

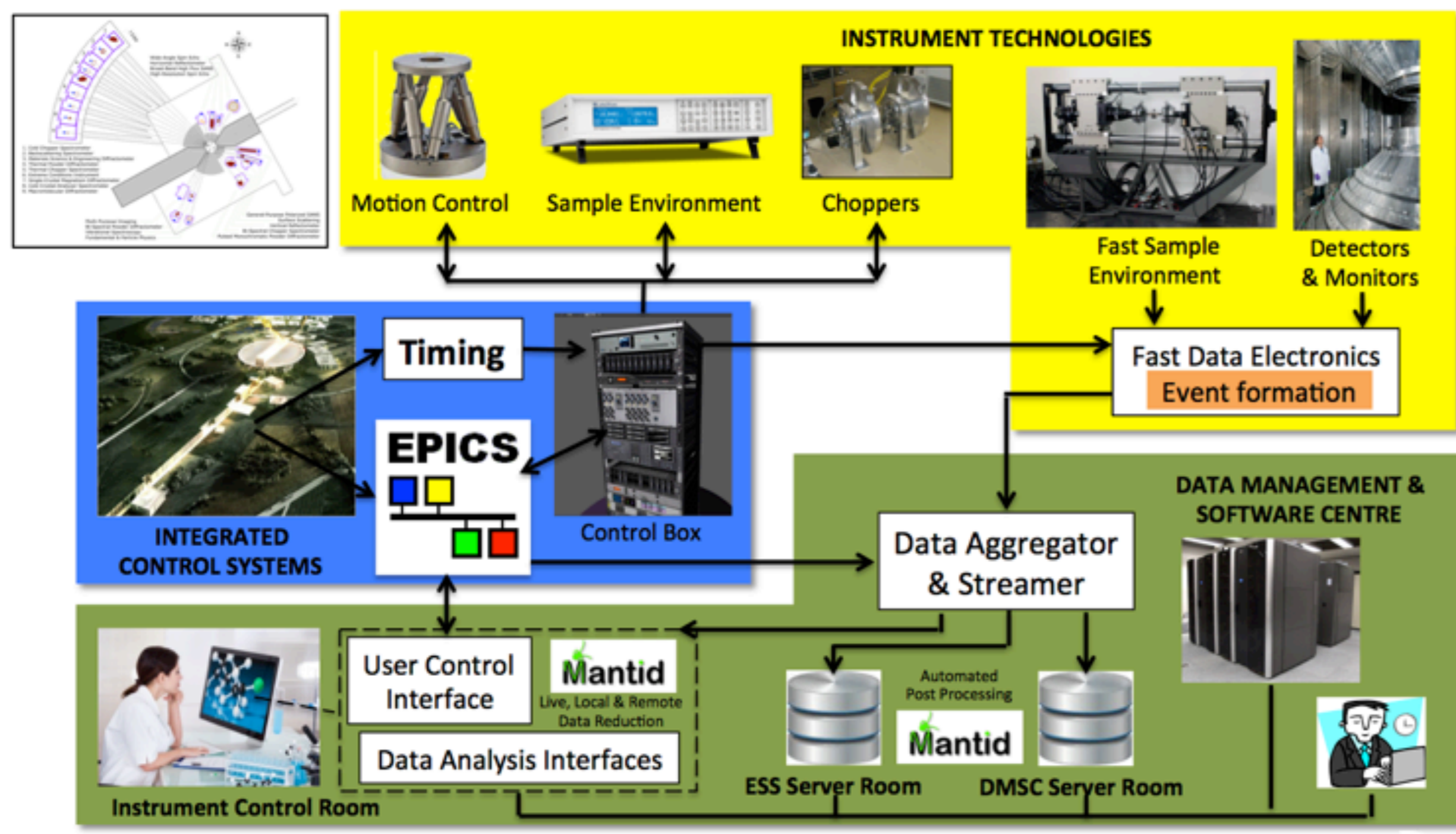
Development, testing and deployment of the ESS data aggregation and streaming software

Afonso Mukai¹, Tobias Richter¹, Mark Koennecke², Michele Brambilla², Dominik Werder², Matthew Jones³, Matt Clarke³, Frederick Akeroyd³

