
AWS
re:Invent

WPS215

Modernizing engagement for the 2020 United States Census

Nicci Neal
Sr. Manager
AWS Government Regions, Civilian and Security

Michael Thieme
Assistant Director
Decennial Census Programs
U.S. Census Bureau

Utpal Amin
Chief Engineer
T-Rex Solutions LLC

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

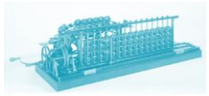
Agenda

1. **Modernizing engagement** for 2020 U.S. Census
   - **Brief history** of Census
   - **Use of technology** in 2020 U.S. Census
2. **Census case study:** Internet Self-Response (ISR)
   - ISR Architecture
3. **Resiliency and scalability** using AWS
   - Examples
   - AWS Services used in 2020 U.S. Census
4. **Successful delivery**



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The U.S. Census continues to advance America's computing industry



1880

1st Commercial Calculator in 1850



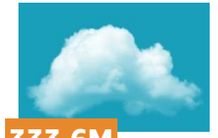
1890

1st Tabulating Machine



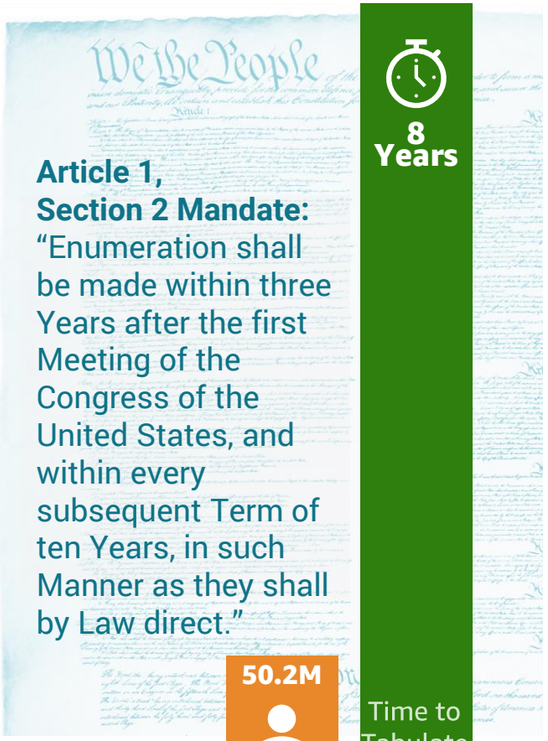
1960

1st General Purpose Electronic Digital Computer in 1951



2020

1st Census to use "Public Cloud" Technology



Article 1, Section 2 Mandate: "Enumeration shall be made within three Years after the first Meeting of the Congress of the United States, and within every subsequent Term of ten Years, in such Manner as they shall by Law direct."

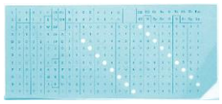
8 Years

Time to Tabulate

50.2M

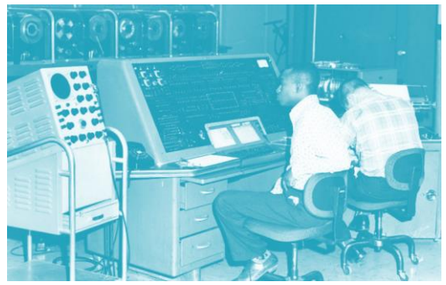


Herman Hollerith created an electronic tabulating machine for use on the 1890 Census





62.9M

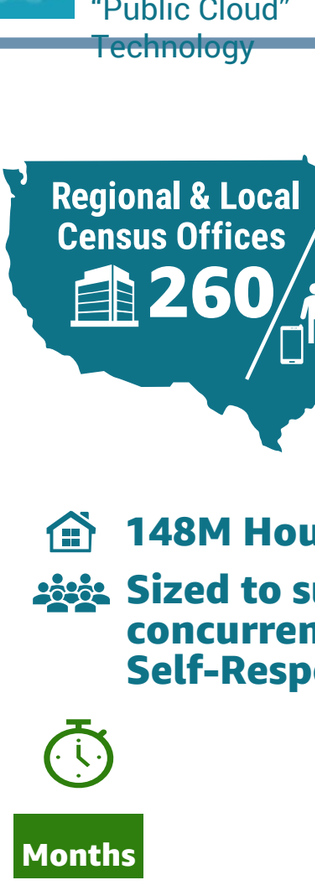


189.3M

UNIVAC I:

- 1st UNIVAC computer delivered to the U.S. Census Bureau in 1951
- 1ST used to process and tabulate data for 1960 Decennial Census

333.6M



148M Households

Sized to support 600K concurrent Internet Self-Respondents

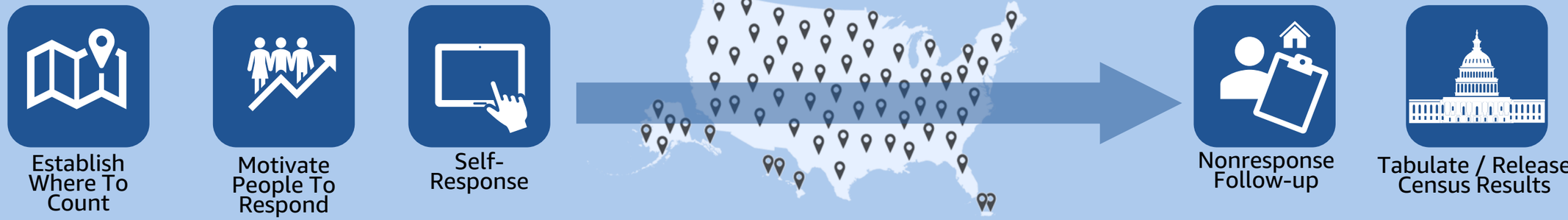
Months

https://www.census.gov/history/www/innovations/technology/tabulation_and_processing.html

T-REX

How technology is being used for the 2020 U.S. Census

Census overview: Complete count of people and housing



Key innovation areas

Re-Engineering Address Canvassing

Reduce nationwide In-Field Address Canvassing by using innovative in-office methodologies

Optimizing Self-Response

Encourage the population to respond to the 2020 Census using the Internet, reducing the need for more expensive paper data capture

Utilizing Admin Records and 3rd-Party Data

Use data the public has already provided and data available from commercial sources, allowing savings

Reengineering Field Operations

Use sophisticated operational control systems to optimize the routes for Census field employees

T-REX

Decennial 2020 U.S. Census by the numbers

Mission Critical Systems

52 **20**

7000+ Instances

W E **2 Region Deployment**

Load Tested for 1.2M Peak Concurrent Users

Regional & Local Census Offices

260 / **17K Office Workers**

Decennial Systems Collect Information From:

