

# Grid and Cloud computing, a place for virtualisation?

Dr. Fabrizio Gagliardi

EMEA and LATAM Director

Technical Computing

Microsoft Research



e-IRG workshop, Zurich, April 2008





# Outline

- Introduction
- Virtualisation
- Definition - Pros & Cons - Market trends
- Grid computing for e-Science and virtualisation?
- Cloud computing adopting virtualisation
- Virtualisation: “Green” IT?
- Conclusions



# White paper overview

1<sup>st</sup> eIRG meeting

Athens

June, 10<sup>th</sup> 2003

Fotis Karayannis

GRNET - Greek Research & Technology Network

fkara@grnet.gr, <http://www.grnet.gr>

Rome, December 10<sup>th</sup>, 2003



# What is Virtualization?

**Virtualization is the isolation of one computing resource from the others:**



or

## Virtual Applications

Any application on any computer on-demand

## Virtual Presentation

Presentation layer separate from process

## Virtual Machine

OS can be assigned to any desktop or server

## Virtual Storage

Storage and backup over the network

## Virtual Network

Localizing dispersed resources

*Slide Courtesy of Neil Sanderson,  
UK Product Manager,  
Virtualisation and Management,  
Microsoft Ltd.*



# Virtualisation: Pros

## Control costs

- Optimize utilisation of servers
- Reduce costs: hardware, power, space
- Reduce overall system complexity
- Reduce application compatibility conflicts

## Improve Availability

- Disaster recovery capabilities
- Streamline server maintenance, and isolate risk
- Enable access to any desktop application anywhere

## Drive Agility

- Dynamic resource allocation
- Faster provisioning of services / workloads to support business growth and meet SLAs



# Virtualisation: Cons?

- New technology
- Management overhead
- Single Point of Failure
- Software Licensing?
- Performance?
- Policy issues
- Security Concerns



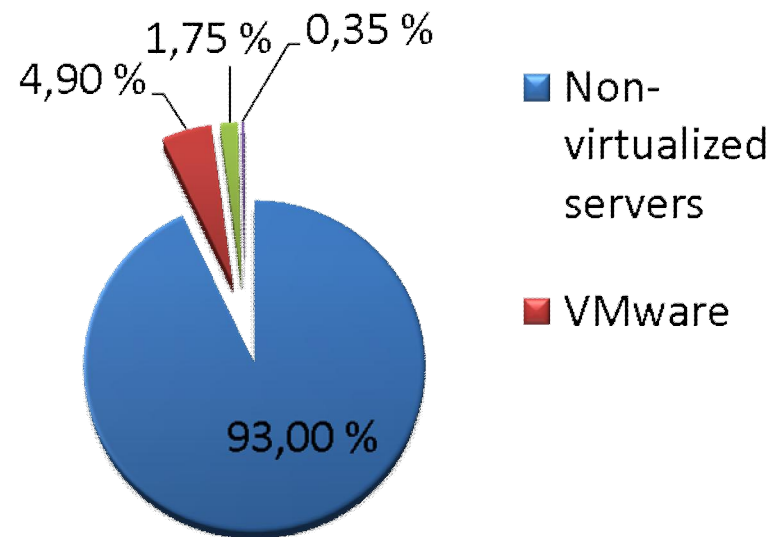


# Virtualisation market: The game is only starting (1/2)

## Only 5% of servers are virtualised

- Computerworld
  - “Although virtualization has been the buzz among technology providers, only 6% of enterprises have actually deployed virtualization on their networks, said Levine, citing a TWP Research report. That makes the other 94% a wide-open market.”
  - “We calculate that roughly 6% of new servers sold last year were virtualized and project that 7% of those sold this year will be virtualized and believe that less than 4% of the X86 server installed base has been virtualized to date.
- Pat Gelsinger, Intel VP Sept. 2007
  - “Only 5% of servers are virtualized.”

**World Wide  
Virtualization Adoption**



# Virtualisation market: The game is only starting (2/2)

## Multiple solutions – New entrants

e-IRG  
worksh  
op,  
Zurich,  
25 April  
2008

SWsoft tries to virtualize

Market Scan

The Competitive Outlook

FTN Midwest Securities analyst Trip Chowdhry said that the problem is that investors wrongly believed that VMware itself was the trend. Now, he says, they are realizing that virtualization is the trend.

