Recommended Resilient Campus Network Design

Best Practice Document

Produced by CESNET led working group
on Network monitoring
(CBPD114)

Authors: Tomas Podermanski, Vladimir Zahorik
March 2010

Document No: GN3-NA3-T4-CBPD114
Version / date: 24. 03. 2010
Original language: Czech
Original title: “Recommended Resilient Campus Network Design”
Original version / date: 1.2 of 3. 12. 2009
Contact: tpoder at cis.vutbr.cz, zahorik at cis.vutbr.cz

CESNET bears responsibility for the content of this document. The work has been carried out by a CESNET led working group on Network monitoring group as part of a joint-venture project within the HE sector in Czech Republic.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 23 8875, relating to the project 'Multi-Gigabit European Research and Education Network and Associated Services (GN3)'.

Table of Contents

Executive Summary

This document describes how to setup a fully resilient network design in a campus. The recommendation for standards and proper technologies are discussed. Descriptions of all the parts - core network, distribution switches and resilient server connections are described.

The main idea of resilient topology is to eliminate downtime during crashes and device upgrades. This document describes how all critical devices are used twice to avoid having a single point of failure. Therefore, any single device can be turned off without significant disruption for the connected applications and users.

The use of standardised protocols is encouraged throughout the document. This allows devices offered by different suppliers to interoperate. A further requirement was to keep the configuration as simple as possible.

1 Basic Components in Resilient Network Design

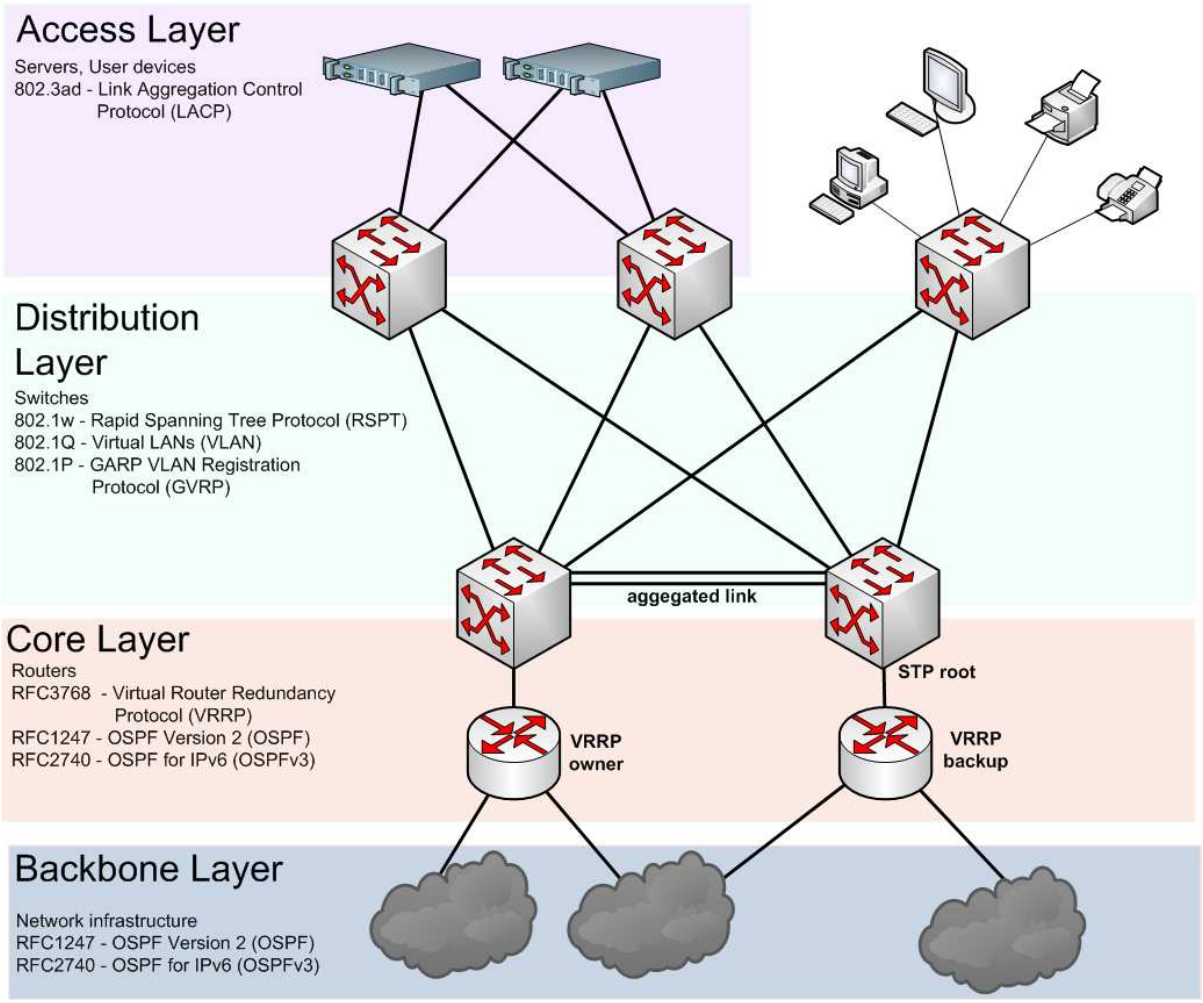

Figure 1: Basic components in the resilient network design

Network topology is divided into several layers. Each one uses a different set of protocols and requires different configuration options. The following layers are used:

1.1 Access layer

This layer is used to connect end devices (servers, user's computers, printers, etc). We presume that servers are connected via two network adapters. For other devices which do not support resilient network connections we terminate the resilient topology on the last switch. For almost all devices which are connected through a single Ethernet port, special protocols are not required.

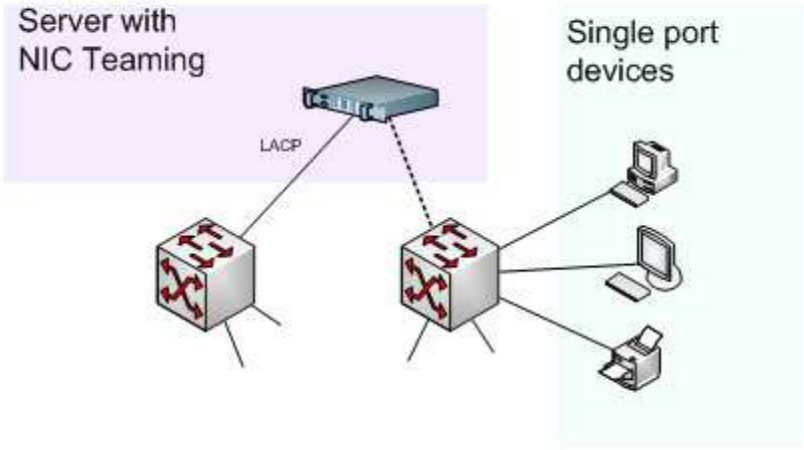

Figure 2: Connecting the end-user devices

A slightly different situation is with a server connection where we would like to connect each one using two independent Ethernet ports. In that case we have two general setup choices on both the server and on the switch side.

