2020 Census Enterprise Architecture and Infrastructure Transition Plan

A New Design for the 21st Century

Issued 4/13/2016
Version 1.0
Prepared by Quyen Nguyen, Scott Fifield, Christopher Carrino, Rainier Suazo-Munoz, Brock Webb, and Curtis Broadway

United States™ Census Bureau

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

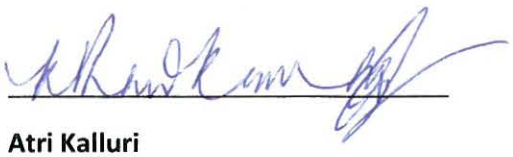
United States™ Census 2020

Table of Contents

Figures

Tables

Approval Signatures



4/8/16

Atri Kalluri

Date Signed

Chief, Decennial Information Technology Division

4/12/16

Lisa Blumerman

Date Signed

Associate Director for Decennial Census Programs

Executive Summary

The 2020 Census Enterprise Architecture and Infrastructure Transition Plan (CEAITP) describes the multi-year plan for the Census Information Technology (IT) evolution from the current 2015 Solution Architecture to the 2020 Target Solution Architecture that will be capable to support the 2020 Census operations. By providing an integrated view and actionable plan of the IT activities and activities milestones, the CEAITP helps to realize the desired Solution Architecture that can fulfill the business goals of achieving cost efficient Census via modern technology, and implementing innovative Census operations as laid out in the 2020 Census Operation Plan for:

- Reengineering Address Canvassing;
- Optimizing Self-Response;
- Utilizing Administrative Records and Third-Party Data; and
- Reengineering Field Operations.

The purpose of the 2020 CEAITP is to communicate and inform the transition phases of all IT solutions comprising the 2020 Census Architecture to the stakeholders, including external oversight bodies, and internal US Census Bureau Directorates having key roles in the 2020 Census, namely, Decennial, Field, and IT Directorates. The 2020 CEAITP is designed to support the 2020 Operational Plan. With a phased approach aligned with the Census Tests, the 2020 CEAITP aims to ensure that the 2020 Census Program is not at risk of having to integrate solutions too late in the planning cycle. The sequencing timeline in the 2020 CEAITP shows that by the start of the 2018 End-to-End test, the majority of the business capabilities and systems must be in place. In addition, the Solution Architecture with its systems and services must be designed to meet the requirements for scalability, reliability, and availability.

In order to achieve its purpose, the 2020 CEAITP is developed according to three strategies:

- First, the transition process is incremental by nature, as the Solution evolves from the current to target state. The Census Tests will demonstrate the progress of the implementation of the Business operations, and the capabilities provided by the Applications, and the IT Infrastructure.
- Second, the sequencing of the 2020 CEAITP has timelines for the architecture domains - Business, Application, Information, Technical, Security, and Quality- , and consists of evolving these domains from the current state to the target state. The technical quality attributes of scalability, availability, reliability, resilience, and security are factored in the evolution of the 2020 Solution and propagated down to the levels of IT systems, services, and infrastructure components.
- Third, the transition maximizes the utilization of Enterprise standards, patterns and Programs, including CEDCaP and CEDSCI, and aims at consolidating similar capabilities into a common service or system. The modernization of the 2020 Solution will be enabled by emerging technologies that are SOA, Cloud, Mobile, and Web, while adhering to Federal directives for Cloud First, API, and Shared Services.

Finally, the 2020 Census Enterprise Architecture and Infrastructure Transition Plan is a living document, that will drive the collaborative process and be refined or adjusted to minimize risk and maximize efficiency, yet meet the ultimate timelines of 2018 End-to-End Test and 2020 Census Day.

1 Introduction

The purpose of the 2020 Census is to conduct a census of population and housing and to disseminate the results to the President, the states, and the American people. The goal of the 2020 Census is to count everyone once, only once, and in the right place, and the challenge is to conduct a 2020 Census at a lower cost per household (adjusted for inflation) than the 2010 Census, while maintaining high quality results.

The 2020 Census has been designed to take advantage of new technology to enable innovative operational approaches that will achieve significant cost savings, while maintaining high quality. There are four key innovation areas that are the focus of the 2020 Census design:

- Reengineering Address Canvassing
- Optimizing Self-Response
- Utilizing Administrative Records and Third-Party Data
- Reengineering Field Operations

Reengineering Address Canvassing is designed to reduce the amount of in-field labor required to update the Master Address File (MAF) and associated technical products used for assessing where to count. In-office Address Canvassing (AdCan) will be completed nationwide using imagery and address files shared by local and tribal governments. In-office AdCan is less expensive than in-field AdCan, but in-field AdCan will still be required in up to 25% of all addresses. The plans for reengineered address canvassing are expected to reduce field workload by up to 75% by adding new addresses to the Census Bureau's address frame using Geographic Information Systems (GIS) and aerial imagery instead of sending Census employees to walk and physically check all the census blocks.

Optimizing Self-Response is designed to maximize the degree to which the respondent pool can successfully self-respond, reducing the cost of paper data capture and in-person Nonresponse Followup (NRFU). By encouraging the population to respond to the 2020 Census using the Internet or the telephone, the need for more expensive options are reduced.

Utilizing Administrative Records and Third-Party data is designed to reduce field workload by improving the efficiency of NRFU operations. By using these alternative sources of data, NRFU operations can eliminate the need for multiple unproductive visits to housing units that are vacant, abandoned, or otherwise unoccupied. It is expected that some nonresponding housing units will be enumerated in the 2020 census by using administrative records. The 2020 Census will enumerate many Group Quarters (GQs) through reference to administrative-type, third-party data. By using data the public has already provided to the government and data available from commercial sources, the Census Bureau can realize savings to focus additional visits in areas that have been traditionally hard to enumerate.

Reengineering Field Operations is designed to increase the efficiency of field operations, allowing managers and field workers to be more productive and effective. Combining new operational control software and case management tools with GPS-enabled devices will improve the efficiency of field workers and allow faster and more accurate management of field worker labor and travel expenses.

Using sophisticated operational control systems, Census employees can follow up with nonresponding housing units and keep better track of the daily progress of field workers.

1.1 Document Purpose

The purpose of the 2020 Census Enterprise Architecture and Infrastructure Transition Plan is to describe the Census Bureau's planned approach to incrementally transition the 2020 solution architecture from the current state to support the planned testing in 2016, 2017, and 2018, and ultimately the 2020 Census. The current state for purposes of this document is the architecture that was in place to support the 2015 National Content Test (NCT) and a separate geographic program infrastructure that is also undergoing evolution simultaneously even though the geographic programs were not part of the 2015 NCT.

The Plan helps ensure that the program is not at risk of trying to integrate too much too late in the 2020 planning cycle. By the start of the 2018 End-to-End test, the majority of the planned business capabilities and systems must be in place. The End-to-End test will ensure that all systems meet the requirements for scalability, reliability, and availability. Finally, the Plan is a living document, intended to be updated over time as capability development and integration continue, and the architecture evolves through the various phases of testing.

1.2 Document Scope

The transition plan documents the transition strategy as well as the business and technical capabilities needed to support each phase of testing for the processing systems only.

The document is a high-level plan that describes implementation strategies and when key elements of the architecture need to be transitioned. This includes the milestones when a solution is introduced or enhanced in the 2020 Solution Architecture, or when an architectural aspect is implemented, e.g. Cloud migration. Detailed integration plan steps can be found in the individual Solution project plans, containing milestones that should be rolled up to the major milestones specified in this 2020 Census Enterprise Architecture and Infrastructure Transition Plan.

Additional specifics of how each solution comprising the 2020 Census Architecture will be transitioned are contained in the individual Solution project plans, which are the responsibility of the solution providers. These detailed plans will include lower level details about the transition from the current state to the target state for each solution. Additional specifics of how the field office infrastructure transitions is covered in the Field Office Infrastructure documentation.

1.3 Document Organization and Relationship to Other Documents

This document contains seven sections and twelve appendices organized as follows:

1. Section 1.0 discusses the document purpose, scope and audience;
2. Section 2.0 contains an Operational Design Overview, Solution Overview, and then lists the Assumptions, Constraints, and Dependencies that apply to the Plan;
3. Section 3.0 discusses the Key concepts, summarizes the Transition Strategy, and lists the architecture considerations that impacted the transition strategy;
4. Section 4.0 describes the current state, including the business capabilities and IT infrastructure elements applicable to the current state;

5. Section 5.0 describes the target state, including the business capabilities and IT infrastructure elements applicable to the target state;
6. Section 6.0 walks through the transition timeline, from the 2015 National Content Test and ongoing Geographic programs up to the 2020 Census;
7. Section 7.0 will discuss ongoing research programs, such as the MAF/TIGER studies;
8. Appendix A contains a list of acronyms and abbreviations used in this document;
9. Appendix B contains a glossary of terms and definitions used in this document;
10. Appendix C contains references used in this document;
11. Appendices D – H contain a diagram for each test, with a high-level list of the capabilities available for each test;
12. Appendix I contains a description of each of the systems used in the 2020 Census;
13. Appendix J contains the outstanding decisions with potential impact on the Census Architecture;
14. Appendix K contains the Architecture Transition Tables organized by domain; and
15. Appendix L contains the TRM domains used in the 2015 NCT.

This document is part of a broader set of documentation for the 2020 Census Program that has been developed or will be developed as the program progresses. Rather than duplicate content covered elsewhere, this Transition Plan incorporates various documents by reference as detailed below:

- 2020 Census Operational Plan, Issued October 6, 2015
  - The Operational Plan documents the current operational design for conducting the 2020 Census. It includes a set of design decisions that drive how the 2020 Census will be conducted. The operational design also drives the requirements for Information Technology (IT) capabilities and acquisitions.
- 2020 Census Architecture, January 7, 2016, Version 1.0
  - The 2020 Census Architecture describes the target solution architecture with systems and their interfaces planned for the 2020 Census operations. It describes the infrastructure needed to handle the large scale of the 2020 Census. It also provides guidance on the development of systems that comprise the solution architecture and communicates the architectural principles which should be considered when developing or providing the capabilities for the 2020 Census.
- 2020 Decennial Census IT Infrastructure Strategy and Roadmap (draft)
  - The IT Roadmap articulates the technology direction for the 2020 Census, along with a vision for how information technology will transition to support the 2020 business and operational requirements. The IT Roadmap contains lower level details about how IT Infrastructure will transition, allowing this transition plan to remain as a high-level plan.
- Integration and Implementation Plan (draft)
  - The Integration and Implementation plan provides a high-level framework and milestones for engineering planning and communicates key dates to project teams and stakeholders.

As shown in Figure 1-1, this document (per the green arrow) is part of a broader set of documentation being developed for the 2020 Census. Note that when the CEDSCI documentation is available, Figure 1-1 will be updated.

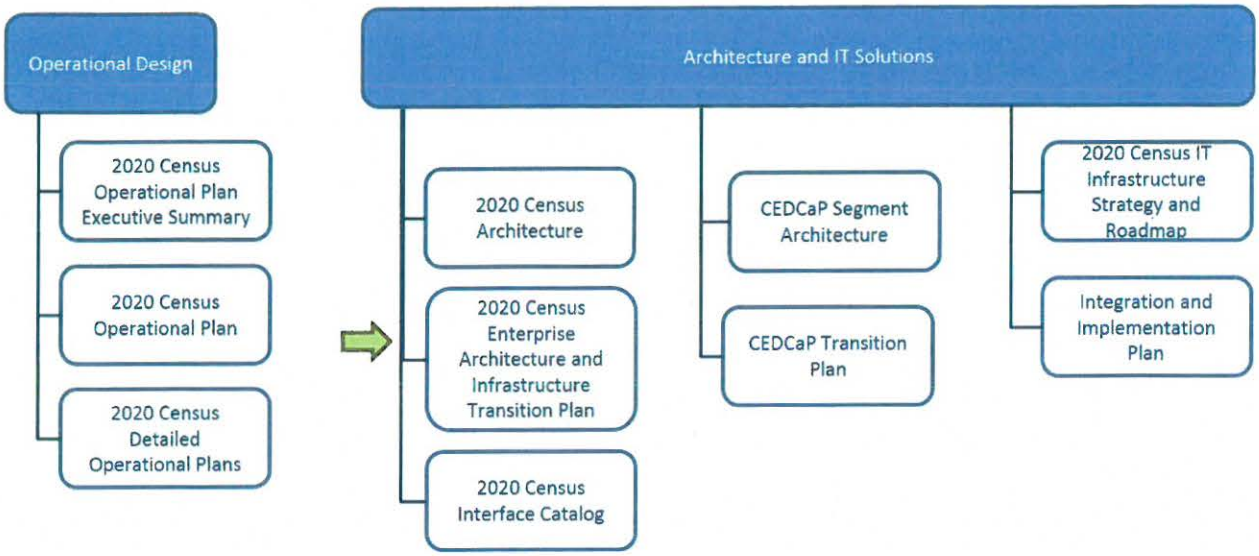


Figure 1-1 2020 Census Document Structure

1.4 Audience

This document is intended for several different audiences:

- IT Directorate
  - Provides information to the IT Directorate on what systems and infrastructure elements need to be ready and when to support the testing for the 2020 Census.
- Census Bureau Leadership
  - Provides an understanding of what systems and infrastructure elements will be implemented to support the phase of testing leading to the 2020 Census.
  - Documents decisions made by the Census Bureau Leadership which impact the 2020 architecture.
  - Documents when replacement systems will be introduced during the transition to the 2020 Census.
- Solution Providers
  - Informs Solution Providers about what systems and technical capabilities are required to support the various phases of testing leading to the 2020 Census.
  - More specifically, provides an understanding for the Census Enterprise Data Collection and Processing (CEDCaP) and Center for Enterprise Dissemination Services and Consumer Innovation (CEDSCI) solution providers about how their respective Transition Plans fit into the 2020 Census Enterprise Architecture and Infrastructure Transition Plan.
- Oversight
  - Provides assurances to oversight (e.g., Government Accountability Office and the Commerce Inspector General), that the Census Bureau has a well-thought-out plan for incrementally implementing the architecture over the phases of testing that will occur in the years leading to the 2020 Census.

2 Program overview

Figure 2-1 provides a high-level overview of how the 2020 Census will be conducted. This design reflects a flexible approach that takes advantage of new technologies and data sources while minimizing risk.

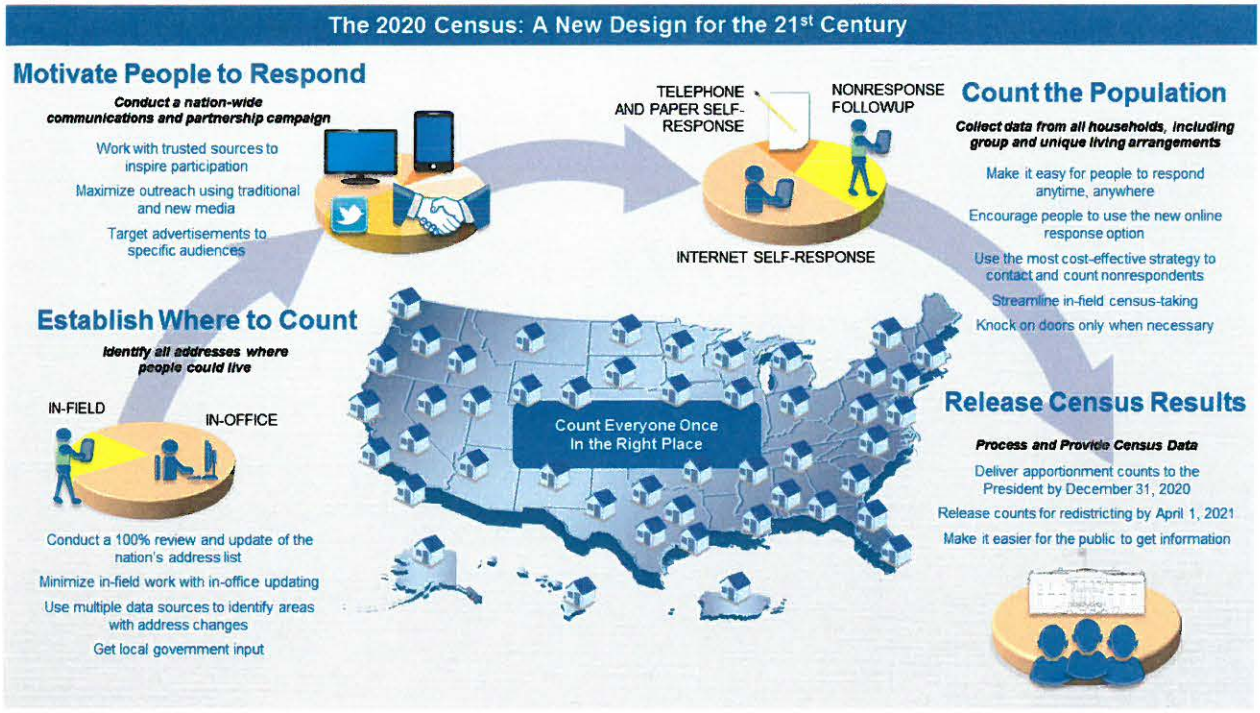


Figure 2-1 The 2020 Census: A New Design for the 21st Century

Successful execution of the 2020 Census will require coordinated efforts across a number of Census Bureau programs and will make use of resources from across the Census Bureau. The capabilities required to compete the 2020 Census are expected to be developed in an evolutionary fashion, with periodic large scale tests designed to exercise these emerging capabilities in realistic settings. Each builds successively from the prior tests, so that by the time of the first 2020 Census operations, the requisite capabilities are in place and have been field tested under realistic conditions.

2.1 Operational Design Overview

The iterative development process described above, which is unlike past decennials, incorporates lessons learned from early research, testing, and analysis. The lessons learned drive updated requirements for capabilities and acquisitions. Another important difference for 2020 is the emphasis on enterprise standards and solutions, such as the CEDCaP and CEDSCI initiatives, both of which are expected to provide substantial support for the 2020 Census.

Another significant difference from earlier Census designs is the focus on innovations to improve efficiencies and reduce costs. The innovations in Reengineering Address Canvassing, Optimizing Self-Response, Utilizing Administrative Records and Third-Party Data, and Reengineering Field Operations could potentially reduce the cost of the 2020 Census by over $5.2B as compared with repeating 2010 Census methods in 2020.

The 2020 Census includes 34 operations that are organized into seven major areas that correspond with the Census Bureau standard work breakdown structure. The term operation refers to both support and business functions. For example, Program Management is considered a support function, and Address Canvassing is considered a business function. The planned operations and major areas are shown in Figure 2-2.

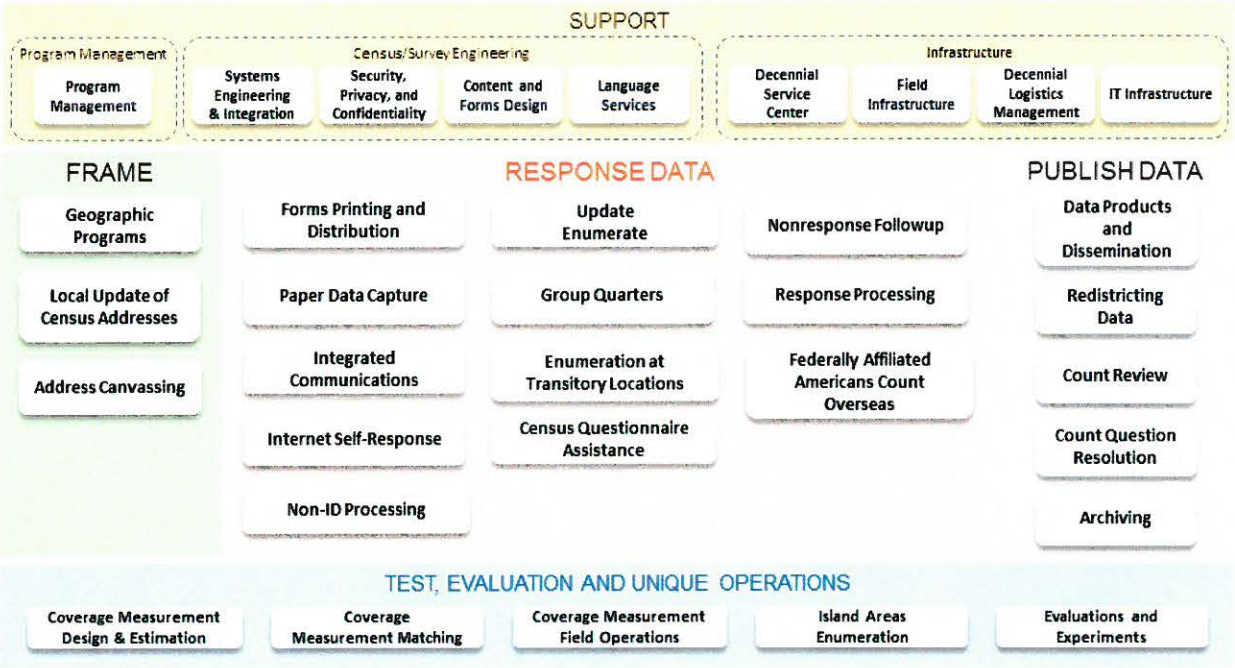


Figure 2-2: Operations by WBS

In addition to the four major innovations mentioned in Section 1, the 2020 Census will include a number of key innovations that impact many of the large, complicated operations. Innovations are considered significant changes to the operational design as compared to the 2010 Census. Figure 2-3 highlights in brown the operations have the most significant innovations. The specific innovations for each of these operations are listed in Table 1 below.

