

# 2020 Systems Readiness

**Michael Thieme**

Assistant Director for Decennial Census Programs, Systems and Contracts



U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU  
*census.gov*



# 2020 Census Systems Readiness

## Agenda

- Update on 2018 End-to-End Census Test
- Update on 2020 Census
- Update on Performance & Scalability Testing

2020 Census Systems Readiness  
2018 End-to-End Census Test  
Key Systems Supporting Ongoing Operation

Operation (Release)	Key Systems Supporting Operation	Conduct Operation Date
Tabulation, Product Creation, and Dissemination	CaRDS, CDL, CEDSCI, DRPS, MAF/TIGER, PEARSIS, Tabulation, SOA	January 7, 2019



# 2020 Census Systems Readiness

## 2020 Census Operational Deliveries

2020 CENSUS OPERATIONAL DELIVERY		Test Readiness Review	Production Readiness Review	Conduct Operation
D1	2020 Census Early Ops Preparation	07/23/2018	10/19/2018	10/23/2018
D2	Address Canvassing	01/04/2019	05/17/2019	08/19/2019
D3	Peak Recruiting & Hiring	01/18/2019	06/07/2019	09/03/2019
D4	Self-Response	06/28/2019	12/13/2019	03/09/2020
D5	Integrated Partnership and Communications (IPC)	05/17/2019	11/01/2019	01/15/2020
D6	Count Review (CR) & Remote Alaska (RA)	05/10/2019	10/25/2019	01/21/2020
D7	Group Quarters Advance Contact (GQAC) & ETL Advance Contact (ETL AC) & Federally Affiliated Count Overseas (FACO)	05/24/2019	11/08/2019	02/03/2020
D8	GQ Enumeration (GQE) & Service Based Enumeration (SBE) & Enumeration at Transitory Locations (ETL)	06/17/2019	01/24/2020	03/30/2020
D9	Update Enumerate (UE) & Update Leave (UL)	07/12/2019	01/03/2020	03/16/2020
D10	Nonresponse Followup (NRFU)	07/19/2019	01/17/2020	04/09/2020
D11	Archiving & Count Question Resolution (CQR)	09/18/2020	03/05/2021	06/01/2021
D12	Response Processing	04/27/2020	06/17/2020	09/15/2020
D13	Data Products / Dissemination	06/15/2020	10/05/2020	11/25/2020
D14	Redistricting	08/28/2020	12/18/2020	02/17/2021
D15	Post Enumeration Survey (PES)	03/05/2021	05/28/2021	07/23/2021
D16	Island Areas	10/16/2020	04/02/2021	06/30/2021

# 2020 Census Systems Readiness

## 2020 Census Operational Deliveries

20 systems are in production for Early Ops Preparation

2020 Website, ATAC, CBS, CDL, CENDOCs, CHEC, CHRIS, CRM, DAPPS, Desktop Services, DPACS, DSC, ECaSE-FLD-OCS, iCADE, IDMS, MAF/TIGER, MOJO Field Processing, R&A, SOA, UTS

