ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 1, Issue 9, November 2012

GRID COMPUTING – AN ALTERNATIVE TO HPC

Sanjeev Narayan Bal

Asst.Prof.Dept.of Comp.Sc.
TACT, Bhubaneswar

***Abstract*-- Grid Computing delivers on the potential in the growth and abundance of network connected systems and bandwidth: computation, collaboration and communication over the Advanced Web. At the heart of Grid Computing is a computing infrastructure that provides dependable, consistent, pervasive and inexpensive access to computational capabilities. By pooling federated assets into a virtual system, a grid provides a single point of access to powerful distributed resources. Researchers working to solve many of the most difficult scientific problems have long understood the potential of such shared distributed computing systems. Development teams focused on technical products, like semiconductors, are using Grid Computing to achieve higher throughput. Likewise, the business community is beginning to recognize the importance of distributed systems in applications such as data mining and economic modeling. With a grid, networked resources -- desktops, servers, storage, databases, and even scientific instruments -- can be combined to deploy massive computing power wherever and whenever it is needed most. Users can find resources quickly, use them efficiently, and scale them seamlessly. We can ensure grid computing an alternative to HPC**

***Keywords:* Grid, HPC, Parallel computing, Sensor network.**

I. INTRODUCTION

A. The Grid Concept

The term 'grid' is variously used to describe a number of different, but related, ideas, including utility computing concepts, grid technologies, and grid standards. In this paper the term 'Grid is used in the widest sense to describe the ability to pool and share Information Technology (IT) resources in a global environment in a manner which achieves seamless, secure, transparent, simple access to a vast collection of many different types of hardware and software resources, (including compute nodes, software codes, data repositories, storage devices, graphics and terminal devices and instrumentation and equipment), through non-dedicated wide area networks, to deliver customized resources to specific applications.

At the most general level Grid is independent of any specific standard or technology. Any practical grid is realized through specific distributed computing technologies and standards that can support the necessary interoperability. Today, there are no universally agreed grid standards, but there are freely available, open source and proprietary grid technologies that implement emerging standards recommendations. Separate web services standards are also emerging which have many grid-like capabilities. Indeed grids are already being built by integrating and enhancing web standards technology.

B. Practical Realizations

Practical grids are generally described in terms of layers (see Fig 1). The lowest layers (the 'platform') comprise the hardware resources, including computers, networks, databases, instruments, and interface devices. These devices, which will be geographically distributed, may present their data in very different formats, are likely to have different qualities of service (e.g. communication speeds, bandwidth) and are likely to utilize different operating systems and processor architectures. A key concept is that the hardware resources can change over time - some may be withdrawn, upgraded or replaced by newer models, others may change their performance to adapt to local conditions - for example restrictions in the available communications bandwidth.

The middle layers (sometimes referred to as 'middleware') provide a set of software functions that 'buffer' the user from administrative tasks associated with access to the disparate resources. These functions are made available as services and some provide a 'jacket' around the hardware interfaces, such that the different hardware platforms present a unified interface to different applications. Other functions manage the underlying fabric, such as identification and scheduling of resources in a secure and auditable way. The middle layer also provides the ability to make frequently used patterns of functions available as a composed higher-level service using workflow techniques.

The highest layers contain the user 'application services'. Pilot projects have already been carried out in user application areas, such as life sciences (e.g. computational biology, genomics), engineering (e.g. simulation and modeling, just in time maintenance) and healthcare (e.g. diagnosis, telematics). These services could include horizontal functions such as workflow (the linkage of multiple services into a single service), web portals, data visualization and the language/semantic concepts appropriate to different application sectors.

