9th Biennial International Conference
& Exposition on Petroleum Geophysics

"HYDERABAD 2012"

P-061

Planning for a HPC Seismic data processing facility

Ajith C, ONGC

Summary

Seismic data processing is a compute intensive exercise. Here we look at components involved in building a high performance computing facility for seismic data analysis.

***Keyword**: Seismic Data Processing, High Performance Computing, Building blocks, Linux, PC Cluster*

Computer systems in Oil exploration

Seismologists were among the first scientists to exploit the capabilities of advanced computing technology. Over fifty years ago in 1962, digital computer was used in field for oil exploration[1]. Mainframe systems marked the entry in mid 1965[2]. By 1968, data search and retrieval system[3] were available for oil exploration. Evolution of standards based computing starting in 1970's. Shortly afterward, computer systems were used for seismological analysis, modeling and visualization.

The field of seismology benefited tremendously and has been dramatically affected from the emergence of high-performance computing. These activities engaged seismologists at the leading edge of technology, spurred the development of new computer technology and enhanced the discovery of energy resources[4]. Due to its very low initial signal-to-noise ratio and the large data size, seismic data processing is extremely demanding in terms of processing power. With advancements in data acquisition, this demand of power is ever increasing.

Mainframes gave way to shared memory Symmetric Multi Processing (SMP) super computers during 90s[5], clients were mostly small time exploration companies. Technology took a u-turn in mid 90s with the introduction of of commercial off-the-shelf (COTS) computers to produce a cost-effective alternative to a traditional supercomputer[6]. Majority of the seismic high performance computing (HPC) data processing centers now use cluster of SMP machines, empowered by best of both worlds.

Today's technology is driven by loosely coupled sub systems, guided by open standards. PC clusters continue to lead HPC benchmarks with advancements every year. PC Clusters continue to gain foothold in latest Top 500[7] world super computing ranking[8]. Futuristic multicore Graphics processing unit (GPU) based parallel computing architecture and OpenCL[9] cross platform frameworks are already available[10]. CUDA powered Nvidia GPU systems[11] are already making large strides in GPU based accelerated technologies. Seismic application vendors have also started their work on these new technologies[12]. They are efficient in terms of real estate, energy and number crunching.

Seismic data processing centers (called DPC from now on) exists through out the world, employing HPC in all stages of data acquisition, analysis and modeling. The goal of this paper is to lay down basic building blocks of a seismic data processing center. This paper is aimed to help data processing managers planning a new DPC.

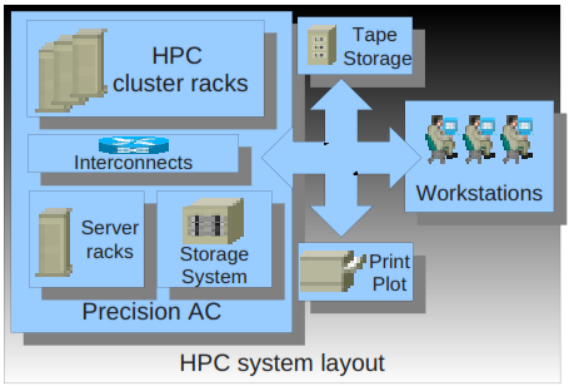

Figure 1: HPC technology building blocks.

RCC, GPS, A & A A Basin, ONGC Ltd, Jorhat.
ajith_c@ongc.co.in

Building blocks of HPC DPC, including hardware, software and other infrastructure are briefed below.

Hardware

There are three major subsystems in HPC. They are Systems, Storage and Communication. We will go through the major building blocks of these subsystems now.

Compute Racks or Blades

Most of the computer hardware majors including national and international companies offer rack mounted compute solutions. Options include energy and space efficient dense blades or 1U/2U/4U form factor high power server/compute nodes. A blade server is a stripped down server computer with a modular design optimized to minimize the use of physical space and energy. Blade servers can equal the performance of its rackables, and the blades are more power efficient and potentially much easier to service. But lack of standardization among blade vendors configurations refrains us from option for blades. In a standard server-rack configuration, 1U (one rack unit, 19"/48cm wide and 1.75"/4.45cm tall) defines the minimum possible size of any equipment. Major integrated cluster vendors[13][14] include IBM, HP, Dell, Oracle/Sun, HCL, Cisco, Fujitsu, SGI. Racks are also supplied by numerous other original equipment manufacturers (OEM) as well.

