Yahoo! Communities Architectures

Ian Flint

November 9, 2007

Agenda

- What makes Yahoo! Yahoo!?

- Hardware Infrastructure

- Software Infrastructure

- Operational Infrastructure

- Process

- Examples

What makes Yahoo! Yahoo!?

- What do these sites have in common?
  - Del.icio.us
  - Flickr
  - Yahoo! Groups
  - Yahoo! Mail
  - Bix

YAHOO!

What makes Yahoo! Yahoo!?

- Accountability at the property level

  – Architecture

  – Application Operations

  – Infrastructure Decisions

- Incubator Environment

  – Properties function independently on a common hardware platform

  – Highly cost-conscious

  – Open-source attitude

What makes Yahoo! Yahoo!?

- Standards at the infrastructure level
    - Hardware/Software platform
    - Configuration Management
    - Operational tools and best practices
- Executive Involvement
    - Cost
    - Robustness
    - Redundancy

Hardware Infrastructure

Common Platform

Hardware Infrastructure

- Shared Components
  - Network, Data Center, NAS
  - Centrally managed by infrastructure team
- Load Balancing
  - DSR is preferred model
 - Proxy load balancing only where necessary

Hardware Infrastructure

- Hardware (x86, RAID/SCSI)
  - Jointly managed by properties and ops
- Hardware Selection
  - Price/Performance is a constant consideration
  - Supply chain and provisioning cost
  - Reliability vs. Price
    - Single-Homed hosts (even databases)
    - Pooling across multiple switches
    - Fast Failover to mitigate risk of switch failure

Hardware Infrastructure Example

- Layered Infrastructure

- Hosts distributed across multiple racks for power/network redundancy at the pool level

- Really Big Load Balancers doing DSR

The Internet

Router

Load Balancer

Aggregation Switch

Rack Switches

Hosts
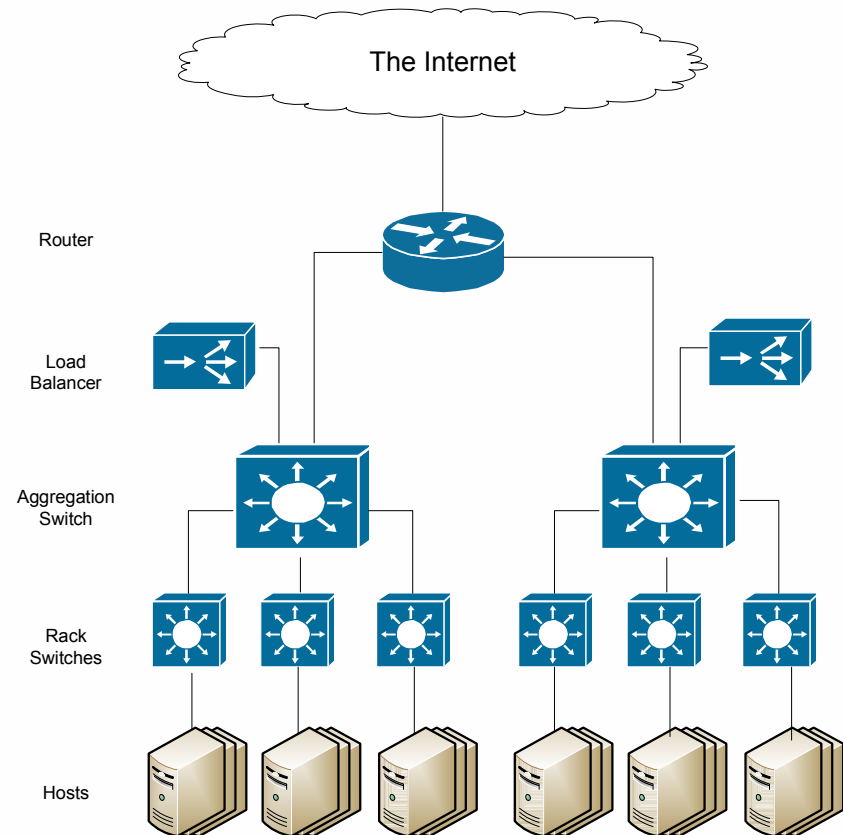
Software Infrastructure

Shared Repository

Software Infrastructure

- OS (FreeBSD, moving to RHEL)

