The transition to the datacenter-based Internet

Eduardo Grampín Castro
grampin@fing.edu.uy
Universidad de la República
Uruguay

based on slides by Geoff Huston [GHuston2018] & others

Outline

UNIVERSIDAD DE LA REPUBLICA URUGUAY

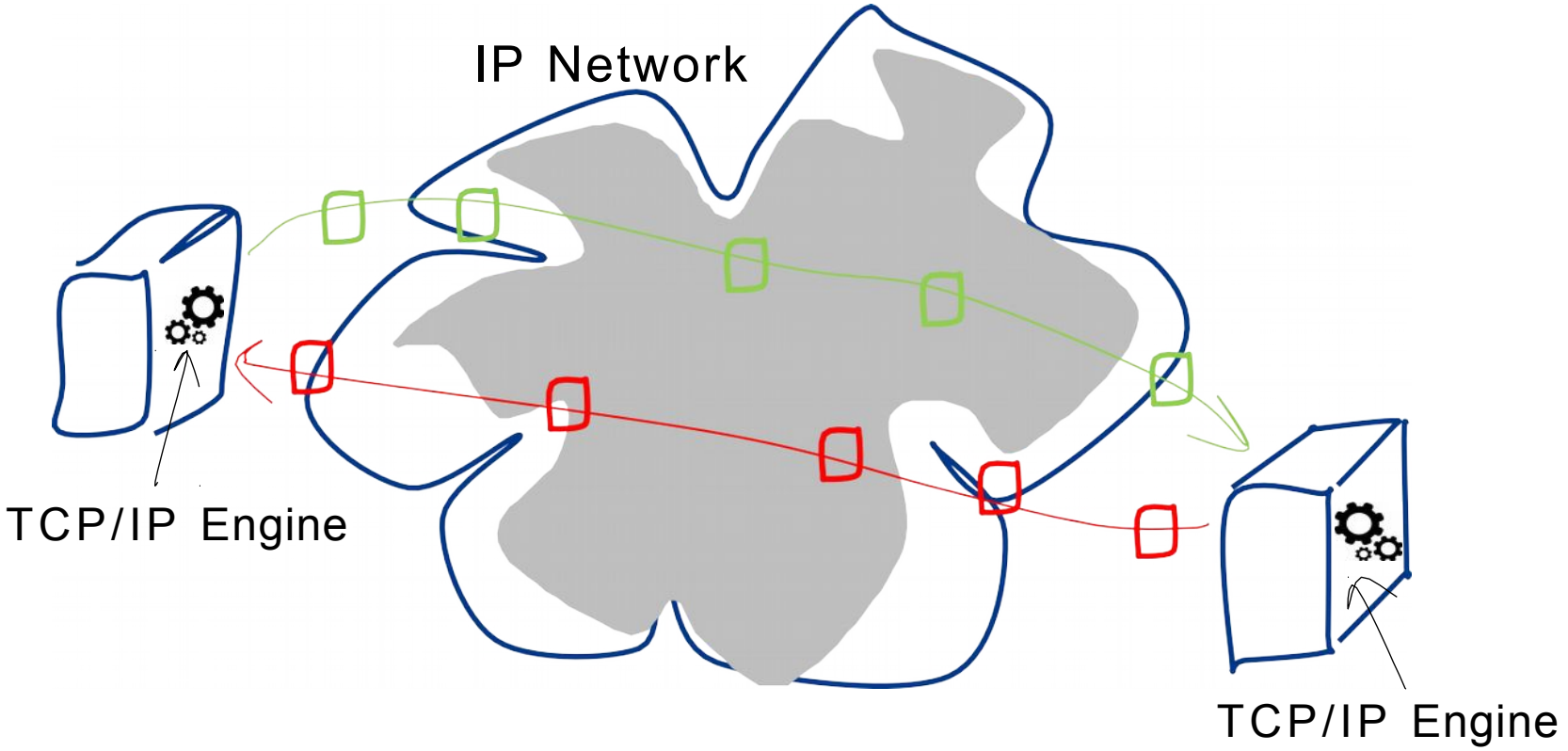
- Internet evolution
- The irruption of Content Providers (OTTs)
  - The end of transit/hierarchical architecture
- The datacenter-based internet
  - OTTs
  - Network Service Providers
    - Intorduction to SDN&NFV

Internet Architecture (1980's)

UNIVERSIDAD DE LA REPUBLICA URUGUAY

- "End-to-End" design:
  - Connected computer to computer
  - The network switching function was stateless
  - No virtual circuits, no dynamic state for packets to follow
  - Single network-wide addressing model
  - Single network-wide routing model
  - Simple datagram unreliable datagram delivery in each packet switching element
  - hop-by-hop destination-address-based packet forwarding paradigm

Internet Architecture (1980's)



The result was outstanding!

UNIVERSIDAD DE LA REPUBLICA URUGUAY

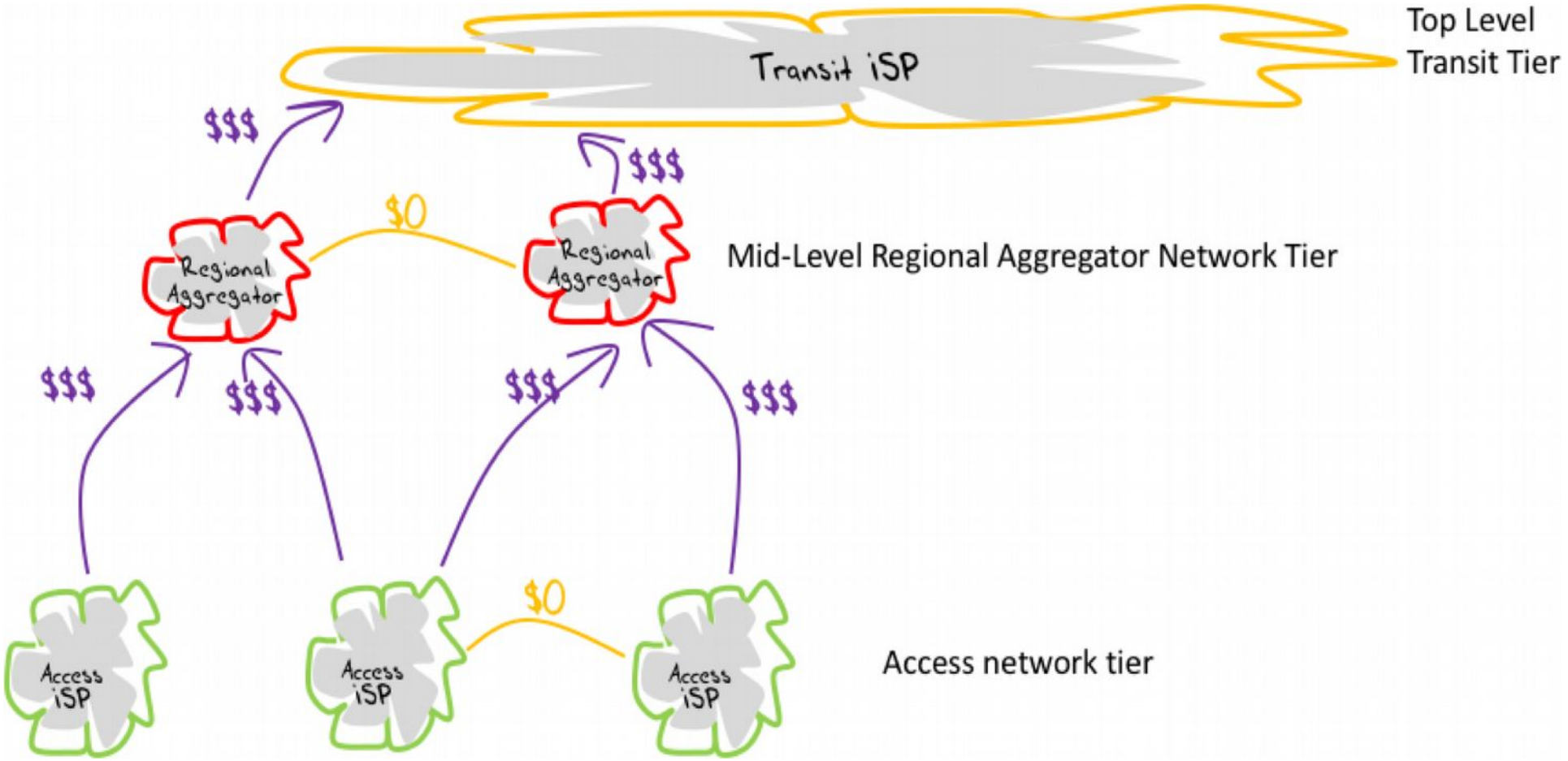
- By stripping out network-centric virtual circuit states and removing time synchronicity (as opposed to the telephone network), the resultant carriage network was minimal in design and functionality