8 more systems will go into production as part of Address Canvassing

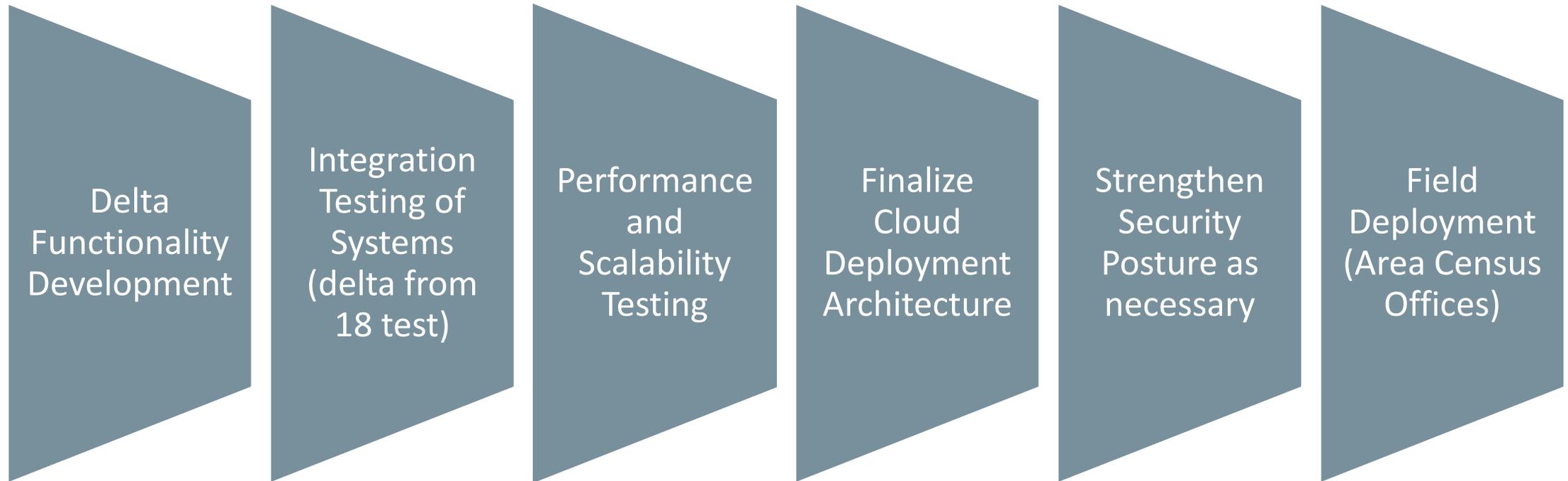
ECaSE-OCS, Geospatial Service, ILMS, LiMA, MCM, MOJO Optimizer, NPC Printing, SMarCS