Rack mounted compute nodes

Modern multi core SMP processors can easily scale to 2/4/6/8/10/12 core 2/4 way SMP architecture on a single node, supporting 512GB shared memory[15]. Form factor range from 1U to 3U, depending on the number of CPU, RAM and internal hard-disks. The most common computer rack form-factor is 42U high. Due to the highly distributed and scalable nature of parallel seismic applications like migration, quad core dual/quad way SMP processor should be sufficient enough. Rather than concentrating on single node's power, application software is designed to make use of distribute compute power. So going for an efficient dense rack node would be advisable. Although we can dedicate DPC HPC system to a application software, it would be efficient to keep the options open. This will allow us to extract maximum efficiency out the HPC, allowing various combination of software and techniques to be used in the same HPC environment. So node specification can be frozen only after merging requirements from all software suites. Most of the integrated cluster vendors have there offering in this segment.

Rack mounted servers

In a HPC DPC, we will have two kinds of servers, IO front-end and application servers. A server(s) can do both the jobs or dedicated single purpose, depending on load conditions and hardware availability. Data path to storage system is usually fibre channel (FC). Due to dense nature of rack nodes, FC is terminated at limited number of servers. These servers pump data to compute nodes as well as workstations which are not on SAN. Front end configuration is depended on the application and/or file system/storage requirements. Again, like compute nodes, most of the integrated cluster vendors have there offering in this segment. Compute nodes and server nodes together will be addressed as racks here after.

CPU options

X86 compatible 64 bit[16] (x86_64) is platform of choice due to backward compatibility and availability of application software. All major chip vendors offer processors in this category under various branding like AMD64, EM64T, x64. Vendors include Intel[17] and AMD[18]. There are neumorous processor models under Xeon, Opteron brands.

Memory

Although nodes can scale up to 512GB or more, application slaves will have a limit to the amount of RAM required. At the same time servers and workstations with caching application servers and interactive jobs can make maximum use of available RAM. So having more RAM is good at the server side, but no at compute node side. DDR-3 remains the technology of choice. DIMM slots from 4GB to 16GB are available now, scaling to 640GB or more. RAM comes as OEM component in the rack node/server.

Local disk

Compute nodes/servers and workstations make use of local disk for caching data files, as scratch disks and also during computation. Though booting from network is possible, operating system (OS) and client components of the application software are locally installed to speed up booting and run-time performance. OS and application client disk requirement range from 1GB to 30GB. Remaining space is used for local compute scratch. It will be interesting to note that a highly tuned compute node OS can be well under 1GB in size. SAS/SATA disks with 2.5" and 3.5" form factor are available. SAS disks with 10K+ rpm are preferred because of high IO and durability criteria. Again disk is also OEM component inside the rack node/server.

Gigabit Ethernet network

TCP/IP remains as a viable choice of communication technology. Fibre channel and InfiniBand provides optimized hand bandwidth interconnects. Gigabit Ethernet network (GbE) was initially proposed in 1999. 10GbE was studied in 2002 and bulk commercial shipping and use started by 2007. 40GbE and 100GbE standards are evolving. For the time being 1/10G at client side and 20/40G backbone can be suitable for of commercial HPC interconnect. High bandwidth interconnect to all servers and workstations are required. Standard based, lossless, Low Latency, data center network convergence ready ethernet fabric technologies are also evolving[19]. Direct IO is possible from workstations due to capabilities of parallel file-systems. Multiple IO paths to racks can also help. Vendors include[20] BLADE Network Technologies, Cisco, Force10 Networks, LG Ericsson, Voltaire etc.

Storage area network switches

Bulk data movement between storage system and clients require a high bandwidth, flexible path. Storage area network (SAN) switch provides this dedicated network. SAN, together with 10GbE acts as the backbone of HPC system. Though SAN or NAS remains as a matter of debate, with lower FC infrastructure costs, SAN can be a performance system to supplement HPC. Alternate technologies like NAS, iSCSI are getting popular in other verticals. But due to sheer raw space requirements and high volume data movement, SAN remains as a good choice. But with 10GB end to end network connectivity, NAS boxes can also offer viable bandwidth. Most of the GbE vendors provide switching infrastructure here. Notable names[21] include Brocade, Cisco, Qlogic etc.

