State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

Data Analysis Infrastructure for DLS MX/CX

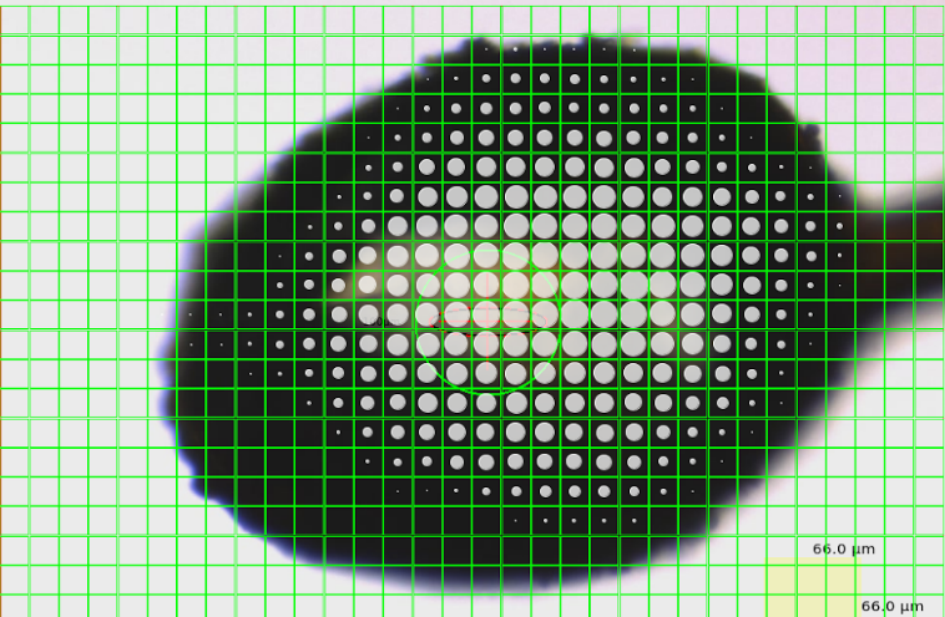M. Gerstel   R. Gildea   K. Levik   A. Ashton   G. Winter

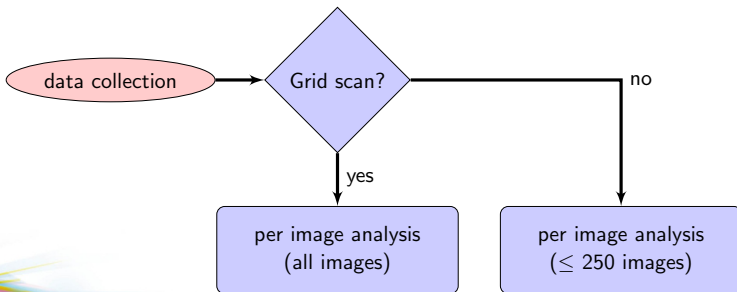Diamond Light Source, Harwell, Oxfordshire, UK