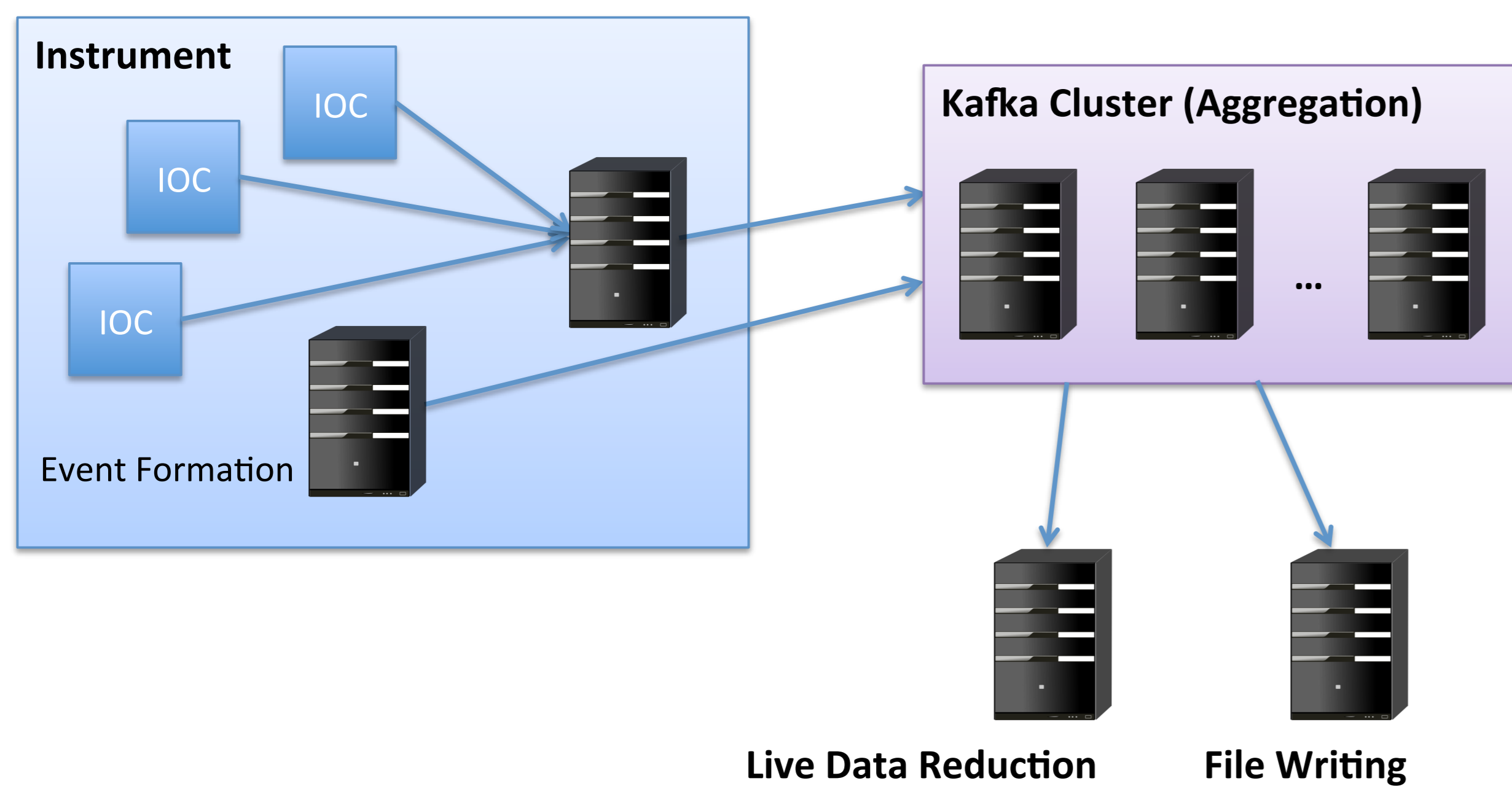
¹Data Management and Software Centre, ESS ERIC, Denmark, ²SINQ, Paul Scherrer Institut, Switzerland, ³ISIS, Science and Technology Facilities Council, United Kingdom

System overview

The data aggregation and streaming software system for ESS instruments is being developed under a BrightnESS Work Package at DMSC and SINQ, and with ISIS as an in-kind partner. It will aggregate neutron events from detectors and experiment metadata from sources such as choppers, sample environment and motion control, allowing clients that subscribe to the data to perform tasks including file writing and live data reduction and visualisation. Apache Kafka was chosen as the technology for data aggregation and streaming, using Google FlatBuffers for serialisation.