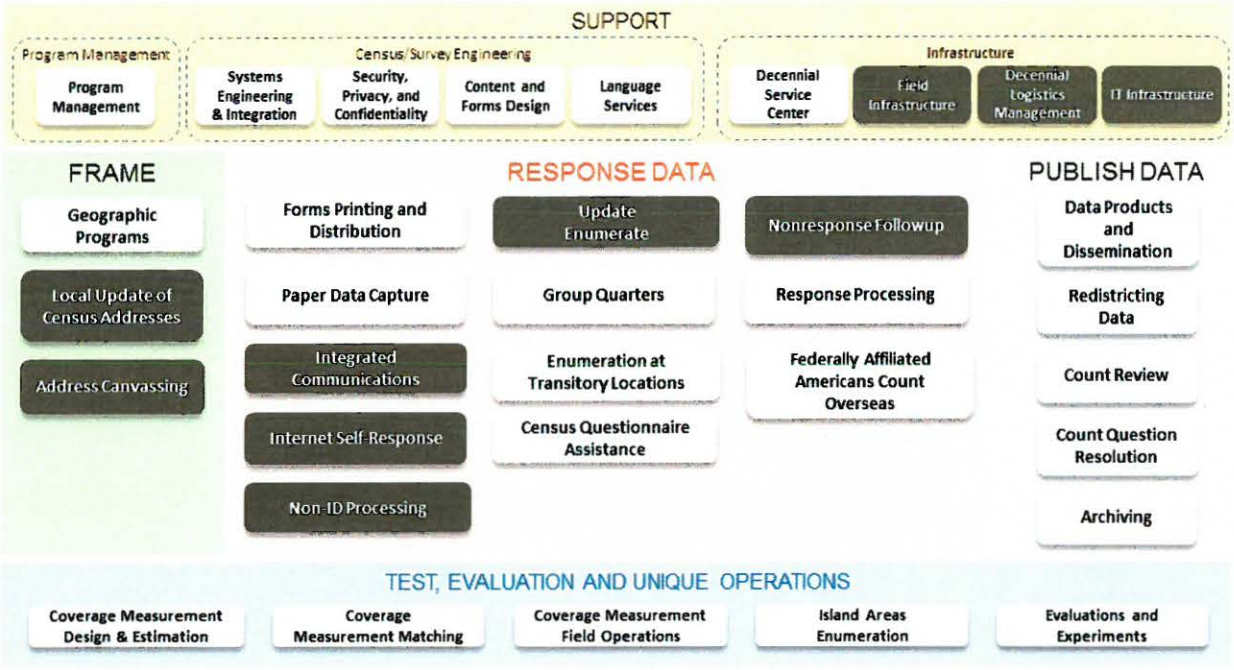


Figure 2-3 Operations with Significant Innovations Since 2010

Table 1 Summary of Key Innovations by Operation

Local Update of Census Addresses ▪ Reduced complexity ▪ Elimination of the full address list submission options to improve quality and reduce burden and cost
Address Canvassing ▪ Use of a combination of In-Office and In-Field methods to achieve a 100 percent address canvassing (target of 25 percent of addresses going to In-Field) ▪ Use of automation and data (imagery, administrative records, and third-party data) for In-Office Address Canvassing ▪ Ongoing MAF Coverage Study to validate In-Office procedures, measure coverage, and improve In-Field data collection methodologies ▪ Use of reengineered field management structure and approach to managing fieldwork, including new field office structure and new staff positions
Integrated Partnership and Communications ▪ Micro-targeted messages and placement for digital advertising, especially for hard-to-count populations ▪ Advertising and partnership campaign adjusted based on respondent actions ▪ Letters, postcards, and questionnaires to motivate self-response

- Expanded predictive modeling to determine the propensity to respond by geographic areas
- Expanded use of social media

Internet Self-Response

- Internet data capture, providing real-time edits, ability to capture household entries, and multi-access methods across different technologies (e.g., computers, phones, tablets, kiosks)
- Online questionnaires available in multiple languages and non-Roman alphabets
- Multi-mode contact approach tailored to demographic or geographic area
- A phone option (via Census Questionnaire Assistance) will be available for respondents, who need help or without Internet access, to provide their census data.
- Paper questionnaires will be provided in the first mailing to some areas of the country, and will also be sent to all nonrespondents after a few weeks.

Non-ID Processing

- Ability for public to respond anytime, anywhere
- Real-time matching and geocoding of responses
- Validation of response data
- Use of administrative records and third-party data used to validate identity and validate and augment address data submissions

Update Enumerate (planned innovations dependent on funding of this operation)

- The 2010 Census Update Leave and Update Enumerate Operations combined into a single operation
- Single visit with enumeration or push to Internet Self-Response
- Use of single device for both listing and enumeration
- Use of reengineered field management structure and approach to managing fieldwork, including new field office structure and new staff positions
- Assignment and route optimization
- Automated training for field staff
- Automation of the field data collection
- Automation of administrative functions such as recruiting, onboarding, and payroll
- Reengineered quality assurance approach

Nonresponse Followup

- Use of administrative records and third-party data to remove vacant housing units from the NRFU workload
- Use of administrative records and third-party data to enumerate some nonresponding occupied housing units from the NRFU workload
- Use of reengineered field management structure and approach to managing fieldwork
- Use of a variable contact strategy and stopping rules to control the number of attempts made for each address (based on paradata)
- Assignment and route optimization
- Automated training for field staff

▪ Automation of the field data collection ▪ Automation of administrative functions such as recruiting, onboarding, and payroll ▪ Reengineered quality assurance approach ▪ The notice of visit will provide the Internet URL to still encourage self-responses
Field Infrastructure ▪ Reduced number of Regional Census Centers managing a reduced number of local field offices tasked with managing field operations and support activities ▪ Automated job application and recruiting processes, payroll submission and approval process, and other administrative processes resulting in reduced staffing requirements ▪ Automated training ▪ Reduced number of enumerators and supervisors due to reengineered design for field operations
Decennial Logistics Management ▪ Implementation of an on-line, real-time Enterprise Resource Planning (ERP) system with extended access for the Regional Census Centers and field offices ▪ Implementation of a wireless network and bar code technology that will automate inventory transactions
IT Infrastructure ▪ Early development of solutions architecture ▪ Use of enterprise solutions as appropriate ▪ Iterative deployment of infrastructure aligned with and based on testing ▪ Implementation of alternatives to providing Government Furnished Equipment such as Device as a Service ▪ Use of demand models to help predict Internet response volume, Census questionnaire assistance center staffing, etc. ▪ Scalable design ▪ Agile development of applications

2.2 Solution Overview

At the solution level, the 2020 Census Architecture addresses all areas of the Survey Life Cycle (SLC) and Mission Enabling Systems (MES). The SLC, a Census enterprise effort, includes stages such as developing and designing the survey, developing the Frame, instrument and sample design, data collection, editing, estimation, review and dissemination. The MES are systems or solutions that support the SLC activities throughout the execution of the 2020 Census. From the functional perspective, the business requirements and capabilities are aligned to either existing systems enhancements or new solutions. From data/information perspective, quality is ensured at the collection points of entry. Applications rely on using a standard data taxonomy coupled with implemented data management procedures and security protocols which ensure data quality and integrity in not compromised while in transition and/or when at rest. From a physical architecture perspective, the 2020 Census Architecture will require the

use of technologies that are new to Census, such as Services-Oriented Architecture (SOA), Enterprise Service Bus (ESB), cloud services, and mobile applications. This plan defines the solution architecture's phased approach and requires the functional and technical consideration of each application and their integration within the overall architecture, enabling the program to leverage the enterprise capabilities including CEDCaP and CEDSCI, and non-enterprise capabilities. Below are the known capabilities that will be leveraged:

- Services-Oriented Architecture (SOA); including a common infrastructure and efficient system interfaces through an ESB to allow IT applications to communicate without the need for costly system re-writes
- Centralized, multi-mode, operational analysis and control capability
- Adaptive survey design set of capabilities, including statistical modeling and easy access to administrative records
- Unified interview operational control capability, including all data collection modes
- Modernized address listing and mapping capability
- Enterprise, integrated paradata repository
- Enhanced questionnaire design, fully leveraging content metadata
- Robust, scalable Internet and mobile data collection set of capabilities
- Electronic correspondence portal
- Modernized, integrated paper data capture capability
- Modernized, enterprise survey sample design and survey response processing set of capabilities
- IT infrastructure scale-up for the 2020 Decennial based on the forecast demand
- Centralized development and testing environment
- Use of a centralized system (PEARSIS) to manage administrative records for production

2.3 Assumptions, Constraints, and Dependencies

The following assumptions, constraints, and dependencies were identified by the architecture team. This list was made based on the current state architecture. Additional assumptions, constraints, and dependencies will be identified as the architecture progresses towards the end state.

2.3.1 Assumptions

- The current systems as described in the CEDCaP Segment Architecture will be implemented for 2020.
- This document will be updated based on:
  - Results of the CEDCaP COTS Analysis of Alternatives (AoA)
  - Major test results.
- All divisions and stakeholders being impacted and having an impact on the transition of the 2020 Census Architecture will communicate planned upgrades that include both server and software versions.
- The current architecture defined for the 2020 Census has been verified and validated by stakeholders.
- This plan is based on the current versions of various documents as referenced in Section 1.3 (e.g., 2020 Architecture Document - baseline 1.0 version, the 2020 Census Operational Plan, the

2020 IT Roadmap, and requirements/diagrams). The transition plan will be maintained to be consistent with updates made to these documents.

- The 2018 End-to-End Test will demonstrate the 2020 Census architecture. The 2016 Census Test, the Address Canvassing Test, the 2017 Census Test, and the series of small tests all include research efforts and each could introduce solution changes. These changes can impact and/or support the development of the 2018 End-to-End Test Architecture.
- Transition planning is an ongoing exercise
- The acquisition of the integrated communications contract in August 2016 could introduce potential solution changes.

2.3.2 Constraints

- The timing constraints arising from each area include but are not limited to:
    - Census tests and outcomes
    - Related activities and milestones by each operation
    - New solutions development and tests
    - Development, testing, and deployment of old and new capabilities.
- Security ATO process and timelines
    - The Census Bureau requirement for applications to receive an ATO is for the application to have a risk score of less than 5% as documented through the Risk Management Framework. The CIO and co-Authorizing Officials use this information to determine whether they are willing to accept the risk at the time of the ATO.
    - OIS requires 4 to 6 months to assess solutions before granting an ATO; to prepare for the ATO approval process, system owners need to plan the project's schedule to account for security review.
        - *Note: the security team requests that if this process starts at the beginning stages of the SDLC, this effort can be conducted in parallel to the development of the solution, resulting in shortened timeframes.*
- The Transition Plan will provide the required dates for divisions to make upgrades and enhancements at specific times, but several divisions support initiatives other than 2020 Census, so those divisions may not be able to follow the recommended timelines. CSVD, TCO and LTSO (IT Operations) need lead time (4-6 months) to procure, install and configure infrastructure to meet system requirements. This is a significant issue in for Census testing in 2016, and actions are being taken to mitigate this in 2017 and 2018.
- There are potential policy constraints with contractors having access to Title 13 data to combine with third-party data on an offsite system (DMP). This can be alleviated if we build the system in-house.

2.3.3 Dependencies

- Operations must define the scope of the test.
- The Transition Plan is dependent on the baseline IT Roadmap, which is partially complete and scheduled for final completion of July 2016.
    - The Transition Plan is a living document that will be revised to incorporate the enhancements made to the IT Roadmap as it is completed.
- The scale-up portion of the Transition Plan is dependent on input from the cloud initiative, and demand models.

- For the development of the plan (Jan-Feb) and review of the plan (Mar), staff availability will be essential for timely completion. Staff availability for June-August scale-up content development may also be needed to conduct reviews.

2.3.4 Issues and Decisions to be made

- The infrastructure for integrating Census Questionnaire Assistance (CQA) will be determined. The CQA contract will be awarded in June 2016, and the system will come online in 2017.
- The timing for introducing C-SHaRPS.
- The infrastructure for integrating the Communications Contract will be determined. The contract will be awarded in August 2016. The system will support the two tests in 2017.

3 Transition Strategy

3.1 Overview

The 2020 Census Architecture Transition strategy is developed based on the standard architecture domains - Business, Application, Information, Technical, Security, Quality- , and consists of evolving these domains from the current state to its target state. The current state is defined as the architecture used in the 2015 National Content Test, and the target state is the architecture that will support the 2020 Census Operations from the architectural aspects of business process, system quality, attributes of scalability, availability, reliability, and security.

The 2020 Census transition strategy delineates the 2020 Census Solution Architecture approach as it incrementally transitions from the current to target state. The individual census tests and activities demonstrate the progress in implementing the Business, Application, Information, and IT Infrastructure, Security, and Quality domains transition. The transition of each domain is shaped by business, operations, EA standards, US Census policies and guidelines, and needed business and operational capabilities.

The 2020 Census Architecture document captures the overall decennial scope, including the list of systems, and the interface catalog used for the integration between systems. The 2020 Census Architecture consists of multiple enterprise programs, including CEDCaP and CEDSCI and non-enterprise systems, which are independent of the two programs. The systems for each program are identified in the diagrams and in the systems list. Please refer to Appendix D: Product Release 2015 National Content Test, Appendix E: Product Release 2016 Census Test, Appendix F: Product Release Address Canvassing Test, Appendix G: Product Release 2017 Census Test, Appendix H: Product Release 2020 Census, and Appendix I: 2020 Census Systems List to view the list of systems and the solution architecture diagrams.

Moving towards a target solution that meets the business requirements, the 2020 Census Architecture's transition milestones need to closely align with the CEDCaP and CEDSCI program milestones and transition activities, as well as other IT infrastructure roadmap timelines. These timelines are further detailed in section 6 of this document.

Currently, the CEDCaP program is analyzing possible COTS solutions for the major capability areas of the data collection segment. The CEDCaP's Analysis of Alternatives (AoA) effort can result in one of two possible outcomes that can impact the 2020 Solution Architecture. One possible outcome is that CEDCaP will implement a COTS solution to support the data collection segment. Providing, this path meets the 2018 End-to-End test timelines imposed by the 2020 Census Architecture milestones; the 2020 Census solution architecture will leverage the implementation of the COTS solution. The second option is for CEDCaP to continue developing in-house custom solution, which also must adhere to the 2020 Census milestones and timelines. It is important to recognize that in either case of COTS vs. Custom development, the major capabilities, solutions, and interfaces will remain unchanged, while adhering to Enterprise Architecture standards such as SOA, ESB, Mobile Technologies, and Cloud First policy.

3.2 Business Architecture

The business architecture relies on the phased approach tests driven by a comprehensive list of the 2020 Census Architecture business requirements. Each test re-uses the existing solution and technologies to the extent possible, introduces new enhancements of existing capabilities and/or replaces older legacy systems with new solution initiatives. The general goals and objectives of each test are to support, evaluate, and modernize the existing solutions.

The business capabilities of the 2020 Census Solution Architecture align with the 34 Census business operations. The Census Bureau and the Decennial Program will conduct various tests that drive the incremental transition of the solution architecture. These tests and Censuses serve as effective tools in verifying the capabilities of each solution, while facilitating the design progression of the infrastructure to support the 2020 Census scalability requirements. The specific capability requirements and Business Process Model (BPM) diagrams for each Test as well as their business requirements can be found in the requirements repositories, and are managed by the Decennial Architecture Requirements Team (DART).

3.3 Application Architecture

The Open Group Architecture Framework (TOGAF) defines the Application Architecture as a description of the structure and interaction of the applications as groups of capabilities that provide key business functions and manage the data assets.

The current state (2015 NCT) application architecture is a highly complex integration of existing and new solutions having different platforms, hardware and software, multiple data sources with some manual, and little automated process.

The 2020 Census Architecture target state is based on the SOA paradigm where each application will provide services to the overall solution. In providing these services, these applications, can dictate development of technical enhancements and defining design patterns, APIs, Web Services, use of ESB, mobile, and cloud technologies. An example of technical enhancement of the future state is the use of ESB and the SOA enterprise methods vs. database links. While in the past, applications communicated often via manual processes to manage large transfer of file/data, the future state will allow applications to use integrated enterprise data models to communicate with other systems and share data. The 2020 Census future state architecture will modernize how data is shared and transferred between the systems. The modernization will consider the interoperability and interfacing elements such as data format, type, size, frequency and performance elements such as throughput, response time, and quality of service. In the current state, files were transferred manually using ad-hoc file transfer solutions, such as File Transfer Protocol (FTP), while the future state will utilize Enterprise Integration Patterns based on API, ESB, and Managed File Transfer (MFT) software to securely and efficiently share data across systems.

The solution architecture diagrams for each test leading up to the 2020 Census lists the systems for each test and identifies their interfaces to each other. The arrows and their makings depict the type interface and of data being transferred as well as whether the transaction takes place via the ESB. Please refer to Appendix D: Product Release 2015 National Content Test, Appendix E: Product Release 2016 Census Test, Appendix F: Product Release Address Canvassing Test, Appendix G: Product Release 2017 Census Test, and Appendix H: Product Release 2020 Census to view the diagrams as well as the Interface

Catalog referenced in Appendix C: References to view the interface catalog detailing the application interfaces.

The general approach to updating the artifacts for each test and census will be similar. When the existing systems or new capabilities are tested the findings will be recorded as we anticipate that the findings will drive the need for additional enhancements or development of new solutions.

3.4 Information Architecture

The information architecture will leverage the efforts of Enterprise Data Model (EDM), as applied to the 2020 Census Architecture applications. The Census Bureau Enterprise Data Team is currently developing an EDM as well as a 2020 Architecture data model. Initially, the EDM will produce and identify the conceptual entities used by MOJO, a dictionary of terms defined for Content related concepts and define the terms for Paradata and Operational related concepts. By year 2017 the EDM will propose the EDM as the standard for National Information Exchange Model (NIEM) statistical domain. Other services and capabilities produced by the EDM team, is a machine readable specification by year 2017 to be used in the 2017 Census Test. This will include a 2020 Cycle specification defined in XML/XSD and an XML validation of each mode/instrument specification.

The Enterprise Data Services will produce variable names based on upon existing content and terms defined in the NIEM and 11179, establish a prototype metadata registry, and by 2017, the metadata registry will be available for content creation. The NIEM design includes both generic information concepts such as Date, Person, and Location that can be reused across multiple federal information architecture domains, as well as a number of domain-specific libraries of terms applicable to certain federal lines of business such as Biometrics, International Trade, Immigration and Maritime.

Finally, the EDM will introduce the EDM concepts in an Interface Control Document (ICD) where ICD's are re-written as NIEM Information Exchange Package Documentation's (IEPD's).

One of the objectives of the enterprise data model(s) development effort is to facilitate efficient data exchange between applications and the environments using a standard taxonomy and approach.

3.5 IT Infrastructure

Based on the Information Technology Infrastructure Library (ITIL) v3, IT infrastructure is a combined set of enablers, tools, procedures, and methods such as hardware, software, networks, and facilities that help develop, test, deliver, monitor, control or support IT services. ITIL provides a consistent approach to managing IT services resulting in cost reduction, improved IT services, customer experience, and improvement in the staff and technology productivity, enhanced delivery of third-party services. The US Census Bureau Decennial program in transitioning to the target 2020 Census Architecture focuses on using the industry best practices and seeks to leverage the approach in developing the IT Infrastructure.

This document and the architecture document address the IT Infrastructure and at a high level and identify the systems that will use the enterprise cloud services. The enterprise cloud implementation strategy is currently under development. It will need to address the readiness of the cloud solution and the detailed strategy in meeting the 2020 solution architecture needs. Furthermore, a number of the interfaces are expected to use the ESB and web services in order to communicate and share files securely and efficiently. The solution architecture and the infrastructure relies on the US Census Security

team to meet the security requirements and continue to be engaged with the program/project team in order to ensure that each system will meet the security requirements. To view a list of systems that will use the cloud technology and ESB please refer to Appendix K.4: Infrastructure Architecture.

The enterprise IT Infrastructure follows a plan for periodic technology refresh with its timelines, including compute resources, storage, network devices, system monitoring, and security appliances. In the current planning, there is also a strategy to do research for utilizing Cloud computing for IT Infrastructure. With the use of Cloud computing, system management and monitoring footprint must be extended to manage systems deployed in Cloud environments. In order to successfully support the 2020 Census, the technology refresh timelines must be aligned with the Census Tests timelines and 2020 Architecture Transition Plan so to ensure readiness of the infrastructure components for Census Tests and 2020 Census. This readiness is critical to the dependable operation of the IT applications within the 2020 Census Solution Architecture.

3.5.1 Service Oriented Architecture (SOA)

A service-oriented architecture (SOA) is an architectural pattern in computer software design in which application components provide services to other components via a communications protocol over a network. Service-orientation principles are vendor, product or technology agnostic.

The Census Bureau has adopted a Service-Oriented Architecture (SOA) as a means to help deliver on its mission requirements. Developing services as "building blocks" that can be reconfigured, will meet evolving business and technical needs of the agency. Building key functionality out of building blocks, utilizing SOA, increases the agency's ability to adapt to changing requirements more rapidly. With a focus on business results, the SOA strategy will enable the US Census Bureau to:

- Decrease complexity in application environments and increase the ability to integrate applications quickly and efficiently.
- Provide Data and application services through APIs, which are accessible internally by systems across the enterprise.
- Deliver application solutions faster, with reduction to system integration costs
- Provide support for application services reuse, and enable business functions to operate more efficiently

Evolving requirements challenge the Census Bureau to meet mission critical needs with greater flexibility and speed. Gaining better efficiency in our IT solutions includes developing tools and services that are both repeatable and reusable. Beginning with the 2016 Census test and moving forward, we look to provide solutions which can be easily integrated into a SOA-based environment and which facilitate centralized API management.

