Installation and Operations Guide

Release Dell R720xd

June 10, 2013

Google Global Cache
ggc@google.com

Contents

CHAPTER 1

Installation and Commissioning Process Overview

The following diagram shows the steps to deploy and commission a Google Global Cache node. This document will provide additional detail required to complete each step.

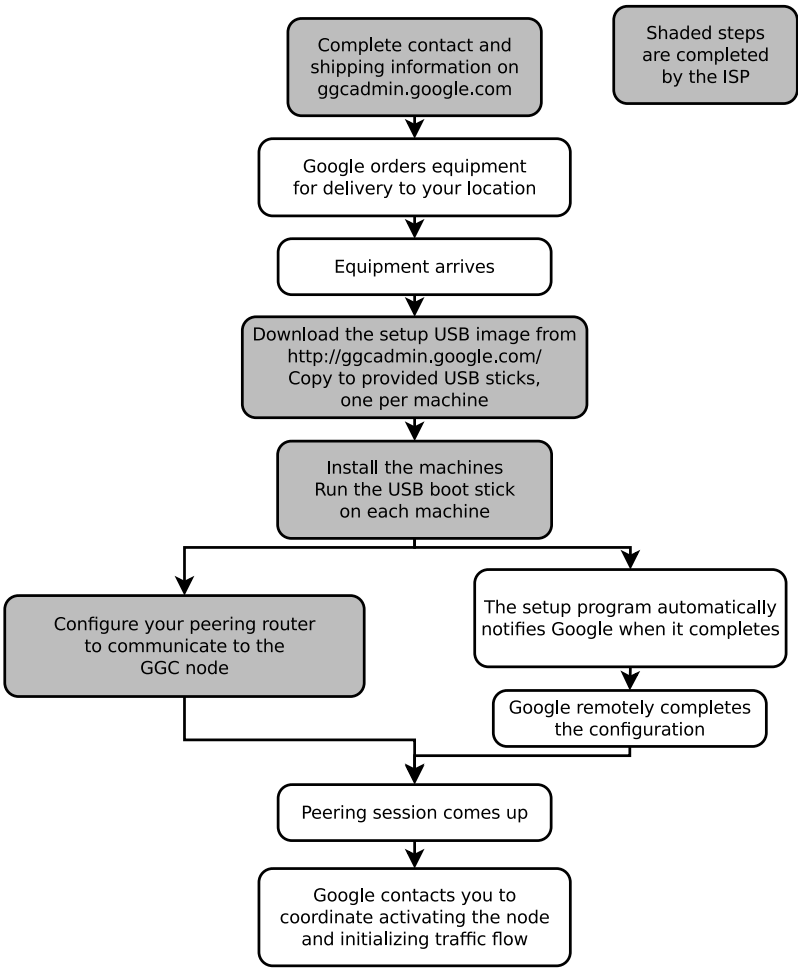

Figure 1.1: Process Overview.

CHAPTER 2

Hardware Installation

2.1 You will need

- Help to lift servers into position
- Rack mount installation kit and instructions (included in server box)
- Philips (crosshead) screwdriver
- Copper Ethernet cables, 4 per server
- Dual AC power feeds (see table below)
- Dual locally required power cords (Google provides C13/C14 power cords only)

Table 2.1: Power Requirements

# Servers	Rack Space	Nominal Power	Peak Power	Amps @ 100 VAC	Amps @ 220 VAC
3	6RU	900W	1200W	9A	5A
4	8RU	1200W	1600W	12A	6.5A
6	12RU	1800W	2400W	18A	10A
8	16RU	2400W	3200W	24A	13A

2.2 Procedure

1. Install the servers in the rack according to the included instructions.
2. Connect the network switch to the ports labeled Gb1, Gb2, Gb3 and Gb4 (see figure *Cabling scheme*).
3. Connect power, *but do not turn the system on yet*.

Note: Both power supplies must be connected. It is strongly recommended that you connect each power supply to an independent power feed (i.e., A and B power). However, both can be connected to the same circuit if a second circuit is not available. This will at least protect from failure of a single power supply.

2.3 Disk layout

In case a disk is showing errors, the GGC operations will contact you and ask you to re-seat or replace a disk. A disk slot number will be provided.

The layout below can help you locating the correct disk.

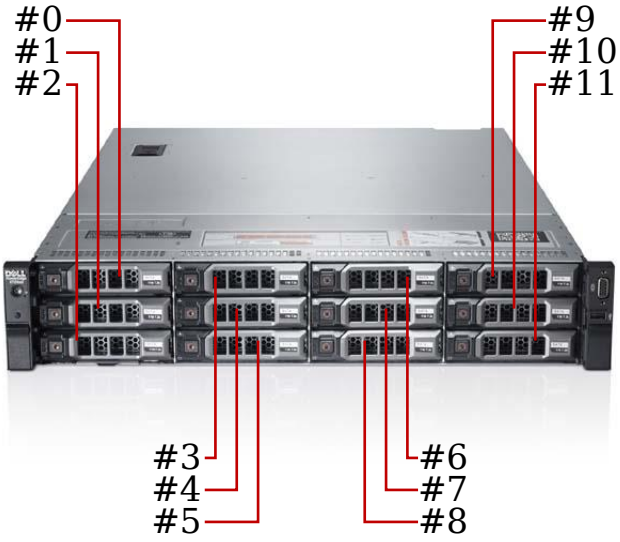

Figure 2.1: Dell PowerEdge R720xd disk layout

2.4 More information

- Dell PowerEdge R720 and R720xd Technical Guide
- Dell PowerEdge R720 and R720xd Owner's Manual

Switch Configuration

3.1 You will need

- Administrative access to the Ethernet switch connected to the GGC servers
- Switch port numbers connected to the GGC servers

3.2 Procedure

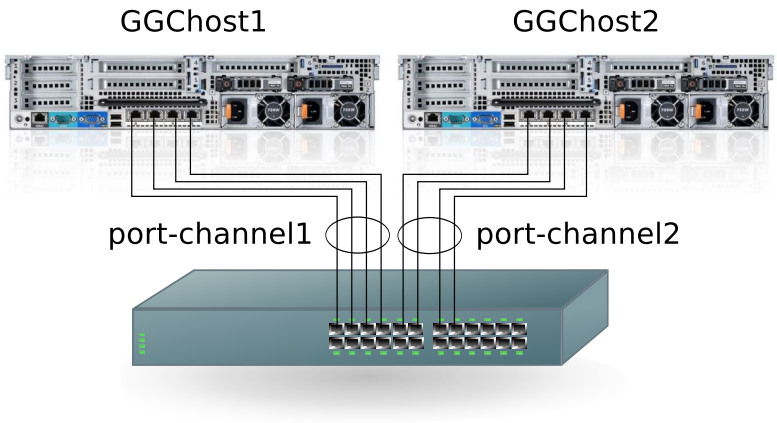

Figure 3.1: Cabling scheme

Please refer to your switch's documentation for the specific commands to configure the Ethernet ports facing the GGC machines as follows:

- 1000Mbps full duplex
- Set to auto-negotiate
- Link Aggregation Control Protocol (LACP) enabled, with:
  - Passive mode
  - Load balanced on source/destination layer 3 information
  - Switch/Layer 2 mode
  - Standalone mode (aggregated link should remain up, even if a physical port is down)

- All machines in the GGC node must be in a single, dedicated layer 2 broadcast domain

See the illustration *Cabling scheme*.