17 systems for Peak Recruiting & Hiring

Same as Early Ops Prep minus ATAC, CRM, iCADE

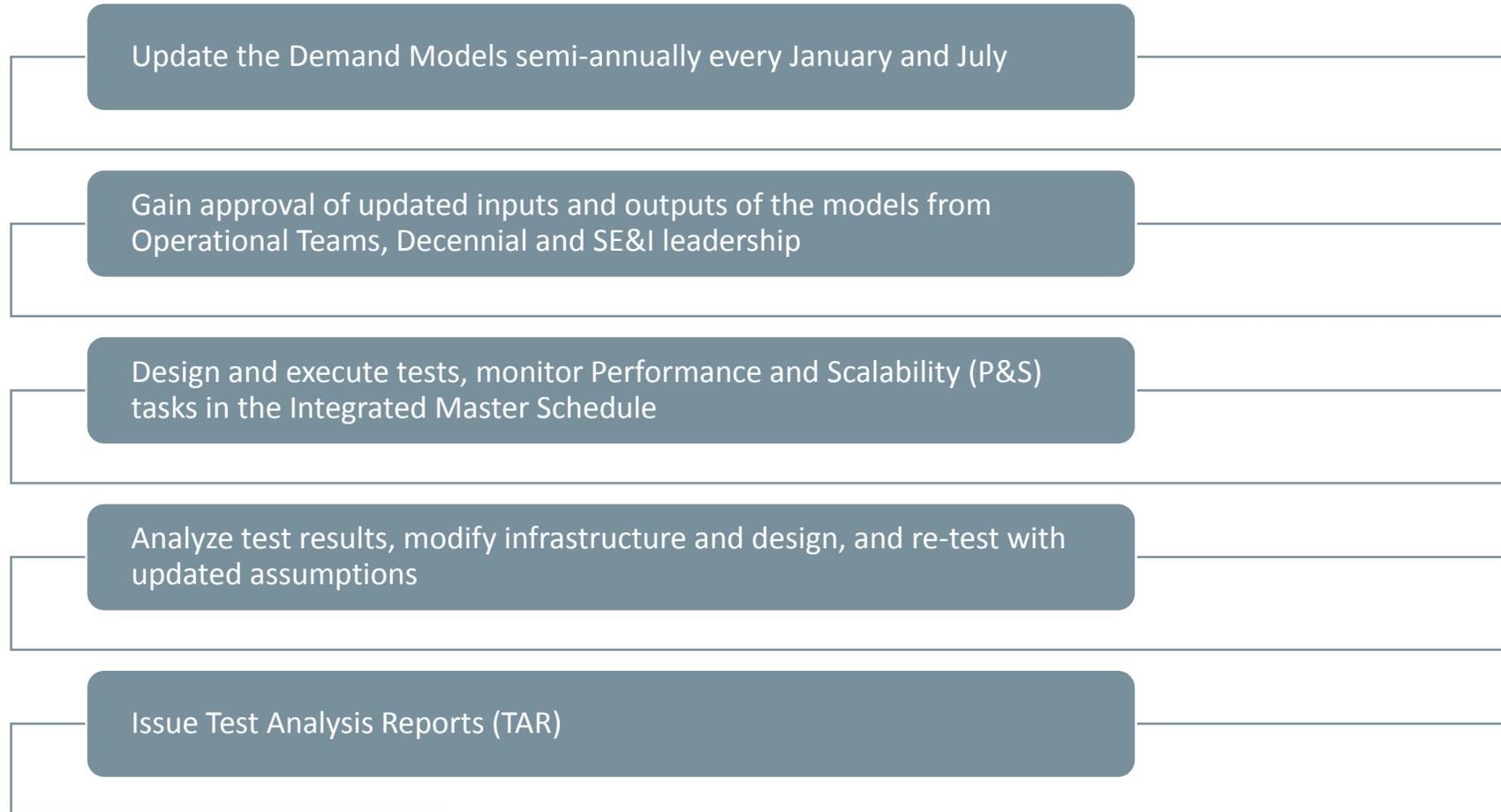
# 2020 Census Systems Readiness

## 2020 Census Systems Readiness – Remaining Key Work



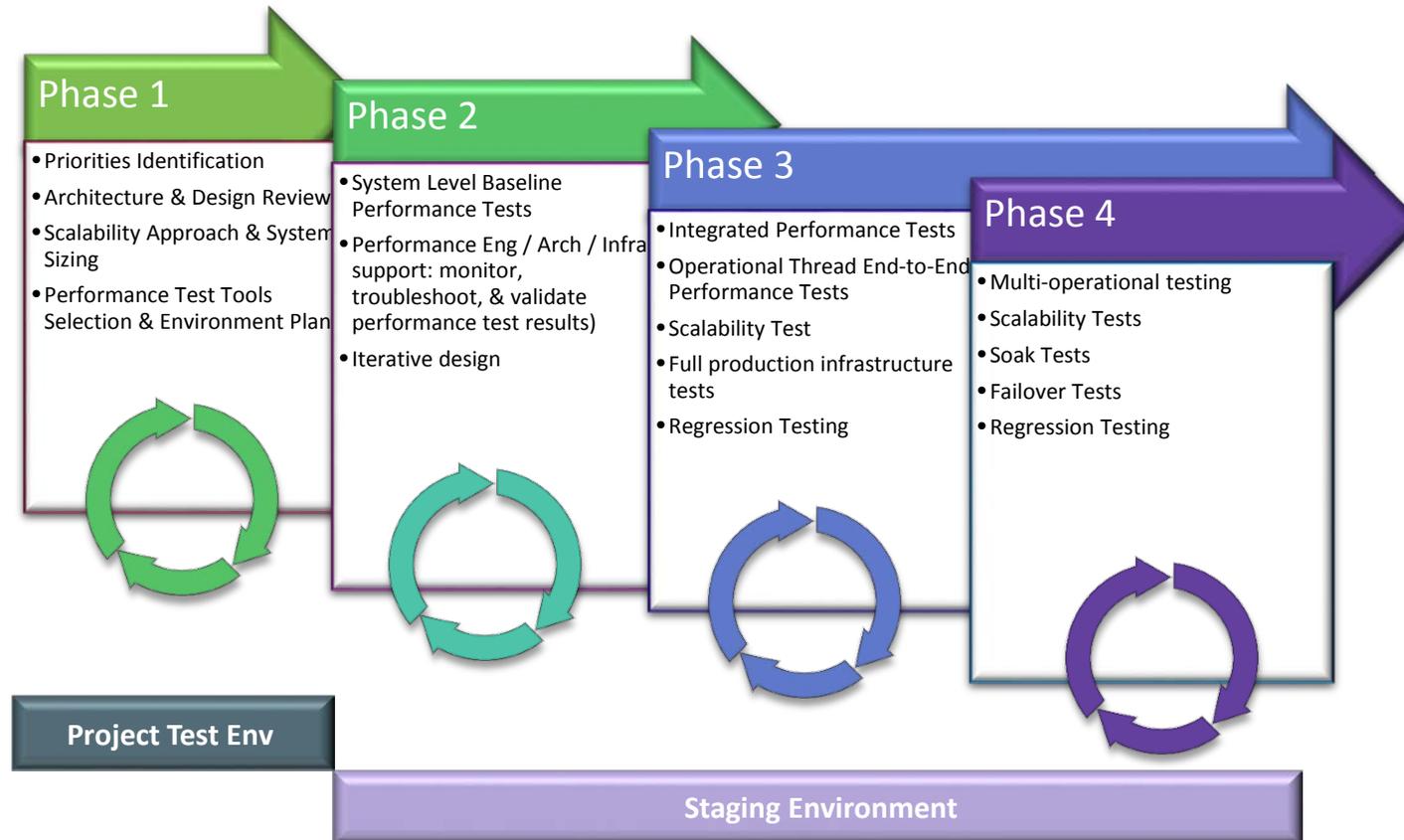
# 2020 Census Systems Readiness

## Key Performance and Scalability Efforts



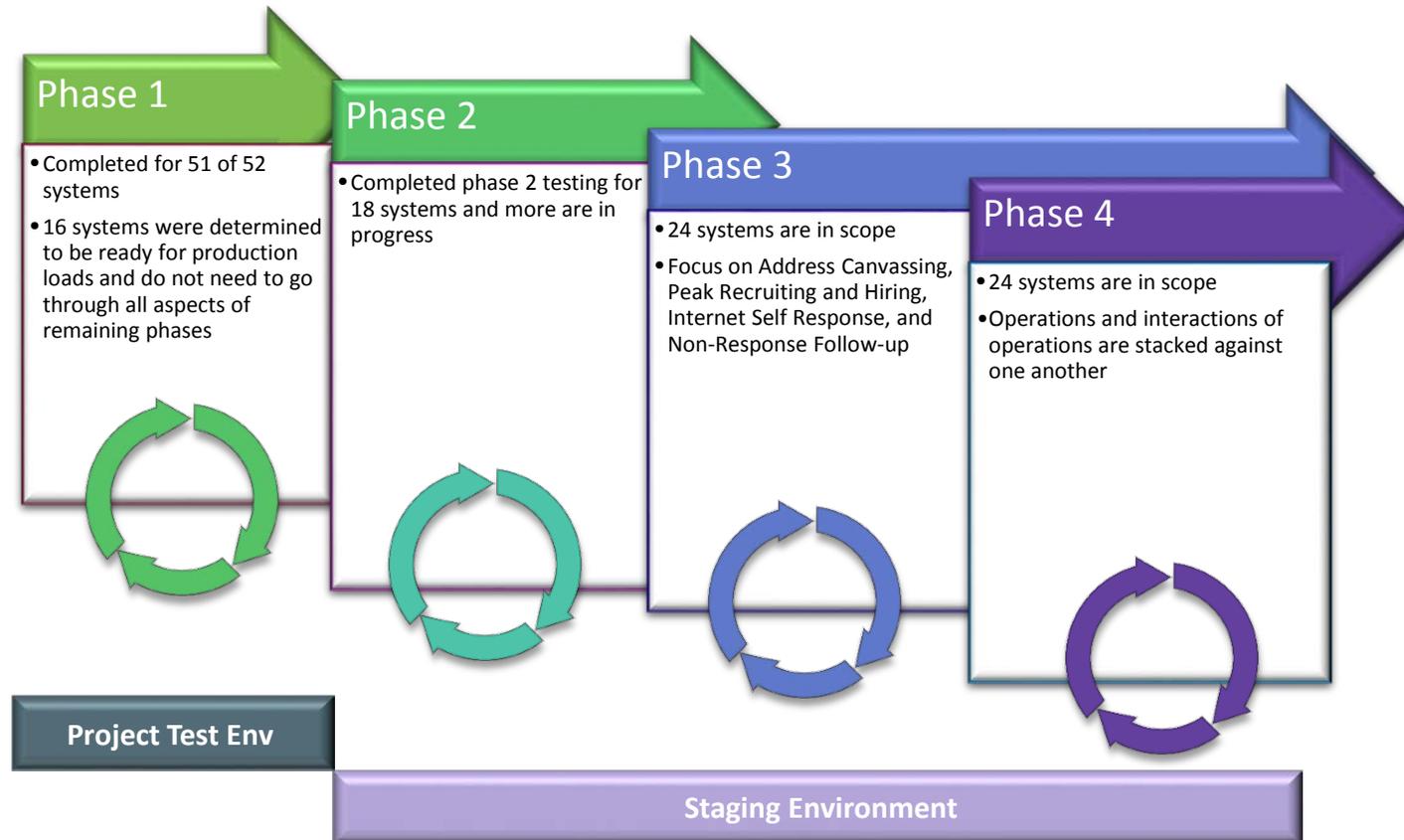
# 2020 Census Systems Readiness

## Performance and Scalability Readiness Plan



# 2020 Census Systems Readiness

## Performance and Scalability Readiness Status



# Performance & Scalability (P&S) Testing

## Example: Internet Self-Response (ISR)

	Demand Model	System Design	Performance & Scalability (P&S) Testing	Resulting Actions																																							
<b>Description</b>	Projected peak concurrent users accessing the ISR application in 2020 is ~120K users	A single ISR computer cluster in the Amazon Gov Cloud can comfortably handle 50K users. Our initial ISR design provides five times the capacity of the demand model to handle up to 600K concurrent users.	