Enterprise class disk storage

Space, space and more space. Outer space or disk space, there is no limit to it. Storage space requirements are growing day by day. We can see terra bytes getting filled in no time. Seismic DPC does not require majority of the enterprise class storage features. Enterprise grade storage vendors include EMC2, NetApp, Hitachi, IBM, HP. It will be interesting to note the emergence of OpenStorage market as alternative to proprietary technologies like Nexenta and few others[22].

Robotic tape drive

Automated (ATL) or Robotic (RTL) tape library is the back-end workhorse for data archival and backups. It provides bulk and long term second tier storage option. Initial data loading will require industry standard drives like IBM3592, LTO3/4. Libraries will help help interim storage, long term vault of final migrated sections and route backups. RTL resource can be optimally shared by HPC data processing and data archival section (DAS) sections. Vendors include IBM, EMC2, HP, Oracle/Sun/StorageTek, Quantum. Although ATLs are becoming increasingly smarter, an external software or data management framework is almost always required to put the library to real work[23].

Interactive Workstations

Workstations are domain experts window to HPC environment. RAM, Disk, CPU and Graphic card requirements are decided as per data size, type of processing jobs, application software requirements etc. Again decision should consolidate multi-vendor multi-application software and multi functional requirements. Ideally a system with 4/8 core with 20/40GB RAM, a GPU with 4GB Video-RAM, 10GbE NIC and Dual port HBA for SAN connectivity would be good. Interactive processing will get good support from dual 32" LCD and 4TB local disk, which can double up as local data store and tire1 backup/archive.

External network hooks

Secure internet access and external network connectivity using VPN are present in most of the DPC environments. Network connectivity will help processing analysts working from remote places, be it home or on tour. It can also help in client interactions. Fire-walled VPN gateway with dedicated ISP hookup will be ideal. Security and access control can be tightened using smart cards or bio- metrics in addition to login credentials.

Precision AC for HPC environment

Precision AC is required to control temperature within ambient limits. Heat generated due to dense nodes need to be controlled to increase hardware life. Specially made curtains and heat flow solutions can reduce energy requirements and carbon footprint. Cooling requirements are normally specified by cluster solution provider. Vendors include Liebert, Emerson, APC, Stulz, Bluestar, Voltas

General Air conditioning

Central Air conditioning for rest of the sub systems is also required. Maintaining dust free conditioned environment will reduce hardware maintenance costs, downtime and increase life of computer workstations and IO sub systems.

Power conditioning and backup

Power conditioning and backup can be achieved using online UPS. Major vendors are Liebert, APC, Sentinel Power, Numeric

Green computing

Day by day, hardware are getting leaner, meaner and greener. Coupled with smart software, HPC DPC can reduce carbon footprints, compared to behemoth super computer environments. Majority of hardware and system software vendors have jumped to this bandwagon, exploring better ways to increase energy efficiency. Evolving standards, certifications like Energy Star, DPC policies and better practices can also add value to this initiative.

System Software

Operating environment

Linux based OS distributions remain as choice of platform here. Options are subscriptions based Redhat and SUSE Enterprise Linux variants or free as in freedom distributions like CentOS, Scientific Linux etc. Though application software vendor can dictate the exact OS and version requirements, we should consider a multi vendor, multi application environment in mind. Most of the time, OS version remains flexible. We can opt for vendor supported variants like RHEL for servers and free options for workstations and compute nodes.

Authentication and management

This include LDAP or NIS. Though LDAP is very scalable, NIS will be easy to manage in a private HPC environment. Both the sub systems are naively available in Linux distributions.

File systems

HPC SAN network enable us to have shared disk file systems[24]. This will help in faster data access at workstations and servers. All FC connected servers and workstations will see the same logical storage units (luns). Available options are GPFS, CIFS, GFS, MPFS, CFS. Racks can be served by native NFS. There are few interesting distributed parallel fault-tolerant file systems including commercial GPFS and free open source GlusterFS as well as LustreFS. Vendor interoperable Parallel NFS (pNFS)[25] is also viable for use now.

Backup

Backup should cover system, user and project data. Once configured correctly, change in system data is not expected. Snapshots can be taken after any major system change. Daily incremental backup policy can be implemented for user and project data. Weekly off-site mirroring of backup can be practiced. Coupled with RTL and HSM (hierarchical storage management) aware applications, we can deploy a multi-tier storage system to augment HPC DPC. Available backup management options include[26] NetWorker, Tivoli, NetVault, NetBackup, and free solutions that include Amanda, Bacula.