Note: When LACP is enabled, at least one single interface (Gb1) on each server must be connected to the switch. The Gb2, Gb3 and Gb4 interfaces can be connected at a later stage, they will automatically aggregate with the Gb1 interface.

3.3 Switch Configuration Examples

The following example is for illustration purposes only. Your configuration may vary. Please contact your switch vendor for detailed configuration support for your specific equipment.

3.3.1 Cisco Switch Configuration Fragment

```
!
interface GigabitEthernet1/1
  description GGChost1-Gb1
  switchport mode access
  channel-protocol lacp
  channel-group 1 mode passive
!
interface GigabitEthernet1/2
  description GGChost1-Gb2
  switchport mode access
  channel-protocol lacp
  channel-group 1 mode passive
!
interface GigabitEthernet1/3
  description GGChost1-Gb3
  switchport mode access
  channel-protocol lacp
  channel-group 1 mode passive
!
interface GigabitEthernet1/4
  description GGChost1-Gb4
  switchport mode access
  channel-protocol lacp
  channel-group 1 mode passive
!
interface Port-channel1
  description GGChost1
  switchport
  switchport mode access
  no port-channel standalone-disable
!
interface GigabitEthernet1/5
  description GGChost2-Gb1
  switchport mode access
  channel-protocol lacp
  channel-group 2 mode passive
!
interface GigabitEthernet1/6
  description GGChost2-Gb2
  switchport mode access
```

```
  channel-protocol lacp
  channel-group 2 mode passive
!
interface GigabitEthernet1/7
  description GGChost2-Gb3
  switchport mode access
  channel-protocol lacp
  channel-group 2 mode passive
!
interface GigabitEthernet1/8
  description GGChost2-Gb4
  switchport mode access
  channel-protocol lacp
  channel-group 2 mode passive
!
interface Port-channel2
  description GGChost2
  switchport
  switchport mode access
  no port-channel standalone-disable
end
```

3.3.2 Juniper Switch Configuration Fragment

```
set interfaces ge-0/0/1 description GGChost1-Gb1
set interfaces ge-0/0/1 ether-options 802.3ad lacp force-up
set interfaces ge-0/0/1 ether-options 802.3ad ae0
set interfaces ge-0/0/2 description GGChost1-Gb2
set interfaces ge-0/0/2 ether-options 802.3ad lacp force-up
set interfaces ge-0/0/2 ether-options 802.3ad ae0
set interfaces ge-0/0/3 description GGChost1-Gb3
set interfaces ge-0/0/3 ether-options 802.3ad lacp force-up
set interfaces ge-0/0/3 ether-options 802.3ad ae0
set interfaces ge-0/0/4 description GGChost1-Gb4
set interfaces ge-0/0/4 ether-options 802.3ad lacp force-up
set interfaces ge-0/0/4 ether-options 802.3ad ae0
set interfaces ge-0/0/5 description GGChost2-Gb1
set interfaces ge-0/0/5 ether-options 802.3ad lacp force-up
set interfaces ge-0/0/5 ether-options 802.3ad ae1
set interfaces ge-0/0/6 description GGChost2-Gb2
set interfaces ge-0/0/6 ether-options 802.3ad lacp force-up
set interfaces ge-0/0/6 ether-options 802.3ad ae1
set interfaces ge-0/0/7 description GGChost2-Gb3
set interfaces ge-0/0/7 ether-options 802.3ad lacp force-up
set interfaces ge-0/0/7 ether-options 802.3ad ae1
set interfaces ge-0/0/8 description GGChost2-Gb4
set interfaces ge-0/0/8 ether-options 802.3ad lacp force-up
set interfaces ge-0/0/8 ether-options 802.3ad ae1
set interfaces ae0 description GGChost1
set interfaces ae0 aggregated-ether-options lacp passive
set interfaces ae0 unit 0 family ethernet-switching port-mode access
set interfaces ae0 unit 0 family ethernet-switching vlan members vlan10
set interfaces ae1 description GGChost2
set interfaces ae1 aggregated-ether-options lacp passive
set interfaces ae1 unit 1 family ethernet-switching port-mode access
set interfaces ae1 unit 1 family ethernet-switching vlan members vlan10
```

IP Addressing

The GGC node requires a dedicated layer 3 subnet. Each server has a management IP address (IPv4) statically assigned to the ethernet interface, or bonded set of interfaces in the case of LACP. Additionally, each server has a number of virtual IP addresses or VIPs (IPv4 and IPv6, if enabled). User traffic is served from the VIPs. In the event of a server failure, other servers in the node pick up the failed server's VIPs.

Note: Regardless of the IPv6 configuration, the GGC node requires IPv4 addresses.

4.1 IPv4

4.1.1 Addressing scheme

The GGC node requires a dedicated /26 IPv4 subnet (netmask 255.255.255.192).

The general guidelines for the assignment of IPv4 addresses with the GGC subnet:

1. Assign the first usable address in the subnet to the subnet gateway
2. If required, use the next two addresses for HSRP or GLBP
3. Assign the 4th IP address in the subnet to the Gb1 interface of the first server
4. The server addresses must be contiguous, do not assign or reserve addresses for the Gb2, Gb3 and Gb4 interfaces as these will be aggregated with the Gb1 interfaces via LACP
5. The 12th address in the subnet is reserved for the first VIP
6. The 16th address is used both as a machine address and as the source for the BGP peering session with your network

Address Number	Use	8 Server Node Example
0	Subnet address	10.10.10.64/26
1	Gateway set on cache servers	10.10.10.65
2,3	HSRP/GLBP gateways (optional)	Unused
4	First GGC server (*)	10.10.10.68
5	Second GGC server (*)	10.10.10.69
continue for each server	Last GGC server (*)	10.10.10.75
12 to (last address - 1)	Virtual IPs (**)	10.10.10.76 - 10.10.10.126
16	This virtual IP is also used for BGP peering (**)	10.10.10.80
Last address in subnet	Broadcast	10.10.10.127

(*) configured manually during setup

(**) configured remotely by Google

4.1.2 Server Naming / Reverse DNS

Please configure reverse DNS entries for all servers' IP addresses (both real and virtual addresses) to *cache.google.com*.

The following example is for illustration purposes only (`bind` configuration):

```
$TTL 1D
$ORIGIN 10.10.10.in-addr.arpa.
@                   IN SOA ...

$GENERATE 65-126 $ IN PTR cache.google.com.
```

4.2 IPv6

4.2.1 Addressing scheme

The GGC node requires a dedicated /64 IPv6 subnet.

The general guidelines for the assignment of IPv6 addresses with the GGC subnet:

1. The 12th address in the subnet is reserved for the first VIP
2. The 16th address is used as the source for the BGP peering session with your network

---

Note: The node relies on IPv6 Router Advertisements (RA) for the configuration of the IPv6 default gateway

---

Address Number	Use	8 Server Node Example (*)
0	Subnet address	fec0:1234::/64
1-11	Unused	Unused
12 to last address (*)	Virtual IPs (**)	fec0:1234::c - fec0:1234::3b
16	This virtual IP is also used for BGP peering (**)	fec0:1234::10

(*) this particular example uses 6 VIPs per host

(**) configured remotely by Google

4.2.2 IPv6 Enablement

Adding IPv6 support to a GGC node is easy for both new or existing nodes.

