

Finding Subatomic Particles and Nanoscopic Gold with Open, Shared Computing

David Swanson, Director, Holland Computing Center, University of Nebraska

Modern data intensive research is advancing knowledge in fields ranging from particle physics to the humanities. The data sets and computational demands of these pursuits put ever-increasing strain on University resources. The history of resources at the Holland Computing Center (HCC) at the University of Nebraska (NU) reflects these dramatic changes. The need to constantly evolve and grow resources economically requires careful strategy and provides incentive to combine and share resources wherever possible. The remainder of this paper provides a brief overview of HCC and outlines some recent highlights and strategies that have emerged over time.

HCC Overview

Facilities

HCC has two primary locations directly interconnected by a pair of 10 Gbps fiber optic links. The 1800 sq. ft. HCC machine room at the Peter Kiewit Institute (PKI) in Omaha can provide up to 500 kVA in UPS and genset protected power, and 160 ton cooling. A 2200 sq. ft. second machine room in the Schorr Center at the University of Nebraska-Lincoln (UNL) can currently provide up to 60 ton cooling with up to 400 kVA of power. A cooling upgrade to match the maximum power load is in process. Both locations have 10 Gbps connections to Internet2 in addition to a pair of 10 Gbps links connecting the machine rooms.

HCC's resources at UNL include two distinct offerings: Sandhills and Red. Sandhills is a linux "condominium" cluster provided for general campus usage with 1500 compute cores interconnected by low-latency Infiniband networking. A recent extension provides a

higher RAM/core ratio. The largest machine on the Lincoln campus is Red, with over 3100 cores interconnected by less expensive, but also higher-latency, gigabit Ethernet. More importantly, Red serves up over 1.2 PB of storage. Red is integrated with the Open Science Grid (OSG), and serves as a major site for storage and analysis in the international high energy physics project known as CMS (Compact Muon Solenoid).

In late May of 2009 the University was donated Firefly, a 1152 node, primarily dual-core Opteron cluster which became the largest resource in the University system. It is connected by Cisco SDR Infiniband and supports 150 TB of Panasas storage. Originally capable of 21.5 TFlops, it is located at PKI, and is nearing retirement. The latest cluster named Tusker is also located at PKI. Tusker offers 6,784 cores interconnected with Mellanox QDR Infiniband along with 360TB of Terascale Lustre storage.

Each individual node contains 64 compute cores with a total of 256 GB RAM.

Campus Grid

HCC maintains a Campus Grid that is able to transparently submit large batches of jobs first to the campus, then to Big 10 peers Purdue and Wisconsin, and finally to the OSG. Over 4M cpu hours have been processed via the campus grid framework over the last year. Processing is occurring significantly in Nebraska, but also at Fermilab, Wisconsin, Purdue, Michigan and many other locations as far away as Caltech and Brazil. This method of computing is also able to access resources on the Amazon EC2 Cloud. These have been minimally utilized due to the fact of current pricing but the spot pricing option may become sufficiently affordable soon. Recent studies by the DOE concerning storage have continued to conclude that Cloud resources, even if capable of the scales demanded by data centric science -- not yet confirmed -- are currently far too expensive compared to locating resources at Universities and Labs.

Open Science Grid

"The Open Science Grid (OSG) advances science through open distributed computing. The OSG is a multi-disciplinary partnership to federate local, regional, community and national cyberinfrastructures to meet the needs of research and academic communities at all scales."

[<https://www.opensciencegrid.org/bin/view/Documentation/WhatIsOSG>]

HCC's relationship with OSG has grown steadily over the last several years, and HCC is committed to a vision that includes grid computing and shared national CI. Recently, an HCC VO was formed, and researchers affiliated with

HCC are now able to run on all HCC resources as well as resources across the globe through OSG protocols. The experience with OSG to date has proven to be a great benefit to HCC and NU researchers, as evidenced by the HCC VO's usage of over 11 million cpu hours in the last calendar year.

Data Intensive High Throughput Computing @ HCC

US CMS Tier2 Computing

HCC is a substantial contributor to the LHC (Large Hadron Collider) experiment located at CERN, arguably the largest cyberinfrastructure effort in the world. Researchers affiliated with the LHC are probing previously unreachable energy regimes to expand our understanding of the very fabric of the universe. In particular, they have been searching for the Higgs, or "god," particle, which will help to explain why matter has mass.

One of the experiments at the LHC is CMS. 30 PB of disk space (in an archival, "write once, read many" mode) and 38 PB of tape were required for the complete analysis of just the first year of LHC data. The storage and processing requirements are at unprecedented scales for scientific computing. In the US, the underlying grid middleware is provided by the OSG.

