5/01/2013

BESTBUY.COM’S
CLOUD ARCHITECTURE
WHO WE ARE

- Best Buy is the world’s largest multi-channel consumer electronics retailer with stores in the United States, Canada, China, Europe and Mexico.
- 11th largest online retailer
- More than 1.6 billion visitors to our stores and BestBuy.com each year
- Reward Zone largest loyalty program in the U.S. – more than 40 million active members
- Provide customers with outstanding choice, unbiased advice and unmatched support for the tech needs
A UNIQUE CUSTOMER PROMISE

• THE LATEST DEVICES AND SERVICES, ALL IN ONE PLACE
• IMPARTIAL & KNOWLEDGEABLE ADVICE
• COMPETITIVE PRICES
• THE ABILITY TO SHOP WHEN AND WHERE YOU WANT
• SUPPORT FOR THE LIFE OF YOUR PRODUCTS
JOEL CRABB

• Chief Architect, BestBuy.com
• Building BestBuy.com’s Ecommerce Platform
• B.S. EE – Washington University in St. Louis
• M.S. NE – University of Wisconsin – Madison
• MBA – University of Minnesota – Carlson School
ARCHITECTURE OVERVIEW

• What’s included
  — Non-functional goals
  — Component system views
  — Real examples
  — Measurements and Volumes

• What’s not included
  — Specific implementation details
  — Products and company names
  — Security information
TERMINOLOGY – HOME PAGE

CUSTOMIZE THE WAY YOU WORK AND PLAY

EXPLOR

EXPLORE

EXPLORE WINDOWS 8


Shop Now

Save up to $100 on Samsung Galaxy S III with 2-year agreement. Plus free shipping.

Shop now

Receive Double Reward Zone® Program Points with select Samsung tablets. Plus free shipping.

Shop now

LOW PRICE GUARANTEE
We'll match prices on qualifying products. See details 

FREE SHIPPING
on everything at BestBuy.com. See details Exclusions apply.
WHAT WE INHERITED

• Monolithic 10 year old architecture
• Minor UI changes took months
• 10,000 line JSPs was normal
• Long standing defects
• Chaotic build and deployment
• Unsatisfied business customers
CLOUD EXPERIMENTS - 2010

• Created a browse-only site to cover outages
  — Always on cloud resident application
  — Elastically scaled and available in around 10 minutes

• Smaller web properties in cloud
  (myrewardzone.bestbuy.com)

• Test environments in cloud
CLOUD RE-ARCHITECTURE

• In 2011 we began planning a new Ecommerce Platform
• Our traffic profile features an approximately 7X peak around Thanksgiving
• Majority of traffic is browse and search
• Re-architect browse tier to cloud for elasticity, scalability and reliability
• Served ~25% of traffic in 2012
OVERARCHING CONCEPTS

• Split Traffic
  — Browse is > 90% of traffic
  — Commerce is most important traffic

• Cache everywhere

• Increase use of Content Delivery Network

• System Isolation
STAKEHOLDER REQUIREMENTS

• Scalability
• Flexibility
• Reliability
SCALABILITY GOALS

• Near-Infinite
• Bursts

• 7X traffic spikes
• Bursts > 50,000 rps
• #3 in eCommerce traffic during holiday

Graph: (from Apr 10, 2012 to Apr 7, 2013) (in millions of hits per day)
(7-day moving averages)
FLEXIBILITY GOALS

- Low cost of change
- Fast concepts to site
- Daily releases
- Multiple versions

One day of work vs. 2 months
RELIABILITY GOALS

• 100% availability
• Zero defects
• ~2s response times

• Achieved 100% cloud uptime during Holiday
CLOUD DOMAIN

- Browse and Search traffic
- Non-transactional
- Non-persistent
- Non-sensitive data
CLOUD ARCHITECTURE CONCEPTS

• Clouds fail, plan for it
  — Multiple availability zones
  — Multiple regions
  — Multiple vendors

• Datacenter connections fail, plan for it
  — Serve pages completely from cloud
  — Browse-only fallback mode
CLOUD ARCHITECTURE CONCEPTS

CDN: Global Traffic Manager

Browse Cloud
Vendor 1

Best Buy Datacenter

Browse Cloud
Vendor 2
CLOUD ARCHITECTURE

- Cloud Load Balancer
- Service Aggregator
- Product Data
- Persistent Cache
- Web App
- Legacy Services and Product Data
- Datacenter
FRONT END APPLICATION

- HTML
- Javascript
- CSS
- Templating framework
- JSON Data Contract with Service Layer
- No Java/JSP in front end
- No MVC framework
Decoupling from back end allows freedom to iterate and deploy independent of back end.

Swapped in Hottest Deals carousel in December 2012 in one day.
FRONT END ARCHITECTURE

CDN: Global Traffic Manager

Browse Traffic

Web Application

JSON Data Contract

Service Aggregator

Legacy Commerce Engine

Commerce Traffic
FRONT END: CDN USAGE

• CDN = Content Delivery Network
• Base page caching
• Edge Side Includes (ESI)
• Image caching
• Traffic routing to multiple clouds and DC
• Overall ~ 80% of requests served by CDN
SERVICES TIER

• Aggregation services
  — Specific to each page
  — Gathers data from 30-50 services
  — Asynchronous
  — Smart caching

• Granular services
  — Specific related data
  — Reusable across views
SERVICES AGGREGATOR - PDP

- Target of 1000ms for 99% of requests
DATA TIER

- NoSQL Product Catalog
  - Schema flexibility
  - Simple REST API
  - Replication – Datacenter to Cloud
  - Pulled from legacy product catalog
  - Data available to all of Best Buy
  - Reliable and Scalable
  - ~20 applications using data since inception
BEST PRACTICES

• Use a distributed service framework
  —Netflix - Hystrix
  —LinkedIn – Rest.li

• Understand velocity of change
  —Front end is high velocity
    • Lightweight
    • Disposable
  —Back end is low velocity
    • Heavyweight
    • Scalable
BEST PRACTICES

• Staged Deployments
  — Started with Home Page
  — Ramped up traffic from 10 – 100% over weeks
  — Added Product Detail Pages in phases

• Have a backup plan
  — Legacy system left intact
  — Occasional rollback to fix defects

• Automate Everything
  — Consistently deploy 100s of cloud VMs
  — Easily create test environments
PROBLEMS ENCOUNTERED

• Data replication is hard at scale

• Scaling takes time
  — Getting from 95% to 99% of calls to our expected SLAs took months

• Legacy data issues
  — Old data formats were difficult to use in redesigned UI
WHAT YOU SEE

Old PDP ~ 4 - 20 seconds

New PDP ~ 2.5 seconds