Availability, the Cloud and Everything
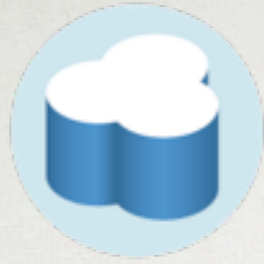
Joe Williams

CLOUDANT

Me

- Joe Williams
  - Infrastructure Engineer
  - Cloudant
  - @williamsjoe
  - joeandmotorboat.com

- Distributed database built on CouchDB
- Real-time Search and Analytics
- Sign Up for our Beta!
- cloudant.com

Bias

- Distributed Databases

- Amazon EC2

- Chef

- Erlang

Availability

Availability

- What is Availability?

Availability

- Uptime / Downtime
  - Perceived
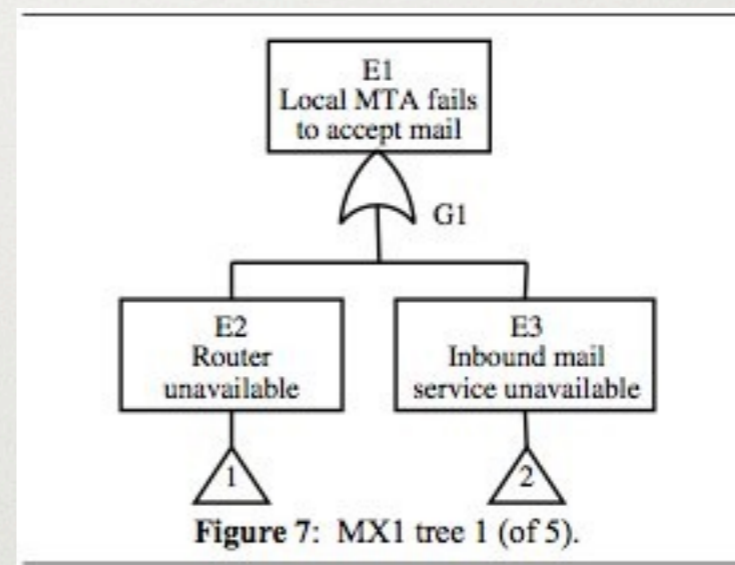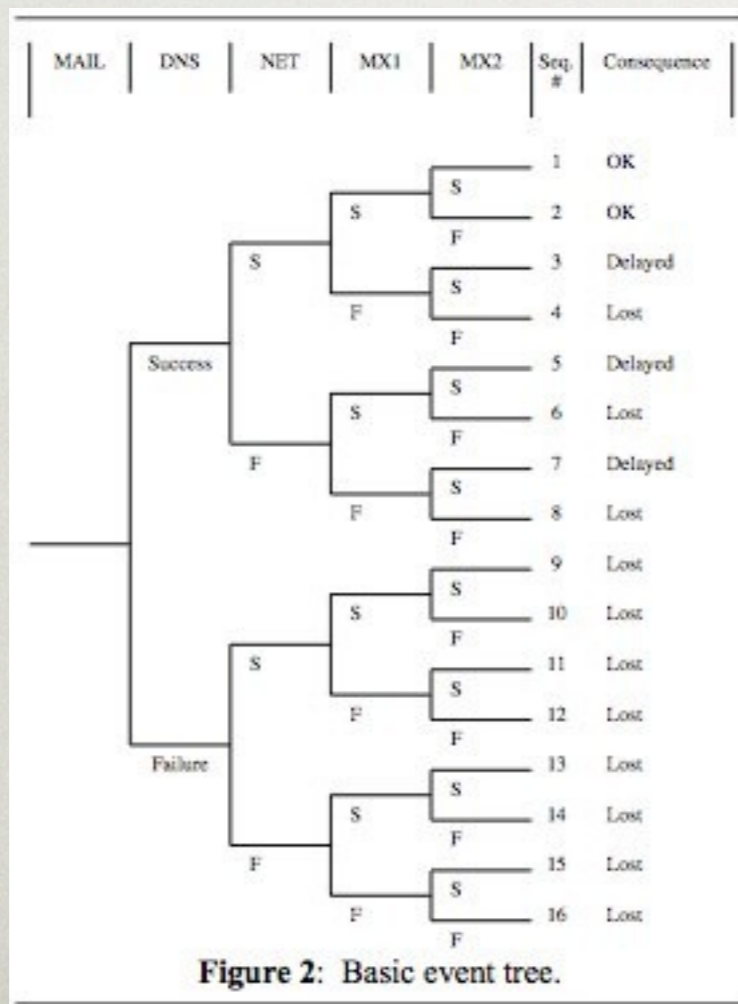 - Actual

Availability

- Reliability

  - Failure rate / Mean time between failures

  - Failure Resistance / Fault Tolerance

  - Durability

Availability

- Probabilistic Risk Assessment

  - Event Tree Analysis

  - Fault Tree Analysis

Availability

- ## Future Work

  - ## Erlang Supervision Trees and FTA / ETA

Figure 2: Basic event tree.