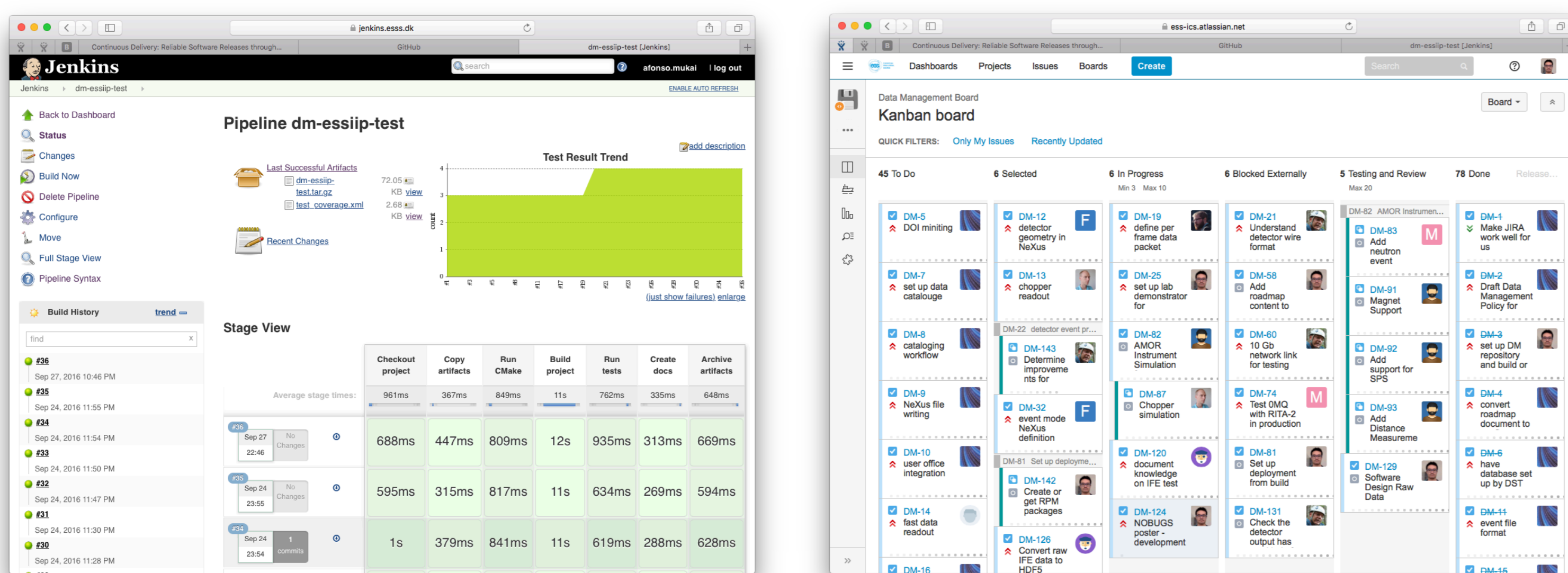
Apache Kafka is a distributed message broker, and the components generated by the project will possibly be deployed to several machines serving different instruments. Moreover, development is being done at three different locations. For testing, deployment and maintenance to be manageable, components must be traceable and auditable, and automation should be employed to the largest reasonably feasible extent.