- More complex functions, such as flow control, jitter stability, loss mitigation and reliability, were pushed out to the attached devices on the edge

No regulation...

- In the regulated world of national telephone operators every telephone network was “equal”
- But we rapidly started differentiating between Internet networks – Internet networks were not all the same
- We started differentiating on roles and services and differentiating by the flow of revenues between networks

Network role segmentation



Content services

UNIVERSIDAD DE LA REPUBLICA URUGUAY

- Breaking the edge into clients and servers
  - Access networks service the needs “clients”
  - Clients are not directly reachable by other clients
  - Clients connect to services

- The role of the network here is to carry clients to the service access point
  - The assumption here is that there are many more clients than service points

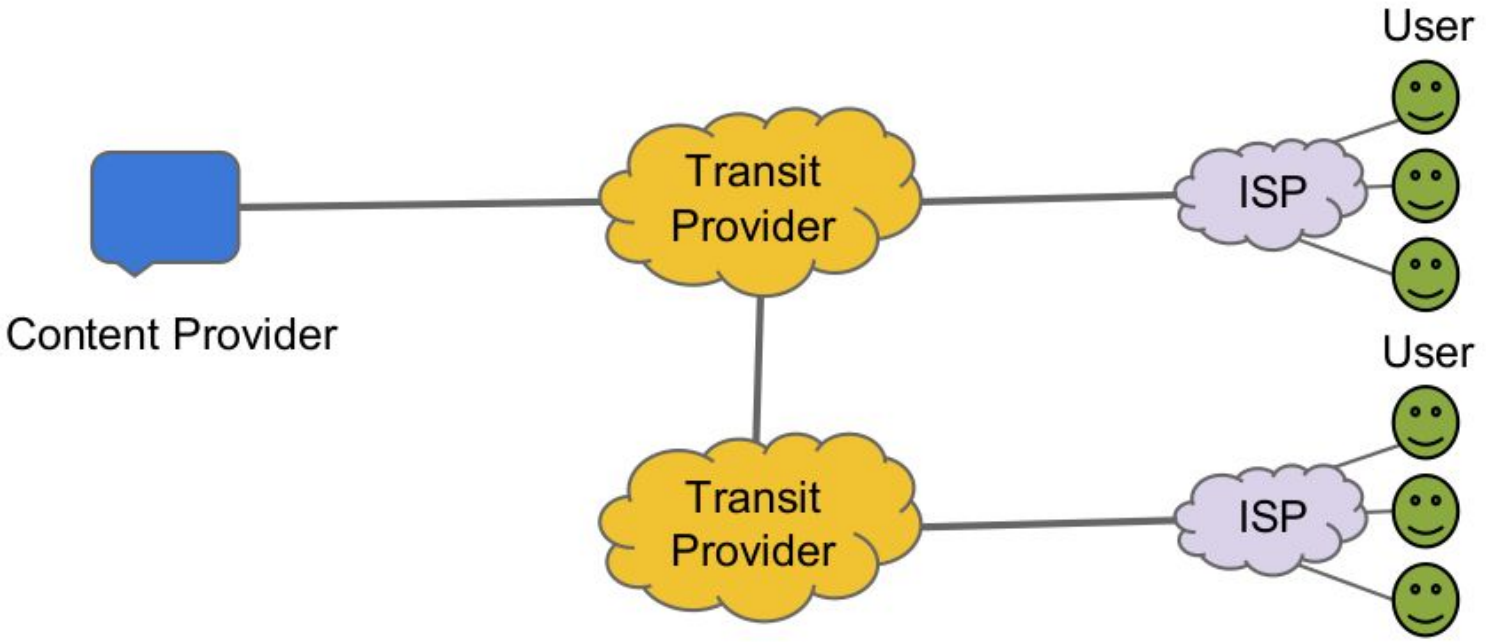
Content vs. carrier providers

UNIVERSIDAD DE LA REPUBLICA URUGUAY

- Who pays whom?
  - The only reason why access networks have clients is because there are content services that clients want to access
    - Therefore carriers should pay for content
  - There is no “end-to-end” financial settlement model in the Internet; both “ends” pay for access and network providers settle between themselves. To a carrier network, content is just another client
    - Content should pay the carrier for transport, just like any other client

