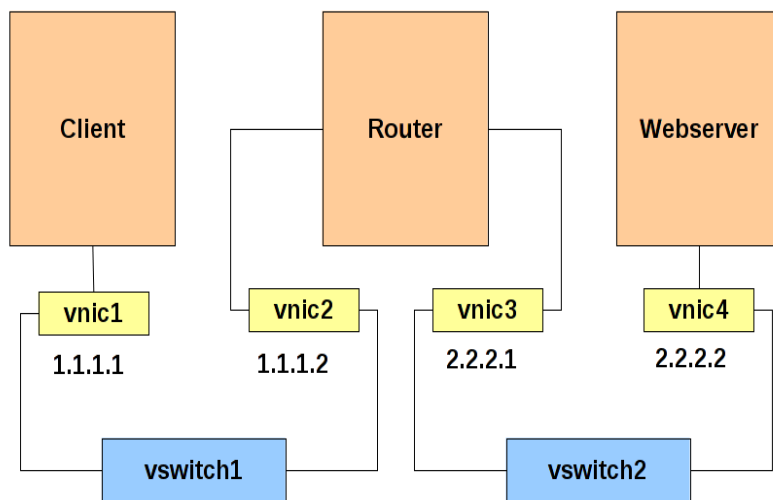


更新日	2012/05/20 shoji.haraguchi@oracle.com
難易度/危険度	★★★★☆/★★★★☆
root 権限の必要性	有り

Solaris 11 仮想ワイヤ どのように見える化?



Oracle Solaris 11 で仮想ワイヤの構築とネットフロー制御のデモを行います。上図のような仮想ネットワークを Zone と Crossbow を使って構築します。Client Zone からネットワークベンチ (netperf) を流し、DTrace や flowstat コマンドをつかってフロー制御の見える化に挑戦してみます。

<事前の仕込み>

短時間でデモ構成を作成するため Zone を複製で作成します。複製の元になる templatezone 作成だけ事前に仕込んでおきます。

```
$ sudo zfs create rpool/Zones
$ sudo zonecfg -z templatezone "create;set
zonepath=/rpool/Zones/templatezone;end;commit;exit"
$ sudo zoneadm -z templatezone install
```

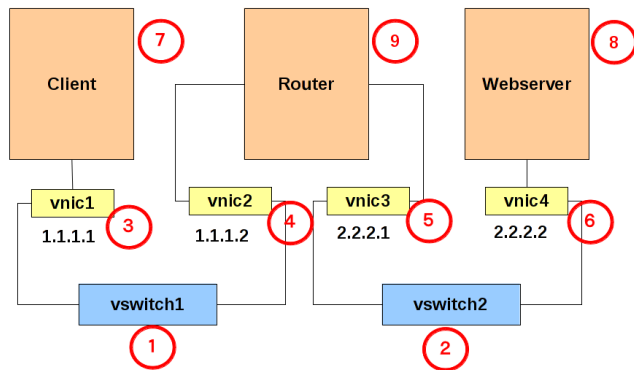
templatezone には、Webserver を構築するため apache-22 を追加インストールしてください。

```
templatezone $ sudo pkg install web/server/apache-22
```

また templatezone には、netperf/netserver また、各 Zone を自動再構成するために、/var/tmp の下にシステム構成のプロファイルもあらかじめセットしてあります。各プロファイルは巻末の付録を参照してください。

Solaris 11 仮想ワイヤ

どのように見える化？



<仮想ワイヤの構成手順> 上図の番号の順番で設定していきます。

(1)-(2) 仮想スイッチの設定

```
$ sudo dladm create-etherstub vswitch1
$ sudo dladm create-etherstub vswitch2
```

(3)(4)(5)(6) VNIC の設定

```
$ sudo dladm create-vnic -l vswitch1 vnic1
$ sudo dladm create-vnic -l vswitch1 vnic2
$ sudo dladm create-vnic -l vswitch2 vnic3
$ sudo dladm create-vnic -l vswitch2 vnic4
```

(7) Client Zone の作成 vnic1 をアサインします。

```
$ sudo zonecfg -z client "create;set
zonepath=/rpool/Zones/client;add net;set
physical=vnic1;end;commit;exit"
$ sudo zoneadm -z client clone templatezone
$ sudo zoneadm -z client boot
```

