

Instrument Data Scientist Report

Gregory Tucker

October 20, 2022

New McStas components enable testing of data pipelines and workflows with realistic simulated data. A McStas instrument model of the BIFROST secondary spectrometer enables validation of a detector calibration procedure. Computing hardware for instrument users will be acquired soon for BIFROST and CSPEC.

1 Interfacing with the event pipeline

Neutrons absorbed in a detector will be converted into timestamped events and recorded to Event NeXus [1] files in a process referred to as the event data pipeline. The exact method and hardware which converts the analog signals from each absorption to one or more digital signals is instrument specific, but the remainder of the pipeline will be standardized into the following steps: 1. The digital information will be collected and combined into network packets by a Readout Master. 2. The Readout Master will send the packets to the Event Formation Unit (EFU) over a redundant network link. 3. The EFU then determines a detector element identifier and proton pulse time for each event and forwards the data to the Kafka Cluster. 4. Within the Kafka cluster, the identified event joins a Kafka stream to be written to disk by a File Writer. This process is visualized in figure 1.

The software components within the event data pipeline are within the purview of the **Experimental Control and Data Curation** (ECDC) group of the DMSC. The individual components and the combined pipeline have thus far been tested using synthetic data with only the proscribed data format but no further meaning. Such a testing regimen can only verify that the components and pipeline behave as engineered, and therefore can not expose faults in assumptions made about the data produced by a working instrument. Until such time as each instrument is in operation we can not test that all assumptions made in the component designs were appropriate, but by producing more-realistic input data that includes the scattering physics in addition to the correct data format we can better approximate the working instrument and hope to expose any flaws.

The neutron ray-tracing program **McStas** can simulate

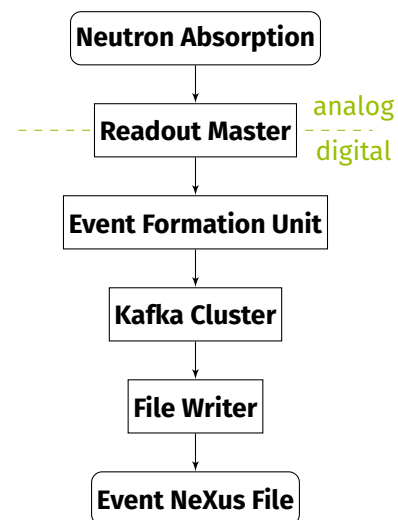


Figure 1: The data pipeline from neutron absorption to written Event NeXus file. The components Readout Master, Event Formation Unit, and File Writer are instances of per-instrument programs performing specific tasks. The Kafka Cluster handles instrument and sample environment data for all instruments simultaneously and therefore has many more inputs and outputs (not shown).

the flight path of a neutron from the source to the detector [2]. The simulated neutron interacts with each beamline component in series and, at each, may have its probability of continuing reduced (such that if the probability reaches zero, that neutron certainly does not pass the nullifying component) [3]. This technique allows efficient simulation of detected neutron intensities since the probability is proportional to the scattering cross section; it does not, however, directly produce individual neutron events as would be recorded by an instrument. Previous efforts to connect simulations to reduction workflows have focused on the Event NeXus file as intermediary [4], and side stepped any potential issues with the event data pipeline.

A new component for McStas has been produced which more-closely simulates the first digital readout from the configuration of three linear proportional detectors in series as will be used by BIFROST [5]. The BIFROST detector array will use triplet tubes in which the three tubes share one digitization chain. Each neutron that reaches the triplet detector component is tagged with the interaction time and two values representing the accumulated voltages at each end of the tube. A probability-weighted Monte Carlo choice can then be used to convert the tagged neutron rays into discrete neutron events. A second new McStas component, produced in collaboration with ECDC, is then used to mimic the Readout Master by collecting the tagged information from events into packets and sending those packets over a network interface to the Event Formation Unit [6]. Through the use of these new components, the first two blocks in figure 1 can be replaced by a McStas instrument producing a simulated event data pipeline, illustrated in figure 2.

Using the new McStas components to more-accurately simulate the output of the Readout Master will enable more-robust testing of the remainder of the data pipeline. Importantly, this will also enable production of realistic Event NeXus files from simulations for the ESS instruments. These realistic Event NeXus files can be used to construct and validate the data transformation pipeline which converts the recorded neutron events from detector element index and event time to $S(Q, E)$. In the future the same simulated event data pipeline can be used to support a digital twin of each instrument that would run in parallel to experiments.

2 Spectrometer simulations

Some developments have been made in the area of instrument simulations, while others remain as future goals.

2.1 BIFROST secondary spectrometer

A McStas instrument model of the BIFROST secondary spectrometer has been constructed and validated against the design specifications of the physical analyzer-detector

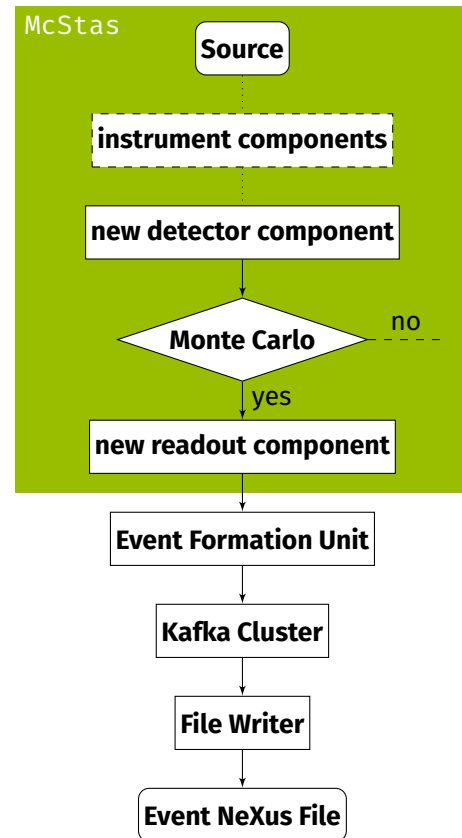


Figure 2: The simulated data pipeline from McStas source emission to written Event NeXus file. The components Event Formation Unit, File Writer and Kafka Cluster are as in figure 1. The shaded box at the top represents a McStas simulation using the new components described in the text which emulates the Readout Master output in figure 1.

