Carnegie Mellon University
Software Engineering Institute

Planning and Design Considerations for On-Premises Computing Environments

Lyndsi Hughes
David Sweeney
Mark Kasunic

July 2021

TECHNICAL NOTE

CMU/SEI-2021-TN-002

DOI: 10.1184/R1/XXXXXXX (Your TC editors will add the DOI to the cover of this report before it is published. https://servicedesk.sei.cmu.edu/jira/servicedesk/customer/portal/6/RS-601)

CERT Division

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

http://www.sei.cmu.edu

Copyright 2021 Carnegie Mellon University.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

The view, opinions, and/or findings contained in this material are those of the author(s) and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by Carnegie Mellon University or its Software Engineering Institute.

This report was prepared for the SEI Administrative Agent AFLCMC/AZS 5 Eglin Street Hanscom AFB, MA 01731-2100

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

Internal use:* Permission to reproduce this material and to prepare derivative works from this material for internal use is granted, provided the copyright and “No Warranty” statements are included with all reproductions and derivative works.

External use:* This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other external and/or commercial use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

* These restrictions do not apply to U.S. government entities.

Carnegie Mellon® and CERT® are registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

DM21-0589

Table of Contents

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

List of Figures

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

List of Tables

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Acknowledgments

The authors would like to thank Timothy Chick, John Stogoski, and Frank Latino for their technical input and comments that improved this report. The authors are also grateful for the help received from the editors, Sandy Shrum and Barbara White.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Abstract

Computing environments that provide access to acutely sensitive data often have a business requirement to restrict unauthorized access to that data. In these cases, hosting an on-premises data center can be the preferred method of providing necessary computing resources to end users to achieve this business requirement. In other instances, cost or special use cases of the system are drivers for building an on-premises data center over using a cloud-based approach. Security concerns can likewise lead to such a decision.

At a high level, deploying a computing environment to support a new capability can seem like a fairly straightforward endeavor. Purchase some hardware, install some software, and enable access to users. However, such a deployment can be a major challenge for small- to mid-size organizations. When establishing an on-premises data center, the initial half of the process—the Plan and Design phase—provides the greatest risks for mistakes and oversights. Most defects that turn up in the later stages of a data center deployment are caused not by problems with the physical components of the system but rather by oversights or misinformed decisions during planning and design. Without appropriate planning and design, the computing environment cannot support the required use cases, and ultimately business operations cannot function.

This report shares important lessons learned from establishing small- to mid-size data centers. These data centers were established within their own organization and for client organizations within the United States government to support development and operations. Their current focus is to establish on-premises data centers that support modern DevSecOps practices and enabling technologies.

This report is intended to help information technology (IT) personnel and management who are responsible for designing and deploying data center technology to become familiar with topics that must be addressed for a successful outcome. While it is beyond the scope of the report to delve into all the details associated with implementing data center operations, it will help IT personnel and management get started.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

1 Introduction

As part of its role as a Federally Funded Research and Development Center (FFRDC), the Software Engineering Institute (SEI) has deployed on-premises data center computing environments within its organization to provide the level of control and security required for certain research and development projects conducted on behalf of the U.S. government.

In addition to establishing organization-based data centers within its own organization, other entities have asked the SEI to provide guidance and consulting for establishing these types of environments within various organizations. We have witnessed many pitfalls first-hand that can cause these types of projects to be delayed or become derailed. This report outlines a framework for planning on-premises data center deployment to help organizations avoid these types of pitfalls.

1.1 Why an On-Premises Data Center?

On-premises data centers are essentially organization-based environments that provide computing, storage, and network capabilities that can be adapted to the internal needs of the organization. They are not subject to the risks, threats, and vulnerabilities associated with moving to a cloud environment. For example, Morrow described a number of vulnerabilities that are associated with moving to the cloud [Morrow 2018, 2019a, 2019b]:

1. Consumers have reduced visibility and control.
2. On-demand self-service simplifies unauthorized use.
3. Internet-accessible management APIs can be compromised.
4. Separation among multiple tenants fails.
5. Data deletion is incomplete.

An on-premises data center provides a high level of control and security and is therefore often used in government and other regulated environments (e.g., financial or medical institutions) to host DevSecOps pipeline environments. This type of data center is also used by other medium-to-large organizations with business-critical operations.

An additional advantage of an on-premises computing environment compared to a cloud-based environment is cost savings. Once an organization invests in building an on-premises data center, it can be used to address future needs for many years to come without additional significant funding. Another cost-related issue is related to the transition from a purchasing model to a service model, which is occurring with many vendors. Many cloud-based licenses become a yearly charge (license and support). A lower cost can be realized by building an on-premises solution because it is a one-time purchase with a lower cost for maintenance support.

Finally, specialized hardware devices may be needed within the computing environment, and it may be impractical or even impossible to integrate specialized hardware into a cloud environment. However, hardware customization can be readily accommodated with an on-premises data center solution.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

1.2 Audience

The primary audience of this report consists of individuals who are responsible for establishing and/or managing a small- to mid-sized data center. Our assumption is that audience members have an information technology (IT) background but may be novices when it comes to the tasks involved in setting up a data center.

1.3 Scope

This report provides a general overview of planning and design considerations involved in establishing a small- to mid-size data center. This report does not offer detailed planning and design advice; instead, it provides a framework for organizing and understanding the logical sequence of the main considerations of data center design. Additional resources are cited about particular topics.

The report provides guidance for both establishing the data center infrastructure and setting up the software-based services that are foundational to data center operations.

1.4 Important Terms

A *data center* is a facility used to host computer systems and associated components, which include telecommunications and storage systems. It must be designed to optimize space and with environment control to keep equipment within specific temperature/humidity ranges.

The *core components* of a data center include equipment and software that support the operations and storage of software applications and data. These components include servers, storage systems, and network infrastructure such as routers, switches, and cabling.

Data centers include a *support infrastructure* that sustains the highest service availability possible (e.g., uninterruptible power sources [UPS] in the form of battery banks, generators, and redundant power supplies). Support infrastructure also includes computer room air conditioners (CRAC), and heating, ventilation, and air conditioning (HVAC) systems. In most cases, support infrastructure should include physical security systems (e.g., badge-entry locks and/or video surveillance system).

1.5 Document Organization

Section 1 of this report provides an introduction and the context for the remainder of the document. Section 2 focuses on the importance of taking a structured approach to planning, designing, and documenting the work. Section 3 shifts to what must be considered when planning and designing the data center. Finally, Section 4 provides a summary of the report.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

2 Getting Organized

2.1 Roles and Responsibilities

Establishing a data center requires the effective collaboration of a number of contributors. Table 1 describes the primary roles and responsibilities of those who participate in the process.

Table 1: Establishing a Data Center

Role	Responsibility
Business/Mission Leader	Provides the business case and criticality of the data center; establishes facility constraints and parameters of a growth plan; approves the budget
IT Management	Collaborates with business/mission leaders to advise and support decision making; participates in or reviews the data center design
Data Center Manager	Leads/facilitates the design of the data center working with data center administrators
Data Center Administrators	Collaborate with the Data Center Manager on aspects of planning and design

2.2 The Planning and Design Approach

From a high level, deploying a computing environment to support a new capability can seem like a fairly straightforward endeavor. Purchase some hardware, install some software, and enable access to your users. However, such a deployment can be a major challenge for small- to mid-size organizations. Executing the deployment is an incredibly complex undertaking, and the path from conceptualization of a need to realization of a capability requires many decisions. The complexity of the deployment is compounded because each decision has the potential to influence later tasks and activities. In fact, early planning activities offer the greatest potential for down-stream mistakes due to oversights or misinformed decisions.

For example, organizations often start searching for the data center space before their design criteria and performance characteristics are defined. This premature search can turn into a real problem if the site cannot meet the design requirements. Similarly, hardware procurement might take place independently of software procurement, which can result in blockers that prevent downstream activities. Even seemingly innocuous items, like the standard power cables shipped with servers or network devices can create problems. If they are not compatible with the existing power infrastructure, they can prevent an organization from even being able to power up their hardware. Fortunately, you can avoid these problems by carefully considering the ramifications of decisions and developing an ordered plan for deployment and decision making.

Figure 1 illustrates where planning and designing fit in the context of establishing a data center. The Plan and Design phase lays the foundation for everything that happens later. Taking a thoughtful and careful approach during this phase is critical to the successful implementation of the data center.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

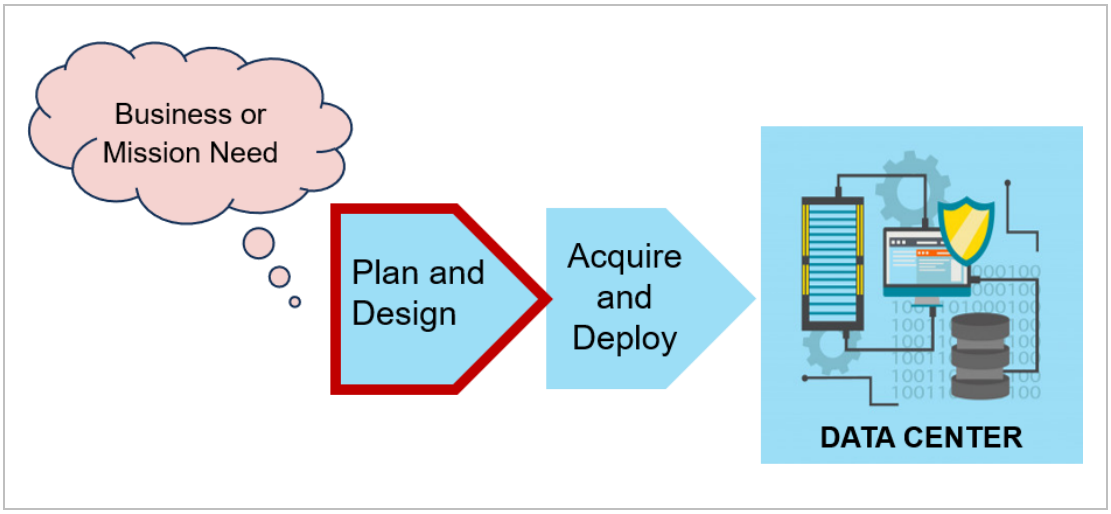

Figure 1: Plan and Design in the Context of Establishing a Data Center

The Plan and Design phase is informed by an existing business or mission need. During the earliest stages of the project, the business/mission leader(s) supply the vision for the project as expressed in the need, and document and communicate the vision and mission to all stakeholders participating in the data center implementation. At this point in the process, the conversation is rather abstract and framed in language that expresses the need in terms of the business or mission. However, it is important that the discussion is at the appropriate level of abstraction without getting into technical details that must be worked out later in the Plan and Design phase. The following are examples of appropriate statements of need that focus on the business/mission.

- Our company develops custom applications that are highly proprietary and that contribute to competitive advantage. Therefore, these mission-critical systems and the intellectual property must be maintained in house.
- Our service level agreement states that we must provide 24/7 service to end users.
- We need to support 25% more users after the first three years of operation.
- Our budget requires the data center to be housed within this building.
- We are contractually/legally obligated to be HIPAA compliant.
- Our timeline for deployment is constrained by the terms of our existing lease agreement(s).

These statements are not detailed in technical implementation terms; instead, they reflect the business needs at the appropriate level of abstraction. The requirement inferences and technical implementation decisions are left to subsequent downstream planning activities by individuals who are competent at analyzing the various tradeoffs that will lead to an implementation that meets your organization's business/mission needs. It is critical that technical personnel, when working out the solutions that meet the business needs, are able to map their technical decisions to the business requirements for the system. This mapping allows technical personnel to both effectively communicate and justify their technical decisions to business/mission leaders.

Figure 2 is a conceptual depiction of how interdependencies affect planning and design decisions. On the right side of the figure, *design drivers* is based on business/mission needs and overlays the various decisions made throughout the Plan and Design phase.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

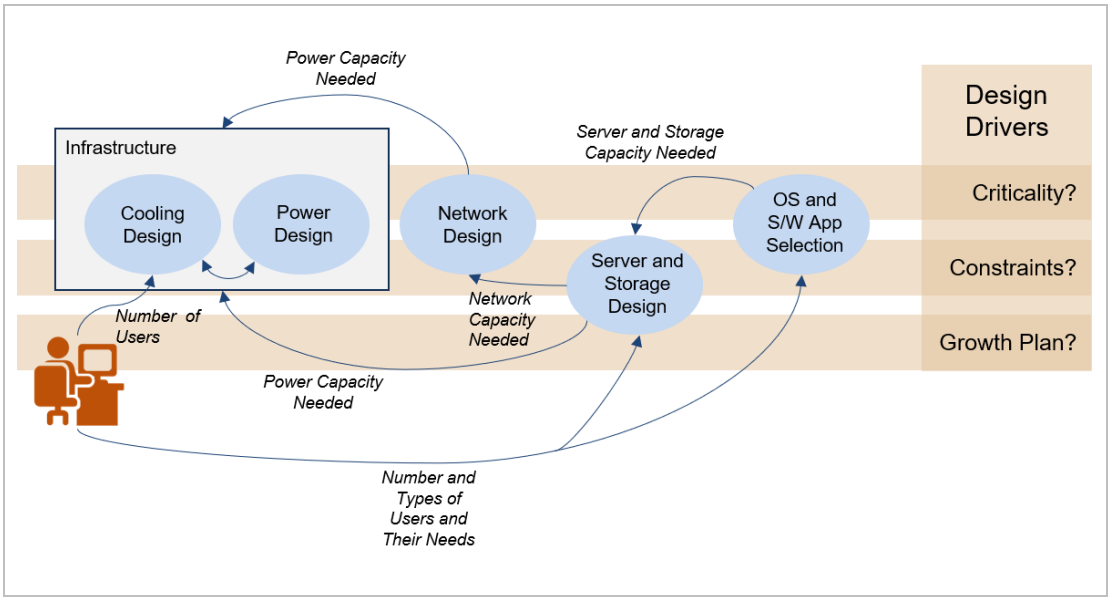

Figure 2: The Interdependencies of Planning and Design Decision Making[1]

The design drivers represent areas that must be taken into consideration throughout the Plan and Design phase of the project. They are essentially high-level requirements that impact lower level design decisions of the data center. Design drivers are discussed in more detail in Section 3.1.1.