(8) Webserver Zone の作成 vnic4 をアサインします。

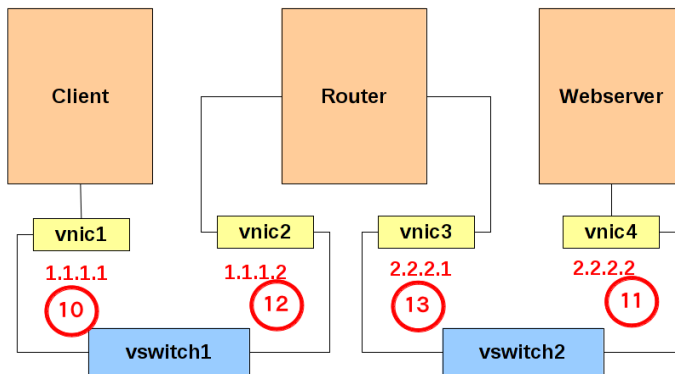
```
$ sudo zonecfg -z webserver "create;set
zonepath=/rpool/Zones/webserver;add net;set
physical=vnic4;end;commit;exit"
$ sudo zoneadm -z webserver clone templatezone
$ sudo zoneadm -z webserver boot
```

(9) Router Zone の作成 ルータには2つの VNIC vnic2 vnic3 をアサインします。

```
$ sudo zonecfg -z router "create;set
zonepath=/rpool/Zones/router;add net;set physical=vnic2;end;add
net;set physical=vnic3;end;commit;exit"
$ sudo zoneadm -z router clone templatezone
$ sudo zoneadm -z router boot
```

Solaris 11 仮想ワイヤ

どのように見える化？



- (10) sysconfig(1M) を使って Client Zone を自動構成します。vnic1 に IP 1.1.1.1 を設定します。

```
$ sudo zlogin client sysconfig configure -g system -c /var/tmp/client.xml
$ sudo zoneadm -z client reboot
```
- (11) sysconfig(1M) を使って Webservice Zone を自動構成します。vnic4 に IP 2.2.2.2 を設定します。

```
$ sudo zlogin webserver sysconfig configure -g system -c /var/tmp/webserver.xml
$ sudo zoneadm -z webserver reboot
```
- (12) sysconfig(1M) を使って Router Zone を自動構成します。vnic2 に IP 1.1.1.2 を設定します。

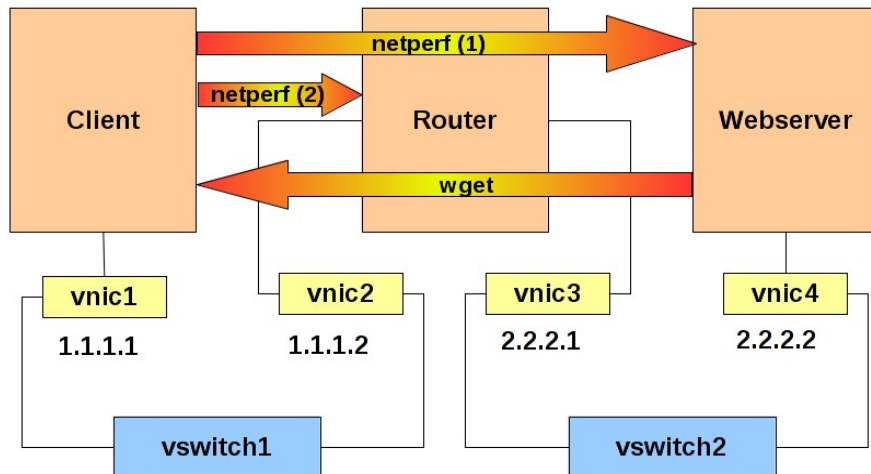
```
$ sudo zlogin router sysconfig configure -g system -c /var/tmp/router.xml
$ sudo zoneadm -z router reboot
```
- (13) 再構成後、Router Zone で、vnic3 の設定を行い、ルータ設定を行います。

```
$ sudo zlogin router ipadm create-ip vnic3
$ sudo zlogin router ipadm create-addr -T static -a 2.2.2.1/24 vnic3/v4
$ sudo zlogin router routeadm -e ipv4-forwarding
$ sudo zlogin router routeadm -u
```

<フロー制御>

仮想ワイヤが構築できたら、フロー制御のデモを行います。

Solaris 11 仮想ワイヤ どのように見える化?



