Lessons Learned from a Global DB

May 23, 2011

Patrick Quaid

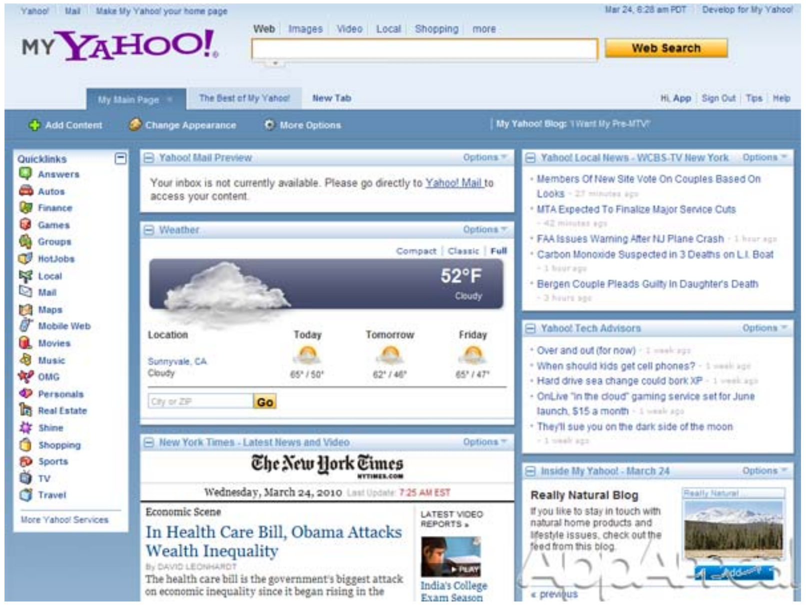
patrickq@yahoo-inc.com

Yahoo! is the premier digital media company

Building the most engaging experience for each individual visitor inevitably depends on data – what our visitors have told us, what we’ve observed, and what we’ve figured out about them and similar visitors.

History

- 1996: My Yahoo
  - First personalized application at Yahoo
- Quickly followed by Mail
- … and then many more applications, large and small


Yahoo! | Mail | Make My Yahoo! your home page
Mar 24, 6:28 am PDT | Develop for My Yahoo!
MY YAHOO!
Web | Images | Video | Local | Shopping | more
Web Search
My Main Page | The Best of My Yahoo! | New Tab
Hi, App | Sign Out | Tips | Help
Add Content | Change Appearance | More Options
Quicklinks
Answers
Autos
Finance
Games
Groups
HotJobs
Local
Mail
Maps
Mobile Web
Movies
Music
OMG
Personals
Real Estate
Shine
Shopping
Sports
TV
Travel
More Yahoo! Services
Options
Compact | Classic | Full
City or ZIP
Go
LATEST VIDEO REPORTS »
PLAY
Add
« previous

My Yahoo! Blog: 'I Want My Pre-MTV!'

Yahoo! Mail Preview

Your inbox is not currently available. Please go directly to Yahoo! Mail to access your content.

Weather

52°F
Cloudy

Location	Today	Tomorrow	Friday
Sunnyvale, CA Cloudy	65° / 50°	62° / 46°	65° / 47°

New York Times - Latest News and Video

The New York Times
NYTIMES.COM

Wednesday, March 24, 2010 Last Update: 7:25 AM EST

Economic Scene

In Health Care Bill, Obama Attacks Wealth Inequality

By DAVID LEONHARDT

The health care bill is the government's biggest attack on economic inequality since it began rising in the

India's College Exam Season

Yahoo! Local News - WCBS-TV New York

- Members Of New Site Vote On Couples Based On Looks - 27 minutes ago
- MTA Expected To Finalize Major Service Cuts - 42 minutes ago
- FAA Issues Warning After NJ Plane Crash - 1 hour ago
- Carbon Monoxide Suspected in 3 Deaths on L.I. Boat - 1 hour ago
- Bergen Couple Pleads Guilty In Daughter's Death - 3 hours ago

Yahoo! Tech Advisors

- Over and out (for now) - 1 week ago
- When should kids get cell phones? - 1 week ago
- Hard drive sea change could bork XP - 1 week ago
- OnLive "in the cloud" gaming service set for June launch, $15 a month - 1 week ago
- They'll sue you on the dark side of the moon - 1 week ago

Inside My Yahoo! - March 24

Really Natural Blog

If you like to stay in touch with natural home products and lifestyle issues, check out the feed from this blog.

