

## Search Results for: vDS

# Vmware vSphere 4.1 Network IO control (NetIOC) understanding

27/07/2010 – 9:03 am

Posted in [Vmware Technology](#), [vmware vSphere 4.1](#)

Tagged [Vmware](#), [vSphere](#), [Network Distributed Switch](#), [Configuration](#), [Troubleshooting](#), [4.1](#), [Netioc](#), [Network Control](#), [qos](#)

Comments (1)

One of biggest changes for vSphere 4.1 is introduction of Network I/O control and Storage I/O.

This post will give you an introduction and understanding of what Network I/O control (NetIOC) is. This is a new technology and we still need to wait and see more real case in the future. But for now, Let's see what Network IO control is.

## Why do we need to have Network IO Control (NetIOC)?

1. 1Gbit network is not enough

As you may know, we have more and more demanding on the network traffic. FT traffic, iSCSI Traffic (Don't you team up?) and NFS Traffic, vMotion Traffic etc. Although you can team up multiple physical nics together, but from a single VM perspective, it can only allow to use one physical nic at one time no matter what kind of teaming method you are using. Plus, network team has already started to talk about 10Gbit network and it's about time to push 10Gbit network into public.

2. Blade server demands

All new blade server has 10Gbit ports switch in the blade. The architecture of Blade server has changed from each blade has it's own ports to central ethernet Module. It saves a lot of resource and traffic can be easily Qos and scaled.

## Prerequisites for Network IO control

1. You need Enterprise Plus license

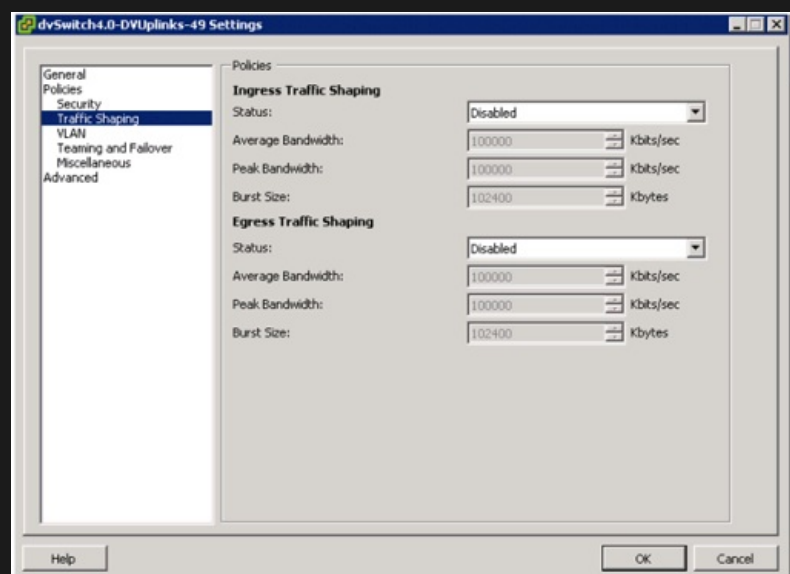
The reason for that is NetIOC is only available for vDS. For vSS, you can only control outbound traffic.

2. You need vSphere 4.1 and ESX4.1

With vSphere 4.0, you do can control traffic by port group. But you can't preconfigure traffic by type (or you can call it by class). This is fundamental architecture change. We will talk about it later. ESX4.1 is also required otherwise you won't see the new tab in the vCenter.

## How does Network IO Control (NetIOC) work?

If you recall vSphere 4.0, we also have ingress and egress traffic control for vDS. (for vSS, we only have outbound control) Traffic shaping is controlled by Average Bandwidth, peak bandwidth and burst size. You have manually divide dvUplinks by functions. Like this dvUplink is for FT, this is for vMotion, this is for iSCSI etc. Everything is done by manual.



