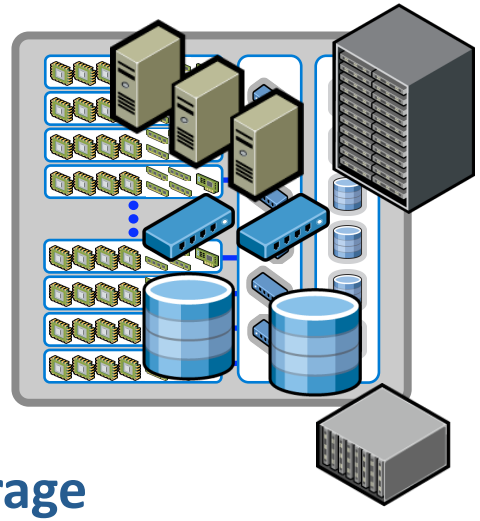
Data Center Evolution and the "Cloud"

Paul A. Strassmann
George Mason University
November 5, 2008, 7:20 to 10:00 PM

Hardware Evolution

Where is hardware going?



- **x86 continues to move upstream**
- **Massive compute power available**
- **Power usage grows even more critical**
- **Highly-connected compute, networking, and storage**

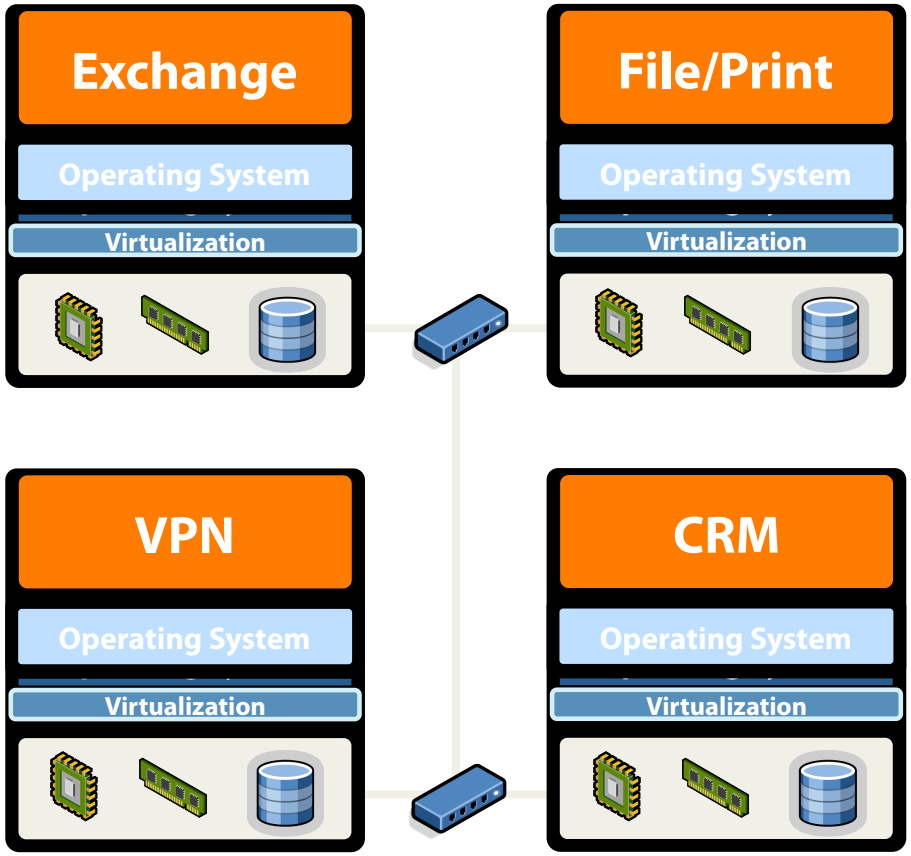
  Emergence of high-speed, shared interconnects

  Shared storage becomes less expensive and more pervasive

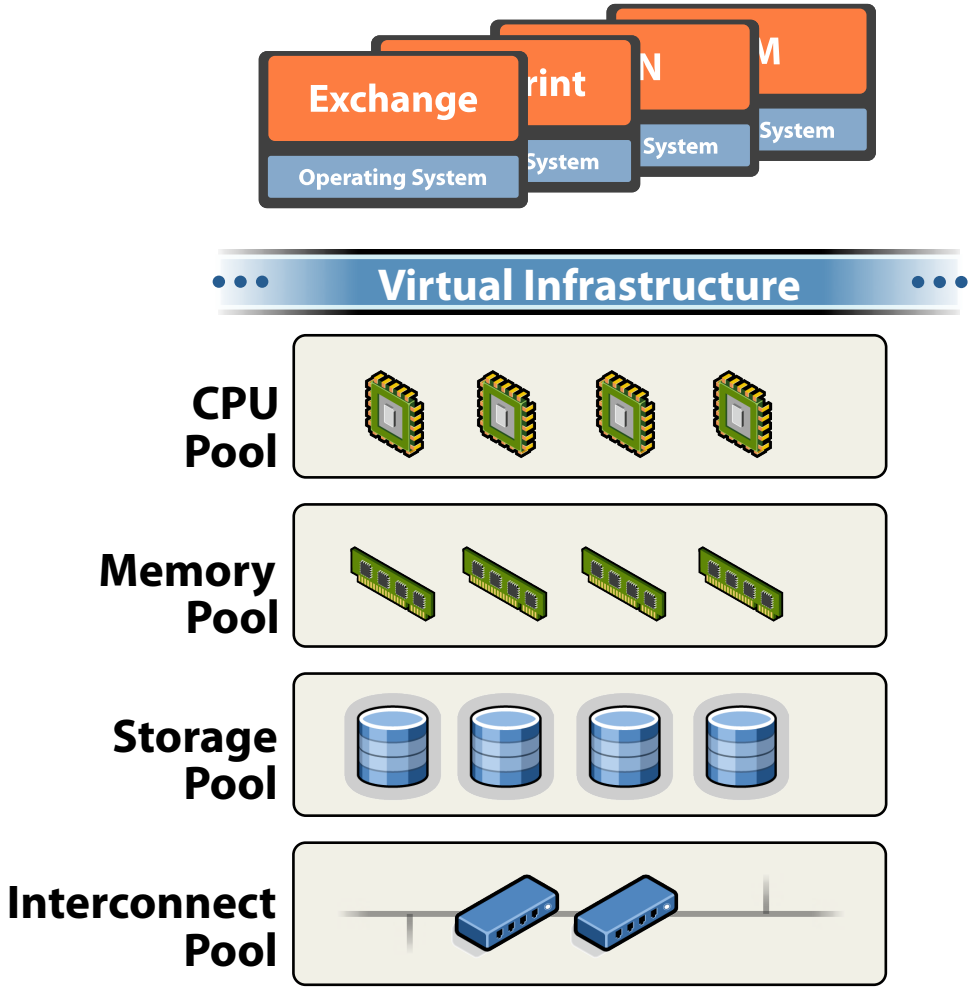
  Networking and storage riding x86 improvement curve

