Efficient loading of reduced data ensembles produced at ORNL SNS/HFIR neutron time-of-flight facilities

William F Godoy
*Computer Science and
Mathematics Division,*
Oak Ridge National Laboratory
Oak Ridge, TN, USA
Email: godoywf@ornl.gov

Andrei T Savici
Neutron Scattering Division,
Oak Ridge National Laboratory
Oak Ridge, TN, USA
Email: saviciat@ornl.gov

Steven E Hahn
*Computer Science and
Mathematics Division,*
Oak Ridge National Laboratory
Oak Ridge, TN, USA
Email: hahnse@ornl.gov

Peter F Peterson
*Computer Science and
Mathematics Division,*
Oak Ridge National Laboratory
Oak Ridge, TN, USA
Email: petersonpf@ornl.gov

Abstract—We present algorithmic improvements to the loading operations of certain reduced data ensembles produced from neutron scattering experiments at Oak Ridge National Laboratory (ORNL) facilities. Ensembles from multiple measurements are required to cover a wide range of the phase space of a sample material of interest. They are stored using the standard NeXus schema on individual HDF5 files. This makes it a scalability challenge, as the number of experiments stored increases in a single ensemble file. The present work follows up on our previous efforts on data management algorithms, to address identified input output (I/O) bottlenecks in Mantid, an open-source data analysis framework used across several neutron science facilities around the world. We reuse an in-memory binary-tree metadata index that resembles data access patterns, to provide a scalable search and extraction mechanism. In addition, several memory operations are refactored and optimized for the current common use cases, ranging most frequently from 10 to 180, and up to 360 separate measurement configurations. Results from this work show consistent speed ups in wall-clock time on the Mantid LoadMD routine, ranging from 19% to 23% on average, on ORNL production computing systems. The latter depends on the complexity of the targeted instrument-specific data and the system I/O and compute variability for the shared computational resources available to users of ORNL's Spallation Neutron Source (SNS) and the High Flux Isotope Reactor (HFIR) instruments. Nevertheless, we continue to highlight the need for more research to address reduction challenges as experimental data volumes, user time and processing costs increase.

Index Terms—experimental data, reduction, workflows, metadata, indexing, Mantid, NeXus, HDF5, neutron scattering

I. INTRODUCTION

Oak Ridge National Laboratory (ORNL) neutron science facilities, the Spallation Neutron Source (SNS) and the High-Flux Isotope Reactor (HFIR), produce large amounts of experimental raw data from a variety of instruments [1]. Raw data at SNS and HFIR is stored in "event mode" [2], in which each neutron is tagged with its arrival time, thus recording what is known as "time-of-flight" information. In addition, detector and measurement characteristics annotations must be stored, to capture meaningful physical information under certain conditions [3]. The result is a large database stored using the metadata-rich standard NeXus schema [4], built on top of the self-describing hierarchical data format, HDF5 [5], hosted at ORNL's SNS and HFIR computing facilities [6]. Raw event data needs to be post-processed into higher-level data products, to extract physical quantities of interest. Moreover, single experimental runs only expose limited information on complex material physics studied with neutron scattering techniques. Hence, experiments from a particular instrument need to cover a wide range of observational parameters (*e.g.* angular coverage, different temperatures or magnetic fields, etc.), which could also incur in significant overlap [7]. The latter results in large data set ensembles of different statistical significance, posing major challenges in post-processing and interpretation workflows.

In this paper, we tackle the data loading challenges for single-crystal inelastic experiments, which require many sample orientations [8]. The raw data and metadata from each orientation is pre-processed into a single "multidimensional" data set ensemble. It contains all the information required to visualize different slices of the dynamic structure factor, typical in neutron scattering techniques. Loading this single, common, "multidimensional" data set from disk and processing the results into a reduced data structure in memory for slicing the entire experimental domain is a major bottleneck in the analysis workflows at SNS and HFIR available to users.