- 334M Residents
- 148M Households
- 500k Mobile Enumerators

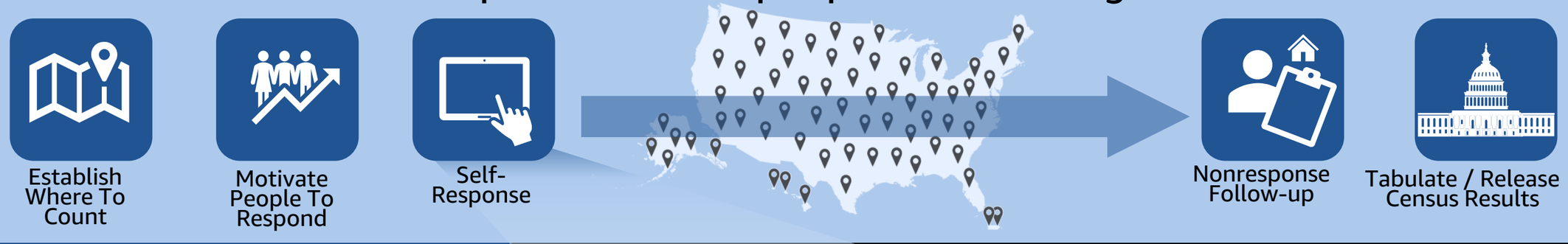
18 Months to Stand up

End-to-End Test

T-REX

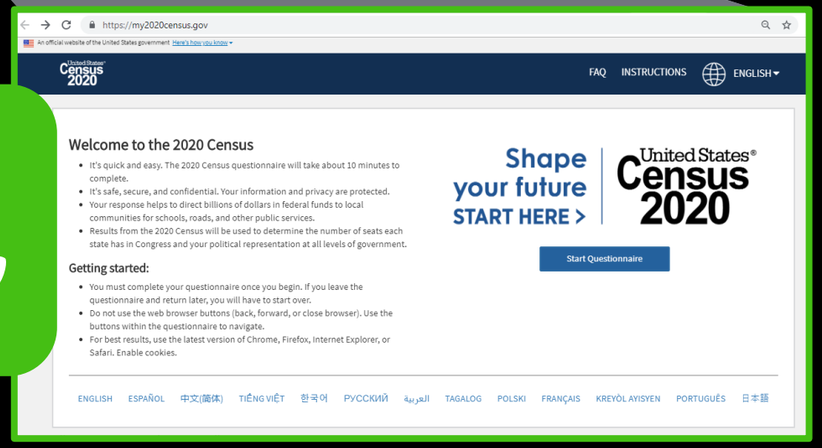
Case study: Internet Self-Response System

Constitutional mandate: Complete count of people and housing



United States® Census Bureau

aws



T-REX

1st Place Gameday 2020
aws Public Sector

T-REX aws partner network
AWS PUBLIC SECTOR
Partner Awards 2020

Bloomberg Government FedCiv #1 Cloud Provider 2019

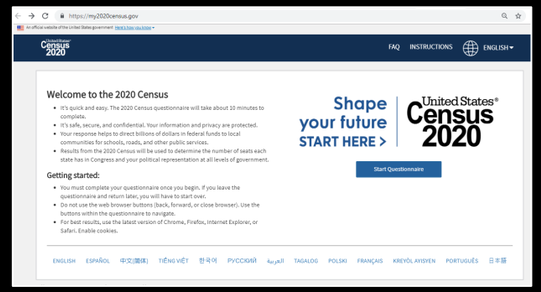
GovWin from Deltek FedCiv Cloud Top Vendor 2017-2019

App layer: AWS GovCloud (US)

Service Identification

High availability service to U.S. public

- Minimal impact to respondent during failure



- Multiple language support
- Responsive design
- 508 Compliant

Design

Resiliency

1. ***Multiple software apps*** (different tech) for single business function
2. ***Redundancy:*** AWS GovCloud (US) Regions; multiple Availability Zones
3. ***Data replication across Regions*** (Census ID, Session Information)
4. ***Native AWS resiliency features*** (Amazon DynamoDB, AWS Transit Gateway, AWS Lambda, Multi-AZ AWS RDS)

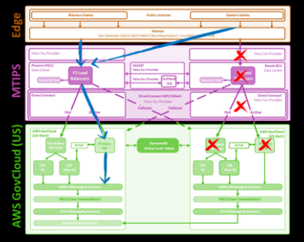
Scaling

1. ***Primary ISR:*** AWS Auto Scaling
2. ***Secondary ISR:*** Cluster-based scaling
3. ***Both apps:*** "Contain and Sustain" strategy

Service Transition

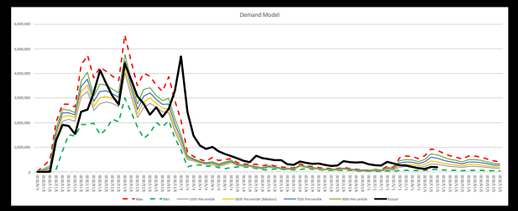
Resiliency

1. ***Execution / practice*** of failover procedures
2. ***Production validation*** conducted before go-live



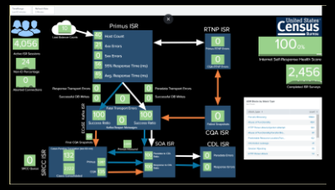
Scaling

1. ***P&S tested to support 1.2M*** concurrent users in each region
2. ***Determined breaking point*** by simulating 2.4M concurrent users



Operation

1. **Monitoring each service component**
2. **Splunk ITSI** as primary aggregator of system metrics
3. **Advanced monitoring dashboards** developed w/key stakeholders
   - Dashboard dry run during production validation
   - Highly relevant / well understood KPIs**:**



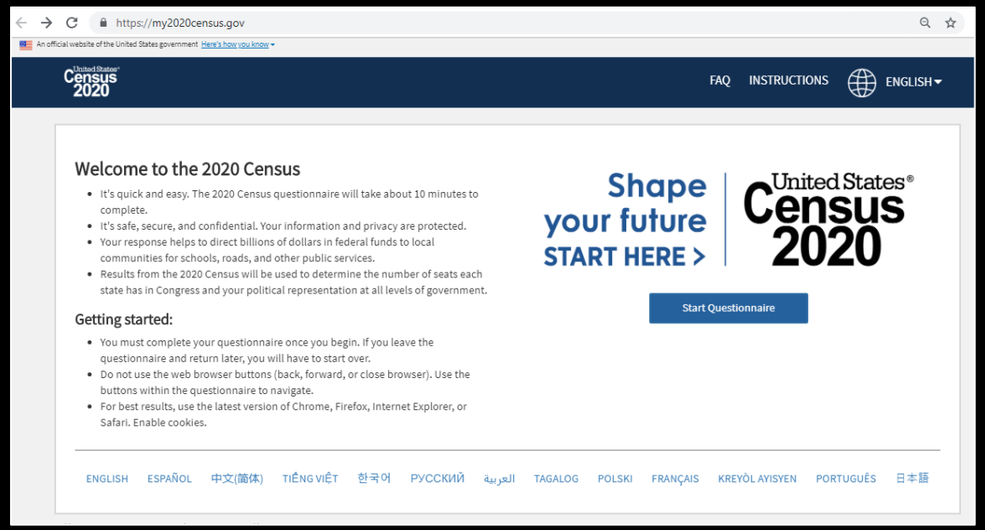
4. **Examples of resiliency / scaling during ISR operations**
   - Network Outage
   - Patching
   - Failover for CQA Call Centers

T-REX

App layer: AWS GovCloud (US)

Service Identification

High availability service to US public
- Minimal impact to respondent during failure



- Multiple language support
- Responsive design
- 508 Compliant

Design

Resiliency

1. ***Multiple Software Apps*** (different tech) for single business function
2. ***Redundancy:*** AWS GovCloud Regions; multiple Availability Zones
3. ***Data Replication across Regions*** (Census ID, Session Information)
4. ***Native AWS resiliency features*** (DynamoDB, Transit Gateway, Lambda, Multi-AZ RDS)

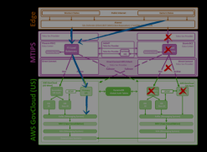
Scaling

1. ***Primary ISR:*** AWS Auto Scaling
2. ***Secondary ISR:*** Cluster based scaling
3. ***Both Apps:*** "Contain & Sustain" strategy