HCC operates one of the seven CMS "Tier-2" computing centers in the United States. Tier-2 centers are the primary sites for user analysis of CMS data; they host datasets that are of interest to physicists, who submit jobs over the grid to process those datasets. They are also the primary sites for generating collision simulations that are used by physicists throughout the experiment; these jobs

	1999	2002	2005	2009	2012	Increase
Personnel	2	3	5	9	13	7x
WAN bandwidth (gbps)	0.155	0.155	0.622	10	30	200x
CPU cores	16	256	656	6956	14492	900x
Storage (TB)	0.108	1.2	31.2	1200	3250	30,000x

are submitted over the grid by a central team of operators. The Nebraska Tier-2 currently has 3000 processing cores and hosts 1200 TB of CMS data with redundant replication.

"I think we have it. We have discovered a particle that is consistent with a Higgs boson." This quote from CERN Director-General Rolf Heuer was the remarkably understated announcement that a decade-long effort by an international team of thousands of scientists using arguably the largest cyberinfrastructure environment in world history

had made a major breakthrough. Although a myriad of fine details remain to be determined, something that is "consistent" with a Higgs field -- and a mechanism that gives rise to mass for all matter -- has been confirmed.

Movement of data sets of these sizes requires the use of high performance networks. UNL has partnered with other regional Universities in the Great Plains Network (GPN) to share the cost of obtaining and maintaining connectivity to Internet2. The Tier2 site at HCC routinely moves data at rates approaching 10

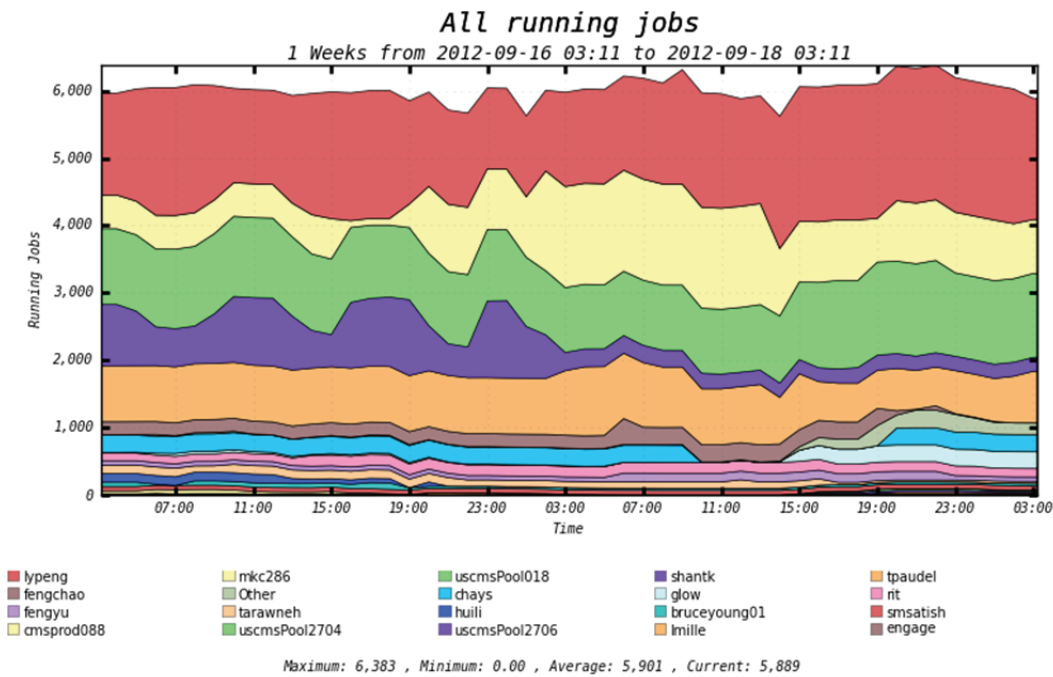


Figure: Mixture of Jobs at HCC

The figure above shows a 2-day window of running jobs at HCC. The salient feature is that the 6400-cores available during this window are efficiently utilized. The green swatch near the middle is composed of grid jobs, surrounded by layers composed of more traditional HPC jobs.

GigaBits per second to and from other LHC collaborators, often via this shared Internet2 link.

HPC @ HCC

While shared high throughput computing (HTC) is a major thrust at HCC, it should be emphasized the majority of computing done at HCC is still traditional high performance computing (HPC), since this is what most HCC researchers currently need to advance their research. The fact that HPC coexists so well with HTC at HCC is strong evidence this model of shared computing can be extended to other locations. Here we briefly highlight several recent high profile findings from the Zeng group at UNL that relies upon HPC.