NOBUGS 2016

diamond

State of Affairs
●○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

Diamond Light Source

State of Affairs
○●○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

GDA grid scan

66.0 μm

66.0 μm

State of Affairs
○○●○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

Grid scan processing

GDA grid scan runs a script

State of Affairs
0000●0000

ActiveMQ
00000

Messaging Infrastructure
000000000000

Scripts run more scripts

```
#!/bin/sh
#
#script that is run at the end of every collect in MX
# original author of this script is Alun Ashton 240608
#
# 110831 added archiving to the end - awa
#
# input to script:
# automaticProcessing_Yes 4315 /dls/i03/data/2008/0-0 test2_MIS1_4_###.img /dls/i03/data/2008/0-0/

#Create A full and proper log/script for re running reprocessing, AWA 131022
echo '#' `date` >> ${3}/spool/RunAtEndOfCollectLog.sh
processing_dir=`/dls_sw/apps/mx-scripts/misc/RtnProcessingDir.sh ${2}`
echo "test -d "${processing_dir}" && mv "${processing_dir} ${processing_dir}"orig" >> ${3}/spool/RunAtEndOfCollectLog.sh
echo $0 $* >> ${3}/spool/RunAtEndOfCollectLog.sh
echo "sleep 120" >> ${3}/spool/RunAtEndOfCollectLog.sh

#
# Pause for nfs
#
umask 002

Pause=60

Visit=`basename $3`
ImageDirectory=${5}
[[ $ImageDirectory != */ ]] && ImageDirectory="$ImageDirectory"/
echo RHEDC ImageDirectory $ImageDirectory

st BEAMLINE env or set
   test $BEAMLINE ; then
       LINE=`echo $3 | cut -d/ -f3`

      /dls_sw/apps/mx-scripts/misc/isItADataCollection $2`

      ion == "True" ]] ; then

      ppear in a separate job for monitoring reasons
      ti/auto-xia2/run_are_all_images_there.sh $1 $2 $3 $4 $ImageDirectory $6

#
# Run XIA2

if [[ $DataCollection == "True" ]] ; then

   if [[ $BEAMLINE != "i23" ]] ; then
      /dls_sw/apps/mx-scripts/auto-xia2/run-fast-dp-refactor.sh $1 $2 $3 $4 $ImageDirectory $6 > /tmp/`date +%s`2>&1 & #/dev/null 2>&1
   fi

   if [[ $(L) == "automaticProcessing_Yes" ]] ; then
      rundirectxia2='/dls_sw/apps/mx-scripts/auto-xia2/RunDirectFromGDA.sh'
      runautoPROC='/dls_sw/apps/mx-scripts/autoPROC/RunDirectFromGDA.sh'
      eval $rundirectxia2 ` $1 $2 $3 $4 $ImageDirectory $6> /dev/null 2>&1 &`
      eval $runautoPROC ` $1 $2 $3 $4 $ImageDirectory $6> /dev/null 2>&1 &`

   echo "`date +%Y` `$rundirectxia2 $1 $2 $3 $4 $ImageDirectory $6`" >> /dls_sw/$BEAMLINE/logs/RunAtEndOfCollectCommand.log
   echo "`date +%Y` `$runautoPROC $1 $2 $3 $4 $ImageDirectory $6`" >> /dls_sw/$BEAMLINE/logs/RunAtEndOfCollectCommand.log

fi
#
# Auto Run EDNA
#
if [[ $BEAMLINE != "i23" ]] ; then
```
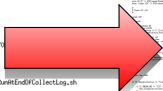
State of Affairs
○○○○●○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

Scripts run more scripts


```
#!/bin/sh
#
#script that is run at the end of every collect in MX
# original author of this script is Alun Ashton 240608
#
# 110831 added archiving to the end - awa
#
# input to script:
# automaticProcessing_Yes 4315 /dls/i03/data/2008/0-0 test2_MIS1_4_###.img /dls/i03/data/2008/0-0/0

#Create A full and proper log/script for re running reprocessing. AWA 131022
echo "#" date -> ${3}/spool/RunAtEndOfCollectLog.sh
processing_dir=`/dls_sw/apps/mx-scripts/mx-scripts/RtnProcessingDir.sh ${2}`
echo "test -d "${processing_dir}" && mv "${processing_dir} ${processing_dir}"orig" >> ${3}/spool/RunAtEndOfCollectLog.sh
echo $0 $* >> ${3}/spool/RunAtEndOfCollectLog.sh
echo "sleep 120" >> ${3}/spool/RunAtEndOfCollectLog.sh

# Pause for nfs

umask 002

Pause=60

Visit=`basename $3`
ImageDirectory=${5}
[[ $ImageDirectory != */ ]] && ImageDirectory="$ImageDirectory"/
echo RWEDC ImageDirectory $ImageDirectory

#set BEAMLINE env or set
# test $BEAMLINE ] then
#    BEAMLINE=`echo $3 | cut -d/ -f3`

    `/dls_sw/apps/mx-scripts/misc/isItADataCollection $2`
       ion == "True" ]] ; then

   appear in a separate job for monitoring reasons
      bin/auto-xia2/run_are_all_images_there.sh $1 $2 $3 $4 $ImageDirectory $6

# Run XIA2

if [[ $DataCollection == "True" ]] ; then

    if [[ $BEAMLINE != "i23" ]] ; then
       /dls_sw/apps/mx-scripts/auto-xia2/run-fast-dp-refactor.sh $1 $2 $3 $4 $ImageDirectory $6 > /tmp/`date +%s` 2>&1 & #/dev/null 2>&1
    fi

    if [[ $(L) == "automaticProcessing_Yes" ]] ; then
    rundirectxia2='/dls_sw/apps/mx-scripts/auto-xia2/RunDirectFromGDA.sh'
    runautoPROC='/dls_sw/apps/mx-scripts/autoPROC/RunDirectFromGDA.sh'
    eval $rundirectxia2 ` $1 $2 $3 $4 $ImageDirectory $6> /dev/null 2>&1 &`
    eval $runautoPROC ` $1 $2 $3 $4 $ImageDirectory $6> /dev/null 2>&1 &`
    fi

    echo "`date +%S`  $rundirectxia2 $1 $2 $3 $4 $ImageDirectory $6" >> /dls_sw/$BEAMLINE/logs/RunAtEndOfCollectCommand.log
    echo "`date +%S`  $runautoPROC $1 $2 $3 $4 $ImageDirectory $6" >> /dls_sw/$BEAMLINE/logs/RunAtEndOfCollectCommand.log

fi

# Auto Run EDNA

if [[ $BEAMLINE != "i23" ]] ; then
```