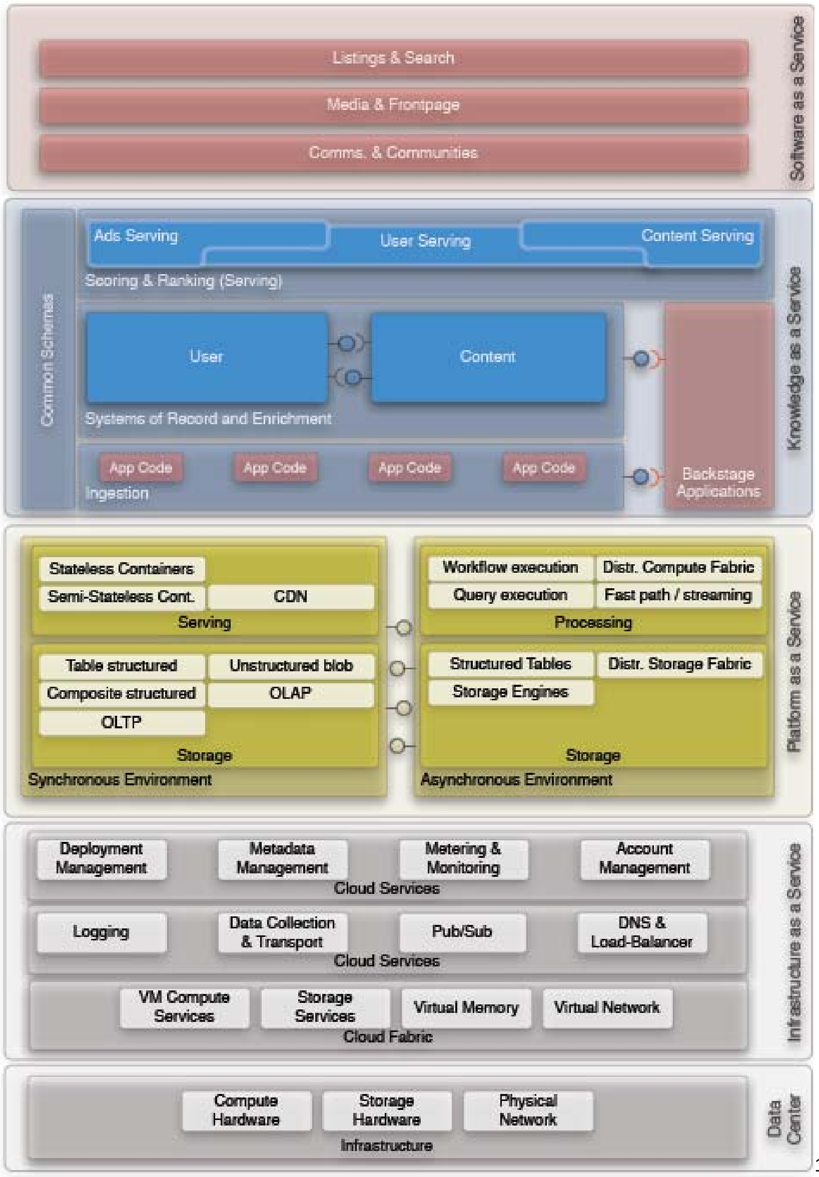
Really Natural

Evolution

- Initial design: A set of machines mounting a filer
- Layers inserted to minimize mounts
- Keys hashed across front-ends to optimize cache
- Replicated as Yahoo expanded
- Storage pulled up into commodity servers
- International deployments drive large-scale data model
  - Unified namespace with partial replicas
- Variations developed to serve specialized use cases
  - Broker presents a unified view

Modern View

- KaaS layer provides aggregated, annotated views of user data
- PaaS layer provides a small number of differentiated onstage storage systems, linked to asynchronous analysis capabilities
- IaaS provides foundational storage



Scale

- Billions of records
  - 600 million+ active users
  - Multiple overlapping databases
- 1 million+ lookups/second
- 500k+ updates/second
- Per-user data size growing fast
- Per-visit requests growing fast

Geography

- Yahoo! is global
- Dozens of datacenters, dozens of markets
- Some properties are global, some regional
  - Example: Formula 1 in the UK, Cricket in India
- Localized access would be nice, but:
  - Global properties require a consistent data view
  - Users just won’t stay put

Problem

- Latency is non-negotiable
  - One of the keys to good user experience
  - Ad exchanges force hard limits
- Everyone wants everything everywhere
  - Fewer, larger datacenters would help
- All dimensions increasing at once
  - Number of records
  - Number of reads and writes
  - Size of records (and reads and writes)

Constraints

- There's only so much bandwidth
- Big-market infrastructure would be wasted in developing markets
  - Reads would be limited, but writes must be applied everywhere
- Data can be sensitive
  - Jurisdictional constraints
  - Terms of service
  - European privacy regulations (for example)

Replication gets complicated

- Per-user replication to relevant regions
  - Moving toward attribute-level replication
  - Relevancy varies by application
- Challenging to verify data consistency
  - “replicas” aren’t
- Chain replication
  - Trades control of bandwidth and fan-out for replication latency

Lessons

- Key to reliability is consistency and architectural simplicity
  - But you can’t have that
- Be prepared for the long haul
  - Replacing large-scale serving systems is hard work
  - Keep it green
- Can’t retrofit consistency
  - Need consistency by design
  - … so you’ll only have to deal with bugs and Byzantine failures
  - … which means you’ll need tools to continuously audit and patch

Questions?