Fig. 1

Fig. 2

1. The OAI-PMH Proxy purx

(cf. Fig. 1)

Markus Demleitner
msdemlei@ari.uni-heidelberg.de

(cf. Fig. 2)

purx = PUblishing Registry Proxy

- The problem
- Our solution
- A sample registration

(cf. Fig. 3)

2. The Problem

To get your service registered, you so far have two options:

- Run an OAI-PMH endpoint (and register it with the RofR – fine for a big data center with lots of resources and frequent changes)
- Use a web interface at ESAC or STScI (and go back there every time your record needs to change – fine if you have one service that rarely changes).

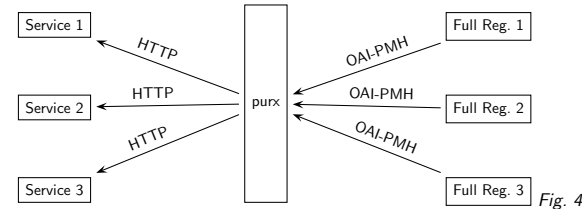. . . but what if you're in between?

Fig. 4

3. What If

If you just have a few TAP services, what if you could

- slap together a quick script
- combining your VOSI data
- with a bit of template XML,
- dump that into an apache-published directory,
- run that script each time your metadata changes,
- and have that reflected in the Registry?

Of course, your server could have built-in support for that, too.

4. Enter purx

(cf. Fig. 4)

Essentially, purx regularly does http requests (with if-modified-since headers) to registered URLs to establish if any changes have been made to the resource records prepared by the services. If there are any, they will be included in incremental harvests from the full registries.

5. Why bother?

- Not all full registries need to worry about the complexities of HTTP harvesting, vanishing documents, etc.
- Operators just fill out a bit of XML and don't need to worry about authorities and most anything else registry-related.
- Thanks to if-modified-since the overhead can be very low.
- Thanks to Resource/@updated the full registries are still isolated from purx client mishaps.
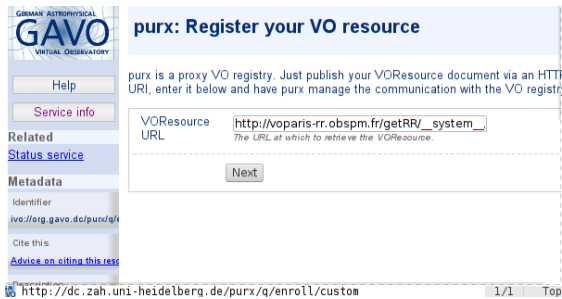
Fig. 5

6. In Practice I

DaCHS, for instance, spits out appropriate files itself:

```
$ curl http://voparis-rr.obspm.fr/getRR/__system__/tap/run | \
  xmlstarlet fo | head -25 | tail -15
<ri:Resource
  created="2009-12-01T10:00:00"
  status="active" updated="2017-03-08T14:59:34"
  xsi:type="vs:CatalogService">
  <title>PADC TAP Server on voparis-rr.obspm.fr TAP service</title>
  <shortName>PADCC TAP</shortName>
  <identifier>ivo://vopdc.obspm/tap</identifier>
  <curation>
    <publisher>Paris Astronomical Data Centre</publisher>
    <creator>
      <name>PADC</name>
      <logo>http://voparis-srv-paris.obspm.fr/logos/PADC_small.png</logo>
    </creator>
    <date role="updated">2017-03-08T14:59:34</date>
    <contact>
      <name>PADC support team</name>
      <address>77 av. Denfert Rochereau, 75014 Paris, FRANCE</address>
```

7. In Practice II

So, submit the URL where the VOResource XML is located:

(cf. Fig. 5)

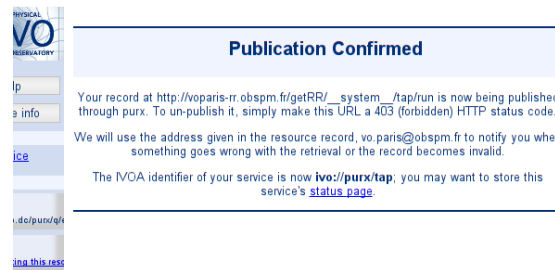(this will also validate your resource record)

Fig. 6

Fig. 7

8. Receive a Mail

(cf. Fig. 6)

This went to contact/email.

9. Confirm Publication

When you retrieve the URL in the mail, you get:

(cf. Fig. 7)

That's it.

purx will warn contact/email after a while when the document becomes unavailble or invalid.

You can immediately get rid of the publication by returning a 403 from the XML URL.

And you can inspect the status of your publication at any time.

10. Do you want to know more?

purx info (including sample records) at

http://localhost:8080/purx/q/enroll/info

Thanks.