

Autoreduction service: installation and maintenance

Mathieu Doucet
Oak Ridge National Laboratory

ORNL is managed by UT-Battelle, LLC for the US Department of Energy



U.S. DEPARTMENT OF
ENERGY

Plan for the next few weeks

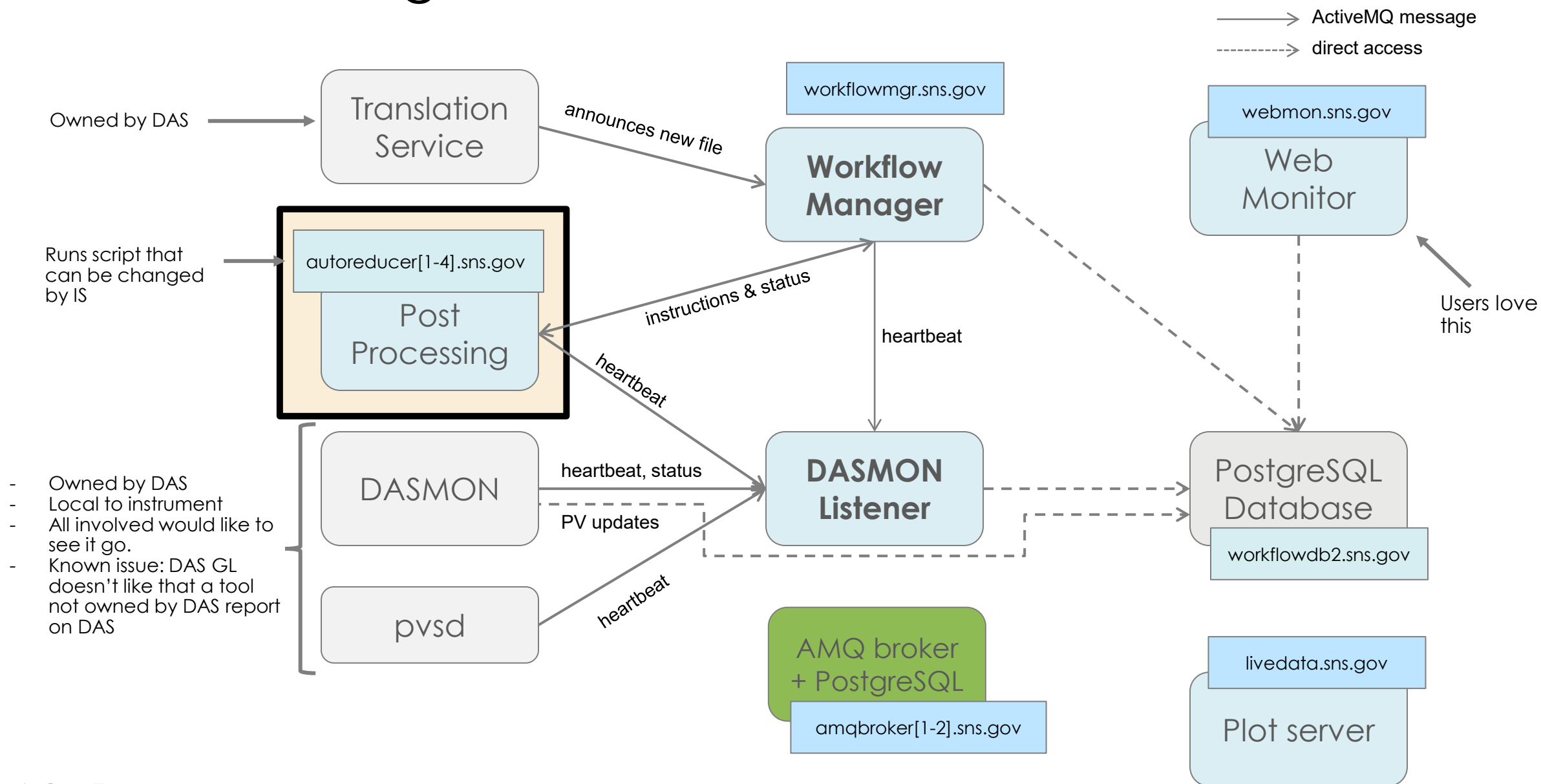
Test environment:

New RHEL8 machines are being set up so we can install them together

Topics to cover:

1. General overview
2. Workflow manager and DASMON listener – Installation & maintenance
3. Web monitor – Installation and maintenance
4. **Autoreduction service – Installation and maintenance [this presentation]**
5. Autoreduction setup through webmon – how-to and future vision
6. The IHC call – when things go wrong & recovery strategies
7. Vision for the future – what I would do differently

Post-Processing Architecture



Overview

- There help on the repo (probably outdated)
- This is an old refactor (2014) of old software (2012-ish).
- As we migrated away from ICAT for cataloging, some code has been cleaned up. Some cleaning up didn't get done yet.

https://github.com/neutrons/post_processing_agent

📁	SPECS	add ONCat reduced cataloging	15 months ago
📁	configuration	Add json data type	5 years ago
📁	postprocessing	add ONCat reduced cataloging	15 months ago
📁	scripts	Fix script name	4 years ago
📄	.gitignore	Adding .gitignore	5 years ago
📄	Makefile	add links	14 months ago
📄	README.md	Update README.md	5 years ago

This is where the newer
job classes are

Some code is no longer
needed

This is the service script

Branch: master post_processing_agent / postprocessing /

Go to file Add file

mdoucet committed 872dd2e on Mar 29, 2019 History

..		
📁	processors	add ONCat reduced cataloging 15 months ago
📄	Configuration.py	Add publisher to config 4 years ago
📄	Consumer.py	Do not log every ping 5 years ago
📄	PostProcessAdmin.py	add ONCat reduced cataloging 15 months ago
📄	__init__.py	release version 6 years ago
📄	ingest_nexus.py	Hand written function to drop Pyxml 4 years ago
📄	publish_plot.py	update 2 years ago
📄	queueProcessor.py	Configure for AR nodes 6 years ago
📄	reduction_script_writer.py	Log the template changes to file 6 years ago
📄	time_conversions.py	Hand written function to drop PyXml 4 years ago

Processors

Job_tree is no longer in use.
It allowed for job hierarchy.