qsub -cwd

qsub -cwd

qsub -cwd

◆ diamond

State of Affairs
○○○○○●○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

Data set processing

State of Affairs
○○○○○○○●○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

Data set processing II

State of Affairs

○○○○○○○●○

ActiveMQ

○○○○○

Messaging Infrastructure

○○○○○○○○○○○○

Limitations

State of Affairs
ooooooooeo

ActiveMQ
ooooo

Messaging Infrastructure
oooooooooooo

Limitations

- Maintainability

State of Affairs
○○○○○○○●○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

Limitations

- Maintainability
- Testing

State of Affairs
ActiveMQ
Messaging Infrastructure
○○○○○○○●○
○○○○○
○○○○○○○○○○○○

Limitations

- Maintainability
- Testing
- Monitoring

Limitations

- Maintainability
- Testing
- Monitoring
- Data reprocessing

State of Affairs
○○○○○○○○●○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

Limitations

- Maintainability
- Testing
- Monitoring
- Data reprocessing
- Efficient use of computing resources

diamond

DECTRIS EIGER X

State of Affairs
○○○○○○○○○●

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

DECTRIS EIGER X

100 Gbit/s

State of Affairs
○○○○○○○○○●

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

DECTRIS EIGER X

State of Affairs
○○○○○○○○○

ActiveMQ
●○○○○

Messaging Infrastructure
○○○○○○○○○○○○

ActiveMQ

State of Affairs
○○○○○○○○○

ActiveMQ
●○○○○

Messaging Infrastructure
○○○○○○○○○○○○

ActiveMQ

State of Affairs
○○○○○○○○○○

ActiveMQ
○●○○○○

Messaging Infrastructure
○○○○○○○○○○○○
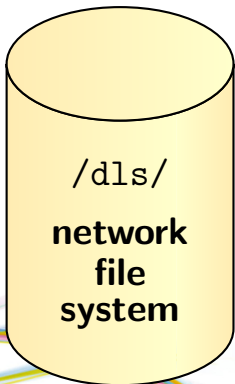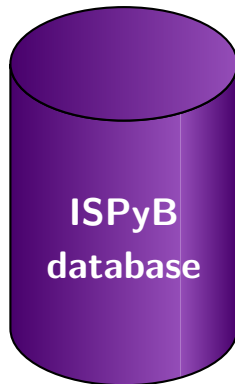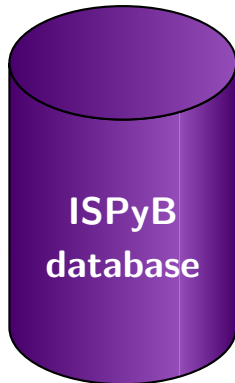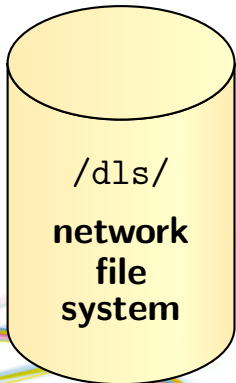
ActiveMQ

+ Fast
+ Lightweight
+ Not invented here
+ Broad language support
+ Already used by other groups
− Documentation

