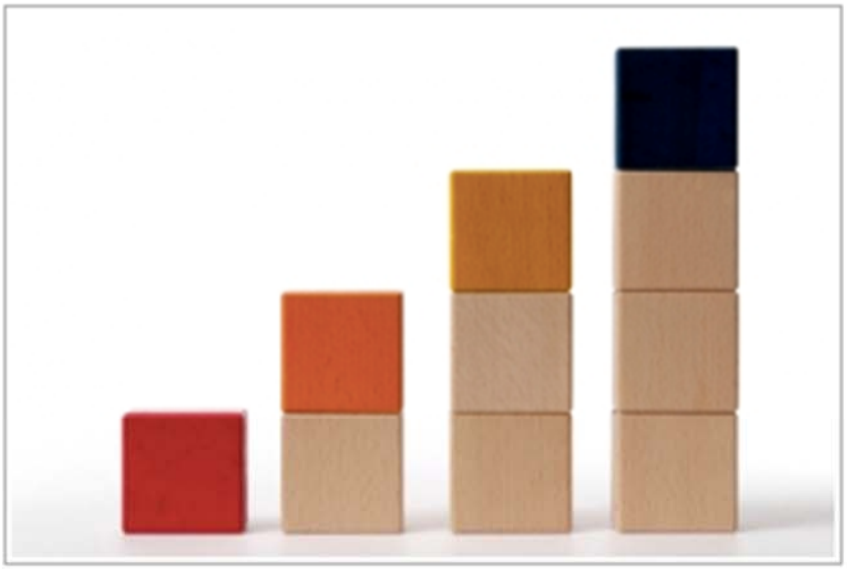
Cloud storage: trends and questions

Geoff Arnold

Yahoo!

gma@yahoo-inc.com

Thought for the day…



In the IT industry, technology and the usage evolves faster than in perhaps any other industry. As a rule of thumb, systems can grow 10 times under their current architecture or paradigm, then they must be re-architected. This 10X effect causes old technologies to become obsolete and new ones to emerge. It also underlies the massive shift to cloud computing.

The last major computing infrastructure paradigm shift happened in the '80s when "client/server" was introduced as the new way to design business applications. Those applications typically ran on x86 computers – aka PCs.

Then, in the '90s, the "client" part of this overall design was disrupted and changed with the

PART 1: TRENDS

Major trends: storage

- From enterprise-class to consumer hardware
- From filers and arrays to commodity JBODs
- From disk to RAM (or SSD if necessary)
- From RAID to decentralized erasure coding

Major trends: connectivity

- From SANs to converged networks
- From FC (and FCoE) to Ethernet and iSCSI
- From NFS to REST

Major trends: systems

- From POSIX & SQL to objects & NoSQL (key-value)
- From transactions to eventual consistency
- From schemas & XML to open content & JSON
- From synchronous to asynchronous operations

Major trends: operations

- From MTBF to MTTR
- From predictable business models to chaotic innovation
- From human operator processes to autonomics and self service

PART 2: QUESTIONS

Some questions

- Are the networks up to the job?
  - Designed for N-S traffic; explosion in E-W
  - Segregating control & data places; QoS
- Are the programming models keeping up?
  - Eventual consistency
  - Parallelism
- How do we build the *right* set of multi-tenant services?
  - The big trade-off: efficiency for all vs. efficiency for each

