

Tracking provenance information with the OPUS job controller

Mathieu Servillat

Observatoire de Paris - LUTH
Paris Astronomical Data Centre



Laboratoire Univers et Théories



OPUS: Observatoire de Paris UWS Server

- **Main features**
- IVOA standards
 - Universal Worker System (**UWS**)
 - **Provenance** data model
- REST architecture
 - Python micro-framework: bottle.py
- Collaborative development
 - Git server at PADC (gitolite)
 - GitHub: <https://github.com/mservillat/OPUS>
- Tests and quality
 - Unit tests with `unittest` and `webtest`
 - Activity history with `logging`
- **Prototype available**
 - <https://voparis-uws-test.obspm.fr>



Job management at PADC

Available structure

- Work cluster (Tycho)
- Job scheduler (SLURM)



PADC projects

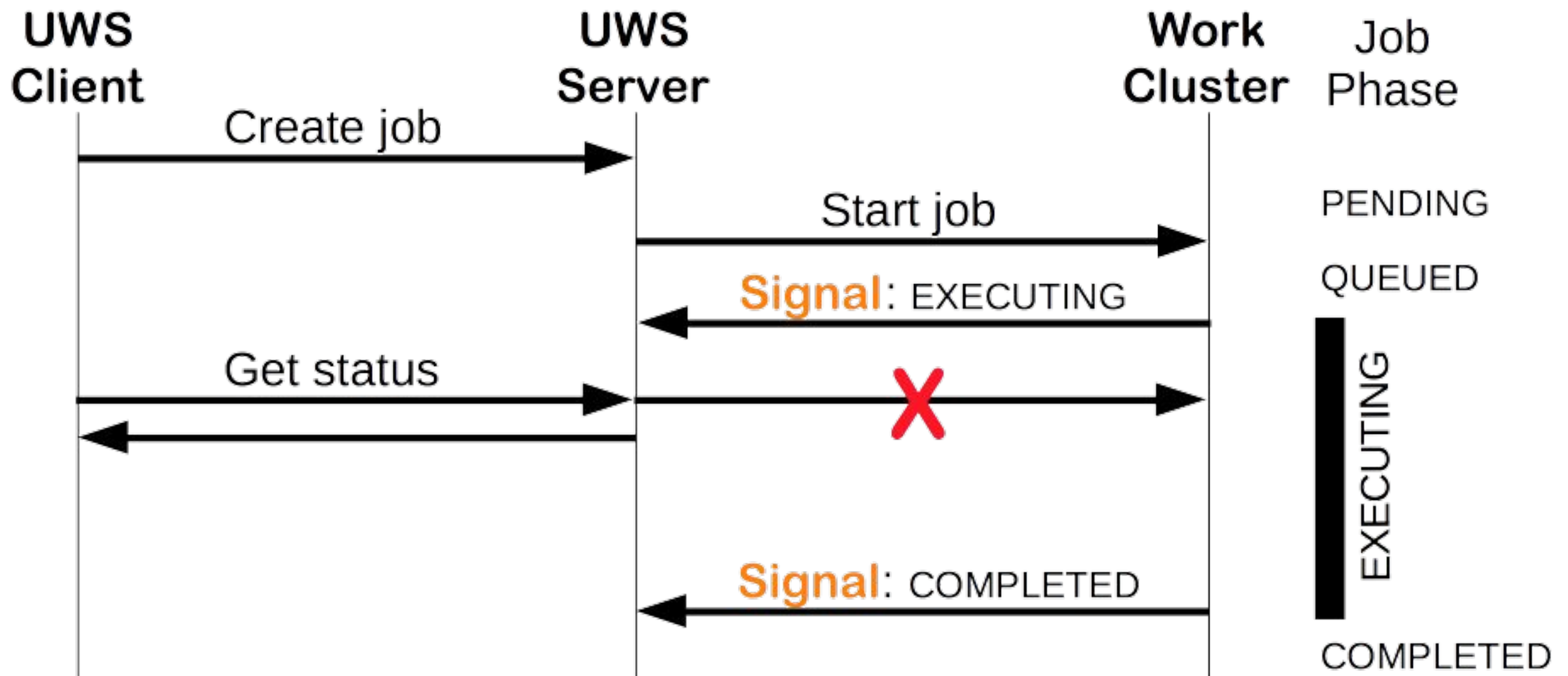
- Web based clients
 - Data access
 - Online data processing
 - Wrap simulation codes



