EPA

U.S. Environmental Protection Agency

DATA CENTER CONSOLIDATION
DRAFT PLAN

April 2017

TABLE OF CONTENTS

Table of Contents

1 Executive Summary

In 2016, the Office of Management and Budget (OMB) launched the Data Center Optimization Initiative (DCOI) as a successor to the Federal Data Center Consolidation Initiative (FDCCI). On August 1st 2016, the Federal Chief Information Officer (CIO) issued memorandum M-16-19 which provided the framework for agencies to achieve data center consolidation and optimization requirements. Under memorandum M-16-19 agencies are required to develop and report on data center strategies to consolidate inefficient infrastructure, optimize existing facilities, improve security posture, achieve cost savings, and transition to more efficient infrastructures, such as cloud services and inter-agency shared services. In addition, OMB set a closure goal for the U.S. Environmental Protection Agency of 34 non-tiered data centers by Fiscal Year 2018.

The EPA expects to achieve OMB's DCOI targets through the identification, consolidation and closure of all non-tiered data centers not designated for retention within a geographic area or unless it is proven infeasible to close the data center due to technical limitations that compromise agency mission essential functions. The EPA is in the process of procuring a Data Center Infrastructure Management (DCIM) tool and expects to have a certified Data Center Energy Practitioner (DCEP) assigned to all remaining data centers by FY 2018. Due to previous EPA projects such as Physical-to-Virtual (P2V) and preexisting energy efficiency initiatives, the EPA expects to meet or exceed all OMB requirements for energy efficiency and server utilization pending resource impact resulting from realized FY18 budget cuts.

2 Introduction

Consolidation and optimization of federal data centers has been a long-standing priority for the federal government and the Environmental Protection Agency (EPA). In August 2016, the Office of Management and Budget (OMB) issued memorandum M-16-19 which supersedes the previous Federal Data Center Consolidation Initiative (FDCCI) and defines the Data Center Optimization Initiative (DCOI). DCOI requires agencies to:

- Meet new data center closure and optimization targets and reassess cost savings.
- Assign certified Data Center Entergy Practitioners (DCEP) for each data center.
- Implement automated energy metering and asset management capabilities at all tiered data centers.
- On a quarterly basis, report to OMB Agency progress with respect to data center closure/consolidation targets, optimization metrics, and cost savings and avoidance targets.

At the inception of DCOI, EPA had 58 non-tiered data centers and computer rooms that supported business operations for EPA facilities located across the country. Per OMB criteria, four of EPA's data centers are classified as tiered data centers. These include the following:

- National Computer Center (NCC) in Research Triangle Park, North Carolina.
- Potomac Yards Data Center in Washington, DC.
- Region 5 Data Center in Chicago, Illinois.
- Region 8 Data Center in Denver, Colorado.[1]

To improve efficiency and consistency of data center operations, EPA implemented an Agency-wide P2V policy requiring offices to convert existing physical servers to virtual servers wherever possible. EPA also defined server and software standards for virtualized platforms and established an enterprise platform for infrastructure monitoring. EPA implemented strategic sourcing initiatives to pool resources and negotiate optimal pricing for IT products and services.

EPA is pursuing opportunities to establish shared data center services and encourage consolidation of data center functions. EPA established centralized resources for continuity of operations planning (COOP) and disaster recovery (DR) in the tiered data centers. Region 9 is currently leveraging these services and the Region 6 implementation is being operationalized. The NCC was designated by the General Services Administration (GSA) as an approved Inter-Agency Shared Service Provider (ISSP) data center. GSA recently established a five-year agreement with EPA to leverage NCC facilities and hosting services. Reference Section 4 – Data Center Optimization Metrics for additional detail.

---

[1] EPA is currently in the process of replacing the Region 8 data center with the National Enforcement Investigations Center (NEIC).

To meet DCOI reporting requirements, EPA is in the process of performing market research on Data Center Infrastructure Management (DCIM) tools. None of EPA's tiered data centers currently have automated monitoring software. EPA expects to conduct proof-of-concept testing in up to two tiered data centers in FY2017 followed by deployment in remaining tiered data centers in FY 2018.