- Base the links between access switches and servers on the *LACP* protocol. This solution allows more intelligent detection of the failed link and better opportunities to achieve load balancing between links. The disadvantage of this solution is a more complicated configuration on the switch side. It is necessary to create a virtual *LACP* trunk interface on each port where a server is connected.
- Using interface teaming (bonding) on the server side without using a link control protocol mechanism. In this case the failure detection is based only on a single link. There is some support for load balancing, but the balancing can only be done in the transmit direction. The main advantage of this solution is that there is no configuration required on the switches. You can have the same configuration of interfaces for both resilient and non-resilient devices.

A special feature is required on the server side. This is usually represented by a virtual interface which joins two or more physical interfaces. This feature is usually known as Ethernet Teaming or Ethernet Bonding. Almost all modern unix-like systems support it natively. Unfortunately, Microsoft™ systems do not support it, but there are many solutions developed by the NIC producers to enable this within an MS Infrastructure.

Virtualisation technology (e.g. VmWare) allows another form of simplification. The Virtual servers are usually connected through a virtual switch. You can create more virtual switches within the virtual appliance, but such solutions do not present any real benefits. A better solution is to connect the virtual switch via teaming and provide a single virtual interface to the virtual host environment. In this case the virtual host is no longer

concerned with issues related to resilience in the network. Virtualisation can also be used to solve problems where systems do not support interface teaming.

- http://en.wikipedia.org/wiki/Link_aggregation, Basic information
- http://www.linux-corner.info/bonding.html, Ethernet Bonding in Linux
- http://www.freebsd.org/doc/en/books/handbook/network-aggregation.html, Link aggregation in *FreeBSD*
- http://www.vmware.com/files/pdf/virtual_networking_concepts.pdf, Networking concepts in VmWare

1.2 The distribution layer

This layer primarily provides *L2* distribution through the switched network. This distribution can be solved by any *L2* loop-free protocol. Among the most widely used solutions are the Spanning Tree Protocol, or some of the more developed protocols (e.g. *RSTP, MSTP*).

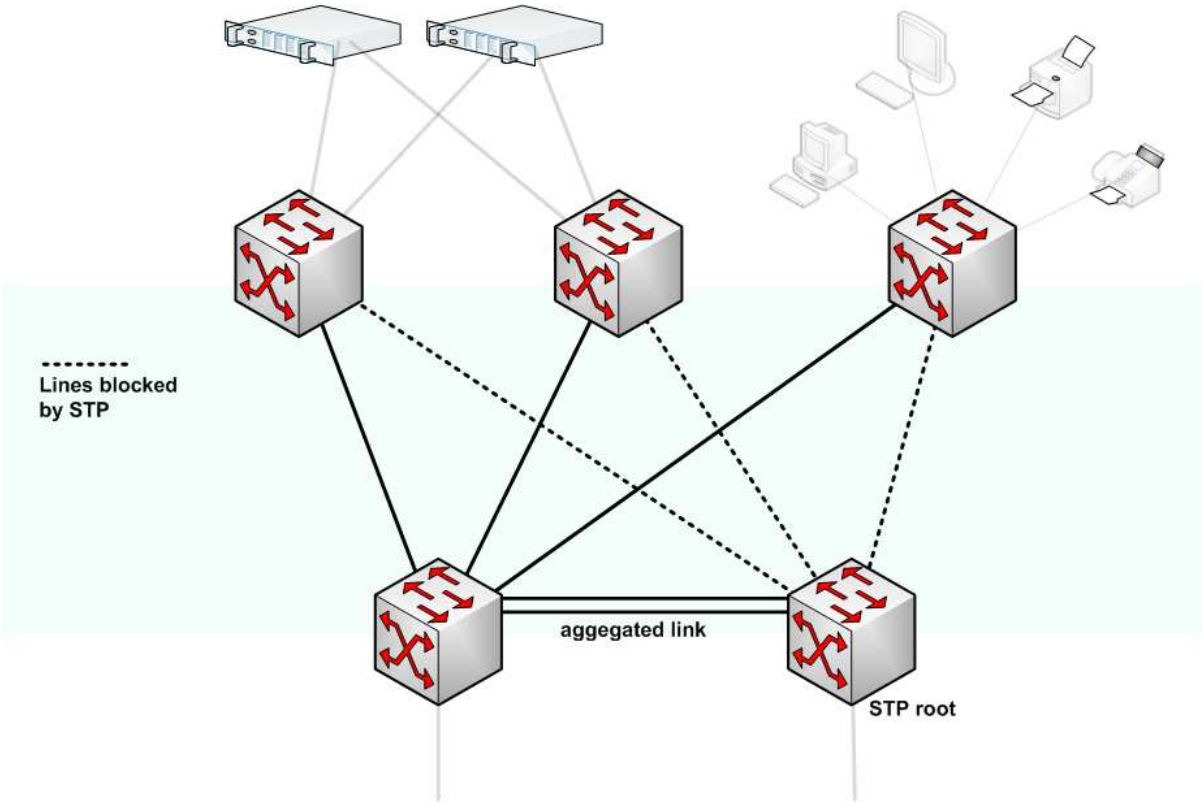

Figure 3: The distribution layer components

The configuration example described in this document uses the Rapid Spanning Tree Protocol (*RSTP*) to get rid of topology loops. The *GVRP* protocol is used to simplify the spread of the 802.1Q VLAN configuration.

The link aggregation protocol can also be used to aggregate multiple uplinks between switches. In these cases *LACP* or port trunking can be used.

The topology of the distribution layer should be designed so that there is no ring containing more access switches. The obvious solution is to design a topology where each access switch is connected to their core

switch by two independent uplinks. That allows the creation of a more transparent topology and minimises the number of *VLANs* on the access switches.

The following table summarises protocols used on the distribution layer:

Protocol	Function
RSTP	Solves Ethernet loops and prunes links to get rid of loops. If you need to split traffic more effectively you can use *MSTP*. The downside is a more complicated configuration.
GVRP	*GVRP* is used to simplify *VLAN* configuration. This protocol can replace the static *VLAN* configuration on devices.
LACP	*LACP* or port trunking is used to create aggregated links between switches. The protocol can be also used to make both link and traffic management more effective between access switches and servers.

Recommendation:
To avoid unrequested transit traffic on uplinks to switches, it is necessary to have *STP* root access on the device which acts as the *VRRP* owner.