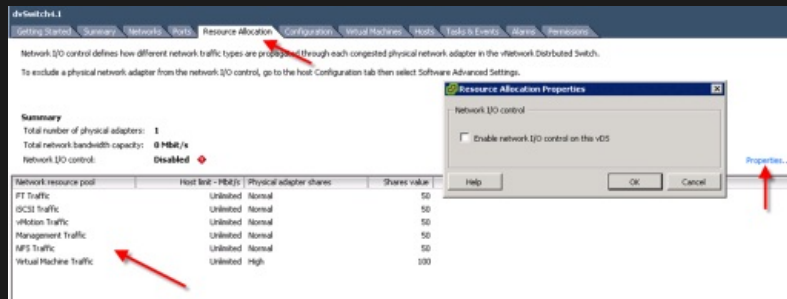
With new vSphere 4.1, we are not only able to control traffic by port group, we are also control traffic by class.

The NetIOC concept revolves around resource pools that are similar in many ways to the ones already existing for CPU and Memory.

NetIOC classifies traffic into six predefined resource pools as follows:

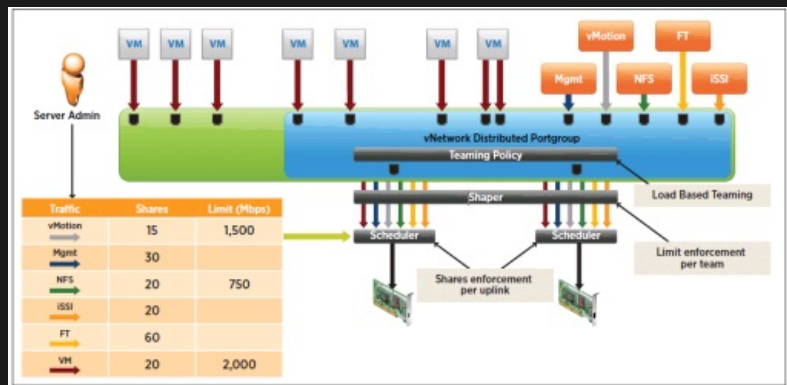
- vMotion
- iSCSI
- FT logging
- Management
- NFS
- Virtual machine traffic

If you open vCenter, you will see the new tab of dvSwitch in your ESXi 4.1 server.



This means all traffic go through this vDS will be under Qos by these rules. Remember, it only works for this vDS.

Now, let's see the architecture picture first and then, we talk about how this thing work.



As you can see, there are 3 layers in NETIOC. Teaming Policy, shaper and Scheduler. As what my previous post mentioned, vDS is actually a combination of special hidden vSS and policy profiles downloaded from vCenter.

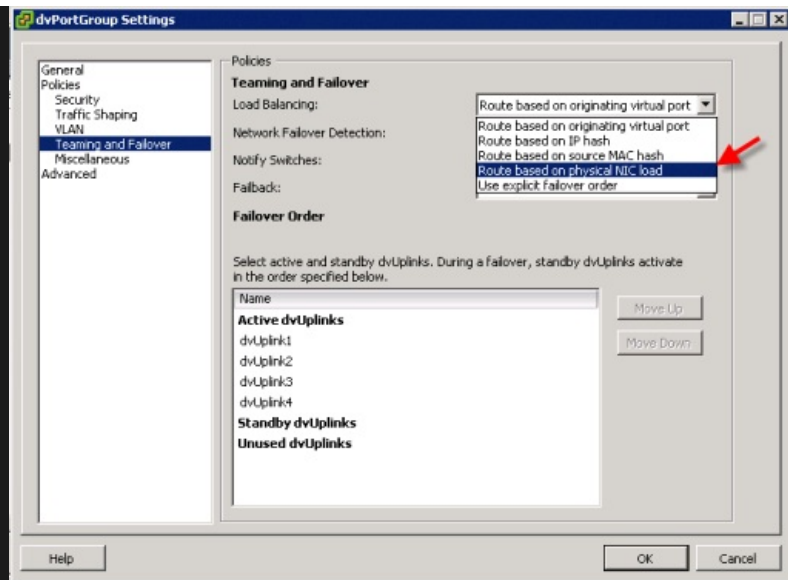
Teaming policy (New policy, LBT)

There is a new method of teaming called LBT(Load base teaming). It basically detect how busy those physical nics are, then it will move the flows to different cards. LBT will only move a flow when the mean send or receive utilization on an uplink exceeds 75 percent of capacity over a 30-second period. LBT will not move flows more often than every 30 seconds.

Best practice 4: We recommend that you use LBT as your vDS teaming policy while using NetIOC in order to maximize the networking capacity utilization.