<p><b>Phase 1:</b> Research compares design to industry benchmarks and best practices (paper exercise)</p> <p><b>Phase 2:</b> Tests and confirms capacity of design (cluster &amp; full capacity)</p> <p><b>Phase 3:</b> Tests loads with interfacing systems</p> <p><b>Phase 4:</b> Tests all 2020 system running together as they would during an operation</p>	<p><b>Phase 1:</b> Confirm or adjust models</p> <p><b>Phase 2:</b> Confirm or revise cluster/full capacity</p> <p><b>Phase 3:</b> Confirm or adjust number of clusters</p> <p><b>Phase 4:</b> Confirm or adjust number of clusters</p>																																							
<b>Math</b>	120K Concurrent users	Design contingency = Demand (120K) x 5 = 600K users Single AWS Cluster handles 50K users $600K / 50K = 12$ AWS Clusters	<p><b>Phase 1:</b> Research validates design</p> <p><b>Phase 2:</b> Isolated ISR handles x users</p> <p><b>Phase 3:</b> Testing indicates interfaces decrease concurrent users capacity to Y</p> <p><b>Phase 4:</b> Testing indicates concurrent operations decrease capacity to Z</p>	<table border="1"> <thead> <tr> <th rowspan="2">P&amp;S Phase</th> <th rowspan="2">Plan</th> <th colspan="2">Concurrent Users</th> <th rowspan="2">Pre-T&amp;S Test</th> <th colspan="2">AWS