

Pluggable Storage Architecture, exploring the next version of ESX/vCenter

March 19th, 2009 | 10 Comments |

The next version of ESX has a totally different architecture for storage. The new architecture is called “Pluggable Storage Architecture”. For my own understanding I wanted to write down how this actually works and what all the different abbreviations/acronyms mean:

- PSA= Pluggable Storage Architecture
- NMP = Native Multipathing
- MPP = Multipathing Plugin
- PSP = Path Selection Plugin
- SATP = Storage Array Type Plugin

At the top level we have “Pluggable Storage Architecture”. This is just the name of the new concept, but it’s a well chosen name cause that’s what it is... a new storage architecture that uses plugins. Let’s start with the native VMware plugins.

The Native Multipathing(NMP) module is the default module that ESX(as of vSphere) uses. If your array is listed on the compatibility list then VMware pre-decides which multipathing “algorithm” will be used. ESX will natively do: Fixed, MRU, Round-Robin. Of course you can change this “algorithm” if you want or need to, but be careful.

The NMP associates paths with the LUN and/or array. The NMP uses rules to decide which Storage Array Type Plugin(SATP) will be used and which Path Selection Plugin(PSP) will be used. If you want to edit these rules you can use “esxcli”, and you can find the rules in the “/etc/vmware/esx.conf”. For instance the default PSP for the EMX Symmetrix array is Fixed:

```
/storage/plugin/NMP/config[VMW_SATP_SYMM]/defaultpsp = "VMW_PSP_FIXED"
```

Now that these basics have been decided the two “sub” plugins take over. The SATP handles path fail-overs, it’s as simple as that. This of course will be reported back to the NMP, because the NMP is responsible for the i/o flow. The SATP will monitor the health of a path and take the required action depending on the type of array. Nothing more, nothing less.

Now you are probably wondering what the PSP does, or when it will be utilized. The PSP determines which path will be used for an I/O request, in other words the PSP deals with the Fixed, MRU and Round-Robin algorithms. Keep in mind that the type of PSP that will be used has already been pre-configured, it’s not the PSP that decides if it will do Fixed or MRU, it’s the NMP that decides this based on pre-defined rules of which an example has been showed above. Both the SATP and PSP are plugins of the NMP and are controlled by the NMP.

You can of course change the default behavior. In my post on iSCSI load balancing we’ve seen the “esxcli” in action. This command also interacts with the NMP. For instance you can easily check which SATPs are available at the moment:

```
[root@esx01.yellow-bricks.com ~]# esxcli nmp satp list
Name                Default PSP      Description
VMW_SATP_ALUA_CX    VMW_PSP_FIXED   Supports EMC CX that use the ALUA protocol
VMW_SATP_SVC        VMW_PSP_FIXED   Supports IBM SVC
VMW_SATP_MSA        VMW_PSP_MRU     Supports HP MSA
VMW_SATP_EQL        VMW_PSP_FIXED   Supports EqualLogic arrays
VMW_SATP_INV        VMW_PSP_FIXED   Supports EMC Invista
VMW_SATP_SYMM       VMW_PSP_FIXED   Supports EMC Symmetrix
VMW_SATP_LSI        VMW_PSP_MRU     Supports LSI and other arrays
VMW_SATP_EVA        VMW_PSP_FIXED   Supports HP EVA
VMW_SATP_DEFAULT_AP VMW_PSP_MRU     Supports non-specific active/passive arrays
VMW_SATP_CX         VMW_PSP_MRU     Supports EMC CX that do not
                    use the ALUA protocol
VMW_SATP_ALUA       VMW_PSP_MRU     Supports non-specific arrays that use
                    the ALUA protocol
VMW_SATP_DEFAULT_AA VMW_PSP_FIXED   Supports non-specific active/active arrays
VMW_SATP_LOCAL      VMW_PSP_FIXED   Supports direct attached devices
```



This example also illustrates that a specific SATP is linked to a specific PSP. You can also check which PSPs are



ABOUT YELLOW-BRICKS.COM

Welcome to Yellow-Bricks.com a blog about virtualization, with a strong focus on VMware, HAand vCloud solutions.

Some facts:
 999 Blog Articles
 1950 Twitter Followers
 5503 RSS Subscribers

SPONSORS

ADVERTISEMENTS


Improve ESX Performance
 Quickly Identify Performance Issues Get VMware Capacity Analyzer Now.
www.VKernel.com/Performance

Storage White Paper
 15 Questions to Ask When Evaluating Storage Virtualization Technologies
www.HDS.com

NEC Enterprise Storage
 Fibre and iSCSI SANs with nationwide support by NEC
www.pacstorage.com



RECOMMENDED



vSphere 4.0 Quick Start Guide
 Shortcuts down the path of Virtualization
 Buyit now just \$15.99.

ADVERTISEMENTS



available at the moment:

```
[root@esx01.yellow-bricks.com ~]# esxcli nmp psp list
Name                Description
VMW_PSP_MRU         Most Recently Used Path Selection
VMW_PSP_RR          Round Robin Path Selection
VMW_PSP_FIXED       Fixed Path Selection
```

At the moment there are only three PSPs available, but you can imagine that vendors develop their own PSP to optimize load balancing for their array. The vendors can even write their own SATPs. You can also use esxcli to change the rules or add a rule, although I would not recommend doing this.

What's next? That would be a Multipathing Plugin(MPP). An MPP can be seen as "NMP+SATP+PSP". An MPP combines all the three functionalities into one module which, these MPP's will be developed by the vendors of the arrays. EMC is one of the vendors that already demonstrated their MPP. You can see a cool demo on the [blog](#) of Chad Sakac, which is EMC's MPP also known as Powerpath for ESX.

VMware I/O Performance

VM to LUN I/O optimization for your mission critical apps
www.virtualinstruments.com



Posted in Server | Tags: performance, Storage, vSphere, vstorage

« Update to several SRMSRA's

Minor Update: RVTools 2.4.1 »

You can skip to the end and leave a response. Pinging is currently not allowed.

10 Responses to "Pluggable Storage Architecture, exploring the next version of ESX/vCenter"



Darlin says:

Friday, March 20, 2009 at 08:40

A Plugin that does Host-Based Mirroring would be cool.



Tim Jacobs says:

Friday, March 20, 2009 at 09:39

Hi Duncan,

Thanks for the elaboration. Do you know at what granularity the NMP will work? I am thinking about mixed SAN environments where you use active/active arrays in combination with active/passive arrays on the same host (for example where you offer high-end high performance vs. low end archiving storage to VMs). Will this load different SATP's on a LUN base? Or will the ESX host fall back to a generic plugin for all LUN's?

Thanks & best regards,

Tim



David says:

Friday, March 20, 2009 at 15:20

Tim,

Based on what I've seen it is per LUN.

David



Duncan Epping says:

Friday, March 20, 2009 at 15:42

As far as I know it's per LUN indeed. The NMP decides which SATP/PSP to use per LUN.



virtualgeek says:

Friday, March 20, 2009 at 18:13

It's by LUN, and the default SATP is configured based on the array ID

Zmanda.com

Ads by Google

RECENT COMMENTS

- ▶ Greg on HACli
- ▶ Duncan Epping on Changing your dvPortgroup settings? (need your input!)
- ▶ Duncan Epping on VMworld San Francisco Run 2010
- ▶ Richard Boswell on Changing your dvPortgroup settings? (need your input!)
- ▶ Keith Norbie on VMworld San Francisco Run 2010

TAGS

2.5 3.0.x 3.5 BC-DR Blogging Book

Bugs cloud design drs **ESX** esxi ha

Howto kb linux networking **News** patches

performance powershell **Scripting**

scripts security service console software srm

Storage Tools training u2 update

upgrade VCB vcdx vcenter vdi

VirtualCenter vmtn **VMware**

vmworld vSphere vstorage whitepaper

Workstation

DISCLAIMER

The views expressed anywhere on this site are strictly mine and not the opinions and views of VMware.



Jon Owings says:

Wednesday, March 25, 2009 at 13:57

With vendor written plugins will there be a support fight about bugs in the plugins?
Or will they get a "gold star version" and vmware will agree to support the vendor code?



Duncan Epping says:

Wednesday, March 25, 2009 at 14:30

I don't think VMware will support the vendor code. Same goes for the SRA with Site Recovery Manager, which is supported by the vendor. But I'm not sure. I haven't found anything to support this, I'm just guessing at his point.



Shane Wendel says:

Wednesday, March 25, 2009 at 17:41

My understanding was that ESX4 would no longer have a service console in any version, like ESX3i.
Did this change or are you accessing an unsupported mode to run these commands? It doesn't appear that you are using VIMA in your article.



Duncan says:

Wednesday, March 25, 2009 at 20:00

there will still be an ESX and ESXi version as far as I know/seen...



sharmelin says:

Wednesday, April 29, 2009 at 00:45

Will this be supported for NFS storage arrays with multiple IPs and gateways?

Leave a Reply

Name (required)

Mail (will not be published) (required)

Website

Submit Comment

Copyright © Yellow Bricks - Building blocks for virtualization...

Powered by WordPress | Modern Style theme by FlexThemes