

The Emergence of Cloud Computing



Susanne M. Balle

SW Architect

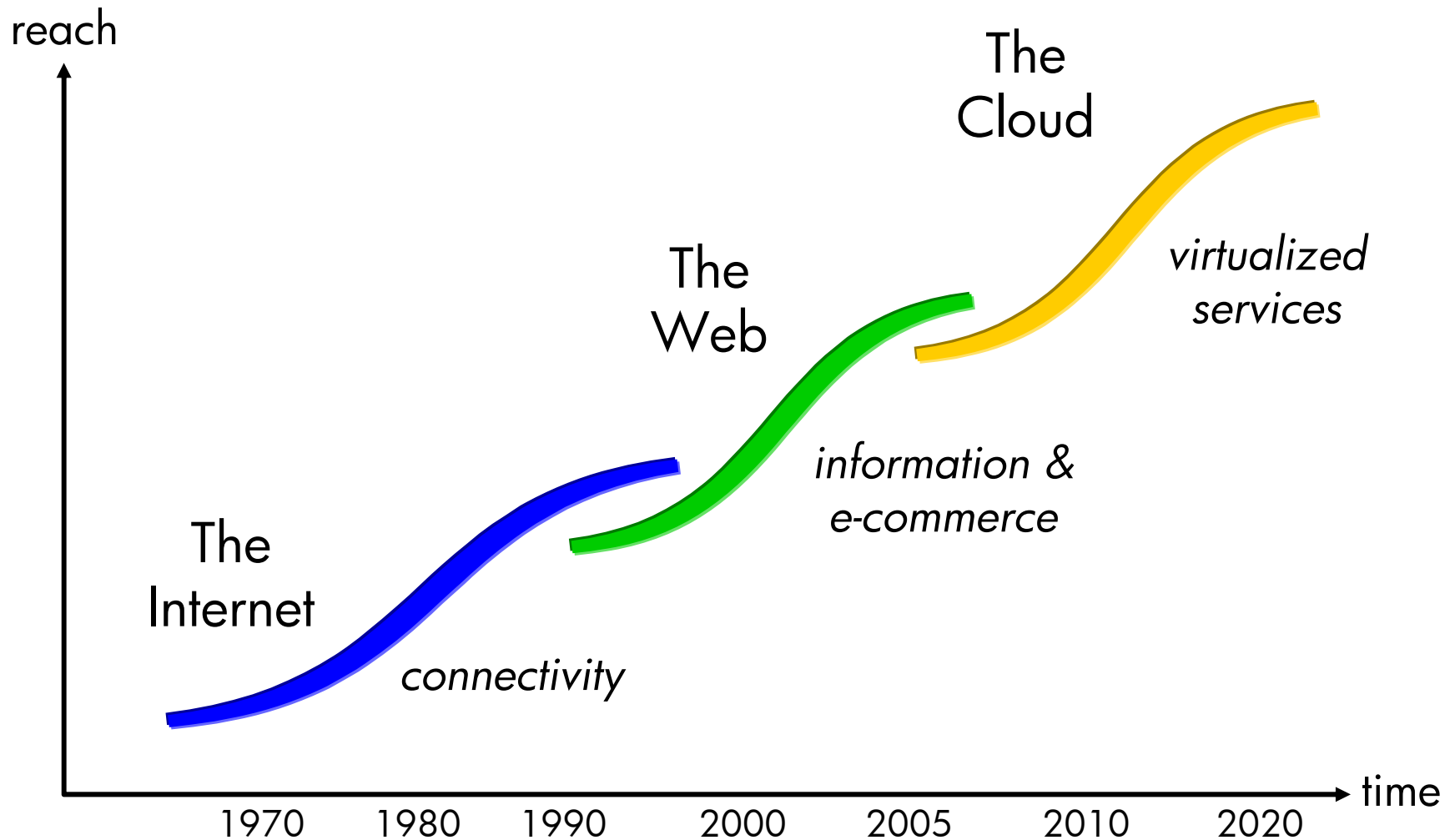
Scalable Computing and Infrastructure

Technology for better business outcomes

Outline

- Emergence of Cloud Computing
- HP and Cloud Computing
- Delivering with the biggest names in the industry
- Pioneering delivery of IT as a service
- HP/Intel/Yahoo! Cloud Computing Research TestBed

The Third Generation



So, What is Cloud Computing?

The definition is suitably vague. Some attempts:

The 451 Group: “The cloud is IT as a Service, delivered by IT resources that are independent of location”

