U.S. Environmental Protection Agency

DATA CENTER CONSOLIDATION PLAN

September 2016

TABLE OF CONTENTS

Contents

1 Introduction

EPA has initiated a series of successful initiatives embracing data center consolidation, industry best management practices, and virtualization across its data centers. EPA has consolidated small data centers and computer rooms in various locations across the country.

The Agency completed a phased virtualization program across its primary data center which included optimizing the efficient use of floor space, turning off air handlers and raising the data center operating temperature towards the upper end of manufactures' specifications. EPA currently hosts individual Agency business applications in an innovative shared hosting environment and IT systems from several external federal organizations in a co-location hosting arrangement offering many of the features of commercial and cloud hosting services.

EPA had 83 data centers and computer rooms located in 66 buildings across 48 cities in 31 states and territories. These data centers and computer rooms support business operations for 25 major EPA facilities, including EPA Headquarters (DC Metro Area), EPA Research Triangle Park (RTP) research center, 10 EPA regional offices, and 13 major research centers. The remaining facilities are smaller field offices and continuity of operations (COOP) sites. Four of the 83 data centers are classified as tiered data centers, and EPA designated these four data centers for its major information technology (IT) operations. EPA's primary data center is the EPA National Computer Center (NCC) located in RTP, North Carolina.

For fiscal years 2016 through 2018, EPA has implemented a revised and enhanced approach for meeting or exceeding OMB's established data center cost savings and avoidance targets. For geographic areas where EPA has multiple data centers, our revised approach establishes a single facility for data center IT assets to be consolidated into. For each data center selected for retention, necessary upgrades are being made to address any facilities, network capacity or operational issues. Data Center Optimization Initiative (DCOI) POCs have identifying "in scope" data center IT assets; are in the process of reducing their server "foot print" through decommission, consolidation and/or virtualization activities; and will then execute consolidation plans to move applicable data center IT assets into the cloud, an EPA tiered data center, or the geographic data center being retained for that given area. The final phase of this initiative requires DCOI POCs to surplus or excess data center IT assets will be decommission and appropriately scaled residual data center assets to be aligned with reduced capacity requirements. As of Q4 FY16, EPA has closed 25 non-tiered data centers and has targeted another 19 for closure through fiscal year 2018 closing 55% of Agency non-tiered data centers.

In some cases, EPA's regional offices, research centers, labs and field offices host local infrastructure services such as file and print services, specialized lab and research support, local program implementation, emergency response (ER) and COOP. These data centers provide localized computing services, or specific scientific computing services (e.g., laboratories), making further consolidation challenging. The distributed nature of these offices and the costs associated with high-speed network connectivity require EPA to balance consolidation efforts with network cost and application performance requirements.