Disaster Recovery

Disaster recovery (DR) plan and practice is recommended by audit agencies. Quick recovery from eminent disaster is made difficult due to strict vendor licensing policies, multilevel user & application project structures, access control, huge storage volume etc. Although full fledged DR implementation is not practical in DPC environment, organisation can define the BCP/DRP and practice the same. Backup of predefined processing stages can be preserved for the project's life time. This practice can involve off-site mirroring to take care of some DR aspects.

Other mandatory software

There are some more disconnected software system required for the DPC. Software is required to mange SAN storage and fabric are mostly bundled with hardware. If native multi-pathing is not supported by storage vendor, required software need to be installed for optimum performance. Software is also required to monitor and manage the systems and network. Examples for cluster system tools are xCAT, Ganglia, Nagios, bWatch, ClusterProbe, OVIS. Network tools include Ciscoworks, Solarwinds, OpenNMS, Ethreal, MRTG, Munin, Cacti, Zabbix, Zenoss. Standard job scheduling and accounting software accompanying the HPC solution may not cater to seismic HPC environments. Majority of the application software vendors have their own queue management or no queue management facility. CPU, network and disk IO resource accounting head aches are left to the operations team. To comply with statutory requirements, DPC will need software for payroll and financial accounting, if not on organization's ERP system. Web access require a secure server cloud with content filtering, proxy and firewall. VPN system software will help us manage external access policies.

Application Software

Signal processing

Due to specialized nature of our business, number of vendors providing end-to-end software systems are limited. Some vendors/software in this vertical are CGGVeritas/Geovation/Geocluster, Paradigm/Epos/Echos, SLB-WGC / Omega, PGS /Tensor / CubeManager, Geocenter / SeisUP, Landmark/ProMax. Notable free and open source software like CWP/SeisUnix, Amoco BP/FreeUSP, Conoco Phillips/CPSeis, Stanford/SEPLib also exists in this segment.

Special processing

Companies developing and providing consultancy services with specialized software packages and plugins are on the rise. Concept of software as a service is also catching up faster. Processing/study/analysis includeThousands of software and service companies exists in this segment. Some notable names include Geotomo/Thrustline, Norsar/2D/3D. Majority of the signal processing vendors provides advanced techniques and solutions for AVO, modeling, VSP, 3D3C, inversion etc.

Data archival

Though not strictly under HPC environment, data archival section (DAS) functions as a independent custodian of data. Archival will help in indexed storage of seismic data with all relevant cultural and non-binary accessory information. This data bank is aimed at long term benefit of the organization. Input to HPC processing systems is from from DAS. Intermediate and final output also goes to DAS with proper indexing. An integrated real time database can help in compliance and integrity pacts. Archival will include transcribing (inter media) and trans- coding (inter format) facilities as well. Available E&P open data model and database services-solutions include SLB-Seabed & Finder, Halliburton-Petrobank & Power Explorer, CGGVeritas-PECC & Petrovision, SpaGeo-Tigress, GeoTrace & GeoBrowse, Geplan-I2AM. Open source and open standards based GIS frameworks[27] are available, if customized software development is in plan..

Certifications

Industry certifications can improve regulatory requirements and standards compliance, streamline operational procedures and improve efficiency over a period of time. DPC can go for ISO 9001:2008[28] for Quality management systems, ISO 14001:2004[29] for Environmental management systems, BS OHSAS 18001[30] for Occupational Health and Safety, ISO/IEC

27001:2005[31] for Information technology security techniques and Information security management systems, ISO/IEC 24762:2008[32] for information and communications technology disaster recovery services

Physical infrastructure

So far we were discussing IT infrastructure. Before concluding, we will take a quick look at non-IT, but critical infrastructure requirement for our DPC.

