HALCYON

UNIFIED HPC CENTER OPERATIONS

PURDUE UNIVERSITY

Defining The Problem

Cluster Management For Faculty

Under a shared campus cluster model, with many research groups investing distinct amounts and annual new hardware acquisitions, management of users and resources can become very complex.

At Purdue, the Research Computing division set out in 2011 to design a cluster management solution to empower faculty to manage access to their own purchased resources.

Immediate Results

Time & Effort Savings

Staff time

6.7 hrs / wk

New storage space

8 days ➡ 1 hour

In its initial incarnation, portal provided faculty, students, and center staff an immediate reduction in time and effort to request, provision, and manage allocations. User allocation alone was estimated as saving 6.7 hours of center staff time per week over the first 18 months. New storage spaces were able to reduce time-to-completion from 8 days to 1 hour within the first year of adoption of storage management.

Portal Evolution

More Features, More Problems

In 2011 a cluster management solution to empower faculty to manage access to their own purchased resources is designed and built.

Time-savings and preference for “self-serve” utilities were immediately apparent.

As operations at Purdue expanded, the internal portal evolved and took on many aspects of the operation of an HPC center beyond resource allocation and management. This included:

- HPC & storage resource management
- User management
- Customer relationship
- Communications
- Documentation
- Purchasing

While the original Purdue Research Computing portal served its purpose well over the years, making changes and adding features could be a cumbersome task. Dependencies were becoming increasingly complex and tangled.

The decision was made to re-architect the underlying code.

Re-Architecture

Planning For The Future

01 Abstraction of assumptions & settings

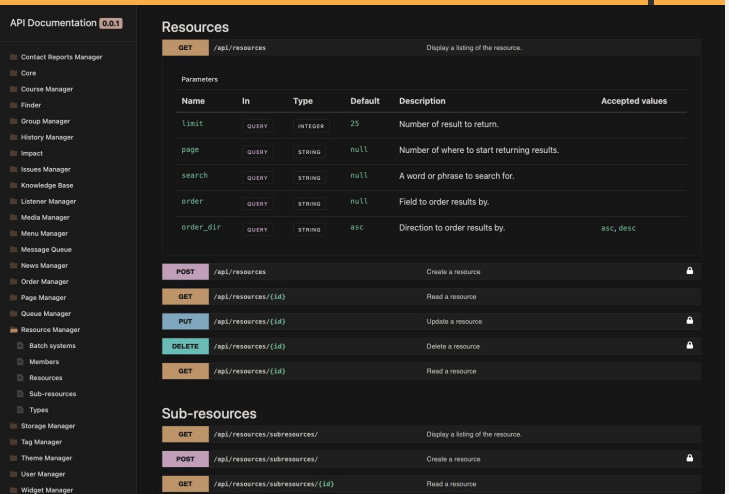
Standard local language for many aspects and various hard-coded site settings were pulled into the internal database, and interfaces were created for editing these. Other environment site settings were incorporated through new configuration files.

02 Restructure for modularity & extendability

The portal code was separated into modules—logical groupings based on the data being handled, tasks, and interfaces. 3rd-party services are integrated as plug-ins through the an event dispatcher system, allowing communication with sources such as LDAP or REST APIs.

Re-Architecture

External Interactions



REST API

Interaction points for external services and resources, modules implement REST APIs, documented according to the OpenAPI v3 spec.

Message Queue

External resources such as compute clusters or storage can make use of a message queue to make changes such as new directories or alter permission.

