

Deploying ITD: Server Traffic distribution Using Direct Server Return

Table of Contents: Deploying ITD: Server Traffic distribution Using Direct Serve	er Return1
1. Introduction :	1
2. Traditional DSR mode using External Load Balancers	1
3. DSR mode using Nexus – ITD	2
4. Deploying ITD for SLB-DSR	3
4.1 Server configuration	4
4.2 Nexus ITD configuration	5
4.3 Verification	6
5. Guidelines and Limitations	8
6. References and Further reading	8

1. Introduction :

Intelligent Traffic Director (ITD) is an ASIC based multi-terabit layer 4 traffic distribution and clustering solution on the Cisco Nexus 5000/6000/7000/9000 Series Switches. ITD provides scalable traffic distribution of client requests to a group of servers or service appliances.

This document describes a typical deployment scenario for distributing traffic with ITD using Direct Server Return (DSR) mode using Nexus switches.

2. Traditional DSR mode using External Load Balancers

With Server Load Balancing (SLB), incoming requests for a particular service are distributed across multiple servers for reasons such as load sharing, capacity and redundancy.

In a traditional DSR mode deployment, the incoming traffic from the client-facing network device is destined to an external Load-balancer appliance via a Virtual IP (VIP). The same IP address is also configured on all the servers (also termed nodes henceforth) as a loopback-IP-address, which enables the servers to reply to the client directly using the VIP. This ensures that the traffic bypasses the Load-balancer on the return path thus eliminating any bottlenecks in flow setups or overall throughput. This makes Direct Server Return(DSR) an ideal choice for stateless services like DNS load-balancing and for services with a large amount of server return data to the client like video services.





Figure 1 DSR Mode using External Load-Balancer(s).

However this approach requires an external Load Balancer with its separate configuration and management, in addition to the Routed/Switched network devices. Redundancy considerations also require that multiple Load-balancers be deployed for HA or Clustered designs.

3. DSR mode using Nexus – ITD.

With ITD, the Cisco Nexus Series Switches provide traffic distribution to servers, without the need for external load-balancers as seen in figure 2.







The Virtual IP, which was on the load-balancer earlier, is now configured within the ITD policy and can also be advertised to clients.

ITD also provides probes for node failure detection and sophisticated node-failureactions, which can be customized as per network requirements.

Using ITD provides many benefits such as:

- Reduced Cost: No external SLBs or Application Delivery Controllers (ADCs) are required.
- Unhindered Performance: ITD leverages ASIC based Line-rate traffic distribution to provide Multi-Terabit capacity without increasing latency.
- Scalability: ITD can scale to support a large number of Servers.

4. Deploying ITD for SLB-DSR

The deployment described below involves the following devices:

- 1) Nexus 7700 Two switches configured in vPC mode running 7.2(0)D1(1).
- 2) 2 Virtual Machines (servers) running Ubuntu Linux 14.04
- 3) 2 Virtual Machines (clients), also running Ubuntu Linux 14.04
- 4) Layer-2 switches (providing connectivity between VM's and Nexus switches)
- 5) The Server nodes host a simple HTTP service for testing/verification .

The same configurations can be applied for Server Load-balancing with DSR mode using ITD on Nexus 9000 series switches as well. The Nexus 5000/6000 switches currently do not support ITD probes.



Figure 3: ITD-DSR Implementation.



4.1 Server configuration

Direct Server Return requires the server to respond to clients directly by using the Virtual IP address as the source IP address. This necessitates that the server be configured with the VIP in addition to the regular IP address of the server. On typical Linux machines, this is done by using a Loopback address. A new Loopback interface and its address can be configured on Linux as below:

sudo ifconfig lo:1 172.16.1.100 netmask 255.255.255.255 -arp up

This command configures a virtual loopback interface "lo:1" with the IP address 172.16.1.100. Since the same Virtual IP is configured on multiple devices, there is a possibility of duplicate IP detection if the Servers also respond to ARP for the VIP. Hence the configuration uses '-arp' to disable ARP on this interface.

However, Linux also responds to ARP on wrong and unassociated interfaces by default. Hence the below configuration is required to ensure that the server does not respond to ARP for this IP on any other interface as well:

Append and save the following configuration to the file "/etc/sysctl.conf" net.ipv4.conf.all.arp_ignore=1 net.ipv4.conf.all.arp announce=2

Every server in the setup needs to be configured with the same Virtual IP address with the corresponding ARP changes using the steps above.

The VIP configuration on the server can be verified using 'ifconfig' from the terminal.

