



SUMMIT
ONLINE

O P E 0 5

How Amazon's culture affects incident management at AWS

Oren Nachman

Software Development Manager
Amazon Web Services

If you care about
building and operating
reliable and available
applications

👏 clap 👏 your 👏 hands!



“We would not have got to this stage so quickly without support from AWS. The reliability and redundancy of our solution on AWS has ensured almost 100 percent availability.”

Andrew Stevens

Technical Analyst, Qantas

<https://aws.amazon.com/solutions/case-studies/qantas/>

We understand that ...

Your availability depends on

1. Our reliability

2. Your resiliency

Our approach to making you resilient

Availability Zones

- We have shown you our fault domains

Resilient service offerings, e.g. AWS Lambda

- We abstract you from certain faults

Guidance

- Amazon Builders' Library: <https://aws.amazon.com/builders-library/>
- Well-Architected Framework: <https://aws.amazon.com/architecture/well-architected/>

This talk!

Agenda: Let's talk about incident management

Introduction

- Culture: Amazon leadership principles
- Mechanisms

1. Culture: Ownership

- Mechanisms: 2 Pizza teams, Ops meeting / wheel, Tenets

2. Culture: Bias for action

- Mechanism: TechOps, call leaders

3. Culture: Insist on the highest standards, dive deep

- Mechanisms: COE, ORR



Get to know the Leadership Principles at Amazon

Our Leadership Principles aren't just inspirational quotes on the wall. Amazonians use them every day, whether we're discussing ideas for new projects or deciding on the best approach to solving a problem. It is just one of the things that makes Amazon peculiar.

Customer Obsession

Leaders start with the customer and work backwards. They work vigorously to earn and keep customer trust. Although leaders pay attention to competitors, they obsess over customers.

Learn and Be Curious

Leaders are never done learning and always seek to improve themselves. They are curious about new possibilities and act to explore them.

Hire and Develop the Best

Leaders raise the performance bar with every hire. They recognize exceptional talent and do everything to ensure that the

Customer obsession

Leaders start with the customer and work backwards. They work vigorously to earn and keep customer trust. Although leaders pay attention to competitors, they obsess over customers.

... No matter how much
they are paying us.

LOWFLYINGHAWK

Early on, the most vocal and critical AWS customer went by "low-flying-hawk" on the discussion forums. We always looked to see what he thought of new features, pricing, and other issues. After a year of this, someone finally asked: "How much does low-flying-hawk spend with us?" The answer was around \$3 a month.

Treating his feedback as importantly as our biggest spenders showcases our commitment to making the customer experience fantastic for all individual developers and helped us raise our standards to serve the all-important (and viral) technorati well.

Leadership (cultural) principles + Mechanisms

You choose which is which!