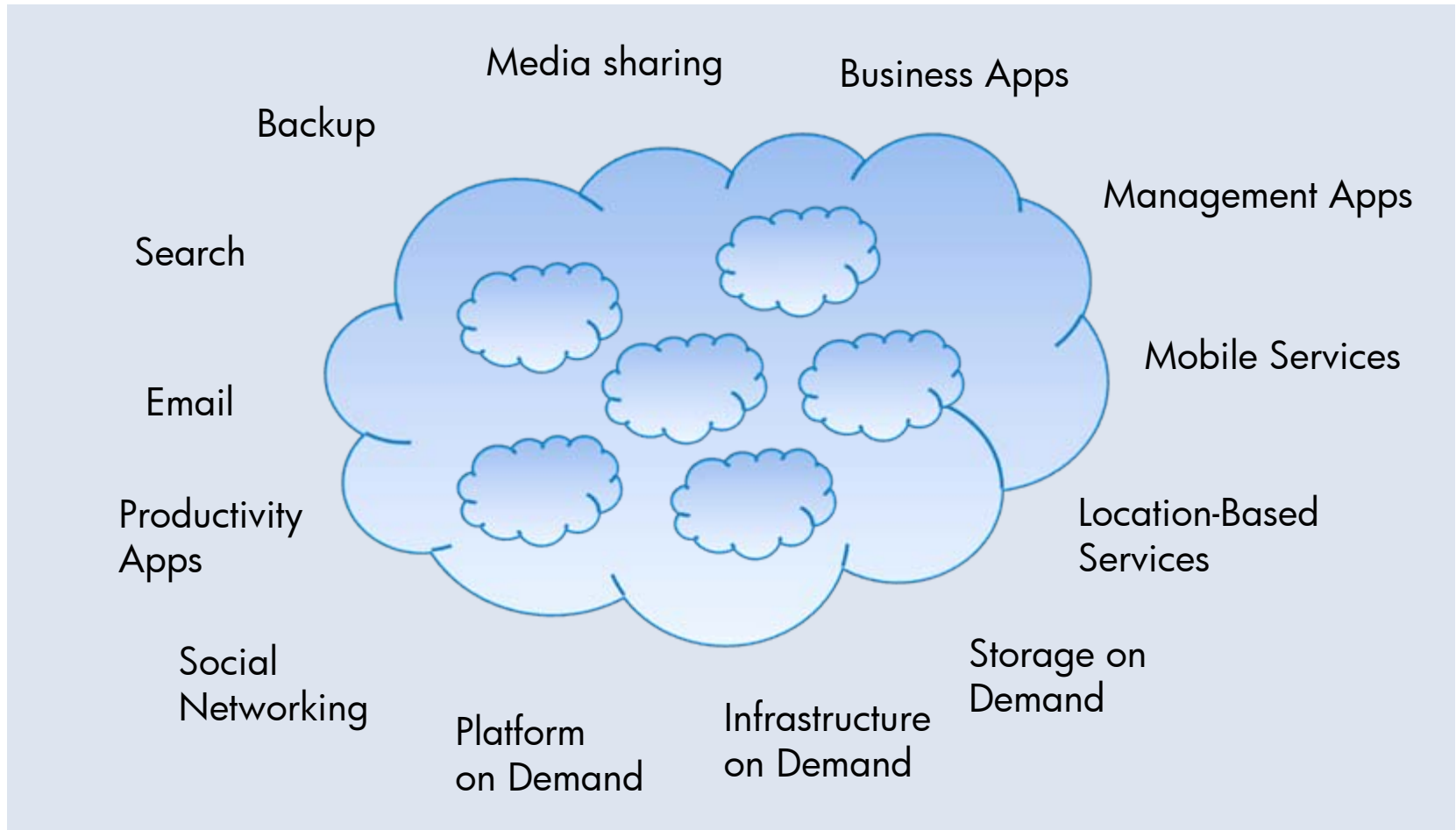
Gartner: “Cloud computing is a style of computing where massively scalable IT-related capabilities are provided ‘as a service’ across the Internet to multiple external customers”

Forrester: “A pool of abstracted, highly scalable, and managed infrastructure capable of hosting end-customer applications and billed by consumption”

Aspects of Cloud Computing

- **Massive scale-out environment** (e.g. Google, Yahoo, Amazon)
 - Highly horizontally scalable infrastructure (10-100 thousand servers)
 - Host “internet-scale” applications
 - Relentless focus on cost
- **Everything as a Service**
 - SaaS (e.g. Salesforce, GoogleApps)
 - IaaS, DaaS, PaaS, ...
 - Delivery over the Internet
- **New business models**
 - e.g. pay per use, subscription, ...
- **Users as co-developers**
 - Social networking effects (e.g. Facebook)
 - Data sources get richer as more people use them
 - Leverage large number of users to rapidly test new ideas

IT as a Service, Delivered by the Cloud

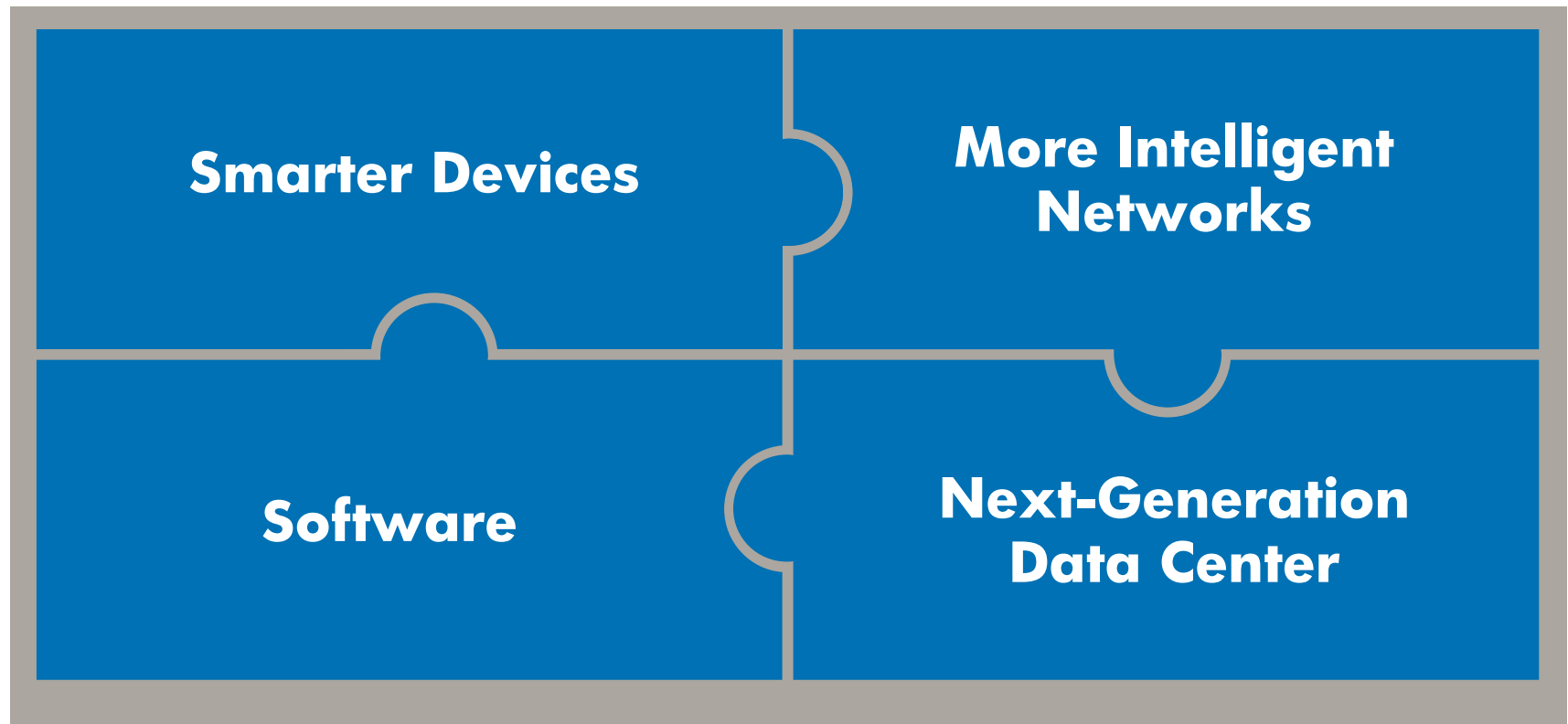


And ...