- Databases (MySQL, Oracle)

- Development Platforms

  – PHP (most properties)

  – C/C++ (primary infrastructure platform)

  – Java

  – Python

Software Infrastructure

- Installable components
  - Managed through yinst package manager
  - Stored on common distribution server
  - Examples: yapache, yts, yfor, ymon, yiv, vespa

Software Infrastructure

- More about yinst
  - Robust Package Manager
    - Installation, Versioning, Scripting
  - Implementation
    - Software installed on distribution cluster (package repository)
    - Hosts then pull software (via proxies)
    - Software stored under a common root
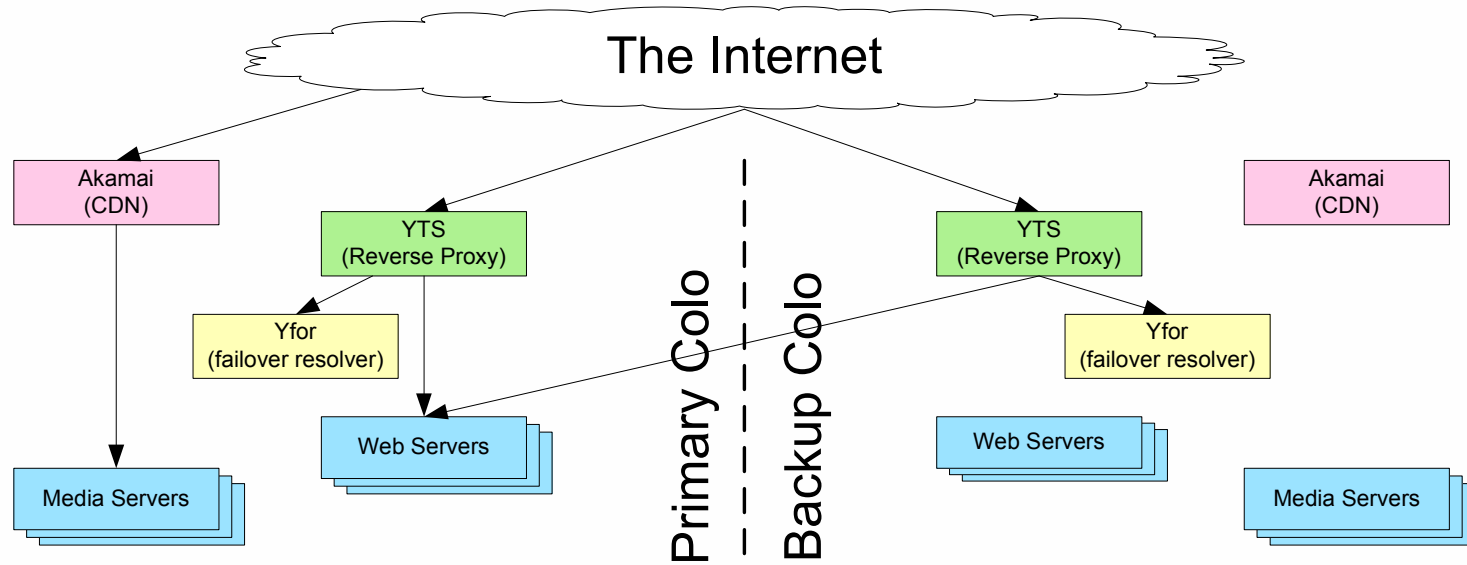   - Used for everything from perl modules to common components to applications

Software Infrastructure

- Shared Infrastructure enables rapid integration of acquisitions
  - UDB
  - SSO
  - SDS
  - YMDB

