Calculating Space and Power Density Requirements for Data Centers

White Paper 155

Revision 0

by Neil Rasmussen

Executive summary

The historic method of specifying data center power density using a single number of watts per square foot (or watts per square meter) is an unfortunate practice that has caused needless confusion as well as waste of energy and money. This paper demonstrates how the typical methods used to select and specify power density are flawed, and provides an improved approach for establishing space requirements, including recommended density specifications for typical situations.

Introduction

Two key design parameters for a data center are the IT load rating in kW and the physical size of the IT and equipment rooms. In principle, these are related by the concept of power density, which vaguely relates the building size to the IT load. Historically, it has been common to describe and specify data centers with phrases like “2,000 square meters at 1,000 watts per square meter”. Terminology like this results in needless confusion and ambiguity in the data center specification. Furthermore, this historic approach often results in underutilized power and cooling equipment, which leads to reduced electrical efficiency and excessive first-time costs.

This paper describes an improved method for specifying power density. Furthermore, this paper makes specific recommendations for power density for new data centers, based on a few simple data center characteristics.

Why the past approach does not work

There are four major problems with the historic practice of describing power density in terms of watts per square foot or watts per square meter.

1. What is included in the area calculation, or how it relates to the number of IT cabinets or devices is not defined.
2. What is included in the power calculation is not defined.
3. It provides no information about the variation in power across a population of IT cabinets; is it a peak number? An average over area? An average over time? Or some other value?
4. It is not clear how this number is used in a data center that has a changing growth plan or is modular or built out over time.

In principle, the first two problems could be improved by establishing standard definitions for power and area. However, the third and fourth problems are very important and cannot be solved by improving current definitions. **A better approach to specifying density considers the reality that IT power varies among cabinets as well as over time, and comprehends the issues of modularity and growth.**

To illustrate the problems with a vague density specification, consider the case of a data center that has the typical specification of 120 W/ft^2 (1,292 W/m^2). To understand what this means for a particular server cabinet, this density specification must be translated to the cabinet level, where, depending on assumptions (like space consumed per cabinet), it equates to somewhere between 3 and 5 kW per cabinet. The middle of this range, or 4 kW per cabinet, may seem reasonable as it is a typical power density measured in existing data centers today. There are some significant undefined variables, however, including:

- If the data center is built to 4 kW per cabinet, what happens when an isolated cabinet has a 6 kW, 12 kW, or 20 kW load?
- If some cabinets that have less than the 4 kW load installed, is the under-used power and cooling capacity available at other cabinets? If so, at which cabinets?
- If some cabinets are greater than 4 kW, do I need to leave unused space around them?
- If some cabinets are greater than 4 kW, can they be located in close proximity to each other or must they be spread out?

With the increasing functionality of server power management features which enables workloads to vary with time, a vague power density specification can have even greater implication. An effective data center density specification must be able to answer the questions posed above.

At first glance, one must question why this cannot be solved simply by specifying a very large power density for a data center, such as 30 kW per cabinet or 1000 W/ft^2 (10,764 W/m^2). It is true that such an "overkill" approach would eliminate most of the problems that have just been described. However, this creates new problems that are very expensive and, of course, wasteful, including:

- A 1000 W/ft^2 (10,764 W/m^2) data center costs about 8 times the cost of a 100 W/ft^2 data center (per unit of floor area). So if all of that density capability is not used there will be a massive waste of capital investment.
- If a 1000 W/ft^2 (10,764 W/m^2) data center actually ends up operating at 100 W/ft^2 (1,076 W/m^2) (3 kW/cabinet), then its operating PUE value is likely to be in the range of 3-5, which reflects a tremendous waste of energy.
- If a 1000 W/ft^2 (10,764 W/m^2) data center is actually populated with IT equipment at lower density, the data center will run out of physical space before it runs out of power and cooling capacity, so much of the capacity of the data center may be stranded or unusable.

These problems can be summarized with the following statement:

Specify too low a density and performance becomes unpredictable with various overload and overheating problems occurring; specify too high a power density and first cost and operating expenses are needlessly increased.

To solve this planning problem, a better way to specify density is needed. It is also necessary to provide guidance on how to choose the most appropriate density specification for a given situation, *even in situations where future density is uncertain.*

