

## Open Connect Appliance Deployment Guide

V 2.4, May 29, 2012



# Contents

---

OVERVIEW .....	3
OPEN CONNECT APPLIANCE PLANNING .....	4
SHIPPING AND RECEIVING OPEN CONNECT APPLIANCES .....	7
RACKING/INSTALLING OPEN CONNECT APPLIANCES.....	7
CONFIGURING OPEN CONNECT APPLIANCES .....	7
ENABLING THE OPEN CONNECT APPLIANCE.....	8
TESTING THE OPEN CONNECT APPLIANCE.....	8
MONITORING AND MAINTENANCE .....	9
HARDWARE SUPPORT .....	9
APPENDIX 1: FAQ .....	10
APPENDIX 2: OPERATING LEVEL AGREEMENT .....	12
APPENDIX 3: BGP CONFIGURATION.....	15

## Overview

---

The Open Connect appliance is a *directed cache appliance*. Its purpose is to allow participating ISPs the ability to manage the flow of Netflix content across their networks. As Netflix traffic is currently more than 30% of peak traffic on US ISP networks, the ability to localize this traffic can represent a significant reduction in the utilization of network transport, as well as significant offload from upstream peer and transit networks, and correspondent reduction in capital and operational expense. It is a *directed cache appliance* in that the manner in which traffic is directed to and cached on the appliance is determined by the ISP and by Netflix, but *not* by the appliance itself. More specifically:

- ISPs control (by BGP advertisement) the mapping of their customers' IP netblocks to particular clusters of appliances
- Netflix control servers direct content requests from Netflix client devices within the advertised netblocks to the appliance when:
  - The requested content is cached on the appliance
  - The health/load of the appliance is sufficient to satisfactorily handle the request

In this way, ISPs can control the sources from which Netflix traffic flows across their network, including planning for capacity and availability, but are relieved from all other aspects of appliance management, including maintenance, repair, and replacement.

## How Traffic is directed to the Open Connect Appliance

The Open Connect appliance is a passive device, which only serves Netflix video content in response to requests sent to it by the Netflix Content Delivery System. Unlike a transparent or other type of proxy, or a demand-driven caching appliance, it is not an active network device and does not attempt to inspect or intercept traffic flowing across the network; rather it only handles traffic purposely directed to the appliance. Requests are only sent to the appliance for content that already exists in the cache on the appliance.

- Netflix places only Netflix audio and video files in the cache
- The operation and performance of the cache affects only the set of ISP subscribers who also subscribe to Netflix

## Peering with the Netflix Open Connect Network

Open Connect appliances are typically deployed in conjunction with peering with the Netflix Open Connect network at common Internet Exchange locations. The percentage of traffic served via peering can vary substantially, depending on the manner in which appliances are deployed within an ISP network. For extensive ISP deployments, which include sufficient capacity to store and serve the entire Netflix catalog, peering is only required to ingest new content. For other types of deployments, peering can accommodate appliance fill traffic, as well as client access to very "long tail" content

which is not cached locally. For ISPs who do not choose to peer with the Netflix network, content will be filled from any of several CDN and transit networks, and Netflix will not be responsible for any costs associated with appliance fill traffic.

## Open Connect Appliance Hardware

The Open Connect appliance is a 4U device pre-configured with Netflix content delivery software and pre-loaded with content. The appliance you receive is the same field-proven appliance Netflix uses in its own data centers and peering locations.

- Netflix remotely monitors the health of each appliance
- Health and load information is automatically incorporated into Netflix content routing decisions
- Appliances are designed to tolerate certain types of hardware failures without affecting performance.
- These include redundant hard disks, power supplies, and network interface ports

Netflix will arrange next day delivery of a replacement appliance in the event of any production-affecting failure.

## Open Connect Appliance Software

The Open Connect appliance is a high-performance web server integrated with the Netflix content delivery system.

- The Operating System is FreeBSD
- The Web Server is nginx
- The BGP daemon is BIRD
- The remaining software on the system manages content and communicates system health and other statistics to the Netflix content delivery system
- IPv4 and IPv6 are fully supported
- Ports which need to be allowed to the Open Connect appliance are:
  - TCP 22
  - TCP 80
  - TCP 179
  - TCP 443
  - UDP 161
  - UDP 514
  - ICMP (echo request and reply)

## Open Connect Appliance Planning

---

When you decide to deploy a an Open Connect appliance, Netflix will work directly with your network capacity planning team to analyze traffic patterns within the various regions of your network, and determine the appropriate number of appliances for each location.