The arrows in Figure 2 relate to the interdependencies of planning and design decisions among the various constructs of the data center. Most importantly, a sequence is implied within the diagram.

1. Once the design drivers are characterized, determining your users' needs is a key activity that results in identifying the types of applications to be hosted, and the platforms that are best suited to hosting those applications. Applications under consideration must include both those accessible on the users' endpoint systems and those hosted in the data center. The applications under consideration should also include evaluating how the applications integrate with one another.
2. Requirements derived from #1 inform the required capacity to be considered for computation, storage, and networking.
3. Design decisions associated with #2 imply both a network design and a power capacity needed to drive the devices for both user endpoint systems and data center systems.
4. Decisions up to this point lead to infrastructure design decisions regarding the power and cooling designs that must be accommodated to maintain the range of temperature and humidity required to safely operate the network, storage, and computing components.

There is an important sequence of decision-making events that must be carried out to avoid downstream problems. Unanticipated changes during implementation can have major implications, including creating downstream calamities.

[1] In the diagram, *OS* is operating system and *S/W* is software.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

During the Plan and Design phase, some requirements might conflict with design drivers or other requirements. In those cases, you might be required to ask management to intervene and resolve the conflicts through targeted tradeoffs leading to subsequent modifications to the plan.

2.3 Documenting Your Plan and Your Progress

As mentioned in the previous section, while devising the data center design and the plan for implementation, there are many inputs from various stakeholders, ranging from abstract or theoretical ideas to detailed technical implementation requirements. It is critical to ensure alignment toward a common goal given this input. Figure 3 is a conceptual diagram showing the progression of planning and design that is required before moving to the phase where data center components are selected, acquired, and deployed to establish the data center.

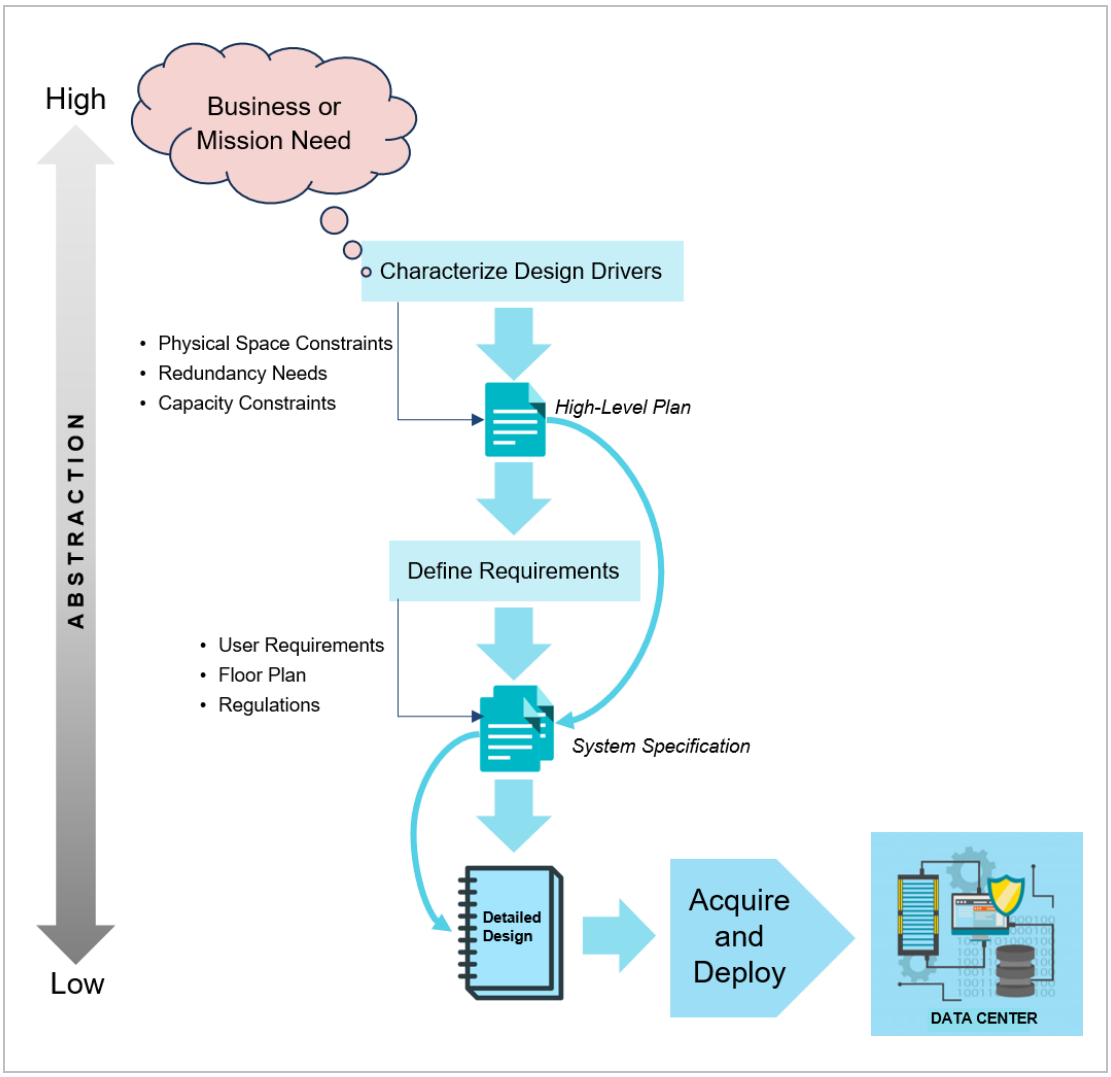

Figure 3: The Progression of Data Center Planning and Design Documentation

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

The key takeaway from Figure 3 is that you must document at all levels of abstraction and make that documentation accessible to all project participants. This documentation includes the abstract definition of the business need, the user requirements/stories, the detailed trade-off analyses when selecting applications, and the implementation details of a particular piece of software. It is crucial to accurately document these decisions since the Planning and Design phase evolves from a concept to a detailed design. Accurate documentation contributes to a clear and shared understanding of the project's mission for all stakeholders, ultimately improving the team's alignment with the end goal.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

There are other important takeaways from Figure 3.

1. Documenting needs, requirements, trade-off analyses, and decision making is crucial as planning and design evolves from a concept to the detailed design needed to move forward with acquisition and deployment of the data center.
2. There is a progression of data that flows and transforms as the abstractions advance from vague to detailed. New data should be traced to earlier decision to ensure that the rationale for design decisions is available to reference, review, and audit if necessary at any point in the sequence.
3. Once the data center is deployed, the rationale and intentions of earlier planners and implementors is preserved by the detailed design documents in the event of personnel turnover during the project.
4. If downstream problems occur, they can be traced back to previous data and/or decision making to aid in troubleshooting or re-evaluation.

The importance of creating and maintaining documentation throughout the project—and referring to it frequently—cannot be understated.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

3 Planning and Designing Your Data Center

This section contains guidance for the Plan and Design phase of establishing a data center. Figure 4 provides an overview of the planning and design process.

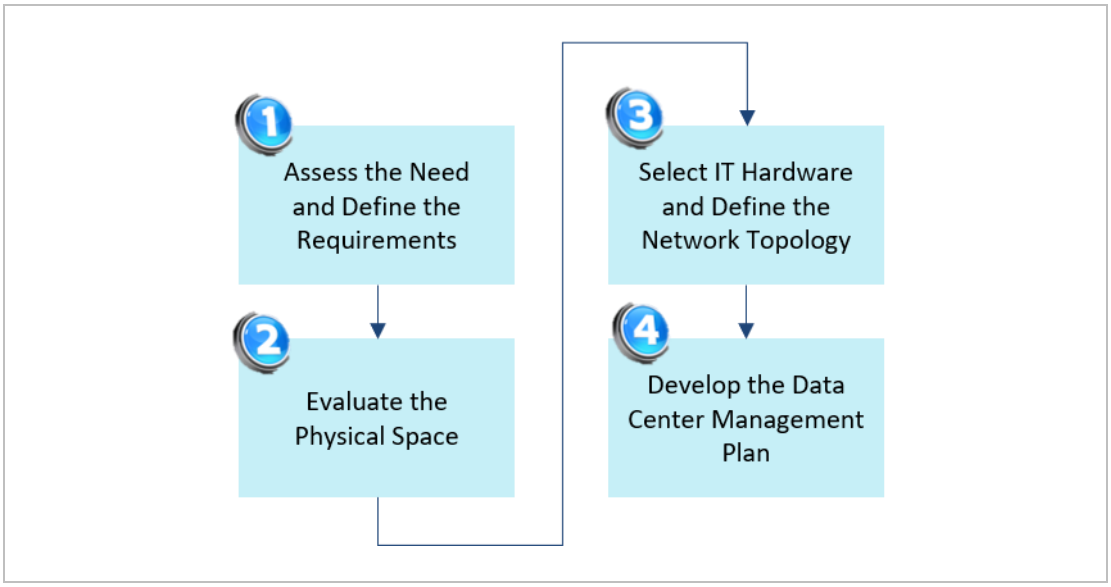

Figure 4: Overview Framework for Planning and Designing a Data Center

The following are the stages of planning and designing a data center:

1. Kick off the Planning and Designing phase with the all-important assessment of the stated business needs and requirements analysis.
2. Analyze the physical constraints of the data center location, which leads to important decisions about space limitations and facility layout.
3. Based on both design drivers and user requirements, demarcate the hardware components and the supporting infrastructure.
4. Develop a plan that specifies the manual and automated processes required to properly deploy, service, and manage the data center.

The flow from stage 1 to stage 4 is critical because each stage prepares the foundation for the next stage. In some cases, you might have to iterate to a previous stage to revisit an earlier decision in light of new information. For example, if you identify a conflict with a previous design decision during stage 3, you might have to return to stage 1 to modify a requirement so that the conflict is resolved.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

3.1 Assess the Business Needs and Define the Requirements

Figure 5 illustrates the main activities that kick off the project:

1. Define and document the design drivers.
2. Collect, analyze, define, and document user needs.
3. Identify and document any regulatory requirements that you must address in your data center design.

At this point, the expressed business needs begin to transform into the detailed information the technical personnel require to deploy new capabilities.

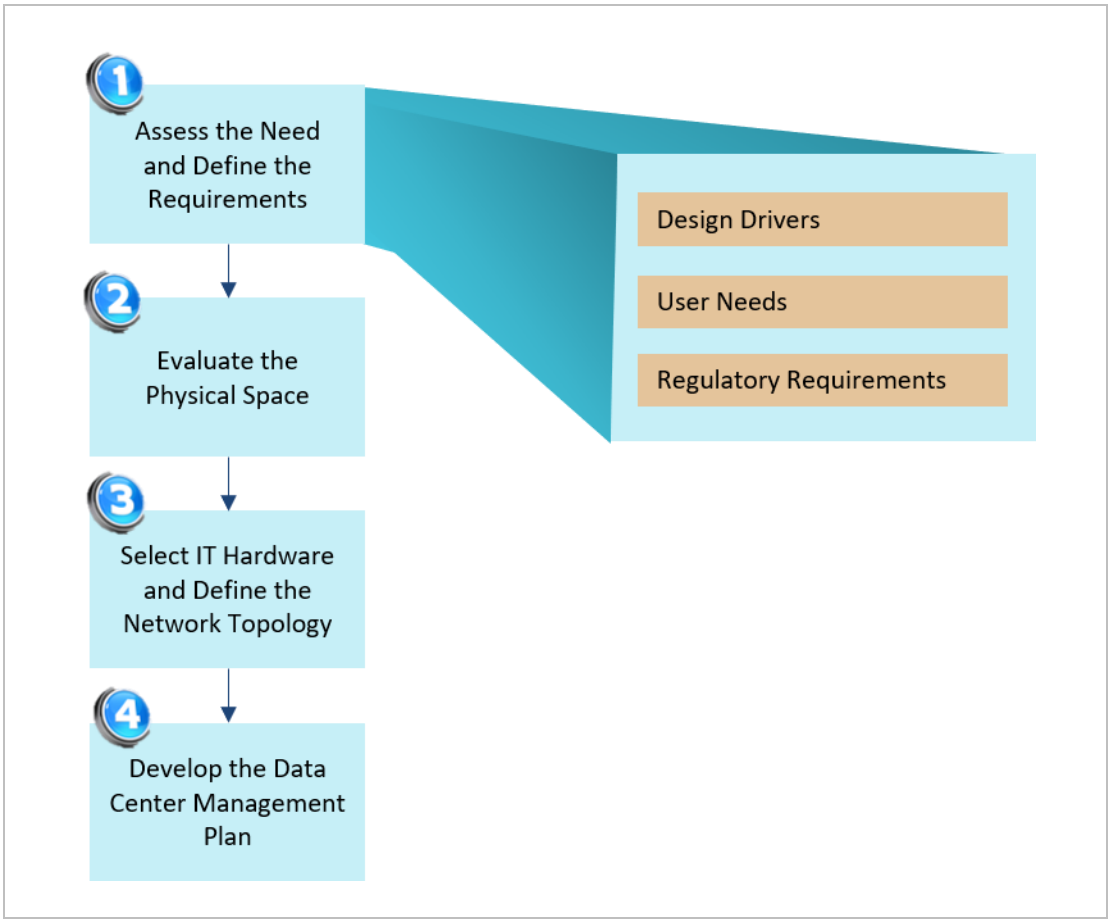

Figure 5: Assessing the Need and Defining Requirements

3.1.1 Determine Design Drivers

As a technical person on the project, it's likely that your organization's leadership has approached you to organize the project because there is a business or mission need for the capability that the facility will provide. While leadership likely does not have the technical expertise to specify the details that will need to be worked out (that's your job), they certainly have interests that they communicate up front. These interests establish the vision and scope for the data center. From

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

your perspective, the first step of the Plan and Design phase is to understand the vision provided by leadership and translate it into the project's *design drivers*.

Design drivers are significant considerations that inform and guide the design of the data center. In some cases, design drivers are related to the goals of the data center or are the primary high-level requirements for the data center. Design drivers are typically related to business/mission needs that have been ordained by the organization's leadership.

As illustrated in Figure 6, we identify the key types of design drivers: criticality, constraints, and growth plan.

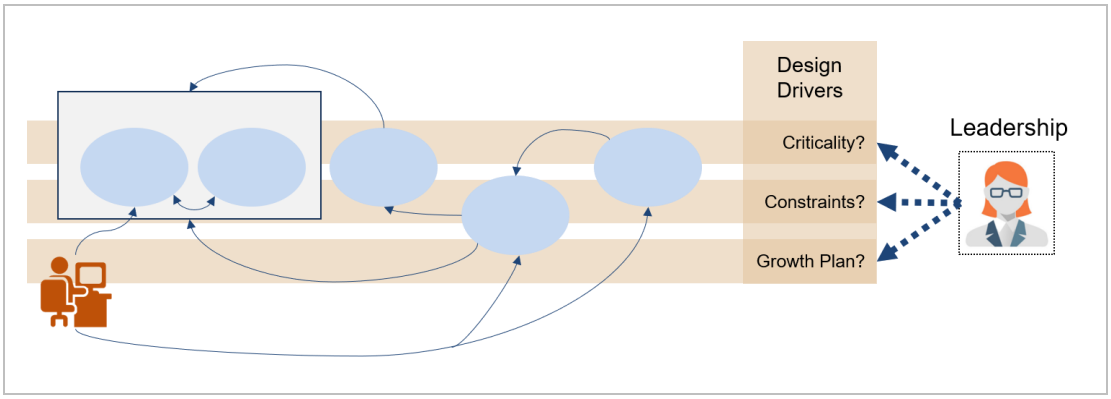

Figure 6: The Influence of Design Drivers on Establishing the Data Center

3.1.1.1 Criticality

Broadly speaking, *criticality* is driven by the purpose of use or value proposition of the data center. Understanding the purpose of use and the mission need is the key to understanding and deconflicting issues as you progress through the definition of data center requirements.

With respect to data center design, criticality specifically refers to how indispensable and vital the components of systems must be. Different systems within a data center, and even separate components of a single system, may have different criticalities. This type of design driver might be manifested in a business need that email servers must be operated with 99.99% uptime. This type of service requires a design that includes sufficient redundancy to operate effectively during both scheduled maintenance periods and unscheduled environmentally caused failures. It may even imply a requirement that some email servers be hosted in a secondary remote location.

3.1.1.2 Constraints

Constraints are factors that compel certain decisions or limit the scope of available solutions. In our experience, this design driver has the largest potential scope of consideration. Failing to adequately consider design drivers that fall into this category has the largest negative impact on data center deployments. Constraints include budget, the geographical location of the facility, the space reserved for the data center within the facility, organizational policies, contractual/service level agreements, and regulatory requirements.

This type of design driver might be manifested in a business need to adhere to a compliance standard, such as the Health Insurance Portability and Accountability Act (HIPAA). Complying

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

with this type of regulation implies the implementation of many security requirements, both virtual and physical.

Another design driver in this category is a limited timeline for deploying an application. This type of constraint can imply a requirement that only the components required to achieve a minimally viable product be deployed first. Another example of a design driver in this category is the facility owner's rules for the use of the building, which may limit how your organization is permitted to modify the facility.

3.1.1.3 Growth Plan

The growth plan is the anticipated increase in capacity of some factor over a defined period of time. Relevant factors include the number of users, customers, or products, and having an idea about how these relate to the amount of storage capacity, computing capacity, or network capacity required to support them.

This type of design driver might be manifested in a business need that user storage quotas must be doubled in two years. Supporting this type of growth implies requirement(s) for storage technology that supports some degree of scalability.

3.1.1.4 Summary

The design drivers for your data center should be well understood to derive accurate requirements for the data center and its infrastructure. An incomplete understanding of these drivers will likely result in technical debt (e.g., duplication of effort, incorrectly specified hardware and/or software). Accurately defining design drivers for a complex system such as a data center is not a trivial task.

Design drivers, their descriptions, and their implications are documented as part of the Plan and Design phase of establishing the data center. Refer to this documentation as necessary to keep the project properly scoped and on track.

3.1.2 Determine User Needs and Requirements

Figure 7 illustrates the relationship of user requirements to the business needs and design drivers (leadership view) and the system requirements that represent both the functional requirements (i.e., what the data center does) and non-functional requirements (i.e., how well the data center does it).

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

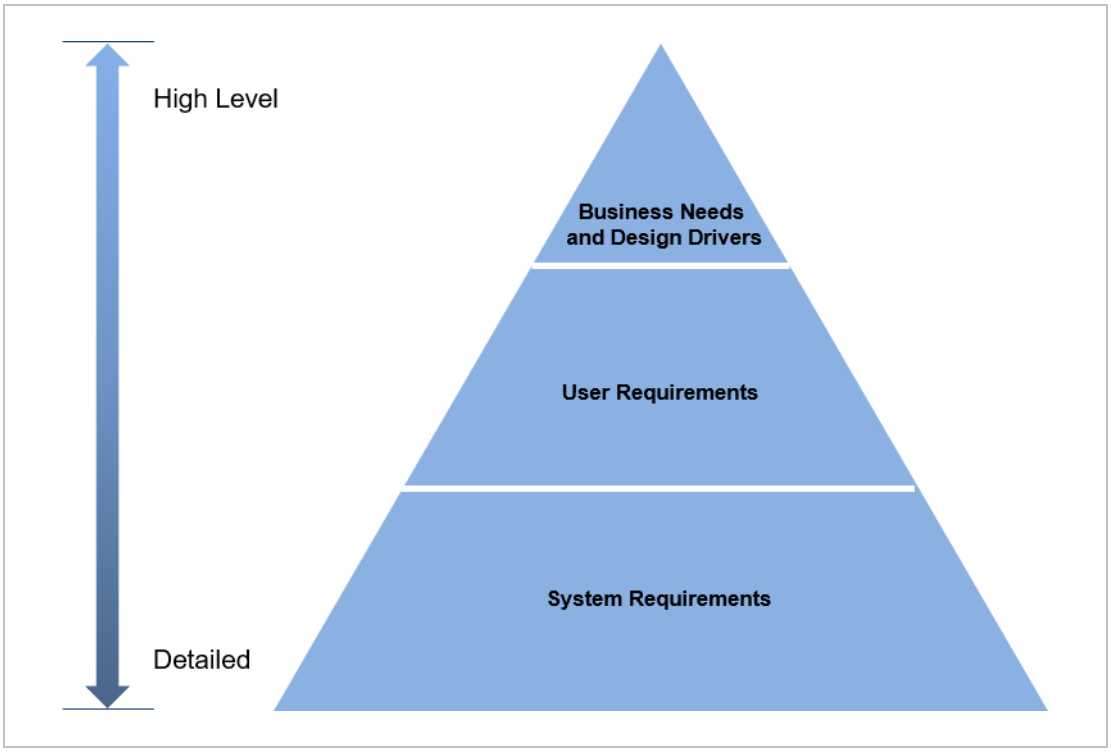

Figure 7: Requirements Gathering Pyramid

This task involves collecting user requirements and evaluating them to determine whether they are valid and should be prioritized or if they are user *preferences* that are subject to change to avoid cost issues with the budget and technical compatibility issues.

User requirements can be things such as features and options and can include facility and/or logistical constraints. Determining user requirements requires discussions with the anticipated users of the data center. Table 2 lists the areas that you should explore with the anticipated user community at this point in the process.

Table 2: Example User Requirement Areas to Investigate

Requirement Area	Questions to Ask
Facility	How many users will you have in the short term and, if there is a growth plan, in the longer term? Will users work in the same facility that houses the data center? Will users work directly in the data center, in separate physical spaces, or remotely? What are the requirements for how your users interact with their endpoint systems? Do end users need physical access to endpoint systems? If so, do they require access 24/7? What accommodations (e.g., kitchenette, coffee maker, microwave) do your users require?

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Requirement Area	Questions to Ask
Use Types	How will the data center be used (e.g., application development and testing, research and analysis, operational use, mixed use)? Will users be divided into multiple tenant groups? What types of endpoint workstations are required (laptops, thick clients, thin clients, zero clients, special use embedded systems)? What types of software applications do the users need to support their work? Will the application(s) be purchased or licensed? Will any special-use hardware (e.g., cell phones, test benches, antennas, robots, vehicles) need to be integrated with the data center? Will software and/or hardware needs change over time?
Power Consumption	What is the power consumption of the users' workstations and peripherals? Can the building/facility support the power consumption needs of the user workstations and the data center, or will an upgrade be required?
Data Storage	How much data storage is required to address the user needs? What types of data structures will be used (e.g., database, large file storage, small file storage, network storage, object storage)?
Security	What are the physical security needs (e.g., access, screen hiding, individual offices versus cube farms, alarms) if any? Are private offices required for some/all users? Are windows to the outside permitted in end-user workspaces?
Communication	Will users work remotely (i.e., off site), or in the same building as the data center, or will they need to occupy the data center when using its features? Will an Internet connection be required? Will Wi-Fi be required? Will the data center need to communicate with other networks? If so, which ones? Are access points currently available in the facility? Will users or administrators need to be available outside of working hours in case something goes awry or to get status updates? Will the data center provide users with updates or status outside of the data center (e.g., email, text, other messaging service)?
Timeline	By what date do the users need to be able to access the computing environment's resources? What features need to be made available to end users first? What features can be implemented at a later date?

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Requirement Area	Questions to Ask
Disaster Recovery [Ready 2021]	What is the plan to safeguard data, critical IT infrastructure, systems, and networks? What are the possible disasters? Is there a mitigation procedure for each possible disaster that is likely to occur? Is there an inventory of IT assets and an assessment of the criticality of each asset? What is the data backup plan?

Consider developing a standard questionnaire for user requirement solicitation to support consistency. Various methods can be used to collect requirements including individual surveys, focus group interviews, or phone interviews.

Once you conduct the requirements solicitation process, develop *personas* that characterize each type of user. For example, personas might include the following:

- Data Center Manager
- Software Developer
- Security Personnel
- IT Operations Personnel
- Data Analyst
- Researcher

After developing appropriate personas, develop scenarios to assist in elaborating the user requirements. Consider having members of each persona review your elaboration to confirm or modify the requirements.

Requirements analysis involves the following steps:

1. Compare requirements to design drivers to determine if there are mismatches in terms of expectations.
2. Determine which requirements are actually preferences versus those where a real need is being specified.
3. Categorize and prioritize the requirements.

After this analysis, feed the pertinent information into the design and tactfully set the expectations of both the end users of the data center and the implementors/maintainers of the data center, ensuring that their expectations align with the leadership vision of the data center.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

3.1.3 Determine Regulatory Requirements

For many organizations, regulatory compliance is a topic that cannot be ignored. Handling confidential data has become an essential task in almost every organization. Table 3 is a non-exhaustive list of regulatory standards that some organizations must comply with. When planning and designing the data center, be sure to consider all regulatory requirements that apply to your organization and the types of work conducted within the data center. These requirements can range from organizational to national to state to local regulations.

Table 3: Examples of Regulatory Standards and Requirements

Standard	Description
ISO/IEC 27001:2013	The International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) standard evaluates risk to information assets. Complying with this standard demonstrates that an organization has a functioning Information Security Management System (ISMS) in place to identify, analyze, and address risks [ISO/IEC 2013].
SSAE 18	The Statement on Standards for Attestation Engagements (SSAE) standard addresses internal controls pertaining to financial reporting and how organizations evaluate and report risks associated with their vendors. The standard is important for service organizations that are obligated to protect client data that they or their subcontractors handle directly [AICPA 2016].
HIPAA	The Health Insurance Portability and Accountability Act (HIPAA) sets the national standards for how organizations should manage protected health information [ASPE 1996].
SOC 2 Type II	Conformance to System and Organization Controls (SOC) 2 Type II demonstrates that the appropriate information security policies and procedures are in place to protect customer data [AICPA 2021].
DISA STIGs	DoD information technology teams must comply with technical testing and hardening frameworks referred to as *STIGs* (Security Technical Implementation Guides). The Defense Information Systems Agency (DISA) STIGs contain technical guidance on how to "lock down" information systems that might otherwise be vulnerable to malicious computer attacks [DoD 2021].
DFARS	The Defense Federal Acquisition Regulation Supplement (DFARS) provides a set of adequate security controls to safeguard information systems that host contractor data. Manufacturers must implement these security controls through all levels of their supply chain [NIST 2021a].
Risk Management Framework	This risk-based approach to security control selection and specification considers effectiveness, efficiency, and constraints due to applicable laws, directives, executive orders, policies, standards, or regulations [NIST 2018, 2020].

3.2 Evaluate Physical Space Constraints

In some cases, organizations have a large budget with little or no constraints associated with the physical location and space for the data center. These data centers are referred to as "Greenfield" data centers where the project starts with a piece of land with no building on it, and—from the

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

very start—the facility is designed to be a data center. In these cases, the system specification can be based on the optimal characteristics that satisfy all user and regulatory requirements. However, in many cases for small- to mid-size data center deployments, an *acceptable* physical space is provided, and the data center designer(s) works within the constraints of the space to provide the optimal capability within the limitations.

In this section, we assume the latter situation, where an adequate physical space is provided for the data center. Assuming that, the challenge is to learn about constraints that will affect procuring appropriate equipment to realize the capability specified in the documented requirements for the project. When the physical space poses limitations on previously defined requirements, it may be necessary to modify the requirements or seek management intervention to address the conflict.

Figure 8 illustrates the issues that must be considered and characterized as part of evaluating the physical space. In particular, three of the five issues pose potential limitations to the amount and types of equipment that can be accommodated within the data center, including (1) power capacity, (2) cooling capacity, and (3) floor space layout.