Virtualization transforms powerful, disparate hardware into Virtual Infrastructure

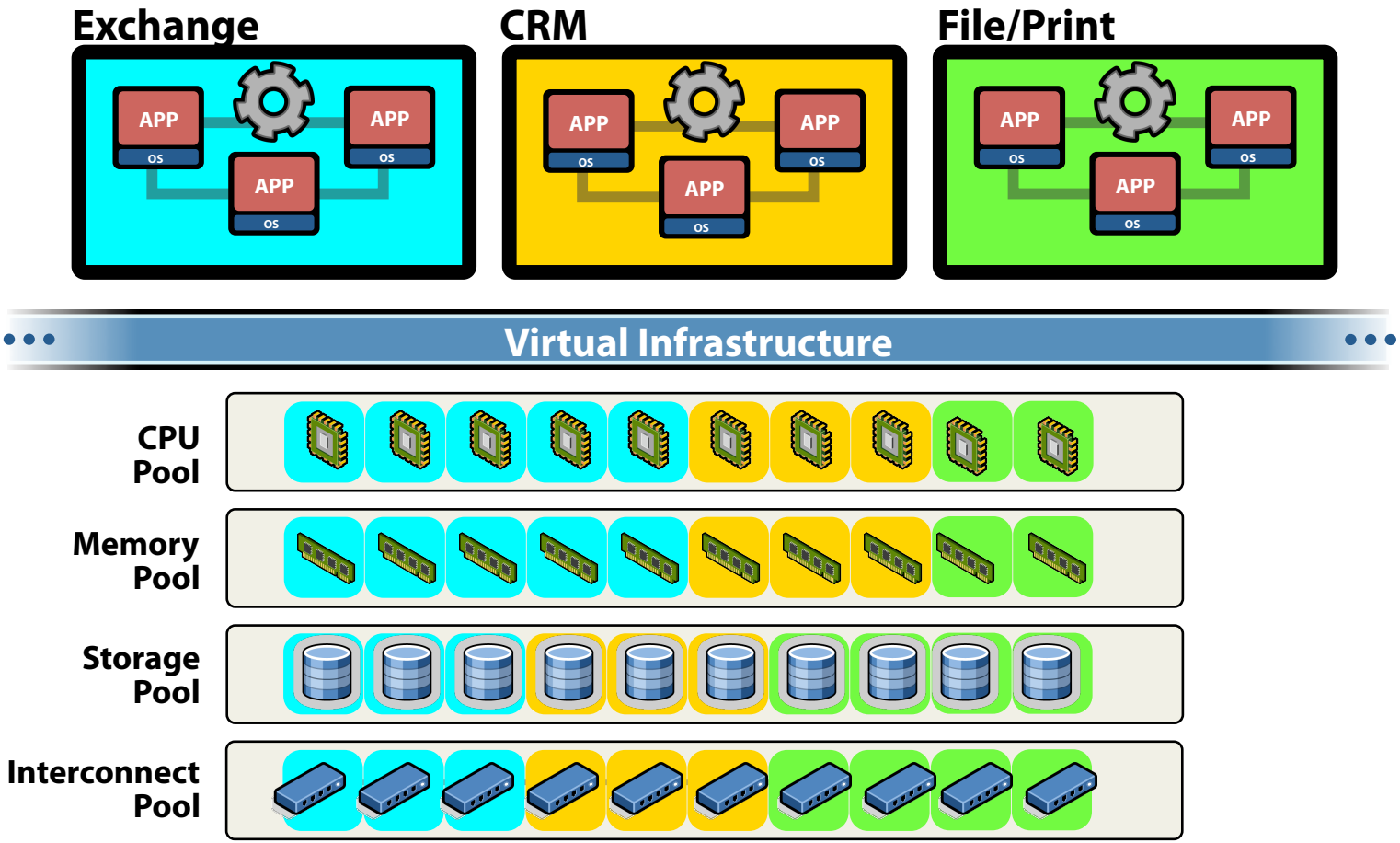
Non-Virtualized World



Enables the Virtual Datacenter



