Cloud Management:

Nigel Cook, HP ISS
Dejan Milojicic, HP Labs
Vanish Talwar, HP Labs
SIMPLY EXPLAINED - PART 17: CLOUD COMPUTING
Why is Management important?

Efficiency, SLAs & costs!
WW Spending on Servers, Power and Cooling, and Management/Administration

- **Power & Cooling**
- **Mgmt & Administration**
- **New Server Spending**

**Workloads View**

$0, $25,000, $50,000, $75,000, $100,000, $125,000, $150,000, $175,000, $200,000

'96, '97, '98, '99, '00, '01, '02, '03, '04, '05, '06, '07, '08, '09, '10, '11, '12

Installed Base:
- 45,000,000
- 40,000,000
- 35,000,000
- 30,000,000
- 25,000,000
- 20,000,000
- 15,000,000
- 10,000,000
- 5,000,000
- 0

©2009 IDC | 12
Why is management important?

Amazon Outage Crashes Reddit, Quora, and Other Websites

By David Daw, PCWorld   Apr 21, 2011 11:05 AM

A disruption at one of Amazon’s datacenters has led to service disruptions in the company’s EC2 or Elastic Cloud computing service. The outage, which started at 1:41 a.m. PDT, in turn brought down major websites such as Reddit and Foursquare.

Amazon’s EC2 service allows business to rent processing time on Amazon’s computers, which is useful for sites that need a lot of processing power. The disruption seriously affected sites that are constantly working through a lot of data. Dynamic sites like Hootsuite and Quora that are constantly processing new data are down completely. Other sites remain up but are still affected; Reddit, for instance, is operating in read-only mode as it cannot process new stories or votes.

VMware Cloud Foundry Suffers Service Outage

As the beta development platform was recovering from a minor power supply problem, human error worsened the setback.

By Charles Babcock InformationWeek   May 04, 2011 10:20 AM

Assembling a distributed computing architecture in the cloud isn’t easy to do. The more resources you try to bring together, the more that can go wrong. No, you’re not hearing about Amazon’s recent EC2 outage again. This time it’s VMware.

VMware recently launched a development platform as a set of services in its CloudFoundry.org, a new developer’s hosting service. On April 25, the Cloud Foundry experienced service disruption. In trying to recover later that day, it suffered an outage that continued into April 26.
1. Cloud Management Introduction

2. State of the Art
   - Managing Private Clouds: CloudSystem Matrix
   - Managing Public Clouds: Internet Data Centers
   - Managing HPC in the Clouds: Towards Exascale

3. IT Industry Trends

4. Future of Cloud Management: Requirements and Research Challenges

5. Summary
Management Definition

The collective processes of deployment, configuration, optimization, and administration during the lifecycle of IT systems and services
Life cycle of a Managed Object

- Bring-up
- Operation
- Failure/Changes
- Retire/Shutdown
Levels of Management

1. Cloud services
2. Services
3. Servers
4. OSes
5. Systems
6. Platforms
7. Virtualization
8. Automation
9. Storage
10. Networking
11. Systems
12. Business apps service mgmt
13. Enterprise sys. mgmt
14. Data center mgmt

HP
The multiple aspects of management

- Overlapping Management Domains
  - BSM: Business Service Management
  - ITSM: IT Service Management
- Evolution of Management Functionality:
  - Integrated Functionality
  - Automation
  - SOA
  - Security
  - Info Management
- Manageability of Infrastructure Elements:
  - Servers
  - Storage
  - Networks
  - Other Elements
Managing Clouds and Cloud Services

- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)
- Physical Resources

Hybrid Clouds

Private Cloud

Public Cloud

Firewall
1. Cloud Management Introduction
2. State of the Art
   - Managing Private Clouds: CloudSystem Matrix
   - Managing Public Clouds: Internet Data Centers
   - Managing HPC in the Clouds: Towards Exascale
3. IT Industry Trends
4. Future of Cloud Management: Requirements and Research Challenges
5. Summary
Private (Enterprise) Clouds

- **Virtualization management** provides life-cycle mgmt of a set of virtualized resources
  - examples: VMware vCenter, MS System Center VM Manager, or HP Insight Control.
  - can be applied to private and public cloud, e.g. Open-Stack or Amazon EC2
  - manages life-cycle of virtualized resources, and info about resource consumption