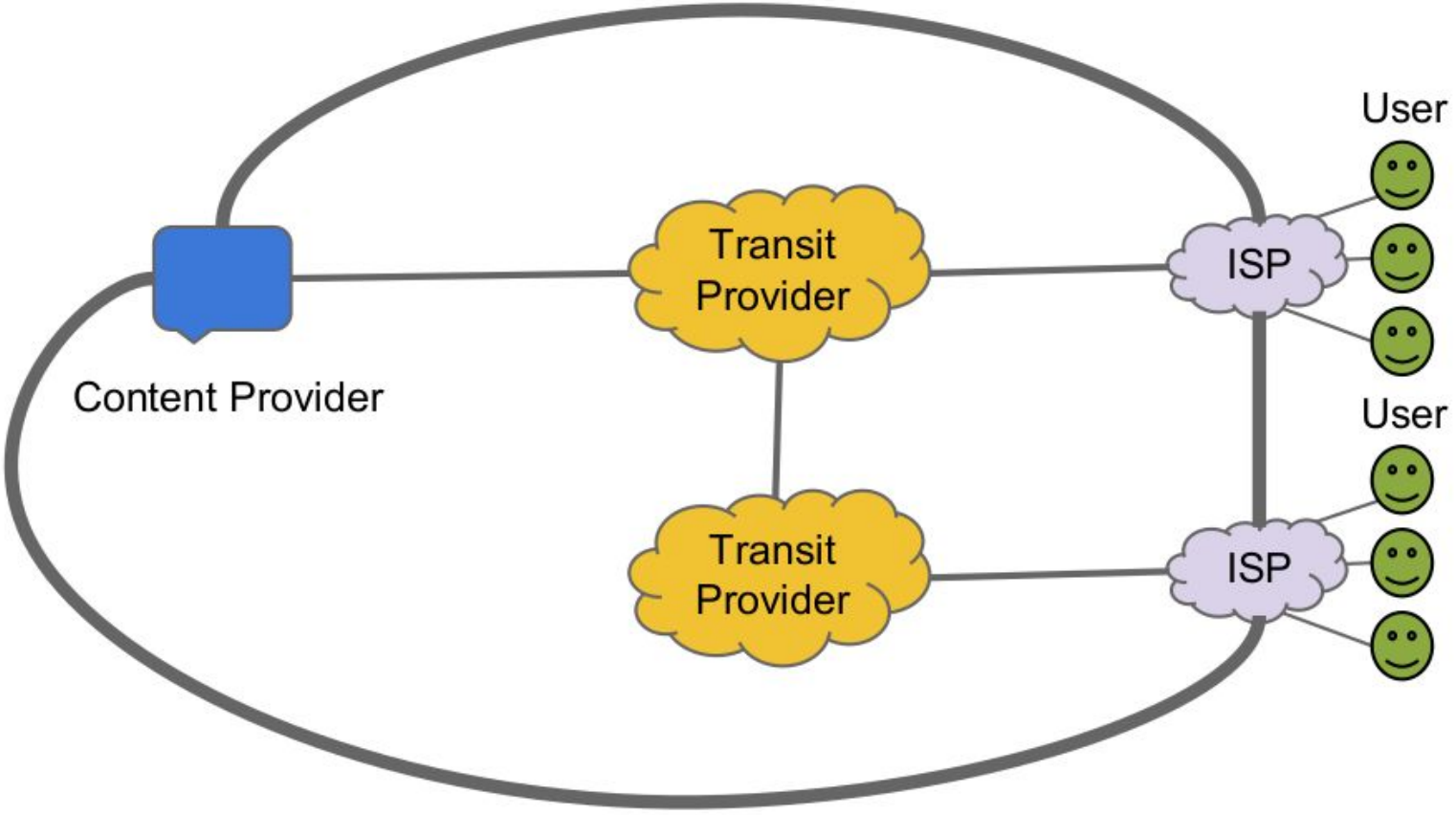
Content vs. carrier providers

- Content providers solved the problem establishing direct relationships with clients
  - From transit to peering



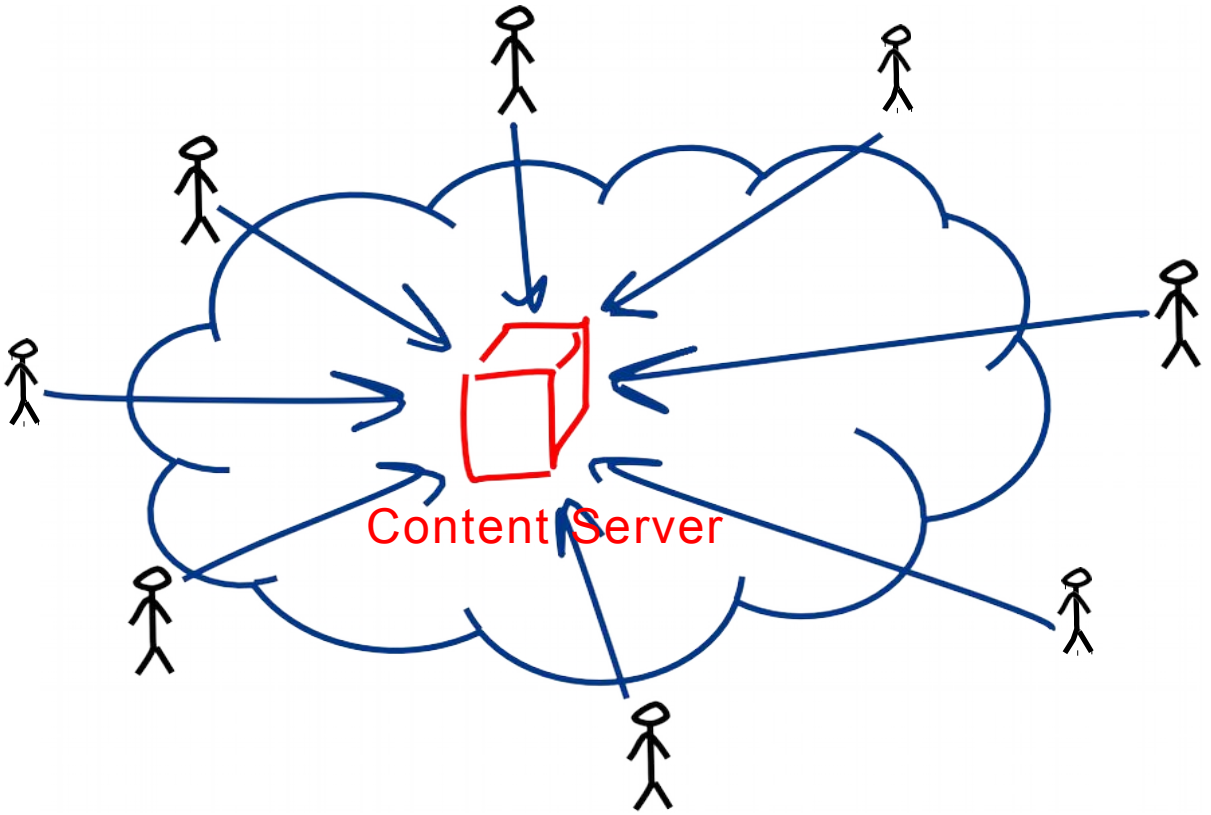
Content vs. carrier providers

- Content providers solved the problem establishing direct relationships with clients
  - From transit to peering