Recommendation:
On the access switches, setup the option “unknown-vlan” to “blocked”. This configuration minimizes the number of *VLANs* on the access switches and creates a more transparent topology.

1.3 Core layer

Core switches and core routers are placed in the core layer. This ensures *L3* availability of the gateway IP address when the one of the routers goes down. In most cases the routing and switching capability will be integrated into one device.

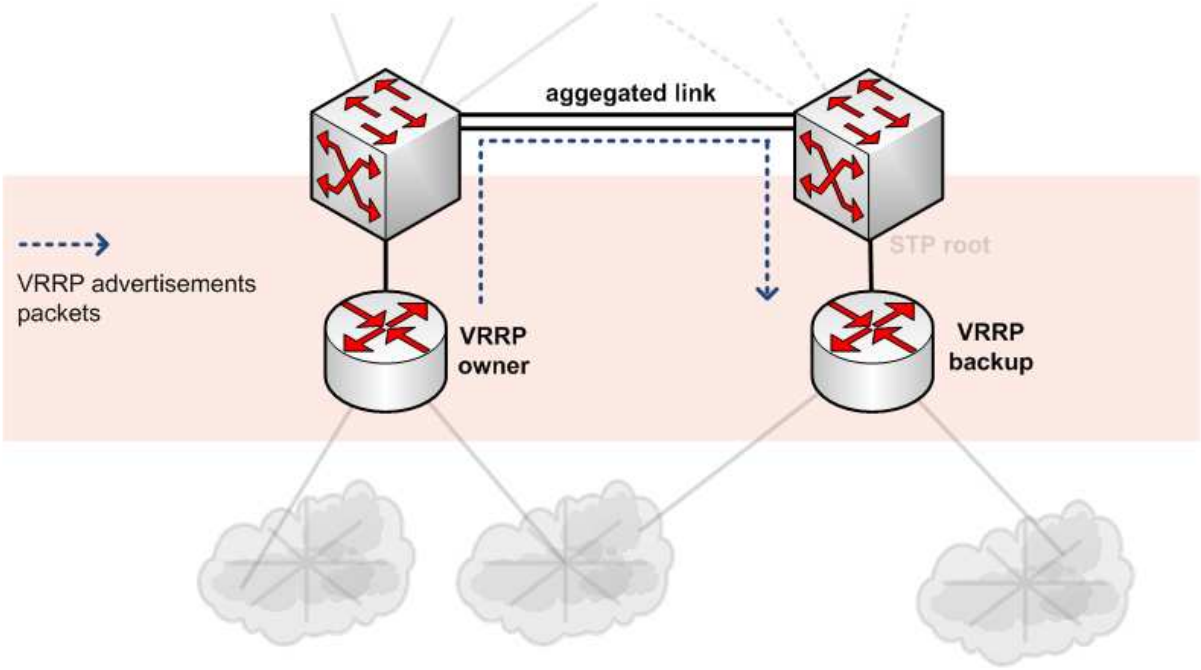

Figure 4: The core layer components

The *VRRP* protocol (*RFC 3768*) is used to backup the subnet gateway address. In *VRRP* there is a Master and a Backup router. The router which acts as Owner (i.e. the router which is currently sending the advertisement packets and acting as the gateway etc, normally the Master router) periodically sends advertisement packets to the network and the Backup router listens for them. If the Backup router does not receive three consecutive advertisement packets from the Owner then it assumes the gateway address and acts as the Owner router until a new advertisement packet is received. Therefore, the current active gateway router is known as the Owner, and this can apply to either the Master or the Backup router, depending on whether the Master router is functioning properly.

Recommendation:
Use the first *IP* address from the network range for the Master router. That address will also act as the gateway address. The second address should be used for the Backup router.

There must also be coordination with the routing protocol. Network changes when the gateway address is moved between routers must be correctly propagated. This problem is particularly visible when the Master router goes back to the "Owner" state. The Master can assume the gateway address, but the routing protocol is not still converged. There is an option called "preempt-delay-time" which allows one to define a time period to postpone the gateway address transfer.

Recommendation:
Do not forget to set up "preempt-delay-time" to postpone transfer of the gateway address. The value should be consistent with the routing protocol used on the backbone layer.

The interoperation with the routing protocol brings another problem. The only mechanism by which the Backup router detects failure is by failing to receive the advertisement packets from the Owner. If the connectivity between the Backup and Master is somehow interrupted the Backup router automatically starts to act as the Owner. This is a problem because two routers are now acting as "Owners" at the same time, which can lead to unintentional results. To avoid this and to ensure better stability of the link, it is recommended to use two physical links between the Master and Backup core switches.

Recommendation:
Robust communications between the Master and Backup routers is significant to ensure proper detection of the Master's failure. To ensure better stability of the link, it is recommended to use two links between the Master and Backup core switches.

Warning:
The *VRRP* protocol in *IPv4* is already supported. Support for *IPv6* is currently in progress. Until the *VRRP* standard (*VRRPv3*) is completed and accepted by manufacturers we have to use the router advertisement mechanism to solve this problem.

1.4 Backbone layer

The Backbone layer is a common part of the routed network. To distribute routing information a dynamic routing protocol (e.g. *OSPF, RIP, BGP*) is used. The routing protocols support the mechanism of resilience and topology changes by default.

2 Testbed configuration

This section describes configuration on the testbed.

The following image shows the testbed where the resilience functionality was verified.

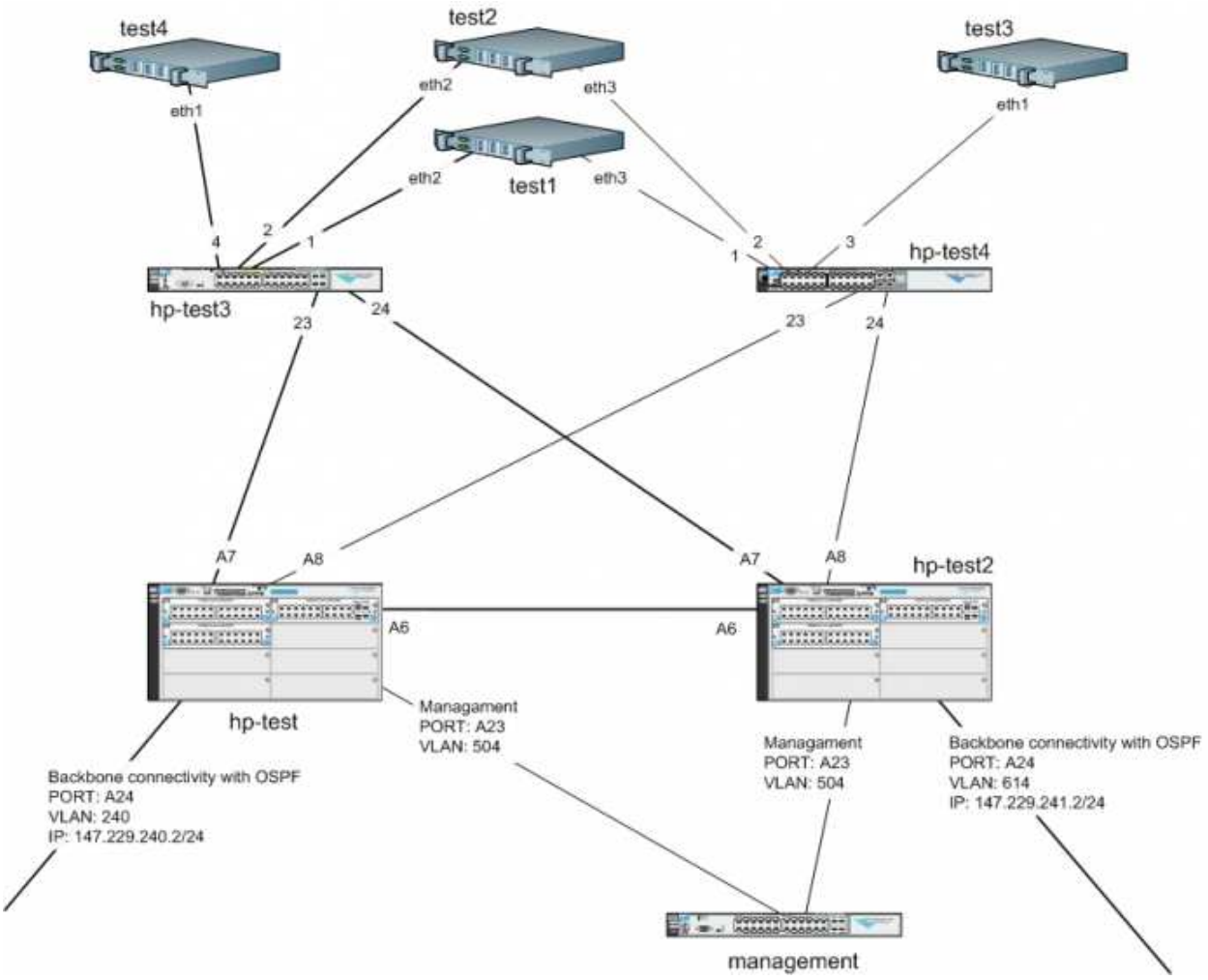

Figure 5: The testbed configuration

2.1 Network device configuration

The actual configuration is shown in the image above. In the testbed infrastructure, HP Procurve devices are used. On the server side, Linux servers are used. The following table shows the types of devices and their function in the testbed configuration.

device	type	description
hp-test	HP 5406	Master router for *VLAN 241*, backup router for *VLAN 242*
hp-test2	HP 5406	Master router for *VLAN 242*, backup router for *VLAN 241*
hp-test3	HP 3500yl	Main distribution switch for servers
hp-test4	HP 2500	Backup switch for servers (only 100Mbps ports)
management	HP 2900	Dedicated switch for management purposes. This switch is dedicated to the distributed management *VLAN* and is part of the independent management network.
test1-4	Linux based servers	Linux based servers which represent the end devices. Test1 and test2 are connected through two interfaces

And the follow VLANs are used in the example

TAG	name	primary router (default GW)	backup router	description
240	ext240	147.229.240.2/24 on hp-test		first external connectivity with OSPF enabled
241	ext241	147.229.241.2/24 on hp-test2		second external connectivity with OSPF enabled
224	int224	147.229.224.1/24 on hp-test	147.229.224.2/24 on hp-test2	a VLAN where end devices are placed
225	int225	147.229.225.1/24 on hp-test2	147.229.225.2/24 on hp-test	a VLAN where end devices are placed
504	mgmt-vlan			VLAN dedicated to management purposes

2.1.1 Core switches configuration

The configuration of hp-test and hp-test2 is very similar. Both of them are configured as a routing switch.

2.1.1.1 *hp-test*

This device works as primary router for *VLAN 224* (147.229.224.0/24) and backup router for *VLAN 225* (147.229.225.0/24). With the following configuration:

- enable basic functionality**,**
- enable routing with *VRRP*, *OSPF* and set area
- allow redistribution to directly connected networks to *OSPF*
- enable spanning tree protocol, enable *RSTP*

```
hp-test(config)# ip routing
hp-test(config)# router ospf
hp-test(ospf)# area 0.0.0.2
hp-test(ospf)# redistribute connected
hp-test(ospf)# exit
hp-test(config)# spanning-tree
hp-test(config)# spanning-tree force-version rstp-operation
hp-test(config)# gvrp
```

External connectivity The network ensures external connectivity. The network is joined to the backbone where *OSPF* routing protocol is run.

```
hp-test(config)# vlan 240
hp-test(vlan-240)# name "ext240"
hp-test(vlan-240)# untagged A24
hp-test(vlan-240)# ip address 147.229.240.2 255.255.255.0
hp-test(vlan-240)# ip ospf 147.229.240.2 area 0.0.0.2
hp-test(vlan-240)# exit
```

Check *OSPF* neighbors

```
hp-test(config)# show ip ospf neighbor

 OSPF Neighbor Information

                                                    Rxmt         Helper
  Router ID      Pri IP Address      NbIfState State    QLen  Events Status
  -------------- --- --------------- --------- -------- ----- ------ ------
  147.229.255.1   1   147.229.240.1   DR        FULL     0     13     None
```

Check routing tables Check if the routing tables contains any networks obtained through *OSPF*.

```
hp-test(config)# show ip route
                                                    IP Route Entries

  Destination       Gateway         VLAN Type      Sub-Type   Metric  Dist.
  ----------------- -------------   ---- --------- ---------- ------  ----
  0.0.0.0/0         147.229.240.1   240  ospf      External2  1        110
  127.0.0.0/8       reject               static               0        0
  127.0.0.1/32      lo0                  connected            1        0
  147.229.0.0/25    147.229.240.1   240  ospf      External2  10       110
  147.229.0.128/25  147.229.240.1   240  ospf      External2  10       110
  147.229.1.128/26  147.229.240.1   240  ospf      External1  1002     110
  147.229.1.192/26  147.229.240.1   240  ospf      External1  1002     110
  147.229.2.0/24    147.229.240.1   240  ospf      External2  10       110
  147.229.3.0/24    147.229.240.1   240  ospf      External2  10       110
  147.229.4.0/23    147.229.240.1   240  ospf      External2  10       110
  ....
```

Internal network/vlan

Setup internal *VLAN* (ID 224) and assign an IP address.

```
hp-test(config)# vlan 224
hp-test(vlan-224)# name int224
hp-test(vlan-224)# ip address 147.229.224.1 255.255.255.0
```

Setup router as VRRP owner for VLAN 224

In this step we'll set the *VRRP* router *ID* and the virtual address.