2 Agency Goals for Data Center Consolidation

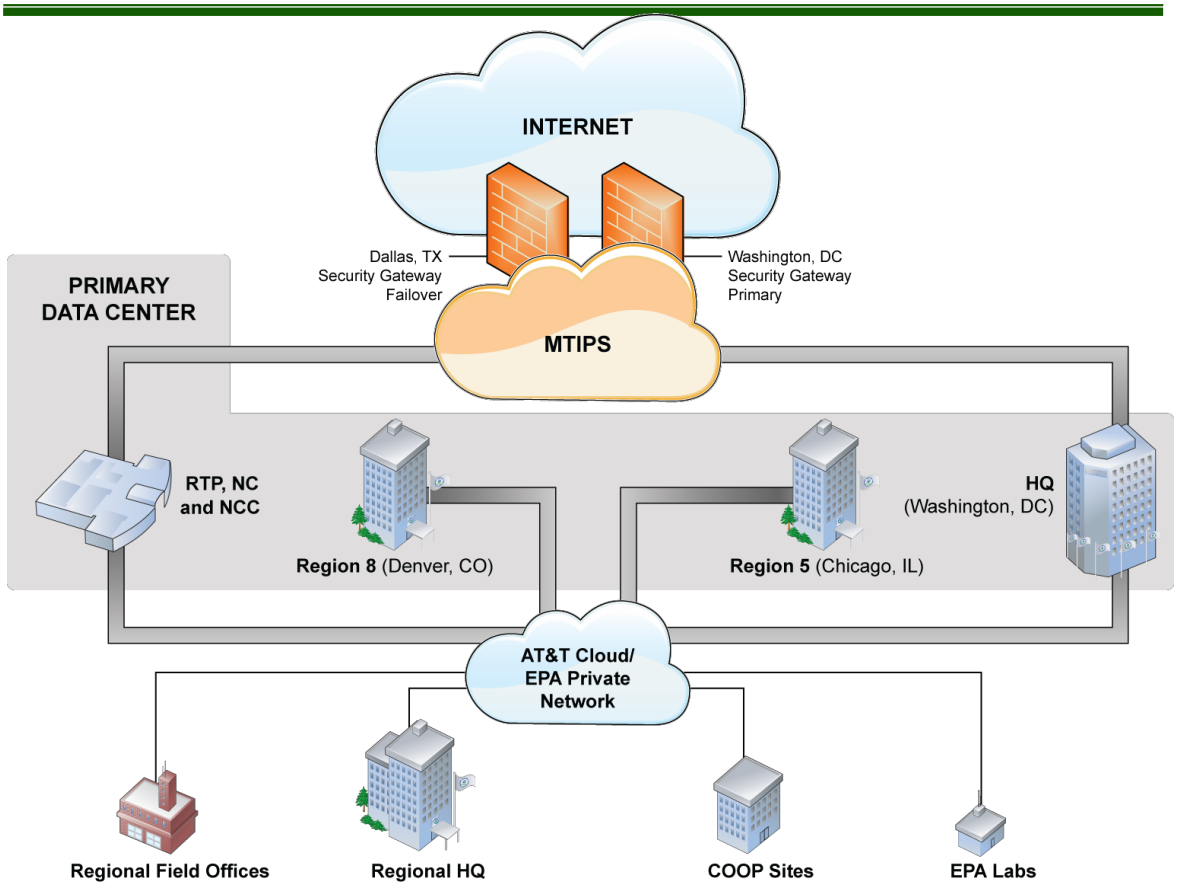
EPA's revised consolidation initiative seeks to achieve the following goals:

- Establish Agencywide contract(s) to facilitate application hosting in the public and/or private cloud environments.
- Establish private cloud services for enterprise application hosting within the primary data centers and to facilitate hybrid cloud deployments.
- Consolidate data center IT assets from 29 data centers across nine distinct geographic areas into nine data centers (one per each geographic area).
- Continue to maximize server virtualization and consolidation across EPA data centers to reduce the number of physical servers.
- Align IT refresh cycle with consolidation activities.
- Establish service level agreements, operating level agreements, and performance standards to deliver computer room and data center services reliably.
- Refine standards for backup, disaster recovery, COOP and ER support.
- Institute "virtual COOP" to the most practical extent possible.
- Incorporate green IT approaches across Agency data centers and computer rooms to maximize data center and computer room energy efficiency.

3 Agency Approach, Rationale and Timeline

The basis of EPA's data center consolidation plan focuses on near-term efforts EPA intends to accomplish through fiscal year 2018 and include: network upgrades or augmentation; server consolidation and virtualization; computer room consolidation at local campuses; establishing a single data centers within geographic areas with multiple data centers; and utilization of Agency primary data centers to establish private cloud services.

The following diagram depicts EPA's primary data centers, computer rooms and network topology.



EPA has 83 computer rooms in 66 buildings located in 48 cities within 31 states.

EPA is in the process of replacing the Region 8 data center with the National Enforcement Investigations Center (NEIC) data center which will serve as EPA's western presence data center. Both facilities are located in Denver, Colorado. However, the NEIC facility is a federally owned building and is optimal for our consolidation initiative. All primary data centers will be operated in a standardized fashion and will implement configurations that maximize power and cooling efficiency. Since 2008, EPA has collaborated with the Department of Energy and EPA's ENERGY STAR program, and will continue to adopt applicable energy saving techniques as they are developed and certified.