The cost of density specification errors

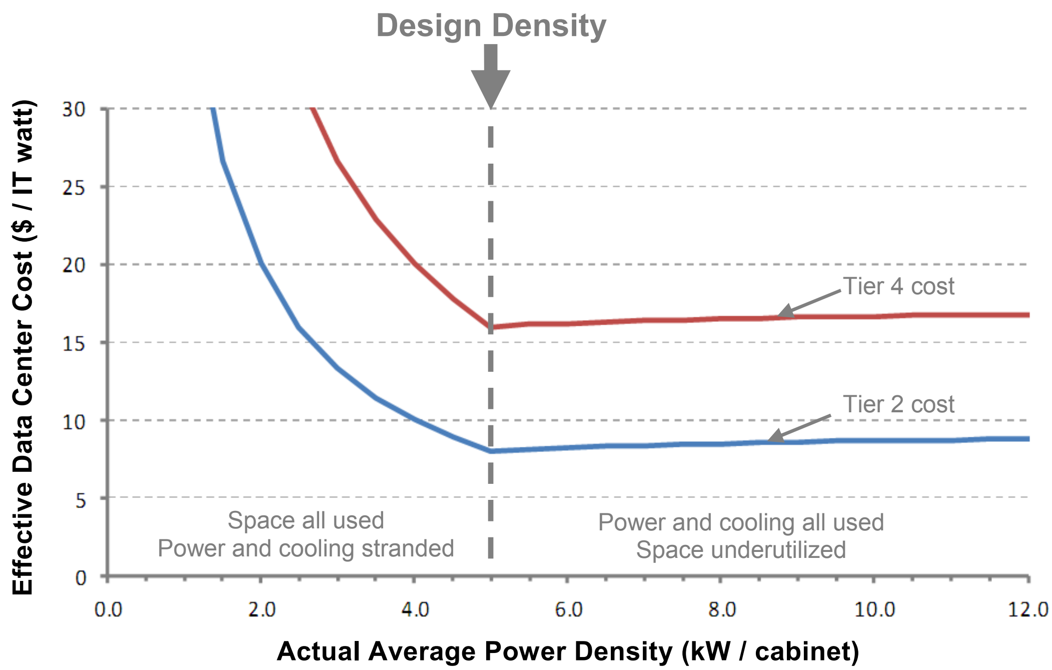
Every data center has a design target average density. A data center also has an IT load power capacity and a cooling capacity (which ideally should be the same), and it has an IT space capacity (cabinets or square feet). The ratio of the watt capacity to the space is the design density (for this discussion, we assume the power, cooling, and space are all built out to maximum design capacity). An example of an overall target design density is 5 kW per cabinet, roughly equivalent to 160 W/ft^2 (1,722 W/m^2).

If the IT equipment is deployed in a way that completely utilizes all of the power, cooling, and space in a data center, then no infrastructure is underutilized. This is the ideal case of 100% utilization. However, this ideal is nearly impossible to achieve, because the actual power use of individual IT cabinets is generally not known in advance. If the actual operating density of a data center and the design value for density do not match, some resource, either power, cooling, or space, cannot be fully utilized. **This wasted infrastructure effectively increases the cost of the data center since the resource is paid for, but not used.**

If the actual density of IT equipment is lower than the design value, then space will be used up before power and cooling are used up; that is, all of the power and cooling capacity cannot be used. If the actual density of IT equipment is higher than the design value, then the power and cooling will be used up before space is used up; so some of the space will go unused. To understand the economic effect of the error between actual density and design density for a data center, a model of the cost of the unusable, or stranded, infrastructure capacity on the overall effective cost of the data center (capex $ per *usable* IT watt) was created. Such an analysis provides the data shown in **Figure 1**.

Figure 1

Variation of effective cost as the actual power density varies from the design density of 5 kW / cabinet



The figure shows the effective data center capital cost, in $ per usable watt, as a function of the actual deployed power density in a data center. The lower curve represents a typical lower cost Tier 2 design data center, and the upper curve is a Tier 4 design. This is a model for a data center with a 5 kW per cabinet design density (approximately 160 W/ft^2 or 1,722 W/m^2).

Figure 1 shows that the effective cost per watt is lowest when the actual average power density is equal to the design power density. As the actual density falls below the design value, the effective cost rises rapidly. This is because the data center is constrained by space causing power and cooling capacity to be stranded and not utilizable. This unusable capacity has a cost, which must be assigned to a smaller IT load than originally planned, increasing the cost per usable watt of IT load. As the actual density rises above the design value, the effective cost rises slightly because of the cost of space that cannot be used because all of the power and cooling is used up. This leads to an important finding:

It is much more costly to deploy IT *below* the data center design density, than to deploy *above* the design density.

This is true because **the cost of space per unit of IT is always much lower than the cost of power and cooling**[1] **per unit of IT**. Given the fact that the actual density of IT equipment in a data center is difficult to predict in advance, this leads us to a critical key conclusion:

When the density of IT equipment is uncertain, a data center should always be constructed for a design density *less* than the mean expected value of IT density.

[1] This holds for typical power densities deployed today. Space costs become comparable to power and cooling costs only for densities well below 1 kW per cabinet.

In this way, the data center operator avoids the steep penalty curve on the left side of **Figure 1**. This leads us to another key result, which helps resolve a long-standing source of misunderstanding between data center operators and organizational management, namely:

A well-designed data center, when filled to power and cooling capacity, is *expected* to have spare or unutilized IT space.

This is a surprising result that at first appears counter-intuitive. The spare space is provided to deal with current or future uncertainty regarding IT equipment density. This unused IT space is insurance to prevent the expensive stranding of power and cooling capacity that might result if the actual density falls below the planned design value. **The more uncertainty there is about future density, the greater the spare IT space is needed.**

A new approach

The new approach to the specification of space requirements and power density has four key features:

- The unit of physical space in the density specification is the IT cabinet, NOT floor area. Floor area is determined during the design *as an output of the process* using per cabinet power and other factors.
- The specification is hierarchical and modular, so that different rooms and zones can have different density requirements.
- The specification comprehends that IT cabinets within data centers have different power requirements, and that these requirements may not be well-defined in advance.
- The specification comprehends that IT equipment cabinets may have power requirements that vary with time.

Before introducing the procedure for documenting a density specification, each of the above features is explained in turn.

Use of cabinet as a measure of physical space

The most common measure of deployment of IT space is the IT cabinet. There are other types of devices such as storage arrays and mainframes that are not cabinets, but they can, in most cases, easily be described as roughly equivalent to one or more cabinets in size. Therefore, we establish the cabinet as a measure of deployment of IT space, and use power per cabinet as the standard measure of density.

Unfortunately, space in buildings is measured in floor area (square feet or square meters) and is not readily expressed in “cabinets”. At some point we must be able to convert cabinet space into physical space. The problem is that this conversion from cabinets to floor area is dependent on a number of key factors that are independent of the per-cabinet power and must be specified separately and explicitly, including:

- Amount of cabinet locations reserved, if any, for future staging or migration
- Space required for power and cooling infrastructure devices inside the IT room, such as PDUs, air conditioners, and UPSs
- Space required for egress, raised floor ramps, and columns (if any)
- Presence of physical partitions, such as cages, subdividing the IT area
- Amount of space reserved, if any, for patch panels or network equipment
- Amount of space reserved, if any, for equipment storage cabinets or cages

- Amount of space reserved, if any, to allow for the actual density being lower than the design specification

These factors have a major impact on the space required and the density per unit area, **AND MUST BE EXPLICITY DEFINED IN THE DESIGN ALONG WITH THE DENSITY PER CABINET**. While determining these characteristics appears much more complicated than simply setting a "watt per square foot" specification, it will be shown that these factors can be determined very quickly while greatly improving the quality and clarity of the specification of the data center. How this is done will be shown later in this paper.

What is a data center "pod"?

A data center pod is a cluster of IT cabinets combined with power and cooling infrastructure that is deployed as a unit. Rooms are planned in advance for a number of pods, but the pods may be separately deployed or upgraded over time. Pods typically assembled on-site in a room to a standard design, but may be partially or extensively pre-manufactured.

In its most common form, a pod is a pair of rows of cabinets sharing a hot aisle.

Pod-based design is a recommended best-practice for larger data centers.

A modular, hierarchical density specification

We require the ability to specify density differently for different parts of a data center. In the general case, a data center can be viewed with the following hierarchy[2]:

Data center facility, comprised of one or more units of

IT rooms, comprised of one or more units of

IT pods[3], comprised of one or more units of

IT cabinets

Since t**here are attributes of data centers that are affected by the density specification at each of these levels, all four levels must be specified in order to control the design and predict its performance.** For example, the power density of a pod affects the ratings of the power sub-feeds to the pod as well as the airflow distribution system to the pod.

While it will be possible to "roll up" the density specifications of the parts to a single, facility-wide, density value, the single density number obtained this way, such as watts per square foot, **will not be enough information to control the design and achieve a predictable result.** Later in this paper a tabular method is provided for creating the hierarchical density specification.

Variation of density across the data center

Density can vary among a group of IT cabinets, across a number of pods, or across different rooms. This variation occurs physically, from cabinet to cabinet or room to room, but the variation also occurs over time because IT devices are added and removed over time, and also because the power drawn by IT devices varies with IT workload. The operating density of every cabinet can be different, and every cabinet's density can vary from moment to moment. Given this, it almost sounds futile to try to define density. However, when we consider that the reason we are defining density is to specify a data center design that will support a population of IT equipment, we can identify a number of key statistical parameters of that population that are sufficient to form a density specification that handles density variation.

For a given level in the hierarchy of the data center, such as a room comprised of pod "units", or a pod comprised of cabinet "units", density can be specified with five key parameters as shown in **Table 1**.

[2] Note that this is a general hierarchy and for smaller data centers, not all may apply. For example, in a small server room the "facility" may only comprise a single room with a single pod, so the density specification becomes simple.

[3] See sidebar for a definition of a data center pod. Pods are sometimes referred to as zones, clusters, or rows; these alternate terms have various meanings in data center design and are not preferred. Small data centers may omit this level and simply deploy IT cabinets in rooms.

Note that floor area is not a key input parameter for the density specification. Floor area is a *computed output* given these specification parameters PLUS the other design factors related to floor space use (described in the previous section "Use of cabinet as a measure of physical space").

Table 1

Five key parameters necessary to allow space and density specifications to be established

Specification parameter	Definition	How it is used in the design
Number of units (#)	Number of cabinets in a pod, pods in a room, or rooms in a facility	To convert the per unit (per cabinet, per pod, per room) values of power, cooling, and space to total values
Design target average power per unit (kW)	Expected full load (rated) power per unit averaged across the population	To size the bulk power and cooling systems for the level (pod, room, facility)
Peak power per unit (kW)	Maximum expected power of the highest unit in the population	To size power distribution and cooling distribution system requirements
Unit power uncertainty (%)	Quantifies the expected uncertainty of the actual power compared to the design target average power	To determine the reserved space needed to ensure that low density deployment will not strand costly power and cooling capacity
Managed power ratio (%)	Power reduction factor (% of design target average) due to power management functions in IT equipment	To establish the operating load points for power and cooling systems to determine efficiency and energy use

While it should be apparent that the number of units and design target average power per unit are necessary to form a density specification, peak power, unit power uncertainty, and managed power ratio are new concepts introduced here. These are necessary because the **design target average power does not provide the information necessary to size the power and cooling distribution systems, nor does it provide enough information to determine operating points for efficiency calculations**.