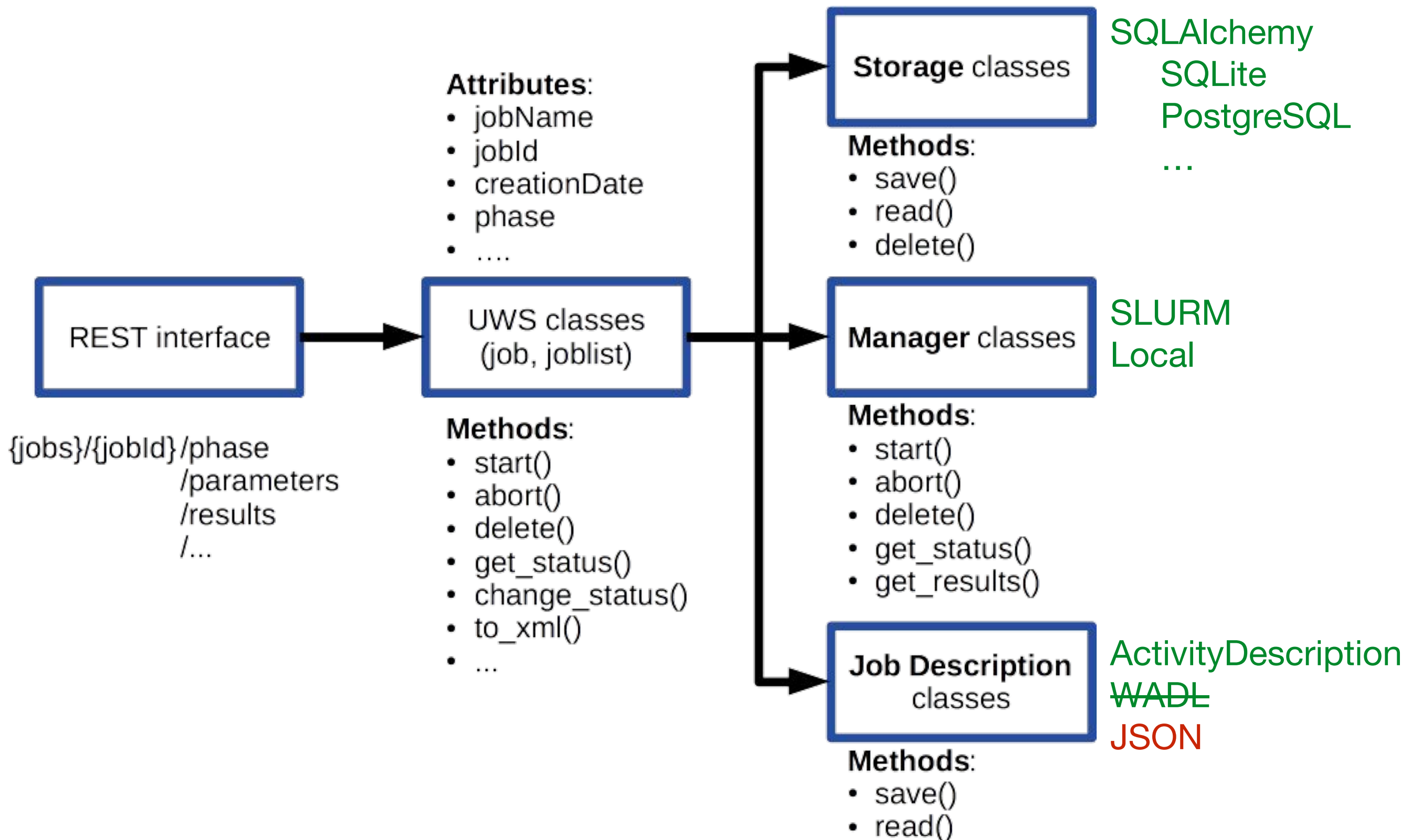
Need a simple interface to computational resources