仮想ワイヤで可能なフロー制御は2種類あります

- 1) dladm(1M) 仮想 NIC のリンクプロパティによる制御
mtu,maxbw,priority 値の変更で制御可能
- 2) flowadm(1M)による制御
IP アドレス、プロトコル、ポート番号毎に、maxbw 値を制御可能

準備: Webserver Zone で Apache22 を起動します。

```
$ sudo zlogin webserver svcadm enable apache22
```

準備 Webserver Zone で netperf 用の netserver を起動

```
$ sudo zlogin webserver netserver
```

準備: Router Zone で netperf 用の netserver を起動

```
$ sudo zlogin router netserver
```

準備: DTrace (Chime) で、各 VNIC の受信量を監視しておきます。

```
$ sudo /opt/OSOL0chime/bin/chime -n 'ip:::receive  
{@receive[args[3]->if_name] = sum(args[2]->ip_plength); }' &
```

Client Zone から Webserver Zone に netperf を流してみます この例では 3000bps ほど出ています

```
$ sudo zlogin client netperf -H 2.2.2.2  
TCP STREAM TEST to 2.2.2.2 : histogram  
Recv  Send  Send  
Socket Socket Message Elapsed  
Size  Size  Size  Time  Throughput  
bytes bytes bytes secs.  10^6bits/sec  
128000 49152 49152 10.00 3146.37
```

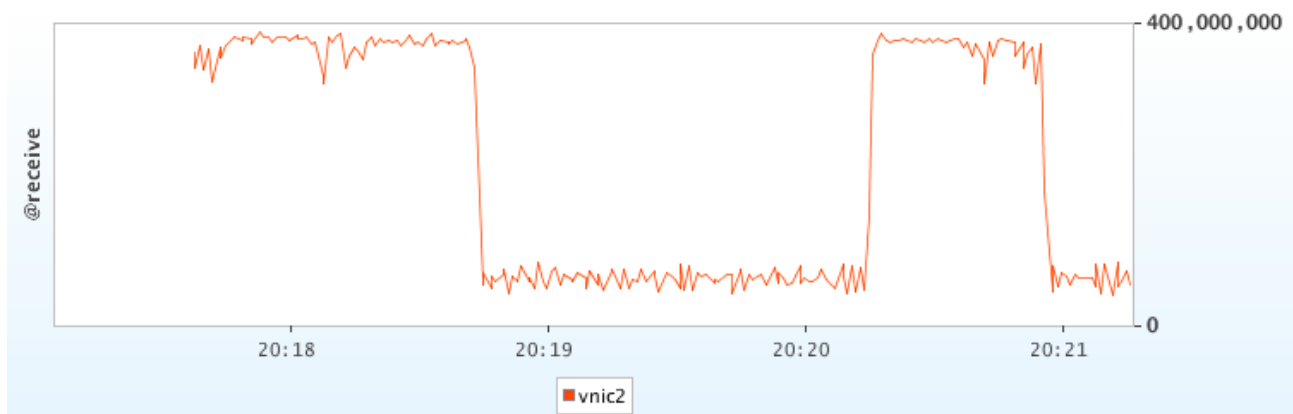
dladm で、Webserver の vnic4 を 500bps に制御してみます。

```
$ sudo dladm set-linkprop -p maxbw=500 vnic4
$ sudo zlogin client netperf -H 2.2.2.2
TCP STREAM TEST to 2.2.2.2 : histogram
Recv  Send      Send
Socket Socket  Message  Elapsed
Size  Size    Size      Time      Throughput
bytes bytes   bytes     secs.     10^6bits/sec

128000 49152 49152    10.40     517.43
```

netperf を 2000 秒流し、制御の変化を動的に監視します。

```
$ sudo zlogin client netperf -H 2.2.2.2 -l 2000 &
$ sudo dladm set-linkprop -p maxbw=500 vnic2
```



netperf を Router Zone にも流します。

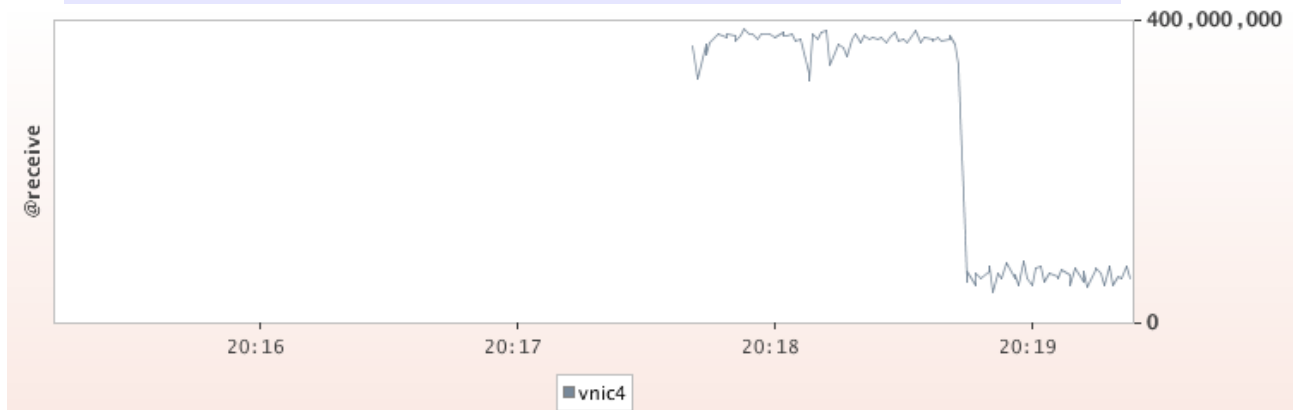
```
$ sudo dladm set-linkprop -p maxbw=5000 vnic2
$ sudo zlogin client netperf -H 1.1.1.2 -l 2000 &
```