- At massive scale
 - Millions of users
- With unprecedented flexibility
 - Mash-ups, aggregation, enhancing services, flexing up and down, ...
- Offering evolving APIs to exploit and extend
- At breakthrough cost levels
 - Economies of scale
 - New revenue models
 - Eliminating old sources of cost (SaaS vs. CD)

Scalable Computing to Power the Cloud

Technology building blocks – Everything as a Service



Delivering with the Biggest Names in the Industry



Including some of the largest collaboration platforms in the world





"Home to the world's leading Internet sites for social networking, games, sports photos and more, Fox Interactive Media helps millions of people around the globe stay informed and connected through online media accessible 24 hours a day, 7 days a week.

Each month, billions of pages are viewed, hundreds of millions of photos and videos are uploaded and watched and those numbers just keep growing, constantly pushing the boundaries of our infrastructure.

For us to meet these growing demands we need high-performance platforms that minimize energy consumption and have high density and capacity. HP's computing team understands our unique requirements and offers us end-to-end cloud computing and storage solutions that allow us to more easily scale as we grow."

Ian Hurlock-Jones, Chief Information Officer, Fox Interactive Media

HP ProLiant BL2x220c G5

Doubling down with the world's first 2-in-1 server blade

2x compute while cutting data center blade infrastructure cost in half

- First industry standard server w/1024 cores per standard 42u rack
- Lowest cost per GigaFLOPS/s
- Up to 12.3 TeraFLOPS/s per rack

Designed for maximum power & cooling efficiency

- 60% more work per watt than competitive product based on specjbb business operations per watt. 5.3KW per 32 nodes at 100% utilization – compared to ~7400W to competition



Broad choice of industry-standard servers, interconnects, Operating Systems

- Choice: processor architectures, interconnects (i.e. 1 GigE, 10GigE, IB), OS (Linux and Windows)

Builds on enormous success of HP BladeSystem c-Class

- Market share leader in x86 blades and ~35% share of TOP500 supercomputing list
- At 232 MegaFLOPS/s/watt – equivalent to #6 on the current Green500* list

Scale-out computing just got 2x as good!

EKA: One of the world fastest supercomputer

- Industry standard parts
- Computational Research Laboratories (CRL), Tata Sons Ltd. (India)
 - First completely privately funded shared resource HPC center in the world*
- 175 Tflops peak performance
 - HP Cluster Platform 3000 BL460c
 - Over 1800 HP BladeSystem c-Class server blades
 - 14,400 processors, 28 terabytes of memory, 140 terabytes of storage
- Yahoo and CRL agreement to jointly support cloud computing research

Substantially more processors than any other supercomputer available for cloud computing research



Defense Information Systems Agency (DISA)

US Department of Defense Cloud Computing Infrastructure



Objective	Approach	Results
<ul style="list-style-type: none">• Rapid access to computing resources• Eliminate the need to procure physical infrastructure• Self service portal through a single, secure interface• User configurable server environments• Automated provisioning• Flexible billing options• Meet DoD security requirements	<ul style="list-style-type: none">• HP C&I development of shared services utility for Rapid Access Computing Environment (RACE)• HP Server Automation and HP Operations Orchestration for provisioning and configuration management• HP Operations Manager for monitoring and control• HP Service Manager to automate incident & problem management• HP Systems Insight Manager and HP Proliant Essentials• Cluster Resources Moab for intelligent orchestration and Gold for billing	<p>Business outcomes</p> <ul style="list-style-type: none">• Reduced costs• Consolidated simplified processes• Shortened time to delivery <p>IT improvements</p> <ul style="list-style-type: none">• Flexible development platforms for Web, Application or Database• User can allocate own resources through Web interface• Can provision a server in a few minutes• CPU, memory, storage, virtual environment provided in one simple solution



Uncertainty and Agility

Challenge – UNCERTAINTY

“UNCERTAINTY is the defining characteristic of today’s strategic environment.”

(National Defense Strategy)

- Adjust to an era of surprise and uncertainty

Response – AGILITY

“We have set about making US forces more AGILE and more expeditionary.”

(Quadrennial Defense Review)

- Enterprise-wide: Battlefield Applications; Defense Operations; Intelligence Functions; Business Processes
- Emphasis Shift: From moving the user to the data – to moving data to the user

Confront Uncertainty with Agility



Adaptive IT Infrastructure for Cloud Computing

- Virtualized Environment – Fast access to computing resources
- Rapid provisioning & repurposing across multiple data centers
- Self-service web site
- Multi-vendor, multi OS
- Policy-driven Task Automation
- Single point of management and user access
- Dynamic inventory
- Billing and chargeback for use
- Platform for sharing IT services across the enterprise (multiple data centers)
- Reduce space & environment requirements
- Manage power and cooling
- Audit and verification

