Answers to the Scaling Challenge:
A Case Study with Answers.com on Scaling with Memcached and MySQL

Presenters:
Dan Marriott, Director - Production Operations, Answers.com
Joaquín Ruiz, EVP Products, Gear6
Answers to the Scaling Challenge:
A Case Study With Answers.com on Scaling With Memcached and MySQL
April 14, 2010

Dan Marriott
Director - Production Operations
danm@answers.com
Overview

Answers.com

The world’s leading Q&A site
Overview

• **Answers.com**
  - Combines power of WikiAnswers® community-driven content with expert licensed content on ReferenceAnswers™
  - Users can ask anything and automatically receive the best available answer.

• **WikiAnswers**
  - Community generated social knowledge Q&A platform, leveraging wiki-based technologies - *answers improve over time.*

• **ReferenceAnswers**
  - Content on millions of topics from over 250 licensed dictionaries and encyclopedias from leading publishers.
Answers.com Sites

• Rank in top Web properties
  ▪ #18 in the U.S. (02/2010)\(^{(1)}\)
  ▪ #31 worldwide (02/2010)\(^{(1)}\)

• Unique monthly visitors
  ▪ 50 million in the U.S. (02/2010)\(^{(1)}\)
  ▪ 72 million worldwide (02/2010)\(^{(1)}\)

• WikiAnswers domain *fastest growing* in U.S.:
  ▪ **2007** unique visitors grew 573% (#1 of top 1,500 Web domains) \(^{(1)}\)
  ▪ **2008** unique visitors grew 154% (#1 of top 200 Web domains) \(^{(1)}\)
  ▪ **2009** unique visitors grew 74% (#2 of top 50 Web domains) \(^{(1)}\)

(1) Source: comScore – Hybrid Measurement Methodology (U.S. only) beginning August 2009
Answers.com Sites

50MM Unique Monthly U.S. Visitors = #18

Feb 2010

Source: comScore – Hybrid Measurement Methodology for Answers sites beginning August 2009
asthma

Did you mean: Asthma (Lyrics - P.O.D.)

Medical Encyclopedia: Asthma

Learn About Asthma
Asthma Has Two Main Causes - Learn How To Treat Both.
www.GoBreath.com

Asthma Information
Get Facts About Asthma Including Symptoms & Trigger Information.
www.Manage-Your-Asthma.com

Definition
Today asthma is viewed as a chronic (long-lasting) inflammatory disease of the airways. In those susceptible to asthma, this inflammation causes the airways to narrow periodically. This, in turn, produces wheezing and breathlessness, sometimes to the point where the patient cannot speak or walk. Obstruction to air flow either stops spontaneously or responds to a wide range of treatments, but continuing inflammation makes the airways hyper-responsive to stimuli such as cold air, exercise, dust mites, pollen, and more. People with asthma experience asthma symptoms during times of stress and anxiety.

Description
Q: Do you claim workman's comp on your income taxes?

A: Workers comp payments should be reported to you via a form 1099 which is also furnished to IRS. If payments are reimbursements for medical treatment and or travel expenses for medical treatment then the answer is no. However, if the workers comp payments are for lost wages or a lump sum settlement for permanent injury then yes, you will need to report these payments to IRS and they are to be included in taxable income for the year they are paid. Remember that IRS is furnished with the total amount paid to you by the Workers Comp agency.

Workers compensation law firm
Speak to a Worker Comp Injury Lawyer in your area for Free
www.InjuryHelpLineLawyer.com

Workers Compensation Claim
NY Workers Comp Lawyers, Free Consultation. No Fee Unless We Win.
www.EBNRLaw.com
Some key infrastructure elements

- 2 Data Centers – Active/Active
- HP Blade Servers (BL460) and 2U Servers (DL380)
- Fusion-io SSDs in all large MySQL DB servers
- VMware
- HP/LeftHand SAN
- Hardware Load-Balancers
Software

Linux – mainly CentOS 5.x/64bit

- Apache/JBOSS
- Memcached
- Lucene / Solr
- MySQL 5.0.x
- Apache-Tomcat
- Memcached
- MySQL 5.0.x
Secret to Web 2.0

Everything runs from memory in Web 2.0.

Evan Weaver - Twitter
Questions: How do I Scale????

Four answers to the scaling challenge from Answers.com:

1. Use Enterprise-grade PCI-SSDs in your MySQL servers
2. Enhance Memcached tier
3. Controlling DB Slave server clusters
4. Hardware Load-Balancers
Scaling Answer #1

*Use Enterprise-grade SSDs in your MySQL servers*
Using SSDs in your MySQL servers

- If you have a large DB read clusters
- Can improve DB server queries/sec by 900%
  - and thereby reduce # servers needed per cluster
- Can improve overall app speed
- Use PCI-based SSDs
  - *Much* faster than SAS/SATA-based SSDs
- Give yourself at least 50% headroom
  - (to handle uncompressing copies locally on SSD)
Quick plug :-)  