Example density specification

To demonstrate the use of the improved density specification method and to show how it relates to floor area, two examples are presented: a small 40 kW server room and a more complex 2 MW multi-room facility with a growth plan.

Example 1: Small server room

A small 40 kW server room is a simple case chosen because the facility has only a single IT room, with a single pod. In this case, the level of specification is the room, which is also the pod, and which contains a group of IT cabinets. The structure of the specification is simple in this case, and the complete specification for this design is shown in the **Figure 2** worksheet. It can be downloaded in the resources section of this white paper.

This specification provides clear and unambiguous guidance to the design. The yellow boxes in the table are user inputs, and the summary performance is calculated.

Figure 2

Worksheet organizing key parameters needed to establish space and density specifications, showing computation of summary values

Density Parameters

	Cabinets in room	
Number of cabinets	12	
Design target average power per cabinet	4	kW
Peak cabinet power	8	kW
Cabinent power uncertainty +/-	15%	(80% confidence)
Managed power ratio	70%	

Space Use Parameters

Area per cabinet	16	sq ft	

	Cabinets	sq ft	subtotal
Cabinet area requirement	12	0	192 sq ft
Space reserved for staging	2	0	32 sq ft
Suggested space for density uncertainty			*34 sq ft*
Space for density uncertainty	2	0	32 sq ft
Space reserved for power	2	0	32 sq ft
Space reserved for cooling	2	0	32 sq ft
Space reserved for ancillary systems	2	0	32 sq ft
Space reserved for storage	0	25	25 sq ft
Space for egress, ramps, and columns	0	40	40 sq ft
			417 sq ft

Room Performance Summary

Rated system power	48	kW
Expected IT operating power	33.6	kW
Peak rated power per cabinet	8.0	kW
Nominal power per cabinet	4.0	kW
Average expected power per cabinet	2.8	kW
Room size	417	sq ft
Expected unused IT space	15%	of total space
Room power density	115	W per sq ft

The density parameters in the specification are determined using the following simple procedure:

1. The number of cabinets is established by the IT requirement.
2. The design target average power per cabinet is determined by IT vendor specifications or by choosing typical average design values for the application. In this case, a typical value for a corporate server room of 4 kW per cabinet was chosen.
3. The peak power is chosen by establishing the maximum expected or allowable cabinet power. In this case we have specified an 8 kW max capability.[4]
4. The cabinet power uncertainty is estimated by considering different scenarios for IT deployments or by choosing typical design values for the application. In this example, the expected deployment density was set to +/- 15% from the design target average of 4 kW.
5. The managed power ratio is estimated based on the expected power management functionality of the IT load. In this example, the power management functions are expected to reduce the actual average power of the IT loads to 70% of the design target average value.

To determine the room space requirement, first the area required by the IT equipment (including front and rear cabinet access)[5] is defined, and then other space requirements are

[4] Note again that the peak capability is provided to accommodate some racks at the peak power. The total power of all racks cannot exceed the power computed from the design average value.

explicitly added to determine the total area requirement. For each non-IT space use, the worksheet is set up so space can be reserved in terms of either cabinet locations or in square feet. This is convenient when accounting for space consuming devices such as power, cooling, or patch panels that come in a cabinet form factor. A "Suggested space reserved for density uncertainty" is calculated based on the power uncertainty specified by the user. The user then explicitly reserves either square feet or cabinet locations to meet the suggested space reserve. In the example of **Figure 1**, the suggested reserve is 34 ft^2 and the user can nearly achieved this value by reserving two cabinet locations which translates to 32 ft^2.

At this point, it must be acknowledged that the specifier may not have all of the above information, because the IT requirements are only vaguely known, or the exact configuration is not yet finalized. For this reason, a table of typical values for different applications is provided in **Appendix 1** of this paper. **The key idea here is NOT that the specification exactly mirrors any specific detailed device-by-device IT plan (which is almost never known in advance), but rather that the specification ensures that the data center will have a known, predictable performance.**

In this example, the specification explicitly defines a design that supports any combination of up to twelve racks where the design target average power is 4 kW and the peak of any one rack is less than or equal to 8 kW. Furthermore, the average power when power management in considered is expected to be 70% of 4 kW per cabinet or 34 kW total, so any performance guarantee for efficiency of the main power and cooling plant should be made at that power density. To ensure the peak power is not exceeded, this data center would have an up-front IT deployment policy that states the maximum per cabinet power is 8 kW and greater loads must be split among cabinets. The additional reserved space ensures all of the available power and cooling can be utilized if the actual IT power density is up to 15% less than the design target average 4 kW value. ***Note that none of this key information is understood if the classical watt per square foot density rating is specified.***

Example 2: Large data center

In this 2 MW example, the data center is defined with the following hierarchy:

Data center facility, comprised of

4 IT rooms, each comprised of:

9 IT pods, comprised of:

10 IT cabinets

There are two basic approaches to the specification of a large data center:

1. Start the specification at the cabinet or pod level and build up the facility specification
2. Start the specification at the facility level, and cut up the specification into rooms, then pods, then cabinet specifications

Ideally, the first approach should be used, however, in many cases this is not practical, because the facility level constraints have been defined first, such as the available mains power, or the physical size of the building. Given a known facility power requirement, the

