Cloud Computing Tutorial

Jens Nimis

Tutorial, IPE-Klausurtagung
30. Juli 2009, Freudenstadt
Sources

  http://www.slideshare.net/jamesbroberg/introduction-to-cloud-computing-ccgrid-2009

• [MM] Michael Maximilien, IBM

• [MK] Dr. Marcel Kunze und Christian Baun (comics), SCC Karlsruhe

• Stefan Tai, Alex Lenk, Markus Klems, Sebastian Schmidt & many more...
Agenda

• Part 1: What is Cloud Computing?
• Part 2: The Cloud Ecosystem
• Part 3: Current research questions and interesting directions
Agenda – Part 1

• Part 1: What is Cloud Computing?
  • Definitions
  • Cloud vs. Grid
  • Challenges and Opportunities

• Part 2: The Cloud Ecosystem

• Part 3: Current research questions and interesting directions
Some remarks on Cloud Definitions

• Anonymous:
  „[…] unfortunately the marketing guys got hold of the term before the technicians had known what Cloud Computing is […]“

• A lot of semi-serious definitions:
  • Cloud = Grid made right / Grid made easy
  • Grid: from Science for Science
  Cloud: from Business for Business

• Let’s get serious (first…)}
Some serious definitions

- **UCBerkeley RADLabs**: “Cloud computing has the following characteristics: (1) The illusion of infinite computing resources… (2) The elimination of an up-front commitment by Cloud users… (3). The ability to pay for use…as needed…” → business perspective

- **McKinsey**: “Clouds are hardware-based services offering compute, network and storage capacity where: Hardware management is highly abstracted from the buyer, Buyers incur infrastructure costs as variable OPEX, and Infrastructure capacity is highly elastic” → only one kind of Cloud

- **Wikipedia**: “.. a style of computing in which dynamically scalable and often virtualized resources are provided as a service over the Internet” → technical perspective

[JB]
Our definition

“Building on compute and storage virtualization, cloud computing provides scalable, network-centric, abstracted IT infrastructure, platforms, and applications as on-demand services that are billed by consumption.”

Common ground:
- Web Service and Web Portal access
- Scalability
- Pay per use
- Virtualisation/abstraction
- XaaS

→ Technical enablers:
- WS-Technology: SOAP, REST,…
- Virtualization: VMWare, XEN, Virtual Box,…
Cloud has replaced Grid in public visibility, but for the last time: Cloud <> Grid V2 !!!

Foster’s Grid Definition “What is the Grid? A Three Point Checklist”
- Computing resources are not administered centrally
- Open standards are used
- Nontrivial quality of service is achieved

Big differences in definitions, but unfortunately promises and the metaphor are similar…
Cloud Computing provides solutions to a variety of challenges and opportunities

The classical problem
- Under-utilized server resources waste computing power (and energy)
- Over-utilized servers cause interruption or degradation of service levels

...today in an Internet setting
- Resource demands are increasingly of highly dynamic nature and Internet-scale
- On-demand resources are a means for faster time-to-market, and cost-effective innovation processes

...and tomorrow in the next-gen Web
- Leveraging the Web as a combined technology, business, and people collaboration platform:
  - Making effective use of sophisticated infrastructure which is increasingly available as (Web) services
  - Enabling dynamic (trans-)formation of open service and business networks
This was our starting point: Cloud TCO (single consumer viewpoint)

1. Collect real-world use cases and identify typical scenarios
2. Examine key aspects from business and IT perspective
   - **business objectives**
     - foster innovation
     - rapid prototyping
     - leverage Web as platform
   - **demand behavior**
     - seasonal
     - temporary spikes
     - unpredictable
   - **IT requirements**
     - scalability
     - reliable and stable platform
     - high availability
3. Understand and valuate benefits from cloud computing
   - **Estimate costs**
     - variable costs
     - fixed costs
     - time to market
   - **Estimate value**
     - Business value
     - Economic value
   - **Derive strategies**
     - Decision processes
     - Recommendations
     - Business transformation

1. Business Scenario
   1.1 Business Domain
   1.2 Business Objectives
   1.3 Demand Behavior
   1.4 Technical Requirements

2. Cloud Computing service
   2.1 Resource usage
   2.2 Utility Computing model
   2.3 Direct costs
   2.4 Indirect costs

3. Reference model
   3.1 Resource usage
   3.2 Reference model
   3.3 Direct costs
   3.4 Indirect costs

Describe the scenario that captures project key characteristics.

Calculate the TCO of a suitable Cloud Computing service.

Calculate the TCO of the next best alternative IT infrastructure solution (opportunity costs).

Agenda – Part 2

• Part 1: What is Cloud Computing?

• Part 2: The Cloud Ecosystem
  • Cloud Architecture
  • Cloud Players
  • Change ahead!

