# University of Hawai‘i Cyberinfrastructure Plan

**Introduction:** The University of Hawai‘i (UH) is already one of the nation’s top research universities, with distinctive strengths in astronomy, earth and ocean sciences. In developing the new Hawai’i Innovation Initiative ($HI^2$), which calls for bold expansion in these and other strategic areas such as agriculture and the health sciences, it became apparent that stronger capabilities would also be needed in informatics and cyberinfrastructure. This plan is intended to provide the necessary platforms for a doubling of research at the University of Hawai‘i.

The University of Hawai‘i cyberinfrastructure plan is designed to support modern research, which increasingly relies on: “big” data collected and generated locally or remotely; comprehensive data management and stewardship; high performance computation (HPC), analytics and simulation; and advanced visualization. The plan is also designed to support global collaboration among teams that cross traditional disciplinary, institutional, and geographic boundaries. These capabilities all require advanced communication networks, which are a particular challenge given Hawai‘i’s location, as well as new education and training programs for researchers and support staff. The University of Hawai‘i cyberinfrastructure plan is intended to address all these elements in a comprehensive and holistic manner.

**Computation / HPC:** As at most universities, many researchers and research groups operated their own small and mid-ranged clusters. From 2009-2012 UH researchers also had access to The Hawai’i Open Supercomputer Center (HOSC), a collaboration with the U.S. Air Force Research Lab at the Maui High Performance Computing Center through which approximately 50 UH faculty and staff have utilized a cluster with 1280 compute cores and 90 TB of scratch space. This facility was closed in December 2012 and a new HPC computing facility will be developed and housed in the new UH Information Technology Center scheduled to open in 2013, which includes an 8000 sq. ft. energy efficient hardened data center. Over $1m in startup funds has already been identified to create the first phase of a “condominium” model for HPC computing at UH that will improve cost efficiencies, operating efficiencies, security and energy utilization over previous approaches, which are not sustainable or scalable. The expected initial startup system is expected to provide capability in the range of 100 TeraFLOPs, far beyond any locally available open capability.

**Data storage and management:** For Hawai’i, the challenges of managing and storing big data are especially acute. Current generation telescopes already produce terabytes of data per day. Ocean, earth and atmospheric observing systems integrate massive global datasets for the study of environmental change, and the study of disease trends in human populations using modern techniques in genomics and bioinformatics generate huge volumes of data. As with HPC, most data storage is decentralized in research labs around the University. This approach presents substantial inefficiencies and risks, and cannot support the planned growth in research at UH. The UH Information Technology Services (ITS) unit has developed an initial data storage resource, which will be relocated to the new IT Center during 2013. The resource is designed to scale with campus needs. The current capacity of the storage resources in use currently is approximately 109TB of faster SAS/FC storage and 145TB of slower SATA storage. In building out its new HPC capability, UH plans to develop petabyte-scale storage capacity to support

researchers. ITS is in the early stages of exploring a collaboration with the UH Library to develop a suite of data management services to support researchers.

In addition to "big data" and HPC support, ITS has recently developed an array of computing, networking, and support services for researchers. Currently, ITS supports a virtual server (VMware) service to provision servers to meet researcher needs. Data storage services include bulk storage of research data, digital media, and other digital assets. A number of research groups use this environment to create data portals, typically using the centrally supported open source Drupal web software, through which data can be uploaded, managed and accessed. This storage is handled by an on-campus storage cloud, which provides advanced RAID technology to ensure the safety, security, and integrity of the stored data. Back-ups of bulk storage occur nightly and are kept until the next backup occurs. More work needs to be done to routinize setup of these lightweight portals and enable them with federated identity and access management.