Service Transition

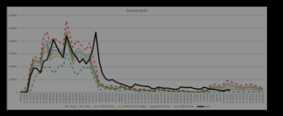
Resiliency

1. ***Execution / practice*** of failover procedures
2. ***Production Validation*** conducted before go-live



Scaling

1. ***P&S tested to support 1.2M*** concurrent users in each region
2. ***Determined breaking point*** by simulating 2.4M concurrent users



Operation

1. **Monitoring each service component**
2. **Splunk ITSI** as primary aggregator of system metrics
3. **Advanced monitoring dashboards** developed w/key stakeholders
   - Dashboard dry run during production validation
   - Highly relevant/well understood KPIs:

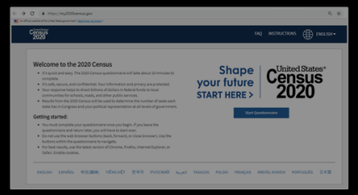


4. **Examples of resiliency/ scaling during ISR operations**
   - Network Outage
   - Patching
   - Failover for CQA Call Centers

App layer: AWS GovCloud (US)

Service Identification

High availability service to US public
- Minimal impact to respondent during failure



- Multiple language support
- Responsive design
- 508 Compliant

Design

Resiliency

1. ***Multiple software apps*** (different tech) for single business function
2. ***Redundancy:*** AWS GovCloud (US) Regions; multiple Availability Zones
3. ***Data replication across Regions*** (Census ID, Session Information)
4. ***Native AWS resiliency features*** (Amazon DynamoDB, AWS Transit Gateway, AWS Lambda, Multi-AZ AWS RDS)

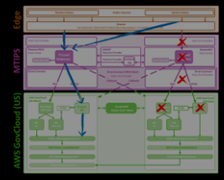
Scaling

1. ***Primary ISR:*** AWS Auto Scaling
2. ***Secondary ISR:*** Cluster-based scaling
3. ***Both apps:*** "Contain and Sustain" strategy

Service Transition

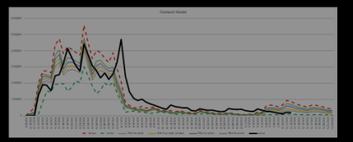
Resiliency

1. ***Execution / practice*** of failover procedures
2. ***Production Validation*** conducted before go-live



Scaling

1. ***P&S tested to support 1.2M*** concurrent users in each region
2. ***Determined breaking point*** by simulating 2.4M concurrent users



Operation

1. **Monitoring each service component**
2. **Splunk ITSI** as primary aggregator of system metrics
3. **Advanced monitoring dashboards** developed w/key stakeholders
   - Dashboard dry run during production validation
   - Highly relevant/well understood KPIs:



4. **Examples of resiliency/ scaling during ISR operations**
   - Network Outage
   - Patching
   - Failover for CQA Call Centers

T-REX

The ISR solution

Edge

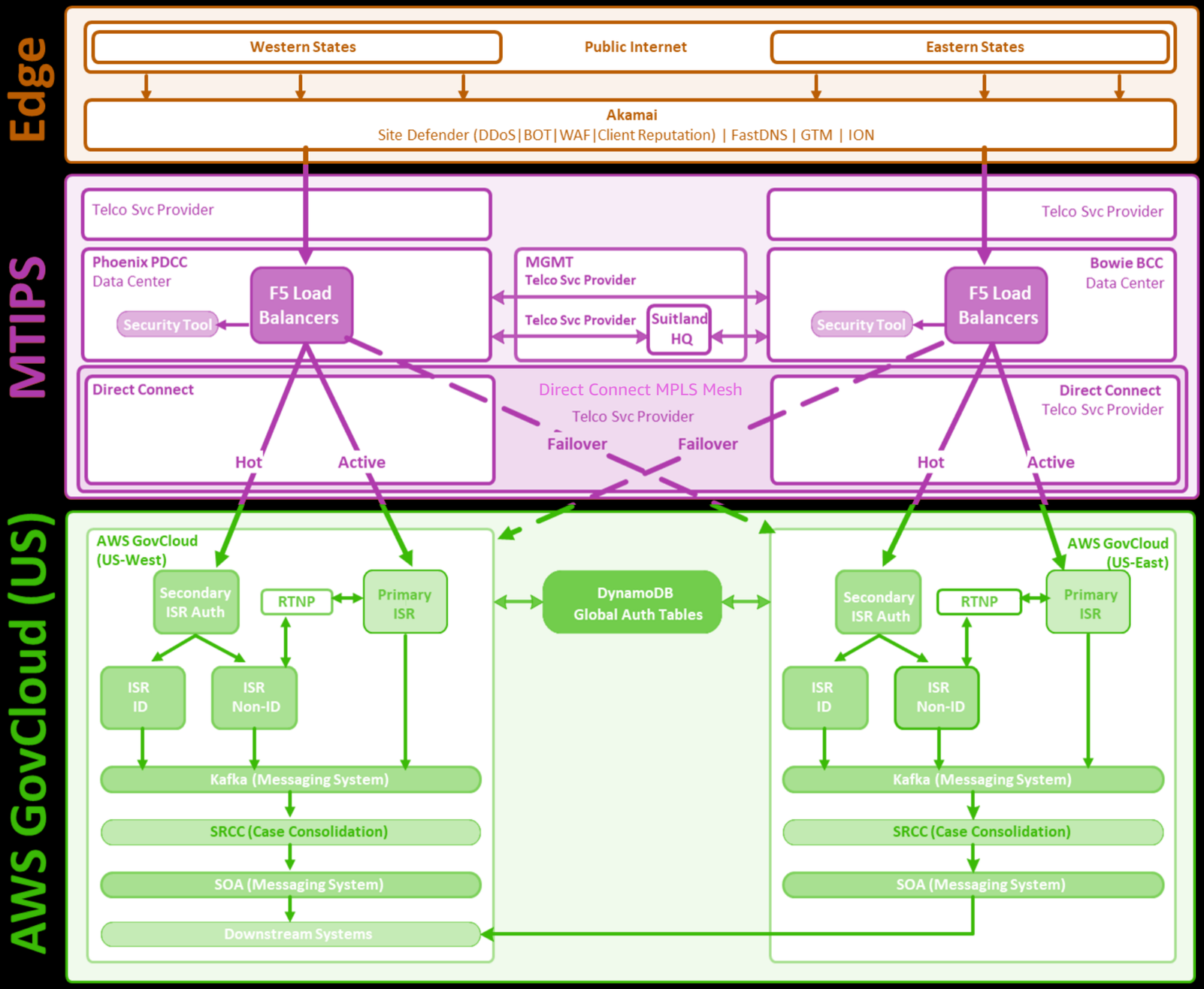
- ***Eastern states:*** BCC
- ***Western states:*** PDCC
- managed traffic pattern

MTIPS

- ***Traffic Traverses*** 40gbps MTIPS
- ***Load Balance:*** On-premises F5 clusters across Primary ISR nodes
- ***Traffic Delivery:*** 40 Gbps AWS Direct Connect circuits to AWS GovCloud

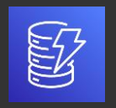
App Layer: AWS GovCloud (US)

- ***Distributed regionally*** across 3 Availability Zones each
- ***Auto-scaling*** resources
- ***Redundancy*** via multiple applications
- ***Global tables*** for Region sync



