

The background of the slide features a wide-angle photograph of a lake at sunset. The sky is filled with large, billowing clouds in shades of deep blue, purple, and orange. Several bright, jagged lightning bolts strike down from the clouds towards the horizon. The lake's surface is calm, reflecting the warm colors of the sky. In the foreground, dark silhouettes of trees and bushes are visible along the shore. A small, dark structure, possibly a pavilion or bridge, is visible across the water.

# CULTURAL REVOLUTION REQUIRED

Bryson Koehler  
EVP, CITO



## Karman Line

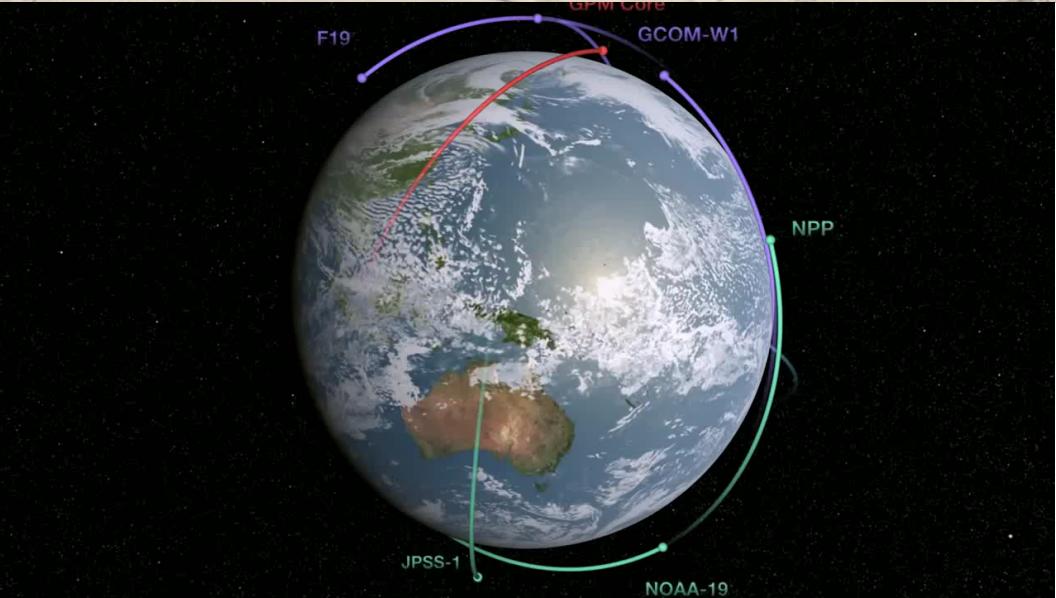
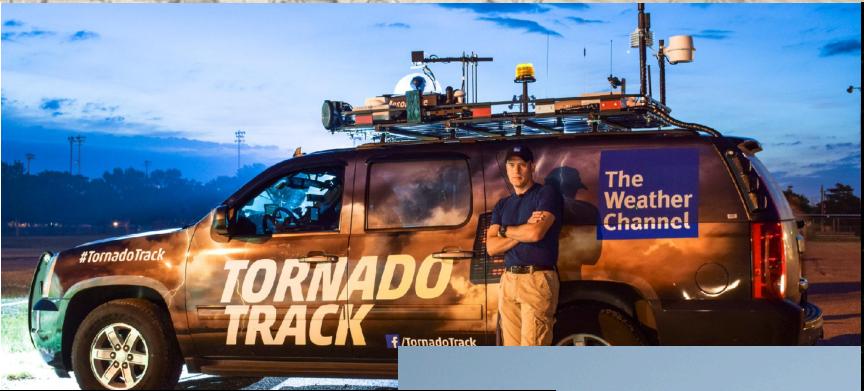
Absorbs ultraviolet radiation, retains heat and reduces temperature extremes - basically it sustains life on Earth..

## SR-71 Blackbird - My Favorite Airplane

Even though it hasn't flown in 15 years, it's still the official record holder for fastest manned jet-powered aircraft.

# We are the real time dashboard for the Karman Line

We use every tool available to push for accuracy

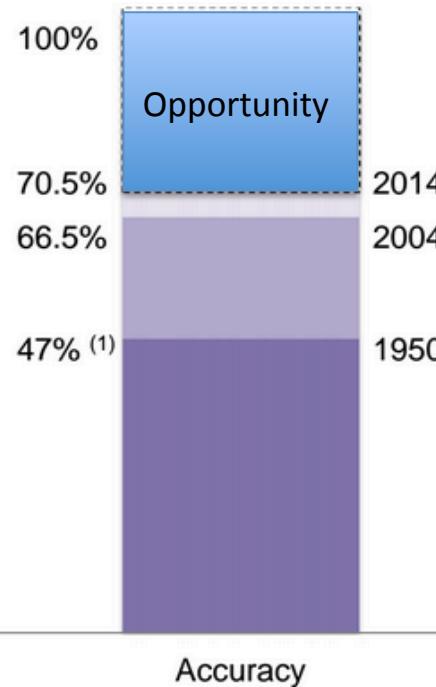


Planes, Trains, Automobiles, Phones, Satellites , 108,000+ Personal Weather Stations, IoT data sets of all types...

# Imagine a PERFECT forecast

What would your life be like if we were 95% accurate?