Inner workings

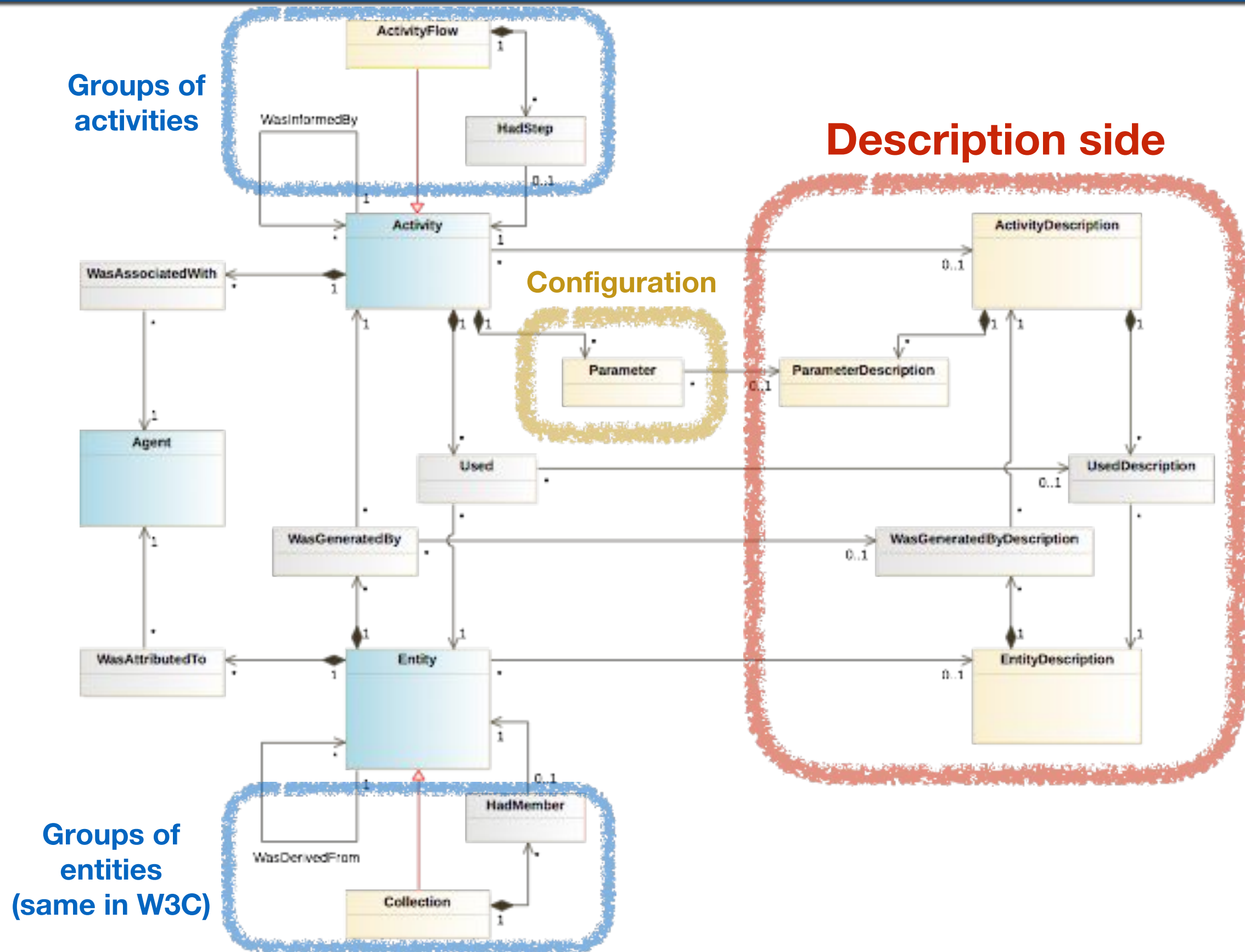
- Separate **job controller** from **work cluster**
 - Wait for work cluster **signals**
 - Avoid (too many) status queries to work cluster



Server main classes



IVOA Provenance Data Model diagram



Serializations - ActivityDescription

```
<RESOURCE ID="gammapy_maps" name="gammapy_maps" type="meta" utype="voprov:ActivityDescription">
```

```
  <DESCRIPTION>Use gammapy to generate a count map from a list of observations</DESCRIPTION>
```

```
  <!-- Service Descriptor -->
```

```
  <PARAM name="accessURL" datatype="char" arraysize="*" value="https://voparis-uws-test/rest/gammapy_maps" />
```

```
  <PARAM name="standardID" datatype="char" arraysize="*" value="ivo://ivoa.net/std/SODA#1.0" />
```

```
  <!-- Activity Description -->
```

```
  <PARAM name="type" datatype="char" arraysize="*" value="None" utype="voprov:ActivityDescription.type"/>
```

```
  <PARAM name="subtype" datatype="char" arraysize="*" value="None" utype="voprov:ActivityDescription.subtype"/>
```

```
  <PARAM name="annotation" datatype="char" arraysize="*" value="Use gammapy to generate a count map from a list of
```

```
  <PARAM name="version" datatype="char" arraysize="*" value="None" utype="voprov:ActivityDescription.version"/>
```

```
  <PARAM name="doculink" datatype="char" arraysize="*" value="https://luthgitlab.obspm.fr/jlefaucheur/hess_release
```

```
  <PARAM name="contact_name" datatype="char" arraysize="*" value="Julien Lefaucheur" utype="voprov:Agent.name"/>
```

```
  <PARAM name="contact_email" datatype="char" arraysize="*" value="" utype="voprov:Agent.email"/>
```

```
  <!-- UWS job attributes -->
```

```
  <PARAM name="executionDuration" datatype="int" value="600" utype="uws:Job.executionDuration"/>
```

```
  <PARAM name="quote" datatype="int" value="120" utype="uws:Job.quote"/>
```

```
  <!-- UWS parameters (Provenance Entities or Parameters) -->
```

```
  <GROUP name="InputParams">
```

```
    <PARAM ID="obs_ids" arraysize="*" datatype="char" name="obs_ids" value="47802 47803 47804
```

```
      <DESCRIPTION>List of runs</DESCRIPTION>
```

```
    </PARAM>
```

```
    <PARAM ID="RA" datatype="double" name="RA" value="329.7169379" unit="deg"...>
```

```
    <PARAM ID="Dec" datatype="double"
```

```
    <PARAM ID="nxpix" arraysize="*" da
```

```
      <DESCRIPTION>Number of pixels
```

```
      <VALUES>
```

```
        <MIN value="0"/>
```

```
        <MAX value="1000"/>
```

```
      </VALUES>
```

```
    </PARAM>
```

```
    <PARAM ID="nypix" arraysize="*" da
```

```
    <PARAM ID="binsz" datatype="float"
```

```
  </GROUP>
```

```
  <!-- Used Entities -->
```

```
  <GROUP name="Used">
```

```
    <GROUP name="obs_ids" utype="voprov:UsedDescription" ref="obs_ids">
```

```
      <PARAM arraysize="*" datatype="char" name="role" utype="voprov:UsedDescription.role" value="DL3"/>
```

```
      <PARAM arraysize="*" datatype="char" name="location" utype="voprov:EntityDescription.location" value="" />
```

```
      <PARAM arraysize="*" datatype="char" name="content_type" utype="voprov:EntityDescription.content_type" v
```

```
    </GROUP>
```

```
  </GROUP>
```

```
  <!-- Generated Entities / UWS results -->
```

```
  <GROUP name="Generated" utype="voprov:WasGeneratedBy">
```

```
    <GROUP name="count_map" utype="voprov:EntityDescription">
```

```
      <DESCRIPTION>Count map</DESCRIPTION>
```

```
      <PARAM arraysize="*" datatype="char" name="role" utype="voprov:UsedDescription.role" value="DL4 image"/>
```

```
      <PARAM arraysize="*" datatype="char" name="default" utype="voprov:Entity.id" value="count_map.fits"/>
```

```
      <PARAM arraysize="*" datatype="char" name="content_type" utype="voprov:EntityDescription.content_type" v
```

```
    </GROUP>
```

```
    <GROUP name="count_preview" utype="voprov:EntityDescription">
```

```
      <DESCRIPTION>Count map preview</DESCRIPTION>
```