Figure 7: MX1 tree 1 (of 5).

Availability

- Performance, Logging and Monitoring

  - Baracus, Rsyslog, Munin

Availability

Distributed Systems

FIG. I — Centralized, Decentralized and Distributed Networks

Baran (http://www.rand.org/pubs/research_memoranda/RM3420/)

Distributed Systems

* Abstraction

Distributed Systems

- Redundancy
  - Duplication
  - Distribution

Distributed Systems

- RAID ain't as redundant as it used to be.

Leventhal (http://queue.acm.org/detail.cfm?id=1670144)

Distributed Systems

- Alphabet Soup
  - ACID, CAP, BASE, 2PC, MVCC
  - Vector Clocks, Eventual Consistency
  - Dynamo, Paxos, Chandra, Byzantine

Distributed Systems

- CAP == Availability

Distributed Systems

- Effects on Availability

The Cloud

The Cloud

- Abstraction

- Commoditization

The Cloud

- Costs
  - Loss of Control
  - Single Points of Failure
  - Network Partitions / Data Locality
  - Unreliable
  - Performance

The Cloud

- Benefits

  - API to everything

  - Fast and Flexible Resource Mgmt

  - "Unlimited" Resources

The Cloud

Total Bootstrapping Time in Weeks
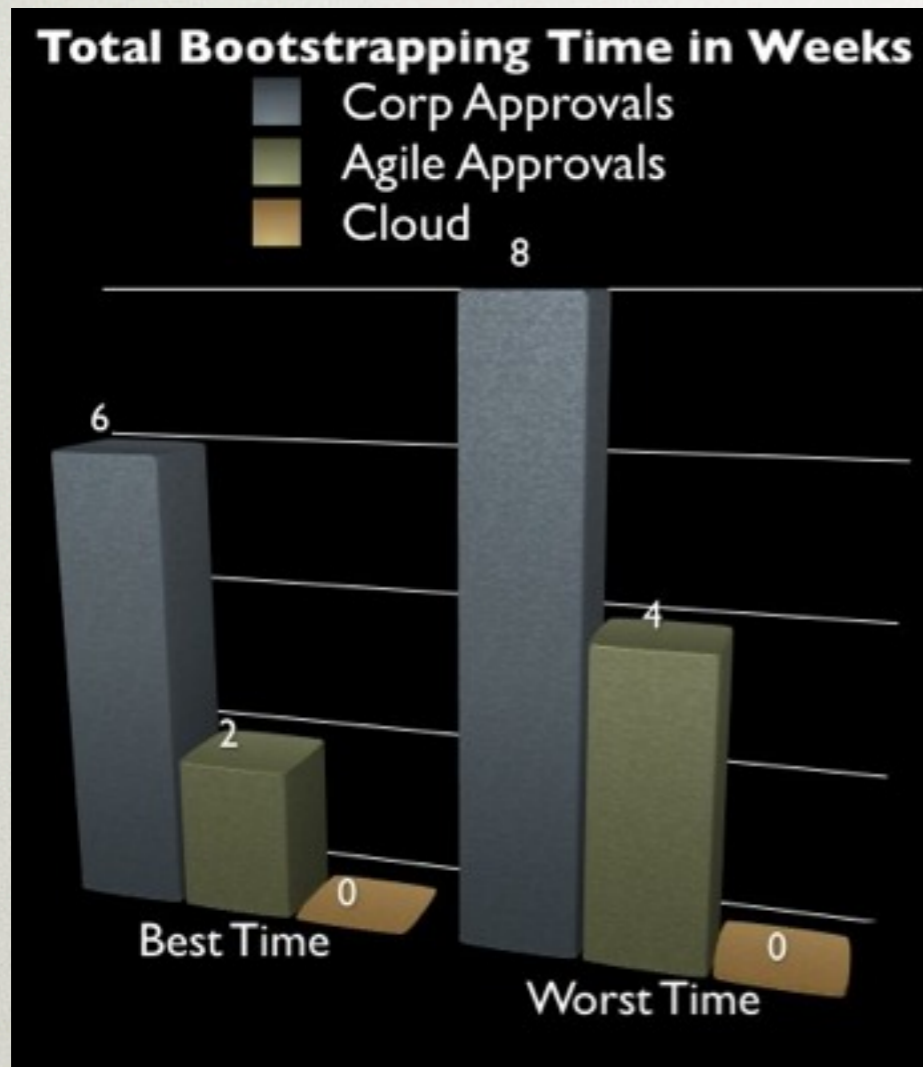
- Bootstrapping
  - Time and Effort

Adam Jacob and Ezra Zygmuntowicz (http://blip.tv/file/2285124/)

The Cloud

- Nodes are stateless and disposable.

