

**Report Title:** ERCAP Requests Details  
**Run Date and Time:** 2023-09-29 11:18:24 Pacific Daylight Time  
**Run by:** PDF Generator User  
**Table name:** u\_ercap\_requests

**ERCAP Requests**

This request is a renewal.:	true		
ERCAP Number of Request to Renew:	ERCAP0023620		
ERCAP Number:	ERCAP0027135	Allocation Year:	2024
Project Title:	Analysis and Simulation for the GlueX Detector	State:	Draft
Label:	GlueX	Revisions required:	
PI Name:	Austregesilo, Alexander (aust)	Rejection Reason:	
PI Name Company:	Jefferson Lab	Project Class:	DOE Mission Science
PI Name Email:	aaustreg@jlab.org	Program:	NP - Nuclear Physics
PI Name Business phone:	7572696982	Sub Program:	
		Science Category:	Physics : Nuclear Physics (Experimental)
		Project:	m3120

**Personnel**

**Senior Investigators:**  
 David Lawrence  
 Alexander Austregesilo  
 Igal Jaegle

**Authorized Preparers:**  
 Lawrence, David (davidl), Larrieu, Christopher (larrieu), Brei, Nathan (nbrei), Jaegle, Igal (ijaegle), Austregesilo, Alexander (aust)

**Funding**

**Primary Funding Source:**  
 DOE Office of Science (SC)  
 DOE Office of Science (SC):  
 true

**Funding Office:**  
 DOE Office of Science Nuclear Physics (NP)

**Funding Program Manager:**  
 Rai, Gulshan (Gulshan.Rai@science.doe.gov)

**DOE/SC Grant, PAMS, FWP numbers:**  
 Contract number DE-AC05-06OR23177

**Federal Agency other than DOE/SC:**  
 false

**Other Funding/Agency Source(s):**

**Other Federal Agency grant numbers:**

LDRD Funding:

false

Funding Laboratory:

LDRD grant numbers:

State or local government or agency:

false

State/Local Govt/Agency name and grant numbers:

Foreign Government or Agency:

false

Foreign Government/Agency name and grant number:

University:

false

University name and grant numbers:

Non-profit Organization:

false

Non-profit Organization and grant number:

Other:

false

Other organizations and grant numbers:

Office of Science relevance:

**Security**

I attest that this project adheres to these guidelines. :  true

I request an exception to these policies, based on the following::  false

Please Explain Policy Exception:

**Project Details**

Project Summary and Goals:

The primary project goal will be to analyze data taken with the GlueX detector from a set of ongoing nuclear physics experiments at the Jefferson Lab accelerator. The analysis will consist of extracting timing and energy deposition information from the data in order to reconstruct individual particle interaction events. The analysis will produce the momentum, direction, and type (particle ID) of each particle detected in each reaction. The statistics of these reaction particles can then be used to measure fundamental physical properties of the excited states of the target + photon system, leading to an understanding of the underlying particles (quarks and gluons) and the forces among them.

Detailed Description for DOE Managers:

We plan to perform the first stage reconstruction of the data from the GlueX experiment at NERSC. This will require transferring the raw experimental data from JLab to NERSC, producing the "data summary tapes" (DST) files, and transporting them back to JLab for further analysis. Our current plan is to try and focus larger scale full passes over the data at NERSC allowing us to use our time on the local JLab cluster for the smaller campaigns that require quick turn around. The NERSC jobs will do the most CPU intensive part of the analysis. Specifically, charged particle tracking, calorimeter cluster finding, and matching reconstructed objects between detectors.

Website URL:

[https://halldweb.jlab.org/wiki/index.php/GlueX\\_Project\\_Overviews](https://halldweb.jlab.org/wiki/index.php/GlueX_Project_Overviews)

Accomplishments Summary:

The GlueX experiment at the Thomas Jefferson National Accelerator Facility is part of a global effort to study the properties of the strong interaction, which binds the fundamental quarks into subatomic particles like protons and neutrons. In 2018, the experiment completed its initial phase, recording a worldleading data set of more than 250 billion events which corresponds to about 5 petabytes of raw data. Given the limited local resources, the collaboration depends on HPC farms like NERSC to condense this massive data sample into physical quantities like particle trajectories and electromagnetic showers. About two thirds of this initial data set were processed at NERSC, and the elaborate analysis is presently ongoing. We expect several high-impact publications to emerge from this program.

In parallel, the experimental setup was upgraded for the second phase of GlueX. From Spring 2020 onward, GlueX is taking data with a considerably higher rate, opening the door for the exploration of rare processes. Within the first few months, we have essentially doubled the data set and have already started to employ NERSC resources for its reconstruction campaign. We invested manpower to optimize the usage and are planning to continue this successful collaboration for the next few years.

In 2023, numerous presentations at conferences and workshops were based on data that was at least partially processed at NERSC. We acknowledge the successful collaboration with NERSC on our website (<http://gluex.org/thanks/>) and in every presentation. For a complete list of talks given in 2023, please see here:

[https://halldweb.jlab.org/wiki/index.php/Gluex\\_Talks#Talks\\_in\\_2023](https://halldweb.jlab.org/wiki/index.php/Gluex_Talks#Talks_in_2023)

Refereed Publications:

S. Adhikari et al. (The GlueX Collaboration) Measurement of the  $J/\psi$  photoproduction cross section over the full near-threshold kinematic region, Phys.Rev.C 108 (2023) 2, 025201

<https://doi.org/10.1103/PhysRevC.108.025201>

Several Nuclear Physics publications are currently under internal collaboration review that use data processed at NERSC. They will be published in the course of the next 12 months.