**Visualization.** Scientific visualization is a key growth area for UH and the $HI^2$. Current visualization facilities are limited, and UH lacks faculty expertise in this area. Fortunately, UH has a small base of capability to build on. The most advanced visualization capabilities are at the new Center for Microbial Oceanography: Research and Education (C-MORE) Hale ("Hale" is Hawaiian for "house") building at UH-Manoa, home of the C-MORE program, an NSF-sponsored Science and Technology Center. C-MORE Hale has two OptiPortal systems. The OptiPortal is a cluster-based visualization and collaboration tool developed at UCSD, CalIT2 and UIC with NSF support that provides support for remote collaboration and data analysis by bringing together a variety of data sets from the partner institutes onto one "canvas" for analysis. A less sophisticated and lower-cost system was designed in-house and deployed in 2011 by the School of Ocean Earth Science and Technology (SOEST). This provides a 4x3 matrix of 55" 1080pixel Planar display-wall class monitors that is used for interactive viewing of large-scale images.

Moving forward, the Department of Computer and Information Sciences (ICS) at UH Manoa has an open faculty search for a professor with expertise in scientific visualization. This individual will join the faculty in Fall 2013 and is expected to provide leadership on infrastructure and applications across multiple disciplines. New visualization facilities will be installed, maintained and made available to the entire campus in the new IT Center, with guidance and leadership expected from the new faculty leader.

**Collaboration:** In addition to the data portals mentioned above, UH has implemented the Sakai community source Collaboration and Learning Environment, locally branded as "Laulima." Laulima is available to every course section of every course on every UH campus. In addition, it is widely used for collaboration within UH for committee, task forces and research projects. This provides a readily available asynchronous collaboration environment that is well known to UH faculty and students. For synchronous collaboration, UH has desktop, room-based and hybrid capabilities. UH has licensed the widely used Adobe Connect software, branded locally as Halawai, and makes it available for all UH research and educational groups and meetings. In addition, UH has been a pioneer in interactive video for distance learning across the islands and has upgraded its statewide distance learning system to utilize standards-based high definition IP videoconferencing technology. The current Polycom technology in wide use interoperates with

all other standards-based videoconferencing (H.323 and SIP) across the state, nationally and globally. Software licensed from Polycom allows interoperability with suitable equipped desktop and even mobile devices. In addition, Cisco donated two midrange telepresence units to UH. This technology will be made generally available in the new IT Center and is configured to interoperate with the Internet2 and NLR telepresence networks for national and global reachability. Future work must focus on proceduralization of access methods across gateways and between systems, and improving access requests and scheduling.

**Campus, Statewide, National and International Networking:** The UH GigaPOP provides connectivity to national and international Research and Education networks for the entire state of Hawai‘i. The University has two 10Gbps circuits that connect directly back to the Pacific Northwest GigaPOP (PNWGP) in Seattle and the Corporation for Education Network Initiatives in California (CENIC) in Los Angeles. In addition, the University has two 10Gbps connections shared with the Australian Advanced Research Network (AARNet) as part of an NSF IRNC grant (OCI-0962931) that connect both to Australia and the U.S. West Coast (Seattle and LA). Planning is underway to upgrade these circuits to 40Gbps during 2013. The UH statewide network is composed of 10Gbps links on and between islands, extending to all of public higher education in Hawai‘i. This network is also being extended to major research facilities on all islands, including the Mauna Kea and Haleakala observatories. An interisland OC-3 microwave network provides interisland backup capacity.

Using physical fiber architecture initially designed in the late 1980s, the UH-Manoa campus network has evolved to a routed topology based on a three-tier distribution model – Core, Distribution Hubs and Buildings. There are currently two core routers and thirteen fiber distribution hub buildings, which feed all other buildings on campus. Each of the fiber distribution hubs is connected to both of the core routers. To further increase reliability, key building routers are directly connected back to the core routers, allowing a fiber distribution hub location to be offline without impacting services to the building. With a few exceptions, 1Gbps connections are utilized from the core to the fiber distribution hubs and from the distribution hubs to almost all of the campus buildings. Most backup connections from buildings directly to the core are at 100 Mbps. Interior building wiring is a mixture of voice-grade copper up to Cat-5e, depending on when funding has been available for upgrades or major renovations performed.

