



DATA CENTER STRATEGIES

**Simplifying high-stakes, mission critical decisions
in a complex industry.**

July 2011

John Rath
Rath Consulting

Brought to you by



Executive Summary

Business climate changes, business models change, Information Technology certainly changes, and data centers change. Wait - data centers change? Yes, the view that the data center is *just* a room or facility that is a large capital investment and built every 10 to 15 years, has changed. This doesn't mean throw more money into it with greater frequency, it means approaching it in new ways and making intelligent decisions about the data center. The data center strategy is just as important as the IT strategy it protects and provides for and the business strategy that both empower.

A data center strategy is anything but a panacea. It must address the exclusive requirements of the business and take special consideration for the growing complexity of choices available. Strategic planning is not an attempt to eliminate risk or to forecast what the data center will look like in fifteen years, it is taking action to understand what risks to take and what paths will align with the business.

This paper explores three actions to take in order to formulate a pragmatic data center strategy:

- Leverage the knowledge of data center trends and technologies to apply to your business.
- Engage all stakeholders and apply appropriate cost modeling.
- Evaluate capacity, cost and capabilities to build an optimized data center strategy.

WHO SHOULD READ THIS PAPER?

- **Anyone responsible for leading data center strategies.**
 - **Chief Executive Officer**
 - **Chief Financial Officer**
 - **Chief Information Officer**
 - **Chief Technology Officer**
 - **Chief Security Officer**
 - **VP/Director of Data Center Operations**
 - **VP/Director of IT**
 - **VP/Director of Facilities**

Flux Capacitor – A term used to confuse the reader. See this and other technical terms defined in the Glossary beginning on page 23.

Contents

- Executive Summary 2
- The Mission Critical Decision Process 4
 - The Data Center Industry 4
 - Data Center Landscape 5
 - Data Center Taxonomy 7
 - The New Data Center 8
 - New Paradigms 9
 - Business Model Shift in the Data Center 9
 - Cloud Computing..... 10
 - The Opportunity 12
- Creating Value – Internal and Financial Perspectives 13
 - Partner Perspectives 13
 - Comprehensive TCO..... 14
 - The Economics of Data Center Operations 15
- Road Map to the Perfect Strategy 16
 - Own Versus Lease: Implications, Actions and Energy..... 16
 - IT Infrastructure Outsourcing..... 17
 - Wholesale Data Center 17
 - Two essential factors framing an effective data center strategy..... 20
- Summary 21
- Glossary 23
- References 26

The Mission Critical Decision Process

The data center paradigm has served as the foundation for Information Technology that either runs business or is the business. That paradigm has been evolutionary throughout the last several decades and transformational in the past 5 to 10 years. Formulating a data center strategy in the past was relatively straight-forward. A typical data center lifecycle was upwards of 15 years. As business and IT strategies iterated throughout those 15 years the data center strategy was stuck and unable to align its objectives.


Today the disruptive technologies that an IT strategy brings combined with the technologies and business options in the data center make for some complex decisions in strategic planning. Incorporating build or lease decisions and evaluation criteria into a data center strategy call for a mission critical decision process. Allotting the suitable amount of time for proactive planning is paramount and will yield a pragmatic data center strategy.

Taking three actions will aid in assembling the foundation of a sound data center strategy.

1. Leverage the knowledge of data center trends and technologies to apply to your business.
2. Engage all stakeholders and apply appropriate cost modeling.
3. Evaluate capacity, cost and capabilities to build an optimized data center strategy.

The Data Center Industry

The description of a data center has almost always been preceded with 'mission critical', because that is the service it provides – the mission critical hardware and software where maximum uptime is required. The data center is a fortress, dedicated to achieving maximum reliability at any cost. While reliability is still the key factor, the data center has evolved and advancements in the past five years have accelerated the pace of innovation.



Action --Leverage the knowledge of data center trends and technologies to apply to your business.

The definition of a data center:

Pre-1990: Large computer rooms where computer bugs were actually bugs inside of large computers.

1990 - 2000: The advent of client-server computing and the Internet began the definition for what a data center was. Large data center companies such as Exodus thrived in the dot-com era when Internet companies faced rapid growth. The data center focused on reliability.

2000 - 2007: Data centers in large cities fulfilled communication needs as a central place where networks exchanged critical information. Financial services with high-speed trading software drove the need for a central place for networks to interconnect, that was also milliseconds (and later microseconds) away from financial districts. Demand for data center services and outsourcing advanced the industry. ISP's, MSP's and ASP's, oh my! Shared and dedicated hosting, coupled with managed services and application hosting flourished. Internet companies grew larger and businesses looked to outsourcing.

Businesses that outsourced had growing footprints within colocation facilities and desired to take back some control over the security and financial flexibility to support their expanding business and IT needs. The concept of turnkey data center space being sold on a wholesale agreement was developed by Digital Realty Trust shortly after 2000. They took the idea of a REIT (Real Estate Investment Trust) to the data center industry by developing properties with advanced technologies and giving businesses a financial strategy for building a data center, without owning a data center. Around 2005 the concept of a modular, portable data center design was introduced.

2008 - 2011: In recent years there have been many innovations in power and cooling technologies and management of facilities. Efficiencies have been integrated into every aspect of the data center and building design, covering everything from bunkers to chicken coop design and mobile data centers to using the building as an air handler. Green technologies and environmental awareness have also been a large part of the industry in the past 3 years. No longer just a choice between build or lease, the data center can be owned, placed in colocation, wholesale, put in a public or private cloud, or a hybrid strategy of options.

