



# Nuclear Science User Facilities High Performance Computing

February 2023

*Enabling the Write-Once Read-Many (WORM)  
Dashboard and Web Interface for HPC Data  
Storage*

Matthew W. Anderson  
*HPC Manager*



*INL is a U.S. Department of Energy National Laboratory  
operated by Battelle Energy Alliance, LLC*

#### **DISCLAIMER**

This information was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness, of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. References herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

# **Nuclear Science User Facilities High Performance Computing**

## **Enabling the Write-Once Read-Many (WORM) Dashboard and Web Interface for HPC Data Storage**

**Matthew W. Anderson  
HPC Manager**

**February 2023**

**Idaho National Laboratory  
High Performance Computing  
Idaho Falls, Idaho 83415**

**<http://www.inl.gov>**

**Prepared for the  
U.S. Department of Energy  
Office of Nuclear Energy  
Under DOE Idaho Operations Office  
Contract DE-AC07-05ID14517**

*Page intentionally left blank*

## EXECUTIVE SUMMARY

Idaho National Laboratory (INL), which is supported by the Department of Energy Office of Nuclear Energy (DOE-NE) through the Nuclear Science User Facilities, provides the broader DOE-NE user community with access to supercomputer systems and data storage, along with support staff for system management, software installation, cybersecurity, and other types of support. A key element of this support is the provision of long-term or perpetual data storage as part of a larger effort to improve reproducibility in modeling and simulation activities, in addition to supporting simulations by helping meet regulatory compliance requirements. The present milestone details the deployment of a new type of storage option for users who require tamper-resistant perpetual data storage and the associated support for data control and release using the data storage web interface.

The perpetual data storage system—known as the write-once read-many (WORM) data storage system—is integrated within the Open OnDemand framework deployed at INL. The Open OnDemand science gateway is utilized by over 70% of users of the high-performance computing (HPC) system, and serves as the principal channel for accessing HPC resources. It also hosts the WORM dashboard for utilizing the perpetual data storage capabilities. Furthermore, it provides an interface to a specialized partition of storage that uses cryptographic hashes to ensure the data are not tampered with or compromised, with fully redundant storage in two physically separate locations.

The WORM dashboard provides visibility into datasets that were previously added to the WORM storage system. This is achieved independently, without requiring a command line interface for attaining access. This dashboard is compatible with most browsers. The WORM filesystem, WORM dashboard, cryptographic hash protection, and data storage redundancy capabilities – all added as part of this milestone – provide a new perpetual data storage capability that functions in tandem with the HPC systems supported by the Nuclear Science User Facilities and open up new capabilities for use by the broader DOE-NE user community.

*Page intentionally left blank*

## CONTENTS

EXECUTIVE SUMMARY .....	iii
ACRONYMS.....	vii
1. INTRODUCTION.....	1
2. WORM.....	1

## FIGURES

Figure 1. Open OnDemand science gateway.....	1
Figure 2. WORM data storage is accessed in the Open OnDemand science gateway by opening the Interactive Apps drop-down menu and selecting the Research Data Management option. ....	2
Figure 3. Form for submitting data to the WORM.....	2
Figure 4. Valid group options that appear whenever a data uploader requests that access to the data be restricted to a specific group.....	3
Figure 5. WORM dashboard.....	3

*Page intentionally left blank*



## ACRONYMS

DOE-NE	Department of Energy Office of Nuclear Energy
HPC	High Performance Computing
INL	Idaho National Laboratory
WORM	Write-Once Read-Many

*Page intentionally left blank*

## 1. INTRODUCTION

Idaho National Laboratory (INL), which is supported by the Department of Energy Office of Nuclear Energy (DOE-NE) through the Nuclear Science User Facilities, provides the broader DOE-NE user community with access to supercomputer systems and data storage, along with support staff for system management, software installation, cybersecurity, and other types of support. A key element of this support is the provision of long-term or perpetual data storage as part of a larger effort to improve reproducibility in modeling and simulation activities, in addition to supporting simulations by helping meet regulatory compliance requirements. The present milestone details the deployment of a new type of storage option for users who require tamper-resistant perpetual data storage and the associated support for data control and release using the data storage web interface.

