

ITD の導入: Direct Server Return を使用したサーバトラフィック分散

目次: ITD の導入: Direct Server Return を使用したサーバトラフィック分散	1
1. はじめに:	1
2. 外部ロード バランサを使用する従来の DSR モード	1
3. Nexus を使用する DSR モード: ITD	2
4. SLB 向けの ITD の導入: DSR	3
4.1 サーバ コンフィギュレーション	4
4.2 Nexus ITD 設定	5
4.3 確認	6
5. 注意事項と制約事項	8
6. 参考資料と詳細記事	8

1. はじめに:

Intelligent Traffic Director (ITD) は、Cisco Nexus 5000/6000/7000/9000 シリーズ スイッチに搭載された ASIC ベースのマルチテラビット規模のレイヤ 4 トラフィック分散およびクラスタリング ソリューションです。ITD は、サーバ グループまたはサービス アプライアンス グループへのクライアント要求のスケラブルなトラフィック分散を可能にします。

このマニュアルでは、Nexus スイッチを使用した Direct Server Return (DSR) モードにて、ITD を用いてトラフィックを分散する一般的な導入シナリオについて説明します。

2. 外部ロード バランサを使用する従来の DSR モード

サーバロード バランシング (SLB) を使用すると、特定のサービスの着信要求がロード シェアリング、容量、および冗長性などの理由により、複数のサーバに分散されます。

従来の DSR モードの導入では、クライアント側のネットワーク デバイスからの着信トラフィックは仮想 IP (VIP) 経由で外部ロード バランサ アプライアンスに送信されます。すべてのサーバ (ノードとも呼ばれる) で同じ IP アドレスをループバック IP アドレスとして設定すると、サーバは VIP を使用してクライアントに直接応答できます。これにより、トラフィックがリターンパスでロード バランサをバイパスするため、フロー設定や全体的なスループットのボトルネックが解消されます。したがって、DNS ロードバランシングのようなステートレス サービスや、ビデオ サービスのようにサーバからクライアントに大量のデータが戻されるサービスの場合、Direct Server Return (DSR) が最適な選択肢となります。

