Technical Report:

CC* Compute: Accelerating Compute Driven Science Through a Sharable High Performance Computing Cluster in Kent State Multi-Campus System: Kent State ICE Sci.DMZ Onboarding Process

Prepared by:

Dr. Javed I. Khan
& Phil Thomas

Media Communications and Networking Research Lab
Department of Computer Science
Kent State University, Kent OH 44242
javed@kent.edu

Technical Report 2024-03-01
Internetworking and Media Communications Research Laboratories
Department of Computer Science, Kent State University
http://medianet.kent.edu/technicalreports.html

CC* Compute: Accelerating Compute Driven Science Through a Sharable High Performance Computing Cluster in Kent State Multi-Campus System: Kent State ICE Sci.DMZ Onboarding Process

Javed I. Khan[1] and Philip Thomas[2]
e-mail: Javed@ kent.edu | plthomas@ kent.edu

[1] Internetworking and Media Communications Research Laboratories
Department of Computer Science

[2] Information Systems Division
Kent State University
233 MSB, Kent, OH 44242
December 2023

1. Introduction

The funding for the ICE project was provided by NSF CCI* program. The program was created to support the NSF's vision for a holistic cyberinfrastructure (CI) ecosystem outlined in the "Transforming Science Through Cyberinfrastructure: NSF's Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century" [1]. This particular program's goal is to help campuses drive toward a 21st century realization of an integrated CI for enabling distributed science. That can address science-driven needs in data networking both intra-campus and externally. It also demanded innovation in institutional practices so "*multiple under-resourced campuses can have access to the integrated CI through partnerships with regional entities and small institutions with experience in high-performance Research & Education (R&E) networking*" [2]. The NSF funded investment also seeks innovation in CI to meet its boarder impacts goals. This document summarized the ICE (sci.DMZ) onboarding process.

2. Mission of ICE

The technical goal of the CC* project is to establish a sharable computing cluster with 600+ core compute node and about a petabyte of storage within the science DMZ for Kent State University- the Integrated Computing Cluster (ICE). The distinction of the ICE project is that it will not be a standalone HPCC facility as existed as island in many institutions. In contrast, it will also be a gateway to a locally and remotely sharable computing infrastructure. Here are some goals:

1. Once established, at least 20% of the NSF funded bare-metal assets resources to be provided for federated computing-resource sharing through a new generation of emerging platform technologies such as Open Science Grid (OSG).

2. Symmetrically, the facility will also act as an access point for other federated resources while creating greater support for responsive workflows locally. The local researchers whose workflows are suitable for federated remote resources will be able to seamlessly use other remote resources.

3. On the other hand, the specific science workflows, which are less suitable for remote processing, such as scienceware as-a-service (SAS), science-data-lakes (SDL), and intense real-time computing (RTC), will be accommodated with greater priority on the local resource. Normally these are applications which require real-time response or require tight geo-cohesion of different parts of its workflows.

4. Further the resources can be used by computing researchers from all areas, from all campuses including regional campuses, and other faculty researchers from the region.

5. Further, the facility also remains as an applied research instrument for networking and cyberinfrastructure researchers to advance the science and technology of distributed scientific cyberinfrastructure and systems for its lifetime.

6. Further, to engage undergraduate students in the science cyberinfrastructure innovation in alignment with the CC* program goals of NSF.

7. Further, an implicit aim of the project is to catalyze emphasis of research computing into the institutional business practice of Kent's Information Systems and Services.

8. Further, the project also aims at developing appropriate principles, policy, and culture conforming to the above seven goals. This also includes defining the management process for ICE, building the process automation, communication, and automation tools supporting the above.

This document outlines our very first version of the process, the management teams, and the experience report that we have built for ICE.