This summer Xiao Zeng published three high-profile papers, all of which required complex computations completed with substantial usage of HCC. All told, Zeng's group used over 11 million cpu hours over the last year. Zeng and colleagues created a model potassium channel -- a nanoscale sieve -- much like found in almost all living cells. Their computations helped to guide the experimental synthesis of this assembly, which was published in Nature Communications in July.

The next month Zeng's group published in the Journal of Chemical Physics a novel 7-coordinate form of carbon CTi7(2+). Finally, Xiao Zeng published in Science an "ordered amorphous carbon cluster" so hard it can dent diamonds, the hardest known substance. These were examples of computational muscle merged with chemical imagination. HCC was built to accelerate this type of success; HTC has not hindered these advances.

Overall, HCC facilities usage is composed of both Nebraska research and international collaboration. A total of 10 Million cpu

hours are used monthly at HCC between our various machines. Top usage includes groups belonging to Xiao Zeng (UNL Chemistry), Anthony Starace and Stephen Ducharme (UNL Physcis), Clint Rowe (UNL Earth and Atmospheric Science), Renat Sabirianov and Wei-Ning Mei (UNO Physics). These represent traditional computationally intensive processing. Interleaved with these users are various Virtual Organizations (VOs) from the Open Science Grid (OSG), led by CMS, which is a direct collaborator with HCC. Others include local "grid" and "HCC" usage, which are direct usage from UNL scientists, including bioinformatics and mathematics. OSG usage is opportunistic, and only uses HCC machines when capacity is available due to fluctuating demand from NU researchers. While it is expensive to operate computational hardware, the annual operational expenses of HCC are dwarfed by the amount of depreciation of the hardware itself.

HCC changes over the last decade

HCC resources have changed dramatically over the last 13 years. Data prior to 2009 is from the previous organization known as the Research Computing Facility (RCF). The following table gives estimates of the amounts of various cyberinfrastructure resources from several years during this time period.

While details will be different for other locations, the qualitative trends are commonly observed. The relative increase in storage capacity at HCC far surpasses that of the other categories. Data sizes are increasing 33 times as fast as the number of CPU cores, and 150 times faster than the WAN bandwidth needed to move the data. For comparison, it takes a month to move 3 PB of data at a sustained 10 gbps. It is thus

not surprising to see 100 gbps initiatives in the research community. Power usage has increased more slowly than the other hardware increases, although precise numbers are difficult to obtain prior to 2009; it has been roughly constant over that time.

Observations and Guidelines

The experiences listed above lead to several guidelines that aid in the management of HCC.

We need to use what we buy. These pieces of cyberinfrastructure are linked, but improve asynchronously. It is exceedingly difficult, if not impossible, to future-proof any hardware investment. Depreciation, on the other hand, is immediate. Leasing, such as purchasing compute time and storage from Amazon Web Services, is not yet cost effective in most cases. Especially because hardware improves dramatically over time, one can increase the return on investment by buying at regular intervals. This is manageable if one purchases on an annual basis. Space, power and cooling infrastructure tends to outlive any particular hardware purchase regardless.

We should share what we aren't using. By giving opportunistic access to researchers from other organizations, we do not relinquish control of our local resources. We do, however, serve a greater good within the broader scientific community, and utilize otherwise squandered idle resources which are rapidly depreciating. In this process, we have the potential to contribute and gain collaborators. These collaborators may in turn share their resources with us, allowing us to burst usage to greater scales when a situation demands it. There are always

many more resources off campus than those currently possessed on campus.

A data deluge is upon us. While other surrounding hardware is increasing at remarkable rates, the amount of data storage capacity demanded is currently outstripping processing and data transfer advances.

Expert support is essential. If we build it, they may not come. Support personnel can be increased at rates much lower than that of the hardware they maintain and facilitate. By acquiring hardware incrementally and regularly, one buys time for the user community to be trained to effectively exploit resources.

Links to regional and national infrastructure are critical. The scientific advances that have involved HCC over the last several years would not have been possible without collaborations at all scales. Regionally, our network connectivity has been more affordable due to planning and cost-sharing with partners in the Great Plains Network. This has facilitated greatly improved access to Internet2 and ultimately the world. Collaborations with OSG as well as XSEDE have facilitated access to resources but have also provided a broader context for the development and contributions of researchers at HCC. These personal contacts have been as valuable as any hardware access. Finally, the LHC is one of the largest international collaborations in history. HCC has been privileged to participate in this effort, and has recently been able to celebrate with colleagues worldwide at the first major breakthrough this project has produced. Through further collaboration, we look forward to future advances and discovery.