Managing the Pace of Innovation

Behind the Scenes of AWS

Constantin Gonzalez
Principal Solutions Architect, AWS Germany
“Amazon is a technology company that happens to sell books online”

– Dr. Werner Vogels, CTO Amazon.com
1 Million Active Customers

11 Regions
28 Availability Zones
53 Edge Locations
### Service Breadth & Depth

<table>
<thead>
<tr>
<th>Support</th>
<th>Professional Services</th>
<th>Platform Ecosystem</th>
<th>Training &amp; Certification</th>
<th>Solutions Architects</th>
<th>Account Management</th>
<th>Security &amp; Pricing Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual Desktop</td>
<td>Sharing &amp; Collaboration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analytics</td>
<td>App Services</td>
<td>Developer Tools &amp; Operations</td>
<td>Mobile Services</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hadoop</td>
<td>Queuing &amp; Notifications</td>
<td>Transcoding</td>
<td>Deployment</td>
<td>Resource Templates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real-time Streaming Data</td>
<td>Workflow</td>
<td>Email</td>
<td>DevOps</td>
<td>Containers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Warehouse</td>
<td>App Streaming</td>
<td>Search</td>
<td>Application Lifecycle Management</td>
<td>Event-driven Computing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Pipelines</td>
<td>Data Pipeline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identity Management</td>
<td>Access Control</td>
<td>Resource &amp; Usage Auditing</td>
<td>Key Management &amp; Storage</td>
<td>Monitoring &amp; Logs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compute (VMs, Auto-scaling &amp; Load Balancing)</td>
<td>Storage (Object, Block and Archival)</td>
<td>CDN</td>
<td>Databases (Relational, NoSQL, Caching)</td>
<td>Networking (VPC, DX, DNS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regions</td>
<td>Availability Zones</td>
<td></td>
<td>Points of Presence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Technical & Business Support
- **Network Management**
- **Business Email**

#### Enterprise Applications
- **Identity Management**
- **Sync**
- **Mobile Analytics**
- **Push Notifications**

#### Platform Services
- **Identity**
- **Sync**
- **Mobile Analytics**
- **Push Notifications**

#### Administration & Security
- **Monitoring & Logs**
- **Networking (VPC, DX, DNS)**
- **Points of Presence**

#### Core Services
- **Network Management**
- **Business Email**

#### Infrastructure
- **Account Management**
- **Security & Pricing Reports**

#### Analytics
- **Hadoop**
- **Real-time Streaming Data**
- **Data Warehouse**
- **Data Pipelines**

#### App Services
- **Queuing & Notifications**
- **Workflow**
- **App Streaming**
- **Email**
- **Search**

#### Developer Tools & Operations
- **Deployment**
- **DevOps**
- **Application Lifecycle Management**
- **Event-driven Computing**

#### Mobile Services
- **Identity**
- **Sync**
- **Mobile Analytics**
- **Push Notifications**

#### Infrastructure Services
- **Compute (VMs, Auto-scaling & Load Balancing)**
- **Storage (Object, Block and Archival)**
- **CDN**
- **Network Management**
- **Points of Presence**
How?
People
Amazon Leadership Principles

Whether you are an individual contributor or the manager of a large team, you are an Amazon leader. These are our leadership principles and every Amazonian is guided by these principles.

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.

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."

Invent and Simplify
Leaders expect and require innovation and invention from their teams and always find ways to simplify. They are externally aware, look for new ideas from everywhere, and are not limited by "not invented here." As we do new things, we accept that we may be misunderstood for long periods of time.

Are Right, A Lot
Leaders are right a lot. They have strong business judgment and good instincts.

Hire and Develop the Best
Leaders raise the performance bar with every hire and promotion. They recognize exceptional talent, and willingly move them throughout the organization. Leaders develop leaders and take seriously their role in coaching others.

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 driving their teams to deliver high quality products, services and processes. Leaders ensure that defects do not get sent down the
Amazon Leadership Principles

- Customer Obsession
- Think Big
- Ownership
- Invent and Simplify
- Dive Deep
- Are Right, A Lot
- Deliver Results
- Bias for Action
- Vocally Self Critical
- Frugality
- Earn Trust of Others
- Hire and Develop the Best
Bias for Action

Insist on the Highest Standards

Customer Obsession

Think Big

Invent and Simplify
Ownership
Deliver Results

- Hire and Develop the Best
- Frugality
- Vocal Self-Criticism
- Have Backbone; Disagree and Commit
- Earn Trust of Others
- Dive Deep
- Backing Commit
- Frugality
Builder
Mechanisms
If you want something done right …

create a

single-threaded
team
Two pizza teams
Fitness Functions
Narratives
Writing a narrative helps you make **best use of time of everyone** at the meeting.
The process of **writing your ideas** helps you refine your thoughts and articulate them effectively, while exposing **key gaps** that you can refine before the meeting.
Presentations are not the best medium for consumption of highly analytical information.
Battle station is heavily shielded

- Fire power greater than ½ Star Fleet.
- Defenses are designed around direct large-scale assaults.
- Small one-man fighter should be able to penetrate the outer defense.