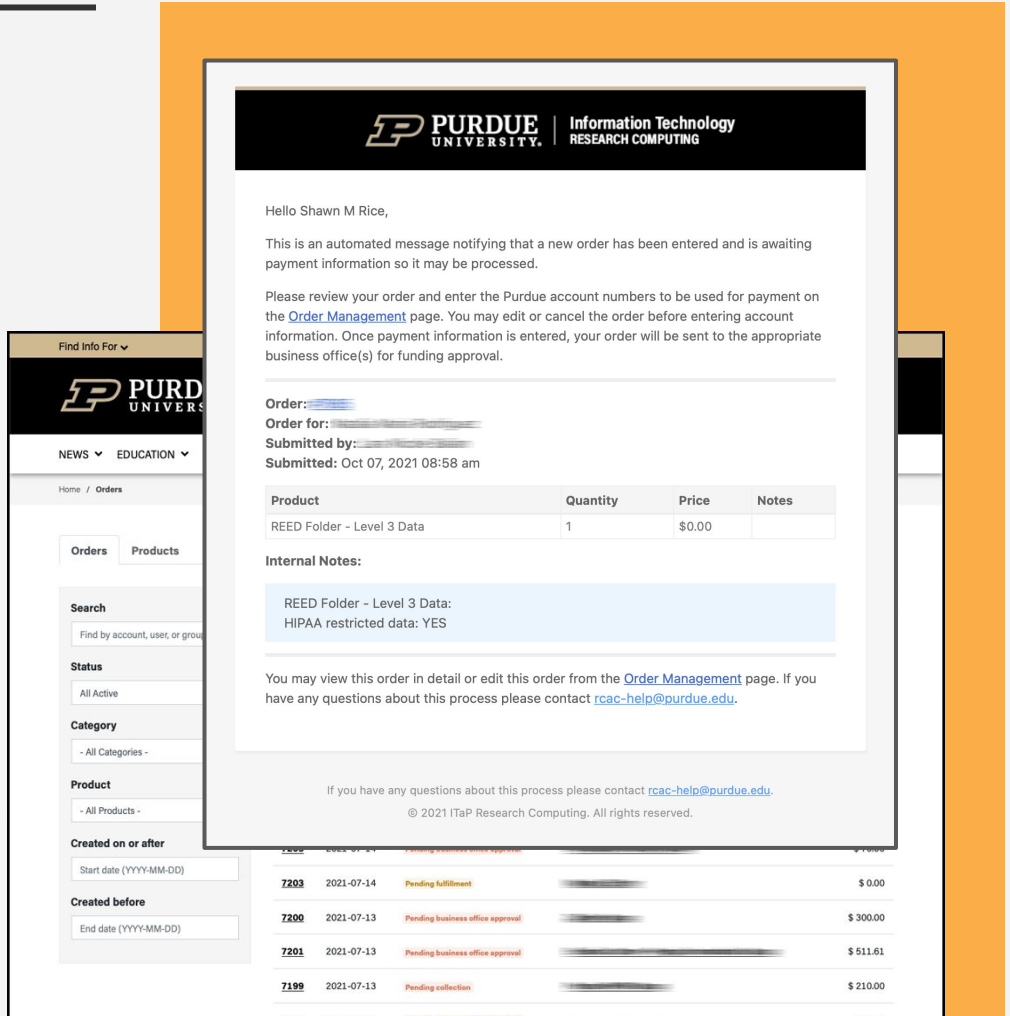
Command Line Utility

A robust command line utility is provided with numerous commands for inspecting and interacting with data.

Self-Serve Interfaces

Make Purchasing Easy

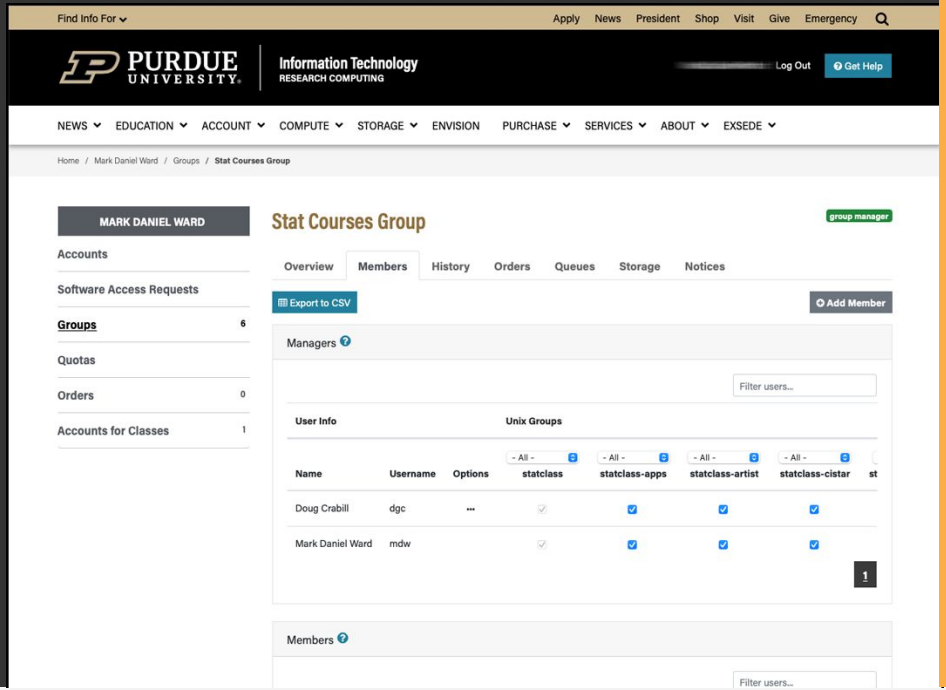
Re-architecture did not stop with the codebase and included careful scrutiny and re-thinking of user interfaces. Allocations can be ordered through a familiar online shopping experience and users are notified of status changes via automated emails. Payment accounts can be added to an order with payment split across various accounts and the relevant business offices assigned as approvers. Much of this is self-serve, minimizing direct staff involvement.

