

VMdamentals Blog

VMware Technical Deepdive - By Erik Zandboer

Product Review esXpress

August 2009

Erik Zandboer

<http://erikzandboer.wordpress.com/>

For a long time now I have been a fan of PHDs esXpress. It is still the only VMware backup solution I know that scales, has no single points of failure and works reliable with VMware snapshots. The solution has always been "other than others": At first it appears to be a really weird piece of software that creates its own appliances to perform its backups. Once you get to know it, esXpress's way of working is great. So great in fact, that VMware themselves are now adopting this very way of working with their Disaster Recovery feature in vSphere 4, maybe even stepping away from their beloved VCB (VMware Consolidated Backup).

VCB in my opinion has never been that great, apart for some special uses in special environments. esXpress fits all, from single ESX hosts to large clusters. In contrast to VMware's Disaster Recovery, this is still buggy at the time of this blog, esXpress has been on this train for years now, and definitely knows the drill. esXpress 3.1 is not the holy grail though. Some features were just not easy to use, there was no global GUI to manage all nodes easily and there was no Data de-duplication available (not that I am that big a fan of data de-dupe for backup, but hey, everybody does it!)

Enter esXpress 3.5

To make up for most of the shortcomings, esXpress version 3.5 has been introduced. The engine itself still is pretty much the same. And exactly there lies the power of esXpress: It still WORKS. It just works, it always works. Extra features have been added in such a smart and incredible simple way, that the product remains rock stable. No "waiting for the point 1 release" needed here!

I was over at a client who suffered a SAN failure (when upgrading firmware). They were in progress of failing over to their recovery site, when the administrator got an email from one of the production ESX hosts: esXpress had successfully completed its backups. What? All LUNs appeared unavailable at the production site. This host did not have its storage devices rescanned; it still kept on ticking.

I think things like this are major plusses for both VMware ESX esXpress showing their enterprise readiness.

Finally: A working global GUI

From the initial ESX 3.5 release, PHD also released a GUI to manage all esXpress instances from one central portal. In the old 3.1 (and before) days, you ended up copying config files between hosts; working, but not very user friendly. You might think that adding a central GUI took a lot of deep digging in the code of esXpress.

But, they surprised once again: The GUI just holds the config files and, could it be more simple, the GUI appliance introduces a small NFS store. The NFS store is automatically mounted to the ESX servers, and presto! That is where the config files can be found. esXpress itself just has to check the share for a new config, something already (partly) in existence in the previous version.

Even better: the GUI does a great job. I had some trouble with the first version; some manual labor was needed to get it going (like manually needing to change the time zone and not being able to add a second DNS server). All these issues are fixed now, but even those early versions were already very effective. And things have become only better since then!

Because "everybody has it": De-duplication

What should we do without de-duplication now-a-days? It is a major hype around storage and backup. If you don't have it, you're out of business it seems. But who ever thinks about the risks and limitations involved (see: The Dedup Dillema).

The idea of de-duplication is brilliant, but the implementation has to be right. I must admit, I am not a big fan of de-duplication. It is still your vital data you are talking about! Admt nr.2: esXpress 3.5 managed to change my opinion to de-dupe a little.

The de-duplication implementation of esXpress is in style with PHDs way of working: both effective and simple. A separate appliance is installed (which is in fact the same one as the GUI appliance. At first boot of the appliance you choose what the appliance will become. Smart!). The de-dupe appliance (called PHDD for PHD Data De-dupe) can mount a data-store or an NFS store for storing its de-duped data. It performs quite well, saving disk-space as you backup more of the same (or alike) data. It is now much "cheaper" to keep more backups of your VMs.

Only few changes appear to have been made to esXpress itself to allow PHDD as a backup target, so once again, stability guaranteed.

So now all your data lives inside the PHDD appliance. Now how do I get out this data the way I want it? PHD did something clever: They added a CIFS/SAMBA interface to the appliance, allowing you to browse, copy and backup your VMs as if they weren't de-duped at all! This last feature makes the mix of backup and de-dupe more acceptable, even effectively useable.

When will the fun EVER stop? File level restore!

The best feature of the PHDD de-dupe target in my opinion, next to de-dupe itself, is the ability to perform file level restores. At last you can get out that one single file of a full VM without having to restore the whole thing. This option is so cool, you simply browse to the appliance, select your files, and save the collection you marked as a single zip file! Couldn't be easier, another bulls eye for PHD, even in their first release of this piece of software.

Scaling esXpress 3.5 with de-dupe

Not all is bright and shiny with de-dupe. I found it hard to scale the solution: If there is only one PHDD target, scaling ends somewhere, and a SPOF (single point of failure) is introduced. Not good (although PHD is working on a way to link the de-dupe appliance to a secondary one). Still, one may consider using two or more PHDD appliances in parallel. This will work, but the de-dupe effectiveness will drop sharply, especially when you use DRS and all VM backups end up on all PHDD targets in time (this happens when you design the often used strategy where one assigns for a backup target to each ESX server individually with fail-overs to others).

You can make it somewhat more effective by specifying a backup target for each VM (in the local config), a best practice that also stands when using multiple FTP targets btw. This will ensure that a backup of a particular VM will always end up on the same backup target, making things clearer and making de-dupe more effective (although far from ideal – Every PHDD target has its own library of data, meaning that identical blocks still get stored on EACH PHDD target instead of just one).

The limitations mentioned above are not a limit of esXpress though, but more a limitation of de-dupe in itself. PHD choose to use online de-dupe (basically you de-dupe while you write), which will use CPU power during backup and restores. CPU power might even be the limiting factor in your backup speed. Luckily CPU power is usually available in abundance now-a-days.

I will dive deeper into performance and scaling of de-duped installations in the next blog-post, which will hopefully prove that de-dupe really performs (like the setup using multiple FTP targets simultaneously described in my blog-post Scaling VMware hot-backups using esXpress).

Conclusion

The new version of esXpress 3.5 is in terms of speed and reliability on par with its predecessor version 3.1. It is still the only backup solution I know that has no Single Point of Failure, scales (REALLY scales) up to whatever size you want without any issues, and best of all: Once it works it **KEEPS** working with hardly any problems around VM snap-shooting like some other backup solutions do have.

On top of all the good things that already were, a global GUI is added which manages all esXpress installs at the same time, and there is a Data De-duplication appliance which features a very well working single file restore option. I would like to have seen a file restore option in a non-de-dupe target as well. From what I've seen, online de-duping costs a lot of CPU power, and the backup speeds go down because of this. Once the already backup up in the de-dupe appliance). Still, calculations have to be done.

In a smaller environment, the de-dupe appliance is no match for a set of non-de-duping FTP targets. This is a drawback from which any de-dupe system suffers... It is just the way the "thingy" works. Still I see a solid future for esXpress's PHDD de-dupe targets where speed is not of the utmost importance.

Make no mistake on backup speeds: IF esXpress and its backup targets are designed and configured properly, it is by far the fastest full-VM backup solution I've seen. It does not mess with taking backups through the service console network; it creates Virtual Appliance runtime that perform the backups- and many in parallel. If you want to see real backup speed from esXpress, do not test it on a single VM like some people tend to do when comparing. If you do, speeds are about on par with other 3rd party vendors. But when scaled up to make 8 or more backups in parallel to several backup targets with matched bandwidth, esXpress will start to shine and leave the competition far behind.

You can view Eriks Blog post here:

<http://erikzandboer.wordpress.com/2009/08/06/the-new-esxpress-3-5/>