Infrastructure and tools

The infrastructure available for tests consists of virtual machines at the DMSC computing infrastructure, and three servers with 10Gbit Ethernet interfaces at the ESS Instrument Integration Project (ESSIIP) Laboratory. Adopted tools include:

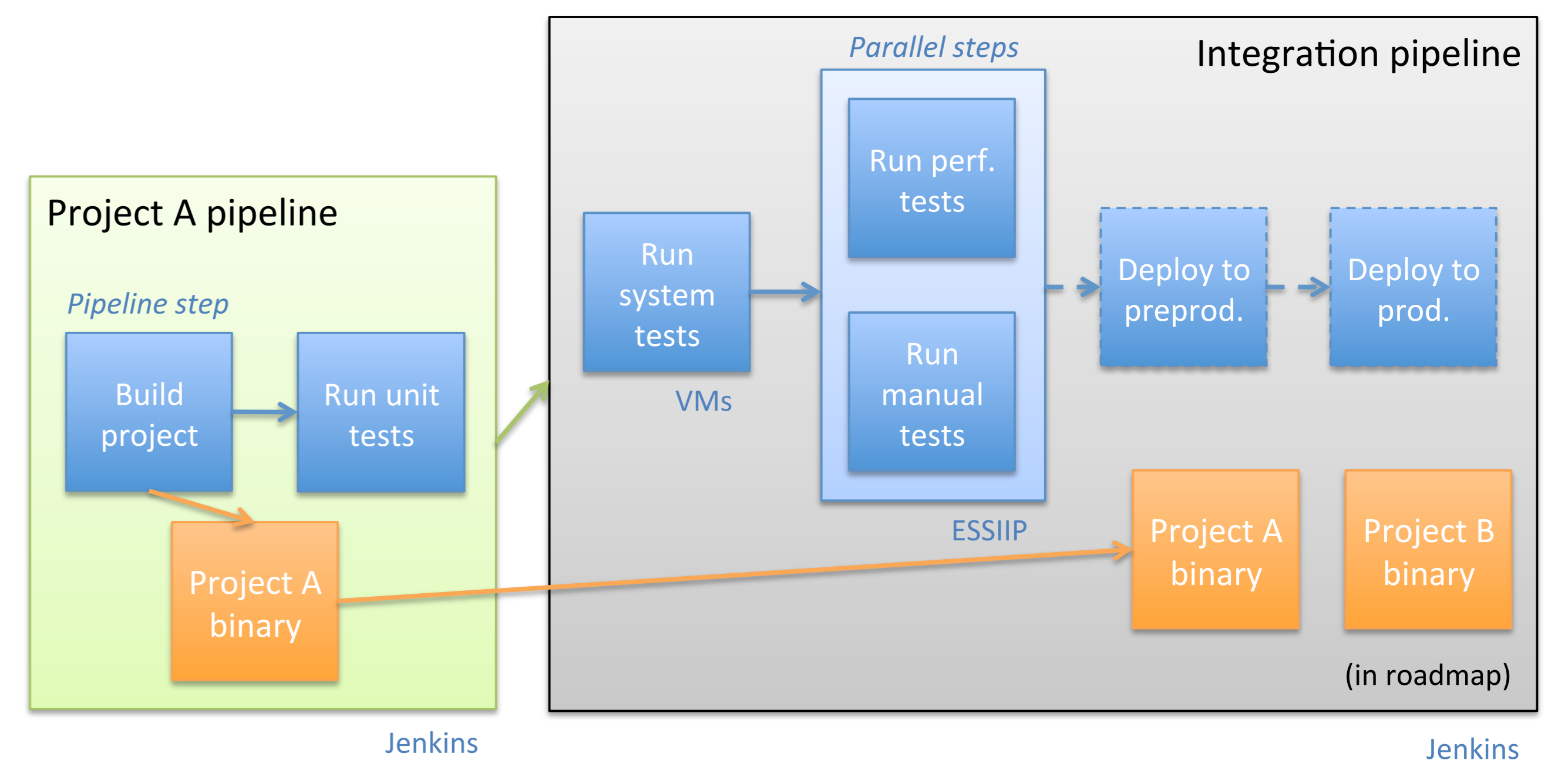
- Atlassian JIRA, Confluence and Bitbucket for collaboration
- CMake for build scripting: support for multiple platforms and IDEs
- Jenkins as a build server with pipelines under version control, C++ testing with googletest
- Ansible for deployment automation and configuration management: parameterisation and reproducibility