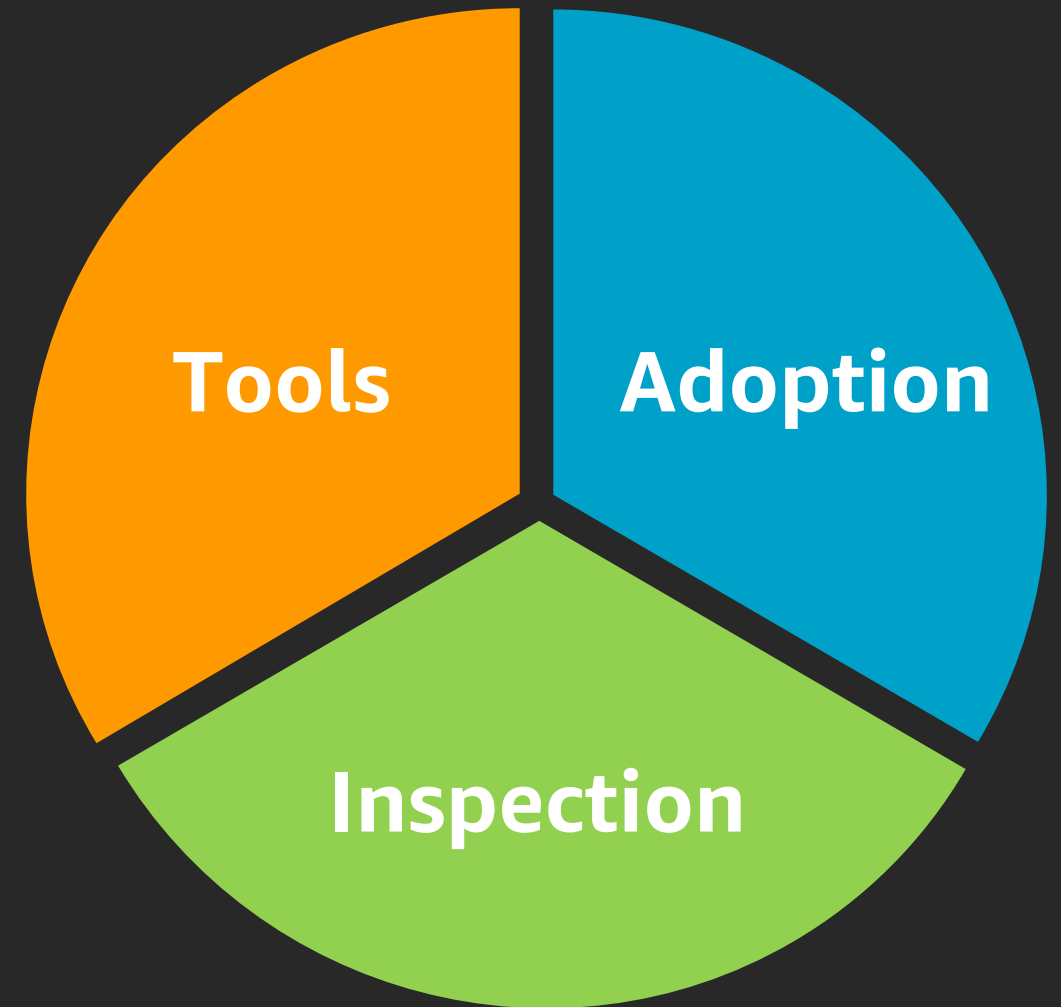
"Good intentions never work,
[...]"

Jeff Bezos
CEO Amazon



"Good intentions never work,
you need **good mechanisms**
to make anything happen."

Jeff Bezos
CEO Amazon





Leadership principles + Mechanisms + People

The journey to a modern, cloud-enabled enterprise is
mostly about transforming the people and how they operate.

“Ownership”

“Giving developers operational responsibilities has greatly enhanced the QUALITY of the services, both from a customer and a technology point of view.

... You build it, you run it.”

Werner Vogels, 2006

CTO, Amazon.com

Ownership

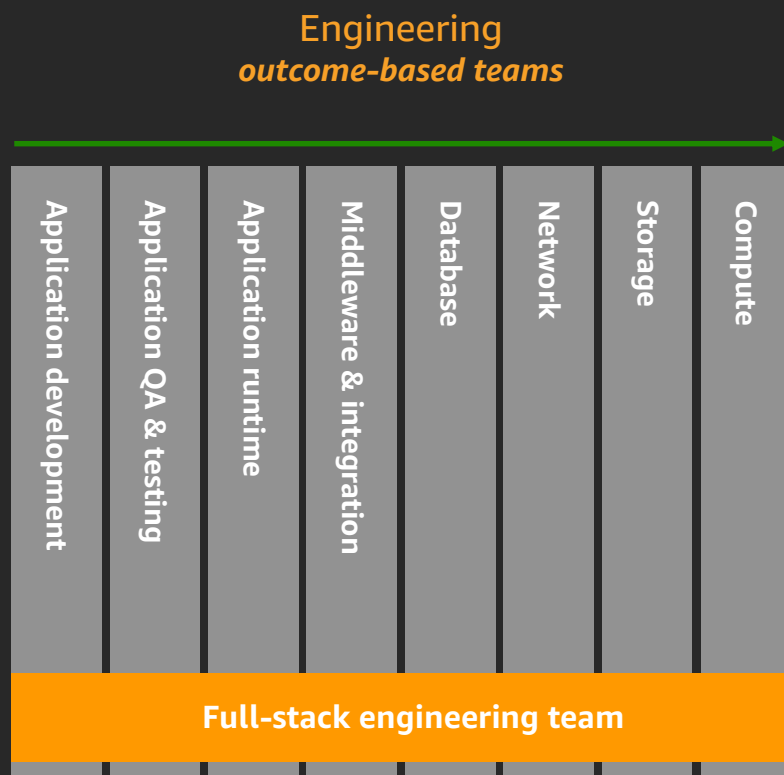
Leaders are owners. They think long term and don't sacrifice long-term value for short-term results. They act on behalf of the entire company, beyond just their own team. They never say, "that's not my job".