VOTable

DataLink Service Descriptor

UWS Job Description Language

Provenance ActivityDescription

Client main features

Javascript based + Python Flask

- `UwsLib.js`: sends AJAX requests to the server
- `uws_manager.js`: handles and displays responses
 - Integration with Bootstrap3
 - HTML pages with specified `<div>` elements
(`id=joblist`, `parameters`, `results...`)

Job definition editor

- Interface to create the ActivityDescription file
 - parameters, used and generated entities
 - bash execution script

Job manager

- list jobs, create jobs, control jobs, view results

Definition of the `gammapy_maps` job

OPUS

Job Definition

Job List

Signed in as **testuser**

Job Definition

Name

Load JDL

Get JDL

Job name.

Description

Job description.

URL

Contact name

Contact email

Parameters

obs_ids	=	47802 47803 47804 47805	Req.? ☒	xs:string	↑	↓	×
Desc.	List of runs						
Options	List of possible choices (comma-separated values)						
Attr.	unit=... ucd=... utype=... min=... max=...						
RA	=	329.7169379	Req.? ☒	xs:double	↑	↓	×
Desc.	Target Right Ascension						
Options	List of possible choices (comma-separated values)						
Attr.	unit=deg						

List of parameter values, type, and units. Specify if checking parameter values on client and server (if to be used). A list of options (comma-separated values). Additional attributes (unit, ucd, utype, min, max).