The Mantid framework [9] is an international collaboration

between several neutron sciences facilities around the world; including ORNL's SNS and HFIR [1], the ISIS Neutron and Muons Source [10], and The Institut Laue–Langevin (ILL) [11]. Mantid is used by several production workflows for data reduction and analysis at these facilities. The present work is the result of the optimizations applied to the Mantid loading algorithm for multidimensional data ensembles named "LoadMD" [9]. The latter has been identified as a bottleneck in the process of data visualization for single crystal experiments. Two major changes are applied to this algorithm, the first one is the caching of an "in-memory index" [12] data structure, which reuses similar strategies outlined in our previous work addressing raw neutron data reduction bottlenecks [13]; the second change is the reduction of "LoadMD" memory allocation requests which, in a shared computational resource, can lead to increased variability in the overall loading operation. Impact is obtained from the wall-clock time scalability of "LoadMD" operations on SNS/HFIR production computational systems[1] as a function of the number of orientations, between 10 to 360 experimental measurements, stored in data sets ensembles that are typical of the targeted applications [7].

The paper is organized as follows, Section II provides a description of the neutron multidimensional data ensembles produced from raw event-based data and the loading and processing workflow for analyzing a full experiment. Section III describes the proposed methodology to address the identified current bottleneck in loading operations from profiling the current operations in "LoadMD", while Section IV presents results on SNS/HFIR data analysis computing systems, including the expected speed ups, scalability from comparing actual experimental data set sizes, and the variability in wall-clock times due to the shared-resource nature of the system. Finally, Section V presents the conclusions from our study outlining the need to continue improving existing data reduction workflows as data volumes continue to increase in future generation instruments at ORNL, thus resulting in a wealth of research opportunities for tackling scientific data challenges [14].

II. Neutron Multidimensional Data Ensembles

The targeted single-crystal neutron inelastic scattering instruments at SNS and HFIR produce raw event-based data for every experimental setup and conditions. In order to get information on an entire experiment, to construct a meaningful map of the dynamic structure factor $S(\mathbf{Q}, E)$, several measurements need to be taken at different sample orientations. A schematic description is shown in Figure 1 for constructing the resulting data ensemble. We use the Mantid terminology of "workspaces" [9] for various data structures in memory.

The raw data for each sample orientation (event workspaces) contain information about neutron detection events (time and position in a detector), together with some metadata about the sample orientation and experiment setup. One transforms this information into a coordinate system relevant to the physics of the materials to be studied, a three dimensional momentum
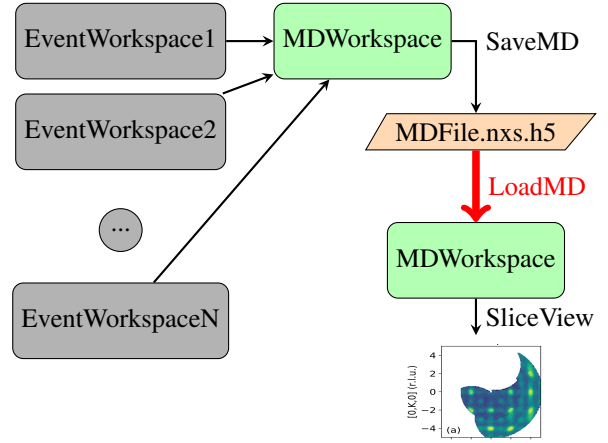
Fig. 1: Schematic overview of the multidimensional event workspace, "MDWorkspace", workflows used at SNS/HFIR using Mantid's "SaveMD" to save on disk (once) and the identified "LoadMD" bottleneck operation (executed many times) to retrieve the entire experimental ensemble containing several orientation outputs.

transfer $\mathbf{Q}$ and an energy transfer E. This is stored as a multidimensional workspace, "MDWorkspace". Information from multiple sample orientations (or other metadata parameters, such as sample temperature) can be converted to the same coordinates, and stored in the same structure. The resulting "MDWorkspace" container is then stored in a single HDF5 file using the standard NeXus schema.

For visualization and further analysis purposes, researchers typically take multiple one or two dimensional slices out of this workspace, aligned along some arbitrary directions, relevant to the particular sample they study. Pre-processing the raw data into the "MDWorkspace" in Fig. 1 is a computationally expensive process, but it is enough to perform it only once. The slicing operation occurs multiple times, with different slicing parameters. It is therefore desirable to store the intermediate data and load it several times using the "SaveMD" and "LoadMD" algorithms in Mantid. Typical 2D slices for various number of sample orientations are shown in Fig. 2.

The entries from a typical file ensemble are illustrated in Table I. Thus showing the scalable nature of the file size, hence operations, as a function of the number of experiments conducted and stored at the targeted SNS/HFIR instruments. For example, based on the current number of entries listed in Table I, a file containing 180 measurements can contain close to 200K entries. This leads to a scalability challenge when processing the metadata and data in "LoadMD".