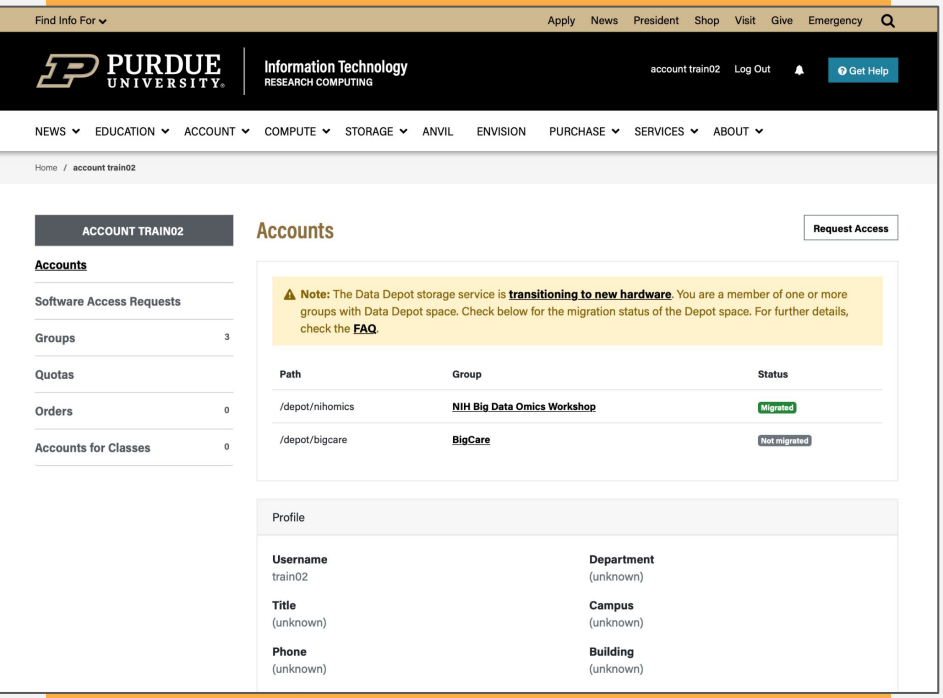


Self-Serve Interfaces

Giving Control To Users

Resource allocations are centralized under research groups and options for control over membership, unix groups, storage directories and access are continually expanded, allowing for an ever increasing self-serve experience. This frees up time for staff and hastens a number of processes as the user no longer must rely on a middle-man.





Self-Serve Interfaces

Centralized Information

Information relevant to a user can be found from their account page. This reduces time and effort on both the customer and staff in looking up details on interactions, allocations, and more.

Each menu and its related page is injected via an event dispatcher. Modules or plugins can subscribe to events and add information or entire pages as needed, allowing for integration of new features but avoiding hard-coding of dependencies.

All The Things

Unified HPC Center Operations

Resources

Manage resources and allocations.

Purchasing

Products, carts, and orders. Allow users to purchase resources with a familiar online shopping experience.

Documentation

How to use the resources provided by the HPC center.

News

News stories, events, maintenance and outage notifications.

Contact Reports

Customer relationship. Document when and what was talked about.

