Status of the Development of the Experiment Data Acquisition Pipeline for the European Spallation Source

A. H. C. Mukai[1], M. J. Christensen[1], J. Nilsson[1], M. Shetty[1], T. S. Richter[1], M. Brambilla[2], M. Könnecke[2], D. Werder[2], F. A. Akeroyd[3], M. J. Clarke[3], M. D. Jones[4]

[1]European Spallation Source ERIC, Denmark,
[2]Paul Scherrer Institut, Switzerland, [3]Science and Technology Facilities Council, United Kingdom, [4]Tessella, United Kingdom

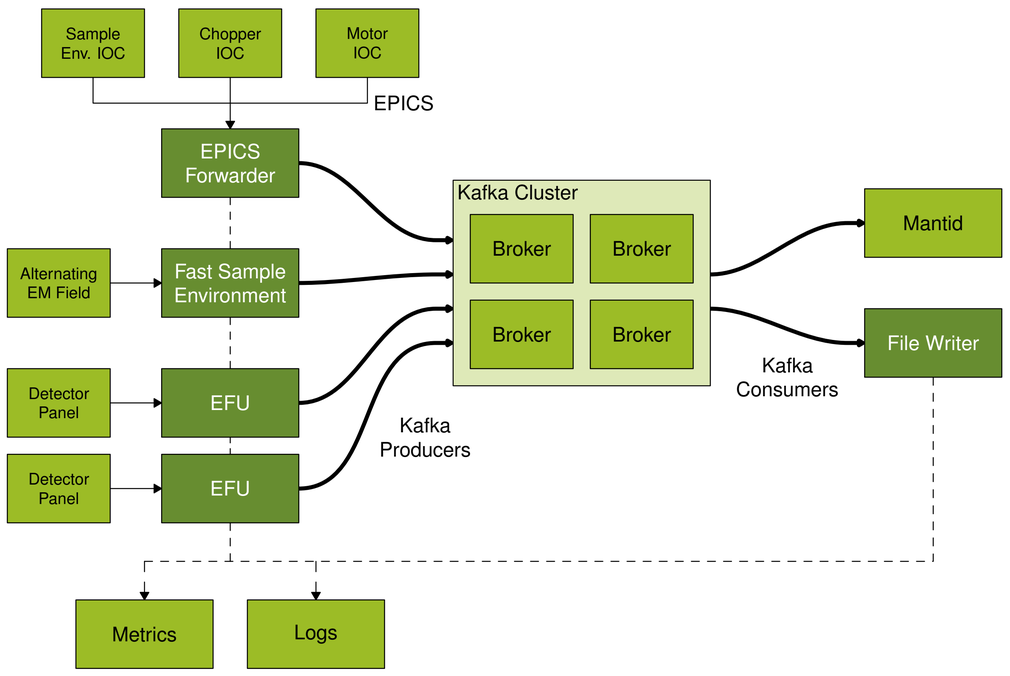
Event Mode Acquisition

Instrument data at ESS will be acquired mainly in event mode with no hardware veto. Accelerator pulse information and chopper top dead centre (TDC) signals will be acquired and attached to the datasets.

Anticipated rates for some early ESS instruments, considering 64 bits per event.

Instrument	Rate on Sample [n/s]	Rate on Detector [Hz]	Data Rate [MB/s]
BEER	10^9	2×10^5	1.6
C-SPEC	10^8	2×10^5	1.6
DREAMS	3.4×10^8	10^7	80
ESTIA		10^8	800
FREIA	5×10^8	1.2×10^7	96
HEIMDAL	2×10^9	8×10^6	64
LOKI	$\leq 10^9/cm^2$	4×10^7	320
SKADI	$\leq 10^9/cm^2$	4×10^7	320
T-REX	10^8	2×10^5	1.6

Data Pipeline Architecture



System architecture and data flow. The components in dark are being developed as part of the pipeline; dashed lines represent metrics and log messages.

ESS instruments will use an aggregator-based data acquisition pipeline. Apache Kafka was chosen as the central technology for aggregation and streaming, using Google FlatBuffers for serialisation.

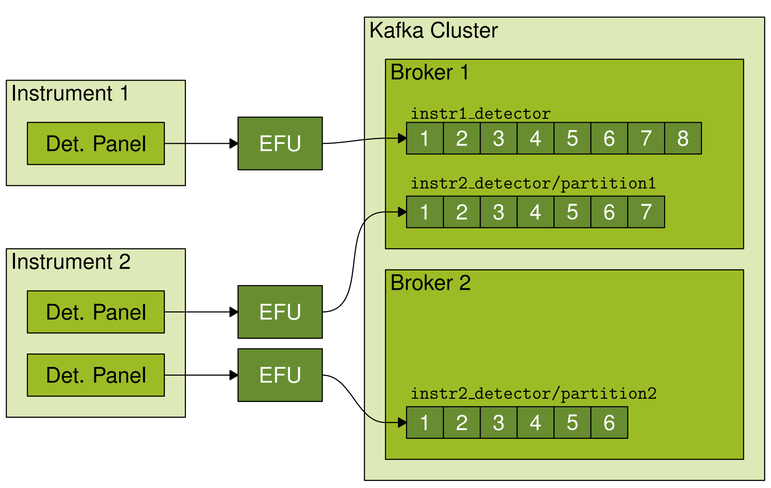
In this architecture, producers and consumers of data are decoupled and exchange data through the aggregator.

Kafka can also be used as a command bus, and dedicated topics have been created for control and status messages to be exchanged by applications, using JSON. These messages are currently used for commands like starting and stopping EPICS data forwarding and file writing, and also to send configuration to some of the applications at runtime.

Status of the Pipeline

Event Formation Unit

Event Formation Units (EFUs) will have a direct fibre connection to detector panels and process their output to generate detector ID and timestamp pairs. Processing happens in software, allowing different methods to be easily prototyped and tested.

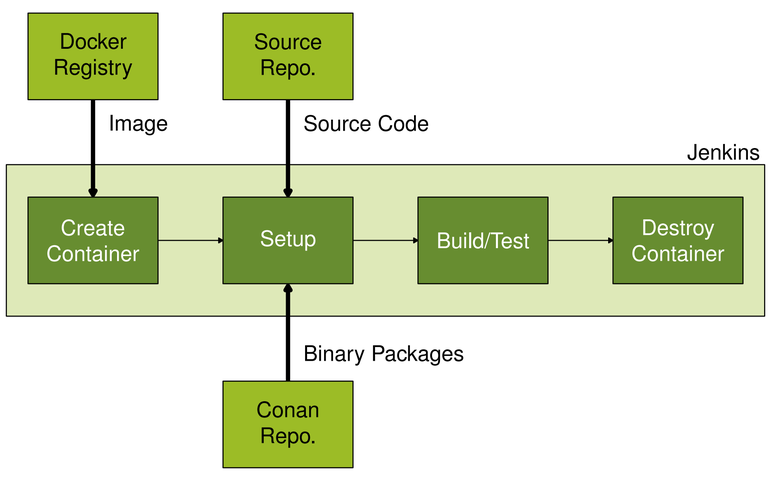


Event Formation Units can send events from different detector panels to different Kafka topic partitions.

Building and Testing

An automated integration test is run twice a day, deploying applications to a set of virtual machines and running a series of commands and checks. Build and test automation are done using Jenkins.

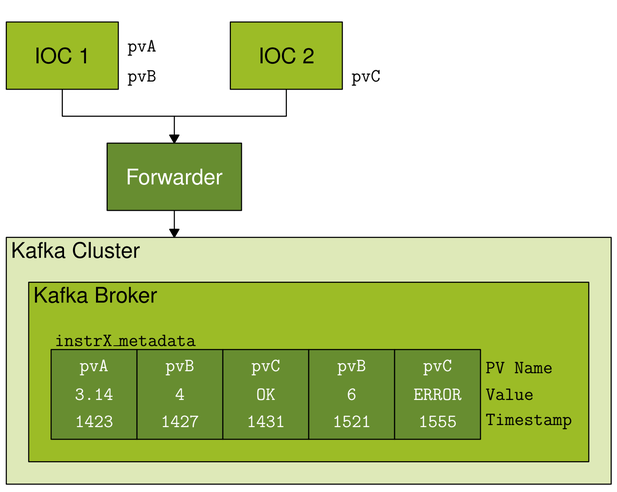
Work is going on to improve reproducibility of the automated software builds in Jenkins by executing them inside Docker containers and using the Conan C++ package manager.



