Google
Global Cache

Multinode Deployment Guide

Release

August 27, 2013

Google Global Cache
ggc@google.com

Contents

Introduction

This document will help guide you through the configuration decisions related to adding multiple GGC nodes to your networks and explain how your users will be served when you have multiple GGC nodes deployed.

To fully understand this document, it is necessary to be familiar with the concept of a GGC Network Location (GNL). **A GNL is a logical grouping of multiple GGC nodes, which serve the same users; typically these are nodes within the same physical location or Metropolitan area, within 10ms RTT of each other.**

You can have multiple GGC Network Locations within your network, but a GGC node can only be a member of *one* GGC Network Location.

When adding a second or subsequent node to your network the most important thing to consider is if the node will serve the same users, or if it will be deployed regionally inside your network (serving different sets of users). Once you understand this, your deployment options are simple.

Note: The details in this document are subject to change. This document and associated activity are commercially confidential.
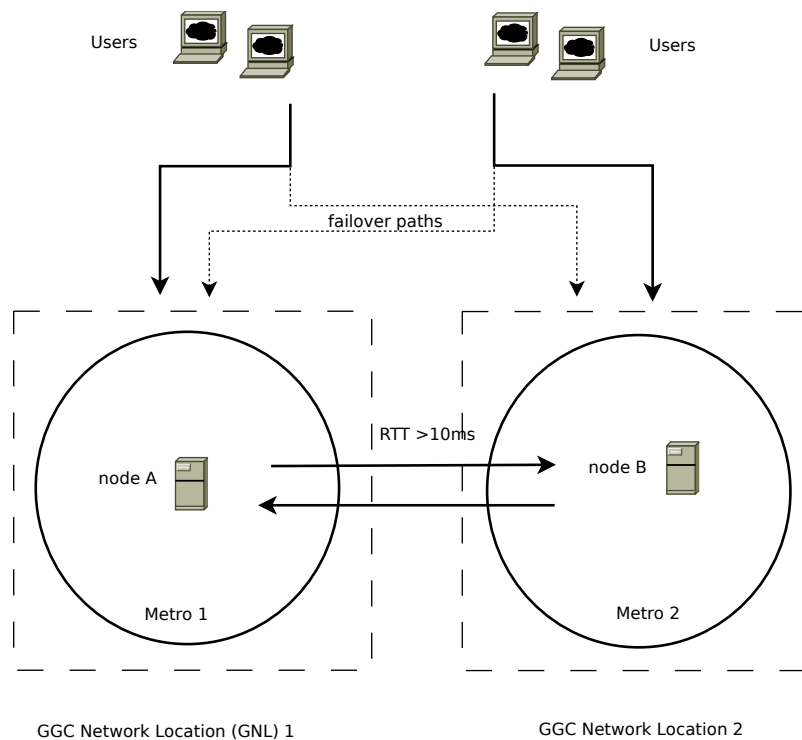

Multiple GGC Network Locations

GGC nodes intended to serve distinct sets of users should be deployed in separate GNLs. This allows you to regionalize your traffic so that it closely resembles your user topology.

Figure 1

In this example GNL 1 and GNL 2 are each serving distinct sets of users. They are also optionally configured as failover locations for each other.

What's the benefit?

This will regionalize your user traffic to the desired serving location (GNL), reducing traffic across your backbone.

What prefixes should I announce?

With this configuration, you must announce unique user & resolver prefixes to each GNL.

If you would like a GNL to act as a failover for another GNL, you should announce the same set of prefixes to both GNLs, but with a longer AS Path at the failover location.

What if I only have one resolver prefix for all of my users?

In this scenario, you should announce your resolver prefix to all your relevant nodes, and use prepending in the following way to ensure service continuation in the event of node failure.

On the GNL which serves the largest set of users, the Resolver prefix AS path should be as short as possible. On the remaining GNLs, the AS path should be longer.

Please note that even though you announce the same resolver prefix to more than one node, it will only get mapped to one GGC GNL at a time.

What happens when a node fails?

Regionalized GNLs can also be configured to act as failover locations for each other. This is achieved by announcing the same prefixes to each node/GNL, and prepending those announcements to the failover locations.

When a node (or part of a node) fails, any portion(s) of the catalog that can no longer be accessed, will automatically continue to be served from caches in your chosen failover locations. To provide the best user experience at all times, we may *occasionally* serve some user traffic from caches on Google's Network.

When no further failover locations are available for that portion of the catalog, that content will continue to be served from caches within Google's network.

This method of service continuation is automatic and transparent to your users, and ensures that your users continue to be served during any server or system maintenance.

Overflow traffic

If user traffic exceeds the capacity of the node(s), any additional traffic will automatically overflow to other cache locations.

If you have configured the GNL's to also act failover locations, overflow traffic will follow the same rules as described in the "failover" section above.

Peer to Peer cache fill

It is possible to configure GGC nodes for Peer to Peer cache fill, which can further reduce load on your peering and transit links.

When this is configured, cache fill will take place between nodes in your selected GNLs. Should the content not be available from those locations, it will be requested from Google's upstream servers.