800	Server 1
ifconfig eth0	<pre> egrep -A 3 "lo:1 eth" Link encap:Ethernet HWaddr 00:50:56:b8:82:d2 inet addr:192.168.1.10 Bcast:192.168.1.255 Mask:255.255.255.0 inet6 addr: fe80::250:56ff:feb8:82d2/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1</pre>
lo:1	Link encap:Local Loopback inet addr:172.16.1.100 Mask:255.255.255.255 UP LOOPBACK RUNNING MTU:65536 Metric:1
	Server 2
ifconfig eth0	egrep -A 3 "lo:1 eth" Link encap:Ethernet HWaddr 00:50:56:b8:2e:43 inet addr:192.168.1.20 Bcast:192.168.1.255 Mask:255.255.255.0 inet6 addr: fe80::250:56ff:feb8:2e43/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
10:1	Link encap:Local Loopback inet addr:172.16.1.100 Mask:255.255.255.255 UP LOOPBACK RUNNING MTU:65536 Metric:1

Figure 4: Server verification.



4.2 Nexus ITD configuration

This deployment example uses vPC's on the Nexus switches for connectivity towards the Server side. vPC is expected to be already configured on the Nexus switches and is not covered here. The below section shows the configuration of switch 'Nexus-1'. The vPC peer 'Nexus-2' needs to be configured identically except for the SVI IP addresses.

The following features are prerequisites for ITD and need to be enabled in addition to other features already in use:

```
feature pbr!Enables redirection functionality required for ITDfeature sla sender!Enables probing mechanism used by ITDfeature sla responder!Enables probing mechanism used by ITDfeature interface-vlan!Enables users to create SVIsfeature itd!Enables the ITD functionality itself
```

The ingress-interface SVI Vlan 10 and the 'server-group' SVI Vlan100 are created.

```
interface Vlan10
  description ITD-DSR Ingress interface
  no shutdown
  ip address 172.16.1.2/24
interface Vlan100
  description ITD-DSR Server VLAN
  no shutdown
  ip address 192.168.1.2/24
```

A device group 'server-group100' is created and the physical-interface IP addresses of the servers are added to it. ICMP probes are also configured for this device group with default values. The probe timers can be tweaked as required (see guidelines and limitations section for probes on Nexus 5000/Nexus 9000)

```
itd device-group server-group100
  probe icmp
  node ip 192.168.1.10
  node ip 192.168.1.20
```

Finally, an ITD Service "vip-dsr100" is configured to use the created device-group "server-group100". The Virtual IP 172.16.1.100 is configured for the ITD Service with advertisement enabled to advertise the VIP to upstream routing protocols. The load balance method of '*Source IP*' is selected. Note that the command "*Failaction node reassign*" needs to be configured explicitly to enable reassigning traffic buckets of a failed node(see references for details). Finally, statistics need to be enabled in order to view ITD traffic distribution statistics.

```
itd vip-dsr100
  device-group server-group100
  virtual ip 172.16.1.100 255.255.255.255 advertise enable
  ingress interface Vlan10
  failaction node reassign
  load-balance method src ip
  no shut
itd statistics DSRService
```



4.3 Verification

The configured ITD Service can be verified as below:

PSK_N77	00_1-ITD-DSR(config-if)# sh	n itd						
Legend: ST(Sta	tus): ST–Stan	dby,LF-Li	nk Fa	ailed,	PF-Pro	be Fai	lled,PD)-Peer [Down,	IA-Inactive
Name	LB Sc	heme Sta	tus	Buck	ets					
vip-dsr	100 src-i	р АСТ	IVE	2						
Exclude	ACL									
Device	Group						Probe	Port		
server-	group100						ICMP			
Pool			Int	terfac	e S	tatus	Track_	_id		
vip-dsr	100_itd_pool		Vla	an10	U	 Р	3			
Virtual	IP				Netma	sk/Pre	efix Pr	otocol		Port
172.16.	1.100 / 255.2	55.255.25	5				IF)		0 0
Node	IP	Cfg-S	WGT	Probe	Port	Pr	obe-IF	STS	Trk#	Sla_id
1	192.168.1.1	0 Active	1	ICMP				 0К	1	10001
В	ucket List									
v	ip-dsr100_itd	1_bu	cket_	_1						
Node	IP	Cfg-S	WGT	Probe	Port	Pr	obe-IF	o sts	Trk#	Sla_id
2	192.168.1.2	0 Active	1	ICMP				ок ОК	2	10002
В	ucket List									
v	ip-dsr100_itd	1_bu	cket_	_2						

Figure 5: Nexus ITD verification

From the Client VMs, the Virtual IP's reachability and the HTTP service can be verified as shown below. For differentiation, a folder named with the Server number and IP address has been created on each Server.

Since the ITD by default uses the last octet(or least significant bits (LSBs)) for bucketing, skewed IP addresses .50 and .250 have been used for the clients to show the load-distribution in action.



Client 1: Ping and HTTP test to the VIP.