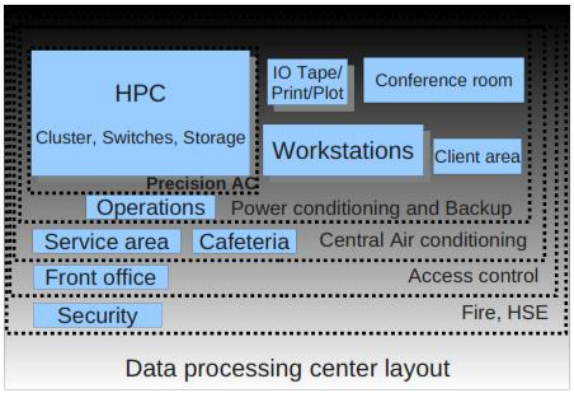

Figure 2: Data processing center layout

As you can see, following functional areas are present in DPC,

Management functions

Front office and security form the outer ring of data center. Cafeteria, support and service areas need to have separate entrance and access control. There can be a client area where discussions and meetings with clients can occur. Conference room with video conference facility can aide in quick meetings and remote sessions.

Technical functions

With in the data-center, we can identify 3 functional areas, namely Systems Administration and Operations area, User area and HPC systems area.

Other subsystems

There are more components in HPC eco system, not directly involved in computing. They demand in-depth multidisciplinary analysis before finalizing the DPC plan[33]. Sub system redundancy requirements should be kept in mind while planning.

Planning for a proper Security system with physical access control systems is must.

ISP network hooks might be required to link the DPC to other DPC sites. Remote office network equipment can help Geo-scientists on the move or at home to remotely connect to DPC systems for monitoring the progress of their jobs.

Layout of the Work stations cum user space and suitable furniture are vital in having a productive environment. Console area for Job monitoring, System health Check etc Is mostly found as part of system administration area.

Some other aspects with long term impact are given below. Many aspects can be outsourced for operational efficiency.

Ambience related topics include Lighting, environment friendly designs, Humidity control, General AC and Precision AC systems

Power supply related topics also come into picture during the initial phases of DPC planning. Some points include Redundant UPS systems, Redundant Power Distribution Units & Transformers, Diesel Generators for standby power, Diesel storage & day tanks, pumps

Few civil engineering related aspects that come into mind include Fire monitoring & suppression, Raised flooring and concealed cabling, Security, Electrical, Fire, Network & Fiber cabling, Cable management planning during construction, possibly structural steel (depending on the UPS, and the location, you may need a beam to spread the load)

Other operations, safety and regulatory requirements might include Insurance, licenses, permits, on-site technical and non technical support staff. Disaster an emergency planning, shutdown/boot up procedures Emergency contact nos. Etc also comes into picture.

Future

Requirements for a dedicated HPC seismic data center is explained in this paper. Although most of the traditional E&P companies prefer to have such a processing center, there is a wind of change. Integrated Data Processing, Interpretation and Reservoir studies and visualization practices are getting popular. Sharing the HPC resources and multidisciplinary approach can improve efficiency and productivity. Hosted compute facility is available now. This will help new exploration companies to outsource the computing component to external agency, who can manage the operations part. Similar hosted seismic services may also catch up soon. HPC in the Cloud[34] and Nvidia-CUDA[11] powered multicore super computer by your desk side is not distant dream. Provided enough disk storage and network bandwidth are made available, entire work flow can be executed from your deskside supercomputer[35].

Acknowledgments

Author had attempted to document what he understood of building blocks of a Seismic data processing facility. Author would like to take this opportunity to thank ONGC management, Head-RCC, Jorhat in particular for the encouragement extended by him in this endeavor. Valuable time, comments and inputs given by Dr. JVSSN Murty and Dr. K. Ramamurthy, both from Database, A & A A Basin, ONGC, Jorhat and Mr. K. Subramanyam, Mr. Alok Roy, Mrs. BVL Suhasini, Mr. T. Madhukar, Mrs. Shyamali Saha, colleagues at Software Group, RCC, GPS, Jorhat are highly acknowledged. Author had attempted to recollect and record his experience at various computing facilities of ONGC spread across the country. Author gratefully remembers his association with previous assignment locations and teams.

Views, suggestions and comments expressed are solely by the author and ONGC Ltd is in no way responsible or liable for the content.

