Resource virtualization and optimization via Grid and Cloud Computing

Moon J Kim
IBM Senior Technical Staff Member/Chief Architect/Master Inventor
moonkim@us.ibm.com
## Grid Computing

### Grid Motivations

- **Accelerate Business Processes**
  - Faster, more accurate decision making
  - Productivity and Collaboration
  - Access to distributed data, information insight
- **IT Optimization**
  - Improve efficiency and cost structure

### Grid Focus Areas

<table>
<thead>
<tr>
<th>Industry/Sector</th>
<th>Focus Area</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Industrial</td>
<td>Public Financial</td>
<td>Create large-scale IT infrastructures to drive economic development and/or enable new government services</td>
</tr>
<tr>
<td>Industrial</td>
<td>Industrial</td>
<td>Optimize computing and data assets to improve utilization, efficiency and business continuity</td>
</tr>
<tr>
<td>Financial Industrial</td>
<td>Financial Industrial Public</td>
<td>Enable faster and more comprehensive business planning and analysis through the sharing of data and computing power</td>
</tr>
<tr>
<td>Financial</td>
<td>Public</td>
<td>Share data and computing power, for computing intensive engineering and scientific applications, to accelerate product design</td>
</tr>
<tr>
<td>Research and Development</td>
<td>Public Industrial</td>
<td>Accelerate and enhance the R&amp;D process by enabling the sharing of data and computing power seamlessly for research intensive applications</td>
</tr>
<tr>
<td>Engineering and Design</td>
<td>Industrial</td>
<td>Access to distributed data, information insight</td>
</tr>
<tr>
<td>Business Analytics</td>
<td>Government Development</td>
<td>Enable faster and more comprehensive business planning and analysis through the sharing of data and computing power</td>
</tr>
<tr>
<td>Enterprise Optimization</td>
<td>IT Optimization</td>
<td>Optimize computing and data assets to improve utilization, efficiency and business continuity</td>
</tr>
</tbody>
</table>

---

*Graph and images are not transcribed.*
An on demand operating environment is an integrated infrastructure aligned to business goals and processes in a resilient and secure manner.

Results: Simplification and Optimization of IT to meet the needs of the business responsively.
IBM is the only solution provider that can help clients realize value from the full spectrum of grid computing solutions.
Realizing Increasing Levels of Business Value

- Application Acceleration
  - Time to Results
  - Higher Quality

- Aggregating Information
  - Business Insight and Collaboration

- Infrastructure Flexibility
  - Enabling Business Resiliency

- IT Simplification for Enterprise Optimization
  - Asset Utilization
  - Workload Prioritization
IBM’s Business Drivers

- Competitive pressure to decrease time to market
  - Grids can help reduce cycle time, improve quality, and enable more effective collaboration
- Many business processes require long compute time
  - Grids can cut processing time through effective management of heterogeneous resources
  - Reduced compute time allows additional iterations for greater accuracy and/or reduced cycle time
- Reduce total cost of ownership (TCO) for traditional data center resources
  - Grids help drive down TCO through optimized use of existing compute resources and centralized system management
Grid Value Summary

Business Value
- Improve Operating Efficiency/ROI
- Reduce Capital Expenses
- Accelerate Business Processes
- Enhance Employee Productivity
- Quickly Adapt to Changing Requirements

Technical Value
- Improve Asset Optimization
- Integrate Heterogeneous Resources
- Enable Data Access, Integration and Collaboration
- Strengthen Redundancy and Resiliency
- Quickly Respond to Variable Demands

Staffing Value
- Embracing "hot" technology attracts best technical talent.
- College I/T students involved with Grid technologies
- Next generation distributed computing

Marketing Value
- Grid is open, a key part of on demand
- Leverages the full range of IBM products and services (hardware, software and services)
- Grid is solutions focused
- Grid has mindshare in the marketplace
How can I simplify my IT infrastructure if my complexity is accelerating?

