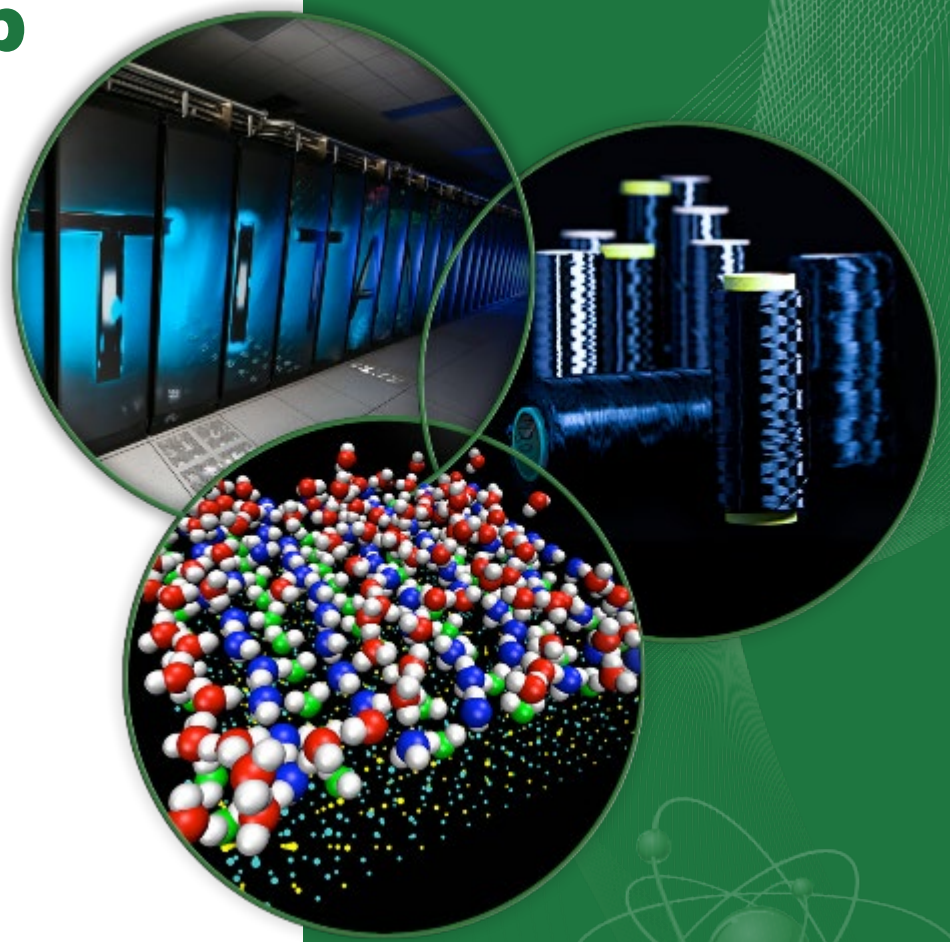
Workflow Management and Instrument Web Monitoring