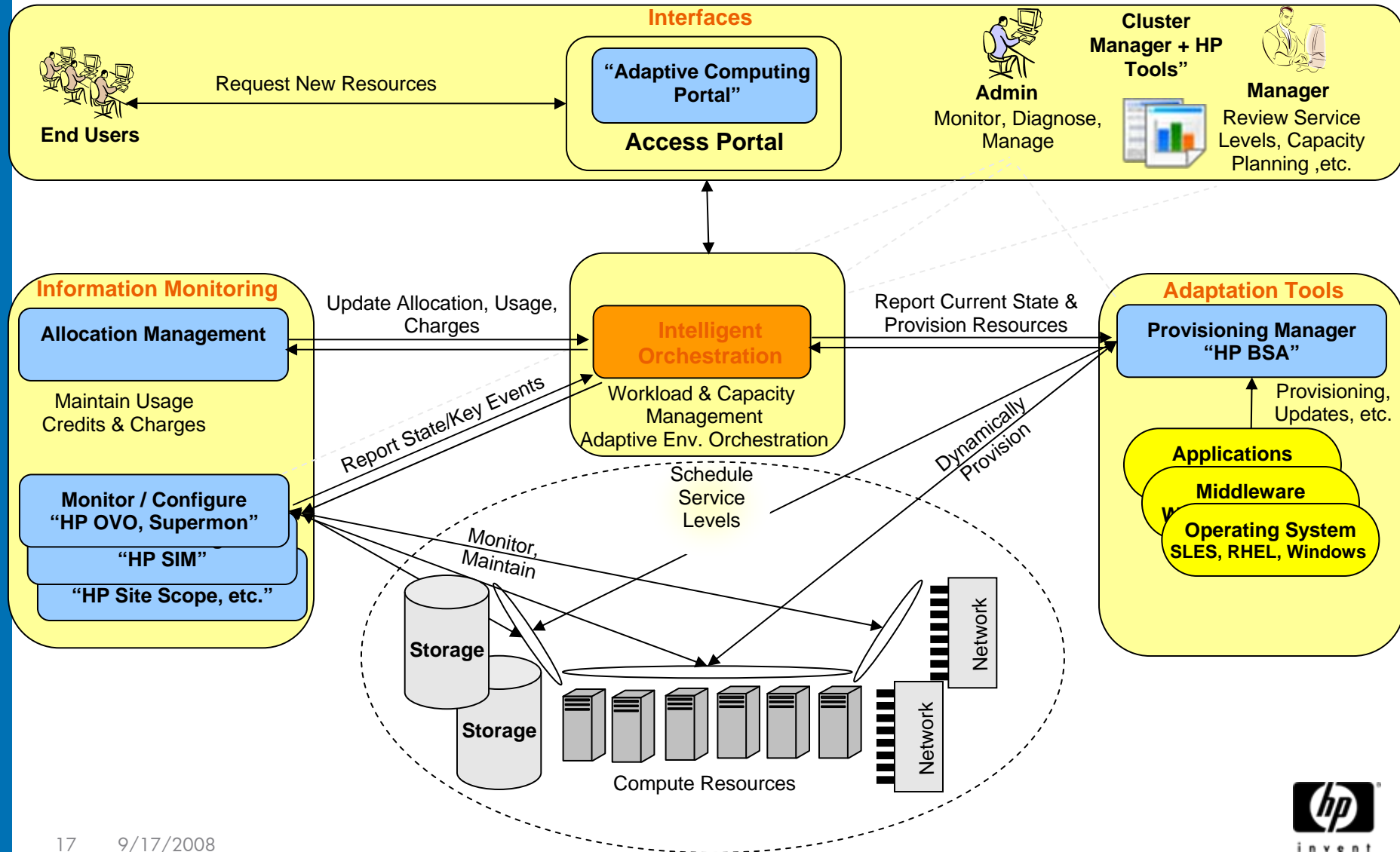
Mission Value

- Reduce costs / Consolidate / Shorten time to delivery / Simplify
- From cost centre to IT business shared services (chargeback)

Platform as a Service

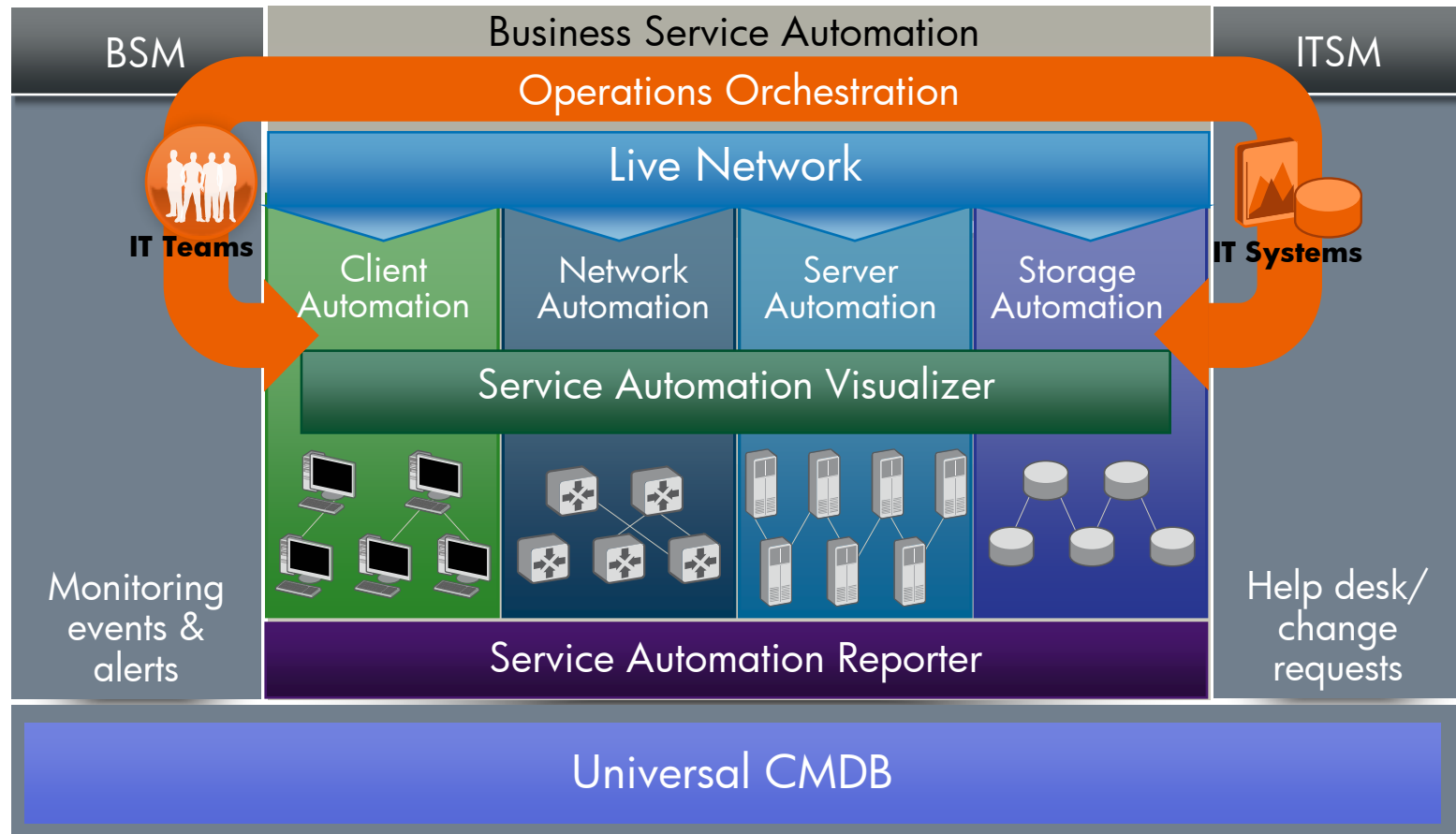
- Self Service Provisioning and Secure web-portal
- Create virtual private resource pools
- Pre-defined infrastructure templates (includes HW + software and application stacks)
 - Defined by administrators
 - Customizable by users
- Architecture-independent
 - Support bare-metal and virtualization based provisioning
- Integrated with accounting