EPA's mid-term efforts will consolidate 29 data centers into nine and establish centralized resources for COOP, disaster recovery and enterprise backup in the primary data centers. With the full implementation of these initiatives, EPA expects to realize several efficiencies:

- Minimize data center energy consumption
- Minimize escalation in IT operations costs
- Maximize server and storage utilization
- Maximize standardization and agility
- Reduce long-term growth of IT infrastructure costs

A summary of the non-tiered data centers targeted for closure is provided in table 3.0 below:

Table 3.0: Summary of EPA's Data Center Consolidation Plan by Geographic Area

Geographic Area Data Center	Data Center Consolidation Target	Summary Status
Ann Arbor - OAR Ann Arbor Lab Campus Server/Telcom Rm. (306B)	3 → 1	Plan received. CIO Portfolio review to be scheduled.
Athens - ORD Athens Lab Campus Server/Telcom Rm. (B13B)	3 → 1	Plan received. Work with OEI to review plan and expedite schedule. CIO Portfolio review to be scheduled.
Cincinnati, OH & Erlanger, KY - OARM-CINCI AWBERC Campus Server/Telcom Rm.	7 → 1	Plan received. Work with OEI to review plan and expedite schedule. CIO Portfolio review to be scheduled.
Denver, CO & Helena, MT - R08-LKCO-NEIC 1B-235-Campus Server/Telcom Rm.	4 → 1	Plan received. Work with OEI to review plan and expedite schedule. CIO Portfolio review to be scheduled.
Edison - R02 Edison Campus Server/Telcom Rm.	2 → 1	Consolidation completed and FDCCI-DC-45594 closed.
Houston & Dallas - R06 HQ Campus Server/Telcom Rm. (S1200)	3 → 1	Region 6 will employ configuration which utilizes an EPA core data center for their disaster recovery solution. Planning phase underway.

Kansas City - R07 HQ Campus Server/Telcom Rm. (2.1-N08)	3 → 1	Plan under development. OEI will review upon submission. CIO Portfolio review may be scheduled.
Las Vegas - OAR - NCRFO-LV NCRFO La Plaza C Server Rm. (LPC621)	2 → 1	Initial plan received. Worked with all Las Vegas stakeholders to advise on course of action to be undertaken. OCFO will submit plan for consolidating in OAR data center; ORD will revise plan to consolidate in NEIC tiered data center; OAR will become hosting data center for EPA's Las Vegas location.
RTP & Chapel Hill - RTP EPA Primary Data Center & Agency NSOC (NCC-N238)	2 → 1	Plan received. Preliminary/preparatory activities to minimize security risk and test secure transport solution prior to production are underway. Current timeline for completion is NLT December 2016.

EPA is pursuing the following activities to achieve OMB's data center optimization metrics.

A. Network Optimization: EPA utilizes General Services Administration's (GSA) Networx contract to provision EPA's wide area network (WAN), Internet and for management of the Trusted Internet Connection (TIC) services. EPA closely monitors Networx contract offerings to determine the optimal circuit configuration and pricing scheme for provisioning EPA WAN circuits. This ensures EPA sites receive required network performance for the least cost. EPA will transition to the upcoming GSA Networx contract.

B. Standardization and Enterprise Procurement: EPA established server and software standards for x86/64 virtualization platforms and established an enterprise platform for infrastructure monitoring. We'll continue to rollout these standards in the upcoming years and verify adherence to drive toward achieving the targeted optimization metrics. EPA has implemented multiple strategic sourcing initiatives to pool monetary resources so as to negotiate optimal pricing for IT products and services. EPA also appropriately scrutinizes all IT procurements as part of the FITARA process thereby ensuring organizations are aware of and adhere to Federal and Agency initiatives and remain aligned with our overall strategic direction.

C. Storage Optimization: EPA will continue to evaluate and leverage its investment in Microsoft's collaboration suite, Office 365. By the end of CY 2016, EPA intends to finalize planning and begin more aggressive deployment of OneDrive. This effort paired with implementation of a cloud storage solution for archived files will reduce desktop support costs and significantly curtail investments required for expansion of storage area network costs.

D. Energy Optimization: EPA has implemented several initiatives that have contributed towards achieving energy optimization for one of its primary data centers, the NCC. Use of new floor tile designed to optimize air flow, turning off several computer room air conditioning units and raising the data center operating

temperature has reduced energy consumption and achieve a PUE of XXX. EPA will incorporate these best practices in the other primary data centers where applicable. We'll also take advantage of cold aisle computer room configuration to further enhance our PUE numbers.