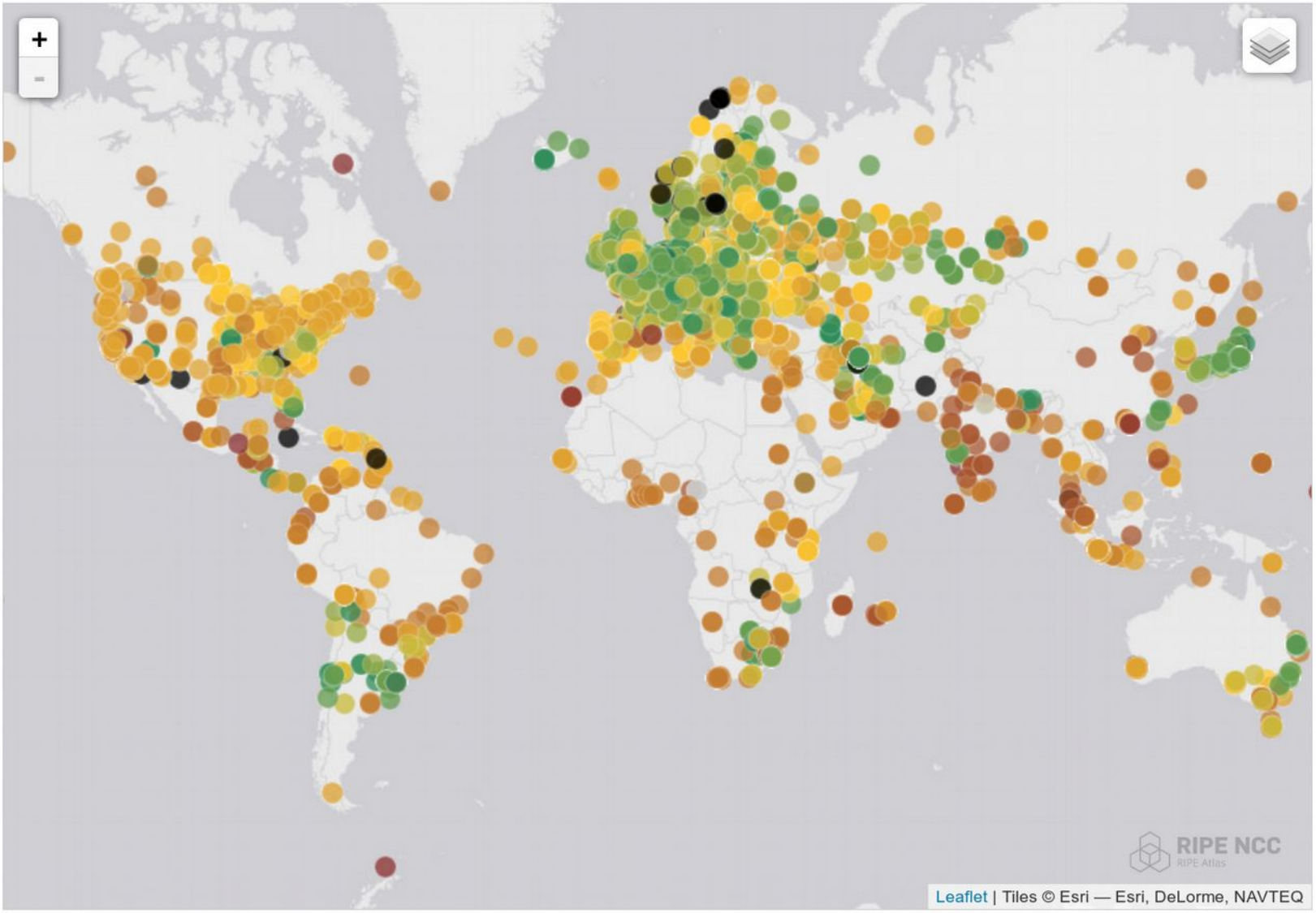


Content wins

- Therefore, the logical view is something like this:

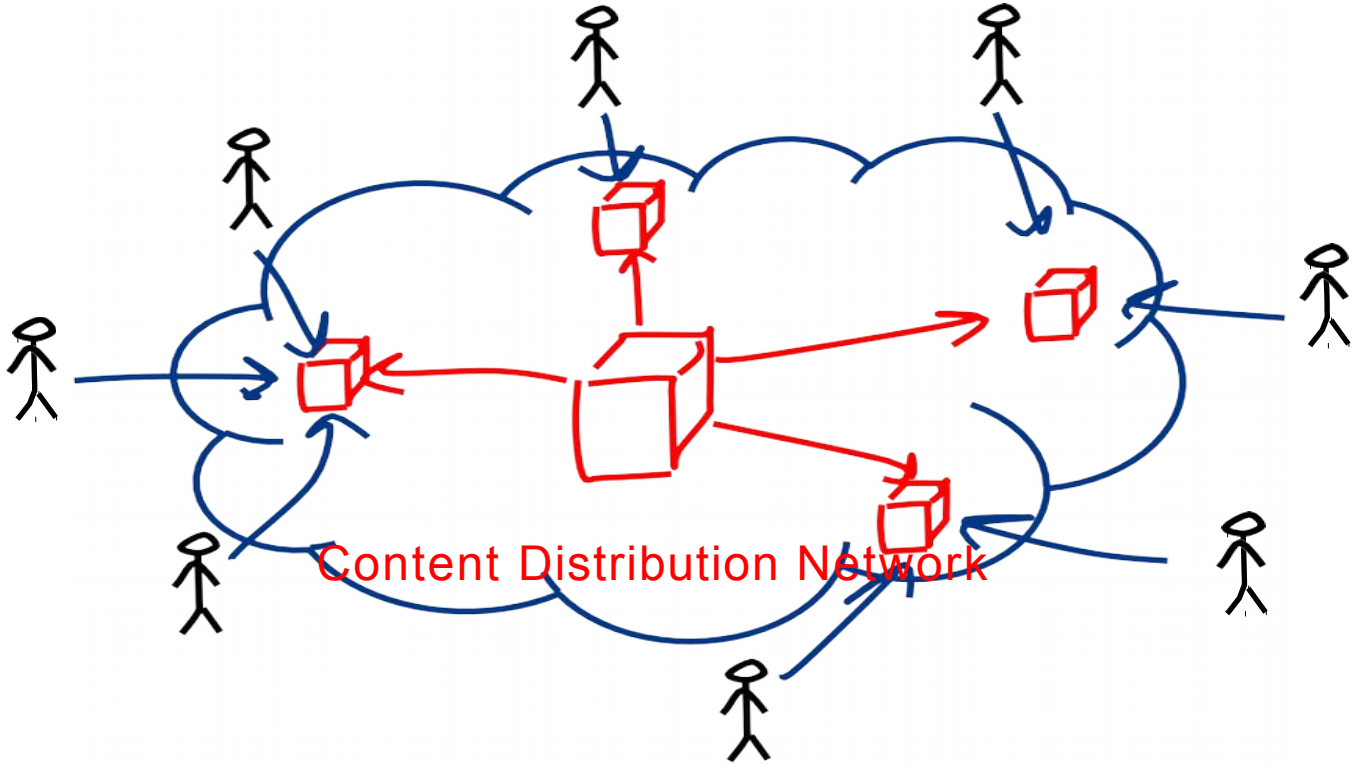


What about RTT?