Self-Healing



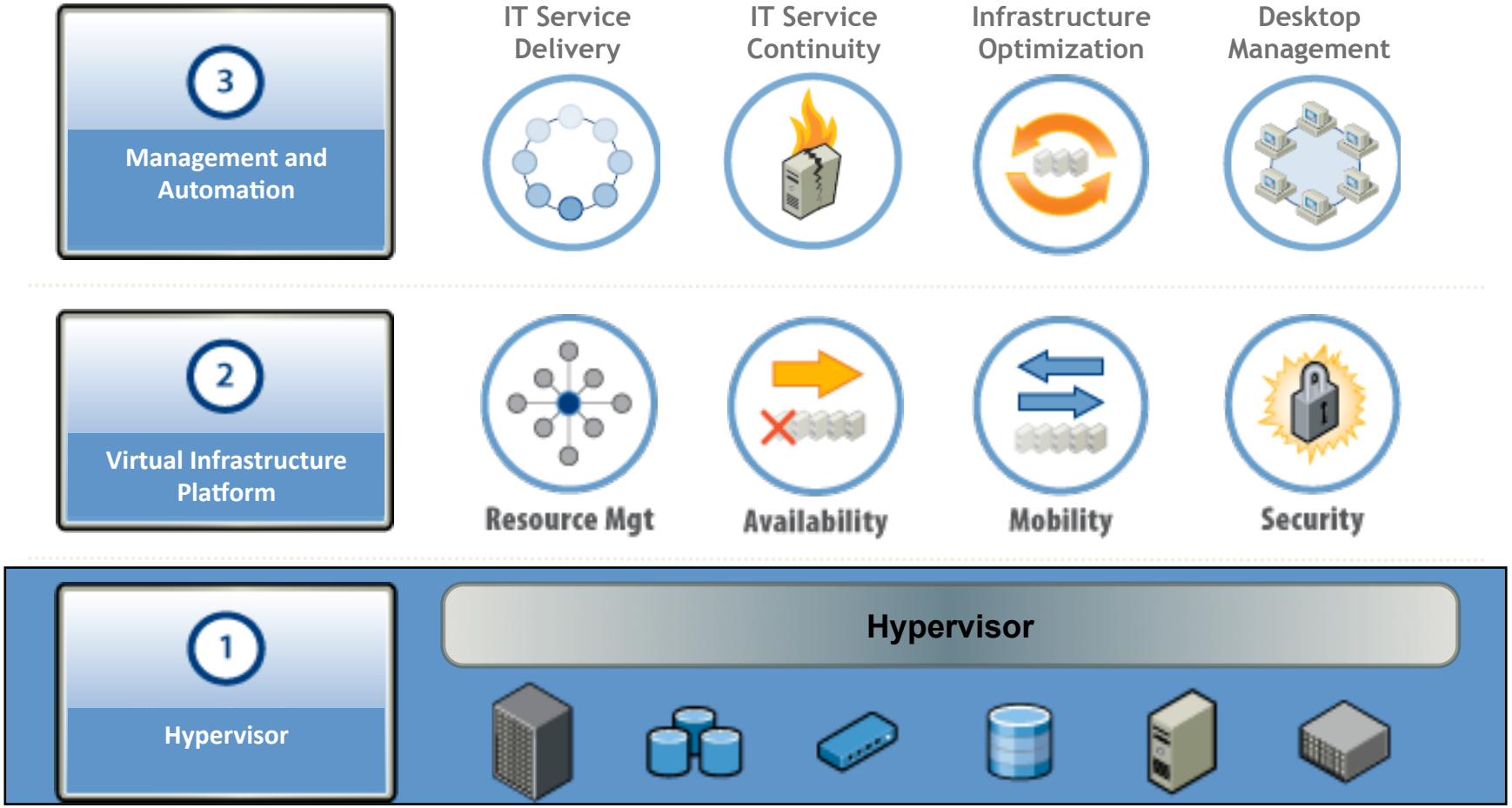
Self-Optimizing, Self-Protecting Datacenter

Moves functionality traditionally considered “management” into the infrastructure