- The *VRRP* router *ID* identifies the group of routers which maintains the same subnet. It must be set to the same value on all routers in the subnet. In most cases it will be set to “1”.
- The virtual address must be the same as the primary address configured on the *VLAN*.

```
hp-test(vlan-224)# vrrp vrid 1
hp-test(vlan-224-vrid-2)# owner
hp-test(vlan-224-vrid-1)# virtual-ip-address 147.229.224.1 255.255.255.0
hp-test(vlan-224-vrid-1)# enable
hp-test(vlan-224-vrid-1)# exit
hp-test(vlan-224)# exit
```

Setup router as VRRP owner for VLAN 225

The router/switch hp-test will act as the Backup router for *VLAN 225*. In accordance with the recommendation, the second address is used.

```
hp-test(config)# vlan 225
hp-test(vlan-225)# name int225
hp-test(vlan-225)# ip address 147.229.225.2 255.255.255.0
```

The next lines are used to configure the Backup router. The virtual address must be set to the same as the owner address.

```
hp-test(vlan-225)# vrrp vrid 1
hp-test(vlan-225-vrid-1)# backup
hp-test(vlan-225-vrid-1)# virtual-ip-address 147.229.225.1 255.255.255.0
hp-test(vlan-225-vrid-1)# enable
hp-test(vlan-225-vrid-1)# exit
hp-test(vlan-225)# exit
```

2.1.1.2 *hp-test2*

This device has a very similar configuration to hp-test. The switch/router acts as the Master router for VLAN 225 (147.229.225.0/24) and the Backup router for VLAN 224 (147.229.224.0/24). Because most of configuration steps are the same as on hp-test only the truncated configuration is shown below.

```
hp-test2(config)# ip routing
hp-test2(config)# router ospf
hp-test2(ospf)# area 0.0.0.2
hp-test2(ospf)# redistribute connected
hp-test2(ospf)# exit
hp-test2(config)# spanning-tree
hp-test2(config)# spanning-tree force-version rstp-operation
hp-test2(config)# gvrp
```

External VLAN

```
hp-test2(config)# vlan 241
hp-test2(vlan-241)# name "ext241"
hp-test2(vlan-241)# untagged A24
hp-test2(vlan-241)# ip address 147.229.241.2 255.255.255.0
hp-test2(vlan-241)# ip ospf 147.229.241.2 area 0.0.0.2
hp-test2(vlan-241)# exit
hp-test2(config)# show ip ospf neighbor
hp-test2(config)# show ip route
hp-test2(config)# vlan 224
hp-test2(vlan-224)# name int224
hp-test2(vlan-224)# ip address 147.229.224.2 255.255.255.0
hp-test2(vlan-224)# vrrp vrid 1
```

Set VRRP backup for VLAN 224

```
hp-test2(vlan-224-vrid-2)# backup
hp-test2(vlan-224-vrid-1)# virtual-ip-address 147.229.224.1 255.255.255.0
hp-test2(vlan-224-vrid-1)# enable
hp-test2(vlan-224-vrid-1)# exit
hp-test2(vlan-224)# exit
hp-test2(config)# static-vlan 225
hp-test2(config)# vlan 225
hp-test2(vlan-225)# name int225
hp-test2(vlan-225)# ip address 147.229.225.1 255.255.255.0
```

Set VRRP backup for VLAN 225

```
hp-test2(vlan-225)# vrrp vrid 1
hp-test2(vlan-225-vrid-2)# owner
hp-test2(vlan-225-vrid-2)# virtual-ip-address 147.229.225.1 255.255.255.0
hp-test2(vlan-225-vrid-2)# enable
hp-test2(vlan-225-vrid-2)# exit
hp-test2(vlan-225)# exit
```

```
hp-test2(config)# show VRRP

 VRRP Global Statistics Information
  VRRP Enabled           : Yes
  Protocol Version       : 2
  Invalid VRID Pkts Rx   : 0
  Checksum Error Pkts Rx : 0
  Bad Version Pkts Rx    : 0
```

```
VRRP Virtual Router Statistics Information

  Vlan ID                  : 224
  Virtual Router ID        : 1
  State                    : Backup
  Up Time                  : 18 hours
  Virtual MAC Address      : 00005e-000101
  Master's IP Address      : 147.229.224.1
  Associated IP Addr Count : 1          Near Failovers            : 0
  Advertise Pkts Rx        : 269        Become Master             : 1
  Zero Priority Rx         : 0          Zero Priority Tx          : 0
  Bad Length Pkts          : 0          Bad Type Pkts             : 0
  Mismatched Interval Pkts : 0          Mismatched Addr List Pkts : 269
  Mismatched IP TTL Pkts   : 0          Mismatched Auth Type Pkts : 0


 VRRP Virtual Router Statistics Information

  Vlan ID                  : 225
  Virtual Router ID        : 2
  State                    : Master
  Up Time                  : 9 mins
  Virtual MAC Address      : 00005e-000102
  Master's IP Address      : 147.229.225.1
  Associated IP Addr Count : 1          Near Failovers            : 0
  Advertise Pkts Rx        : 0          Become Master             : 1
  Zero Priority Rx         : 0          Zero Priority Tx          : 0
  Bad Length Pkts          : 0          Bad Type Pkts             : 0
  Mismatched Interval Pkts : 0          Mismatched Addr List Pkts : 0
  Mismatched IP TTL Pkts   : 0          Mismatched Auth Type Pkts : 0
```

2.1.2 Distribution switches configuration

2.1.2.1 *hp-test3, hp-test4*

Enable rapid spanning tree, GVRP

```
hp-test3(config)# spanning-tree
hp-test3(config)# spanning-tree force-version rstp-operation
hp-test3(config)# gvrp
```

At this time, if you have the configuration correct you should check if the VLANs were obtained from the core switch. Here is an example output.

```
hp-test3(eth-24)# show vlans

 Status and Counters - VLAN Information

  Maximum VLANs to support : 256
  Primary VLAN : DEFAULT_VLAN
  Management VLAN :

  VLAN ID Name                | Status     Voice Jumbo
  ------- ------------------- + ---------- ----- -----
  1       DEFAULT_VLAN        | Port-based No    No
  224     GVRP_224            | Dynamic
  225     GVRP_225            | Dynamic
  504     mgmt-vlan           | Port-based No    No
```

Now we can assign ports to particular *VLANs* and connect end devices:

```
hp-test3(config)# static-vlan 224
hp-test3(config)# vlan 224
hp-test3(vlan-224)# tagged 1,3,4
hp-test3(vlan-224)# untagged 2
hp-test3(vlan-224)# exit

hp-test3(config)# static-vlan 225
hp-test3(config)# vlan 225
hp-test3(vlan-225)# tagged 1,3,4
hp-test3(vlan-225)# exit
```

2.2 Host devices configuration

pc-name	vlan 224	vlan 225	ip address	switch port
test1	tagged	tagged	147.229.224.101	1
test2	untagged		147.229.224.102	2
test3	tagged	tagged	147.229.224.103	3
test4	tagged	tagged	147.229.224.104	4

For connecting important devices, it is important to connect these devices to the network through two independent links. You need a server equipped with two network Ethernet cards and special driver support. On Linux based systems the new virtual interface is created and real interfaces are connected to this interface.