## Overall Average % Correct 0–10 Day Forecast



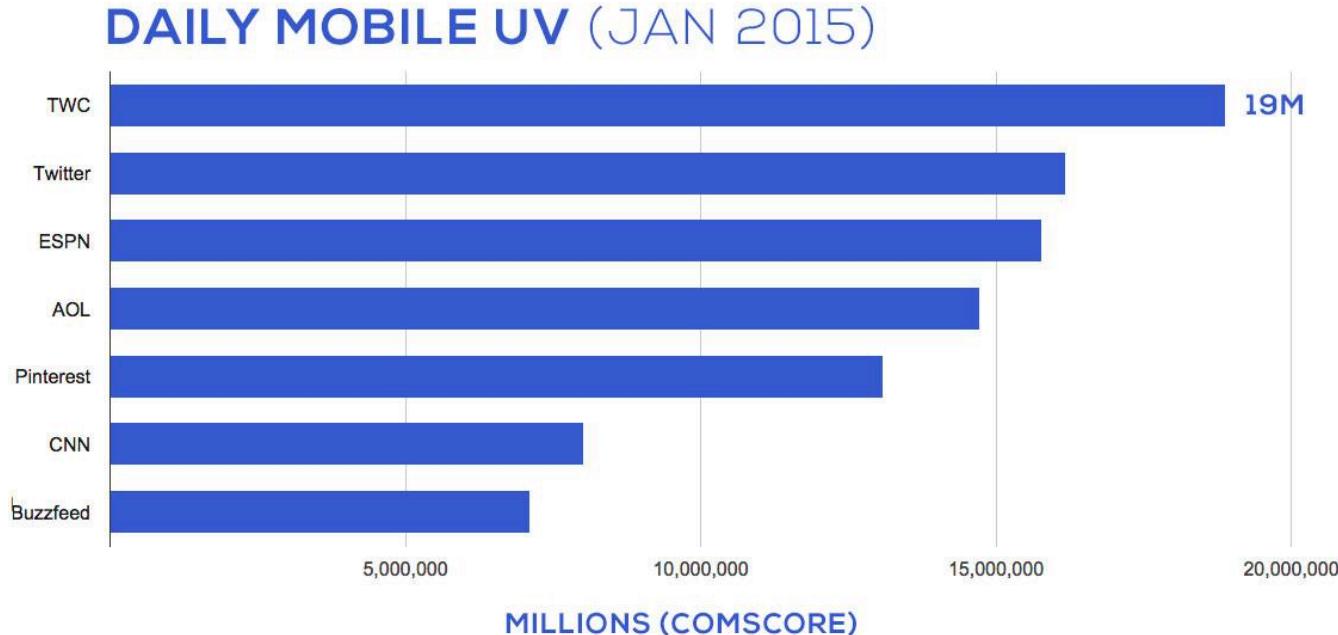
What are the implications of a perfect weather forecast, everywhere on Earth?

- Save millions of lives impacted by hurricanes, tornadoes, typhoons, etc.
- Improve crop yields and dramatically decrease world hunger
- Demonstrate the impact of climate change
- Make better decisions in every aspect of our daily lives
- A More Livable Planet

# Mobile Is Not Our Future – It's Right Now

Over 1.2 Billion video plays in 2014

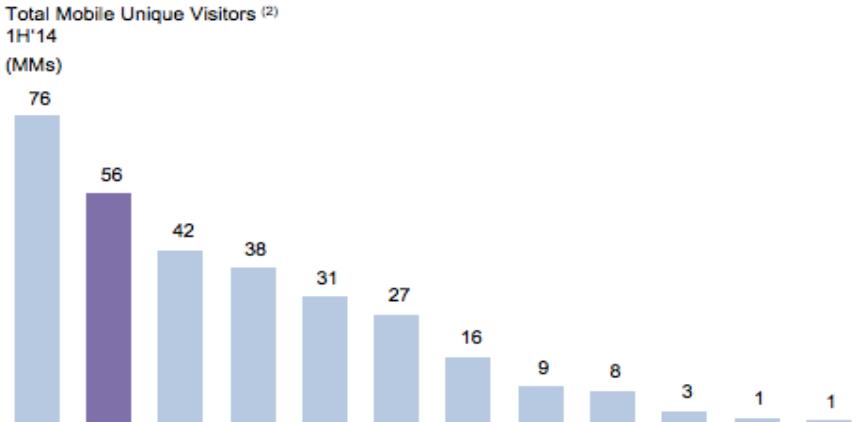
- Daily Mobile UVs are 65% higher than desktop (19M vs 11M)
- Bigger than ESPN, CNN, Pinterest and Twitter
- Mobile video users watch 2X to 3X more than desktop



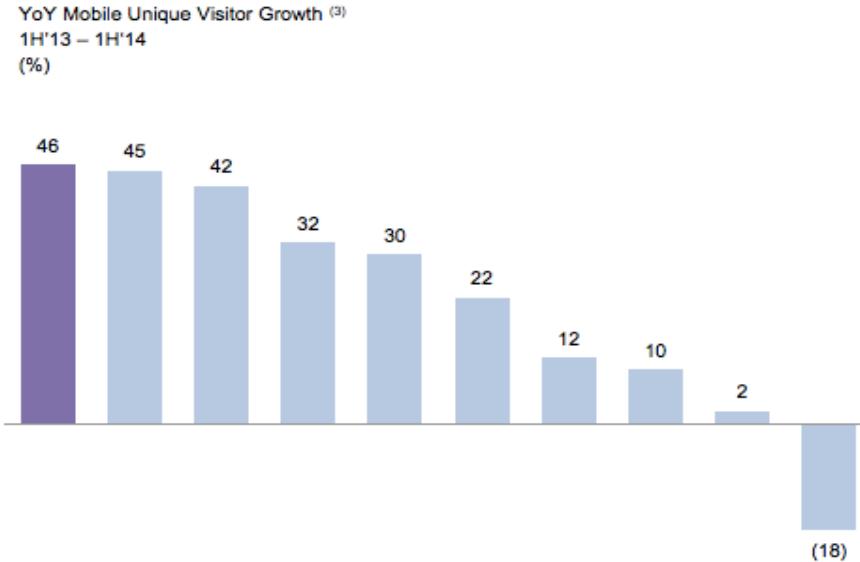
# We are the 2<sup>nd</sup> Largest location based services company in the world

**Mobile weather touch points grew 57% from 2013 to 2014**  
**Tablet weather touch points grew 92.7% from 2013 to 2014**

## One of the Most Used Location-Based Platforms and ...



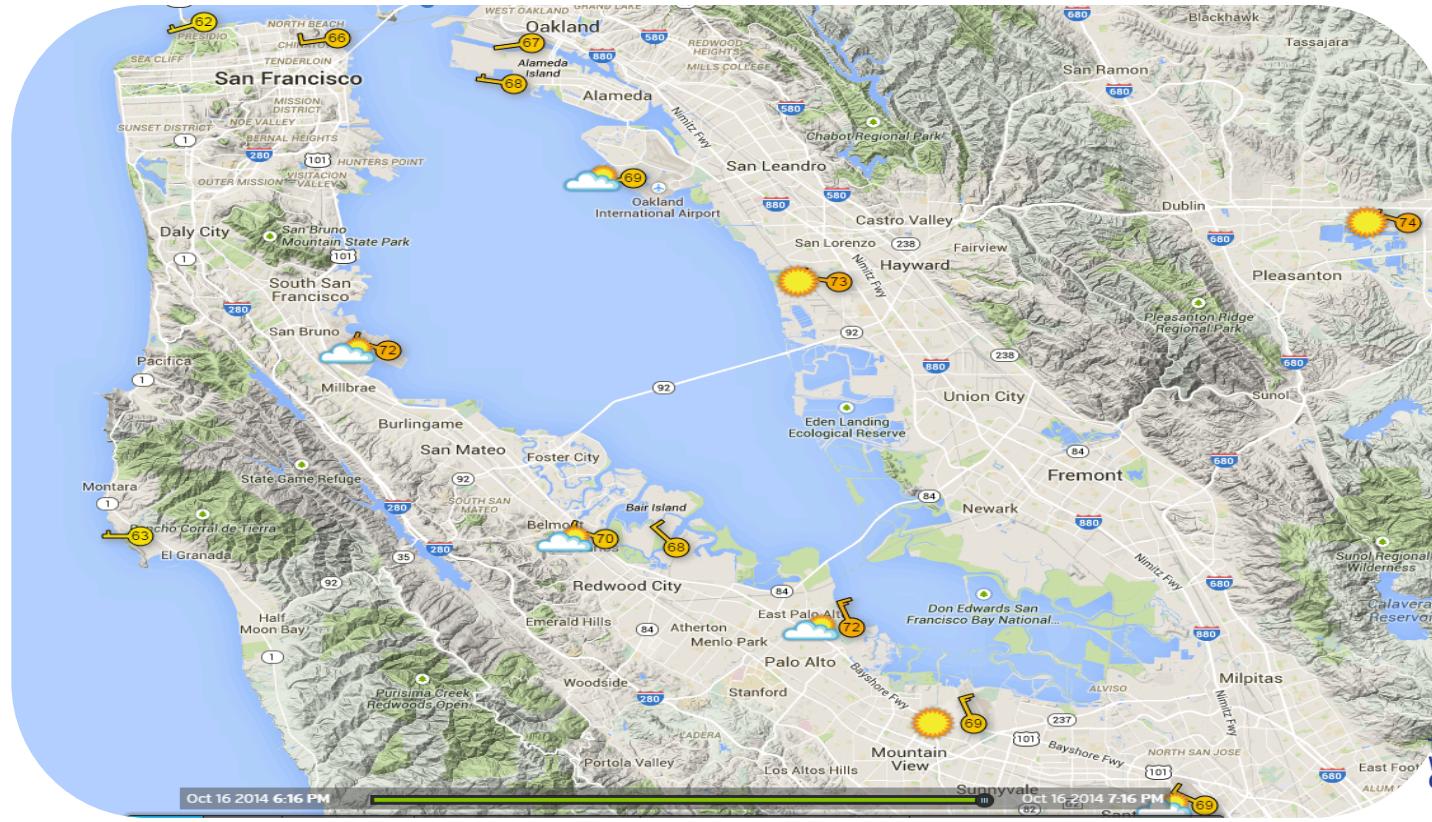
## ...Growing Faster than Any Other Local Media



Leadership in localization will allow continued expansion into additional local services, directly and through partnerships (e.g., Hearst, Tribune)

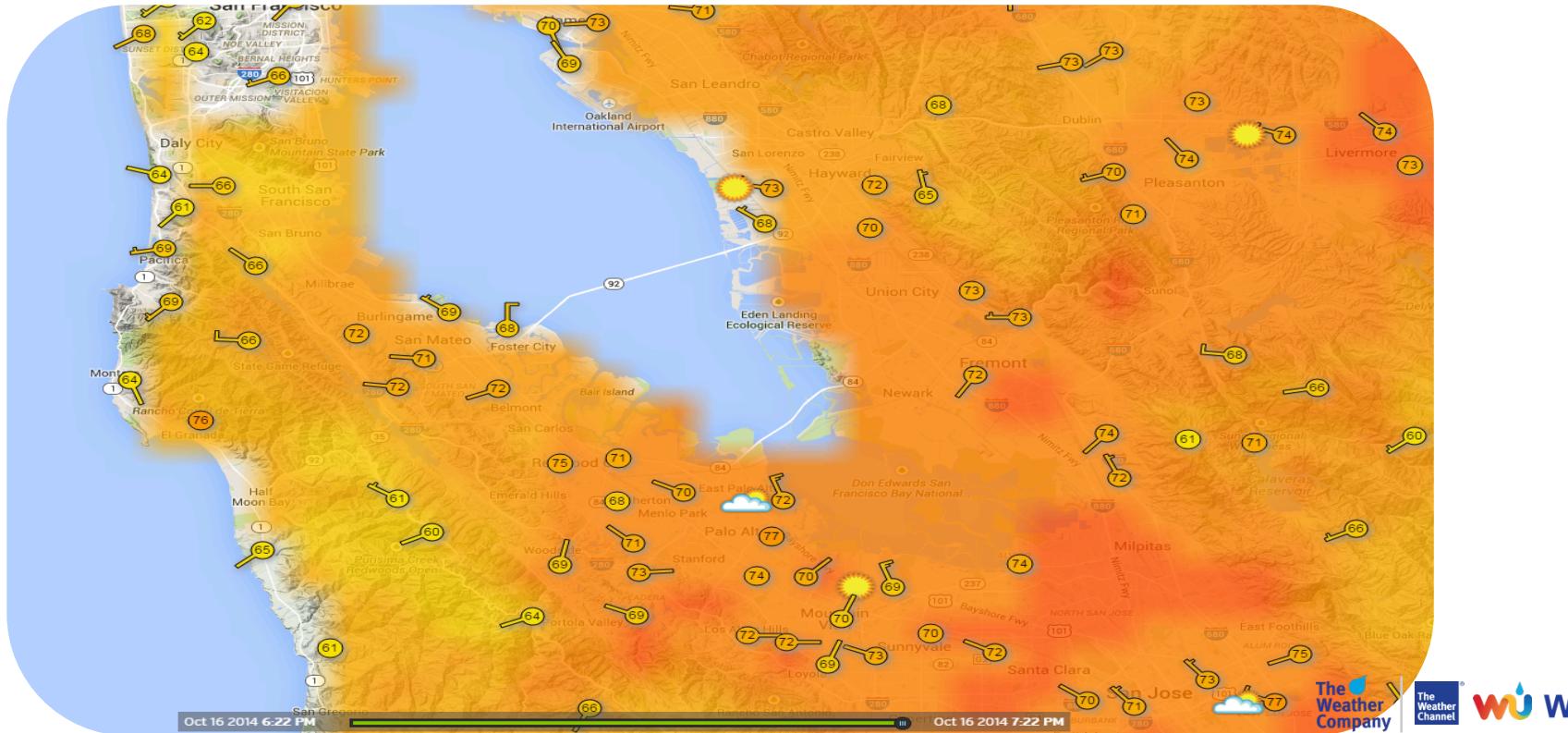
# Traditional Data (NWS, FAA, DoD) View

Provides 14 observations points near Mountain View, CA

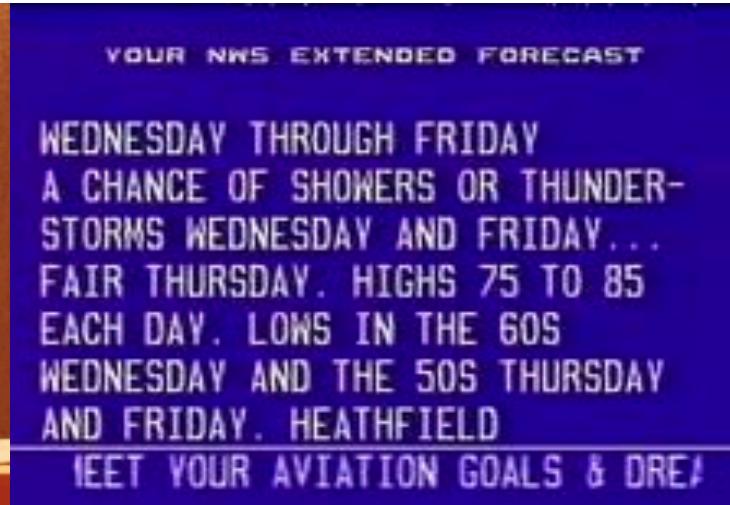


# Our proprietary Surface Observations (PWS)

Provides for 96 locations near Mountain View, CA



# In 1982, there was a Vision



TWC originally gathered its national region forecasts from the National Oceanic and Atmospheric Administration (NOAA) and its local forecasts from the National Weather Service.



In 2013

We started a *Revolution*



# Our Journey from Cable TV – Big Data Technology Leader

The world's most advanced broadcasting system

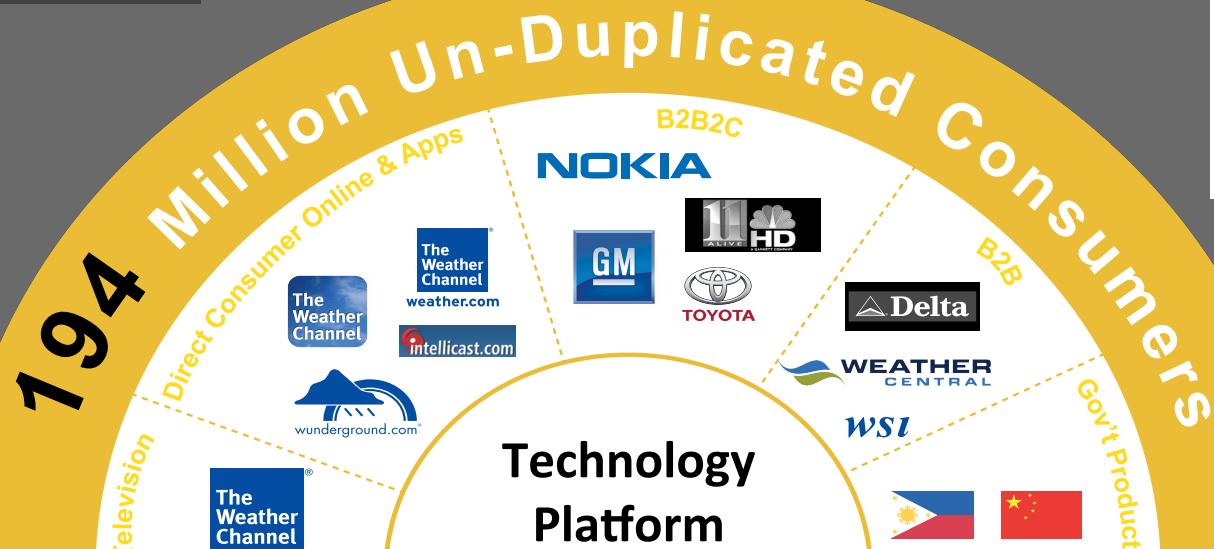
The world's most accurate forecaster

100% Cloud based forecasts



The most viewed media platform

The world's largest API platform



# World's Leading Weather Media & B2B Company

4,600+ clients world-wide



## World's Leading Aviation Weather Service

50,000+ flights daily with embedded forecasters, systems and tools

19 out of the top 20 airlines in North America

46 of top 100 airlines globally

## World's Leading Energy Forecaster

Over 300 power trading clients worldwide

## Fastest Growing Insurance Weather Service

7 of the top 10 P&C insurance customers

Connecting Marketing, Underwriting, & Claims with systems & tools

# However, we still had roots of our 1982 era tech

Through acquisitions, lack of architecture governance and under investment we were broken...

- 13 Data Centers interdependent on each other
- 4 forecasting systems
- Duplicated data platforms and data ingestion points
- SQL Server, Oracle, Mongo, MySQL, PostgreSQL, Hadoop, Redis, Riak and Cassandra
- Over 400 pages of documented single points of failure
- VAX still in production
- < 75% of production changes successful
- Unmonitored and unreportable system performance data

# We needed a global data platform built for scale

Scale to over 150,000 transactions a second to over a billion devices

Java  
Scala  
C  
C++  
RabbitMQ  
Redis  
Riak  
Cassandra  
R Studio  
Redshift  
Varnish  
Akamai  
Akamai GTM  
Puppet  
Gradle  
MySQL  
AWS  
Etc...

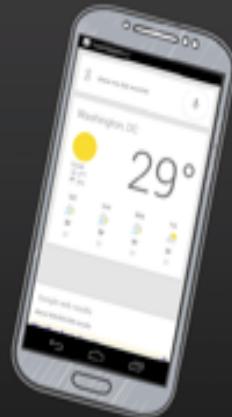
We Are on Every iOS8 Device

182 Million+ Devices



We Are on Google Now

900 Million+ Devices



170 Million+ Downloads  
The Weather Channel Apps

*At our core we believe in the power, capabilities and quality of the open source community.*

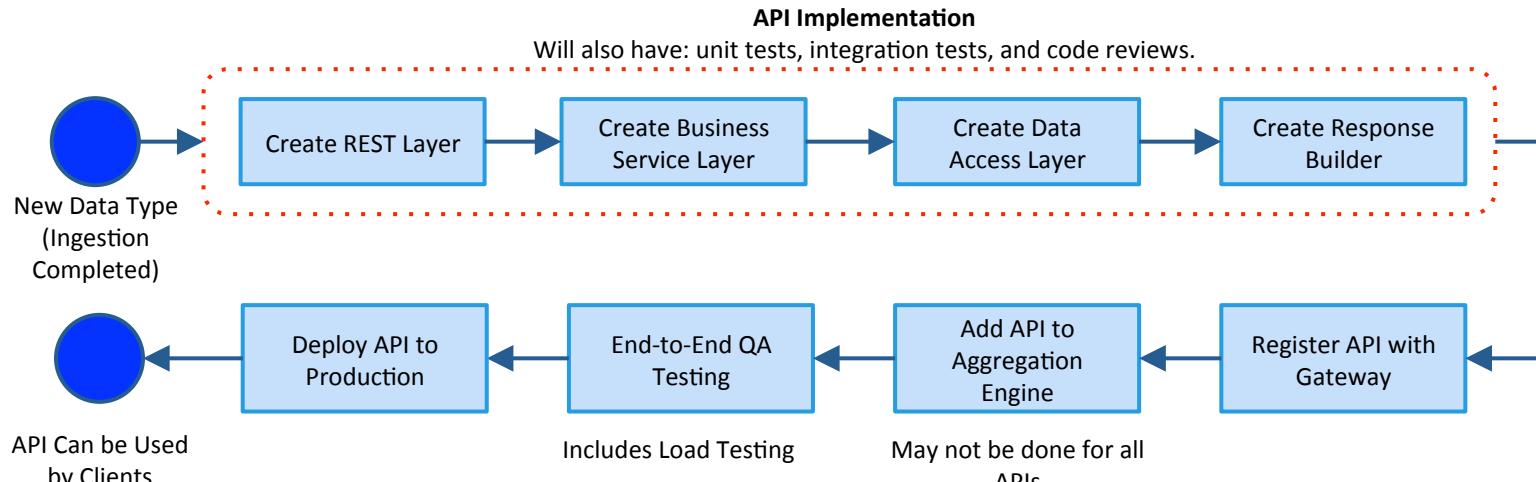
# Data Ingestion

Our goal was to be able to ingest, store and expose any new data type within one week through a new API

- Always on, Always flowing
- High availability, low latency and horizontal scalability
- Loosely coupled layered systems
- Smart publishers to handle queuing semantics
- Exactly once delivery through queuing system
- Message pump / Listener container for asynchronous processing of inbound messages
- Provides support for Message-driven POJOs (Cartridges)
- Graceful handling of outages through error queues
- Process-store-archive-distribute paradigm with extendable workflow model

# Data Ingestion

Our goal was to be able to ingest, store and expose any new data type within one week through a new API

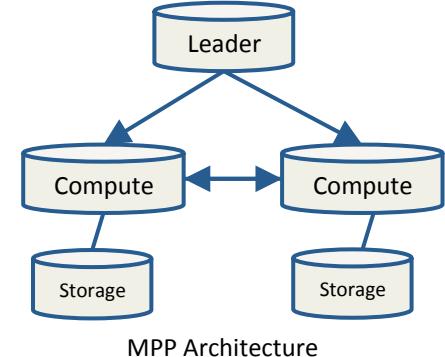
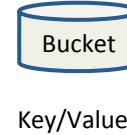
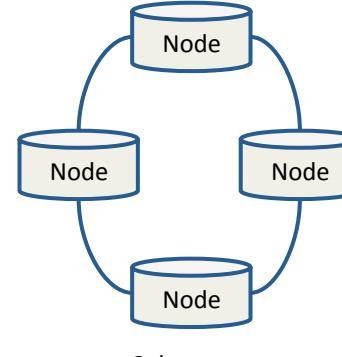
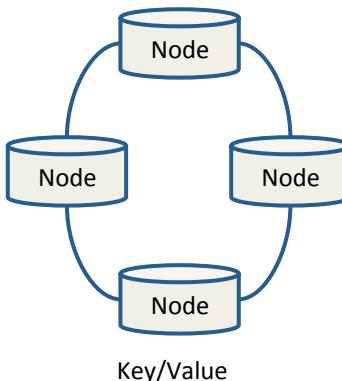
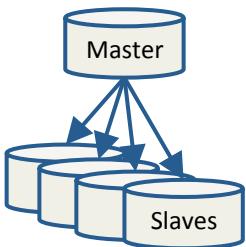


- 900+ Data Types
- 20 TB / Day of data volume
- Updates from 108,000+ PWS stations every 3 seconds
- Telemetry from aircraft
- Etc...

- Multi step processing for complex tasks with configurable cartridges
- Generic geospatial cartridges to collect smart dust
- Cartridges to support scientific datasets
- Image processing cartridges with mip-mapping and interpolation techniques

# Data Storage

Simple global replication for all data sets is critical



Real-time  
Data and  
Caching

Archive  
Data

API Gateway  
Data  
and Analytics

Archive  
Data

Analytics

# A RESTful API Centric Distribution System

On a mission to achieve 150,000 transactions per second with sub 10ms latency

**Fine Grained** - one data type for one location (top level resource is location)

- Allows for targeted selection and caching of individual data types
- Can be composed via aggregation

**Aggregate** - multiple data types for one location

- Reduces chattiness and latency from multiple network hops with backend
- Simplifies client code by only having to react to a single HTTP call

**Polling** - customer hits consistent URL to get updates to a product

- Allows client to determine how fast they wish to consume the data
- Simplifies integration technology requirements to only requiring HTTP

Additionally have message queues for a small number of internal clients only.

# Data Distribution

On a mission to achieve 150,000 transactions per second with sub 10ms latency

## Availability

- Deployed across 4 global regions (US-Virginia, US-Oregon, Singapore & Ireland) on 8 data centers (2 data centers in each region)
- 1 management data center (US-California) (Monitoring, Configuration management, Automation framework, etc)

## Scalability

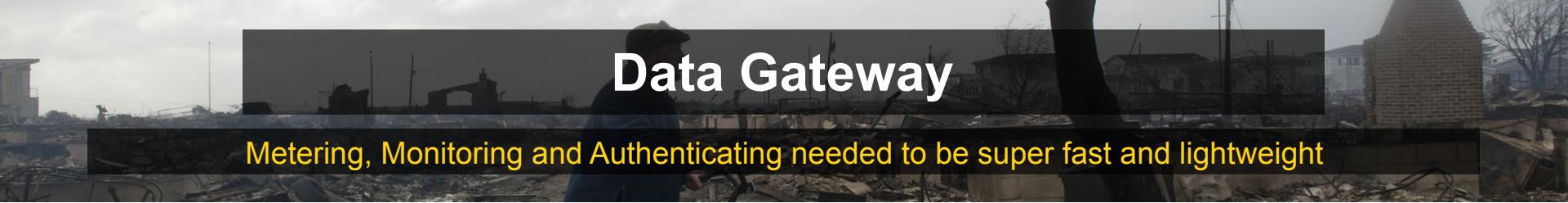
- Auto scaling (up & down) based on load
- Each region has been load tested to handle 150K req/sec (Can be increased based on needs)

## Distributed

- Persistent data stores (Riak and Cassandra) are globally distributed
- 50+ Languages

## Performance

- Multi-tiered caching for low latency



# Data Gateway

Metering, Monitoring and Authenticating needed to be super fast and lightweight

- Eventual Consistency is Sufficient for IoT data sets
- We sacrificed absolute precision in throttling to allow for an API which could respond with very low latency at scale
- The authentication and metering adds .25 milliseconds to each HTTP transaction
- Authentication is truly horizontally scalable, not limited to backend DB bandwidth/scale

# Data Analytics

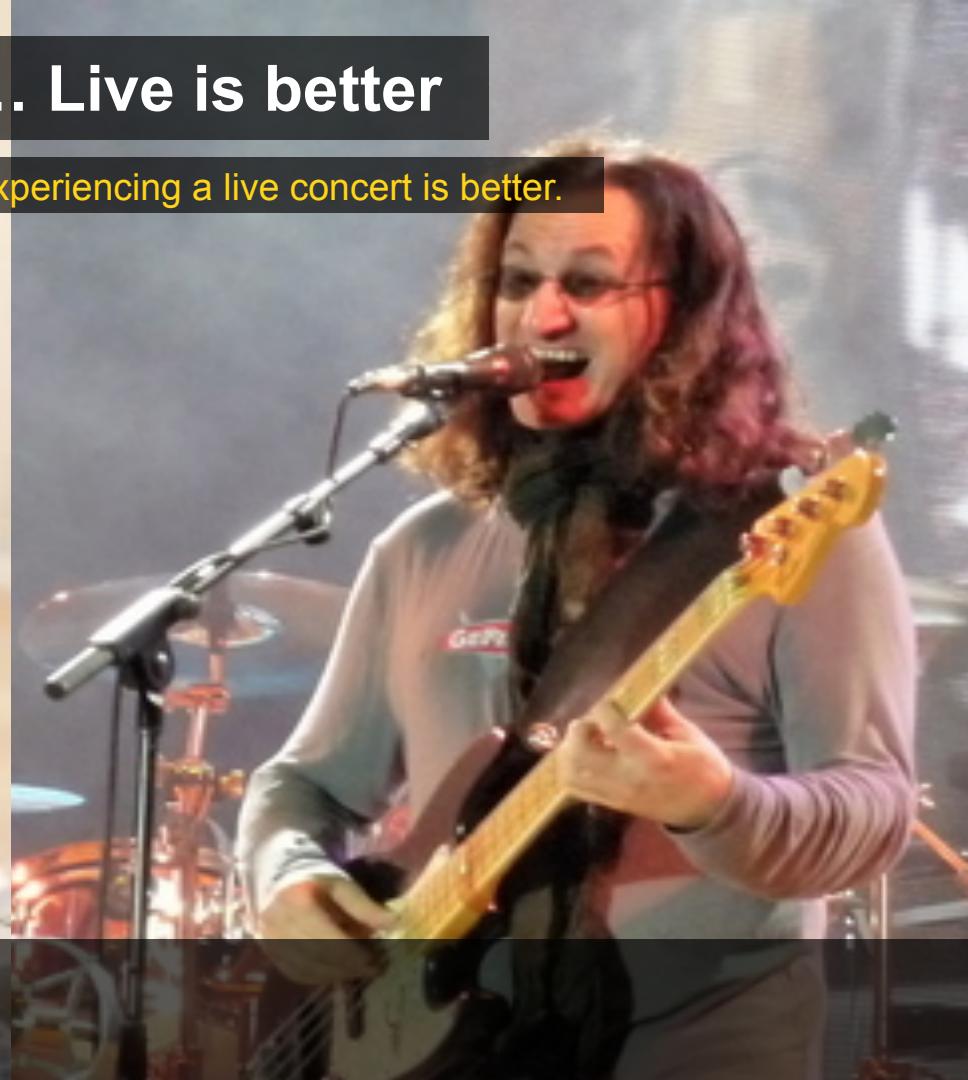
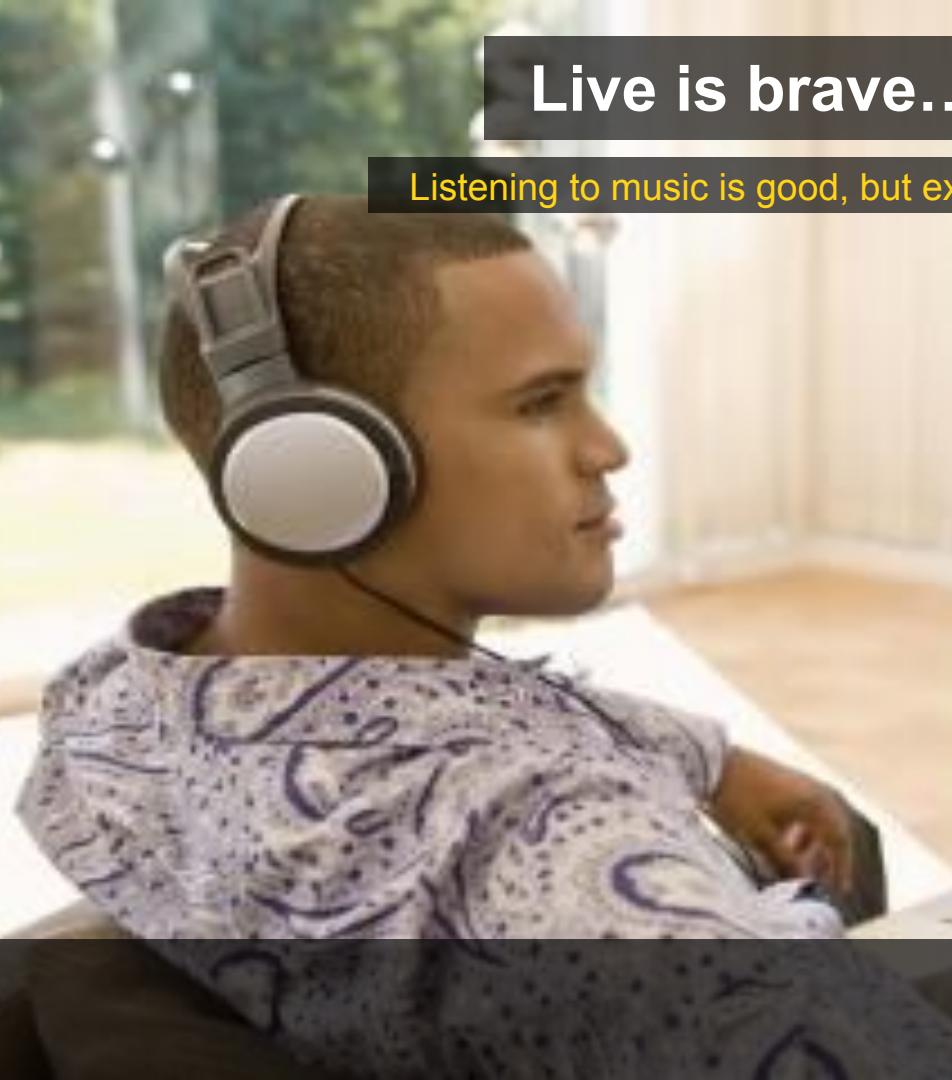
The real time intersection of weather and human/consumer behavior

- Based on **Spark + Cassandra**
- Supports both **streaming** and **batch** analysis via Spark/Spark Streaming
- Accepts a **generic time-series** event schema
- Ingestion API supports **filtering and pipelining** based on event type/content
- **SQL queries** using SparkSQL
- **Workload separation** achieved using Cassandra's data center replication features
- Analytics DC runs Cassandra, HDFS, and Spark co-located for **data local processing**
- Analysis output is written to **aggregate tables**, then **replicated in real time**
- Live and analysis workloads **don't fight for the same resources**
- **No ETL** in either direction



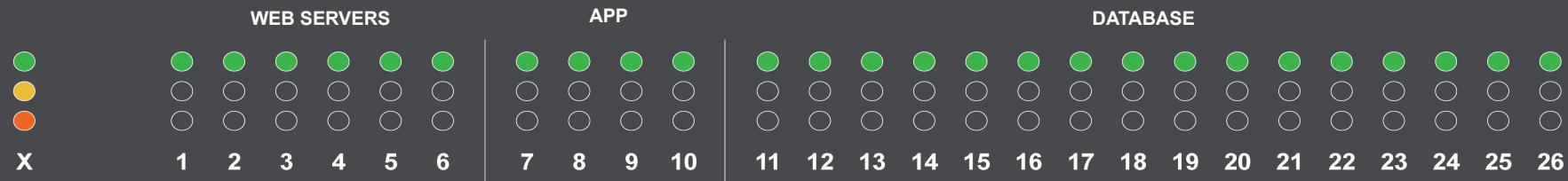
# Live is brave... Live is better

Listening to music is good, but experiencing a live concert is better.