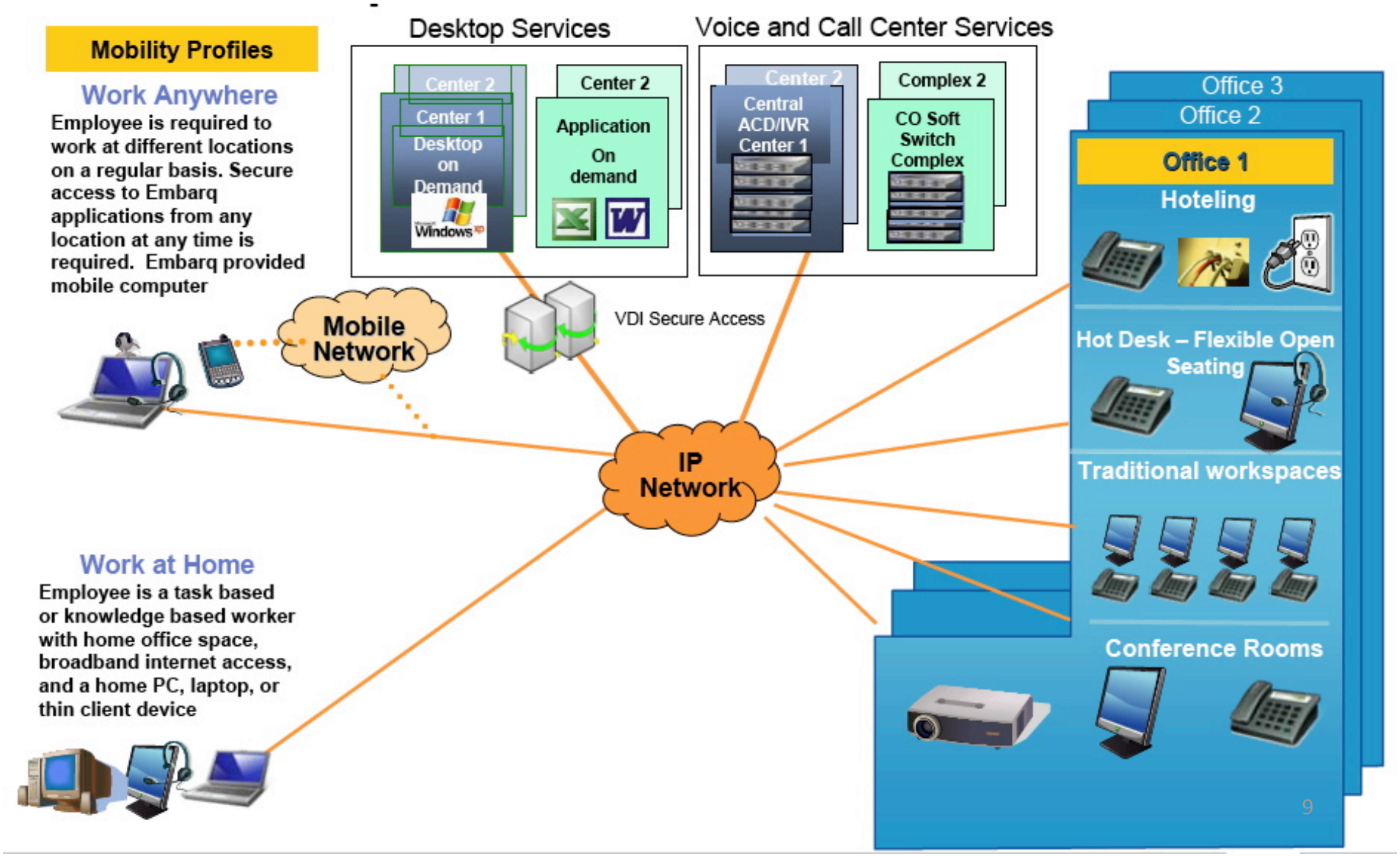
Any application gains the benefits

Provides new reasons to go virtual

Evolution of a Virtual Datacenter

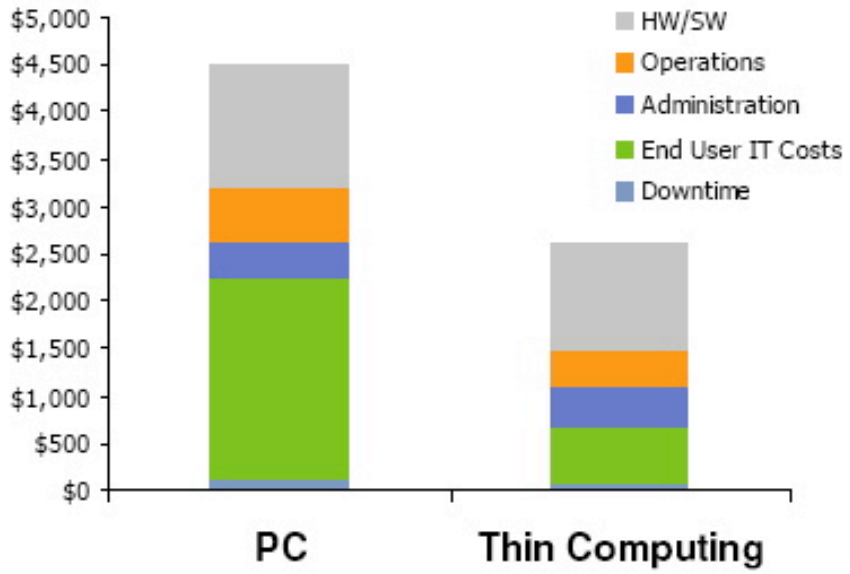


The Mobile Workplace



Why Use Thin Client?

Cost Benefits



40% Reduction HW and SW costs

29% Reduction IT operations costs

88% Reduction worker downtime

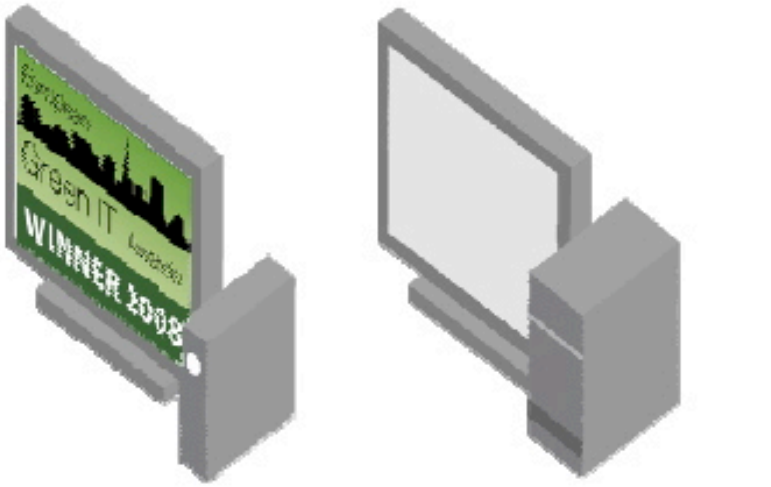
78% Increase IT staff productivity

Business Benefits

- Security / Privacy
- Compliance
- Manageability
- Reliability
- Rapid Deployment
- Power/Noise/Cooling