3. Outline of The ICE Management Process

Fig-1 describes our first onboarding process management plan for bringing in science driver projects over the ICE. It is handled by eight teams. The idea is to give our researchers a white glove onboarding experience with the intention to lower the barrier for entry into our system for those with limited cluster/IT experience while perpetually fulfilling the NSFs objectives for this

investment. A team of 17 experts from various units are graciously working for the inaugural ICE team. The teams and their responsibilities are described below.

3.1. Requirement Gather Team (RGT)

The RGTs purpose is to meet with the researcher, understand what they are trying to accomplish and pick a solution that best meets the researcher needs based on the resources we have. Data is collected to specify items such as number of cores, memory, operating system, GPU need, software requirements, security, and storage requirements.

3.2. Engineering Design Team (EDT)

The EDTs purpose is to review request and verify that all needed information has been gathered and confirm request is doable with the ICE constraints. This group contains experts from all the other required groups and teams.

3.3. Policy & Executive Review Board (PERB)

PERVs task is to arbitrate and ensure the match between the demand and supply of general resources, and priorities between projects when there is contention. It normally handles requests from the researcher that are outside the scope of this project. They handle issues such as researchers requesting maybe requesting large 200PB storage which may monopolize the entire cluster for one project's individual purpose. At current stage, we are focusing on onboarding s many possible projects and our expectation is that after saturation EAT will delve into policies for contention of resource as more such situation arises.

3.4. Network Team (NET)

NET is part of the deployment team. NETs purpose is to configure and insure that the network Layer 2 and 3 are configured correctly and add new VLANs for each new project.

3.5. Security Teams: (SEC)

SEC is also part of the deployment team. SECs purpose is to handle VPN access and control all required firewall rules for each project.

3.6. Server Management Team (SER)

SMT is also a part of the deployment team. ST builds the servers with appropriate operating system and predetermined versions and patches following standards set up by security. ST also works on automation projects with Ansible to make deployments as automated and consistent as possible.

3.7. Federated Resource & Access Team (FED)

The task of the FED is to ensure that federated computing systems are integrated into ICE. Its job is to ensure that external scientific and computing resources such as OSC, OSG and various national Labs are available and accessible to ICE users. Also, it will ensure that part of ICE hardware is available to remote researchers.

3.8. Application and Support Team (AST)

Once the deployment is ready, AST meets with researchers at the end of the process to walk them through VPN access and connecting to their requested resources and installing any applications that they might need help with. They also perform follow up and continued support on any issues or questions that may come up. The AST also is expected to provide time-to-time consultation to researchers if there is any issue.

3.9. NSF Goal Conformance Reporting, Broder Impact (NSF) Team

This team's task is to ensure that projects committed goals to National Science Foundation are met throughout the lifetime of the NSF funded resources. It includes reporting, outreach coordination, and advocacy ensuring its broader impacts goals are met. These including equitable availability of the NSF funded resources [3,6] to all faculty and student researchers from all KSUs campuses as well as to faculty researchers from other regional institutions, availability of minimum 20% resources to federated cyberinfrastructure, support for broadening participation in computing (BPC) goals by the ICE facility as well as projects supported by ICE. Assurance of Broader impact goals [2] including wide accessibility, commitment towards federated and shared use of resource, and broadening participation in computing (BPC) [5], and support for undergraduate research.

4. Inaugural Team

We are glad to report that a team of 17 experts have joined to serve on our inaugural ICE Management Body. They will serve for a period of 5 years. The teams have been purposefully built with overlapping membership. For example, implementers (SEC+NET+SER) also serve in the EDT to ensure all aspects are reviewed and addressed in the design phase.

