# Scaling Internet TV Content Delivery

ALEX GUTARIN
DIRECTOR OF ENGINEERING, NETFLIX

NETFLIX


#netflixeverywhere
• Inventing Internet TV
• Available in more than 190 countries
• 104+ million subscribers

# Lots of Streaming == Lots of Traffic

- **Streaming 125+ million hours per day of TV shows and movies**
- **Responsible for over one-third of North American peak internet traffic**

# Architecture

DEVICE RUNNING
A NETFLIX CLIENT APP


VIDEO STREAM

DATA PLANE

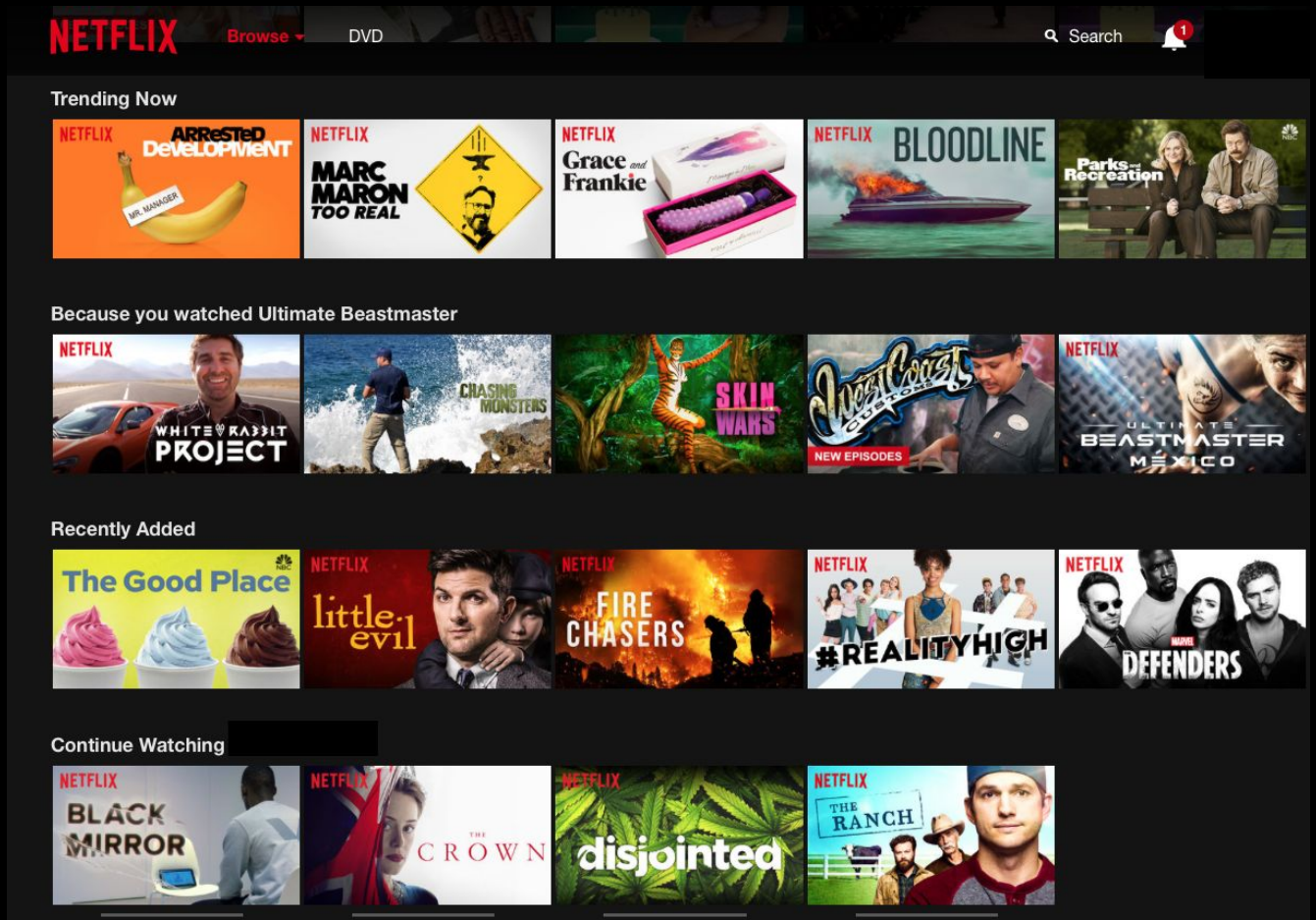
CONTROL PLANE

WEB SERVICES

WEB SERVICES

amazon
web services

# before streaming starts = control plane = AWS




streaming = data plane =