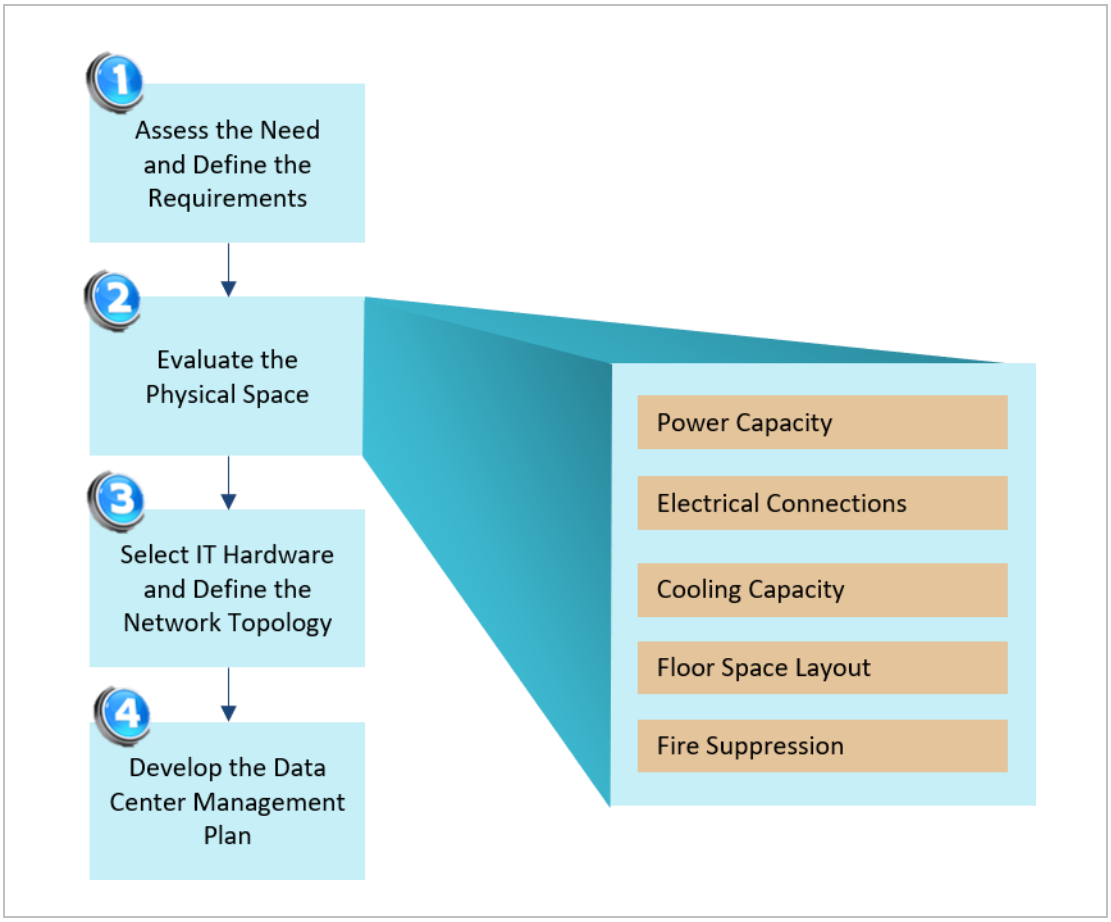

Figure 8: Evaluate Physical Space

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

3.2.1 Power Capacity

The power requirement for data centers is typically much greater than what is needed for a home or office environment. The electrical system supports many different types of equipment within the data center. Examples include the IT components (e.g., servers, storage devices, monitors) but also air conditioners, fans, pumps, lighting, environmental monitors, etc. Estimate the power requirements for the data center and ensure that you have sufficient power to support system availability. One way of doing this is to build a spreadsheet listing the maximum power specifications for each physical system you plan to install in the data center. If your estimate exceeds what the location can provide, you will need to either scale back your plans to fit within the constraint, or investigate if the power can be upgraded.

A power distribution unit (PDU) is a device fitted with multiple outputs that distribute electric power to racks of servers and networking equipment. It is designed to handle the voltage/amperage requirements of data center devices. The PDU that you plan to use in your data center must be compatible with the electrical service provided to the facility. For multiple circuits, you must ensure that each circuit does not exceed the amperage supported on the circuit.

Power redundancy is a key consideration. If reliability is an important aspect of your data center operations, design how power will be maintained during an outage. Options include (1) a backup engine-generator and (2) an uninterruptible power supply (UPS). A UPS differs from a backup generator in that it provides instantaneous protection from input power interruptions by supplying energy stores in batteries or flywheels. Maintaining uninterrupted power is key to preventing data loss during a power outage of any length. Power redundancy provides crucial time to enable either waiting for the power to return or following a procedure to prevent data loss.

3.2.2 Electrical Connections

The design of the power distribution system must also consider how connections to the various data center equipment will be accomplished. Care must be taken with the connector appliance and the wiring itself.

There are two types of wiring: (1) one-phase wire (sometimes referred to as single-phase wire or split wire) and (2) three-phase wire. One-phase wiring has three wires located within the insulation—two hot wires and one neutral wire. Each hot wire provides 120 volts of electricity. Three-phase wiring provides power supplied by four wires—three hot wires carrying 120 volts and one neutral wire. Three-phase wiring is more efficient than one-phase wiring.

Care must be taken when considering connectors. Some connectors appear to have the same shape when inspected visually, but their operating specifications are different. An example of this is the L6-20R and L5-20R connectors. These connectors look the same, but they are not interchangeable. Be forthcoming with your hardware vendors about the type of power supplied in your data center and how it is distributed. The vendors can work with you to specify the types of cords and connectors you should purchase with your equipment so that you can successfully operate the equipment at your facility.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

3.2.3 Cooling Capacity

Operating the electrical equipment in the data center generates heat. Since the equipment has specified temperature operating ranges, the heat must be removed to prevent equipment failure. If the heat is allowed to build up, servers will either shut down or, if they operate at higher-than-recommended temperatures for extended periods of time, their lifespan will be shortened. Heat is removed from IT equipment by running a stream of cooling air across the electronics.

Various cooling architectures can be designed to move heat from inside the data center to outside [Evans 2004]. They include the following:

- air-cooled systems (two piece)
- air-cooled self-contained systems (one piece)
- glycol-cooled systems
- water-cooled systems
- chilled water systems

Be sure that you understand the limits of the cooling capacity in your data center and that the cooling needs of your IT equipment do not exceed the data center's cooling capacity. Supplemental cooling devices are available and can be considered as backup systems if needed. Consider using environmental monitoring devices in the data center design to generate alerts when environmental conditions, such as temperature and humidity inside the data center, approach less than ideal levels.

3.2.4 Floor Space Layout

The floor layout plan specifies the boundaries of the room(s) and the layout of IT equipment within the room. These preliminary floor layouts do not require you to specify the specific location of IT equipment. Instead, you only need to consider the location of equipment racks, other cabinets, and user space (if needed). The floor layout includes the structural layout of the empty room and the equipment layout of what will go in the room. Since this report assumes a pre-existing space, the only option is to lay out the equipment locations (i.e., racks, monitors, cabinets).

To begin the process, identify and locate any room constraints such as the following:

- columns (their dimensions)
- doorways
- existing room equipment (e.g., pipe connections, cooling equipment, fire suppression equipment)

The floor plan strongly impacts the number of rack locations that are possible in the space and hence the amount of IT equipment that can be accommodated. The plan also affects the cooling distribution, which is critical for operating IT equipment.

A well-established layout for a data center uses the hot-aisle/cold-aisle rack layout. The basic idea is to separate IT equipment intake air and exhaust air by establishing *cold aisles* where equipment is positioned with the intake facing the aisle. Conversely, *hot aisles* are established where the

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

equipment on the racks is positioned with the exhaust facing the aisle. The basic concept is illustrated in Figure 9.

The objective of this layout is to reduce the amount of hot exhaust air that is drawn into the IT equipment intakes. Additional guidance can be found in the publication, *Data Center Power Equipment Thermal Guidelines and Best Practices* [ASHRAE 2016].

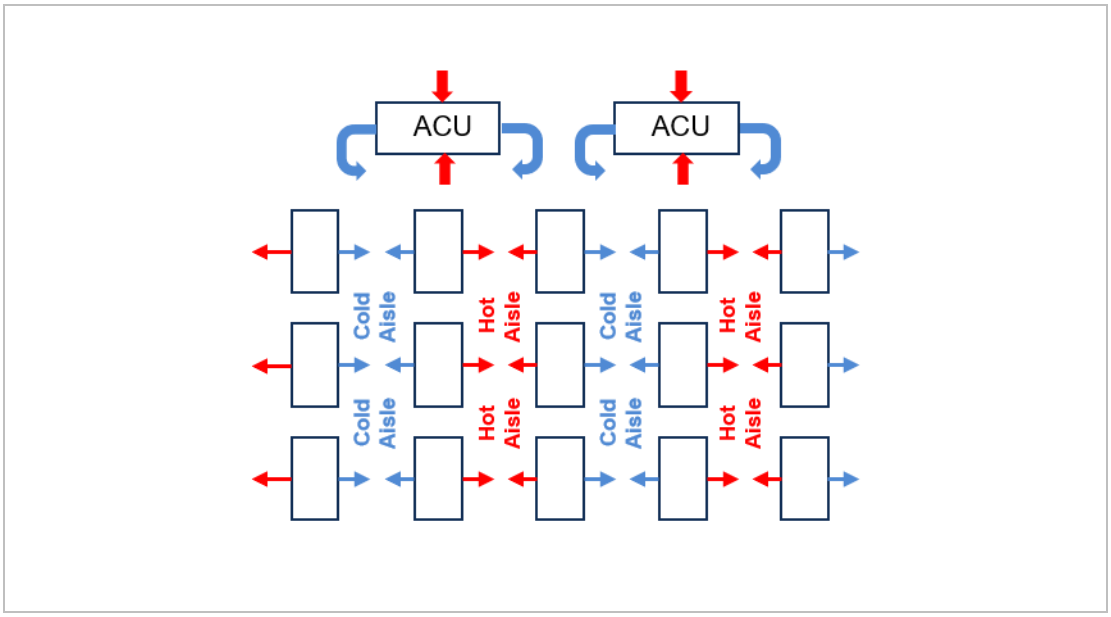

Figure 9: Basic Layout for a Hot-Aisle/Cold-Aisle Configuration[2]

3.2.5 Fire Suppression

Fires are not a common cause of data center downtime. However, they do occur and can cause major disruptions. Unless you are building a data center from scratch, you will have little or no control over the design of the fire suppression system. That said, you need to understand what is in place and how it will protect the data center systems.

Table 4 describes the components of a fire suppression system [Ingram 2020].

Table 4: Components of a Fire Suppression System

Component	Description
Detection System	The detection system should detect both smoke and heat to signal a fire.
Fire Alarms	Fire alarms should be integrated with the fire suppression system so that the timing of the release of the suppression agent allows for orderly exit from the data center.

[2] In the diagram, ACU is air conditioning unit.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Component	Description
Emergency Power-Off Switch	The emergency power-off switch automatically cuts the power in the event of a fire. An additional switch should be considered for cutting all power manually.
Portable Fire Extinguishers	Portable fire extinguishers should be provided, but it is important that staff members understand that in the event of a fire, they must leave the room. They should not attempt to fight the fire on their own with portable fire extinguishers.
Suppression Agent	Suppression agents can vary from a water-based system to different types of gaseous suppression systems.

Regarding the suppression agent, a water-based system is the least attractive of the alternatives because of the damage that can result from its use. Also, condensation from the sprinkler pipes and leaks can drip on equipment and invite disaster. Gaseous fire suppression agents work relatively quickly and cause minimum damage to equipment. Whatever your situation, your data center operations plan must include contingencies for what you will do in the aftermath of the fire suppression system being triggered.

3.3 Identify Hardware Components and Design the Network

The next stage of the Plan and Design phase involves determining what hardware to install and how that hardware is connected via a network. Figure 10 highlights the topics to be discussed in this section.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

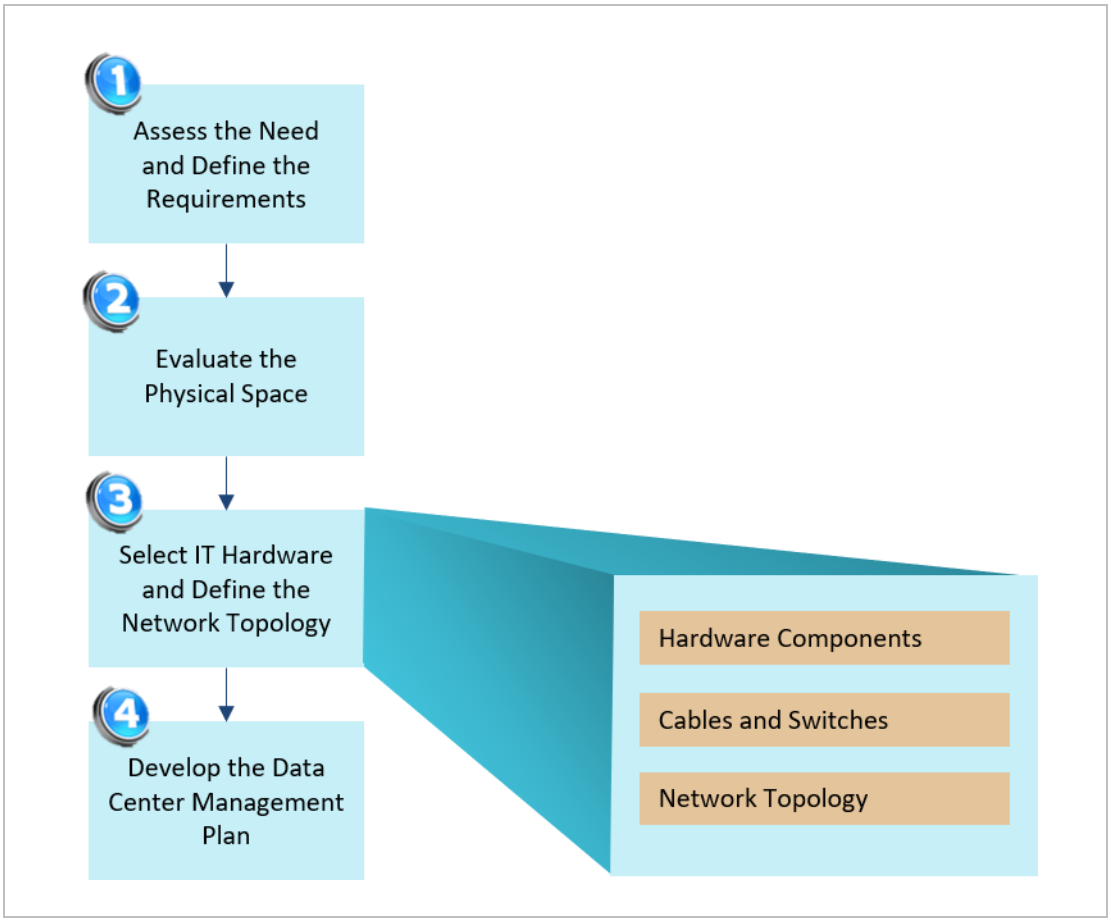

Figure 10: Overview of Select IT Hardware & Define Network Topology

As hardware components are selected and information is fed into the planning and design of the data center, it is important to keep in mind the amount of time required to acquire hardware components. Be sure to understand the lead-time required to acquire hardware and software—both your organization's procurement lead time and the lead time the vendors need to deliver their products. During acquisition and deployment, this information is invaluable for devising reasonable estimates for when tasks can be completed. Since many later tasks depend on the completion of earlier tasks, having this information ahead of deployment greatly aids planning.

