AN INFLUXDATA CASE STUDY

Turning Cloud Metric into Results with InfluxData

Rob Frohnapfel

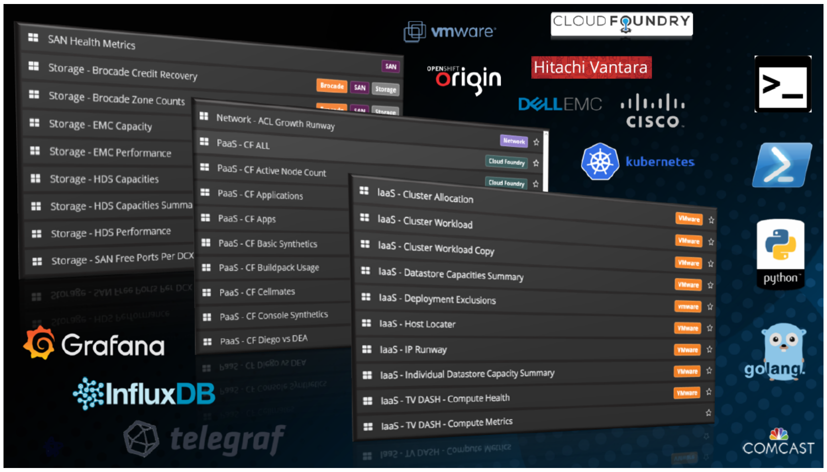
Director of Product Engineering, Comcast

FEBRUARY 2018

Company in brief

Comcast Corporation (Nasdaq: CMCSA) is a global media and technology company with two primary businesses, Comcast Cable and NBCUniversal. Comcast Cable is one of the nation's largest video, high-speed Internet and phone providers to residential customers under the XFINITY brand and also provides these services to businesses. NBCUniversal operates news, entertainment and sports cable networks, the NBC and Telemundo broadcast networks, television production operations, television station groups, Universal Pictures and Universal Parks and Resorts.

Comcast is shaping the future at the intersection of media and technology. It is a leader in the worlds of media, entertainment and technology and its story has unfolded due to an entrepreneurial spirit that is the foundation of everything it does.


SAN Health Metrics
Storage - Brocade Credit Recovery
SAN
Brocade
SAN
Storage
Storage - Brocade Zone Counts
Storage - EMC Capacity
Storage - EMC Performance
Storage - HDS Capacities
Storage - HDS Capacities Summa
Storage - HDS Performance
Storage - SAN Free Ports Per DCX
Network - ACL Growth Runway
Network
PaaS - CF ALL
Cloud Foundry
PaaS - CF Active Node Count
Cloud Foundry
PaaS - CF Applications
PaaS - CF Apps
PaaS - CF Basic Synthetics
PaaS - CF Buildpack Usage
PaaS - CF Cellmates
PaaS - CF Console Synthetics
PaaS - CF Diego vs DEA
IaaS - Cluster Allocation
IaaS - Cluster Workload
IaaS - Cluster Workload Copy
IaaS - Datastore Capacities Summary
IaaS - Deployment Exclusions
IaaS - Host Locater
IaaS - IP Runway
IaaS - Individual Datastore Capacity Summary
IaaS - TV DASH - Compute Health
IaaS - TV DASH - Compute Metrics
VMware
vmware
CLOUD FOUNDRY
OPENSHIFT origin
Hitachi Vantara
DELL EMC
CISCO
kubernetes
python
golang
Grafana
InfluxDB
telegraf
COMCAST

Monitoring strategy providing full visibility into Comcast’s environment.

Case overview

Comcast wanted to develop a monitoring strategy for its Cloud Foundry and Kubernetes architecture and provide metrics around those stacks. Their internal teams today demand metrics-driven visibility into various aspects of platform performance. Comcast is using InfluxDB, Kapacitor and Telegraf to capture and report on real-time metrics and to generate impactful, tangible business results. Comcast applied their multi-tenant cloud metrics gathered with InfluxData to identify huge opportunities, partner with the business, and drive outcomes that include reliable cloud performance and the return of real dollars to the business, not to mention happy customers. Taking the time series data collected, aggregated, and stored using open source InfluxData tools, Comcast was able to deliver transformative results.