Platform as a Service: Instantly Create Project and Test Environments



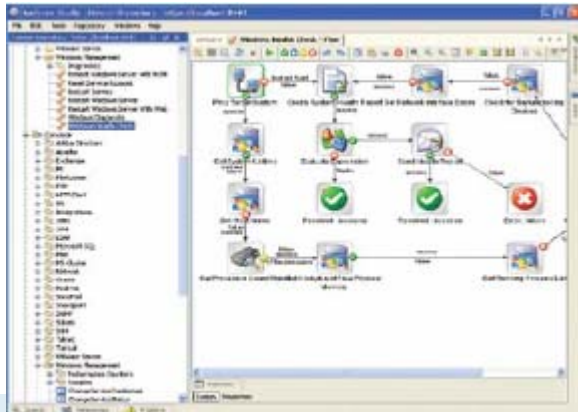
HP Business Service Automation (BSA)

Automate management of client, server, software, network, storage & process



Resource management Lifecycle

Operations Orchestration



Operations Orchestration (Opware Process Automation System™ PAS) is an innovative run book automation platform for IT Operations.

Cost Efficiency

Enables IT departments to eliminate manual and error-prone procedures by automating end-to-end IT processes

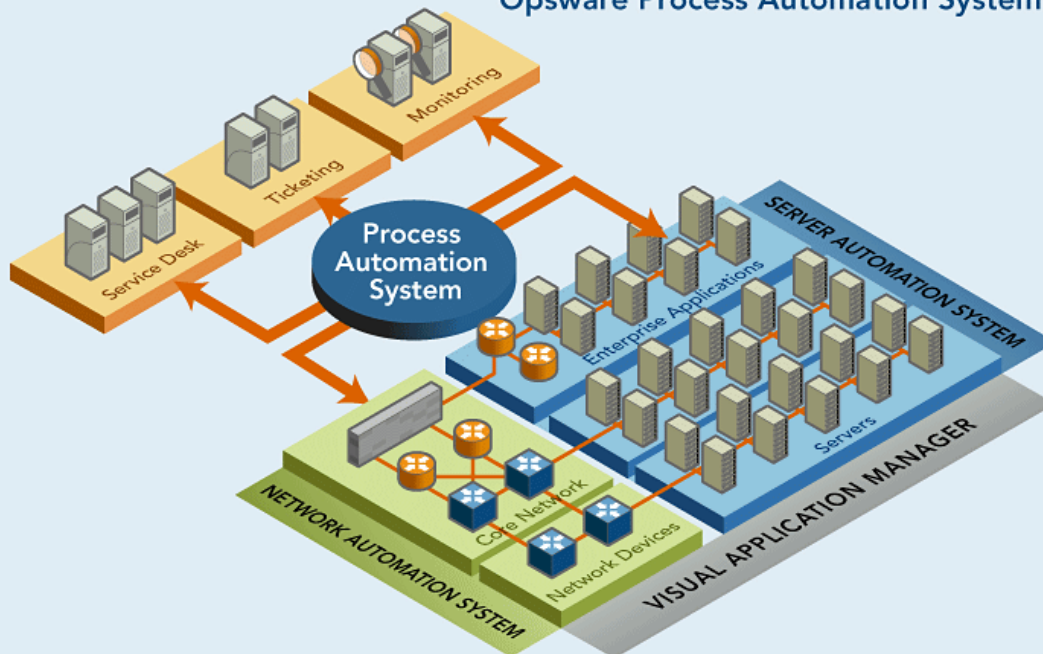
Quality and Consistency

Reduces server and application downtime by eliminating the inconsistency and misconfigurations that are the root cause of most problems