For a *new node*, IPv6 can be enabled prior to installation by specifying an **IPv6 subnet** and **IPv6 Router for BGP Sessions** when you supply the other technical information required for node activation in the GGCAdmin Portal.

For an *existing node* that is already serving IPv4 traffic, IPv6 can be enabled through the GGCAdmin Portal URL: http://ggcadmin.google.com/v6 and selecting the node you wish to enable IPv6 on. A screen similar to *IPv6 enablement* appears.

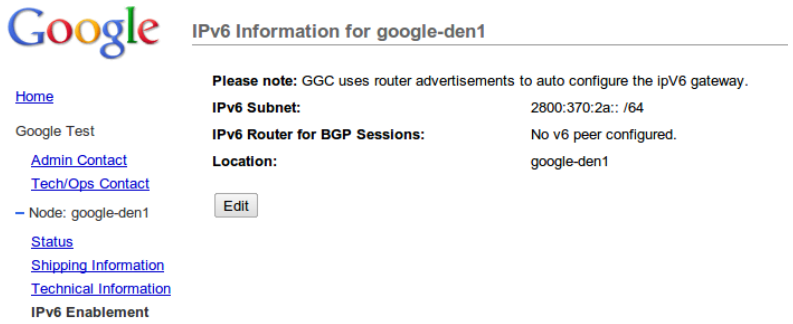

Figure 4.1: IPv6 enablement

Click the 'Edit' button to enter or change the IPv6 subnet of the node and the IPv6 address of the BGP peer.

Note:

- the IPv6 subnet should be entered in CIDR notation, including the '/64'
- use of IPv6 link-local addresses for the BGPv6 peer is not allowed

---

4.3 Proxies and Filters

No transparent proxies or filters may be placed in the path of communications between the GGC Node and Google's back-end servers.

CHAPTER 5

Software Installation

This section describes the steps to install the initial setup software on the machine. After completing this step the installer will automatically signal to Google to remotely begin the next step in the process.

These are the high level installation steps:

1. Download the USB image
2. Create the USB boot sticks (or drives)
3. Boot each machine from a USB stick
4. Enter the network configuration and wait for the installer to complete
5. Reboot the server

5.1 You will need

- A monitor and keyboard
- IP information provided to Google in the GGCAdmin portal
- Labels to mark the IP address on each server
- A USB stick with a capacity of at least 256 MB. One USB stick per server is provided.
- Access to a computer with appropriate permissions to
  - download and save the install image file
  - download, save and run the tools required to create a bootable USB stick

Attention: Each server should have its own USB boot stick, as server specific configurations are stored on the stick.

5.2 Preparing the USB stick (drive)

Download the install image ggc-setup2_0.img. You can find the link in the GGCAdmin Portal (http://ggcadmin.google.com/), by clicking on the link 'Setup Image (R720)' in the 'Downloads' section at the bottom of the page.

Once the image is downloaded, the USB boot sticks need to be created. This should be repeated for each USB stick for each server.

Warning: All data on the USB stick will be erased when you load the GGC image.

5.2.1 Create the USB boot stick on Microsoft Windows

1. Download 'Image Writer for Windows' (`win32diskimager-binary.zip`) from http://sourceforge.net/projects/win32diskimager/ and extract the archive. The executable is '**Win32DiskImager.exe**'.
2. Insert the USB stick. Wait till all appropriate drivers are installed, if necessary.
3. Start '**Win32DiskImager.exe**'.

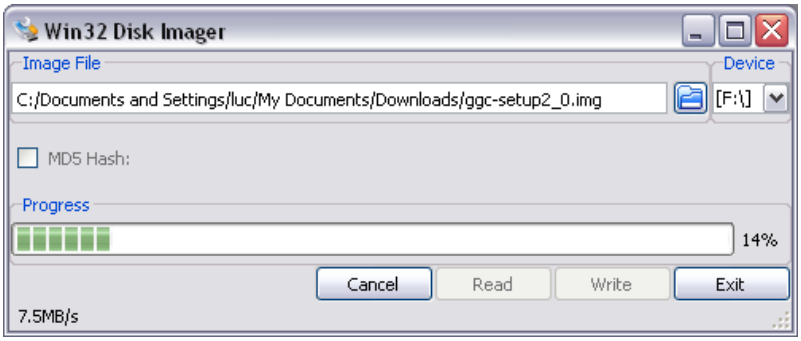

Figure 5.1: Image Writer for Windows

4. The 'Device' combo box automatically selects the USB drive letter. Verify if this is the correct device.
5. Use the button in the 'Image File' group box to select the downloaded setup image file (`ggc-setup2_0.img`)
6. Press the 'Write' button to write the image to the USB stick and confirm the operation.
7. After a few seconds, a message box should appear, stating that the write operation was successful. The USB stick can be removed.

5.2.2 Create the USB boot stick on Mac

1. Open a terminal.
2. Insert the USB stick, a new device (e.g., `/dev/disk2s1`) will appear.
3. Check if a partition on the device is mounted:

```
$ df -h
Filesystem      Size   Used  Avail Capacity  Mounted on
/dev/disk1     112Gi   33Gi   78Gi    30%    /
devfs          203Ki  203Ki    0Bi   100%    /dev
map auto.auto    0Bi    0Bi    0Bi   100%    /auto
map auto.home    0Bi    0Bi    0Bi   100%    /home
map -hosts       0Bi    0Bi    0Bi   100%    /net
/dev/disk2s1   7.5Gi  1.5Gi  6.0Gi   21%    /Volumes/Cruzer
```

4. You will need to unmount the usb stick as follows:

```
$ diskutil umount /Volumes/Cruzer
Volume Cruzer on disk2s1 unmounted
```

5. Verify it is gone:

```
$ df -h
Filesystem      Size   Used  Avail Capacity  Mounted on
/dev/disk1     112Gi   33Gi   78Gi    30%    /
devfs          201Ki  201Ki    0Bi   100%    /dev
map auto.auto    0Bi    0Bi    0Bi   100%    /auto
map auto.home    0Bi    0Bi    0Bi   100%    /home
map -hosts       0Bi    0Bi    0Bi   100%    /net
```

6. Provided the USB stick is device `/dev/disk2`, the following command creates the bootable USB stick:

   ```
   $ dd if=/path/to/ggc-setup2_0.img of=/dev/disk2
   ```

   After this command has completed, the USB stick can be removed.