Newer jobs are done this way

Branch: master post_processing_agent / postprocessing / processors /

Go to file Add file

mdoucet committed 872dd2e on Mar 29, 2019 History

..		
__init__.py	Add processor plugin mechanism	6 years ago
base_processor.py	better logging	4 years ago
job_handling.py	Improve error processing	5 years ago
job_tree.py	remove err log if the job succeeded.	4 years ago
oncat_processor.py	Catch file handling errors	3 years ago
oncat_reduced_processor.py	add ONCat reduced cataloging	15 months ago
test_processor.py	Added processor to run local scripts.	6 years ago

```
class ONCatProcessor(BaseProcessor):
    """
    Define post-processing task
    """
    ## Input queue
    _message_queue = "/queue/CATALOG.ONCAT.DATA_READY"
    STARTED_QUEUE = '/queue/CATALOG.ONCAT.STARTED'
    COMPLETE_QUEUE = '/queue/CATALOG.ONCAT.COMPLETE'
    ERROR_QUEUE = '/queue/CATALOG.ONCAT.ERROR'
    SCRIPT_PATH = "/opt/postprocessing/scripts/oncat_ingest.py"

    def __init__(self, data, conf, send_function):
        """
        Initialize the processor
        @param data: data dictionary from the incoming message
        @param conf: configuration object
        @param send_function: function to call to send AMQ messages
        """
        super(ONCatProcessor, self).__init__(data, conf, send_function)

    def __call__(self):
        """
        Execute the job
        """
```

Configuration & Installation

/etc/autoreduce/post_processing.conf

- make rpm
- Copy rpm to packages.sns.gov
- Check configuration and update it as needed

```
{
  "failover_uri":
  "failover: (tcp://amqbroker.sns.gov:61613)?randomize=false,startupMaxReconnectAttempts=100,
  initialReconnectDelay=1000,maxReconnectDelay=5000,maxReconnectAttempts=-1",
  "amq_queues": ["/queue/REDUCTION.DATA_READY",
                  "/queue/REDUCTION.CREATE_SCRIPT", "/queue/CATALOG.DATA_READY"],
  "amq_user": "wfclient",
  "amq_pwd": "XXXXX",
  "sw_dir": "/opt/postprocessing",
  "python_dir": "/opt/postprocessing/postprocessing",
  "start_script": "python",
  "python_exec": "/opt/postprocessing/scripts/mantidpython.py",
  "task_script": "PostProcessAdmin.py",
  "task_script_queue_arg": "-q",
  "task_script_data_arg": "-d",
  "log_file": "/var/log/SNS_applications/postprocessing.log",
  "postprocess_error": "POSTPROCESS.ERROR",
  "catalog_started": "CATALOG.STARTED",
  "catalog_complete": "CATALOG.COMPLETE",
  "catalog_error": "CATALOG.ERROR",
  "reduction_started": "REDUCTION.STARTED",
  "reduction_complete": "REDUCTION.COMPLETE",
  "reduction_error": "REDUCTION.ERROR",
  "reduction_disabled": "REDUCTION.DISABLED",
  "reduction_catalog_started": "REDUCTION_CATALOG.STARTED",
  "reduction_catalog_complete": "REDUCTION_CATALOG.COMPLETE",
  "reduction_catalog_error": "REDUCTION_CATALOG.ERROR",
  "heart_beat": "/topic/SNS.COMMON.STATUS.AUTOREDUCE.0",
  "dev_output_dir": "",
  "webmon_url_template": "https://monitor.sns.gov/files/$instrument/$run_number/submit_reduced/",
  "catalog_data_ready": "CATALOG.DATA_READY",
  "reduction_data_ready": ["REDUCTION.DATA_READY", "REDUCTION.AR4.DATA_READY"],
  "reduction_catalog_data_ready": "REDUCTION_CATALOG.DATA_READY",
  "communication_only": 0,
  "max_procs": 10,
  "jobs_per_instrument": 4,
  "remote_execution": 0,
  "processors": ["job_tree.JobTreeProcessor", "oncat_reduced_processor.ONCatProcessor",
                 "oncat_processor.ONCatProcessor"],
  "publisher_username": "plot_publisher",
  "publisher_password": "xxxxxxx",
  "publish_url_template": "https://livedata.sns.gov/plots/$instrument/$run_number/upload_plot_data/",
  "exceptions": ["Unverified HTTPS", "InsecureRequestWarning", "plotly", "operands could not be
broadcast", "Publish plot failed", "cann
ot import name factorial", "tcmalloc: large alloc", "dPS_array", "wks_utility", "The proton charge is
zero", "Error in logging framework"
, "Found no integrated charge value in gd_prtn_chrg"]
}
```