For non-tiered data centers, EPA is evaluating how it can expand its existing EM7 tool to support system monitoring requirements. EPA estimates that it is currently at a Power Usage Effectiveness (PUE) ratio of 1.6 and that the virtualization and other optimization activities will enable the Agency to reach a PUE of $\geq$ 1.5 by FY 2018. See Section 4 – Data Center Optimization Metrics for additional detail.

EPA's efforts to optimize and consolidate its data centers has been significant. However, it is important to acknowledge that these updates involve considerable resource commitments to consolidate and upgrade data center infrastructure and procure and implement DCIM tools. Investing in these improvements and tools will enable EPA to better monitor and manage its energy consumption and report on OMB's metrics, but may not produce cost savings in the near-term. Additionally, EPA must also balance consolidation efforts with network costs and application performance requirements. For example, some EPA regional offices, research centers, labs and other facilities host local infrastructure such as file and print services, telecommunications infrastructure, specialized lab and research systems, and COOP and DR. The distributed nature of this infrastructure and the need to support these critical requirements can make further consolidation challenging. As EPA moves forward to achieve DCOI requirements, it will need to continue to balance the benefits of consolidation with operational requirements and costs. The remainder of this document provides more detail on EPA's progress and plans for meeting DCOI requirements.
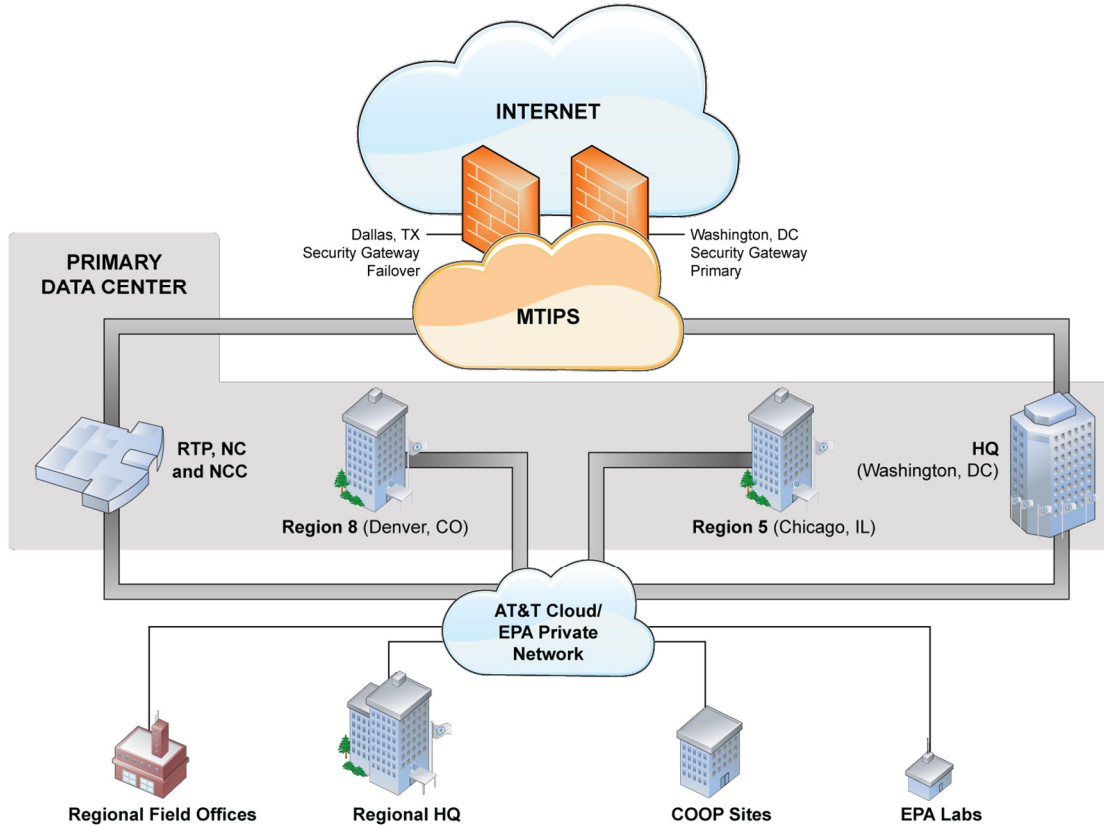
3 Data Center Consolidation and Closure Targets