NOTE: As LBT moves flows among uplinks it may occasionally cause reordering of packets at the receiver.

I haven't done any tests on how much extra CPU cycles are required to run LTB, but we will keep eyes on it.



## Limits

There are two attributes (Shares and Limit) you can control over traffic via Resource Allocation. Resource Allocation is controlling base on vDS and only apply to this vDS. It applies on vDS level not on port group or dvUplink level. Shaper is where limits apply. It limits traffic by the class of traffic. Be noticed at this 4.1, each vDS has it's own resource pool and resource pool are not shared between vDS.

A user can specify an absolute shaping limit for a given resource-pool flow using a bandwidth capacity limiter. As opposed to shares that are enforced at the dvUplink level, limits are enforced on the overall vDS set of dvUplinks, which means that a flow of a given resource pool will never exceed a given limit for a vDS out of a given vSphere host.

## Shares

Shares apply to dvUplink Level and each share rates will be calculated base on traffic of each dvUplink. It controls share value of traffic going through this particular dvUplink and make sure share percentage is correct.

the network flow scheduler is the entity responsible for enforcing shares and therefore is in charge of the overall arbitration under overcommitment. Each resource-pool flow has its own dedicated software queue inside the scheduler so that packets from a given resource pool won't be dropped due to high utilization by other flows.

## NetIOC Best Practices

NetIOC is a very powerful feature that will make your vSphere deployment even more suitable for your I/O-consolidated datacenter. However, follow these best practices to optimize the usage of this feature:

Best practice 1: When using bandwidth allocation, use "shares" instead of "limits," as the former has greater flexibility for unused capacity redistribution. Partitioning the available network bandwidth among different types of network traffic flows using limits has shortcomings. For instance, allocating 2Gbps bandwidth by using a limit for the virtual machine resource pool provides a maximum of 2Gbps bandwidth for all the virtual machine traffic even if the team is not saturated. In other words, limits impose hard limits on the amount of the bandwidth usage by a traffic flow even when there is network bandwidth available.

Best practice 2: If you are concerned about physical switch and/or physical network capacity, consider imposing limits on a given resource pool. For instance, you might want to put a limit on vMotion traffic flow to help in situations where multiple vMotion traffic flows initiated on different ESX hosts at the same time could possibly oversubscribe the physical network. By limiting the vMotion traffic bandwidth usage at the ESX host level, we can prevent the possibility of jeopardizing performance for other flows going through the same points of contention.

Best practice 3: Fault tolerance is a latency-sensitive traffic flow, so it is recommended to always set the corresponding resource-pool shares to a reasonably high relative value in the case of custom shares. However, in the case where you are using the predefined default shares value for VMware FT, leaving it set to high is recommended.

Best practice 4: We recommend that you use LBT as your vDS teaming policy while using NetIOC in order to maximize the networking capacity utilization.

NOTE: As LBT moves flows among uplinks it may occasionally cause reordering of packets at the receiver.

Best practice 5: Use the DV Port Group and Traffic Shaper features offered by the vDS to maximum effect when configuring the vDS. Configure each of the traffic flow types with a dedicated DV Port Group. Use DV Port Groups as a means to apply configuration policies to different traffic flow types, and more important, to provide additional Rx bandwidth controls through the use of Traffic Shaper. For instance, you might want to enable Traffic Shaper for the egress traffic on the DV Port Group used for vMotion. This can help in situations when multiple vMotions initiated on different vSphere hosts converge to the same destination vSphere server.

Let me know if you have more questions.

Reference:

<http://h18000.www1.hp.com/products/blades/components/ethernet/10-10gb-f/specifications.html>

[http://kb.vmware.com/selfservice/microsites/search.do?language=en\\_US&cmd=displayKC&externalId=1022585](http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1022585)

[http://www.vmware.com/files/pdf/techpaper/MMW\\_Netioc\\_BestPractices.pdf](http://www.vmware.com/files/pdf/techpaper/MMW_Netioc_BestPractices.pdf)

## What is Vmware Future look like in 2011?

16/07/2010 – 11:01 am

Posted in Microsoft, Vmware Technology, Vmware vSphere 4.01

Tagged Vmware, vSphere, Understanding, vmworld 2010

[Leave a Comment](#)

This is not exactly a technology post. But I would like to use this opportunity to discuss few interesting things developing in the Vmware and virtualization world.