Table-1 Inaugural Teams & Team Members for KSU sci.SDZ Integrated Research Computing Facilities

Team Member	Position	Teams							
		NSF	RGT	EDT	PERB	SEC	NET	SER	AST
Dr. Javed I. Khan	ICE Principal Investigator & Professor of CI, CS	x			x				
Philip Thomas	ICE Co-PI & Research Architect, IT	x	x	x	x				x
Jeanne Tan	ICE Project Director & Special Assistant, CS	x							
Jim Raber	Assoc Chief Info Officer, IT				x				
Dr. Michael Kavulic	Asst VP, Research Adminm RED				x				
Jeff Bailey	Sr. Research Engineer, CS			x				x	
Roy Heath	Lead System Admin, Sr. Research Engineer, CS		x	x				x	x
Jeremiah Schilens	Lead Systems Admin, IT			x				x	
Joshua Talbott	Dir Information Technology, CAS		x	x				x	x
Keith Hodar	Lead Systems Admin, IT			x				x	
Mike Geist	Dr. Information Technology, IT			x					
Nick Lupica	Lead Systems Admin, IT			x				x	
Tom Schindler	Assoc Director, IT					x			
Tom Bordonaro	Sr Network Engineer, IT			x			x		
Walt Bainey	Sr. Dir Information Technology,IT			x					
William Wroblewski	Sr. Systems Admin, IT			x				x	
AJ Chiccarino	Sr Manager, IT			x		x			

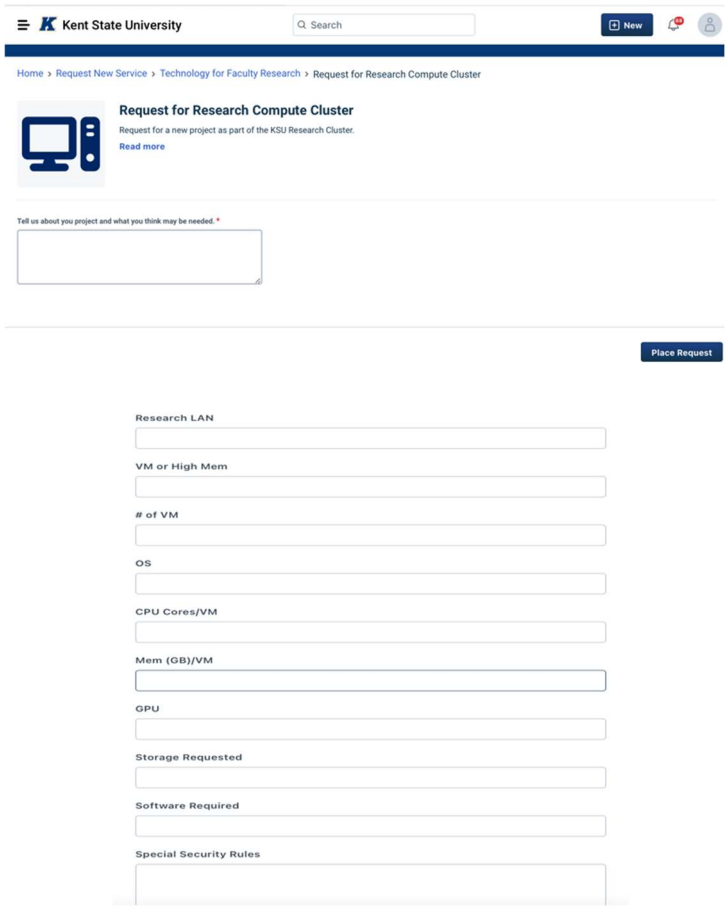
** CS, Department of Computer Science*
** RED, Division of Reseaech and Economic Development*
** IT, Division of Information Technology*

5. Onboarding

We have implemented the first web-based ticketing system to smoothly manage the onboarding process. The first current system can be found here. It is currently accessible to only Kent's edu-persons.

https://kentstate.freshservice.com/support/catalog/items/367

The simplified ticket asks basic questions and also researcher does not need to fill up all to proceed. Below are two snapshots.



6. The ICE Onboarding Process & Ticketing System

Once the onboarding ticket is created the overall process flow is shown in Fig-1. It first alerts the RGT to contact the researchers. The RGT meets the researcher and extracts required information about the requirements. Then the ticket moves to EDT. If there is anything unusual in the request, then the EMT is consulted. Otherwise, it moves to the deployment teams. After completing the onboarding finally, it is given to the Application and Project support team. Who finally provides the researcher with the required information and training for him/her to effectively start using the project-computing.