If you wish to ensure that Peer to Peer cache fill is configured between specific GNLs in your network, please communicate this to the ggc@google.com team during the turnup phase of the new node.
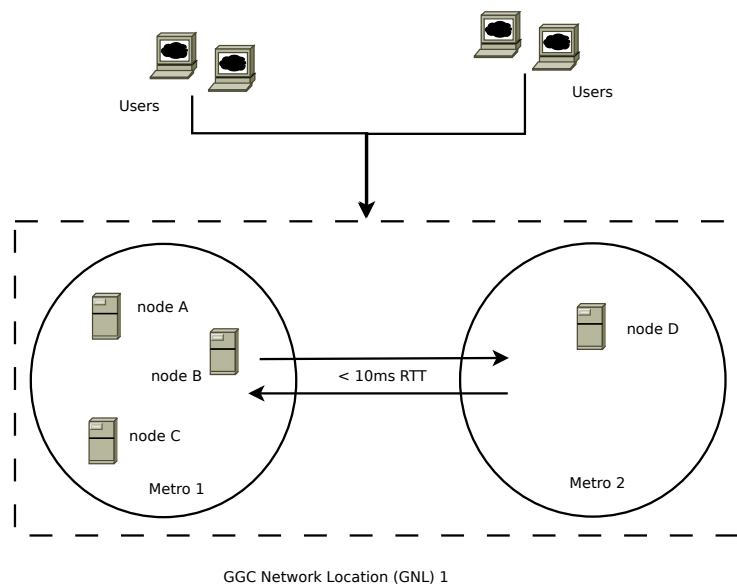
Multiple nodes in the same GGC Network Location

In order to provide additional serving capacity to users in a specific location, two or more nodes can be deployed within the same GNL. The cache catalog is automatically spread across the nodes in that location.

Figure 2

GGC Network Location (GNL) 1

In this example, Nodes A, B, C, and D are *all* within 10ms RTT of each other, so they are part of the same GNL.

What's the benefit?

This will increase the size of the cache served from the GGC nodes within that location, meaning more requests can be served locally, which can help to further reduce traffic on your peering and transit links.

What prefixes should I announce?

With this configuration, you must announce the same user & resolver prefixes to each node (with equal AS path length).

What happens when a node fails?

If you have multiple GNLs, they can be configured to act as failover locations for each other. This is achieved by announcing the same prefixes to nodes in other GNLs, and making sure to prepend those announcements.

When a node (or part of a node) fails, any portion(s) of the catalog that can no longer be accessed, will automatically continue to be served from caches in your chosen failover locations. To provide the best user experience at all times, we may *occasionally* serve some user traffic from caches on Google's Network.

When no further failover locations are available for that portion of the catalog, that content will continue to be served from caches within Google's network.

This method of service continuation is automatic and transparent to your users, and ensures that your users continue to be served during any server or system maintenance.

Overflow traffic

If user traffic exceeds the capacity of the node(s), any additional traffic will automatically overflow to other cache locations. You can express a *preference* on where you would like traffic to go to in this scenario, by prepending your prefix announcements to other nodes within your network (nodes in other GNLs).

Google will take this *preference* into account as it dynamically determines which cache locations will provide the best user experience, before seamlessly redirecting that user traffic. Note that this might not always be to a cache within your network, and could instead be a cache from within Google's network.

Frequently Asked Questions

How does overflow and failover work?

You can view your current failover configuration on the GGC Portal (https://peering.google.com/failover/).

We use AS Path length to determine where you would prefer traffic to overflow *and* failover to. You can specify multiple failover and overflow locations.

For example:

```
GNL         Prefix          AS Path length
---         ------------    ------------------
 A          10.10.10.0/24   12345-12345
 B          10.10.10.0/24   12345-12345-12345
```

GNL B is designated overflow and failover location for GNL A.

```
GNL         Prefix          AS Path length
---         ------------    ------------------
 A          10.10.10.0/24   12345-12345
 B          10.10.10.0/24   12345-12345-12345
 C          10.10.10.0/24   12345-12345-12345-12345
```

Now GNL B is the *primary* overflow/failover location, and GNL C is the *secondary* overflow/failover location. GNL C is only used when GNL B is no longer able to handle traffic.

You should also be aware that the prefixes would need to be identical for this to work. Overlapping prefixes will cause unexpected behaviour regarding where traffic is served in general (see next point regarding Prefix Length).

AS Path length, or Prefix length?

Both of these matter, although in differing circumstances.

When identical prefixes are advertised to multiple GNLs, AS Path length is used to determine the preferred GNL to serve those users. Please see the section above for examples.

If overlapping prefixes are advertised to distinct GNLs, the more specific prefix length is used to determine where those users will be served (in this case AS Path length is ignored).

For example:

```
GNL         Prefix          AS Path length
---         ------------    ------------------
 A          10.10.10.0/23   12345-12345
 B          10.10.10.0/19   12345-12345-12345
 C          10.10.10.0/24   12345-12345-12345-12345
```

GNL C is the prefered location to serve the users.