## 2. WORM

The perpetual data storage system—known as the write-once read-many (WORM) data storage system—is integrated within the Open OnDemand framework deployed at INL. For accessing the high-performance computing (HPC) systems, the Open OnDemand science gateway (illustrated in Figure 1) functions as the primary channel for over 70% of the userbase. When using the Open OnDemand science gateway, access to the WORM data storage is attained by opening the Interactive Apps drop-down menu and selecting the Research Data Management option (see Figure 2).

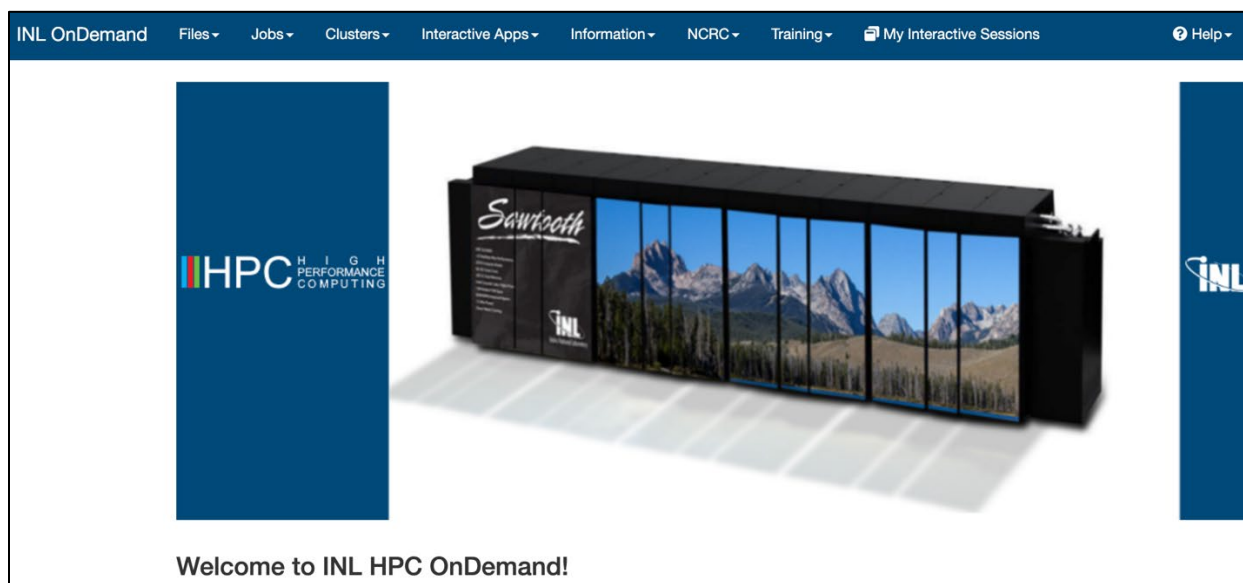


Figure 1. Open OnDemand science gateway.

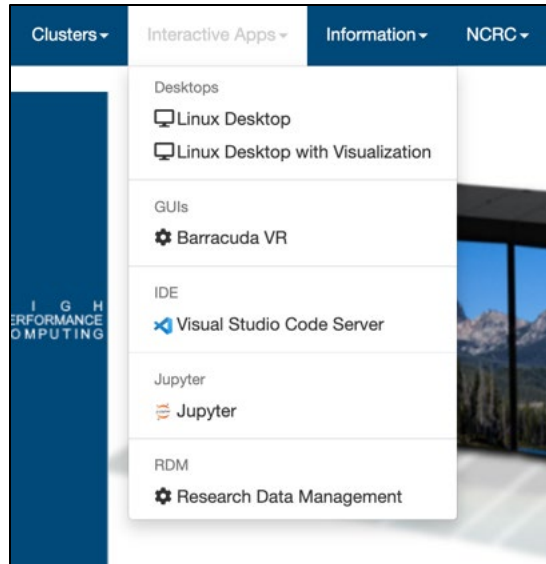


Figure 2. WORM data storage is accessed in the Open OnDemand science gateway by opening the Interactive Apps drop-down menu and selecting the Research Data Management option.

