

Massachusetts Green High Performance Computing Center
100 Bigelow Street, Holyoke, MA



Request for Information AI Compute Resource Operations and Engineering

March 17, 2025



Contents

- 1 Introduction..... 3**

- 2 The Operations and Engineering Stack..... 3**
 - A. System Software..... 4
 - B. Operations..... 5
 - C. Service Delivery..... 5
 - D. Services and Applications..... 5

- 3 Responses..... 6**
 - A. General Guidance..... 6
 - B. AICR Project Year 1 Phasing..... 7
 - C. Response Elements..... 8
 - D. Response Evaluation..... 9
 - E. Submission Details..... 9

- Appendix A - Example Stack Components..... 10**

1 Introduction

This Operations and Engineering RFI is a companion to the AI Compute Resource (AICR) Infrastructure System RFP. The AICR project is developing a solution to support the Massachusetts AI Hub. The Massachusetts AI Hub focuses on accelerating AI adoption, innovation, and economic development in Massachusetts. More information about the Massachusetts AI Hub and the AICR project can be found in the AICR Infrastructure System RFP document.

The solution described by the companion AICR Infrastructure System RFP provides a computing foundation but requires additional software, applications, and personnel to be useful to researchers and developers. This Operations and Engineering RFI seeks proposals from qualified vendors to provide these additional software, applications, and personnel.

We plan to operate the AICR platform using a combination of full-time AICR staff augmented by services proposed in the response to this RFI. AICR plans to hire a full-time Executive Director as well as select senior systems, operations, and service delivery staff. The initial mix of third-party and in-house personnel and solutions is not currently defined and will be influenced by the quality and depth of responses to this RFI. The long-term composition will evolve as the cluster evolves.

Respondents should propose solutions that address the Operations and Engineering Stack components identified in figure 1 below. Proposals must address the *Systems Software* and *Operations* portions of that diagram. Respondents with experience delivering solutions corresponding to the *Service Delivery* and the *Service and Applications* components of the stack are encouraged to propose solutions for those as well.

Note, the lowest two levels in figure 1, *Cluster Hardware* and *Datacenter Infrastructure*, are shown for illustration purposes only and are not part of this RFI.

2 The Operations and Engineering Stack

Figure 1 represents the full AICR solution. The bottom two infrastructure layers consist of the physical datacenter and hardware resources. The components running above these layers work in conjunction with one another and implement the Operations and Engineering stack.

Operations and Engineering Stack components are segmented as follows:

- System Software
- Operations

- Service Delivery
- Services and Applications

Each of these components is shown in Figure 1; within those components, some examples are called out. These examples are not exhaustive nor are they meant to be prescriptive (with the exception of the applications specifically identified in the "Services and Applications" section). For more examples please see Appendix A.

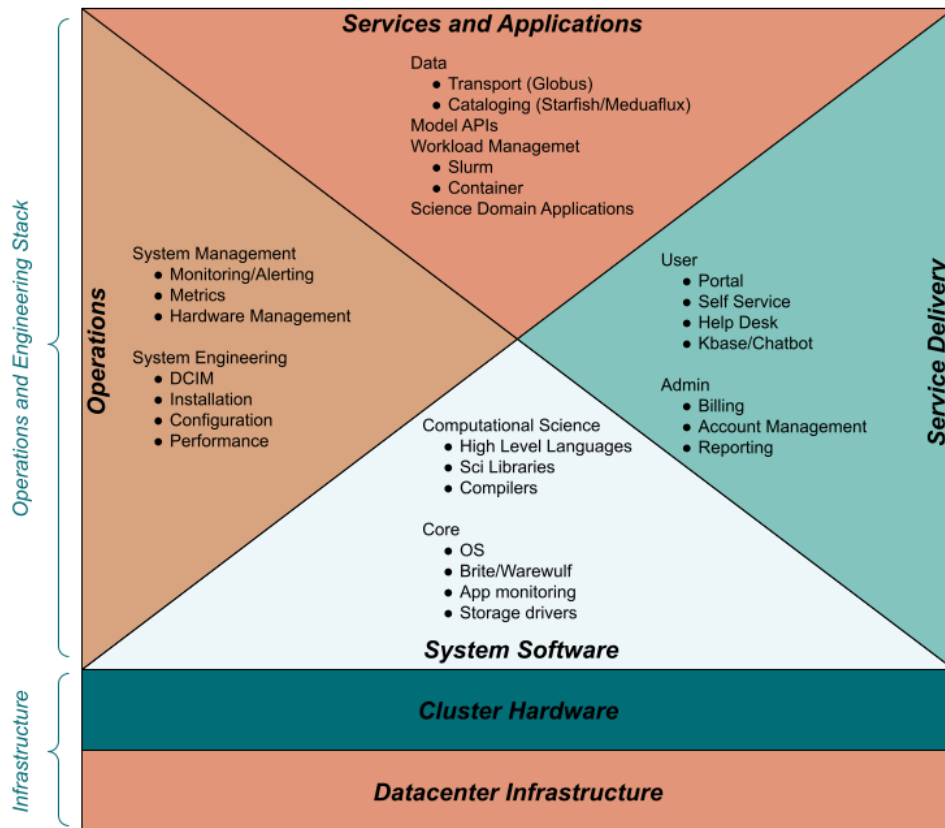


Figure 1: Operations and Engineering Stack

A. System Software

System software turns networking, storage, and computing hardware devices into AI research tools.