The 2020 Census solution architecture is based on the SOA paradigm. The interactions and interfaces needed by the architecture requires each solution to provide services and share data via interfaces to one or more systems in the solution. The solution architecture requires certain (e.g. software, hardware and enterprise tools) technology components. To that extent, the US Census Bureau Chief Technology Officer (CTO) Office is responsible for the assessment, evaluation, and adoption of technologies. Moreover, with the introduction of Cloud services, the SOA architecture will encompass not only

services provided by in-house solutions, but also services delivered via the PaaS and Software as a Service (SaaS) models.

In March 2013, the CTO office conducted an ESB and Rule Engine need analysis. The team selected the Oracle SOA Stack that includes ESB (OSB) and the Oracle Policy Automation (OPA). Currently, the CTO office is leading the planning and conducting the AoA for the Enterprise SOA, API, and BPM Tools. It is expected that by FY 2017, the AoA is complete and the production SOA stack integration will be ready. The utilization of the SOA principles and Enterprise Tools will mature as the analysis continues to establish reliable levels of service and availability.

One critical of the SOA is the ESB, which facilitates the integration of loosely coupled services within the architecture. ESB is a common implementation pattern for SOA and its objective is to find a standard, structured, and general solution for implementing loosely coupled software services such as event handling, data transformation and mapping, message and event queuing and sequencing, security or exception handling, protocol conversion and enforcing proper quality of communication. Services are expected to be independently deployed, running, heterogeneous, and disparate within a network. Like concurrent operating systems an ESB caters for common and enterprise services in addition to adoption, translation and routing of a client request to the appropriate service.

The 2020 solution architecture goal is to transition into the ESB and be able to effectively and efficiently transfer messages and data files across the platform. The initial ESB implementation will take place for the 2016 Census Test, where the API management and Architectural Patterns (i.e. messages and Manage File Transfer (MFT)) will be used. To view transition of the 2020 solution architecture to ESB, refer to Appendix E: Product Release 2016 Census Test, Appendix F: Product Release Address Canvassing Test, Appendix G: Product Release 2017 Census Test, and Appendix H: Product Release 2020 Census where the interfaces that will use ESB are identified. Note that the solution architecture calls for API First policy established by the Enterprise Architecture as directed by the Digital Government Strategy.

3.5.2 Mobile Technology - Device as a Service (DaaS)

The 2020 Census Architecture will transition to Device as a Service (DaaS). The 2016 Census Test will be the first opportunity to test DaaS. During the 2016 Census Test, field enumerators will use mobile devices provided by a vendor via DaaS, and loaded with applications to perform enumeration, listing (except in the 2016 Census Test), and payroll reporting. The US Census Bureau will provide mobile enumeration devices to each field enumerator. These devices will be fully operational with a uniform vendor platform installation and provisioning for accessing the Census network. The implementation of DaaS, using commercially available mobile phones, implementation will allow a transition to being directly connected to the US Census network to record secure electronic data collection and transmission of Census results.

3.5.3 Enterprise Cloud

Enterprise Cloud Computing environment resides behind a firewall and delivers software, infrastructure and platform services to an enterprise. Cloud computing eliminates the typical risk and loss of system utility arising from power grid interruptions, physical data loss due to catastrophic events and malicious on site attacks to the IT infrastructure within an organization. The cloud computing framework provides an optimal environment for faster, safer and cheaper delivery of IT services within an enterprise.

The US Census Bureau approach to developing or replacing IT Infrastructure elements at each refresh or end of lifecycle, is Cloud first; meaning any component of the infrastructure included in the Technical Reference Model (TRM), can be evaluated and considered for migration to the Cloud. Key strategic steps used to evaluate fitment for Cloud migration or adoption are described below and presented by Figure 3-1: Key Strategic Steps for Consideration to Enterprise Cloud.

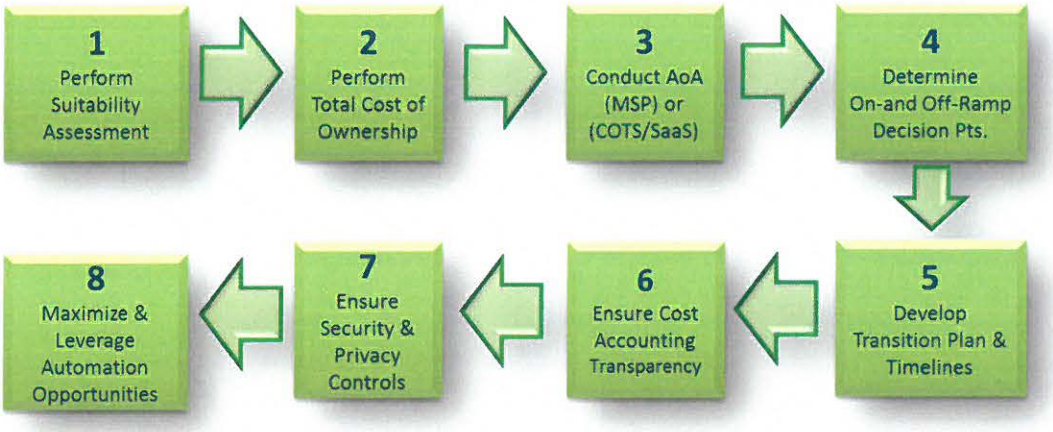


Figure 3-1: Key Strategic Steps for Consideration to Enterprise Cloud

- **Perform suitability assessment** on technology/architecture to determine readiness or potential for a cloud solution while recognizing the importance of application design, interfaces and connections are keys to a successful cloud solution implementation.
- **Perform a Total Cost of Ownership (TCO)** including an AoA to see if a commercial managed service provider (MSP) or COTS/SaaS is more appropriate to deliver the capability based on time/cost/resources (human)/maturity of technology.
- **Determine appropriate on-ramp and off-ramp,** either in a tech refresh, or a move to a different cloud service provider (CSP), while considering alignment with other technologies that have dependencies or a specific application is dependent on.
- **Ensure appropriate security and privacy controls** including records retention can be achieved with the solution.
- **Ensure cost accounting and transparency** can be achieved with the solution. Consider what to purchase, how to monitor and control costs, and ensure appropriate levels of service(s) are provided for payments made.
- **Develop transition plan** and schedule to include a WBS that aligns resources to the various activities.
- **Maximize and leverage automation opportunities –** It is important to recognize possible areas and opportunities to automate and optimize technical and business operations. Some of the areas are lifecycle management or provisioning and patch automation, using automation tools to install, scan

for security, development environments movements to test and production environments, and incorporating automated workflows in the applications business process.

3.5.4 Disaster Recovery and Continuity of Operation Plan (COOP)

The US Census Bureau places a high value on being prepared for disasters and operation continuity. Failing to be prepared to quickly recover from a disaster or to prevent it initially can result in a loss of public confidence, putting public safety at risk, and service downtime that can be very costly. Having services and operations available is critical to the US Census Bureau organizational mission.

Note: This section is under development. The content to be included here will elaborate on the strategy for developing customized Disaster Recovery and COOP plans and timelines. These plans can help the US Census Bureau and the 2020 Census operations quickly recover from system failures, and keep the systems and operations running with a minimal loss of data, and productivity. We will begin planning this activity in 2016, with design and implementation of the overarching solution planned for late 2016 in order to ensure quick recovery, high availability, and reliability during the Census operation period.

3.6 Security Architecture

The Census Bureau Office of Security (OIS) has established security policies and guidelines at the program, enterprise and system levels. In adherence with the security policies, the system owners and program offices are required to register each system where they are responsible for its operation into the eSDLC program. Project teams are guided to develop and publish a number of artifacts such as a System IT Contingency Plan and a baselined project schedule as preliminary steps to obtaining an Authorization to Operate (ATO). The CIO also recommends having security team activities that are integrated into the system development schedule and timelines as early as possible.

The security architecture details are described in the 2020 Census Architecture document. As an indication and verification of system and solution's readiness, each system must comply with existing enterprise security policies and authorization process described in the document. Any authorizations, verifications, notices and agreements must be completed for validation by the 2018 End-to-End Test.

The CIO maintains a systems Authorization to Operate (ATO) status tracker which is updated at a minimum bi-weekly, to record the changes and any scheduled re-authorization of a given system. The systems are categorized based on a CEN where it allows the CIO to conduct re-authorizations for multiple systems by conducting one CEN group authorization.

The transition to the 2020 Census architecture involves validating the security of nearly 60 systems. While some systems will be replaced by new solutions for the 2020 Census, the security evaluations will continue on a yearly basis for all systems. As the 2020 solution architecture matures by conducting each test and the possible introduction of new enhancements, new solutions, new technologies such as Cloud and Mobile technologies, the use of an Enterprise Service Bus (ESB), and interfaces with external entities and agencies, the CIO will continue to update and establish the necessary guidelines in order to prevent unsafe use and unauthorized access to and sensitive and non-sensitive data, applications, and environments.

The strategy to deliver the necessary security to the 2020 solution calls for the security team to work with the architecture solution team to identify the security needs of the solution. This collaboration should result in determining the security requirements and an implementation model on which the security team can demonstrate the approach of delivering security to the solution.

More information will be provided regarding the timelines and plans for implementing the appropriate levels of protection for the systems involved in the Census Tests and 2020 Census. However, it is suggested that the ATO process of an in-house system with little constraints, will require sixty business days to complete. A table containing a list of systems, timelines, and their latest ATO statuses will be provided and included in the appendix section of this document.

3.7 Program-Enterprise-System Quality

Achieving high quality in business, providing service to the public, and solutions are essential to the success of the US Census Bureau. Quality data gathering, processing and storing is a high-priority goal of the 2020 Census. The US Census gathers data of variety nature and size some of which is considered personal and confidential. Therefore, quality in gathering such information securely is critical. The 2020 solution architecture considers four categories of quality; they are:

- Business/Program quality - Ensures governance processes, timelines, and milestones are realistic. It ensures the needs of the business are met and that collaboration with end users to achieve the optimum business outcome is completed. At the business/operation level, substantial effort is placed on making business processes more efficient such as CAES adjusting workloads based on business rules, applying admin records to improve the quality of samples. We need to continue working and identifying ways to measure and effectiveness of the steps we are taking prior to using them in production.
- Enterprise quality - Ensures enterprise level services, integrations, and tools operate as designed and intended. Some quality attributes related to the enterprise as well as systems include performance, security, availability, scalability, usability, interoperability, portability, adaptability, reliability, maintainability, dependability, elasticity, testability, extensibility.
- Data Quality – The overall Census data quality approach is to ensure the overall Census Count is accurate. To that extent, systems like Coverage Measurement (CM) use a statistical method, matching an independent survey to a sample of census returns, and matching an independent listing to a set of canvassing results, to assess the overall quality of census results. By carefully designing a sampling approach, the census returns can be compared to sample returns and compared for alignment or variation. Depending on the degree of observed differences between the CM sample-based results and the Decennial Census-based results, the statistical measures of quality can be determined.
- System Quality - Various systems and initiatives, such as the development of CAES, SMaRCS and other systems within the solution, are designed/will be designed to ensure higher-quality data capture via entry by human or machine (e.g. scanning solution), or via interfacing with other internal and external systems. These efforts support the user to efficiently gather data, the system to capture data, and promote accurate modeling of the information gathered during the next census. The internal process quality parameters make sure that each system controls the quality of their output. For example, iCADE will sample work from data entry operators and

reprocess poor quality work; Telephony solution will sample and record phone calls, and review them to make sure operators are performing up to standard. Likewise, other systems are expected to have integrated controls to ensure quality input and output.

The aspects of quality, which accompany the expansion of the business functionalities, are further described below and are presented by Figure 3-2: Quality Aspects vs. Expanding Capabilities and Solutions, which indicates that as each solution's capabilities expand, the quality aspects of the target architecture will increase. The architecture as a whole will improve in delivering better solution performance, better reliability, decreasing downtime for higher availability, and high scalability in order to meet the business and traffic demands of the 2020 Census.

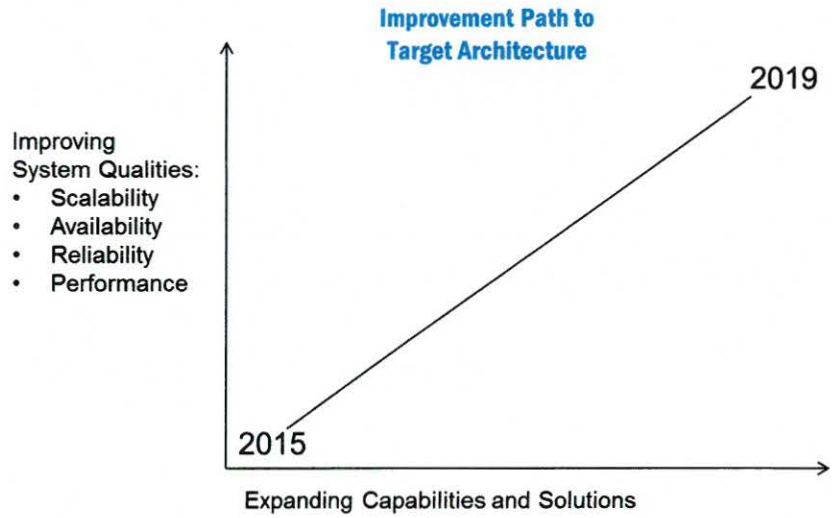


Figure 3-2: Quality Aspects vs. Expanding Capabilities and Solutions

3.7.1 Scalability & Performance

Scalability refers to the ability of a system, network, or process to adapt and handle change in workload capacity and process demands. Elasticity is the ability to dynamically provide additional resources when load increases, and to tear down unused resources when demand wanes. While scalability is important to the system, processes and network, elasticity is mostly important in business scenarios such as Census, where high surge in traffic is expected. Moreover, elasticity is cost-effective since cloud environment can provide services on a pay-per-use arrangement.

The demand for a scalable systems and processes in 2020 Census Architecture is very high, due to recognition of collecting, storing, and saving vast amount of data via various modes of operation.

Currently a 2020 Census Architecture scale-up initiative is underway led by a three stage model, the conceptual, analysis, and implement/test. The conceptualize stage starts the process by producing several artifacts. They are the demand models, and the Non-functional requirements. The analysis stage realizes the overall solution, system, data flow, and interfaces while the Implement/test focuses on execution and testing of the analyzed models, using optimization, partitioning, replication, design

patterns, scaling and continuous testing techniques. This model is shown in Figure 3-3: 2020 Scale-Up Three Stage Process. The scale-up team has conducted various internal and external demand studies and has developed models to establish the scale-up and scale-out needs of the 2020 Census Architecture. This approach seeks to identify the scalability needs at various stages in the solution. The goal is to achieve true scalability that will improve system performance.

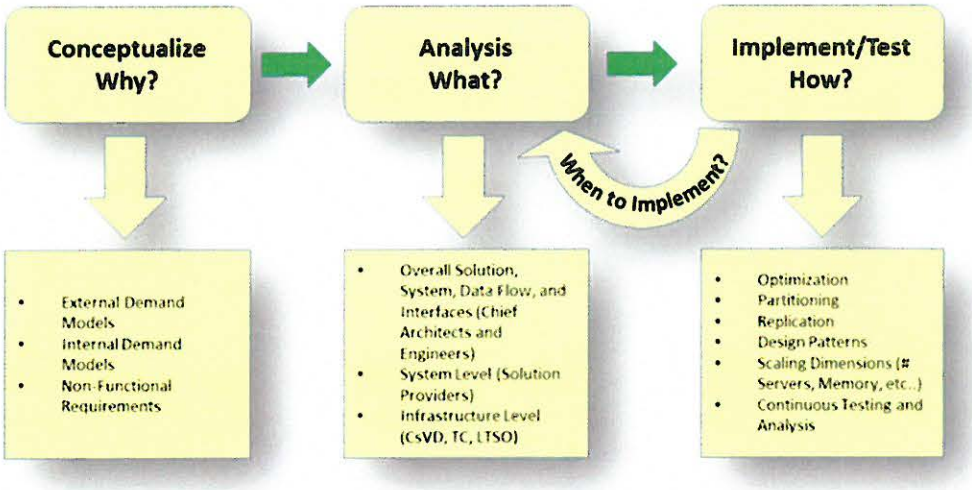


Figure 3-3: 2020 Scale-Up Three Stage Process

Although the workload during the 2015 NCT did not require scalability of the systems/architecture due to the size of the test, as we transition to 2020 Census and target state, the need to scale up will increase and during the transition period the goal is to test and demonstrate the ability to scale up to full 2020 Census size. Figure 3-4: Scale-up Initiative Transition Timeline, illustrates the Scale-up readiness timelines that the team has developed and when a scalable solution will be ready.

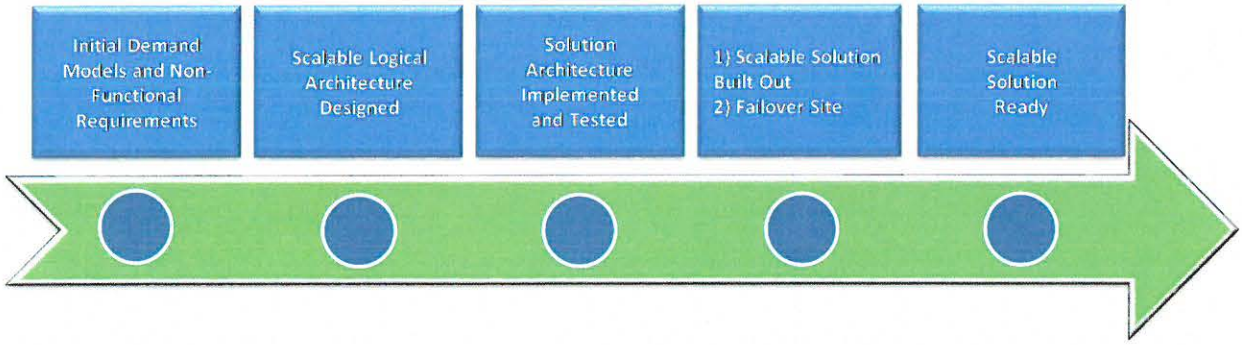


Figure 3-4: Scale-up Initiative Transition Timeline

3.7.2 Availability & Reliability

A system is considered reliable when it continues to operate as designed and availability is the probability that a system will work as required for the duration of its mission. A system is not considered

to be available during non-operational periods; scheduled downtime is not included in the computation of downtime. Availability of a solution is represented by the ratio of the expected value of the uptime (operational time) of a system to the sum of the expected values of the up and down times (operational) + (non-operational) as shown below.

$$A = \frac{E[\text{Uptime}]}{E[\text{Uptime}] + E[\text{Downtime}]}$$

The solution architecture team continues to collaborate with stakeholder teams (i.e. Scale-up working group and Infrastructure IPT) to establish the accepted levels of these architectural aspects. These collaborations aim to improve not only the availability and reliability but also of the solution as a whole, its interfaces, technology stack, and increasing the solution's fault tolerance.

During the transition period, the goal is to determine the accepted level of system availability for each system. This will help determine the thresholds of availability for mission critical and non-critical systems and help determine the overall architecture functional availability to conduct the census. While availability was not a critical factor during the 2015 NCT, the goal is to collect metrics, to test and to demonstrate higher system and architectural availability as we approach the timelines for the 2020 Census.

3.8 Transition Approach

The approach to an architecture transition differs for each implementation and or organization. However, there are two standard approaches to developing an architecture based on the industry standard. The baseline first approach is used where organizational units have had a high degree of autonomy. In this style, an assessment of the baseline landscape is used to identify problem areas and improvement opportunities. This process is most suitable when the baseline is complex, not clearly understood, or agreed up on. The other approach is a target first style. In this style, the target state solution is elaborated in detail and then mapped back to the baseline, in order to identify change activity. This process is suitable when a target state is agreed at a high level and where the enterprise wishes to effectively transition to the target model.

The 2020 Census Architecture was developed using a hybrid approach and its transition requires an in-depth view of the needed capabilities, inventory of systems, technologies and innovations, business drivers and requirements, the IT infrastructure as well as security requirements, contracts, and integration needs. It is also imperative to define a starting point or establishing a current state in the life-cycle of architecture from where to transition to the future state.

Successful transitioning of the 2020 Architecture depends on a number of facts, factors, practices, standards and guiding principles. They include:

- Some solutions will reach their end of life, or be replaced by COTS, or CEDCaP.
- 2020 Logical Architecture remains the same.

- The data exchange model should be standardized and follow enterprise standards and patterns.
- Wherever the final solutions are not yet available for testing, legacy solutions will be used. Adequate infrastructure support must be extended to the legacy solutions. Interfaces between legacy and new systems must be developed where needed.

In determining the current and target states the Architecture team teamed up with various teams and stakeholders to ensure a current state is precisely established. The diagram below depicts the communication and collaboration that will continue to take place among all stakeholders and program/project teams as well as the security and information engineers. The objective of these outreach activities includes communicating the architecture design and requirements to each team, as well as discussing business and technology perspectives in all architectural domains as the tests and census are conducted. Figure 3-5 below illustrates the collaboration with the stakeholders, subject matter experts, architects and engineers across the directorates in order to address each architectural domain via IPT(s) and Working Groups. This collaborative effort will produce a convergent and orchestrated result for the overall Solution Architecture to support the 2020 Census operations.