- External Infrastructure
 - Akamai CDN and DNS
  - Gomez & Keynote

Software Infrastructure - Bix

The Internet

| Akamai (CDN) | | | Akamai (CDN) |

YTS (Reverse Proxy) — Primary Colo | Backup Colo — YTS (Reverse Proxy)

Yfor (failover resolver)   Yfor (failover resolver)

Web Servers   Web Servers

Media Servers   Media Servers

- Global Server Load Balancing between sites
- YTS provides Reverse Proxy and Connection Management
- Yfor provides fast failover from colo to colo
- Media is served via a content delivery network for performance and to reduce load on servers

YAHOO!

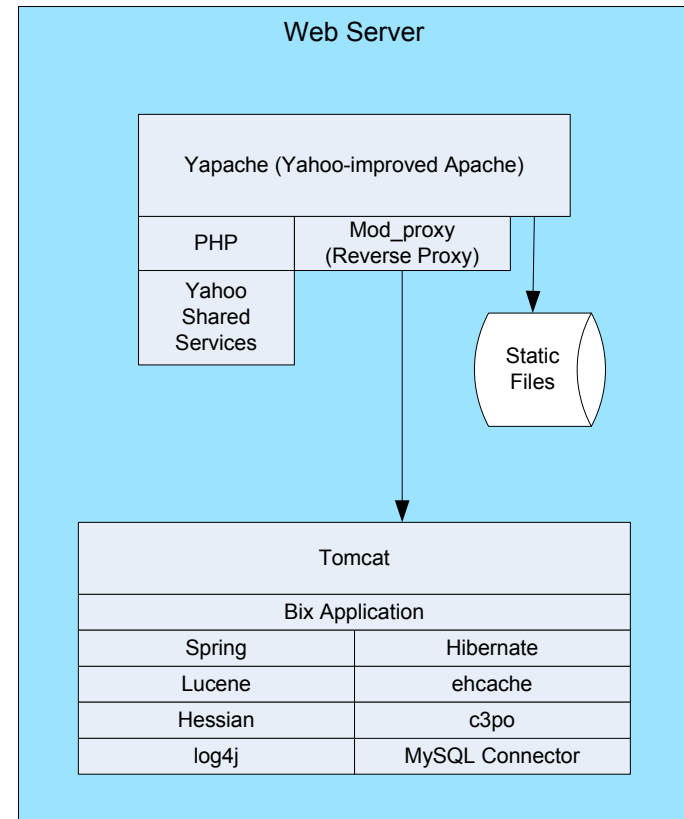# Software Infrastructure - Bix
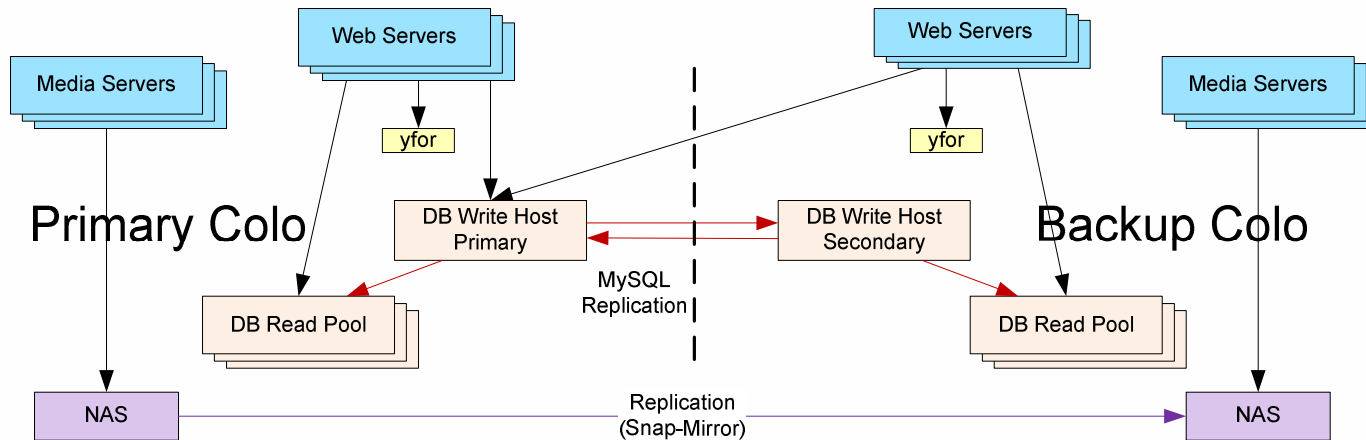
- Yfor Failover Resolver used for fast failover of database connections
- Dual Master MySQL setup for write hosts
- Media storage on NetApp NAS device, with snap-mirroring to backup data center
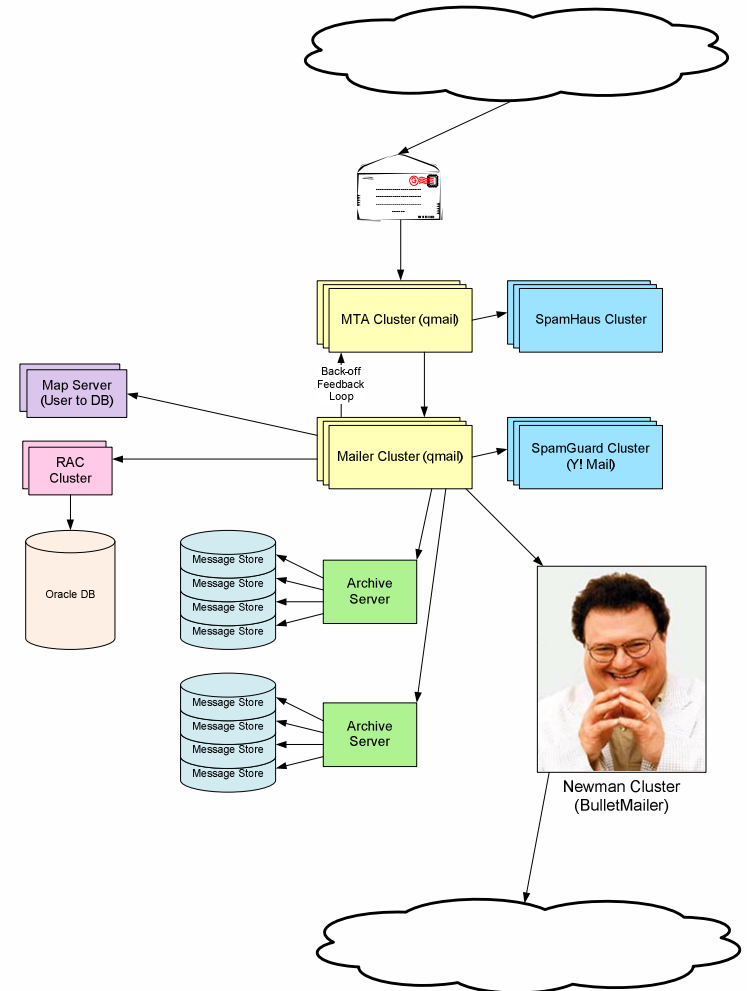
Software Infrastructure - Bix

