DATA CENTER KNOWLEDGE

THE INFORMATION SOURCE FOR THE DATA CENTER INDUSTRY

EXECUTIVE GUIDE SERIES – PART 1

Data Center: Build vs Buy

by Julius Neudorfer

This is the first of a six part series of our Executive Guide whitepapers:

1. Data Center: Build vs Buy
2. Data Center Total Cost of Ownership
3. Data Center Energy Efficiency
4. Creating Data Center Strategies with Global Scale
5. Custom Data Centers
6. Data Center Designs

Brought to you by

DIGITAL REALTY
Data Centre Solutions

Executive Summary

Computing architecture and systems continue to constantly evolve by their very nature. The latest technologies, such as virtualization and cloud computing, are the current darlings and buzzwords of the industry, nonetheless, the need to support them still depends on a physical data center.

Build vs Buy is an age-old question for many industries. It is sometimes driven by hard number-crunching business analytics, while in some cases it is controlled by human emotion and ego or simply out of habit. Some issues can be clearly defined, such as cost and availability of capital and time to market, while other factors such as market reputation and perceived strength, are a little fuzzier.

There is an emerging trend which is growing rapidly toward outsourcing the data center as many organizations re-evaluate the need to build, own and operate their own data centers. They are either outsourcing it in whole or in part, in an effort to concentrate their resources on their core business.

While building, owning and operating a data center could be seen by some as a normal part of any business with a P/L statement, it differs significantly for an Enterprise customer vs. a Colocation, Hosting, Managed Services Provider or Cloud Provider.

For Enterprise data centers, the equation may not be as clear cut as an item on their normal P/L statement, since the "Profit" related to the data center is somewhat subjective, however reliable computing is critical to the operation of the business, while the "Loss" represents a potentially large financial risk factor should the data center experience an unplanned outage.

For Colocation, Hosting or Managed Services Providers, the P/L somewhat resembles a hybrid of traditional commercial real-estate operation, mixed with the seating revenue, economics and fuel efficiency issues of an airline. These are clearly understood and definable OpEx numbers to those organizations. However, even those organizations may not have the all in-house expertise to cost effectively build their data centers and have outsourced building and operating the facility.

Depending on your own organization's type and size, performing an honest self-evaluation of core strengths and the level of design-build expertise and experience of internal resources is sometimes a difficult, but necessary task in order to help make the best Build vs Buy decision.

However, for all types of users the designing, building and operating of a data center is somewhat different than other buildings and requires a highly specialized level of expertise and experience. Even very experienced data center operators may not have the design and build resources necessary to undertake a major building project. The nuances and subtleties of a sophisticated data center design are somewhat similar to some other types of specialized facilities such as an Oil Refinery or Chemical Plant, all of which require highly specialized design skills and experience and may not be part of your internal team's core skill set.

This whitepaper examines some of the major issues in more detail that should be part of your due diligence in making the Build vs Buy decision.

Defining "Build vs Buy"

The term "Build" in this case implies a "Build It Yourself" aspect. Your organization will be engaging outside professionals and general contractors to design and build your data center. It should be noted that not all architects, engineers and contractors are knowledgeable about the latest technology and practices in designing and building data centers, so care must be taken to evaluate their experience. Ultimately however, they will be working under your direction, and within limits, they will build whatever your organization's in-house team requests, for better or worse, depending on their design decisions.

The term "Buy" in most cases represents the flexibility to lease a facility from a provider whose core business is designing, building and operating data center facilities for their customers. Since this is their primary business, they can be very cost effective, as well as being able to rapidly build a data center using a well proven design. It is therefore important to carefully examine the various potential data center providers' offerings and their other existing facilities, as well as their operating history and financial strength.

Expectations of Continuous Availability and Understanding Tier Levels

The expectation of continuous availability, redundancy, fault tolerance and "tier" levels is an often debated and misunderstood subject by many, leading to the expectation by senior management that data centers can be "bullet proof".