Data Center Landscape

Looking at the data center landscape, technological advancements may have shaped how data centers were built and optimized, but the economics behind them define the strategies taken. Big Internet companies built big data centers. Enterprises and federal agencies with numerous legacy data centers embarked on massive consolidation projects, and many more took advantage of colocation or wholesale data center agreements to further their data center strategy.

The changes that have come about have even altered the meaning of what constitutes a data center. To Google, Microsoft, or Yahoo it is a hyper-scale facility with tremendous innovation engineered into it. For the consolidation projects it means taking what they once considered to be data centers and bringing them into a small number of new, large-scale facilities. To others, their definition of a data center was transformed by the advances in IT equipment that required more power, more cooling and a more advanced facility to support it.

Many facets of site selection for data centers were shaped by the proliferation of fiber networks and the need to both avoid natural disasters and achieve extreme power and cooling efficiencies. In the U.S. several regions became data center hubs, where Internet companies and enterprises selected to build new data centers. Silicon Valley continued to prosper and grow and regional hubs developed in Quincy, Washington, Chicago, Dallas/Fort Worth, North Carolina and the New York/New Jersey region.

Consolidation

One of the most visible consolidation plans in past years has been the Federal Government. In 2010 Federal CIO Vivek Kundra began an ambitious plan to dramatically reduce IT operations distributed among more than 1,100 data centers. As a part of the initial inventory phase it was discovered there were actually over 2,000 data centers in existence. Several other cloud projects were initiated at government branches, such as the FCC utilizing a Terremark (later acquired by Verizon) Infrastructure as a Service (IaaS) offering that would give them on-demand access to computing resources. Other projects from government organizations or large technology companies like HP and Intel have demonstrated the benefits of using innovative data center design to consolidate data centers.

Virtualization

While many technologies have had a profound impact on the data center, it can be argued that no other IT factor has contributed more than virtualization. Virtualization became a component of the larger cloud computing concept, which encompasses many technology and business facets.

With these technology innovations came an increased requirement of the data center facilities that were to support them. Data center operations faced a similarly complex environment as it adapted to not only ensuring reliability, but responding to increasing power and cooling demands.

Many assumptions about traditional IT environments that had followed the data center for years were now being challenged and data center technology saw innovation like no other time in history.

Data Center Taxonomy

Innovations in both technology and business have compounded the definition of what a data center is. The following types and use cases of data centers exist:

- **Warehouse-scale:** In 2009 Google engineers Urs Hozle and Luiz Andre Barroso coined this term in their book *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines*. It refers to the unique business case a large scale Internet company like Google has when faced with providing software to millions of users, and managing massive amounts of servers and storage. Google and Microsoft have spent billions of dollars on data centers in the past five years to keep up with these large scale dynamics. If your IT is adding or retiring thousands of servers every week, this model might be the right fit for your strategy. These facilities draw a large amount of power and will have special consideration and planning taken for site selection and energy sources.
- **Supercomputer data centers:** The facilities that house supercomputers have different requirements than most data centers. The supercomputer facility must provide electrical and mechanical systems for extremely dense compute nodes. To properly dissipate heat from these servers water-cooling can be delivered all the way to the chip level. While those facilities are typically a single-owner, purposed facility there have been cases of supercomputers or HPC (High Performance Computing) systems operated at a colocation or wholesale provider's data center.
- **Single-owner dedicated data center:** Owned, operated and controlled by the business. This category accounts for the majority of data centers in the market and can range from a small server room to large facilities with megawatts of capacity. Faced with the burden of maintaining data center technology in their own facility, many are seeking colocation or wholesale arrangements.
- **Public or private cloud:** While not physically a data center it can be referred to as a virtual data center in such forms as Platform-as-a-Service (PaaS), Infrastructure-as-a-Service (IaaS), or Software-as-a-Service. These are used in hybrid strategies where certain applications or business functions are a good fit for cloud computing. A private cloud has the option of being hosted at a company data center or through colocation or wholesale data center space.
- **Retail Colocation:** This ranges from the initial outsourcing needs to provide for a small IT footprint all the way up to numerous racks of IT equipment as customer need grows.
- **Wholesale lease:** Wholesale agreements with data center providers allow companies to lease floor space in a facility that is designed and operated to support their needs. Similar to a commercial building lease, the tenant is able to take liberties with use of their own electrical, mechanical and security solutions. The economies of scale in wholesale arrangements bypass customer requirements seen in colocation terms and conditions. Power requirements within a leased space for wholesale are typically greater than 300-500 kilowatts.

The New Data Center

Industry surveys place the age of an average U.S. data center around ten years old. Much of this can be attributed to a late 90's focus on preparing for Y2K and a build-out in response to the dot-com era. The data center landscape has been radically transformed since the days when those facilities were constructed, both from the technology within and the technology in data center infrastructure and operating philosophies. Since 2000 there have been a number of innovations in the data center that expanded business models, altered its very appearance and delivered benefits in operations and efficiency.

In 2008 Microsoft engineers Christian Belady and Sean James took on ingrained principles of the data center environment and asked why a cooling infrastructure was even necessary. In an experiment they took a rack of HP servers and placed it in a tent inside the company's fuel yard at one of their data centers. The result was that for the better part of a year the servers ran with zero failures.