## Traffic Analysis and Sizing

The number of appliances deployed is determined by two factors: the availability required at a given location, and the percentage of Netflix traffic you wish to originate from that location

- An individual appliance can offload approximately 70% -90% of content requests, depending on country catalog size
- An individual appliance can deliver up to 8 Gbps of peak network throughput
- Deploying additional appliances can increase the total throughput as well as the percentage of content requests served, as content can be sharded across the disk footprint of available appliances
- Appliances are typically deployed in an n+1 configuration for availability
- Multiple tiers of appliances can be deployed within an ISP network, depending on density of traffic in individual locations, as well as connectivity between metro regions (see BGP Configuration, [p. 15](#) for more detail on how tiers are configured).

## Hardware Options

Appliances are available with several different hardware options to suit the specific needs of each location.

Options	
Rack Size	2 or 4 post rack (19" wide)
AC or DC Power Supply (Peak power draw: 560 watts)	<p>AC:</p> <ul style="list-style-type: none"> <li>• C14 connector</li> <li>• Voltage: 90 ~ 264 VAC Full Range</li> <li>• Freq: 47 ~ 63Hz</li> </ul> <p>DC:</p> <ul style="list-style-type: none"> <li>• PLB06M connector</li> <li>• Voltage: -36 ~ -72 V DC</li> <li>• Input current: 16A at -48V D</li> </ul> <p><b>Note:</b> For configurations requiring DC power supplies, Netflix provides 20-ft long power cables with PLB06M connectors.</p>

Open Connect appliance with AC power supplies:



Open Connect appliance with DC power supplies:



## Typical Deployments

Within an ISP network, appliances are generally placed in clusters sufficient to the demand of the local metro. As a rough rule of thumb, at least one appliance should be placed in any location which has 5 Gbps or greater of Netflix traffic. The Netflix capacity team can help you identify how much Netflix traffic is on your network, and how to properly deploy and configure appliances to optimally manage Netflix traffic on your network.

## Shipping and Receiving Open Connect Appliances

---

After planning is complete, Netflix will arrange to ship appliances directly to your datacenter locations identified for deployment during the planning process. Your datacenter (or other designated location) must be able to receive the appliance shipment as follows:

- Estimated appliance weight: 100 pounds
- Estimated weight of pallet and shipping materials: 50 pounds

### Replacements Appliances and Returns

In the event of a performance impacting hardware failure, Netflix will send a replacement appliance. The replacement will include return shipping forms for the defective unit. Ship the defective unit back using the pallet and container the replacement arrived in.

## Racking/Installing Open Connect Appliances

---

Netflix ships the appliance to your designated location ready to deploy. You provide space, power and a 10 Gbps optical network connection.

- The appliance is shipped ready to operate.
- Rack the appliance, plug in power and network

After racking the appliance, proceed to [Configuring Open Connect Appliances, p.3](#).

## Configuring Open Connect Appliances

---

You control traffic routing to a single appliance or a cluster of appliances, as well as tiering, failover and redundancy among multiple sets of appliances and peered connections to the Netflix network. See [BGP Configuration, p.15](#) for details.

### Firewall Considerations

See ports specified in [Open Connect Appliance Software, p.4](#).

## Appliance Fill

For the US Netflix catalog, a fill rate of approximately 5 Gbps for approximately 2 hours (or equivalent) is currently required. A lower speed link will require a longer fill window. When multiple appliances are present in a single location, appliances can fill from one another, reducing the upstream network load.

An appliance updates its cached content nightly with the appropriate content. The exact time window is determined for each appliance during the planning process, and is generally chosen by the ISP as an off-peak period, such as early morning hours in the local time zone.

## Enabling the Open Connect Appliance

---

After you have established the BGP session with the appliance (described in BGP Configuration, *p. 15*), contact Netflix at [cacheops@netflix.com](mailto:cacheops@netflix.com) to enable clients to start accessing your appliance.

## Testing the Open Connect Appliance

---

Netflix extensively tests every revision of the appliance hardware and software with every supported Netflix client (as incorporated into various consumer electronics devices such as digital televisions, Blu-ray players, game consoles, iPads, etc.). There is no need for ISPs to perform client-based testing or analysis as this is an integral part of the Netflix service offered to our subscribers.