OPUS
Observatoire de Paris
UWS Server

Parameters

obs_ids	=	47802 47803 47804 47805 47806 47807 47808 47809 47810 47811 47812 47813 47814 47815 47816 47817 47818 47819 47820 47821 47822 47823 47824 47825 47826 47827 47828 47829 47830 47831 47832 47833 47834 47835 47836 47837 47838 47839 47840 47841 47842 47843 47844 47845 47846 47847 47848 47849 47850 47851 47852 47853 47854 47855 47856 47857 47858 47859 47860 47861 47862 47863 47864 47865 47866 47867 47868 47869 47870 47871 47872 47873 47874 47875 47876 47877 47878 47879 47880 47881 47882 47883 47884 47885 47886 47887 47888 47889 47890 47891 47892 47893 47894 47895 47896 47897 47898 47899 47900 47901 47902 47903 47904 47905 47906 47907 47908 47909 47910 47911 47912 47913 47914 47915 47916 47917 47918 47919 47920 47921 47922 47923 47924 47925 47926 47927 47928 47929 47930 47931 47932 47933 47934 47935 47936 47937 47938 47939 47940 47941 47942 47943 47944 47945 47946 47947 47948 47949 47950 47951 47952 47953 47954 47955 47956 47957 47958 47959 47960 47961 47962 47963 47964 47965 47966 47967 47968 47969 47970 47971 47972 47973 47974 47975 47976 47977 47978 47979 47980 47981 47982 47983 47984 47985 47986 47987 47988 47989 47990 47991 47992 47993 47994 47995 47996 47997 47998 47999 48000 48001 48002 48003 48004 48005 48006 48007 48008 48009 48010 48011 48012 48013 48014 48015 48016 48017 48018 48019 48020 48021 48022 48023 48024 48025 48026 48027 48028 48029 48030 48031 48032 48033 48034 48035 48036 48037 48038 48039 48040 48041 48042 48043 48044 48045 48046 48047 48048 48049 48050 48051 48052 48053 48054 48055 48056 48057 48058 48059 48060 48061 48062 48063 48064 48065 48066 48067 48068 48069 48070 48071 48072 48073 48074 48075 48076 48077 48078 48079 48080 48081 48082 48083 48084 48085 48086 48087 48088 48089 48090 48091 48092 48093 48094 48095 48096 48097 48098 48099 48100 48101 48102 48103 48104 48105 48106 48107 48108 48109 48110 48111 48112 48113 48114 48115 48116 48117 48118 48119 48120 48121 48122 48123 48124 48125 48126 48127 48128 48129 48130 48131 48132 48133 48134 48135 48136 48137 48138 48139 48140 48141 48142 48143 48144 48145 48146 48147 48148 48149 48150 48151 48152 48153 48154 48155 48156 48157 48158 48159 48160 48161 48162 48163 48164 48165 48166 48167 48168 48169 48170 48171 48172 48173 48174 48175 48176 48177 48178 48179 48180 48181 48182 48183 48184 48185 48186 48187 48188 48189 48190 48191 48192 48193 48194 48195 48196 48197 48198 48199 48200 48201 48202 48203 48204 48205 48206 48207 48208 48209 48210 48211 48212 48213 48214 48215 48216 48217 48218 48219 48220 48221 48222 48223 48224 48225 48226 48227 48228 48229 48230 48231 48232 48233 48234 48235 48236 48237 48238 48239 48240 48241 48242 48243 48244 48245 48246 48247 48248 48249 48250 48251 48252 48253 48254 48255 48256 48257 48258 48259 48260 48261 48262 48263 48264 48265 48266 48267 48268 48269 48270 48271 48272 48273 48274 48275 48276 48277 48278 48279 48280 48281 48282 48283 48284 48285 48286 48287 48288 48289 48290 48291 48292 48293 48294 48295 48296 48297 48298 48299 48300 48301 48302 48303 48304 48305 48306 48307 48308 48309 48310 48311 48312 48313 48314 48315 48316 48317 48318 48319 48320 48321 48322 48323 48324 48325 48326 48327 48328 48329 48330 48331 48332 48333 48334 48335 48336 48337 48338 48339 48340 48341 48342 48343 48344 48345 48346 48347 48348 48349 48350 48351 48352 48353 48354 48355 48356 48357 48358 48359 48360 48361 48362 48363 48364 48365 48366 48367 48368 48369 48370 48371 48372 48373 48374 48375 48376 48377 48378 48379 48380 48381 48382 48383 48384 48385 48386 48387 48388 48389 48390 48391 48392 48393 48394 48395 48396 48397 48398 48399 48400 48401 48402 48403 48404 48405 48406 48407 48408 48409 48410 48411 48412 48413 48414 48415 48416 48417 48418 48419 48420 48421 48422 48423 48424 48425 48426 48427 48428 48429 48430 48431 48432 48433 48434 48435 48436 48437 48438 48439 48440 48441 48442 48443 48444 48445 48446 48447 48448 48449 48450 48451 48452 48453 48454 48455 48456 48457 48458 48459 48460 48461 48462 48463 48464 48465 48466 48467 48468 48469 48470 48471 48472 48473
---------	---	---

Submission of a `gammapy_maps` job

- OPUS reads the **ActivityDescription** file to generate a form
- This form also carries the Datasets **Obscore description**