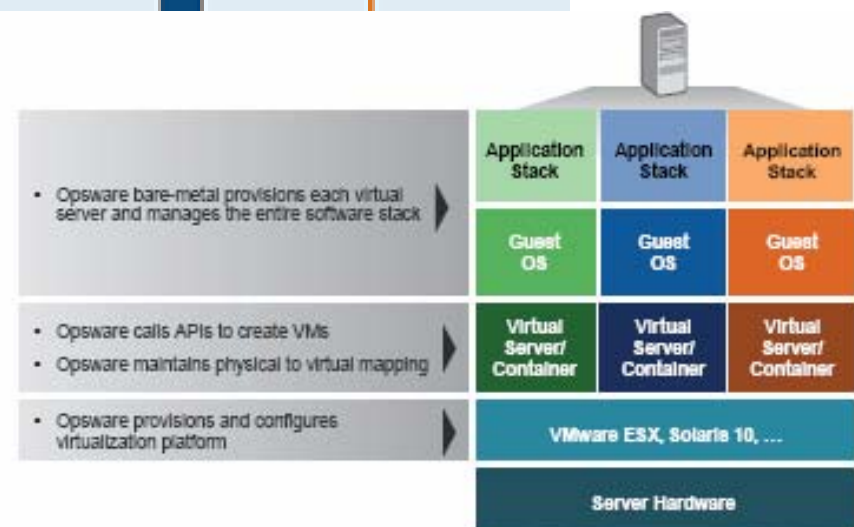
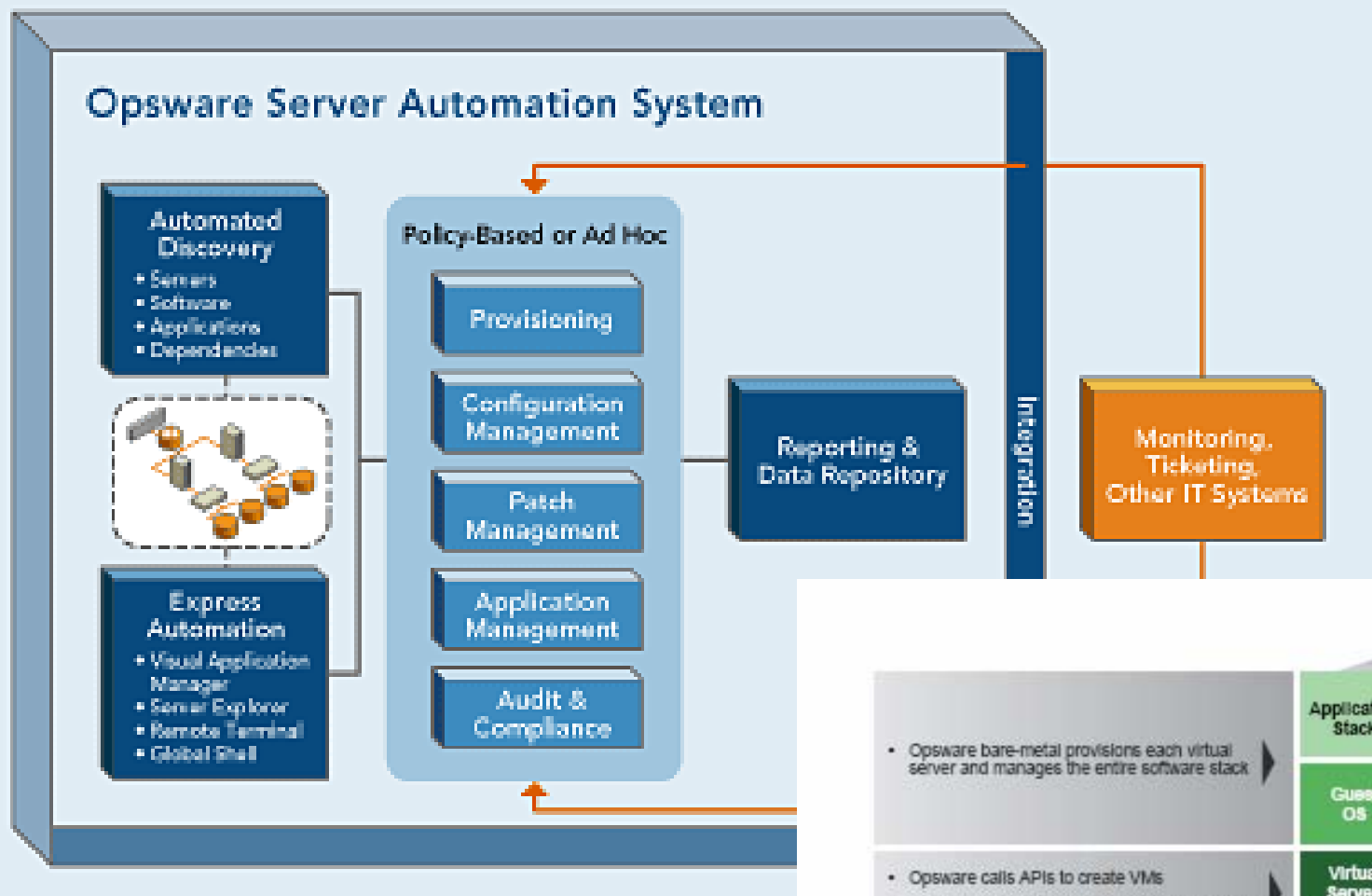
Agility and Responsiveness

Allows IT departments to react more quickly to changing business needs by reducing the time to deploy new infrastructure

Opware Process Automation System



Server Automation



HP pioneering delivery of IT as a service

HP Flexible Computing Services

Secure utility computing services that interoperate with customers' in-house data centers

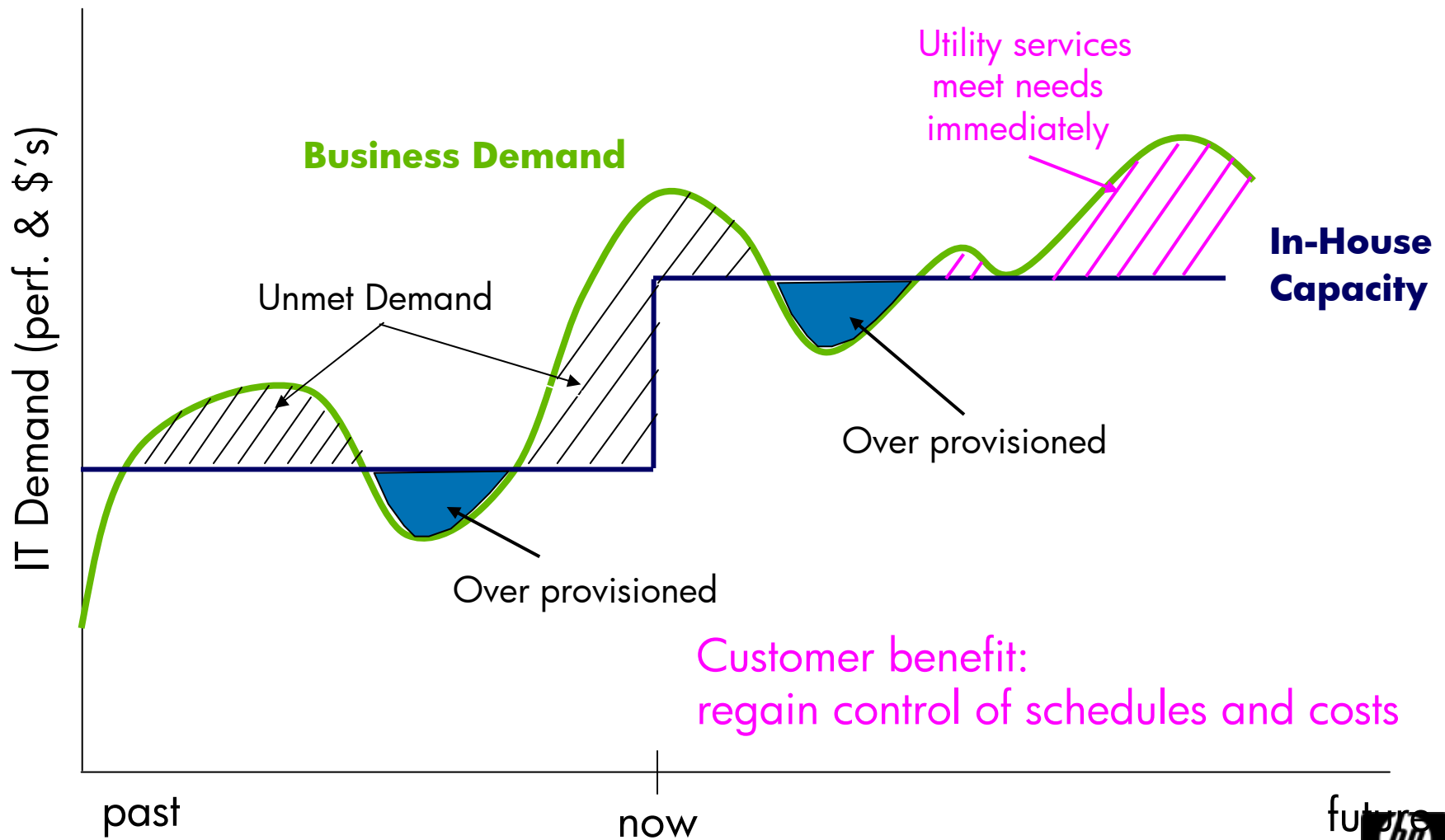
HP Adaptive Infrastructure as a Service

Delivers Microsoft® Exchange, SAP® applications and other critical business applications

HP Software as a Service

Managed service offering of HP Software Products, delivered over the internet

HP Flexible Computing Services Complements In-House Capacity

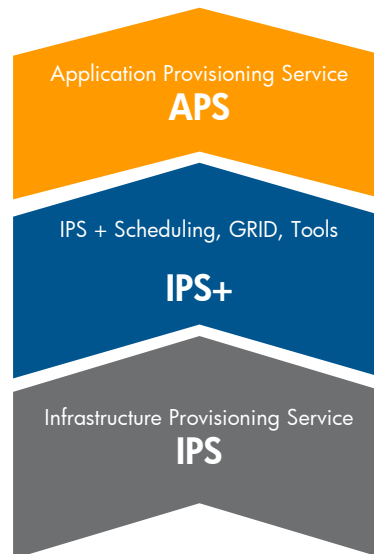


HP Flexible Computing Services



- A pay-per-use public computing utility, charged on a "cpu-hour" basis
- HP-owned assets housed in HP's highly secure data centers, dedicated per customer for the time reserved
- Choice of standards-based technologies and operating systems
- Access to industry leading applications and middleware on a utility basis