---

[5] It is useful to include front and rear access space as part of the cabinet footprint, because these required access and egress areas are then automatically included when racks are added or removed from a pod, so that a separate calculation of the total egress area is not necessary.

specification must be broken down into rooms and pods, and then roll them back up to the facility level, using the following procedure:

1. Determine the number of rooms in the facility, establishing the room power
2. Determine the number of pods in a room, establishing the pod power
3. Determine the number of cabinets in a pod, establishing the cabinet power
4. Establish the facility, pod, and room space use parameters
5. Determine remaining density parameters
6. Roll up the total specification and validate against the design constraints
7. Adjust and repeat until the design meets the requirements

For simplification, this example assumes that the specifications for all rooms are the same, and all pods are the same, and variation only exists at the cabinet level. This is an appropriate design assumption for many cases. **Figure 3** illustrates the worksheet used for this example 2 MW specification. It can be downloaded in the resources section of this white paper.

Figure 3

Example worksheet for a 2 MW data center

Density Parameters

	Room Units in Facility System		Pod Units in Room System		Cabinet Units in Pod System	
Number of units	4		10		10	
Design target average power per unit	500	kW	50	kW	5.0	kW
Peak power per unit	500	kW	50	kW	12.5	kW
Unit power uncertainty +/-					24%	
Managed power ratio	80%		80%		80%	
Area per unit	4480	sq ft	280	sq ft	14	sq ft

Space Use Parameters

	Rooms	sq ft	subtotal	Pods	sq ft	subtotal	Cabinets	sq ft	subtotal
Area requirement for units	4	0	17920 sq ft	10	0	2800 sq ft	10	0	140 sq ft
Space reserved for staging	0	500	500 sq ft	1	0	280 sq ft	0	0	0 sq ft
Suggested space for density uncertainty			*283 sq ft*			*324 * sq ft*			*44 sq ft*
Space reserved for density uncertainty	0	0	0 sq ft	1	0	280 sq ft	2	0	28 sq ft
Space reserved for power	0	2000	2000 sq ft	0	80	80 sq ft	1	0	14 sq ft
Space reserved for cooling	0	2000	2000 sq ft	0	80	80 sq ft	3	0	42 sq ft
Space reserved for ancillary systems	0	400	400 sq ft	0	80	80 sq ft	0	0	0 sq ft
Space reserved for storage	0	500	500 sq ft	0	80	80 sq ft	0	0	0 sq ft
Space for egress, ramps, and columns	0	2000	2000 sq ft	0	800	800 sq ft	0	56	56 sq ft
			25320 sq ft			4480 sq ft			280 sq ft

Performance Summary

	Facility	Room	Pod
Rated system power	2000 kW facility	500 kW / room	50 kW / pod
Expected # IT cabinet requirement	400 cabinets / facility	100 cabinets / room	10 cabinets / pod
Max # IT cabinet capability	576 cabinets / facility	144 cabinets / room	12 cabinets / pod
Expected IT operating power	1600 kW facility	400 kW / room	40 kW / pod
Peak rated power per Unit	500 kW / room	50 kW / pod	12.5 kW / cabinet
Nominal rated power per Unit	500 kW / room	50 kW / pod	5.0 kW /cabinet
Average expected power per Unit	400 kW / room	40 kW / pod	4 kW / cabinet
System size	25320 sq ft / facility	4480 sq ft / room	280 sq ft / pod
Expected unused unit space	2% of facility	13% of room	10% of pod
System power density	79 W per sq ft (facility)	112 W per sq ft (room)	179 W per sq ft (pod)

** This field needs to sum the space from the various pods, if more than one type of pod exists*

This compact worksheet contains a large amount of information about the design. The worksheet is grouped into three input columns: the left column describes how the facility is comprised of rooms, the center column describes how a room is comprised of pods, and the right column describes how pods are comprised of cabinets. Data center attributes that are defined from the worksheet for this 2 MW example are:

- A pod is comprised of 12 IT cabinets plus 4 cabinet locations dedicated to power distribution and in-row cooling, with a pod footprint of 20 ft by 14 ft (6 m x 4 m).
- The design target average power per cabinet is 5 kW.
- The peak power allowed in any cabinet is 12.5 kW as long as the pod power does not exceed 50 kW for all 12 cabinets combined.
- The total indoor space required by this design is 25,320 ft^2 (2,352 m^2).
- Using conventional metrics for W/ft^2 at the IT room level, this data center has a density of 112 W/ft^2 (1,206 W/m^2).
- In each pod, 2 spare cabinet locations have been provided, to allow the pod power and cooling to be utilized in case the deployed average power is less than the specified 5 kW per cabinet.
- In each room, 2 spare pod locations have been reserved, one is for staging of new pods without disrupting an existing pod, and one to allow the room power and cooling to be utilized in case the deployed average power is less than 5 kW per cabinet.

When compared to the previous single room example, this worksheet is more sophisticated in how it keeps track of space reserved for density uncertainty. The uncertainty is captured at the IT level in individual cabinets, so the user need not enter uncertainty for the pod or room levels; these levels are just computed roll-ups of the lower level. However, the user can reserve space to account for uncertainty at different levels in the design. The user can reserve for density uncertainty by leaving extra space in a pod for more cabinets, or by leaving extra space in a room for more pods, or by leaving extra space in the facility for more rooms, or a combination of these three approaches. Which method for reserving space is preferred is often controlled by room geometry or other factors. The worksheet keeps track of the total space requirement and allows the user to reserve space with any combination of extra cabinets, pods, or rooms.

