

Data Center News:

DATA CENTER MANAGEMENT ADVISORY NEWSLETTER

Four mistakes to avoid during data center consolidation

By Amichai Lesser, Contributor
02 Sept 2005 | SearchDataCenter.com

What you will learn from this tip: What not to do when embarking on a server or data center consolidation project.

Data center moves and consolidations can deliver cost savings, enhanced business continuity, and optimized service management. But moving servers away from your end-users. ((Content component not found.)) That's because users who were local to servers become remote users, and back end servers may wind up remote from each other during interim relocation stages.

The impact of this physical displacement should not be underestimated – but it often is. In fact, if you don't adequately understand the issues that arise when you put more physical distance between users and servers, you can set yourself up for serious pain and potential failure. Here are four specific mistakes you should be particularly careful to avoid:

1. Confusing network latency with application latency

When you move servers away from users, you introduce network latency. Physical distance between users and servers causes a delay in the signal between the two. But adding 100 milliseconds of network delay doesn't mean that your application response times will only increase by 100 milliseconds. On the contrary, most applications require many back-and-forth interactions between user and server (often referred to as application "turns") to perform even the most basic tasks. Thus, the addition of just 50 milliseconds of network delay can cause an action that only took 3 seconds to complete locally to take a full 30 seconds to complete after a server move.

This network-related latency is usually regarded as the network manager's problem, even though application design (including the number of "turns" it requires) is the real issue. But network managers can't change the speed of light, or make Tokyo closer to New York. So it doesn't make sense to lay the problem on them. In fact, because application design issues are often responsible for poor response times after a server move, additional investments in the network may be of little use whatsoever.

2. Failing to realize how network latency impacts server performance and scalability

Network latency can substantially degrade server performance and scalability. Servers allocate resources to each concurrent client session. Local clients complete these sessions quickly, because their application turns are subject to minimal network-related delay. Remote sessions, on the other hand, take much longer to complete because each application turn takes so much longer.

It's important to note that servers lock resources up for the duration of a process, freeing them only after the process is completed. Thus, when remote users communicate with a server, they keep its resources busy for a longer period of time. This prevents the server from releasing its resources for use by other clients – severely limiting its performance and ability to scale. That's why IT organizations have to consider the possibility that network latency will require them to invest in additional server infrastructure.

3. Putting distance between servers – even temporarily – can crush performance

It can take weeks or months to move the dozens or hundreds of servers in a data center to their new location. During this process, some systems will operate from their original location while others operate from the new location. The impact of this server separation on application performance can be even more dramatic and unexpected than the introduction of latency between users and servers, because computing processes are almost never designed to accommodate significant inter-server latency.

Any IT organization planning a data center move must therefore ask a variety of questions. What happens when servers with critical inter-dependencies are temporarily separated? Which servers must be moved with other servers? When should Active Directory servers be moved? Which servers will need to be replicated for the duration of the move?

4. Not dealing with users' performance expectations until after the move

It's critical to address users' service level expectations up front. If you wait until after the move and tell users they have to live with what you can deliver, you're setting yourself up for a battle. But if you can get buy-in beforehand as part of the planning process, you can avoid such hassles and ensure that no one has unrealistic expectations.

Sometimes, it doesn't make sense to set a post-relocation Service Level Objective (SLO) that is identical to the SLO before the move. If it originally took a local user three seconds to execute a task, it is very unlikely that the task will take the same amount of time after servers are moved across the country. So an SLO of seven seconds, for example, may be more reasonable.

To achieve this pre-relocation acceptance, IT must be able to predict and simulate post-relocation performance. These predictive and simulation capabilities enable IT to set up "acceptance environments" where users can experience post-relocation performance before

the move is actually executed.

In fact, IT organizations can avoid all of these mistakes. But to do so, they must take a disciplined approach to planning that leverages the expertise of the application team, systems managers and network architects. The ability to create virtual models of both the pre- and post-relocation enterprise environment – as well as all transitional phases – can be particularly useful for anticipating and addressing any potential application performance problems that may result from the movement of servers from one location to another. All participants in the planning process, including business users, need concrete information about how network issues will impact application performance as a result of the data center move.

So if you're planning a data center consolidation or other type of server move, consider investing in simulation technology that allows you to experiment with alternative scenarios and determine in advance what will work and what won't. It's a great way to ensure that your business reaps the full benefits of the move – without suffering any of its potentially disastrous consequences.

Amichai Lesser is the director of product marketing at Shunra Software, a company that delivers award-winning solutions that recreate a replica of any production network environment for testing the functionality, robustness, performance and scalability of applications and services - before rollout. Amichai is responsible for product marketing, market analysis, and field marketing programs, and has extensive experience in real-time engineering, performance management and security. He regularly presents at industry conferences, seminars and events. Amichai can be contacted at amichai.less@shunra.com. For more on Shunra, see www.shunra.com.