- http://linux-ip.net/html/ether-bonding.html
- http://www.linuxhorizon.ro/bonding.html
- http://www.linuxquestions.org/questions/linux-general-1/network-bonding-questions-525194/
- http://www.devco.net/archives/2004/11/26/linux_ethernet_bonding.php

If you want to set up server bonding on a Linux system you have to take the follow steps:

Add the following line to */etc/modprobe.conf*

```
# vim /etc/modprobe.conf
alias bond0 bonding
options bond0 miimon=100 mode=1 primary=eth2
```

Edit /etc/sysconfig/network-scripts/ifcfg-eth2

```
# vim /etc/sysconfig/network-scripts/ifcfg-eth2
DEVICE=eth2
ONBOOT=yes
MASTER=bond0
SLAVE=yes
```

Edit /etc/sysconfig/network-scripts/ifcfg-eth3

```
# vim /etc/sysconfig/network-scripts/ifcfg-eth3
DEVICE=eth3
ONBOOT=yes
MASTER=bond0
SLAVE=yes
```

Set up the IP address for ***bond0*** interface by editing

/etc/sysconfig/network-scripts/ifcfg-bond0

```
# vim /network-scripts/ifcfg-bond0
DEVICE=bond0
ONBOOT=yes
BOOTPROTO=static
IPADDR=147.229.224.101
NETMASK=255.255.255.0
GATEWAY=147.229.224.1
```

After rebooting, ifconfig should output something like this:

```
 [root@test2 ~]# ifconfig
bond0     Link encap:Ethernet  HWaddr 00:30:48:5F:3A:A6
          inet addr:147.229.224.101  Bcast:147.229.224.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MASTER MULTICAST  MTU:1500  Metric:1
          RX packets:42 errors:0 dropped:0 overruns:0 frame:0
          TX packets:11 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:4252 (4.1 KiB)  TX bytes:704 (704.0 b)

eth2      Link encap:Ethernet  HWaddr 00:30:48:5F:3A:A6
          UP BROADCAST RUNNING SLAVE MULTICAST  MTU:1500  Metric:1
          RX packets:39 errors:0 dropped:0 overruns:0 frame:0
          TX packets:11 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:100
          RX bytes:4060 (3.9 KiB)  TX bytes:704 (704.0 b)
          Base address:0x3000 Memory:d8200000-d8220000

eth3      Link encap:Ethernet  HWaddr 00:30:48:5F:3A:A6
          UP BROADCAST RUNNING SLAVE MULTICAST  MTU:1500  Metric:1
          RX packets:3 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:192 (192.0 b)  TX bytes:0 (0.0 b)
          Base address:0x4000 Memory:d8300000-d8320000

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
          RX packets:3 errors:0 dropped:0 overruns:0 frame:0
          TX packets:3 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:267 (267.0 b)  TX bytes:267 (267.0 b)
```

Now you can test your connectivity by using ‘ping’ to a host outside of the testbed

```
[root@test2 ~]# ping 147.229.3.10
PING 147.229.3.10 (147.229.3.10) 56(84) bytes of data.
64 bytes from 147.229.3.10: icmp_seq=1 ttl=62 time=0.258 ms
64 bytes from 147.229.3.10: icmp_seq=2 ttl=62 time=0.095 ms

--- 147.229.3.10 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 999ms
rtt min/avg/max/mdev = 0.095/0.176/0.258/0.082 ms
```

Now you can display the bonding status

```
[root@test2 ~]# cat /proc/net/bonding/bond0
Ethernet Channel Bonding Driver: v3.0.3 (March 23, 2006)

Bonding Mode: fault-tolerance (active-backup)
Primary Slave: None
Currently Active Slave: eth3
MII Status: up
MII Polling Interval (ms): 100
Up Delay (ms): 0
Down Delay (ms): 0

Slave Interface: eth2
MII Status: up
Link Failure Count: 0
Permanent HW addr: 00:30:48:5f:3a:a6

Slave Interface: eth3
MII Status: up
Link Failure Count: 0
Permanent HW addr: 00:30:48:5f:3a:a7
```

Of course, if you need to, you can change the active interface

```
[root@test2 ~]# ifenslave -c bond0 eth2
[root@test2 ~]# cat /proc/net/bonding/bond0
Ethernet Channel Bonding Driver: v3.0.3 (March 23, 2006)

Bonding Mode: fault-tolerance (active-backup)
Primary Slave: None
Currently Active Slave: eth2
MII Status: up
MII Polling Interval (ms): 100
Up Delay (ms): 0
Down Delay (ms): 0

Slave Interface: eth2
MII Status: up
Link Failure Count: 0
Permanent HW addr: 00:30:48:5f:3a:a6

Slave Interface: eth3
MII Status: up
Link Failure Count: 0
Permanent HW addr: 00:30:48:5f:3a:a7
```

2.3 Testing

It is important to have a way to test the configuration. The simplest tool is the ping command. This is a short script which sends test packets across the network and reports the time interval between the time sent and the time the packet is received back and computes the length of the interval.

The script transmits *ICMP* requests to the destination address with a very short interval (0.01 seconds). The computation of the downtime is based on the ICMP responses.

Both tests start with the infrastructure where all protocols are converged. The active paths and master routers are shown in the following image.