- Yapache reverse proxy in front of Tomcat instance

- PHP used to access Yahoo shared services

- Static files served from disk

- Fairly standard Java environment (Spring, Hibernate, ehcache, c3po, log4j, etc.)
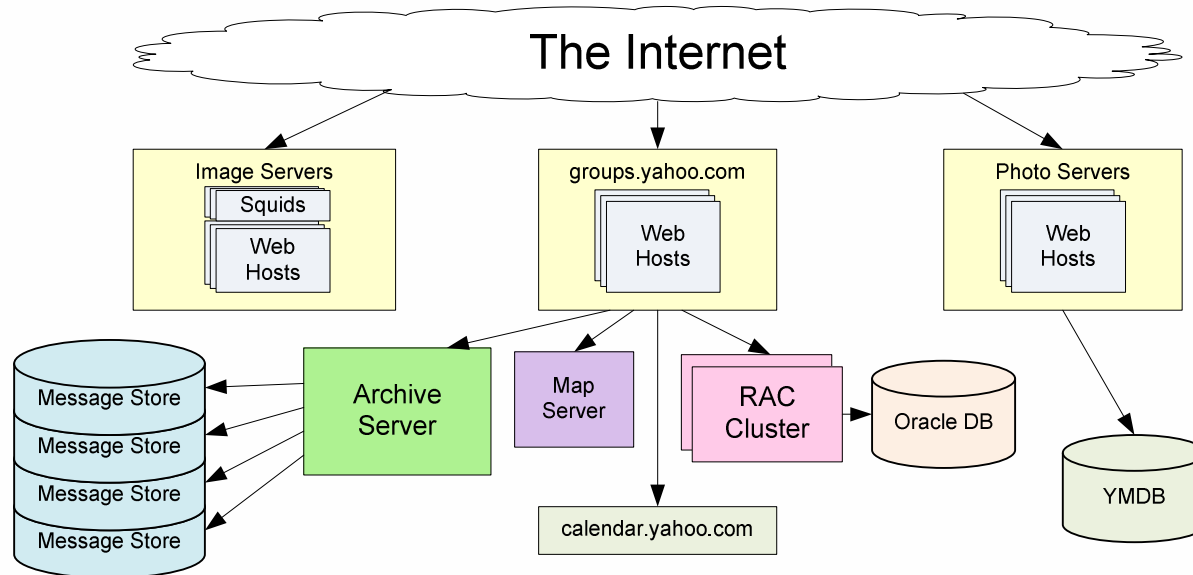
Software Infrastructure - Groups

- Inbound Groups mail hits a qmail cluster

- Mail filtered against real-time blacklist

- Mail forwarded to second qmail cluster

- Proprietary anti-spam algorithms applied

- Mail forwarded to group members

- Mail stored on archive servers

- Oracle RAC clusters store metadata

- Periodic "Electric Potato" measures QoS

MTA Cluster (qmail)

SpamHaus Cluster

Back-off Feedback Loop

Map Server (User to DB)

Mailer Cluster (qmail)

SpamGuard Cluster (Y! Mail)

RAC Cluster

Oracle DB

Message Store
Message Store
Message Store
Message Store

Archive Server

Message Store
Message Store
Message Store
Message Store

Archive Server

Newman Cluster (BulletMailer)

Software Infrastructure – Groups

- Dynamic content served via web pool running python/c++ application
- CSS and images served via a squid-fronted pool
- Group photos on Y! photos infrastructure backed by Yahoo! Media DB (YMDB)
- Database feature implemented as sleepycat DB hosted on message store
- Calendar feature implemented via API calls to calendar.yahoo.com