### What's going on with Vmware world 2011?

I have read some content lists of Vmworld 2010. If you pay attention enough you would discover Vmware finally started to focus on optimizing their system and develop real piratical applications instead of spending time to build new infrastructure of future Operating system. If you can recall what happened last year in vmware world 2009, Vmware pushed out Cloud system but it was not well accepted by most of companies. Most people had doubts in their mind and confusing on their faces. The idea of if system is not broken, don't touch it and also falling global economy environment ring bell much louder than anything else near IT Managers ears. So this year of Vmworld 2010, Vmware noticed how important education of Cloud and also practical applications are. They started to turning back to pick up lost customers who they left behind before and help them with vShield, Charge Back and Capacity IQ, AppSpeed.



### vSphere 4.1 is out!

Oh, yes. The great 4.1 is out. It's not 5.0 or new cloud system. It's just a service pack according to Vmware. Like what I mentioned before, the major feature of vSphere 4.1 is turning their system and try to add more practical functions. Why? Because vSphere 4 is not that great comparing with ESX 3.5. Yes, some hardware performance get 10% up and some new features like vDS, FT or Data Recovery sounds pretty. However, is this a real necessary to "must upgrade" or it's just another "nice to have"? If you went through the list of difference between vSphere and ESX 3.5, you would see the whole point. Please count how many exactly new features vSphere will bring to you. Yes, ESXi is better than ESX, but it's not a must. New backup method is interesting, but current backup is doing ok with third party software. DR does sound nice. But it does also require a whole new budget and SAN level support. vDS is simply not necessary. vSS is great for normal job. View is not as good as Citrix Xen

desktop. ThinApps is not bad but it can't compete with native support from Microsoft vApp. Especially, MS has lots of experience on managing and deploying Applications. Don't mention vShield and VMsafe, those 2 applications are not practical at all. Let's hope vShield 2.0 will be better.



## Battle with in Cloud War

Oh, yes. There is War coming. Ms is attacking from both hardware level (selling servers with Cloud OS in built) and purely software level (just hosting your MS office application , Email and sharepoint online). VMware shake hands with Google, EMC, Cisco and allies has been formed to fight this giant software company. Microsoft , VMware, and Google have already hosting their cloud internally and Virtual applications will soon spreading everywhere and hit every corner on this earth. 2010 is a quiet year. But if you look up in the sky, the gunpowder colour cloud rolling from the edge of sky, bleached lighting merging and swimming in deep layers of clouds. Are you ready to choose side? Have you embrace yourself and ready for impact?

Reference:

<https://vmworld2010.wingateweb.com/scheduler/catalog/catalog.jsp>

## vDS (vNetwork Distributed Switch), My Understanding Part 2

21/05/2010 – 4:40 pm

Posted in VMware Technology, VMware vSphere 4.01

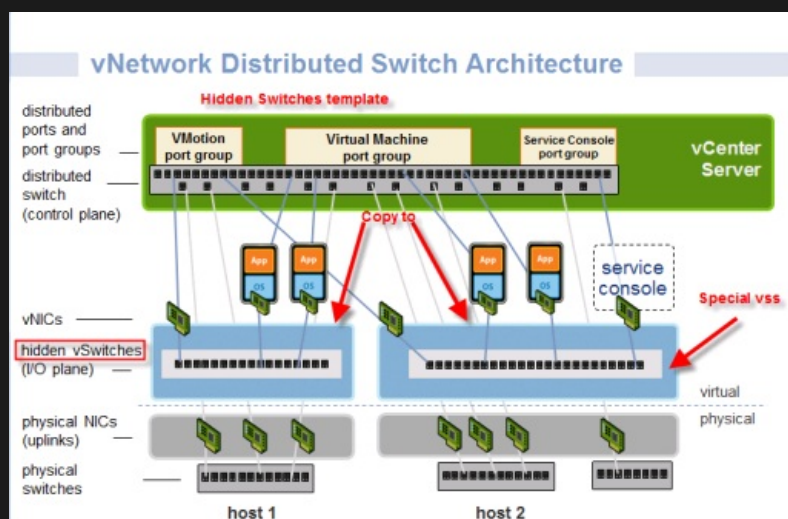
Tagged VMware, vSphere, vds, Network Distributed Switch, Understanding

[Leave a Comment](#)

This is part 2 of vDS (vNetwork Distributed Switch), My Understanding.