Eggert Apple  
lets Leopard  
virtualize

Tired of thinking  
different

1 Nov 18:33

13 Nov 01:38

XenSource preaches the joy of platform virtualization

Interop New management, new credo, same old enemy

By Paula Rooney • Wednesday 24 Oct 2007 18:21

Slide Courtesy of Nell Sanderson



# The experience of the Grid <sup>1/3</sup>

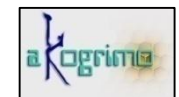
- Grids for e-Science:
  - a success story so far?
    - Several mature Grid Middleware stacks
    - Many HPC applications using the Grid
      - Some of them (HEP, Bio) in production use
      - Some of them still in testing phase: more effort still required to make the Grid their day-to-day workhorse
    - e-Health applications also part of the Grid
    - Some industrial applications:
      - CGG Earth Sciences
- Virtualisation not much exploited in e-Science



**HEALTHGRID**



EGEODE VO





# The experience of the Grid <sup>2/3</sup>

- Grids beyond e-Science?
  - Slower adoption: different priorities and requirements, different tools and different business models and TCOs
    - Intra grids, dedicated clusters, cloud computing
  - e-Business applications
    - Finance, Enterprise Resource Planning, SMEs
  - Industrial applications
    - Automotive, Aerospace, Pharmaceutical industry, Telecom
  - e-Government applications
    - Earth Observation, Civil protection:
    - e.g. The Cyclops project
- Virtualisation exploited beyond e-Science?





# The experience of the Grid <sup>3/3</sup>

- Industry also demonstrated interest in becoming an HPC infrastructure provider:
  - On-demand infrastructures:
    - Cloud and Elastic computing, pay as you go...
    - Data centers: Data getting more and more attention
  - Service hosting: outsourced integrated services
- Virtualisation being exploited in Cloud and Elastic computing (e.g. Amazon EC2 virtual instances)
- “Pre-commercial procurement”
  - Research-industry collaboration in Europe to achieve new leading-edge products
    - Example: PRACE building a PetaFlop Supercomputing Centre in Europe





# Cloud computing and storage on demand (1/2)

**Cloud Computing:** [http://en.wikipedia.org/wiki/Cloud\\_computing](http://en.wikipedia.org/wiki/Cloud_computing)

**Amazon, IBM, Google, Microsoft, Sun, Yahoo, major 'Cloud Platform' potential providers**

**Operating compute and storage facilities around the world**

**Have developed middleware technologies for resource sharing**

**First services already operational - Examples:**

**Amazon Elastic Computing Cloud (EC2) -Simple Storage Service (S3)**



# Cloud computing and storage on demand (2/2)



## Sun's Grid Utility Expands Beyond the United States

Published: May 3, 2007

by Timothy Prickett Morgan

Nearly two months ago, [Sun Microsystems](#) and a raft of software partners put a jukebox of sorts out onto its [network.com](#) Sun Grid compute and storage utility. Since end users don't have to get utility licenses, the problem that Sun had, however, is that only a limited capacity.

Starting today, however, the Sun Grid utility, by default, is now available in 24 additional countries: Australia, Austria, Belgium, Canada, China, Czech Republic, Greece, Hungary, India, Ireland, Italy, Japan, Korea, Sweden, and the United Kingdom. To help spread the word, Sun is giving away 200 CPU hours of compute time to each individual—who starts up an account. This is a significant step when it has enough new users on the system.

<http://www.itjungle.com/bns/bns100807-story02.html>

<http://www.itjungle.com/tug/tug050307-story05>

[http://www.computerworld.com/action/article.do?command=viewArticleBasic&taxonomyName=mainframes\\_and\\_supercomputers&articleId=9073758&taxonomyId=67&intsrc=kc\\_top](http://www.computerworld.com/action/article.do?command=viewArticleBasic&taxonomyName=mainframes_and_supercomputers&articleId=9073758&taxonomyId=67&intsrc=kc_top)



## Google, IBM Partner on Utility Computing Cloud

Published: October 8, 2007

by Timothy Prickett Morgan

Search engine and Web advertising giant [Google](#) and server maker and supercomputing giant [IBM](#) are working with some prominent American universities to set up a cloud computing utility to assist those universities in teaching and developing applications that represent the future of corporate computing some day.

For Internet-based applications, Google perhaps knows more about utility computing (or cloud computing) than most. It has been utilized on the fly as workloads change) or cloud computing. It seems to imply utility-style infrastructure as well as the size of the computing resources dedicated to the application. It could encompass many different kinds of utilities, with varying degrees of complexity, and a workload would make use of the computing resources in a way to refer to a network of machines connected by a network. For instance, a Linux-X64 cluster might be used to run a vector parallel computer might be used to run data and then a separate virtualization system might be used to run the application. Similarly, large-scale parallel systems can be used to search for spam (as well), perhaps looking for spam lurking in wiki and forum posts in business transactions.