5.2.3 Create the USB boot stick on Linux

1. Open a terminal.
2. Insert the USB stick, a new device will appear (e.g., `/dev/sdb`). The device name can be checked using '**dmesg**':

   ```
   $ dmesg
   usb 1-4: new high speed USB device using ehci_hcd and address 5
   scsi7 : usb-storage 1-4:1.0
   scsi 7:0:0:0: Direct-Access     Kingston DataTraveler G3  1.00 PQ: 0 ANSI: 0 CCS
   sd 7:0:0:0: Attached scsi generic sg2 type 0
   sd 7:0:0:0: [sdb] 7567964 512-byte logical blocks: (3.87 GB/3.60 GiB)
   sd 7:0:0:0: [sdb] Write Protect is off
   sd 7:0:0:0: [sdb] Mode Sense: 0b 00 00 08
   sd 7:0:0:0: [sdb] Assuming drive cache: write through
   sd 7:0:0:0: [sdb] Assuming drive cache: write through
   sdb: sdb1
   sd 7:0:0:0: [sdb] Assuming drive cache: write through
   sd 7:0:0:0: [sdb] Attached SCSI removable disk
   ```

   In this particular example, the device is `/dev/sdb`.
3. Make sure no partition on this device is mounted. The command '**mount | grep /dev/sdb**' should not return any output. If it does, unmount the partition(s). Here is an example:

   ```
   $ mount | grep /dev/sdb
   /dev/sdb1 on /media/DEBIAN_LIVE type vfat
   ```

   In this particular instance, the partition `/dev/sdb1` is mounted. To unmount the partition:

   ```
   $ sudo umount /dev/sdb1
   ```

   Now, the command:

   ```
   $ mount | grep /dev/sdb
   ```

   returns no output, meaning no partition on this device is mounted.
4. Provided the USB stick is device `/dev/sdb`, the following command creates the bootable USB stick:

   ```
   $ sudo dd if=/path/to/ggc-setup2_0.img of=/dev/sdb
   ```

   After this command has completed, the USB stick can be removed.

5.3 GGC Software Installation

Warning: the GGC software installation will potentially destroy data on the server's disks

The installation needs to be done on every server.

1. Connect the monitor and keyboard on the server to be installed
2. Insert the setup USB boot stick in any USB port (front or back). Please be aware that the USB stick should be left in the machine after installation. Using the front USB port may prevent you from easily connecting a keyboard in the future, or from closing the rack doors.
3. Start the server from the power switch.
4. After a couple of minutes, the server boots from the USB stick. A screen similar to the *Installer Start Screen* appears.

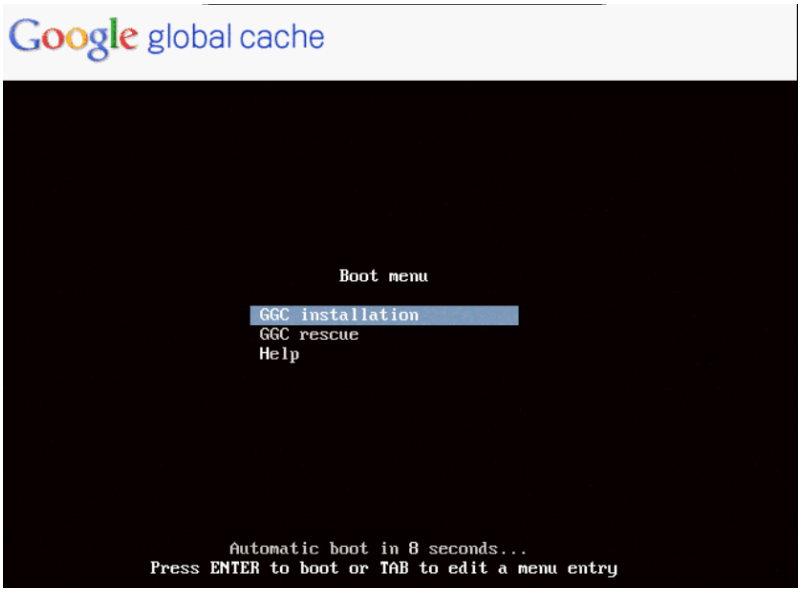

Figure 5.2: Installer Start Screen

5. Press `enter` or wait for 10 seconds for the 'Boot Menu' to disappear. The system boots up and starts the installation program.

   The installer will examine the hardware configuration. Depending on the current configuration, modifications to the BIOS and/or RAID controller configuration may be applied. This potentially requires a reboot of the system. The installation program will restart automatically.

6. Enter the IPv4 configuration for this particular server in the *IPv4 Information* screen. The configuration should match the IP information provided to Google in the GGCAdmin portal:
   - Enable LACP [Y]: unless you plan not to use LACP, enter 'N'. otherwise just press `enter`. Please note that you can enable LACP even when there are not yet multiple network cables connected.
   - Enter the IPv4 address
   - Enter the netmask. Only '255.255.255.0' (/24), '255.255.255.128' (/25), '255.255.255.192' (/26) and '255.255.255.224' (/27) are allowed.
   - Enter the IPv4 address of the default gateway. If this address is the first usable address in the GGC node subnet, you can just press `enter`.
   - Confirm your configuration when prompted.

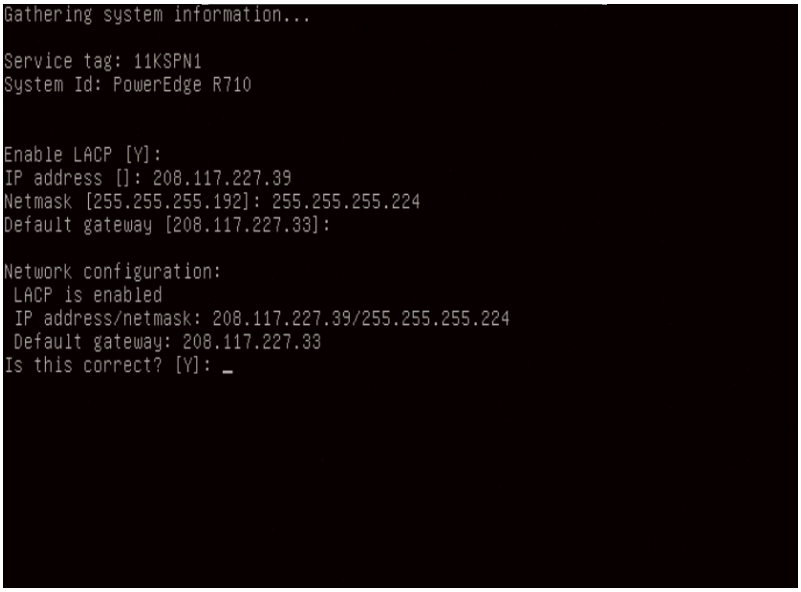

Figure 5.3: IPv4 Information

7. Upon validation of the IP information and connectivity, the server will begin the local software installation. This step will take a couple of minutes. Please be patient and allow it to finish.
8. When the installation process completes successfully, you will see a screen similar to *Successful installation*. Press enter to reboot the server.

   In case you notice warnings or error messages on the screen, please do not reboot the server (See section '*When things go wrong*').

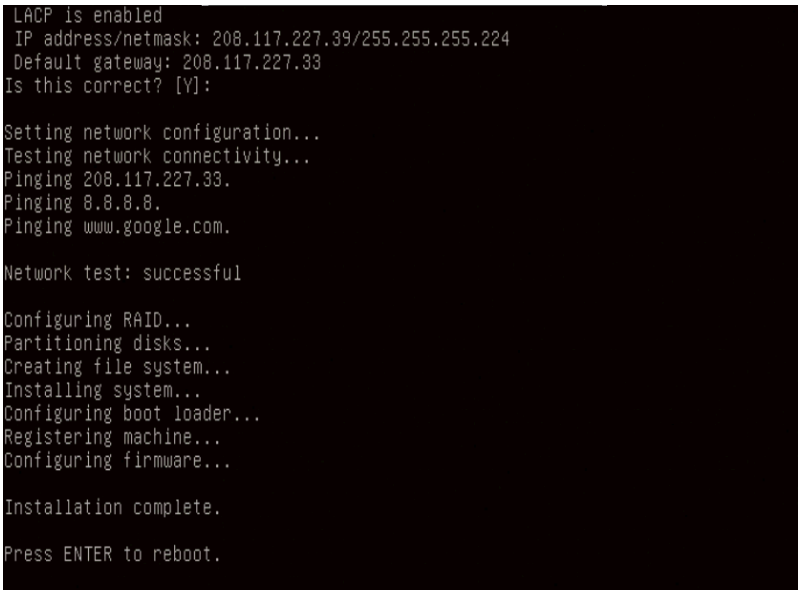

Figure 5.4: Successful installation

   The USB stick should be left in the machine, in case a re-installation is needed (See section '*GGC Software Reinstallation*').
9. When the server reboots after a successful installation, it will boot from disk. The machine is now ready for remote deployment.