The Cloud

- Compute vs Data

- Performance

The Cloud

- Costs ($$) - EC2
  - Different prices in different zones
  - Plus data transfer, EBS, and extras
  - Base Instance Price Range:
    - m1.small ~ $62/mo
    - m2.4xlarge ~ $1750/mo

http://aws.amazon.com/ec2/#pricing

The Cloud

- Effects on Availability

The Cloud

"Clouds are systems ... and with systems, you have to think hard and know how to deal with issues in that environment. The scale is so much bigger, and you don't have the physical control. But we think people should be optimistic about what we can do here. If we are clever about deploying cloud computing with a clear-eyed notion of what the risk models are, maybe we can actually save the economy through technology."

~ Security in the Ether By David Talbot - MIT Technology Review Jan/Feb 2010

Automation and Configuration Mgmt

ETTEETOODE-A

Etteetoode-a is a leettle-a theeng thet mekes a beeg deefffference-a. Bork Bork Bork!

Automation and Configuration Mgmt

- Abstraction

Automation and Configuration Mgmt

- Implementations

  - Chef, Puppet, bcfg2, cfengine

Automation and Configuration Mgm

- Chef
  - Cookbooks/Recipes
  - File Templates
  - Resources/Providers
  - Provides a DSL to every part of your system

Automation and Configuration Mgmt

- Erlang + Chef (v0.8)
  - erl_call Provider

```
erlang.rb #

1   version = "1.0"
2
3   erl_call "erlang hot upgrade" do
4     cookie "erlangcookie"
5     node_name "upgrade@hostname"
6     name_type "sname"
7     code <<-EOH
8     release_handler:unpack_release("yourapp_#{version}.tar.gz"),
9     release_handler:install_release(version),
10    release_handler:make_permanent(version).
11    EOH
12  end
```

Automation and Configuration Mgmt

- Erlang
  - Hot Code Upgrades
  - Distributed Upgrades are HARD

Automation and Configuration Mgmt

- Effects on Availability

Sources and Further Reading

Availability

http://en.wikipedia.org/wiki/Fault_Tree_Analysis
http://en.wikipedia.org/wiki/Probabilistic_risk_assessment
http://www.usenix.org/events/lisa01/tech/apthorpe/apthorpe.ps

The Cloud

http://cseweb.ucsd.edu/~hovav/dist/cloudsec.pdf
http://cloudscaling.com/blog
http://open.eucalyptus.com/wiki/presentations
http://www.morganclaypool.com/doi/pdf/10.2200/
S00193ED1V01Y200905CAC006

Configuration Mgmt

http://wiki.opscode.com/
http://kallistec.com/2010/01/23/the-chef-way/
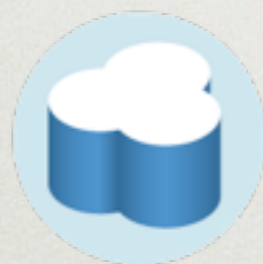http://onlamp.com/pub/a/onlamp/2004/04/15/cfengine.html
http://blog.loftninjas.org/2010/01/22/configuration-management-vs-
meatcloud-5-reasons-cm-wins/
http://cfwiki.org/cfwiki/index.php/Singlecopy_Nirvana

Distributed Systems

http://en.wikipedia.org/wiki/Byzantine_agreement
http://en.wikipedia.org/wiki/Paxos_algorithm
http://en.wikipedia.org/wiki/Multiversion_concurrency_control
http://en.wikipedia.org/wiki/ACID
http://www.julianbrowne.com/article/viewer/brewers-cap-theorem
http://en.wikipedia.org/wiki/Dynamo_(storage_system)
http://en.wikipedia.org/wiki/2PC
http://queue.acm.org/detail.cfm?id=1394128
http://it.toolbox.com/blogs/oracle-guide/acid-vs-base-25938
http://devblog.streamy.com/2009/08/24/cap-theorem/
http://www.infoq.com/presentations/availability-consistency
http://citeseer.ist.psu.edu/652189.html
http://labs.google.com/papers/bigtable-osdi06.pdf
http://database.cs.brown.edu/sigmod09/benchmarks-sigmod09.pdf
http://www.allthingsdistributed.com/2007/10/amazons_dynamo.html
http://www.ctlab.org/documents/How%20Complex%20Systems%20Fail.pdf
http://labs.google.com/papers/mapreduce-osdi04.pdf
http://www.rand.org/pubs/research_memoranda/RM3420/
http://www.cs.jhu.edu/~jak/docs/paxos_for_system_builders.pdf
http://citeseer.ist.psu.edu/652189.html
http://blog.basho.com/2010/01/29/why-vector-clocks-are-easy/

Friday, March 26, 2010

Questions?

Joe Williams - @williamsjoe