Mathieu Doucet

Neutron Data Analysis and Visualization Division

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

NDAV Training / Nov 2016

OAK RIDGE National Laboratory

Post-Processing Overview

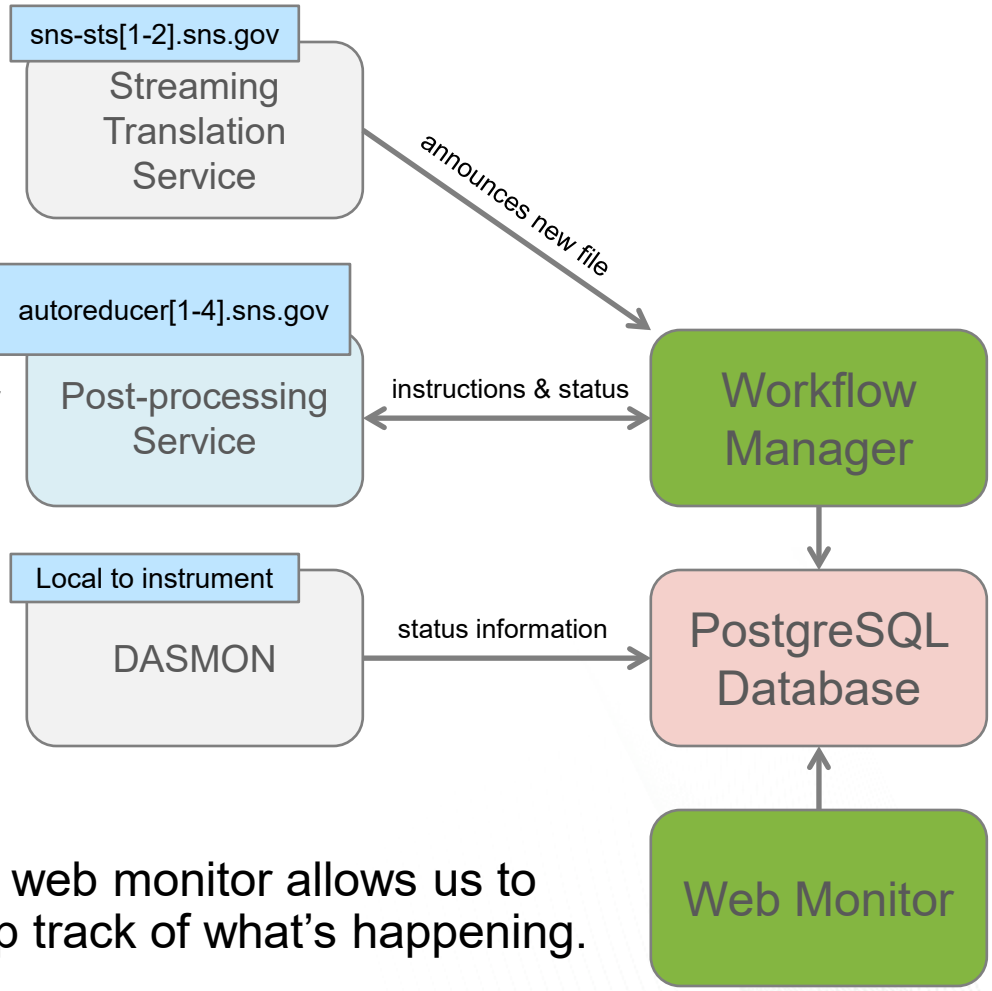
Translation service notifies the workflow manager when a new data file is ready.

The auto-reduction service takes care of reduction and cataloging.

➔ Reduction script can be modified by instrument staff.

/SNS/REF_L/shared/autoreduce/reduce_REF_L.py

DASMON reads in the stream and performs diagnostics.

