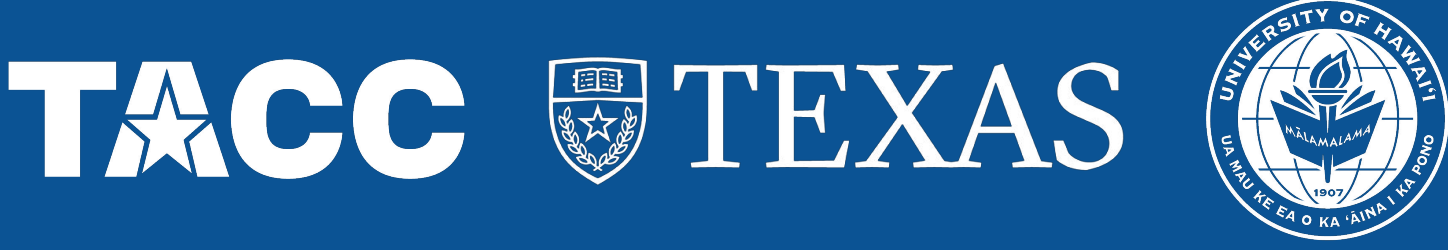




Collaborative Framework: Project Tapis: Next Generation Software for Distributed Research

Joe Stubbs (PI, University of Texas, Austin)
Gwen Jacobs (PI, University of Hawaii, Systems)
Sean Cleveland (co-PI, University of Hawaii, Systems)
Maytal Dahan, (co-PI, University of Texas, Austin)



Overview

Tapis is a hosted, multi-tenant, RESTful API framework for distributed computational research, providing support for data management and code execution across institutional boundaries. Using Tapis APIs, a user can program portable, reproducible workflows that leverage sensors, instruments, experimental facilities and a network of national and regional cyberinfrastructure.

Tapis builds on NSF investments into the Abaco, Agave, and CHORDs projects to provide a production-grade framework for use across a wide range of domains of science and engineering. Including support for streaming/sensor data, portable, reproducible workflows leveraging containerized applications, and a novel, decentralized security kernel, Tapis will be the underlying cyberinfrastructure for modern research.

Early Adopters

Real Time Climate Data for the Hawaiian Islands - Dr. Giambelluca’s research group at the University of Hawaii is working on automated workflows to provide ready access to a spatially comprehensive, high quality, reliable climate data set and data analysis products covering the Hawaiian Islands.

Receptor Plus: Human Immunological Data Storage, Integration and Controlled Sharing - The iReceptor Plus platform provides a science gateway linking to federated, geo-distributed repositories of NGS data conforming to adaptive immune receptor repertoire (AIRR) community standards. Co-funded by the European Union and the Canadian Institutes of Health Research, iReceptor Plus groups plan to leverage Tapis for Metadata management and analysis.

Planet Texas 2050 - The Planet Texas 2050 group wants to leverage spatially dense and ever-increasing Geoscience temporal datasets generated by Arduino-based microcontrollers as ground-truth inputs for integrated models of water-land-atmosphere-urban systems. Tapis will be used for streaming data management and analysis job management.

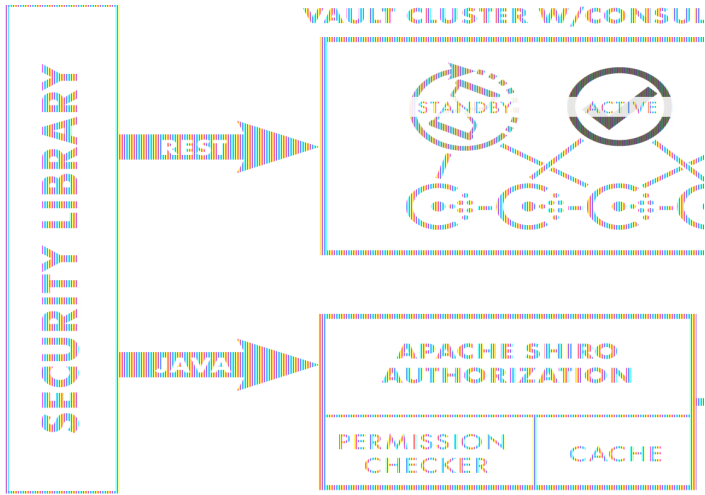
Join Us!

Tapis Early Adopter Workshop at PEARC 2020 Portland, OR	Join us for a full day workshop of talks from Tapis core team members and participants in our Early Adopters program.
Read More On Our Documentation Sites	Tapis Project: https://tapis-project.org OpenAPI Live docs: https://tapis-project.github.io/live-docs/ Developer Docs: https://tacc-cloud.readthedocs.io
 slack	Get questions answered by our support staff and members of our user community: http://bit.ly/join-tapis

Primary Capabilities

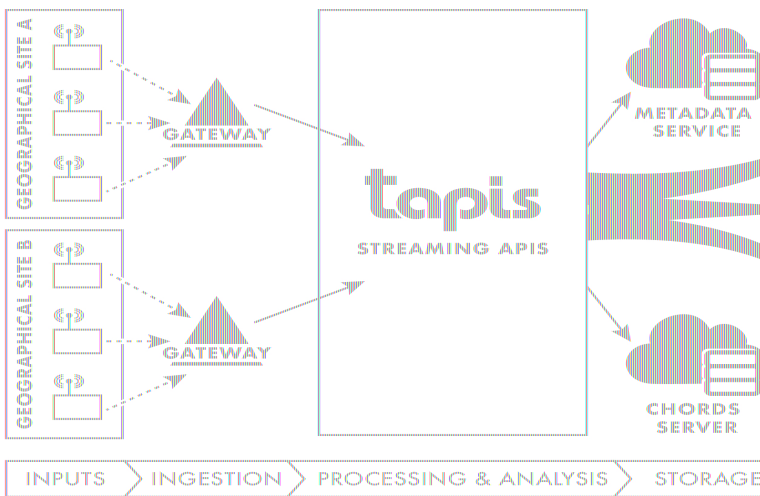
Data Management & Code Execution - Manage files on remote server, run applications on HPC, HTC and cloud systems, with first-class support for containerized codes, including Docker and Singularity, and container image registries such as Docker Hub.

Identity, Authorization, Security and Tenancy - Tapis introduces a novel, decentralized security kernel allowing institutions to manage secrets and authorization data on premise while simultaneously leveraging the Tapis primary site to minimize the administrative effort required to keep the platform running.



The Tapis platform supports multi-tenancy and a modular authentication system, enabling projects to incorporate local campus credentials or a federated identity system such as InCommon, EGI or other eduGAIN.

Support for Streaming Data - Tapis provides APIs to support storing and managing data from IoT devices and integrates this capability into its batch and real-time compute APIs to enable users to program a variety of analysis workflows against the data.



DevOps Tooling for Automated Platform Management - The Tapis Platform provides tooling to automate the deployment and management of its components. All services are packaged as Docker containers and tooling targets the Kubernetes orchestration platform.

Tapis Architecture

Tapis supports a multi-datacenter deployment configuration, including a primary site and one or more associate sites. Associate sites run a relatively small number of Tapis services to support data locality and maintain security compliance.