https://atlas.ripe.net/results/maps/rtt-fixed/

Content Distribution



Content Distribution

UNIVERSIDAD DE LA REPUBLICA URUGUAY

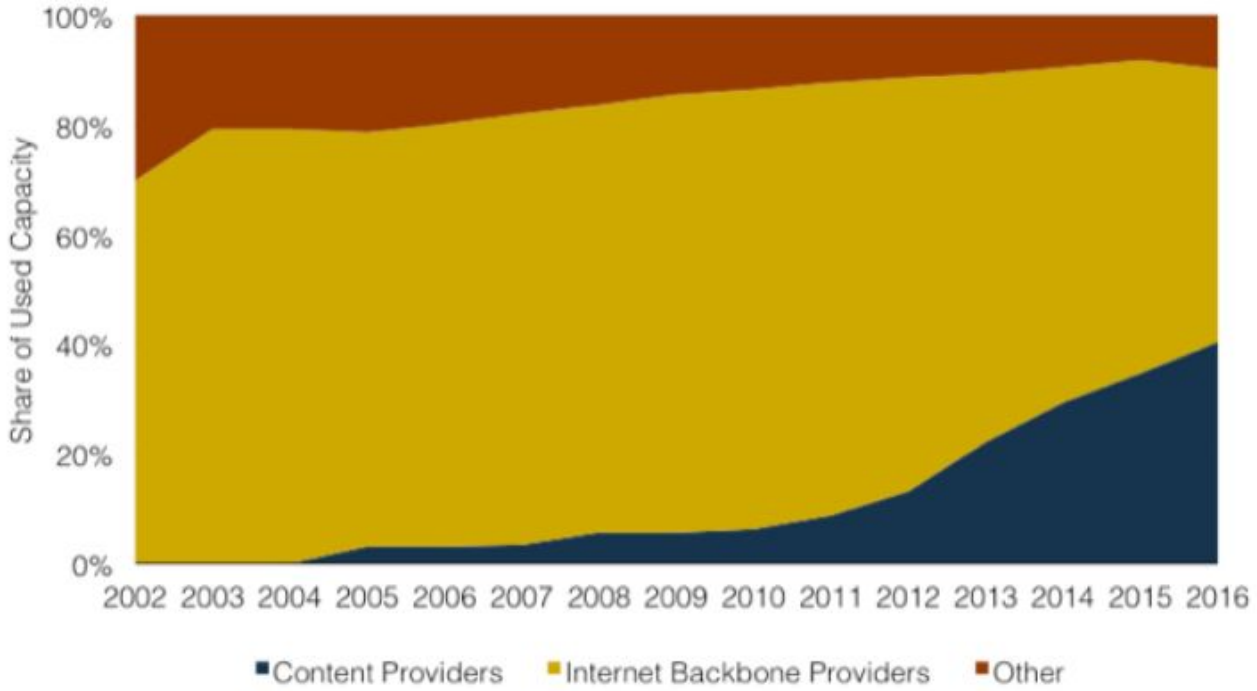
- The rise of the Content Distribution Network
  - Replicate content caches close to large user populations
  - The challenge of delivering many replicant service requests over high delay network paths is replaced by the task of updating a set of local caches by the content distribution system and then serving user service requests over the access network
  - Reduced service latency, increased service resilience, happy customers!

Content Distribution: role reversal

UNIVERSIDAD DE LA REPUBLICA URUGUAY

- Service portals are increasingly located adjacent to users and that means changes to the network:
  - Public Networks no longer carry users' traffic to/from service portals via ISP carriage services
  - Instead, Private Networks carry content to service portals via CDN services
- This shift has some profound implications for the Internet

Used International Bandwidth by Source, 2002-2017



"Unlike previous submarine cable construction booms, content providers like Amazon, Google, Facebook, and Microsoft are taking a more active role in this recent surge.

These companies alone have such incredible demand for data center traffic that they're driving projects and route prioritization for submarine cables, which is why we thought it made sense to compile a list of their submarine cable holdings."

https://blog.telegeography.com/telegeographys-content-providers-submarine-cable-holdings-list
https://blog.telegeography.com/content-providers-google-submarine-cables-bandwidth-market-junior-curie

Today's Internet Architecture

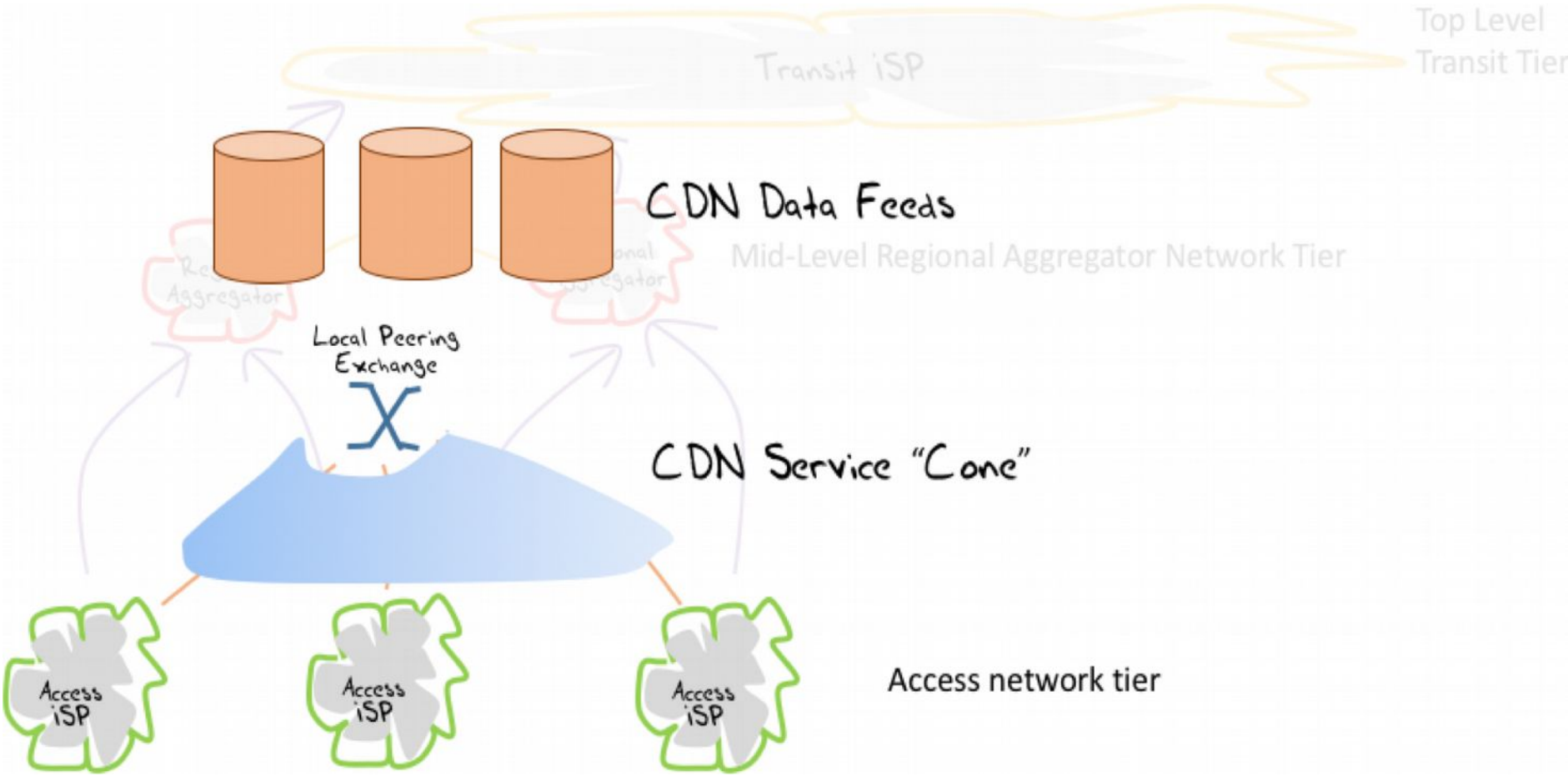
UNIVERSIDAD DE LA REPUBLICA URUGUAY

- We've split the network into clients and servers
  - Web servers
  - Streaming servers
  - Computing & Storage servers (Cloud Computing)
  - Mail servers
  - DNS servers

