ?



New statefull ExecScheduler API

When*CanIDoThis?*

Useful for interactive analysis

- Jobs scheduled on a calendar

D.Morris
Institute for Astronomy,
Edinburgh University

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

ESCAPE OSSR | Open-source Scientific Software and Service Repository

ESCAPE ESAP | ESFRI Science Analysis Platform

ESAP plugin

openstack python client

Task metadata

```
{
  "type": "uri://openstack-deployment",
  "data-resources": [
      ....
      ]
  "compute-resources": [
      ....
     ]
  "storage-resources": [
      ....
     ]
}
```

openstack

Apache Zeppelin

Apache Spark

D.Morris
Institute for Astronomy,
Edinburgh University

IAA-CSIC access policy

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

ESCAPE OSSR | Open-source Scientific Software and Service Repository

ESCAPE ESAP | ESFRI Science Analysis Platform

ESAP plugin

openstack python client

Task metadata

```
{
  "type": "uri://openstack-deployment",
  "data-resources": [
      ....
      ]
  "compute-resources": [
      ....
     ]
  "storage-resources": [
      ....
     ]
}
```

openstack

Apache Zeppelin

Apache Spark

D.Morris
Institute for Astronomy,
Edinburgh University

IAA-CSIC access policy

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

ESCAPE OSSR | Open-source Scientific Software and Service Repository

ESCAPE ESAP | ESFRI Science Analysis Platform

Task metadata

```
{
  "type": "uri://openstack-deployment",
  "data-resources": [
      ....
      ]
  "compute-resources": [
      ....
     ]
  "storage-resources": [
      ....
     ]
}
```

ESAP plugin

The Astropy Project

UWS client

UWS service

openstack.

python client

openstack.

Apache Zeppelin

APACHE Spark

D.Morris
Institute for Astronomy,
Edinburgh University

IVOA interop
April 2022

IAA-CSIC access policy

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

IVOA

ESCAPE OSSR | Open-source Scientific Software and Service Repository

ESCAPE ESAP | ESFRI Science Analysis Platform

Task metadata

```
{
  "type": "uri://openstack-deployment",
  "data-resources": [
      ....
      ]
  "compute-resources": [
      ....
     ]
  "storage-resources": [
      ....
     ]
}
```

ESAP plugin

docker

The Astropy Project

UWS client

docker

UWS service

openstack

python client

openstack

Apache Zeppelin

APACHE Spark

D.Morris
Institute for Astronomy,
Edinburgh University

IVOA interop
April 2022

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

ESCAPE OSSR | Open-source Scientific Software and Service Repository

ESCAPE ESAP | ESFRI Science Analysis Platform

Task metadata

```
{
  "type": "uri://openstack-deployment",
  "data-resources": [
      ....
      ]
  "compute-resources": [
      ....
     ]
  "storage-resources": [
      ....
     ]
}
```

docker

The Astropy Project

UWS client

docker

UWS service

openstack.

python client

openstack.

Apache Zeppelin

Apache Spark

D.Morris
Institute for Astronomy,
Edinburgh University

IVOA interop
April 2022

ESCAPE
European Science Cluster of Astronomy & Particle physics ESFRI research Infrastructures

ESCAPE OSSR | Open-source Scientific Software and Service Repository

ESCAPE ESAP | ESFRI Science Analysis Platform

Task metadata

```
{
  "type": "uri://openstack-deployment",
  "data-resources": [
      ....
      ]
  "compute-resources": [
      ....
     ]
  "storage-resources": [
      ....
     ]
}
```

docker

The Astropy Project

UWS client

docker

UWS service

EUROPEAN OPEN SCIENCE CLOUD

python client

EUROPEAN OPEN SCIENCE CLOUD

Apache Zeppelin

APACHE Spark

D.Morris
Institute for Astronomy,
Edinburgh University

IVOA interop
April 2022