…it may be easier than you may think, it is about simplifying your environments to align to your business goals. This means making your IT environments interoperable, fully integrated; and automated for greater efficiency, productivity and service quality.
Today’s complex infrastructure create challenges

- Management of complex, heterogeneous environments too hard
- IT asset utilization is too low and no interoperability in multiplatforms
- Privacy, security and business continuity
- Swamped by the proliferation of technology and platforms to support
- Operational speed is too slow; IT flexibility too limited
- Inability to manage the infrastructure seamlessly
# Server / Storage Utilization

<table>
<thead>
<tr>
<th></th>
<th>Peak-hour Utilization</th>
<th>Prime-shift Utilization</th>
<th>24-hour Period Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mainframe</strong></td>
<td>85-100%</td>
<td>70%</td>
<td>60%</td>
</tr>
<tr>
<td><strong>UNIX</strong></td>
<td>50-70%</td>
<td>10-15%</td>
<td>&lt;10%</td>
</tr>
<tr>
<td><strong>Intel-based</strong></td>
<td>30%</td>
<td>5-10%</td>
<td>2-5%</td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>N/A</td>
<td>N/A</td>
<td>52%</td>
</tr>
</tbody>
</table>

*Source: IBM Scorpion White Paper: Simplifying the Corporate IT Infrastructure,*
Utilization-week (AFS)

For data displayed here: Average = 20.00% Peak = 43.00%
Optimize IT Through Infrastructure Management: Virtualization and Automation

**Infrastructure Management**
Sense and respond to changes, manage and optimize for the needs of the business via Automation and Virtualization

**Virtualization**
- Build enterprise-wide application fabrics simplifying and exploiting internal and external resources

**Automation**
- Increase flexibility and reduce costs by automating your IT best practices and extend across process disciplines and resources
Virtualization is a fundamental imperative

“Virtualization is the process of presenting computing resources in ways that users and applications can easily get value out of them, rather than presenting them in a way dictated by their implementation, geographic location, or physical packaging. In other words, it provides a logical rather than physical view of data, computing power, storage capacity, and other resources.”

Jonathan Eunice, Illuminata Inc.

...Virtualization is far more than just partitioning or single products
Virtualization - levels of technologies

Create multiple images of a resource within a physical resource
Virtual Memory - LPAR’s – Virtual Machines – Logical Channels

Create functions and facilities which appear to be real although they do not exist within the physical resource
Virtual Networks, Hipersockets, Virtual Disk and Caches

Pool multiple separate “distributed” resources so they appear as a single resources from the user point of view
Clusters – GSA Grid - SAN Volume Controller - SAN File Sharing

Create SERVICE’s using open interfaces and standards, where the services are “unaware of” and “have no dependency” on the underlying “distributed”, “heterogeneous” and “shared” physical infrastructure
Convergence of Systems Management across Enterprise
Enterprise-wide Grid of Resources
WebServices Resource Framework (WSRF)
Virtualization Capabilities - the foundation

- **Servers**
  - Virtual Resource
  - Server Virtualization
    - PR/SM
    - Dynamic LPARs
    - MicroPartitioning
    - zVM, VMware
    - Virtual I/O
    - Physical Partitions
  - Shared Infrastructure
    - Blades
  - Clustering
    - Parallel Sysplex
    - HACMP
    - Linux Clusters
  - Workload Mgmt.
    - Work Load Managers
    - Intelligent Resource Dir.
    - Partition Load Manager
    - Enterprise WLM
    - Resource & Goal based
    - Policy based
    - System scope
    - End-to-End scope

- **Storage**
  - Total Storage Virtualization
    - GSA Grid
    - SAN Block Virtualization
    - Common File system
    - IBM & non-IBM
  - Storage Server Virtualization
    - POWER5 partitioning
  - SAN Volumes
    - Storage Pools
    - Centralized mgnt
  - Total Storage Management
    - Manage storage according to policy

- **Distributed Systems**
  - GRID
    - Globus Toolkit
    - IBM OGSA Toolbox
  - Virtual Engine is not here yet
  - Server Allocation for Web Application Servers
    - Computation heavy, parallel applications
    - Manage multiple applications across multiple server clusters
    - ISV Grid middleware
    - Provide services such as data services, scheduling, etc