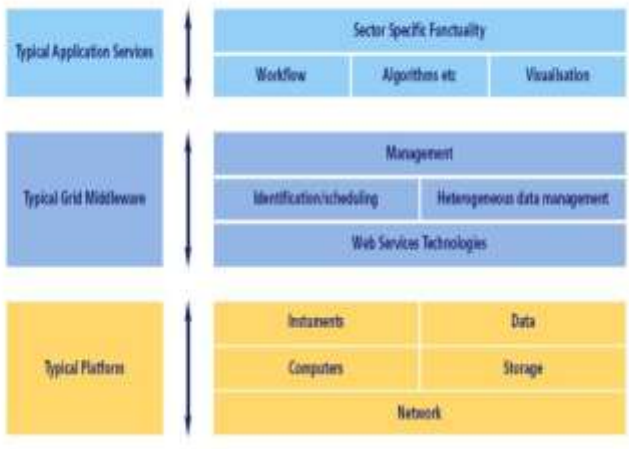

Figure 1: Simplified Grid Architecture

C. Grid Developments and Deployment

A key issue facing the industry is the timing and mode of deployment of Grid technology to ensure that it is sufficiently mature to deliver the expected business benefits. There is emerging evidence that the technology can achieve significant operational benefits (e.g. in telemedicine), improvements in performance (e.g. in climate modeling and genomics) and a significant reduction in costs. Nevertheless, current grid technologies are not yet viewed as sufficiently mature for industry scale use, and remain largely unproven in terms of security, reliability, scalability, and performance.

Short term

For the short term (within the next two years), Grid is most likely to be introduced into large organizations as internal 'Enterprise grids', i.e. built behind firewalls and used within a limited trust domain, perhaps with controlled links to external grids. A good analogy would be the adoption into business of the Internet, where the first step was often the roll out of a secure internal company 'Intranet', with a gradual extension of capabilities (and hence opportunity for misuse) towards fully ubiquitous Internet access. Centralized management is expected to be the only way to guarantee qualities of service. Typically users of this early technology will be expecting to achieve IT cost reduction, increased efficiency, some innovation and flexibility in business processes. At the same time the distinction between web services and grid services is expected to disappear, with the capabilities of one merging into the other and the interoperability between the two standards being taken for granted.

Medium Term

In the mid term (say a five year timeframe) expect to see wider adoption – largely for resource visualization and mass access. The technology will be particularly appropriate for applications that utilize broadband and mobile/air interfaces, such as on-line gaming, 'visualization-on-demand' and applied industrial research. The emphasis will move from use within a single organization to use across organizational domains and within Virtual Organizations, requiring issues such as ownership, management and accounting to be handled within trusted partnerships. There will be a shift in value from provision of computer power to provision of information and knowledge. At the same time open standards based tooling for building service oriented applications are likely to emerge and Grid technology will start to be incorporated into off-the-shelf products. This will lead to standard consumer access to virtualized compute and data resources, enabling a whole new range of consumer services to be delivered.

Long term--In the longer term, Grid is likely to become a prerequisite for business success -central to business processes, new types of service, and a central component of product development and customer solutions. A key business change will be the establishment of trusted service providers, probably acting on a global scale and disrupting the current supply chains and regulatory environments.

II. Grid Computing Advantages

A. Grid Computing Business Advantages

As described, a grid is essentially a set of computing resources shared over a network. Grids differ from more traditional distributed systems, such as the classic n-tier systems, in the way its resources are utilized. In a conventional environment, resources are dedicated: a PC or laptop has an owner, and a server supports a specific application. A grid becomes useful and meaningful when it both encompasses a large set of resources and serves a sizable community. The large set of resources associated with a grid makes it attractive to users in spite of the overhead (and the complexity) of sharing the resource, and the grid infrastructure allows the investment to be shared over a large community. If the grid were an exclusive resource, it would have to be a lot smaller for the same level of investment.

In a grid environment, the binding between an application and the host on which it runs begins to blur: the execution of a long-running program can be allocated to multiple machines to reduce the time (also known as wall clock time or actual time) that it takes to run the application. Generally, a program designed to run in parallel will take less time to run as more nodes are added, until algorithmic or physical bottlenecks develop or until the account limits are reached.

Two assumptions must hold for an application to take advantage of a grid:

• Applications need to be re-engineered to scale up and down in this environment.

• The system must support the dynamic resource allocation as called by applications.

As technology advances, it will become easier to attain both these conditions, although most commercial applications today cannot satisfy either of them without extensive retrofitting[1].

