Below the Clouds
A Berkeley IT Operator's View of Cloud Computing

Jon Kuroda – UC Berkeley EECS
jkuroda@eecs.berkeley.edu

Cloud: (n) A Pay-As-You-Go computing service with no upfront cost to the user that provides on-demand scaling with the illusion of infinite resources.
What does this mean? It is deceptively simple yet carries several implications

The Cloud User's View From Above the Clouds

Interface for computation
- VMs: EC2 /AWS – flexible, generic
- Domain Specific: App Engine, SalesForce, Cloudera
- In-Between: Azure

Storage
- Gotta have a "Storage Story," a good one.
- How to get Software Stack into the Cloud
- How to get data into and out of The Cloud
- S3, HDFS, MS Cosmos, Others?

Provisioning and Management
- How to request and return resources
- Virtualized Network Topologies
- Resource Constraint Policy
- Automated control?

Business Support
- Billing
- Tech Support
- Legal
- What Else?

The Cloud Operator's View From Below the Clouds

Datacenter(s)
- Shelter, Physical Security, Power, AC
- Network, Storage
- O(10,000) Bare Metal Systems
- MgmtM/onitoring (Ganglia, Nagios, etc)

Virtualization
- More generally, it's Resource Abstraction
- What sort? Xen/VMWare? Higher Level?
- Provisioning? VMware VI, Eucalyptus, EC2
- Storage, Network, what else?

Software Stack
- Platform SW/OS
- Cluster Infrastructure
  MapReduce/Hadoop, Dryad
- End-user Applications

Pervasive Infrastructure
- Identity Management/AuthNZ
- Usage Accounting
- Monitoring/Instrumentation
  Chukwa, X-trace, LogMining

Datacenter(s): Solved Problem(s)?

- O(10-100K) systems = lots of engineering
- Modular/Containerized DCs are not a panacea
  - Still need a physical shelter
  - They encourage density but is it a good idea?
    - Tight physical constraints (HW and Operator)
    - Racks on Wheels? Try that with 500+kg
    - Network HW and Airflow requirements
  - A rack of 35 systems means:
    - 2 network switches
    - 3 "smart" power control strips.
    - 9-10 dozen cables (power and network)
    - Lots of Velcro, "Creative" Cabling Topologies
    - 24 doz. rack screws (and three drill batteries for them all)
  - Density × Service Accessibility = Constant
- Little rack level engineering, let alone at DC Level
  - Blade servers go partway
  - Cabinet sized box – the wrong direction?
  - Need holistic design from node through DC scale

And that's just the Bare Metal Infrastructure

From Bare Metal System to Prêt-à-Utiliser

- System Installation
  - Would like to think that this is 'solved'
  - How to handle it for O(10-100K) systems
  - How to handle different install images
- Change and configuration management (CCM) essential
  - Essential at the O(1000)scale already
  - Integrate with base installation
- Monitoring Systems
  - Nagios/Ganglia – do they scale past O(1000)?
  - Need more than service check/system metrics
    - X-trace – instrumenting software
    - Chukwa / LogMining – post event data collection/analysis
  - Automated response – speed matters
  - Need more unified interface to these systems
    - Separate systems – hard to keep a global view
    - Over lap in data collection
    - Multiple ways to view data, but one framework
  - Operators need toolkits, not monolithic 'solutions'

The Dial Tone has become a lot more complex

Potential New Trends

- Impact of many-core (10+) systems
  - CPU/RAM usage isolation exists
  - Network BW guarantees possible
  - Memory/IO BW guarantees?
    - Can memory/IO keep up?
    - Turn off cores to aid IO isolation?
    - Use some only for low IO BW work
    - Does it matter? Maybe only for VM
- Need for debugging tools
  - Some tools exist
  - Virtualization confounds debugging
  - How to deal with non-determinism?
    - Tackle with statistics/ML
    - Expose lower layers to higher
- VM network topology still simple
  - Is this a bad thing?
  - Will they (over)grow in complexity?
  - Exposing limits of "real" network

Clouds change with the weather

Security: Not just keeping10k systems secure

Beyond the VMM
- VMMs are shrinking tip of the iceberg
- Is it just Multi-Tenancy?
- Middleware, Mgmt, CCM SW?
- Entropy doesn't virtualize

Forensics
- Bad things still happen
- What audit trail exists?
- End-user access?
- Cloud provider support?

Approaches
- Treat like Internet "just a transport"?
- Crypto is part of but not the answer
- How to verify work done on the cloud?
- Be like Mulder – Trust no one?

"Private" Clouds?
- More Business Continuity Control
- Still have to keep up internal walls
- What happens when you "Surge"?
- Still need to self-audit

Does this really change IT operations?

Example: Surge Computing
- Internal /Private Cloud Service
- Shed load to Public Cloud
- **Concerns**
- Inter-cloud compatibility?
- Cross-domain auth?
- Data Integrity/Security
- **Reactions**
- Oh No! Lack of data control!
- Have to take all these measures!

Reality

Those measures we might take to deal with any perceived dangers of putting non-trivial data into the loud are ones we should be taking anyway even 'on the reservation'.

Where one might get 'lazy' when everything is kept local, one has to clean up one's act when one lets data go off the reservation.

Major players (Google, Microsoft, Amazon, others) have figured out (some, a lot, most?) of this but in ways specific to their needs, and they haven't shown all (or even all that many) of their 'latest and greatest', In many ways, the "answers" themselves are not so important as the questions and challenges faced and why they chose to deal with them as they did – a handbook There may not be a lot of new technology – putting it all together is the challenge.