3.3.1 Hardware Components

Hardware components include servers and storage systems.[3] The number and type of servers that you select is driven by the requirements (as described in Section 3.1).

When it comes to servers, there are different types that can be considered. Table 5 briefly describes the different types of servers by function.

[3] This list is not exhaustive. Depending on the need, other specialized hardware could be included in a data center. An example is an Application Delivery Controller.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Table 5: Different Types of Servers by Function

Server Type	Description
Database Server	A server computer that runs database software
Application Server	A server computer that runs a specific application that requires a server (e.g., an accounting application)
Web Server	A server that runs software that enables hosting an Internet website
File Server	A server computer that provides centralized disk storage that can be shared by client computers on the network
Print Server	A server computer that collects information being sent to a shared printer by client computers
Mail Server	A server computer that handles the network's email needs (It can also handle audio and videoconferencing, chat rooms, and instant messaging.)

For a data center that services applications that require large amounts of data, a file server may be connected to networked-attached mass storage. (See Figure 11.)

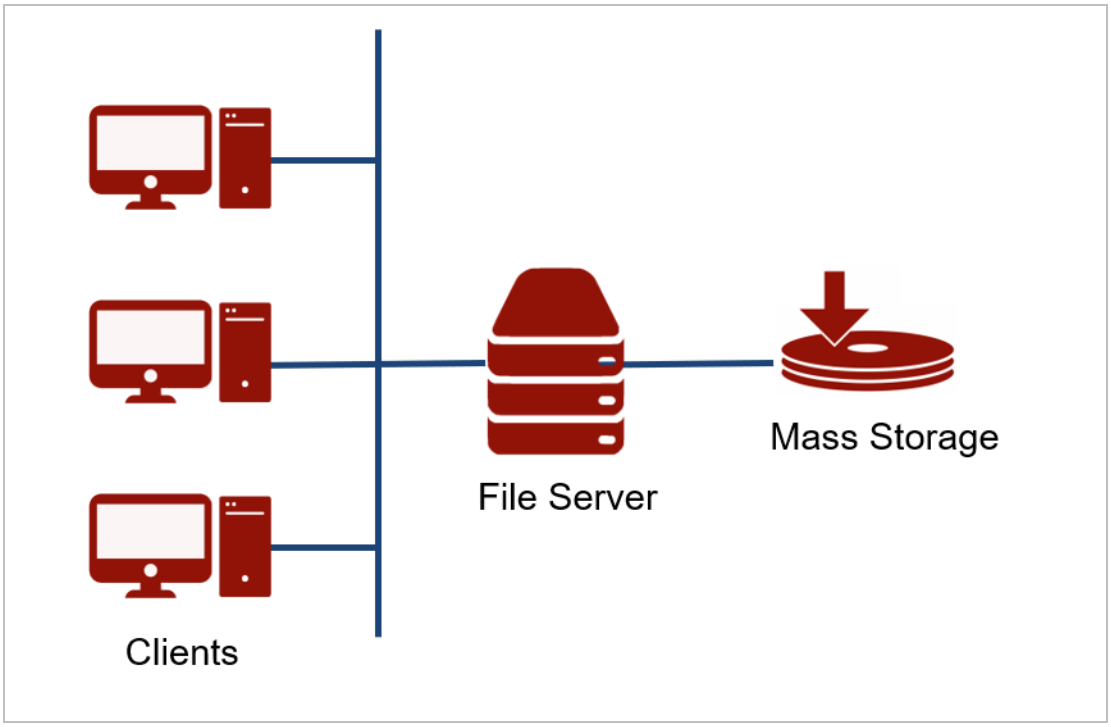

Figure 11: File Server Connected to Mass Storage

Server racks are specifically designed to hold and organize servers and storage devices. A server rack cabinet also holds equipment just as a normal server rack does, but it is enclosed on all

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

sides.[4] When organizing equipment within the racks, consider optimal placement. For example, storage systems placed near each other and near client systems make it easier to manage those systems and can reduce network latency. A practical approach is to place heavier hardware on the bottom of the rack to prevent tipping and struggling with the equipment when deploying or decommissioning the hardware.

3.3.2 Cables and Switches

Organizations strive to achieve their networking objectives with the most efficiency and value possible. *Bandwidth* measures how much data can pass through a network during a fixed period of time. Understanding the organization's bandwidth requirements is an essential part of managing a data center.

Several key factors impact your bandwidth: (1) the bandwidth capability of the servers, (2) the anticipated number of users of the system, and (3) the applications that will be running on network devices. Bandwidth estimates must also consider changing requirements and the growth plan for the data center. When estimating your bandwidth requirements, overestimating is always better than underestimating.

3.3.2.1 Cables

Cables are the plumbing of your network. Ethernet is the most common system of network cabling for small networks.[5] Among other things, it specifies what cables to use, how long a type of cable needs to be, and how to connect the cables.

The three main types of cables are (1) copper twisted-pair, (2) fiber optic, and (3) twinaxial (twinax). Organizations use all three cables in their data center deployments.

Most networks use twisted-pair cable. ANSI/EIA 568 is a standard that specifies the bandwidth of each category of cable. **Error! Reference source not found.** lists the characteristics of various ategories of copper twisted-pair cable.

[4] Server rack positioning is discussed in Section 3.2.4.

[5] You can use wireless technology to create networks without cables, but most networks use cables to physically connect devices to the network.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Table 6: Network Cable Categories, Bandwidth, and Maximum Throughput [[ANSI/TIA/EIA 2021]

Cable Category	Maximum Supported Speed	Bandwidth	Ethernet Standard	Description
CAT 5E	1000 Mbps	100 MHz	1000Base-T Ethernet	This cable is the minimum requirement for LAN networks.
CAT 6	10 Gbps	250 MHz	10GBASE-T Ethernet	This cable uses a plastic core to prevent cross-talk between twisted-pair. It also uses a fire-resistant plastic sheath.
CAT 6a	10 Gbps	500 MHz	10GBASE-T Ethernet	This cable reduces attenuation and cross-talk. It can also potentially remove the length limit.
CAT 7	10 Gbps	600 MHz	Not yet drafted	This cable uses multiple twisted pairs and shields each by its own plastic sheath.

Cable categories 1-5 are not listed in **Error! Reference source not found.** since they are not suitable for today's data centers. CAT 6a is the recommended cable for all modern Ethernet LAN networks [CompNetNotes 2021].

Once you select the cables required for your data center, determine the various lengths of cable needed (both within each rack and from the racks). Cable can be purchased in predetermined lengths or in bulk, which you must cut to size. When purchasing cable of predetermined lengths, the cable may be too short (requiring coupling) or too long (in which case, you have to tie up the excess and tuck it away). To avoid network clutter, you may want to consider purchasing cable that you can cut yourself [Chiappetta 2012] using a proper crimping and wire cutting tool. Cables are connected to devices using RJ45 connectors.

Twisted-pair cable can be purchased as shielded or unshielded. While unshielded cable is the least expensive of the two, be careful not to route the cable close to air conditioners, fluorescent light fixtures, or electric motors that have a lot of electrical interference. Your facility or regulatory requirements may dictate the type of cable that must be used in your environment.

Twinaxial cables can provide high speeds over short distances. The type of twinax cable used in these networking applications is called *SFP+ direct-attach copper* and can support network speeds of 10Gb, 40Gb, and 100Gb/s [Wikipedia 2021d].

Fiber optic cables are used to span long distances at high speeds. The transmission capacity of optical fiber cable is 26,000 times higher than that of twisted-pair cable [FSCom 2013]. Table 7 compares the characteristics of fiber optic cable versus copper twisted-pair cable.

Table 7: Characteristics of Fiber Optic Cable vs. Copper Twisted-Pair Cable

Cable Type	Speed	Bandwidth	Distance
Fiber optic cable	10/100/1000 MBPS, 10/40/100/200 GBPS	Up to 4700 MHz	Up to 80 km
Twisted-pair cable	10 GBPS	250 MHz	Up to 100 m

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

There are many options for connecting fiber optic cables. In all, approximately 100 different types of fiber optic connectors have been introduced to the market [FOA 2021]. These options do not include how the fiber optic cable connects to switches and routers. For a fiber optic cable to connect, an interface called an SFP (Small Form-factor Pluggable) transceiver is needed.

SFP performs conversions between optical and electrical signals.[6] There are variations of SFP, as listed in Table 8 [CABLExpress 2021]. Select an SFP variant based on your user requirements—specifically, (1) the amount of traffic that you anticipate your network will handle, (2) the physical length of your network, and (3) the capabilities of the device connected to the SFP. However, there are variants within variants of SFP that are not listed in Table 8. Additional listings of variants can be found in the literature [Wikipedia 2021a].

Table 8: Variations of SFP

Type	Description	Transmission Speed	Cable
SFP	Small Form-Factor Pluggable	1 Gbit	Typically, LC fiber connector
SFP+	Small Form-Factor Pluggable Plus	10 Gbit	Typically, LC fiber connector
QSFP	Quad Small Form-Factor Pluggable	40 and 100 Gbit	MPO
OSFP	Octal Small Form-Factor Pluggable	200 and 400 Gbit	MPO or LC
QSFP-DD	Quad Small Form-Factor Pluggable – Double Density	200 and 400 Gbit	MPO or LC

3.3.2.2 Switches

Devices within the network are connected by switches [Wikipedia 2021b]. A switch is a device that receives incoming packets of information from the network and determines where each packet should be sent. Extremely small networks might only require an unmanaged switch or hub (plug and play). Larger networks need the ability to monitor and configure the behavior of each switch in the network to control and restrict the flow of data between various network segments. Devices that enable this ability are referred to as *managed switches*. If your network requires more than a single switch, you should use managed switches that can support multiple virtual LANs (VLANs).

[6] In some instances, SFP is used to connect copper cable.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

3.3.3 Network Topology

Network topology is the arrangement of elements within the network. Any given node in the network has one or more physical links to other devices in the network. Graphically mapping these links results in a geometric shape that can be used to describe the physical topology of the network.

Many models have emerged for data center topology. One popular topology is a tree-based topology called three tier or three layer. The layers, illustrated in Figure 12, are (1) core, (2) distribution, and (3) access.

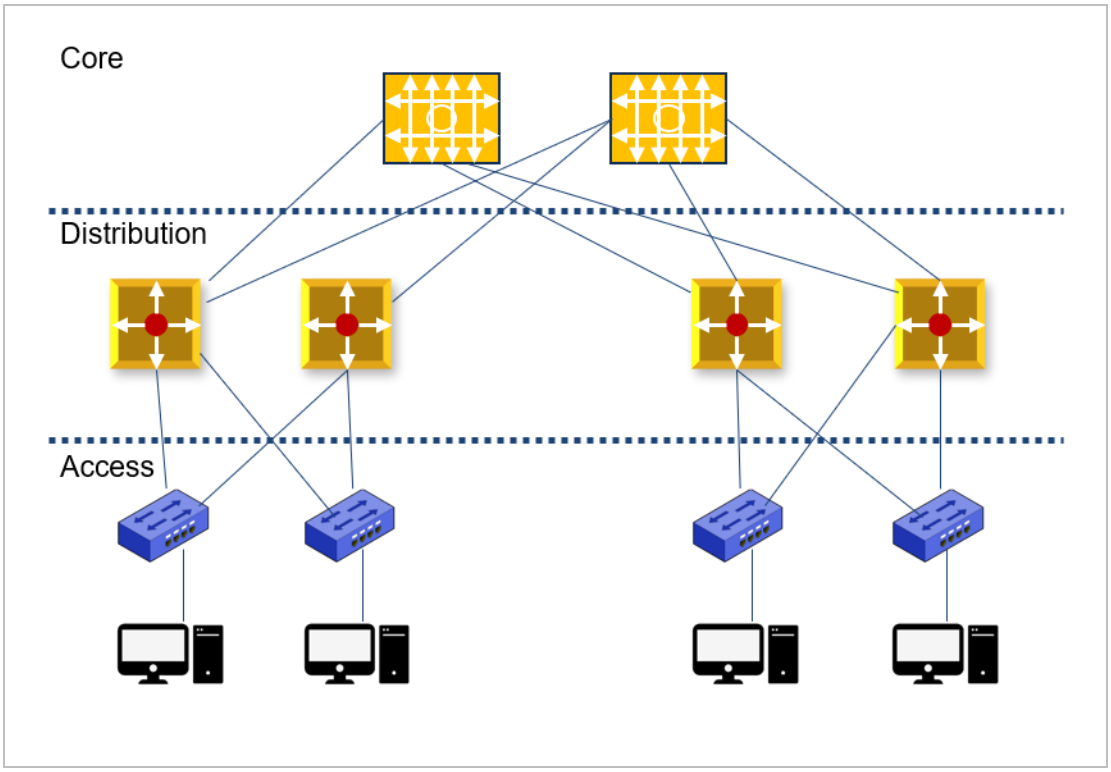

Figure 12: Three-Layer Network Topology

Core switches are large modular chassis switches with very high throughput. They are typically used to merge geographically separated networks.