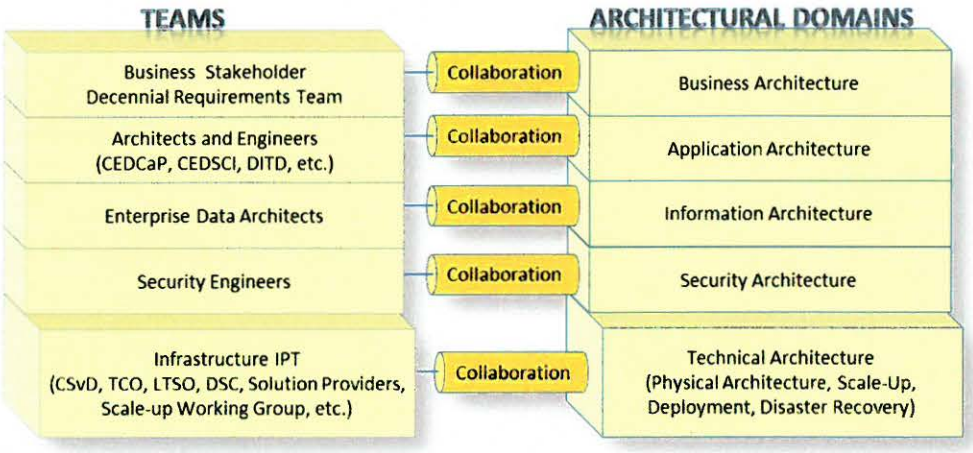


Figure 3-5: Architectural Domains and Team Collaboration

3.8.1 Transition Steps

Transitioning to the 2020 Census Architecture target state requires practical steps and a phased approach that considers the complexity and diversity of the solutions within the Architecture. The steps to defining a transition path are depicted in the Figure 3-6 below.

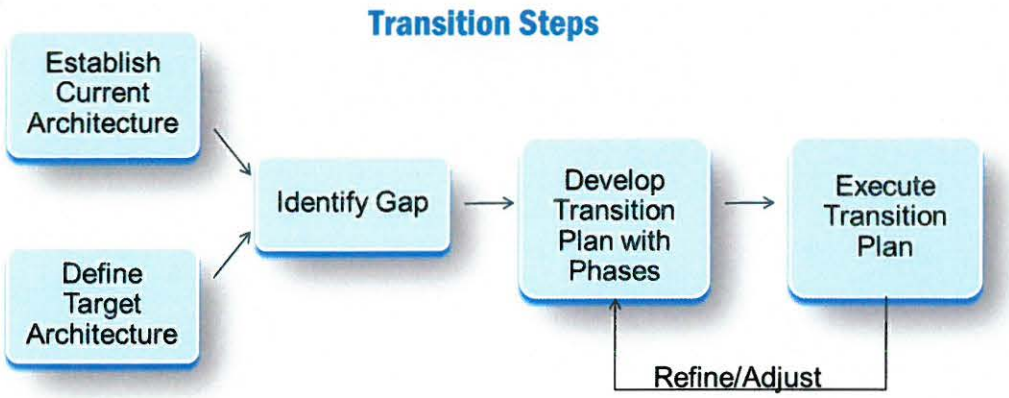


Figure 3-6: 2020 Census Architecture Transition Steps

The phased approach consists of the various anticipated Census Tests; their results will validate a designated part of the 2020 Census solution, while the 2018 End-to-End Test shall validate the major components of the solution. The diagram below graphically represents the transition to the 2020 Census Architecture.

During the various transitional phases, multiple solutions will be introduced in the scheduled tests prior to the 2020 Census while some solutions continue to undergo enhancements leading up to 2020 Decennial Census Day. The major tests of systems ensuring solution readiness are the 2015 National Content Test, 2016 Census Test, AdCan Test, 2017 Census Test, Puerto Rico Test and finally the 2018 End-to-End Test. These approximate 90% of what the actual 2020 Census will be. The transition plan introduces the systems and describes where enhancements lead to a final state and production.

3.8.2 Phased Testing

At the result of communications and discussions with various stakeholders and systems owners, and, in consideration of the business requirements and the needed capabilities, the 2020 Census Architecture current state began with the 2015 NCT. The transition is phased and promoted incrementally from its current state of to the target state. For each phase test and census, a specific solution architecture diagram is developed, verified and validated by the appropriate stakeholders. Leveraging the outcome of the tests include introduction of new capabilities, making enhancements, system performance, scalability while aiming to meet milestones for each test. Figure 3-7: 2020 Census Architecture Incremental Transition Model illustrates the phased testing approach supported by detailed operation plans, GOSC and business requirements as the control and the enabling technologies such as SOA, Cloud and Mobile as the mechanisms.

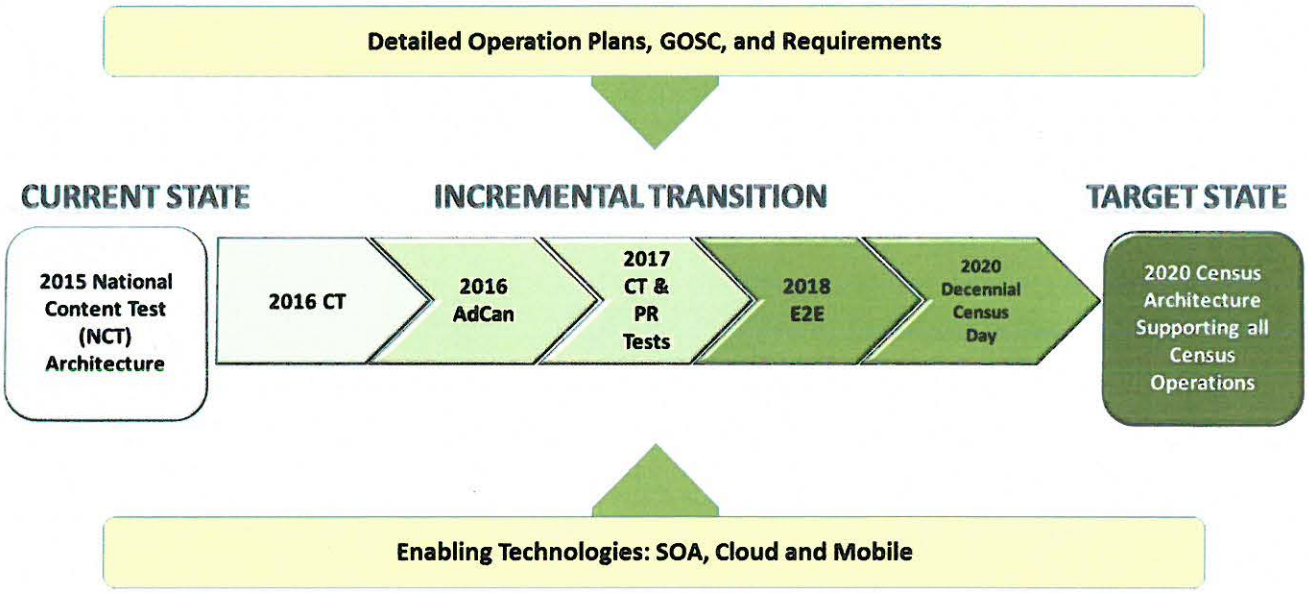


Figure 3-7: 2020 Census Architecture Incremental Transition Model

3.8.3 Transition Tasks & Activities

To ensure timely execution of each test and census, the Architecture Team will continue to execute on on-going tasks and activities. Some of these activities have been completed as part of an architectural deliverable or analysis; however, understanding the CQA, C-SHaRPS, and the Enterprise Cloud capabilities is incomplete. The architecture team and leadership plan to continually collaborate with other enterprise architecture programs and non-enterprise system owner to ensure the 2020 Census Architecture requirements are understood and the solution capability will deploy as scheduled. Table 2 documents the transition tasks and future activities.

Table 2 Transition Tasks and Future Activities

Activity	Description
Finalize Target State Application Architecture	1) Collaborate with Business Stakeholders and Business Requirements Team to analyze the Operations that are in the process of developing the Detailed Operational Plans, including requirements, Business Process Models, and Integrated Operation Diagrams. One venue is to participate in the IPT sessions. 2) Identify solutions that are needed for new capabilities, by leveraging existing systems, enhancing existing systems, or designing new systems, such as Island Areas, Coverage Measurement (CM), and Group Quarters (GQ), etc.
Disaster Recovery and COOP	Collaborate with the COOP Team to establish and analyze the COOP and Disaster Recovery requirements.

Activity	Description
Scale-up Project Next Steps & Continuous Quality Improvement	Establish a Working Group to plan and execute the Scalability Framework encompassing the overall 2020 Solution Architecture, and all layers of the technology stack.
IT Infrastructure Support	a) Working Group established to design the IT Infrastructure, including Field offices, in order to support the 2020 Census Operations and systems in terms of capacity, scalability, reliability and system administration and monitoring. b) This effort will include the analysis and design of a failover site for Disaster Recovery and COOP. c) Ensure the readiness of IT Infrastructure by aligning the Census Test dates with the Technology Refresh phases.
Cloud Transition	Collaborate with Enterprise Services Framework Team and IT Divisions to establish a systematic plan to transition suitable systems that support 2020 Census to the Cloud environment.

4 Current State Architecture

The current state 2020 Census architecture is based on the 2015 National Content Test (NCT) technical baseline and the infrastructure that supports the ongoing Geographic Programs. This architecture serves as the starting point for subsequent architecture developments, deployed to support a series of large-scale Census tests planned for 2015-2018 and to support upcoming changes to Geographic Programs during those years. The 2015 NCT architecture consists of mostly fielded legacy application enhancements, with some newer application development efforts introduced for test evaluation. The 2015 NCT was selected as a baseline since it was the most current. As a result, the architecture was based on the needs of the 2015 NCT and the systems available.

The 2015 NCT was used for content testing, testing different contact strategies aimed at optimizing self-response, and testing different approaches for offering in-language materials. The test also included Puerto Rico but did not include Nonresponse Followup (NRFU) or other field operations. Note that there is a separate geographic programs infrastructure that is also undergoing evolution simultaneously, and where relevant, the linkages between the Current State changes and geographic programs will be noted.

4.1 Business Architecture

The information flows among the primary business operations for the 2015 NCT are highlighted in the Figure 4-1. Major interactions and flows are shown via the arrows in the diagram and the key external interfaces are depicted via labeled icons.

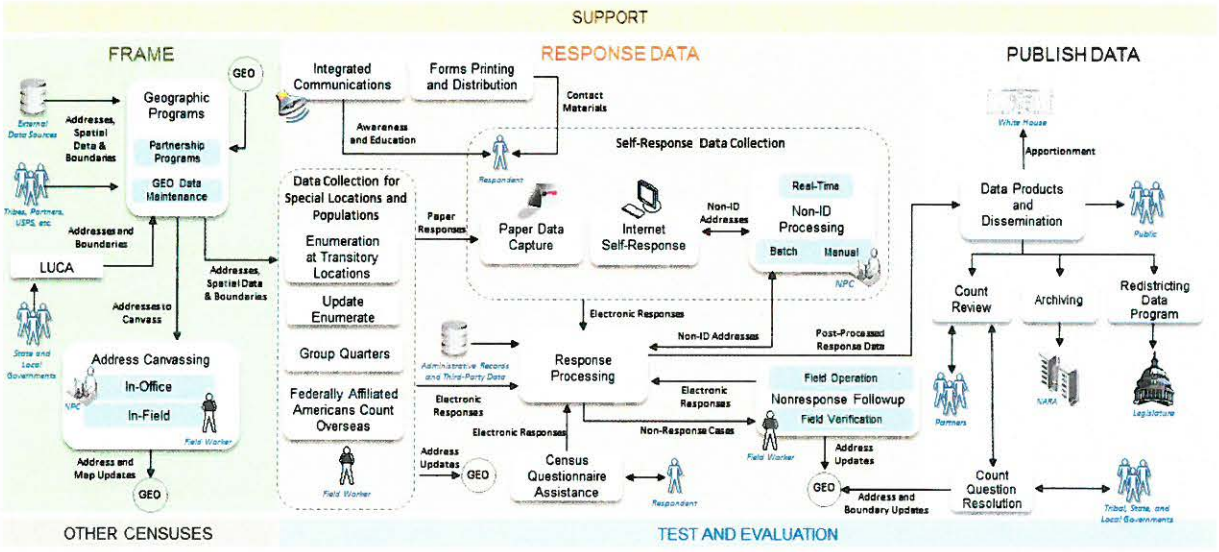


Figure 4-1 High-Level Integration of Operations for 2015 NCT

The operation that was primarily being informed by the 2015 NCT was the Content Forms and Design (CFD) and the contact strategy which is currently developed in Internet Self-Response (ISR), but executed through Forms Printing and Distribution (FPD). Refer to Appendix K: Architecture Transition

Table for a complete list of operations that supported this test. The other operations that participated in this test were implemented to support this test, but were not the focus of the test in terms of answering the design questions.

The 2015 NCT includes multiple operational areas that collectively provide the anticipated capabilities that map to the end-to-end survey lifecycle and are grouped into five categories: support, frame development, response data collection, disseminate data, and test, evaluation, and other censuses.

- The Support segment addresses the needs for HR, IT, and Infrastructure support services of the 2020 Census. Support is comprised of the program management, survey engineering, and infrastructure operational categories.
- The Frame development segment entails activities within the geographic programs, LUCA, and Re-Engineered Address Canvassing. These activities support the task of developing administrative records frame, geographic delineations, and address canvassing. LUCA was not part of the 2015 NCT, although it is part of the Geographic Programs.
- The Response Data Collection segment is the largest segment within the 2020 Census Architecture where thirteen distinct operations work together to achieve the data/response information collection goals. The response data collection segment includes a multitude of enterprise and non-enterprise systems, the majority of which are scoped within CEDCaP.
- The Disseminate data segment is responsible for activities such as accurate production, review, and dissemination of the data collected by the response data collection segment and shared with the stakeholders. The disseminate data segment was not involved in the 2015 NCT.
- The test, evaluation, and other censuses segments include five operational areas with activities that address coverage measurements, matching, follow-ups of the housing units and interviews, enumerations, and experiments. The test, evaluation, and other censuses segments were not involved in the 2015 NCT.

4.2 Application Architecture

The application architecture is the set of application areas identified to support the 2015 NCT. It provides a context for system development or enhancement. Application areas represent a useful subdivision of activities, services, and data that can be linked to other objects in the architecture. They are a group of activities and entity types with strong interdependencies such that a single application or more than one application can support the area.

The list of systems participating in the 2015 NCT can be found in Appendix K.2: Application Architecture: Systems. Figure 4-2 shows the 2015 NCT Solution Architecture.

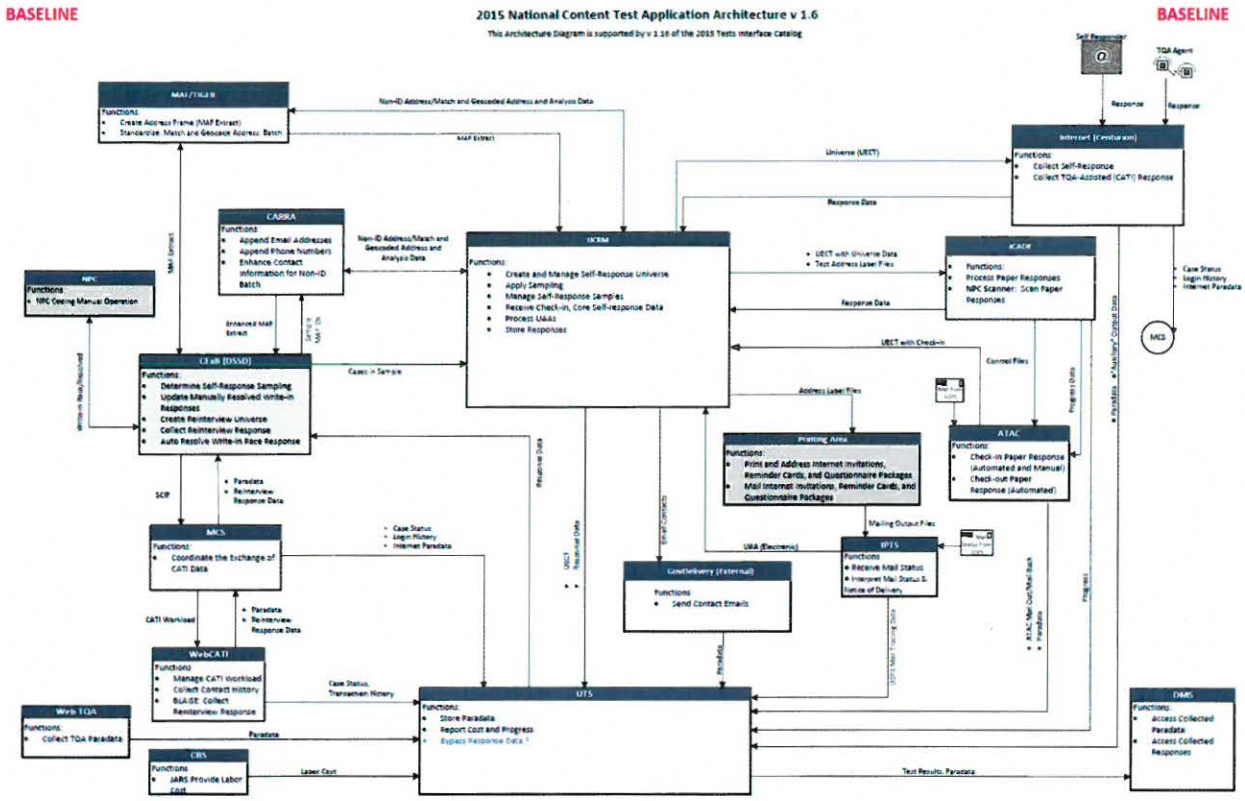


Figure 4-2 2015 NCT Solution Application Architecture Model

In the geographic programs infrastructure, there were upgrades to the geographic analytics, and three services were introduced: address matching, address analyzing and geo-coding. In CY 2015, GUPS was introduced as a new tool to interface with partners and supported the ongoing geographic programs such as BAS and SDRP.

4.3 Information Architecture

The 2015 NCT information architecture is designed to ensure that data is efficiently and effectively managed, maintained and kept consistent across the survey lifecycle.

Information in the 2015 NCT was managed in the native formats of the component systems used in the test. This includes the questionnaire instrument data, respondent data, and metadata about the respondent data. In addition, paradata was collected and stored in the Unified Tracking System (UTS), for use by Census researchers in the analysis of 2015 field test performance. This includes data about the cost and progress of the survey, and detailed information about contact events between field representatives and respondents.

4.4 Infrastructure Architecture

The 2015 NCT conforms to the Technical Reference Model (TRM) to guide the design of operational and development environments and communication services for the 2020 Census solutions.

Infrastructure encompasses several distinct domains of functionality and technology. The domains are used in the TRM to categorize the different technologies. Please refer to Table 31 in Appendix L: TRM Domains used in 2015 NCT which shows the TRM domains used by the systems in the 2015 NCT architecture.

4.5 Security Architecture

The 2015 NCT architecture was designed to be in compliance with all Census Bureau, Department of Commerce, and applicable government security standards for operational systems, including Federal Information Security Management Act (FISMA) which requires that all federal agencies document and implement controls for information technology systems that support their operations and assets, and the National Institute of Standards and Technology (NIST) Special Publication (SP) 800-37 Revision 1, which provides the framework for Risk Management & Security Assessment and Authorization process. The architecture was developed, and recommended for approval by the CIO following a review by the Office of Information Security. The architecture was designed to be compliant with requirements for Privacy and Confidentiality for the handling of Personally Identifiable Information (PII).

The 2015 NCT was designed to be conducted in compliance with all applicable laws and policies. The two primary applicable statutes are Title 13 and Title 5. Title 13 provides the basis for conducting the census and also ensures the privacy of respondents' personal information. Title 5 includes the Privacy Act of 1974 and further reinforces the legal basis for protecting respondent's personal information. In addition, the requirements under Title III of the E-Government Act of 2002, the Federal Information Security Management Act (FISMA) will be met by all IT systems supporting the Census.

4.6 Quality Architecture

Coverage Measurement was not a major focus of the 2015 NCT and will be investigated further in future tests.

Data quality is a major focus of the Geographic Programs infrastructure due to their involvement with checking for indications of growth and change and determining if the resources are available to make updates and capture those changes. In October 2015, in-office AdCan began and BARCA was introduced. Data quality in BARCA involved reviewing blocks to determine the accuracy of the MAF/TIGER database. Beginning in 2013, the Geographic Support System initiative established relationships with local and tribal governments to share address lists and ArcGIS files to capture growth and make periodic updates. This was used to keep the MAF/TIGER database up to date.

5 Target State Architecture

The 2020 Census architecture is described in the context of the seven domains as shown in Figure 5-1.

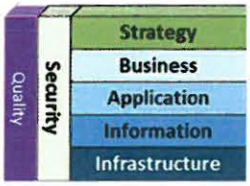


Figure 5-1 2020 Census Architecture Framework

The strategy domain consists of the program vision and mission level business requirements. This domain is not relevant to the Transition Plan, so the reader should refer to the 2020 Census Architecture Document, Section 4.2 for more information on the strategy domain. The other domains are described in the sections that follow.