As expected from being the inverse operation of "SaveMD", "LoadMD" essentially retrieves the stored information by looping through the experiments data and metadata from the overall ensemble. This is illustrated in Fig. 3 showing the steps for loading samples and log info for each experiment configuration in the dataset ensemble. While, multithreaded
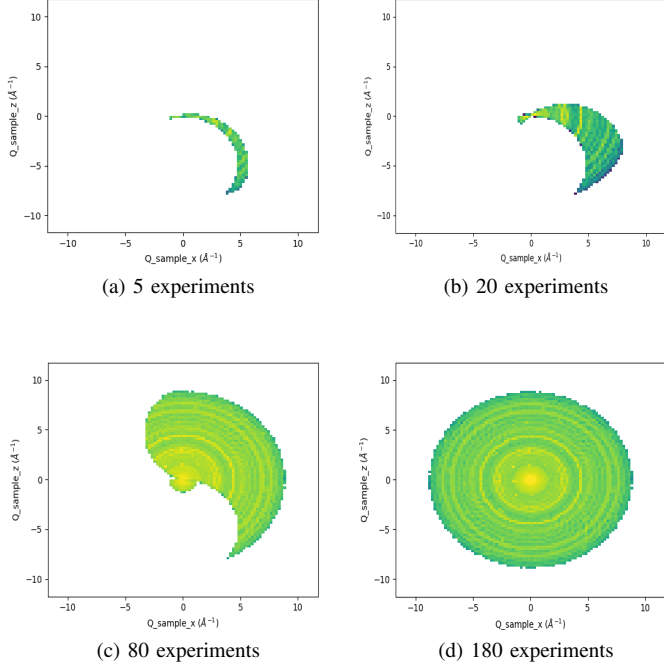
(a) 5 experiments

(b) 20 experiments

(c) 80 experiments

(d) 180 experiments

Fig. 2: Slice view results of the dynamic structure factor $S(\mathbf{Q}, E)$, as a function of two of the momentum transfer components (Q_x and Q_z). The multi-dimensional data is a synthetic one for the ARCS instrument [15], covering different orientation ranges of the sample ($2°$ offsets in the sample rotation).

Data Type	Entry Name
group	/entry
attribute	/entry/NX_class
	...
group	/MDEventWorkspace/experiment0
group	/MDEventWorkspace/box_structure
dataset	/MDEventWorkspace/coordinate_system
dataset	/MDEventWorkspace/dimensions
group	/MDEventWorkspace/event_data
	...
attribute	/MDEventWorkspace/experiment0/NX_class
group	/MDEventWorkspace/experiment0/instrument
	... 20 entries
group	/MDEventWorkspace/experiment0/logs
	... 700 to 1000 entries
group	/MDEventWorkspace/experiment0/sample
	... 17 entries
group	/MDEventWorkspace/process
	...
	...
	...
group	/MDEventWorkspace/experimentN
	...

TABLE I: Schematic representation of the hierarchical NeXus schema [4] for "MDWorkspace" files storing information for several experiments including instrument configuration information (logs, sample, process). Index contents scale up with the total number of experiments (N+1).
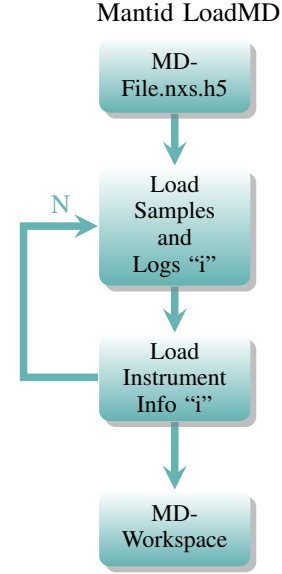
Mantid LoadMD

Fig. 3: Mantid's "LoadMD" loop steps for processing entries of a single input "MDWorkspace" NeXus file containing a "N" number of experiments.

parallelization strategies could be an option to speed up the present loop, it is still desired to load the entire experimental data ensemble. It is expected that more complex slicing patterns than those shown in Fig. 2 will not necessarily match the data ordering in the in-memory "MDWorkspace" data structure. As a result, focusing on the repeated functionality will guide initial efforts to locate bottlenecks that scale with the number of experiments stored in a single ensemble.

III. PROPOSED METHODOLOGY

Similar to our previous effort in speeding up the loading of raw neutron scattering data [13], we first understand the current bottlenecks in "LoadMD" due to existing access patterns previously identified in Mantid [16].