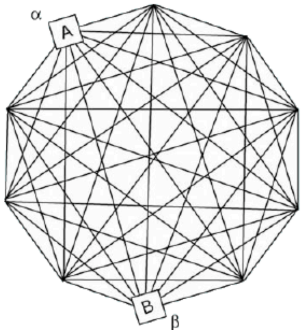
Energy Efficiency – Thin Clients

- Reduce energy usage by up to 90%
- Less materials used in manufacture
- Lower transportation costs
- Longer service life
- Fewer parts to recycle
- 90% reduction in e-waste

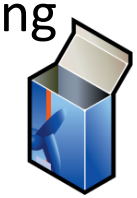


Software Evolution

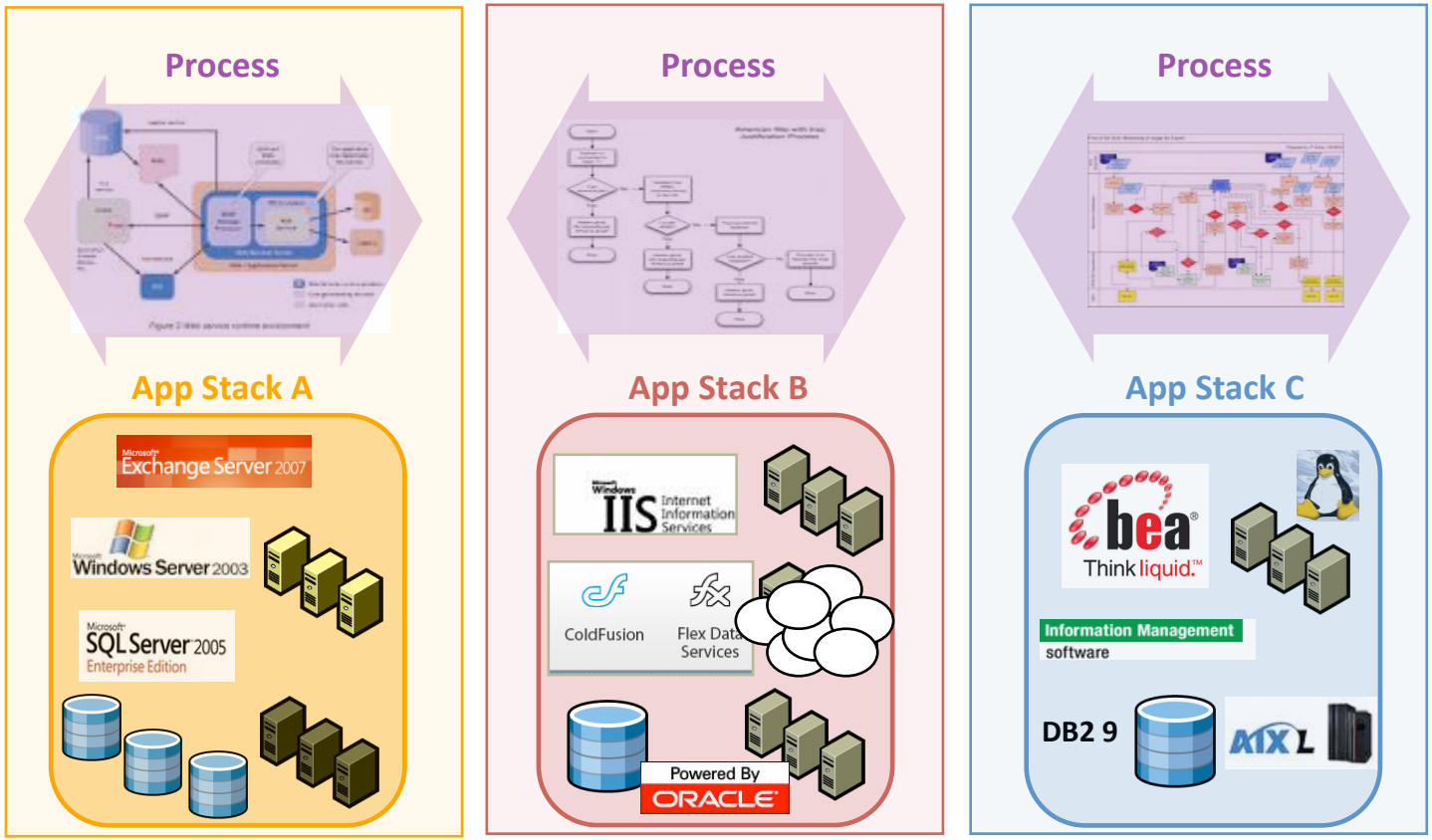
Where is software going?