Additional data center innovations from economizers to evaporative cooling and modular data centers to management and automation tools evolved the data center further. To continue its role as a vital supporting asset for IT the data center had to become more agile and empower the operations team to make it a strategic advantage.

Facebook

Although Microsoft and Google have made tremendous contributions to data center efficiency design and modernizing the data center, the most recent case is of the hyper-growth data center needs of Facebook. As IT requirements skyrocketed they approached the challenge with the goal of scaling computing infrastructure in the most efficient and economical way possible. Facebook released details of its new technology infrastructure, which features custom-built servers, racks and UPS units that will fill its new data center in Prineville, Oregon. Making use of the Open Web Foundation license they have made all server, data center and mechanical drawings available as a part of the Open Compute Project.

The ability of Facebook engineers to design an end-to-end solution in their Prineville data center meant they could innovate every aspect:

- Two tier structure with penthouse cooling – separates servers and cooling infrastructure and eliminated the use of ductwork.
- Custom racks, enclosed in a hot aisle containment system.
- Customized servers, cabling, power supplies and fans to meet their exact needs.
- 48 Volt DC UPS System integrated with a 277VAC power supply.
- 100% airside economization with an evaporative cooling system.

New Paradigms

Efficiency and innovation have been driven into every aspect of the data center. Innovation for the data center changed the paradigm in many ways:

- Modular: the concept of a modular data center has developed over the past five years, with Google and Microsoft as early adopters, deploying modular units of servers in their data centers. By combining modules with advanced power and cooling solutions they achieved high densities in the rack and noted electrical and mechanical efficiencies.
- Cooling technologies: Technologies for heat dissipation in the data center have progressed in products such as air and water economizers but also in core architectural design. Google removed the chillers completely in their Belgium facility, HP used free cooling and an innovative airflow design in their Wynyard England data center, and several companies have used geothermal solutions.
- Energy: Following environmental stewardship initiatives companies have taken advantage of renewable energy sources such as solar, wind, geothermal and hydropower. In some cases it is for the cause of stewardship and offsetting power consumption, and in some cases it is to supplement power being brought in and take advantage of the renewable source that is available where the facility is being built.
- Power: When designing new data centers companies are incorporating efficient power distribution and reducing electrical transmission line power loss. As a part of an expansion project Purdue University placed a data center container within the campus power plant property in order to be closer to the source. They also elected to not protect it with a UPS. Facebook and Vantage Data Centers built data centers next to utility substations. Vantage went one step further in their Santa Clara, California campus by purchasing the substation adjacent to the property. Data center integrations with smart grid and micro grid technologies have enabled *smart* power.
- Data Center Infrastructure Management (DCIM): This software gives operational teams insights into various IT, energy and infrastructure components of the facility through a single interface.

Business Model Shift in the Data Center

At the center of these disruptive technologies and rapidly evolving computing and business models is the data center. Its ability to keep pace and align with business and IT strategies is the difference between being a burden to keep up with or the catalyst to empower. Executives looking at the data center should embrace the challenges presented by this growing complexity and present a data center strategy that adapts to changing business and IT models and delivers on a promise for existing and future needs. The data center is thought of as an IT asset and decision, but ultimately is a business decision.

Sustainability

Innovations in cooling efficiencies and energy have attempted to address the massive power consumptions of data centers. Various government and private organizations have formed to help guide industry metrics and best practices such as Leadership in Energy and Environmental Design (LEED) certification from The U.S. Green Building Council (USGBC). LEED provides a framework for identifying and implementing practical and measurable green building design, construction, operations and maintenance solutions. The Green Grid, an industry consortium, facilitates a collaborative environment to build educational tools, technical resources and common metrics such as Power Usage Effectiveness (PUE).

Cloud Computing

The only possible match for the value that cloud computing has brought is perhaps the hype that has surrounded it. Now that it is on a downward trend from the hype cycle and matured into a variety of approaches, a clear path can be developed for business. In a report on *Sizing the Cloud* Forrester Research estimated a move from \$40.7 billion spent on the global market for cloud computing in 2011 to more than \$241 billion in 2020.

Public and Private Clouds

The distinction between public and private clouds is significant, as it presents the appropriate use cases for each. The public clouds are those from large internet companies such as Amazon and Microsoft and will typically serve the mass-market needs of consumers. A private cloud is better suited for the business with a desire to maintain control over security and management, as well as potentially run the private cloud on their own infrastructure. In either case the NIST (National Institute of Standards and Technology) defines cloud computing as an elastic use of computing power that delivers ubiquitous, convenient, on-demand network access to a shared pool of configurable computing that can be rapidly provisioned and released with minimal management effort or service provider interaction. NIST classifies cloud deployment models into public, private, community and hybrid.

Cloud computing is IT operation outsourcing, with an obvious impact on the physical data center. Faced with growing IT complexity many are lured by cloud concepts to empower their IT strategies to consolidate, streamline and simplify. Cloud computing software is designed to be in tune with the infrastructure it runs on, and thus the efficiencies, automation and agility make it an attractive option over the traditional data center model. As an example of outsourcing of a company's CRM (Customer Relationship Management) software, Salesforce.com delivers their complex CRM solution to 97,000 customers using 3,000 servers, a ratio of 0.031 servers per customer.

