

Who Owns Application Performance?

Net Forecasts – Peter J. Sevcik
BCR Volume 36, Number 7
July 2006

I visited several enterprises recently and was struck by the abysmal state of application performance management. Let me describe two companies.

The first is an international manufacturer where I visited the network group to discuss the performance of a mission-critical application. In fact, I was being introduced to the application users in the office by the folks in the network group. The reason I was at this office is because the application users had been complaining about response time performance.

I always ask a lot of questions about the application, network path to the servers, how they are accessed, and what the users actually do with the application on a daily basis. The network manager replied with a single sentence of about 10 words and then said, “And that is all I know about the application, period.” We then had to wait for a complex diplomatic process within the organization to have the network folks introduce me to the users. The network folks had actually never talked to these users before. It turns out that the users were only two cubicle rows away from the network group!

The second company is a serious research organization with thousands of employees scattered around the globe. In this case, I was talking to an application manager who knows, based upon user reports and personal observation, that some of the users see very poor performance. When I explained some of the performance acceleration technologies that may help the situation, she said, “Why did my network staff not tell me about these options?” In fact, the network group had constantly reported great network performance and simply washed their hands of the problem.

What do these two stories tell us? I am sure you have heard the saying, “I work in a mushroom factory; they keep me in the dark and feed me sh**.” I bet the network managers in both of these companies say that line from-time to time. Here is my reply, “You like to hide in your dark corner and don’t want to be bothered with what is really going on over your network.” This lowest common

denominator approach to application management is called silo management.

Default Performance Management Is Silo Management

When an enterprise buys a new technology, they often purchase an element manager to make the system work. Over time, the element managers are grouped so they align with the organization structure of the company. The result is shown in Figure 1, which becomes the default enterprise management strategy. Oh yes, many enterprises will work hard on defining strategy, policy, procedures, training, help desk, and even SLAs for each silo so the systems work like finely oiled machines.

There are often many silos. For example, the single WAN silo in Figure 1 can easily become separate silos for the private network, IP-VPN service, and the public Internet. I recently saw a big bank operating different management systems for Web, midrange, and mainframe servers. These basic management investments are needed to diagnose and repair broken parts of any single silo, given that you can find the break. And therein lies problem number one: Which silo is broken?

So some enterprises go farther to integrate the information from each silo into a manager of managers, business analytics engines, event aggregation systems (alarms), event correlation (fewer alarms), automated threshold algorithms (just the right number of alarms), and of course a huge repository of data most of which is never read.

Many vendors work hard at trying to improve upon this state of affairs by adding end-to-end or application-centric views to the data gathered by each silo. The thinking goes that if you can just get all of the data from all of the silos into one place and correlate it by some means (typically by time), then you will see an unusual pattern that indicates an anomaly.

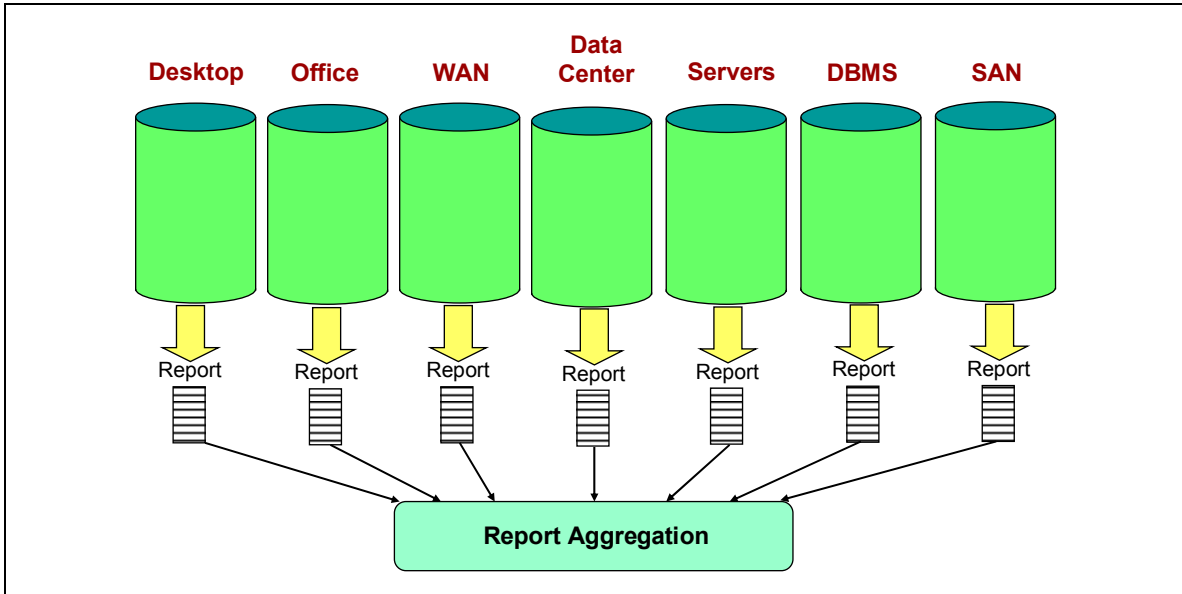


Figure 1 – Silo Management

Why Silo Management Fails

None of the extensions and improvements to the silo data actually gets to the heart of the problem, which is to figure out why a real end-user is experiencing a problem. In fact, the added cleverness is often fuel for the assertion, “There is no problem in my silo.” The person who just made that remark would probably add the following under his/her breath, “Now leave me alone so I can go back to my dark corner.”

One would think that senior management would not stand for such an attitude. The fact is that they do, and here is why: Management thinks it understands how the system works. You wouldn't be a manager unless you convinced yourself and your boss of that fact. They “know” how user tasks operate inside the silos and are then passed from silo to silo. They think the reports they are reading are views into each silo, as shown in the top of Figure 2.

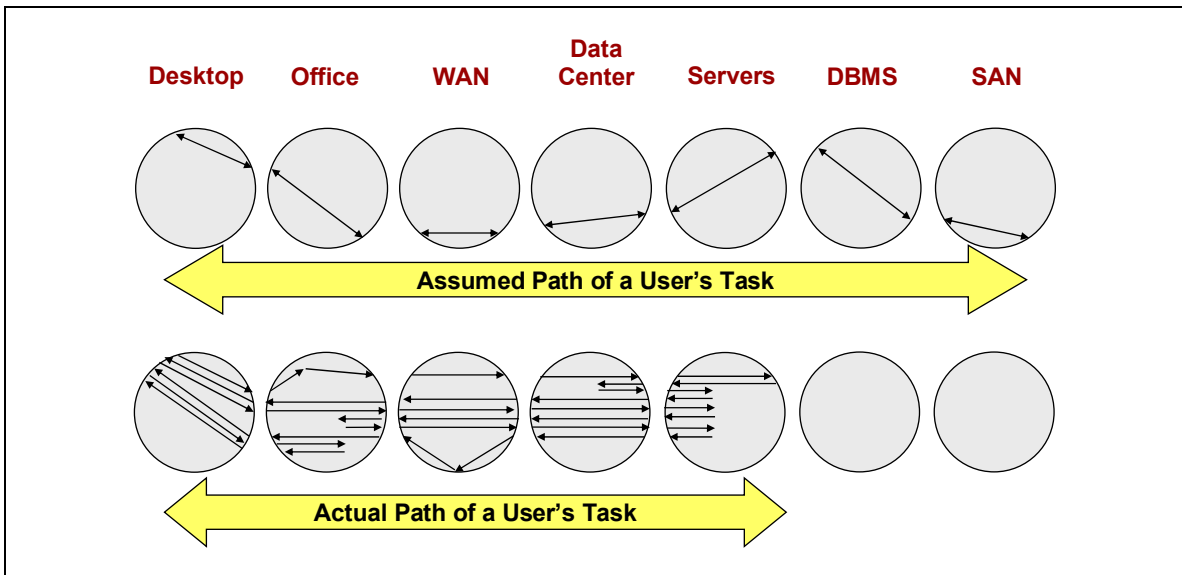


Figure 2 – Why Silo Performance Does Not Equal User Performance

Figure 2 is a view down into each silo of Figure 1. The top view is the assumed path each user's task takes. There is some work to be performed to make sure you align the circles just right, so that you are looking at the proper subsets of each circle and the interfaces between circles. If the observations across all circles are that all is well, then the complaining user must be hallucinating.

But the humans watching all the silos really do not know the path of a user task. The path changes over time as systems change and get upgraded. The user does things that were not foreseen. Complete subsystems move to new locations. Security and acceleration devices add proxies, filters, and even change how the application operates. The result is a user task path as shown in the bottom of Figure 2.

Notice that in the example, the DBMS and SAN systems were never involved, so the reports on those silos were wasted for this problem and may have been misleading. Further notice that the path crisscrosses some silos many times. There is no way to see and understand the real behavior of a user task without following the task explicitly.

You can't find an illegal activity by asking each bank to tell you the total funds on deposit each day. You must follow the money account-by-account and transaction-by-transaction.

I am not proposing that enterprises spend the huge resources required to track the details of every user task. Instead they need to watch the resulting application user response time as the key indicator of overall application performance. If the task response time starts to degrade, then detailed investigations should be performed on the paths those tasks take.

Despite all the talk about business oriented performance metrics for IT (dashboards, business logic, etc) and full end-to-end views of infrastructure (very sophisticated silo integration) few tools truly look at real-time user flows.

Application Performance Manager Needed

Proper performance management is delivered in two parts:

- 1.) Ongoing end-user performance measurement, reporting and tracking followed by
- 2.) flow-based diagnostics to find and solve specific problems.

But most importantly, there is no group within an enterprise that has this responsibility. In my examples above, there was no person whose job was to cut through the silos following a real user's task path.

I recently found a small group within a bank that does have this job. They have excellent tools and methods for finding the real performance problem. They operate the second half of the performance management strategy: flow-based diagnosis. The first half is performed by a business manager listening to his users (no formal measurement and reporting exists) and then going to each silo manager several times with no resolution. After hearing "not in my silo" often enough, he hires the in-house team of cross-silo experts to really figure out the problem.

Enterprises must move from these ad-hoc approaches to a formal application performance management process that is owned and operated by a formal group within the company. We won't really address application performance until we put someone in charge of the problem and give them the resources required to get the job done. This will take more than funding. It will also require new relationships among groups within a company, and new management procedures.

Proper performance will be realized when a company can point to a person in charge and read the ongoing reports from that person documenting real end-user performance. In addition, employees who hide behind a narrow job description and show no initiative to fix problems should be re-educated or shown the door. Companies should not let anyone become a self-made mushroom. Anything less is insufficient for a competent and competitive enterprise.

Peter Sevcik is president of NetForecast and is a leading authority on Internet traffic, performance and technology. Peter has contributed to the design of more than 100 networks, including the Internet, and holds the patent on application response-time prediction. He can be reached at peter@netforecast.com.

NetForecast is an internationally recognized engineering consulting company that assess and improves voice, video and data application performance. We help enterprises align application performance with business needs using a process based on the Apdex standard. NetForecast also advises vendors about customer requirements, technology issues, and the business value of their products. Visit our library of educational reports and articles about performance engineering.
www.netforecast.com