B. Parallel-Computing Grids Productivity Advantages

Another usage model associated with grids is parallel computation[2]. For instance, if it takes one server-node 10 minutes to update 100,000 records, 10 nodes working together (that is, in "parallel") could theoretically do the same

job in one minute. In practice, of course, the time required would be somewhat more than a minute due to overhead:

the input/output (I/O) subsystem may experience interference with 10 nodes doing simultaneous updates, there might be data dependencies, and one processor might have to wait until another is finished. Nevertheless, parallel processing would reduce the time required to perform the work. In some cases, wall-clock time is of primary importance; for instance, a weather simulation done for forecasting purposes needs to be completed on a deadline. If these calculations can be accelerated by applying more CPUs, even if the CPUs interfere with each other, the reduction in execution time can make the difference between success and failure at meeting the deadline.

A similar dynamic applies to simulation and analysis jobs in engineering shops, albeit less dramatically. Because of the potential savings in worker time involved, it is enormously valuable to be able to run jobs that take several CPU hours in a few minutes of clock time. Because design is an iterative process, detecting a flaw more quickly can equate to significant savings in terms of workers' time, increasing productivity. In the late phases of a design cycle, parametric runs (that is, similar runs with slightly different data) may be necessary. With a job that takes eight hours to run on one CPU, a one-CPU workstation running for an entire month will yield about 100 data points. If an unexpected flaw is discovered in the data at the end of that month, and the run needs to be repeated, the project essentially slips by a month. If, instead of one CPU, 100 CPUs can be applied to the same problem in a grid environment, it is very likely that the computation will not be done 100 times faster—perhaps just 25 times faster. Thus, the grid system might yield one data point every 20 minutes or so (at 25 percent efficiency). Furthermore, let us assume that a grid with 4,000 nodes is available. In this case, 40 jobs can be launched in parallel and, hence, the team might be able to deliver the 100 data points in one hour. The productivity implications of being able to do a month's work in one hour are epochal. It might mean saving the production time of a $100M movie by a few weeks and a few million dollars through the use of parallel rendering engines, or the ability to base real-time quotes on complex derivative securities calculations.

C. Grid Computing Technology Transitions

Grid computing has strong HPC roots, perhaps because the early demands of HPC dictated that some of the solutions, technologies and usage models associated with grid computing be investigated in an HPC context first. This is the case with cycle scavenging and distributed parallel computation. Technology developments like multi-core CPUs may become forced functions for the pervasive use of multithreaded and parallel programming techniques that have been in use in the HPC space for more than 20 years, both for grid computing and in other areas of the computing industry. A quantum jump in the beneficial impact of grid computing will take place when the grid gets adopted in a broader context, including the enterprise and consumer spaces.

The architectural advances that are taking place today, and which will continue to develop over the coming years, will set the stage for widespread adoption of grid computing among relatively small users. This sphere of technology, which is currently widely associated with government and university research environments and the largest corporations, may become well within the reach of all businesses by the end of the decade. The parallel distributed computation enabled by this model will enable businesses to undertake computationally intensive operations such as sophisticated rendering and analysis that would otherwise be impossible for them to do directly.

III. Wireless Grid Computing

Wireless grid architectures can be broadly classified into the following four categories [3,4] based on the devices predominant in the grid and the relative mobility of the devices in the grid.

- Fixed Wireless Grids
- Mobile or Dynamic Wireless Grids
- Ad Hoc grids
- Sensor Network Grids

A. Fixed Wireless Grids

The wireless grid extends grid resources to wireless devices of varying sizes and capabilities such as sensors, laptops, special instruments, and edge devices, where these devices are usually static. In wireless grids, wireless devices can act as real grid nodes where part of data processing and storage is taking place. In a special type of wireless grid, all wireless devices are considered pure access devices without processing or storage capabilities [4]: required resources are obtained from a wired, resource-rich backbone arid. Many technical concerns arise when integrating wireless devices into a grid. These include low bandwidth and high security risks, power consumption, and latency. So. several communities, including the Interdisciplinary Wireless Grid Team (www.wirelessgrids.net) are exploring these new issues to ensure that future grid peers can be wireless devices.

Mobile Or Dynamic Wireless Grids: Mobile grids make grid services accessible through mobile devices such as PDAs and smart phones. As the processing power and other capabilities of these mobile devices increases, researchers and commercial organizations are discovering new ways to use and share their resources. When the large numbers of available wireless devices is considered, the potential of these dynamic ad hoc connections becomes vast [5]. Through wireless grid connections, these devices are able to connect to the Internet, provide peer-to-peer networking, take advantage of the resources of wired grid networks and make their own resources available to the wired grids [3]. In emergency situations, such as during natural disasters and on battle fields, wireless mobile devices might be the only available communication and computation services. The mobile devices integrated into grid systems can actively participate and provide computational or data services [6]. These mobile devices also serve as an interface to a stationary grid for sending requests and receiving results . Sometimes this approach is labeled mobile access to grid infrastructure, or simply mobile access grids. Recently, researchers have made numerous efforts toward establishing mobile grids. Researchers have proposed various techniques for implementing the mobile grid vision, including centralized and P2P structure, intelligent mobile agents, mobile grid middleware, and many more [6]-[8]. Existing mobile grid projects include Akogrimo (www.mobilegrids.org). and MADAM (www.inteiinedia.iiio.no/display/madam/Home).

B. Ad Hoc Grids

Mobile ad hoc network consists of devices with a high degree of heterogeneity. These mobile devices range from relatively powerful computing systems carried by a vehicle, to very tiny, low-power sensors that can be implanted in the human body. Although they may know little about the identities and capabilities of each other, a group of mobile devices are able to organize a highly dynamic and infrastructure-less ad-hoc network, in which nodes can communicate in a hop-by-hop manner. We have to integrate the resource aggregation model of grid with mobile ad hoc networks, so as to build a mobile ad-hoc Grid platform that can be instantly constructed anytime, anywhere [9]. Having been constructed from a group of mobile devices, an ad-hoc grid would allow the networked devices to accomplish a specific mission that maybe beyond an individuals computing or communication capacity. Examples of applications of mobile ad-hoc grids can be disaster management, wildfire fighting, and e-healthcare emergency, etc. The general attributes of mobile ad hoc networks can be extended to ad hoc grids and hence ad hoc grids are attributed by bandwidth and power constraints, multi-hop delivery, network partitioning, and infrastructure unpredictability.

C. Sensor Network Grids

Sensor networks are composed of tiny devices that are generally dedicated to a single purpose. Though each device in the grid is dedicated to a single purpose, the grid itself may be comprised of a number of different types of devices in order to accomplish the goals of the particular grid. As described in [10], wireless sensor networks integrate detection, processing and communication into the grid. The sensors taking part in a sensor grid may be stationary once they have been deployed [11] or they may be mobile. Sensor networks are currently in use monitoring environmental factors such as temperature or humidity change, motion and light intensity. The developments in sensor networks is leading to advances in agriculture, physical building security, firefighting, warfare and a number of industrial areas. Wireless sensor network that can be deployed in residential and commercial buildings to monitor human presence and turn off the lights when no people are detected , is under development by Dust Networks of Berkeley. California. RFID technology to track products from manufacture through distribution and delivery and potentially, to the consumers shopping cart leaving the store , is developed.

IV. CONCLUSION

Points of Comparison With HPC--

Grid Computing, a fast-maturing technology, has been called the "silver bullet" that will address many of the financial and operational inefficiencies of today's information technology infrastructure. After developing strong roots in the global academic and research communities over the last decade, Grid Computing has successfully entered the commercial world. It is accelerating product development, reducing infrastructure and operational costs, leveraging existing technology investments and increasing corporate productivity. Today, Grid Computing offers the lowest cost high-throughput solution, enabling companies to migrate from expensive HPC systems.

(1)Leveraging Existing Hardware Investments & Resources

Grids can be deployed on an enterprise's existing infrastructure, including the multitude of desktops and existing servers, thereby mitigating the need for investment in new HPC systems and other hardware.

(2) Reducing Operational Expenses

The operational expenses of a Grid Computing deployment are 73% less than for comparable HPC based solutions. Many of the existing cluster solutions are based on open source cluster management software that is complex and unsupported. Operational expenses associated with these deployments have been so high that many enterprises are being forced to outsource management of HPC systems to the suppliers themselves.

Additionally, both small and large enterprise grids are being deployed in as quickly as two days - with little or no disruption to operations. Cluster system deployments on the other hand are taking 60-90 days, in addition to the days to configure and deploy the applications. Deployments may take longer if the existing enterprise data center is out of capacity. Additionally, HPC installations in data centers can cause substantial disruptions and potential downtime. In fact, 54% of data center site infrastructure failures are coincident with human activities .

(3) Creating a Scalable & Flexible Enterprise IT Infrastructure

Grid Computing allows companies to add resources linearly based on real-time business requirements. These resources can be derived from within the enterprise or from utility computing services. Never again do projects have to be put on hold from lack of computational capacity, data center space or system priority. The entire compute infrastructure of the enterprise is available for harnessing. Grid Computing can help bring about the end of departmental silos by exposing computational assets curtained by server huggers and bureaucracy. Yet, while departments will be making their resources accessible to the whole enterprise, the right Grid Computing solution can allows them to maintain local control.

(4) Accelerating Product Development, Improving Time to Market and Raising Customer Satisfaction

Grid Computing has a direct impact on the top line by accelerating product development at enterprises and helping bring product to market quicker. The dramatic reduction in, for example, simulation times can get products completed quickly. This also provides the ability to perform much more detailed and exhaustive product design - since the computational resources brought to bear by the grid can quickly churn through the complex models and scenarios to detect design flaws.

(5) Increasing Productivity

Enterprises that have deployed Grid Computing are seeing tremendous productivity gains.

REFERENCES

[1] HPC and Intel® Cluster Tools Intel® Developer Forum ttp://softwareforums.intel.com/ids/board?board.id=HPC) is a discussion board for discussing technical issues related to High Performance Computing with industry peers and Intel experts.

[2]Erik Brynjolfsson, "The IT Productivity GAP." Optimize magazine, Issue 21, July 2003. (http://ebusiness.mit.edu/erik/Optimize/pr_roi.html)

[3] Sanjay P. Ahuja. Jack R. Myers, A Survey on Wireless Grid Computing, The Journal of Supercomputing. 37, 321, 2006.

[4] Heba Kurdi, Maozhen Li, and Hamed Al-Raweshidy. A Classification of Emerging and Traditional Grid Systems, IEEE Distributed Systems Online, vol. 9, no. 3, 2008, art. no. 0803-3001.

[5]T. Phan. L. Huang, and C. Dulan. Challenge: Integrating Mobile Wireless Devices into the Computational Grid. Proc. 8th ACM Ami. Intl Conf. Mobile Computing and Networking, ACM Press, 2002, pp. 271278.

[6]K. Ohta et al.. Design and Implementation of Mobile Grid Middleware for Handsets, Proc. of 11th Intl. Couf. Parallel and Distributed Systems (ICPADS 05) ,IEEE CS Press, 2005, pp. 679683.

[7] L. S. Lima et al.. Peer-to-Peer Resource Discoveiy in Mobile Grids, Proc. 3rd Intl Workshop Middleware for Grid Computing (MGC05) ,ACM Press, 2005, pp. 16

[8] S. Kurkovsky. Bhagyavati, and A. Ray. Modeling a Grid-Based Problem-Solving Environment for Mobile Devices, J. Digital Information Management, vol. 2, no. 2, 2004, pp. 109114.

[9] Z. Li, L.Sun and E. C. Ifeachor. Challenges of Mobile Ad-hoc Grids and their Applications in e-Healthcare. Proc. of the 2nd Intl. Conference on Computational Intelligence in Medicine and Healthcare (CIMED' 2005) July 2005, Lisbon, Portugal

[10] G. Xing. C. Lu. R. Pless. and Q. Huang. On greedy geographic routing algorithms in sensing covered networks, Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing ,May 2004.

[11] M. Miller, and N. Vaidya. Power save mechanisms for multi-hop wireless networks. Proceedings of the First International Conference on Broadband Networks (BROADNETS04),Oct. 2004

Author Biographies

Sanjeev Narayan Bal
MBA, MCA MTech(CSE)
Assistant Professor (CSE)
Trident Academy of Creative Technology
Bhubaneswar
Odisha
Life Member – ISTE