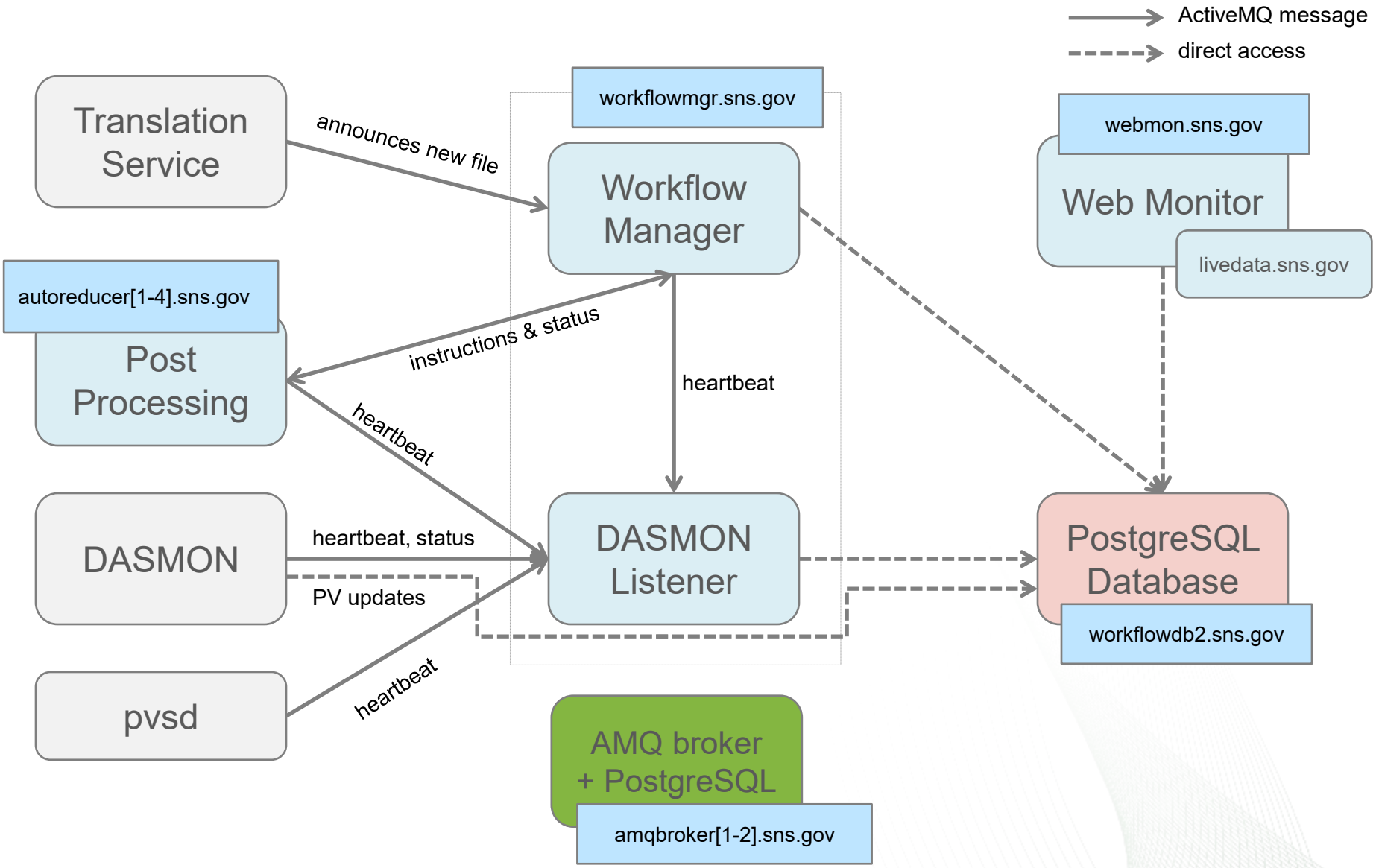


The web monitor allows us to keep track of what's happening.