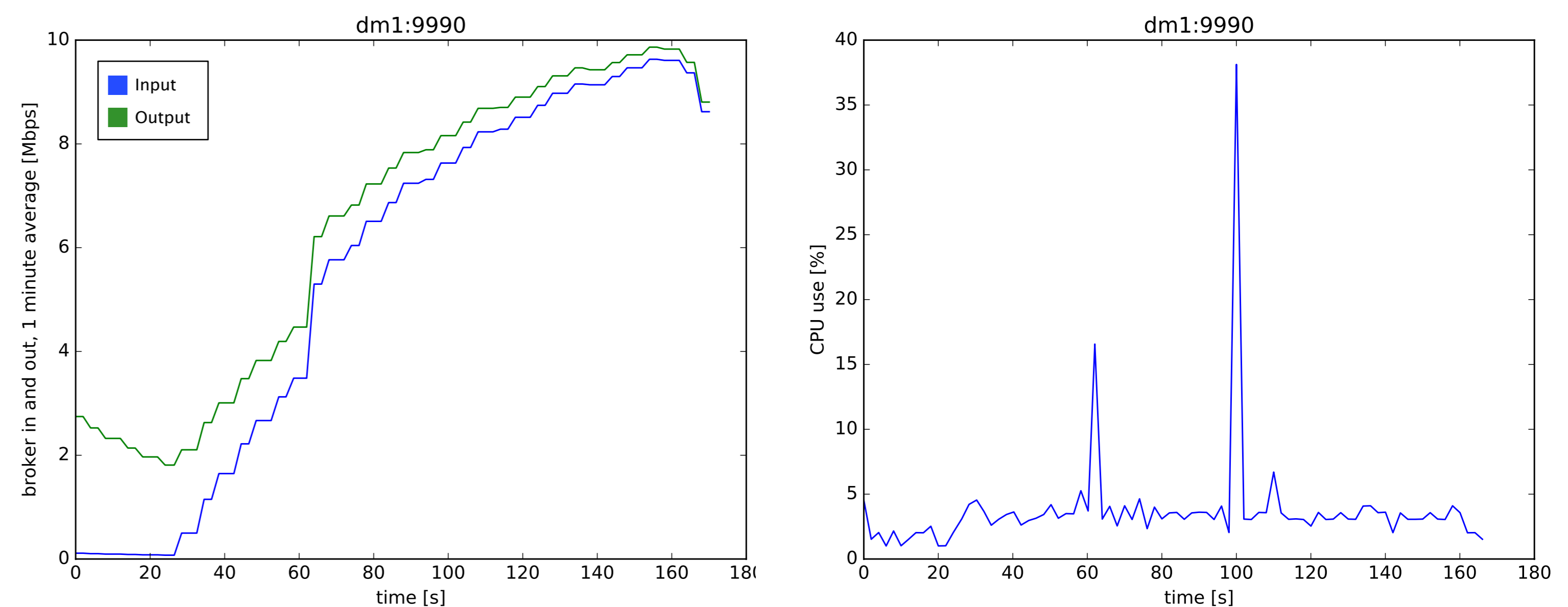
Current setup

In the current setup, part of the pipeline stages have been implemented and are integrated, and some software components have been developed to generate simulated data to be aggregated, so tests can be run.

- Build and unit test stages can be run on Jenkins
- Deployment to virtual machines and ESSIIP Laboratory servers is automated using Ansible, not yet fully integrated with the Jenkins pipeline
- Instrument metadata simulation with EPICS IOCs is available, based on the PSI instrument AMOR
- End-to-end tests have been developed to send and receive experimental data from NeXus files to and from a Kafka cluster, with remote monitoring; not yet automatically triggered and run inside the Jenkins pipeline



Software integration and delivery pipelines



Sample end-to-end test results on virtual machine

Roadmap

Next steps to be addressed are integration of Jenkins builds with the test environments using automated deployments with Ansible, along with automation of integration and performance tests and reporting on the ESSIIP servers. The objective is to achieve the integration of the complete pipeline, from commit to source control to automated deployment to the production environment under request of authorised users.

Acknowledgements

This work is being developed as part of the BrightnESS Work Package 5: Real-Time Management of ESS Data, under Task 5.3: Software to aggregate and make available the neutron event data and sample meta-data. BrightnESS is funded by the European Union Framework Programme for Research and Innovation Horizon 2020, under grant agreement 676548.

References

- [1] BrightnESS Deliverable Report 5.1: Design report data aggregator software
- [2] Humble J and Farley D 2010 *Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation* (Boston: Addison-Wesley Professional)