NETFLIX
OPEN CONNECT

DEVICE RUNNING
A NETFLIX CLIENT APP


VIDEO STREAM

NETFLIX
OPEN CONNECT

DATA PLANE

CONTROL PLANE

WEB SERVICES

WEB SERVICES

amazon
web services

# Open Connect is a Content Delivery Network (CDN)

- A set of content delivery servers (appliances)
  - Geographically distributed
  - Attached to or embedded within ISP networks
- A way of routing requests to the optimal server

We provide free Open Connect Appliances (OCA's) for Internet Service Provider (ISP) data centers.

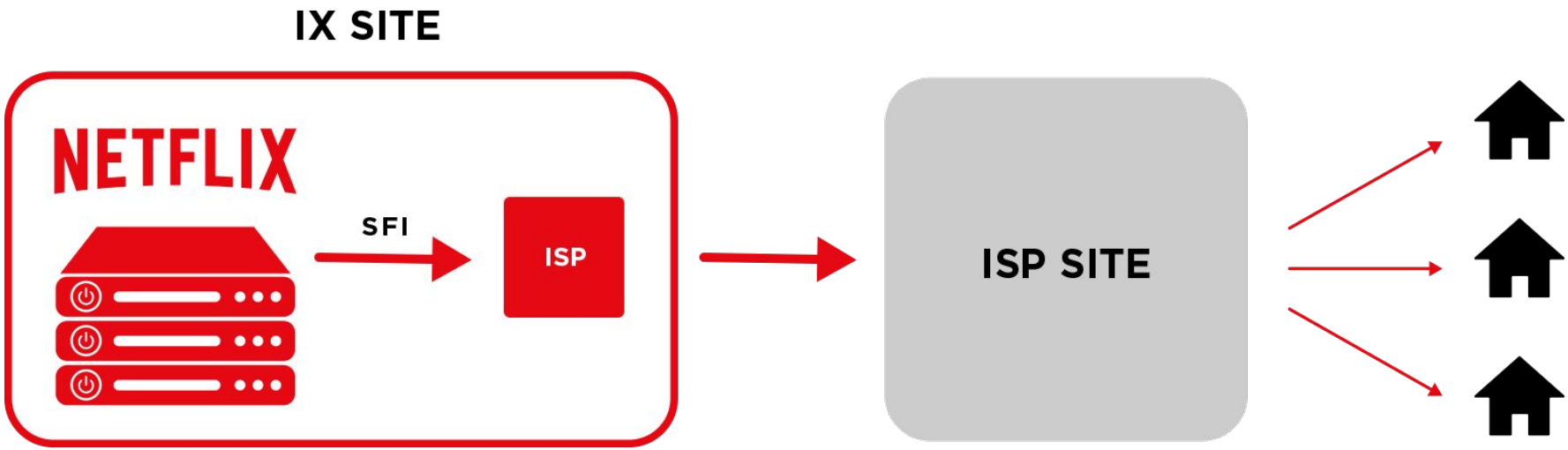
We also provide settlement free interconnection to ISPs at Internet Exchange (IX) locations throughout the world.