T-REX

AWS services for 2020 U.S. Census

Product	Purpose	Performance
 Amazon DynamoDB	… storage of Census ID and session information to allow multiple ISR applications to be able to use same information across regions	• **100+ tables** • **1.2B objects,** Provisioned IOPS, • **Global Tables** with cross region replication
 Amazon EC2	… primary compute for various applications in the cloud	• **7000 + Instances** across 16 VPC using 69 different instance types with 7.5 PB of EBS storage and three OS (RHEL, Windows, Linux)
 Amazon RDS	… database for various application to store information like census count status, mailing information, payroll, etc.	• **750+ instances** using Oracle, Postgres, and MySQL with various licensing models and 500 TB of storage • **MultiAZ,** Read Replicas for reporting
 Elastic Load Balancing	… primary mechanism used by all the apps in the cloud to load balance work across resources	• **1348 Application** Load Balancers • **27 Network** Load Balancers • **20 Classic** Load Balancers
 AWS Lambda	… used by ISR primary application to send data to downstream systems. Automated infrastructure maintenance tasks like disable public read/write access on S3 buckets	• **2 languages** (Node.js and Python) • **100+ functions** across all environments
 Amazon Simple Storage Svc (S3)	… map data for download to enumerator app, storage of simulated data, backups of various systems	• **219** buckets • **4PB** of total storage

T-REX

Additional AWS services for 2020 U.S. Census


Amazon EMR


AWS CloudTrail


Amazon Simple Notification Service


AWS Snowball


AWS Direct Connect


AWS CloudFormation


Amazon Simple Queue Service


Amazon S3 Glacier


AWS Transit Gateway


Amazon CloudWatch


AWS Auto Scaling


AWS Key Management Service


AWS Cost Explorer


AWS Config


AWS Identity and Access Management

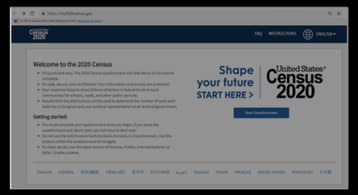

Amazon GuardDuty

T-REX

App layer: AWS GovCloud (US)

Service Identification

High availability service to US public
- Minimal impact to respondent during failure



- Multiple language support
- Responsive design
- 508 Compliant

Design

Resiliency

1. ***Multiple software apps*** (different tech) for single business function
2. ***Redundancy:*** AWS GovCloud (US) Regions; multiple Availability Zones
3. ***Data replication across Regions*** (Census ID, Session Information)
4. ***Native AWS resiliency features*** (Amazon DynamoDB, AWS Transit Gateway, AWS Lambda, Multi-AZ AWS RDS)

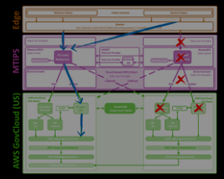
Scaling

1. ***Primary ISR:*** AWS Auto Scaling
2. ***Secondary ISR:*** Cluster-based scaling
3. ***Both apps:*** "Contain and Sustain" strategy

Service Transition

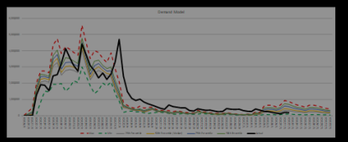
Resiliency

1. ***Execution / practice*** of failover procedures
2. ***Production Validation*** conducted before go-live



Scaling

1. ***P&S tested to support 1.2M*** concurrent users in each region
2. ***Determined breaking point*** by simulating 2.4M concurrent users



Operation

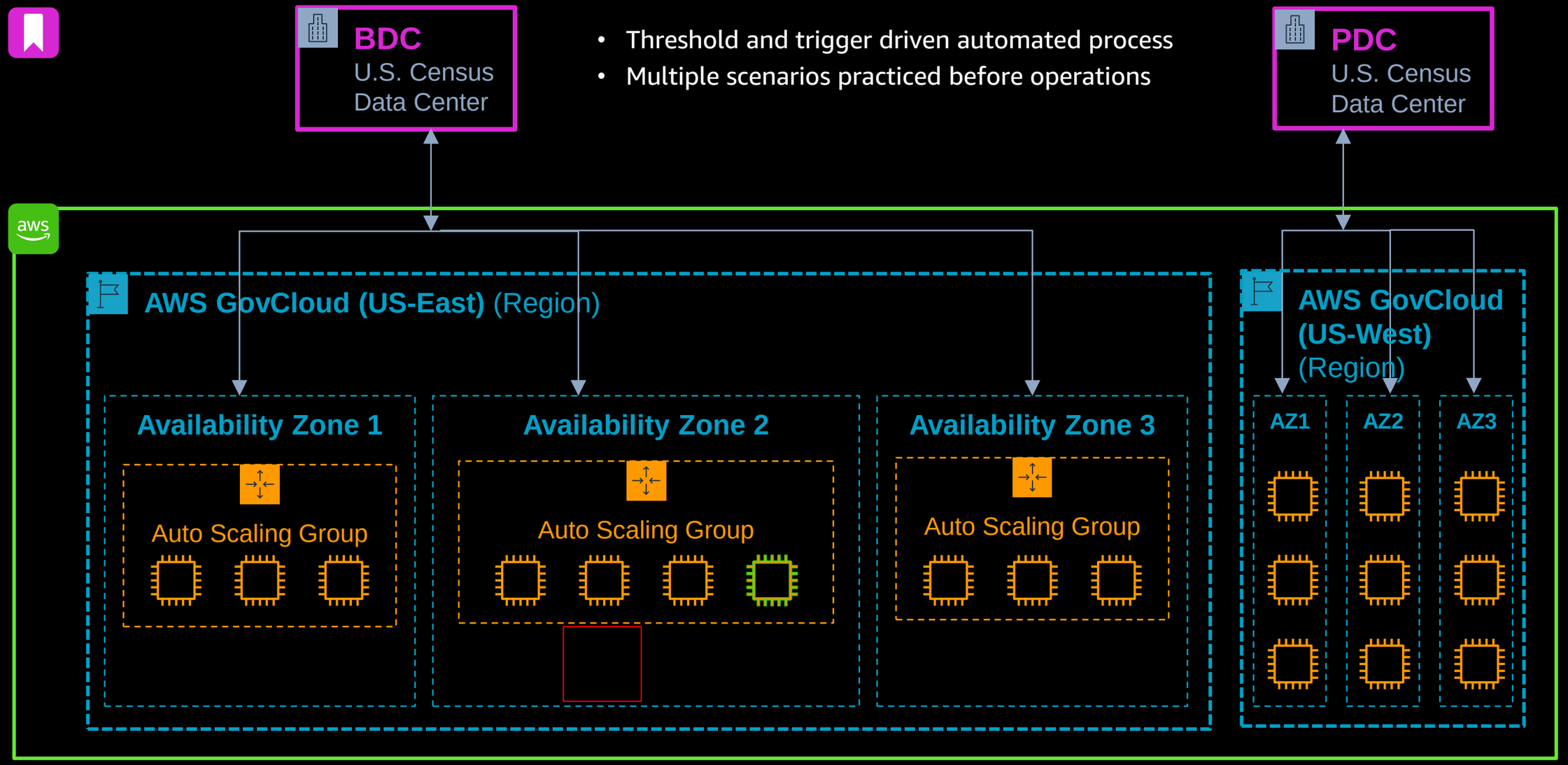
1. **Monitoring each service component**
2. **Splunk ITSI** as primary aggregator of system metrics
3. **Advanced monitoring dashboards** developed w/key stakeholders
   - Dashboard dry run during production validation
   - Highly relevant/well understood KPIs:



4. **Examples of resiliency/ scaling during ISR operations**
   - Network Outage
   - Patching
   - Failover for CQA Call Centers