E. Enterprise COOP and Disaster Recovery (DR): The goal of enterprise COOP and DR is to provide for the COOP and DR services using shared services that are hosted in the primary data centers. COOP and DR services are currently provided using site specific solutions. Enterprise COOP and DR services must provide remotely accessible data and applications to support continued operations and emergency response to EPA Regions or field office. In 2010, EPA began testing virtualization configurations to host COOP and DR services for field sites at the applicable primary data center. EPA has now successfully implemented virtualized COOP for Region 9 and is in the planning phase for a Region 6 implementation. The provisioning of virtual COOP and disaster recovery at the primary data centers is a key component of EPA's data center optimization and server reduction strategy.

The following table outlines planned and achieved performance levels for each optimization metric, by year.

Table 3.1: Planned and Achieved Performance Levels for Optimization Metrics

No.	Optimization Metric	FYE 2018 Target Value	Current Value (estimated)	Planned Value	
				FY 2017	FY 2018
1	Energy Metering	100%	100%	100%	100%
2	Power Usage Effectiveness (PUE)	≤ 1.5	1.6	1.6	≤ 1.5
3	Virtualization	≥ 4	TBD	≥ 4	≥ 4
4	Server Utilization & Automated Monitoring	≥ 65%	0%	≥ 25%	≥ 65%
5	Facility Utilization	≥ 80%	33%	≥ 45%	≥ 80%

Estimated figures are based on EPA's Q3 2016 input. There were some discrepancies discovered regarding DCOI data input (tiered data center count, one non-tiered data center was omitted during initial data entry and the total physical server count appears unaccounted for in current statistics) that OMB's DCOI POC will need to be consulted on. Also note that none of EPA's tiered data centers have automated monitoring software. As such, estimated figures for the *Server Utilization & Automated Monitoring* metric is based on proof of concept implementation within one primary data center in FY 2017 prior to rolling out to the remaining three.

Table 3.2 captures planned and achieved closure by year and has been updated to reflect the previously unaccounted Las Vegas data center which EPA intends to close in FY 2017.

Table 3.2: Completed and Planned Data Center Closures

Data Center Closures Completed to Date							Data Center Closures Planned			
FY2010	FY2011	FY2012	FY2013	FY2014	FY2015	FY2016	FY2017	FY2018	FY2019 and Beyond	Total
0	1	14	3	1	5	1	19			**44**
25 Total							**19 Total**			

4 Optimization Metrics and Data Center Closures

EPA's existing IT governance framework is structured to ensure Agency initiatives have the appropriate management oversight, commitment to performance and project management. These types of initiatives are supported by a cross-Agency team of IT operations managers from EPA's national program offices, Regions and research laboratories.

EPA's previous strategic plan and scheduled data center closure dates were based on an assumed commitment of dedicated federal and extramural resources. These assumptions tend to be relatively accurate the majority of the time. The accuracy of estimating implementation schedules tends to diminish with increased project complexity and/or when the project manager has little historical basis of reference. Both of these factors negatively influenced EPA's ability to achieve targeted data center closure dates. Missed optimization metrics are attributed to the interdependencies of acquiring new technologies in concert with the hardware refresh cycle. The procurement delays have hampered our ability to make more notable progress in achieving optimization metric targets. In summary, specific challenges were as follows:

Competing Priorities - Periodic security incidents and follow-on initiatives have required commitment of resources to address identified risks across the federal sector. For example, EPA had to commit key individuals to focus on last year's Cyber Sprint initiative.

Constrained Resources - IT personnel are primarily focused on day-to-day operations and maintenance activities. As such, resources normally used to support data center consolidation activities are periodically pulled away to address more immediate operational activities.

Procurement delays - Necessary legal and procurement requirements have delayed the implementation of resources needed to migrate IT systems to centralized data centers. For example, new procurement requirements resulting from a protest decision has delayed the acquisition of high capacity networking devices. This further delayed the migration of systems into the centralized data centers.

5 DCOI Spending and Cost Savings

As noted above, EPA's data center consolidation work is structured to promote cost savings in areas of reduced energy consumption, maximized of server and storage use, and reduction in the long-term growth of IT infrastructure costs. Project costs needed to achieve these goals include plan and design activities, updates to data center facilities (uninterruptible power supply, power distribution unit, computer room cooling equipment), and infrastructure upgrades. Costs also include investments in optimizing the remaining data centers, such as cold aisle containment, modified floor tiles, and standard rack implementation.