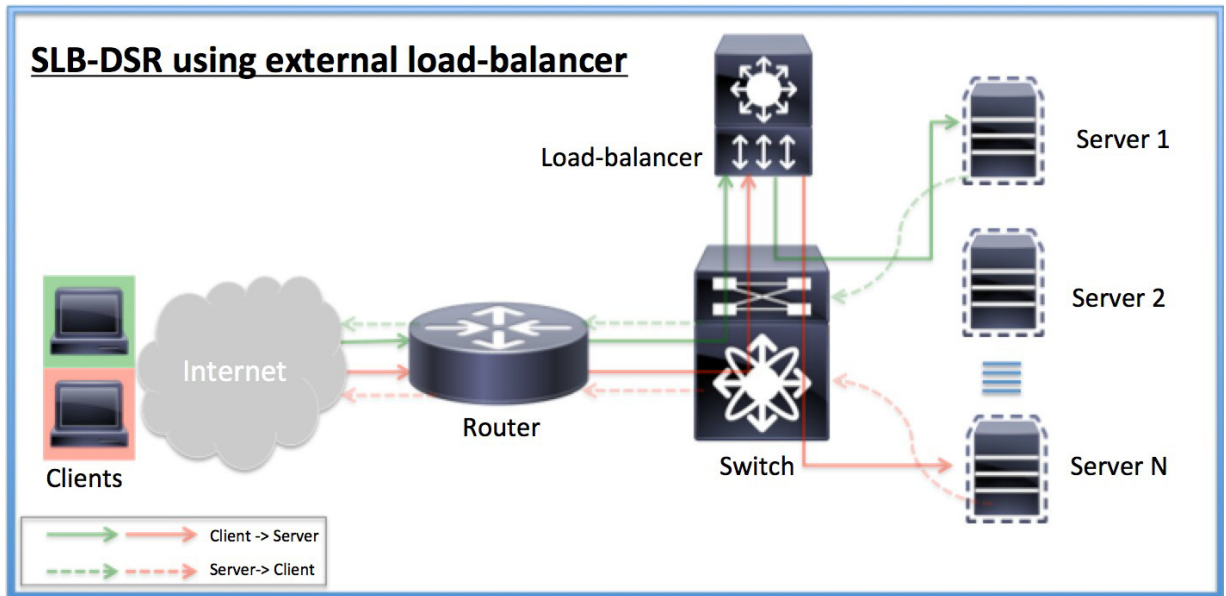


図 1 外部ロード バランサを使用する DSR モード

ただしこの方法では、ルーテッド/スイッチド ネットワーク デバイスに加えて、外部ロード バランサを個別に設定・管理する必要があります。また、冗長性を考慮すると、HA またはクラス タ化設計を実現するために複数のロード バランサを導入する必要もあります。

3. Nexus を使用する DSR モード:ITD

ITD を使用すれば、図 2 に示すように Cisco Nexus シリーズ スイッチでサーバへのトラフィック分散が可能になるため、外部ロード バランサは必要ありません。

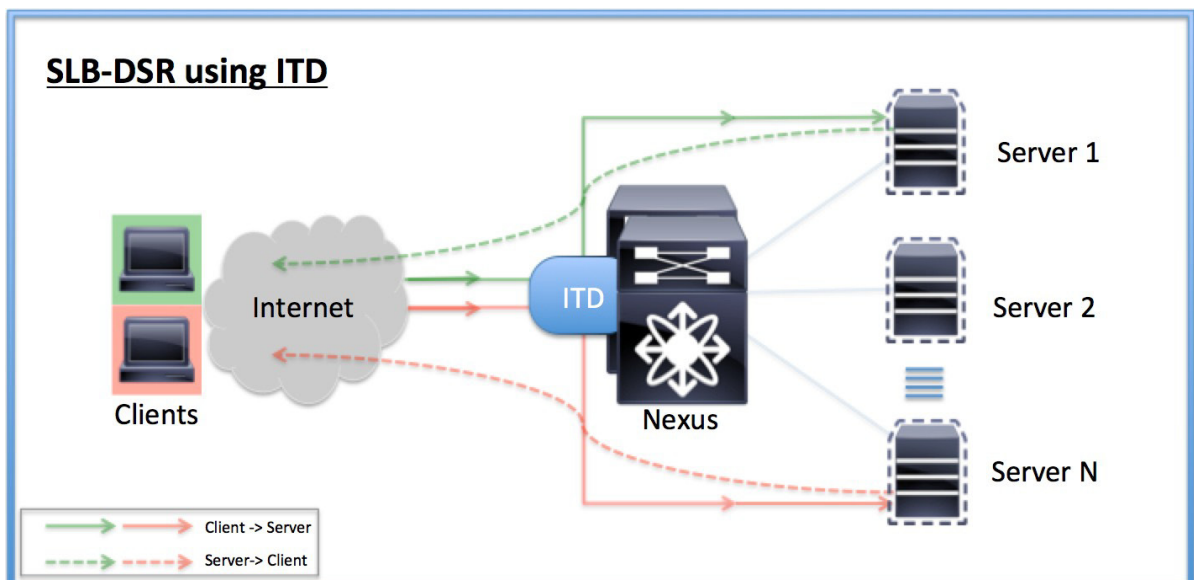


図 2:Nexus を使用する DSR モード:ITD

以前はロード バランサ上にあった仮想 IP が、ITD ポリシー内に設定され、サブネットへのルーティング用にアドバタイズできるようになりました。
また、ITD ではプローブによってノード障害の検出および高度なノード障害処理(ネットワーク要件に応じてカスタマイズ可能)を実現できます。

ITD の使用には次のような多くのメリットがあります。

- コスト削減:外部 SLB やアプリケーション配信コントローラ(ADC)は不要です。
- 制約のないパフォーマンス:ITD では、遅延を増やさずにマルチテラビット規模の容量を確保するため、ASIC ベースのラインレートトラフィック分散を活用します。
- 拡張性:ITD は多数のサーバに対応するように拡張可能です。

