Getting your Netezza System Back in Minutes After a Disaster

Author: Roy Hammett

When talking of Disaster Recovery people often focus on the cost of doing it. By definition, you should have two of everything, possibly in different locations. Logistically that seems expensive, and many will consider it unnecessary because if you don't have a disaster or have some sort of major outage then it seems like a horrific waste of resources in time money and effort.

Realistically, though, the starting point of any kind of contingency planning or disaster recovery discussion should really be the cost of not doing it. Moreover, to focus on the cost of building that resiliency into your business is the wrong way round. Instead, the impact, and ipso facto the cost of not having a Disaster Recovery capability will make the cost of doing it seem irrelevant and relatively meaningless.

In some ways it's like the cost of not addressing global warming. The consequences of global warming are so horrific to the human race that the cost of trying to solve global warming has to be worth doing because the alternative is extinction. And although global warming is an extreme example it is a relevant analogy to the costs of not having a resilient infrastructure because it could mean the extinction of the business, depending on how critical the system is that's not covered by such a disaster recovery planning scenario.

All too often businesses underestimate the mean time to recovery (MTTR) because even for customers that have a DR system, many will never have tried cutting over to it as a planning exercise to prove that it is fit for purpose and take the time to assess what the impact would be of having to do that.

If a business has never fully tested the MTTR, they may not realise that for a Netezza system this can be measured in many days or even weeks. Losing a data warehouse for a day or two may be acceptable to the business but losing complete visibility of how your business is doing for an extended period would be unacceptable not only to business, but also the board and shareholders. Furthermore, if the outage happens to coincide with a financial period end, not being able to produce reports and file returns because your data warehouse his out of action may well have legal ramifications. Not knowing your MTTR means that you may well have a false sense of security even if you have a second system that is sitting idle.

Why does it take so long to recover from a disaster?

Recovering a Netezza system can be very time consuming because even if you can do a full back-up in a relatively short period of time (e.g. 8 hours or so) it's extremely unlikely that your systems or your primary site will go down immediately after doing a full back-up. It would be extremely fortuitous if that were to happen, but it would be extremely unlikely. Typically, businesses do a full back-up once a week, daily incremental back-ups and perhaps an accumulative back-up in the middle of the week. To recover your system, you will need to first apply the most recent full back-up, then all the differentials and deltas that have occurred since this back-up. Even then you then you may be missing some data because after the daily incremental back-up, changes may have occurred throughout the next business day. To further complicate matters, locating, physically moving and then restoring back-ups to your disaster recovery location can be incredibly slow – either because the of the bandwidth of your network, or because the secondary infrastructure does not have the same number of physical back-up devices that were used to back up the primary. Netezza is capable of backing up to 16 devices simultaneously, but if the business has not invested in the requisite number of back-up devices on its disaster recovery system, the restore is going to be incredibly slow. Don't forget that whilst the recovery process is happening, the system is dead in the water and that's probably not the best time to be admitting to a business that is already complaining that they can't see last week's sales or produce any financial reports that it is going to take weeks to recover the system. So, whatever else you do, your DR strategy should include a full disaster recovery test of your Netezza system to see whether the observed MTTR is acceptable to the business.

What are the real costs of losing your Netezza system to an extended outage?

Disasters happen and there are many recent headlined examples of extended outages affecting businesses because of natural disasters or hardware failure. If you've not done so already take a stab at calculating just the wage bill of IT and sales and

marketing professionals whose work would be interrupted by an extended outage. If, for example, you have 200 users of your data warehouse appliance earning an annual salary of 100k your daily wage bill is about $54k. Add to that the work they do which may generate millions of dollars of annual revenue from direct marketing campaigns or pricing price optimization.

If your Netezza system is mission critical for providing services to other organizations, you may also be facing financial penalties for not meeting SLA's. Government penalties for any breach of regulatory requirements such as GDPR, and litigation settlements are very real financial drains. And for companies dealing in physical products, depleted inventory is a significant risk. It is easy to see, therefore, how the true cost of a long outage can very quickly turn into millions. Here's an interesting link for estimating those costs.

The cost of not having a disaster recovery process in place – the loss of business; loss of reputation; loss of good will; financial penalties and potential legal action by various parties is what needs to be factored into the make case for having a mature reliable repeatable disaster recovery capability.

Why having a backup server may not be enough

As we already mentioned there are a number of Netezza users out there with disaster recovery systems thinking they're safe and ready for a disaster to occur, but when it does occur they suddenly realize it takes a lot longer than they had expected.

So just having another server in a different data centre isn't enough - that's our point. You need to think about the mean time to recovery (MTTR) and the recovery point objective (RPO). The latter means the maximum amount of data – as measured by time – that can be lost after a recovery from a disaster, failure, or comparable event that is acceptable to an organization.

Here's a solution for minimising your MTTR and RPO

If you want to be able to minimise your MTTR and your RPO we have the ideal solution for you, and recommend using our Smart Management Frameworks SmartSafe module.

SmartSafe keeps your primary and secondary sites synchronized so the MTTR is extremely low and can be measured in minutes. Furthermore, the frequency of running the synchronisation process is determined by the customer, meaning that the RPO, being determined by the time the most recent synchronisation occurred, can also be reduced to just a few minutes, depending on the volatility of your current system.

Compare this to the scenario that we described earlier where you have to transfer all your back-ups over the network to your secondary site and kick off the full restore which could take days, followed by all the incremental restores which could take more days, and then apply any missing transactions that you couldn't restore.

With SmartSafe there's none of that because the data is always up to date by design. So, in the event of a disaster, only three things need to be done to get the system fully back up and running. Firstly, the secondary site is configured to become the master in SMF. Secondly the secondary database is unlocked so the changes can be applied to it directly. Finally, network administrators at the customer end updates their DNS settings to point the primary system name to the secondary server's public IP address.

By just by doing those 3 simple tasks, voila! all the users will automatically connect to the secondary server, all the ETL processes will connect to the secondary server rather than the primary and the users won't even know it because as far as they are concerned it's the same machine - it just happens to be in a different place now.

Doubling the performance of your Netezza queries

If you are using SmartSafe to keep your Netezza systems synchronized, you can load balance users and workloads across both the primary and the secondary site. Admittedly the secondary site is read only, but as long as it is of no consequence to those users that it's maybe 10 minutes out of date, the

effects of doing so will improve performance for both sets of users. In fact, physically relocating where queries are run, and reports are produced, will result in a perceived doubling of performance. So, this is one of the potential tangible benefits of deploying SMF to provide disaster recovery that is above and beyond the ability to recover from a disaster. It enables what would otherwise be an idle machine that is waiting for a disaster to be put to a good use. Furthermore, because SMF is a multi-master bi-directional replication system it means that you can update and maintain one set of databases on the primary system that you replicate to the secondary, and update and maintain another set of databases on the secondary system that you replicate to the primary, with users distributed across the two systems instead of just the one.

There are more use case scenarios in our SmartSafe FAQ's, and if you would like to discuss further with us, feel free to contact us here.

smartassociates

Smart Associates (Aotearoa) Ltd	Smart Associates Ltd	Smart Associates ApS
203/11 Vinegar Lane Grey Lynne Auckland 1021 New Zealand	Valley View, The Old Quarry Haslemere Surrey GU27 3SS United Kingdom	% 360 Law Firm Lautrupsgade 7, 3. tv DK-2100 København Ø Copenhagen, Denmark
T: +64 (9) 415-8120	T: +44 (208) 133-6008	T: +45 36 98 71 11