- **More diversity in Operating Systems and Applications**
  - Simultaneous use of Linux, Solaris, and Windows farms
  - Customized open source stacks and diverse versions/distributions
- **Services becoming disaggregated and distributed**
  - SOA leads to many more components in use
  - Software as a Service and hosted computing options growing
- **Management task becomes even more difficult**
  - Need a way to normalize management approach
  - Can't stifle software progress along the way

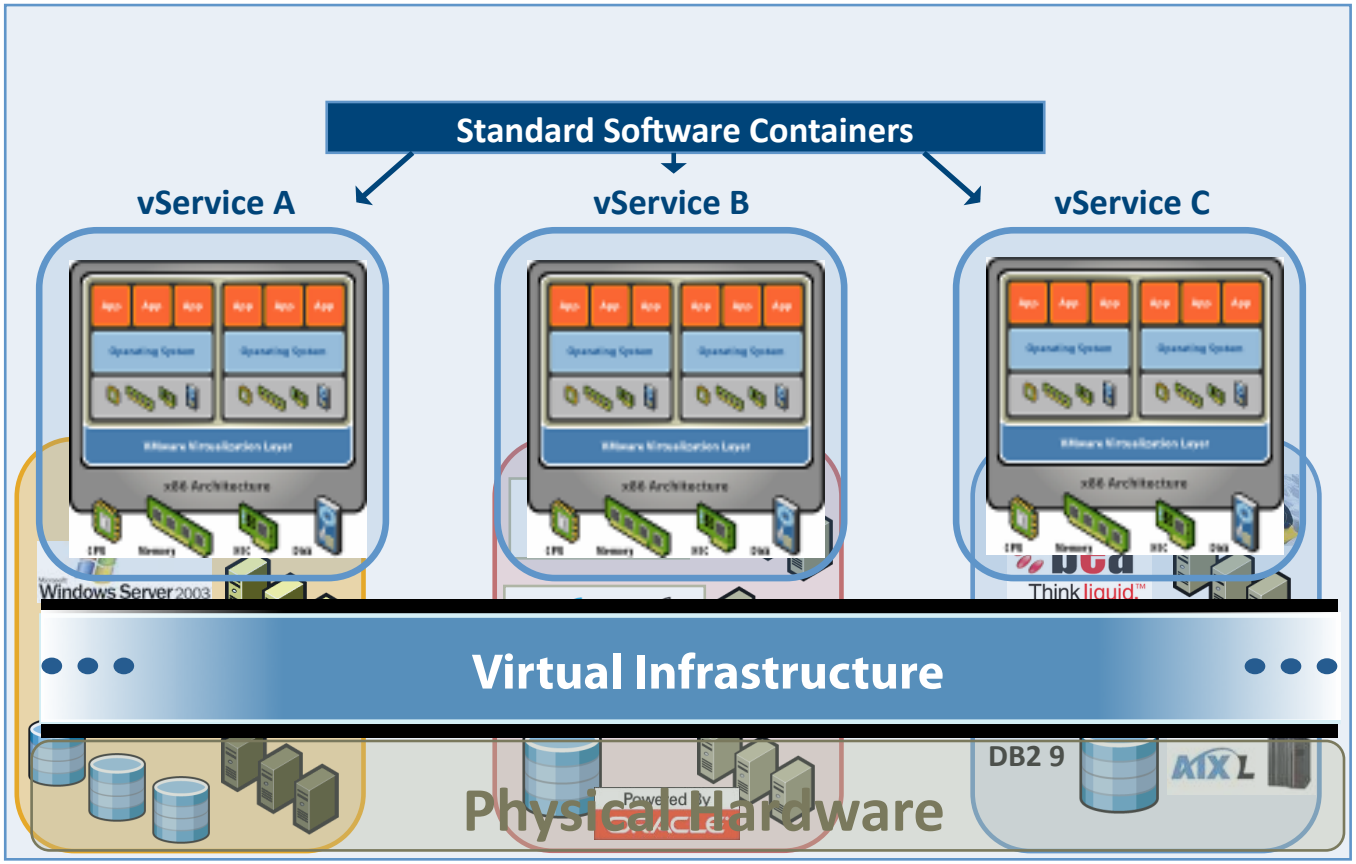


Software's Growing Complexity



Different processes must be executed differently depending on the hardware, Operating System, or application environment

Legacy Processes Can Be Moved



Complex application stacks become standardized software containers

With common processes for management

Next Phase: Connecting the Clouds

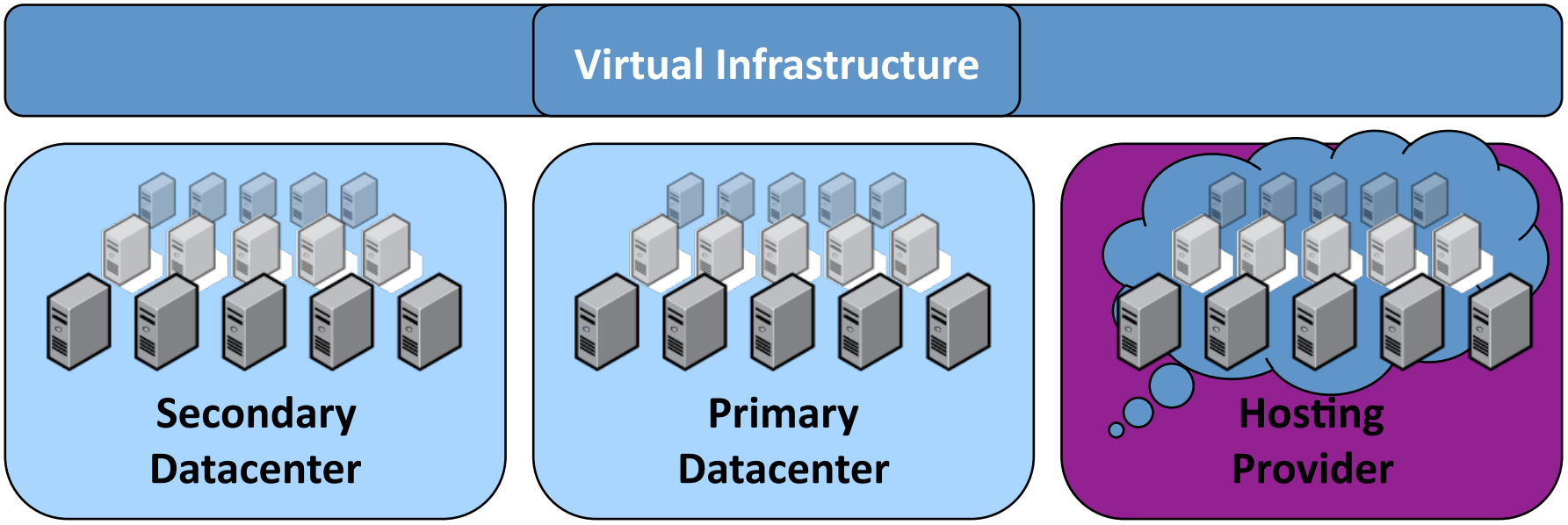
Extends Virtual Infrastructure beyond single datacenter