There are two parts to the WORM interface. The first part is a form, shown in Figure 3, for adding data to the WORM. By either typing in the path or selecting the file(s) via a file browser, the user can specify an entire directory or set of files to be added to the WORM. A brief description of the data is requested but not required. The “Requestor” field reflects the name of the user serving as the principal point of contact for the dataset. This may or may not be the same user who is uploading the data to the system. Finally, the data uploader must indicate whether these data can be made world readable or should instead be limited to a certain group. If the data are set to be restricted to a particular group, valid group options will automatically appear, as shown in Figure 4.

 A screenshot of the 'Research Data Management' form, version 546ad8e. The form is titled 'Copy research data to a Write-Once Read-Many (WORM) location.' and contains several input fields and a 'Launch' button. The 'Directory' field contains the path '/home/andematt/nsuf\_185396'. Below this field is a 'Select File' button. The 'Requestor' field contains the name 'Matt Anderson'. The 'Description' field contains the text 'NSUF test dataset'. The 'Location' section has two radio buttons: 'World Readable Location Only' (which is selected) and 'Group Readable Location Only'. At the bottom of the form is a large blue 'Launch' button. A footnote at the very bottom states: '\* The Research Data Management session data for this session can be accessed under the data root directory.'

Figure 3. Form for submitting data to the WORM.

**Description**

NSUF test dataset

Details describing the data and reason for backing up. Single- and double-quotes are not allowed.

**Location**

World-Readable Location Options

- activity\_detection
- advantg3
- air-dev
- ale3d\_llnl
- alegra
- ✓ andematt
- arc11
- arc112
- bison-conda
- bison-dev
- bison-exec
- bluecrab-conda
- bluecrab-dev
- bluecrab-exec
- carbondioxide-dev
- covid19
- cth
- cubit
- cubit\_src

Figure 4. Valid group options that appear whenever a data uploader requests that access to the data be restricted to a specific group.

When the data uploader has completed the form and then selects “Launch,” a background process begins copying the data. Copies of larger data files can require several hours; the user can monitor the progress by referring to the session logs.

Data in the WORM can be accessed in three ways: via the command line interface, file browser, or WORM dashboard. For both the command line interface and file browser, the WORM data are located in the /rdm partition on the HPC systems. World-readable data are separated from group-readable data by different head directories. Figure 5 shows the WORM dashboard, which lists all the projects in WORM that are accessible to the user who logged into Open OnDemand science gateway. The dashboard also shows who uploaded the original dataset. This dashboard provides an alternative to using the command line interface and is compatible with most browsers.

Home /

**WRITE-ONCE READ-MANY FILES**

*Files may not appear immediately depending on the size of the folder that was uploaded*

[+Upload New File](#)

Show  entries

Search:

Name	Last Modified	Size	Location	Owner
A617_TESTG-7	2023-02-21T14:58:14.364	188 bytes	world	copetm
nsuf	2022-07-13T08:37:37.815	75 bytes	world	andematt
project1	2021-12-14T19:25:21.096	79 bytes	world	evangj
nids	2022-08-23T13:32:23.931	237 bytes	world	sgammir
software_users_report.ipynb	2022-04-01T15:01:29.654	17404938 bytes	world	rothbh
milestone.txt	2022-08-29T13:26:27.052	38 bytes	group	test_website

Showing 1 to 7 of 7 entries

[Previous](#) [Next](#)

Figure 5. WORM dashboard.

In each submitted project, a sha256sum hash is included with all submitted files and directories. This cryptographic hash provides users complete assurance that the data have not been modified or tampered

with since being uploaded to the WORM data storage system. Running the command `sha256sum -c <hashname>` generates a list of each file in the current directory, and whether the cryptographic hash for the file indicates that the file has not been tampered with or changed. Any user can run this command at any time.

The data storage hardware consists of four Dell EMC A300 nodes, each featuring 15x12 TB drives and connected via two S5232F network switches. Data redundancy within the configuration keeps the overall storage capacity at around 500 TB. The data storage capability was integrated into the data center and was first made available to users on October 20, 2022.