Client Zone 上で、flowadm によって 両方の netperf のフローを制御するフローを作成します。

```
$ sudo zlogin client netstat
$ sudo zlogin client flowadm add-flow -l vnic1 -a
transport=tcp,local=port=43318 flow22
$ sudo zlogin client flowadm add-flow -l vnic1 -a
transport=tcp,local=port=64108 flow12
```

flowadm で 二つの netperf のフローを制御してみます。

```
$ sudo zlogin client flowadm set-flowprop -p maxbw=500 flow22
$ sudo zlogin client flowadm set-flowprop -p maxbw=5000 flow22
$ sudo zlogin client flowadm set-flowprop -p maxbw=500 flow12
$ sudo zlogin client flowadm set-flowprop -p maxbw=5000 flow12
```



Webserver 上の http のフローを制御してみます。

```
$ sudo zlogin webserver flowadm add-flow -l vnic4 -a
transport=tcp,local_port=80 httpflow
```

Client Zone から wget で iso イメージをダウンロードしてみます。

```
$ sudo zlogin client wget http://2.2.2.2/sol-11-1111-live-x86.usb
```

http のフローを 5m bps 程度に抑えてみます。

```
$ sudo zlogin webserver flowadm set-flowprop -p maxbw=50 httpflow
```

もとに戻します。

```
$ sudo zlogin webserver flowadm set-flowprop -p maxbw=5000
httpflow
```

<あとかたづけ>

```
$ sudo zlogin webserver flowadm remove-flow httpflow
```

```
$ sudo zlogin client flowadm remove-flow flow12
$ sudo zlogin client flowadm remove-flow flow12
```

```
$ sudo zoneadm -z router halt
$ sudo zoneadm -z webserver halt
$ sudo zoneadm -z client halt
```

```
$ sudo zoneadm -z router uninstall -F
$ sudo zoneadm -z webserver uninstall -F
$ sudo zoneadm -z client uninstall -F
```

```
$ sudo zonecfg -z router delete -F
$ sudo zonecfg -z webserver delete -F
$ sudo zonecfg -z client delete -F
```

```
$ sudo dladm delete-vnic vnic4
$ sudo dladm delete-vnic vnic3
$ sudo dladm delete-vnic vnic2
$ sudo dladm delete-vnic vnic1
```

```
$ sudo dladm delete-etherstub vswitch2
$ sudo dladm delete-etherstub vswitch1
```

<Tips>

1) 仮想ワイヤを構成した場合、仮想スイッチや仮想 NIC(VNIC) の MTU 値のデフォルトは、9000 に設定されています。現実的な速度で検証する際は、576 (ダイヤルアップ) から 1500 (Ether) の値に設定するといいでしょ。尚、VNIC だけ MTU の設定変更しても有効にはなりません、必ず仮想スイッチもおなじ MTU に設定してください。

2) VNIC の mtu や maxbw は通信中でも動的に変更可能ですが、仮想スイッチのプロパティを通信中に変更すると通信が切れてしまうことがあるようです。

3) dladm によるフロー制御は、Global Zone でのみ制御可能です。また flowadm による Local Zone のフロー制御を動的に行うには、Local Zone でのみ制御可能となります

< 参考資料 >

ネットワークの仮想化とリソース管理: http://docs.oracle.com/cd/E26924_01/html/E25834/gdytf.html#scrolltoc

Netperf: <http://www.netperf.org/netperf/>

Chime: <http://hub.opensolaris.org/bin/view/Project+dtrace-chime/>

ブログ: Solaris 3 分クッキング <https://blogs.oracle.com/hara/>

<付録>

sysconfig 用 profile, この設定の場合、default user-id; abc123 passwd: abc123, root passwd: abc123 になります。