Mechanism: "2-Pizza" teams

Anatomy

"Builders"

Customer-focused outcomes



New burdens



Infrastructure



Operations



Security



Finance

Operations

Clear definition of KPIs

First contact for alarms

Automate as much as possible

Safe operations vs. manual/ad hoc

Use existing standard services like CloudWatch, CloudTrail (and new stuff especially)

Ownership

Mechanism: Ops meeting

Weekly operational metrics review

Continuous inspection mechanism

Maintains focus on operations

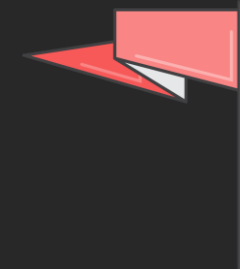
Foundation of a healthy operations program

Accountability & learning, not blaming

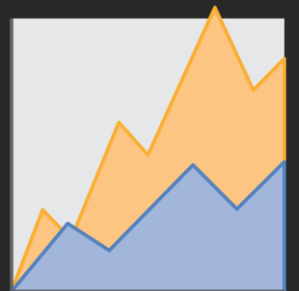
Typical agenda



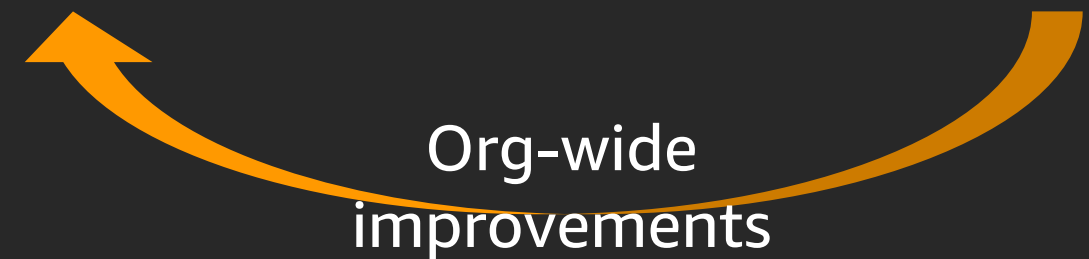
Share and
celebrate
success



Review
operational
events (COE)



Review service
metrics



Org-wide
improvements

Ownership

Spin the wheel!

<https://aws.amazon.com/blogs/opensource/the-wheel/>

The chosen team will:

- Walk through their operational dashboard
- Explain their operational performance
- Answer questions from leaders in the room

"Solutions don't always need to be complicated to be effective."