T-REX

ISR Primary Application: Contain & Sustain



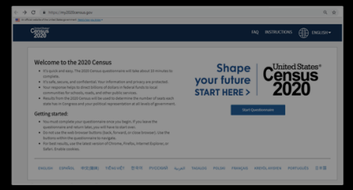
- Threshold and trigger driven automated process
- Multiple scenarios practiced before operations

T-REX

App layer: AWS GovCloud (US)

Service Identification

High availability service to US public
- Minimal impact to respondent during failure



- Multiple language support
- Responsive design
- 508 Compliant

Design

Resiliency

1. ***Multiple Software Apps*** (different tech) for single business function
2. ***Redundancy:*** AWS GovCloud Regions; multiple Availability Zones
3. ***Data Replication across Regions*** (Census ID, Session Information)
4. ***Native AWS resiliency features*** (DynamoDB, Transit Gateway, Lambda, Multi-AZ RDS)

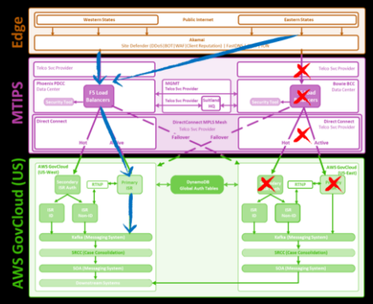
Scaling

1. ***Primary ISR:*** AWS Auto Scaling
2. ***Secondary ISR:*** Cluster based scaling
3. ***Both Apps:*** "Contain & Sustain" strategy

Service Transition

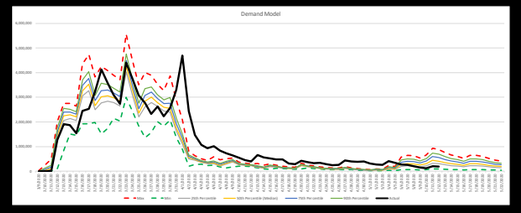
Resiliency

1. ***Execution / practice*** of failover procedures
2. ***Production validation*** conducted before go-live



Scaling

1. ***Performance & Scalability (P&S) tested to support 1.2M*** concurrent users in each region
2. ***Determined breaking point*** by simulating 2.4M concurrent users



Operation

1. **Monitoring each service component**
2. **Splunk ITSI** as primary aggregator of system metrics
3. **Advanced monitoring dashboards** developed w/key stakeholders
   - Dashboard dry run during production validation
   - Highly relevant/well understood KPIs:



4. **Examples of resiliency/ scaling during ISR operations**
   - Network Outage
   - Patching
   - Failover for CQA Call Centers

T-REX

Region failover

Scenario

A system failure in a single region prevents public access to the self-response instrument

Trigger

1. MTIPS circuit failure
2. F5 failure
3. AWS Direct Connect failure
4. Primary or Secondary ISR application failure

Execution

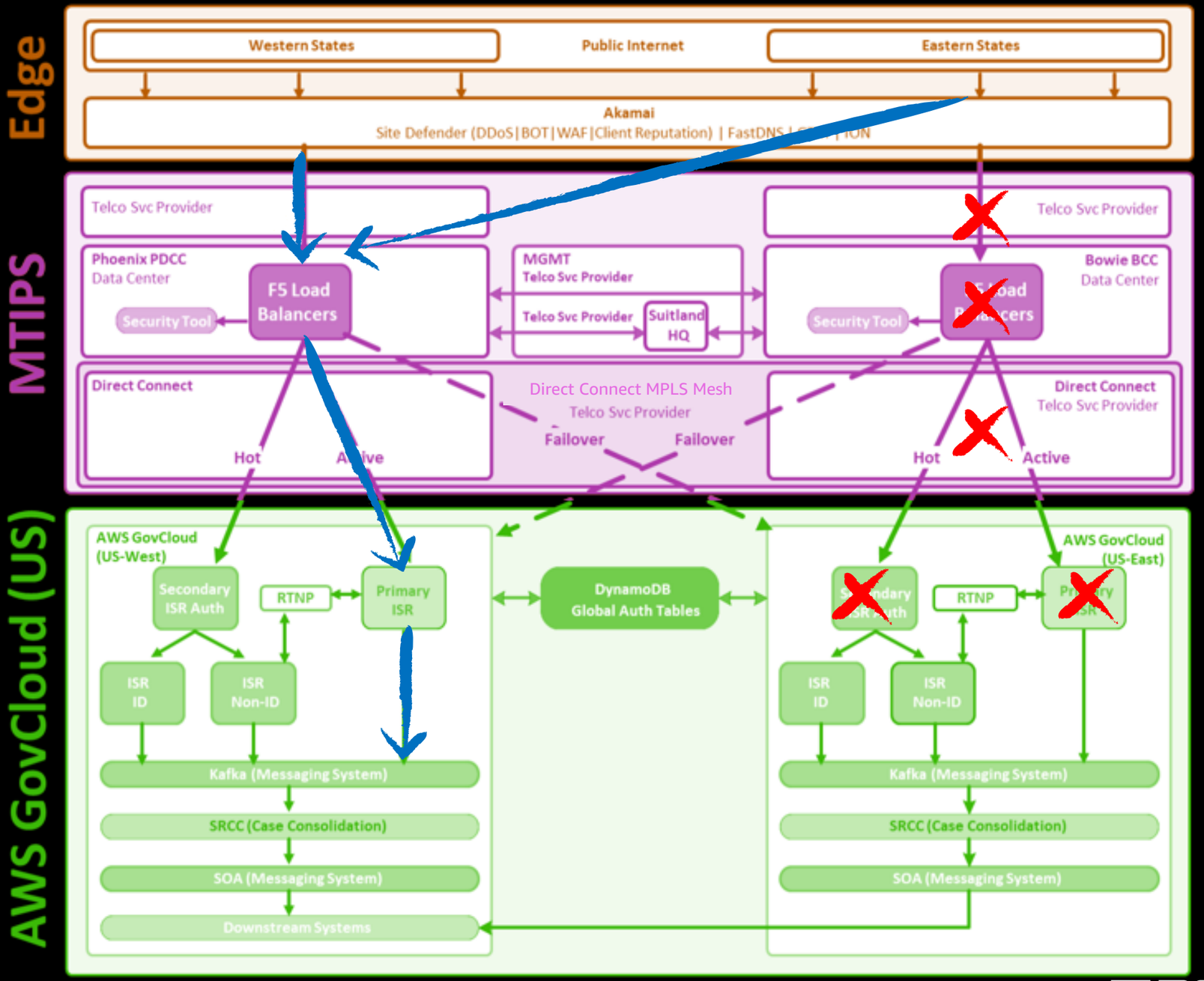
Automated: Edge detects failure
Manual: DNS update