Starting at the lowest level, firmware, drivers and operating systems turn CPUs and GPUs into nodes; above that, provisioning and workload management software turn collections of nodes into clusters; computational science tools like compilers, numerical libraries, and high level languages turn clusters into AI research tools.

Respondents are expected to describe how they would support the AICR system software component in detail. Responses should include system software specifics (e.g., software used for provisioning, directory, config management) and proposed personnel/skillsets.

B. Operations

An AI cluster is a complex system composed of thousands of hardware and software elements. Cluster operations manage all of these to guarantee performance and availability. Operations is responsible for both day-to-day activities required to keep the cluster running as well as lifecycle activities such as inventory tracking, spares management and part replacement. The respondent should describe how they would support AICR Operations, including the tools, processes, and personnel that they would apply. Please provide customer examples.

C. Service Delivery

Similar to how system software turns hardware into researcher tools, service delivery turns those tools into services. Service delivery "wraps" lower level capabilities (storage, computing, connectivity) with software, documentation and personnel so that they can be accessed by and shared fairly among users. Systems supporting service delivery include account creation and management, billing, customer self-service, help desk, kbase/chatbots, and reporting. These systems work together to guarantee predictable and performant access to the AICR. This service wrapper also includes people who help researchers access and use the platform and its applications.

Respondents should describe how they would support service delivery as well as what software they would use. It is anticipated that some of the components (e.g. websites, self-service portal, allocation management) of the solution may not have "off-the-shelf" options and that respondents may need to propose custom solutions or bespoke development to address some features. Additionally, we expect solutions to contain some complement of customer support personnel.

It is currently envisioned that domain-specific, scientific facilitation will be handled *outside* of this RFI. *Please do not include services for end-user science facilitation in your proposal.*

D. Services and Applications

AICR will provide some pre-installed applications for users. We are currently planning on deploying the following specific applications.

- SLURM
- Globus
- Open OnDemand (OOD)

In addition to these specific applications, AICR will likely deploy the following types of applications:

- Data-related
 - Catalog
 - Management
 - Federation
- Science domain applications
- Specific model APIs

Respondents should discuss which of our planned applications (i.e. SLURM, Globus and OOD) they can provide and how they would support day-to-day management of those applications. Please indicate whether you have supported these applications for other customers.

Also, if appropriate, please describe which of the proposed classes of applications (i.e. Data, Science Domain and Models) you can support, how you would support them and examples of customers that you currently support.

3 Responses

A. General Guidance

AICR is soliciting proposals for computing via the AICR Infrastructure System RFP and proposals for operations via the Operations and Engineering RFI. Both these solicitations are being released at the same time so that parties who plan to respond to both can create a streamlined proposal. A separate Operations and Engineering RFI also allows specialized AI/HPC professional services companies to propose service-only solutions that align with the AICR Infrastructure System RFP requirements.

All parties are expected to describe how they have supported the type of system described in the AICR Infrastructure System RFP in prior engagements. If the respondent is also submitting a proposal for the AICR Infrastructure System RFP, explain how the specific system in that submittal will be supported.

The AICR and the Massachusetts AI Hub that it supports are startup activities. We expect the AICR solution to evolve significantly during the first year of operation as we gain experience delivering service. *We will be looking for proposals that anticipate this evolution and remain*

robust as project needs change. Stock/boilerplate solutions are unlikely to match project needs in its initial 12 month period.

B. AICR Project Year 1 Phasing

The engagement is expected to proceed in three phases covering the periods given below; the phases will build sequentially on each other.

1. 0-3 months Initial Phase:
 - a. During this phase the AICR facility expects to operate a system with the following characteristics:
 - Several hundred compute nodes
 - Up to 1000 AI/ML accelerators
 - Up to 1000 network ports on 2-3 networks
 - 500-1000 total accounts / 100-200 active accounts
 - b. This system will have been provisioned with a base software stack. At least one external gateway to other systems within the MGPHCC and to external Internet resources will be fully functional.
 - c. Operating practices and processes will likely be evolved during this period, concurrent with testing by soft launch customer accounts.

2. 3-6 months Stable-v1 Phase:
 - a. During this phase it is expected that initial processes and practices will begin to transition to steady state operations.
 - b. Customer access will ramp up during this phase with associated service level expectations including:
 - robust uptime
 - proactive monitoring of hardware
 - proactive monitoring of runtime environment issues including performance and security
 - continuous monitoring and response, and user notification
 - c. Maturing usage monitoring and usage profile tracking, detailed reporting and feedback for efficiency and cost allocation.
 - d. No new system hardware will be added during this phase.
 - e. Some additional external network connectivity may be introduced during this phase.