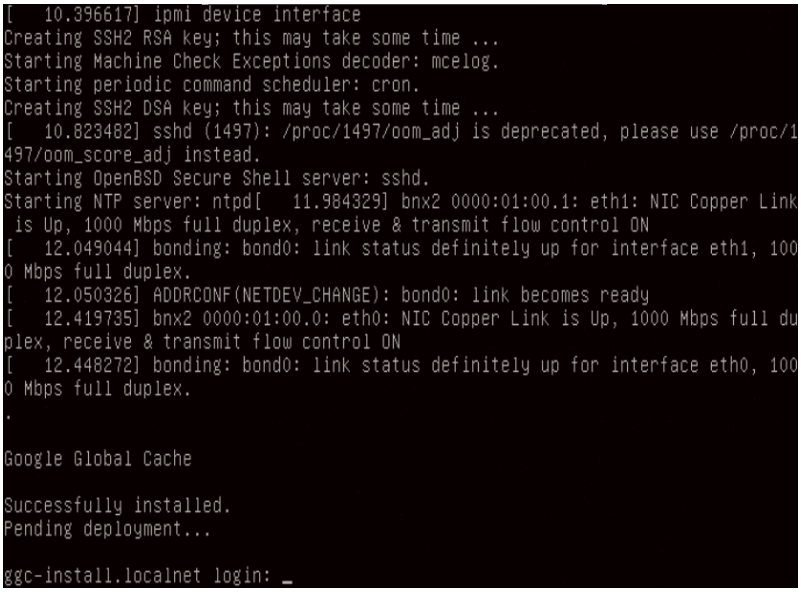

Figure 5.5: Booted from disk

10. Label each server with the IP address you have assigned to it.

Once the field setup process is complete, the setup program will automatically report the configuration to Google so that the installation can be remotely completed and the node can be brought on-line.

5.4 GGC Software Reinstallation

In some cases (e.g., when the root disk has been replaced) a server needs to be re-installed. The procedure is very similar to the one described above, the most important difference is that a manual intervention is required to boot the server from the USB boot stick.

1. Connect monitor and keyboard to the server to be installed.
2. The setup USB boot stick should still be in a USB port, it has the previously entered network configuration stored. If it is not, the USB stick can be recreated as described in section '*Preparing the USB stick (drive)*'. The stored network configuration will be lost and has to be entered again.
3. Start the server. During POST, a menu will appear on the screen, similar to the *POST screen*.

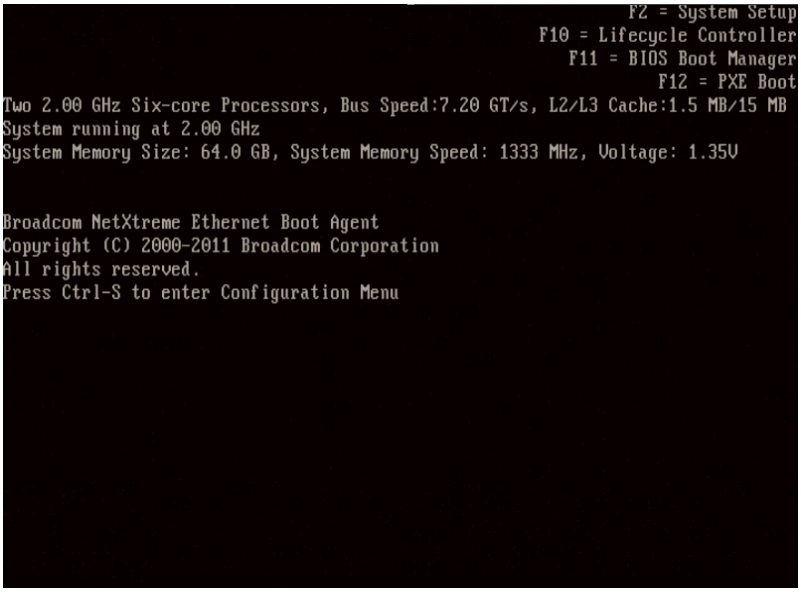

Figure 5.6: POST screen

4. Press `F11` to enter the *Boot manager*. Select the 'BIOS Boot Menu'.

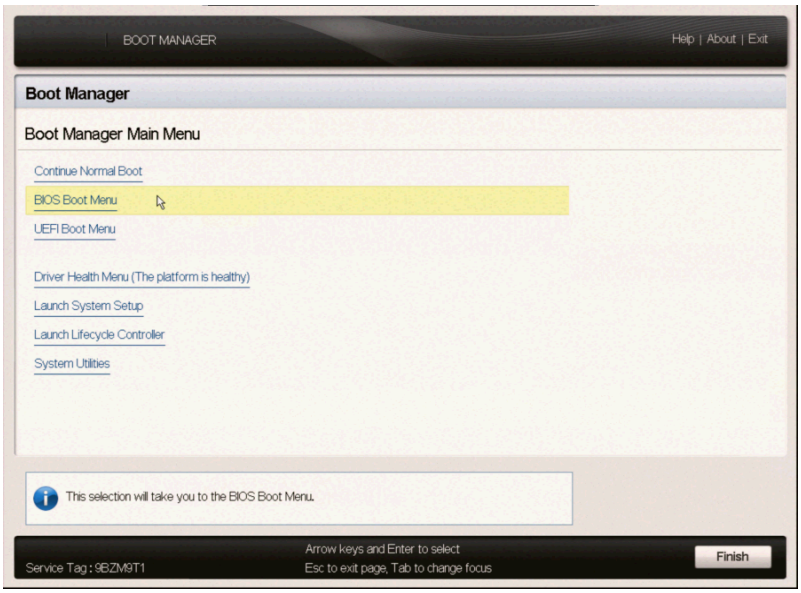

Figure 5.7: Boot manager

5. From the list of bootable devices, select 'Hard disk C:' and then the USB stick, as shown in the illustration *Boot from USB stick*

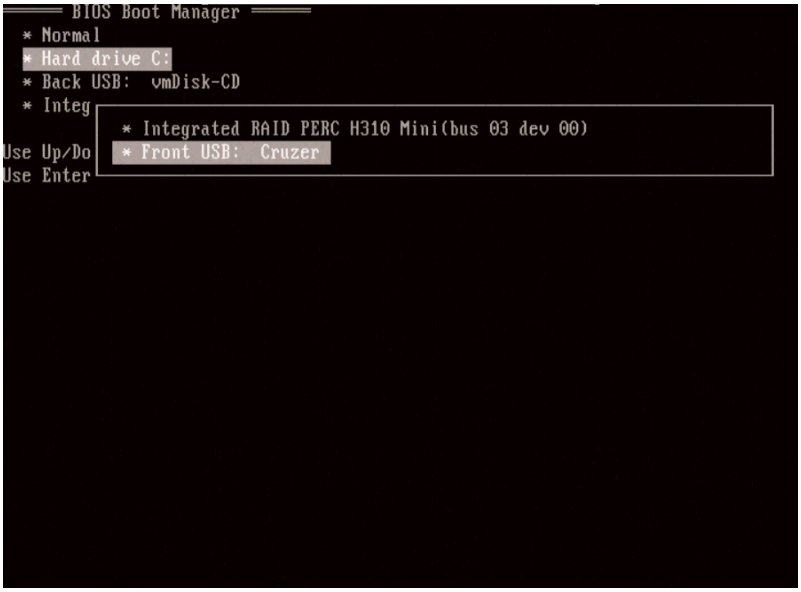

Figure 5.8: Boot from USB stick

6. Once the installation program is launched, you can proceed with the same steps as described in section '*GGC Software Installation*'.