Panel:
How Solid-state Technologies are Transforming MySQL Server Performance and the Datacenter Architectures

» Sumeet Bansal (Fusion-io)
» Ryan White (Cloudmark)
» Dan Marriott (Answers.com)
» Vadim Tkachenko (Percona Inc)
» Jeremy Zawodny (craigslist.org)

Weds 5:15pm - Ballroom D
Scaling Answer #2

*Enhance Memcached tier*
Memcached tier - before

- 10 x 16/32GB RAM Memcached Servers per Data Center
- Divided into several clusters
- Striped instances across each cluster
- 10’s millions items/instance
Memcached tier – Challenges

- No redundancy
- Lose 25% of cache (or worse) on any server failure
- Loss of cache = poor performance/user experience
- Costly (OpEx: Rackspace, power, cooling, maint, admin)
Cache data critical to performance

- 10’s of millions of pages
- Pages are dynamic - always publishing
- Unify data on same topic from different data sources onto one page
trounce

**Dictionary:**

trounce (trouns)

v., trounced, troun-cing, troun-ces.

v.tr.
1. To thrash; beat.
2. To defeat decisively.

v.intr.
To censure something or someone forcefully: “I was out to trounce on every digression and indiscretion conducted (or should I say semiconducted) in this performance” (Robert Maxwell Stern).

[Origin unknown.]

**Thesaurus:**

trounce

verb

Antonyms: trounce

Definition: defeat overwhelmingly
Antonyms: fail, forfeit, lose

WordNet: trounce

The verb has 3 meanings:

**Meaning #1:** beat severely with a whip or rod
   Synonyms: flog, whelk, whip, lather, lash, slam, strap

**Meaning #2:** come out better in a competition, race, or conflict
   Synonyms: beat, beat out, crush, shell, vanquish

**Meaning #3:** censure severely or angrily
   Synonyms: rebuke, rag, reproof, lecture, reprimand, jaw, cress down, scold, chide, berate, bawl out, remonstrate, chew out, chew up, have words, lambaste, lambast

Translations: trounce

Trounce

**Dansk (Danish)**
 v. tr. - banke, prygle, gennemhagle
 v. intr. - barke, prygle

**Nederlands (Dutch)**
 inmaken (overwinnen)

**Français (French)**
 v. tr. - flanquer une radée
 v. intr. - imposer une censure sur (qch/qn)
Cache data critical to performance

- Multiple data sources (entries) per page
- Each entry can have multiple links to other entries
- Cache metadata:
  - List of entries for each page
  - Internal links for each entry
Lose memcached node?

- Lose 25% cached data (or more)
- 2-20 addit. MySQL queries to retrieve metadata and data needed to construct page
- Heavy load on MySQL Servers
- Longer page load times
- Site slows down
Original plan to solve memcached weakness

- Engineering team to develop memcached wrapper
- Mirrored memcached server clusters (2x hardware)
- 6 man-months to develop:
  - Write to both clusters
  - Handle node failure
  - Handle cache data restore when node becomes available again
- Also: Address memcached slab issues
Alternative solution – Gear6

- Built-in redundancy function (pair of boxes)
- Mirroring (writes *everything* to both nodes)
- 200GB Memcached space per box (RAM+SSD)
- Uses standard VIP mechanism
- Graceful failover (no impact)
- Automatic cache-sync on node recovery
- Auto fail-back and rebalance VIP
- *No code changes necessary*
Gear6 GUI config example

Service: wa_i_web (2 instances)

<table>
<thead>
<tr>
<th>Memory Usage</th>
<th>Per Instance</th>
<th>Total For Service</th>
<th>Total In Use</th>
<th>Threads</th>
<th>Interface</th>
<th>TCP Port</th>
<th>UDP Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flash</td>
<td>4,096 MB</td>
<td>8,192 MB</td>
<td>4,986 MB</td>
<td>16</td>
<td>wlan24</td>
<td>11211</td>
<td>11211</td>
</tr>
<tr>
<td>DRAM</td>
<td>8,226 MB</td>
<td>16,452 MB</td>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item Count</td>
<td>85,899,345 items</td>
<td>171,798,690 items</td>
<td>132,148,605 items</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Instances**

<table>
<thead>
<tr>
<th>Module</th>
<th>DNS Name</th>
<th>IP Address</th>
<th>Mask</th>
<th>Memory Usage</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>m04</td>
<td></td>
<td></td>
<td>22</td>
<td>61% 2,493 MB</td>
<td>Remove</td>
</tr>
<tr>
<td>m03</td>
<td></td>
<td></td>
<td>22</td>
<td>61% 2,493 MB</td>
<td>Remove</td>
</tr>
</tbody>
</table>
Testing

• Same blisteringly fast memcached performance
  ▪ use of slower SSD was undetectable

• Tried to break it
  ▪ yank power, kill network ports, pull SSD drive

• Try software upgrade
  ▪ seamless failover/cache-sync on node-up

• Discovered vastly improved slab management
  ▪ plus ability to fine-tune instance memory usage

• Scalable
  ▪ can handle 60,000+ ops/sec
Value-add with Gear6 memcached

- Improve app reliability
- Ensure no SPOF at memcached layer
- Scale up our database infrastructure safely
- Significantly reduce TCO by decommissioning 20 servers
- Save 6 man-months of custom memcached wrapper development
Gear6 Value-add

...cont’d

• Insight into memcached performance

• GUI helped us troubleshoot app issues

• First class support (even at 4:00am!)