### How does vDS work?

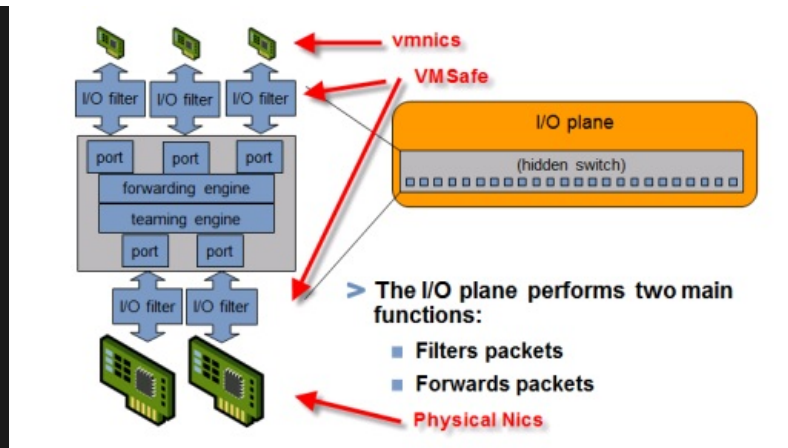
What will your instructor tell you? "Please don't consider vDS is a switch connecting to Hosts. vDS is just a template" Well, that's what you always heard from all your instructors. but template of what? The answer is vDS is template of **HIDDEN vSwitch** sitting on your local host. vDS(the template) is managed by vCenter(high level operation) and your local Host(low level operation). Let's see a diagram.



From this diagram, you can see there are two hosts. Each host has hidden switch which received template (vDS) from vCenter. The local template will be updated every 5 minutes like what I mentioned in Part 1.

Now, let's open this hidden switch and see what's happening in there.





As you can see, the hidden switch has forwarding engine and teaming engine which will be configured and controlled by setting in vCenter. There are two IO filters (not just one) is to be used in VMsafe. So what VMsafe does is let third party software (for example, the Trend Micro) build a VM appliance and be certified by VMWARE to prove it won't do any damage. That special VM will use special API to monitor traffic (like firewall) or check virus. Meaning, if you want to use VMsafe product, you have to use vDS, meaning you have to buy Enterprise Plus license! I guess that's why VMsafe product is not popular.

ok. Back to vDS. Let's make a small conclusion. vDS is also a vSS. But it's hidden in the Host. This hidden vSS is using template made by vCenter and Local Host so you can control traffic and share switch data between hosts.

### Few things you need to know about vDS

vDS is capable to do everything vSS can do because it's basically a super (hidden) vSS. Once you assign a vmnic OR VMkernel, SC or VM to vDS, you won't be able to use them in vSS. It's same thing as vSS.

I won't say there are not much point to use vDS but if you do want to use vDS, you would either use Cisco Nexus to replace vDS or you want to use VMsafe product. Or you have Enterprise Plus license and want to use host profiles.