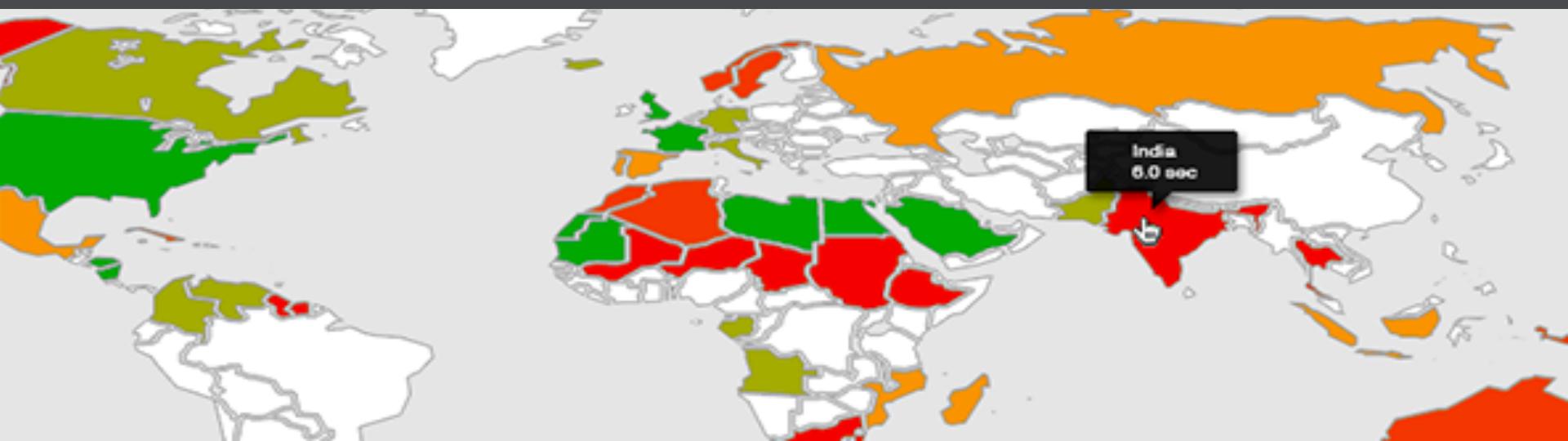


# What Gets Measured

Server performance doesn't tell the right story.



Our teams should be measured on what our customers see



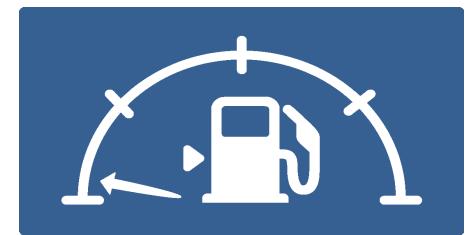
# Focus on the Right Measurement

Similar to performance, costs should be looked at on a unit basis.

Unit Price



Retail Price



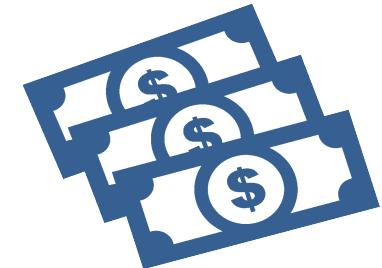
Cost per page view

Cost per API call

Cost per current condition ingested

Cost per forecast model run

Etc.



# The Process of Getting it Done

Working harder isn't a long term solution

**“Follow the Sun”**



[We can't work 24/7 but there is a way that work can happen around the clock.]

Development must become 24/5

# Once the U2 was shot down, there was no choice...

Built through brut force willpower



They produced 6,000 parts for each plane, and of them fewer than ten percent were any good. Titanium was so brittle that if you dropped a piece on the floor it would shatter

Ordinary drills were rendered useless because after about 17 rivet holes the drill would be completely destroyed

# Safe is Risky = Leaking Fuel Tanks



Several design “flaws” were left unsolved due to lack of tech... But not one plane was ever shot down...



Because of the fuel leak it had to refill immediately after takeoff. Every single time.

# My Favorite SR71 Facts

Taking risks, calculated ones, pays off

New York - London 1:54:56.4

London - Los Angeles 3:47:39

West Coast - East Coast USA 1:7:53.6

Los Angeles - Washington D.C. 1:4:19.8

St Louis - Cincinnati 0:8:31.9

At cruise, about 600°F caused the aircraft to grow 3-4 inches length & 1-2 width during cruise.

It could photograph an entire swath of North Korea in seven minutes.

...And it set one record with only one engine...

On the run, one engine shut down at the start of the second run, and the pilot tried to get it relit. The pilot, then-Capt. Al Joersz, said, "By the time we'd gone through the checklist, we'd already passed the second gate [thus officially starting the run]. Still, we exited the gate at Mach 3.2."

Over the course of the plane's service, more than 1,000 missiles were fired in an attempt to take it down. The evasive maneuver was just to hit the throttle and go faster than the missile. Needless to say, the SR-71 took it's 1.000 batting average into retirement..

# Have the Courage to Commit



- Change what gets measured
- Remove barriers and layers
- Single person accountability
- RAPID decision making
- Know your customer base
- Play to your strengths
- Supplement your weaknesses
- Have the courage to decide
- Be the chief disorganizer