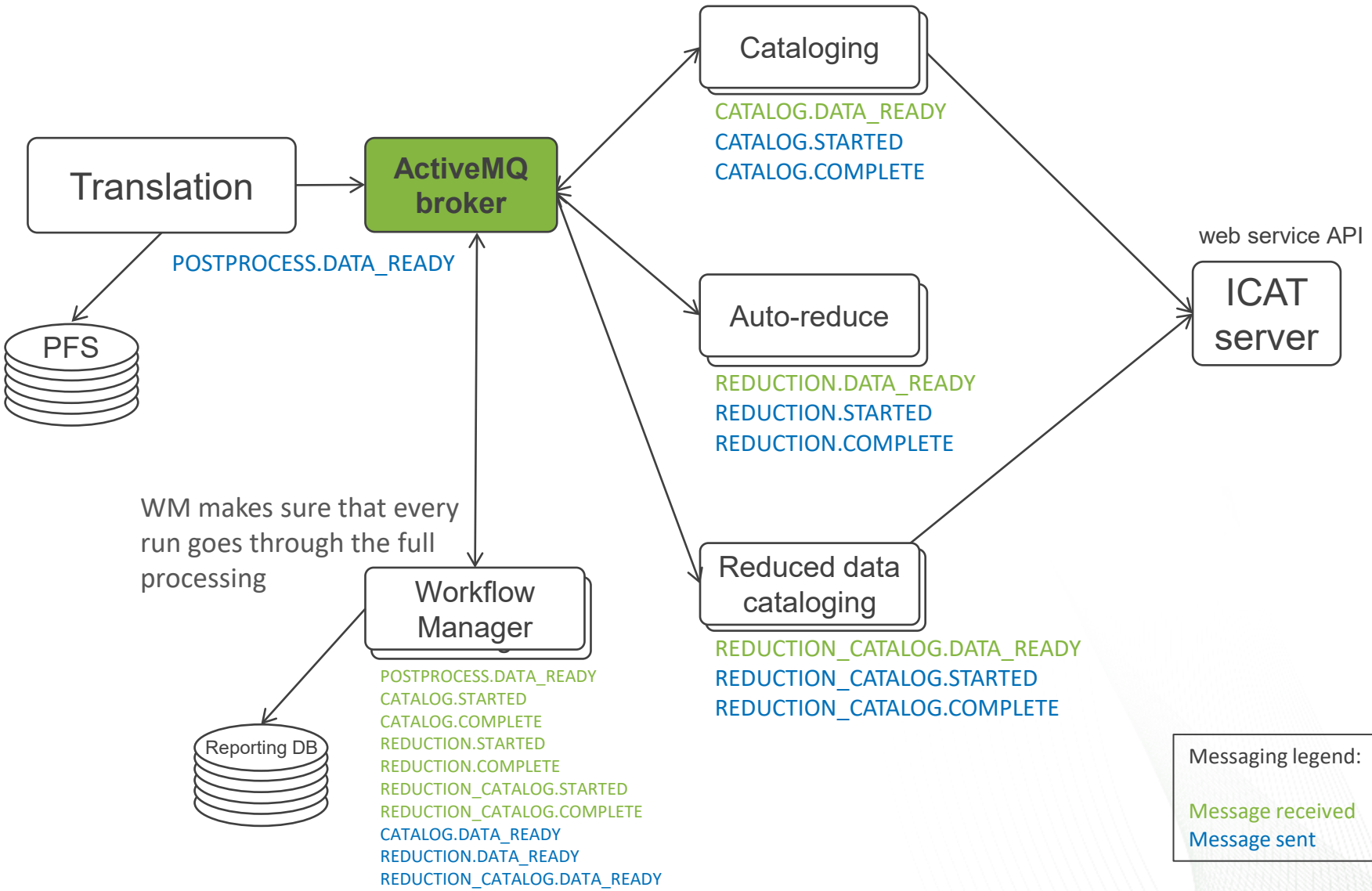
Workflow Management of Post Processing Services

- Manages all aspects of the post processing workflow, which can be individualized for each beam line.
- Updates the database as services complete tasks.
- Translation service can send the message and forget. The workflow manager will be responsible for ensuring that the post processing completes.
- Transacts with other workflow components through **ActiveMQ** using dedicated queues for each task.