5.1 Business Architecture

The information flows among the primary business operations are highlighted in Figure 5-2. Major interactions and flows are shown via the arrows in the diagram and the key external interfaces are depicted via labeled icons.

Although each operation is presented separately, the operations must work together to achieve a successful Census. Information flows among the operations as the census proceeds from frame development through collection of response data to the publishing and release of the data.

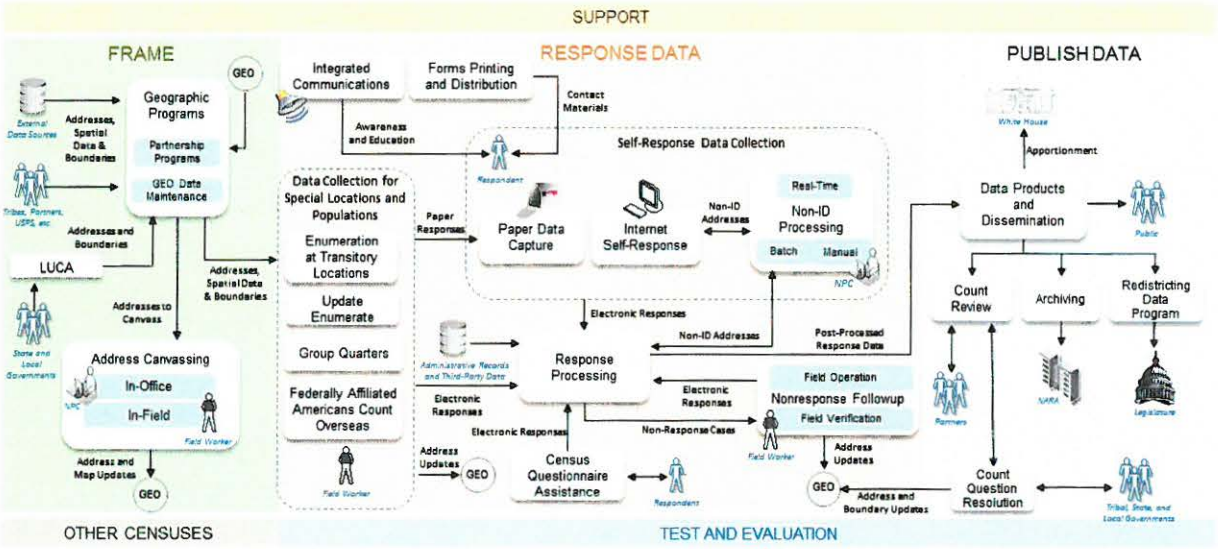


Figure 5-2 High-Level Integration of Operations

The 2020 Census Business Architecture describes the functional business process, the relationship, and the geography aspects of the business environment. It is decomposed into multiple operational areas that collectively provide the anticipated capabilities that meet the end-to-end survey lifecycle phase and is grouped into five categories: support, frame development, response data collection, disseminate data, and test, evaluation, and other censuses.

1. The Support segment addresses the needs for HR, IT, and Infrastructure support services of the 2020 Census. Support is comprised of the program management, census/survey engineering, and infrastructure operational sub-categories.
2. The Frame development segment entails activities within the geographic programs, LUCA, and Re-Engineered Address Canvassing. These activities support the task of developing administrative records input, geographic delineations, and address canvassing.
3. The Response Data Collection segment is the largest segment within the 2020 Census Architecture where thirteen distinct operations work together to achieve the data/response information collection goals. The response data collection segment includes a multitude of enterprise and non-enterprise systems, the majority of which are scoped within CEDCaP.
4. The Disseminate data segment is responsible for activities such as accurate production, review, and dissemination of the data collected by the response data collection segment and shared with the stakeholders, the majority of which are scoped within CEDSCI.
5. The test, evaluation, and other censuses segment includes five operational areas with activities that address coverage measurements, matching, follow-ups of the housing units and interviews, enumerations, and experiments.

Figure 5-3 shows how the operations are divided into the business architecture categories. More details about the Business Architecture can be found in Section 4.3 of the 2020 Census Architecture Document.

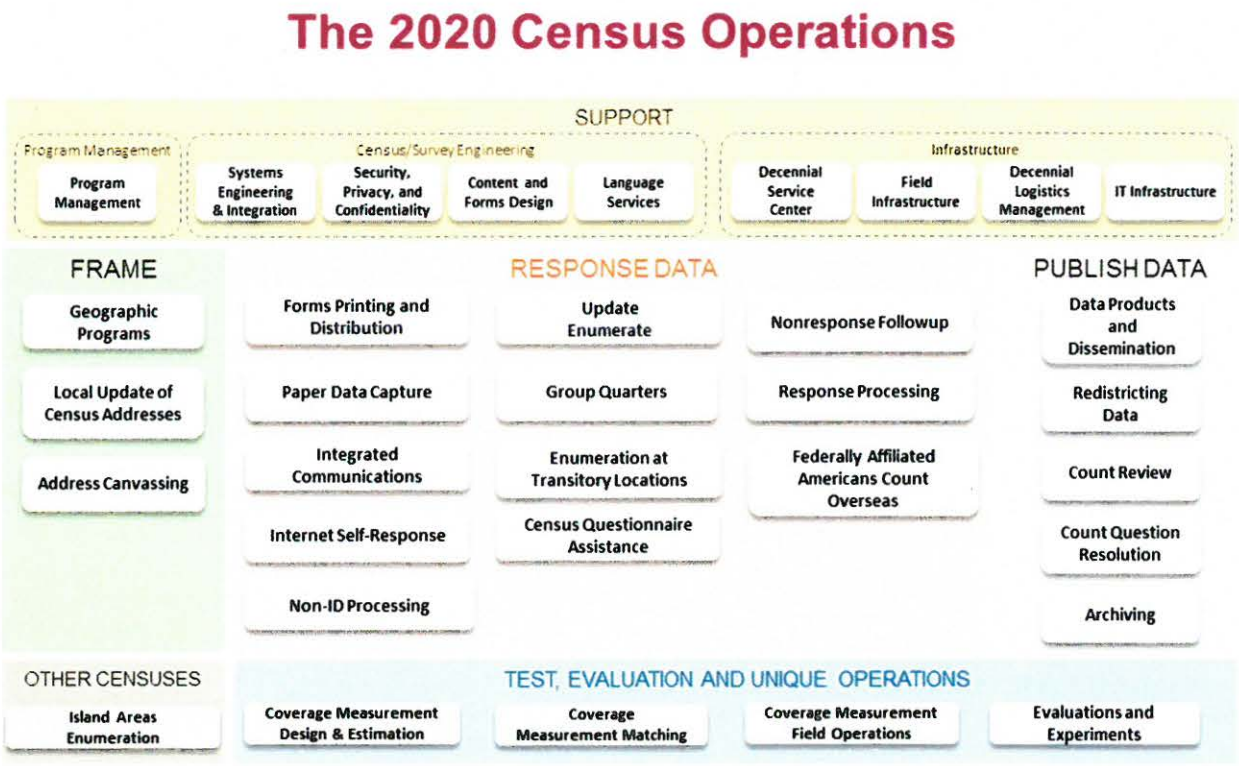


Figure 5-3 2020 Census Operations

5.2 Application Architecture

The application architecture is the set of application areas identified to support the 2020 Census. It provides a context for system development or enhancement. Application areas represent a useful subdivision of activities, services, and data that can be linked to other objects in the architecture. They are a group of activities and entity types with strong interdependencies such that a single application or more than one application can support the area.

The application architecture consists of both legacy application enhancements and new application development efforts. Projects such as CEDCaP will replace multiple legacy systems used during the 2010 Decennial Census, while other legacy applications such as CIRA will remain to support the 2020 Decennial Census.

More details about the Application Architecture are found in Section 4.4 of the 2020 Census Architecture Document. The list of systems participating in the 2020 Census can be found in Appendix I: 2020 Census Systems List. The list of outstanding questions that impact solutions can be found in Appendix J: Outstanding Decisions with Potential Impact on the Census Architecture. Figure 5-4 and Figure 5-5 show the 2020 Census Solution Architecture Model and the 2020 Census Support and Infrastructure Model.

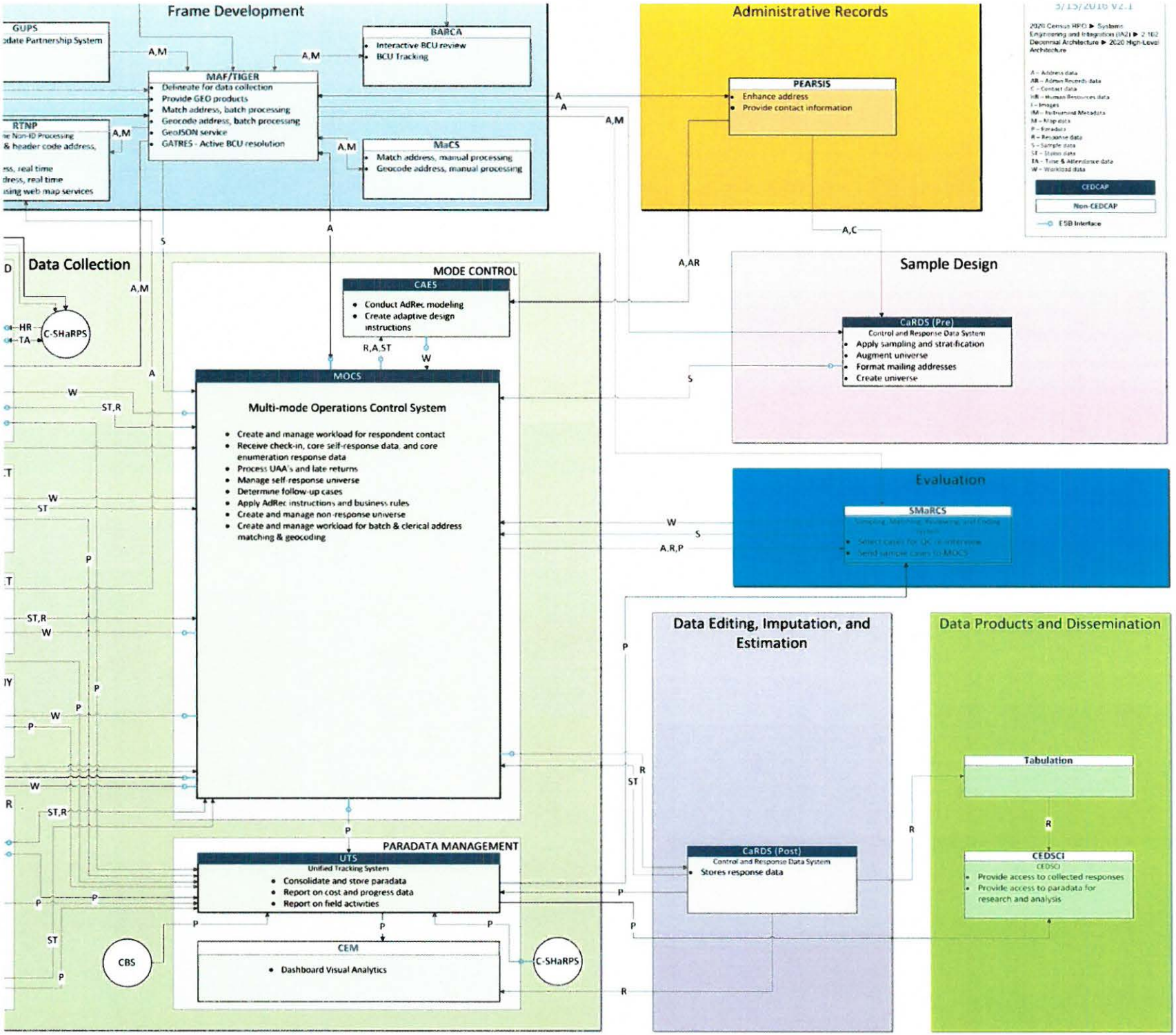


Figure 5-4 2020 Census Solution Architecture Model

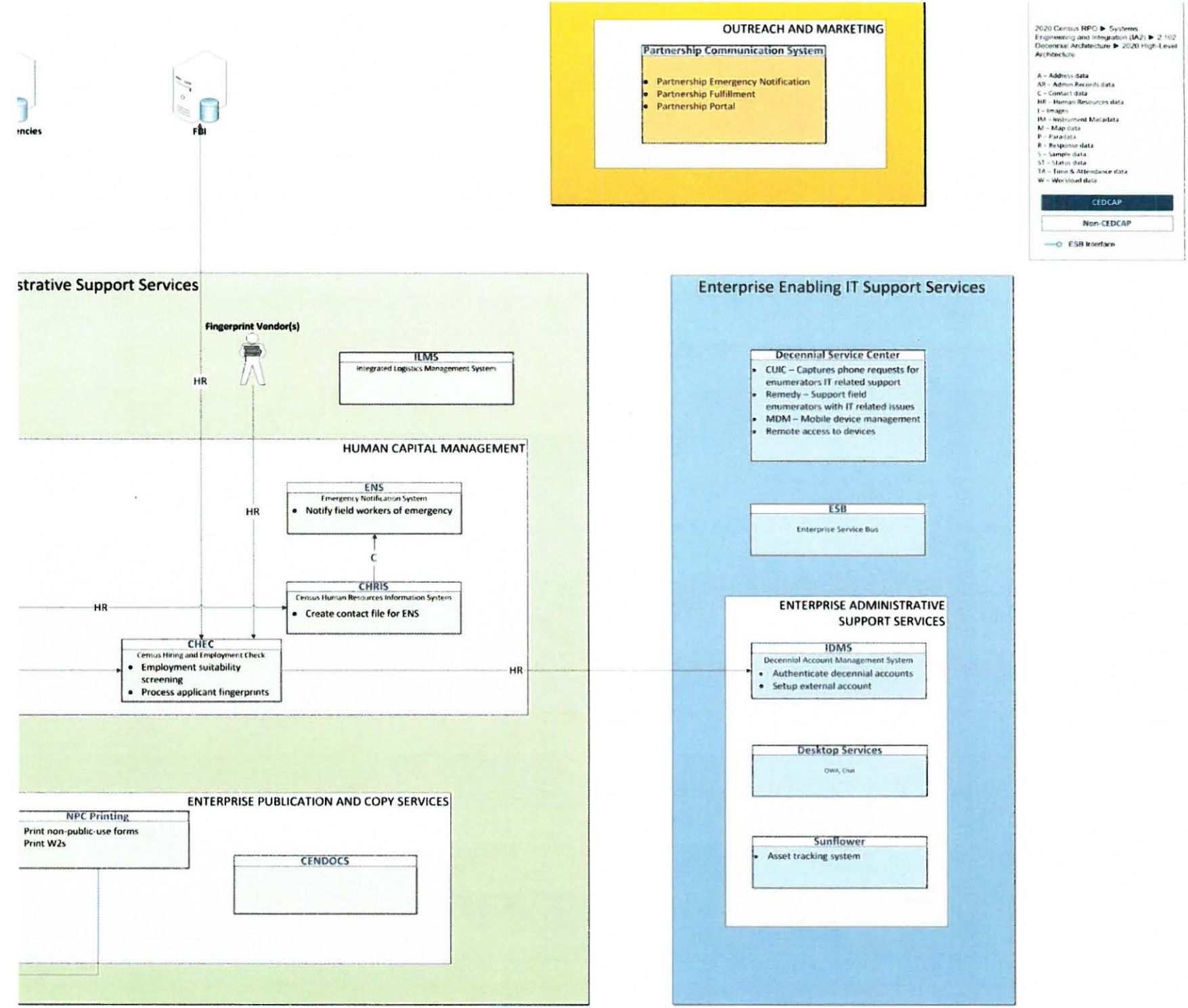


Figure 5-5 2020 Census Support and Infrastructure Model

5.3 Information Architecture

The 2020 Census information architecture focuses on ensuring that data is efficiently and effectively managed, maintained, and kept consistent across the survey lifecycle. The 2020 Census focus is on ensuring the accuracy, completeness, consistency, and reliability of the data to provide stakeholders with accurate and timely data. The data architecture has several key components, at the foundation are the adoption and extension of statistical data standards.

The 2020 Census information architecture systems will adhere to the common Enterprise Data Model (EDM) and use the Enterprise Data Registry. The 2020 Census Data Architecture will adhere to and use the EDM. Migration of legacy data is out of scope for the 2020 Census. Old, existing data will not be retrofitted to the EDM.

All systems will reference the Enterprise Data Registry via a lightweight Application Programming Interface (API) for the representation of Enterprise Data elements. The system interfaces will support the new APIs, and new data will be compatible with this format. At least some portion of the new data model will be used for the 2017 Census Test, and the API-based Enterprise Data registry will be in place for the 2020 Census.

The proposed data architecture of the Census Bureau is built upon the concept of a structured, machine-readable specification at every major segment of the Survey Life Cycle. The specification will be machine readable by implementing the EDM in an XML schema document.

Finally, the proposed data architecture will make use of Interface Control Documents (ICDs) which are a set of artifacts that define the content and structure of a message passed between two systems across the ESB. The Enterprise Data Model will standardize the concepts and relationships in the ICDs.

5.4 Target State Infrastructure Architecture

The Census Bureau's Technical Reference Model (TRM) guides the design of operational and development environments and communication services for the 2020 Census solution.

5.4.1 Infrastructure Domains

Infrastructure encompasses several distinct domains of functionality and technology. The domains are used in the TRM to categorize the different technologies. The domains are as follows:

- Application Technology – Standards and software applications, which support the development, and integration of software applications
- Collaboration and Electronic Workplace – Software applications, which promote collaboration and knowledge sharing
- Data Management Technology – Standards and software applications, which provide or support the capture, storage, retrieval, management, and analysis of data
- Networking and Telecommunications – Standards, software applications, and hardware that provide or support computer networking and telecommunications
- Infrastructure Platforms and Storage – Standards, software applications, and hardware that support digital data computation and storage

- Security – Standards, software applications, and hardware, which provide or support information security, computer security, and information assurance
- Service Management – Standards and software applications, which support the management of software applications and hardware

Figure 5-6 illustrates the infrastructure domains and associated sub-domains within the TRM.

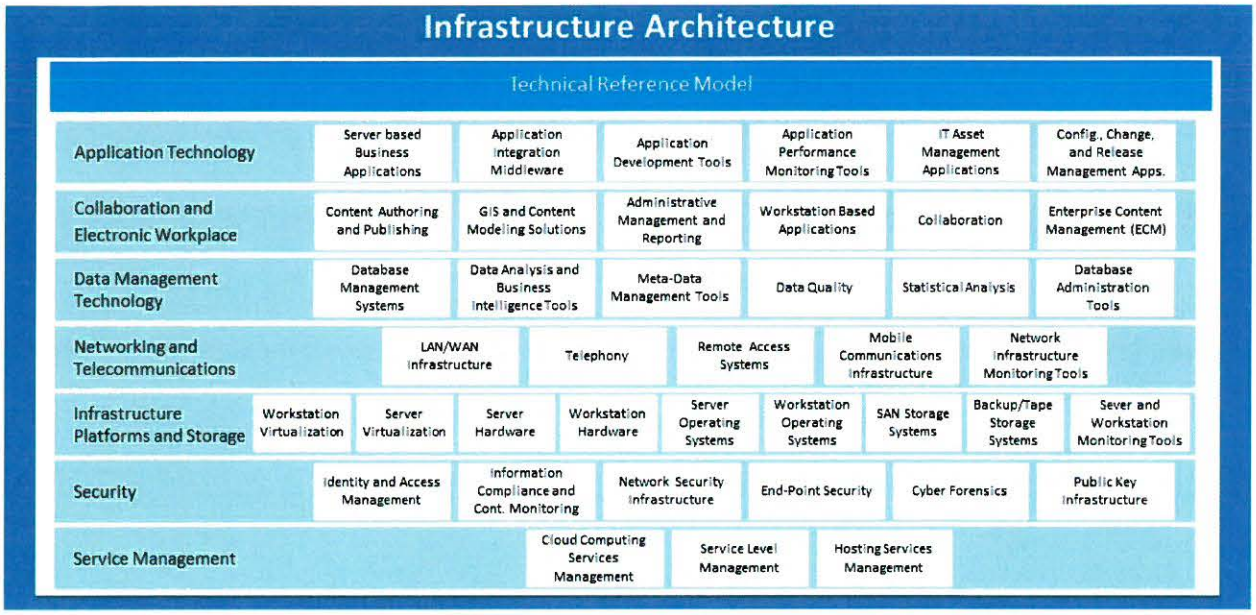


Figure 5-6 TRM Domains and Sub-domains

5.4.2 Key Infrastructure Changes

Between the 2015 NCT and ongoing Geographic Programs current state and the 2020 Census, several key infrastructure changes occur. The following list contains the currently known key infrastructure changes. As the design matures, the following list will be updated to reflect the revised design.

- For the 2020 Census, PRIMUS and RTNP will be deployed in the cloud. MOJO, MOCS, MCM are being considered for the cloud by 2020.
- LiMA will move from being deployed on a laptop to being deployed on a handheld.
- The new CQA infrastructure will be introduced to provide online assistance.
- UTS and CEM will be using Tableau. UTS will be based off of CEDaR.
- CHEC will interface with multiple external interfaces and third-party vendors.

For a full list of changes, refer to the Architecture Table in Appendix K: Architecture Transition Table. Field office infrastructure transitions are covered in the Field Office Infrastructure documentation.