Ownership

[Quick Segway]

BSS: 🍌 Stand Service

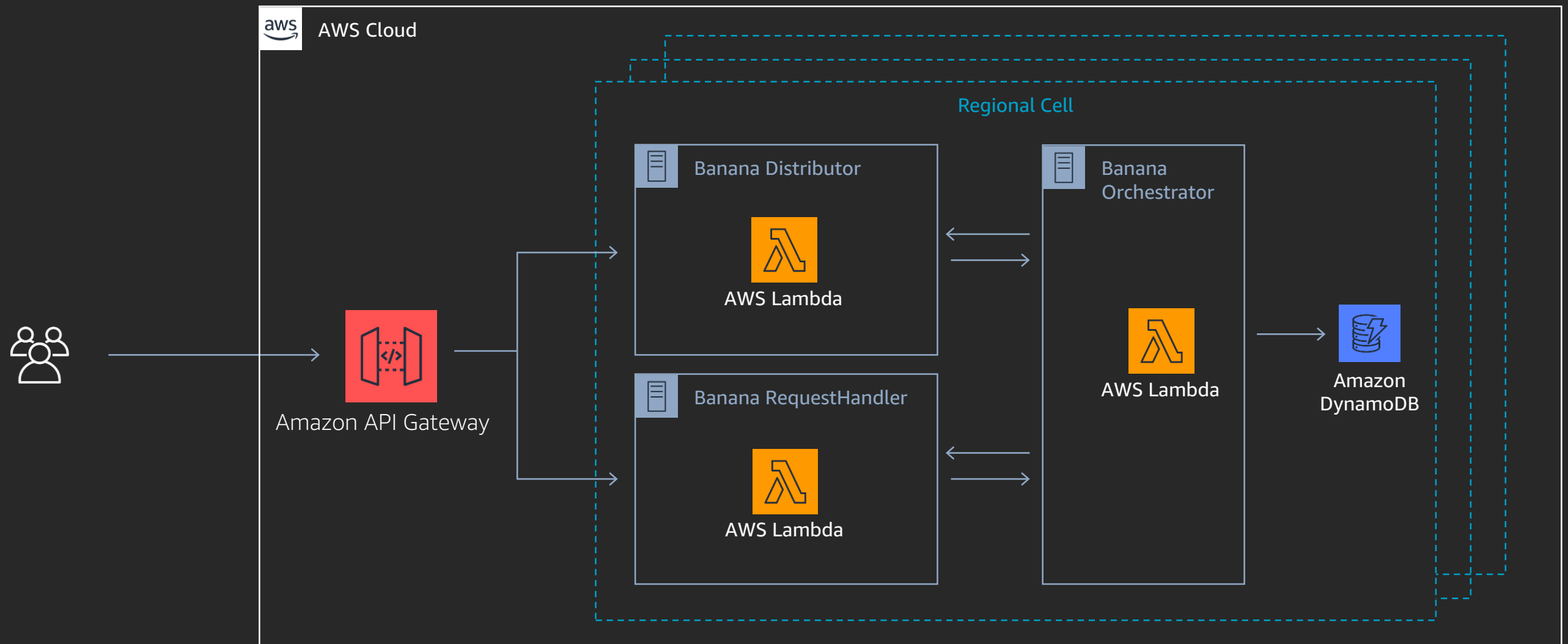
The BSS distributes
free bananas to the
community in Seattle

(it's a real thing, check it out next time
you're in Seattle!)



[Quick Segway] BSS: Banana Stand Service

Mock architecture for context



Mechanism: Tenets

~~Resolve Impact~~
~~Root Cause~~
~~Fixing the Root Cause~~
~~Debrief~~
~~Relationship management~~
~~git blame~~
...