Since the appliance is a passive device, which responds exclusively to requests sent to it from the Netflix Content Delivery system, lab testing of this device is neither effective nor necessary. We have found in practice that the best way for ISPs to become familiar with appliance operation is to work with the Netflix Content Delivery team to initially enable small amounts of production traffic to the appliance.

Netflix collects a wealth of performance data for each subscriber interaction, and we are able to share summarized performance data with participating ISPs during turn-up, and in an ongoing fashion once appliances are in full production.

## Monitoring and Maintenance

---

Netflix continuously monitors all aspects of the quality and availability of the Netflix service, including the performance and availability of the appliance. See Operating Level Agreement, [p.12](#) for details.

### Automatic Software Updates

Netflix manages appliance software, and may make occasionally update the appliance software. Any updates are performed during the cache fill window or at a time mutually agreed upon with the ISP. When updating multiple appliances, “rolling upgrades” are performed to ensure ongoing service availability. All appliance software, including OS, web server and control code is updated as a single image. In the event a new image fails to properly load, the appliance will automatically boot to the last good image.

## Hardware Support

---

If an appliance has defects that cannot be resolved in the field, Netflix will overnight ship a pre-configured replacement appliance. See Shipping and Receiving Open Connect Appliances, [p.7](#) for important shipping details.

## Appendix 1: FAQ

---

### How much network offload will the appliance help me achieve?

The number of appliances deployed and the volume of traffic served through that location determine the percentage of total traffic originating from appliances within a certain location.

A single appliance can offload in the range of 70%-90% of Netflix traffic, depending on country catalog. Multiple appliances operating together in a single region can provide even greater levels of traffic offload, approaching 100% for deployments of sufficient scale.

### What happens in the event of a recoverable hardware fault?

Contact Netflix engineering using the escalation process described in Operating Level Agreement, [p.12](#).

### What happens in the event of a non-recoverable hardware fault?

If the Netflix engineering team cannot recover the appliance, a new one will be shipped to you. Replacement guidelines are described in Operating Level Agreement, [p.12](#).

### How do I determine if an additional appliance will benefit a given data center?

Initially, Netflix will work with your network-planning group to deploy the appropriate number of appliances to optimize offload for your network. Additionally, Netflix will continually monitor your traffic needs and recommend additional appliances as your needs change.

### How can I change the appliance cache-fill window?

The fill window is initially determined based on your input and an analysis of your network conditions. If you need to change it, contact [cacheops@netflix.com](mailto:cacheops@netflix.com).

### Is any field service required for the appliance?

No. In general, you should not attempt to modify, upgrade, or repair the appliance.

### What throughput does the appliance deliver during peak demand?

For planning purposes, we estimate peak throughput for each appliance at 8 Gbps.

### What throughput does the appliance need during fill?

This is a variable that you control within some basic constraints. For the US Netflix catalog, a fill rate of approximately 5 Gbps for approximately two hours (or equivalent) is currently required. A lower speed link will require a longer fill window and vice-versa. When multiple appliances are present in a single location, appliances can fill from one another, reducing the upstream network load.

## What does the appliance cost my organization?

The appliances (and any necessary replacements) are provided to participating ISPs free of charge when used within the terms of the license agreement.

## How can I test the offload performance of the appliance before live deployment?

Netflix can provide estimates for you during planning discussions, but in general, there is very little value to testing in a lab environment. The benefits of the appliance are only realized when they serve video streams to Netflix members. This cannot be replicated in a lab setting.

## How does the appliance determine what content to cache?

Netflix uses a popularity-based algorithm to determine fill content for appliances, and content may be sharded across multiple appliances. A number of factors affect popularity, and the contents of an appliance are expected to change on a regular basis. Currently, in the United States, about 5% of the content on the appliance changes daily.

## Is the popularity calculation used to populate the appliance done on a regional basis?

Testing has shown that a single popularity per country catalog is the most effective model for providing optimal popularity to content change ratios. This analysis continues on a regular basis, and the system allows for continual improvement of popularity algorithms.

## What size IP address blocks should I assign to the appliance?

The appliance supports IPv4 assignments of /31 and larger and IPv6 assignments of /126 and larger. It is acceptable to number the appliance out of space assigned to a larger subnet.