- Servers and services now sit in CDN bunkers with global replication and DDOS hardening

- Users don't reach out to content any more, the CDNs bring content to users

Today's Internet Architecture



Today's Internet Architecture

UNIVERSIDAD DE LA REPUBLICA URUGUAY

- If users don't send packets to users any more…
- If content is now delivered via CDNs to users via discrete service cones…
- If there is no universal service obligation…

Then why do we still need Transit Service providers?

- We see the CDN systems reserve a carriage resource through dedicated bandwidth / wavelength / cable purchase and effectively bypass the open IP carriage infrastructure

- Once the CDN caches sit "inside" the Access ISP then the entire wide area network becomes a marginal activity compared to the value of the content feeds!

Today's Internet Architecture

UNIVERSIDAD DE LA REPUBLICA URUGUAY

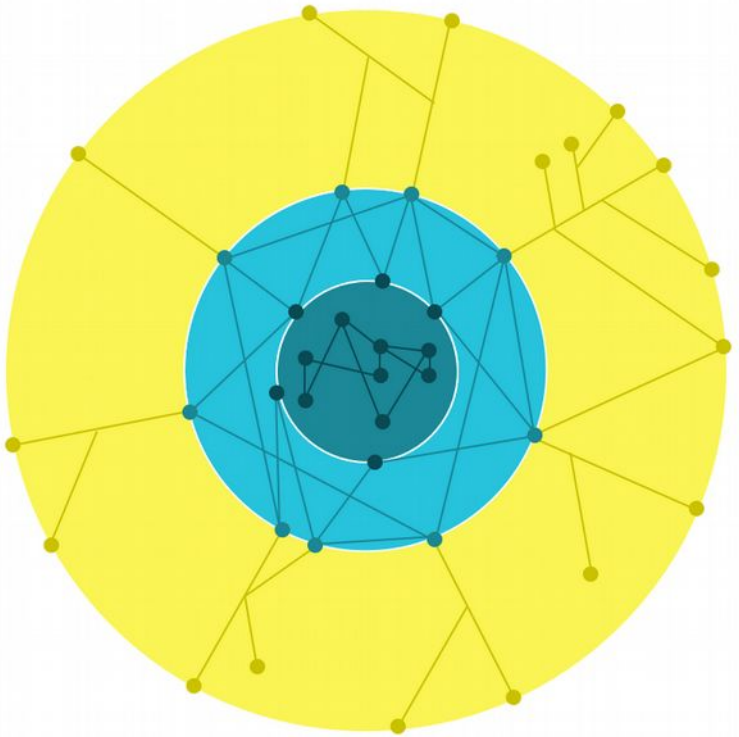
- Re-purposing of the entire network...
  - Service provisioning sits within cloud providers and distributed datacenters
  - Edge computers are now acting as televisions into the clouded world of data
    - No more peer to peer communication
  - The distinction between personal and public data realms is disappearing into the realm of corporately owned private data companies
  - The internal parts of the network are now being privatized and removed from public regulatory scrutiny

Content business

Company	$B USD
Apple	**818**
Alphabet	**727**
Microsoft	**659**
Amazon	**563**
Facebook	**513**
Tencent	**493**
Berkshire Hathaway	489
Alibaba Group	**440**
Johnson & Johnson	375
JPMorgan Chase	371

The world's 10 largest publicly traded companies, as ranked by their market capitalization, 4Q, 2017

An example: google infrastructure



Google aims to deliver its services with high performance, high reliability, and low latency for users, in a manner that respects open internet principles.

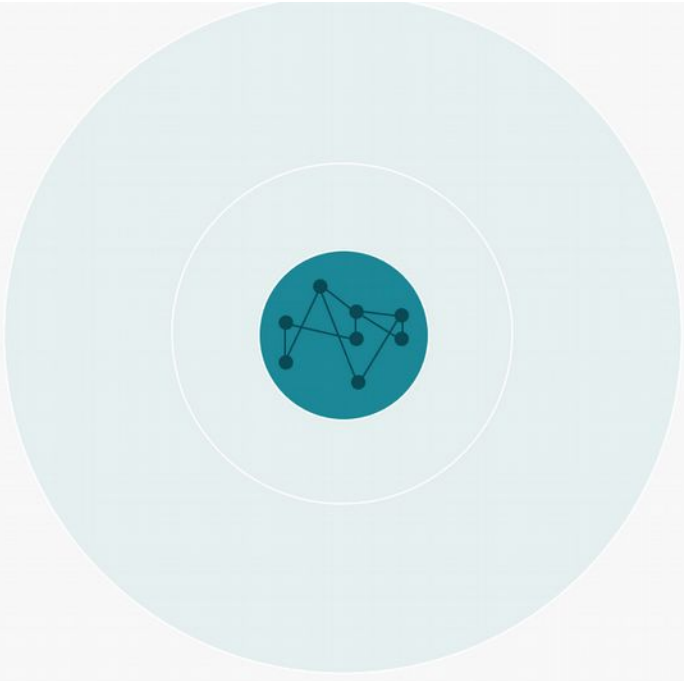
We have invested in network infrastructure that is aligned with this goal and that also allows us to work with network operators to exchange traffic efficiently and cost-effectively.