Uses secondary Data Center site for workload overflow

Leverages geographically distributed resources

Rents resources from Service providers for capacity overflow

Maintains IT Service Service Level Agreements



The “Cloud”

What is a Cloud?

- A “cloud” is a totally reliable, extensible, and manageable software platform that delivers a self-healing and self-managing datacenter.
- It aggregates on-premise servers, storage and network into “an internal cloud” that also federates with “external clouds” of computing capacity that frees I.T. from the constraints of hardware-mapped applications.
- The “cloud” guarantees the high levels of availability, low latency, security and scalability for all applications independent of hardware and location.

How the “Cloud” is Different

- Unlike a traditional Operating Systems, which are optimized for a single server, the virtual datacenter OS – the “cloud” - serves as the OS for the entire datacenter.
- Datacenters of cloud providers and Software as a Service companies can operate the virtual data center using cheap commodity hardware because the computing is redundant.

A Virtual Data Center Delivers

- A set of infrastructure services to seamlessly aggregate on-premise servers, storage and network
- A set of cloud services to federate the on-premise infrastructure with third party cloud infrastructure
- A set of application services to guarantee the right levels of availability, security and scalability to all applications independent of hardware and location.
- A set of management services that allow to proactively manage the virtual datacenter and the applications running on it.

Evolution Towards a "Cloud"

- Virtualization it has evolved to include not just a hypervisor but a comprehensive set of capabilities that minimize planned and unplanned downtime, ensure application service levels and automate routine processes in the datacenter.
- Virtualization is an enabling technology that frees up applications from physical servers and enables hardware independence.
- Virtualization is now evolving to a dynamic datacenter with infrastructure services for aggregating compute capacity and for automated restart of redundant datacenters affected by software failure.

Fault Tolerance

- Solutions to deliver zero downtime, zero data loss and +Six Sigma availability.
- Assure execution of applications based on chosen microprocessors.
- Deliver fault tolerance without the cost and complexity of hardware duplication.
- Quick, simple and cost effective backup as well as recovery for all applications.

Security & Control

- Offers visibility into all machine resources and processes.
- Monitors and controls the execution of all applications.
- Stops viruses, rootkits and malware before they can infect a system.

Scalability

- Enables resource intensive applications to run up to 200,000 I/Os per second.
- Offers hot add capability for increasing hardware resources.
- Enables applications to scale without disruption or planned downtime.

Storage Management

- Dynamic storage allocation enables the reduction of storage by >50%.
- Uses memory only as required and tracks actual usage for billing.
- Reduces storage by sharing common user data.
- Links all storage for managing capacity and improving performance.

Economics

- “Cloud” data centers can be billed in actual usage increments.
- The infrastructure can obtain processing power from “clouds” of external vendors.
- Capacity can be distributed across several data centers for handling of peak loads.

“Cloud” or “Utility” Computing

- The networked infrastructure provides capacity in an on-demand environment.
- “Cloud” computing offers the capacity to pay only for actual usage, as a “utility”.
- We prefer the term “utility” computing to describe what is described as a “cloud”.

Megatrends

- Progress from a device centric world to a person centric world.
- Enables connecting from anywhere, by any means.
- Offers access privileges only to authorized persons.
- Allows purchasing of computer processing power independent of circuit technology.
- Makes it possible to associate computing services according to a person's roles or location.

Directions

- Virtual Infrastructure aggregates hardware into flexible resource pools and creates a truly virtual datacenter
- Virtualization simplifies software management, and extends this even further with central management of software configuration
- Virtual mobility, virtual infrastructure, enables effective use of resources that are local or remote, owned or rented

