Avoiding Cascading Failures

SRECON2016 - Craig Fender - Ravindra Punati
Avoiding Cascading Failures

Craig Fender

Craig is presently a Senior Technical Duty Officer at eBay and is responsible for commanding all types of large scale site incidents. In addition to an undergraduate degree Craig holds numerous professional certifications related to the computer industry (RHCE, SSCA, ITIL and et cetera). Craig has held several roles at multiple start-up and fortune 500 companies such as Senior Systems Engineer, Project Manager, Presenter and Major Incident Commander.

Ravindra Punati

Ravindra Punati is leader of the Site Reliability Engineering team. In other roles at eBay Ravi has been responsible for the infrastructure automation initiatives and cloud operations. Ravi brings extensive expertise in the fields of database engineering, application development and software as a service product lines. In addition to holding multiple degrees in computer science Ravi has held several roles as an engineer, architect, manager and executive in various silicon valley start-ups.
EBAY AT A GLANCE
EBAY AT A GLANCE

- **$8.6B** Revenue in 2015
- **162M** Global Active Buyers *
- **800M** Live Listings *
- **57%** Percentage of eBay Inc. revenue that is international *
- **190** Markets eBay apps are available in *
- **$82B** GMV in 2015

*Q4 2015 data
EBAY LEADS IN MOBILE

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>eBay Application downloads *</td>
<td>304M</td>
</tr>
<tr>
<td>Percentage of GMV closed on Mobile *</td>
<td>43%</td>
</tr>
<tr>
<td>2015 MCV across eBay’s portfolio of apps</td>
<td>$33B</td>
</tr>
<tr>
<td>Every 28 sec. a tablet is bought via mobile in the U.S. *</td>
<td>9.2M</td>
</tr>
<tr>
<td>Every 10 sec. a ladies handbag is bought via mobile in the U.S. *</td>
<td></td>
</tr>
</tbody>
</table>

*Q4 2015 data
EBAY OPERATIONS
Background on eBay Operations
Background on eBay Operations
Background on eBay Operations

• In any mission critical real-time environment where one is highly incentivized toward maximum uptime one must have:
  – redundancy,
  – resiliency and
  – robustness.
• These needs give rise to complexity.
• Complexity can lead to fragility.
  – Multilayered dependencies can cause cascading failure.
• To manage that complexity certain behaviors, culture, principles and technology emerge as the most successful.
Background on eBay Operations

- eBay itself is like a bridge between buyers...
Background on eBay Operations

...and sellers
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

• Pictures and slides about what an SEC is and who we are – see Excellence in engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in engineering presentation.
Background on what SEC and Operations

Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation.

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in engineering presentation
Background on what SEC and Operations

• Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations
• Pictures and slides about what an sec is and who we are – see Excellence in engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

Pictures and slides about what an SEC is and who we are – see Excellence in engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

• Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation.
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
Background on what SEC and Operations

• Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation
Background on what SEC and Operations

- Pictures and slides about what an SEC is and who we are – see Excellence in Engineering presentation

PREVENT CASCADING FAILURES
SO FULL OF FAIL

Creating resiliency through *intelligently* injected failure.
So Full of Fail – Failure is NOT an option. 
*It’s a requirement!*

- The inability of a system or system component to perform a required function within specified limits. A failure may be produced when an error is encountered.
So Full of Fail – Failure is NOT an option.
It’s a requirement!
So Full of Fail – Atomic failures – Overview

• **Software**
  – Database failure
  – Service failures

• **Hardware**
  – LB failures
  – Compute failures
  – Network failures
A cascading failure is a failure in a system of interconnected parts in which the failure of a part can trigger the failure of successive parts or the whole system.
So Full of Fail – Cascading failures
So Full of Fail – Cascading failures
So Full of Fail – Cascading failures – Software

- **Service to Service**
  - Service A talks to Service B using load balancer virtual ip

- **Service B to Database**
  - Service B talks to data store using data access layer
So Full of Fail – Cascading failures – Software
So Full of Fail – Cascading failures – Software

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>half-agg-sl2-001</td>
<td>100% / 100%</td>
<td>H: 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kemmis-app</td>
<td>33% / 16%</td>
<td>H: 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INFRA cmevpqry-app</td>
<td>10% / 3%</td>
<td>H: 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INFRA infciscatalog</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...isckConsumer.ORDER.FUNDING.UPDATE</td>
<td>BES MD: 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...ngCart.TnsEvaluateService.HTTP11</td>
<td>SC MD: 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...io.AsyncHttpClientConfig.HttpUtil</td>
<td>SC MD: 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...vc.production.IafAdmCosSvcClient</td>
<td>SC MD: 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Databases</td>
<td>Suppressed: 116</td>
<td>Snoozed Markdowns: 4798</td>
<td>Snoozed Databases: 3</td>
<td></td>
</tr>
<tr>
<td>MD: 13</td>
<td>VCS Status:</td>
<td>Faulted: 0 Partial: 2 Frozen: 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No alerts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| T3 | CORE_excdata01 | MD: 9 | | |
| T3 | CORE_bes0 | MD: 1 | | |
| T1 | ...okup_read-lookup.scluokup04 | MD: 1 | | |
| T2 | CORE_categoryhost18 | MD: 1 | | |
| T2 | CORE_myerbavlookuphost | MD: 1 | | |