## Appendix 2: Operating Level Agreement

---

The Netflix network Operations Center (NOC) serves as the single point of contact for Netflix on all Open Connect appliance issues requiring support or escalation. The Netflix NOC is staffed 24x7x365 and can be reached by email at the address listed in Escalation Contacts (below).

This document outlines the operating level agreement between the Netflix NOC and its counterpart operations teams at ISP partner locations. Its purpose is to clearly communicate expectations on communications, escalation, and support requirements to ensure maximum service availability for both organizations.

The support organization structure outlined in the information exchange template (below) will help each organization understand the capabilities available at each point of escalation.

There are three types of outage or impact communications: *Notification*, *Escalation*, and *Maintenance Notifications*; these are outlined below.

**Notification of outage** – provides notification and status of impact

When the Netflix NOC (or ISP partner) detects a service-impacting event, they will send an *Issue Detected* notification by email to the designated support mail alias within fifteen minutes. The mail notification will contain severity, ticket tracking number, impact statement, and current status. For all ongoing incidents, the ISP partner *and* the Netflix NOC will send a *Status Update* email every 30 minutes containing severity, ticket tracking number, impact statement, and current status. The resolving party will send a *resolution status* email within fifteen minutes of resolution.

Outage Notification Timeline	
00:00	Issue detected
00:15	<i>Issue detected</i> Notification Email is sent
00:30	<i>Status update</i> emails sent every 30 minutes
<b>Issue resolved</b>	
+00:15	<i>Resolution Status</i> email is sent

**Escalation for unplanned outage** – engage partner for support in resolving outage  
 The Netflix NOC or ISP partner will immediately send an *Escalation Notification* email to the designated support mail alias containing severity, ticket tracking number, impact statement, and repro steps. Phone contact will be initiated once the mail is sent to ensure receipt, clear understanding of the issue, and partner ticket tracking information. The Netflix NOC or ISP partner will escalate internally within ten minutes and will bring engineering support to a conference call within thirty minutes.

For all ongoing incidents, the Partner (or Netflix NOC) will send a *Status Update* email every 30 minutes containing severity, ticket tracking number, impact statement, and current status. The resolving party will send a *Resolution Status* mail within fifteen minutes of resolution.

Escalation Timeline	
00:00	Escalation Sent
00:05	Phone contact
00:10	Internal escalation to engineering
00:30	Engineering resource joins conference call
00:30	<i>Status update</i> emails sent every 30 minutes
<b>Issue resolved</b>	
+00:15	<i>Resolution Status</i> email is sent

**Notification of planned outage or maintenance** – provides notification of all planned maintenance

The Netflix NOC (or ISP partner) will send a *User Impacting Maintenance* email to the designated support mail alias seven days prior to any planned outage or scheduled maintenance activity that will result in users being impacted. For emergency service activities that do not allow for seven days notice, provide notification as early as possible. The Netflix NOC (or ISP partner) will send a *Confirm Maintenance* email two hours before the start of the planned maintenance activity to signal that the maintenance *will* be performed as expected, and will also send *Notification of Maintenance Start* email at the beginning of the maintenance process. The Netflix NOC (or ISP partner) will send email notification when the maintenance activity is completed to signal an all clear.

When scheduling maintenance, please note:

- Whenever possible, any planned maintenance or downtime should happen during the cache fill window.
- For planned work that must happen outside of a cache fill window please allow 2-3 hours after withdrawing routes from the appliance so that any Netflix viewing activity currently associated with the appliance is not interrupted.
- Under normal conditions, appliance software updates will happen automatically during the cache fill window (when the appliance is not serving customers).

Escalation Timeline	
7 days before	First notification of user-impacting maintenance
-02:00	Confirm maintenance
00:00	Notification of maintenance start
Maintenance complete	Email notification

**Appliance Replacement Timelines** – Appliances that are not recoverable will be replaced according to the following schedule:

Appliance Replacement Timeline	
<b>Service impacting</b> - an appliance failure which degrades service	Ship within one business day for next business day delivery
<b>Availability impacting</b> - service not degraded but redundancy is compromised	Ship within one business day for next business day delivery
<b>All other replacements</b>	Delivery within 2 weeks

## Escalation Contacts

Netflix	Partner
Role: <b>Netflix NOC</b> Phone: E-mail: Availability: 7x24x365	Role: Partner NOC Phone: [NOC phone number] E-mail: [NOC email] Availability: [hours] - (time-zone)