State of Affairs
○○○○○○○○○○

ActiveMQ
○○●○○

Messaging Infrastructure
○○○○○○○○○○○○

Terminology

- **Broker**: Central server

diamond

State of Affairs
000000000

ActiveMQ
○○●○○

Messaging Infrastructure
000000000000

Terminology

- **Broker**: Central server
- **Queue**: Producer–Consumer

diamond

State of Affairs
000000000

ActiveMQ
00●00

Messaging Infrastructure
000000000000

Terminology

- **Broker**: Central server
- **Queue**: Producer–Consumer
- **Topic**: Publish–Subscribe

diamond

State of Affairs
000000000

ActiveMQ
00●00

Messaging Infrastructure
000000000000

Terminology

- **Broker**: Central server
- **Queue**: Producer–Consumer
- **Topic**: Publish–Subscribe
- **Transaction**: All or nothing

diamond

State of Affairs
○○○○○○○○○

ActiveMQ
○○○●○

Messaging Infrastructure
○○○○○○○○○○○○

Delivery guarantee with transactions

State of Affairs
○○○○○○○○○○

ActiveMQ
○○○○●

Messaging Infrastructure
○○○○○○○○○○○○○

ActiveMQ

State of Affairs
○○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
●○○○○○○○○○○○○

ActiveMQ

State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○●○○○○○○○○○○
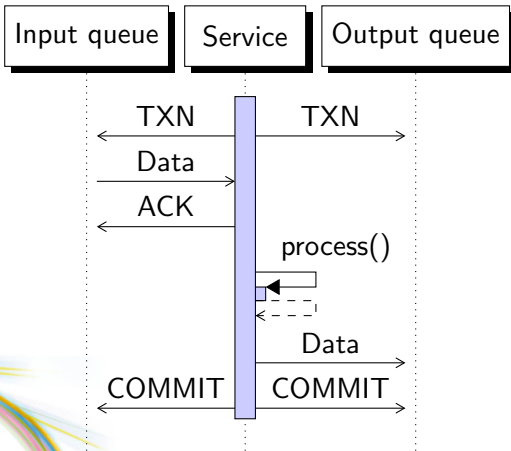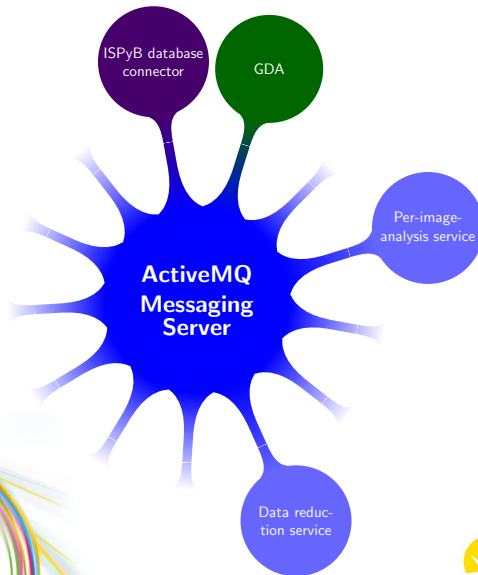
GDA grid scan announces data collection

${DCID}
+ do-this

/queue/data-collection

State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○●○○○○○○○○○

Data collection triggers processing

/queue/fast_dp

${DCID}

/queue/data-collection

${DCID}

/queue/xia2

${DCID}

/queue/multi-xia2

diamond

State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○●○○○○○○○○
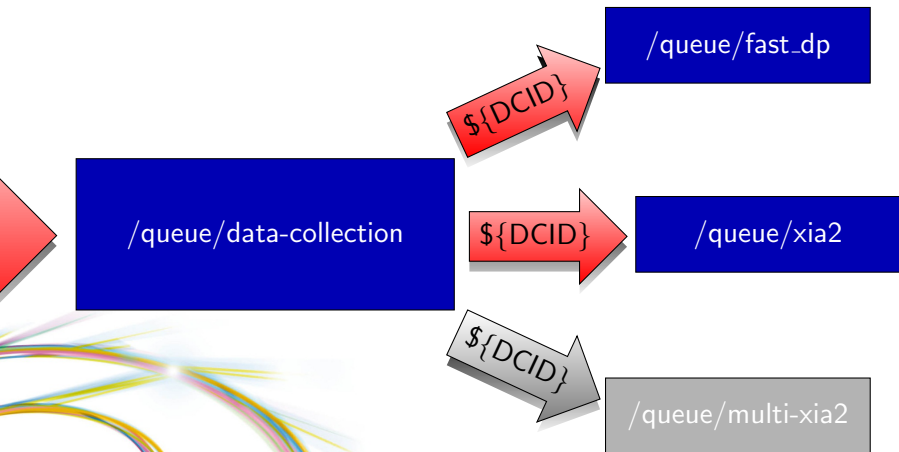
Supervised & encapsulated processing

xia2 manager:
run task
monitor status
publish status & results

/queue/xia2

drmaa
submit

xia2 pipeline=dials...