   Note that, if the same USB stick is used as the one for the original installation, the settings for the network configuration are prefilled with the information entered previously. Provided these data are still valid, you can just press `enter` to proceed.

5.5 When things go wrong

- When network connectivity cannot be established, please check the cables, switch and router configuration, and the IPv4 information entered during installation.
- If the setup process encounters an error *after* network connectivity is established, it will automatically report the error to Google for investigation. If this happens, please leave the server running with the USB stick inserted.
- In other cases, please contact the GGC Operations team: ggc@google.com.

BGP Configuration

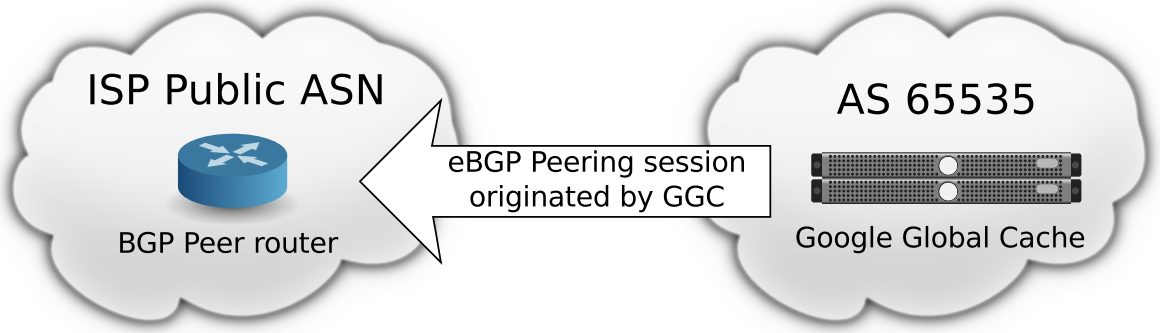

Figure 6.1: BGP overview

6.1 You will need

- IP address of the BGP peer router
- Administrative access to the BGP peer router

Note: Only a single session is permitted to each GGC node. Redundancy is not required as an interruption of this session will not impact traffic flow at the node.

6.2 Procedure

- Your end of the session will be the router specified in the GGCAdmin portal. Use your public ASN as provided in the GGCAdmin portal for your end of the eBGP session.

  Note:

  - BGP multihop is supported
  - IPv6 link-local addresses are not allowed

- The GGC end of the session will be the **16th IP address in the GGC subnet** (see also: *IP Addressing*). This is a virtual address which will be configured after Google completes the remote installation.

Note:

- the GGC ASN is always 65535

- The session should be configured in passive mode. The connection is always initiated by the GGC end.

Note:

- the session will not come up until Google completes the next step of the installation

- Do not configure monitoring on the BGP session. The GGC system does not interpret an interruption of the BGP feed as a loss of the node. The node will continue to serve based on the most recent valid feed received until the session is restored. Google will monitor the availability of the node and automatically shift traffic away in the event of an outage. Normal management activity may briefly interrupt the session at any time.
- For configuration simplicity, MD5 passwords are not recommended. MD5 passwords are supported, if required.

6.3 What to Announce Over the Peering Session

Google Global Cache uses BGP only as a mechanism to communicate the list of users that should be served from a node. It is not used for routing or to determine if the cache is online. An interruption to the BGP session has no effect on the cache.

6.3.1 User and Resolver Prefixes

In order for the GGC node to perform optimally, **both the IP address of the user and the IP address of the DNS resolver they are using must be advertised to the cache**.

While the vast majority of end user traffic is delivered by GGC nodes based on the end user's IP address alone, a small subset of requests use the IP address of the DNS resolver being used by the end user.

Optionally, Google also supports EDNS and you can increase the number of queries based on the user's actual IP address by implementing Client Subnet Information in DNS Requests.

Besides mapping, the user IP addresses are used to build an access control list on the node itself.

Note: GGC will ignore

- any /32 IPv4 prefixes
- private IP addresses (ie: RFC 1918)

6.3.2 Peers and Downstream ASNs

Prefixes from other ASNs can be mapped to the cache as well, providing the following conditions are met:

- Both the DNS resolver and user prefixes must be advertised in order to both map and serve those users from the cache
- If the other AS transits an AS with a peering relationship with Google, their traffic will not be mapped to the cache. If an exception is required, please contact ggc@google.com.
- Do not send the full Internet routing table to the GGC node. Only send the prefixes that should be served from the node.

6.4 Multiple Cache Nodes

There are a few configuration options available when multiple cache nodes are deployed in the same network. There is a brief description below. The accompanying document 'Multi-Node concepts' (GGCMultinodeDeployments.pdf) describes the options in detail.

Please indicate your configuration preference to the GGC support team (ggc@google.com).

6.4.1 Users load balanced over multiple cache nodes

- If traffic can be load balanced across multiple cache nodes, send the same BGP advertisements to all nodes.
- Failure of one node will send traffic to the remaining node(s). If the load exceeds the capacity of the remaining node(s), it will overflow to caches on Google's network.

6.4.2 Users directed to a specific cache node

In some cases, network topology dictates that it is best to serve specific sets of users from specific cache nodes.

- Advertise the user and resolver prefixes to the cache node they should prefer.
- If a DNS resolver serves users from multiple nodes, advertise that resolver to the cache node that will serve the majority of the users. Advertise the user prefixes to the specific nodes where they should be served. Users will be redirected to the desired node.

  Note: for the best user experience, it is strongly recommended to provide dedicated DNS resolvers for each cache node

  ---

- Do not advertise the same prefix(es) to multiple nodes. The mapping system uses a complex set of tie-breakers to determine the preferred location. As a consequence, traffic can shift between nodes unexpectedly and the mapping will not be optimal for a subset of users.

  Note: overlapping prefixes are accepted: the more specific advertisement will determine the preferred node

  ---

- Failover is configured at the node level. In the event of a failure, all prefixes advertised to the failed node will be served from the designated backup node or ultimately from the Google network.

6.5 BGP Peer Configuration Examples

The following examples are for illustration purposes only. Your configuration may vary. Please contact us if you require additional support.