Figure 4 illustrates the result of profiling "LoadMD" from the existing algorithm in Mantid's main branch, as of October 2021, using flame graphs [17] for visualization purposes. The x axis indicates the relative proportion of time for functions inside "LoadMD" in which most time on the central processing unit (CPU) is consumed, while the y axis indicates the call stack down to the underlying libraries outside Mantid (in red). Figure 4a shows that functions associated with reconstructing the "in-memory" metadata index can take a large portion of CPU runtime ($\sim 20\%$), in addition a few cases of memory reallocation per experiment entry have been identified for potential refactoring that lead to savings in memory operations ($\sim 4\text{-}5\%$). The latter becomes important in the targeted shared resource as requests can add to the operating system (OS) overhead for finding available resources (*e.g.* contiguous memory).

To address the bottleneck due to in-memory index metadata reconstruction operations we reuse the binary-tree formulation described in our previous work [13]. This binary-tree is cached in memory from the beginning of "LoadMD" and is kept persistent through the loop operations described in Fig. 3. Table II shows a schematic representation of the proposed in-memory index for the reduced ensemble stored in a single HDF5 NeXus file with "SaveMD". Rather than following a hierarchical approach as the data outlined on disk, the entries are sorted by the `NX_class` attribute to match processing patterns observed in "LoadMD" as profiled in Fig. 4. Large data entries are identified as scientific datasets (SDS) and compose the majority of the entries. The final number of SDS entries scale up with the number of experiments stored in a single file.

Key: NX_class	Value: Sorted binary-tree with absolute-path entry key
NXdata	/MDEventWorkspace/box_structure
	/MDEventWorkspace/event_data
	/MDEventWorkspace/experiment0/sample/material
	/MDEventWorkspace/experiment1/sample/material
	...
	/MDEventWorkspace/experimentN/sample/material
NXentry	/MDEventWorkspace
NXgroup	/MDEventWorkspace/experiment0
	/MDEventWorkspace/experiment0/logs
	...
	/MDEventWorkspace/experimentN
	/MDEventWorkspace/experimentN/logs
	..
NXinstrument	/MDEventWorkspace/experiment0/instrument
	...
	/MDEventWorkspace/experimentN/instrument
NXpositioner	/MDEventWorkspace/experiment0/goniometer
	...
	/MDEventWorkspace/experimentN/goniometer
NXlog	/MDEventWorkspace/experiment0/logs/gd_prtn_chrg
	...
	/MDEventWorkspace/experimentN/logs/gd_prtn_chrg
SDS	/MDEventWorkspace/experiment0/instrument/...
	/MDEventWorkspace/experiment0/logs/...
	/MDEventWorkspace/experiment0/sample/...
	...
	/MDEventWorkspace/experimentN/instrument/...
	/MDEventWorkspace/experimentN/logs/...
	/MDEventWorkspace/experimentN/sample/...
	...

TABLE II: Schematic "in-memory" binary-tree index from MDFile NeXus file entries implemented using using C++'s `map<string,set<string>` data structure in our refactoring. Scientific data set (SDS) entries form the largest sub-tree.

