

JACKIE'S Forum

High Availability + Data Warehousing



Jackie Jansen

“High Availability and Data Warehousing, a match made in heaven”. If that is a bit of a stretch, how about: “two technologies that were meant for each other”. On a more serious note, from a business point of view these two solution areas complement each other extremely well.

I previously addressed this subject in a column a few years ago. Based on the reprints that I am continually asked for, this topic is still very pertinent today and deserves some attention again.

If you have a mission critical application that drives the requirement for a high availability system, then justifying a data warehouse just became a whole lot simpler. If you would like the security of a high availability server, then data warehousing may give you the extra ROI you require to cost justify an additional system.

Let's start by looking at this from a high availability point of view. An obvious requirement for high availability is that you replicate your mission critical data from your production iSeries onto a backup server. This backup server has to be available at a moment's notice to take over should your primary production server experience a failure. What you don't want, is to have your backup server sitting idle. A data warehouse or even a simple query and reporting system is usually not considered to be a mission critical application. It can be shut down in an emergency and is an excellent choice to utilize the additional CPU cycles normally available.

Does having your data replicated on a second server mean that you have a data warehouse? The answer to that is a resounding NO. You will have a system that contains the source data that you

need for your data warehouse and presumably quite a few available CPU cycles. You will also have a system where you can run your operational queries and reports without impacting your production system.

Now from a hardware perspective is this a “free” data warehouse? The answer is: typically not. Although you will usually have enough CPU cycles to run your data warehouse you will still need to add additional disk space. As mentioned earlier, what you now have on the second server is the source data still formatted for your ERP system. One of the characteristics that differentiates a data warehouse or an end-user query/reporting system from your operational

nightly batch window. You can use your high availability server to perform your end of day reporting and potentially your nightly backups depending on physical location and tape access.

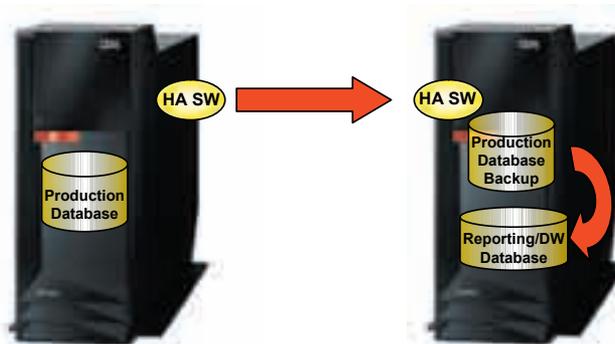
Many of the iSeries excellent high availability software vendors have promoted using the backup copy of your production data to offload batch reporting from the production system. Whether you just want a simple reporting system or a data mart or a more complete data warehouse, you can use your second server to accomplish all this while it is receiving transactions for high availability data recovery.

On Tuesday March 29 at noon Eastern Time, I will be delving into this topic in much more detail during a Webcast sponsored by two IBM Business Partners. In addition to myself there will be various customers

discussing their experiences with utilizing one server for both High Availability and Business Intelligence. You can enroll at www.mimix.com/webcasts. I look forward to seeing you there. **Two technologies meant for each other – High Availability and Data Warehousing!** 

Jackie Jansen is a Senior Consulting IT Specialist. She currently works in the IBM Americas Advanced Technical Support Solutions Centre. Jackie is a frequent speaker at iSeries Technical Conferences and User Group meetings. Contact her at jjansen@ca.ibm.com.

High Availability / Business Intelligence



system is the database design. The design methodology used for a transaction processing system is very different from the design required for a large, potentially unstructured query workload to occur efficiently. You will still need to transform and load your data but you will no longer need to impact your production system. This can all be done within the one server, directly from the replicated version of the data.

Another advantage to using the high availability server as both the source and target system for your data warehouse is the additional time it buys you in your