#### vDS timeout issue

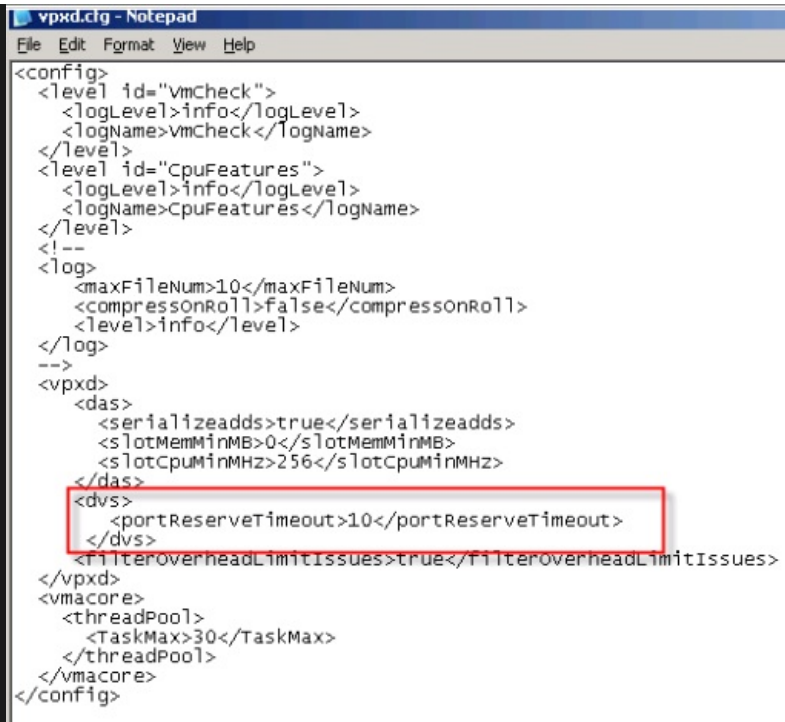
vDS is not as convenient as vSS which only connects to single Host. sometimes, it's not easy to remove vDS switches or even switch physical nic to different vDS switches. If vDS believes a port is busy, vCenter won't allow you to delete vDS or remove a host from it. By the default, vCenter automatically forces all "busy" ports on all distributed switches to time out every 24 hours.

You can make change on vpxd.cfg to make it as 10 minutes.

vpxd.conf is located at c:\documents and settings\all users\application data\VMware\VMware VirtualCenter\vpxd.cfg

In vpxd.cfg, add the

line <vpxd><dvs><portReserveTimeout>10</portReserveTimeout></dvs></vpxd> and save the file.



```
<?xml version='1.0' encoding='UTF-8'>
<config>
  <level id="vmCheck">
    <logLevel>info</logLevel>
    <logName>vmCheck</logName>
  </level>
  <level id="CpuFeatures">
    <logLevel>info</logLevel>
    <logName>CpuFeatures</logName>
  </level>
  <!--
  <log>
    <maxFileNum>10</maxFileNum>
    <compressionRoll>false</compressionRoll>
    <level>info</level>
  </log>
  -->
  <vpxd>
    <das>
      <serializeadds>true</serializeadds>
      <slotMemMinMB>0</slotMemMinMB>
      <slotCpuMinMHZ>256</slotCpuMinMHZ>
    </das>
    <dvs>
      <portReserveTimeout>10</portReserveTimeout>
    </dvs>
    <filterOverheadLimitIssues>true</filterOverheadLimitIssues>
  </vpxd>
  <vmacore>
    <threadPool>
      <TaskMax>30</TaskMax>
    </threadPool>
  </vmacore>
</config>
```

Restart vCenter. The default timeout is now set to ten minutes.

After the port reservation has timed out, remove the vNetwork Distributed Switch or dvPort group.

Reset the default timeout by removing the line you previously added to vpxd.cfg.

Restart vCenter.

#### Best Practice for vDS daily operation

If you run into a problem with vDS, always start checking from vCenter->Networking level. Because it has a general view to tell you all vDS details such as IP on each port group, PVLAN info and which VM or kernel ports it connects. Then, you should drill down to single host->Configuration->Networking to add or remove objects. If you do have an issue, try to remove all objects to another vDS and then, make your change.

For the rest of details, like PVLAN, blocking of individual ports, you can check this file to continue your journey of vDS.

<http://vmware.com/files/pdf/vsphere-vnetwork-ds-migration-configuration-wp.pdf>

Reference:

<http://kb.vmware.com/1010913>

<http://vmware.com/files/pdf/vsphere-vnetwork-ds-migration-configuration-wp.pdf>

## vDS (vNetwork Distributed Switch), My Understanding Part 1

21/05/2010 – 12:00 am

Posted in VMware Technology, VMware vSphere 4.01

Tagged VMware, vSphere, vds, Network Distributed Switch, Understanding

Leave a Comment

vSphere has introduced many new features. One of the new features is vDS (vNetwork Distributed Switch) which always confuses me and lots of people. I'm trying to explain it as simple and easy, also deep to you as I can. If I make any mistakes, please feel free to leave a comment. Thank you.