Google's network infrastructure has three distinct elements:

- Core data centers
- Edge Points of Presence (PoPs)
- Edge caching and services nodes (Google Global Cache, or GGC)

https://peering.google.com/#/infrastructure

An example: google infrastructure



Data centers

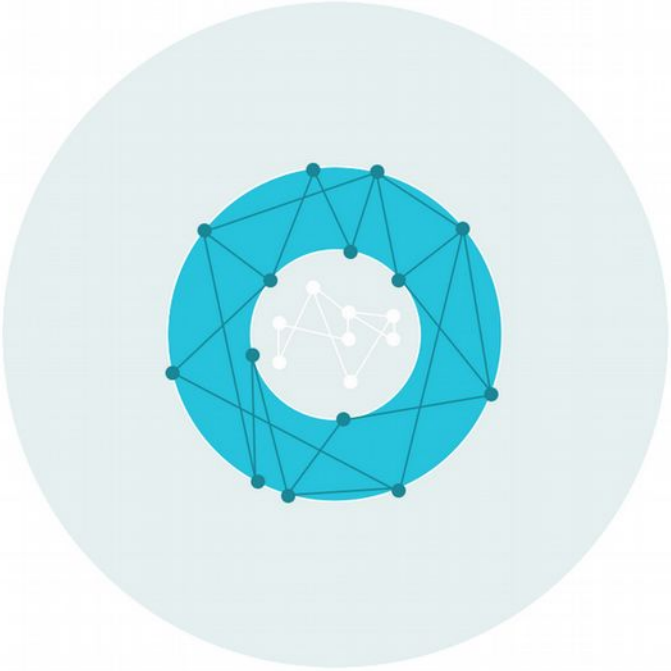
Google operates data centers that we use for computation and backend storage. We have data centers in the Americas, Europe and Asia.

Our data centers are the heart of Google content and services.

Google has built a large, specialized data network to link all of its data centers together so that content can be replicated across multiple sites for resilience, and services can be delivered closest to the end user.

- More information on our data centers.
- More information on our Google Cloud Platform regions.

An example: google infrastructure



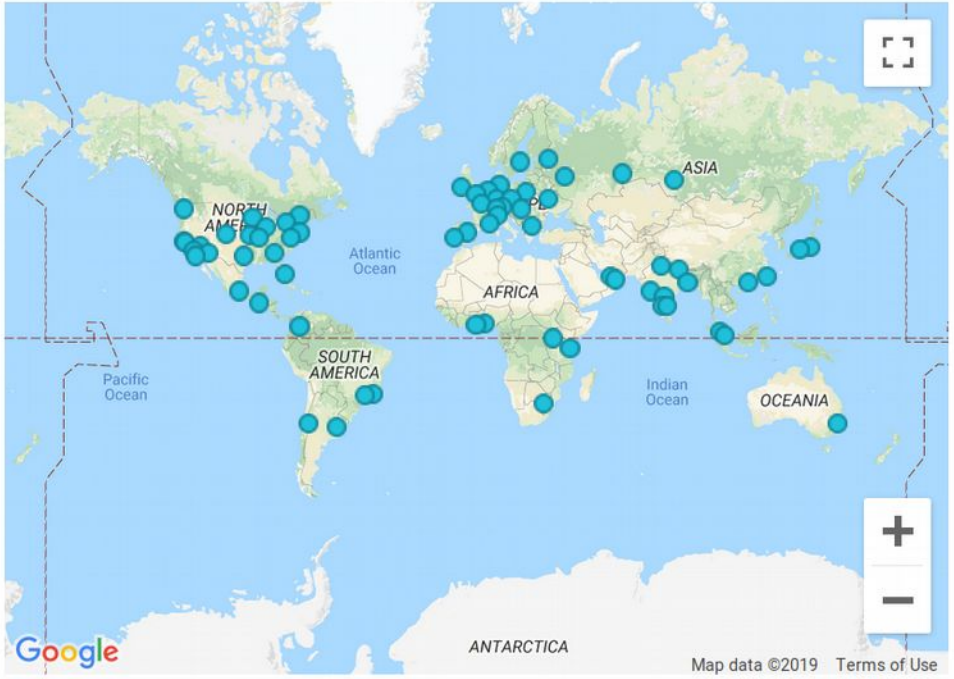
Edge Points of Presence (PoPs)

Our Edge Points of Presence (PoPs) are where we connect Google's network to the rest of the internet via peering. We are present on over 90 internet exchanges and at over 100 interconnection facilities around the world.

By operating an extensive global network of interconnection points we can bring Google traffic closer to our peers, thereby reducing their costs and providing users with a better experience.

Google operates a large, global meshed network that connects our Edge PoPs to our data centers.

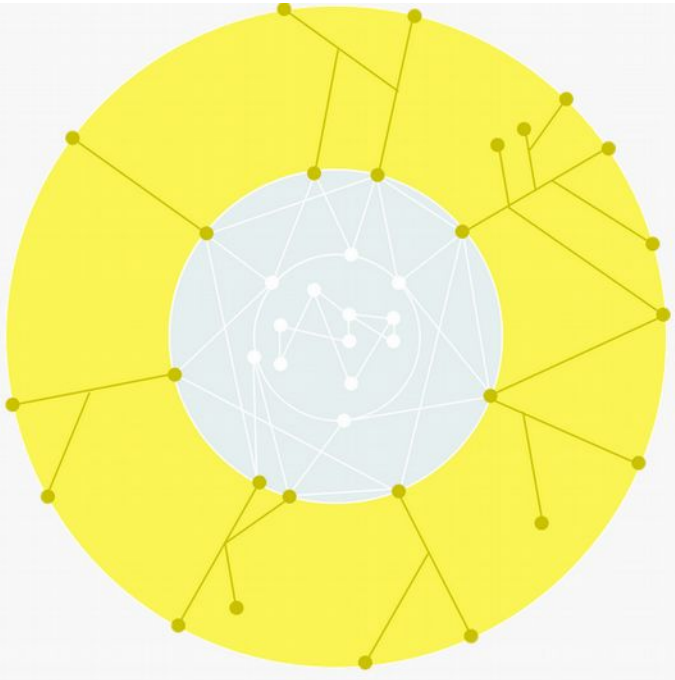
See our record in peeringdb.com



Map of metros where at least one Edge PoP is present.

An example: google infrastructure

UNIVERSIDAD DE LA REPUBLICA URUGUAY

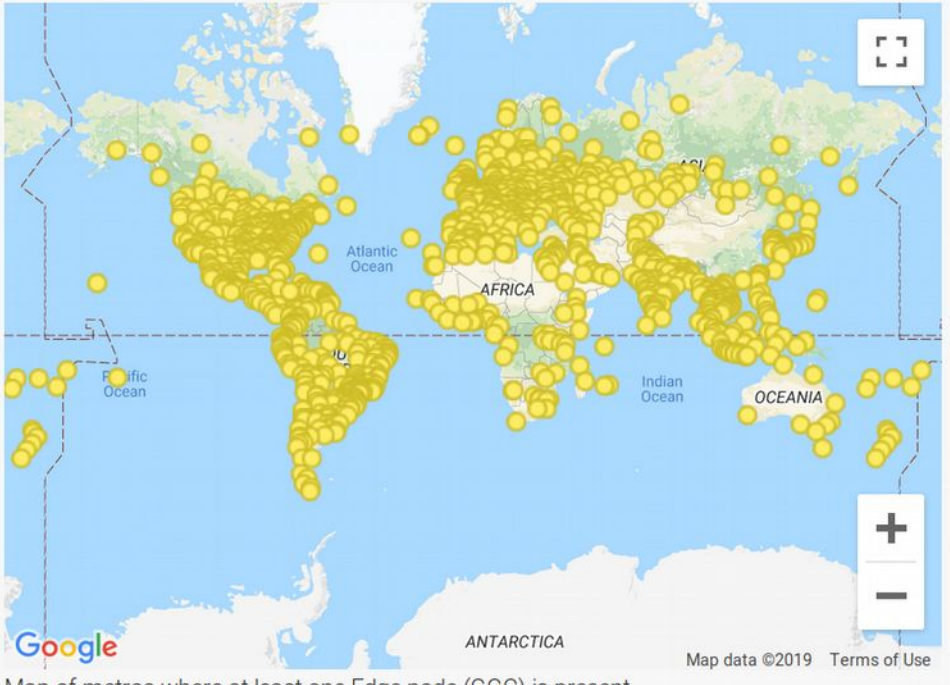


Edge nodes (Google Global Cache, or GGC)

Our edge nodes (called Google Global Cache, or GGC) represent the tier of Google's infrastructure closest to our users. With our edge nodes, network operators and internet service providers deploy Google-supplied servers inside their network.