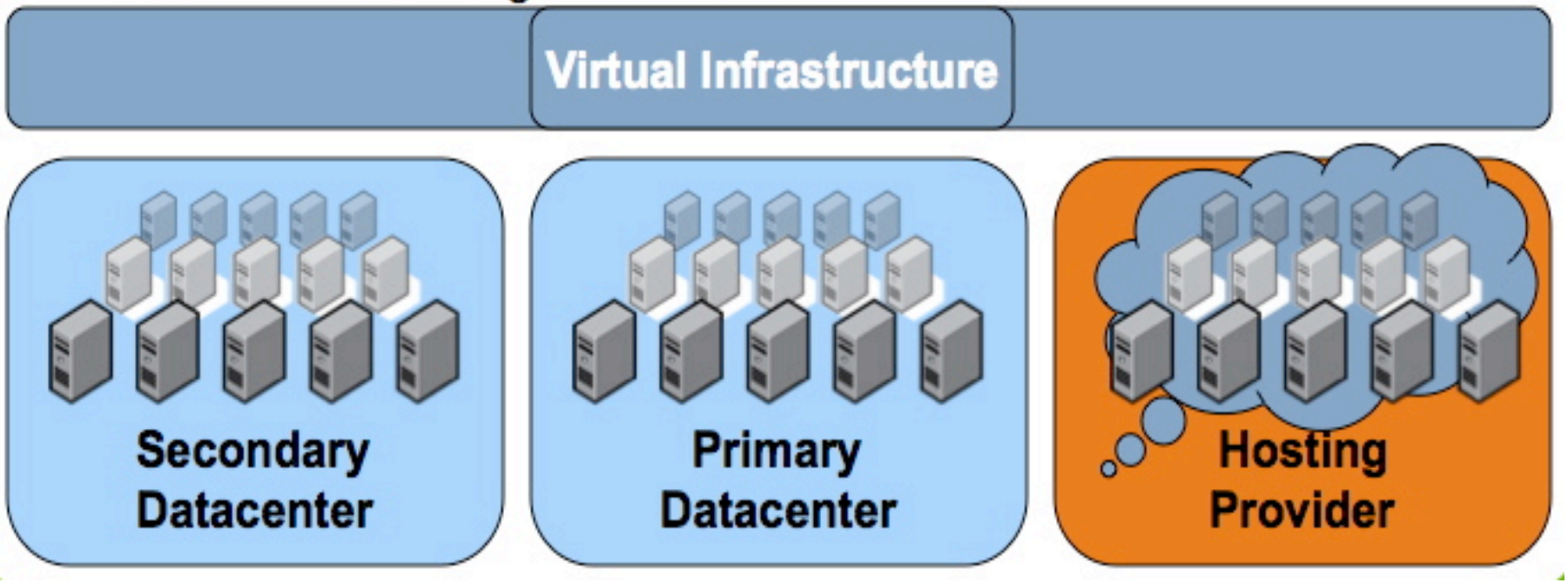
Virtualization to Multiple Datacenters30

Use secondary DR site for workload overflow

Leverage other departments'/LOBs' resources

Rent resources from hosters when appropriate

All while maintaining the IT Service SLA



Summary

- The current proliferation of servers has been costly, unreliable and insecure.
- It has decreased the quality of service.
- Clustered servers cannot offer the fault tolerance, scalability and the favorable economics of “Cloud” computing.
- Migration to a Service Oriented Architecture must include the “Cloud” in its architecture.

Amazon Elastic Compute Cloud (Amazon EC2)

What is a Cloud Service?

- Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the “cloud”. It is designed to make web-scale computing easier for developers.

Example of a Service: Elastic Block Store (EBS)

- Provides block level storage volumes for use with EC2 applications.
- EBS volumes storage that persists independently from the life of an applications.
- EBS is particularly suited for applications that require a database, file system, or access to raw block level storage.

Software Available on EC2

Category: Amazon Elastic Compute Cloud

Documents in Category **Go to Page: [1 2 3 | Next]**

Sort by: Rating - Best Filter by Document Type: Developer Tools

C# Library for Amazon EC2
C# Library for Amazon EC2
Last Modified: Oct 23, 2008 5:14 AM

typica - A java client library for a variety of Amazon Web Services
A Java client library for Amazon SQS, Amazon EC2, Amazon SimpleDB and Amazon DevPay
Last Modified: Sep 21, 2008 2:48 PM

Simple Command-Line Access to Amazon EC2 and Amazon S3
AWS developer Timothy Kay put together this command-line tool that provides easy command-line access to Amazon EC2 and Amazon S3. The tool is designed to be simple to install and simple to use.
Last Modified: Sep 21, 2008 2:48 PM

RightScale Ruby library to access Amazon EC2, S3, SQS, and SDB
RightScale's AWS gems provide Ruby interfaces to four key Amazon Web Services: EC2, S3, SQS and SDB. An optional robust HTTP layer retries and clears transient errors.
Last Modified: Sep 21, 2008 2:02 PM

Code for Mobile Monitoring and Administration of Amazon Web Services
The Apple iPhone has been one of the most anticipated mobile computing platforms in years. It provides a web browser experience previously found only on traditional computing platforms. This project demonstrates the development of an Amazon EC2/Amazon S3 monitoring and administration application, optimized for the iPhone, using familiar web development tools—PHP and JavaScript.
Last Modified: Sep 21, 2008 1:39 PM

Usage Pricing

High CPU Instances	Linux/UNIX	Windows
Medium	$0.20 per hour	$0.30 per hour
Extra Large	$0.80 per hour	$1.20 per hour

Data Transfer In	
All Data Transfer	$0.10 per GB

Data Transfer Out	
First 10 TB per Month	$0.17 per GB
Next 40 TB per Month	$0.13 per GB
Next 100TB per Month	$0.11 per GB
Over 150 TB per Month	$0.10 per GB

Costs Known in Advance

Amazon S3 (US)	Storage:	100	GB-months
	Data Transfer-in:	20	GB
	Data Transfer-out:	500	GB
	PUT/LIST Requests:	10000	Requests
	Other Requests:	100000	Requests

Estimate of Your Monthly Bill

Amazon S3 (US)	Storage	$ 15.00	
	Data Transfer	$ 87.00	
	Requests	$ 0.20	
	Amazon S3 (US) Bill:		$ 102.20

Summary: Key Features of EC2

- Allows you to obtain and configure capacity on demand.
- Provides complete control of computing resources and lets you run in a proven computing environment.
- Reduces the time required to obtain and boot new servers in minutes, allowing to scale capacity as computing requirements change.
- Changes the economics of computing by paying only for capacity used.
- Offers tools to build failure resilient applications.