7. Conclusion and Ongoing Work

This document outlines the inaugural version of Research Computing at Kent State University established. A key objective of the next step of this project is to automate and be able to scale up, secure, standardized deployments, of virtual machines.

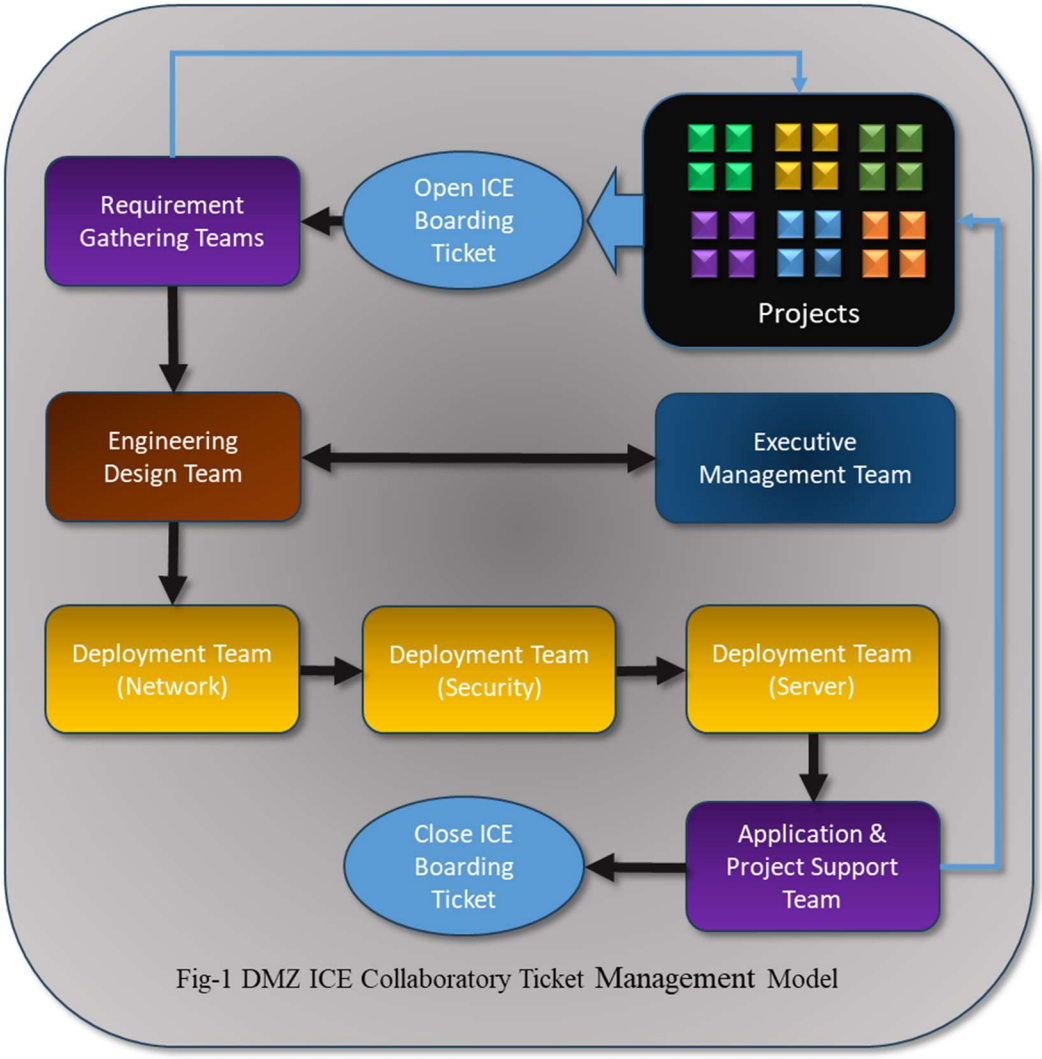


Fig-1 DMZ ICE Collaboratory Ticket Management Model

8. Acknowledgement

The project acknowledges support from many individuals including John Rathje, Jeanne Tan, Jim Raber, Dr. Michael Kavulic, Jeff Bailey, Roy Heath, Jeremiah Schilens, Joshua Talbott, Keith Hodar, Mike Geist, Nick Lupica, Tom Schindler, Tom Bordonaro, Walt Bainey, William Wroblewski and AJ Chiccarino, Dr. Tim Arndt, Cleveland State University, Dr. Michael Strickland, Professor, Physics, Director, Center for Nuclear Research, Dr. Helen Piontkivska, Associate Professor, Biological Sciences, Dr. He Yin, Assistant Professor, Geography, Dr. Barry Dunietz, Associate Professor, Chemistry & Biochemistry, Dr. Robin Selinger, Professor, Physics, Dr. Xiaozhen Mou, Assistant Professor, Biological Sciences, Dr. Robert Clements, Assistant Professor, Biological

Science, Dr. Joseph Ortiz, Professor, Geography, Dr. Michael Carl, Professor, Modern & Classical Languages, Dr. Jong-Yoon Kim, Assistant Professor, Computer Science, Dr. Kwangtaek Kim, Assistant Professor, Computer Science, Dr. Qiang Guan, Scientist, Los Alamos National Lab & Assistant Professor, Computer Science, Dr. CC Lu, Professor, Computer Science, Dr. Ye Zhao, Professor, Computer Science, Dr. Jong-Hoon Kim, Associate Professor, Computer Science, Dr. Xiang Lian, Associate Professor, Computer Science, Dr. Gokarna Sharma, Associate Professor, Computer Science, Dr. Rouming Jin, Professor, Computer Science, Dr. Angela Guercio, Professor, Kent State Stark Campus, Dr. Bathi Kasturiarachi, Professor, Mathematics, Kent State Stark Campus, Dr. Younghun Chae, Assistant Professor, CS, Kent State Stark Campus.

