**MySQL (recap)**

- Which segment of the market is MySQL strong in?
  - Large Companies or Small Companies?
  - Web Applications or Critical Enterprise Data?

- Why would a major enterprise want to pay so much more for an Oracle or IBM DB?

**MySQL: Market**

<table>
<thead>
<tr>
<th>Market Segment</th>
<th>Small 20%</th>
<th>Medium 30%</th>
<th>Large 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterprise web data</td>
<td>Microsoft</td>
<td>Oracle</td>
<td>IBM</td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>Scalability</td>
<td>Support</td>
</tr>
<tr>
<td></td>
<td>Longevity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web Sites</td>
<td>MySQL</td>
<td>Cost</td>
<td></td>
</tr>
</tbody>
</table>

How should MySQL grow in order to meet its stated goal of getting to $100 million in revenue?

---

**MySQL: Growth Strategy**

<table>
<thead>
<tr>
<th>Market Segment</th>
<th>Small 20%</th>
<th>Medium 30%</th>
<th>Large 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterprise web data</td>
<td>Microsoft</td>
<td>Oracle</td>
<td>IBM</td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>Scalability</td>
<td>Support</td>
</tr>
<tr>
<td></td>
<td>Longevity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web Sites</td>
<td>MySQL</td>
<td>Cost</td>
<td></td>
</tr>
</tbody>
</table>

- Lack of brand identity in this segment
- MySQL lacks the organization to offer support
- Large enterprises have high switching costs

---

**MySQL: Growth Strategy**

<table>
<thead>
<tr>
<th>Market Segment</th>
<th>Small 20%</th>
<th>Medium 30%</th>
<th>Large 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterprise web data</td>
<td>Microsoft</td>
<td>Oracle</td>
<td>IBM</td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>Scalability</td>
<td>Support</td>
</tr>
<tr>
<td></td>
<td>Longevity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web Sites</td>
<td>MySQL</td>
<td>Cost</td>
<td></td>
</tr>
</tbody>
</table>

- Not a big enough market to reach stated $100 million goal.
**My SQL: Growth Strategy**

<table>
<thead>
<tr>
<th>Small 20%</th>
<th>Medium 30%</th>
<th>Large 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft</td>
<td>Oracle IBM</td>
<td>Maybe?</td>
</tr>
</tbody>
</table>

- Many of these customers already using MySQL with websites
- Least emphasis on global organization
- Leverage SAP alliance
- Up against Microsoft

**My SQL: Growth Strategy**

<table>
<thead>
<tr>
<th>Small 20%</th>
<th>Medium 30%</th>
<th>Large 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterprise</td>
<td>Web sites</td>
<td>My SQL</td>
</tr>
<tr>
<td>50%</td>
<td>10%</td>
<td>10%</td>
</tr>
</tbody>
</table>

- Builds on existing brand and strengths
- Market not as big

**Networks**

What are some examples of communications networks?

- Public Telephone Network
- Internet
- LANs (Local Area Networks)

**What does a network do?**

1. Transport data from one host to another.

**Network Architecture**

- Network architectures are layered
- Each layer
  - uses the services of the layers below
  - To offer more advanced services to layer above
- Allows layers to be designed independently
- We will talk about 3 layers next...

```
Network
Link
Physical
```
Physical Layer: Convey bits over a wire

Bits: 010110...

- **Link Layer**
  - Make a Frame link out of a bit link
  - Instead of endless sequence of 1s and 0s, we want distinct packages of data that are separate from each other
  - Say we want to send 2 Frames with data
  - 010101010101010 and 101010101010101
  - Concatenate them and send them as a sequence?
  - How can the receiver tell where the new frame begins?
  - Solution: insert a special sequence at the start of frame: for example: 0111110

- **More Link Layer.. -- Ethernet**
  - Want to allow multiple hosts to share a link
  - How do they avoid talking at the same time?
    - Don't transmit if you hear another host transmitting
    - If there is a collision, stop wait a random amount of time, and try again
  - This is a Medium Access Control (MAC) Protocol

- **Physical layer**
  - Other schemes for mapping a bit sequence to a physical sequence are possible.
  - These are called modulation schemes

- **Link Layer (cont'd)**
  - Also does error detection/correction
  - Insert extra information the helps the receiver to determine if the data has been corrupted.
  - Example: parity bit
    - Sender adds a 1 or zero to end of data so number of ones is always odd
    - 10011 or 10000
    - If receiver counts an even number of ones, then it knows the data was corrupted

- **Ethernet Continued**
  - How do the hosts on this Ethernet identify each other?
    - Each host (actually each interface)
      - has a globally unique MAC address
      - Cannot be changed
**Ethernet Hub**
- Hub broadcasts packets on a link to all others
- As if all hosts connected to single link
  - We say it is a Single collision domain
  - Only one host can talk at a time

**Ethernet Switch**
- If switch knows where the destination is
  - Switch forwards an incoming frame to destination only.
  - Otherwise, it broadcasts it to everyone.
- Thus, parallel conversations possible.