/ var/tmp/client.xml

```
<!DOCTYPE service_bundle SYSTEM "/usr/share/lib/xml/dtd/service_bundle.dtd.1">
<service_bundle type="profile" name="sysconfig">
  <service version="1" type="service" name="system/config-user">
    <instance enabled="true" name="default">
      <property_group type="application" name="root_account">
        <propval type="astring" name="login" value="root"/>
        <propval type="astring" name="password"
value="$5$Wjac/s7j$NCcniP2FCn5QlcjxOp.ipGXFmNIcn1sGm8Bb9S78Ck."/>
        <propval type="astring" name="type" value="role"/>
      </property_group>
      <property_group type="application" name="user_account">
        <propval type="astring" name="login" value="abc123"/>
        <propval type="astring" name="password"
value="$5$TbVigXdp$Qt2PwKPzv7Z59mSicfcln7h006zBBt8wFEcD0a7Y5UD"/>
        <propval type="astring" name="type" value="normal"/>
        <propval type="astring" name="description" value="abc123"/>
        <propval type="count" name="gid" value="10"/>
        <propval type="astring" name="shell" value="/usr/bin/bash"/>
        <propval type="astring" name="roles" value="root"/>
        <propval type="astring" name="profiles" value="System Administrator"/>
        <propval type="astring" name="sudoers" value="ALL=(ALL) ALL"/>
      </property_group>
    </instance>
  </service>
  <service version="1" type="service" name="system/timezone">
    <instance enabled="true" name="default">
      <property_group type="application" name="timezone">
        <propval type="astring" name="localtime" value="Japan"/>
      </property_group>
    </instance>
  </service>
  <service version="1" type="service" name="system/environment">
    <instance enabled="true" name="init">
      <property_group type="application" name="environment">
        <propval type="astring" name="LANG" value="C"/>
      </property_group>
    </instance>
  </service>
  <service version="1" type="service" name="system/identity">
    <instance enabled="true" name="node">
      <property_group type="application" name="config">
        <propval type="astring" name="nodename" value="client"/>
      </property_group>
    </instance>
  </service>
  <service version="1" type="service" name="system/keymap">
    <instance enabled="true" name="default">
      <property_group type="system" name="keymap">
        <propval type="astring" name="layout" value="US-English"/>
      </property_group>
    </instance>
  </service>
  <service version="1" type="service" name="system/console-login">
    <instance enabled="true" name="default">
      <property_group type="application" name="ttymon">
        <propval type="astring" name="terminal_type" value="vt100"/>
      </property_group>
    </instance>
  </service>
</service_bundle>
```

```

    </property_group>
  </instance>
</service>
<service version="1" type="service" name="network/physical">
  <instance enabled="true" name="default">
    <property_group type="application" name="netcfg">
      <propval type="astring" name="active_ncp" value="DefaultFixed"/>
    </property_group>
  </instance>
</service>
<service version="1" type="service" name="network/install">
  <instance enabled="true" name="default">
    <property_group type="application" name="install_ipv4_interface">
      <propval type="astring" name="address_type" value="static"/>
      <propval type="net_address_v4" name="static_address" value="1.1.1.1/24"/>
      <propval type="astring" name="name" value="vnic1/v4"/>
    </property_group>
    <property_group type="application" name="install_ipv6_interface">
      <propval type="astring" name="stateful" value="yes"/>
      <propval type="astring" name="stateless" value="yes"/>
      <propval type="astring" name="address_type" value="addrconf"/>
      <propval type="astring" name="name" value="vnic1/v6"/>
    </property_group>
  </instance>
</service>
<service version="1" type="service" name="system/name-service/switch">
  <property_group type="application" name="config">
    <propval type="astring" name="default" value="files"/>
    <propval type="astring" name="printer" value="user files"/>
  </property_group>
  <instance enabled="true" name="default"/>
</service>
<service version="1" type="service" name="system/name-service/cache">
  <instance enabled="true" name="default"/>
</service>
<service version="1" type="service" name="network/dns/client">
  <instance enabled="false" name="default"/>
</service>
</service_bundle>

```

/var/tmp/webserver.xml と /var/tmp/router.xml も同様の内容になります。下記の箇所をそれぞれに変更してください。

```

<propval type="astring" name="nodename" value="client"/>
<propval type="net_address_v4" name="static_address" value="1.1.1.1/24"/>
<propval type="astring" name="name" value="vnic1/v4"/>

```