5.5 Security Architecture

The Federal Information Security Management Act (FISMA) requires that all federal agencies document and implement controls for information technology systems that support their operations and assets. Standards and guidelines have been developed and published by the National Institute of Standards and Technology (NIST) to assist and to guide the federal agencies in adhering to the FISMA. NIST introduced a framework for managing information system-related security risks in NIST SP 800-37, Guide for Applying the Risk Management Framework to Federal Information Systems. The Census Bureau Risk Management Framework (RMF) document describes the application of the NIST Risk Management Framework to information systems within the organizational boundaries of the Census Bureau.

The RMF steps will be performed in conjunction with or as part of SDLC processes. Not only is early implementation of security in the SDLC integral to the RMF process, but the RMF itself provides a framework for managing security considerations throughout the life cycle. As a condition of reaching the target state, a system must be vetted by security and be FISMA-compliant. The interfaces to other systems must have been thoroughly tested and verified, and the agreements with external interfaces must have been established. For more details on Security and the FISMA Risk Framework, refer to the 2020 Census Architecture Document, Appendix C.

The Security Architecture will abide by Census Bureau and Department of Commerce standards, rules and regulations. Please refer to Section 4.7 of the 2020 Census Architecture Document for more information on the security architecture.

5.6 Quality Architecture

5.6.1 Data Quality Management

The primary goal of the 2020 Census is to collect complete, accurate, and quality Census data. While data quality management is a systematic and ongoing effort, it starts at the point of entry from the user and/or through interfaces and continues throughout each system.

5.6.2 Data Quality in Systems

Various systems contribute to the quality aspect of the 2020 Census in two ways. The first way is through systems that support Quality Control. The Sampling, Matching, Reviewing, and Coding System (SMaRCS) application specifically supports quality control operations designed to determine whether field listers and enumerators are using validated procedures and collecting accurate data. The SMaRCS application will facilitate quality control operations by providing a mechanism for selecting quality control samples, validating production interview data against administrative records sources, and by providing a tool for clerical matching to compare the production interview data against re-interview (RI) data. SMaRCS will participate in Listing Quality Control by looking for "using validated procedures" and for falsification. SMaRCS will also serve as a major control component for quality control operations by managing the selection of quality control samples for field follow-up related to Census and Census Coverage Measurement (CM) operations and tracking the progress of the RI work through the matching, field, and resolution processes. Another example of QC is PRIMUS (which replaces Centurion) and MOCS which perform data validation from end to end.

The second way is through systems that have improved algorithms to support quality and efficiency modeling, such as in the Concurrent Analysis and Estimation System (CAES). The CAES will be an enterprise modeling platform which uses adaptive design to store data and uses it to execute statistical models in support of survey flow processing, analysis, and control. Systems such as the Control and Response Data System (CaRDS) (Post processing) will maximize the degree of completeness of the information by utilizing admin records. The CaRDS will act as the front and back door to data collection. It will pull personal or business data from the Census Master Address File and send it to MOCS. The CaRDS will provide Universe determination for the Decennial, and post data collection response processing.

6 Transition Timeline and Phases

This section provides the timeline and phases for the Census transition from the current state as of 2015 to support the planned testing in 2016, 2017, and 2018, and ultimately the 2020 Census. The transition is accomplished via a series of phases, each containing one or more major test events as shown in Figure 6-1.

Each test event has associated system and capability enhancements. This section describes the key features of each major test event and highlights the important aspects of new or transitioned capability associated with each. Additional details about each of the tests can be found in the appendixes.

Note that in addition to transitioning to support the Census Tests, other geographic programs are undergoing transition in support of ongoing programs such as in-office address canvassing, BAS, and SDRP, and in preparation for 2020 programs such as LUCA, PSAP, VTD, and RDP.

The overall Census test plan spans five years, from 2015 to 2020. Refer to the 2020 Census Architecture Document, Section 4.2 for more information on the planned tests for each year. The sequencing of the major test events is presented in Figure 6-1.

For each major test event, the figure lists the associated systems.

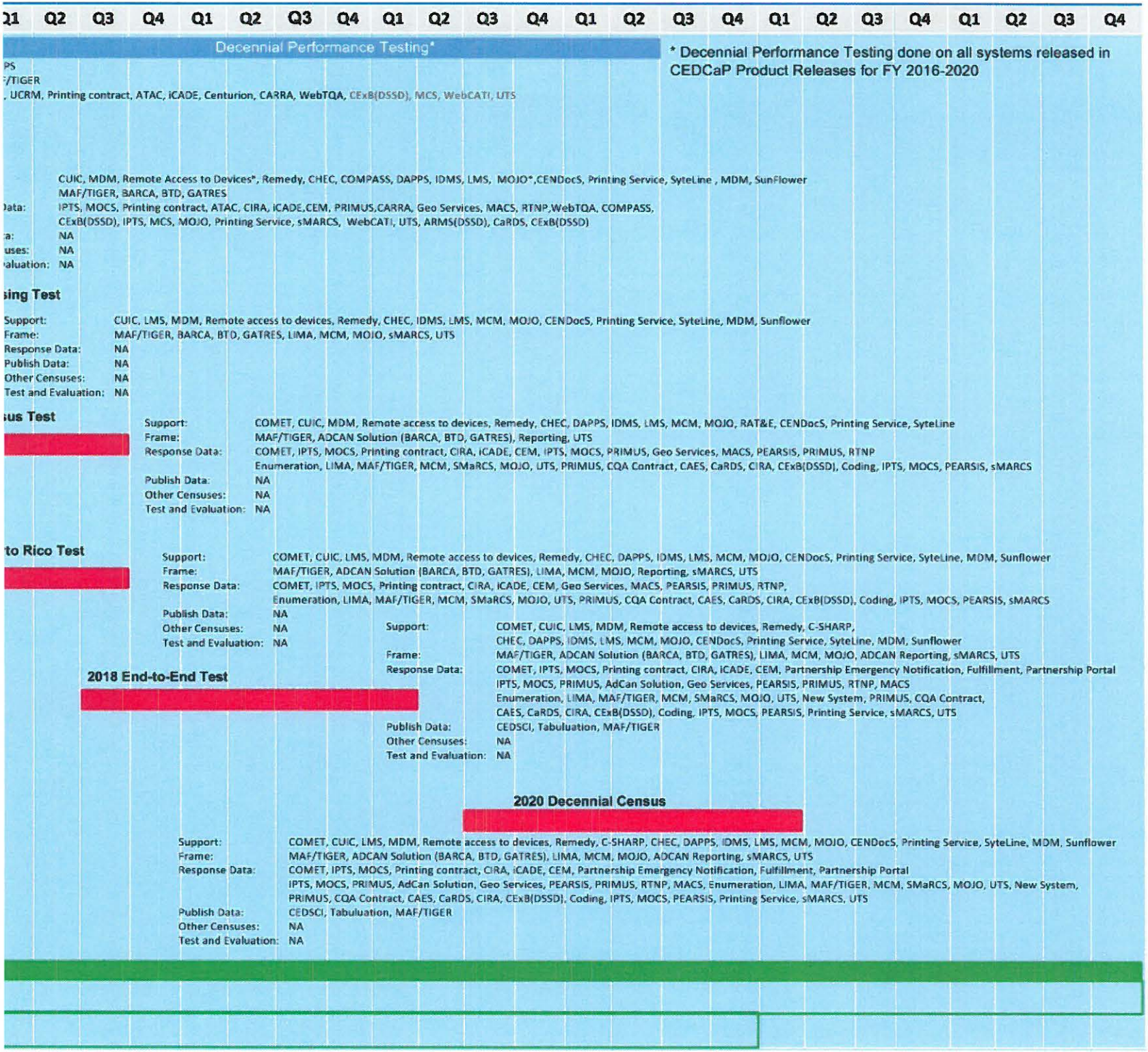
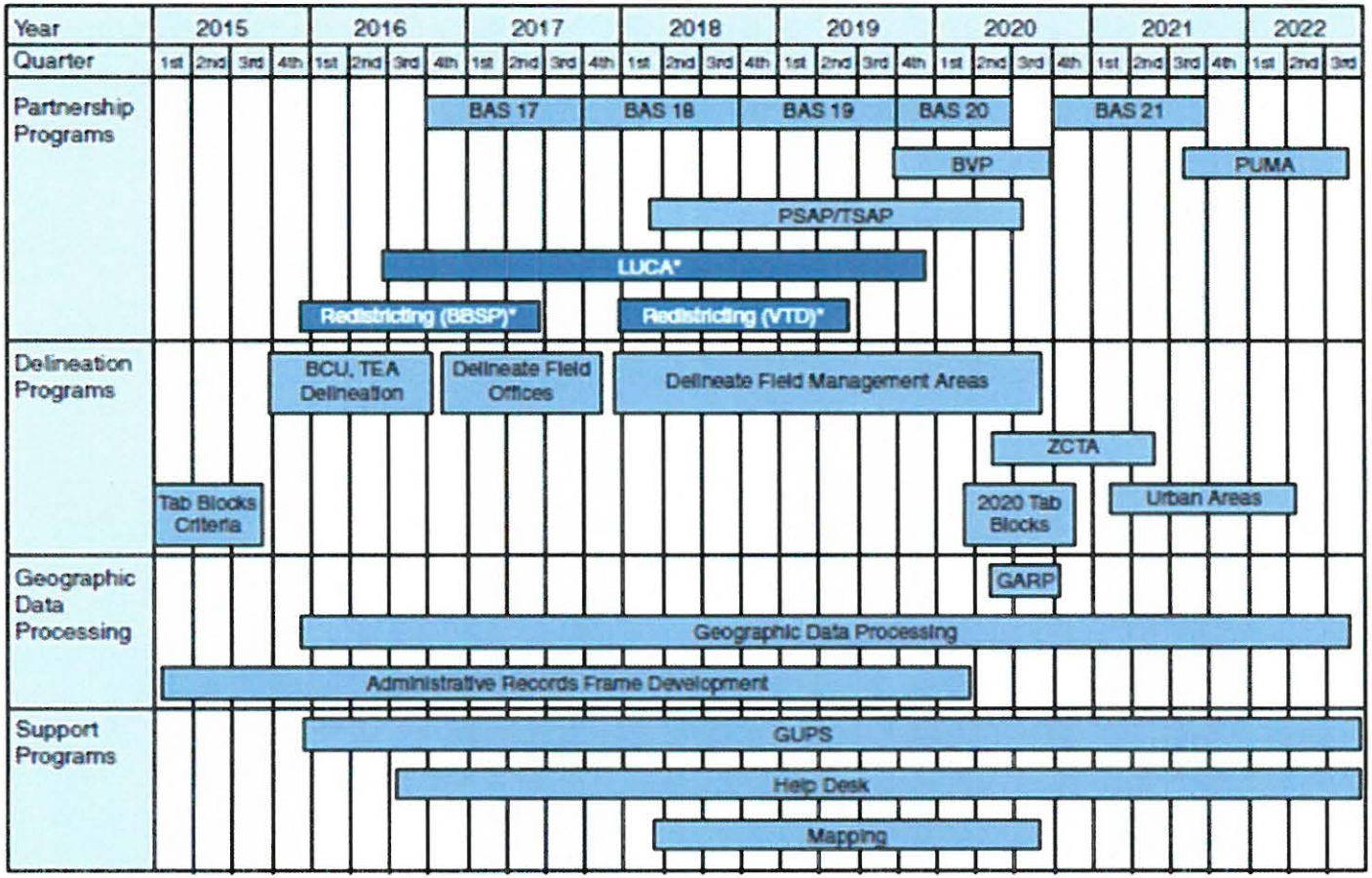


Figure 6-1 2020 Census Sequencing Timeline

During the 2015-2022 timeframe, the Geography Division will be developing and deploying a number of capabilities to support Census tests in preparation for the 2020 Decennial. A complete description of these planned activities can be found in the 2020 Census Architecture Document, Section 5.4.1, Geographic Programs. In particular, Figure 31 of the 2020 Census Architecture Document provides an excellent summary of the planned Geography activities; this figure is reprinted below in Figure 6-2.



*LUCA and Redistricting are Partnership Programs that are managed outside of the Geographic Programs Team.

Figure 6-2 Geographic Programs Timeline

During the transition period, the goal is to design, build, test and demonstrate the ability to scale up to full 2020 Census size. Figure 6-3 depicts the Scale-up readiness timelines that have been developed to show when the fully scalable solution will be ready for 2020. Some 2020 activities may start prior to August 2019, the first operational use of the fielded built out scalable solution will start with the In-Field Address Canvassing in August of 2019. The major milestones for building out the fully scalable and highly available solution are planned as follows:

- By June 2016, the Initial Demand Models and non-functional requirements will be defined to serve as input data to the scale-up process.
- By September 2016, the logical design of the scalable solution will be completed.
- By September 2017, the solution designed in September 2016 will be developed, tested and ready to be used for the 2018 End-to-End Test. While the physical configuration of the solution to be used for 2018 End-to-End Test will be commensurate to the limited scale as determined by the Test's operations, the logical architecture for 2018 will be the same as will be used for 2020. Thus, we will have the opportunity to exercise in production the scalable solution, and obtain insights and validation. The design of the solution will also be ready for a limited COOP test. COOP testing will be focused on systems that will be deployed in the cloud, but not on the entire solution, since the second site will be built out and tested in 2018.
- By September 2018, the fully scalable solution will be built out, along with the second failover site. The entire solution will be then ready for COOP testing. The solution at full production level will undergo testing throughout most of FY 2018.
- By August 2019, the highly available solution with all the capabilities and full scale will be ready to operate for the 2020 Census.

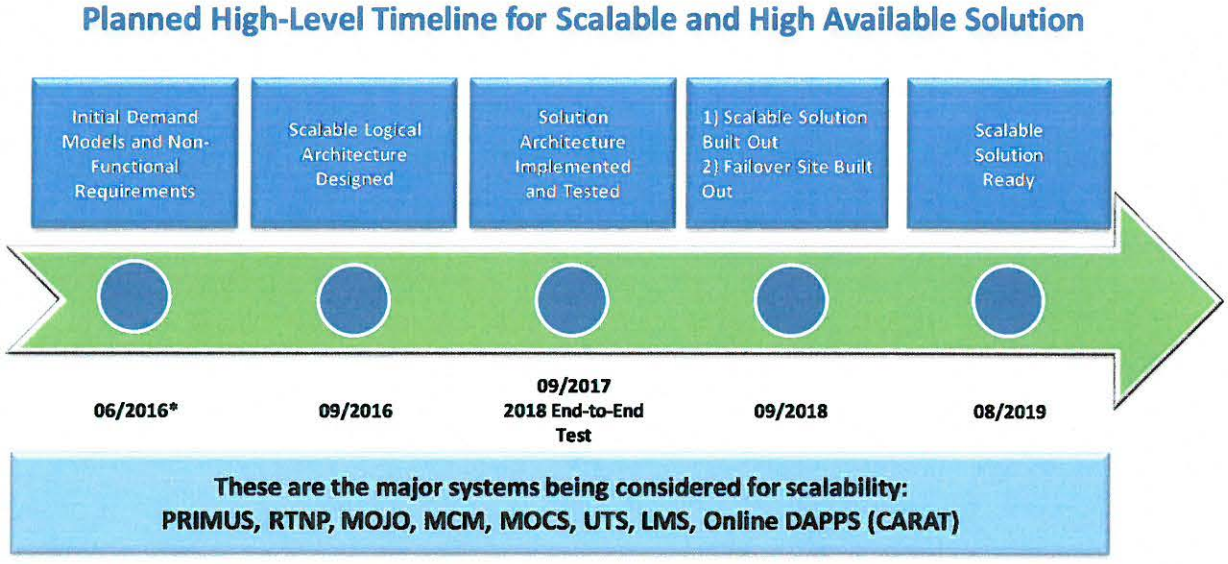


Figure 6-3 Planned High-Level Timeline for Scalable and High Available Solution

Figure 6-4 shows the planned high-level timelines for the Cloud Transition that the team has developed and lists some candidate systems for the cloud transition.

- By June 2016, the main systems to be transitioned to the cloud will have been identified and validated according to the Enterprise Architecture Cloud readiness guidelines. This includes both first and second tier candidate systems.
- By September 2016, the logical architecture of the systems to be transitioned to the cloud will be evaluated, and necessary modifications completed.
- By September 2017, the solution architecture will complete its implementation and testing in the Cloud environment in preparation for the 2018 End-to-End test. The solution will be provisioned in the cloud to satisfy the demand of the 2018 End-to-End Test, instead of full scale.
- By September 2018, the full scale cloud implementation will be available, and undergo testing and monitoring.
- By August 2019, the full cloud solutions are fully tested and available for use at full scale.

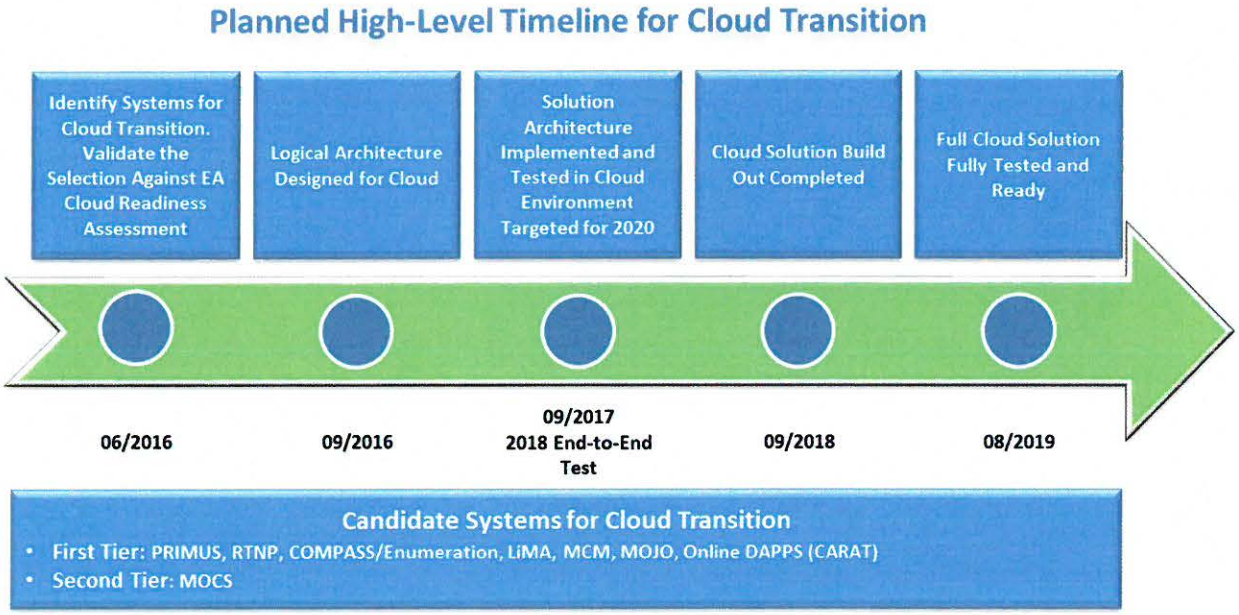


Figure 6-4 Planned High-Level Timeline for Cloud Transition

6.1 2015 National Content Test (Current State)

The 2015 National Content Test (NCT) was conducted as a preliminary mid-sized test of operations and systems. In order to provide a starting point for subsequent phases of the transition plan, the 2015 NCT is considered the baseline for this transition plan, and is considered the "Current State" as of the time of this writing. Note that there is a separate geographic programs infrastructure that is also undergoing evolution simultaneously, and where relevant, the linkages between the Current State changes and geographic programs will be noted.

The 2015 NCT evaluated and compared different census questionnaire content. It assumed a Census Day of September 1, 2015. The test included 1.2 million nationally-representative households, including 20,000 households in Puerto Rico and 100,000 re-interviews.

The operation that was primarily being informed by this test was the Content and Forms (CFD) and the contact strategy, which is currently developed in Internet Self-Response (ISR), but executed through Forms Printing and Distribution (FPD). The other operations were implemented to support this test, but were not the focus of the test in terms of answering design questions. Figure 6-5 showing the operations and associated systems is provided below.