**Network Layer**
- A wants to send some data to C
  - Suppose A knows C's address
  - A sends a packet towards C
  - A marks his packet with C's Address (an IP Address)

**Post Office Analogy**
- Alan Smith
  - 1155 High St
  - Santa Cruz 95064
- Alice
  - 41 Main Ave
- Truck to London
  - Plane to SFO
  - Plane to SFO
- SFO Post Office
  - Truck to Santa Rosa
  - Truck to Santa Cruz
  - Look at address
    - Make forwarding decision

**Network Layer**
- Host A: 128.144.60.202
  - (IP Address)
- Host B: 128.144.60.201
- Host C: 128.144.60.202
- Host D: 128.144.60.203

**Network Layer**
- Host A: 128.144.60.202
- Host B: 128.144.60.201
- Host C: 128.144.60.202
- Host D: 128.144.60.203

- A uses Link 1 to send to B
- B looks at
  - Packet Header
  - Routing Table

**Address:** 128.144.60.202
**Next Hop:** 128.144.60.201
**Link 2**
Routing in the Internet

Many feasible paths from source to destination.

Routing

- Updating the routing table
- Objective: each packet gets closer to destination

Packet forwarding

- Transmitting each packet on the appropriate output link
- Based on routing table

Routing Algorithms

Routers talk to each other to build their routing tables

Routing Table has Wild Cards

Internet Routing is Hierarchical

Routing Concerns

- Long routes
- Circular routes
- Hijacking routes
- Route flapping
IP Addresses vs Mac Addresses
- Hierarchical
  - The beginning bits tell you which network the host is on.
  - Ex: UCSC addresses start with 128194.XX
  - The last bits tell you which host of the network.
- Not Hierarchical
  - Beginning bits tell nothing useful.
- Changeable
  - Changes with location of Host.
- Not Changeable
  - 6 bytes
  - Only 4.2 billion

Issues In Networking
- Sharing of Limited Resources
  - How Should A and B share a link with limited bit rate?

Issues In Networking
- Time Division Multiplexing
  - Gives each connection the use of the link a fixed fraction of time.
  - Fixed fraction of resources reserved for each connection.
  - Technology called circuit switching.
- Problem
  - When A is silent, A’s fraction of link goes unused.

Issues In Networking
- Statistical Multiplexing
  - Link shared in such a way that connections are not assigned fixed fraction of Link.
  - A and B unlikely to offer peak rate at the same time.
  - \( \max(A + B) + \max(A) + \max(B) \)

Statistical Multiplexing
- Because resources aren’t reserved, it’s possible offered load too high.
- Packets are put into a queue.
- If offered load remains too high, queue will fill up and overflow.
Transport Protocols

- The Internet is unreliable
  - It will make a "best effort" to get your packet to its destination
- Packets can be lost because of
  - Congestion
  - Link errors
  - Routing problems

<table>
<thead>
<tr>
<th>Application</th>
<th>Presentation</th>
<th>Session</th>
<th>Transport</th>
<th>Network</th>
<th>Link</th>
<th>Physical</th>
</tr>
</thead>
</table>

Transmission Control Protocol (TCP)

- Retransmit mechanism for reliability
  - Receiver sends acknowledgements to sender
  - If a packet is lost, source fails to get ACK, and then retransmits.
- Congestion control
  - If congestion perceived (by lost packets)
  - Source reduces its send rate
    - When lost, sender reduces send rate by half
    - Otherwise slowly increases

TCP cont’d

- TCP port numbers
  - TCP Header has a "port" number field
  - Helps host sort out how to route packets to applications

```
Your Computer

Port 80 Packet

Port 80

Email Client
```

UDP

- For some applications, packet retransmissions are not worthwhile
  - Why?
- For these applications, we use UDP
- UDP is a transport protocol that
  - Does not do retransmissions
  - Does not do congestion control

Congestion Control

- When networks are congested, certain sessions (Source-destination pairs) should reduce offered rates.
  - Today all TCP sessions slow down when they detect packet losses.
  - UDP sessions do not slow down.

- What are some alternative strategies?
  - Have those whose applications aren’t as sensitive slow down more?
  - How would we know which are less sensitive.

Pricing within the Internet

- Customer pays an ISP
  - Often flat rate per month
- ISP pays a backbone AS
  - Often just flat rate, dependent on access link speed
  - Sometimes based on total usage
- Backbone NSPs peer with each other
  - Often for free if they exchange comparable amounts of traffic
- Overall
  - Internet billing today is much more course-grained than telephone billing.
Domain Names

IP addresses are inconvenient for people
- 32 bits hard to remember
- 128 bits very hard to remember

Domain names
- e.g. argus.eecs.berkeley.edu
- Easier to remember than IP addresses
- However, we need some way of mapping domain names to IP addresses.