Impact

Global Tables keeps active sessions in sync, No impacted to users

Duration

Sub second



App Failover

Scenario

Primary App has experienced a non-recoverable incident

Trigger

1. Primary application is unable to support the operation
2. Census leadership deems failover necessary

Execution

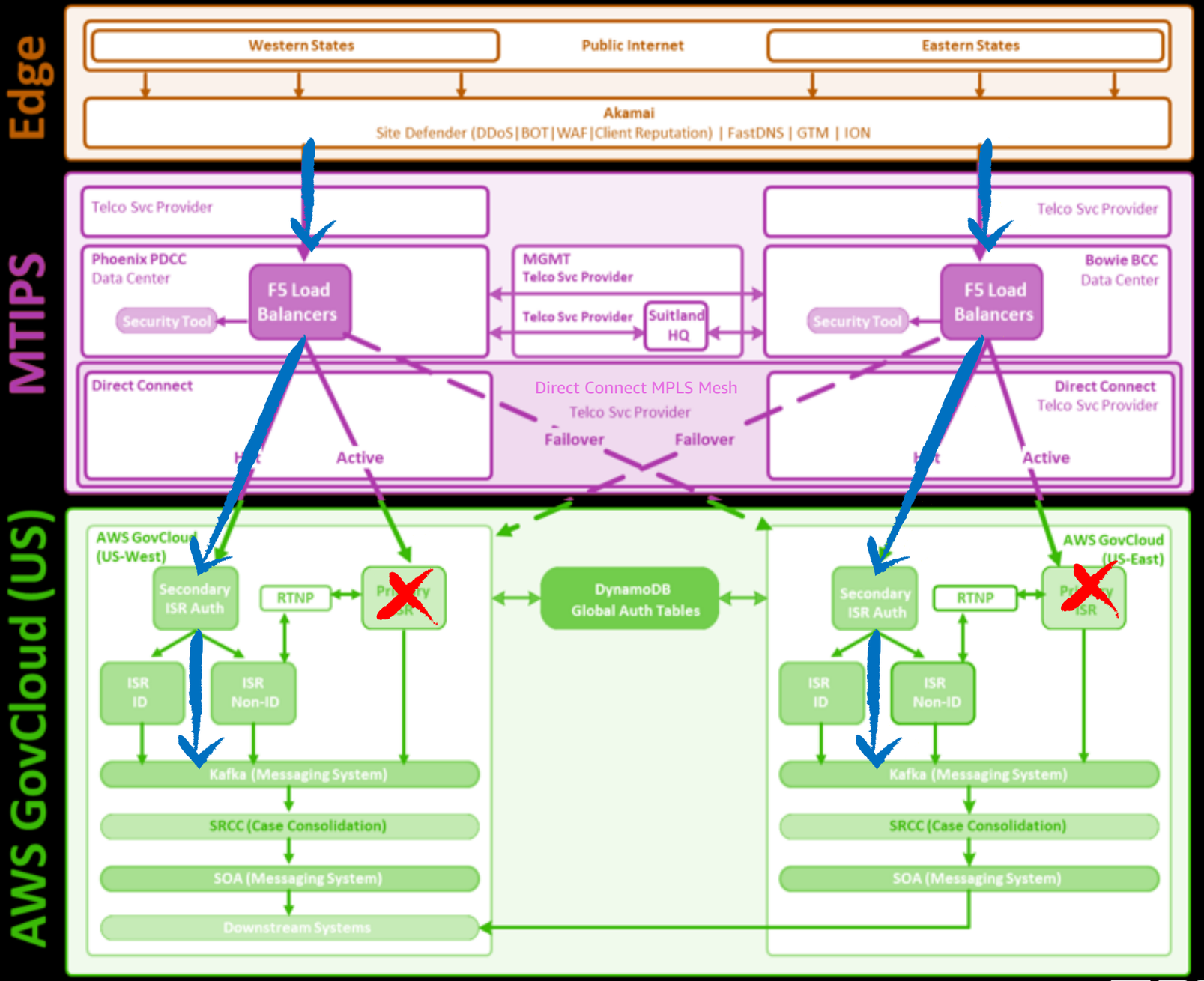
Edge: Move 'my2020census.gov' hostname to Secondary ISR Property

Impact

Active sessions will break, forcing survey restart

Duration

30 minutes



Complex failover

Scenario

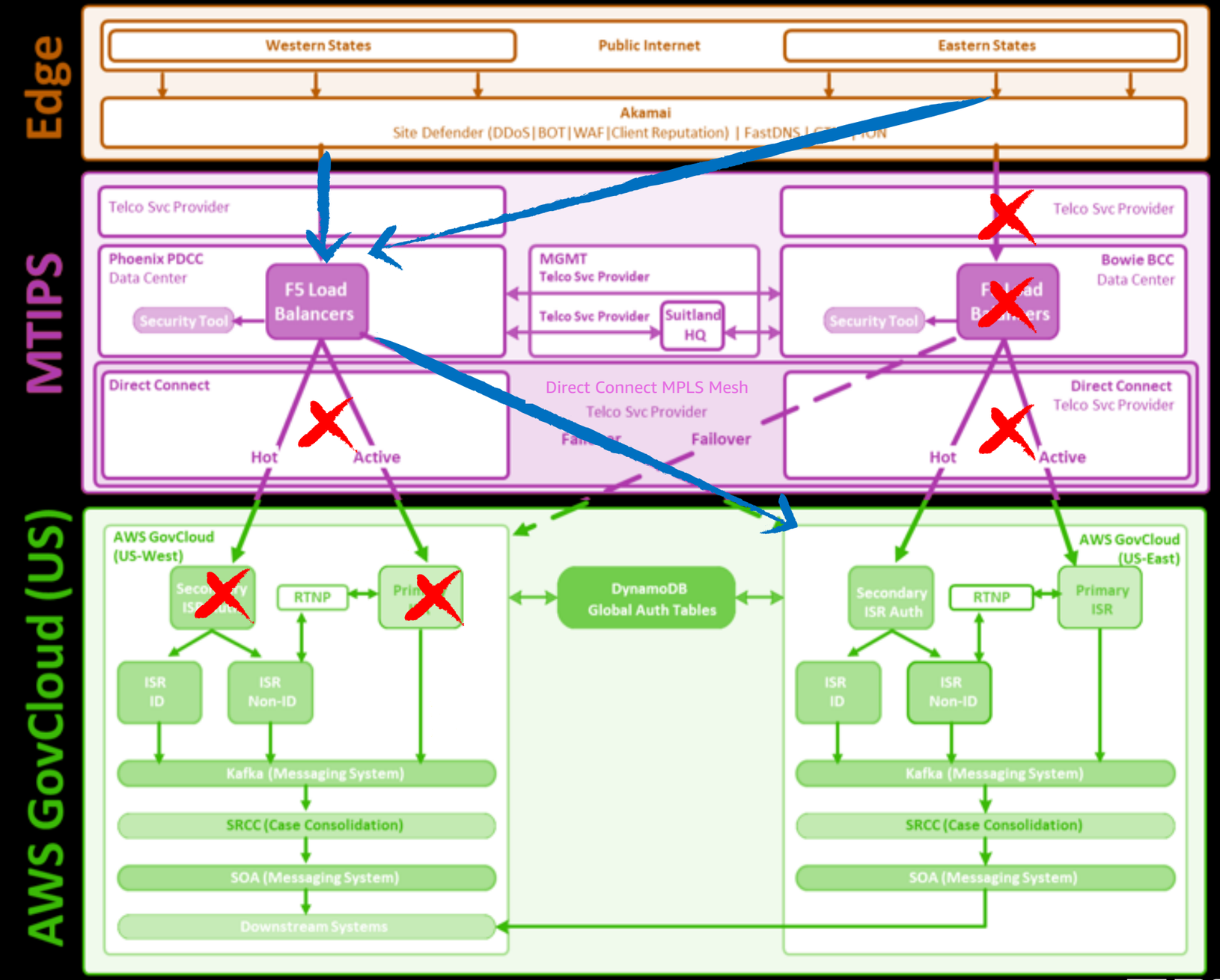
Concurrent failures in East and West Regions make it necessary to cross-route traffic