Mitigate Impact

Drive pre-incident training and meetings (Game Days, Ops Meetings etc) to reinforce this idea

Ownership

“Bias for Action”

Mechanism: Tooling to drive immediacy

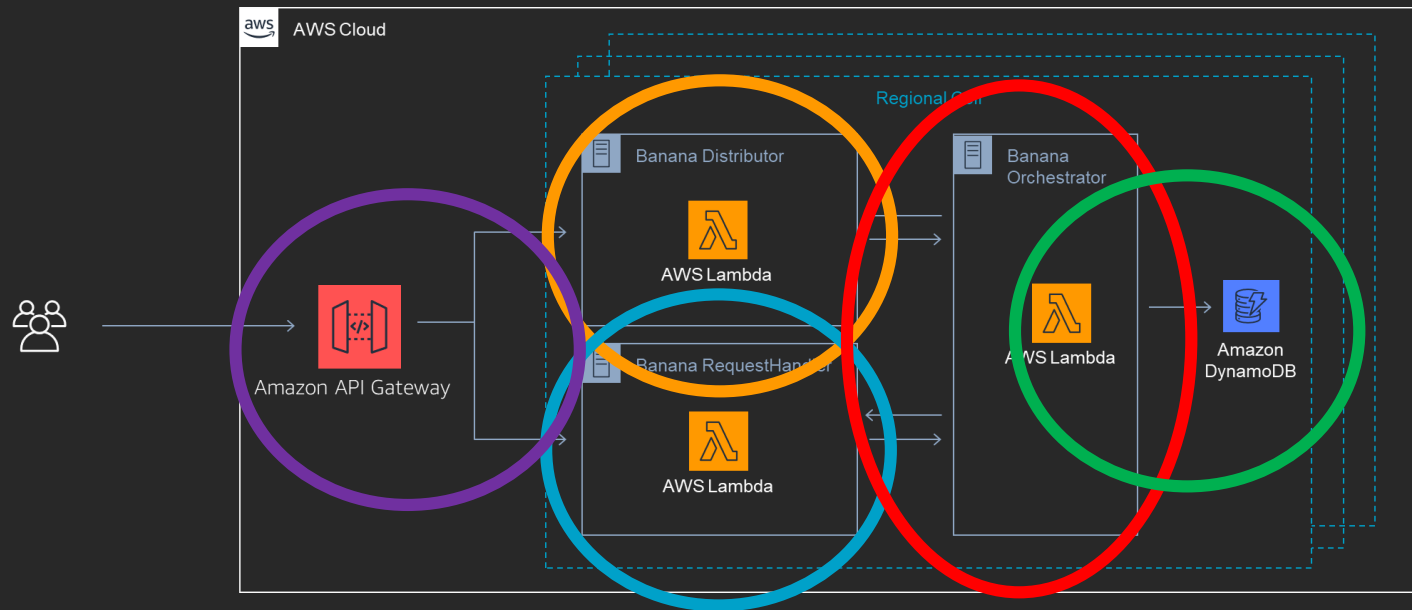
1. Individual Engagement
(Paging, Emails, Texts etc)
2. Global Awareness
(visual cues, chat rooms)



Bias for action

Mechanism: Detection of impact at scale

Use aggregation to detect impact across different, overlapping, spheres



Bias for action

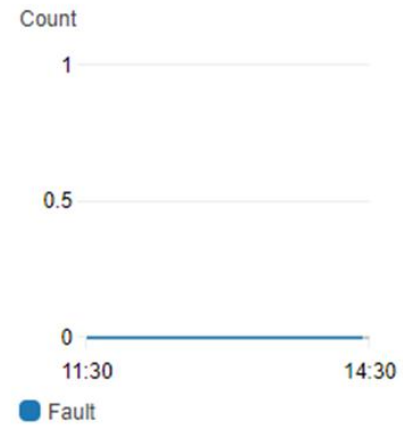
CloudWatch Dashboards

BananaStand Ops Dashboard

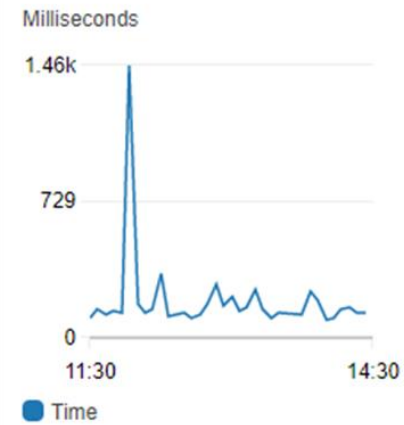
[Sev-2 Runbooks](#) | [Banana Distributor Dashboard](#) | [Banana RequestHandler Dashboard](#) | [Banana Orchestrator](#)