Configuration data powered by CMS. Montage UI Version: 15.0
So Full of Fail – Cascading failures – Software
So Full of Fail – Cascading failures – Software - Database

• Connectivity failure
• Query timeouts
• SQL errors
• Data access layer identifies any of the above issues and marks the connection down.
• When the destination is recovered the path is marked active.
• Time to detect the path state is ~15 ms
• Request for connections to a struggling DB can lead to stacking in upstream
• The fail fast waiter queue logic helps the loaded database
So Full of Fail – Cascading failures – Software

<table>
<thead>
<tr>
<th>Configuration</th>
<th>eBay Kernel ServeTraffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configuration</td>
<td>Value</td>
</tr>
<tr>
<td>Alias</td>
<td>ebay.kernel.ecv</td>
</tr>
<tr>
<td>Description</td>
<td>Serve Traffic to this server.</td>
</tr>
<tr>
<td>Group</td>
<td>ebay.kernel</td>
</tr>
<tr>
<td>ID</td>
<td>ebay.kernel.ServeTraffic</td>
</tr>
<tr>
<td>Initializable</td>
<td></td>
</tr>
<tr>
<td>Last Updated</td>
<td>Thu Mar 24 14:36:30 GMT-07:00 2010</td>
</tr>
<tr>
<td>Persistent</td>
<td>true</td>
</tr>
<tr>
<td>Persistent Location</td>
<td>file:/ebay/cronus/software/service_nodes/.ENV85r92web0.viewitem-app._ENV85r92web0.viewitem-app._ENV85r92web0-LVS-CLhyz5fSDo39kD2kg-10.134.108.134/installed-packages/fomcat/7.0.47-13_raptor_taginstace.unx/cronus/karatis/fomcat/webapps/ROOTWEB-INF/classes/appconfig/Production/serviceconfig/config/temp_persist_config_015.xml</td>
</tr>
<tr>
<td>Site Operations Command</td>
<td></td>
</tr>
<tr>
<td>Validator Class Name</td>
<td>com.ebay.kernel.bean.configuration.adapter.ConfigManagementValidatorAdapter</td>
</tr>
</tbody>
</table>

Click the below values to edit

**Configurable Properties**

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>TrafficEnabled</td>
</tr>
</tbody>
</table>
So Full of Fail – Cascading failures – Software

## Command & Control

<table>
<thead>
<tr>
<th>Database Name</th>
<th>Application Service Selection</th>
<th>Application Services</th>
<th>Node Servers</th>
<th>Sitewide (Ignore DBMap)</th>
</tr>
</thead>
<tbody>
<tr>
<td>caty10stby</td>
<td>all application services</td>
<td>Search...</td>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>
So Full of Fail – Cascading failures – Hardware

• LB failures
• Compute failures
• Network failures
So Full of Fail – Cascading failures – Hardware - Network

- Traffic organized in layers.
- Redundant interconnected paths.
So Full of Fail – Cascading failures – Hardware

\[ P(n) = \text{MAX} \left( 17\%, \frac{1}{S_n}, \frac{1}{F_n} \right) \]

- \( n \) = Failure Domain (e.g. Network)
- \( S_n \) = Number of Service Instances in Failure Domain \( n \)
- \( F_n \) = Number of instances of Failure Domain \( n \)
PREVENTION THROUGH PARADIGM

Prevent and remediate atomic failure in order to prevent cascading failure scenarios.
PREVENTION THROUGH PARADIGM

• Site code is deployed on a common set of platforms which enforce the aforementioned tools.
  – Engineering and development architecture effort is “front-loaded” to ensure that all production services can be restarted in exactly the same way no matter what the underlying command.
  – Uniformity of response
  – Remediation at the node level can be automated.
• State based and event based monitoring of key operating metrics.
  – The 9 or so OS-level metrics we all monitor
• Capacity monitoring for “DR’ compliance
  – Two or more co-locations
  – One feature per “pool”
• Provision on demand
• Regional code roll
PARADIGM BEFORE PROCEDURES.

- For maximum rapid remediation your pool or feature should comply with operational paradigms.
- If the care and feeding of your particular beautiful ‘unicorn’ requires special attention or procedures:
  - You’re asking someone to go against their established paradigm and training.
  - At a high rate of speed
  - As the rarely seen or remembered exception
DEALING WITH DISASTER
DIRECTION AND DECISION

Preventing the cascade failure through redirection and deciding which features to keep.
When a lion tamer holds a chair in front of the lion’s face, the lion tries to focus on all four legs of the chair at the same time. With its focus divided, the lion becomes confused and is unsure about what to do next. When faced with so many options, the lion chooses to freeze and wait instead of attacking the man holding the chair.
DEALING WITH DISASTER - DIRECTION

WEB TIER LB

APP TIER LB

APPLICATION SERVERS

REGION A

REGION B
DEALING WITH DISASTER - DIRECTION

Graphic about two lane bridge with one lane broken
DEALING WITH DISASTER - DIRECTION
DEALING WITH DISASTER - DIRECTION
DEALING WITH DISASTER - DIRECTION

• To prevent total failure of the bridge we have to get the kittens to drive over only one lane
• Connection limits would be required
• Blocking bots
• Scale down to minimal serving modes
• Wiring off features like advertising.
CONCLUSIONS
So Full of Fail—Total Failure Avoidance

• Bubble up anomalies for inspection
• Fail Fast
• Mark down the failed paths
• Preserve the user experiences
• Use system approaches to solve