Bibliography

[1] TI, Digital Field System, http://www.ti.com/corp/docs/company/history/timeline/gps/1960/docs/62-field_crews.htm

[2] IBM, Mainframe, http://www.ibmsystemsmag.com/mainframe/trends/z-os/Early-Mainframe-Aids-Oil-Exploration/

[3] D. L. Stauft, Computer applications in an oil-exploration company, http://bcpg.geoscienceworld.org/cgi/content/abstract/16/1/64

[4] ISC, High Performance Computing in Seismic Processing & Oil Exploration, http://www.supercomp.de/isc09/Program/At-a-Glance/High-Performance-Computing-in-Seismic-Processing-Oil-Exploration

[5] Cray, Cray SMP Super computers, http://en.wikipedia.org/wiki/Cray

[6] Wikipedia, Computer cluster, http://en.wikipedia.org/wiki/ Computer_cluster

[7] top500.org, TOP500 List of World's Supercomputers, http://www.top500.org/

[8] BioInform, PC Clusters Continue to Gain Foothold, http://www.genomeweb.com/informatics/pc-clusters continue-gain-foothold- latest-top-500-supercomputing-ranking

[9] Wikipedia, Open Computing Language, http://en.wikipedia.org/wiki/OpenCL

[10] NVIDIA, CUDA-Accelerated Applications, http://www.nvidia.co.in/object/cuda_app_tesla_in.html [11] Wikipedia, Compute Unified Device Architecture, http://en.wikipedia.org/wiki/CUDA

[12] Bernard Deschizeaux, Imaging Earth's Subsurface Using CUDA, http://http.developer.nvidia.com /GPUGems3/gpugems3_ch38.html

[13] linuxhpc.org, Cluster Quote Request Form, http://www.linuxhpc.org/pages.php?page=Quote

[14] hclinfosystems.in, HCL-HPC SOLUTION SPECTRUM, http://www.hclinfosystems.in/hpc.htm

[15] IBM System specs, IBM System x3755 M3, http://www.03.ibm.com/systems/x/hardware/rack/x3755m3/index.html

[16] Wikipedia, x86-64, http://en.wikipedia.org/wiki/X86-64

[17] intel.com, SERVER PROCESSORS, http://www.intel.com/ products /server/processor/index.htm

[18] amd.com, AMD Products for Servers, http://www.amd.com/US/PRODUCTS/SERVER/Pages/server.aspx

[19] IETF, Lossless and data center network-convergence-ready Ethernet switching, http://www.hpcwire.com /hpcwire/2009-03-24/ethernet_fabrics_offer_way_forward_ for_seismic_applications.html
[20] ethernetalliance.org, Affiliation, http://www.ethernetalliance.org/about_the_alliance/member _roster
[21] Wikipedia, List of Fibre Channel switches, http://en.wikipedia.org/wiki/List_of_Fibre_Channel_switch es
[22] linuxclusters.com, Alphabetical List of Cluster Software/Hardware Vendors, http://www.linuxclusters.com /vendors.html
[23] it.toolbox.com/, Automated Tape Libraries, http://it.toolbox.com/wiki/index.php/Automated_Tape_Libr aries
[24] Wikipedia, Shared disk file systems, http://en.wikipedia.org/wiki/List_of_file_systems#Shared_d isk_file_systems
[25] Multivendor consortium, Parallel NFS, http://www.pnfs.com/
[26] Wikipedia, List of backup software, http://en.wikipedia.org/wiki/List_of_backup_software
[27] Wikipedia, List of geographic information systems software, http://en.wikipedia.org/wiki/List_of_geographic_ information_systems_software
[28] ISO, Quality management systems, http://www.iso.org/iso/catalogue_detail?csnumber=46486
[29] ISO, Environmental management systems, http://www.iso.org/iso/catalogue_detail?csnumber=31807
[30] OHSAS, HEALTH and SAFETY, http://www.osha-bs8800-ohsas-18001- health-and-safety.com/
[31] ISO, IT Security techniques and management systems, http://www.iso.org/iso/catalogue_detail?csnumber=42103
[32] ISO, ICT disaster recovery, http://www.iso.org/iso/catalogue_detail? csnumber=41532
[28] webhostingtalk.com forums, Setting up our own data center, http://www.webhostingtalk.com/showthread.php? t=594681
[29] Vedaprakash Subramanian et.al., Azure Use Case Highlights Challenges for HPC Applications in the Cloud, http://www.hpcinthecloud.com/hpccloud/2011-02-21/ azure_use_case_highlights_challenges_for_hpc_application s_in_the_cloud.h tml
[30] NVIDIA, High Performance Computing - Supercomputing with Tesla GPUs, http://www.nvidia.com/object/tesla_computing_solutions.ht ml