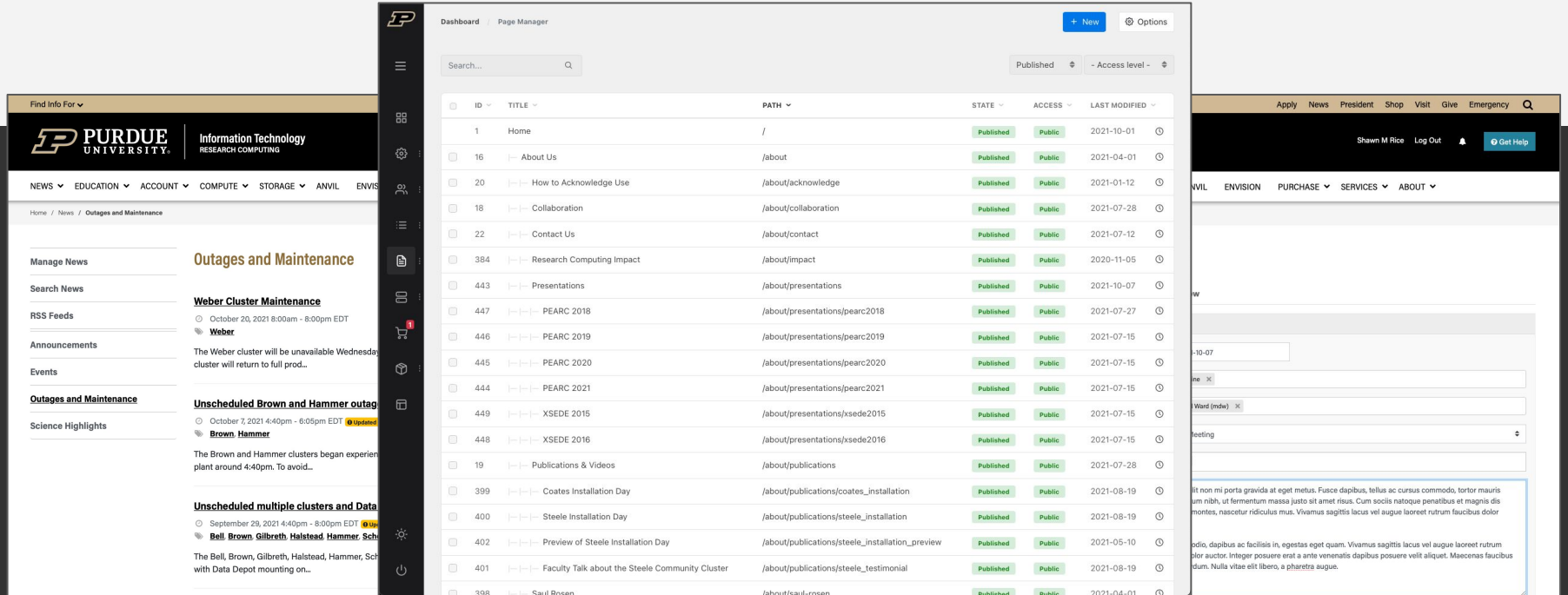
Online Presence

Informational pages, menus, contact forms, and other utilities needed for a HPC center's online presence.

Extensibility

More Features, No Problems

Extending functionality is considerably easier and entire modules can even be replaced with ease if the current functionality doesn't meet current needs.



Future Work

Things To Come

Collaboration

Implementing Halcyon at another organization is a current goal, which will further highlight where flexibility and configuration can be introduced to better accommodate different infrastructure and HPC center policies.

Refining

While there has been significant re-architecture, there still remain a few values and interactions built into code which are specific to the environments and implementations of Research Computing at Purdue. Work on improving UI/UX continues.

With Thanks

References

Shawn Rice
Research Computing
zooley@purdue.edu

Kevin Colby
Research Computing
colbykd@purdue.edu

Icons provided by *Komkrit Noenpoempisut*
https://thenounproject.com/itim2101

1 Kevin D. Colby, Daniel T. Dietz, Preston M. Smith, and Donna D. Cumberland, ”Self-service Queue and User Management in Shared Clusters,” in *Proceedings of the 1st International Workshop on HPC User Support Tools, HUST14, Nov. 2014, New Orleans, LA, USA*, IEEEPress, 2014.

2 Mohammed Zia, et al., ”ColdFront: An Open Source HPC Resource Allocation System,” Presented at the 5th International Workshop on HPC User Support Tools, HUST18, Nov. 2018, Dallas, TX, USA. 2018.

3 Laravel LLC, ”The PHP Framework for Web Artisans”, laravel.com. [Online]. Available: https://laravel.com/. [Accessed: Jul. 16, 2021].

4 iMatrix Corporation, ”Distributed Messaging”, *zeromq.org*. [Online]. Available: https://zeromq.org/get-started/. [Accessed: Jul. 16, 2021].

5 Donna D. Cumberland, ”Six Sigma: Research Account Queue Management Process Improvement”, West Lafayette: Purdue University Press, 2011.

6 Kevin D. Colby, Daniel T. Dietz, and Lev A. Gorenstein, ”Shared Research Group Storage Solution with Integrated Access Management,” in *Proceedings of PEARC17, 2017*, New York: ACM, 2017.