4. SLB 向けの ITD の導入: DSR

以下で説明する導入には、次のデバイスが使用されます。

- 1) Nexus 7700:7.2(0)D1(1) を実行する vPC モードで設定されたスイッチ 2 台。
- 2) Ubuntu Linux 14.04 を実行する仮想マシン(サーバ)2 台。
- 3) 同じく Ubuntu Linux 14.04 を実行する仮想マシン(クライアント)2 台。
- 4) レイヤ 2 スイッチ (VM と Nexus スイッチを接続)
- 5) サーバノードはテストおよび検証用に簡単な HTTP サービスをホストしています。

同じ構成は、Nexus 9000 シリーズ スイッチで ITD を使用した DSR モードでのサーバロード バランシングにも適用できます。Nexus 5000/6000 スイッチは、現在のところ、ITD プローブをサポートしていません。

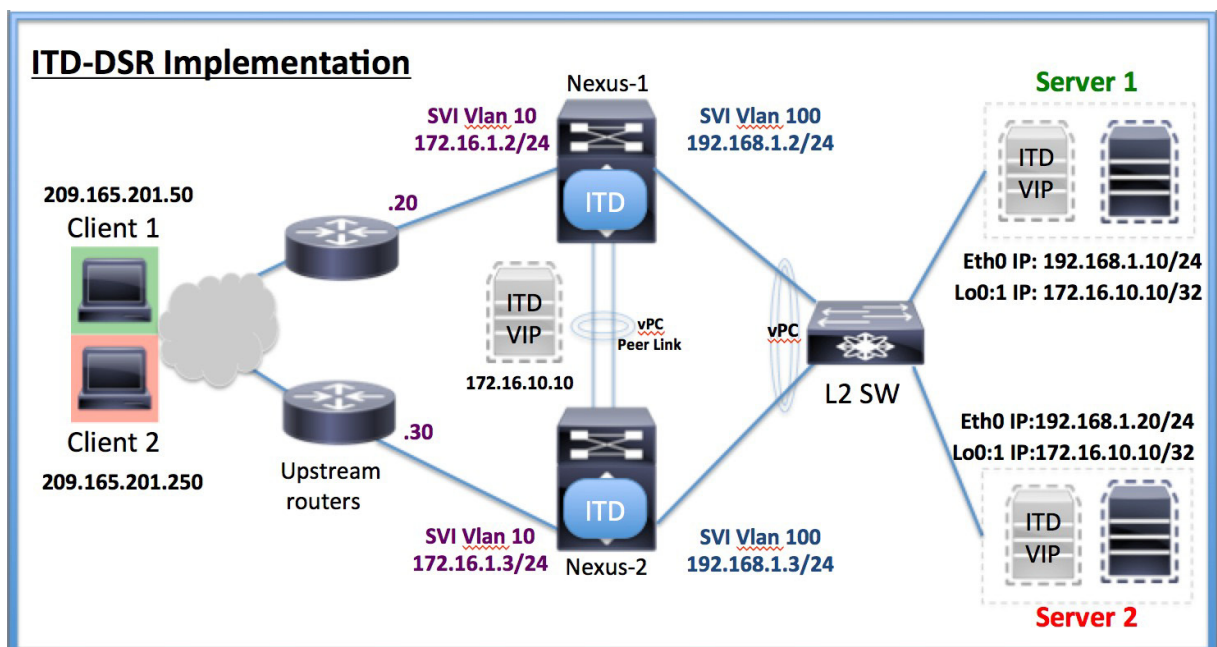


図 3: Nexus スイッチでの ITD による Direct Server Return の実装

注: この場合、クライアントをサーバと同じ VLAN 上に置くことはできません。また、仮想 IP が別のサブネット上にあるため、ARP の問題が回避されます。このサブネットは、ITD VIP「アドバタイズ」機能を使用してルーティング プロトコルによってアドバタイズできます。

4.1 サーバコンフィギュレーション

Direct Server Return では、サーバが仮想 IP アドレスを送信元 IP アドレスとして使用してクライアントに直接応答する必要があります。そのためには、サーバにサーバの通常の IP アドレスだけでなく VIP を設定する必要があります。