Current Progress

The DCOI assigned a target goal for the EPA to close 34 non-tiered data centers by September 2018. At the inception of the DCOI, EPA maintained 58 telecommunication closets and computer rooms that met the classification of a data center per memorandum M-16-19. These data centers support business operations for EPA facilities including Headquarters (DC Metro area), Research Triangle Park (RTP), 10 regional offices and 13 HQ Program Offices. The remaining facilities are small field offices and continuity of operations (COOP) sites.

Per OMB's criteria, EPA classifies four of its data centers as tiered data centers. The tiered data centers support EPA's enterprise information technology (IT) operations. EPA's primary data center, the NCC, is located in RTP, North Carolina. Figure 1 depicts EPA's tiered data centers, computer rooms and network topology. EPA is currently in the process of replacing the Region 8 tiered data center with the National Enforcement Investigations Center (NEIC) which will serve as EPA's western presence data center. All tiered data centers will be operated in a standardized fashion and will implement configurations that maximize power and cooling efficiency.

Figure 1 - EPA's Primary Data Centers

Figure 1 - EPA's Primary Data Centers

In order to meet OMB's closure target, the EPA reviewed its existing data centers and revised its approach for consolidation utilized under FDCCI. For geographic areas where EPA has multiple data centers, EPA identified a single facility into which data center IT assets will be consolidated. Additionally, the EPA intends to consolidate all non-tiered data centers not designated to be retained to serve a geographic area or unless it is proven infeasible to close the data center due to technical limitations that compromise agency mission essential functions. Through these efforts, EPA has identified a pool of candidate data centers for closure to ensure we meet OMB's target of 34. Several validation activities are still required on this candidate pool.

As Table 1 shows, under FDCCI, EPA closed 21 of its data centers. For DCOI, EPA has closed 10 data centers, and 24 are in progress or planned.

Table 1 - Completed and Planned Data Center Closures by Fiscal Year

Completed and Planned Data Center Closures by Fiscal Year										
FY2010	FY2011	FY2012	FY2013	FY2014	Q1/2 2015	Q3/4 FY2015	FY2016	Q1/2 FY2017	Q3/4 FY2017	FY2018
0	1	14	3	1	2	3	3	4	10	14
21 Total FDCCI Closures						10 DCOI Closures Completed			28 DCOI Planned	
21 Total FDCCI						34 Total DCOI				

Activities Underway

EPA has revised its methodology to achieve OMB's closure goal of 34 non-tiered data centers. For those data centers targeted for closure, DCOI points of contact (POCs) are reviewing their data center IT assets and determining which to decommission, consolidate and/or virtualize. They will then execute consolidation plans to move applicable data center IT assets to the Cloud, an EPA tiered data center or the geographic data center retained for that area. The final phase of the effort will require DCOI POCs to surplus or excess data center IT assets and scale remaining data center assets to align with reduced capacity requirements.

Challenges and Risks

In some cases, EPA's regional offices, research centers, labs and field offices host local infrastructure data center assets such as specialized lab and research support, emergency response (ER) and COOP that may have non-severable configurations and potentially, must remain co-located to operate and fulfill mission functions. There are also cases where assets provide localized or specific scientific computing services (e.g., laboratories) and are affixed to the data center. The distributed nature of these offices and the continuation of required localized mission functions require EPA to balance consolidation efforts with cost, security and application performance requirements. There is also budgetary uncertainties that if realized, may significantly impede EPA progress for DCOI implementation.

4   Data Center Optimization Metrics

Current Progress

The DCOI establishes the performance measures and targets for energy consumption, virtualization and server and facility utilization that tiered data centers must meet. The EPA reports current PUE of 1.6, with a FY2018 target of 1.5 PUE. The EPA has made considerable progress optimizing its data centers and utilizing consolidation. These improvements will enable EPA to meet DCOI targets. For example, EPA implemented a "virtual first" policy requiring virtual servers be deployed to the maximum extent possible. At the NCC, the EPA initiated a phased data center modernization program which included a redesigned layout for more efficient use of floor space, turning off air handlers, raising the operating temperature towards the upper end of manufacturers' specifications, and preliminary implementation of cold aisle containment technologies. Similar improvements are planned for the other tiered data centers.

Several program and regional offices pursued similar virtualization initiatives, implemented more efficient LED lighting, and/or made improvements to server room cooling systems to reduce energy consumption. For example, Regions 9 and 10 implemented airside economizers and cold aisle containment systems reducing energy consumption over traditional HVAC systems. Region 9's improvements included raising the operating temperature to minimize energy consumption. As a result of these improvements, Region 10 saw a 47% reduction using the air economizer as compared to prior year HVAC energy consumption. Region 10 estimates that over a 10 year lease, the Region will save at least 915,782 kWh and approximately $89,747 (using 9.8 cents per kWh for Seattle, WA). EPA is implementing metrics to validate and track realized savings.

EPA also completed the following activities to improve efficiency and encourage consolidation:

- **Standardization and Enterprise Procurement:** EPA established server and software standards for the x86/64 virtualization platforms and established an enterprise platform for infrastructure monitoring. The Agency will continue to rollout these standards in upcoming years and verify adherence to drive toward achieving the targeted optimization metrics. EPA also implemented multiple strategic sourcing initiatives to pool monetary resources so as to negotiate optimal pricing for IT products and services. EPA reviews 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 EPA's overall strategic direction.
- **Enterprise COOP and Disaster Recovery (DR):** Many EPA offices currently provide COOP and DR services using site specific solutions. EPA established centralized resources for COOP and DR in the tiered data centers to enable the Agency to offer shared service options and encourage consolidation. Enterprise COOP and DR services provide remotely accessible data and applications to support continuity of operations and disaster recovery to EPA regions or field offices. EPA successfully implemented virtualized COOP for Region 9 and is implementing a similar configuration for Region 6.
- **ISSP Data Center Hosting Services** – As an ISSP, EPA entered into an agreement with GSA to provide data center hosting services at the NCC. This engagement began in September 2016 as an incrementally phased deployment that allows GSA to strategically optimize its data center footprint and operating costs over the next several years. The initial five year agreement will provide GSA with facility hosting space and associated environmental requirements (air conditioning and humidity control, electrical, physical security, etc.) as well as provisions for executing options to continue hosting at the NCC. EPA will develop documentation and procedures to enable the Agency to standardize and reuse its processes for future co-location hosting engagements with other federal organizations.

Moving Forward

For the remaining data centers, data center POCs are working with EPA management and data center staff to identify opportunities for greater efficiency, make necessary upgrades to address any facilities, network capacity or operational issues, decommission and excess unneeded IT assets, and appropriately scale remaining assets to align with reduced capacity requirements. EPA is also in the process of establishing and assigning a certified Data Center Energy Practitioner (DCEP) for each data center.

Per DCOI requirements, EPA is performing market research on Data Center Infrastructure Management (DCIM) tools to support automated collection and reporting of data center hardware/software inventories and optimization metrics in its tiered data centers. None of EPA's tiered data centers currently have automated monitoring software that meets current DCOI requirements. EPA plans to conduct proof-of-concept testing in FY 2017 followed by procurement and deployment to remaining tiered data centers by FY2018.

For non-tiered data centers, EPA is analyzing how it can expand its existing EM7 tool and associated reporting capabilities. EPA plans to leverage EM7 to enable the Agency to improve system monitoring capabilities in non-tiered data centers in order to identify systems that are not

meeting efficiency standards, address deficiencies, and meet power efficiency goals.

In addition to the data center updates already discussed, EPA is also pursuing the following activities on an on-going basis:

- **Network Optimization:** EPA utilizes GSA's Networx contract to provision EPA's wide area network (WAN), Internet and 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 WAN circuits.
- **Storage Optimization:** EPA will continue to evaluate and leverage its investment in Microsoft's collaboration suite, Office 365. By the end of calendar year (CY) 2017, 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 and should significantly curtail investments required for expansion of existing on premise storage solutions.