The benefits of cloud computing should not compromise an overall strategy to the operations of IT. Solutions fit to a public cloud should be balanced with an alternate option, and a private cloud should

have its redundancy capabilities optimized to IT requirements. The benefits of cloud computing should also not overshadow security – a primary concern for this relatively young technology.

A cloud computing strategy plays an important role in developing a business's data center strategy and understanding where the two intersect is important. When looking to investigate this intersection, look at a matrix of application delivery models and infrastructure delivery models available to the organization. There is no single best way to provide for every application or IT initiative.

Modular Data Centers

The reasons that led to modular design in the data center were partially evolutionary and partially the financial aspect of budgets and time to market. The evolutionary part of modular design grew out of an increasing unit of measure. Moving one level up from a focus on individual servers, IT would view an entire rack of servers as a unit. To reduce complexity and time to market for making use of the compute power, the rack became an easily deployable unit. As requirements continued to grow and the concept of hot and cold aisles in a data center caught on, the next modular step was to make an entire set of pre-populated server racks. From a financial view it was a phased approach applied to building out a data center - in modules that enabled a simpler financial justification over a completely new facility and that were built and deployed congruent with demand.

The early years of this movement took the shape of data centers in ISO shipping containers. The shipping container represented an international standard that was mobile, and could have racks of IT equipment pre-populated and shipped to the customer's site. There were numerous applications for this, but it was not practical for most data center needs. What was still attractive was the modular mentality of having a measurable unit with standardized components in it. It also helped remove some of the integration complexities by having an end-to-end, repeatable solution that could be implemented in increments to match demand.

Google and Microsoft were early adopters of the modular approach and engineered their own solutions to drive data center efficiencies. They are scaling thousands of servers and have the benefit of being able to standardize IT components to implement within the data center. As a unit of measure they can simply replace an entire module when a defined percentage of servers have been identified as not functional.

Supply Chain

Even repeatable modular designs for data center have financial and logistical inefficiencies and require the upfront capital to construct. Modular design shifts focus to the larger supply chain of components that make up a standardized, easy to maintain data center. No longer a niche play, large technology players such as HP, Dell, IBM, SGI and others have developed factory-built modular solutions that are constructed in a matter of weeks or months and delivered as a fully functional data center. The data center is built in a factory with standard components and can be customized to customer requirements.

As in supply chain management, a primary objective of a modular deployment strategy is to create net value and leverage logistics. Data center modules enable flexibility to scale the data center in sync with IT. It breaks the data center into pieces that can be financed as demand dictates. Power and cooling technologies have embraced modular design. Power systems and cooling technology modules can be deployed along side of a modular data center to support them.

The Opportunity

Supply chains, clouds, major business model shifts and data center factories – it's a lot to take into consideration for a data center strategy. Leveraging best practices and knowing available options, and more importantly what is feasible for your business is what will make a comprehensive and valuable data center strategy. Having knowledge of where the industry is and where it is headed will facilitate a better structure to the data center strategy and give proper context to those involved in the decisions.

Creating Value – Internal and Financial Perspectives

The data center is a complex ecosystem of many different functions coexisting under one roof. As such, there are many important inputs to make a holistic data center strategy. Information Technology is the primary input and reason for the data center to exist. Its needs are aggregated and IT strategy and forecasts are studied for how to best accommodate its needs.



Action --- Engage all stakeholders and apply appropriate cost modeling.

Partner Perspectives

To achieve a truly valuable data center strategy, seek out perspectives from internal partners:

- Finance: Financial departments play a key role and should be involved prominently and early in the process.
- Infrastructure, Facilities, Operations and Support groups play a vital role in performing maintenance and ensuring the reliability and uptime of the facility.
- Capacity Management: The ITIL (Information Technology Infrastructure Library) process of capacity management ensures that services and infrastructure are able to deliver the agreed capacity and performance targets in a cost effective and timely manner. An effective data center strategy enables flexibility to meet forecasted capacity estimates.
- Environmental Stewardship: Whether thought of as a function of facilities or of IT, the environmental aspect of the data center is important. Environmental considerations should be included in strategic planning early with appropriate personnel or departments collaborating on energy management and use of renewable energy sources.
- Real Estate: Depending on the size and reach of the business there is potential input from those dealing with Real Estate management.
- GRC (Governance, Risk, Compliance): Input is important from those with a stake in adherence to business and technology standards such as COBIT, ISO/IEC 20000 and ITIL. Additionally, consult with those responsible for Information Lifecycle Management (ILM) for protection of corporate data.
- Other Groups: Procurement, Human Resources, Security, Enterprise Project Management Office, Enterprise Architecture, and Networking.

Comprehensive TCO

Costs for operating the data center can be spread across many departments or divisions at a company, which reinforces the need for a data center strategy to include all of these areas. Numerous surveys have shown that managers responsible for the data center either did not have an understanding for what the data center power bill was or the bill was sent to a different department. Aside from ensuring that all costs are known, the TCO (Total Cost of Ownership) analysis must be a comprehensive look at all parts of what is involved in operating a data center. Investigate the total cost of acquiring, installing, using, maintaining, changing and disposing of IT and data center assets across an extended period of time. In a company-owned data center delivery of power and cooling will be the primary cost drivers of either new construction or maintenance of existing facilities. In the case of colocation or wholesale options, the cost structures will most likely center around kilowatts per cabinet or kilowatts or megawatts per module/space.