The terms Availability, Redundancy, Reliability and Uptime have been used almost interchangeably by many as well. Ultimately, the data center's level of system "Redundancy" is the basis of the projected level of "Availability" (effectively its "Uptime"). Essentially, should a component or system fail, a redundant component or system is in place that will continue to support the "critical load" (computing equipment). Reliability is simply a projection of the expected life of a component or system, it is not a guarantee. Even a brand new "highly reliable" component or system can fail a month after installation; a properly designed redundant system will ensure that the data center remains available, even during a component failure. (see Uptime Sidebar, page 4)

There are endless references to the proverbial "Five 9s" (99.999% availability) which is commonly touted in marketing material, yet with very little actual detail in many cases. "Five 9s" refers to the percentage of time in a year that a system or facility is available (99.999% = 5 minutes of annualized downtime). It can represent a projection based on the expected reliability and redundancy of the systems or it can be historic, based on the operating record of a data center.

The expectation of continuous availability, redundancy, fault tolerance and "tier" levels is an often debated and misunderstood subject by many, leading to the expectation by senior management that data centers can be "bullet proof".

Site Selection and Risk Factors – Knowing Where to Build

It takes a very knowledgeable team to fully evaluate the suitability of a location to build a new data center. Many factors come into play, such as the cost and availability of sufficient and diverse sources of power (as well as water), to meet immediate and future growth requirements. In addition, ensuring direct access to multiple fiber network carriers is also crucial. Moreover identifying risk factors, both the obvious ones, such as known seismic or flood zones, or the less obvious ones, such as adjacencies to "invisible" but potential hazards, such as airports and their related flight paths, must be an essential part of the final decision.

Operational Procedures – Keys to Successful Mission Critical Operations

Besides the level of infrastructure redundancy at the facility level, consistency of procedures for operations, maintenance and support of the critical infrastructure systems are key to ensuring continuous availability. In fact, the Uptime Institute created a specific sub-category called Operational Sustainability to define and evaluate operational procedures, as an addition to their well recognized Tier Classification rating system.

Proper facilities management procedures and 7x24 availability of qualified support and maintenance staff for the mission critical systems such as UPS, power distribution, back-up generators and cooling systems are key to ensuring continuous availability. The importance of this factor should not be underestimated, and there are significant ongoing costs of hiring and retaining the required highly specialized and qualified personnel.

It should be recognized that even with the best designs and equipment, systems fail unexpectedly. While in most cases redundancy should prevent an immediate outage, the ability to have immediate access to resources in the form of qualified personnel and on-site back-up components are crucial to reducing your exposure and preventing a wide-scale outage by returning a data center to full operational status. It has been shown time and again that most outages have ultimately been caused by human error, either during routine maintenance or by personnel that were not fully qualified or properly trained to react to an equipment failure or overload condition and responded incorrectly in trying to rectify the problem.

Understanding "Uptime"

Expectations and Tier Levels – Myths vs Reality

The data center industry is laden with inter-related terms such as Uptime, Tier Levels, Availability, Redundancy and Reliability. In order to make an informed decision it is important to understand what they mean and what is the actual significance for any proposals you are considering in your "Build vs Buy" analysis.

The concept of "Uptime" was pioneered by the Uptime Institute which was founded in 1993 and introduced its well defined Tier Classification system: I, II, III and IV, of which Tier IV represents the highest level of projected availability. Today, its Tier Certification system is globally recognized and its members are mostly Fortune 100-sized companies having multiple data centers averaging 50,000 square feet. And while not everyone subscribes to the Uptime Institute officially, marketing references to *Tier 2,3 or 4 are common among those seeking to ascribe a certain level of design or construction to a data center's overall availability or system redundancy, which may, or may not be totally accurate.

The terms "N, N+1 and 2N", typically refer to the number of power and cooling components that comprise the entire data center infrastrure systems. Wherein "N" is the minimum rating to of any component (such as a UPS, generator or cooling unit) required to support the critical load. An "N" system is non-redundant at all, and the failure of any component will cause an outage, effectively describing a tier 1 type facility. N+1 and 2N, represent increasing levels of component redundancies and power paths, roughly mapping to the tiers 2-4, however it is important to understand that redundant components in themselves do not guarantee continuous availability, nor insure compliance with an Uptime Institute certified data center tier level.

Moreover, besides redundancy, the ability to do planned maintenance or emergency repairs on systems may involve the necessity to take them off-line. This involves the key concept of "concurrent maintainability" which permits systems to be bypassed, without impacting the availability of the computing equipment. The Uptime Institute is well recognized in clearly defining concurrent maintainability in their Tier Level system. This is one of the key criteria in the design or certification of Tier III and Tier IV data centers.

