YARN Registry

Steve Loughran– Hortonworks
stevel at hortonworks.com

2015

What's in HDP 2.2

- ZK-based service for clients; implements client API
- Extended for servers; implements async cleanup, user path creation
- RM integration: creation of /users/$user with correct permissions on job submit; cleanup on termination events
- Mini cluster tests including KDC for security
- Slider publishing endpoints for Agents and apps wanting to enum & retrieve site-XML docs.
- Slider python read-only client
- Slider functional tests validate code across platforms & with security.

Slider 0.60 (in HDP 2.2)

- Slider publishing endpoint for Agents
- python read-only client in Agent
- Registered endpoint for enum & fetch XML/JSON/properties configs
- Registered endpoint for list and fetch other generated files
- Ambari to build quicklink lists via config endpoint
- PoC REST view for enum/fetch of service records
- Functional tests validate code across platforms & with security.

Pending YARN commits to branch-2

JIRA	Title	Status
YARN-2683	*registry config options: document and move to core-default*	*patch pending*
YARN-2571	*RM to support YARN registry*	*patch pending*
YARN-2657	MiniYARNCluster to add MicroZookeeper service	patch pending on YARN-2571
YARN-2646	distributed shell & tests to use registry	patch pending on YARN-2646
YARN-2616	Add CLI client to the registry to list/view entries	patch for review (from Sanjay's son)

Why does nobody ever review my patches?

== why should I bother writing hadoop code?

Client API designed for REST

```
public interface RegistryOperations extends Service {
  boolean mknode(String path, boolean createParents)
  void bind(String path, ServiceRecord record, int flags)
  ServiceRecord resolve(String path)
  RegistryPathStatus stat(String path)
  boolean exists(String path)
  List<String> list(String path)
  void delete(String path, boolean recursive)
  boolean addWriteAccessor(String id, String pass)
  public void clearWriteAccessors()
}
```

© Hortonworks Inc

TBD

- REST endpoint (YARN-2948)
- REST implementation of client API (R/O; R/W)
- Long-haul bridge in Knox
- Registration of core Hadoop services
- Self-registration of HBase, Oozie, …
- Standard/HDP client scripts to optionally bind via registry
- Define some standard REST APIs for config publishing, codahale servlets?

Ubiquity

A modern REST client library

1. SPNEGO
2. Setup via UGI library
3. Higher level than `java.net.HttpURLConnection` (e.g. Jersey, Restlet)
4. Proxy & Long-haul
5. Good error propagation
6. In/alongside Hadoop-common
7. Foundation for Slider, Knox, Registry clients in Java

A basic UGI-integrated client is in

```
org.apache.hadoop.hdfs.web.URLConnectionFactory
```

Hortonworks

Actions

- Get existing patches in
- REST server based off Slider PoC
- Implement client API: read-only
- Client API: SPNEGO-authed write
  (will this let AMs register entries?)
- Knox: bridge to REST server and endpoints
- Ambari: register static services under /services
- Enhance client scripts

Need commitments of coding *and review* effort

Questions?

hortonworks.com