How TCO Is Influenced:

1. Energy: Pay particular attention to the cost of energy because failure to evaluate and forecast it properly will have a profound impact on the TCO analysis. The Uptime Institute's Kenneth Brill recently explained that the four year cost of a server's electricity is the same as the cost of the server itself.
2. Trends: Balance IT growth projections with reductions due to virtualization or other technologies. What are the trends for energy consumption, possible energy sources and prices going to be in the future?
3. Efficiency impact: What impact does a 2% efficiency gain at the processor level net in overall power consumption savings? Efficiency gains can be realized at all aspects of data center operations, technical and business alike.
4. Direct, indirect and overhead costs: Evaluate data center costs imputed from all parts of the company.
5. Energy Profile: Project electricity costs, the practicality of using renewables, and realize the impact of a site selection study or ability of a data center vendor to procure energy. Has the vendor thought about power capacity, scalability energy prices and sources for the future? Does it make sense to utilize a microgrid at your data center?
6. Compliance: What are the costs of obtaining required compliance with government regulations or other desired industry standards versus contracting with a provider that assumes those costs?

It all comes down to cost and a detailed look at driving down the total cost of ownership for a data center. While the best data center is one that doesn't have to be built, maintaining the one that does means it should be economically efficient.

The Economics of Data Center Operations

The economics of operating a data center are comprised of many items that factor into the total cost of ownership.

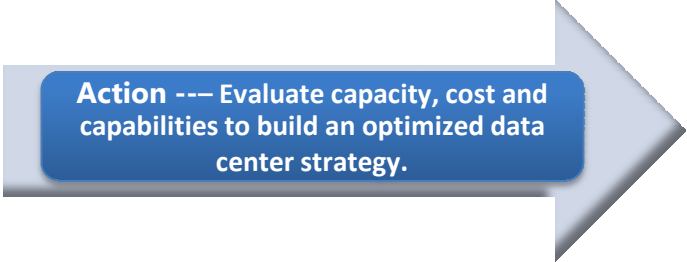
1. Resiliency: Whether building a data center or evaluating provider facilities, cost is derived from the level of redundant infrastructure built into it. The Uptime Institute data center tiers describe criteria to differentiate four classifications of site infrastructure topology based on increasing levels of redundant capacity components and distribution paths.
2. Down Time: The historical cost model for operating IT effectively has been the cost of down time. The data center reliability attribute is a key ingredient to how the data center should be designed and what the requirements of infrastructure are. The cost of down time is drastically different among the different types of businesses and the facility design considerations should reflect as much. The amount of risk a business is willing to assume in maintaining the uptime of their IT has a large impact on the cost of a data center.
3. Staffing is an often overlooked or underestimated factor in determining the cost of data center operations. In addition to IT staff, facilities staff ensure data center reliability and provision and maintain electrical and mechanical systems. Security staff requirements will vary depending on the size of the data center and individual needs of the business, but often require on-site personnel 24 hours a day and 365 days a year. If building a company data center, does the business have the experience in designing, building and operating it?
4. Financial considerations:
 - a. Site Selection: If you have the luxury of selecting a location throughout the U.S. for a data center site, incorporate local utility rates and tax incentives into the overall cost.
 - b. Cost Segregation: Research the use of audit estimating techniques to segregate or allocate costs to individual components of property (e.g., land, land improvements, buildings, equipment, furniture and fixtures, etc.).
 - c. Capital Recovery Factor: When evaluating the true capital cost of a data center, look at capital recovery factor, which is the ratio of a constant annuity to the present value of receiving that annuity for a given length of time.
 - d. Internal Rate of Return (IRR): What is the estimated IRR for the data center build project? IRR is an indicator that is commonly used to evaluate the desirability of investments or projects.
5. Timing: Consider the economics of technological obsolescence if building a data center. Weigh the costs of alignment with business and IT strategies against the risk of obtaining additional funding to increase power and cooling capacity to accommodate higher IT densities down the road.
6. Vertical Scalability: Scale is top of mind throughout most aspects of IT. The significance of scalability in the data center carries a different connotation and has a higher price tag if not considered properly. Vertical scalability means cloud computing-like elasticity capabilities incorporated into data center infrastructure and available floor space. It means turning up the dial

on power and cooling densities without disrupting the business. The gains of turning up that dial equate to agility in operations, adaptability to changing business needs and future cost avoidance in provisioning additional power and cooling to match increased requirements of IT.

An implicit value is derived from a data center strategy that is the coalescence of internal and financial perspectives. The misconception that going high-density in the rack equates to a higher cost is usually not founded with the realities of what efficient infrastructure is capable of. The new data center is equipped to handle and scale with high density servers and will ultimately save money through the power and cooling efficiencies gained. Examine all aspects of the economics of data center operations in order to understand the implications of risk assumption and true costs involved.

Road Map to the Perfect Strategy

With the context of knowing how data center technologies have advanced, who the stakeholders are inside the company and what best practices have yielded, a data center strategy starts to come in to focus.



Action --- Evaluate capacity, cost and capabilities to build an optimized data center strategy.

