

Publication: Data Management Review

Author: Douglas Hackney

Title: Ensure A Good Beginning By Forcing A Bad Ending

[begin copy]

Good beginnings are critical to the success of new products and services, be they Data Warehouses, or magazine columns. I'm going to cover a topic that's critical to having a successful beginning for your data warehouse project, but first I'll give you an overview of what to look for here in coming issues.

This column will focus on the challenges and solutions that face those of us tasked with the real world implementation and sustenance of Data Warehouses and their associated Information Delivery Systems. My experience and audience feedback at numerous presentations have made it clear that there is a dearth of information and advice available concerning these issues.

The column will address issues such as: why data warehouses fail - from the end user point of view, sustenance issues for Data Warehouses and Information Delivery Systems, end user meta data and how to provide it, effective communication methods with end users, sustainable data warehouse team management and structure, mid and long term data warehouse management issues, critical end user access tool features, end user access methodologies and techniques, delivering ROI analysis on the desktop, finding and delivering common mandates with users, and effectively marketing the Data Warehouse.

Readers will learn ways to move beyond the discussion of tools and architectures to the design, packaging, delivery and implementation of effective and sustainable Data Warehouses and Information Delivery Systems.

One critical sustenance issue is a disaster recovery plan for your Data Warehouse. If your experience is typical, when it comes time for roll out you'll be behind plan, and one of the first things that will slip off the agenda is a test of the disaster recovery plan.

Considering the common characteristics of the many sites I've visited, your project will need a disaster recovery plan at least as detailed, and better rehearsed than any other mission critical system of the business. Typically, the data warehouse is one of the business's first forays into client server architecture. It is likely that your team's level of Server OS and RDBMS expertise is relatively low. I find it fairly common for the Server OS Admin and the RDBMS Admin to have had little to no previous experience with their relative products prior to the Data Warehouse project. This guarantees that your key players are going up a steep learning curve at the same time you're struggling to understand, design, and implement a new methodology of decision support in the enterprise.

In these low experience situations a common recovery plan is "We've installed the RDBMS several times as part of the development of the project, along with the OS a couple of times. We've got a list of the things to do from the RDBMS Vendor/OS Vendor/Consultant in case something bad happens." Don't be seduced by this siren song of readiness for the worst case. As you draw near to roll out you'll be under severe pressure from management, users, and your team to get the Data Warehouse on line. The last thing you'll want to think about is taking the server down to rehearse a recovery. In many cases, you literally won't have a minute to spare, since you'll have extracts, loads, and summaries scheduled up to the millisecond you turn on the users.

As dark a prospect as it may seem, you cannot afford to not take the server completely down and completely rebuild it. As you know, circumstance can sometimes be cruel, and if you don't rehearse a recovery you'll practically guarantee an outage fatefully synchronized with month end reporting.

Do not believe any assurance of readiness until the team has passed the following test:

- completely take down the server, including the OS
- re-install and re-configure the OS
- re-stripe the drives
- re-install and re-configure the RDBMS, monitors and middleware
- re-load and re-index the data

The first time the team tries this they will find many, many errors and gaps in those lists left behind by the vendors and consultants. They will probably have to re-do many steps. Don't be satisfied until the whole process happens from beginning to end in one continuous, contiguous process. Do not be satisfied with a report of "we found a bunch of stuff wrong with the lists, we fixed it along the way and we're ready now." Make sure that the entire process is fully documented, and can be executed even if a key player is not available.

For best results, arrange to take the server down again, unexpectedly, just like when it drops out from under your Data Warehouse. This is the true acid test, and will guarantee that your team, and your Data Warehouse are ready for a good beginning with your enterprise and its users.

[end copy]