Non-refereed materials:

Resources

CPU Node Hours Used:	5,088	CPU Node Hours Requested:	350,000
GPU Node Hours Used:	0	GPU Node Hours Requested:	0

GPU Readiness:

What is the typical number of nodes your individual jobs will use concurrently?:

What is the maximum number of nodes your individual jobs could use concurrently?:

Justification for Compute Request:

Values are calculated based on the GlueX Computing model. This model is implemented in a python script using input from XML files representing different data sets. Both can be found here:

[https://github.com/JeffersonLab/hd\\_utilities/tree/master/comp\\_mod](https://github.com/JeffersonLab/hd_utilities/tree/master/comp_mod)

The xml files contain extensive comments detailing how numbers that are input to the model were determined. For the coming year we plan to process the data for the Fall 2022 run period (RunPeriod-2022-08.xml) and the Spring 2023 run period (RunPeriod-2023-01.xml) which approximately add up to the 350kh request.

During 2023, the data production was significantly delayed while waiting for the calibration of the Spring 2022 run period (RunPeriod-2022-05.xml). While constituting a comparably small sample, these data were taken with an unusual configuration and revealed shortcomings in the calibration workflow. The discovered issues are now largely solved and we will soon resume the reconstruction campaign. We expect to use 100% of our remaining allocation (40,000 node hours) by the end of 2023 to process the Spring 2022 run period. No such delays are expected for the data sets provisioned for 2024, since the calibration was done in parallel and is nearly complete already. In addition, we used the time to benchmark the different datasets on the Perlmutter system in order to accurately predict the node hours required for their processing.

Archival Storage Used (TB):	0	Archival Storage Requested (TB):	1
CFS Storage Used (TB):	0	CFS Storage Requested (TB):	20

Mostly large files:  true

Mostly large numbers of smaller files:  false

Increase in CFS files?:

No

Justification for Storage Request:

I plan to use resources consistently throughout the allocation year:  true

My usage will vary throughout the allocation year:  false

1st Quarter:	<input type="checkbox"/> 25 %	3rd Quarter:	<input type="checkbox"/> 25 %
2nd Quarter:	<input type="checkbox"/> 25 %	4th Quarter:	<input type="checkbox"/> 25 %

Key Events or Deadlines:

Need real-time computing?:

false

Explanation for Realtime Computing Needs:

Experimental or Observational project?:

true

Special Requirements:

**Codes**

Software Dependencies:

How many codes will you use? (select up to 5):

2

Code 1 Name:

JANA

Code 1 URL:

<https://www.jlab.org/JANA/>

Code 1 Description:

GlueX data analysis

Code 1 is GPU Enabled?:

false

Code 2 Name:

Hall-D Reconstruction Code

Code 2 URL:

[https://github.com/JeffersonLab/hall\\_d\\_recon](https://github.com/JeffersonLab/hall_d_recon)

Code 2 Description:

Reconstruction of GlueX data

Code 2 is GPU Enabled?:

false

Code 3 Name:

Code 3 URL:

Code 3 Description:

Code 3 is GPU Enabled?:

false

Code 4 Name:

Code 4 URL:

Code 4 Description:

Code 4 is GPU Enabled?:  
false

Code 5 Name:

Code 5 URL:

Code 5 Description:

Code 5 is GPU Enabled?:  
false

**Supporting Information**

Other HPC Support:  
Some portion of the data will also be processed on the JLab Scientific computing farm, where our fairshare amounts to roughly 60M thread-hours per year. This is divided into different production tasks (data calibration, monitoring, reconstruction) and individual user analysis. We are efficiently using the Open Science Grid for Monte Carlos simulation, since it does not require the transfer and intermediate storage of large raw data file. We are also exploring ways to use this additional resource for data recontruction.

Additional Information:

Feedback:

**Usage Agreement**

Usage Agreement Initials:  
AA

**Award Information**

Approval State:  
Not Yet Requested

CPU Node Hours Requested:	350,000	CPU Node Hours Awarded:	
GPU Node Hours Requested:	0	GPU Node Hours Awarded:	
Archival Storage Requested:	1	Archival Storage Awarded (TB):	1
CFS Storage Requested:	20	CFS Storage Awarded (TB):	20

DOE PM Notes:

Approver:

Award Status:  
Draft

**Record History**

Computational Repo:	m3120	Project ID:	61031
Computational Current Allocation:	45,044	Archival Current Quota (TB):	1
Computational Repo ID:	61032	Archival Repo ID:	61033
GPU Account:	m3120_g	HPSS Only:	false
GPU Current Allocation:	38	Current Project Storage Quota (TB):	535
GPU Account ID:	67592	CFS Max Projdirs:	10
Renewed by ERCAP Request:		CFS Max Files:	20,000,000
AY Year Start:	January 17, 2024	Sponsoring Organization:	Jefferson Lab

---

AY Year End: January 15, 2025

Does PI work at a federal agency or national lab?: true

Funding Sources:

**Related List Title:** Attachment List

**Table name:** sys\_attachment

**Query Condition:** Table name = u\_ercap\_requests AND Table sys ID = 86b742701b6db91059b0dc6ce54bcb7a

**Sort Order:** Created in descending order

None