Besides the level of infrastructure redundancy at the facility level, consistency of procedures for operations, maintenance and support of the critical infrastructure systems are key to ensuring continuous availability. Moreover, the Uptime Institute has now also established a related sub-category called Operational Sustainability to define and evaluate data center's operational procedures, as an addition to their well recognized Tier Classification rating system.

Whether you chose to build or buy, you should examine all of these factors very closely to understand what is being promised if it is a brand new building and/or what the operating history of a proposed data center facility provider has been.

Footnote:

*"It should be noted that besides the Uptime Institute, the Telecommunications Industry Association "TIA" also has a similar tier structure rating (1-4). It is a design specification referenced by some in the industry; however there is no certification process available from the TIA to assure adherence to its precepts."

Important Factors to Consider for Your Internal Build vs Buy Evaluation

- Your own organization's present level of experience operating a data center.
- Do you have an on-going existing relationship with a trusted design-build firm?
- The scale of the proposed site and its relative size to your largest or last site.
- The scale and scope of the project – whether 10,000 sf or 100,000 sf – also is a significant factor.
- Have you done this before or is this the first time you are considering building your own site of this magnitude.
- If you have built previously were you satisfied with the results i.e.
- Did the projected budgetary costs match the final actual cost?
- Was the project completed on time?
- Did the facility's system and overall performance meet the design criteria?
- How many design changes were made during the build-out phase?
- What was the cost of those changes, in both time and money?
- If mistakes were made, would you be able to avoid them in your next build out?

Decisions made in the design and build stages will not only affect the total CapEx of the data center, it will forever impact its energy efficiency and long term operating costs, as well as limiting the functional life of the data center.

Energy Efficiency and Sustainability

Many studies have shown that energy costs are now the largest expense related to operating a data center. In fact, in older less efficient data centers the energy costs for a typical server exceed the cost of the server over its typical 3 year operational life. A new building offers the greatest opportunities to incorporate the latest sustainable and energy saving LEED building practices, as well as installing the most efficient power and cooling systems. Designing and building for energy efficiency is no longer a political nicety to demonstrate your organization's commitment to social responsibility, it is an operational necessity in this constrained economy.

Expected Life of a Data Center

It should be noted that in the past, traditional data centers were fairly rigid and unchanged for several decades and were expected to last 15-25 years. That traditional static model and thinking is no longer valid. Some of those sites that continued to follow that model are now functionally obsolete, either because they are unable to handle the higher power densities of newer IT equipment or simply because they were not very energy efficient and could not be cost effectively upgraded once they were put in to production.

Functional obsolescence, even for some newer data centers built by less experienced organizations can occur far sooner than expected, since computing equipment changes constantly and very rapidly, which impacts the supporting infrastructure equipment. In some cases organizations were forced to compromise the better long-term design to meet short-term requirements, due to limited capital resources. Newer designs, incorporating flexible capacity, are based on extensive long-term capacity modeling that was done well before the facility was constructed. Moreover, capacity planning, when used in conjunction with a pre-planned phased design and modular build-out, can greatly extend the life of the data center well beyond 10 years. In addition, it will optimizing the site's overall energy efficiency over its lifetime.

Choices When Considering a "Buy" Option

As noted previously the "Buy" option typically represents a range of leasing options offered by data center facility providers. Typical lease commitments range from 5-10 years and average 8 years. They generally fall into three categories:

Powered Base Building

A powered building shell typically includes the main utility feeds and primary switchgear installed (may be available with or without generators and basic cooling systems), ready to have the end-user take possession and furnish all the interior systems themselves. This can accelerate the site's development timeframe and complete the basic build-out of your facility on a committed schedule. A powered base building usually provides a facility that includes power planning and entitlement approval with a dedicated utility infrastructure and switchgear hand-off, including power procurement and utility load letters. This is only recommended for the most highly experienced operators.

Move-In Ready

Move-In Ready is the next level beyond a Powered Base Building. As its name implies, it is a standardized predesigned facility ready to accept your computing equipment. It comes complete with the basic power and cooling infrastructure. Your organization can then have the option of installing its own racks, as well as power distribution and network cabling infrastructure or have the facility provider furnish and install it as per your specifications.