- **Cloud service composition** manages composition to aggregate cloud service infrastruc.
  - uses a heterogeneous mix of virtualized resources to realize a service offering
  - requires a model of service components, relationships, and capacity
  - Scheduling algorithms take account of QoS, (availability, isolation, compliance)

- **Application management models** business app and relationship to the infrastructure
  - infrastructure needs of application in life-cycle, or due to varying workload demand
  - during development phase, app may reside on a testbed on a public cloud;
    during production it may reside both on private and one or more external clouds
  - while app is running, responsiveness is monitored, and scaling adjustments made

- **QoS management**: scaling of app instances, maintenance cycles, and automatically removing unneeded capacity from a service
  - application conform to patterns supported by PaaS layer; as a result PaaS manages scalability and availability rather than each application development team
Self Service at the top of Service Management

Templates

Cloud APIs (personalize)
- Billing
- Approval
- Processes

Resource Pool (Dynamically provisioned)
- network
- storage
- compute, etc.

Capacity Optimization
- shift workload
- consolidation
- cut energy, etc.

Self-Service portal

Data Warehouse
- Mail
- File, Print
- CRM
- ERP
Service Templates & Workflows

Components
- Physical Server Group
- Virtual Server Group
- Physical Storage
- Virtual Storage
- Network

Existing Templates
- Published (20)
- Working (1)
- New Templates (1)

Pustak Portal
- Published
- Validation Status
- Notes: Consumers can use templates to store and restore images.
Private (enterprise) Cloud Challenges

• For the composition layer: *Algorithms for distributed placement and scheduling of virtualized resources* into the distributed capacity pools, particularly for requests targeted at times in the future.

• For the application management and scaling, a key issue is understanding the *scaling model of an application*, and interpreting the *root cause of application service level changes*.

• *Automated elasticity and SLA guarantees, security, and availability in shared environments* are hard to support.

• *Unified and integrated management across compute, storage, and network* does not exist, preventing end-to-end management of applications and cloud services.

• *Federated management across clouds instances* is hard to achieve for independently managed private clouds.
1. Cloud Management Introduction

2. State of the Art
   - Managing Private Clouds: CloudSystem Matrix
   - Managing Public Clouds: Internet Data Centers
   - Managing HPC in the Clouds: Towards Exascale

3. IT Industry Trends

4. Future of Cloud Management: Requirements and Research Challenges

5. Summary
Public Clouds

- **Massive scale in terms of users, machines, data.** Hundreds of thousands of machines
- Challenges for automated deployment, load balancing, fail-over, capacity planning
- Constraints for CAPEX/OPEX, sustainability, SLOs
- System logging, monitoring, and analysis for abnormal system behavior
- Growth of data -> dynamic partitioning, migration, replication, archiving and backup

- **Services built from open source frameworks** reduce time to market and ensure scalability: Hadoop, Cassandra, Thrift, Storm, Hive, HBase, MySQL, PHP, Flume, etc.
- Automated configuration management across multiple tools is a challenge.
- Gluing together pieces written by different developers requires careful integration
- Ad hoc processes for configuration, tuning, end-to-end diagnosis, e2e self-mgmt

- **DevOps,** developer and sys-admin operations are merging: Internet companies develop in-house services and operations is also done by in-house system administrators
- In contrast to previous models where software used to be packaged and shipped.
- Releases happen more frequently (weekly/daily) do not require physical packaging.
- Functions: modeling/configuration mgmt; infrastructure provisioning; app deployment; infrastructure/app monitoring; embedded WLM
The DEVOPS role

Agile development, continuous integration, agile operations, infrastructure as code

With increase in complexity of composite web application development and the frequency and velocity of release and deployment, a new cross department function has evolved.

**Dev Ops**
- Breaking through the walls that traditionally separate software developers and IT operations
- Enabling the agile releases of today’s applications across hybrid cloud environments
- Bringing together how applications are deployed and monitored on top of dynamic IT infrastructure
Application Release & Management disciplines

**Package**
Model driven, platform agnostic application model that can be deployed in different Hybrid IT environments

**Publish**
Taking a composite application model and create a deployment model for a specific environment

**Provision & Deploy**
Determine optimized infrastructure for the application and deploy the application

**Workload Management**
Managing workloads in a Hybrid IT environment on business and IT policies

Feedback from Ops to Release Mngt
Feedback back to Dev and Test (PAL)
Public Cloud Challenges