A new profiling flame graph is shown in Fig. 4b after applying the proposed algorithmic changes to reduce the identified bottlenecks in "LoadMD". It can be seen that the index reconstruction functions have been replaced by the proposed cached large binary-tree. Also, refactoring memory allocations and deallocations in fewer operations lead to a less fragmented profile for memory operations. As a result, the redesigned
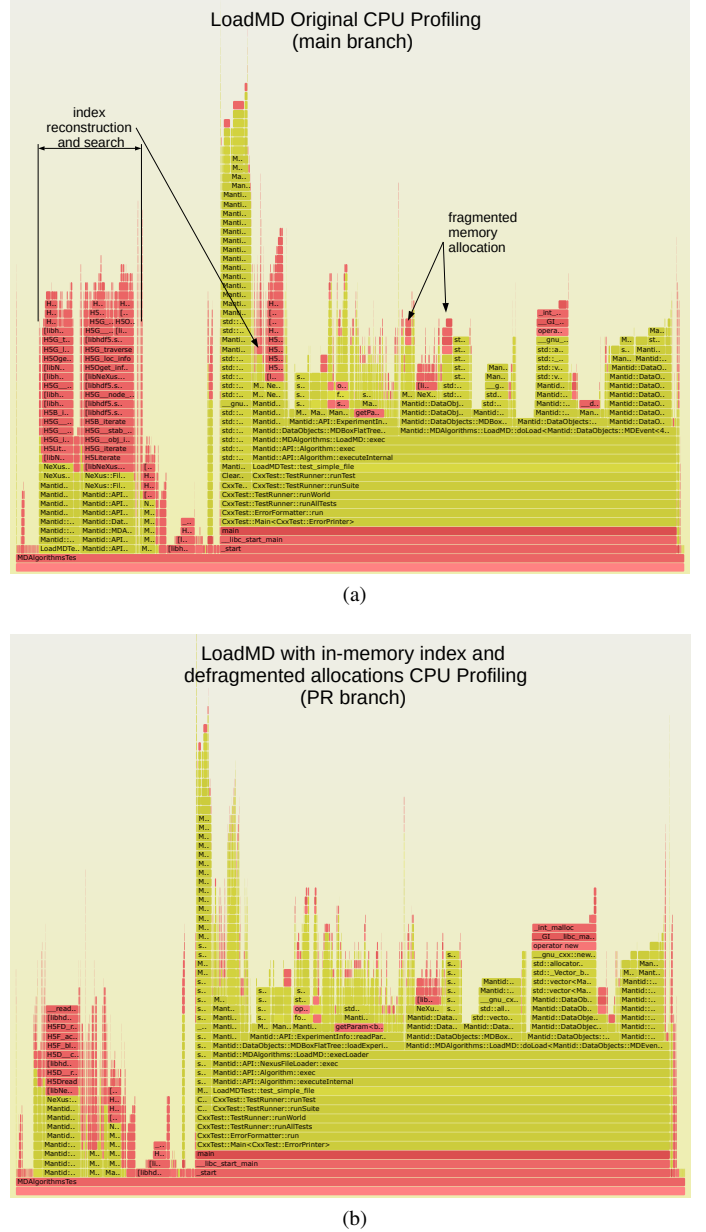

(a)

(b)

Fig. 4: Mantid's "LoadMD" CPU profiling flame graph representation for (a) Mantid v6.0 main branch, and (b) proposed PR branch[2] on Mantid. The reduction of metadata-related CPU operations bottlenecks and memory allocation fragmentation is illustrated.

"LoadMD" algorithm is expected to make fewer calls to the underlying HDF5 library binary tree search functions and the OS "malloc/free" functions for memory management. The expectation is that these algorithmic changes would impact the overall wall-clock times occupied by users of SNS/HFIR experimental facilities and computational resources[1].

IV. PERFORMANCE RESULTS AND IMPACT

This section presents the performance gains that result from our proposed implementation in Section III on SNS/HFIR

production computing systems[1] that are accessible to users of the neutron science facilities at ORNL [6].

After applying the algorithmic changes described in Section III, we compare the performance of "LoadMD" for the current Mantid "main" branch and our proposed implementation, accessible through a pull request (PR)[2], for a wide range of files with different number of experiments stored with "SaveMD". These files are described in Table III showing the typical ranges for the synthetic database resembling products from the Wide Angular-Range Chopper Spectrometer, ARCS instrument [15] at SNS. It can be seen that both, metadata entries and data sizes, scale up perfectly linearly since the number of stored parameters and data per experiment are fixed when selecting a particular instrument. Nevertheless, these numbers might change as we move to other instruments, which is currently out of the scope of the present work.

| Number of Experiments	Number of Entries	File Size (GB)
10	11,000	2.1
40	43,910	8.1
80	88,887	17.0
180	198,587	37.0

TABLE III: Description of the synthetic file ensembles based on the ARCS instrument [15] products used for the performance and scalability tests in this study.

To obtain a meaningful benchmark statistic, we ran 500 computational experiments for each case in Table III on a single node on SNS/HFIR computational systems[1]. Each node is powered by a 48-core Intel Xeon E5-2670 CPU and 512 GB of random-access memory (RAM). It is important to mention that system variability must be considered in our results due to the shared nature of the resources, thus rather than a deterministic analysis we obtain a distribution from our benchmarks tests capturing overall "LoadMD" wall-clock times. The obtained wall-clock histograms are presented in Fig. 5 for the largest cases in Table III, (40, 80 and 180 experiments). Overall, it can be seen that even when system variability is considered, the modifications introduced in this work result in consistent speed ups for any number of experiments. As expected, Fig 5 also shows that system variability has a noticeable influence on "LoadMD" wall-clock times as the stored number of experiments, thus operations, increase per target file. The system variability is more likely to be influenced by the state of the system when performing this calculations: CPU load, memory availability, file disk requests. Hence, we observe the effect that reducing the number of requests to lower-level memory and I/O operations also reduces the variability introduced by other users requesting system resources.