The *distribution layer* defines policy for the network by ensuring that packets are properly routed between subnets in the network.

The *access layer* includes switches that are connected to end devices such as servers, computers, and printers.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Figure 13 illustrates a network topology that is catching on for data centers that experience especially heavy data transfer, such as server-to-server and storage area network data traffic. This topology, referred to as the *spine-leaf topology*, minimizes latency and bottlenecks because each payload only has to travel to a spine switch and another leaf switch to reach its endpoint [Zaharoff 2021]. These types of networks are useful in highly virtualized, multi-tenant deployments with lots of east-west traffic.

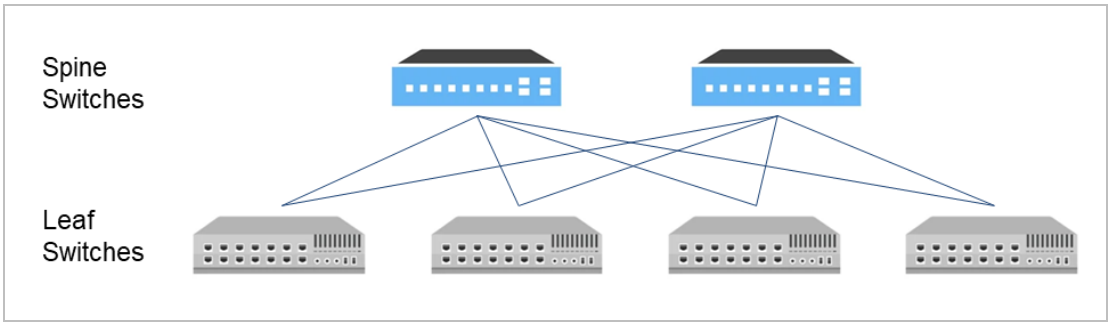

Figure 13: Spine-Leaf Network Topology

Properly planning the data center network topology is critical, and performance, resiliency, and scalability need to be carefully considered.

When selecting a network topology for your data center design, do the following:

- Strongly consider the platform you intend to use to deploy your applications. Different platforms can have varying networking preferences and can benefit from the deployment of one network topology over another.
- Consider how your network will connect to outside resources, if that is a requirement for your environment. External connections necessitate special consideration for security requirements dictated by compliance standards and/or acceptable risk levels for your organization.

3.4 Develop a Data Center Management Plan

Data center management refers to the role of individuals that are tasked within a data center to oversee technical and IT issues. Figure 14 emphasizes the processes that must be designed to realize the final major stage of your data center. These processes include managing server and computer operations, software services and applications, large amounts of data, and the security of the data. In the initial stages of data center deployment, tasks are managed manually; down the line, automation plays a large part in system management and configuration. These automation tools are needed to provide resiliency, improve uptime, and reduce risk across data center operations, including the network, power, IT equipment, applications software, and applications services.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

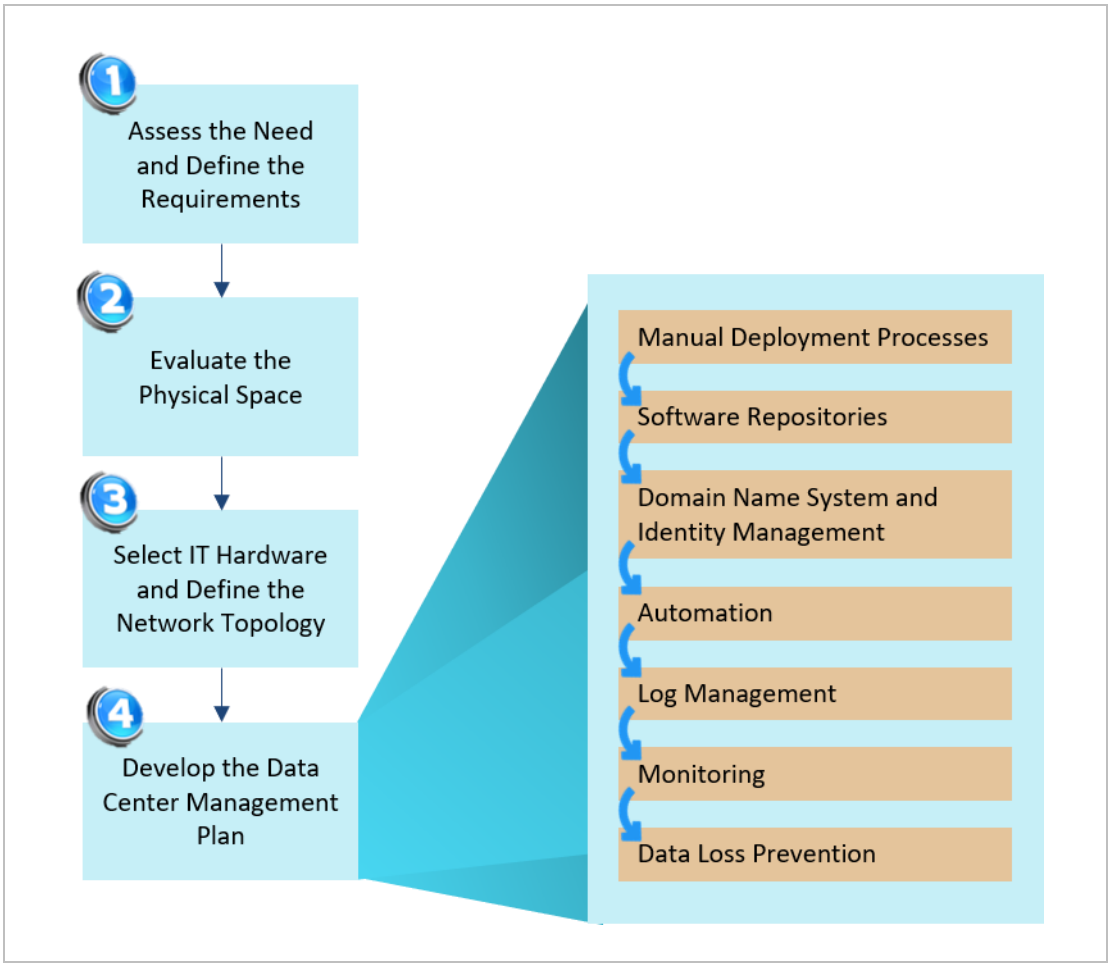

Figure 14: The Develop Data Center Management Plan Process

As with all items considered within data center design, evaluate the various vendor offerings that address your needs. Evaluate the products against your requirements, and conduct trade-off analyses to arrive at decisions that lead to your solution.

While this section describes planning and designing these features, deployment and implementation should be carried out in the sequence implied by the curved arrows in Figure 14 due to the interdependencies among the features.

3.4.1 Design the Manual Deployment Processes

In the earliest stages of deploying your computing environment, you'll likely find yourself in a room filled with new equipment. While it may be your intention to automate *all* operations immediately, automation will need to evolve over time because the requisite systems required to run automation tools have not yet been deployed.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Your data center management plan must consider a method for installing several foundational pieces: the operating systems (OSs), hypervisors, and network device images on your core systems; this method will likely be manual. Consider how and where you will get the files and/or media required to install these foundational pieces.

Before you install the operating system or hypervisor on a bare metal system, ensure that your server meets the minimum requirements of the operating system. If your design indicates a conflict, reevaluate your hardware and/or operating system selections. Carefully compare the release notes associated with the operating system software with the server specifications. Check whether there are specific warnings that may apply to your situation. Also, determine whether your compliance standards (e.g., FIPS standards) dictate installation requirements, and then make your operating system selections accordingly.[7]

Generally speaking, plan to have an Internet connection available to download operating system images and software installation files. If a connection is unavailable, you will likely install the operating systems from a vendor-supplied CD, DVD, or USB drive.

When installing the operating system manually, a set-up program leads you through the installation process. Most operating system installation media contain a basic set of software that can be used to implement basic system functions. For example, Windows Server is bundled with the software for Active Directory Domain Services. CentOS media contains the curl and wget packages.

In these early deployment stages, where tools for automation are not yet available, be prepared to take detailed notes as you execute installation processes because there can be unforeseen steps that you did not anticipate, and you might need to repeat the process manually. Keep a detailed record of all modifications made during your manual deployment processes. This record can be extremely useful once automation tools are available, because you can use it to comprehensively include all changes in your configuration management solution once it is deployed.

3.4.2 Design the Software Repositories

An important task for the data center management team is to host the various software applications required by network users and computer systems (Figure 15). The required software will likely include packages/binaries/executables for hardware components, applications, operating systems, and network devices. If your data center provides development environments, software developers will require access to libraries and development tools; system administrators will require access to management tools.

[7] FIPS is Federal Information Processing Standards.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

Figure 15: Software Repositories

You will need to have repositories of software available once the operating systems are deployed so you can begin to install and configure specific services and applications in your data center.

Consider, in particular, what software you'll need to complete the first steps in your deployment plan and which pieces are not needed until later.

The types of repositories you deploy in your environment are directly related to the operating systems you deploy. Think about how you intend to populate the repositories for the first time. If you expect to have Internet access, you may be able to download the repositories directly from the vendor into your environment. If not, you may need to use sneakernet.[8] Also plan how software applications will be kept up-to-date with the latest patches and updates from software vendors.

Part of the software application management task is managing the licenses associated with the copies of software that are deployed in the network. Familiarize yourself with your organization's licensing requirements and the license agreements associated with applications to ensure you understand any restrictions. For example, activating a Windows 10 system requires a product key or access to a key management server. Access to Red Hat content requires a valid license subscription. Seat licenses are required for various user endpoint applications.

Document all customizations implemented on software repository systems so that, in an emergency, they can be recreated and repopulated. Ultimately, these customized configurations are controlled by a configuration management system.

3.4.3 Design the Domain Name System and Identity Management Approach

The next phase of your deployment plan should address identity management (IdM). IdM is a framework of policies and technologies for ensuring the proper people in an enterprise have the appropriate access to technology resources [Wikipedia 2021c].[9] As illustrated in Figure 16, IdM

---

[8] Sneakernet is a jargon term for the method of transmitting electronic information by personally carrying it from one place to another on removable media. The idea is that someone is using their shoes (possibly sneakers) to move data around rather using the network.

[9] Identity management is also known as *identity and access management (IAM).*

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

identifies, authenticates, and authorizes individuals or groups of people to have access to applications, systems, or networks by associating user rights and restrictions with established identities. This process is typically enabled using specialized software. Therefore, as in the case with the procurement of all software tools that you need in the data center, you will need to evaluate your user requirements and budget, and then conduct research and tradeoff analysis to select the IdM solution that best meets your needs.

There are several types of access control strategies. Two popular high-level strategies are

- Role-Based Access Control (RBAC)
- Attribute-Based Access Control (ABAC)

RBAC is access control based on user roles (i.e., a collection of access authorizations a user receives based on an explicit or implicit assumption of a given role). Role permissions may be inherited through a role hierarchy and typically reflect the permissions needed to perform defined functions in an organization. A given role may apply to a single individual or several individuals [NIST 2021b].

ABAC is an access control approach in which access is mediated based on attributes associated with subjects (requesters) and the objects to be accessed. Each object and subject has a set of associated attributes, such as location, time of creation, access rights, etc. Access to an object is authorized or denied depending on whether the required (e.g., policy-defined) correlation can be made between the attributes of that object and of the requesting subject [NIST 2021c].

Similarly, you must ensure that your information management solution is compatible with your environment's components and can manage identities to meet compliance requirements.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

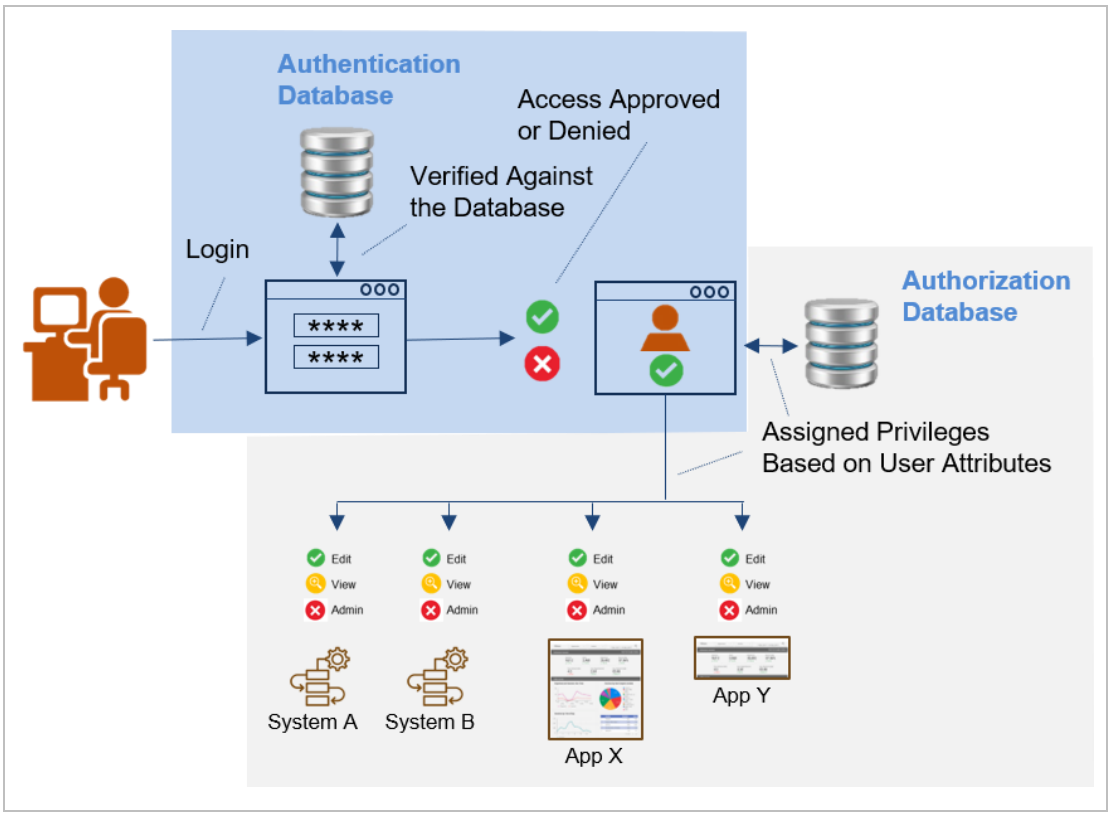

Figure 16: Identity Management Example

Security-related requirements inform the authentication method that you will use (e.g., single sign-on [SSO] versus multifactor authentication [MFA]). When designing your IdM solution, keep in mind that it should be expandable to accommodate the data center's growth plan. These security-related requirements also influence how the IdM solution is set up for authorizing users accessing resources while co-located with the data center environment in addition to authorizing users accessing resources remotely.