Operational Infrastructure

Managing the Platform

Operational Infrastructure

- Common Monitoring Infrastructure
  - Nagios
    - Main monitor for clusters
    - Numerous standard plugins
    - Standards/Best Practices around custom plugins
  - Ywatch
    - Basic monitoring of machines over SNMP
    - Heartbeats plus fundamental metrics (IO, CPU, Disk, etc.)
  - Ymon
    - NRPE/NSCA on steroids
    - Automated forwarding of active and passive checks
    - Scripted setup
  - Drraw
    - Data Visualization
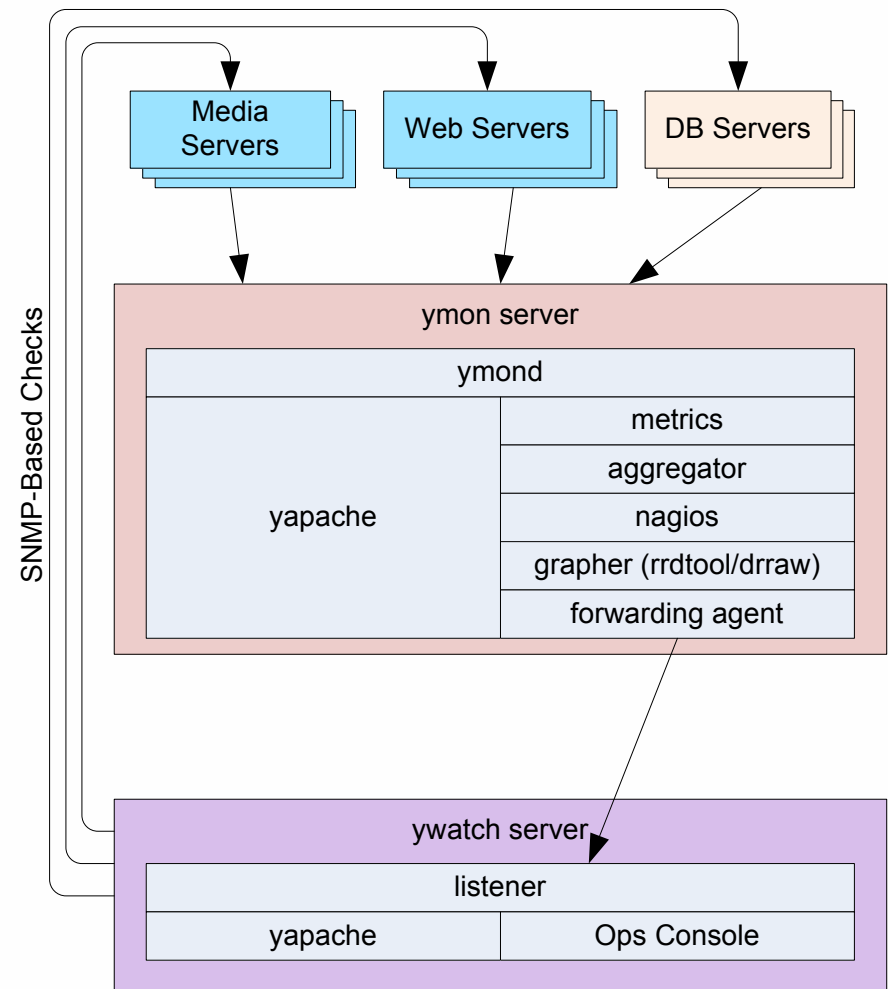   - Deep integration with Nagios and ymon

Operational Infrastructure

- – Rollup Monitoring
  - • Clusters rolled up to centralized monitoring console
  - • Prioritization and correlation of events
- – Internal Site QOS Monitoring
  - • QOS monitoring for sites
  - • Response time and availability
- – "The OC"
  - • 24x7, worldwide operations center
  - • Provides tier 1 and 2 support
- – Centralized CMDB
  - • Configuration Management DB – manages every device
  - • Contact info, escalations, and runbooks included

YAHOO!

Operational Infrastructure Example

- **Application Servers perform checks which are registered by Nagios as passive checks**

- **Metrics are aggregated by metrics module**

- **On-demand graphing is provided by drraw**

- **Nagios alerts are forwarded to central ywatch console**

SNMP-Based Checks

Media Servers

Web Servers

DB Servers

ymon server

ymond

yapache

metrics

aggregator

nagios

grapher (rrdtool/drraw)

forwarding agent

ywatch server

listener

yapache

Ops Console

Processes and Standards

Keeping it sane

Process and Standards

- Hardware Review Committee
  - Strong emphasis on economics
  - Personal attention from David Filo
- Software Review Committee
  - Thinking through major licensing decisions
- Business Continuity Planning
  - Required of all properties
  - Must have <u>and test</u> backup data center
- Paranoids
  - Ongoing site scans
  - Enforcement of standards

YAHOO!

Questions?