Own Versus Lease: Implications, Actions and Energy

More and more the CFO or finance department is gaining influence over data center decisions. With the traditional data center being a large investment and the utility bill going higher, financial executives are taking over the responsibility of the data center. Not to drive what technology is used for the company inside the data center, but to ensure sensible business decisions are made and that efficiencies are used where ever possible to reduce costs.

Once the costs involved in a data center are completely understood, the decision can be parsed into the implications of actions taken and the energy attributes of own versus lease.

IT Infrastructure Outsourcing

The implication of leasing a data center instead of owning one comes down to the benefits of cost savings, flexibility and scalability and finding a partner worthy of trust. Uncertainty is the major deterrent to outsourcing and providing for IT is nothing to be uncertain about.

Assess the layers of outsourcing available:

1. Profiles: As IT infrastructure evolved, the roles that outsourcing is able to perform have changed. It can be utilized in cloud computing (public, private, PaaS, or IaaS), colocation or wholesale data center services. If a major IT vendor such as IBM, HP or Dell is the preferred solution for IT needs, they are still able to service equipment in the provider's facility. Other outsourced or offloaded profiles pair with various cloud computing offerings and take away additional layers of complexity in managing applications and data center operations.
2. Network: Network connectivity is a major influence to any data center decision. If deciding where to build or expand, the ability to bring carriers in with the capacity needed now and in the future is something to research. If looking at colocation or wholesale data center services, the provider will most likely have a broad mix of carriers already available and it will be a line item addition to make cross connects to them. When comparing colocation or wholesale, research what carrier connections and capacity are available and match that with IT requirements.
3. Staffing: Colocation or wholesale data center services will all include (or should) staff to support and operate the facility a 24 hours a day, 7 days a week and 365 days a year. If comparing between own or lease, the comparable staffing levels should be built into the evaluation. An Uptime Institute tier 3 data center dictates that at minimum security personnel are to be on-site 24x7x365.

Wholesale Data Center

A small number of businesses are aware that a wholesale lease is possible and available in many locations. Wholesale data center services present a unique opportunity for a variety of businesses needs. The wholesale provider focuses on real estate, efficient infrastructure and energy, which deliver flexibility and scalability for their tenants within to make a long-term home for their infrastructure to grow. Wholesale is an attractive option for enterprises that are looking to quickly expand their footprint of existing company-owned facilities without the risk of financing and building another facility. It is a flexible choice that can deliver a greater value than colocation when looking to outsource. The flexibility of wholesale comes from the ability to add capacity on demand without incurring further costs or performing a data center move of equipment to some place that can handle the additional capacity. Typical parameters for what defines a wholesale data center customer vary, but follow a progression of where the service is a good fit. On the low end when a business approaches, or is forecasting a total IT load around 300 kilowatts they are a

candidate for wholesale. When a business approaches several dozen megawatts of IT power required they should explore the possibilities of building a dedicated data center.

No More Forklift Upgrades

Part of the pain of dealing with technology advances is the forklift upgrade – where either the owned or provider data center can no longer deliver the power and cooling needed, or the facility is out of space. After migrating from one solution to another to house IT the wholesale data center opportunity can offer a longer term solution, by applying vertical scalability. The wholesale tenant is buying space that can be expanded as demand dictates, and if the provider data center has engineered it, power can scale as well. Put another way – you can optimize almost every aspect of your IT, why not optimize your data center instead of planning and budgeting for yet another move. While the best data center is one that doesn't have to be built, many would amend that to be one that you don't have to move out of every 5 to 10 years.

The way we built data centers in the past limited our capability for quick expansion, future flexibility, highest efficiency, sustainability or lowest total cost.

Dean Nelson & Mike Lewis

Timing and Risk

A traditional data center is estimated to have a 10 to 15 year life. Traditionally, this has meant that every 10 to 15 years a large capital outlay was set aside to provide for an updated facility. The IT lifecycle can be 3 to 5 years, and the business strategy anywhere in between. Contracts for clouds, colocation or wholesale data center services can range from 3 to 5 years. The innovation and options available with data centers today presents a strategy synchronization opportunity to those that properly plan and take advantage of the technology and partner opportunities available.

The colocation or wholesale provider assumes the risk of technological obsolescence. They must build and maintain highly efficient and flexible facilities in order to foster an environment that benefits from innovative data center technologies. In doing so they are able to extend the useful life and future adaptability of the facility, which extends future revenue streams. The customer buys power, cooling and space today, and the assurance that they won't have to perform a forklift upgrade in the future.

The importance of selecting a vendor means balancing their ability to provide reliability and uptime, while having the business acumen to have flexible, scalable facilities and intelligent energy strategies. The role that data center design can play in a colocation or wholesale provider has a direct correlation on their service offering and ability to calm uncertainties about future scalability and reliability. As providers are evaluated, look into their ability to scale density and how prepared they are for future power scalability.



With the aim of delivering a more compelling business model for their tenants Vantage Data Centers set out to build a highly efficient data center campus in Santa Clara, California that incorporated numerous innovations and would take on the burden of time, capital and risk for their enterprise clients.

Instead of constructing a single building for a data center Vantage built an entire campus with the goal of enabling their customers to occupy areas of variable sizes and power densities, getting exactly what they want when the need it. Delivering vertical scalability means reducing overall operational costs while power densities increase. Some of the electrical, mechanical and physical innovations Vantage built in to their 18 acre campus include:

- ❖ Dual fed, 50 megawatt (MW) utility substation (37MW IT load)
- ❖ Dedicated (Vantage-owned) utility substation
- ❖ 2N electrical redundancy from substation to the server
- ❖ Air-side economizers
- ❖ (In the V2 building) elastic cooling: unique penthouse architecture.
- ❖ 1.2 PUE

Two essential factors framing an effective data center strategy

1. Scalability

What scalability truly means to a business has grown right along side of the advances in technology and data centers. The advent of virtualization impacted the context of scalability to shrink the server footprint, yet increase server density. An increase in IT demand, whether through Internet businesses serving millions of end users or enterprises with increasing dependencies within IT, meant trying to scale energy consumption. Growing energy demands put a focus on efficiency, alternative energy resources and the financial responsibility for powering the data center. Vertical scalability implies that adequate power is installed in the data center for today's forecast, and the ability to scale total facility power capacity exists. This makes a future upgrade possible without completely starting over with power equipment and utility feeds. Having a strong foundation like that built in to the design of the facility effectively doubles the life of the data center.

2. The Economics of the Data Center Decision

Cost underscores data center decisions and as such should be a focus during strategic planning.

Economics of Energy Efficiency

A predominant metric for measuring energy efficiency is Power Usage Effectiveness (PUE). Developed by The Green Grid in 2008, it calculates how much power is being consumed by IT in relationship with the power being consumed to cool and run the data center. For example a PUE of 2.5 indicates that the data center demand is two and one half times greater than the energy necessary to power the IT equipment. Measuring energy efficiency is important, because it is a large cost element of an overall energy model being applied to the data center. If building a data center or optimizing an existing one, energy prices and sources should be evaluated to determine the impact on capital expenditures up front and operational spending on the utility bill. For an outsourced solution, evaluate the provider's PUE their ability to provide energy efficient power and cooling systems.

Achieving a lower PUE does not always equate to a higher cost of data center operations. A net cost savings over time can be seen through implementation of high-efficiency infrastructure to power and cool a data center. Any level of power capacity can have efficiencies built in to how the data center components operate and save money. A larger power draw has a greater opportunity for savings.

Automating collection and analysis of data from data center infrastructure will give insights to operational teams and enable them to make intelligent decisions about energy usage. DCIM (Data Center Infrastructure Management) software has grown tremendously in recent years in response to the desire of data center managers to aggregate and analyze data across facility systems.

Financial Modeling

Most discussions investigating the true cost of own versus lease focus on the advantages of getting rid of capex (Capital Expenditures) in favor of opex (Operational Expenditures). What is really being referenced in these discussions is cash flow timing differences, and not the accounting definitions of capex and opex. If owning or building a data center there is a large up-front capital investment made that is then depreciated over its life. Alternatively if you receive data center services from a provider, you are paying a monthly amount and at the end of the term you can either leave the data center space or renew.

The true costs of what is involved in up-front purchases versus regular payments to a provider are not mentioned in the typical cost analysis. Consider how the cost of capital, return on capital, tax savings and other economic incentives can contribute to the financial analysis. Evaluate direct, indirect and overhead costs to get a true picture of the **total** cost of ownership.

Summary

Peter Drucker's definition of strategic planning

"Strategic planning is the continuous process of making present entrepreneurial (risk-taking) decisions systematically and with the greatest knowledge of their future; organizing systematically the efforts needed to carry out these decisions; and measuring the results of these decisions against the expectations through organized, systematic feedback."

Know how you can best align with business strategies, gather input from all stakeholders, determine the square feet of space and energy needs, and evaluate short and long term forecasts. Then look a level deeper to determine which data center model will provide the greatest financial benefit to your business and drive the greatest amount of agility in powering IT.

Owned	Leased
<ul style="list-style-type: none"> • What is the total cost of designing, building and maintaining a data center facility? • What are the long term costs of upgrading data center infrastructure based on forecasted IT needs? • Do you have the proper amount of qualified staff to support data center operations? • Develop a site selection plan for locating the data center in a safe area with connectivity and sufficient power. • Build power and cooling redundancies according to IT requirements. • Is there proper power vertical scalability available for forecasted future needs? 	<ul style="list-style-type: none"> • Compare colocation and wholesale offerings to determine the best match with your IT and business strategies for space, power and providers. • Investigate how forward-thinking the providers are with the technology and efficiencies employed at their facilities. • Is there proper power vertical scalability available for forecasted future needs?
How agile does the model choice make your data center strategy?	

An effective strategy is one that will merge business and technology objectives through a thorough planning process and careful analysis of technologies and options available. To the CEO and CFO the data center can be shown to be an agile asset that is able to strategically align with the business. The CIO plays a transitional role today, managing more business processes than technology.

Don't let best practices, or stories of super-efficient designs cloud (pun intended) a data center strategy that should align with the business. With a spotlight on resiliency, environmental and energy impact and most of all cost, the best strategy is an insightful, aligned plan to drive the business.

While due diligence and proper planning are required, as Peter Drucker said, the best way to predict the future is to create it. Take action, and build the perfect data center strategy.

Brought to you by



Glossary

Airside economizer: A mechanical device used to take outside air into the data center when the air temperature is at or below the set cooling point.

ASP: Application Service Provider (preceded Software as a Service).

Capacity Planning: An ITIL component that supports the optimum and cost-effective provision of IT services by helping organizations match their IT resources to business demands.

Capital Recovery Factor: See the equation for calculating capital recovery factor at http://www.wolframalpha.com/entities/calculators/capital_recovery_factor_calculator/tq/pq/5l/

Cloud Computing: A model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

COBIT: Control Objectives for Information and related Technology. A framework for IT management and IT Governance created by ISACA (Information Systems Audit and Control Association).

Colocation (retail): Placement of IT equipment in a provider's shared data center space. Targets smaller requirements for floor space and power requirements (typically less than 500 kW for total IT load). Colocation agreements can range from one to five year terms and may include other items such as caged space or managed services.

Container: Originated from the use of an ISO (International Standards Organization) intermodal shipping container being used to accommodate data center racks. Many variations of the 20 and 40 foot long containers exist and fulfill needs such as quick deployment and mobility. Power and cooling plants are also available in ISO shipping containers.

(hot or cold aisle) Containment: Placing isolation barriers at the front, back and top of aisles to eliminate air mixing.

DCIM: Data Center Infrastructure Management. The integration of IT and facility management disciplines to centralize monitoring, management and intelligent capacity planning of a data center's critical systems.

Disaster Recovery: The process, policies and procedures related to preparing for recovery or continuation of technology infrastructure critical to an organization after a natural or human-induced disaster. Disaster recovery is a component of a Business Continuity Plan (BCP) which focuses on aspects of a business functioning in the midst of disruptive events.

Evaporative cooling: Also known as swamp cooling, this is a way to cool air that takes advantage of the drop in temperature that occurs when water that is exposed to moving air begins to vaporize and change to gas.

Flux Capacitor: A device featured in the movie *Back to the Future* that, as described by Doc, made time travel possible.

HPC: High Performance Computing. The use of parallel processing for running advanced application programs efficiently, reliably and quickly. The term applies especially to systems that function above a teraflop or 10^{12} floating-point operations per second.

IaaS: Infrastructure as a Service. The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

IRR: The Internal Rate of Return on an investment project is the 'annualized effective compounded return rate' or discount rate that makes the net present value (NPV) of all cash flows (both positive and negative) from a particular investment equal to zero.

ISP: Internet Service Provider.

ISO/IEC 20000: the first international standard for an IT Service Management System. It is based on and supersedes BS 15000.

ITIL: Information Technology Infrastructure Library. A set of concepts and practices for Information Technology Services Management (ITSM), IT development and IT operations.

MSP: Managed Service Provider.

MicroGrid: a localized grouping of electricity generation, energy storage, and loads that normally operates connected to a traditional centralized grid (utility) connected to it - generators, consumers and those that do both – in order to efficiently deliver sustainable, economic and secure electricity supplies.

Modular: A conceptual definition for a method of deploying components of, or a complete data center. Modularity implies a pre-packaged, sometimes manufactured module of data center components that through its standardized and pre-defined parameters enable scalability and rapid delivery schedule.

PaaS: Platform as a Service. The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

Private cloud: The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise.

Public cloud: The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

PUE: Power Usage Effectiveness. A metric developed by The Green Grid (thegreengrid.org) to determine the energy efficiency of a data center. PUE is calculated by dividing the amount of power entering a data center by the power used to run the IT load within it.

REIT: Real Estate Investment Trust. A security that sells like a stock on the major exchanges and invests in real estate directly, either through properties or mortgages.

SaaS: Software as a Service. The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser (e.g., web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

UPS: Uninterruptible Power Supply. The UPS is intended to provide clean undisturbed stabilized AC voltage, within strict amplitude and frequency limits, to protect the IT load from any utility power disturbances and irregularities, including outages for a limited time dictated by the capacity of the battery bank. The term UPS refers generally to AC Static systems, other types include DC and Rotary UPS.

Water Economizer: A mechanical device used to take cold air and cool an exterior water tower. The water from the chilled tower is then used in the air conditioners inside the data center instead of mechanically-chilled water.

Wholesale data center service: Leasing of a fully-built data center space. The wholesale tenant gains more control over how their space is used for security and supporting infrastructure. Having the built-out space that is supplied with electrical and environmental infrastructure allows the tenant quick time-to-market in getting their IT infrastructure up and running.

References

- ❖ <http://blog.irvingwb.com/blog/2009/08/the-data-center-in-the-cambrian-age.html>
- ❖ <http://www.datacenterknowledge.com/archives/2008/09/22/new-from-microsoft-data-centers-intents/>
- ❖ <http://perspectives.mvdirona.com/2010/09/18/OverallDataCenterCosts.aspx>
- ❖ <http://www.morganclaypool.com/doi/abs/10.2200/S00193ED1V01Y200905CAC006>
- ❖ <http://blogs.technet.com/b/msdatacenters/archive/2011/04/19/the-disappearing-datacenter.aspx>
- ❖ <http://www.datacenterknowledge.com/archives/2010/07/29/hp-pods-speed-purdue-research-expansion/>
- ❖ <http://www.usgbc.org/LEED>
- ❖ <http://opencompute.org/>
- ❖ <http://www.datacenterknowledge.com/archives/2011/04/07/facebook-unveils-custom-servers-facility-design/>
- ❖ <http://nist.gov>
- ❖ <http://blog.uptimeinstitute.com/2011/05/large-data-center-operators-moving-to-colo-third-party-providers/>