University of Glasgow

Networks and Internet Routing

Networked Systems (H)

Lecture 9

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow | This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nd/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Lecture Outline

- Networks and Internet Routing:
  - Content distribution networks
  - Inter-domain routing
  - Routing security
  - Intra-domain routing
- Wrap-up

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

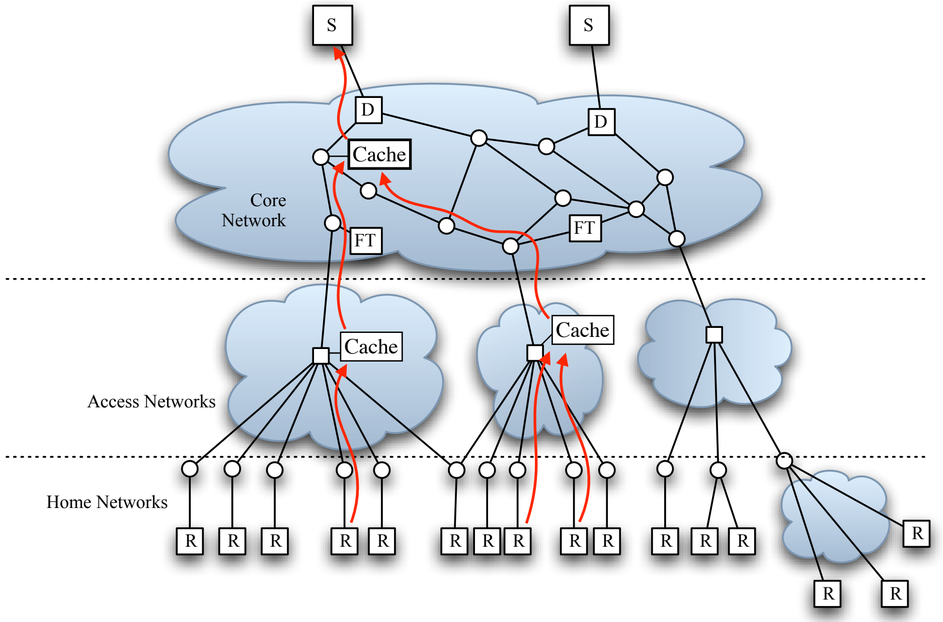
Content Distribution Networks (CDNs)

- Load balancing and latency reduction
- Implementation using DNS
- Implementation using anycast routing

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Content Distribution Networks (CDNs)

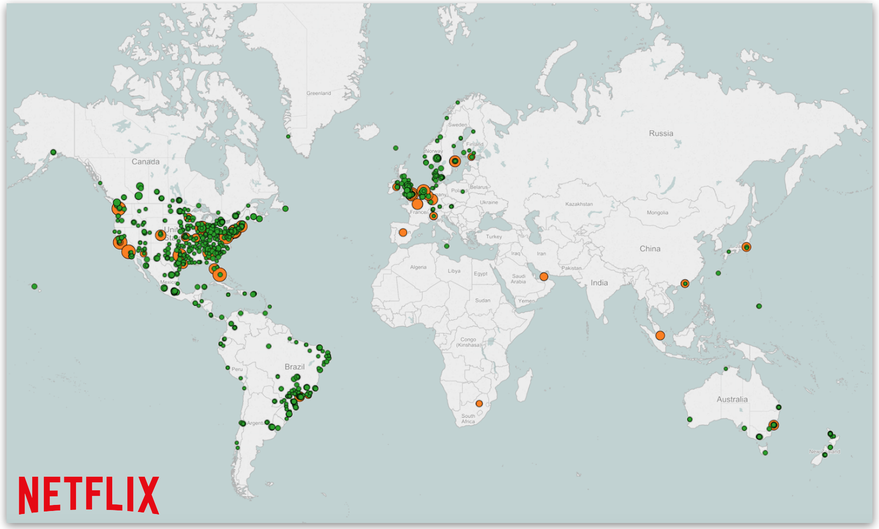
- **Content distribution networks (CDNs)** provide scalability, load balancing, and low latency
  - Host content for their customers in web caches spread around the world in edge networks and data centres
  - Reduces load on the main servers – rather than keep a local copy of the files, they link to the CDN-hosted copy
  - Reduces latency to respond to requests
  - Reduces chances of successful denial-of-service attack
- Many commercial CDNs available
- Many large organisations run their own
  - **Hypergiants** such as Google, Facebook, Netflix, …



Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Using CDNs to Distribute Load

- CDNs distribute load by locating web caches around the world and answering most requests from a local cache
- Needs extensive investment, large-scale cooperation with ISPs, Internet Exchange Points, etc.
- Mutual benefit for an ISP to host the servers for a CDN
  - Reduces the load on the ISP's network
    - One copy of the content sent to cache; the cache then distributes many times
    - Avoids overload of link from ISP to outside world
    - e.g., Netflix distribute "tens of terabits per second" of video – not possible from a single data centre; needs local fanout
  - Increases the reach and robustness of the CDN