API Gateway

Faults

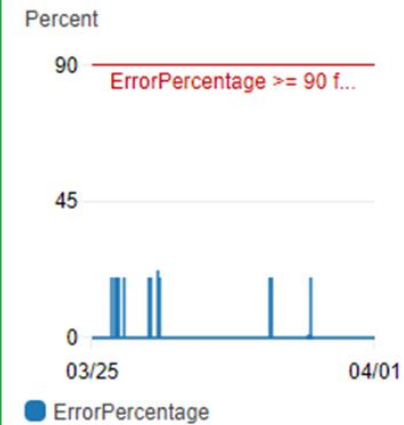


Time (p99)

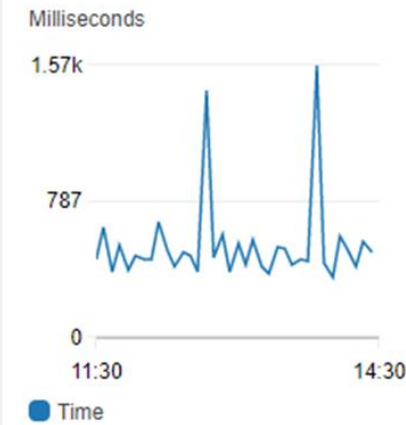


Banana Distributor

Error Percentage

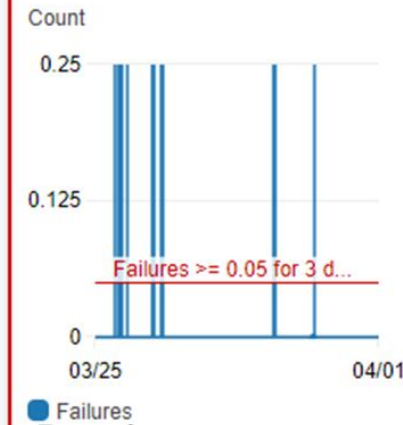


Time (p99)

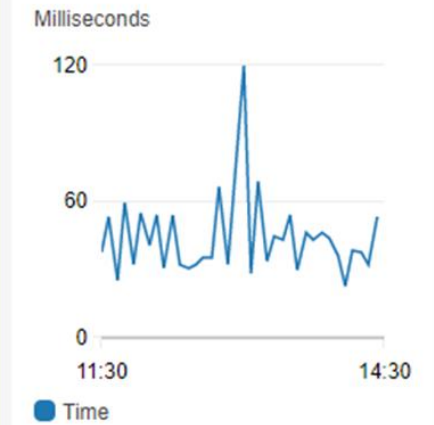


Banana RequestHandler

Faults

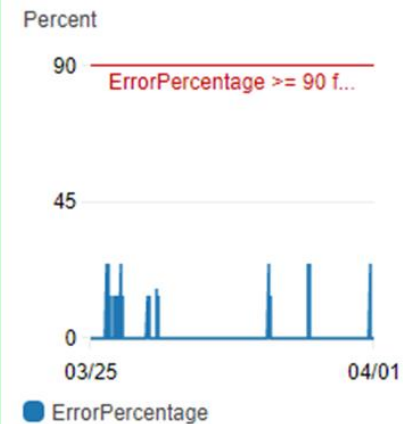


Time (p99)