Clusters</th> </tr> <tr> <th>Actual</th> <th></th> <th>Post P&amp;S Test Indication</th> <th></th> </tr> </thead> <tbody> <tr> <td>1</td> <td>600,000</td> <td>600,000</td> <td></td> <td>12</td> <td>12.0</td> <td></td> </tr> <tr> <td>2</td> <td>50,000</td> <td>60,000</td> <td></td> <td>1</td> <td>0.8</td> <td></td> </tr> <tr> <td>3</td> <td>600,000</td> <td>540,000</td> <td></td> <td>12</td> <td>13.3</td> <td></td> </tr> <tr> <td>4</td> <td>600,000</td> <td>500,000</td> <td></td> <td>12</td> <td>14.4</td> <td></td> </tr> </tbody> </table> <p>(Next Slide)</p>	P&S Phase	Plan	Concurrent Users		Pre-T&S Test	AWS Clusters		Actual		Post P&S Test Indication		1	600,000	600,000		12	12.0		2	50,000	60,000		1	0.8		3	600,000	540,000		12	13.3		4	600,000	500,000		12	14.4	
P&S Phase	Plan	Concurrent Users		Pre-T&S Test			AWS Clusters																																				
		Actual			Post P&S Test Indication																																						
1	600,000	600,000		12	12.0																																						
2	50,000	60,000		1	0.8																																						
3	600,000	540,000		12	13.3																																						
4	600,000	500,000		12	14.4																																						