OPUS

Job Definition

Job List

Signed in as **testuser**

Create new `gammapy_maps` job

← Back to job list

obs_ids

47802 47803 47804 47827 47828 47829 33787 33788 33789 33790

List of runs

RA

329,7169379

Target Right Ascension

Dec

-30,2255883

Target Declination

nxpix

400

Number of pixels on the X axis

nypix

400

Number of pixels on the Y axis

binsz

0,02

Pixel size

Add control parameters

Chose parameter

Submit

Reset

Show optional parameters

Results and Provenance

OPUS

Job Definition

Job List

Signed in as user

Job List for

gammapy_spectra

Refresh Job List

Create Test Job

Create New Job

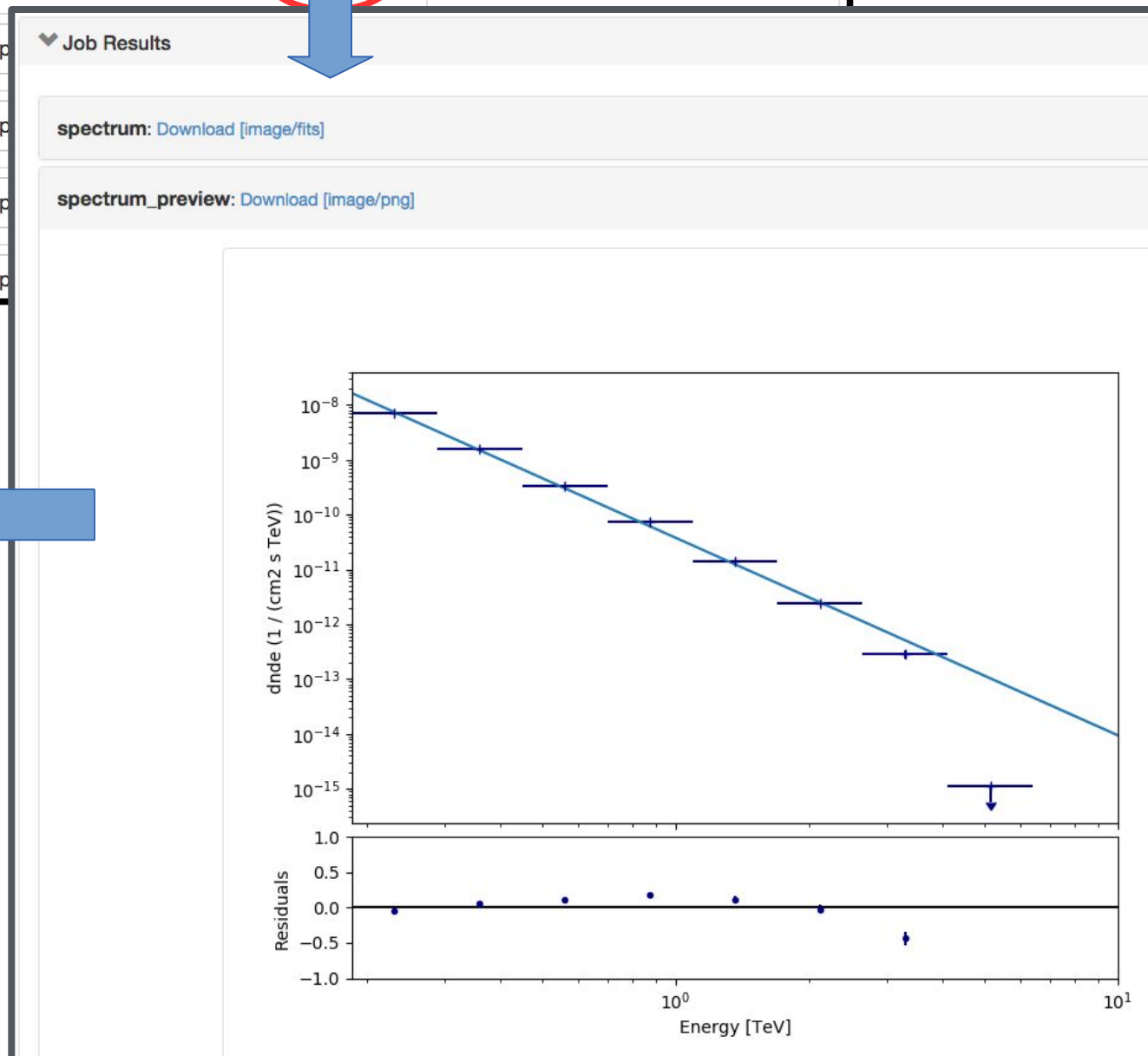
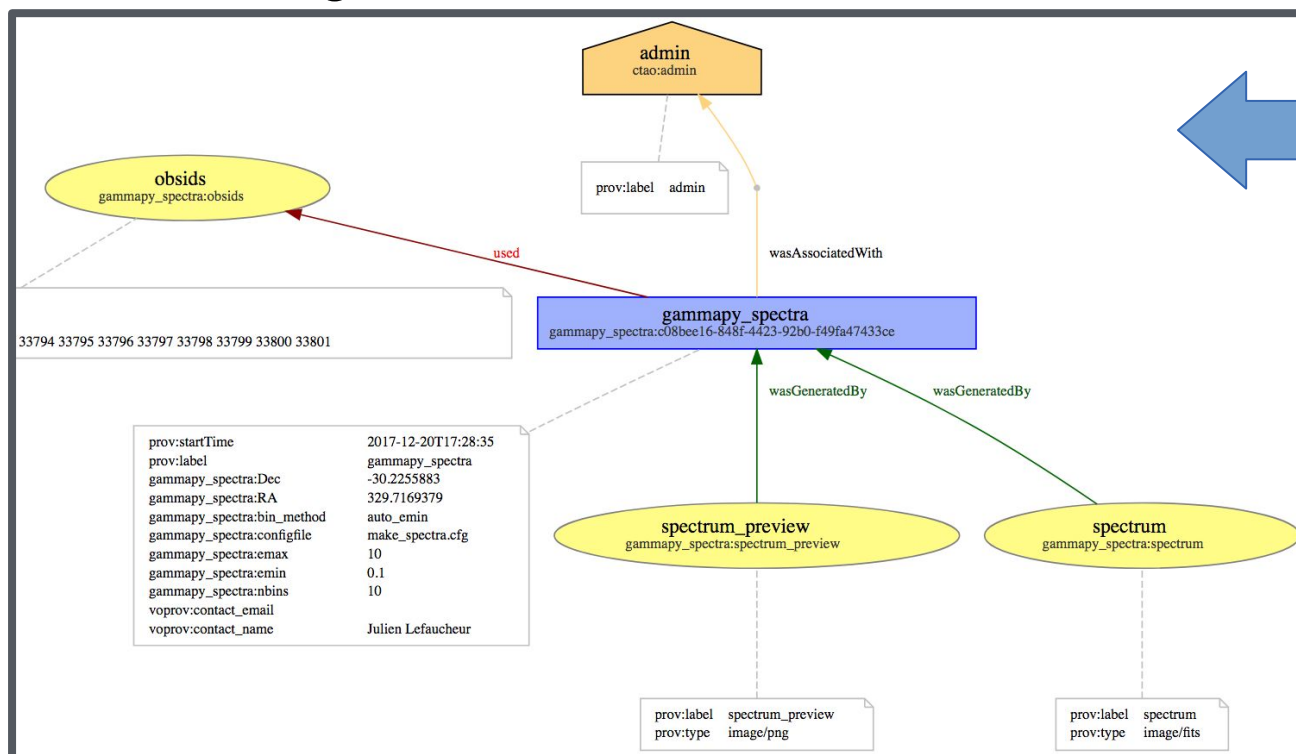
Type	Start Time	Destruction Time	Phase	Details	Control
gammapy_spectra	2017-10-02 10:47:07	2017-11-01 10:47:05	COMPLETED	Properties Parameters Results	Start Abort Delete
gammapy_spectra		2017-11-01 10:47:03	PENDING	Properties	
gammapy_spectra	2017-09-29 15:07:52	2017-10-29 15:07:51	COMPLETED	Properties	
gammapy_spectra	2017-09-29 14:55:10	2017-10-29 14:55:09	ABORTED	Properties	
gammapy_spectra	2017-09-29 14:21:20	2017-10-29 14:21:19	COMPLETED	Properties	