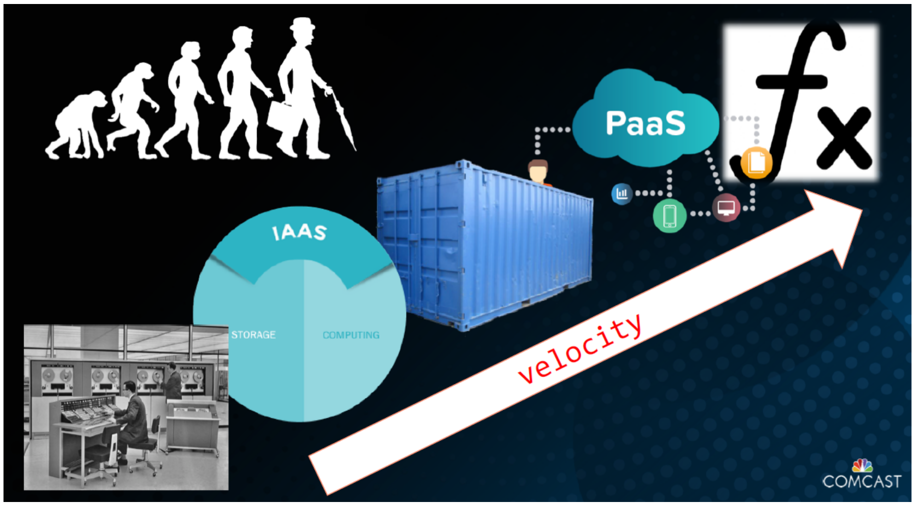
> *“Metrics equate to balance, and balance equates to efficiency.”*
>
> ***Rob Frohnapfel****, director*

The business problem

Comcast knew they had to transform how they deliver and manage infrastructure to their developers in order to help them bring their products to market faster. To do so, they needed to remove the infrastructure heavy lifting for their developers, partner with their dev and product teams (their customers) and enable them a frictionless experience to deploy their applications. In support of this, they also needed to develop a monitoring strategy for its internal IaaS and PaaS that provide metrics around those stacks as well as provide their developers with metrics on how their applications are consuming (and wasting) these resources.

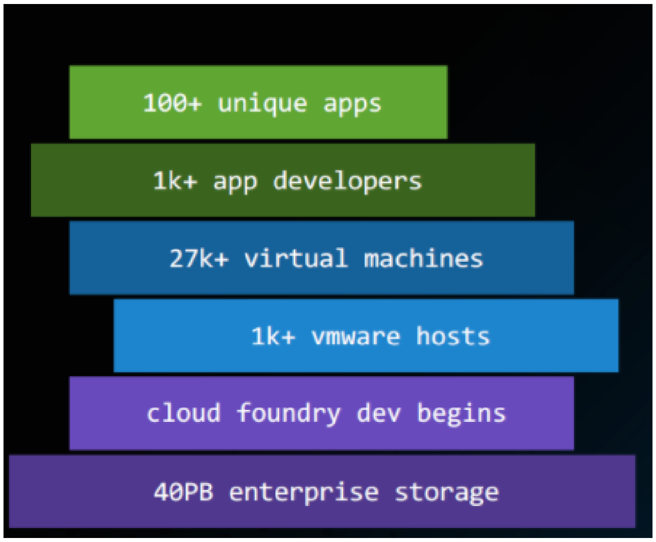
The team responsible for application platform delivery for Comcast product and development teams has been shaped by the industry, technology evolution, and the requirements of their product and dev teams. About 15 years ago, many team members were working on physical server hardware and were Microsoft and Linux administrators at Comcast. Over time, they moved into virtualization using VMware, ultimately delivering self-service for VMware. That was the first iteration of Infrastructure as a Service (IaaS) at Comcast, resulting in a sizeable VMware infrastructure.

As they evolved through this progression of platforms from the ground to the clouds, they realized that developer velocity increases significantly and that the ability to run in a multi-cloud configuration and move your workloads where you need them, whether public or private cloud, is a gamechanger for their development teams.