Empire doesn't use small one-man fighters to be effective or they'd have a fighter defense.
At many points during its investigation, the Board was surprised to receive similar presentation slides from NASA officials in place of technical reports. The Board views the endemic use of PowerPoint briefing slides instead of technical papers as an illustration of the problematic methods of technical communication at NASA.

The vaguely quantitative words "significant" and "significantly" are used 5 times on this slide, with de facto meanings ranging from "detectable in largely irrelevant calibration case study" to "an amount of damage so that everyone dies" to "a difference of 640-fold." None of these 5 usages appears to refer to the technical meaning of "statistical significance."

The low resolution of PowerPoint slides promotes the use of compressed phrases like "Tile Penetration." As is the case here, such phrases may well be ambiguous. (The low resolution and large font generate 3 typographic orphans, lonely words dangling on a separate line.)

This vague pronoun reference "it" alludes to damage to the protective tiles, which caused the destruction of the Columbia. The slide weakens important material with ambiguous language (sentence fragments, passive voice, multiple meanings of "significant"). The 3 reports were created by engineers for high-level NASA officials who were deciding whether the threat of wing damage required further investigation before the Columbia attempted return. The officials were satisfied that the reports indicated that the Columbia was not in danger, and no attempts to further examine the threat were made. The slides were part of an oral presentation and also were circulated as e-mail attachments.

In this slide the same unit of measure for volume (cubic inches) is shown a different way every time. 3 cu. in. vs. 1920 cu. in. is rather than in clear and tidy exponential form 1920 in³. Perhaps the available font cannot show exponents. Shakiness in units of measurement provokes concern. Slides that use hierarchical bullet-outlines here do not handle statistical data and scientific notation gracefully. If PowerPoint is a corporate mandated format for all engineering reports, then some competent scientific typography (rather than the PP market-pitch style) is essential. In this slide, the typography is so choppy and clumsy that it impedes understanding.

The analysis by Dr. Edward Tufte of the slide from the Debris Assessment Team briefing. [SOFI=Spray-On Foam Insulation]
Slides have **choppy transitions** that make it very difficult to share a continuous stream of thought.
Slides are open for interpretation, and the same slides can be used to present completely different stories.
Most interesting details are often hidden in **sub-sub-sub-bullets**.
At Amazon, we always work backwards from the customer.
1. Press Release
2. FAQ
3. Build
Primitives
Primitives behind hardened APIs
Primitives behind hardened APIs help us innovate independently
Once we have the right primitives, we ask ourselves:

“Can we simplify?”
We eat our own dog food.
We eat our own dog food. This puts ourselves in the shoes of our customers.
We eat our own dog food. This puts ourselves in the shoes of our customers, and compels us to be vocally self-critical to innovate on behalf of our customers.
Consuming our own APIs allows us to build **primitives** on top of primitives.
Amazon S3 for highly durable storage.
Amazon EC2 for computing.
Amazon EBS for block storage.
Amazon Route 53 health checks as a monitoring and failover primitive.
EC2 + EBS + S3 + Route 53 = Amazon RDS
When we added **PIOPs** to EBS, they were just **inherited** by RDS.
Complete the innovation loop with customers.
Metrics
Weekly ops metrics meeting.
A scorecard for each service team.
A graph for every **metric** that **customers** care about.
Each graph has a line.
Any metric going beyond the line is considered a breach worthy of correction.
Correction of Error (COE) process
Five Whyss
DevOps
Continuous Integration
Continuous Integration

Small, frequent changes, constantly integrating into production.
## Deployments at amazon.com

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean time between deployments (weekday)</td>
<td>11.6s</td>
</tr>
<tr>
<td>Max number of deployments in a single hour</td>
<td>1,079</td>
</tr>
<tr>
<td>Mean number of hosts simultaneously receiving a deployment</td>
<td>10,000</td>
</tr>
<tr>
<td>Max number of hosts simultaneously receiving a deployment</td>
<td>30,000</td>
</tr>
</tbody>
</table>
Canary testing and phased rollouts
Automated rollbacks and health checks
Summary

- Culture
- Single-threaded two pizza teams
- Narratives
- Start with the customer
- Primitives behind hardened APIs
- Metrics, COEs, five whys
- Continuous integration
Recommended Reading

**Building Microservices**

Sam Newman

**The Lean Enterprise**

Trevor Owens & Obi Fernandez
Thank You

Constantin Gonzalez
glez@amazon.de
@zalez