6.5.1 Cisco Option 1: Prefix list based route filtering

```
neighbor <IP address of GGC> remote-as 65535
neighbor <IP address of GGC> transport connection-mode passive
neighbor <IP address of GGC> prefix-list deny-any in
neighbor <IP address of GGC> prefix-list GGC-OUT out

ip prefix-list deny-any deny 0.0.0.0/0 le 32

ip prefix-list GGC-OUT permit <x.y.z/24>
ip prefix-list GGC-OUT permit <a.b.c/24>
```

6.5.2 Cisco Option 2: AS-PATH based route filtering

```
neighbor <IP address of GGC> remote-as 65535
neighbor <IP address of GGC> transport connection-mode passive
neighbor <IP address of GGC> filter-list 1 in
neighbor <IP address of GGC> filter-list 2 out

ip as-path access-list 1 deny .*

ip as-path access-list 2 permit _100_
ip as-path access-list 2 permit _200$
ip as-path access-list 2 permit ^300$
```

6.5.3 Juniper Option 1: Prefix based policy

```
neighbor <IP address of GGC> {
  description "GGC";
  import no-routes;
  export export-filter;
  peer-as 65535;
  passive;
}

policy-statement no-routes {
  term default {
    then reject;
  }
}
```

6.5.4 Juniper Option 2: AS-PATH based policy

```
neighbor <IP address of GGC> {
  description "GGC";
  import no-routes;
  export export-filter;
  peer-as 65535;
  passive;
}
policy-statement no-routes {
  term default {
    then reject;
  }
}
policy-statement export-filter {
  term allow-routes {
    from {
      from as-path-group GGC;
    }
    then accept;
  }
}
as-path-group GGC {
  as-path AS-PATH-NAME-1 "^100.*";
  as-path AS-PATH-NAME-2 "^200.*";
}
```

CHAPTER 7

Server Operating Temperature

7.1 R720xd Operating Temperatures

The default configuration of these servers is calibrated toward reducing power consumption. This is accomplished by allowing the servers to run hotter than you might expect from other data center equipment. This configuration will allow the exhaust temperature to rise to 50°C (122°F), likely 10°C - 15°C (18°F - 27°F) higher than you might be used to. This is normal.

You may notice the fans spinning more slowly than you might expect. We have instrumentation in place that collects this data, and have seen fans running at approximately 25% of their rated speed in warm environments, with no difficulty in keeping the server to its desired temperature.

This change allows you to save power and cooling costs, with no adverse effects on the servers.

CHAPTER 8

Operations and Troubleshooting

8.1 Shutdown and Traffic Drain

In the event you need to shut the node down for scheduled maintenance of your data center or network, we ask that you inform us ahead of time so that we know the downtime is expected. You may still receive automated alert messages generated by our monitoring systems. If the outage is temporary and expected, these can safely be ignored.

Note: Please do not simply shut down the node without a graceful traffic drain: users might experience service interruptions.

A **graceful traffic drain** can be initiated either by you or by the GGC Operations team:

- You can perform a graceful traffic drain by shutting down the ethernet interfaces facing **all but one** of the servers. When the Google monitoring system detects only a single server in a node is reachable, the mapping system will stop sending traffic to the node. As a consequence, the Google authoritative DNS servers will no longer hand out the node's IP addresses. Due to DNS response caching, it can take up to 30 minutes for user traffic to fully drain away from the node. The one remaining server will redirect excess load, so no users will be denied service while the drain is taking effect.
- You can request the GGC Operations team to drain the node for you (ggc@google.com). For the reasons described in the paragraph above, coordinate with the GGC operations team as it takes up to 30 minutes before the node is fully drained.

If you need to power down a GGC machine, you can initiate a graceful shutdown by pressing the server's power button once. There will be a several second pause before the system shuts down.

To restore traffic to the node, simply re-enable all switch interfaces or restore power to all machines. Once the monitoring system detects more than one machine is reachable, the mapping system will start sending user requests to the node again. This can take a while, so please be patient.

8.2 Hardware Monitoring and Repair

Google's monitoring system will remotely identify hardware failures. Your technical contact will be notified if we require any local assistance, troubleshooting, or RMA coordination. If you believe that hardware is not operating properly, please contact us at ggc@google.com.

Note: Keep the technical contacts section of the GGCAdmin portal (ggcadmin.google.com) up to date. The GGC operations team relies on this information in the event you need to be contacted.

8.3 Local Monitoring

While no monitoring is required, some local monitoring can be helpful. However, it is important to understand the following considerations:

- It can be helpful to monitor the availability and performance of the path between the GGC node subnet and Google's network. A sample host for video content origin is `v1.cache1.googlevideo.com`.
- If you monitor egress traffic from the node, bear in mind that traffic at the node will be impacted by `youtube.com` maintenance and availability.
- Binary and configuration changes are regularly pushed to machines in the node in a rolling fashion. If you are monitoring egress per machine, you will see occasional interruptions of service during the associated restarts. The GGC software ensures that the load for the machine under service is spread around the node during these events.
- The BGP session to the node is different from typical peering sessions. It is not used for routing or to establish the availability of the node. Brief interruptions of the session are normal and will not impact user traffic. If you are monitoring this session, you should not consider it an actionable alert unless the session is down for longer than an hour.

8.4 Playing a Test Video

See illustration *Developers Tools in Google Chrome*.

Using Google Chrome

1. open the 'Developers Tools' (Wrench Button > Tools > Developer Tools or `Shift+Ctrl+I`)
2. open the 'Network' tab
3. point your browser to www.youtube.com
4. play a popular video
5. watch the item 'videoplayback', this shows the name of server the video is actually played from:

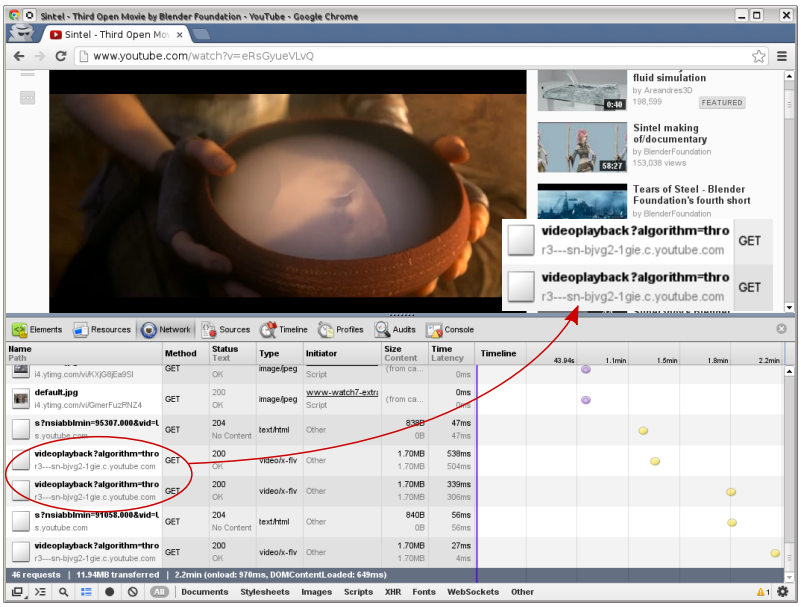

Figure 8.1: Developers Tools in Google Chrome

6. resolve this name using `nslookup` or similar tools:

```
$ nslookup r3---sn-bjvg2-1gie.c.youtube.com
Server:        <your_name_server>
Address:       <your_name_server IP>

Non-authoritative answer:
r3---sn-bjvg2-1gie.c.youtube.com
   canonical name = r3.sn-bjvg2-1gie.c.youtube.com.
Name:   r3.sn-bjvg2-1gie.c.youtube.com
Address: 193.142.125.14
```

If the resulting address (193.142.125.14 in this example) is an address in the subnet allocated to the GGC node, the video is playing from the cache.

