Building Security at Scale

Stanford Computer Forum 2014

Alex Stamos
Yahoo!
Who am I?

- 36 days as CISO of Yahoo
- Founder of Artemis Internet
- Co-Founder of iSEC Partners
- @stake, Loudcloud
- Cal BS EECS ‘01, worked on Patterson team
Take-Aways from Today

1. Internet-scale companies have unique economic security drivers
2. The security industry does not serve us well
3. Most academic research does not help
4. There is a huge opportunity for both academia and industry to work with us
5. Our problems will be everybody’s problems soon
When you think of an industry that is subject to online attacks, what first comes to mind?
How are these firms related?

*Millions* of customers pay dozens to *hundreds* of dollars visit *rarely* and *have* meat-space identities.
How about for the Web Scale Companies?

*Billions* of customers pay *nothing* (but click on ads) visit *often* and have *no* link to real-life
<table>
<thead>
<tr>
<th></th>
<th>Big Banks</th>
<th>Online Payments</th>
<th>Web Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Customers</td>
<td>$x 10^7$</td>
<td>$x 10^8$</td>
<td>$x 10^9$</td>
</tr>
<tr>
<td># of Concurrent Users</td>
<td>$x 10^4$</td>
<td>$x 10^5$</td>
<td>$x 10^8$</td>
</tr>
<tr>
<td># of FE Servers</td>
<td>$x 10^2$</td>
<td>$x 10^3$</td>
<td>$x 10^4$</td>
</tr>
<tr>
<td># of Total Servers</td>
<td>$x 10^4$</td>
<td>$x 10^4$</td>
<td>$x 10^5$</td>
</tr>
<tr>
<td>Customer Value</td>
<td>$100's$</td>
<td>$10's$</td>
<td>$.01s</td>
</tr>
<tr>
<td>Cust Stickiness</td>
<td>High</td>
<td>Medium</td>
<td>Low-Medium</td>
</tr>
<tr>
<td>Meat-Space Identity</td>
<td>Strong</td>
<td>Moderate</td>
<td>Weak</td>
</tr>
<tr>
<td>Post-Facto Action?</td>
<td>Yes</td>
<td>Yes</td>
<td>Rarely</td>
</tr>
</tbody>
</table>
Two totally different problems:

Banks are protecting real customers from attack

Web companies are trying to figure out which users are assets and which are liabilities
Things people try to sell us
Things people try to sell us: Smart Firewalls!

**Description**
A true masterpiece of engineering. Delivers the highest density, lowest power, and fastest Ethernet switching system.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switching Capacity</td>
<td>30Tbps</td>
</tr>
<tr>
<td>Linecard Capacity</td>
<td>3.84Tbps</td>
</tr>
<tr>
<td>10Gbe Interfaces</td>
<td>1152</td>
</tr>
<tr>
<td>40Gbe Interfaces</td>
<td>268</td>
</tr>
<tr>
<td>100Gbe Interfaces</td>
<td>96</td>
</tr>
<tr>
<td>Forwarding Rate</td>
<td>14.4Gbps</td>
</tr>
<tr>
<td>Total Buffer</td>
<td>144GB</td>
</tr>
<tr>
<td>Rack Units</td>
<td>11</td>
</tr>
<tr>
<td>Nominal Power Draw</td>
<td>5050W</td>
</tr>
</tbody>
</table>

- 120 Gbps firewall throughput (App-ID enabled\(^1\))
- 100 Gbps threat prevention throughput (DSR\(^2\) Enabled)
- 60 Gbps threat prevention throughput
- 24 Gbps IPSec VPN throughput
- 24,000,000 max sessions
- 720,000 new sessions per second
- 25/225 virtual systems (Base/Max\(^3\))
Things people try to sell us: DB Backed SIEM

“Just dump your data into Oracle and put it on a SAN”

1PB=$1.4M

1PB=$80K
Things people try to sell us: IDS Appliances

300Gbps x $15/Mbps = $4.5M
Things people try to sell us: Reputation Services

“Call our web service with the data and we’ll return a result in only 2000ms.”

In an ad-supported business, latency is death.
Aren’t we a special case?

Not really…

- Big data means that power efficiency is becoming a competitive advantage for many
  - Finance
  - Biotech
  - Logistics and Operations

- Latency is also more important than ever
  - See “Flash Boys” by Michael Lewis
Where security needs to go
Collapse the perimeter

Security services need to be as close as possible to the data you are protecting:

- Anomaly/Intrusion Detection
- Data Encryption
- AAA
- Network access control

Only sell software. Pizza boxes are great for pizza.
False Positives are Death

.01% False Positive Rate x 800M MAU = 80,000 alerts

- Alerting isn’t my problem
- The response funnel needs to narrow quickly
Latency is Death

Security needs to move towards asynchronous reactions

DRM world provides good examples
Better Mousetraps
Freemium Key Management and App Auth

Dual-auth TLS is the future of app auth

- Conceptually simple
- Open-source foundation
- Decentralized failure modes
- X.509 reasonably flexible

Why isn’t there a MySQL for Auth?
Bug Bounty with Automated Verification

Bug bounties create huge problems for companies.

Why can’t the reporter upload a selenium script that verifies the issue?

Solving this would open SME market.
Reputation Services that Work

Industry is moving to “slow auth”

What I want is:

● Open-Source
● To benefit from other sensors with privacy
● Realistic geo-based tracking
● Accuracy with IPv6
Hadoop Based SIEM

Proprietary distributed file systems will die

- Let me figure out how I store my data

Give me:

- Fast scrubbing/tokenization
- Natural language search
- Useful visualization
- Pre-defined but tunable anomaly models
ARM Based Secure Systems

ARM is going to take over the datacenter

If you could go back and build the x86 datacenter, what would you do?

- Lightweight containers
- Aggressive anti-exploit
- Trusted, diskless boot
Thank you!

alex@stamos.org
stamos@yahoo-inc.com