Scale & Cost Issues for Cloud Computing
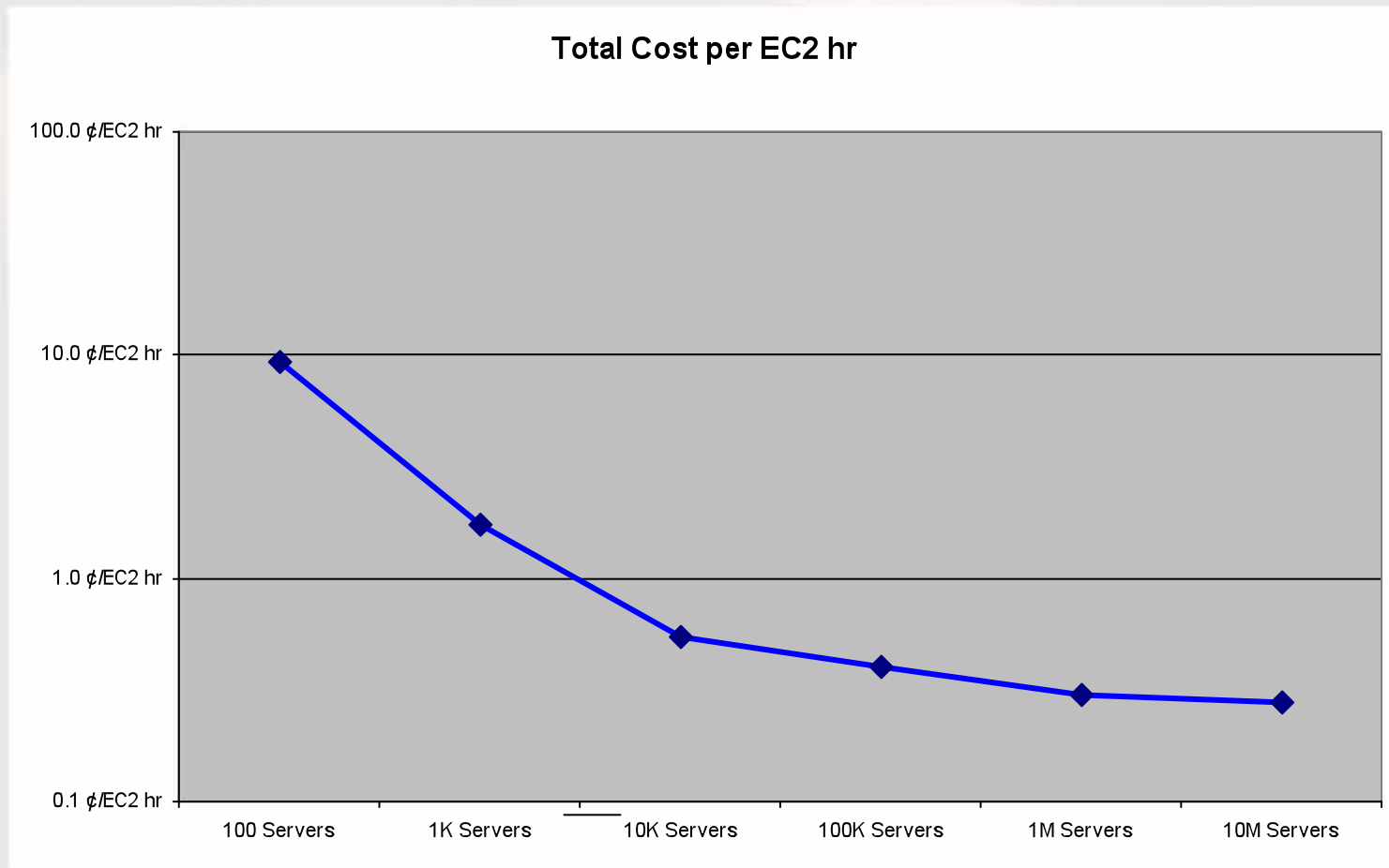
and the Dependability Implications