一般的な Linux マシンの場合、ループバックアドレスを使用して設定できます。

```
# sudo ifconfig lo:1 172.16.10.10 netmask 255.255.255.255 -arp up
```

このコマンドは、IP アドレスが 172.16.10.10 の新しい仮想ループバック インターフェイス「lo:1」を設定します。同じ仮想 IP が複数のデバイスに設定されているため、サーバが VIP に対する ARP にも応答した場合は重複 IP が検出される可能性があります。したがって、設定では「-arp」を使用してこのインターフェイスの ARP をディセーブルにします。

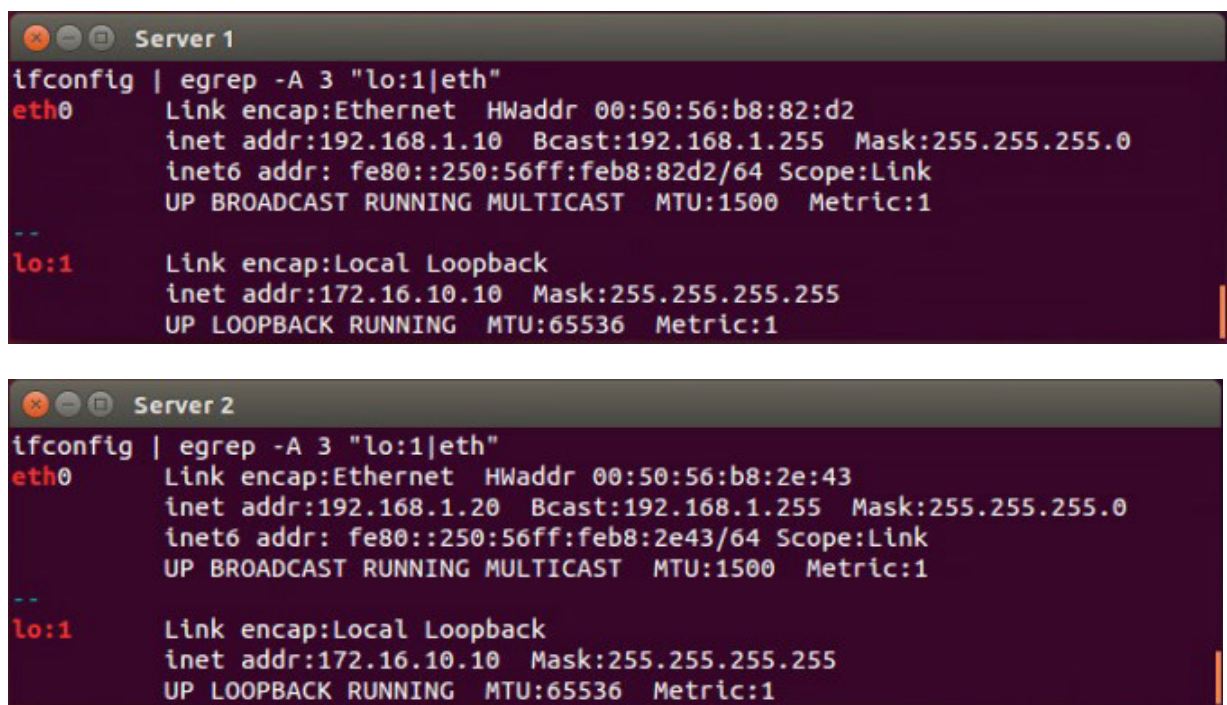
ところが Linux は、関連付けのない誤ったインターフェイスの ARP にもデフォルトで応答します。そのため、他のインターフェイス上のこの IP に対する ARP にサーバが応答しないように、次の設定も必要です。

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
```

上記の手順を使用して、設定内のすべてのサーバに同じ仮想 IP アドレスと対応する ARP の変更を設定する必要があります。

サーバの VIP 設定は、端末から「ifconfig」を使用して確認できます。



```
Server 1
ifconfig | egrep -A 3 "lo:1|eth"
eth0    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
--
lo:1    Link encap:Local Loopback
        inet addr:172.16.10.10  Mask:255.255.255.255
        UP LOOPBACK RUNNING  MTU:65536  Metric:1