Trigger

1. System failures in both East and West Regions

Execution

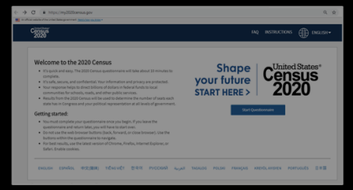
1. ***Automated:*** Edge directs traffic to alternate origin
2. ***F5 member pools updated*** to send to alternate region



App layer: AWS GovCloud (US)

Service Identification | **Design** | **Service Transition** | **Operation**

High availability service to US public
- Minimal impact to respondent during failure



- Multiple language support
- Responsive design
- 508 Compliant

Resiliency

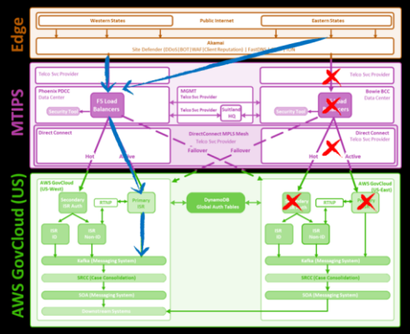
1. ***Multiple Software Apps*** (different tech) for single business function
2. ***Redundancy:*** AWS GovCloud Regions; multiple Availability Zones
3. ***Data Replication across Regions*** (Census ID, Session Information)
4. ***Native AWS resiliency features*** (DynamoDB, Transit Gateway, Lambda, Multi-AZ RDS)

Scaling

1. ***Primary ISR:*** AWS Auto Scaling
2. ***Secondary ISR:*** Cluster based scaling
3. ***Both Apps:*** "Contain & Sustain" strategy

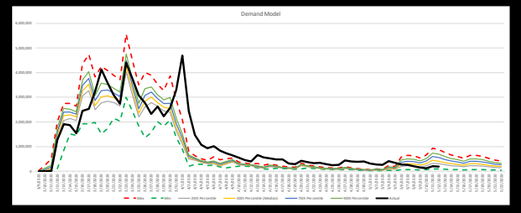
Resiliency

1. ***Execution / practice*** of failover procedures
2. ***Production validation*** conducted before go-live



Scaling

1. ***P&S tested to support 1.2M*** concurrent users in each region
2. ***Determined breaking point*** by simulating 2.4M concurrent users



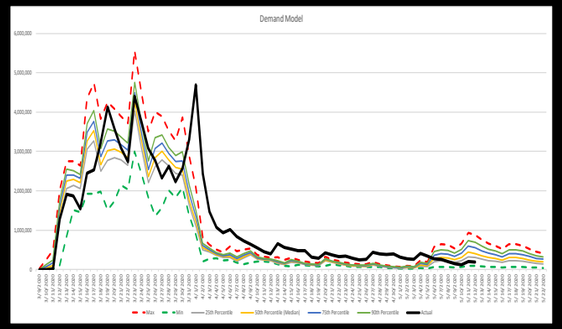
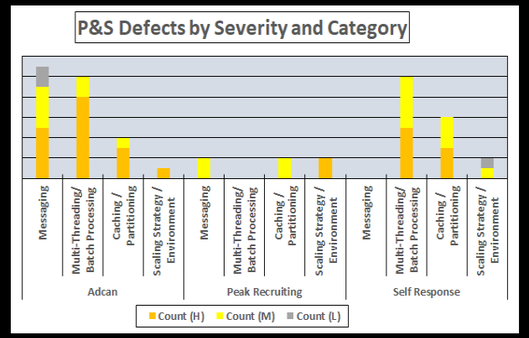
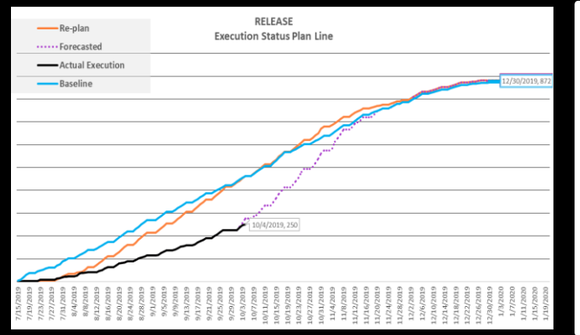
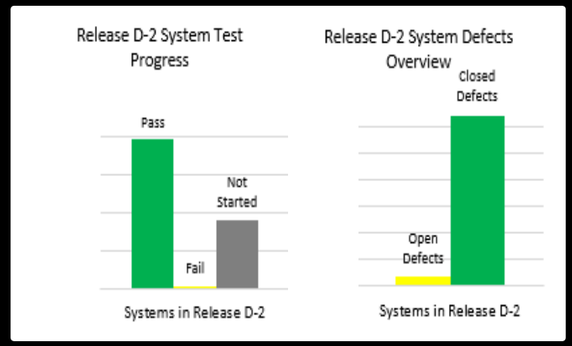
1. **Monitoring each service component**
2. **Splunk ITSI** as primary aggregator of system metrics
3. **Advanced monitoring dashboards** developed w/key stakeholders
   - Dashboard dry run during production validation
   - Highly relevant/well understood KPIs:



4. **Examples of resiliency/ scaling during ISR operations**
   - Network Outage
   - Patching
   - Failover for CQA Call Centers

T-REX

Extensive performance and scalability testing

	Phase 1	Phase 2	Phase 3	Phase 4
	• Architecture and design reviews	• Baseline performance tests • Isolated performance tests	• Integrated performance tests • Business thread end-to-end tests	• Chaos tests • Failover tests
				
AWS Cloud	• **Architecture assessment by AWS Team** • AWS Services to be used • System Infrastructure initial sizing – Servers, Storage, Bandwidth	Updated / Refined Sizing • Resource utilization • Engineering error / buffer • Error in model assumptions • Initial impact of varied input	• Scaling Plan – Triggers, thresholds, auto scale limits • **Discussion with AWS Service Teams**	• Finalized Sizing • Capacity Reservations across AZ / Regions • Contain and Sustain Capacity
Issues discovered	High-level system issues	Low-level system issues	Low-level integration issues	Issues that manifest over time

High confidence going into ISR operation for meeting scalability goals

T-REX

App layer: AWS GovCloud (US)

Service Identification | Design | Service Transition | Operation

High availability service to US public
- Minimal impact to respondent during failure



- Multiple language support
- Responsive design
- 508 Compliant

Resiliency

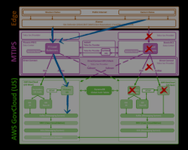
1. *Multiple Software Apps* (different tech) for single business function
2. *Redundancy:* AWS GovCloud Regions; multiple Availability Zones
3. *Data Replication across Regions* (Census ID, Session Information)
4. *Native AWS resiliency features* (DynamoDB, Transit Gateway, Lambda, Multi-AZ RDS)

Scaling

1. *Primary ISR:* AWS Auto Scaling
2. *Secondary ISR:* Cluster based scaling
3. *Both Apps:* "Contain & Sustain" strategy

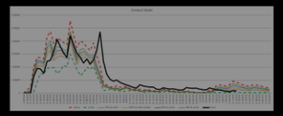
Resiliency

1. *Execution / practice* of failover procedures
2. *Production Validation* conducted before go-live



Scaling

1. *P&S tested to support 1.2M* concurrent users in each region
2. *Determined breaking point* by simulating 2.4M concurrent users