Post-Processing Architecture



ActiveMQ Communication Flow



ActiveMQ Messaging

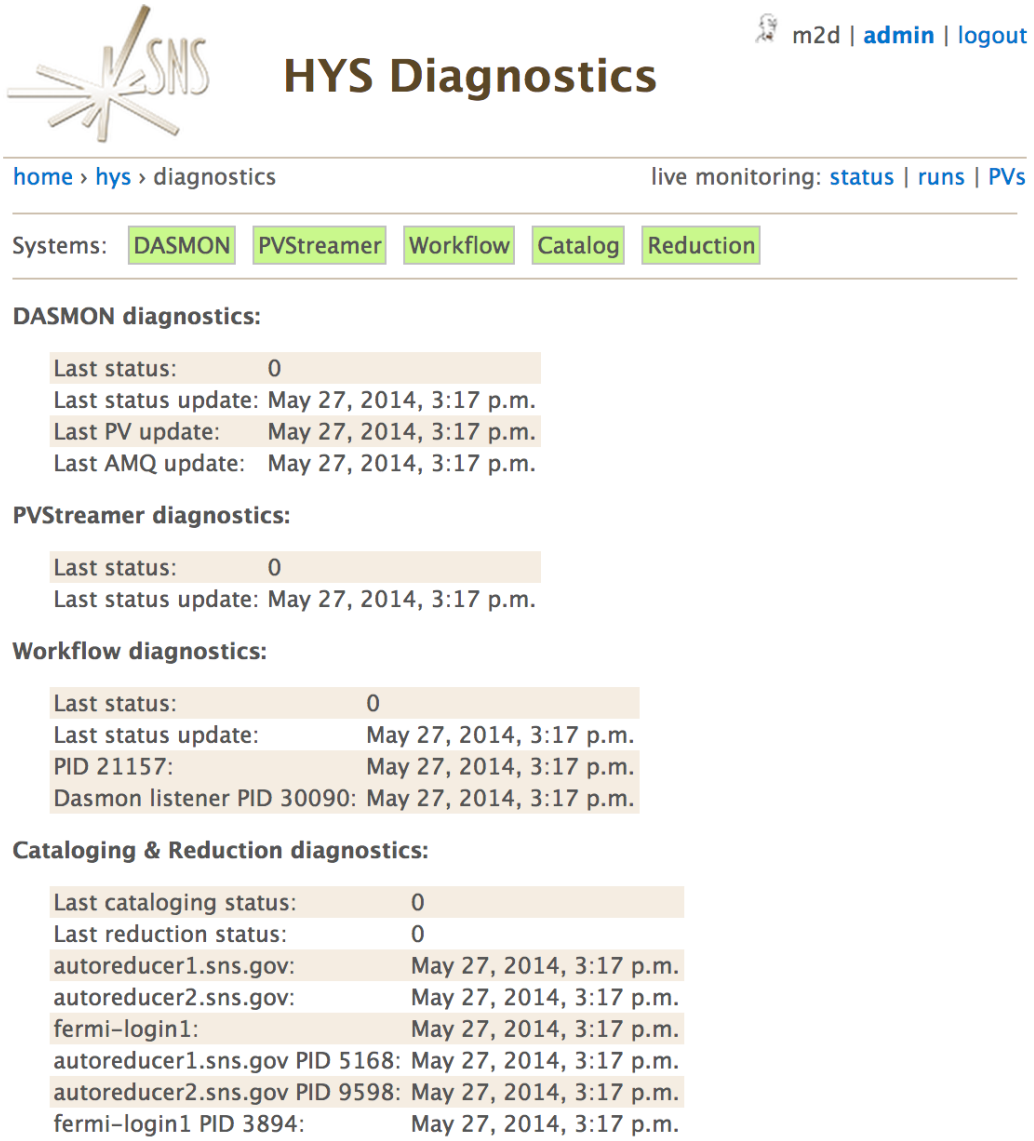
The workflow manager and the services communicate through ActiveMQ.

http://activemq.apache.org/

- Messages delivered to brokers are durable and will be preserved should a broker fail.
- Clustered brokers provide failover functionality to support high-availability.
- Load balancing across all components is managed automatically.
  - Separate queues are used for each task, and balancing occurs as worker nodes pick up the messages.

Online Diagnostics

- A diagnostics page allows us to verify the health of the system.
- Pinpoints the issue for common problems.
- The system tries to self-diagnose as much as possible.



m2d | admin | logout

HYS Diagnostics

home › hys › diagnostics — live monitoring: status | runs | PVs

Systems: DASMON | PVStreamer | Workflow | Catalog | Reduction

DASMON diagnostics:

Last status:	0
Last status update:	May 27, 2014, 3:17 p.m.
Last PV update:	May 27, 2014, 3:17 p.m.
Last AMQ update:	May 27, 2014, 3:17 p.m.

PVStreamer diagnostics:

Last status:	0
Last status update:	May 27, 2014, 3:17 p.m.