3. 6-12 months Preliminary Enhancement Phase:
 - a. In this phase, the initial system will be in a production mode with all services adequately monitored and supported.
 - b. It is expected that during this phase a small secure enclave component will be introduced to the system through a hardware addition. The scale of the hardware

addition will likely be relatively small, but software and operating practices for fully regulated data (likely designed to meet NIST800-171/NIST800-53 levels) will begin to be developed.

C. Response Elements

Responses are requested for an engagement starting approximately August 1st, 2025 and running up to 12 months. For each of the phases previously described, respondents should provide estimates and material that cover the following:

1. Services Proposed

- What services the respondent can support during that phase.

2. Staffing

- Staffing estimates to support those services.
 - Staffing should clearly indicate what FTE resources would comprise dedicated staff solely focused on this project and what resources would be partially committed staff.
 - Skill levels/grade of the different resources
 - For dedicated staff, a monthly rate range for different skill/grade levels should be provided.
 - For partially committed staff, the basis for charging those staff and the range of cost rates by relevant skill level should be provided
 - Staffing should clearly indicate which support would be handled remotely and which requires on-site presence or coordination.
- A summary table showing month by month per FTE estimated costs, and total estimated costs during each phase. This table should include an estimated minimum FTE commitment and an optimal level of FTE commitment.
- A summary table mapping FTE estimates to effort on different areas of the Figure 1 Operations and Engineering Stack.

3. Engagement Model

- Respondents should also describe their process(es) for working with the MGHPCC AICR staff to coordinate work, assign tasks, and track progress day to day, within a week, weekly, and monthly.
- Processes for emergency/on-call coverage to address urgent operational needs and any associated limits/extra charges on those services should be described.
- Practices and processes for documenting operations material and playbooks, and for professional software engineering and DevOps procedures to ensure repeatable, versioned, and well documented test and production solutions.
- The respondent should outline their approach to periodic overall service review and ongoing tracking of work items in a manner that the MGHPCC staff have full visibility and

adequate control of work items and scheduling.

4. Customer References

- Examples and project scope summaries from at least 3 prior engagements that most closely align with this project that the responding organization has successfully undertaken. This document should include a section describing approaches to coordinating and scheduling contract work with client operational and engineering teams.

D. Response Evaluation

The evaluation process is expected to involve review of the response for

- cost competitiveness
- respondent's demonstrated credibility in terms of quality of answers to response elements labeled A through I as described above

For respondents that are cost competitive and demonstrate credibility, we expect to follow up with a detailed discussion when the overall system architecture is finalized in May, 2025.

E. Submission Details

RFI Responses:

- Responses are due by April 17th 2025
- Responses should be submitted by email as an attached zip file or a zip file download link to mghpcc-ioc-aicr-rfp@mghpcc.org

RFI Questions:

- Questions concerning the RFI may be submitted to mghpcc-ioc-inquiries@mghpcc.org
 - Responses to questions of general interest will be posted to a public FAQ page on the <https://mghpcc.org> web site.

Appendix A - Example Stack Components

The contents of the "Examples" column are meant to be exemplar and are not intended to be prescriptive with the exception of the applications specifically identified in the "Services and Applications" section.

OM Stack	Component	Examples
System Software	Provisioning	Bright, Warewulf
System Software	Internal	DNS, DHCP, PXE, NTP, SSHD
System Software	OS	Rocky, RHEL, Ubuntu
System Software	Drivers	NVIDIA, CUDA, Lustre
System Software	Compilers	C, C++, Fortran, GO, Rust
System Software	High Level Lang	R, python, julia, matlab, mathematica, octave
System Software	Libraries	seq/parallel libs, MPI
System Software	Comp. Virtualization	OpenStack, OpenShift, Kubernetes
System Software	Container	Aptainer/Singularity, CharlieCloud, fakeroot
Operations	Config Management	Ansible, Tower, SaltStack, Puppet, Foreman, Chef, xcat
Operations	Observability	Prometheus, Telegraf, Influx, Grafana, ossec, aide
Operations	Forensics	syslog, ELK stack, Splunk
Operations	Alerting	Nagios, Zabbix, wazuh
Operations	DCIM	Netbox
Operations	Remote Hands	FRU replacement, rack/stack
Operations	Automation	Bash, python, cron
Operations	Tech Support	ZenDesk, RT, FreshDesk, Service Now, osTicket/SupportSystem
Service Delivery	Documentation	
Service Delivery	Kbase	
Service Delivery	Help Desk	ZenDesk, RT, FreshDesk, Service Now, osTicket/SupportSystem

Service Delivery	Account Management	Coldfront
Service Delivery	Self Service Portal	Bespoke (Django, etc.) Coldfront, Keycloak
Service Delivery	Billing	Xdmod, bespoke
Service Delivery	Reporting	
Service Delivery	Web Site	Wordpress
Applications	Workload Mgmt	SLURM, Kubernetes
Applications	Access	Open OnDemand, SSH, Cloudflare, Pomerium
Applications	File Transfer	Globus
Applications	Environments	Environment Modules, conda, virtualenv
Applications	Frameworks	Pytorch, tensorflow
Applications	Data Federation	Starfish, Mediaflux