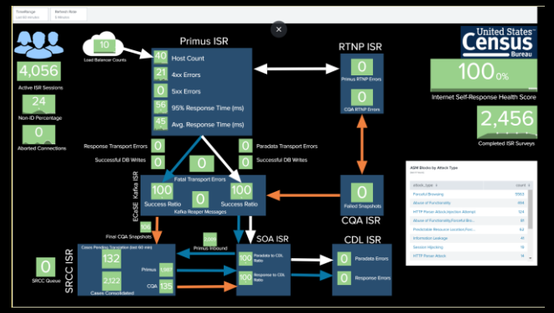


1. Monitoring each service component

2. Splunk ITSI as primary aggregator of system metrics

3. Advanced monitoring dashboards developed w/key stakeholders

- Dashboard dry run during production validation
- Highly relevant / well understood KPIs:

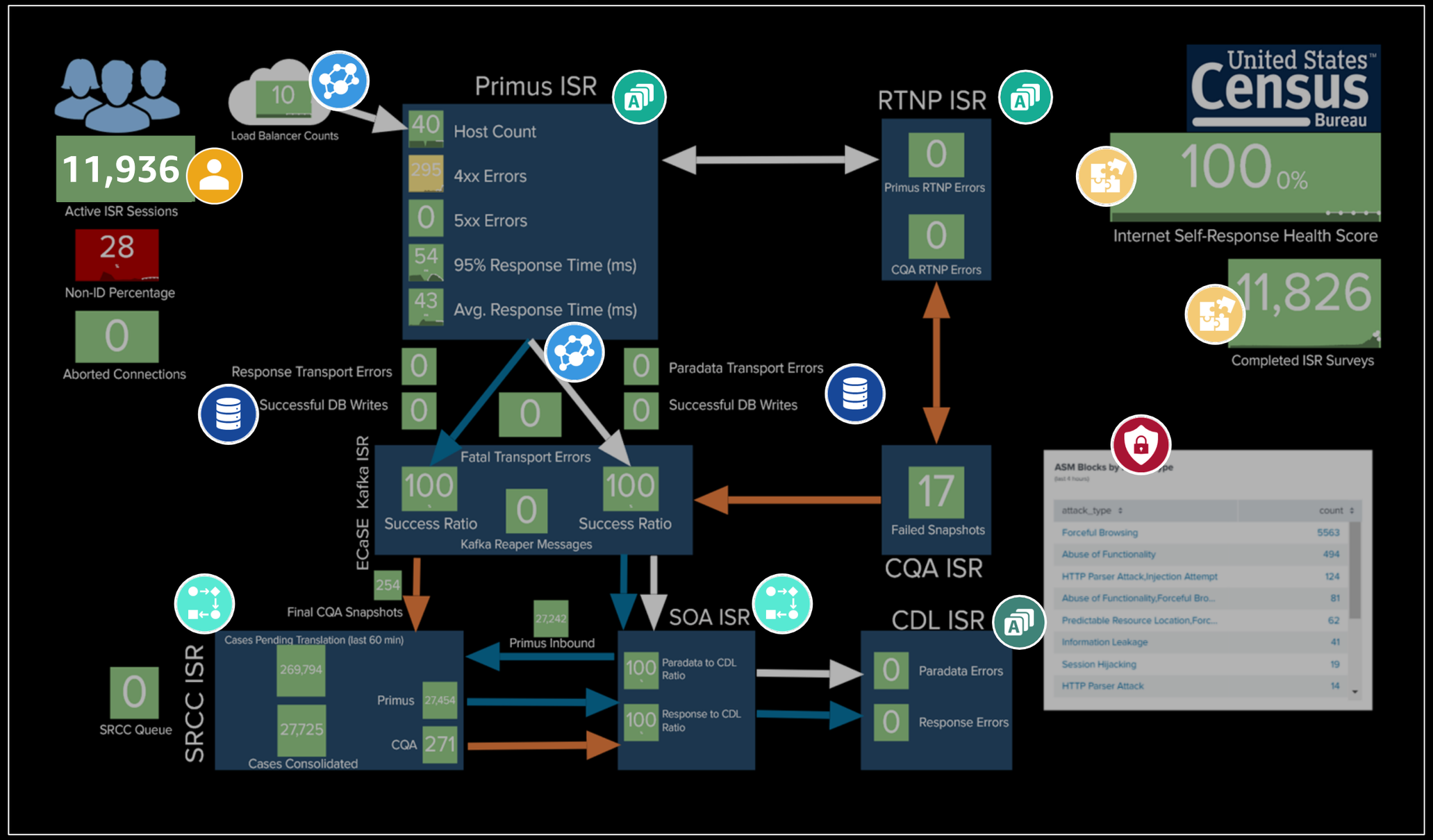


4. Examples of resiliency / scaling during ISR operations

- Network Outage
- Patching
- Failover for CQA Call Centers

T-REX

Internet Self-Response Dashboard



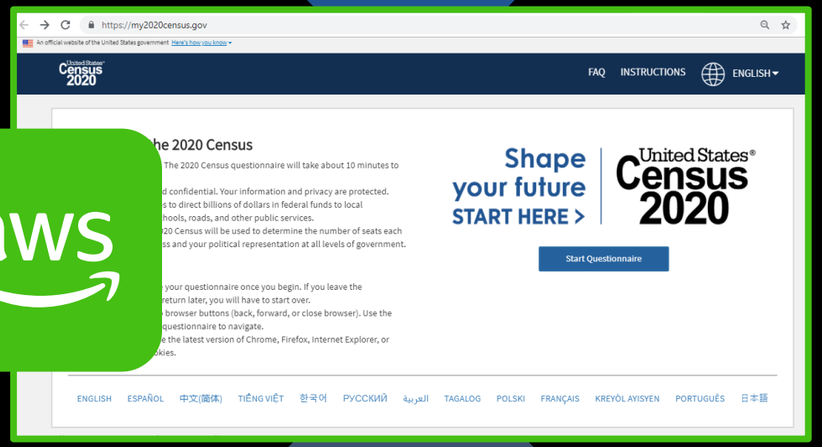
Areas of visibility

- Apps
- Active users
- Networking
- Data flow
- Security
- Database
- Operational context

T-REX

Successful delivery of 2020 U.S. Census by T-Rex

United States® Census Bureau



aws

T-REX

Key accomplishments

1. **COVID resilience** Collected 53% of responses via the internet during COVID-19, helping meet the critical mission of the U.S. Census of counting everyone
2. **18 months to delivery** Accelerated technology adoption, established and held end-to-end production testing and secured Census Decennial 2020 cloud in less than 18 months
3. **Relentless automation** Modeled workloads, leveraged on-demand capacity with automated cost management, to establish a budget and trigger investigation when exceeded
4. **DevSecOps** Utilized automation and DevSecOps approach to maintain operations achieving zero unplanned downtime on the Internet Response System
5. **ACD integration** Leveraged cloud-native services integrated with an experienced active cyber defense solution (using Threat Modeling) to ensure a secure movement of mission critical systems to the cloud

T-REX

2020 U.S. Census key success factors

- Achieved a National Enumeration Rate of **99.98%**
- Automation and increased use of technology like enumerator use of iPhones for case routing optimization, which contributed **to 90% increased enumerator productivity**
- **Online training:** new technology adapted for work in COVID environment
- **Data collection completed** in a challenging environment involving a global pandemic, wildfires, hurricanes, and civil unrest
- **79M responses** collected via Internet

"Despite the pandemic, the 2020 Census **self-response has been a tremendous success**…

… Our response system **has not had a single minute of downtime** since we first invited people to respond online beginning in March."

Dr. Steven Dillingham
Director; U.S. Census Bureau before
House Committee on Oversight and Reform



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!



AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.