Server 2
ifconfig | egrep -A 3 "lo:1|eth"
eth0    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
--
lo:1    Link encap:Local Loopback
        inet addr:172.16.10.10  Mask:255.255.255.255
        UP LOOPBACK RUNNING  MTU:65536  Metric:1
```

図 4:サーバ検証

4.2 Nexus ITD 設定

この導入例では、サーバ側への接続に Nexus スイッチの vPC を使用します。vPC はすでに Nexus スイッチに設定されているものとし、ここでは説明を省略します。スイッチ「Nexus-1」の設定を以下に示します。vPC ピア「Nexus-2」は、SVI IP アドレスを除き同様に設定する必要があります。

次の機能は ITD の前提条件であり、すでに使用されている他の機能に加えてイネーブルにする必要があります。

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

入力インターフェイス SVI VLAN 10 と「server-group」SVI Vlan100 が作成されます。

```
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
```

デバイスグループ「server-group100」が作成され、これにサーバの物理インターフェイス IP アドレスが追加されます。また、ICMP プロブがデフォルト値でこのデバイスグループに設定されます。プローブ タイマーは必要に応じて調整できます。Nexus 5000/Nexus 9000 のプローブについては、注意事項と制約事項の項を参照してください。

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

作成したデバイスグループ「server-group100」を使用するように ITD サービス「vip-dsr100」を設定します。アップストリーム ルーティング プロトコルに VIP をアドバタイズするには、アドバタイズメントをイネーブルにして仮想 IP 172.16.10.10 を ITD サービスに設定します。「送信元 IP」のロード バランス方式を選択します。コマンド「*Failaction node reassign*」を明示的に設定して、障害が発生したノードのトラフィック バケットの再割り当てをイネーブルにする必要があることに注意してください(詳細については参考資料を参照)。最後に、ITD トラフィック分散統計情報を表示するために、統計情報をイネーブルにする必要があります。

```
itd vip-dsr100
  device-group server-group100
  virtual ip 172.16.10.10 255.255.255.255 advertise enable
  ingress interface Vlan10
  failaction node reassign
  load-balance method src ip
  no shut
itd statistics DSRService
```

4.3 確認

次のように、設定した ITD サービスを確認できます。

```
PSK_N7700_1-ITD-DSR(config-itd)# sh itd
Legend:
ST(Status): ST-Standby,LF-Link Failed,PF-Probe Failed,PD-Peer Down,IA-Inactive

Name          LB Scheme  Status  Buckets
-----
vip-dsr100    src-ip     ACTIVE  2

Exclude ACL
-----

Device Group          Probe  Port
-----
server-group100      ICMP

Pool                  Interface  Status  Track_id
-----
vip-dsr100_itd_pool  Vlan10    UP      3

Virtual IP            Netmask/Prefix  Protocol  Port
-----
172.16.10.10 / 255.255.255.255          IP        0

Node  IP          Cfg-S  WGT  Probe  Port  Probe-IP  STS  Trk#  Sla_id
-----
1     192.168.1.10  Active  1    ICMP          OK    1    10001

Bucket List
-----
vip-dsr100_itd_vip_1_bucket_1

Node  IP          Cfg-S  WGT  Probe  Port  Probe-IP  STS  Trk#  Sla_id
-----
2     192.168.1.20  Active  1    ICMP          OK    2    10002

Bucket List
-----
vip-dsr100_itd_vip_1_bucket_2
```

図 5:Nexus ITD 検証

クライアント VM から、次に示すように仮想 IP の到達可能性および HTTP サービスを確認できます。区別するために、名前にサーバ番号と IP アドレスが付いたフォルダがサーバごとに作成されています。

ITD ではデフォルトで最後のオクテットまたは最下位ビット (LSB) がバケットに使用されるため、スキューされた IP アドレス .50 および .250 は、クライアントが動作中の負荷分散を示すために使用されています。

クライアント 1:VIP への ping および HTTP テスト