Extending the approach to data centers with pod and room variations

Ideally, the pod and room designs within a facility are uniform and standardized. This provides a number of benefits to the design, including:

- Simplicity of scaling
- Standardization of management tools, methods, and procedures
- Simplicity of planning and design

However, this is not always appropriate or even feasible, due to the following:

- Known different types of IT equipment with greatly differing requirements will be deployed
- The room dimensions are already defined and cannot be standardized
- Different areas have different availability requirements which will affect the amount of space taken by redundant power and cooling equipment

A recommended best-practice is to define a **minimal set** of standard cabinets, standard pods, and in very large data centers, standard rooms. For example, a large data center might define three different standard pods *with the same footprint* for low density, high density, and storage. The data center would be designed for an expected mix of these pods, but flexibility could be maintained to adjust the mix during a long term deployment.

The worksheet shown in **Figure 3** uses a statistically average cabinet, deployed in a standard pod, into standard rooms. To use this method with a practical mix of pre-defined pod or room types, the worksheet must be extended.

Choosing density parameter values

Since the density parameters can drive significant costs, guidance is provided in this section for selecting values. **Appendix 1** provides suggested density parameter values for a number of common applications. This is an excellent starting point for developing a density specification for any type of installation. They can be used "as is" or adjusted to suit special requirements.

Number of units

For a simple, small data center, the number of "units" is the number of IT cabinets in the room. For a large data center this takes on three values: the number of IT cabinets in a pod, number of pods in a room, and number of rooms in the facility. Most of these values are established by facility constraints. However, one key design value is the number of IT cabinets in a pod, which affects many aspects of the design. This leads to the question as to whether there are preferred values for the number of cabinets in a pod.

Since a pod by definition typically includes power and cooling distribution systems, it often becomes impractical to deploy a pod at less than 20 kW, which translates to between 2 and 6 cabinets, depending on density. Since a pod is generally considered a contiguous group of cabinets, and legally-mandated personnel egress around pods is generally required, a maximum number of cabinets is around 24, which translates to between 75 and 500 kW depending on density. A key factor in determining pod size is the best-practice of doing pod-by-pod refreshes and retirements (as opposed to cabinet-by-cabinet refreshes). Smaller pod sizes allow smaller scale refreshes. A 500 kW pod roughly translates to 1000 servers which is appropriate only for large cloud providers. For many customers a pod size in the 50-100 kW range, corresponding to 100 to 200 servers is more practical.

A recommended number of IT cabinets in a pod is from 8 to 24. Large data centers and low density suggest a higher number, while smaller data centers and high density suggest a lower number.

Design target average power per unit

Choosing the design target average power per cabinet can be quite confusing and controversial. It has a very large effect on the data center design and cost, and there are conflicting recommendations in the literature. **First, it is important to note that most of the discussions in the literature do not correctly separate the ideas of uncertainty and peak values which are important to the density specification.**

Recall that the design target average value is the target average (rated) cabinet power averaged across the population of cabinets. If a 5 kW design target average cabinet power is specified, 20 kW cabinets can still exist in the pod as long as peak power specification accommodates them. There is a tendency to over-specify the design target average power per cabinet for a data center, to accommodate future IT equipment which is widely said to be increasing in power requirements, or to provide an apparent safety margin. Therefore, picking a high number appears to be prudent. However, as explained earlier in this paper, this is NOT the case, and in fact specifying a higher number for this value than is actually eventually deployed can move the data center into the left side of the curve in the earlier

Figure 1, resulting in extreme waste and inefficiency. This unfortunate condition has been reached by many data center operators who specified a high design target average cabinet power value and thought they were designing a data center to accommodate future needs. A better approach is to choose a total power in kW for the data center and then choose a best guess of the expected cabinet power. **Supporting future densities different from the design value are dealt with using the uncertainty and peak parameters in the following sections**.

The range of power per cabinet in actual data centers today is from 2 kW to 30 kW per cabinet. However, average values of over 12 kW are quite rare and are mainly achieved in high performance computing (HPC) or high density cloud computing applications. The vast majority of mixed use data centers within organizations exhibit average values in the range of 4 kW to 8 kW per cabinet. However, in general, the existing average values in a data center are not a good predictor of future values. Consolidation, standardization, new server technology, and virtualization are driving up per cabinet densities. As a general rule, new data centers designed for mixed IT use should be specified to a design target average cabinet power at least 50% greater than the organization is currently running.

Unit power uncertainty

An effective density specification will almost always have a non-zero value for unit power uncertainty. The only exception would be in a case such as HPC where the facility may be constructed for a very specific standard IT deployment where the power requirements of the IT equipment are exactly known.

It is important to remember that this uncertainty value is used to describe how the **deployed average power density across all cabinets** is expected to vary from the design target average design value. This number is NOT used to deal with variations between individual units (between cabinets, rooms, etc).

