DCC: Data Centric Compute Resource at the NCI-NF

Authors:
{Jie Cai, Muhammad Atif, Michael Chapman, Joseph Antony, David Singleton, Ben Evans}*,
{Peter Strazdins}^

NCI National Facility*, College of Engineering and Computer Science^,
Australian National University, Canberra, ACT 0200, Australia

Abstract:

The NCI National Facility operates Australia's largest academic supercomputer. With increasing storage scalability, we are facing significantly increasing demands of data intensive workloads e.g. processing of climate simulation results towards the IPCC programme, implementation of custom bioinformatics pipelines to aid in immunology research, data parallel processing of satellite imagery.

Aided by investment programs by the Australian government, the NCI NF has invested in both data storage and data centric compute infrastructure (Figure 1). The two major programs are called Research Storage Data Initiative (RDSI) and the National eResearch Collaboration Tools and Resources (NeCTAR). Funding from the RDSI program was sought to expand the underlying data infrastructure, whereas funding from NeCTAR is sought for the compute cloud component. A unified view of the NCI NF's cloud assets showing the relative positioning of RDSI and NeCTAR RC nodes is shown in Figure 1. From the top, Figure 1 depicts user-facing virtual laboratories (VLs) run on IaaS platforms (OpenStack—providing research cloud cycles—and VMWare ESX server—providing mission critical application cycles).

The 'Physical Infrastructure Layer' is physically distributed, initially across two with future expansion to three data centers. This layer provides physical compute (blade servers), a storage subsystem, unified networking (SAN and IP) services across sites (Huxley/Crisp/new NCI Data Centre) and cloud-wide monitoring. Networking to the NCI NF EIF supercomputer, AARNet and other dedicated 10GigE point-to-point links enter the NCI NF cloud at this layer.

A 'Services Layer' runs on top of the distributed physical infrastructure layer, effectively virtualising the location of physical resources, i.e., user facing VLs and compute services are unaware of the physical locality of the compute servers, storage and networking. Mission critical applications are run using VMWare ESX server, while the NeCTAR RC are provided by OpenStack, and, within this same layer, NCI-NF RDSI services will also be run (i.e, DaSH, ReDS services from the RDSI program; Lustre [LAN/WAN][1]; relational databses, HSM access; specialised access to Hadoop HDFS, SciDB and Rasdaman high-performance data management systems). Internal networking will be configured for dual requirements: to sustainable performance levels consistent with a HPC methodology; and to ensure secure access

[1] The NCI-NF RDSI Lustre installation has been architected to allow access to Lustre from the NCI-NF (EIF funded) petascale supercomputer, from mission critical VMs, and from the NeCTAR RC node VMs (i.e. LAN access). If VLs are run remotely and require access to data residing in the NCI-NF RDSI Lustre filesystem, data will be made available using wide-area Lustre (i.e., WAN access).

that provides both the robustness and protection of sensitive information in the cloud.

We note the OpenStack instance will use QDR IB for fast intra-node communication and access to Lustre/HSM resources in aid of data intensive tasks. The NCI NF also has in-house tools to provide cloud-wide accounting for its scheduling/software stack. Finally, end-user VLs (run from either VMWare ESX server or OpenStack) are exposed to users via the 'User-Facing Layer'.

Currently, as a pilot project, we have deployed a virtual data centric compute (DCC) cluster using VMWare ESX. It consists of 64 compute cores across 8 physical machines with 40GB RAM each. They are connected via channel bonded 10GigE interconnects. Different file systems, including online persistent storage, short-term scratch storage and tape storage, are directly accessible from individual virtual compute node with secured read/write permissions. An ANU in-house implementation of OpenPBS (ANUPBS) is used as a resource management tool and user compute job scheduler. All system and user applications are identical with those on the NCI NF peak system (VAYU), including OpenMPI, Intel compilers, MPI profilers, Guassian, NCL and much more.

At the meantime, DCC has been fully utilized to handle data processing jobs for several research communities and groups, including Earth System Grid (ESG), IPCC, immunogenomics group of John Curtin School of Medical Research at ANU etc. The pilot program of DCC has shown great success in terms of meeting users' demands on both flexibility and performance. In the short future, DCC will be expanded to cooperate with OpenStack to launch virtual compute nodes upon requests and over different physical places, including NCI NF data centers, Amazon.com and more.