In 2014, their focus shifted to Cloud Foundry PaaS (Pivotal & Open Source CF), Kubernetes, and other "above the value line" services for Comcast. Comcast's platforms host thousands of unique applications, tens of thousands of virtual machines and containers, dozens of petabytes of storage and an expanding portfolio of homegrown telemetry and automation. To their internal customers, they are a direct competitor to the top public cloud providers.

Comcast's Challenges in 2014



The team's business goals were:

- Transform their mindset — how they deliver and manage infrastructure and how their developers bring their products to market
- Remove infrastructure heavy lifting for their developers, partner with their dev and product teams (their customers) and enable them a frictionless experience to deploy their applications
- Consolidate their tooling
- Provide transparency into their environment's performance
- Drive efficiencies and identify opportunities
- Provide application metrics for all developers that deploy on their infrastructure

The technical problem

To meet their business goals and provide visibility into their environment, Comcast needed a new monitoring strategy suited to their modern cloud and container-based environment. Cloud Foundry — given its multi-cloud portability, autohealing, and autoscaling — has had an exponential adoption within

Comcast’s product and dev teams. They were also working on container orchestration as they viewed containers and Kubernetes as a means to fill the gap between IaaS and PaaS.

Comcast faced a handful of challenges with monitoring and telemetry across modern platforms:
Such platforms had very ephemeral workloads. Whereas old monitoring solutions relied on static infrastructure components, containers and functions constantly spin up and down and are elastic by nature. Sporadic workloads made it very difficult to use a traditional pull-based approach within a monitoring solution. Data also needed to be collected from a variety of systems and a simple interface was needed to push that data in. The team’s approach shifted from the traditional pull approach associated with older traditional monitoring solutions towards a push approach. That's when they started to refine their monitoring and telemetry strategy and found InfluxDB and Kapacitor to be the perfect centerpieces for that strategy.

The solution

Why InfluxDB?

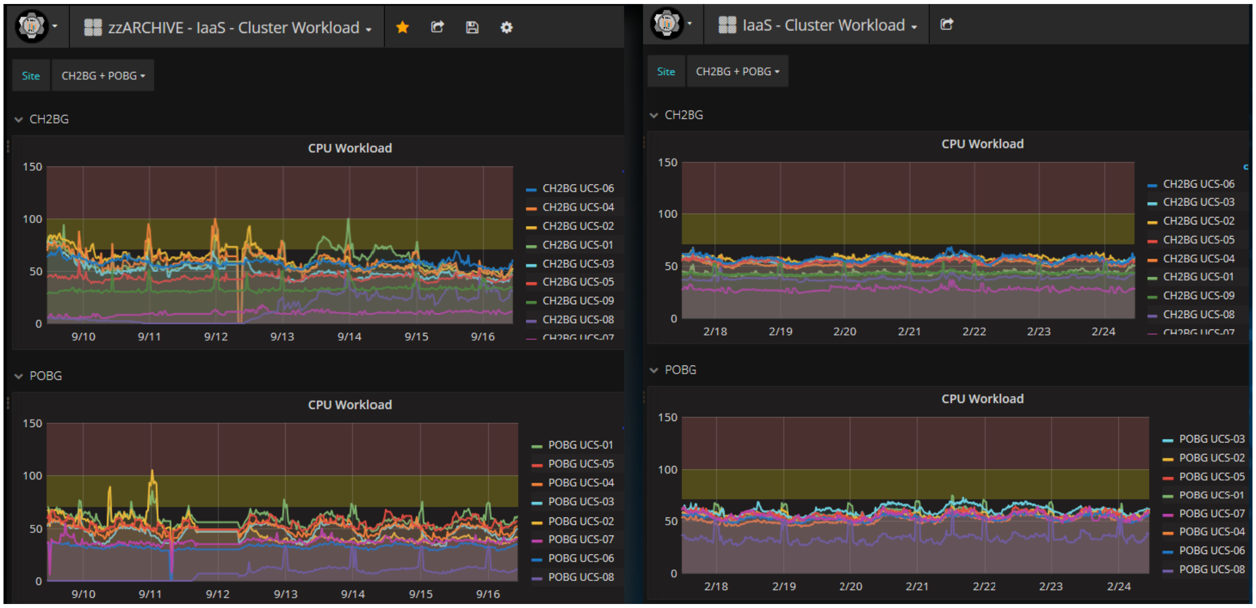
While searching for a way to monitor Cloud Foundry and provide metrics around the Cloud Foundry stack, the Comcast team discovered InfluxDB and Grafana and soon realized they could use InfluxDB to monitor everything they support — VMware, storage, and cloud platforms. They focused on the idea of creating a product out of all of those metrics that they could use to meet their business goals. The team’s traditional sysadmins, who were typically Windows or Linux administrators, started dabbling in Python and Golang and learning DevOps tools such as Concourse and Ansible, and it wasn't long before they started delivering tangible results. The team began to use InfluxDB, Telegraf, and Grafana — and were thereby able to target exactly what they were looking for to improve the following process:

- **Capacity management:** This was a manual and reactive process involving logging into different systems, pulling data into spreadsheets, aggregating it, and using that as a single-point-in-time snapshot for use that day. The process had to be repeated anytime they wanted to buy something or forecast a budget. With data pulled into InfluxDB, capacity management became a simple process done in minutes.

- **Performance management:** They were not practicing performance management due to lack of data evidence and the time and effort required to collect and pull together the needed data. Using InfluxDB, they gained a very quick snapshot at a very high level of aggregated performance across their environment.

- **VMware-based performance balancing:** Such balancing required VMware teams to know what their environments were doing. For example, aggregating the data using InfluxData provided evidence of a problem impacting customer experience — that they had suspected but never had the data to prove. Once visualization showed them that clusters were spiking at 100% CPU utilization for sustained periods of time, and once they solved that issue, they began to make data-driven decisions on how to balance workload across clusters. This resulted in greater customer satisfaction, proactive balancing, and the advantage of cost deferral, whereby their old methodology was to operate in a reactive state and repeatedly buy VMware equipment.

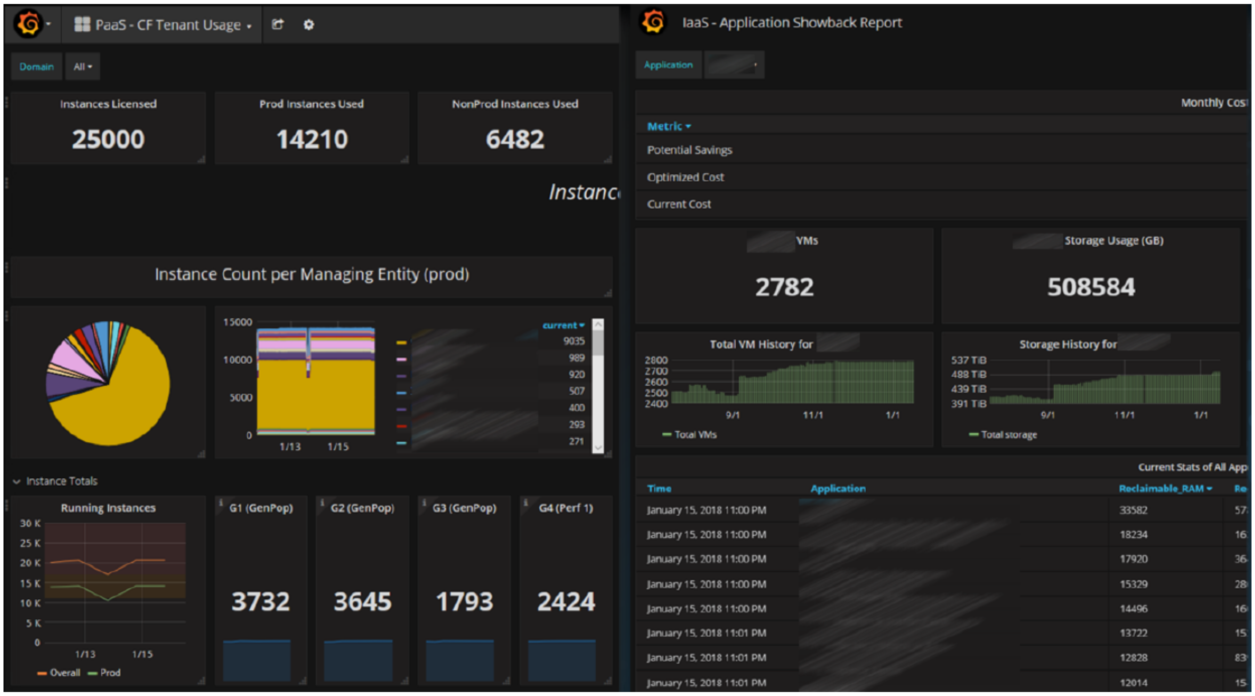
Gaining Data Evidence to Balance Workloads Across Clusters