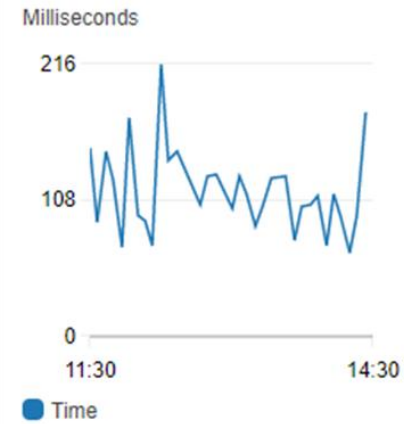


Banana Orchestrator

Error Percentage

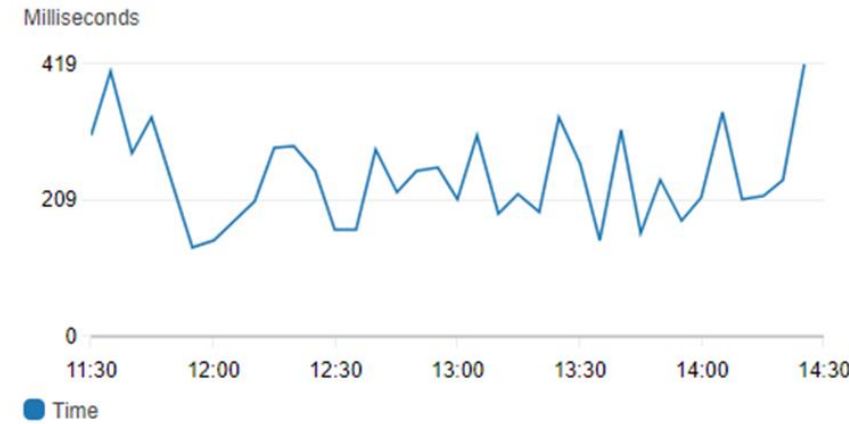


Time (p99)

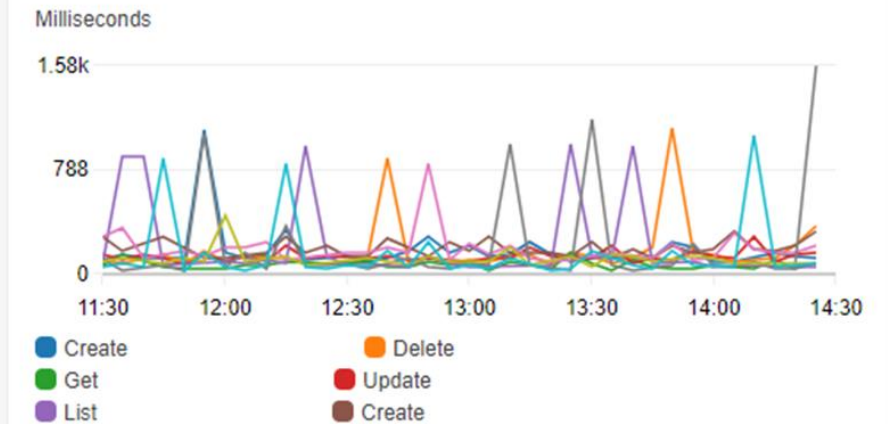


DynamoDB

Time (p99)



Latency (p99)



CloudWatch Composite Alarms

CloudWatch > Alarms > Create composite alarm

Step 1
Specify composite alarm conditions

Step 2
Configure actions

Step 3
Add name and description

Step 4
Preview and create composite alarm

Specify composite alarm conditions

Conditions

Composite alarm conditions
This alarm will go in alarm when the following rule is met. Configure by using AND/OR in the text editor.

Add another alarm ▼

```
1 ALARM("BananaStand-Distributor-Error-Percentage") OR
2 ALARM("BananaStand-RequestHandler-Faults") OR
3 ALARM("BananaStand-Orchestrator-Error-Percentage")
```

CloudWatch
Dashboards
* EmbedTest
Alarms
ALARM 2
INSUFFICIENT 0
OK 2
Billing
Logs
Log groups
Insights
Metrics
Events
Rules
Event Buses

CloudWatch > Alarms

Switch to your original interface

Alarms (4) ☐ Hide Auto Scaling alarms Clear selection ↻ Create composite alarm Actions ▼ Create alarm