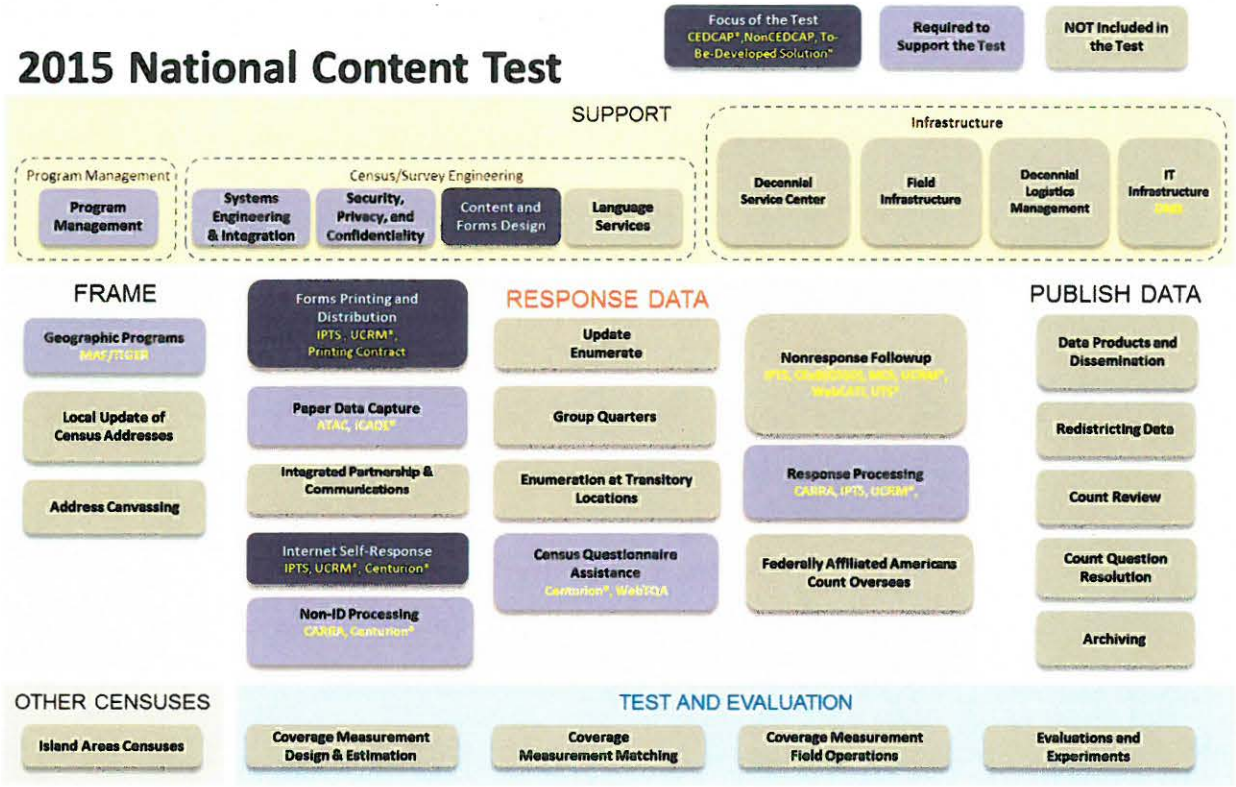


Figure 6-5 2015 National Content Test

6.1.1 Main Features

The main focus of the 2015 NCT architecture was to test out multiple content forms with multiple mail out strategies. Table 3 shows the main features of the test.

Table 3 Main features of the 2015 NCT

System	Description	Program
ATAC	Check-In (Automated and Manual)	N/A
CARRA	Append Email addresses Append Phone numbers Enhance contact information for Non-ID Batch	N/A

System	Description	Program
CBS	Provide labor cost (JARS)	N/A
Centurion	Collect Self-Response Collect TQA-Assisted (CATI) Response	CEDCaP
CexB (DSSD)	Determine Self-Response Sampling Update Manually resolved Write-in Responses Create Re-interview Universe Collect Re-interview Universe Conduct data analysis	N/A
DMS	Access Collected Paradata Access Collected Responses	N/A
GovDelivery	Send Confirmation Emails	External
iCADE	Capture Paper Responses (OCR/OMR/KFI) Reverse check-in for Blank Responses	CEDCaP
IPTS	Receive Mail Status Interpret Mail Status & Notice of Delivery	N/A
MAF/TIGER	Delineate Collection Boundary Create Address Frame (MAF Extract) Standardize, Match and Geocode Address, Batch	N/A
PhoneTree	Support outbound calls	N/A
Printing	Spray address (Labeling Equipment) Merge static forms and variable data, print (DocuTech)	N/A
UCRM	Create and Manage Self-Response and Non-Response Universes Apply Sampling Manage Self-Response Samples Receive Check-in, Core Self-response Data Process UAAs and Late Returns	CEDCaP
UTS	Store Paradata Report Cost and Progress	CEDCaP
Web TQA	Collect TQA Paradata	N/A
WebCATI	Collect contact history Manage CATI workloads Used for Re-interviews	N/A

The 2015 NCT included upgrades to the geographic analytics. Three services were introduced: address matching, address analyzing and geo-coding. In 2015, GUPS was introduced as a new tool to interface with partners and supported the ongoing geographic programs such as BAS and SDRP.

6.2 2016 Census Test

The 2016 Census Test is designed to build on the 2015 NCT results and previous Census Tests and introduce new systems and capabilities into the operational suite. The 2016 Census Test focuses on the

integration of Self-Response and Nonresponse Followup operations. A diagram of the 2016 Census Test operations and systems is provided below in Figure 6-6.

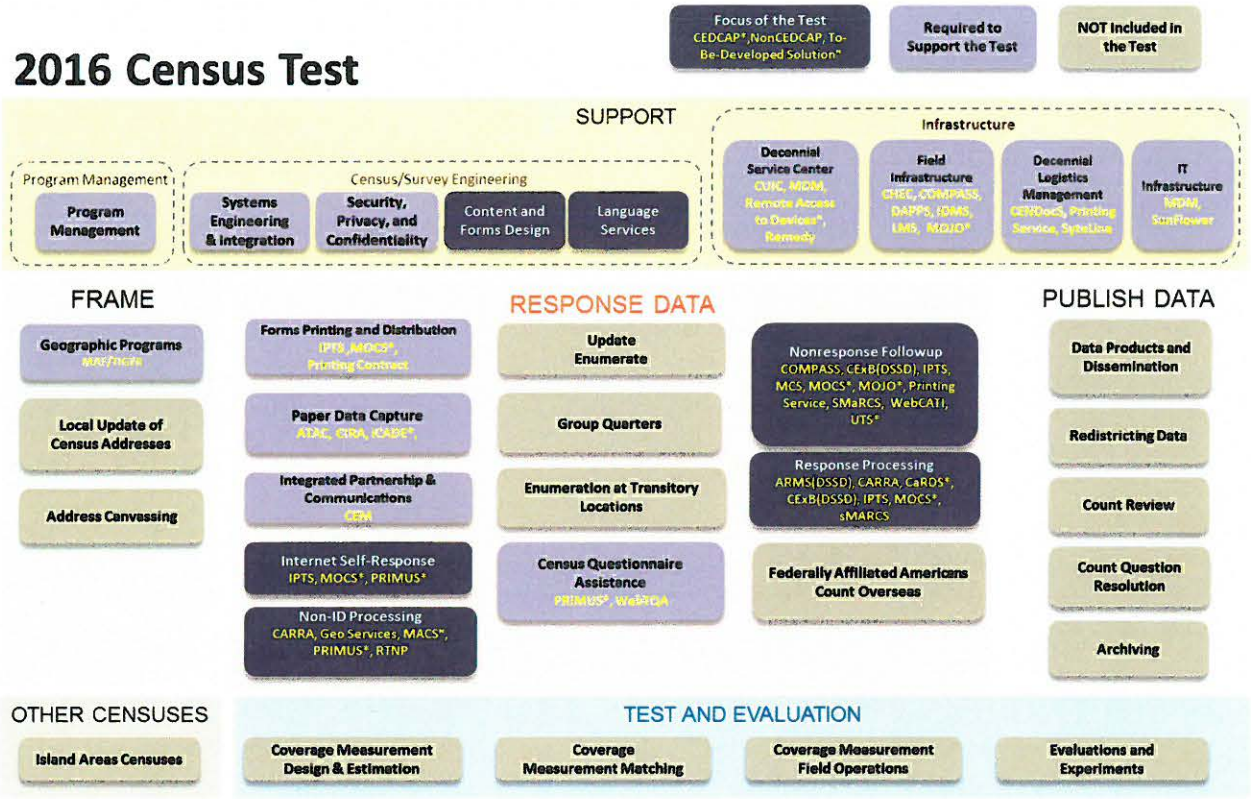


Figure 6-6 2016 Census Test

6.2.1 Main Features

The main features of the 2016 Census test include refining the NRFU strategy and introducing CEDCaP. Table 4 shows the main features of the test.

Table 4 Main features of the 2016 Census Test

System	Description	Program
MOCS	Controls data collection modes for all field operations	CEDCaP
MOJO	Controls mobile data collection	CEDCaP
PRIMUS	Replaces Centurion	CEDCaP
BARCA	First use during a test, although it is part of the on-going Geographic programs Basic Collection Unit (BCU) introduced as a new unit of geography. Some systems will be required to support the new BCU format.	N/A
DAPPS	Used as-is	N/A
LMS	Used as-is	N/A
CHEC	Used as-is	N/A
CaRDS	Introduced for sampling and response processing	CEDCaP

System	Description	Program
SMaRCS	Introduced, based on the MaRCS 2010 system, performs re-interview QC sample selection and re-interview case matching, to detect falsification.	N/A
CARRA	Used as-is, supports use of administrative records for modeling and optimizing NRFU workload operation	N/A
CEM	Introduced for data analytics and reporting about customer experience. Plans to import data from UTS (or some move towards a common data source) and will be importing response data from CaRDS	N/A
MaCS	Introduced to support manual matching and address geocoding of Non-ID cases.	N/A

The CEDCaP product releases for the 2016 Census test are detailed in the CEDCaP Transition Plan, Section 4.3.3.

6.3 Address Canvassing (AdCan) Test

The Address Canvassing Test is designed specifically to exercise new features for address canvassing and to refine the in-office and in-field methods. This test is not intended as a full scale demonstration of end-to-end capabilities but rather focuses only on AdCan technologies. It will begin in the fall of 2016 and will continue into 2017. This test will cover various sites across the nation.

The Address Canvassing Test is now independent of other tests and will be conducted at two AC Test sites within the Continental US containing 4000 contiguous, non-zero-Housing Unit (HU) BCUs, one of which has been updated by GSS-I and one of which is experiencing growth. The results of the fieldwork will be used to update a copy of the MAF/TIGER DB.

A diagram of the operations and associated systems is provided below in Figure 6-7.

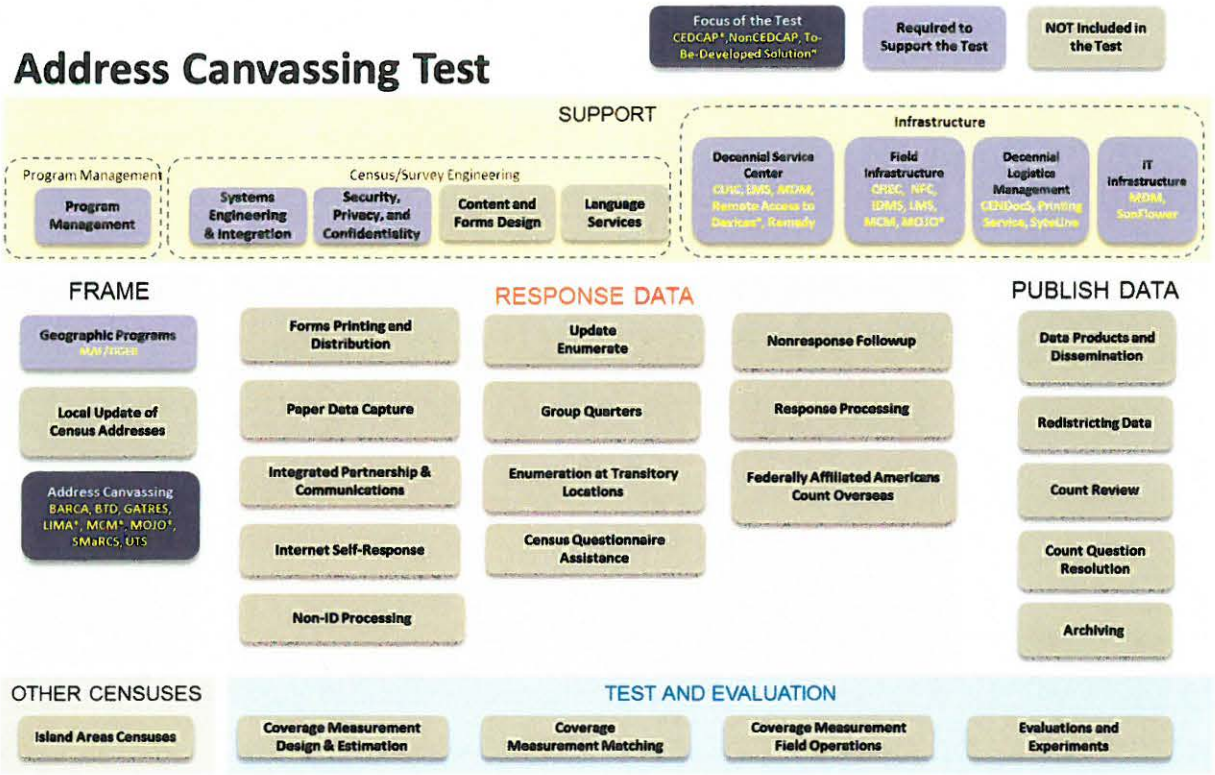


Figure 6-7 Address Canvassing Test

6.3.1 Main Features

The main features of the Address Canvassing test include the introduction of LiMA, CEDCaP systems coming on line, and testing of in-office canvassing. Table 5 shows the main features of the test.

Table 5 Main features of the Address Canvassing Test

System	Description	Program
LiMA	New platform for use of LiMA introduced. LIMA was initially built as a laptop system, but will be introduced on handheld for AdCan and for the AdCan portion of Puerto Rico Test. LiMA captures the GQ data for AdCan test that has not previously been captured, including listing not only the GQ addresses but also interacting with staff from the GQ to capture GQ type and contact information.	CEDCaP
National Finance Center	The National Finance Center performs payroll/personnel functions for Field Staff.	N/A
MCM	First used for mobile case management.	CEDCaP
SMaRCS	Introduced, used to perform sampling for the relisting operation.	N/A

The CEDCaP product releases for the 2016 Census test are detailed in the CEDCaP Transition Plan, Section 4.3.5.

6.4 2017 Census Test

The 2017 Census Test is another large scale test designed to exercise operations and systems. The 2017 Census Test is planned to be an operational study of only Update Enumerate systems and procedures. It will also introduce new systems using administrative records in a support capacity. A diagram of the operations and the associated systems is provided below in Figure 6-8.

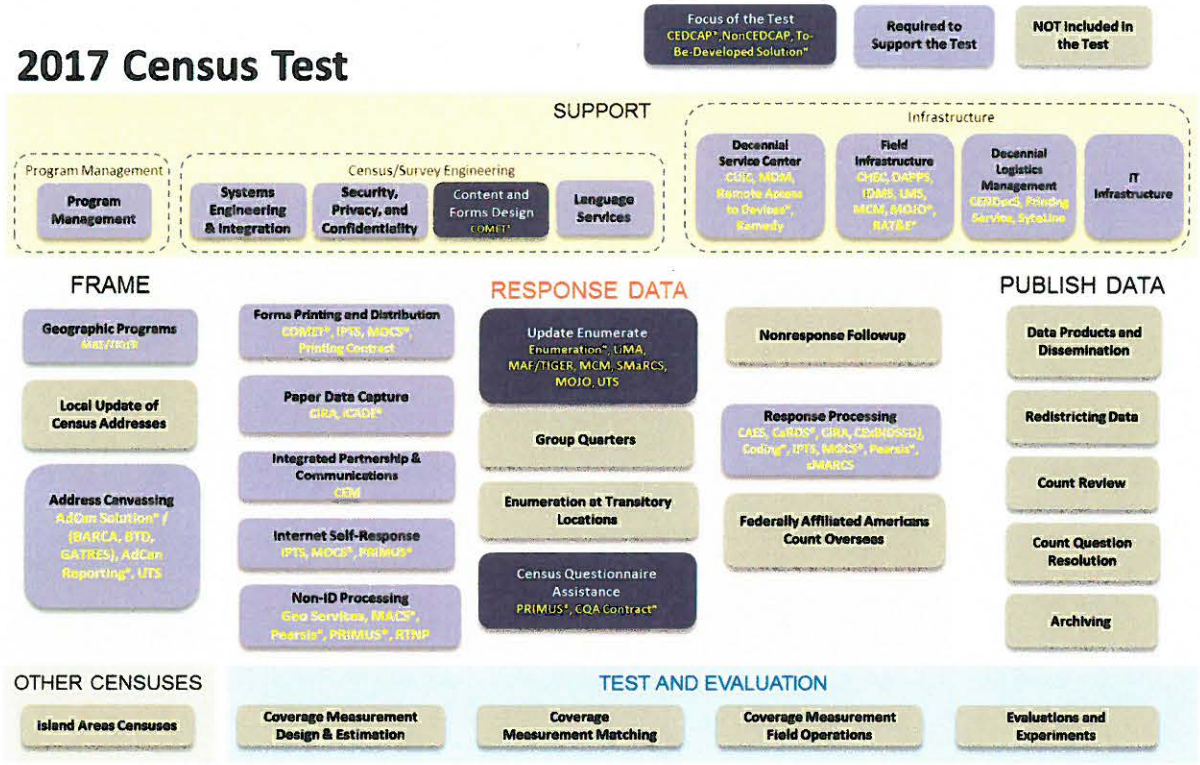


Figure 6-8 2017 Census Test

6.4.1 Main Features

The main features of the 2017 Census Test are very similar to the objectives of the Puerto Rico Test, and include the following:

- Transition some capabilities to the cloud
- Introduce Update Enumerate operation
- Test Update Enumerate
- Accelerate development time schedules, create testing window of sufficient length to do adequate testing

Table 6 shows the main features of the test.

Table 6 Main features of the 2017 Census Test

System	Description	Program
CAES	Introduced to perform adaptive design modeling to optimize field contact strategy. Stores data and uses it to execute statistical models in support of survey flow processing, analysis, and control.	CEDCaP
CQA	Use telephone center for re-interview workload. New CQA vendor introduced, contract awarded in June 2016, replaces TQA • Their systems need to interface with PRIMUS • Provides online assistance	N/A
PRIMUS	Needs to be enhanced to allow CQA staff to login and collect data for respondents.	CEDCaP
PEARSIS	Replaces CARRA for managing administrative records.	N/A
LiMA	LiMA enumeration interface updated, introducing the interface between LiMA and the enumeration instrument (currently COMPASS). LiMA translation to Spanish introduced.	CEDCaP
COMPASS	Needs to be enhanced to allow mobile survey instrument and listing application (LiMA) to run on same mobile device. Note: In the future, this will be the CEDCaP Enumeration System	N/A
SMaRCS	Introduced to provide quality control listing	N/A
CHEC	Interface to 3rd party vendor model tested	N/A
CARAT	New capability applicants to apply on-line as well as taking the assessment test.	N/A

The CEDCaP product releases for the 2017 Census Test are detailed in the CEDCaP Transition Plan, Section 4.3.7.

6.5 Puerto Rico Test

The Puerto Rico Test combines the goals of the 2016 Census Test, the Address Canvassing Test, and the 2017 Census Test. It is designed to test census operations and systems on a large scale, and include features to handle non-English language processing as well. The Puerto Rico Test will test the utility of the updated Puerto Rico MAF, including frame updating, self-response, and field enumeration. Address Canvassing will take place, but will be unlinked from the enumeration part of the test similar to the Address Canvassing Test and the 2017 Census Test. The Puerto Rico Address Canvassing start date is now planned for Feb 2017. A diagram showing the operations and associated systems is provided below in Figure 6-9.

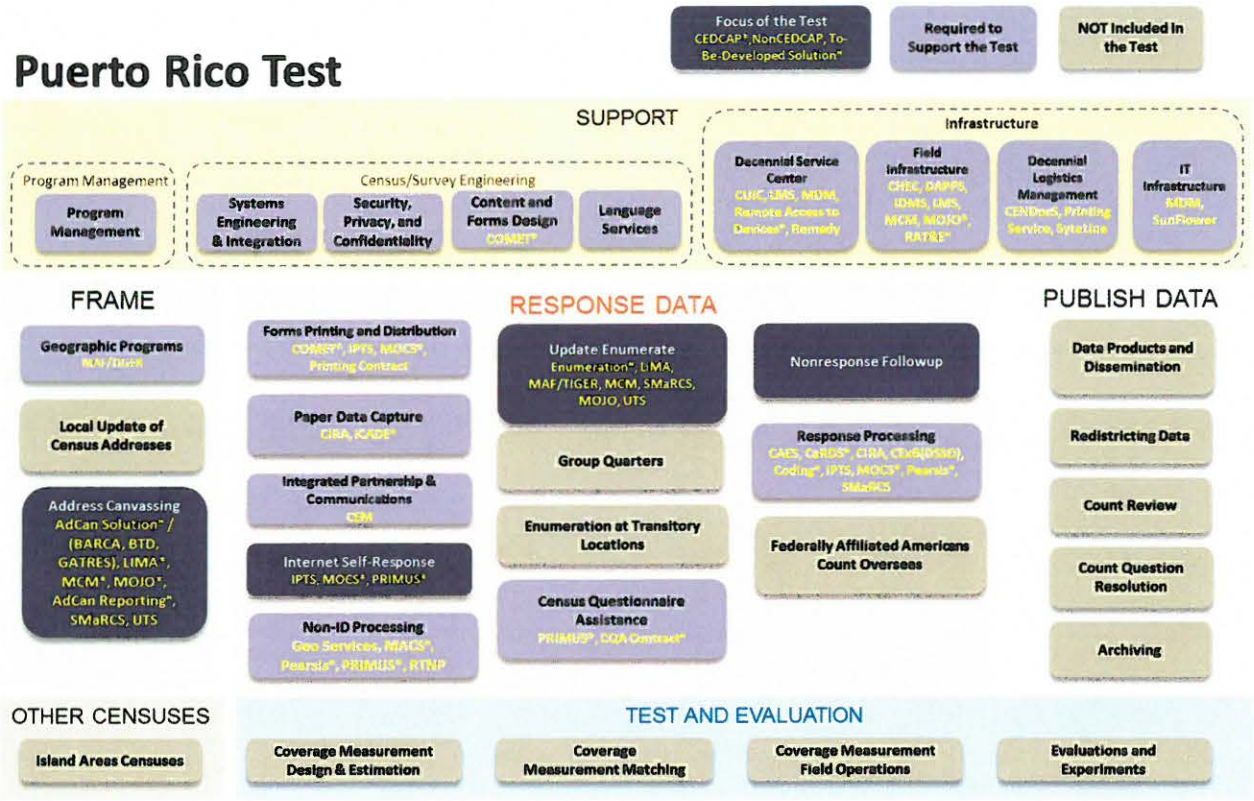


Figure 6-9 Puerto Rico Test

6.5.1 Main Features

The main features of the Puerto Rico Test include the following:

- Transition some capabilities to the cloud
- Introduce Update Enumerate operation
- Combine aspects of 2016 Census Test, AdCan Test, and 2017 Census Test all together
- Flexibility to handle PR differences
- Test different mail-out strategies

Table 7 shows the main features of the test.