For completeness, we present the statistical box plots showing the consistent speed ups in terms of median values of the obtained wall-clock times on SNS/HFIR systems in Fig. 6. This includes improvements for even the smaller case for files containing 10 experiments. To further illustrate the overall

[2]https://github.com/mantidproject/mantid/pull/32529
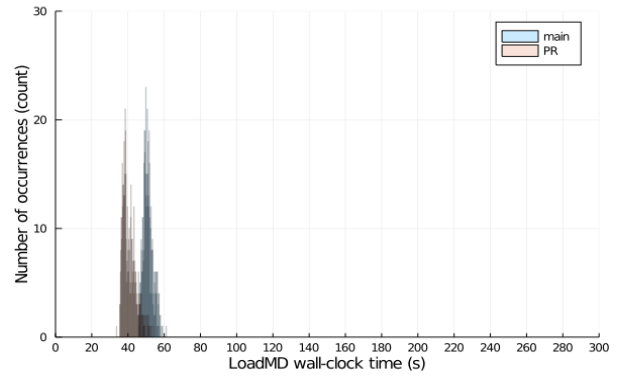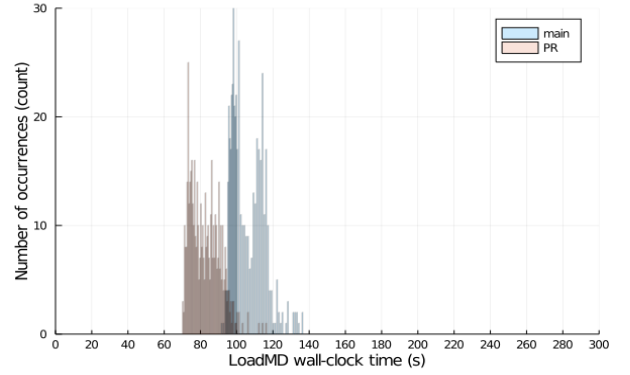
(a) 40 experiments

(b) 80 experiments
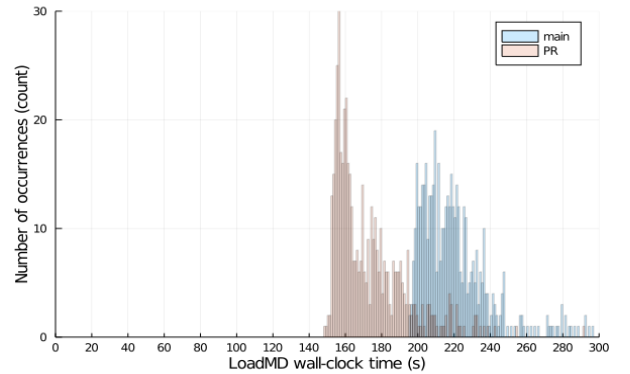
(c) 180 experiments

Fig. 5: Wall-clock times histograms for current Mantid "main" branch and the speed ups proposed changes from our pull request (PR), showing a 19% to 23% improvements on average for all cases on SNS/HFIR computing systems[1].

speed ups, we present a summary of the obtained wall-clocks time median, standard deviation, and speed ups from applying "LoadMD" on files up to 360 experiments in Table IV. It is interesting to note that while we confirm the speed ups in terms of the median, the standard deviation indicating the dispersion of the measurements might not correlate with the saving in wall-clock times. We infer that this might be due to system interference and more runs might be necessary for capturing different states of the system. Nevertheless, these are
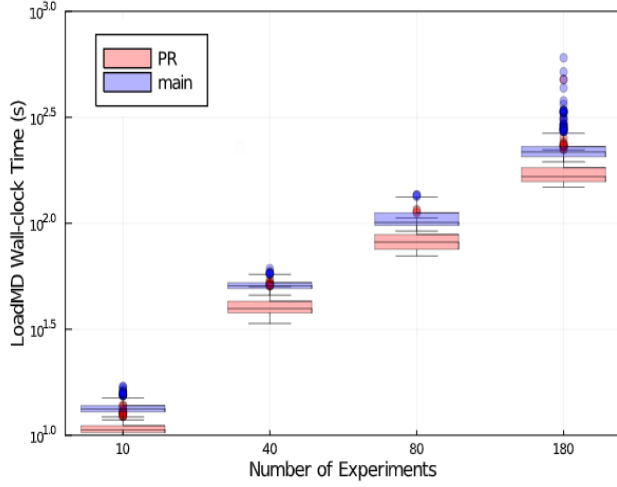
Fig. 6: Boxplots showing a consistent reduction in the overall wall-clock times for Mantid's "LoadMD" operation on a single node on SNS/HFIR computational resources[1] for a wide range of "MDWorkspace" files with a different number of experiments from SNS/HFIR instruments.