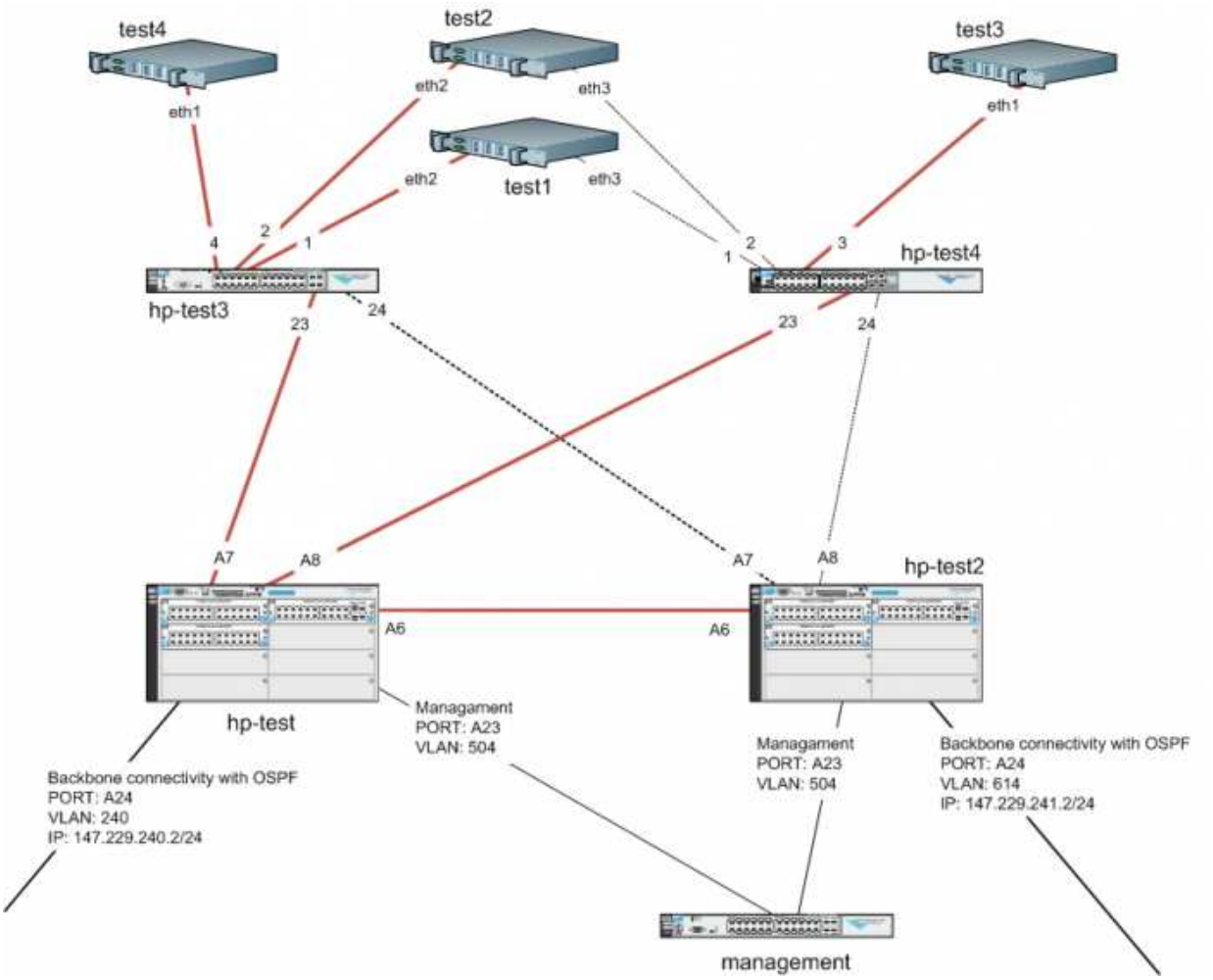

Figure 6: Active paths when the protocols are converged

2.3.1 Access switch failure test

The first test simulates a situation when the one of the distribution switches fails. We did it by removing the power cord from the switch. The following figure shows the situation after shutdown.

The test was performed on the test3 server. During the failure the follow changes were expected:

- The change of the active interface - the active interface should be changed to the Backup one.
- *STP* root changed.
- Unblock the lines previously blocked by *STP.*