• New in v3.0: Ability to dynamically resize memcached instances with no cache loss!!
Scaling Answer #3

Controlling DB Slave server clusters
Controlling DB Slave server clusters

- Maint. required sending command to Load Balancer to disable slave node in cluster
- On DB node failure, lost ‘000s queries before monitoring noticed and issued LB command

Solution: Install lighty (lighttpd) on every DB server
Using Lighty to Control DB Slave clusters

• Config lighty to make lightweight “healthcheck” call to local DB (SHOW STATUS LIKE 'Threads_running')

• Can set maint. flag on any DB server

• Have LB monitor make http call to lighty (http://servername:3300/dbcheck?p=3309)

• Works for multiple MySQL instances on same box
Lighty DB health-check

VIP Status: Green
up-3312.txt: Present.
DB Status: Up (Threads_running: 23)
if maint_flag = TRUE
    return “red”
elseif dbcall not good
    return “red”
else
    return “green”
end
Scaling Answer #4

HW Load Balancers
Hardware Load Balancers

- Avoid making firewall your bottleneck
- Immense flexibility [CLI, stats, ecv]
- Quick reconfigurations, adds/changes/deletes
- Highly secure
- Caching, Compression, GSLB, throttle for safety
Which Hardware Load Balancers?

• **Citrix & F5**
  - Leaders of the pack *(Gartner Magic Quadrant)*
  - Amazing performance
  - Supremely flexible
  - Advanced GSLB
  - Expensive

• **A10 Networks**
  - Up and coming
  - Great performance
  - Decent feature set
  - Cost-effective *(certainly for internal load-balancing)*
Thank you.

Slides: http://tinyurl.com/mysqlconf2010-scaling-tips
Gear6: *Perspectives on the scaling challenge*
MASSIVE increase in:

Population

Traffic

Content & Applications

- digg
- Answers.com
- Facebook
- YouTube
- Twitter
- Yelp
Essentially the SAME Major Building Blocks

- Web
- App
- RDBMS
Shift to Web 2.0 services puts strain on origin sites
Can’t Accelerate Using CDN
LATENCY: The Web 2.0 Killer

Latency: Time is the Web 2.0 killer
Hierarchy of Needs: Survive!

Accelerated Data: i.e. CACHE = Survival!

Web Stack

Data Cache

Storage

RDBMS

App

Web

Net Interface

Proxy

Load Balancer

Internet

Clients

CDN

Clients
Web 2.0 Needs: Memcached

In LAMP/JAVA/Ruby World
Data Cache = Memcached
Memcached = Survival
Is there more to life?
Scaling Needs: More than Memcached

- Sharing?
- Ops?
- Query?
- Persistence?

Acceleration
Foundation **Acceleration** Services for Gear6 Memcached:

- Memory on-boarding
  - Hybrid storage: DRAM and Flash
- Clustering
- Replication
- Elastic pool sizing
- Data management
  - Enhanced slab management
  - Enhanced key space management

*These basic features have been built into the Gear6 Memcached Server*
**Gear6: Beyond Acceleration**

- **Persistence**
  - Local or In-network snapshots
  - Rapid restore and cache warm-up
  - Offline analytics
  - Primary store: Unstructured datasets

- **Query**
  - Native Memcached Regex matching
    - Key-based
    - Value-based
    - K-V based
  - Deletes: K, V, and K-V matching delete operations
Gear6: The near term

- **Operations**
  - Advanced atomic operations based on integration with Redis
  - Memcached interface
  - Elements: Strings, Sets and Lists
  - Atomic support
    - Push, pop, add, remove
    - Unions, Intersections, Diffs

- **Sharing**
  - Cloud to origin synchronization services
  - Multi-site replication
  - “Dynamic Data Bus”
Gear6: Beyond Memcached

Web Stack

Storage

RDB MS

App

Web

Net Interface

Proxy

Load Balancer

Internet

Clients

Gear6

Storage Interface
Gear6:
• Enhanced Memcached for accelerating MySQL / RDBMS
• KV operation store for unstructured data sets
Gear6 products available as:
- Cloud images: free and paid on Amazon and GoGRID
- DEB and RPM packages: free and paid
- Dedicated data center server images: paid
- HP solutions

More resources:
- http://www.gear6.com/memcached-resources/query
- http://www.gear6.com/memcached-resources/export
- http://www.gear6.com/memcached-resources/tools
QUESTIONS?
Thank You

Presenters:
Dan Marriott, Director - Production Operations, Answers.com
Joaquín Ruiz, EVP Products, Gear6