Virtualization/Vagrant/Cloud Computing

Amol Deshpande
University of Maryland
amol@cs.umd.edu
Cloud Computing

- Computing as a “service” rather than a “product”
  - Everything happens in the “cloud”: both storage and computing
  - Personal devices (laptops/tablets) simply interact with the cloud

- Advantages
  - Device agonstic – can seamlessly move from one device to other
  - Efficiency/scalability: programming frameworks allow easy scalability (relatively speaking)
    - Increasing need to handle “Big Data”
  - Reliability
  - Cost: “pay as you go” allows renting computing resources as needed – much cheaper than building your own systems
Cloud Computing

- Basic ideas have been around for a long time (going back to 1960’s)
  - Mainframes + thin clients (more by necessity)
  - Grid computing a few years ago
  - Peer-to-peer
  - Client-server models
  - …

- But it finally works as we wished for…
  - Why now?… A convergence of several key pieces over the last few years
  - Does it really? … Still many growing pains
Virtualization

- Virtual machines (e.g., running Windows inside a Mac) etc. has been around for a long time
  - Used to be very slow...
  - Only recently became efficient enough to make it a key for CC

- Basic idea: run virtual machines on your servers and sell time on them
  - That’s how Amazon EC2 runs

- Many advantages:
  - Security: virtual machines serves as almost impenetrable boundary
  - Multi-tenancy: can have multiple VMs on the same server
  - Efficiency: replace many underpowered machines with a few high-power machines
Virtual Machines

https://stalk-calvin.github.io/blog/2017/04/15/vagrant-docker.html
Vagrant makes it easy to create and configure virtual environments.
Virtual Machines vs Containers

https://stalk-calvin.github.io/blog/2017/04/15/vagrant-docker.html
Data Centers

- The key infrastructure piece that enables CC
- Everyone is building them
- Huge amount of work on deciding how to build/design them
Data Centers

- Amazon data centers: Some recent data
  - 8 MW data center can include about 46,000 servers
  - Costs about $88 million to build (just the facility)
  - Power a pretty large portion, but server costs still dominate

“Every day, Amazon Web Services adds enough new capacity to support all of Amazon.com’s global infrastructure through the company’s first 5 years, when it was a $2.76B annual revenue enterprise”

[source: James Hamilton Presentation]
Putting it together

- A Cloud Computing Provider builds and manages “Data Centers”, often with millions of servers
- They may rent you:
  - Raw hardware (not that common)
  - Virtual machines in those data centers (Infrastructure-as-a-service)
  - A software service that does something specific for you (Software-as-a-service)
  - Something in-between (Platform-as-a-service)
<table>
<thead>
<tr>
<th>On-Premises</th>
<th>IaaS (Infrastructure as a Service)</th>
<th>PaaS (Platform as a Service)</th>
<th>SaaS (Software as a Service)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applications</td>
<td>Applications</td>
<td>Applications</td>
<td>Applications</td>
</tr>
<tr>
<td>Data</td>
<td>Data</td>
<td>Data</td>
<td>Data</td>
</tr>
<tr>
<td>Runtime</td>
<td>Runtime</td>
<td>Runtime</td>
<td>Runtime</td>
</tr>
<tr>
<td>Middleware</td>
<td>Middleware</td>
<td>Middleware</td>
<td>Middleware</td>
</tr>
<tr>
<td>O/S</td>
<td>O/S</td>
<td>O/S</td>
<td>O/S</td>
</tr>
<tr>
<td>Virtualization</td>
<td>Virtualization</td>
<td>Virtualization</td>
<td>Virtualization</td>
</tr>
<tr>
<td>Servers</td>
<td>Servers</td>
<td>Servers</td>
<td>Servers</td>
</tr>
<tr>
<td>Storage</td>
<td>Storage</td>
<td>Storage</td>
<td>Storage</td>
</tr>
<tr>
<td>Networking</td>
<td>Networking</td>
<td>Networking</td>
<td>Networking</td>
</tr>
</tbody>
</table>
Amazon Web Services

- Leader in this space for a while

Amazon Dominates Public Cloud Market

Global market share of public cloud infrastructure service providers in Q3 2019*

- Amazon Web Services: 39%
- Microsoft Azure: 19%
- Google Cloud: 9%
- Alibaba Cloud: 5%
- Salesforce: 4%
- IBM Cloud: 3%
- Others: 21%

Worldwide spending on public IaaS and PaaS in Q3 2019

$20 billion

* Includes platform as a service (PaaS) and infrastructure as a service (IaaS)
Source: Synergy Research Group
AWS has 175 Services Today.. More every year...
Microsoft Azure not far behind