- **Insight into storage balancing:** Upon learning from the metrics what their storage was doing, the team found two distinct use cases for storage balancing:
    1. To the storage team, the storage arrays seemed full or at predefined capacity threshold, when in reality, there was a lot of stranded storage within VMware data stores. By balancing storage throughout the VMware data stores, they identified storage that they could reclaim.
    2. There was a severe lack of balance across storage arrays themselves. Once the team visualized that, they identified stranded capacity and performance issues. They wrote some automation and established a process to help balance storage arrays so the provisioning storage teams could follow best practices and provision more efficiently.

- **Multi-tenant reporting:** They did not have visibility into usage, waste levels, and cost of multi-tenant private cloud platforms within their infrastructure. By leveraging time series data, they could visualize who is using what and how much they're using. This reporting enabled them to reference the data evidence to demonstrate usage and accelerate their ability to procure capacity. In addition, publishing waste information and cost information was inspirational to the teams consuming it.

Monitoring Cloud Foundry Tenant Usage and VMware-Based Metrics



PaaS - CF Tenant Usage

Domain: All

Instances Licensed: 25000

Prod Instances Used: 14210

NonProd Instances Used: 6482

Instance Count per Managing Entity (prod)

Instance Totals

Running Instances

G1 (GenPop): 3732

G2 (GenPop): 3645

G3 (GenPop): 1793

G4 (Perf 1): 2424

IaaS - Application Showback Report

Application: [redacted]

Monthly Cost

Metric

Potential Savings

Optimized Cost

Current Cost

[redacted] VMs: 2782

[redacted] Storage Usage (GB): 508584

Total VM History for [redacted]

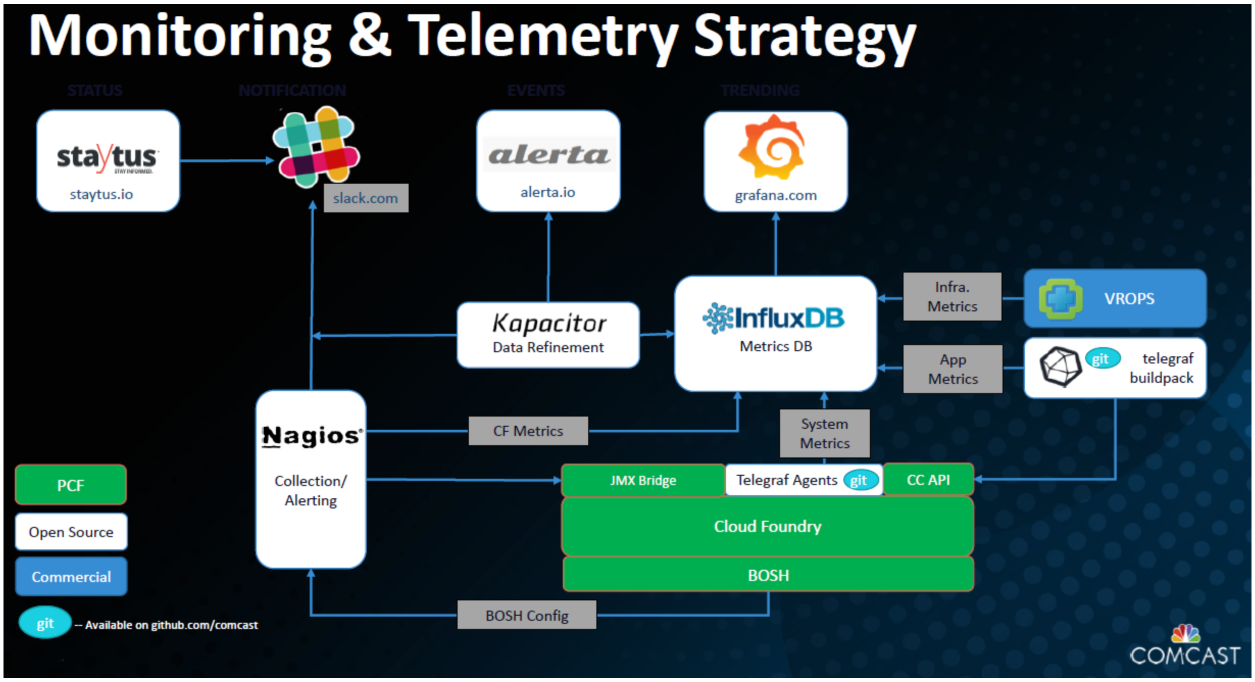
Storage History for [redacted]