This network has been more-or-less adequate for two decades of general campus use, but is not positioned to support high-capacity data flows that characterize advanced cyberinfrastructure-empowered research. The campus network core, along with the GigaPOP, will be relocated into the new IT Center in 2013, which will also house the UH IT Operations Center, which serves as a consolidated NOC and monitoring center for systems and services. A major upgrade of the campus fiber and wiring infrastructure is funded and underway. Intra-building wiring is being upgraded to provide fiber to all wiring closets and Cat5e to the jack. All buildings will have fiber connectivity to one or more of the distribution hub buildings, which are being consolidated. In addition, new Power over Ethernet (PoE) gigabit Ethernet switches are being deployed across the campus to support a concurrent VoIP implementation. With the new infrastructure in place, the campus is being positioned well for 10Gbps links to distribution hubs, buildings and labs across campus. Upgrades are required to support high-end science, starting with buildings and labs engaged in data-intensive science and to support global and national end-to-end flows.

IPv6 is an integral end-to-end capability of the entire UH network. The campus network provides IPv4 and IPv6 connectivity to every building and room within the campus and IPv6 is fully supported on the wide area as well. UH was an early and active adopter of IPv6, driven by our relationships with Asian R&E networks and our longstanding close connections with the Defense Research and Engineering Network (DREN), the early IPv6 adopter within the U.S. government. UH founded and facilitates the Hawai‘i IPv6 Task Force, which promotes IPv6 across the public and private sectors statewide.

UH has also embraced the perfSONAR framework, primarily through our work in the NSF International Research Network Connection (IRNC) program. UH network engineers provide active leadership in the "perfSONAR club." Strategic deployment of perfSONAR nodes around the UH network has proven to be invaluable in diagnosing performance problems and anomalies and needs to be expanded on campus to support network monitoring and management.

Moving forward, UH has been selected for deployment of a GENI rack during the next spiral. GENI, and more broadly, software defined networking, must be explored to bring additional capabilities in support of advanced research as well as network management.

**Identity and Access Management:** The UH has deployed the UH Identity Management System (UHIMS) for enterprise identity and access Management utilizing a mix of home grown, Internet2, and InCommon software. UHIMS [UHIMS] currently provides the following features:

- Person Registry - the Person Registry is integrated with all ERP systems that contain person information
- Directory Information Provisioners: Directory information is provisioned to AD, LDAP, Grouper, as well as to a number of administrative services.
- Authentication and Authorization Services - single sign-on authentication and authorization functions are provided via our deployment of the JA-SIG (now Apereo) Central Authentication Service (CAS) and the Internet2 Shibboleth Identity Server.
- Federation: UH is an InCommon member and participates in identity federation, enabling access to resources such as NSF's research.gov website via UH institutional credentials.

Future plans to enhance federated identity services via UHIMS include the following:

- "Shibbolizing" key research services, such as Drupal data portals and collaboration servicers to enable full participation by collaborators at other InCommon institutions and organizations.
- Utilization of Grouper for centralized entitlement management (authorizations).
- Implementation of multifactor authentication and the business practices required for assertion of InCommon Bronze and Silver Levels of Assurance.

**Education and Training:** The transformation of research to embrace cyberinfrastructure-empowered approaches will not simply occur when the physical and logical cyberinfrastructure is provided. UH must fully integrate its support of active researchers with its human resource development activities. UH has just hired its first Director of Cyberinfrastructure, who will serve

as the lead for cyberinfrastructure training programs. With limited resources, UH must leverage existing and emerging national-scale programs. The Director of Cyberinfrastructure will be the UH focal point for engagement in programs such as XSEDE through its Campus Champions program, opportunities available with and through EPSCoR to develop cyberinfrastructure capacity throughout its jurisdictions, CASC, and others. Hawai‘i must also engage in development of curriculum in computational sciences, support undergraduate and graduate students engaged in cyberinfrastructure-enabled research under the direction of active investigators, share information and collaboration opportunities within Hawai‘i regarding effective uses of cyberinfrastructure in UH research programs.