Job Results

spectrum: Download [image/fits]

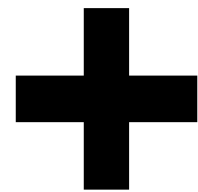
spectrum_preview: Download [image/png]

Tracking of Provenance informations



Provides Provenance files

UWS
database

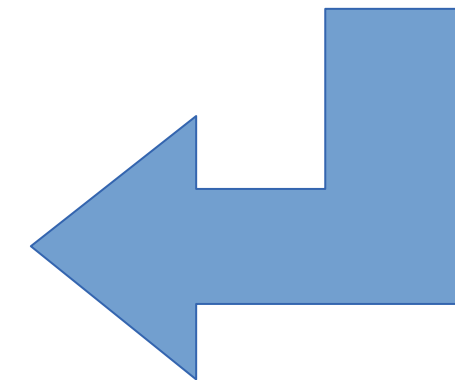
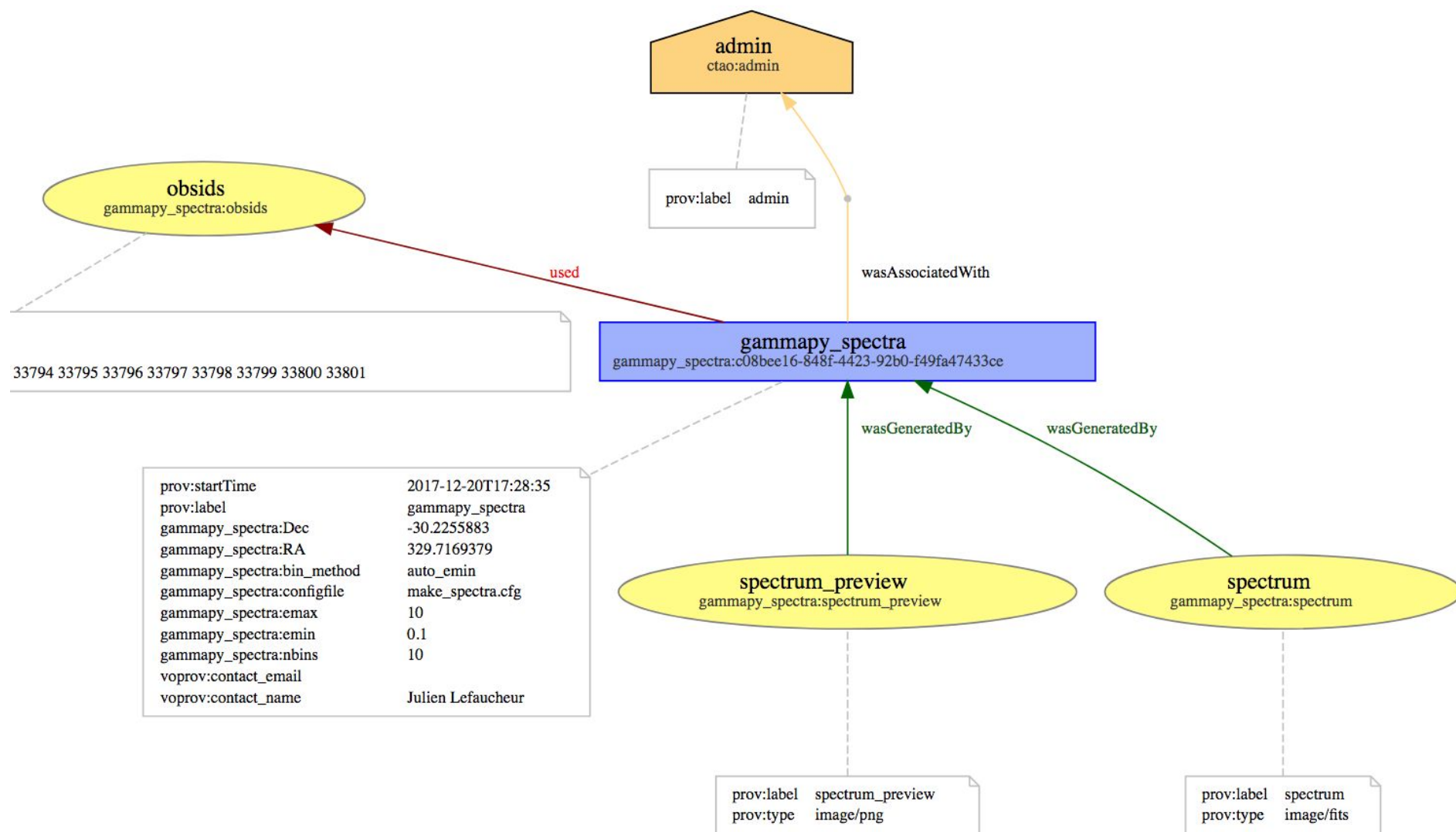