- **Heterogeneity** of deployment environments, e.g. multiple infrastructure choices, databases, or hypervisors, as well as working across private and public clouds

- **Automated release and testing**, to enable stable products (as the versions of managed objects change and the deployed base grows substantially)

- **Support and documentation**, to resolve production environment issues with performance life-cycle management; enough information to enable support to identify problems and provide feedback through DevOps to developers to diagnose and fix issues

- **Modeling for automated configuration management**, to address complex configurations of service compositions

- **Maintaining stringent service level guarantees**: ensure continuous availability of global Internet services with low latency response time even in presence of flash crowds
1. Cloud Management Introduction
2. State of the Art
   • Managing Private Clouds: CloudSystem Matrix
   • Managing Public Clouds: Internet Data Centers
   • Managing HPC in the Clouds: Towards Exascale
3. IT Industry Trends
4. Future of Cloud Management: Requirements and Research Challenges
5. Summary
Managing HPC in the Clouds

• Use of clouds for HPC is growing, but it is limited to small scale, test and dev

• Amazon built a top-500 supercomputer in its cloud
  • 7,000 cores, 41.82 teraflops, 231st fastest supercomputer (at the time)
  • with Linux on Intel Xeon X5570 with a 10 Gig Ethernet interconnect.
  • de-provisioned soon after, demonstrated supercomputer at $1.60/node/hour

• At high-end HPC, US Department of Energy is preparing Exascale program, and so are governments of other countries, such as in Europe, China and Japan
  • these are boundaries of high-end HPC, and are evolving as high-end data centers
  • major differences are slower interconnects and less powerful computation nodes,
  • similarity is in power, cooling, and packaging

• In the future, clouds will contain improved interconnects, such as photonics, that will enable more HPC applications to be executed in the cloud
## HPC Evolution

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Year 2010</th>
<th>Year 2015</th>
<th>Year 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power</td>
<td>6MW</td>
<td>15MW</td>
<td>20MW</td>
</tr>
<tr>
<td>Nodes #</td>
<td>18,700</td>
<td>5,000</td>
<td>100,000</td>
</tr>
<tr>
<td>Node concurrency</td>
<td>12</td>
<td>~1,000</td>
<td>~10,000</td>
</tr>
<tr>
<td>Interconnect BW</td>
<td>1.5GB/s</td>
<td>1TB/s</td>
<td>2TB/s</td>
</tr>
<tr>
<td>MTTI</td>
<td>Day</td>
<td>~Day</td>
<td>~Day</td>
</tr>
</tbody>
</table>
HPC Applications and Target Platforms

- Of interest is differentiating which applications are best suited to which platform.
- Applications that are less latency sensitive are suited for clouds.
- A platform that can perform matching automatically will benefit HPC cloud adoption.

$ vs. Performance and Latency Sensitivity

- $ Cost
- Run on SC
- Run on Cloud
- App’s Latency Sensitivity
- GFLOPS/sec
Remaining Challenges for HPC in Cloud

- **Latency**: current interconnects deployed in cloud data centers do not offer sufficient performance for HPC applications. Photonics offers some promise for the future.

- **Cost**: to enable clouds for HPC, managing cost/pricing is essential. Existing pricing models need to be expanded for physical clusters, job submissions, future reservations.

- **Power**: as HPC grows in performance, power will continue to be one of the main obstacle. (power capping, server consolidation, migration, etc.)

- **Virtualization**: while overheads are of less concern for cloud applications, they limit use for HPC applications. For example, in HPC apps I/O virtualization is not used.

- **Security**: it will be unacceptable to execute some apps globally due to national security. Privacy and export rules limit the use to specific regions.

- **Automated management of regulatory compliance**: will be a key differentiator.
1. Cloud Management Introduction
2. State of the Art
   • Managing Private Clouds: CloudSystem Matrix
   • Managing Public Clouds: Internet Data Centers
   • Managing HPC in the Clouds: Towards Exascale
3. IT Industry Trends
4. Future of Cloud Management: Requirements and Research Challenges
5. Summary
IT Industry Trends

New products result in faster, bigger, more reliable devices, however, today we are at a point where new technology will have a lasting impact on mgmt

- **NVRAM**: persistency, low latency storage access
  - drive the need for low-latency and lightweight management stacks
  - new management models (e.g., new WBEM), hardware monitoring and tools
- **Architecture**: novel memory hierarchies, multi-core, photonics, networking
  - management stacks will need to be optimized, lightweight, and decentralized
- **Power and cooling** dominate OPEX/CAPEX.
  - to limit these costs, I/F will have to be exposed for system and app power mgmt
- **Operating systems** will get redesigned with built-in mgmt in various components
  - there will be multiple components in the architecture that will contribute to mgmt
  - integration and federation of management domains will become important
- **Data-intensive computation**, continuous production of data
  - ability to archive, and manage the data lifecycle, different from computation elasticity
  - management will be intertwined with functional support
- **New app models** (social netw., big data) require new mgmt architectures, algor.
  - this will result in new management models, which will be application-driven
1. Cloud Management Introduction
2. State of the Art
   • Managing Private Clouds: CloudSystem Matrix
   • Managing Public Clouds: Internet Data Centers
   • Managing HPC in the Clouds: Towards Exascale
3. IT Industry Trends
4. Future of Cloud Management: Requirements and Research Challenges
5. Summary
Future Cloud Management Requirements

• *Global scale* (7-8B users), mobile access by most users, elasticity at this scale

• *Ease of use* resulting in *short time-to-manage*, using visual tools, analytics, what-if analysis, predictions, etc.

• *Cost efficiency*, understanding the costs of hosting services (infrastructure, services, and business objectives)

• *Support for SLAs with multiple objectives*, ability to make tradeoffs in an easy and predictable way

• *Availability and business continuity*. Managing replication at resources and service level; trading off replication for availability

• *Automated regulatory compliance*. Due to the global nature export and privacy rules need to be verified automatically
Future Cloud Mgmt Research Challenges

- Management@scale
  - global/mobile access -> unpredictable scale up/down
  - elasticity of access -> elasticity of management
  - federation a way to address scalability, connect independently managed Clouds

- Sustainability
  - environmental awareness increasingly regulated, a requirement
  - power limitations will drive cost and scale as data centers continue to grow

- Reliability and Support
  - failure rates will increase with scale, leaving no choice but to automate support
  - support will also move away from reactive towards deferred and proactive
  - supportability and reliability will be built into the design across all layers

- QoS
  - SLA management always hard, will grow in complexity with global access,
  - a variety of standard non-standard interfaces, and different APIs for SLA mgmt
  - Multiple objectives will result in further complexity
Future Cloud Mgmt Research Challenges, cont.

- **Data management**
  - continuous generation of new data from sensors, multimedia data formats
  - ability to manage this data, and compress, deduplicate, archive, and dispose of
  - regulatory compliance, will be a huge challenge

- **Integration** of management components, and run time *composition*
  - increasingly more integrated services will result in even higher complexity
  - versioning, compatibility, coordination among multiple mgmt components

- **Quantifying Cloud Manageability**
  - checklist of manageability functions
  - number of steps to manage towards desired state
  - time to manage (including time to insight)
  - documentability (e.g. lines of management code)
  - elasticity of management (manage at scale)
  - availability and continuity of management
  - ease of use (GUls, visualization, analytics, etc.)
Research Challenges, *Management@scale*

- global/mobile access -> unpredictable scale up/down
- elasticity of access -> elasticity of management
- *federation* a way to scalability, connect independently managed Clouds

A 50-rack datacenter

- X 64 blades per rack
- X 2 sockets per blade
- X 32 cores per socket
- X 10 VMs per core
- X 10 Java objects per VM

20,480,000 managed objects

Bring-up
Tuning
Diagnostics

Operations
Retirement
Research Challenges, *Sustainability*

- Environmental awareness increasingly regulated, a requirement
- Power limitations will drive cost and scale as data centers continue to grow

### An Illustrative Dashboard

<table>
<thead>
<tr>
<th>Open Cirrus Sites</th>
<th>Economical ($)</th>
<th>Ecological</th>
<th>Social</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IT</td>
<td>cooling</td>
<td>ntwk</td>
</tr>
<tr>
<td>Site 1</td>
<td>$0.72</td>
<td>$0.35</td>
<td>$0.16</td>
</tr>
<tr>
<td>Site 2</td>
<td>$1.27</td>
<td>$0.59</td>
<td>$0.21</td>
</tr>
<tr>
<td>Site 3</td>
<td>$1.05</td>
<td>$0.47</td>
<td>$0.12</td>
</tr>
<tr>
<td>Site 4</td>
<td>$0.75</td>
<td>$0.35</td>
<td>$0.12</td>
</tr>
<tr>
<td>Site 5</td>
<td>$0.27</td>
<td>$0.13</td>
<td>$0.05</td>
</tr>
<tr>
<td>Site 6</td>
<td>$1.82</td>
<td>$0.77</td>
<td>$0.11</td>
</tr>
<tr>
<td>Site 7</td>
<td>$1.23</td>
<td>$0.54</td>
<td>$0.11</td>
</tr>
<tr>
<td>Site 8</td>
<td>$0.55</td>
<td>$0.26</td>
<td>$0.10</td>
</tr>
<tr>
<td>Site 9</td>
<td>$1.01</td>
<td>$0.44</td>
<td>$0.10</td>
</tr>
</tbody>
</table>
Research Challenges, **Reliability and Support**

- failure rates will increase with scale, leaving no choice but to automate support
- support will also move away from reactive towards deferred and proactive
- supportability and reliability will be built into the design across all layers

---

**Serviced Objects**

Customer’s environment ← HP’s environment

**Self-healing** (local)

Parts if needed

**Runbook automation break-fix**

**Customer or CE**

**Parts/CE Dispatch**

**New events**

**Known events**

**Analysis Engine**

**Trends, Failure forecast**

**Serviceability Enhancements**

**HP Support**

**Workflow / KM**

**Product R&D**

**Automated Learning**

**Actionable knowledge**

**Serviceability Analysis**

**Incident related data (internal + external)**

**Parts if needed**

**Customer’s environment** ← **HP’s environment**

---

**Serviceability Enhancements**

**Product R&D**

**Automated Learning**

**Actionable knowledge**

**Serviceability Analysis**

**Incident related data (internal + external)**

**Parts if needed**

**Customer’s environment** ← **HP’s environment**
Research Challenges, QoS

- SLA management always hard, will grow in complexity with global access,
- a variety of standard non-standard interfaces, and different APIs for SLA mgmt
- Multiple objectives will result in further complexity
Research Challenges, Data management

- continuous generation of new data from sensors, multimedia data formats
- ability to manage this data, and compress, deduplicate, archive, dispose of
- regulatory compliance, will be a huge challenge.

5 exabytes in 2002 to 280 exabytes in 2009 of online data
“Physics of data, Myers, Google”

56X in seven years

16X Moore’s law
Static rich media
e.g., wiki pages, images, video, imager, ...
Map information

- Satellite imagery
- 3D models
- Static photos
- Static web pages
- Transit
- Visualizers

Synthetic data superpose
e.g., 3D models, timelines, trendalyzers...
Real-time data

e.g., traffic, weather, news, feeds, …
Map information

- Satellite imagery
- 3D models
- Static photos
- Static web pages
- Transit

Live traffic

- Twitter feeds
- Visualizers
- Live weather
- Live news
- Live videos
- Personal history
- Personal preferences

Personalization and contextual responses

e.g., GPS = location? Last two searches? Last search for this location? Social networks
Map information

- Satellite imagery
- 3D models
- Static photos
- Static web pages
- Transit

Live traffic

- Prescriptive recommendations
- Twitter feeds
- Visualizers
- Live weather
- Live news
- Live videos

Personal history

Inferred preferences

Personal preferences

Social-intelligence-based recommendations

e.g., google-squared table of restaurants, sightseeing
Research Challenges, Integration, Composition

- increasingly more integrated services will result in even higher complexity
- versioning, compatibility, coordination among multiple mgmt components

Research Challenges, Quantifying Manageability

- checklist of manageability functions
- number of steps to manage towards desired state
- time to manage (including time to insight)
- documentability (e.g. lines of management code)
- elasticity of management (manage at scale)
- availability and continuity of management
- ease of use (GUIs, visualization, analytics, etc.)
## State of the Art & Research Direction of Cloud Mgmt

<table>
<thead>
<tr>
<th>Management Functionality</th>
<th>State of the Art</th>
<th>Research Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mgmt at Scale &amp; Federation</strong></td>
<td>x100k nodes in DCs; zones, service-level integration, increment. scalability; simple visualization</td>
<td>Domain hierarchies, federations of independently managed DCs &amp; Cloud; visualiz. analytics at full scale</td>
</tr>
<tr>
<td><strong>Sustainability</strong></td>
<td>Track power, CO₂, water usage, minimize environmental impact; introduce end-to-end sustainability</td>
<td>Trade off sustainability for QoS, automate sustainability &amp; SLA mgmt, sustainable mobile services delivery</td>
</tr>
<tr>
<td><strong>Support and reliability</strong></td>
<td>Reactive at high end (field engin.) deferred at low end with minimal human use; semi-automated</td>
<td>Preventive, substantially automated, self-healing and rejuvenation; field engineers only at very high end</td>
</tr>
<tr>
<td><strong>QoS: SLA management</strong></td>
<td>Simple SLOs. Lack of compliance and enforcing SLAs. No integration with business models</td>
<td>Multi-objectives, business obj (pricing, costing). Autom. enforce compliance. Hierarch. decompose SLAs</td>
</tr>
<tr>
<td><strong>Data management</strong></td>
<td>DC data deduplication, petascale struct/unstructured data; disk/tape for backup; regul. compliance</td>
<td>Global deduplication, exascale largely unstructured data; storage hierarchies NVRAM, disks; global compliance.</td>
</tr>
<tr>
<td><strong>Integration of management components</strong></td>
<td>Integrate single layer, local feedback loops; rapid deployment, config. mgmt; orchestrate global services</td>
<td>Choreographies, closed loops of loosely coupled domains (power, perform., availab.) individual/tradeoff</td>
</tr>
<tr>
<td><strong>Quantifying manageability</strong></td>
<td>Checklist of management functions, documentation, time and steps to manage objects/services</td>
<td>Measuring Quality of Mgmt (QoM), elasticity of management (match mgmt to function supported), ease of mgmt</td>
</tr>
</tbody>
</table>
# Summary, Trends Impacting Future of Cloud Mgmt

<table>
<thead>
<tr>
<th>Layer of the Stack</th>
<th>State of the Art</th>
<th>Research Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud users</td>
<td>Internet users, increased mobile access (developed and emerging areas). Some mash-up ability, Some ability to customize, personalize</td>
<td>Dominantly mobile access, developing countries growth towards 8B users, especially mobile. Extensive mash-ups, user composed services. Extensive personalization and customization</td>
</tr>
<tr>
<td>Cloud services developers</td>
<td>Small # for traditional, mobile services Few releases annually, careful testing some service location awareness New services through development</td>
<td>Through composition, integration, large % of developers; continuous roll-out of new releases, agile development. Full location awareness; integrate w local services available ubiquitously</td>
</tr>
<tr>
<td>Cloud mgmt operators</td>
<td>Cloud infra/service operators (small %) increasing updates to mobile devices some high level dashboard, analytics reporting, some prediction</td>
<td>Merging role with Cloud developers (large %) frequent update to mobile services, devices detailed business dashboards, visual analytics what if analysis, prediction business outcomes</td>
</tr>
<tr>
<td>HW &amp; impact on support</td>
<td>Disks, early SDDs, 10Gb/s Eth., early adoption of optical interconnect, 16-24 Core CPUs, 100,000 server DCs, air cooling, limited use of water high redundancy, reactive &amp; delayed support, field eng., complex SW repair</td>
<td>NVRAM adoption, optical interconnects photonics off-on chips, 1000+ Core CPUs, 10^{12}+ server DCs, ambient cooling (commodity), liquid (high end) self-healing, proactive support, customer self-repair, moving up the stack, restartable services</td>
</tr>
</tbody>
</table>
1. Cloud Management Introduction
2. State of the Art
   • Managing Private Clouds: CloudSystem Matrix
   • Managing Public Clouds: Internet Data Centers
   • Managing HPC in the Clouds: Towards Exascale
3. IT Industry Trends
4. Open Challenges and Road Ahead
5. Future of Cloud Management: Requirements and Research Challenges
6. Summary
Summary

• Cloud computing has a fundamental role in future, as IT is migrating towards cloud
• As mobile services find their way into the cloud, it will become even more ubiquitous
• The role of cloud management will become essential—particularly in regard to
  • scale, DevOps, and QoS
• Data-intensive ops/mgmt will become dominant compared to compute intensive
• Sustainability and support will substantially evolve in the future
Thank you!

Contact: dejan.milojicic@hp.com