**COMPUTERWORLD**  
Hardware

SEARCH Google Custom

» Five Technologies Simplifying Infrastructure Management  
» Powering Change in the Data Center  
» Time Warner Cable of New York, Case Study

### Supercomputing at less than the cost of a gallon of gas

New service allows users to offload computational work to supercomputer

By Patrick Thibodeau    Comments 3    Recommended 45    Share

April 1, 2008 (Computerworld) A supercomputing, pay-per-use service that would make supercomputing affordable to small and midsize businesses, as well individuals who need added computation power, was announced today.

Interactive Supercomputing Inc., a Waltham, Mass.-based company said its Star-P On-Demand lets users tap into 168 processor cores on Intel Xeon processors.

This is not a software-as-a-service offering that gives users access to the cloud, but rather a hardware-as-a-service offering that gives users access to the cloud.



# Amazon EC2 and S3

- EC2 Beta Service: Web-Services based  
<http://www.amazon.com/gp/browse.html?node=201590011>
  - \$0.10 per hour - Small Instance (Default)
    - 1.7 GB of memory, 1 EC2 Compute Unit (1 virtual core with 1 EC2 Compute Unit), 160 GB of instance storage, 32-bit platform
    - EC2 Compute Unit (ECU) - One EC2 Compute Unit (ECU) provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor
- S3 storage services: WS-based (REST and SOAP)  
[http://www.amazon.com/S3-AWS-home-page-Money/b/ref=sc\\_fe\\_l\\_2?ie=UTF8&node=16427261&no=5440661&me=A36L942TSJ2AJA](http://www.amazon.com/S3-AWS-home-page-Money/b/ref=sc_fe_l_2?ie=UTF8&node=16427261&no=5440661&me=A36L942TSJ2AJA)
  - *Storage: \$0.15 per GB-Month of storage used*
  - *Data Transfer: \$0.10 per GB - all data transfer IN*
    - \$0.18 per GB - first 10 TB / month data transfer OUT*
    - \$0.16 per GB - next 40 TB / month data transfer OUT*
    - \$0.13 per GB - data transfer out / month over 50 TB*

*Services may be given below actual cost for various reasons*



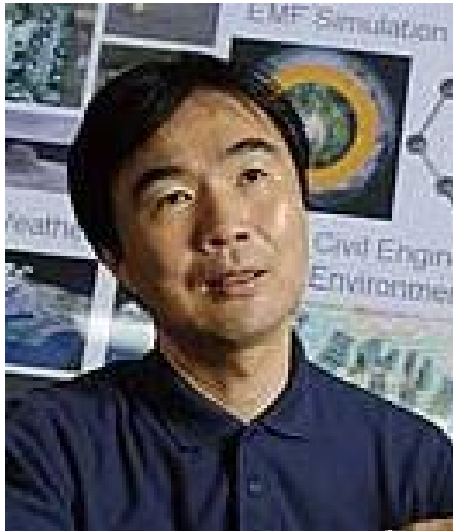
# Virtualization and Cloud Computing



- ***“There’s Grid in them than clouds!”***
  - I. Foster’s blog, ANL & UC, Jan. 8, 2008
- Clouds have a very simple user API effectively hiding all the complexity of an *ad hoc grid* on the back-end
  - e.g., Amazon’s EC2 & S3, IBM’s Blue Cloud and others ...
- ***If so, will this enable mass-market grids?***
  - Users don’t have to be aware of using “a grid”
- ***If so, what does “cloud interoperability” require?***
  - Is *virtualization* a means of achieving this?
- ***Major opportunity for synergy***

# Switching Gears:

## “To Distribute or Not To Distribute”



- Prof. Satoshi Matsuoka, TITech
- Keynote at Mardi Gras Conference, Baton Rouge, Jan.31, 2008
  - **In the late 90s, petaflops were considered very hard and at least 20 years off ...**
  - **... while grids were supposed to happen right away**
  - **After 10 years (around now) petaflops are “real close” but there's still no "global grid”**
  - ***What happened?***