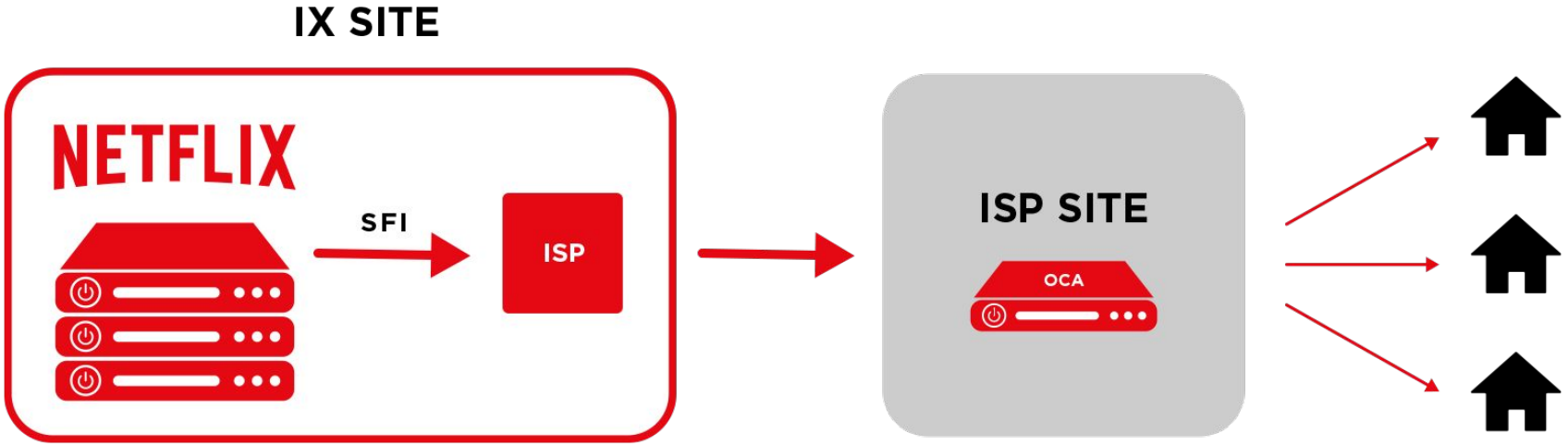
# Wait, what's an IX?

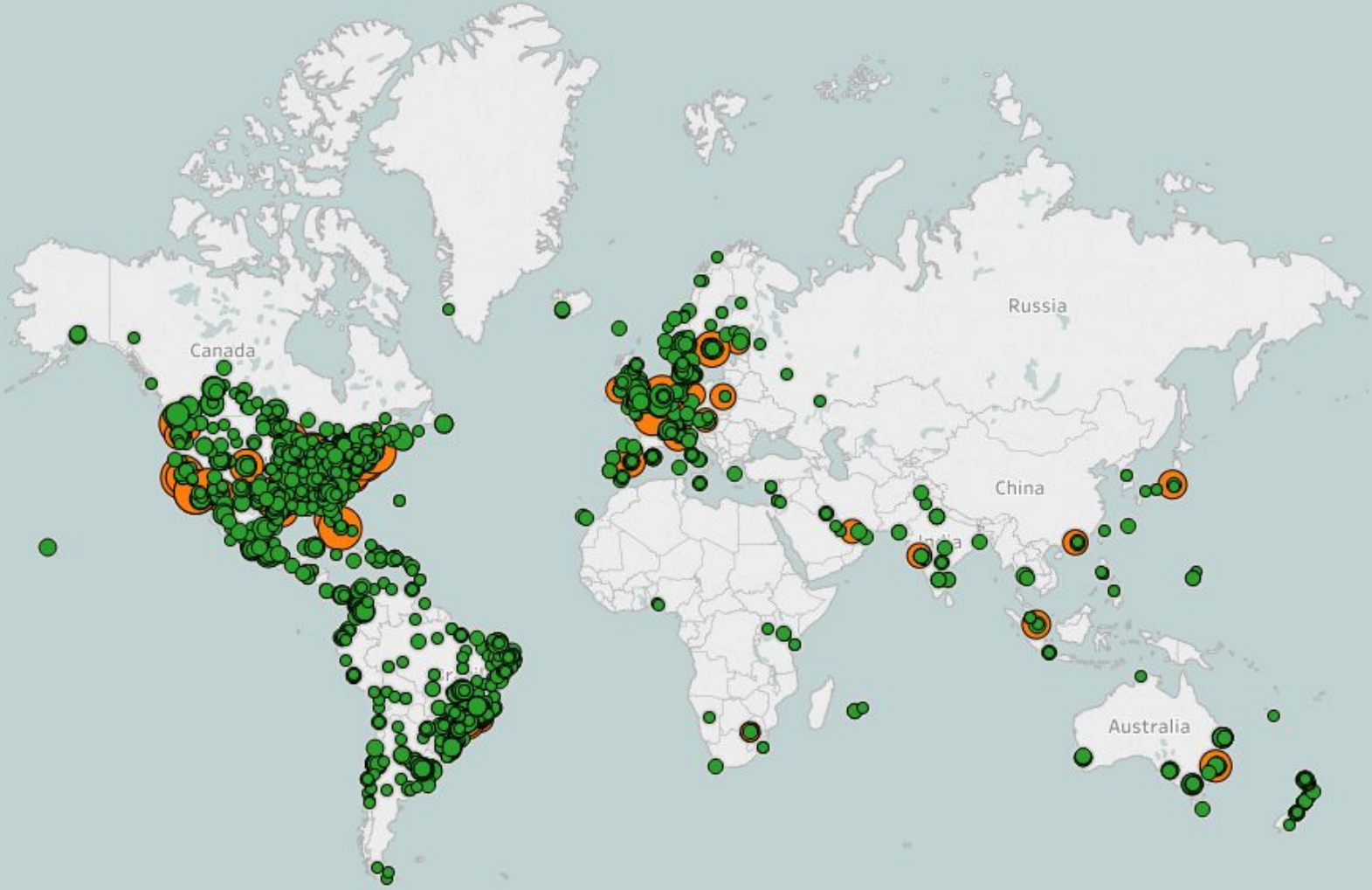
An Internet exchange point (IX or IXP) is a physical infrastructure through which Internet Service Providers (ISPs) and Content Delivery Networks (CDNs) exchange Internet traffic between their networks