Number of	WC median (s)		WC stddev (s)		WC median
Experiments	main	PR	main	PR	speed up
10	13.3	10.6	0.7	0.6	20.3%
40	50.7	39.6	2.6	3.6	21.9%
80	101.0	81.5	8.7	8.1	19.2%
180	216.0	166.0	42.1	24.8	23.1%
360	357.0	285.0	59.4	32.9	20.3%

TABLE IV: Statistics comparing "LoadMD" wall-clock times from 500 experimental runs on a single node of SNS/HFIR computational resources[1] using Mantid's main branch and the current PR effort[2] for different numbers of experiments on a single file.

measurements worth considering in order to be confident that the present algorithmic improvements will have a very high probability to provide a consistent 19% to 23% speed up when compared to the current Mantid "LoadMD" implementation.

V. CONCLUSIONS

The present work reuses efficient strategies for addressing bottlenecks when processing reduced data ensembles stored using a single HDF5 file and the NeXus schema for a range of experimental setups. The latter is a typical use-case on time-of-flight instruments at ORNL neutron science facilities: SNS and HFIR. Bottlenecks due to index reconstruction and memory allocation and deallocation operations have been identified and addressed on Mantid's "LoadMD" function that loops through different experimental configurations. "LoadMD" serves as a backend for loading reduced data ensembles in several reduction workflows that provide facility users with a complete picture of the dynamic structure factor, $S(Q, E)$, that is typical in neutron scattering experiments. Two algorithmic changes on "LoadMD" are proposed: i) reusing a cached in-memory index for efficient entry search (15-20% savings), and ii) re-

ducing the number of memory allocations and deallocations in "LoadMD" (4-5% savings). The proposed algorithmic changes were tested across files stored at SNS/HFIR covering a typical range for the data ensemble size (10 to 360 experiments). Tests ran on a single node powered with a 48-core Intel Xeon E5-2670 CPU and 512 GB of RAM on the SNS/HFIR hosted computational resources available to facility users[1]. Results show consistent speed ups in the range of 19% to 23% for the median of the measured wall-clock times distributions expected from the shared nature of the resources. The proposed improvements have been recently integrated into the Mantid open-source framework, thus expecting further impact from our work in other neutron science facilities using "LoadMD" for retrieving data ensembles from multiple experiment configurations.

REFERENCES

[1] Oak Ridge National Laboratory, "Neutron Sciences." [Online]. Available: https://neutrons.ornl.gov/
[2] G. E. Granroth, K. An, H. L. Smith, P. Whitfield, J. C. Neuefeind, J. Lee, W. Zhou, V. N. Sedov, P. F. Peterson, A. Parizzi, H. Skorpenske, S. M. Hartman, A. Huq, and D. L. Abernathy, "Event-based processing of neutron scattering data at the Spallation Neutron Source," *Journal of Applied Crystallography*, vol. 51, no. 3, pp. 616–629, Jun 2018.
[3] P. F. Peterson, S. I. Campbell, M. A. Reuter, R. J. Taylor, and J. Zikovsky, "Event-based processing of neutron scattering data," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 803, pp. 24 – 28, 2015.
[4] M. Könnecke, F. A. Akeroyd, H. J. Bernstein, A. S. Brewster, S. I. Campbell, B. Clausen, S. Cottrell, J. U. Hoffmann, P. R. Jemian, D. Männicke, R. Osborn, P. F. Peterson, T. Richter, J. Suzuki, B. Watts, E. Wintersberger, and J. Wuttke, "The NeXus data format," *Journal of Applied Crystallography*, vol. 48, no. 1, pp. 301–305, Feb 2015.
[5] The HDF Group. (1997-NNNN) Hierarchical Data Format, version 5. [Online]. Available: http://www.hdfgroup.org/HDF5/
[6] S. I. Campbell, S. D. Miller, J.-C. Bilheux, M. A. Reuter, P. F. Peterson, J. A. Kohl, J. R. Trater, S. S. Vazhkudai, V. E. Lynch, and M. L. Green, "The SNS/HFIR web portal system for SANS," *Journal of Physics: Conference Series*, vol. 247, p. 012013, oct 2010.

