VO Feed Introduction

Marian Babik
VO Feed

- FWIR started as initiative within SAM during design and development of Aggregated Topology Provider (ATP)
  - Back then, SAM reported to both WLCG and EGI while also receiving results from OSG/RSV, which required processing 6 different topological views
  - VO feed was proposed as a common schema binding EGI/OSG (Sites) view with WLCG (Experiments) views
    - Grouping GOCDB/OIM sites and services into Experiments sites, tiers and clouds
- VO feed became important concept for many applications requiring processing or displaying experiments topologies
  - The core applications are still in monitoring
Basic Concepts

<title>CMS Topology Information for ATP</title>
<description>
List of CMS site names for monitoring and mapping to the SAM/WLCG site names provided by the Dashboard
</description>
<feed_responsible dn="/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=psaiz/CN=542764/CN=Pablo Saiz" name="Pablo Saiz"/
<last_update>2016-11-10T12:56:49Z</last_update>
<vo>cms</vo>
<atp_site name="CSCS-LCG2">
  <service hostname="arc01.lcg.cscs.ch" flavour="ARC-CE"/>
  <service hostname="arc02.lcg.cscs.ch" flavour="ARC-CE"/>
  <service hostname="arc03.lcg.cscs.ch" flavour="ARC-CE"/>
  <service hostname="storage01.lcg.cscs.ch" flavour="SRMv2"/>
  <group name="Tier-2" type="CMS_Tier"/>
  <group name="T2_CH_CSCS" type="CMS_Site"/>
  <group name="T2_CH_CSCS" type="AllGroups"/>
  <group name="T2_CH_CSCS" type="Tier2s"/>
  <group name="T2_CH_CSCS" type="Tier3s + Tier2s"/>
  <group name="T2_CH_CSCS" type="Tier2s + Tier1s + Tier0"/>
  <group name="T2_CH_CSCS" type="Tier3s + Tier2s + Tier1s"/>
</atp_site>
Optionals

```xml
<atp_site infrast="LCG" name="AM-04-YERPHI">
  <service endpoint="srm://se.yerphi-cluster.grid.am:8446/srm/managerv2?SFN=" flavour="SRMv2" hostname="se.yerphi-cluster.grid.am">
    <spacetoken base_path="/dpm/yerphi-cluster.grid.am/home/atlas/atlaslocalgroupdisk/" name="AM-04-YERPHI_LOCALGROUPDISK/>
    <spacetoken base_path="/dpm/yerphi-cluster.grid.am/home/atlas/atlassproddisk/" name="AM-04-YERPHI_PRODDISK/>
    <spacetoken base_path="/dpm/yerphi-cluster.grid.am/home/atlas/atlasscratchdisk/" name="AM-04-YERPHI_SCRATCHDISK/>
  </service>
  <service endpoint="ce.yerphi-cluster.grid.am:8443" flavour="CREAM-CE" hostname="ce.yerphi-cluster.grid.am"/>
  <service endpoint="http://atlas-ce.yerphi-cluster.grid.am:3128" flavour="Squid" hostname="atlas-ce.yerphi-cluster.grid.am"/>
  <service endpoint="psaanl.yerphi-cluster.grid.am" flavour="PerfSonar" hostname="psaanl.yerphi-cluster.grid.am"/>
  <group name="AM-04-YERPHI" type="ATLAS_Cloud_NL"/>
  <group name="AM-04-YERPHI" type="ATLAS_Federation_NL-Tier3s"/>
  <group name="AM-04-YERPHI" type="ATLAS_Site"/>
  <group name="Tier-3" type="ATLAS_Tier"/>
  <group name="AM-04-YERPHI" type="All sites"/>
  <group name="AM-04-YERPHI" type="Tier3s"/>
</atp_site>
```
## Implementation

<table>
<thead>
<tr>
<th>VO</th>
<th>Maintained by</th>
<th>Deps</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS</td>
<td>AGIS</td>
<td>AGIS (GOCDB/OIM, etc.)</td>
<td><a href="http://atlas-agis-api.cern.ch/request/atp/xml/">http://atlas-agis-api.cern.ch/request/atp/xml/</a></td>
</tr>
<tr>
<td>CMS</td>
<td>SSB (IT/MM)</td>
<td>PhEDEX, pilot factories</td>
<td><a href="http://wlcg-sam-cms.cern.ch/dashboard/request.py/cmssitemapbdii">http://wlcg-sam-cms.cern.ch/dashboard/request.py/cmssitemapbdii</a></td>
</tr>
<tr>
<td>ALICE</td>
<td>SSB (IT/MM)</td>
<td>ALICE LDAP</td>
<td><a href="http://wlcg-sam-alice.cern.ch/dashboard/request.py/alicesitemap">http://wlcg-sam-alice.cern.ch/dashboard/request.py/alicesitemap</a></td>
</tr>
</tbody>
</table>
Remarks

- **VO feed more than just common API**
  - In the WLCG/EGI SAM times, single tool facilitated common understanding on different attributes and their meaning across EGI/OSG/Experiments
    - e.g. What does GOCDB flag is_monitored mean today? What about OIM InterOp?
    - Recently scheduled/unscheduled downtime was discussed

- **Coordination on terminology and processes important**
  - Between experiments
    - Service types now more or less common across the board, but this might not last
    - HT-Condor-CE service type had 5 different names before this was raised at WLCG ops coordination, finally we agreed on HTCONDOR-CE
  - Between sites and experiments
    - What does topology mean in the context of VO feed?
      - Production vs non-production services
Evolution

● In Nov/2015 an extension of VO feed was proposed and was agreed (in Feb/2016), but then was put on hold due to unclear situation with CRIC:
  ○ new atp_site attributes lat/long and contacts
  ○ ce_resource element under <service> providing supported batch system and queue
  ○ se_resource element under <service> listing paths and space tokens
  ○ new flavours, e.g. XROOTD, VAC, GSIFTP, etc.
  ○ dropping OSG notation (OSG-CE/OSG-SRMv2) replacing them with regular service types
  ○ adding storage area/compute area distinction to understand which endpoints are part of the same storage/batch system

● Different experiments are now in various stages of developments of this extension since things need to evolve
  ○ ATLAS has almost production ready feed at http://atlas-agis-api.cern.ch/request/atp2/xml/
  ○ CMS and LHCb have also started developments in this direction
Summary

● VO feed - common schema/API binding site and experiments views
  ○ Currently used by many monitoring applications - use cases next
  ○ Provides very important functionality that will be needed

● Transition of the VO feed to CRIC needs to be clarified
  ○ Ownership and maintenance currently the “hottest” topics
    ■ Not all VO feeds maintained by VOs
  ○ Evolution is highly needed (now)
  ○ We have many dependencies, transition needs to be adiabatic

● Coordination likely needed as part of the process
  ○ CRIC steering meeting ?
  ○ WLCG operations coordination ? Other ?

● Finally, current VO feed shows its age and would benefit from refresh
  ○ Likely coming as part of CRIC
  ○ Support for JSON as an example