Current Stats of All App

Time	Application	Reclaimable_RAM	Re
January 15, 2018 11:00 PM	[redacted]	33582	57
January 15, 2018 11:00 PM	[redacted]	18234	16
January 15, 2018 11:00 PM	[redacted]	17920	36
January 15, 2018 11:00 PM	[redacted]	15329	28
January 15, 2018 11:00 PM	[redacted]	14496	16
January 15, 2018 11:01 PM	[redacted]	13722	15
January 15, 2018 11:01 PM	[redacted]	12828	83
January 15, 2018 11:01 PM	[redacted]	12014	15

Technical architecture

> *"We've got InfluxDB and Kapacitor at the heart of everything. Kapacitor is enriching our data and helping us process alerting. And then we've got a ton of feeds from all different directions pushing data into that solution."*



Monitoring & Telemetry Strategy

STATUS
NOTIFICATION
EVENTS
TRENDING

staytus.io
slack.com
alerta.io
grafana.com
Kapacitor
Data Refinement
InfluxDB
Metrics DB
Infra. Metrics
VROPS
App Metrics
telegraf buildpack
git
Nagios
Collection/ Alerting
CF Metrics
System Metrics
JMX Bridge
Telegraf Agents
CC API
Cloud Foundry
BOSH
BOSH Config

PCF
Open Source
Commercial
git -- Available on github.com/comcast

COMCAST

Comcast's new monitoring and telemetry strategy, with InfluxDB as their metrics database, now works very well for them within their Cloud Foundry and Kubernetes infrastructures. This strategy's flow is as follows (they adopt the same approach depicted for Cloud Foundry in the above picture, for Kubernetes and other modern platforms):

- vRealize Operations Manager (VROPS) pushes in VMWare and infrastructure metrics into InfluxDB.
- Nagios is tied into the Cloud Foundry JMX Bridge to push in all Cloud Foundry platform metrics.
- Telegraf collects application metrics as well as some of Comcast's own platform metrics and pushes them in.
- Then Comcast uses Riemann processing to process and aggregate logs, and push that into InfluxDB.
- Staytus.io is used to publish maintenance updates to Comcast customers, with Slack used for alerting, and Alerta used to process the alerting.
- Grafana is used for visualizing and publishing metrics.

This approach is not a single monolithic solution but a collective of different tools that are performant for their purpose, enabling Comcast at any time to take a component out and slide a new one in as the next greatest tool comes out. It also lets Comcast retire old components, so it's a very modular approach.

Comcast has also made open-source contributions to Telegraf for Cloud Foundry users. These Telegraf enhancements, which were written by Sergey Matochkin (part of the brain trust that created Comcast's monitoring-telemetry strategy) are:

- Cloud Foundry/Telegraf buildpack, which lets Comcast's application team quickly and easily deploy Telegraf and a Telegraf config against their app
- Telegraf release for BOSH, which allows BOSH to manage Telegraf in a very self-healing fashion

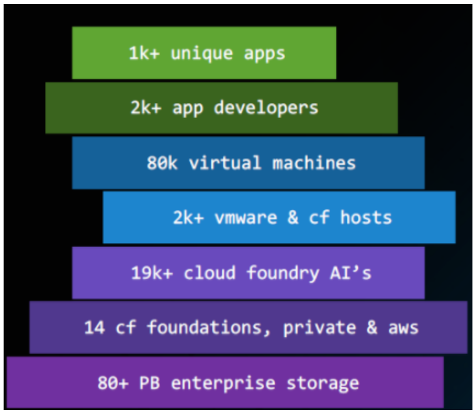
Results

> *"We had just discovered great things like InfluxDB, Grafana, and Telegraf. Shortly thereafter, we were able to deliver consolidated metrics dashboards that summarized our entire environment."*