# What Happened?

- It was easier to put together massive clusters than to get people to agree about how to share their resources
- For tightly coupled HPC applications, tightly coupled machines are still necessary
- Grids are inherently suited for loosely coupled apps (e.g., Monte Carlo, Parameter Sweep), or enabling access to machines and data, and the integration of the two
  - **With Gilder's Law, bandwidth to the compute resources will promote thin client approach**
  - **Example: *Tsubame* machine in Tokyo**

# Microsoft virtualisation: More Green IT?

- Production Use
- Virtual Server in heavy use for 18+ months
- 1,250+ virtual machines
- Consolidation Ratios
  - 8 servers to 1 server
- Test/Development Use
- Virtual Server in heavy use for 18+ months
- 500+ virtual machines
- Consolidation Ratios
  - 16 servers to 1 server

## • Test/Development Savings:

Item	Physical System Cost	Virtual Server Build Cost	Savings
Number of servers required	477 systems @ a cost of \$5k each Total \$2.3 million	16 physical host systems @ \$20k each Total: \$320 thousand	Just under 2 million dollars
Hard drive space	19 terabytes	8 terabytes	11 terabytes
Rack space	30 racks	2 racks	28 racks
Power	525 amps	8 amps	517 amps

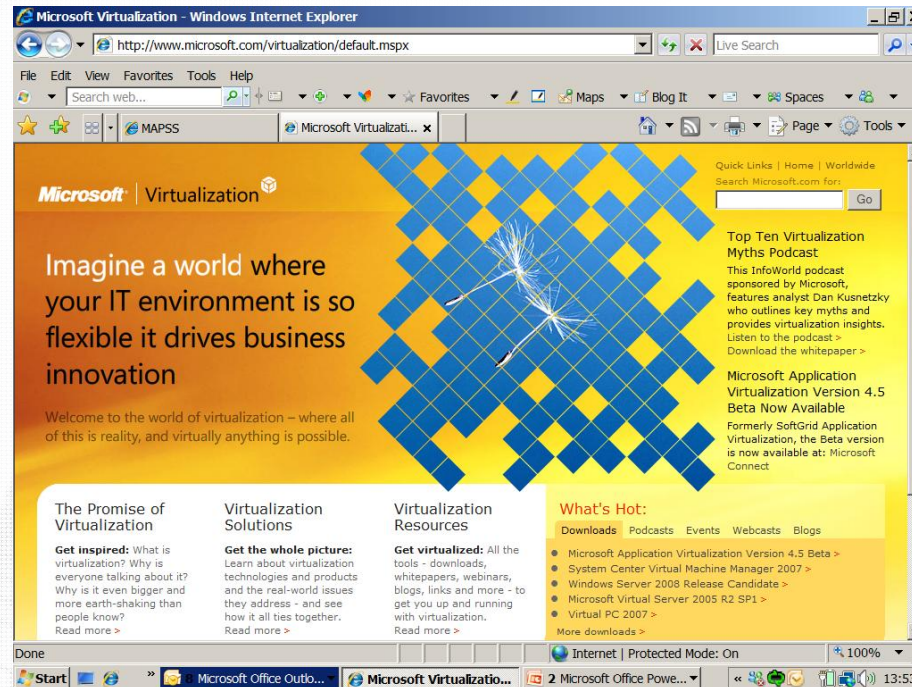


# Conclusions

- Virtualisation: A promising technology
- It's not only about virtual machines...
- Virtualisation not much exploited in e-Science so far
- However exploited in Cloud computing and hosted services
  - Which emerge as the next incarnation of distributed computing
- Virtualisation is a better Green technology!
  - Less space, less energy consumption
- Microsoft: a player in the field
  - HyperVisor



## •Resources



<http://www.microsoft.com/virtualisation>



# Online Resources

- Microsoft Virtualization Home:
  - <http://www.microsoft.com/virtualization>
- Windows Server Virtualization Blog Site:
  - <http://blogs.technet.com/virtualization/default.aspx>
- Windows Server Virtualization TechNet Site:
  - <http://technet2.microsoft.com/windowsserver2008/en/servermanager/virtualization.mspix>
- TS Blog:  
<http://blogs.msdn.com/ts>
- Windows Server 2008 TS site:  
<http://www.microsoft.com/windowsserver2008/>
- TS TechNet Site:  
<http://technet.microsoft.com/ts>
- TS 2008 Resources  
<http://technet2.microsoft.com/windowsserver2008/en/servermanager/terminalservices.mspix>



# Thanks

Thanks to e-IRG for the kind  
invitation and to all of you  
for your attention

Fabrizio.Gagliardi@microsoft.com