Netflix Open Connect CDN locations (source: https://media.netflix.com/en/company-blog/how-netflix-works-with-isps-around-the-globe-to-deliver-a-great-viewing-experience)

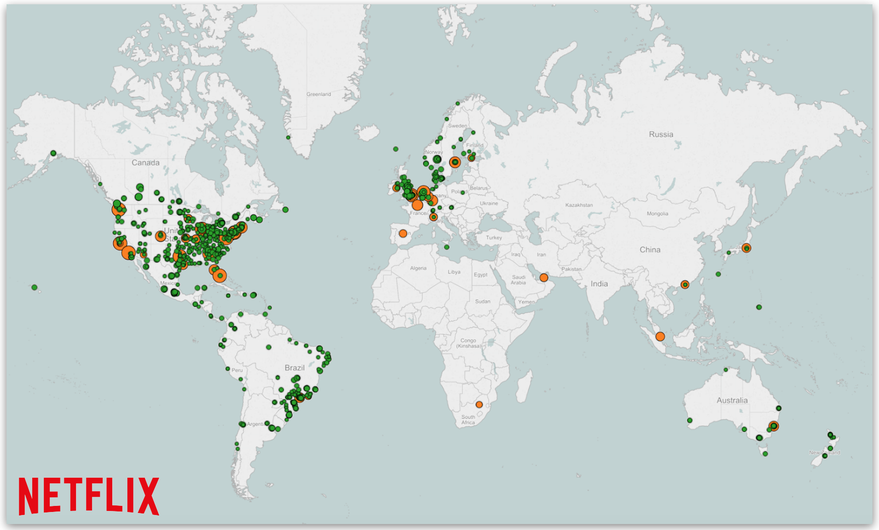
Akamai

"more than 240,000 servers in over 130 countries"

CC BY ND

Using CDNs to reduce latency

- Key CDN benefit is reducing latency → content is delivered from nearby caches
  - e.g., requests to a US-based service are answered from a CDN cache in Europe, rather than having to cross long-distance path to the US
- Requires global distribution of CDN proxy caches
  - Are CDNs effectively serving the entire world?
  - Providing effective Internet access to developing regions requires more than just connectivity; also needs data centres and infrastructure to host the CDN nodes



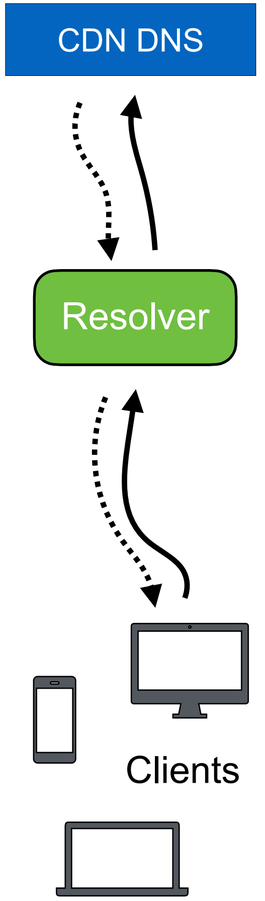
Netflix Open Connect CDN locations (source: https://media.netflix.com/en/company-blog/how-netflix-works-with-isps-around-the-globe-to-deliver-a-great-viewing-experience)

- Effective for cacheable static content – video, software updates, images – may need edge compute infrastructure to support other applications

CC BY ND

Locating the Nearest CDN Node using DNS

- Locate nearest CDN node using the DNS
- Each resource hosted by the CDN has a unique DNS name – each resource is given a different domain name
  - e.g., the CDN hosts **`https://example.com/images/kitten.jpg`** as **`https://9BC1C10B7947A890B9.cdn-example.com/kitten.jpg`**
- DNS server for the CDN returns different A or AAAA records for a name, depending on where it's requested from, what CDN caches have the data
  - Directs to local cache based on IP address of resolver making the query
    - DNS client subnet extension [RFC7871] for remote resolvers
  - Doesn't need to be particularly accurate – trying to figure out if you're in UK rather than US, to direct you to data centre in London rather than New York
  - Allows very fine-grained control, but puts high load on DNS



Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Locating the Nearest CDN Node using Anycast Routing

- Locate nearest CDN node using **anycast routing**
- Each resource hosted by the CDN has a unique filename
  - e.g., CDN hosts **`https://example.com/images/kitten.jpg`** as **`https://cache.cdn-example.com/9BC1C10B7947A890B9.jpg`**
  - The DNS name always maps to the same IP address – that IP address is a load balancer at the entrance to a data centre
- **The CDN uses multiple data centres – all using the same IP addresses**
  - Each data centre advertises its addresses via BGP
  - Inter-domain Internet routing will ensure traffic goes to the closest data centre to source



Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Content Distribution Networks (CDNs)

- Load balancing and latency reduction
- Implementation using DNS
- Implementation using anycast routing

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow