# CIS 520 – Operating Systems I

## Lecture 26: Cloud Computing

**Instructor: Mitch Neilsen**

**Office: N219D**

* NOTE: Some lecture notes taken from CIS 520 – D. Andresen, Kansas State University, and CIS 140 – Operating Systems - Stanford University

# Two Technologies for Agility

**Virtualization:**

*The ability to run multiple operating systems on a single physical system and share the underlying hardware resources**

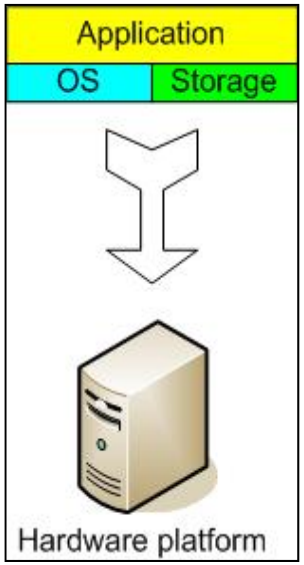
**Cloud Computing:**

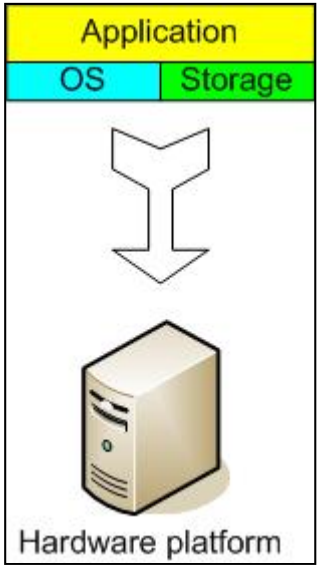
*"The provisioning of services in a timely (near on instant), on-demand manner, to allow the scaling up and down of resources"***

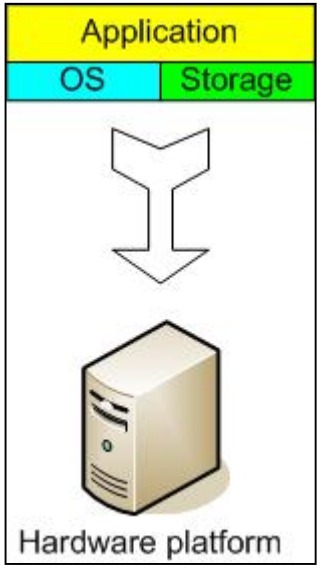
* VMware white paper, *Virtualization Overview*

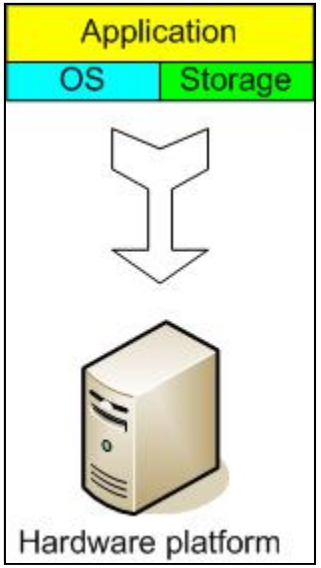
** Alan Williamson, quoted in *Cloud BootCamp March 2009*

# The Traditional Server Concept



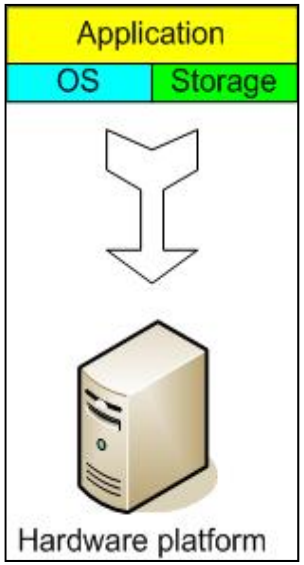




| Web Server | App Server | DB Server | EMail |
|---|---|---|---|
| Windows | Linux | Linux | Windows |
| IIS | Glassfish | MySQL | Exchange |

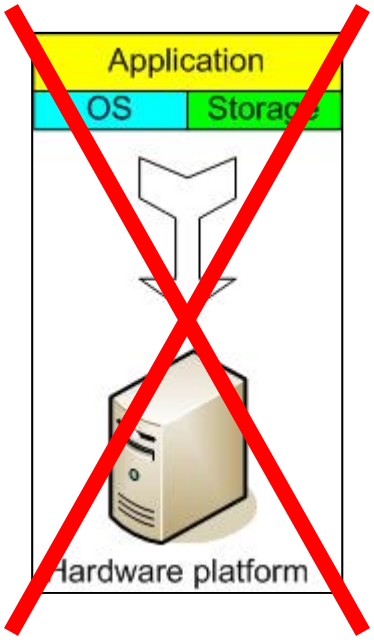
# And if something goes wrong ...




**Web Server**

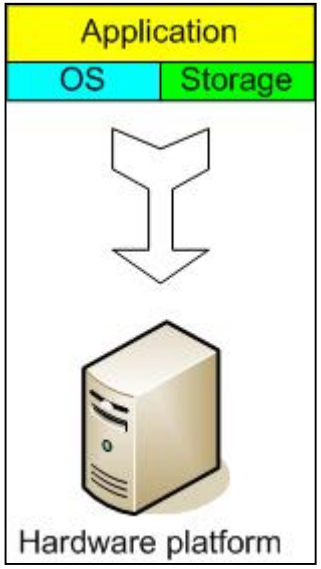
**Windows**

**IIS**




**App Server**

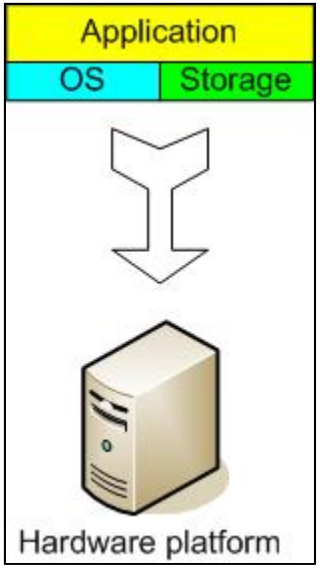
**DOWN!**




**DB Server**

**Linux**

**MySQL**




**EMail**

**Windows**

**Exchange**

# The Traditional Server Concept

- System Administrators often talk about servers as a whole unit that includes the hardware, the OS, the storage, and the applications.
- Servers are often referred to by their function; e.g., the SQL database server, the file server, etc.
- If the file server fills up, or the database server becomes overtaxed, then the System Administrators must add a new server.

# The Traditional Server Concept

- Unless there are multiple servers, if a service experiences a hardware failure, then the service is down.
- System Administrators can implement clusters of redundant servers to make them more fault tolerant. However, even clusters have limits on their scalability, and not all applications work in a clustered environment.

# The Traditional Server Concept

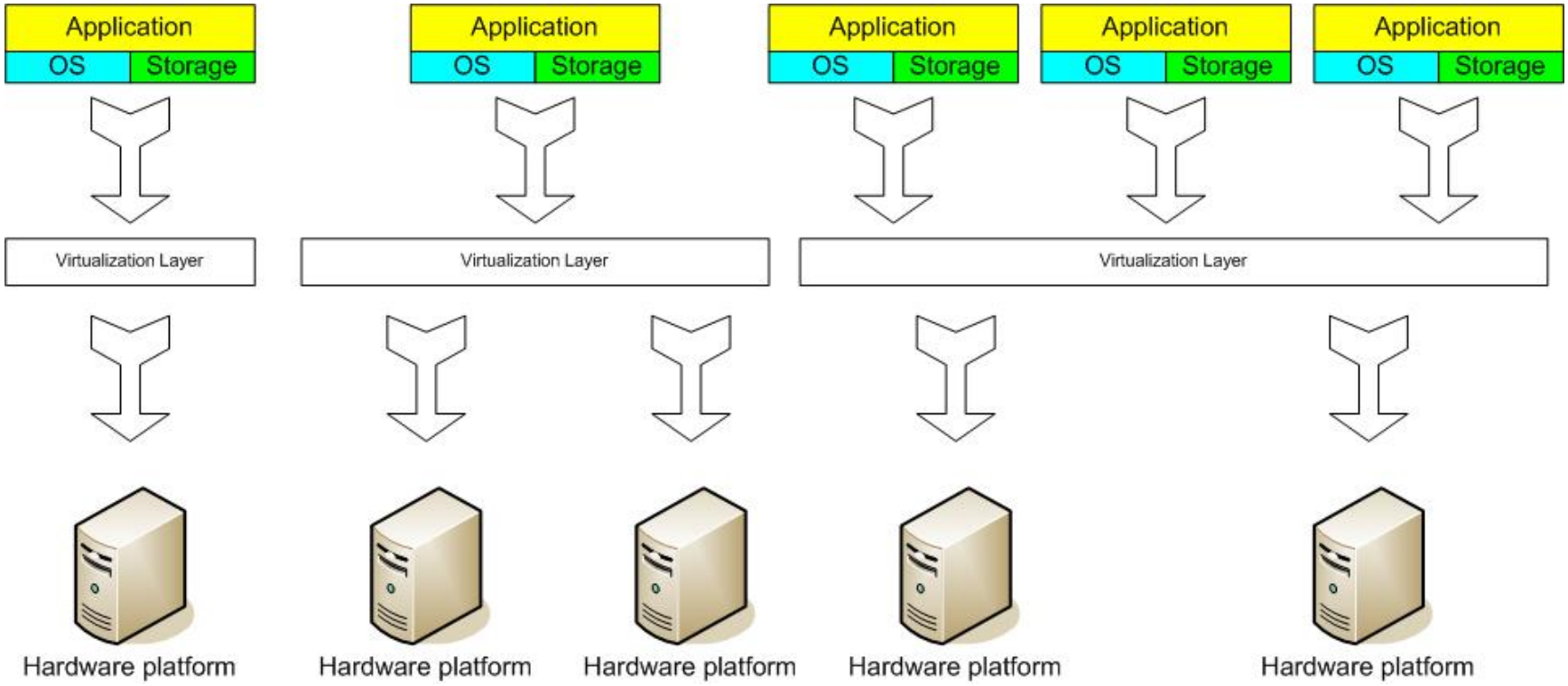
**Pros**

- Easy to conceptualize
- Fairly easy to deploy
- Easy to backup
- Virtually any application/service can be run from this type of setup

**Cons**

- Expensive to acquire and maintain hardware
- Not very scalable
- Difficult to replicate
- Redundancy is difficult to implement
- Vulnerable to hardware outages
- In many cases, processor is under-utilized

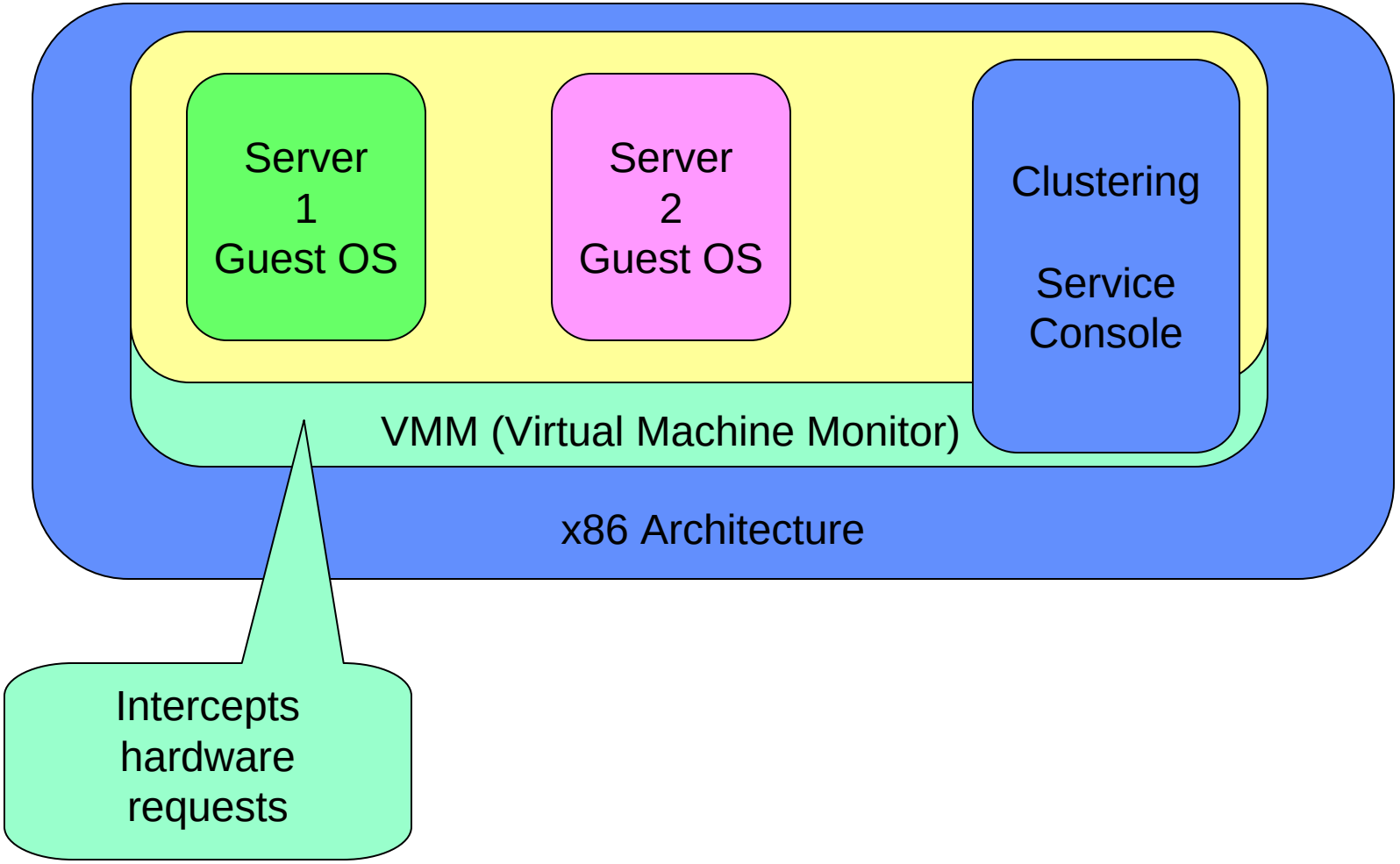
# The Virtual Server Concept




**Virtual Machine Monitor (VMM) layer between *Guest OS* and hardware**

# Close-up*

* adapted from a diagram in VMware white paper, *Virtualization Overview*



# The Virtual Server Concept

- Virtual servers seek to encapsulate the server software away from the hardware. This includes the OS, the applications, and the storage for that server.
- Servers end up as mere files stored on a physical box, or in enterprise storage.
- A virtual server can be serviced by one or more hosts, and one host may house more than one virtual server.

# The Virtual Server Concept

- Virtual servers can still be referred to by their function i.e. email server, database server, etc.
- If the environment is built correctly, virtual servers will not be affected by the loss of a host.
- Hosts may be removed and introduced easily to accommodate maintenance.

# The Virtual Server Concept

- Virtual servers can be scaled out easily. If the administrators find that the resources supporting a virtual server are being taxed too much, they can adjust the amount of resources allocated to that virtual server.
- Server templates can be created in a virtual environment to be used to create multiple, identical virtual servers.
- Virtual servers themselves can be easily migrated from host to host.

# The Virtual Server Concept

**Pros**

- Resource pooling
- Highly redundant
- Highly available
- Rapidly deploy new servers
- Easy to deploy
- Reconfigurable while services are running
- Optimizes physical resources by doing more with less

**Cons**

- Slightly harder to conceptualize
- Slightly more costly (must buy hardware, OS, Apps, and the abstraction layer)

# Virtualization Status

Offerings from many companies

- e.g. VMware, Microsoft, Sun, ...

Hardware support

- Fits well with the move to 64 bit (very large memories) multi-core (concurrency) processors.
- Intel VT (Virtualization Technology) provides hardware to support the Virtual Machine Monitor layer

**Virtualization is now a well-established technology**

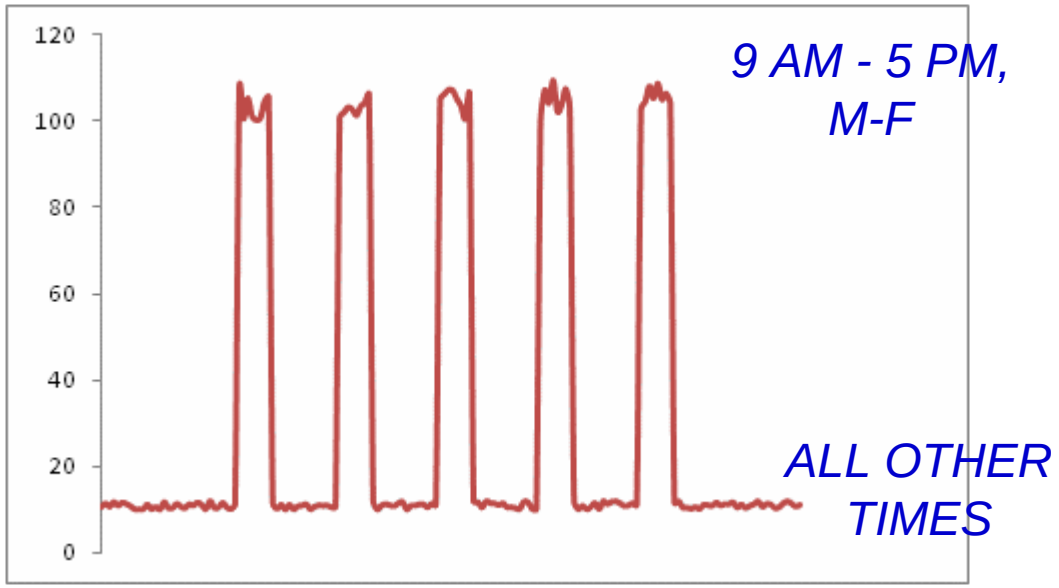
# So what about Cloud Computing?

# Suppose you are Forbes.com

You offer on-line real time stock market data

Why pay for capacity weekends, overnight?

**Rate of Server Accesses**



# Forbes' Solution

Host the web site in Amazon's EC2 *Elastic Compute Cloud*

Provision new servers every day, and deprovision them every night

Pay just $0.10* per server per hour

- * more for higher capacity servers

Let Amazon worry about the hardware!

# Cloud computing takes virtualization to the next step

**You don’t have to own the hardware**

You “rent” it as needed from a cloud

There are public clouds

- e.g. Amazon EC2, and now many others (Microsoft, IBM, Sun, and others ...)

A company can create a private one

- With more control over security, etc.

# Goal 1 – Cost Control

Cost

- Many systems have variable demands

  Batch processing (e.g. New York Times)

  Web sites with peaks (e.g. Forbes)

  Startups with unknown demand (e.g. the *Cash for Clunkers* program)

- Reduce risk

  Don't need to buy hardware until you need it

# Goal 2 - Business Agility

More than scalability - ***elasticity***!

- Ely Lilly in rapidly changing health care business

  Used to take 3 - 4 months to give a department a server cluster, then they would hoard it!

- Using EC2, about 5 minutes!

  And they give it back when they are done!

Scaling back is as important as scaling up

# Goal 3 - Stick to Our Business

Most companies don't WANT to do system administration

- Forbes says:

  *We are is a publishing company, not a software company*

*But beware:*

- *Do you really save much on sys admin?*
- *You don't have the hardware, but you still need to manage the OS!*

# How Cloud Computing Works

Various providers let you create virtual servers

- Set up an account, perhaps just with a credit card

You create virtual servers ("virtualization")

- Choose the OS and software each "instance" will have
- It will run on a large server farm located somewhere
- You can instantiate more on a few minutes' notice
- You can shut down instances in a minute or so

They send you a bill for what you use

# Any Nasty Details?
## *(loads of them!)*

How do I pick a provider?

Am I locked in to a provider?

Where do I put my data?

What happens to my data when I shut down?

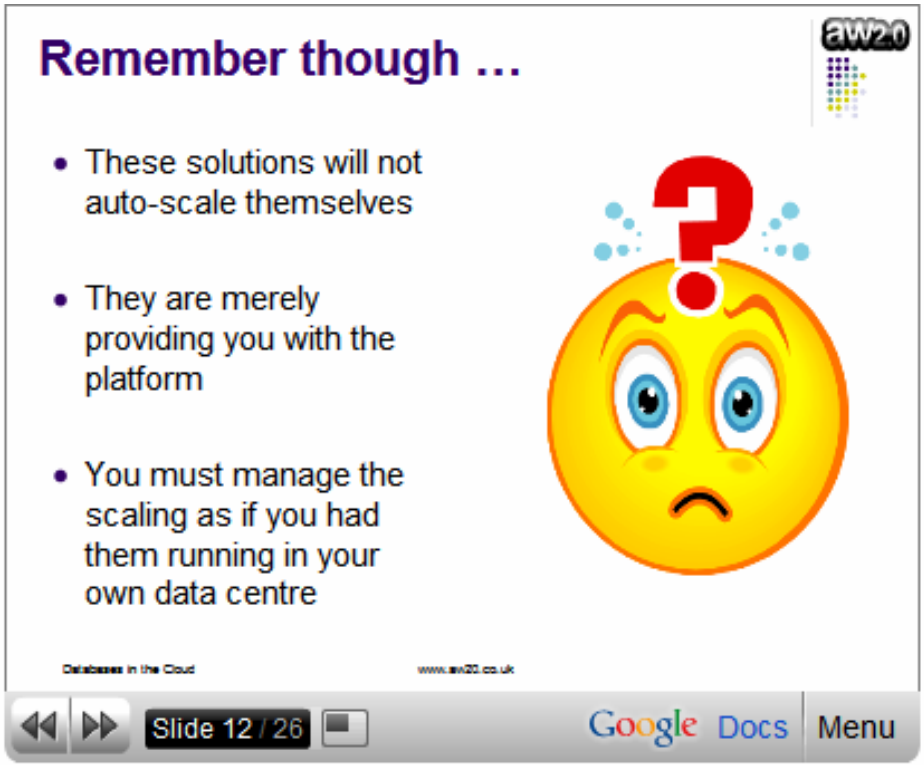
How do I log in to my server?

How do I keep others from logging in (security)?

How do I get an IP address?

Etc.

# And One Really Important Caveat*




** Cloud BootCamp March 2009*

# (footnote)
# How come Amazon?

Grew out of efforts to manage Amazon's own services

- (Each time you get a page from Amazon, over a hundred servers are involved)
- See reference *Amazon Architecture* on their service design concepts

They got so good at it that they launched Amazon Web Services (AWS) as a product

# Cloud Computing Status

Seems to be rapidly becoming a mainstream practice

Numerous providers

- Amazon EC2 imitators ...
- Just about every major industry name

  IBM, Sun, Microsoft, ...

Major buzz at industry meetings

# References

(links are current as of September, 2009)


VMware Inc., *Virtualization Overview*, http://www.vmware.com/pdf/virtualization.pdf

Todd Hoff, *Amazon Architecture*, http://highscalability.com/amazon-architecture, Sept. 18, 2007

Intel Corp., *Technology brief: Understanding Intel® Virtualization Technology*, http://download.intel.com/technology/virtualization/320426.pdf

aw2.0 Ltd, *Cloud BootCamp March 2009*, http://www.aw20.co.uk/help/cloudbootcamp_march2009.cfm

also online: CloudBootCamp2009.ppt