As a result of its new monitoring and telemetry strategy using InfluxData open source tools, Comcast achieved tangible results:

- Major transformation in their mindset, in how they manage infrastructure, and how they deliver products
- Removed infrastructure heavy lifting
- Consolidated tooling
- Full transparency throughout the environment enabling internal teams and customers to see how everything is running at any given time
- Exercised numerous opportunities and drove numerous efficiencies including major cost savings
- App metrics being available the minute the app is "born" within one of their platforms

A Modern Infrastructure Today Allowing Visibility and Control



1k+ unique apps

2k+ app developers

80k virtual machines

2k+ vmware & cf hosts

19k+ cloud foundry AI's

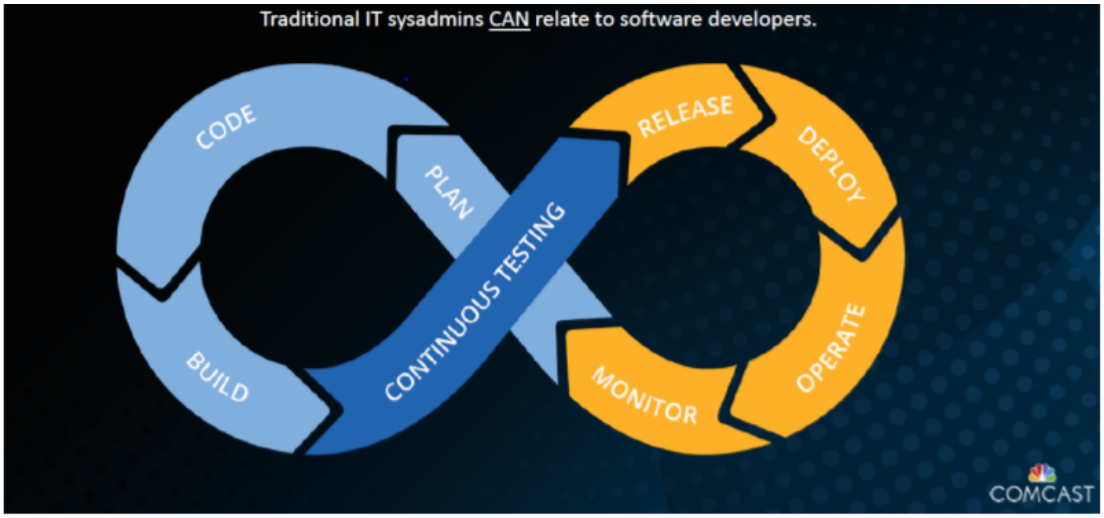
14 cf foundations, private & aws

80+ PB enterprise storage

Some of the concepts that helped Comcast drive successful results are:

- The DevOps mindset – inspiring each other to write code and automate
- Performance goals – setting goals for creating automation mechanisms
- Partnering with stakeholders – such as with developers consuming and deploying the platform
- Community-building – customers asking questions that get answered by other customers
- Transparency – sharing metrics to help build trust with stakeholders
- Run it like a business – soliciting and incorporating feedback from customers into the product roadmap

Nurturing the DevOps Mindset


Traditional IT sysadmins CAN relate to software developers.

Using InfluxData for monitoring and managing its infrastructure and apps, Comcast is fulfilling its mission of “Connecting You to What Matters”

About InfluxData

InfluxData is the creator of InfluxDB, the leading time series platform. We empower developers and organizations, such as Cisco, IBM, Lego, Siemens, and Tesla, to build transformative IoT, analytics and monitoring applications. Our technology is purpose-built to handle the massive volumes of time-stamped data produced by sensors, applications and computer infrastructure. Easy to start and scale, InfluxDB gives developers time to focus on the features and functionalities that give their apps a competitive edge. InfluxData is headquartered in San Francisco, with a workforce distributed throughout the U.S. and across Europe. For more information, visit influxdata.com and follow us @InfluxDB.