More questions

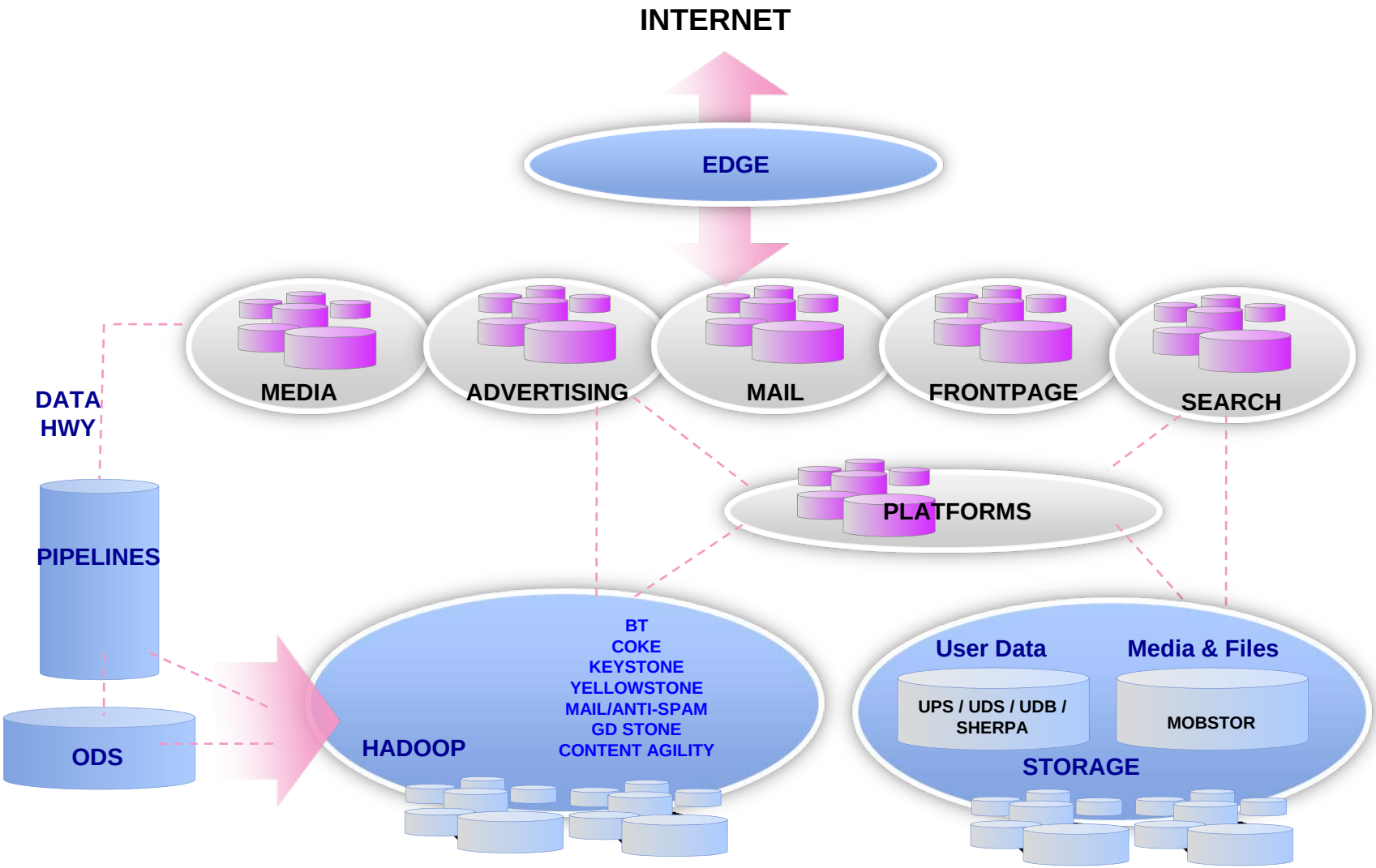
- What are the consequences of automating systems designed for human interactions?
  - Time constants; checks and balances
- We only have a short history with these technologies; we have no basis for trusting SLA, performance, reliability. Are we measuring things enough?
- Are autonomic self-healing & “fail in place” really well matched to the actual systems (hardware & software)?