9/3/10

Basil Smith
March 24, 2010

Summary Results - Achievable Cost vs Scale

Total Cost per EC2 hr

There is close to a 100X cost gradient between traditional IT delivery, and IT delivery from a mega scale cloud/utility provider!
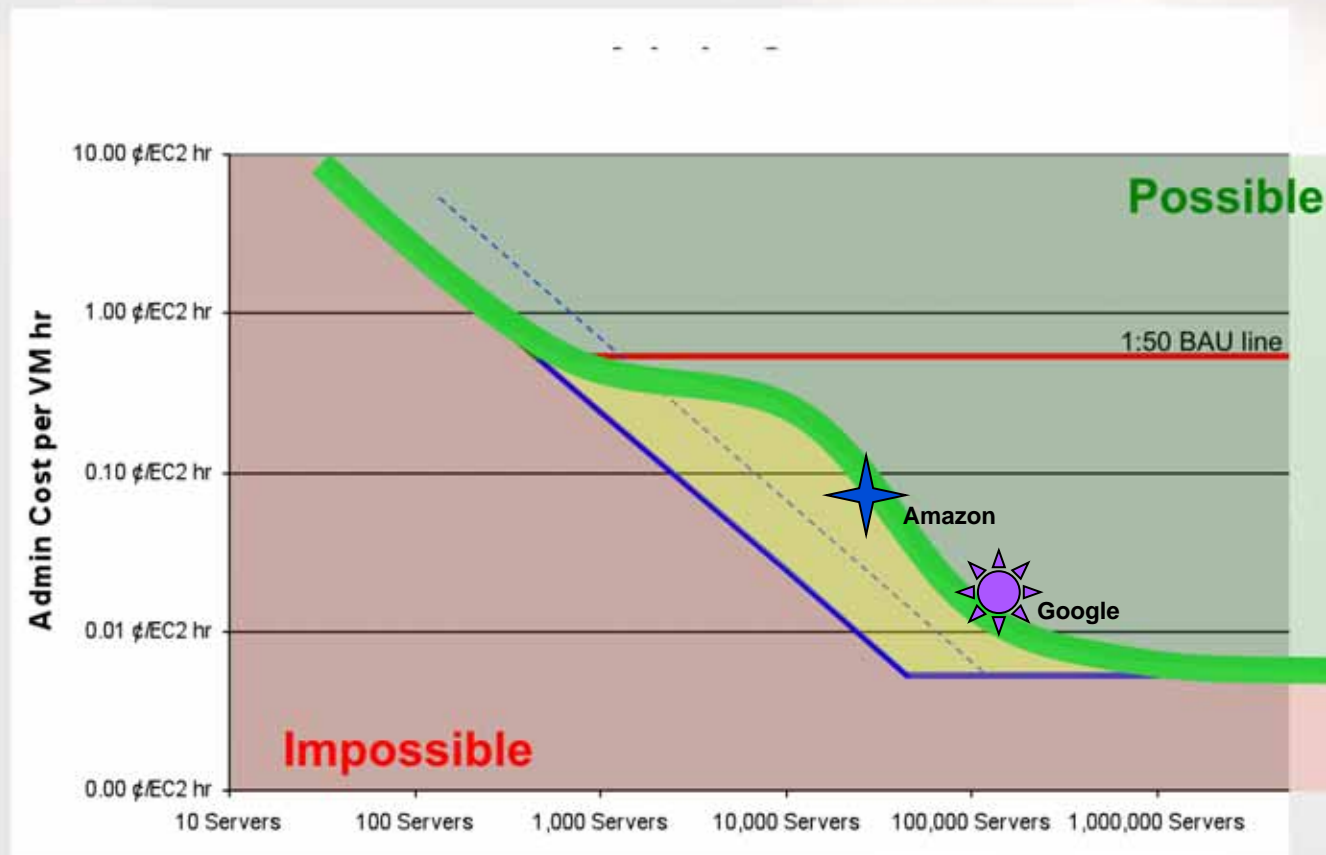
Scale Implications

- **50,000 Servers each able to support up to 64 EC2-small images (128? by end 2010)**
  - An EC2-small image scale (as unit of measure for compute metric:
    - Compute capacity equivalent to Intel Atom processor (a netbook)
    - 64 EC2-small images / contemporary server (YE 2009, dual socket, 8 cores, 32 GB memory)
    - 1.7 GB memory (slightly smaller than a 2GB netbook)
    - 160 GB of memory (typical for a netbook HDD)
  - Larger image types consumer more resources (80% of images are larger than EC2-small[1], and could not run on Atom)
  - With EC2 images types matching suspected distribution[1], there are ~11 images per contemporary server

- **800K VM's active at peak times** (10M? VM images inclusive of hibernating/suspended/archived images)
  - Observed Amazon provisioning rate (50K VM's per day) suggests average endurance of ~10 days with 3 minute start time:
    - ~55 VM's launched and terminate each minute on average (~80K per day)
    - Expect normal peak's >5X this (~275 VM's per minute launch rate)
    - >800 VM's could routinely be in launch/provisioning process at one time
    - "launch storms" will inflate this number by 10X if complete data center reboot is to take <1 day

- **Storage scale**
  - @ 3 volumes / image → ~3 M virtual volumes attached to active VM's
  - @ 100 GB average volume size → 300 PB of storage for active volumes
  - With inactive volumes storage would exceed EB
  - Sparse volumes will shrink this considerably, but still a very large number.
  - For example; with average need to create/attach 3 volumes to each VM, ~2,400 volumes must be attached/detached/created/released per min (peak), 24K in launch storm

- **Labor**
  - With 2,000:1 Server/Admin ratio, total of 25 admins to cover 3 shifts X 365 days
  - Suggests only about 4 or 5 people max operationally managing the center at one time 1st shift (fewer in 2nd,3rd) plus additional labor associated with manual operations such as ordering, install, repair, and retire of equipment

[1] Randy Bias, Founder, Cloudscaling, *Amazon's EC2 Generating $220M Annually,* March 4, 2010

Administration Cost Model with Scale

Summary - Dependability Implications

- **At this scale, massive economic value is associated with continuous and secure operation of a provider's cloud.**

- **This as a minimum implies continuous, secure, and correct operation of the management system (or cloud OS if you prefer)**
  - But the management system (cloud OS) itself is an extraordinarily large system
    - Resources needed to run (mature) cloud OS will consume several 1,000 servers (with similar demands on storage, networking)
    - There could easily be ten's of thousands of processes or management threads active, databases, files, multiple versions of almost every component.
    - There are likely more management servers than admins.
  - Providing dependable secure and correct operation of such a system OS is close to a grand challenge class problem
  - Then there is of course the problem of the providing for dependable execution of client workloads on the remaining 200,000 or so servers spread over several data centers.

- **At this scale you can assume that:**
  - There are tens of 1,000's of SW errors in the implementing management code (Heisenbugs)
  - There are at least 1000's of HW faults, most often only intermittently or situationally revealing themselves as transient errors (Heisenfaults)
  - On rare occasions (by comparison), some obvious easy to diagnose and repair faults will occur (Bohrfaults).

- **At this scale how do you even get things to work at all, let alone dependably.**
  - Yet by many measures this is no more difficult that several other examples of very complex systems that operate dependably.
  - Waiting until we get them to work before thinking about dependably working likely to be fatal.

Paradox:
Admin of cloud OS could be more costly than admin of cloud itself