diamond
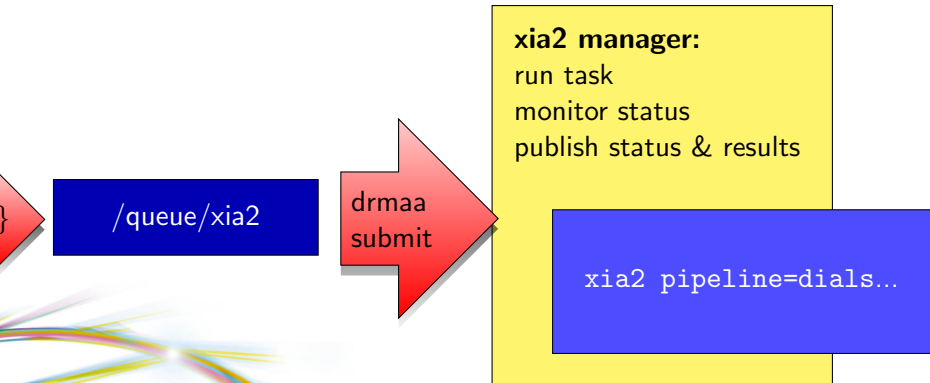
Reporting results

xia2 manager:
run task
monitor status
publish status & results

xia2 results → /queue/xia2-results

xia2 pipeline=dials...
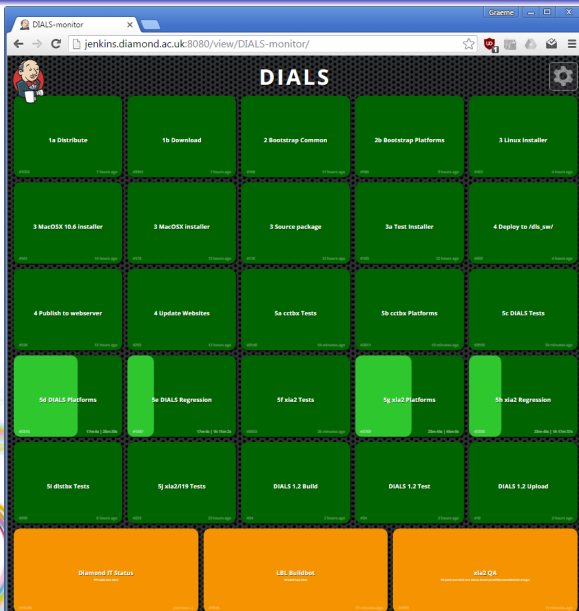
State of Affairs
○○○○○○○○○
ActiveMQ
○○○○○
Messaging Infrastructure
○○○○○●○○○○○○

Reporting results

ISPyB

results

/queue/xia2-results

results

GDA results panel

results/status

QoS monitor

diamond

State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○●○○○○○

Status monitor

State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○●○○○○
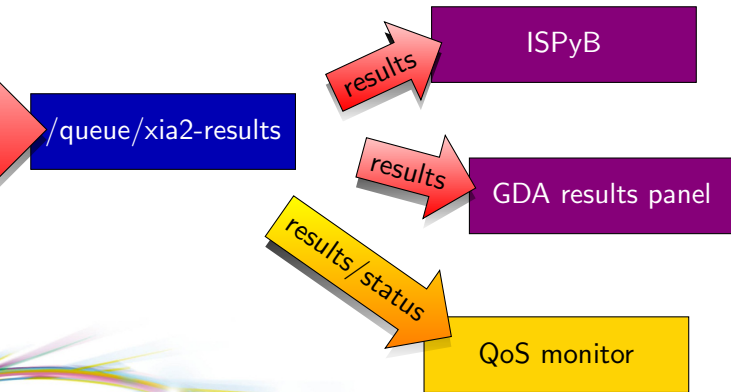
Status monitor

Past Day   Past Week   Past Month

APP SERVER AVAILABILITY
100.0%
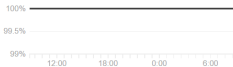
PAGES BUILDS FAILURE RATE
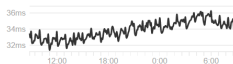## 0.1618%