Jenkins build job workflow using Docker containers and a local Conan package repository.

EPICS Forwarding

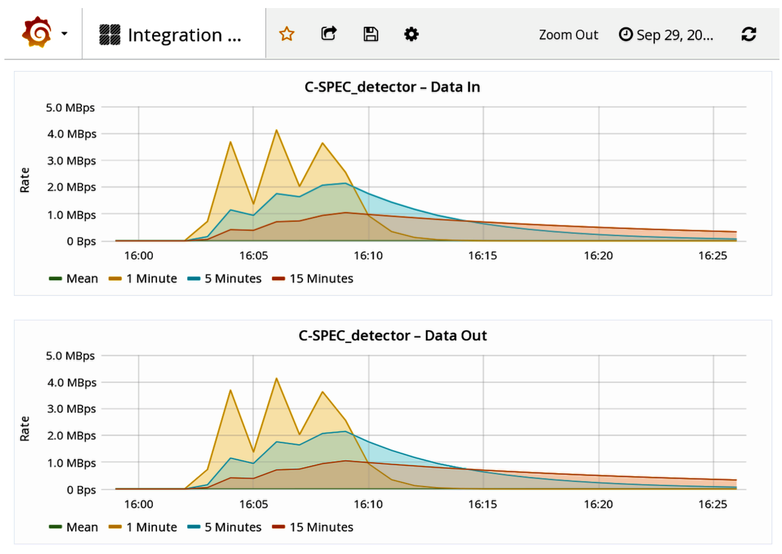
Non-neutron metadata will arrive through EPICS from sources such as the instrument choppers, motors and sample environments. The EPICS to Kafka Forwarder is being developed to monitor PVs of interest and forward their values to the Kafka cluster.



The EPICS Forwarder multiplexes updates from multiple PVs over a Kafka topic.

Metrics and Logging

Graphite is currently being used for storing metrics and Grafana is used for visualisation. Log messages are sent to a Graylog server, using the graylog-logger library developed at DMSC.



Grafana screen with Graphite metrics for the integration test Kafka cluster.

Fast Sample Environment

Measurements of alternating electric and magnetic fields, and some strain and pressure sensors are also expected to generate data at high rates, with anticipated frequencies on the order of 1 kHz to 1 MHz. Possibilities for handling these data include using EPICS and the Forwarder, a directly integrated Kafka producer, or sending data to an EFU.

NeXus File Writer

Experiment data will be persisted to files conforming to the NeXus standard for offline reduction, analysis and archiving. The NeXus File Writer subscribes to the complete set of Kafka topics containing neutron detector events and metadata for one instrument, deserialises the FlatBuffers messages and writes data to files according to a specified hierarchy and naming scheme.

Mantid

Mantid will be used for data reduction at ESS and work is in progress at DMSC and partner institutes to adapt it to consume data from Kafka in real time. This will allow it to be used to provide live visualisation and feedback for experiments.

Roadmap

Development of the pipeline components is progressing as planned, with regular automated tests being performed individually and in an integrated setup. Planned future work includes continuing the development of pipeline components and their integration with the experiment control program. Existing builds and tests will be improved by adding more checks and better reproducibility, with automated test environment configuration. Performance tests on physical machines will also be performed to evaluate the pipeline and tune the Kafka cluster configuration.

Acknowledgements

This work is partially funded by the European Union Framework Programme for Research and Innovation Horizon 2020, under grant agreement 676548.

EUROPEAN SPALLATION SOURCE

PAUL SCHERRER INSTITUT PSI

Science & Technology Facilities Council

brightness