# Performance & Scalability (P&S) Testing

## Example: Internet Self-Response (ISR) (continued)

P&S Phase	# of Concurrent Users		# of AWS Clusters	
	Plan	Actual*	Pre-P&S Test	Post P&S Test Indication*
1	600,000	600,000	12	12.0
2	50,000 & 500,000	60,000 & 600,000 (X)	1 & 12	0.8 & 12.0
3	600,000	540,000 (Y)	12	13.3
4	600,000	500,000 (Z)	14	14.4

\*Data for demonstration purposes only. ISR results are still pending.

- Though the expectation is that this contingency number (600K concurrent users) will not be reached, a set of 6 pre-built clusters will be available to be turned on with little turnaround time once we hit a pre-determined threshold (~400K) of concurrent users.
- P&S Testing enables us to confidently configure cloud apps so they are capable of serving peak loads on day one. However, dynamic scaling will be available if we unexpectedly go beyond the contingency levels.

# QUESTIONS?

# Back Up Slides

# 2020 Census Systems Readiness

## 2020 Census Systems List

	System	Description
1	<b>2020 Website</b>	2020 website provides a public web interface and allows partners to publish content.
2	<b>ATAC (Automated Tracking and Control)</b>	ATAC provides customer, employee, and workflow management. It also provides outbound call tracking for Geographic Partnership Programs and material tracking and check-in.
3	<b>BARCA (Block Assessment, Research, and Classification Application)</b>	BARCA is an interactive review tool, which is designed to assist an analyst in assessing present, past, and future housing units for a set of geographic work units. It is used for in-office address canvassing.
4	<b>CAES (Concurrent Analysis and Estimation System)</b>	CAES is a high speed processing environment used to make fast decisions during the data collection phase of the Census.
5	<b>CaRDS (Control and Response Data System)</b>	CaRDS provides sample design and universe determination.
6	<b>CBS (Commerce Business System)</b>	CBS is an electronic financial tracking and approval system. CBS collects and reports labor hours and costs for activities that the National Processing Center (NPC) performs.
7	<b>CDL (Census Data Lake)</b>	CDL serves as the repository for paradata and response data. It is built on a distributed, scalable platform to support data ingest, storage, and access for reporting and analytics applications.
8	<b>CEDSCI (Center for Enterprise Dissemination Services and Consumer Innovation)</b>	CEDSCI is an enterprise dissemination platform with search capability and access to surveys and censuses data.
9	<b>CEM (Customer Experience Management)</b>	CEM is an enterprise platform, which is used to provide dashboards providing insight into customer-focused performance metrics.
10	<b>CENDOCS (Census Document System)</b>	CENDOCS is a web-based system for requesting forms design services, publications and graphics services, and printing services which includes CD-ROM replication.
11	<b>Centurion</b>	Centurion provides for the upload of GQ eResponse data, LUCA address updates, and FACO data.