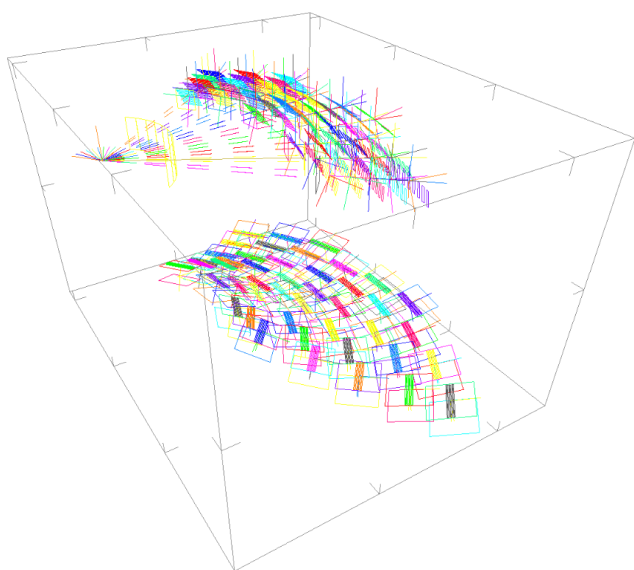


Figure 3: A representation of the McStas instrument model of the BIFROST secondary spectrometer.

construction. A view of the instrument model is shown in figure 3. In addition to enabling the creation of realistic Event NeXus files for validation of the data transformation pipeline as described above, this model will be used to validate a detector calibration procedure using Read-out Master data and to study the instrumental resolution of BIFROST along with collaborators at the University of Copenhagen.

The detector calibration will account for any detector misalignment. It will also be used to set the EFU parameters to correctly identify pixel indexes for every detector tube.

2.2 Multiple scattering effects

Within the idealized spectrometer, neutrons interact only with the components of the instrument and the sample no more than one time. Their interactions with sample environments or non-optical components, such as vacuum windows, in the beam path are typically ignored, as are scattering events which interact with the sample more than once. Understanding such secondary scattering effects may be crucial to extracting useful information from measured data in some cases.

The Union components of McStas can be used to simulate multiple scattering events in complex geometry instruments [7]. In order to understand potentially-confounding signals in data from the ESS spectrometers, robust Union McStas simulations should be prepared in the near future. As such simulation models can be complex to construct, expert input will be required from within the DMSC and the instrument teams of CSPEC and BIFROST.

3 Instrument user computing hardware

The hardware components that comprise each instrument will communicate over a per-instrument EPICS network [8]. These independent networks will send data directly to components on the digital side of figure 1, which will all be located in a single server room in building H01 on the ESS campus in Lund, under the purview of the Data Systems and Technology (DST) group within DMSC. DST plans to co-locate instrument control client machines as well as data reduction and data analysis machines in the H01 server room to simplify operation and management of the computing hardware. User interaction with these systems will be via web-accessible virtual machines and/or ‘thin-client’ graphical terminals located in the instrument control cabins.

DST will soon acquire the user-interface hardware for nine instruments including BIFROST and CSPEC, and together with the instrument scientists, we are assembling a list of requirements which the hardware must fulfill to be used by DST in their procurement. The currently envisioned list of user hardware for each instrument control cabin is: one instrument control terminal, with four monitors (per ECDC requirements for the instrument control software); at least one analysis terminal, each with a minimum of two monitors; status monitors for the proton accelerator, spallation target, instrument components, and sample environment; and a video conferencing system for efficient communication with off-site collaborators.

References

- [1] M. Könnicke et al., “The NeXus data format”, *J Appl Cryst* **48**, Number: 1 Publisher: International Union of Crystallography, 301 (2015).
- [2] P. Willendrup, E. Farhi, and K. Lefmann, “McStas 1.7 - a new version of the flexible Monte Carlo neutron scattering package”, *Physica B: Condensed Matter, Proceedings of the Third European Conference on Neutron Scattering* **350**, E735 (2004).
- [3] P. K. Willendrup and K. Lefmann, “McStas (i): Introduction, use, and basic principles for ray-tracing simulations”, *Journal of Neutron Research* **22**, Publisher: IOS Press, 1 (2020).
- [4] T. R. Nielsen, A. J. Markvardsen, and P. Willendrup, “McStas and Mantid integration”, *Journal of Neutron Research* **18**, Publisher: IOS Press, 79 (2015).
- [5] G. S. Tucker, (2022) <https://github.com/g5t/mcstas-detector-triplet> (visited on 2022-10-18).
- [6] G. S. Tucker and M. J. Christensen, (2022) <https://github.com/g5t/mcstas-broadcast-component> (visited on 2022-10-18).

- [7] M. Bertelsen, (2022) https://mads-bertelsen.github.io/tutorial/Union_tutorial_1_processes_and_materials.html (visited on 2022-10-20).
- [8] L. R. Dalesio et al., "The experimental physics and industrial control system architecture: past, present, and future", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **352**, 179 (1994).