Workflow diagnostics:

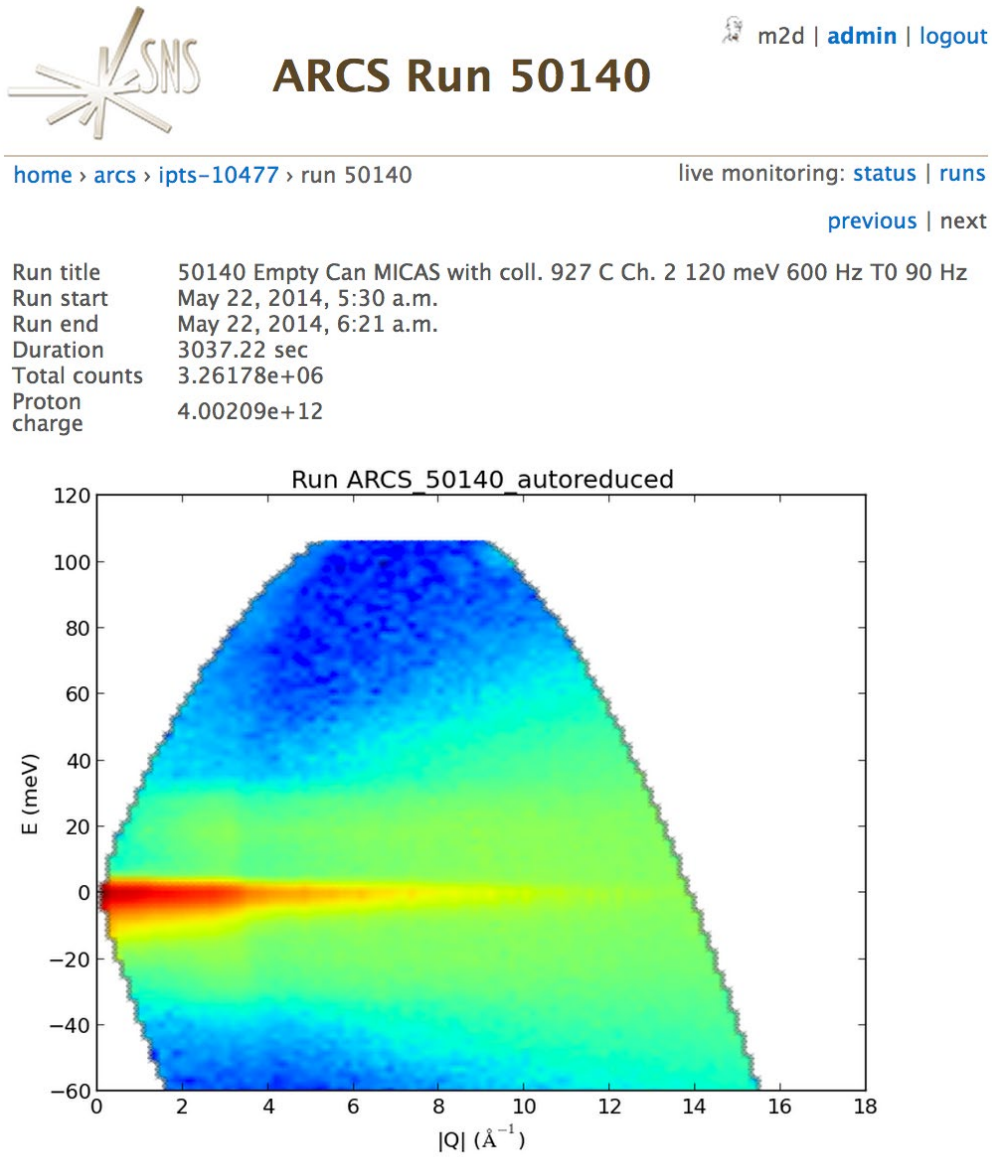
Last status:	0
Last status update:	May 27, 2014, 3:17 p.m.
PID 21157:	May 27, 2014, 3:17 p.m.
Dasmon listener PID 30090:	May 27, 2014, 3:17 p.m.