# 2020 Census Systems Readiness

## 2020 Census Systems List

	System	Description
12	<b>CHEC (Census Hiring and Employment Check System)</b>	CHEC is an administrative system that automates clearance processing of all personnel at the Census Bureau Headquarters, the Bureau of Economic Analysis (BEA), the Regional Offices (ROs), the National Processing Center (NPC), and Computer Assisted Telephone Interview (CATI) sites. This system supports fingerprint processing with the Federal Bureau of Investigation (FBI), the Office of Personnel Management (OPM), the Department of Homeland Security (DHS), the Office of Management and Budget (OMB), and the Department of Commerce (DOC).
13	<b>CHRIS (Census Human Resources Information System)</b>	CHRIS is a web-based personal information tool which provides personnel and payroll information on desktops. It creates a contact file for the Emergency Notification System (ENS).
14	<b>CIRA (Census Image Retrieval Application)</b>	CIRA stores decennial census images as well as raw and edited data. It provides a user interface for reviewing the images and data and is used for age match research at NPC.
15	<b>CQA (Census Questionnaire Assistance)</b>	CQA is a system that collects response data through inbound respondent-initiated calls, assists respondents with completing their questionnaires, including capturing their responses over the phone, and answering questions. CQA also provides outbound calling and data capture for reconciliation of response inconsistencies and follow-up QC operations.
16	<b>CRM (Customer Relationship Management)</b>	CRM is a suite of applications that provide the management of partnership activities, and case management for customer service.
17	<b>DAPPS (Decennial Applicant, Personnel and Payroll Systems)</b>	DAPPS is the system of record for the C-SHaRPS system of systems. It supports selection, hiring, personnel actions, time and expense processing, and payroll administration for temporary Census Bureau employees.
18	<b>Desktop Services</b>	Desktop Services is a set of office automation applications and tools for field office staff and remote workers.
19	<b>DMP (Data Management Platform)</b>	DMP provides centralized data analytics service to support decision-making for the Census advertising campaign.
20	<b>DRPS (Decennial Response Processing System)</b>	DRPS provides autocoding, clerical coding, data editing, and imputation for the post-data-collection response processing.
21	<b>DPACS (Decennial Physical Access Control System )</b>	DPACS maintains information relating to physical access control to facilities.

# 2020 Census Systems Readiness

## 2020 Census Systems List

	System	Description
22	<b>DSC (Decennial Service Center)</b>	DSC is the help desk service for field representatives.
23	<b>ECaSE Enum (Enterprise Censuses and Surveys Enabling Platform – Enumeration)</b>	ECaSE Enum supports field enumeration, provides the employee availability, and time & expense capabilities.
24	<b>ECaSE FLD OCS (Enterprise Censuses and Surveys Enabling Platform – Field Operational Control System)</b>	ECaSE FLD OCS manages field assignments and leverages the Optimizer for routing field staff to and between assignments.
25	<b>ECaSE ISR (Enterprise Censuses and Surveys Enabling Platform – Internet Self-Response)</b>	ECaSE ISR supports self-response data collection via the Internet. It also supports data collection by call center agents.
26	<b>ECaSE OCS (Enterprise Censuses and Surveys Enabling Platform – Operational Control System)</b>	ECaSE OCS manages the data collection case history for all enumeration operations and maintains operational workloads as data collection proceeds.
27	<b>Geospatial Services</b>	Geospatial Services provide mapping and imagery services.
28	<b>GUPS (Geographic Update Partnership Software)</b>	GUPS allows partners to provide geographic data updates.
29	<b>iCADE (Integrated Computer Assisted Data Entry)</b>	iCADE captures and tracks respondent information from scanned paper questionnaires. The iCADE system processes digital images of respondent questionnaires through three functions. Optical Mark Recognition (OMR) automatically detects and captures checkmark responses. Optical Character Recognition (OCR) automatically captures the presence of pre-selected numeric answer fields. Key From Image (KFI) directs keyers to all pre-identified, non-OCR write-in answer zones for keying hand-written data.
30	<b>IDMS (Identity Management System)</b>	IDMS provides enterprise authentication, authorization, identity, and access management.

# 2020 Census Systems Readiness

## 2020 Census Systems List

	System	Description
31	<b>ILMS (Integrated Logistics Management System)</b>	ILMS is used to manage logistics and resource planning.
32	<b>IPTS (Intelligent Postal Tracking System)</b>	IPTS is a mail tracking system that traces individual mail pieces through the USPS mailstream.
33	<b>LiMA (Listing and Mapping Application)</b>	LiMA provides address and map feature collection in support of surveys and censuses.
34	<b>MaCS (Matching and Coding Software)</b>	MaCS provides interactive clerical matching and coding. This application supports Non-ID processing.
35	<b>MAF/TIGER (Master Address File/Topologically Integrated Geographic Encoding and Referencing Database)</b>	MAF/TIGER is a repository of spatial and address data and is used to provide spatial and address products and services in support of survey and census operations.
36	<b>MCM (Mobile Case Management)</b>	MCM provides mobile device-level survey case management.
37	<b>MOJO Optimizer/Modeling</b>	MOJO Optimizer/Modeling is a route optimizer for field case assignments.
38	<b>MOJO Field Processing</b>	MOJO Field Processing provides a dashboard for recruiting and performance reports.
39	<b>NPC Printing (Printing at the National Processing Center)</b>	NPC Printing provides printing services.
40	<b>OneForm Designer Plus</b>	OneForm Designer Plus is a tool used to create paper forms including decennial questionnaires, letters, envelopes, notice of visit, language guides, and other decennial field and public materials.
41	<b>PEARSIS (Production Environment for Administrative Records, Staging, Integration, and Storage)</b>	PEARSIS manages administrative records and services associated with these records.

# 2020 Census Systems Readiness

## 2020 Census Systems List

	System	Description
42	<b>PES Clerical Match and Map Update (Post-Enumeration Survey - Clerical Matching System and Map Update)</b>	PES Clerical Match and Map Update provides clerical address and person matching, and residential unit location updates to support the Coverage Measurement operation.
43	<b>PES Imputation and Estimation (Post-Enumeration Survey - Imputation and Estimation System)</b>	PES Imputation and Estimation provides coding, imputation of responses, and estimation of coverage for the Coverage Measurement operation.
44	<b>PES PCS (Post-Enumeration Survey - Processing and Control System)</b>	PES PCS provides sampling, automated matching, and workload control for the Coverage Measurement operation.
45	<b>R&amp;A (Recruiting and Assessment)</b>	R&A provides online job application capability for temporary positions, online training, and applicant pre-selection assessments.
46	<b>RTNP (Real Time Non-ID Processing)</b>	RTNP standardizes, matches, and geocodes addresses in real-time.
47	<b>SMaRCS (Sampling, Matching, Reviewing, and Coding System)</b>	SMaRCS applies quality control algorithms to the field address canvassing and enumeration operations.
49	<b>SOA (Service Oriented Architecture)</b>	SOA, along with the enterprise service bus, is a web service infrastructure that facilitates data transport, communication, and integration of systems.
49	<b>SRQA (Self-Response Quality Assurance)</b>	SRQA identifies fraudulent cases of self-response using analytical models.
50	<b>Tabulation (Decennial Tabulation System)</b>	Tabulation produces tabulated response data for redistricting and other purposes.
51	<b>UTS (Unified Tracking System)</b>	UTS provides operational and cost & progress reports.
52	<b>WebTQA (Web Telephone Questionnaire Assistance)</b>	WebTQA provides telephone assistance for Geographic Partnership Programs.