For example, if it is known in advance that the IT load will consist of 80% 4 kW cabinets and 20% 20 kW cabinets, then we know with certainty that the design target average power will be 7.2 kW/cabinet (the weighted average). Even though there is a mix of racks ranging from 4 kW to 20 kW in this example, the unit power uncertainty is **zero** for this population of cabinets. If instead, the percent of 4 kW cabinets is not definitely 80%, rather is between 70% and 90%, and the remaining cabinets are 20 kW, then the unit power uncertainty becomes +/-1.6 kW or 22%.

To establish the unit power uncertainty, assume the data center is fully populated and take the IT deployment assumptions that lead to the lowest average power per cabinet, and the assumptions that lead to the highest average power per cabinet, and use half the difference between these figures. Because designing for uncertainty has a cost, it is a good practice not to use absolute worst case assumptions about extremes of high or low density, but rather to consider those assumptions that establish the density range with 80% confidence.

Peak power per unit

Almost every data center has some variation of power among cabinets. It is common to find cabinets operating from 50 watts (a network switch with patch panels) up to 30 kW (fully loaded high performance blade servers). This represents a range of 60 to 1 in power consumption.

If a data center is required to deal with power variation among cabinets, the power and cooling distribution systems must be able to provide for the **peak** power values of the individual cabinets. Therefore, power and cooling **distribution** must be sized larger than would be expected from the design target average cabinet power. **The design target**

average cabinet power drives the bulk power and cooling plant ratings, but the peak cabinet power drives the power and cooling distribution ratings. Oversizing power and cooling distribution has a cost, but provides the ability to handle power variations among cabinets. When the ratio between the peak cabinet power and the design target average cabinet power is 3X or more, it may be desirable to manage the cost by attempting to reduce this ratio. Two techniques for optimizing the peak to design target average power ratio are:

1. Group cabinets of similar power together into pods, and define pods for different densities. The design target average power per cabinet will be different between pods, but the ratio between peak and design target average for the individual pods will fall.
2. Control the maximum power per cabinet by policy. Require IT deployments above a certain cabinet density to break up equipment among cabinets. By capping the peak power it is not necessary to provision for extreme power or cooling distribution equipment. This is very effective if the data center is expected to have a small fraction of blade server cabinets.

To assist with establishing appropriate peak values, **Appendix 1** provides typical values for different applications.

Managed power ratio

The power management features in modern IT systems cause the average power drawn over time to be less than the value at full computational load. For capacity purposes, power and cooling systems must be designed for the IT power at full computational load. However, for establishing electrical efficiency, the average power is the more important number, because it establishes the average expected operating power level of the power and cooling systems.

The smaller the power ratio, the more the power and cooling systems are forced to operate at light load conditions, where their efficiency is typically reduced. The managed power ratio is approximately 95% in typical data centers today, but is projected to fall to between 40% and 80% for many applications in the next few years.

Note that the managed power ratio does not specifically modify either the area of a data center, or the ratings of the bulk power and cooling systems or the power and cooling distribution systems. However, it strongly impacts the efficiency or PUE of a data center, **which should influence the choice of system architecture**. Low managed power ratio values suggest data center designs that are modular and scalable, or have excellent light load efficiency, leading to major life cycle energy cost savings. A projected data center energy efficiency that is modeled or calculated without consideration of the managed power ratio, will result in over-optimistic and suspect calculations.

IT density policies

In many cases, the data center operator may have choices about cabinet density. For example, when rack mount servers are deployed, cabinet density can be limited by simply leaving blank spaces in a cabinet. Furthermore, an operator may have discretion in the deployment of individual devices, and may be able to mix high and low-density devices within cabinets to control the watts per cabinet. These decisions may be the responsibility of the data center operator or they may be under the control of users or other parties.

Every data center should have policies regarding density. The peak cabinet power should not be established by guessing worst case maximum power of future IT devices, but rather by establishing a reasonable limit and forcing deployments to remain within that limit. Data centers with a wide projected range of cabinet operating power benefit from having cabinet power limits specified by pod, establishing pods specifically for either high or low density, and then establishing policies by pod type.

A common problem observed in many existing data centers is that some or most of the IT deployment is at average or peak density beyond the capabilities of the data center. This condition leads to overloads and overheating as the distribution capabilities are stressed and can cause the facility to run out of bulk power or cooling capability. These are obvious conditions. However, as we have pointed out earlier, deployment at *low density* is also a problem because it may cause power and cooling capacity to become stranded as the data center fills up. Unnecessarily poor utilization of U space in cabinets is a common contributor to low density. **Therefore, both minimum and maximum cabinet deployment power should be monitored and subject to policy**. A sample density policy is provided in **Appendix 2**.

Application to modular data center design

The techniques and methods described can be applied to any unique data center project. However, they also lend themselves to the specification of density for standard modular data centers. Standard pre-engineered modules or reference designs for pods, rooms, and facilities can and should have density specifications using this approach, greatly simplifying the design of data centers.

For further discussion of modular architecture and how density and other specifications apply, please consult White Paper 160, *Specification of Modular Data Center Architecture.*

Conclusion

When a data center power density is specified by a single number, such as W/ft^2 or W/m^2, many important performance characteristics are not defined. This can lead to considerable confusion during the specification, design, and commissioning processes, and leaves operators with a limited understanding of the capabilities of the data center.