Ensure that the process of installing and configuring the IdM system is documented as a repeatable process.

3.4.4 Design the Automation Approaches

At this point in your data center deployment plan, you have accounted only for manual installation and configuration tasks. The approach is designed this way because there is value in prioritizing identity management over automation in the beginning stages of data center deployments. This sequence of operations allows you to control which users and groups are authorized to access certain segments of the infrastructure.

For example, when IdM is addressed before automation, you can lock down your Linux deployments to your Linux team and Windows deployments to your Windows team. You can also limit the scope of access for your novice system administrators until they gain additional experience.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

As you begin to design your approaches to automation, pay attention to two main considerations:

1. how you wish to deploy and install new systems and new software
2. how you wish to manage the configurations of the systems and software that you have deployed

Your response to these considerations is based on the platform you are using and the OS distribution you are using. For example, deployment of an OS onto bare metal is slightly different than the deployment onto a virtual machine. Likewise, software installation and configuration procedures differ based on whether you are running Windows, Linux, or MacOS.

It is possible to automate the operating system installation process to ensure consistency and reduce human error. One method for semi-automatic installation is to create a generalized image of the installed system that can be deployed on other systems. With Windows installations, this is called *Sysprep* [Microsoft 2017b]. For Linux virtual systems, the tool is called *virt-sysprep.* Another method for automated system installation is to supply a text file containing the appropriate set of system configuration parameters. For Linux, this is called a kickstart file [Red Hat 2021], and for Windows, this is called an Answer file [Microsoft 2017a].

A Preboot Execution Environment (PXE) is a client-server interface that allows computers in a network to be booted from the server before deploying the operating system to other servers that are part of the system. Other tools are available to aid in the automated deployment of virtual machines into a virtual infrastructure. It is beyond the scope of this report to describe all steps involved in setting up and configuring servers. However, significant resources are available in the documentation that accompanies your operating system software product. These tools can also be used to ensure that the desired software is installed on systems during initial deployment.

No matter how large or small your network environment, change is inevitable. Although change is almost always a good thing in the end, bad things can happen. One of the greatest responsibilities of data center administrators is to make sure that the computing environment is functioning properly at any given moment, particularly if the environment has a high criticality as a design driver.

Network configuration management is a process for maintaining servers, networking devices, computer systems, and software in a desired consistent state. To achieve consistency, configuration management ensures that both small and large changes to the system are documented and that a change control process is in place to support stability and uptime. The change management database contains locations and IP addresses or network addresses of all hardware devices as well as data about the default settings, programs, versions, and updates installed on network-connected computers.

In practice, configuration management involves the following tasks:

1. Maintain documentation that describes the current configuration of all network devices.
2. Maintain documentation that describes the rationale and details of any changes.
3. Maintain an archive of older configurations that can be deployed in an emergency.
4. Implement policies that control who may perform changes.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

A primary feature of network configuration management is the ability to replace functions of a network device in the event of failure. Configuration management tools make system changes and deployments faster due to automation while also removing the potential for human error during deployments. If a deployment is unsuccessful, the changes can be rolled back to a previous configuration. If a server goes down for some unknown reason, a new one with exactly the same functionality can be quickly deployed. When the data center is used to develop software applications, configuration management can be used to ensure that the test and production environments match, thus avoiding questions about which environment is the root cause of a problem.

The need for configuration management comes into focus when multiple system administrators make changes to data center equipment. How does person A know what person B has been up to? Even if person A and person B communicate well, the chance that critical details are left out is significant.

Many tools are available for both automated system deployments and network configuration management (both open source and vendor provided). A single tool might meet all your deployment and operational automation needs, or several tools working with one another might provide the best solution.

At this point, you can begin deploying software to hosts using automation, and the manual installation processes that you used earlier will no longer be necessary.

3.4.5 Design the Log Management Approach

Administrators of a data center must rely on collecting information about the status, health, and actions of the network infrastructure and software applications. *Logging* refers to the process of collecting data from various sources in the data center. The data is information about events, which are the discrete pieces of information that inform administrators about what is happening on the network, on a server, or with specific applications.

Most networks, applications, and servers have base level and additional logging capabilities that might need to be tailored to show more or less information, depending on your needs. These logging configurations are in the product's documentation. A log event typically includes the time, occurrence, and details that explicitly relate to the event in the environment that may help explain or understand the event's cause or effects. In some cases, government regulations require certain types of information to be captured as part of a log event and specifies how long log data must be retained [Kent 2006].

Security logging makes it possible to see and trace nefarious actors causing problems (referred to as incidents) on the network. For this reason, log data is typically accessible by data center administrators only. The security and privacy of data logs must be maintained because log data can sometimes contain personal data, highly sensitive information, or data that is subject to security regulations.

Log storage should be centralized (Figure 17); centralized logs reduce the complexity and risk of maintaining policies in multiple locations. The centralized logging system should be set up early during data center deployment. This centralization allows you to baseline your system early and observe how the system evolves over time.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

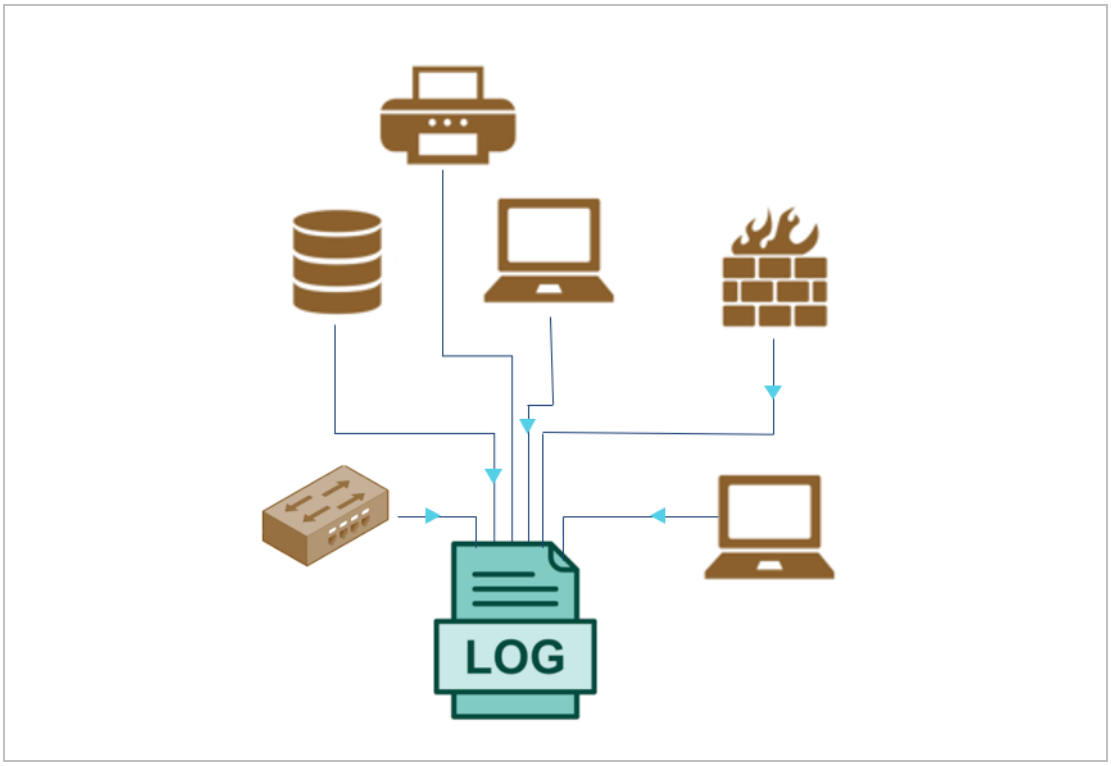

Figure 17: Centralized Logging System

Log data must be succinct but must also provide the complete story. Data should be selective and provide the correct context to facilitate troubleshooting. There are many software tools that can help you monitor log files and/or perform automated analysis.

3.4.6 Determine Data Center Monitoring Needs and the Approach

While logging tells you *what* happened in your network (and provides the raw data to track down an issue), monitoring tells you *how* an application is performing and can alert you to issues. Data center monitoring applications can aggregate and analyze network data. They can provide visualization dashboards to improve your understanding of how your system is behaving and provide alerts to notify system administrators when a measured parameter falls outside of its normal operational profile.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

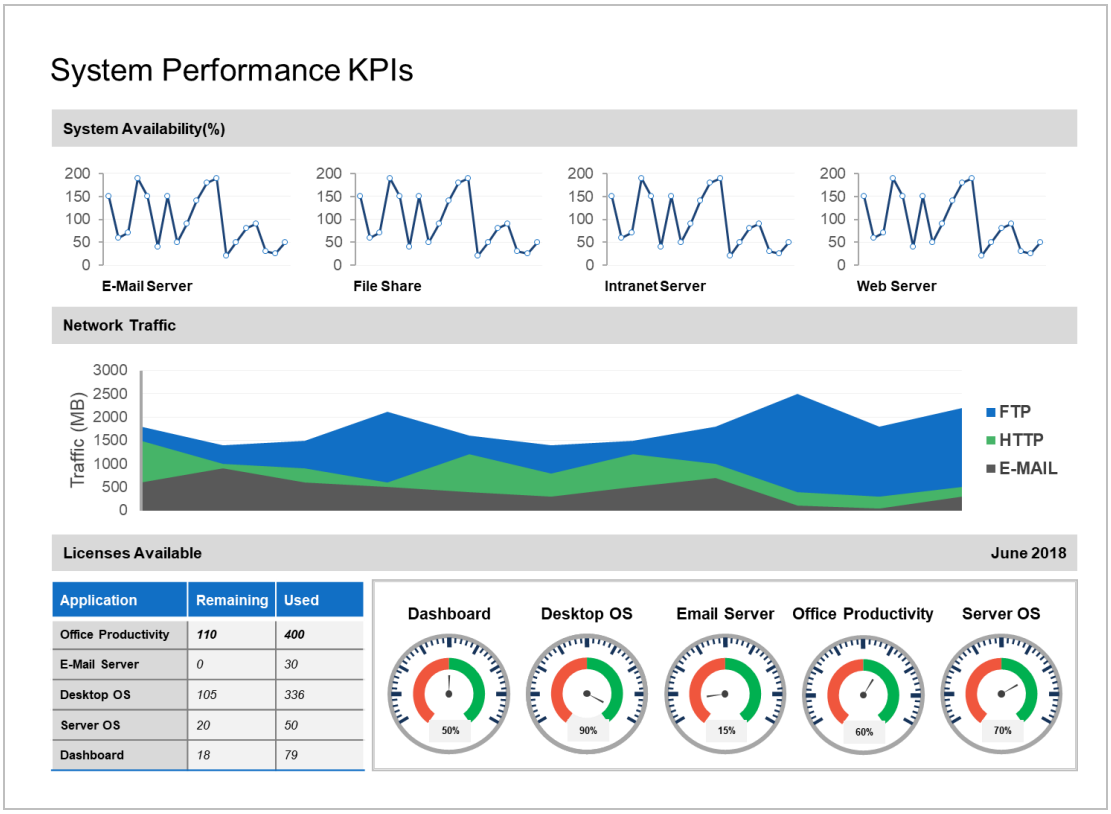

Figure 18: Example of a Visualization Dashboard for Tracking Network Performance

Figure 18 illustrates an example of a network performance dashboard. Dashboards such as this depict large amounts of information in a graphical format. They can be considered an at-a-glance version of a text file of log data. Some monitoring systems can even provide automated responses based on triggers assigned by the administrator.

3.4.7 Design the Data Loss Prevention Approach

The data center administrator is primarily responsible for the safety of the data on the network. Design your data loss prevention approach based on your disaster recovery requirements. (See Table 2.)

Backups are duplicates of important data necessary for the organization's operations. Without backups, a hard drive failure can set your organization back for days or weeks while the lost data is reconstructed. Have a solid plan for backing up your data. Your backup approach is a cornerstone of your disaster recovery plan. For example, if a water-based fire suppression system destroys your equipment, you will must know how to restore your backed-up data to replacement equipment.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

A popular approach to data backup is the 3-2-1 rule, which was developed in 2012 by the United States Computer Emergency Readiness Team (US-CERT) [Ruggiero 2012]. As illustrated in Figure 19, the 3-2-1 rule has three conditions:

1. Keep **three** copies of each important file that is important to the organization (i.e., the original plus two backups).
2. Keep the copies on **two** different types of media. (If all copies are of the same media, then all three copies are subject to the same risks.)
3. Store **one** copy outside the data center. (If all copies are stored together, a calamity such as a fire can destroy all three copies.)