Cataloging & Reduction diagnostics:

Last cataloging status:	0
Last reduction status:	0
autoreducer1.sns.gov:	May 27, 2014, 3:17 p.m.
autoreducer2.sns.gov:	May 27, 2014, 3:17 p.m.
fermi-login1:	May 27, 2014, 3:17 p.m.
autoreducer1.sns.gov PID 5168:	May 27, 2014, 3:17 p.m.
autoreducer2.sns.gov PID 9598:	May 27, 2014, 3:17 p.m.
fermi-login1 PID 3894:	May 27, 2014, 3:17 p.m.

Run Details

- Information about each run is available.
- Allows for image upload by the auto-reduction.



Run Details

- Shows you a list of all cataloged files.
- Both raw and reduced files are cataloged.
- Gives you a timeline of the workflow progress.
- Allows instrument team to reprocess their data.

Data files:

- /SNS/ARCS/IPTS-10477/0/50140/preNeXus/ARCS_50140_pulseid0.dat
- /SNS/ARCS/IPTS-10477/0/50140/preNeXus/ARCS_50140_runinfo.xml
- /SNS/ARCS/IPTS-10477/0/50140/preNeXus/ARCS_50140_neutron3_event.dat
- /SNS/ARCS/IPTS-10477/0/50140/preNeXus/ARCS_50140_neutron0_event.dat
- /SNS/ARCS/IPTS-10477/0/50140/preNeXus/ARCS_50140_bmon2_histo.dat
- /SNS/ARCS/IPTS-10477/0/50140/preNeXus/ARCS_50140_pulseid3.dat
- /SNS/ARCS/IPTS-10477/0/50140/ARCS_50140.meta.xml
- /SNS/ARCS/IPTS-10477/0/50140/preNeXus/ARCS_50140_cvinfo.xml
- /SNS/ARCS/IPTS-10477/0/50140/preNeXus/ARCS_50140_bmon1_histo.dat
- /SNS/ARCS/IPTS-10477/0/50140/NeXus/ARCS_50140_histo.nxs
- /SNS/ARCS/IPTS-10477/0/50140/NeXus/ARCS_50140_event.nxs

Reduced files:

- /SNS/ARCS/IPTS-10477/shared/autoreduce/ARCS_50140_autoreduced.nxs.png
- /SNS/ARCS/IPTS-10477/shared/autoreduce/ARCS_50140_autoreduced.nxspe
- /SNS/ARCS/IPTS-10477/shared/autoreduce/ARCS_50140_autoreduced.nxs
- /SNS/ARCS/IPTS-10477/shared/autoreduce/reduction_log/ARCS_50140_event.nxs.log

Message	Information	Time
reduction_catalog.complete	autoreducer1.sns.gov	May 22, 2014, 6:28 a.m.
reduction_catalog.started	autoreducer1.sns.gov	May 22, 2014, 6:28 a.m.
reduction_catalog.data_ready		May 22, 2014, 6:28 a.m.
reduction.complete	autoreducer1.sns.gov	May 22, 2014, 6:28 a.m.
catalog.complete	autoreducer1.sns.gov	May 22, 2014, 6:27 a.m.
reduction.started	autoreducer1.sns.gov	May 22, 2014, 6:27 a.m.
catalog.started	autoreducer1.sns.gov	May 22, 2014, 6:27 a.m.
reduction.data_ready		May 22, 2014, 6:27 a.m.
catalog.data_ready		May 22, 2014, 6:27 a.m.
postprocess.data_ready		May 22, 2014, 6:27 a.m.

Submit for post-processing: catalog | reduction | all post-processing

Failure Rates

- Web monitor: never
- AR clients: never
- Workflow manager: < 2/yr
- AMQ brokers: once per 2-3 months
- Workflow DB: frequent high IO due to system. No actual problem. The issue is being resolved.
- ICAT: has to be restarted once every 6 weeks
- DASMON listener: once a month, due to high traffic and WorkflowDB IO problems.

Where's the code?

Web monitor and workflow manager:

https://github.com/neutrons/data_workflow

Auto-reduction clients:

https://github.com/neutrons/post_processing_agent

Reduction script repo:

https://github.com/mantidproject/autoreduce/tree/master/ReductionScripts/sns