Note: The base web pages of `www.youtube.com` may not be served from the cache. These host names will typically not resolve to the GGC node.

You can use FireFox as well to perform this test, but you will need to install the *FireBug* extension.

8.5 Videos Not Playing From the Cache

There are several possible reasons why a video may not play from the cache.

8.5.1 The user's DNS resolver is not in the BGP feed to the GGC node.

One of the mechanisms the mapping system uses to send requests to a node is DNS. The DNS request from the user will go to your resolver, which will then come to Google's authoritative resolvers. If your resolver's IP address is in a prefix that is being advertised to the node, the IP address returned to your DNS resolver (and then to the user) should be from the GGC node subnet. To determine the resolver Google is seeing, execute the following command from your test client:

```
nslookup -q=txt o-o.myaddr.l.google.com
```

Use *o-o.myaddr.l.google.com* verbatim, do not substitute the *myaddr* part. This is a special host name that will return the IP address of the DNS resolver as seen by Google. You should see a response similar to:

```
Non-authoritative answer:
o-o.myaddr.l.google.com.google.com
text = "<IP_address>"
```

Confirm that *IP_address* is in the BGP feed to the GGC node. A common error is using a test resolver that forwards requests to another DNS server whose IP address is not in the BGP feed.

The mapping file is updated periodically, so it takes some time before changes in the advertised BGP feed are picked up by the mapping system. If the address was added to the BGP feed within the last 24 hours, please contact ggc@google.com to confirm that the change has been pushed to our production servers.

8.5.2 The client's IP is not in the BGP feed to the GGC node

If the requested video is not playing from the cache, it is possible that the BGP feed does not include the test client's IP address. If this is the case, the cache will get the request and then redirect it to a cache outside your network. Verify that the test client's IP address is in the feed and has been there for at least 1 hour.

8.5.3 The cache is overloaded and overflowing

If the requested video plays from the cache sometimes, but not every time, the cache may be overflowing. As the cache reaches its configured serving capacity, it will begin redirecting requests to external caches. The service

capacity of the cache is based on a combination of several factors:

- number of servers in the node
- number of interfaces connected on each server (LACP)
- available bandwidth provisioned between the cache and your network (reported on the GGCAdmin portal)
- manually configured limits

You can determine if the cache is overflowing by reviewing the '*Traffic Graph*' on the GGCAdmin portal (follow the 'Status' link). An example graph is shown below *GGCAdmin Traffic Status Graphs*. If you suspect that the node should not be overflowing at the current traffic level, it is possible that an out of date limit is configured. Contact ggc@google.com to confirm that the capacity is set correctly.

8.5.4 The video is not popular enough to be in the cache

The cache will store the most popular videos your users are requesting. There is an admission mechanism that can prevent a video from being cached on the first play. If this is the case, a second playback should come from the cache.

8.6 Node Status in the GGCAdmin Portal

During the installation process, the 'Status' tab in the GGCAdmin portal will provide information on the fulfillment, shipping, and turn-up status of the node.

After the node is activated, the 'Status' tab will provide several graphs that can be useful for evaluating node performance.

See illustration *GGCAdmin Traffic Status Graphs*.

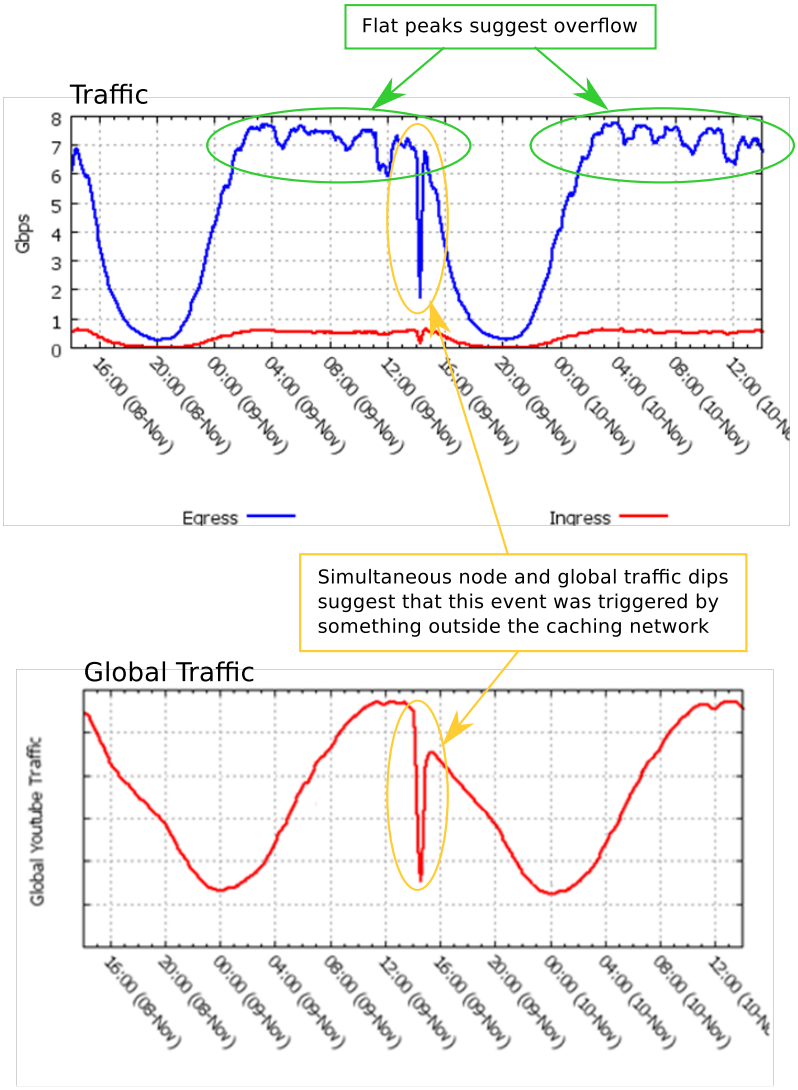

Figure 8.2: GGCAdmin Traffic Status Graphs

8.6.1 Traffic and Global Traffic Graphs

The *Traffic Graph* shows ingress and egress traffic:

- ingress traffic is cache fill coming from Google's origin servers
- egress is traffic from the cache sent towards the users

The ratio between the two will show you the cache's efficiency. If the egress line appears flat at the peaks, it is possible that there is not enough cache capacity and traffic is overflowing. Please contact ggc@google.com to find out what can be done.

The *Global Traffic Graph* shows the global egress traffic from the Google Global Cache network. This graph is useful in determining if a traffic interruption at your node was part of a larger, global event. If so, then the most common explanation is an error or maintenance on the `youtube.com` site, which can instantly reduce global demand from the caches.

8.6.2 Packet Loss

The *TCP Retransmits Graph* shows transmit packet loss (measured by retransmissions) from the node. If you are observing slow video playbacks or significant rebuffering, this graph can tell you if the node is having difficulty

reaching your users. High packet loss is most often caused by congestion or faults in the access network between the cache and the users. The issue can sometimes be traced to a network bottleneck, such as improper load balancing across aggregated links, a faulty circuit, or a malfunctioning interface.