The pilot DCC cluster is measured using number of benchmarks to evaluate its compute capability, interconnect speed and I/O performance. These benchmarks include SPEC CPU2006, OSU MPI Benchmark suite, an in-house designed I/O performance testing scheme and some well-known real applications. The details of the evaluation will be presented in the final poster.

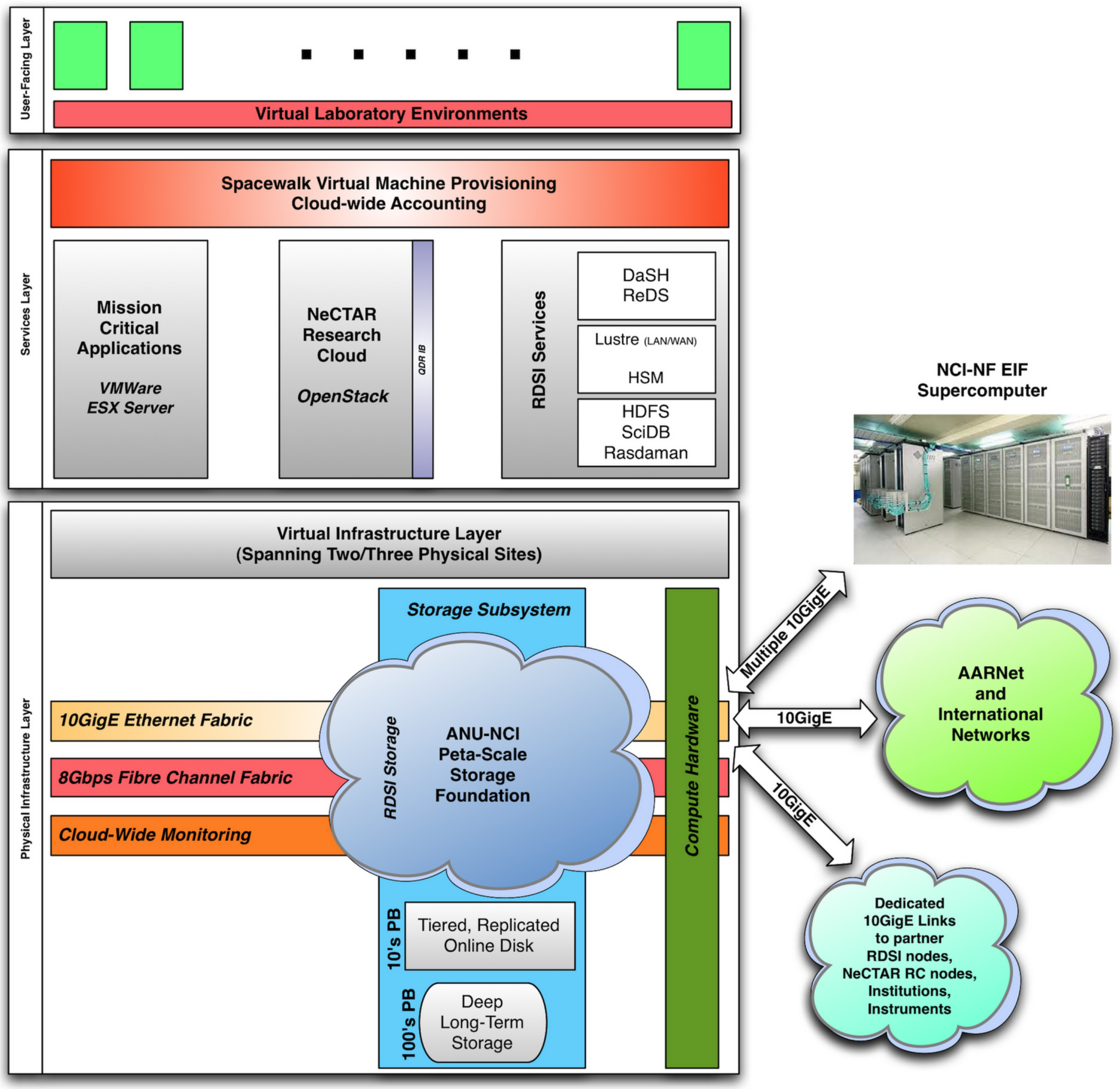

Figure 1: The NCI-NF's Data Cloud Environment