- **Network**
  - Hipersockets/ Virtual Ethernet
    - Optimized inter-partition communications, virtual network
  - VLANs
    - Isolate/prioritize traffic on shared network, 802.1
  - Differentiated Services
    - Prioritize network traffic
    - Network QOS, IP TOS
  - Vendor Alliances
    - Cisco

- **Shared Infrastructure**
  - Blades
  - Clustering
    - Parallel Sysplex
    - HACMP
    - Linux Clusters
  - Workload Mgmt.
    - Work Load Managers
    - Intelligent Resource Dir.
    - Partition Load Manager
    - Enterprise WLM
    - Resource & Goal based
    - Policy based
    - System scope
    - End-to-End scope

- **Server Virtualization**
  - PR/SM
  - Dynamic LPARs
  - MicroPartitioning
  - zVM, VMware
  - Virtual I/O
  - Physical Partitions

- **Storage Server Virtualization**
  - POWER5 partitioning

- **SAN Volumes**
  - Storage Pools
  - Centralized mgnt

- **Total Storage Management**
  - Manage storage according to policy

- **GRID**
  - Globus Toolkit
  - IBM OGSA Toolbox

- **Virtual Engine** is not here yet

- **Server Allocation for Web Application Servers**
  - Computation heavy, parallel applications
  - Manage multiple applications across multiple server clusters
  - ISV Grid middleware
  - Provide services such as data services, scheduling, etc
Server Grid Infrastructure

Grid Management Center
- Resource Registration
- Project Registration

Applications and resources registered at

Grid Services
- Scheduling
- Load Balancing

Manages applications on

Runs applications off of

Virtual Software Storage
- Software and I/O files storage
- Create File access control

Runs locally on

Grid Resource Controller
- Securing and partitioning projects
- Monitoring/accounting
Server Grid Architecture
IPC Example on intraGrid

The "stress test" script has been invoked which generates a large load on the system.

Load begins to increase but is still, right now, only running locally on the superior node.
IPC on intraGrid

Load hits high water mark level (1500) and spills over to the first of the inferior nodes:

**trexnode2.itso.ibm.com [9.12.6.129]**

This inferior node has 2 CPU's available:
- one CPU already has the IPC process **“running”**
- the other CPU is currently **“starting”** up another IPC process to share the load

From the graph you can see the average load is dropping, but is still above the 1500 "high water mark"...
IPC on intraGrid

A second inferior node:


is recruited to help reduce the average load
IPC on intraGrid

The average load has leveled off and is now below the “high water mark” of 1500, so no more resources need to be recruited.
Data Grid

- A scalable, reliable and secure system that provides an efficient and adaptable download service.
- Features
  - Parallel downloads from multiple servers
  - Load balancing
  - Resumable downloads
  - Quality of Service
  - Public / Private files with centralized file access control.
- Benefits:
  - 4 to 5+ time faster downloads
  - Non-dedicated resources
  - Better network load distribution
  - Reduced administration
Clients will simultaneously download package segments from local Depot Servers and Peers.

Overall staging server infrastructure will be significantly reduced.

WAN traffic will greatly reduced.
Cloud Computing

- Cloud computing also describes applications that are extended to be accessible through the Internet.

- These *cloud applications* use large data centers and powerful servers that host Web applications and Web services.

- Anyone with a suitable Internet connection and a standard browser can access a cloud application.
**Definition**

- A cloud is a pool of virtualized computer resources.
- A cloud can:
  - Host a variety of different workloads, including batch-style back-end jobs and interactive, user-facing applications
  - Allow workloads to be deployed and scaled-out quickly through the rapid provisioning of virtual machines or physical machines
  - Support redundant, self-recovering, highly scalable programming models that allow workloads to recover from many unavoidable hardware/software failures
  - Monitor resource use in real time to enable rebalancing of allocations when needed
Grid Computing


*Systems virtualization and management*

- Virtualized computer resources.
- Autonomic systems management.
- Robust platform for next-generation Web 2.0 applications.
Grid Computing

