TAOS

Where Your Servers Call Home: Data Center, Computer Room or Telco Closet?
Part 2 in a 4 Part Series on Data Centers

By Bruce Whetstone, Executive IT Management Consultant with Taos' Office of the CIO Practice

"I wouldn't call it a data center," remarked the Director of IT. "At this point, I think it's better to refer to it as a computer room." We hear similar comments from IT managers regularly. What is the distinction between a computer room and a data center? The terminology contributes to the confusion, as a high-end computer room would be classified as a Tier 1 data center. Large companies typically have a variety of locations for servers, including server rooms, development labs, computer rooms, data centers, and yes, unfortunately, sometimes servers get stuffed into the communications closet, or under someone's desk.

The data center tiering structure that Taos most often uses is cited by The Uptime Institute and is consistent with publications from the Telecommunications Industry Association. Tier 1 data centers have basic site infrastructure support, may or may not have a raised floor, UPS or backup generator, and have a single path for power and cooling distribution. As you move up the scale to Tiers 2, 3 and 4, the redundancy of supporting systems in the data center increases, contributing to higher availability and reliability (and of course, increased cost as well). Note that the tier numbering for data centers, where Tier 4 is the highest, is the opposite of network provider tiering, where Tier 1 is the highest (Network Tier Wiki).

I tend to think of server rooms, computer labs and computer rooms, as part of the company's primary or satellite facilities, and data centers as separate buildings specifically and exclusively built for computer and network equipment (either for a single company, sourced to a third party or shared via a third party, like co-location centers). To simplify the discussion in this article, I will refer to all types of computer space as simply that.

I am a big proponent for putting computer equipment in space designed to accommodate computer equipment (seems logical). Why? Because the worst enemy of semiconductors is heat. Equipment will break down more frequently when subjected to overheating. While the obvious issues are effective heat dissipation, which is directly related to cooling capacity in the space housing the equipment, the not so obvious is dust accumulation. Dust acts as an insulator, and accumulates inside the equipment on the surface of the boards and electronic components. Heat gets trapped inside the semiconductors, contributing to eventual breakdown and reliability issues. If the design of the space was not originally for supporting computer equipment, the facilities team typically inherits the problems with insufficient power and cooling capacity. How many of you have seen a computer space with the door propped open, and a consumer-grade window box fan conscripted to help cool the room?

Small companies typically place their desktops and laptops at a workstation, and their servers in a server room (computer space typically in a small to medium-sized room, outfitted with racks, cable routing trays, extra electrical and air condition/air handling). Fair enough. As companies grow, the number of servers proliferates, as do the types of spaces in which they can be housed, and things rapidly get very complicated. Server location depends a lot upon a company's stage of growth, rate of growth and process maturity. Some of the major issues we encounter in helping clients with their computer space issues include:

> ***Fast growing companies*** – These companies can't acquire or build out computer space quickly enough to keep up with demand. Lots of servers end up in inappropriate environments, including workstations (under desk) and telecommunications closets. Servers at workstations become a headache for facilities managers, as they can draw a lot of power and generate a lot of heat in areas that weren't designed to handle the load. Security is also an issue, as the servers are typically in uncontrolled environments, and easily 'move around,' making them hard to track, maintain and secure. Servers in communication closets are the worst, as the rooms tend to be small to begin with and often have inadequate power, UPS and cooling to support the additional equipment.
>
> ***Start-ups*** – Working on a shoestring budget, these companies don't have the money to invest in infrastructure. They put servers in any corner available, including unsuitable office areas. As the company starts to grow, the infrastructure has to be revamped, or as we commonly observe, the start-up mentality and processes carry over, resulting in reliability issues, operational support issues, and server "sprawl."

contactus@taos.com www.taos.com (888) 826-7686

Large companies – These organizations have a very bad habit of filling all space available for the retention of equipment, files – you name it – and at the same time, are seldom disciplined in their approach to retiring, recycling and disposing of unneeded information technology assets. The mentality is, "We might need it someday, so until someone says we have to get rid of it, let's keep it for safety's sake." Unfortunately, the cost to maintain more systems than needed (including very high maintenance costs for legacy systems and equipment) can be extreme.

Companies of all sizes – With little or no process rigor or maturity around server acquisition, procurement and provisioning, development and project teams tend to focus on the cost of the server, but not the facilities, power, software licensing and operations resources that inevitably come with it. When a new project arises, they go out and buy new servers, and bury the acquisition cost in the project budget. If servers came with a total cost of ownership (TCO) price tag, instead of a hardware-cost only price tag, more effort might be invested in server rationalization. This would force organizations to answer the question, "What are other options for supporting our computing needs on this project, other than acquiring new?" Insufficient process and review of project hardware plans mean that organizations rarely investigate whether other teams have servers that can be reused, or whether they could use virtualization technology to increase server utilization. The ultimate result is a dramatically higher cost to the business.

A major planning obstacle for most companies is an inability to forecast their needs for computing capacity. If a product or service "takes off", they can't build fast enough to keep up with demand. If the economy screeches to a halt (recent history, as the prime example), the prior business forecasts will be off target, resulting in excess expense for facilities. However, implementing forecasting processes in conjunction with tight processes around server rationalization can make a major impact on data center expenses. This is arguably the **best** way to control expenses, because it helps gauge and throttle demand "upstream," before the problem gets out of control.

Another way to control computer space costs is to stratify the needs, and come up with distinct, but standard, service offerings. Depending upon the business' needs for uptime, availability and customer satisfaction, development systems can go in a lower-cost (Tier 1 type) computer room and production systems can be moved to a Tier 3 or Tier 4 data center. Unfortunately, no two situations are exactly alike, so what might work for the company down the street, may or may not work for you. In general, there will be trade-offs around expense and availability/reliability. For example:

High availability customer-facing systems – These systems need to be in a high availability data center and very often require deliberate connection to more than one data communications network provider for redundancy, best performance and customer experience.

High availability corporate systems – Since they are not external customer-facing, but still critical to the business, these systems need to be placed in a high-availability data center, but not necessarily a Tier 4, carrier neutral, data center, depending upon the business' tolerance for disaster recovery and business continuity (return to operation time).

General corporate systems – These can be placed in a lower tier computer space to save money, but caution should be applied to ask the business what systems they consider critical. There are a rising number of companies that now consider their e-mail systems, voice over IP (VoIP) phone systems and file sharing systems, to be mission critical.

Development systems – These can also be placed in lower tier computer space to save on costs. Depending upon how critical some development systems are to the core business (e.g., the software business), the tier level for housing these systems may need to be bumped up.

While most companies or organizations will relate to the above scenarios, there are always other situations and considerations to take into account in making data center decisions. Some of these include:

- Stratifying the type of workload, availability needs and risk tolerance (for outages) can help ensure that servers are placed in computer space that is "fit for use," but at the same time, is as cost-efficient as possible.

contactus@taos.com www.taos.com (888) 826-7686

- In companies with an insufficient number of units to justify creating different spaces for several strata, it may be more cost-effective to keep mission critical systems in a high-availability data center and everything else in a less expensive computer space (thereby reducing complexity by having only two choices, with little impact on overall support cost).
- The more locations that have to be linked together across geography, the higher the expenses to the business for data communications (circuits, equipment and support). So, at some point, consolidation and aggregation of computer space makes a lot of sense when looking at the overall company support costs (not just facility by facility).

If all of this is beginning to sound very complicated, it is. And there is no substitute for experience to handle the more difficult situations and design problems. But a basic understanding of computer space, and business trade-offs relative to computer space, is not beyond the reach of anyone interested and motivated to sit and talk with an engineer, designer or consultant that does this work every day.

Bruce Whetstone is an Executive IT Management Consultant with Taos' Office of the CIO Practice. Contact him at bwhetstone@taos.com