[7] M. B. Stone, J. L. Niedziela, D. L. Abernathy, L. DeBeer-Schmitt, G. Ehlers, O. Garlea, G. E. Granroth, M. Graves-Brook, A. I. Kolesnikov, A. Podlesnyak, and B. Winn, "A comparison of four direct geometry time-of-flight spectrometers at the spallation neutron source," *Review of Scientific Instruments*, vol. 85, no. 4, p. 045113, 2014.

[8] R. Ewings, A. Buts, M. Le, J. van Duijn, I. Bustinduy, and T. Perring, "Horace: Software for the analysis of data from single crystal spectroscopy experiments at time-of-flight neutron instruments," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 834, pp. 132–142, 2016.

[9] O. Arnold, J. Bilheux, J. Borreguero, A. Buts, S. Campbell, L. Chapon, M. Doucet, N. Draper, R. F. Leal], M. Gigg, V. Lynch, A. Markvardsen, D. Mikkelson, R. Mikkelson, R. Miller, K. Palmen, P. Parker, G. Passos, T. Perring, P. Peterson, S. Ren, M. Reuter, A. Savici, J. Taylor, R. Taylor, R. Tolchenov, W. Zhou, and J. Zikovsky, "Mantid—data analysis and visualization package for neutron scattering and μ sr experiments," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 764, pp. 156 – 166, 2014.

[10] J. Thomason, "The ISIS Spallation Neutron and Muon Source—The first thirty-three years," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 917, pp. 61 – 67, 2019.

[11] P. Ageron, "Cold neutron sources at ILL," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 284, no. 1, pp. 197 – 199, 1989.

[12] W. Zhang, S. Byna, C. Niu, and Y. Chen, "Exploring metadata search essentials for scientific data management," in *2019 IEEE 26th International Conference on High Performance Computing, Data, and Analytics (HiPC)*, 2019, pp. 83–92.

[13] W. F. Godoy, P. F. Peterson, S. E. Hahn, and J. J. Billings, "Efficient Data Management in Neutron Scattering Data Reduction Workflows at ORNL," in *2020 IEEE International Conference on Big Data (Big Data)*, 2020, pp. 2674–2680.

[14] Klasky, Scott and Wolf, Matthew and Ainsworth, Mark and Atkins, Chuck and Choi, Jong and Eisenhauer, Greg and Geveci, Berk and Godoy, William and Kim, Mark and Kress, James and Kurc, Tahsin and Liu, Qing and Logan, Jeremy and Maccabe, Arthur B. and Mehta, Kshitij and Ostrouchov, George and Parashar, Manish and Podhorszki, Norbert and Pugmire, David and Suchyta, Eric and Wan, Lipeng and Wang, Ruonan, "A View from ORNL: Scientific Data Research Opportunities in the Big Data Age," in *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*, 2018, pp. 1357–1368.

[15] Abernathy,D. L. and Stone,M. B. and Loguillo,M. J. and Lucas,M. S. and Delaire,O. and Tang,X. and Lin,J. Y. Y. and Fultz,B. , "Design and operation of the wide angular-range chopper spectrometer ARCS at the Spallation Neutron Source," *Review of Scientific Instruments*, vol. 83, no. 1, p. 015114, 2012.

[16] W. F. Godoy, P. F. Peterson, S. E. Hahn, J. Hetrick, M. Doucet, and J. J. Billings, "Performance Improvements on SNS and HFIR Instrument Data Reduction Workflows Using Mantid," in *Driving Scientific and Engineering Discoveries Through the Convergence of HPC, Big Data and AI*, J. Nichols, B. Verastegui, A. B. Maccabe, O. Hernandez, S. Parete-Koon, and T. Ahearn, Eds. Cham: Springer International Publishing, 2020, pp. 175–186.

[17] B. Gregg, "The Flame Graph," *Commun. ACM*, vol. 59, no. 6, p. 48–57, May 2016. [Online]. Available: https://doi.org/10.1145/2909476