Grid Example:
China Nation Grid (CNGrid)
CNGrid Team

- CNIC, CAS (Beijing, major node)
- Shanghai Supercomputer Center (Shanghai, major node)
- Tsinghua University (Beijing)
- Institute of Applied Physics and Computational Mathematics (Beijing)
- University of Science and Technology of China (Hefei, Anhui)
- Xi’an Jiaotong University (Xi’an, Shaanxi)
- NUDT (Changsha)
- Hong Kong University (Hongkong)
- The CNGrid Operation Center (based on CNIC, CAS)
CNGrid Resources

- Aggregated computing power:
  - 18 TFlops
- >200TB disk storage
- Application software
- Databases
CNGrid Key Systems

- Equipped with Dawning 4000A, 10.2 TFlops
- Utilization >80%

- Equipped with Lenovo DeepComp 6800, 5.324 TFlops
- Average utilization 86%
国内首台IBM Cell刀片服务器集群投入运行 中计在线

国内首台IBM Cell刀片服务器集群投入运行
作者：佚名 文章来源：中计在线 更新时间：2007-4-16 17:26:21

【中计在线消息】（2007年4月16日，北京）今天，中国科学院（CAS）计算机网络信息中心和IBM（NYSE：IBM）在国内合作的第一台基于Cell BE（Broadband Engine）WIDE系统开始投入运行。这一系统安装在位于北京的中国科学院超级计算中心内，中国科学院、多所大学以及一些国内科研和商业机构将很快能够通过中国国家网格（CNGrid）系统。在早期的系统测试中，CNGrid和IBM联合组建的测试团队进行了包含化WIDE BE的硬件潜能开发，将地震预报的计算速度提高了60倍。地震预报专家可以在短短数小时内得到计算结果。该系统在超大规模超算领域也取得了的成就（1），而且类似的成就也将在金融模拟、多种类型的建模、航天计算、网络通信等领域得到验证。支持CNGrid开发的中国国家高科技项目组关键项目主任钱学松表示：“这一最新的CNGrid节点为国内行业和研究机构提供了一个令人兴奋的机遇。”

中国科学院计算机网络信息中心和IBM在2006年10月启动的这一联合项目将为CNGrid提供一个基于Cell BE的节点，并鼓励开发各种能充分利用其独特性能的应用。该服务的中标运行将使国内科研、教育和工程界迈出至关重要的一步。

IBM中国研究院高级研究员魏镇表示：“非常高兴有这样的机会，CNGrid IBM Cell BE刀片将与IBM Cell BE刀片进行合作，通过使用这种技术大力推
CNGrid Cell B.E. Node

- IBM Cell B.E.
- 7 Blades in a BladeCenter
- Compute Intensive applications

- Non-homogeneous coherent multi-Processor
  - Dual-threaded PPC control processor
  - 8 independent SIMD/Vector Accelerators
- ~250M transistors; ~235mm²
- Performance - Top frequency >4GHz
  - > 256 GFLOPs (Single Precision)
  - Up to 25.6GB/s memory B/W
  - Up to 75 GB/s I/O B/W
CNGrid Software

Grid Computing

CNGrid Software

- Vega GOS (社区、网程、路由器及系统服务)
- GrisHIELD (网络安全机制)
- GrisDaen (网格文件系统)
- 基于网格软件的网格门户引擎和Web应用

网格服务器
- 上海超算网格节点
- 网格服务器
- 网格证书管理中心 (CA)
- 网格服务器
- 香港大学网格节点
- 专用客户端 (网格应用客户端)
- Web浏览器

网格客户端
- 通用Web浏览器 (IE, Firefox)
- 和/或网格软件管理工具
- 和/或基于网格软件API的网格应用
CNGrid Applications

- Applications from selected areas
  - Resource and Environment
    - NGG, DFG, SeisGrid
  - Research
    - SDG, BAGrid, DDG
  - Services
    - CMAG, TIG
  - Manufacturing
    - AviGrid, SimGrid
CNGrid Applications