Hierarchy in Addresses vs. Names

Addresses hierarchical in topology
- Maximize "wild cards" and distribute address administration

Names hierarchical in administration
- Single administered organizations often distributed topologically (e.g. ibm.com)

Domain Name System (DNS)

Transport Protocols

- The Internet is unreliable
  - It will make a "best effort" to get your packet to its destination
- Packets can be lost because of
  - Congestion
  - Link errors
  - Routing problems

OSI Layers

Some Typical Topologies

Home Network
**Small/Medium Business**

- Ethernet Switch
- T1 Switch
- Router with Firewall
- Web Site Server
- T1 Line

**ISP Topology**

- Telephone Company
- Local Office
- ISP Point of Presence
- DSL Modem
- Leased Line
- ISP

**Network Service Provider**

- Network Access Point

**Large E-Business**

- Load Balancer
- Web Servers
- Application Server
- Databases
- Customers
- Merchandise
- Orders
- Interconnected with Gigabit Ethernet or other technology
- Logic Flow of Interaction
- Presentation Logic (Assembling Web page)

**Web Caching**

- Speed up web page loading by storing previously seen components locally

**Akamai Case**

- Web server
- Web caching
- Cache on Hard Drive
- http://www.ucsc.edu
- UNIVERSITY OF CALIFORNIA
Akamai Case

1) Akamai’s technology can be best classified as
   - A) Content Delivery Network
   - B) Database Management System
   - C) Thin Client

2) Akamai’s customers included
   - A) Home users wanting faster Internet access
   - B) Companies with content-heavy websites
   - C) Companies wanting a simple, inexpensive database

3) Which is _not_ one way Akamai sold its product/service?
   - A) “Partner” firms like companies who do system integration
   - B) Retailers like Fry’s and Best Buy
   - C) A sales force employed by Akamai

Internet Bottlenecks

- First Mile (Server Capacity) – 70% of website performance problems according to one study
- Backbone – Plentiful, but some shortage within metropolitan areas
- Peering – Exchange of traffic between NSPs
- Last Mile to home
  - 56 K modems are slow
  - Shared LAN limitations

Solutions

- Expand Bandwidth
  - Being done
- Mirroring web sites
  - Put exact copy of some web page to multiple servers
  - Tricky to duplicate content
- Caching
  - Problem: Stale Content
  - Problem: Hard to count “click throughs”
- Content Distribution Networks...

Freeflow

- Deployed in 1999
- Akamai Infrastructure
  - 13,000 servers in 954 networks by 2001
- Customers –
  - Large Commercial Websites
- Revenue model - $2000 per mbps served
  - (For comparison, normal Internet access cost 500 mbps at time)

2000 Financials

- $196 Million Loss (before special charges)
- $90 million revenue
- 12% gross margin, after deducting
  - server depreciation
  - payments to network partners
  - Data center space
- But, most expenses of shouldn’t grow at same rate as number of customers; so margins should improve
- $201.5 million SG&A
  - (selling general and administrative)
  - (largely sales force cost)
  - Again, this might not grow at same rate as number of customers
- $40 million R&D
**Competition**

- Hosting firms (substitute)
  - Exodus
- Other CDNs
  - Sandpiper, Adero, Mirror Image
- Content Alliances
  - Akamai's competitors banded together to share networks

**2001 Market Changes**

**Bad**

- Dot-coms bust
- Customers leave
  - "churn rate goes to 22% per quarter"

**Good**

- Hosting firms go bust (exodus)
- Some CDN competitors go bust
- Competing CDN alliances mired in problems

**EdgeSuite**

- Assemble dynamic pages at edges rather than just serve heavy objects
- Value proposition
  - Performance improvement
  - Cost and complexity reduction
  - Scalability
  - Security
- Pricing – higher than old service
- Soon edge suite dominated revenue

**Technology**

Dynamic CDN technology: ESI (edge sides includes)

Develop as open standard why?

Akamai not big and credible enough to force a de-facto standard on market

**Marketing**

- Difference in selling old vs new products:
  - Old product
    - Geared toward speeding up websites
    - Revenues of their clients depended on speed
    - Easier to get sale
  - New Product
    - Simplify company IT function
    - Cost vs. revenue center
    - Harder sell. More data driven.
    - Consequently new product needs more professional sales force
- Channels?
  - Distribution Partners (IBM) credibility
  - Direct Sales Force too
## Recent Performance

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Net Income</strong></td>
<td>$258,051</td>
<td>$18,219</td>
<td>$184,030</td>
<td>$181,239</td>
<td>$85,799</td>
</tr>
<tr>
<td><strong>Total cost and operating expenses</strong></td>
<td>$164,034</td>
<td>$72,041</td>
<td>$172,600</td>
<td>$3,077,062</td>
<td>$806,030</td>
</tr>
<tr>
<td><strong>Net income (loss)</strong></td>
<td>34,004</td>
<td>(51,561)</td>
<td>(284,617)</td>
<td>(2,953,082)</td>
<td>(820,950)</td>
</tr>
<tr>
<td><strong>Net income (loss) attributable to common shareholders</strong></td>
<td>34,004</td>
<td>(51,561)</td>
<td>(284,617)</td>
<td>(2,953,082)</td>
<td>(820,950)</td>
</tr>
</tbody>
</table>

(All amounts are in thousands, except per share data.)