Build-to-Suit

One step up from Move-in-Ready, Build-to-Suit offers complete customization during the design phase, your team works interactively with the facility provider's design team from day one, starting with the site selection. This has the benefit of meeting your organization's specific requirements, while leveraging the facility provider's design expertise and build experience.

Managed Growth and Flexibility vs 100% Initial Build Out

One of the problems for many organizations, both well established and those that are new and rapidly expanding, is managing growth and having flexibility. The ability to design for flexible staged growth is an important part of the initial decision. This can be done with a preplanned modular design, which is built-out in stages. This lowers the initial capitalized costs, yet allows for seamless expansion as needed and avoids the operational and energy inefficiencies that can occur when 100% of all power and cooling systems are installed initially, but are under utilized in a facility that was fully built out up front.

Factors Impacting the Total Cost of Ownership of the Data Center

It is important to understand the difference between physical size and capacity. In the past, size and capacity were nearly synonymous. However, in today's computing environment, IT equipment power densities have changed that significantly. The cost to build a data center was previously based on the number of square feet of white space. Today that is only part of the cost, the power density (watts per square foot) and level of redundancy have a huge impact on cost (both CapEx and OpEx). The scale and Scope of the Project—10,000 sf or 100,000 sf— is also a factor. Experienced builders and operators of larger facilities typically have more expertise and resources to optimize the total cost of ownership than an organization with only occasional or limited data center building experience.

Part of the total cost will be determined by the cost and availability of capital for construction. In many cases, the "Buy" is really a lease commitment that mitigates the need to tie up significant capital in a potentially depreciating asset, which may have been built from an inflexible older or lower tier design. A Tier III or Tier IV facility is concurrently maintainable and allows the facility operator with the ability to continuously maintain and upgrade infrastructure equipment as it ages, with no disruption of service to the end-user client.

Time to Market

Controlling the time line for the project is going to impact the capital costs as well as the time of final delivery. Design and construction can typically range from 18-24 months, but sometimes can exceed 30 months if there are too many changes made (beware of "Feature Creep") which also increases total costs. One of the benefits of using outside providers is that they have a pretty well established formula and usually can complete a project in only 12-18 months. This can have a significant impact on TCO, since a 6-12 month difference saves both up front costs and offers a potential strategic advantage by bringing new computing resources on-line sooner, enabling more aggressive delivery of new applications. In some cases larger data center providers may have a suitable base building in their inventory, or a prepared site which is ready to build on. This could shorten the delivery to as little as 6-12 months.

Leveraging Contractor and Vendors Negotiations and Relationships

The design and building of a data center requires careful scrutiny of costs associated with the bidding of construction and the purchase of major capital equipment such as generators, UPS, power distribution and cooling system components, all of which are going to have a significant impact of the final costs. One of the advantages of using a larger "Buy" provider is their sizeable purchasing power and on-going relationships with both construction firms and equipment vendors. Ongoing relationships with major equipment vendor should not be overlooked, since this will be reflected in shorter delivery times for key components, such a large generators and chiller plants, which typically have long lead times.

Concerns over Ultimate Control and Responsibility For the Data Center

This is a valid issue and should be weighed carefully. One of the advantages of owning and operating your own data center is complete control of all aspects of your site. However, it should be considered cautiously against the ongoing cost and responsibility of maintaining all the properly trained facilities personnel required to support all the various critical systems on a 7 x 24 basis. Moreover, should an emergency occur, such as the failure of a major system, it is likely that a larger provider can draw upon a bigger pool of skilled support personnel to expedite repairs and minimize exposure to potential outages.

The Bottom Line

It should be part of the equation to investigate a number of "buy" options that are offered in today's market by a wide array of facility providers to lease, and operate fully managed sites. While there are a multitude of factors to be considered, in the end you will have to ask if building, owning and operating your own data center is a strategic advantage to your business, or just a burden on internal resources and capital, perhaps best to be outsourced to a qualified external organization. This fundamental question should help you make the ultimate decision, by closely examining your core business goals and the best use of key internal resources in your own organization.

This is the first of a six part series of Executive Guide whitepapers

Next in the series is:

2. Data Center Total Cost of Ownership
3. Data Center Energy Efficiency
4. Creating Data Center Strategies with Global Scale
5. Custom Data Centers
6. Data Center Designs