```
Client 1
64 bytes from 172.16.10.10: icmp_seq=142 ttl=63 time=0.447 ms
64 bytes from 172.16.10.10: icmp_seq=143 ttl=63 time=0.504 ms
64 bytes from 172.16.10.10: icmp_seq=144 ttl=63 time=0.525 ms
64 bytes from 172.16.10.10: icmp_seq=145 ttl=63 time=0.409 ms
64 bytes from 172.16.10.10: icmp_seq=146 ttl=63 time=0.405 ms
64 bytes from 172.16.10.10: icmp_seq=147 ttl=63 time=0.535 ms
64 bytes from 172.16.10.10: icmp_seq=148 ttl=63 time=0.339 ms
64 bytes from 172.16.10.10: icmp_seq=149 ttl=63 time=0.480 ms
```



クライアント 2:VIP への ping および HTTP テスト

```
Client 2
64 bytes from 172.16.10.10: icmp_seq=214 ttl=63 time=0.660 ms
64 bytes from 172.16.10.10: icmp_seq=215 ttl=63 time=0.390 ms
64 bytes from 172.16.10.10: icmp_seq=216 ttl=63 time=0.436 ms
64 bytes from 172.16.10.10: icmp_seq=217 ttl=63 time=0.599 ms
64 bytes from 172.16.10.10: icmp_seq=218 ttl=63 time=0.383 ms
64 bytes from 172.16.10.10: icmp_seq=219 ttl=63 time=0.365 ms
64 bytes from 172.16.10.10: icmp_seq=220 ttl=63 time=0.589 ms
64 bytes from 172.16.10.10: icmp_seq=221 ttl=63 time=0.392 ms
```



図 6:クライアント サービス検証

Nexus 7700 の CLI を使用して、ITD のリダイレクトされたトラフィックの統計情報を確認できます。

```
PSK_N7700_1-ITD-DSR# show itd vip-dsr100 statistics
```

Service	Device Group	VIP/mask	#Packets
vip-dsr100	server-group100	172.16.10.10 / 255.255.255.255	98 (100.00%)
Traffic Bucket		Assigned to	Mode
vip-dsr100_itd_vip_1_bucket_1	192.168.1.10	Redirect	Original Node
Traffic Bucket		Assigned to	Mode
vip-dsr100_itd_vip_1_bucket_2	192.168.1.20	Redirect	Original Node

図 7:ITD 統計情報

5. 注意事項と制約事項

- サーバノード上のループバックの設定は、サーバが他のオペレーティングシステム (Windows など) を実行している場合も DSR モードが機能するために必須です。
- Nexus vPC ピア デバイスが正常に動作するように、ITD サービスの設定は同一である必要があります。これには、ロードバランス方式やプローブの指定などのパラメータだけでなく、ノードの番号、順番、設定も含まれます。
- スタンバイが存在せず、Failaction 再割り当てが設定されていない状態でノードに障害が発生すると、通常、トラフィックは ITD リダイレクションなしでルーティングされます。
- サーバノードがポートチャネルを使用する複数の NIC に対応している場合、Nexus スイッチを vPC 経由でサーバに直接接続できます。
- ノードの重みを使用して、トラフィックを不均一に分散することができます (容量の異なるサーバを使用する場合)。
- リリース 7.2(0)D1(1) 以降では、必要に応じてプローブをノードごとに設定できます。
- Nexus 9000 の ITD-ICMP プローブでは、現在、機能「SLA レスポンダ/送信者」は前提条件として必須ではありません。
- Nexus 5000/6000 シリーズ スイッチは、現在のところ、ITD プローブをサポートしていません。

6. 参考資料と詳細記事

Linux での仮想インターフェイスの作成:

<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 の考慮事項: <http://linux-ip.net/html/ether-arp.html-ether-arp-flux>

Nexus 7000 ITD コンフィギュレーションガイド: [Nexus 7000 / 7.x / ITD Config Guide](#)

Nexus 9000 ITD コンフィギュレーションガイド: [Nexus 9000 / 7.x / ITD Config Guide](#)

Nexus 5500 ITD コンフィギュレーションガイド: [Nexus 5500 ITD Config Guide](#)

Nexus 5600 ITD コンフィギュレーションガイド: [Nexus 5600 ITD Config Guide](#)