## Appendix 3: BGP Configuration

---

IP prefix and metric (MED) data are used to determine which network blocks an appliance or appliances will serve. At the time of initial configuration, a BGP session is established with the appliance. The routes received via this session are summarized and inserted into the Netflix content routing system. Content request from clients in a given netblock are routed to the appliance (or appliances), which have received the route matching that netblock. Routes received by the appliance are *not* used to determine routing for traffic outbound from the appliance; rather the appliance uses a single default route pointed at the gateway of the ISP networks' choosing. Each appliance in the ISP network is required to maintain a BGP session with a router from which it receives routing information.

### BGP Session Details

Three elements are required in order to bring up a BGP session with an appliance:

- Globally routable Autonomous System (AS) number
- IRR (Internet Routing Registry) AS or AS-SET object with associated route objects (see <http://www.radb.net/tutorials/how1.php> for an example), including downstream/customer blocks or references
- IP address of router to establish BGP session with (multihop permitted)

### Traffic Management

#### Setting Client-to-Cache Mapping Preferences via BGP MEDs

ISPs should take care to tag their routes with MEDs corresponding to a weighted tag of 100-0.

- A missing MED or MED of 0 indicates that the appliance should receive all servable traffic for the associated prefixes (if multiple appliances receive the same prefix with the same metric traffic will be load-balanced across those appliances)
- MEDs greater than 100 are normalized to 100
- MEDs greater than zero should be considered “tiers”, and we will select tiers in a weighted manner, e.g.:
  - For prefix 10.10.10.0/24
    - Appliances 1 & 2 have a weight of 10
    - Appliances 3 & 4 have a weight of 20
    - Appliances 5 & 6 have a weight of 30

- For prefix 20.20.20.0/24
  - Appliances 1 & 2 have a weight of 30
  - Appliances 3 & 4 have a weight of 20
  - Appliances 5 & 6 have a weight of 10
- Traffic for 10.10.10.0/24 will normally flow to appliances 1 & 2 unless there is a failure or lack of capacity, and then will bleed to appliances 3 & 4 and potentially 5 & 6
- Traffic for 20.20.20.0/24 will normally flow to appliances 5 & 6 unless there is a failure or lack of capacity, and then will bleed to appliances 3 & 4 and potentially 1 & 2

## AS Path Length

In the event that a prefix is received from multiple sources, it will be treated as load-balanced so long as the AS path is the same. In the event that AS paths are different, the shorter AS path will win.

## Regional Routing

ISPs should also take care to only send routes to appliances for regions that they wish to serve with that specific appliance. Netflix will not attempt to geographically map an ISP's network topology. In the event that an ISP desires that an appliance serve as a backup-only node for a set of prefixes (geographic failover), they should tag the failover prefixes with a MED of 100.

## Health Checking

Once the session has been established, Netflix will use the session state as one of their health metrics to ensure that the appliance is still active. In the event that a session is terminated, the Netflix traffic management system will no longer send new sessions to the appliance and the Netflix and customer Network Operations Centers will be notified. Existing sessions will remain with the appliance until they complete unless the appliance software or hardware is shut down.

## Example BGP Session Configuration - Cisco

```
!  
router BGP <ASN>  
neighbor <IP address of Appliance> remote-as 40027  
neighbor <IP address of Appliance> route-map NETFLIX-OUT out  
!  
route-map NETFLIX-OUT permit 10  
match ip address prefix-list NETFLIX-LOCAL  
set metric 0  
!  
route-map NETFLIX-OUT permit 20  
match ip address prefix-list NETFLIX-BACKUP  
set metric 100  
!  
route-map NETFLIX-OUT deny 30  
!
```

## Juniper

```
neighbor <IP address of Appliance> {  
    export NETFLIX-EXPORT;  
    peer-as 40027;  
}  
  
policy-statement NETFLIX-EXPORT {  
    term NETFLIX-LOCAL {  
        from {  
            prefix-list NETFLIX-LOCAL;  
        }  
        then {  
            metric 0;  
            accept;  
        }  
    }  
    term NETFLIX-BACKUP {  
        from {  
            prefix-list NETFLIX-BACKUP;  
        }  
        then {  
            metric 100;  
            accept;  
        }  
    }  
    then reject;  
}
```