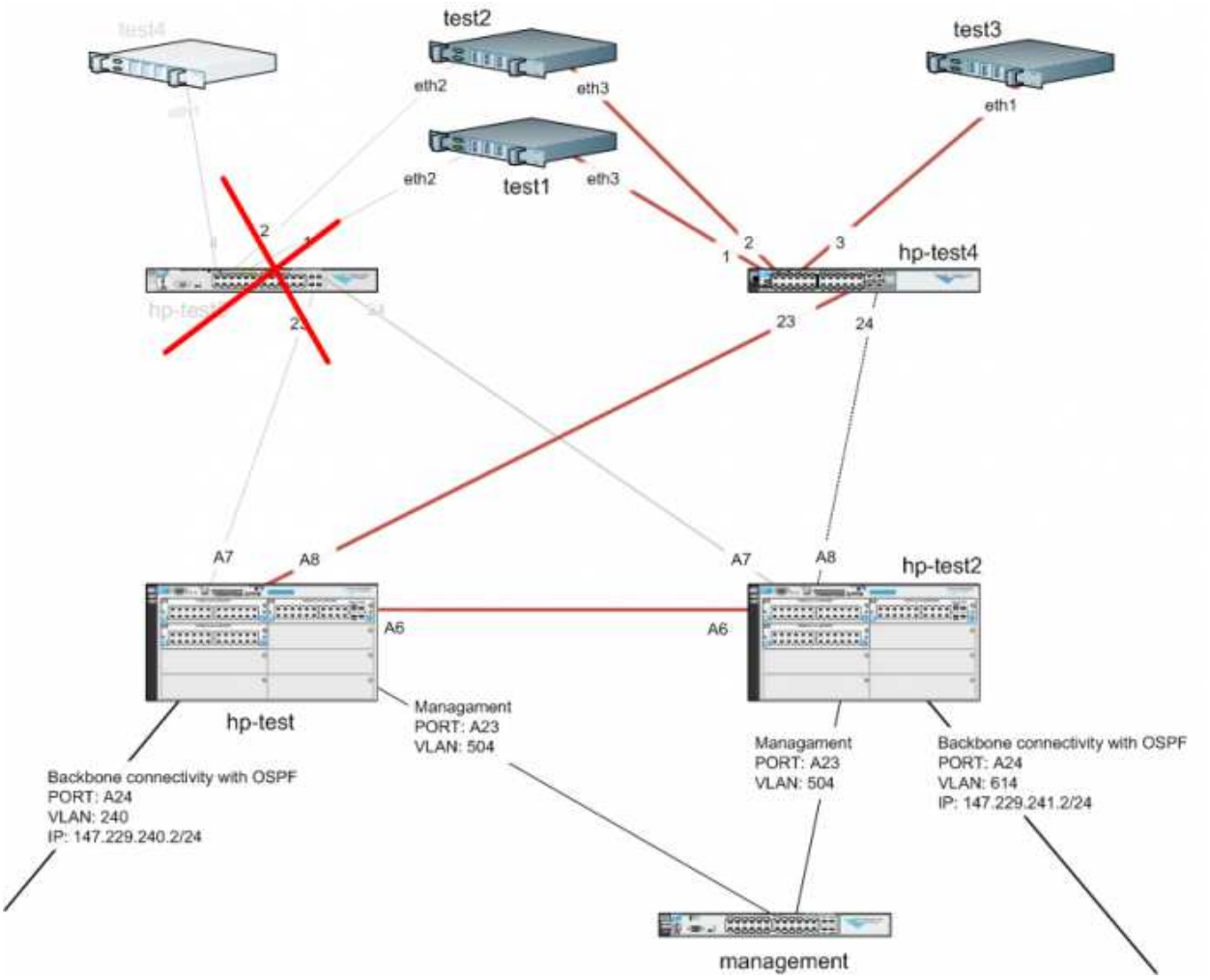

Figure 7: The testbed behaviour when the access switch fails

```
[root@test2 ~]# ./avatest.pl ferret.cis.vutbr.cz
Sending ICMP to ferret.cis.vutbr.cz with interval 0.010000 s
PING ferret.cis.vutbr.cz (147.229.252.11) 1400(1428) bytes of data.
time since start: 18.444675 s, lost packets: 296, down window size: 2.958039 s
time since start: 185.226505 s, lost packets: 20, down window size: 0.198872 s
time since start: 186.156383 s, lost packets: 4, down window size: 0.039849 s
time since start: 206.229157 s, lost packets: -44, down window size: 0.006245 s
time since start: 206.231781 s, lost packets: 41, down window size: 0.000249 s
time since start: 206.232032 s, lost packets: -40, down window size: 0.000251 s
time since start: 206.240775 s, lost packets: 33, down window size: 0.004998 s
```

Conclusion:
The results show that the changeover to the Backup interface and *STP* convergence is very fast. This test shows that only 4 packets were lost. Some packets were delivered in a different order to which they were sent. Most applications should work with that delay without any significant issues.

2.3.2 Core switch failure test

The second test shows the situation when the one of the core switches is powered off. The test was done by removing the power cord from the hp-test switch which had been set up as the *VRRP* Owner/Master.

In this test we expect to see the following changes:

- Change the *STP* root.
- Unblock the lines blocked by *STP*.
- Move Master router to the Backup one.

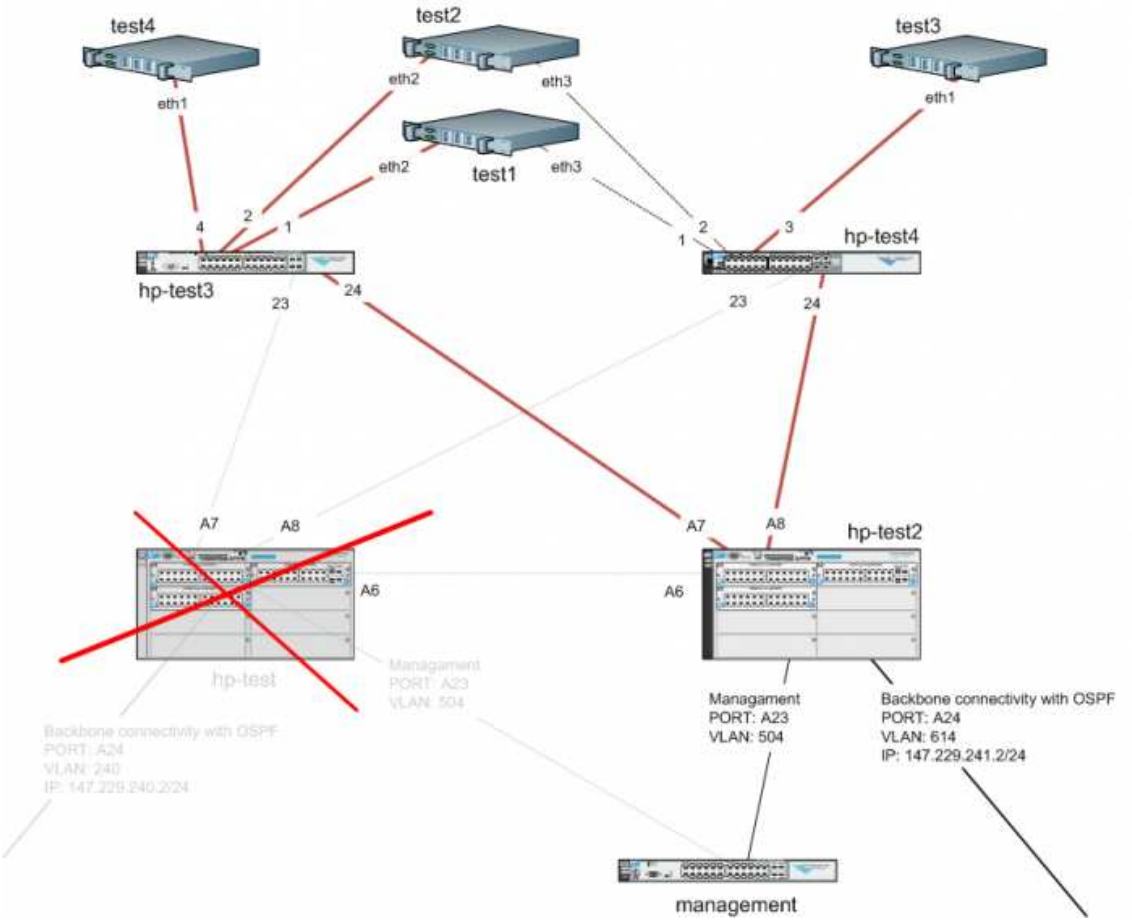

Figure 8: The testbed behaviour when the core switch fails

```
[root@test2 ~]# ./avatest.pl ferret.cis.vutbr.cz
Sending ICMP to ferret.cis.vutbr.cz with interval 0.010000 s
PING ferret.cis.vutbr.cz (147.229.252.11) 1400(1428) bytes of data.
time since start: 18.444675 s, lost packets: 296, down window size: 2.958039 s
time since start: 185.226505 s, lost packets: 20, down window size: 0.198872 s
time since start: 186.156383 s, lost packets: 4, down window size: 0.039849 s
time since start: 206.229157 s, lost packets: -44, down window size: 0.006245 s
time since start: 206.231781 s, lost packets: 41, down window size: 0.000249 s
time since start: 206.232032 s, lost packets: -40, down window size: 0.000251 s
time since start: 206.240775 s, lost packets: 33, down window size: 0.004998 s
```

3 Conclusion:

In the results of the ping test above we can see several holes, the significant one is visible on the first line; the window size is virtually 3 seconds. This is exactly the time which is needed by the backup *VRRP* router to assume the gateway address. This time can be determined with the *VRRP* delay option. The problem is that that option cannot be less than 1 second and *RFC 3768* specifies that the backup router has to wait for three consecutive *VRRP* delay times without receiving an advertisement packet before it assumes the Owner role. Another potential critical situation is when the Master router reverts back to being the the Owner. In the test outputs above, we can see that several packets have been lost and several packets have been delivered in a different order to which they were sent.

4 Figure list