Table 7 Main features of the Puerto Rico Test

System	Description	Program
CAES	Introduced	CEDCaP
CQA	New CQA vendor introduced, contract awarded in June 2016, replaces TQA • Their systems need to interface with PRIMUS • Provides online assistance	N/A

System	Description	Program
PEARSIS	Replaces CARRA	N/A
LiMA	LiMA enumeration interface updated, introducing the interface between LiMA and the enumeration instrument (currently COMPASS). LiMA translation to Spanish introduced.	CEDCaP
COMPASS	Needs to be enhanced to allow mobile survey instrument and listing application (LiMA) to run on same mobile device. Note: In the future, this will be the CEDCaP Enumeration System	N/A
SMaRCS	Updated from the 2016 Census Test; used to provide re-interview functionality only	N/A
CHEC	Interface to 3rd party vendor model tested	N/A
Partnership Communication System	New system that includes: Partnership Emergency System	N/A

There are no specific plans for CEDCaP product releases for the Puerto Rico Test documented in the CEDCaP Transition Plan.

6.6 2018 End-to-End Test

The 2018 End-to-End test is designed to be a large scale test for the 2020 Census. The intent is to fully exercise all major operations and systems in preparation for Decennial workloads and timelines. The goal is to have the operational design for the major operations ready for production – from a systems, operational, and architectural perspective. The 2018 End-to-End test will include significant field data collection components, and the timing of the field operations will mimic the 2020 Census. A diagram of the operations and supporting systems is provided below in Figure 6-10.

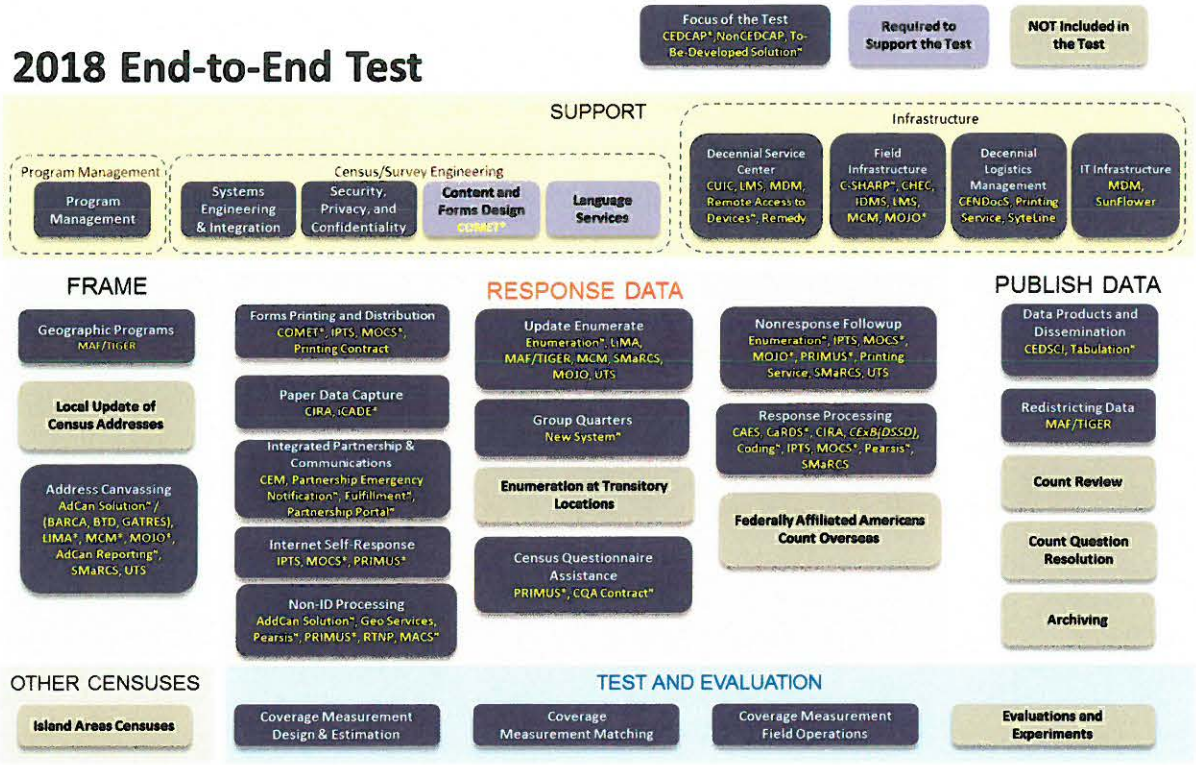


Figure 6-10 2018 End-to-End Test

6.6.1 Main Features

The main features of the 2018 End-to-End test are as follows.

- Listing QC is introduced, with the full LiMA-SMaRCS interface
- The 2018 End-to-End test should mirror the expected system configuration for the 2020 Census. The 2018 End-to-End test will use the systems and the hardware, cloud and ESB services that are planned to be used in 2020.
    - The actual workload of the 2018 End-to-End test will not approach the levels of the 2020 Census. One alternative being considered is to use large sets of synthetic data for load and stress testing to make sure the systems can handle the increased workload.
- Coverage Measurement (CM) introduced
    - Goal is to automate some systems that were heavily manual in 2010
    - CM will attempt to take advantage of the centralized person-matching systems as they are created for 2020 Census efforts
- Tabulation and Dissemination is also being introduced

- Second Data Capture Site: for Census tests up to and including the 2018 End-to-End test, there are no plans for a second data capture site. However, there will be a second data capture site for 2020 Census.

Table 8 shows the main newly-introduced features of the test.

Table 8 Main features of the 2018 End-to End Test

System	Description	Program
Group Quarter Enumeration Portal (TBD)	Group Quarters operation introduced (military bases, college campus, retirement homes, etc.) • Heavily administrative records- based • 2018 Test will need some capability to provide external interfaces where people can upload administrative records	N/A
COMPASS	Needs to be enhanced to allow enumeration of Group Quarters	N/A
Partnership Communication System	Introduces new capabilities to include: Partnership Fulfillment Partnership Portal	N/A
Tabulation (TBD)	System to perform data tabulation	N/A
CEDSCI	System for data dissemination	CEDSCI
LiMA	Introduced the listing QC functionality	CEDCaP
C-SHaRPS	If C-SHaRPS replaces DAPPS, then we need to consider the following impact: • This system is critical to the success of tests and operations to be conducted for the 2018 End-to-End Test and 2020, especially because this is the first time the system will be tested with new interfaces impacting key Census Systems such as MOJO, Geography for Geocoding, CBS, UTS, and others.	N/A

The CEDCaP product releases for the 2018 End-to-End Test Overview are detailed in the CEDCaP Transition Plan, Section 4.3.10. For the 2018 End-to-End test, any CEDCaP capabilities intended for use in the 2020 Decennial need to be fully deployed.

6.7 2019 Testing

According to the 2020 Census Operational Plan, two types of tests are planned for 2019: Defect Resolution Testing and Post End-to-End Performance Testing. These two Tests will be completed by August 2019, so that the complete Solution will be ready for the 2020 Census.

1. The Defect Resolution Testing will ensure that any changes made to fix defects in the applications tested in the 2018 End-to-End Test are correctly resolved for the production of the 2020 Census. Note that not only application defects have to be corrected and tested, but issues related to infrastructure will also be resolved.
    a. Defect testing will first be performed in a development and test environment, then deployed in the to-be Production environment built out as shown in Figure 6-3.
    b. Test cycles and phases will follow the eSDLC (enterprise System Development Lifecycle) methodology.
2. The Post End-to-End Performance Testing will ensure that the Solution as a whole satisfies the system quality attributes, especially high scalability, availability, and reliability in the to-be Production environment built out as shown in Figure 6-3. This will help minimize the risk of system crashes and delays in processing respondent Internet submissions and phone calls during the production period. Note that throughout the eSDLC, and as part of the Scale-up Readiness effort (Figure 6-3), performance testing of the systems and the Solution is performed. Thus, the Post End-to-End Performance Testing can be considered as a regression performance testing, and used to validate the production environment in order to ensure its readiness. Load simulation and synthetic data may be used to inject high load into the Solution.
    a. In terms of scalability, the Post End-to-End Performance Testing will ensure that the Solution can handle the traffic surge, and has sufficient redundancy to satisfy availability and survey reliability.
    b. For Internet Self Response, testing response time will be emphasized to ensure excellent user experience for Census respondents.
    c. The Solution redundancy consists of two levels: a) within one site, with clustered systems and components: b) across two sites, with replicated systems for major functionalities. The second level will most likely be exercised in case of a disaster recovery. Failover is tested to cover scenarios of failure of a certain system/component within one site, and failure of a whole site.
    d. The test plan will encompass all systems and their interactions that comprise the overall Solution. In other words, the Post End-to-End Performance Testing will consist of an Integrated Solution Performance Test.

6.8 2020 Census

The 2020 Census is the actual national census conducted to meet the Constitutional requirement to enumerate the population every 10 years. A diagram showing the operations and supporting systems is provided below in Figure 6-11.

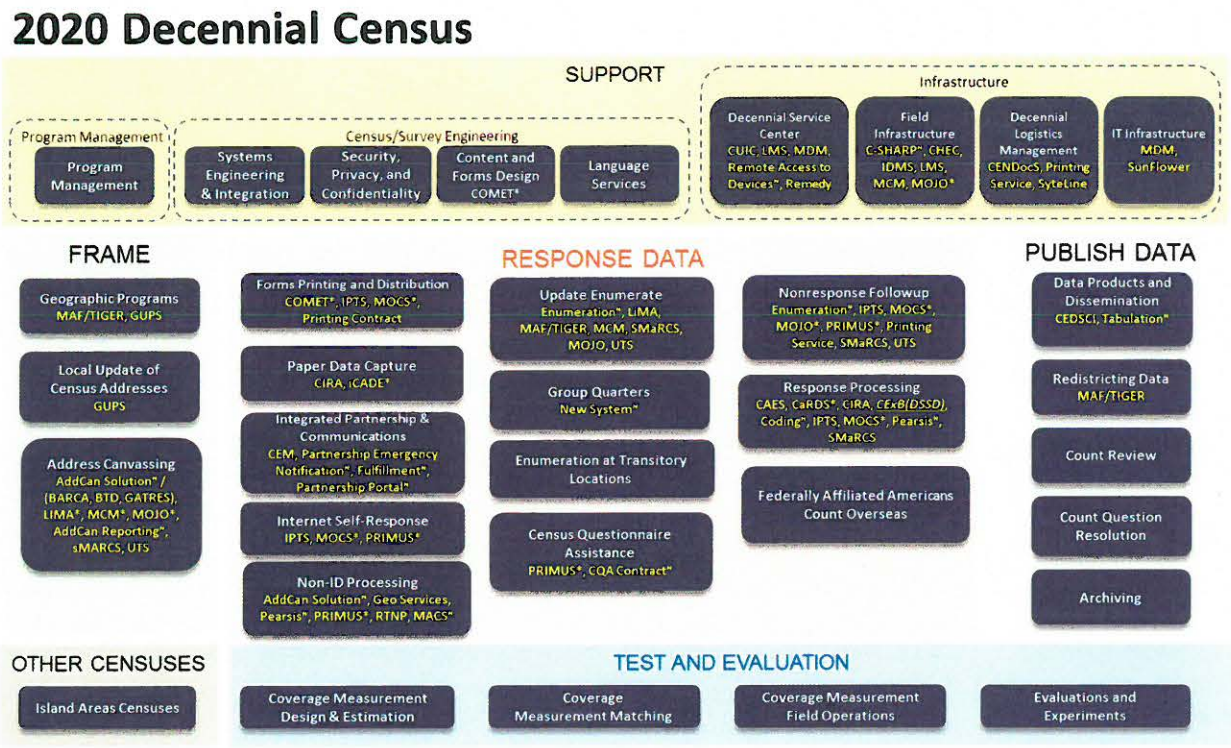


Figure 6-11 2020 Census

6.8.1 Main Features

The main features of the 2020 Census are as follows.

- Island Area Census Infrastructure added. Sharing stateside systems as appropriate, but also developing unique capability.
- Other operations Count Review, CQR, will have some operation-specific systems that need to be identified.
- The scope of the 2020 Experiments and Evaluations program has not yet been determined, but the experiments and/or evaluations could impact various systems (e.g., different content, different mailing strategies, etc). Please refer to Appendix J: Outstanding Decisions with Potential Impact on the Census Architecture for outstanding decisions that may impact the 2020 Architecture.

At this time, details for the CEDCaP Product Releases for 2020 Census are not fully developed. The information that is available can be found in the CEDCaP Transition Plan, Section 4.3.

7 Geographic Activities, Ongoing Operations, and Small-Scale Tests

This section provides details about Geographic programs, ongoing activities, and small-scale tests that provide inputs for the overall 2020 Decennial architecture transition. These inputs can be in the form of new systems, system upgrades to existing capabilities, or increased understanding of operational procedures derived from field testing.

This Geography Division (GEO) in Decennial has an ongoing activity to enhance and maintain the Master Address File (MAF), which forms the sampling frame for the decennial Census and many surveys as well. GEO is also the steward of the Topologically Integrated Geographic Encoding and Referencing (TIGER) system, which is responsible for managing the technical aspects and standards defining geospatial data used by Decennial Census and many surveys.

Census conducts operations on a continuous basis to fulfill certain missions. The results of these ongoing activities may have an effect on transition planning.

Several small-scale studies are planned for 2016 thru 2020. These tests are designed to investigate specific operational capabilities and/or geographies. The results of these tests are used to adjust large-scale Census tests and 2020 Decennial planning.

7.1 GEO Activities

7.1.1 In-Office Address Canvassing

Continual research and updating will be conducted through an In-Office Address Canvassing operation that will begin in September 2015 and continue through the 2020 Census. Clerks will start with the 2015 Census address list and update it based on new information from the United States Postal Service (USPS), and data from tribal, state, and local governments and third parties (i.e., commercial vendors). Clerks will review satellite imagery to determine where changes in addresses are occurring, and based on these changes, the Census Bureau will develop a plan for capturing those changes. This plan will include an In-Field Address Canvassing operation where address updates cannot be obtained or verified or in areas undergoing rapid change. The number of addresses requiring In-Field Canvassing is expected to be approximately 25 percent of the total number of addresses. These design changes have the potential to save the Census Bureau an estimated $900 million.

7.1.2 Local Update of Census Addresses (LUCA)

The LUCA provides the opportunity for tribal, federal, state, and local governments to review and comment on the Census Bureau's address list and maps to ensure an accurate and complete enumeration of their communities. The Census Address List Improvement Act of 1994 (P.L. 103-430) authorized the Census Bureau to provide individual addresses to designated local officials of tribal, federal, state, and local governments who agreed to conditions of confidentiality in order to review and comment on the Census Bureau's address list and maps prior to the decennial census. The basic process for LUCA includes:

- Census Bureau provides address list and maps to the governmental entities.
- Governmental entities review and add, delete, or change address records or features.
- Census Bureau incorporates the updates to MAF/ TIGER System.

- Census Bureau validates the updates via Address Canvassing.
- Census Bureau provides feedback to the governmental entities.
- Governmental entities can appeal the Address Canvassing validation outcomes.

7.1.3 Participant Statistical Areas Program / Tribal Statistical Areas Program (PSAP)

The Participant Statistical Areas Program / Tribal Statistical Areas Program (PSAP/TSAP) is part of the Geographic Partnership Program. These are programs that allow designated participants, following Census Bureau guidelines, to review and suggest modifications to the boundaries of block groups, census tracts, Census County Divisions, and Census Designated Places. Participants can also propose new Census Designated Places based on specific criteria. The 2020 Census PSAP includes all tribal statistical boundaries, which were administered through the TSAP in the 2010 Census, combining the two programs. The TSAP geographies are Oklahoma Tribal Statistical Areas, Tribal Designated Statistical Areas, State Designated Tribal Statistical Areas, tribal census tracts, tribal block groups, statistical tribal subdivisions, Alaska Native Village Statistical Areas, and for administrative purposes, one legal area, state reservations.

7.1.4 Redistricting Data Program (RDP)

The purpose of the RDP operation is to provide to each state the legally required P.L. 94-171 redistricting data tabulations by the mandated deadline of 1 year from Census Day: April 1, 2021.

The RDP Operation provides the 50 states, the District of Columbia, and Puerto Rico with the opportunity to identify, delineate, and update geographic boundaries for data tabulation. It also allows for continuous process improvement through an evaluation of the program with recommendations for the next cycle that is in an official publication called "The View From the States."

The five major components in the 2020 Census RDP include:

- Phase 1—Block Boundary Suggestion Project.
- Phase 2—Voting District Project.
- Phase 3—P.L. 94-171 data and geographic support products design and delivery.
- Phase 4—Collection of changes to Congressional and State Legislative Districts.
- Phase 5—Evaluation of the 2020 Census RDP and recommendations for the 2030 RDP.

7.1.5 Voting Data Program (VDP)

TBD

7.1.6 GEO Partnership and Outreach Programs

GEO plans to conduct several programs in the 2016 to 2020 time frame. The major efforts planned for GEO include the following.

- Geographic Partnership Programs
    - Prior to the 2020 Census, the Census Bureau conducts geographic partnership programs to make the address list as up-to-date as possible and ensure complete coverage of all

housing units. The Partnership Programs also help define statistical geographic area boundaries that will provide meaningful data from the 2020 Census.

- Geographic Programs Delineation
  - The Geographic Delineation component of the Geographic Programs determines, delineates, and updates the geographic area boundaries for 2020 Census data collection and data tabulation. Census data collection relies on the delineation of various geographic areas, known as "collection geography," to support the capture of data during Census activities. This includes both the delineation of the areas where differing methods are used to enumerate households and the delineation of field management areas.

- Geographic Data Processing Programs
  - The Geographic Data Processing component of Geographic Programs includes all activities that relate to the extract, update, and maintenance of the features, boundaries and addresses in the MAF/ TIGER System. Geographic data captured as part of the 2020 Census, including address updates, structure coordinate locations, boundaries, and roads data will be processed to ensure that the MAF/TIGER System is up to date.

- Geographic Support Programs
  - The Geographic Support Programs component includes support for GUPS, Help Desk, and Mapping.
  - GUPS will support all geographic partnership programs (i.e. Boundary and Annexation Survey (BAS), PSAP/TSAP, Boundary Validation Program, and Public Use Microdata Areas, Redistricting Data Program (RDP), Local Update of Census Addresses (LUCA), and Count Question Resolution.

7.2 Ongoing Operations

Ongoing operations are activities that are continuous. These programs may provide inputs for transition planning as an outcome of their ongoing work.

7.2.1 In-Office Address Canvassing

As discussed in section 7.1.1, In-Office Address Canvassing is part of the 2020 Census Reengineering Address Canvassing activity designed to eliminate the need to canvas every block. Continual research and updating will be conducted through an In-Office Address Canvassing operation that will begin in September 2015 and continue through the 2020 Census.

7.2.2 Boundary and Annexation Survey (BAS)

The Boundary and Annexation Survey is part of the Geographic Partnership Program. It provides an ongoing survey for collecting and maintaining information about the inventory of the legal boundaries for, and the legal actions affecting the boundaries of counties and equivalent governments, incorporated places, Minor Civil Divisions, Consolidated Cities, Urban Growth Areas, Census Areas of

Alaska, Hawaiian Homelands, and federally recognized legal American Indian and Alaska Native areas (including the Alaska Native Regional Corporations). This information provides an accurate identification and depiction of geographic areas for the Census Bureau to use in conducting the decennial and economic censuses and ongoing surveys such as the ACS.

7.2.3 School District Review Program

TBD

7.3 Small-Scale Tests

Preparations for 2020 Decennial include a series of large scale tests previously discussed in Section 6. In addition to these large scale tests, Census also plans to conduct a series of small-scale highly focused tests in order to increase understanding of operational field procedures.

7.3.1 Continuous Small Scale Testing

The Continuous Small-Scale Testing is a study of respondent and nonrespondent reactions to new modes of decennial census contact and response. The study focuses on reactions related to privacy and confidentiality of these modes. This study started in January 2014 and is ongoing as needed. It included e-mails to 1,000–2,200 housing units sampled from an opt-in frame.

7.3.2 Service-Based Enumeration (aka the GQ Test)

Service-Based Enumeration is part of the Census Group Quarters (GQ) Operation, designed to perform enumeration of people experiencing homelessness or utilizing transitional shelters, soup kitchens, regularly scheduled mobile food vans, and targeted nonsheltered outdoor locations.