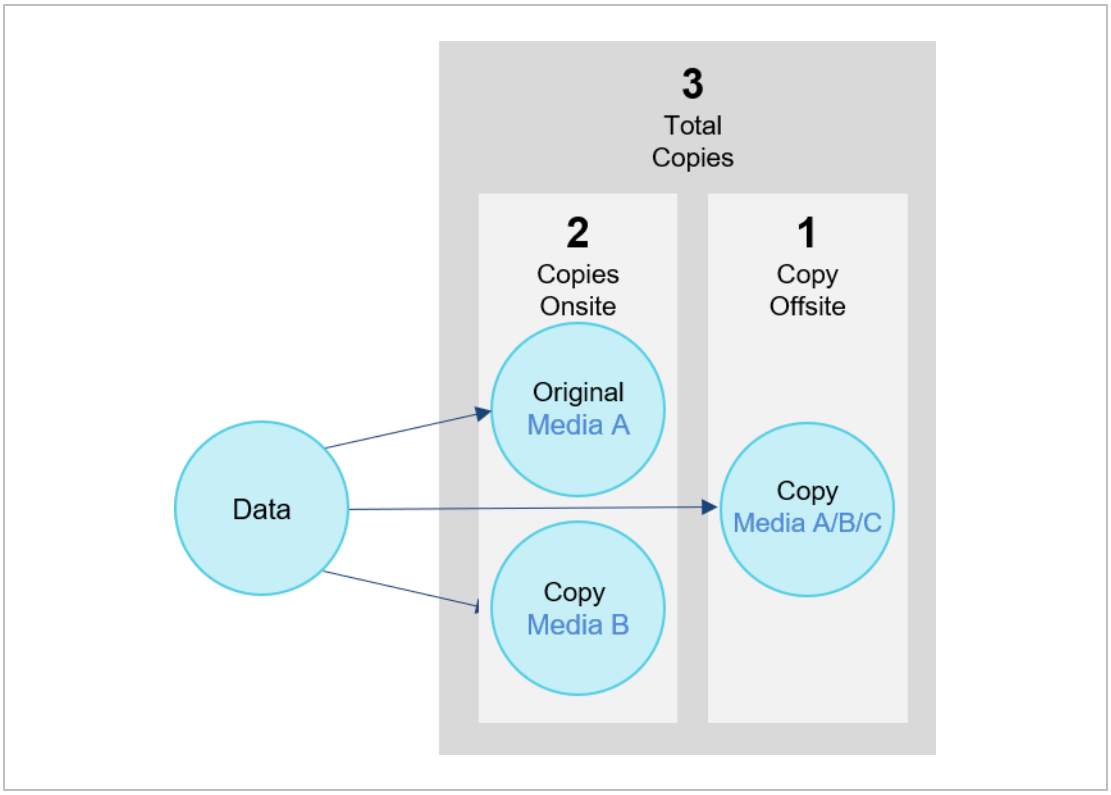

Figure 19: The 3-2-1 Backup Rule

With regard to local storage, Ruggiero and Heckathorn note that one advantage of storing backups on your network drive is being able to "quickly update backup files and maintain a simple file structure" [Ruggiero 2012].

Another factor to consider in your backup plan is the frequency of how often to back up your data. The best way to answer this question is to consider the loss that your organization can afford in the event of a data loss. If your organization can get by if it loses one week's worth of data, then your backup plan should reflect performing a backup every week. If the loss of one day's worth of data is intolerable, then you would need to back up every day.

There is a vast array of vendor software for supporting the automation of network backups. As always, consider your requirements and budget while conducting tradeoff analyses of the various candidates that will serve your needs the best.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

4 Summary

Developing a data center is not a simple task for an organization, and it can be a complex and costly endeavor. When an organization begins to consider such a project, much of the work should take place before the data center is deployed. The Plan and Design phase pays dividends before you engage with vendors or embark on soliciting bids to contractors. Without proper planning and design, organizations risk inefficiencies and rework leading to a result not optimized for efficiency.

The initial decision to build a data center begins with identifying its purpose of use (e.g., why an on-premises environment versus a cloud-based environment), design drivers (e.g., understanding constraints, criticality, and growth plan) and sound requirements and estimates, which include considerations of risks, mitigations, and impact.

This report provides IT personnel and management teams with a basic understanding of the practical planning and design considerations for an on-premises computing environment. This information and other supporting papers and websites (provided in the References section) provide readers a sense of the effort and basic elements required to set up a data center facility reliably and maintain it efficiently over its life span.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

References

URLs are valid as of the publication date of this document.

[AICPA 2016]
American Institute of CPAs (AICPA). Statement on Standards for Attestation Engagements No. 18. April 2016. *AICPA Concepts Common to All Attestation Engagements*. https://www.aicpa.org/research/standards/auditattest/ssae.html

[AICPA 2021]
American Institute of CPAs (AICPA). Report on Controls at a Service Organization Relevant to Security, Availability, Process Integrity, Confidentiality or Privacy. *SOC2 – SOC for Service Organizations: Trust Services Criteria*. June 3, 2021 [accessed]. https://www.aicpa.org/interestareas/frc/assuranceadvisoryservices/aicpasoc2report.html

[ANSI/TIA/EIA 2021]
American National Standards Institute/Telecommunications Industry Association/Electronic Industry Alliance (ANSI/TIA/EIA.) ANSI/TIA/EIA 568-B: Commercial Building Telecommunications Cabling Standard. June 3, 2021 [accessed]. https://www.csd.uoc.gr/~hy435/material/Cabling%20Standard%20-%20ANSI-TIA-EIA%20568%20B%20-%20Commercial%20Building%20Telecommunications%20Cabling%20Standard.pdf

[ASHRAE 2016]
ASHRAE Technical Committee (TC) 9.9 Mission Critical Facilities, Data Center, Technology Spaces, and Electronic Equipment. Data Center Power Equipment Thermal Guidelines and Best Practices. 2016. http://tc0909.ashraetcs.org/documents/ASHRAE_TC0909_Power_White_Paper_22_June_2016_REVISED.pdf

[ASPE 1996]
Office of the Assistant Secretary for Planning and Evaluation (ASPE). Health Insurance Portability and Accountability Act of 1996. *U.S. Department of Health & Human Services.* August 21, 1996. https://aspe.hhs.gov/report/health-insurance-portability-and-accountability-act-1996

[CABLExpress 2021]
CABLExpress. Tips for Determining Transceiver and Fiber Cable Selection. *CABLExpress.* June 3, 2021 [accessed]. https://www.cablexpress.com/education/blog/tips-for-determining-transceiver-and-fiber-cable-selection/

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

[Chiappetta 2012]
Chiappetta, Marco. Make Your Own Ethernet Cables. *PC World*. August 2, 2012. https://www.pcworld.co.nz/article/481582/make_your_own_ethernet_cables/

[CompNetNotes 2021]
Computer Networking Notes. Network Cable Types and Specifications. *Computer Networking Notes*. June 3, 2021 [accessed]. https://www.computernetworkingnotes.com/networking-tutorials/network-cable-types-and-specifications.html

[DoD 2021]
Department of Defense (DoD) Cyber Exchange: Public. Security Technical Implementation Guides (STIGs). *DoD Cyber Exchange: Public.* June 3, 2021 [accessed]. https://public.cyber.mil/stigs/

[Evans 2004]
Evans, Tony. The Different Types of Air Conditioning Equipment for IT Environments. *American Power Conversion White Paper 58.* 2004. https://www.apcdistributors.com/white-papers/Cooling/WP-59%20The%20Different%20Types%20of%20Air%20Conditioning%20Equipment%20for%20IT%20Environments.pdf

[FOA 2021]
The Fiber Optic Association, Inc. (FOA). Guide to Fiber Optics & Premises Cabling. *The Fiber Optic Association, Inc. – Tech Topics*. June 3, 2021 [accessed]. https://www.thefoa.org/tech/connID.htm

[FSCom 2013]
FS Community (FSCom). Fiber Optic Cable vs Twisted Pair Cable vs Coaxial Cable. *FS Community Blog*. March 25, 2013. https://community.fs.com/blog/the-difference-between-fiber-optic-cable-twisted-pair-and-cable.html

[Ingram 2020]
Ingram, Jonathon & Nadgir, Abhay. Data Center Fire Suppression: An Overview of Why and How. *Mission Critical Magazine.* July 14, 2020. https://www.missioncriticalmagazine.com/articles/93098-data-center-fire-suppression

[ISO/IEC 2013]
International Organization for Standardization/International Electrotechnical Commission (ISO/IEC). ISO/IEC 27001:2013 Information Technology – Security Techniques – Information Security Management Systems – Requirements. 2013. https://www.iso.org/standard/54534.html

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

[Kent 2006]
Kent, Karen & Souppaya, Murugiah. Guide to Computer Security Log Management: Recommendations of the National Institute of Standards and Technology (NIST Special Publication 800-92). September 2006. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-92.pdf

[Microsoft 2017a]
Microsoft. Answer Files (unattend.xml). May 2017. https://docs.microsoft.com/en-us/windows-hardware/manufacture/desktop/update-windows-settings-and-scripts-create-your-own-answer-file-sxs

[Microsoft 2017b]
Microsoft. Sysprep (System Preparation) Overview. May 2017. https://docs.microsoft.com/en-us/windows-hardware/manufacture/desktop/sysprep--system-preparation--overview

[Morrow 2018]
Morrow, Timothy. 12 Risks, Threats, & Vulnerabilities in Moving to the Cloud [blog post]. *Software Engineering Institute (SEI) Blog*. March 5, 2018. https://insights.sei.cmu.edu/blog/12-risks-threats-vulnerabilities-in-moving-to-the-cloud/

[Morrow 2019a]
Morrow, Timothy; Pender, Kelwyn; Lee, Carrie; & Faatz, Donald. *Overview of Risks, Threats, and Vulnerabilities Faced in Moving to the Cloud*. CMU/SEI-2019-TR-004. Software Engineering Institute, Carnegie Mellon University. 2019. http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=551354

[Morrow 2019b]
Morrow, Timothy; LaPiana, Vincent; Faatz, Donald; & Hueca, Angel. *Cloud Security Best Practices Derived from Mission Thread Analysis*. CMU/SEI-2019-TR-003. Software Engineering Institute, Carnegie Mellon University. 2019. http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=551466

[NIST 2018]
National Institute of Standards and Technology (NIST) Computer Security Resource Center (CSRC). Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy (SP 800-37 Rev.2). *The National Institute of Standards and Technology.* December 2018. https://csrc.nist.gov/publications/detail/sp/800-37/rev-2/final

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

[NIST 2020]
National Institute of Standards and Technology (NIST) Computer Security Resource Center (CSRC). Security and Privacy Controls for Information Systems and Organizations (SP 800-53 Rev. 5). *The National Institute of Standards and Technology.* September 2020. https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final

[NIST 2021a]
National Institute of Standards and Technology (NIST) Computer Security Resource Center (CSRC). Protecting Controlled Unclassified Information in Nonfederal Systems and Organizations (SP800-171 Rev. 2). *The National Institute of Standards and Technology.* February 2021. https://csrc.nist.gov/publications/detail/sp/800-171/rev-2/final

[NIST 2021b]
National Institute of Standards and Technology (NIST) Computer Security Resource Center (CSRC). Role-Based Access Control (RBAC). *The National Institute of Standards and Technology*. June 3, 2021 [accessed]. https://csrc.nist.gov/glossary/term/role_based_access_control

[NIST 2021c]
National Institute of Standards and Technology (NIST) Computer Security Resource Center (CSRC). Attribute-Based Access Control (ABAC). *The National Institute of Standards and Technology*. June 3, 2021 [accessed]. https://csrc.nist.gov/glossary/term/attribute_based_access_control

[Ready 2021]
Ready.gov. IT Disaster Recovery Plan. *Ready.gov.* June 3, 2021 [accessed]. https://www.ready.gov/it-disaster-recovery-plan

[Red Hat 2021]
Red Hat. Red Hat Enterprise Linux 8: Performing an Advanced RHEL Installation—Installing Red Hat Enterprise Linux 8 using Kickstart. June 2021. https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/8/html/performing_an_advanced_rhel_installation/index

[Ruggiero 2012]
Ruggiero, Paul & Heckathorn, Matthew A. Data Backup Options. *United States Computer Emergency Readiness Team, Carnegie Mellon University*. 2012. https://us-cert.cisa.gov/sites/default/files/publications/data_backup_options.pdf

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

[Wikipedia 2021a]
Wikipedia. Small Form-factor Pluggable Transceiver. *Wikipedia*. June 3, 2021 [accessed]. https://en.wikipedia.org/wiki/Small_form-factor_pluggable_transceiver

[Wikipedia 2021b]
Wikipedia. Network Switch. *Wikipedia.* June 3, 2021 [accessed]. https://en.wikipedia.org/wiki/Network_switch

[Wikipedia 2021c]
Wikipedia. Identity Management. *Wikipedia*. June 3, 2021 [accessed]. https://en.wikipedia.org/wiki/Identity_management

[Wikipedia 2021d]
Wikipedia. Twinaxial Cabling. *Wikipedia.* June 3, 2021 [accessed]. https://en.wikipedia.org/wiki/Twinaxial_cabling

[Zaharoff 2021]
Zaharoff, Sharon. Leaf-Spine (Leaf-Spine Architecture). *TechTarget.* June 3, 2021 [accessed]. https://searchdatacenter.techtarget.com/definition/Leaf-spine

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
(Leave Blank)	July 2021	Final

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Planning and Design Considerations for On-Premises Computing Environments	FA8702-15-D-0002

6. AUTHOR(S)

Lyndsi Hughes, David Sweeney, & Mark Kasunic

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213	CMU/SEI-2021-TN-002

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
SEI Administrative Agent AFLCMC/AZS 5 Eglin Street Hanscom AFB, MA 01731-2100	n/a

11. SUPPLEMENTARY NOTES

12A DISTRIBUTION/AVAILABILITY STATEMENT	12B DISTRIBUTION CODE
Unclassified/Unlimited, DTIC, NTIS	

13. ABSTRACT (MAXIMUM 200 WORDS)

Computing environments that provide access to acutely sensitive data often have a business requirement to restrict unauthorized access to that data. In these cases, hosting an on-premises data center can be the preferred method of providing necessary computing resources to end users to achieve this business requirement. In other instances, cost or special use cases of the system are drivers for building an on-premises data center over using a cloud-based approach. Security concerns can likewise lead to such a decision.

This report shares important lessons learned from establishing small- to mid-size data centers. These data centers were established within their own organization and for client organizations within the United States government to support development and operations. Their current focus is to establish on-premises data centers that support modern DevSecOps practices and enabling technologies.

This report is intended to help information technology (IT) personnel and management who are responsible for designing and deploying data center technology to become familiar with topics that must be addressed for a successful outcome. While it is beyond the scope of the report to delve into all the details associated with implementing data center operations, it will help IT personnel and management get started.

14. SUBJECT TERMS	15. NUMBER OF PAGES
data centers, organization-based data center, on-premises data center, core components, support infrastructure, data center management plan	52

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. Z39-18
298-102

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution.