

U.S. Environmental Protection Agency

# DATA CENTER CONSOLIDATION STRATEGIC PLAN



June

2021

## TABLE OF CONTENTS

### Table of Contents

1	Executive Summary .....	3
2	Background .....	4
3	Introduction .....	4
4	Data Center Consolidation and Closure Targets .....	5
5	Cost Savings .....	6
6	Data Center Optimization Metrics .....	7
7	Conclusion .....	9

## 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. On June 25<sup>th</sup>, 2019 the Federal CIO issued M-19-19 “Update to Data Center Optimization Initiative (DCOI)”. M-19-19 rescinds M-16-19 and requires agencies to develop, monitor and report on data center strategies to:

- Consolidate inefficient infrastructure.
- Optimize existing facilities.
- Improve security posture.
- Achieve cost savings; and
- Transition to more efficient data center infrastructures, such as cloud services and interagency shared services.

The EPA made considerable progress towards achieving the targets as established in M-16-19. This year, the EPA closed one of its tier-2 data centers<sup>1</sup> and expects to maintain a high level of availability at the remaining data centers due to redundancy inherent to a tiered data center. In 2021 the EPA was able to deploy the Data Center Information Management (DCIM) tool to the main data center at the National Computer Center (NCC) in Research Triangle Park, North Carolina. The EPA leverages the DCIM in combination with its deployment of Science Logic’s EM7 monitoring tool to capture information on underutilized servers<sup>2</sup> in its tiered data centers. Using this reporting capability, the EPA is in the process of reducing the number of underutilized servers.

In 2020, the EPA created internal connections to Amazon and Microsoft cloud environments. These connections allow the EPA to continue to leverage virtualization and cloud offerings to reduce physical hardware deployment. Work is on track to complete its initial Data Center Information Management (DCIM) tool deployment at the National Enforcement Investigations Center (NEIC) to increase automation and data collection accuracy this year. The EPA does not expect to close any additional data centers in 2021, however a project is underway to remove all enterprise services from the Chicago tier-2 data center. Once enterprise services are no longer in use at Chicago, the EPA will downgrade the facility from tier-2 to a non-tiered facility. The data center downgrade combined with previous closures will allow the EPA to retain a two tier-2 data center<sup>3</sup> environment. Under OMB guidance, the EPA marked all facilities that were not considered tiered data centers as invalid. While the EPA continues to maintain non-tiered data centers essential to agency operations and mission, the remainder of this strategic plan will focus on the tiered data centers maintained by EPA.

---

<sup>1</sup> The EPA closed the One Potomac Yard data center in Q1 2021.

<sup>2</sup> The EPA defines an underutilized server as a server in which the CPU utilization averages less than 30% over the course of a reporting period.

<sup>3</sup> The EPA expects to maintain tier-2 data centers at the National Computer Center in RTP, NC and National Enforcement Investigations Center in Lakewood CO locations.

## 2 Background

In August 2016, the Office of Management and Budget (OMB) released memorandum M-16-19 that established the Data Center Optimization Initiative (DCOI) which superseded the previous Federal Data Center Consolidation Initiative (FDCCI). In June 2019, OMB released memorandum M-19-19 “Update to Data Center Optimization Initiative (DCOI)” that supersedes M-16-19. Memorandum M-19-19 establishes new optimization requirements and rescinds several M-16-19 requirements, such as, reporting for non-tiered data centers, PUE and DCEP. DCOI’s consolidation and optimization requirements have a targeted completion date of September 2020 and are summarized as follows:

- Continue to close non-tiered data centers where viable
- Increase the use of virtualization
- Continue to replace manual collections with automated tools
- Evaluate and report data center costs, savings, and avoidances

## 3 Introduction

Consolidation and optimization of federal data centers continues to be a priority for the EPA. 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, three 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.
- Region 5 Data Center in Chicago, Illinois.
- National Enforcement Investigations Center (NEIC) in Lakewood, Colorado.

To improve efficiency and consistency of data center operations, EPA implemented a Physical-to-Virtual (P2V) initiative at the primary tiered data center (NCC); 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 (COOP) and disaster recovery (DR) in the tiered data centers. The General Services Administration (GSA) also established a five-year agreement with EPA to leverage NCC facilities and hosting services.

For EPA tiered data centers, the agency leverages its existing EM7 tool to support system monitoring requirements of server utilization and virtualization. As of June 2021, the EPA utilizes EM7 at three (3) tiered data centers<sup>4</sup> and reports underutilized servers at two<sup>5</sup>. EPA’s

---

<sup>4</sup> EM7 Deployed to DCOI-DC-45624, DCOI-DC-45634, and DCOI-DC-45623.

<sup>5</sup> Underutilized server reporting available at DCOI-DC-45624 and DCOI-DC-45623.

efforts to consolidate and optimize its data centers have been significant. However, it is important to acknowledge that these efforts involved considerable resource commitments to upgrade data center infrastructure, consolidate data center assets and perform tool assessments. These efforts will enable EPA to better monitor and manage its energy consumption and report on OMB’s optimization metrics, but they may not produce cost savings in the near-term. Additionally, EPA must 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 telecommunications infrastructure, specialized lab and research, continuity of operations and disaster recovery. 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 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.

#### 4 Data Center Consolidation and Closure Targets

##### Achieved

The EPA worked diligently to consolidate and close data centers. At the inception of DCOI, the EPA maintained 58 data centers as defined in M-16-19. As of March 2019, the EPA successfully closed 22 non-tiered data centers and reclassified one data center as invalid due to modifications to the definition of a data center. Per OMB’s criteria, EPA classified four of its data centers as tiered. The tiered data centers support EPA’s enterprise information technology (IT) operations. EPA’s primary tiered data center, is the NCC, located in RTP, North Carolina. In 2018 the EPA downgraded the Region 8 data center from tiered to non-tiered and promoted the federally owned National Enforcement Investigations Center (NEIC) to tiered data center status. This data center, located in Lakewood Colorado, will serve as EPA’s western presence data center. Under FDCCI, EPA closed 21 of its non-tiered data centers. For DCOI, EPA closed 26 non-tiered data centers. The EPA does not anticipate additional tiered or non-tiered data center closures in FY2021. Table 1 depicts a summary of EPA progress from FDCCI through FY 2021.

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

Completed and Planned Tiered and Non-Tiered Data Center Closures by Fiscal Year											
FY 2010	FY 2011	FY 2012	FY 2013	FY 2014	FY 2015	FY 2016	FY 2017	FY 2018	FY 2019	FY 2020	FY 2021
0	1	14	3	1	2	3	3	13	4	3	1
<b>21 Total FDCCI Closures</b>						<b>27 DCOI Closures Completed</b>					
<b>48 Total Closures, 0 Planned</b>											

##### Activities Underway

The EPA closed the Potomac Yard (DCOI-DC-45621) tier-2 data center in FY2021. This closure reduced EPA’s tiered data centers to a total of three. The EPA anticipates some cost savings / avoidance due to the closure through the reduction of lease payment as the assets will migrate to other federally owned facilities. At the time of this writing, the EPA continues to maintain a lease for the Potomac Yard facility.

In FY 2020, the EPA deployed its first enterprise cloud connections to the Amazon Web Services



(AWS) and Microsoft Azure platforms. Through ongoing evolution and improvement, the EPA expects continued growth of these environments to reduce its physical device footprint.

### **Challenges and Risks**

The EPA experiences challenges balancing consolidation and optimization goals with operational requirements and business needs of the regions/program offices. Additionally, the EPA faces challenges in identifying areas for increased cost savings and avoidances. In some cases, EPA's regional offices, research centers, labs, and field offices host local infrastructure assets in support of specialized lab and research capabilities, emergency response (ER) and continuity of operations that may have non-severable configurations and potentially, must remain co-located to operate and fulfill mission functions. The distributed nature of these offices and the continuation of essential localized mission functions require EPA to balance consolidation efforts with continuity of business functions, application performance requirements, cost, and security. Testing is required to validate operational functionality before migrating these systems away from the local site.

## **5 Cost Savings**

### **Initial Costs for Consolidation and Optimization**

Given the upfront investments for infrastructure upgrades, including tool assessment and validation, the EPA does not expect significant near-term cost savings under DCOI. The EPA structured its data center consolidation work to promote cost savings in areas of increased facility utilization, reduced energy consumption, maximized server and storage use, and reduction in the long-term growth of IT infrastructure. EPA is continuing to refine and enhance DCOI cost savings calculations to ensure accuracy. Though OMB agreed there would be negligible savings in the near-term, the EPA has begun gathering additional information to track any realized costs or savings within quarterly progress data calls.

### **Life Cycle Costs Savings and Other Improvements**

Projected costs needed to fulfill identified DCOI goals include planning and design activities, updates to data center facilities (e.g., uninterruptable power supply, power distribution unit, and computer room cooling equipment), validation testing of local applications across the Wide Area Network (WAN), DCIM and security device procurements, training and/or hiring. Additionally, the EPA requires investment in WAN bandwidth upgrades to accommodate additional traffic that is no longer isolated to the Local Area Network (LAN) infrastructure. These investments, while critical to support long-term DCOI objectives, offset any near-term cost savings. In the longer-term, however, the Agency expects these improvements will reduce energy consumption leading to greater efficiency and demonstrable cost reductions.

**Table 2**  
**Historical Costs, Cost Savings and Cost Avoidances**

<b>Agency</b>	Total Achieved savings for 2016 through 2020	OMB Savings Target	Additional Planned through 2021
Environmental Protection Agency	\$0	\$0	\$0

## 6 Data Center Optimization Metrics

OMB established several optimization metrics under DCOI to include virtualization, advanced energy metering, server utilization and availability. Energy efficiency and facility utilization metrics from M-16-19 were removed in M-19-19 and will no longer be reported by the EPA under DCOI. EPA’s implementation strategy for the DCOI metrics is outlined below.

### 6.1 Updated Metric: Virtualization

OMB updated the virtualization metric under DCOI M-19-19. Moving forward OMB requires the EPA to report the number of servers and mainframes that are currently serving as hosts for virtualized or containerized systems at EPA-managed data centers.

- **Planned**

The EPA plans to continue deployment of virtualization technologies such as containers and utilization of cloud. As these deployments are finalized and mature, the EPA expects that regions and program offices prioritize the use of virtualized servers, containers, and cloud deployments as part of equipment lifecycle planning allowing the EPA to naturally gravitate toward a greater virtualized portfolio.

- **Achieved**

As of July 2021, the EPA maintains roughly a 4-to-1 ratio of total servers to virtual hosts as specified under M-19-19<sup>6</sup>.

### 6.2 Updated Metric: Advanced Energy Metering

Under the M-19-19 DCOI guidance, OMB requires agencies to provide energy metering numbers only for data centers that have advanced energy metering. Additionally, previous requirements for advanced energy metering at all tiered data centers under Executive Order 13693: Planning for Federal Sustainability in the Next Decade, were revoked by the Executive Order 13834: Efficient Federal Operations. The EPA continues to evaluate available tools including their planned return on investment to ensure the agency is maximizing value of taxpayer money.

---

<sup>6</sup> As of July 2021, the EPA reports 758 total servers and 173 virtual hosts at its tiered data centers (DCOI-DC-45623, DCOI-DC-45634, DCOI-DC-45624). These numbers were reported in the July 2021 IDC.

- **Planned**

The EPA has advanced energy metering installed at one<sup>7</sup> of the three (3) tiered data centers. The EPA deployed a Data Center Infrastructure Management (DCIM) tool at the DCOI-DC-45624 NCC data center and is in the progress of expanding its use at the DCOI-DC-45634 NEIC data center. Upon completion, the EPA will evaluate the business case for expansion of DCIM to the other non-tiered data centers.

- **Achieved**

Under M-16-19 the EPA maintained an average 1.6 Power Usage Effectiveness (PUE). As this metric was removed under M-19-19, the EPA does not have an average PUE value for advanced energy metering. Future cost savings and avoidance through efficient energy usage and improvements is expected for FY 2022.

### 6.3 Updated Metric: Server Utilization

M-16-19 set specific guidelines for server utilization metrics that created significant challenges. In June 2019, OMB issued revised Data Center Optimization Initiative (DCOI) guidance (Memorandum M-19-19), which allowed individual DCOI agencies to define their own server utilization optimization metric. The EPA leveraged the guidance outlined in OMB Memorandum M-16-19 and collaborated with EPA internal groups to establish an underutilized server definition as server CPU utilization averaged over the reporting quarter to be less than 30%. Using this definition, the EPA identified its Science Logic tool (EM7) for providing automated reporting for server utilization across EPA data centers.

- **Planned**

In FY2021 the EPA plans to continue use of EM7 to report on and remediate underutilized servers. As additional cloud and container options become available, the EPA will evaluate and recommend additional virtualization opportunities to EPA regional and program offices.

- **Achieved**

In 2021 the EPA successfully deployed EM7 to all three (3) tiered data centers. With this deployment, the EPA can provide reports for underutilized servers and progress with justification or remediation. As of May 2020, the EPA removed more than 10 underutilized servers at its tiered data centers through decommissioning and virtualization.

### 6.4 New Metric: Availability

OMB added data center availability as a new metric for DCOI. The EPA will report planned hours of data center availability, and actual downtime from the previous quarter (in hours).

- **Planned**

EPA expects 100% uptime at the three (3) tiered data centers for FY 2021. Due to the redundant configuration in place at the EPA's tiered data centers to include generators, uninterruptable power supplies (UPS), and network redundancy the EPA does not anticipate its data centers to experience any significant downtime.

---

<sup>7</sup> As of May 2020, only DCOI-DC-45624 maintains advanced energy metering.



- **Achieved**

As of July 2021, the EPA experienced 3 hours of total downtime<sup>8</sup> across all tiered data centers. The EPA expects to maintain 100% availability throughout the upcoming fiscal year.

## 7 Conclusion

Significant progress has been made to optimize and consolidate EPA data centers. The remaining planned improvements will enable the EPA to establish a baseline for achieving the vast majority of OMB's DCOI objectives, such as increased virtualization using cloud and container offerings, and increased visibility towards server utilization using of EM7. The EPA will work to balance the achievement of the DCOI objectives with the business needs, and work to overcome technical issues and resource constraints associated with achieving advanced energy metering at all tiered data centers as recommended by OMB. Investing in further consolidation 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. As some EPA regional offices, research centers, labs and other facilities host local infrastructure such as, telecommunications infrastructure, specialized lab and research systems, continuity of operations, and disaster recovery in their data centers, further consolidation is expected to be challenging. As EPA moves forward to achieve DCOI requirements, it will continue to balance the benefits of consolidation with operational requirements and implementation costs.

---

<sup>8</sup> DCOI-DC-45634 NEIC data center experienced 3 hours downtime shortly after completing the build-out. This was reported in the Q1 2021 IDC.