MEAN WEB RESPONSE TIME
158ms

EXCEPTION PERCENTAGE
0.0001%

MEAN API RESPONSE TIME
33ms

MEAN HOOK DELIVERY TIME
0.85s
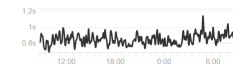
98TH PERC. WEB RESPONSE TIME
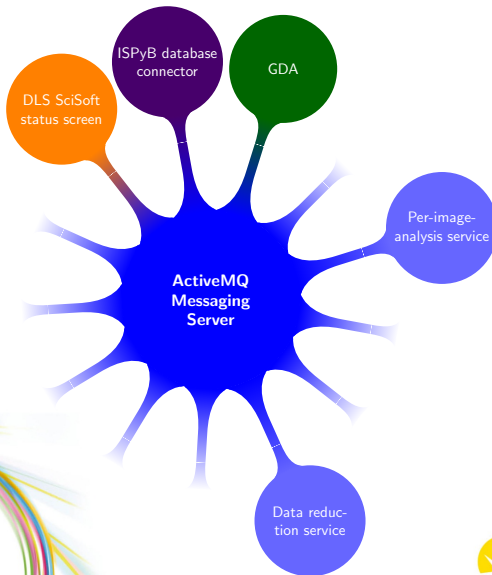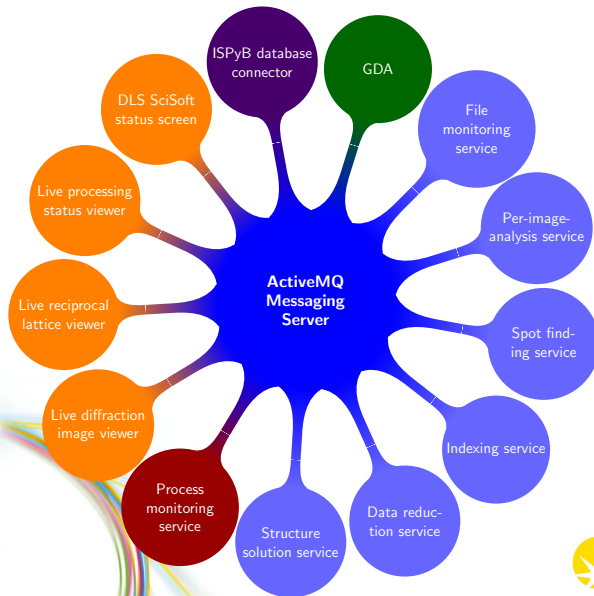## 700ms

98TH PERC. BROWSER TIME TO FIRST BYTE
2458ms

diamond

State of Affairs
ooooooooo

ActiveMQ
ooooo

Messaging Infrastructure
oooooooo●ooo

Infrastructure Overview

State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○●○○

Infrastructure Overview

State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○●○

Infrastructure Overview

State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○
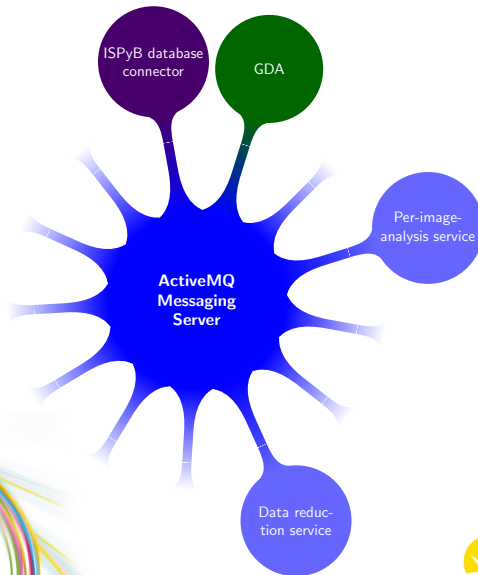
Messaging Infrastructure
○○○○○○○○○○○○●

Timeframe

- Diamond currently in shutdown for ring upgrade
- Short (2 week) run in December 2016
- New infrastructure to be deployed on 2-3 beamlines
- Running in parallel to existing scripts
- Begin replacing existing infrastructure from early 2017

diamond

State of Affairs
○○○○○○○○○

ActiveMQ
○○○○○

Messaging Infrastructure
○○○○○○○○○○○○

Thank you