A data center should be specified in a way that allows consideration of the key design constraints, while providing unambiguous guidance to the engineers and contractors who detail the design. The specification should clearly provide the information that data center operators need to establish operational policies and procedures and to give operators the confidence that the performance of the data center will be predictable.

This paper has introduced a logical and rapid approach to documenting data center space and density requirements that provides sufficient detail to assure that performance is predictable and not left to chance. When a data center is specified in this way, it provides much more complete and clear guidance to the detailed data center design than is provided with the historic methods.

Even data centers with incomplete information and uncertain plans can use this method. To assist users attempting to define a density specification, typical design values have been provided. It is envisioned that standardized, pre-engineered, and modular pods, rooms, and facilities would provide density specifications using this method, simplifying data center design.

About the author

Neil Rasmussen is a Senior VP of Innovation for Schneider Electric. He establishes the technology direction for the world's largest R&D budget devoted to power, cooling, and rack infrastructure for critical networks.

Neil holds 25 patents related to high-efficiency and high-density data center power and cooling infrastructure, and has published over 50 white papers related to power and cooling systems, many published in more than 10 languages, most recently with a focus on the improvement of energy efficiency. He is an internationally recognized keynote speaker on the subject of high-efficiency data centers. Neil is currently working to advance the science of high-efficiency, high-density, scalable data center infrastructure solutions and is a principal architect of the APC InfraStruXure system.

Prior to founding APC in 1981, Neil received his bachelors and masters degrees from MIT in electrical engineering, where he did his thesis on the analysis of a 200MW power supply for a tokamak fusion reactor. From 1979 to 1981 he worked at MIT Lincoln Laboratories on flywheel energy storage systems and solar electric power systems.

Resources

Data Center Projects: System Planning
White Paper 142

Specification of Modular Data Center Architecture
White Paper 160

Browse all white papers
whitepapers.apc.com

Browse all TradeOff Tools™
tools.apc.com

Space & Density Worksheet: Small Server Room Specification
Double click icon to access spreadsheet

Space & Density Worksheet: Large Data Center Specification
Double click icon to access spreadsheet

Contact us

For feedback and comments about the content of this white paper:

Data Center Science Center
dcsc@schneider-electric.com

If you are a customer and have questions specific to your data center project:

Contact your Schneider Electric representative at
www.apc.com/support/contact/index.cfm

Appendix 1: Typical density specification parameters application

Table A1 provides typical starting point values to assist in calculating space and density for a data center. Note that some values cannot be generalized because they depend on the design or business model and must be determined for a specific project.

Table A1

Five key parameters necessary to allow space and density specifications to be established

Specification parameter	Small enterprise	Large enterprise	Colocation	Cloud	HPC
Number of cabinets per pod (#)	4-10	10-14	6-14	10-20	10-20
Design target average power per cabinet (kW): Full load power of the average cabinet	4 kW	6 kW	4 kW	12 kW	16 kW
Peak power per unit(kW): Maximum power of the highest power cabinet	8 kW	12 kW	12 kW	25 kW	25 kW
Unit power uncertainty (%): The uncertainty of the estimate of design target average power	50%	30%	50%	30%	50%
Managed power ratio (%): Average of actual power draw to full load power of the cabinets	90%	80%	90%	70%	90%
Type of space reserved for staging	Cabinet Locations	Pod Locations	Pod& Room Locations	Pod & Room Locations	Pod & Room Locations
Amount of space reserved for staging	5-10% of total cabinet count	1 pod per room	Depends on business model	1 pod per room	1 pod per room
Space reserved for power and cooling	Design dependent	Design dependent	Design dependent	Design dependent	Design dependent
Space reserved for ancillary systems	5% of room	5% of room	10% of room	5% of room	5% of room
Space reserved for storage	10% of room	5% of room	none	none	none
Space reserved for egress, ramps, and columns	50% of room	30% of room	Design dependent	20% of room	20% of room

Appendix 2: Sample density policy statement

To assure predictable and reliable performance of this data center, the following policies guide the installation of IT equipment:

DENSITY POLICY

This pod is designed for an average per-cabinet power of xx kW and a peak of xx kW.

All U locations of all rack enclosure must be occupied with either IT equipment or a blanking panel, in order to maximize efficiency and reduce hot spots.

Individual cabinets may not be configured to be above the peak rating, because the power distribution is not rated to support loads greater than the peak, and the cabinet may not receive appropriate airflow. If a cabinet exceeds the peak limit, one of the following options applies:

- Remove some IT loads and spread them to other available cabinets that are not at the limit.
- Apply for a special accommodation from the data center engineering department, which may require re-wiring and/or limitations on the use of adjacent cabinets.
- Find or commission an alternative pod that has a higher peak cabinet power rating.

If cabinets are configured with low density devices such as patch panels, switches, storage, or other low power devices, try to ensure that the average power per cabinet in this pod is maintained near the average rating. Excess installation of low density equipment can result in stranded unusable power and cooling capacity. If the average power is below the pod rating, one of the following actions is recommended:

- Review the physical mounting of devices to ensure that the arrangement minimizes excess unused U space.
- Consider rear mounting some of the low density devices to reduce U space use.
- Use higher density patch panels if feasible.

To determine the power draw of existing cabinets, equipment in-hand and intended for installation, or proposed equipment, please consult with the data center engineering department.