• Part 3: Current research questions and interesting directions
Organizational Cloud Architecture: Public-/Hybrid-/Private-Cloud
Technical Cloud Architecture: Cloud Computing Stack

- Generic Approach
- Layered architecture
- Everything as a Service concept
  - Standard layers
    - Infrastructure as a Service
    - Platform as a Service
    - Software as a Service
  - Extra Layers
    - Human as a Service
    - Administration/Business Support

Infrastructure as a Service

- **Infrastructure Services**
  - Storage
  - Computational
  - Network
  - Database
  - e.g. Google Bigtable, GoogleFS, Hadoop MapReduce, HadoopFS

- **Resource Set**
  - Machine Images
  - e.g. EC2, Eucalyptus
Platform as a Service

- Programming Environment
  - Programming Language, Libraries
  - e.g. Django, Java

- Execution Environment
  - Runtime Environment
  - e.g. Google App Engine, Java Virtual Machine
Software as a Service

- **Applications**
  - User Interface
  - Frontend Application
  - e.g. Google Docs, Yahoo Email

- **Application Services**
  - Webservices Interface
  - Basic or Composite
  - e.g. Opensocial, Google Maps
Human as a Service

- Crowdsourcing
  - Enabling Collective Intelligence
  - e.g. Mechanical Turk

- Information Markets
  - Prediction of events
  - e.g. Iowa Electronic Markets
Administration/Business Support

- Available on all layers

- Administration
  - Deployment
  - Configuration
  - Monitoring
  - Life cycle management

- Business support
  - Metering
  - Billing
  - Authentication
  - User management
Cloud Architecture → Cloud Players

- **Infrastructure SPs**
  - Basic SPs
  - Intermediaries
  - High-value SPs

- **Basic SPs**
- **Intermediaries**
- **High-value SPs**
Players

Cloud **infrastructure service providers** – raw cloud resources
IaaS (infrastructure-as-a-service)

Cloud **platform providers** – resources + frameworks; PaaS (platform-as-a-service)

Cloud **intermediaries** – help broker some aspect of raw resources and frameworks, e.g.,
  - server managers, application assemblers, application hosting

Cloud **application providers** (SaaS)

Cloud **consumers** – users of the above
Players: Providers

Programmatic access via Web Services and/or Web APIs
“Pure” virtualized resources
  CPU, memory, storage, and bandwidth
  Data store

versus

Virtualized resources plus application framework
  (e.g., RoR, Python, .NET)
  Imposes an application and data architecture
  Constrains how application is built

[MM]
Players: Cloud Intermediaires

Resells (aspects of) raw cloud resources, with added value propositions

- Packaging resources as bundles
- Facilitating cloud resource management, e.g., setup, updates, backup, load balancing, etc.
- Providing tools and dashboards

Enabler of the cloud ecosystem
Players: Application Providers

Software as a Service (SaaS): Applications provided and consumed over the Web
Infrastructure usage (mostly) hidden
Amazon AWS Cloud Offerings:

• Amazon Elastic Compute Cloud (Amazon EC2)
• Amazon Simple Storage Service (Amazon S3)
• Amazon Simple Queuing Service (Amazon SQS)
• Amazon SimpleDB

• Amazon Elastic MapReduce
• Amazon CloudFront
• Amazon DevPay
• AWS Import/Export
Cloud computing by example: Amazon EC2

Typical Workflow:
• Selection of AMI selection
• Selection of instance size and availability zone
• Generation of Key-pair
• Start of Instance
• Definition of Security Zone / Accessibility
• Persistence of States → EBS
• Generation of individual AMIs

→ E.g. GUI tool support
Typical Workflow:

- **Anlegen von Buckets**
  
  ```
  s3cmd mb s3://Bucket
  ```

- **Hochladen von Objekten in einen Bucket**
  
  ```
  s3cmd put LokaleDatei s3://Bucket/EntfernteDatei
  ```

- **Auslesen von Meta-Daten z.B. Bucketinhalten**
  
  ```
  s3cmd ls s3://Bucket
  ```

- **Herunterladen von Objekten aus einem Bucket**
  
  ```
  s3cmd get s3://Bucket/EntfernteDatei LokaleDatei
  ```

- **Löschene von Dateien**
  
  ```
  s3cmd del s3://Bucket/EntfernteDatei
  ```

- **Löschene von (leeren) Buckets**
  
  ```
  s3cmd rb s3://Bucket
  ```

→ E.g. command line tool support
Typical Workflow:

- **CreateQueue**: Anlegen einer Queue im AWSBenutzerkontext
- **ListQueues**: Aufzählung der existierenden Queues
- **DeleteQueue**: Löschen einer Queue
- **SendMessage**: Einstellen einer Nachricht in eine Queue
- **ReceiveMessage**: Auslesen einer (oder mehrerer) Nachrichten aus einer Queue
- **ChangeMessageVisibility**: Einstellen weitere Sichtbarkeit gelesener Nachrichten
- **DeleteMessage**: Löschen einer gelesenen Nachricht
- **SetQueueAttributes**: z.B. Zeit zw. zwei Leseoperationen auf dieselbe Nachricht
- **GetQueueAttributes**: z.B. Anzahl der aktuell in der Queue befindlichen Nachrichten
- **AddPermission**: Freigabe von Queues zum geteilten Zugriff verschiedener Benutzer
- **RemovePermission**: Widerrufen der Freigabe für andere Benutzerkontexte
Typical Workflow:

- **CreateDomain, ListDomains, DeleteDomain**: Domäne ≈ Relation
- **DomainMetadata**: Auslesen z.B. von aktuellem Speicherplatzbedarf
- **PutAttributes**: Hinzufügen oder Aktualisieren eines Datensatzes basierend auf einem Datensatzidentifikator und Attribut/Wert-Paaren
- **BatchPutAttributes**: Gleichzeitiges Anstoßen mehrerer Einfügeoperationen zur Performance-Erhöhung
- **GetAttributes**: Lesen eines identifizierten (Teil-)Datensatzes
- **DeleteAttributes**: Löschen von Datensätzen, Attributen oder Werten
- **Select**: Anfrage in SQL-ähnlicher Syntax (ohne Joins!)
Change Ahead!

Caveat: AWS are not 1:1 replacements for traditional IT infrastructure components – they change the way how to build systems!

E.g. GrepTheWeb
Agenda – Part 3

- Part 1: What is Cloud Computing?
- Part 2: The Cloud Ecosystem
- Part 3: Current research questions and interesting directions
  - In general
  - At IPE
  - „Near“ IPE
Open Issues in General

Table 1: Quick Preview of Top 10 Obstacles to and Opportunities for Growth of Cloud Computing.

<table>
<thead>
<tr>
<th>Obstacle</th>
<th>Opportunity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Availability of Service</td>
<td>Use Multiple Cloud Providers; Use Elasticity to Prevent DDOS</td>
</tr>
<tr>
<td>2 Data Lock-In</td>
<td>Standardize APIs; Compatible SW to enable Surge Computing</td>
</tr>
<tr>
<td>3 Data Confidentiality and Auditability</td>
<td>Deploy Encryption, VLANs, Firewalls; Geographical Data Storage</td>
</tr>
<tr>
<td>4 Data Transfer Bottlenecks</td>
<td>FedExing Disks; Data Backup/Archival; Higher BW Switches</td>
</tr>
<tr>
<td>5 Performance Unpredictability</td>
<td>Improved VM Support; Flash Memory; Gang Schedule VMs</td>
</tr>
<tr>
<td>6 Scalable Storage</td>
<td>Invent Scalable Store</td>
</tr>
<tr>
<td>7 Bugs in Large Distributed Systems</td>
<td>Invent Debugger that relies on Distributed VMs</td>
</tr>
<tr>
<td>8 Sealing Quickly</td>
<td>Invent Auto-Scaler that relies on ML; Snapshots for Conservation</td>
</tr>
<tr>
<td>9 Reputation Fate Sharing</td>
<td>Offer reputation-guarding services like those for email</td>
</tr>
<tr>
<td>10 Software Licensing</td>
<td>Pay-for-use licenses; Bulk use sales</td>
</tr>
</tbody>
</table>

Reliability, Portability, Security/Trust, Scalability, SLAs, Licenses,...
Overview

CC Research Questions@IPE

Business Cases and Cloud TCO
- Research paper: „Do Clouds Compute?“
- Project ICE (T-Labs): CC business cases for T-Com

Cloud Computing Adoption
- MTh & IBM GBS: CC Maturity Model w. online tool

Cloud Value Creation
- Cloud offering value creation esp. for intermediairies
- Architecture of „the Cloud“
  - Research paper: „What's inside the Cloud?“

Cloud Engineering
- Project ICE (T-Labs): Dev. support for IntraCloud-Patterns
- MTh (OpenCirrus/HP): „Cloudification“ of apps
  - Cloud service composition, Cloud application arch.

SAP Landscape Provisioning
- MThs & Project Proposal (ZIM fluidOps): Reliability of VPDC

SAP Cloud-Demo
- Project (SAP CEC): SLA mgmt for complex systems
Condensed Topics at www.eOrganization.de
Cloud Service Engineering

Cloud Services (XaaS)

Cloud Programming Models and Architecture

Service Value Networks

Cloud Service Engineering

- Business opportunities
- Internet-scale service computing
- Enterprise-grade systems management
OpenCirrus™
Cloud Computing Research Testbed

• An open, internet-scale global testbed for cloud computing research
  • Data center management & cloud services
  • Systems level research
  • Application level research
• Structure: a loose federation
  • Sponsors: HP Labs, Intel Research, Yahoo!
  • Partners: UIUC, Singapore IDA, KIT, NSF
  • Members: System and application development
• Great opportunity for cloud R&D

http://opencirrus.org
• See also http://markusklems.wordpress.com/
• and soon the new ICE-Cloud-Feed
And more Information (in German):
http://tinyurl.com/CloudBuch

Christian Baun, Marcel Kunze, Jens Nimis, Stefan Tai:

Cloud Computing: Web-basierte dynamische IT-Services
(Reihe: Informatik Im Fokus)