EPA estimates that it is currently at a Power Usage Effectiveness (PUE) ratio of 1.6. The Agency anticipates that virtualization and planned optimization activities will enable the Agency to reach a PUE of less than 1.5 by FY 2018. Table 4 outlines planned and achieved performance levels for each optimization metric by fiscal year. Please note: The estimated figure for the *Server Utilization and Automated Monitoring* metric is based on proof of concept implementation with one primary data center in FY 2017. Once the DCIM tool is implemented in all data centers, EPA performance metrics will be based on the results of all tiered data centers.

Table 2 - Planned and Achieved Performance Levels for Optimization Metrics

No.	Optimization Metric	FY 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	$\geq 65\%$	0%	$\geq 25\%$	$\geq 65\%$
5	Facility Utilization	$\geq 80\%$	33%	$\geq 45\%$	$\geq 80\%$

Challenges and Risks

EPA has faced procurement delays which have hindered the migration of IT systems to centralized data centers. For example, new procurement requirements resulting from a protest decision have delayed acquisition of high capacity networking devices. This slowed migration of systems into the centralized data centers.

The increased accuracy of optimization metrics and automated reporting of OMB's optimization metrics is dependent on the procurement and implementation of a DCIM tool. While the Agency is working to plan for and mitigate potential procurement and implementation risks, delays in either process could further impact EPA's ability to meet DCOI targets. As automation tools are installed and configured, the EPA anticipates that future PUE numbers may differ from those previously reported due to direct measurement at the source provided by these specialized tools. Additionally, the EPA is still working towards defining and documenting server metrics, idle thresholds and specific server properties for accurately reporting server utilization between different server types.

Investments are also required to support the planning, design and implementation of optimization activities. The cost to make these improvements is substantial and may make it difficult for the Agency to demonstrate near-term cost savings. Potential schedule adjustments from realized budgetary cuts is also a potential outcome that may delay the implementation of DCOI.

5 Cost Savings

Currently, EPA does not expect to realize significant cost savings under DCOI. Due to the low number of tiered data centers (4) utilized by the EPA, OMB did not set a target for tiered data center closures. As per OMB, most cost savings are expected to be realized from the closing of tiered data centers. The EPA structured its data center consolidation work to promote cost savings in areas of reduced energy consumption, maximized server and storage use, and reduction in the long-term growth of IT infrastructure costs. EPA is continuing to refine how it calculates DCOI costs savings to ensure better accuracy. Project costs needed to achieve these goals include planning and design activities, updates to data center facilities (i.e., uninterrupted power supply, power distribution unit, and 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. Additionally, the EPA required investment in Wide Area Network (WAN) bandwidth upgrades to accommodate additional traffic that is no longer isolated to the Local Area Network (LAN) infrastructure. All of these investments, while critical to support long-term objectives, make short-term cost savings difficult to attain. In the longer-term, however, the Agency expects these improvements will reduce energy consumption leading to greater efficiency and demonstrable cost reductions.

6 Conclusion

Significant progress has been made to optimize and consolidate EPA data centers. The remaining planned improvements will enable the Agency to meet OMB's DCOI objectives. Figure 2 provides a high-level depiction of the timeline for completing these key improvements. It is important to note, however that these updates require substantial resource commitments to consolidate and update infrastructure and implement the tools needed to gather the data needed to monitor and report OMB's metrics. Investing in these resources will enable the agency to better monitor and reduce energy consumption, but may not produce cost savings in the near-term.

EPA must balance consolidation efforts with network costs and application performance requirements. Because some EPA regional offices, research centers, labs and other facilities host

local infrastructure such as file and print services, telecommunications infrastructure, specialized lab and research systems, and COOP and DR in their data centers, further consolidation can be challenging. As EPA moves forward to achieve DCOI requirements, it will need to continue to balance the benefits of consolidation with operational requirements and implementation costs.