Customers
choose how to
plug in



Includes IPS+, and HP installs [to a default configuration] application software

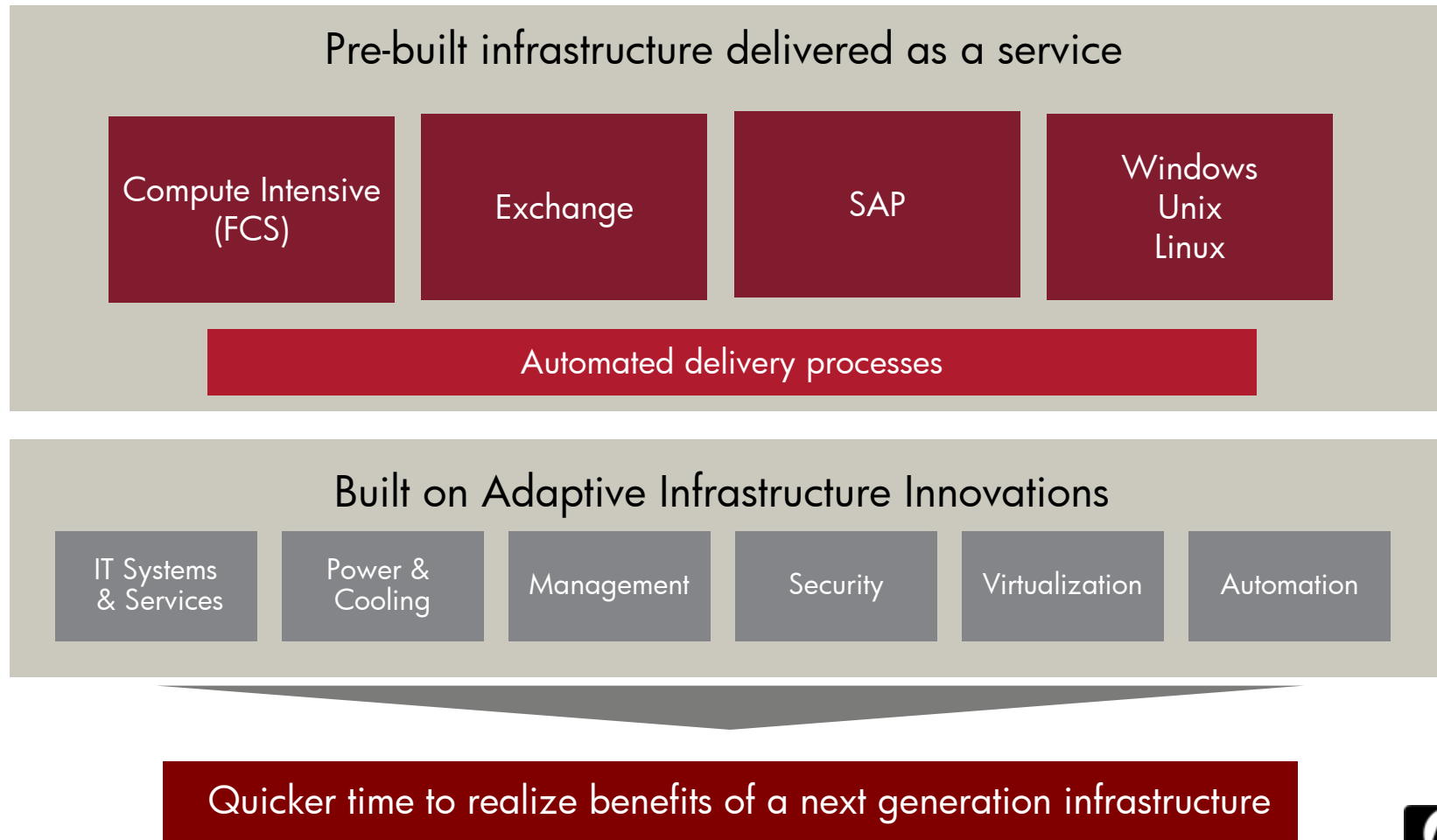
Same as IPS, but HP installs & manages scheduling, compiler, and/or grid software.

Basic infrastructure service. HP provides & manages hardware infrastructure and OS. Customer provides & manages applications.

CPU, Storage, OS, Network, Maintenance

HP Adaptive Infrastructure as a Service

Extending HP FCS - A new approach to outsourcing



HP, Intel, Yahoo Join Government, Academia In Cloud Computing Research

Each of the founding members will host a cloud-computing infrastructure largely based on HP computers and Intel processors in six data centers.

By Antone Gonsalves, [InformationWeek](#)

July 29, 2008

URL: <http://www.informationweek.com/story/showArticle.jhtml?articleID=209800449>

Hewlett-Packard, Intel, and Yahoo on Tuesday said they have joined government and academia in launching a global, multi-data center test bed for experimentation and research in cloud computing, which many experts believe will be the dominant IT delivery model of the future.

The [initiative aims at building a computing network](#) comprised of six data centers spanning three continents. The idea is to have a large-scale [platform](#) for testing all technology -- hardware and [software](#) -- related to delivering application services over the Internet.

"This is a global collaboration that spans the industry, spans academia and government," Prith Banerjee, senior VP for research at HP, told reporters during a teleconference held by the three founding companies.

Open Cloud Computing Research Testbed

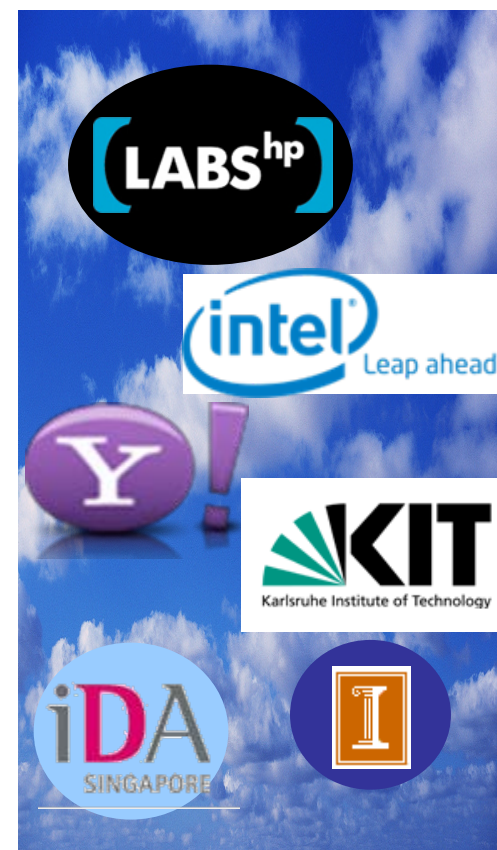
What is it?

- An *open*, scalable, secure, large scale global testbed for cloud computing research
 - Collaborative research focused on data center management & cloud services
 - Scalable, multi-continent, multi-datacenter, cloud computing system
 - A broad base research community including HP, Intel, Yahoo, and other partners and academic researchers

Open Cloud Computing Research Testbed

What is it?

- A loose federation of “Centers of Excellence” around the globe
 - UIUC, Singapore IDA, KIT: 3 initial CoE
 - HP, Intel, Yahoo: 3 initial sponsors with CoE
 - One common research theme
 - Specific research germane to each region
- Sponsors
 - HP Labs, Intel Research, Yahoo!
- Each center: 1000-4000 cores and up to PB storage
 - With experimental cloud management services



Cloud Computing Research @ HPL

Fulfill the “Everything as a service” vision

- **Cells as a Service:** Service Cells are secure “containers” for virtual infrastructure elements. Each Cell can contain an arbitrary assembly of virtual machines, virtual storage volumes and virtual networks
- **Exascale Data Center:** Aims to radically redesign the future data center and its computing components, delivering significantly improved customer experience at dramatically lower costs
- **Scalable Storage:** A versatile automatic enterprise storage service that can scale, across geographically dispersed locations
- **Service Lifecycle Management:** A management system for converting any type of application into a service
- **Sustainable Data Center:** Focus on developing sustainable IT approaches that are more energy efficient and will reduce the environmental impact