Any state ▼ Any type ▼ < 1 > ⚙

<input type="checkbox"/>	Name ▼	State ▼	Last state update ▼
<input type="checkbox"/>	BananaStand-Distributor-Error-Percentage	✔ OK	2020-04-02 02:14:28
<input type="checkbox"/>	BananaStand-RequestHandler-Faults	⚠ In alarm	2020-04-01 01:15:57
<input type="checkbox"/>	BananaStand-Overall-Health	⚠ In alarm	2020-04-01 01:15:57
<input type="checkbox"/>	BananaStand-Orchestrator-Error-Percentage	✔ OK	2020-04-01 01:05:13

Mechanism 3: Centralise handling of large-scale cross team impact

Create a highly technical central team for coordination

During Incident

- Identify aggregate impact
- Engage the correct teams
- Ensure logical flow of investigations
- Incident mitigation pulse checkers

Peace Time

- Use KPIs to accelerate cross-team response
- Improve central tooling

Bias for action

“Insist on the Highest Standards”
“Learn and Be Curious”

“Insist on the Highest Standards”

Leaders have relentlessly high standards - many people may think these standards are unreasonably high. Leaders are continually raising the bar and drive their teams to deliver high quality products, services, and processes. Leaders ensure that defects do not get sent down the line and that problems are fixed so they stay fixed.

Mechanism: "Call Leader"s

"Lead, follow or get the
hell out of the way"

- General George S. Patton

Mechanism: “Call Leader”s

Decentralised non-SME Incident Mitigation SMEs

Responsibilities

- Voice of the customer
- Ensure correct engagements and escalations
- Gate production changes
- Event leadership & lifecycle

Keys to success

- Keep it small (AWS has < 15 Call Leaders)
- Always maintain a list of shadows



“Earn Trust”

Earn trust

Leaders listen attentively, speak candidly, and treat others respectfully. They are vocally self-critical, even when doing so is awkward or embarrassing. Leaders do not believe their or their team's body odor smells of perfume. They benchmark themselves and their teams against the best.

Mechanism: Correction of Errors (COE)

Mechanism to **learn from our mistakes, resolve root cause** and **drive continuous improvement**.

Areas of focus:

- Technical
- Process
- Documentation
- Organisational
- Everything else!

The key to being successful in this process is being **open and transparent about what went wrong**.

Earn Trust

Anatomy of a COE

What happened?

- BananaStand was unavailable for 90 minutes due to a misconfigured API Gateway

What was the impact on customers and your business?

- Customers were unable to collect their bananas for the duration of the impact

What were the contributing factors?

- Don't stop at operators

What data do you have to support this?

- Metrics and graphs

What lessons did you learn?

- Need better e2e testing that encompasses all of BananaStand's infrastructure

What corrective actions are you taking?

- Actions items
- Related items

Insist on Highest Standards

Mechanism: Follow up action items

- Short term items are completed **before the event is resolved**
- Medium term action items are tracked and reported on weekly
 - Avoid “**Good Intentions**”
- Long term actions should be funded and reported on monthly / quarterly
- Drive action items and learnings across organisations
 - Avoid “**Chernobyl**”: Blameless encourages others to adopt learnings



Call To Action



Earn the trust of your customers

- Empower your engineers
 - Provide them with the training and tools they need to mitigate customer impact
-

Use mechanisms

- Avoids “good intentions”
 - Trust but verify
-

Suggested Mechanisms for your consideration

- **BAU:** OpsMeetings, Game Days
- **Incident:** Focus on mitigation, TOS, call leaders, centralise data with decentralised resolvers
- **Post-incident:** Blameless COE, tracked follow ups, problem management

Oren
(just in case!)

This is
Dave!

The actual Banana Stand
experiencing:
`BananaDistributionLimitExceeded`



Thank you!

Oren Nachman

Software Development Manager
Amazon Web Services