Activity
Description
(VOTable)



Provenance
Files

(VOTable,
XML,
JSON, SVG)



Serializations - W3C PROV formats

```
<prov:document xmlns:ctadata="ivo://vopdc.obspm/cta#" xmlns:ctajob
  <prov:activity prov:id="ctajobs:ctbin">
    <prov:startTime> 2016-03-13T23:44:46 </prov:startTime>
    <prov:endTime> 2016-03-13T23:44:56 </prov:endTime>
  </prov:activity>
  <prov:agent prov:id="cta:consortium">
    <prov:type xsi:type="xsd:string"> Organization </prov:type>
  </prov:agent>
  <prov:wasAssociatedWith>
    <prov:activity prov:ref="ctajobs:ctbin" />
    <prov:agent prov:ref="cta:consortium" />
  </prov:wasAssociatedWith>
  <prov:entity prov:id="uwsdata:parameters/inobs" />
  <prov:used>
    <prov:activity prov:ref="ctajobs:ctbin" />
    <prov:entity prov:ref="uwsdata:parameters/inobs" />
  </prov:used>
  <prov:entity prov:id="uwsdata:results/outcube" />
  <prov:wasGeneratedBy>
    <prov:entity prov:ref="uwsdata:results/outcube" />
    <prov:activity prov:ref="ctajobs:ctbin" />
  </prov:wasGeneratedBy>
  <prov:wasDerivedFrom>
    <prov:generatedEntity prov:ref="uwsdata:results/outcube" />
    <prov:usedEntity prov:ref="uwsdata:parameters/inobs" />
  </prov:wasDerivedFrom>
  <prov:entity prov:id="uwsdata:results/logfile" />
  <prov:wasGeneratedBy>
    <prov:entity prov:ref="uwsdata:results/logfile" />
    <prov:activity prov:ref="ctajobs:ctbin" />
  </prov:wasGeneratedBy>
  <prov:wasDerivedFrom>
    <prov:generatedEntity prov:ref="uwsdata:results/logfile" />
    <prov:usedEntity prov:ref="uwsdata:parameters/inobs" />
  </prov:wasDerivedFrom>
</prov:document>
```

```
{
  - wasAssociatedWith: {
    - _id1: {
      prov:agent: "cta:consortium",
      prov:activity: "cta:anactools_v1.1"
    }
  },
  - agent: {
    - cta:consortium: {
      prov:type: "Organization"
    }
  },
  - entity: {
    uwsdata:results/fit_results: { },
    uwsdata:results/configfile: { },
    uwsdata:results/butterfly: { },
    uwsdata:results/spectrum_plot: { },
    uwsdata:results/spectrum: { }
  },
  - prefix: {
    uwsdata: "https://voparis-uws-test.obspm.fr/rest",
    cta: "http://www.cta-observatory.org#",
    voprov: "http://www.ivoa.net/ns/voprov#"
  },
  - activity: {
    - cta:anactools_v1.1: {
      prov:startTime: "2016-04-07T00:26:00",
      prov:endTime: "2016-04-07T00:27:15"
    }
  },
  - wasGeneratedBy: {
    - _id5: {
      prov:entity: "uwsdata:results/butterfly",
      prov:activity: "cta:anactools_v1.1"
    },
    - _id4: {
      prov:entity: "uwsdata:results/fit_results",
      prov:activity: "cta:anactools_v1.1"
    }
  },
}
```