The project has been supported by funding from the National Science Foundation NSF Award#2201558 [3], and NSF Award# 1925678 [4], engineering time contributed by the Division of Information Technology, location and engineering time donated by Department of Computer Science. The ICE cluster has also received supplementary funding from the division of Research and Economic Development.

9. References

[1] "Transforming Science Through Cyberinfrastructure: NSF's

Blueprint for a National Cyberinfrastructure Ecosystem for Science and Engineering in the 21st Century" (see https://www.nsf.gov/cise/oac/vision/blueprint2019/).

[2] NSF Perspectives on Boader Impacts, https://nsf-gov-resources.nsf.gov/2022-09/Broader_Impacts_0.pdf

[3] Javed I. Khan (PI), (with co-PI Phil Thomas) Proposal Title: CC* Compute: Accelerating Compute Driven Science Through a Sharable High Performance Computing Cluster in Kent State Multi-Campus System Amount: $400,00, Sponsoring Agency: National Science Foundation, Cyber Infrastructure Program, NSF Award#2201558, Duration: July 15,2022- June 30, 2025.

[4] Javed I. Khan (PI), (with co-PI Paul Schopios, OARNET, & Phil Thomas) Proposal Title: CC Networking Infrastructure Network for Data Driven Science in Allied 21st Century Smart Multi-Campus System: A Use Case Design Through Kent State's Sharable Science DMZ, Sponsoring Agency: National Science Foundation, Cyber Infrastructure Program, NSF Award# 1925678, Duration: July 15,2019- June 30, 2023.

[5] Broadening Participation in Computing, NSF Publication https://new.nsf.gov/cise/broadening-participation, Last Retrieved 2/2/2024.

[6] Technical Report TR 2023-10-01, A sci.DMZ and Data Transfer Network Node for Kent State University, Javed I. Khan and Philip Thomas, October 2023