*-Wikipedia*

"LINX" - London Internet Exchange


IX SITE
NETFLIX
SFI
ISP
ISP SITE


IX SITE
NETFLIX
SFI
ISP
ISP SITE
OCA


Canada
Russia
China
Australia

# Control Plane

1. Where (on which OCAs) to pre-position the content?

2. Where to steer (which OCAs to hand out to) the client for streaming the content?

# Offload

$$\frac{\textit{bytes served from ideal location}}{\textit{total bytes served}}$$

# Pre-positioning the Content

**For every file, figure out a list of OCAs it should be deployed to:**

- **Predict popularity**
- **Maximize (predicted) offload**
- **Minimize OCA requirements**

# Steering Clients

For every playback session,
return a set of OCAs based on:

- Content availability
- Network proximity
- OCA capacity to serve (health)

# Control Plane: Interesting Problems

NETFLIX

# Predicting Popularity

## Inputs

- Observed popularity
- View hour predictions for new titles

## Goals

- Maximize offload
- Minimize churn (number of bytes that need to be copied as popularity changes)

# Content Sharding

## Inputs

- Cluster of heterogeneous OCAs
  - Very different storage and throughput capabilities

## Goals

- Maximize offload (with fault tolerance)
- Maximize load balance

# Tiered Content Fill

## Inputs

- Network topology

## Goals

- Minimize number of bytes copied over narrow / expensive links
- Minimize deployment latency

# Scaling Issues

## Inputs

- Up to hundreds of thousands of routes per network location
- Hundreds of thousands of files per OCA
- Thousands of OCAs and locations

## Goals

- Perform steering quickly and efficiently

# Content Delivery Appliances

# What is an OCA?

- Massive web server
- Serving static content
- HTTP GET range requests
- (Mostly) TLS connections

# Two Types of Appliances:

## “Storage”

- Lots of storage, medium throughput

## “Flash” (aka “offload”)

- Limited storage, high throughput
- Used to serve popular content where traffic density allows for it

# “Storage” OCA

16 core Intel Xeon E5 CPU, 128 GB RAM

10 x 1 TB SSD; 24 x 10 TB HDD

4 x 10GbE links

# “Flash” OCA

16 core Intel Xeon E5 CPU, 128 GB RAM

4 x 4 TB NVMe flash storage

1 x 100GbE links

# OSS Stack


- FreeBSD OS
- NGINX web server
- BIRD BGP daemon
- Netflix apps to interface with the control plane and fill content

## Flash OCAs

- 90 Gbps of TLS encrypted traffic
  - Memory bandwidth limited
- 100K TCP connections
- 30-40K RPS
- Most of the traffic not served from RAM

## Storage OCAs

- 20 - 30 Gbps (disk IO limited)

# OCA:
# Interesting Problems

NETFLIX

# CPU & Memory Efficiency

- Optimize
  - OS kernel
  - Network adapter drivers
  - Web server
- Remove bottlenecks that prevent us from achieving higher throughput
- Reduce CPU requirements (cost)

# IO Efficiency

- Get more throughput from HDDs
  - Reduce IOPS and increase read sizes via intelligent pre-fetching
- Improve RAM and SSD caching of content

# Network Transport Algorithms

- Improve TCP on the sender side
  - Loss detection and recovery
  - Congestion control
- UDP-based transport deployable to clients
- Goal is achieve higher goodput to clients and improve customer quality of experience

# Open Connect Team 2016

http://techblog.netflix.com

# Thank you.
# Questions?

NETFLIX