Still more questions

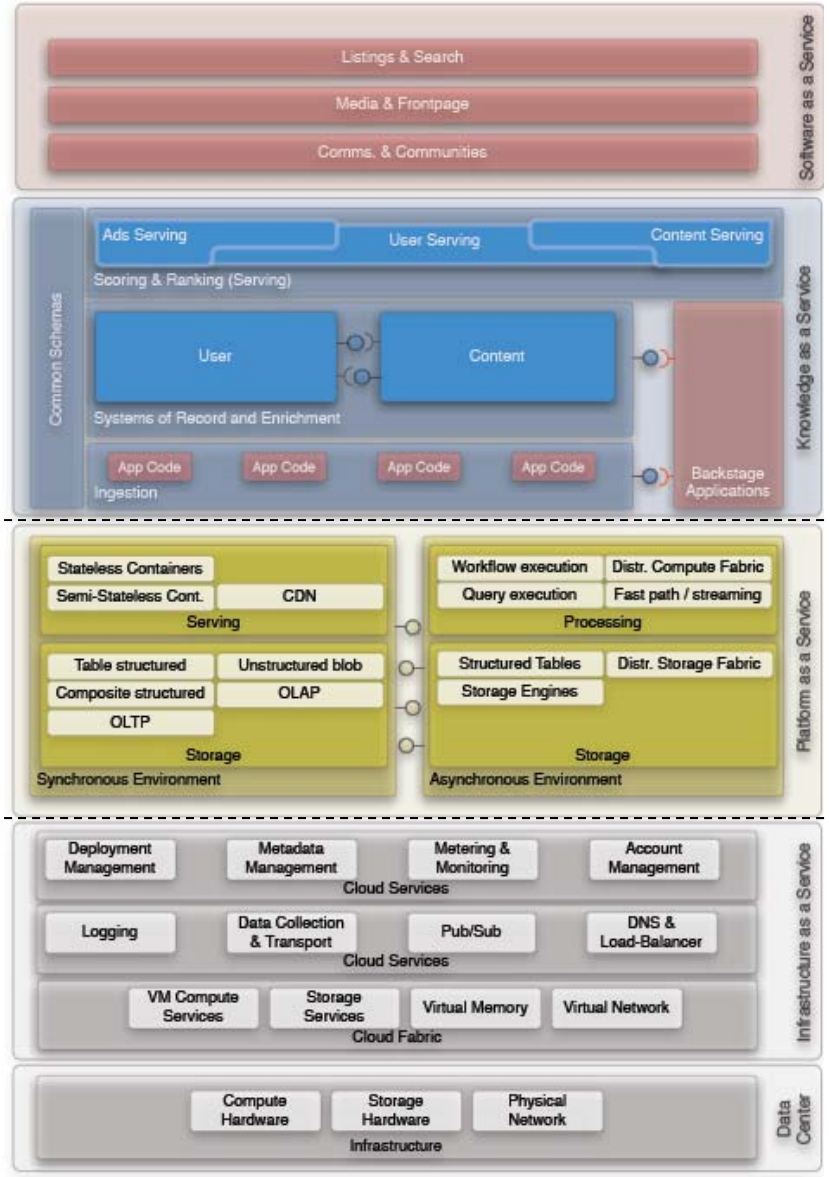
- What happened to operational expertise?
  - If DBAs are obsolete, who provides data modeling, performance analysis, adult supervision?
  - Are users ready for self-service?
- As we refactor the architecture (stove-pipe products to services), how does the human organization keep up?
  - Conway's Law
- Big data outlives hardware. How do we move from *ad hoc* lifecycle management to an "industrialized" model?

PART 3: YAHOO'S CLOUD

Yesterday's (and today's) Yahoo



Architectural Vision



Software as a Service

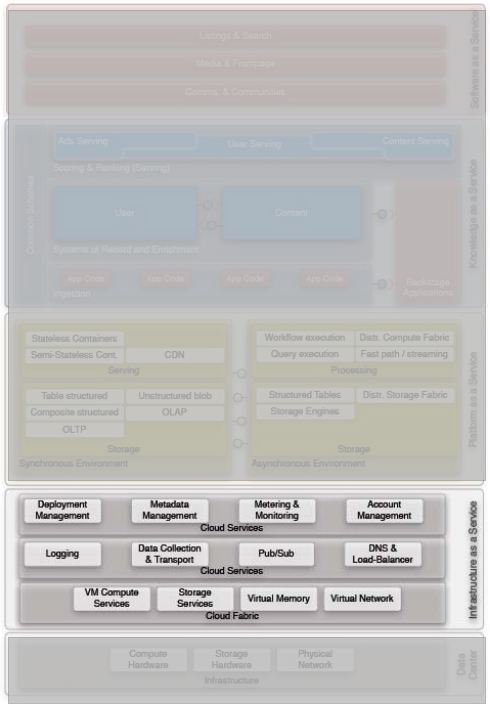
Knowledge as a Service

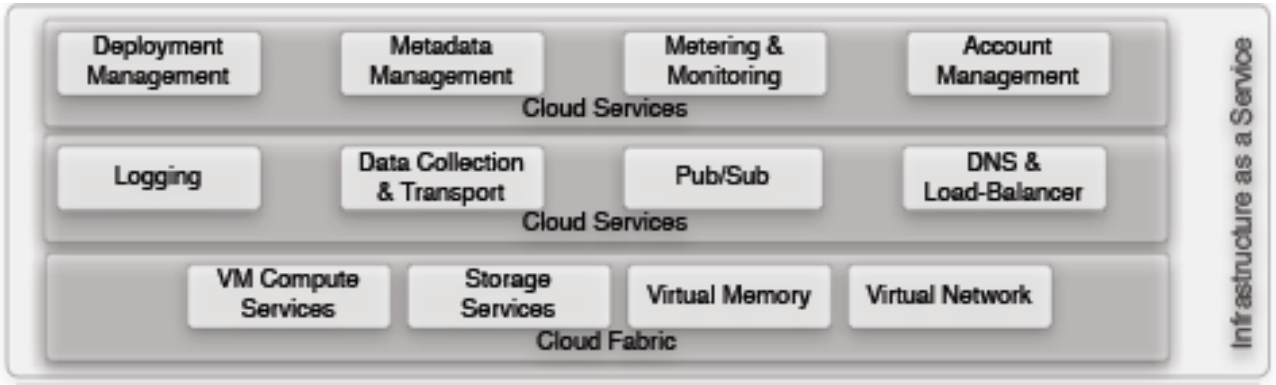
Platform as a Service

Infrastructure as a Service

Hardware

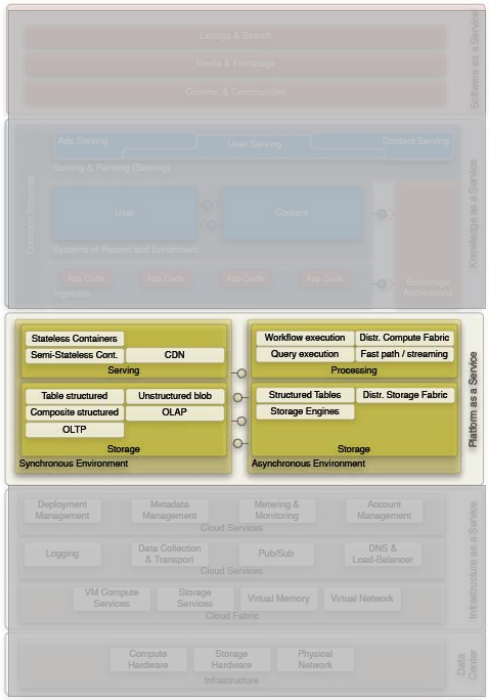
Architectural Vision: IaaS

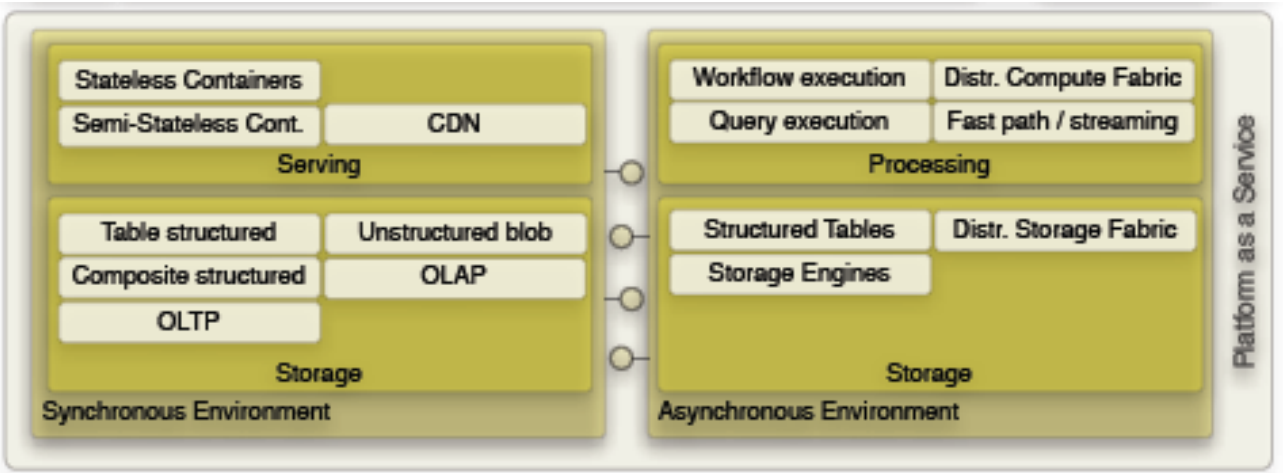




- Planet scale Cloud Fabric that abstracts away all the underlying hardware
- Fundamental Cloud Services that can be assembled to build higher level services

Architectural Vision: PaaS





- On demand, higher level Platform services that support:
  - Interactive onstage applications
  - Offline batch applications
- Sophisticated programming environments that are offered as a holistic Platforms
- Hosts the development and operation of Yahoo applications

~~Space~~ Yahoo is big. You just won't believe how vastly, hugely, mind- bogglingly big it is. I mean, you may think it's a long way down the road to the chemist's, but that's just peanuts to ~~space~~ Yahoo.

Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

Unfortunately, I can't say just how big

- What can I say?
  - 680+M users
  - 200PB of data, adding ~50TB of data per day
  - 100B events a day captured, collected, transported and processed
  - 43k Hadoop servers running 5M+ Hadoop jobs every month
- Let's just say that life is simpler for other folks…

Rackspace Now Has 70,000 Servers

May 10th, 2011 : Rich Miller

Rackspace Hosting added more than 4,400 servers in the first quarter of 2011 to support its managed hosting and cloud computing operations, giving the company a total of 70,473 servers in its data centers. The milestone makes Rackspace only the sixth company to publicly report a server count higher than 70,000.

Why cloud – specifically, why a “private cloud”?

- After all, many people think that “private cloud” is an oxymoron
  - OH: *“You should either be using a public cloud or operating one.”*

Anticipated benefits

- **Business agility**
- Operational consistency
- Interoperability
- Quality
- Tech deduplication
- Efficiency
- Risk reduction
- Cost transparency

Why is agility #1?

- March 11, 2011 – the Japanese earthquake and tsunami
  - Yahoo News spiked to 20.6 million unique visitors
  - We served 371 million page views
  - We added a “Donate Now” button which raised $7M
- Frequently need to spin up new sites with high traffic and short lifetime:
  - Royal Wedding
  - Canadian Election

But why build a private cloud? Why not build (or use) a public cloud?

- Our business is digital media. From our CEO: *“People want to be informed, they want to be entertained, they want to be educated, and they want to communicate. And that's what Yahoo is all about.”*
- There are no public cloud operators that have the scale, geographic presence, and functionality that we would need
- So we have to do it ourselves

PART 4: YAHOO! CLOUD CHALLENGES

Forget “long tail” – the problems are with the head

- Yahoo has a number of “mega-properties”
  - Mail, Front Page, Sports, News, Flickr
- They all want to take advantage of the benefits of cloud
  - Potential for huge gains in efficiency, predictabilty
- However because of their size, they’ve had to radically optimize their technology and operations
- They have stringent performance (latency) requirements and use highly customized technologies
- This means that they aren’t necessarily a good fit for generic, multi-tenant services

Cranking up the rate of change

- Historically, Yahoo properties have acquired new hardware through traditional committee-driven processes
- So there are long lead times, and plenty of time to provision systems like asset management, DNS, monitoring, access control, and so forth
- Many of these systems were not designed to support real-time updates
  - And even if they were, they were rarely stressed
- The effects of introducing on-demand provisioning tend to cascade though many operational systems

Invalidating traditional assumptions

- In a slowly-changing environment, it's easy to assume that some things are constant
- For example, Yahoo has a flat IP network with fixed IP addresses per-box and per-rack
- So it's reasonable to use an IP address as a key for many things
  - It violates current policy, but there are many legacy systems…
- So when a system that uses IP addresses for app instance identifiers is dynamically [re]deployed in a cloud fabric, things break
  - E.g. Reprocessing historic Apache log records

The reality of “on demand”

- “On demand” capacity presumes that each individual customer request is very small compared with the size of the resource pool
  - Capacity management involves careful over-provisioning of the pool
- But we have many large properties…
- Today we overprovision to their planned size/QPS, which:
  - Can waste a lot of resources
  - Has long lead times
- We need to be able to capture the time dimension in our capacity planning and provisioning requests
  - Define the envelope, allow on-demand within that envelope

Automation

- One of the key benefits of cloud computing is improved consistency and predictability through automation
- Scale affects automation in interesting ways
  - Potential interactions between automated deployment and self-healing
  - Very long-running automation, such as the initial population of a new storage farm in a new location
- Issues also arise when automation is designed to exploit an API that was created for a different use-case
  - E.g. an API intended to support a portal GUI with a human operator, which has never been stressed

geoff
arnold

Cloud services wrangler

gma@yahoo-inc.com
Twitter: @geoffarnold
Y! IM: geoff_arnold@yahoo.com
http://geoffarnold.com