0		Client	1			_				
64	bytes	from	172.16.1.100:	icmp_seq=617	ttl=63	time=0.371	ms			
64	bytes	from	172.16.1.100:	<pre>icmp_seq=618</pre>	ttl=63	time=0.458	ms			
64	bytes	from	172.16.1.100:	icmp_seq=619	ttl=63	time=0.416	MS			
64	bytes	from	172.16.1.100:	icmp_seq=620	ttl=63	time=0.602	ms			
64	bytes	from	172.16.1.100:	<pre>icmp_seq=621</pre>	ttl=63	time=0.614	ms			
64	bytes	from	172.16.1.100:	icmp_seq=622	ttl=63	time=0.471	ms			
64	bytes	from	172.16.1.100:	icmp_seq=623	ttl=63	time=0.571	ms			
64	bytes	from	172.16.1.100:	icmp_seq=624	ttl=63	time=0.431	ms			
	-									
8	🗧 🖨 🛛 Fi	e Edit	View History Boo	kmarks Tools Hel	lp.		- †‡	En 📢))) 9:01 PM	ψ
1		Directo	ory listing for / ×	+						



Client 2: Ping and HTTP test to the VIP.

8	00) c	lient	2					-		
64	byt	es	from	172.16.1.100:	icmp_seq=774	ttl=63	time=0.641	ms			
64	byt	es	from	172.16.1.100:	<pre>icmp_seq=775</pre>	ttl=63	time=0.552	ms			
64	byt	es	from	172.16.1.100:	<pre>icmp_seq=776</pre>	ttl=63	time=0.687	ms			
64	byt	es	from	172.16.1.100:	<pre>icmp_seq=777</pre>	ttl=63	time=0.573	ms			
64	byt	es	from	172.16.1.100:	<pre>icmp_seq=778</pre>	ttl=63	time=0.533	ms			
64	byt	es	from	172.16.1.100:	<pre>icmp_seq=779</pre>	ttl=63	time=0.583	ms			
64	byt	es	from	172.16.1.100:	<pre>icmp_seq=780</pre>	ttl=63	time=0.329	ms			
б4	byt	es	from	172.16.1.100:	<pre>icmp_seq=781</pre>	ttl=63	time=0.265	ms			1
	08.0										
8	00	File	e Edit	View History Boo	kmarks Tools He	lp		t,	En 📢))	9:01 PM	华
1	3	1	Directo	ry listing for / ×	+						
	9	•) @ 17	2.16.1.100:8000		▼ C (२ Search	☆	ê 4	~ >>	≡
Ľ		D	irec	ctory listin	g for /						
C	9		• _Tł	HIS_IS_SERVER_	2_IP_192.168.1.2	20/					
Fig	ure 6:	: Clie	ent ser	vice verification.							



Using the CLI on the Nexus7700, statistics can be verified for ITD redirected traffic:

PSK_N7700_1-ITD-DSR(config-if)	# show itd vip-dsr100 statist	ics			
Service	Device Group	VIP/mask		#Packets	
vip-dsr100	server-group100	172.16.1.100 / 255.255	.255.255	2823	(100.00%)
Traffic Bucket	Assigned t	o Mode	Original Node	#Packet	s
<pre>vip-dsr100_itd_vip_1_bucket_1</pre>	192.168.1.	10 Redirect	192.168.1.10	1419	(50.27%)
Traffic Bucket	Assigned t	o Mode	Original Node	#Packet	s
<pre>vip-dsr100_itd_vip_1_bucket_2</pre>	192.168.1.	20 Redirect	192.168.1.20	1404	(49.73%)

5. Guidelines and Limitations

- Configuration of Loopback on the Server nodes is mandatory for DSR mode to work even if the servers run other operating systems(Windows for eg.).
- Nexus vPC peer devices needs to have identical configuration of the ITD service to operate correctly. This includes the number, order, configuration of the nodes as well as all other parameters like load-balance method, probe specifications etc.
- If failaction reassign is not configured with no standby, the traffic will get routed normally without ITD redirection.
- If the Server nodes support multiple NICs with Port-channel/LACP support, the Nexus switches can connect directly to the servers via vPC.
- Using weights for the nodes, traffic can be distributed unequally (when using servers of different capacities).
- For ITD-ICMP Probes on the Nexus 9000, the features "SLA responder/sender" are currently not required as prerequisites.
- Nexus 5000/6000 series switches currently do not support ITD Probes.

6. References and Further reading.

Creating Virtual interfaces in Linux: http://linuxconfig.org/configuring-virtual-network-interfaces-in-linux Linux ARP announce/ARP ignore: http://kb.linuxvirtualserver.org/wiki/Using arp announce/arp ignore to disable ARP Linux ARP Flux considerations: http://linux-ip.net/html/ether-arp.html - ether-arp-flux Nexus 7000 ITD configuration Guide: Nexus 7000 /7.x / ITD Config Guide Nexus 9000 ITD configuration guide: Nexus 9000 / 7.x / ITD Config Guide Nexus 5500 ITD config guide: Nexus 5500 ITD Config Guide

Nexus 5600 ITD config guide: <u>Nexus 5600 ITD Config Guide</u>