#### So What is vDS? What's the difference between vSS and vDS from configuration file structure?

vDS is a new Virtual Switch introduced by VMware. The old vSS is more like a local host property. All switch data is saved in the local host. Other hosts are not aware of what kind of vSS other hosts have. Not only vCenter can't do anything about it, it causes trouble when you do vMotion. vDS is saved in both vCenter and Host. One copy is in the vCenter, vDS is saved in the SQL database. In the local host, vDS has another local database cache copy sitting at /etc/vmware/dvsdata.db. This local cache will be updated by vCenter every 5 minutes.

You can use the following command to help you to get copy to read the local host database.

```
cd /usr/lib/vmware/bin
```

```
./net-dvs > /tmp/dvs.txt
```

then, you can read dvs.txt

```

switch ca 49 07 50 f7 81 c1 fb-da e6 79 e2 aa f1 5f 85 (etherswitch)
global properties:
  com.vmware.common.alias = vDS-Production
  com.vmware.etherswitch.mtu = 1500
  com.vmware.etherswitch.cdp = CDP, advertise & listen
  com.vmware.common.uplinkPorts:
    dvUplink1
  com.vmware.etherswitch.pvlanMap:
    (100, 100) - promiscuous
    (100, 101) - community
    (100, 102) - isolated
  com.vmware.common.pgmap:
    dvportgroup-84 <-> vDS-Production-DVUplinks-83
    dvportgroup-143 <-> PVLAN100
    dvportgroup-144 <-> PVLAN101
    dvportgroup-145 <-> PVLAN102
    dvportgroup-146 <-> Production
host properties:
  com.vmware.common.host.portset = DvsPortset-0
  com.vmware.host.volatile.status = green
  com.vmware.common.host.uplinkPorts:
    130
port 130:
  com.vmware.common.port.alias = dvUplink1
  com.vmware.common.port.block = false
  com.vmware.etherswitch.port.teaming:
    load balancing = source virtual port id
    link selection = link state up; link speed=10Mbps;
    link behavior = notify switch; reverse filter; best effort on failure; shotgun on
failure:
  active =
  standby =
  com.vmware.etherswitch.port.security = deny promiscuous; allow mac change; allow forged f
cmds
  com.vmware.etherswitch.port.vlan = Guest VLAN tagging

```

**dvsdata.db export**

Also, after you configure vDS on your local host, your esx.conf (/etc/vmware) has record shows brief configuration information of vDS.

```

firewall/services/nfsClient = "0"
firewall/services/nisClient = "0"
firewall/services/sshClient = "1"
firewall/services/sshServer = "1"
firewall/services/swISCSIClient = "1"
/nas/NFS01/enabled = "true"
/nas/NFS01/host = "192.168.100.172"
/nas/NFS01/readOnly = "true"
/nas/NFS01/share = "/iso"
/nas/VBCRepository/enabled = "true"
/nas/VBCRepository/host = "vdc-nas-a.vmeduc.com"
/nas/VBCRepository/readOnly = "true"
/nas/VBCRepository/share = "/vdcrepos"
/net/dvswitch/child[0000]/dvsClassName = "etherswitch"
/net/dvswitch/child[0000]/dvsName = "vDS-Production"
/net/dvswitch/child[0000]/name = "DvsPortset-0"
/net/dvswitch/child[0000]/numPorts = "256"
/net/dvswitch/child[0000]/uplinks/child[0000]/connectionId = "1268107261"
/net/dvswitch/child[0000]/uplinks/child[0000]/dvpId = "130"
/net/dvswitch/child[0000]/uplinks/child[0000]/pnic = "vmnic1"
/net/dvswitch/child[0001]/dvsClassName = "etherswitch"
/net/dvswitch/child[0001]/dvsName = "vDS-vMotion"
/net/dvswitch/child[0001]/name = "DvsPortset-1"
/net/dvswitch/child[0001]/numPorts = "256"
/net/dvswitch/child[0001]/uplinks/child[0000]/connectionId = "810138511"
/net/dvswitch/child[0001]/uplinks/child[0000]/dvpId = "130"
/net/dvswitch/child[0001]/uplinks/child[0000]/pnic = "vmnic2"
/net/dvswitch/child[0002]/dvsClassName = "etherswitch"
/net/dvswitch/child[0002]/dvsName = "vDS-15531-NAS"
/net/dvswitch/child[0002]/name = "DvsPortset-2"
/net/dvswitch/child[0002]/numPorts = "256"
/net/dvswitch/child[0002]/uplinks/child[0000]/connectionId = "1132366343"
/net/dvswitch/child[0002]/uplinks/child[0000]/dvpId = "130"
/net/dvswitch/child[0002]/uplinks/child[0000]/pnic = "vmnic3"
/net/pnic/child[0000]/mac = "00:21:5a:ce:87:06"

```

**esx.conf**

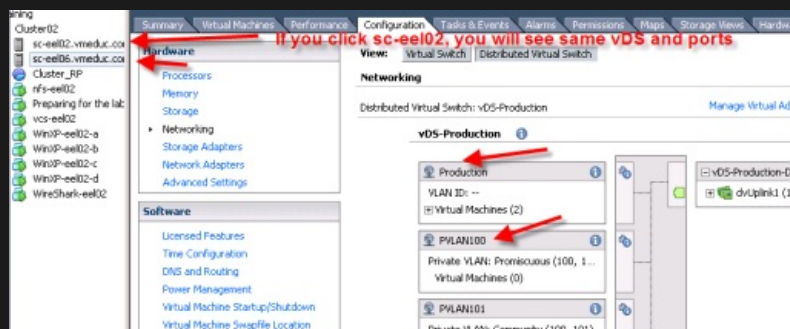
Those 3 configuration combines vDS structure. This also makes vDS can work even after Host disconnected from vCenter.

#### What's difference between vSS and vDS on control level?

With vSS, everything should be controlled on local host. Basically, you go to Local Host->Configuration->Networking. Then, you start everything from there. But vDS is different. vDS divide control into 2 different level. I call them high level and low level.

High Level: High level is to create/remove, management teaming, distribution port group etc. This level sits at vCenter->Inventory->Networking.

Low Level: This level is to connect your vm, vmkernel, and your local physical cards to vDS. Please be aware that your vm, vmkernel, etc are connecting to distribute port group. Unlike local vSS (you have create same vswitch, same vswitch port group on all hosts), vDS is pushed from vCenter to all Hosts. As long as you are connecting to same vDS, you will have same distribute port group.



With local physical nic card, they need to connect to dvUplink side. You can choose any number of local nics to connect or even no nic at all. But what you can't do is to setup teaming (only work for 2 nics from same host), traffic shaping, VLAN because you need to setup on



high level.

To be continued. ....

#### GeekSilver Email Subscription

Enter your email address to subscribe to this blog and receive notifications of new posts by email.

Sign me up!

#### GeekSilver RSS Link



RSS - Posts

#### GeekSilver Twitter

- VMware vSphere 4.1 vs ESX 3.x storage identifier Understanding: <http://wp.me/pVbEv-6R> 16 hours ago
- VMware vSphere 4.1 Memory Compression Understanding: <http://wp.me/pVbEv-6B> 5 days ago
- VMware vSphere 4.1 vStorage APIs for Array Integration (VAAI) understanding: <http://wp.me/pVbEv-61> 1 week ago
- Difference between vSphere Hypervisor 4.1 and ESXi 4.1: <http://wp.me/pVbEv-6t> 1 week ago
- VMware vSphere 4.1 Storage IO Control (SIOC) understanding: <http://wp.me/pVbEv-65> 1 week ago

#### Search

#### Tags

4.1 101 Architecture aso Configuration Control  
Courses difference EMC ESX esxi fast hot add How to hp cim  
agent hypervisor Inform2010 Install IO logs Memory compression  
Microsoft Netioc Network Control Network Distributed Switch  
pvscsi qos SAN SGU Stargate Storage Tools Training  
Troubleshooting Understanding vaai vcp vds VIEW  
vMA VMware vmworld 2010 vSphere  
Windows 2008 sp2 winsxs

#### Search

#### Blogroll

- Development Blog
- Documentation
- Plugins
- Suggest Ideas
- Support Forum
- Themes
- WordPress Planet

#### RSS Feeds

- All posts
- All comments

#### Meta

- Register
- Log in