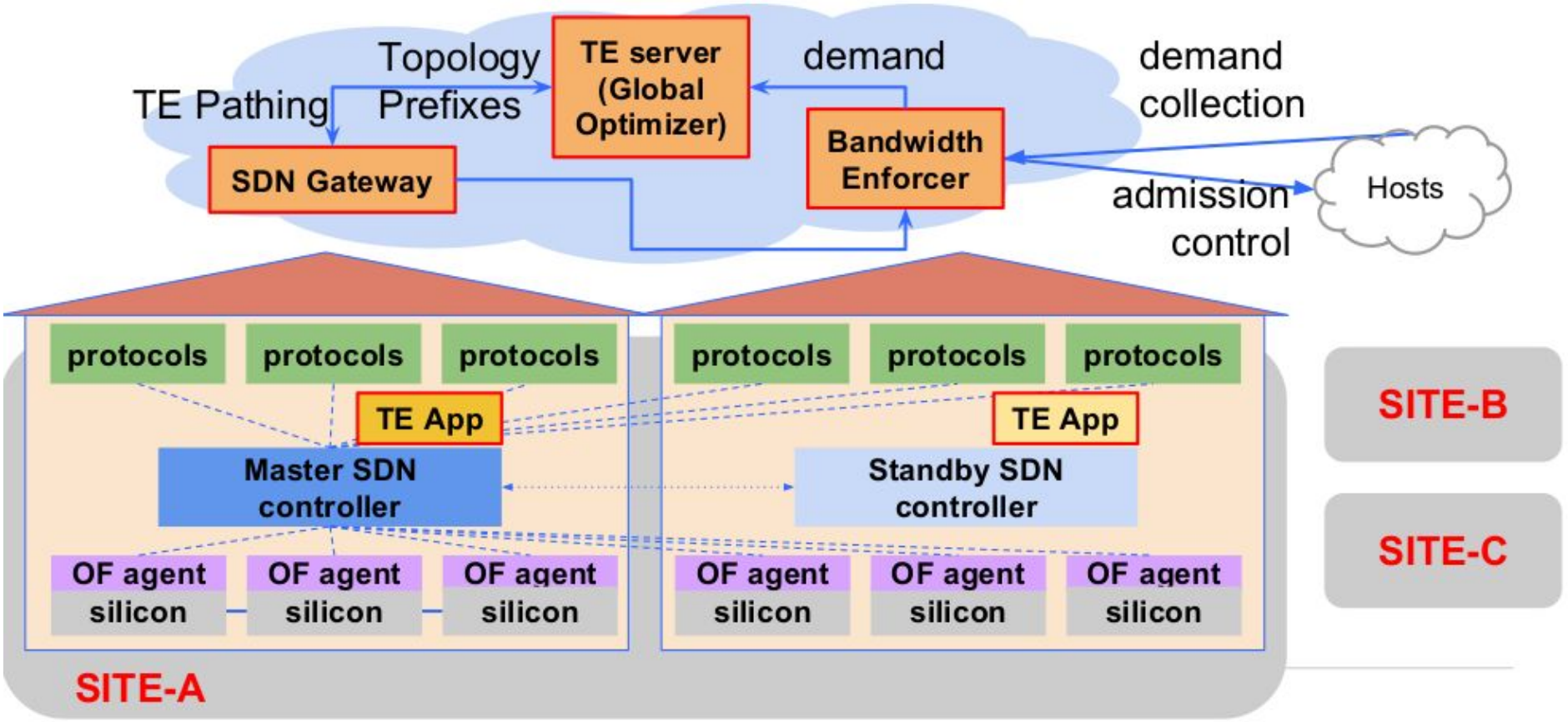
Static content that is very popular with the local host's user base, including YouTube and Google Play, is temporarily cached on edge nodes. Google's traffic management systems direct user requests to an edge node that will provide the best experience.

In some locations, we also use our edge nodes to support the delivery of other Google services, such as Google Search, by proxying traffic where it will deliver improved end-to-end performance for the end user.



Map of metros where at least one Edge node (GGC) is present.

An example: google infrastructure (back in 2013)



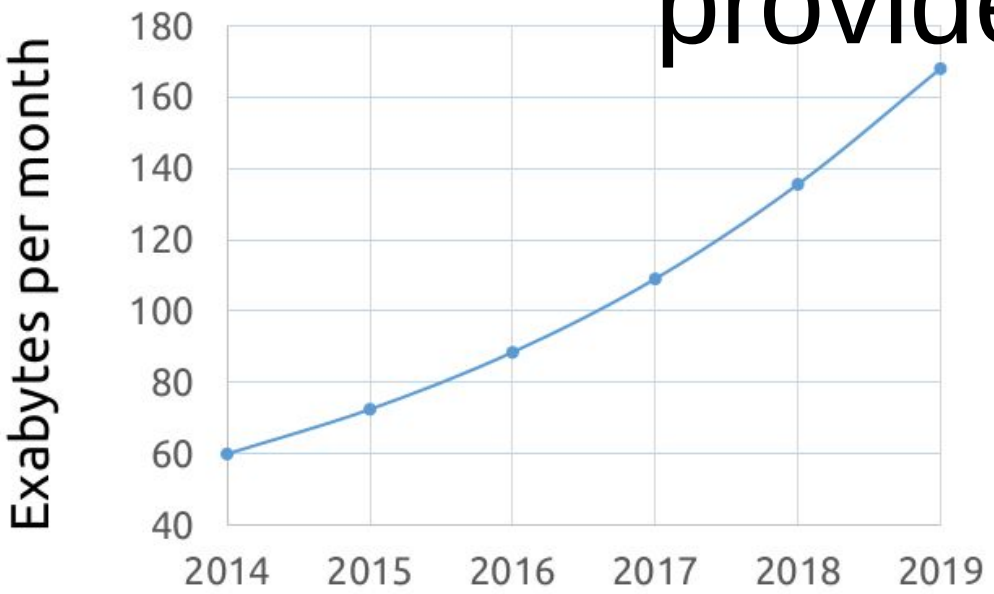
Results published in ACM SIGCOMM 2013 [GoogleB4]

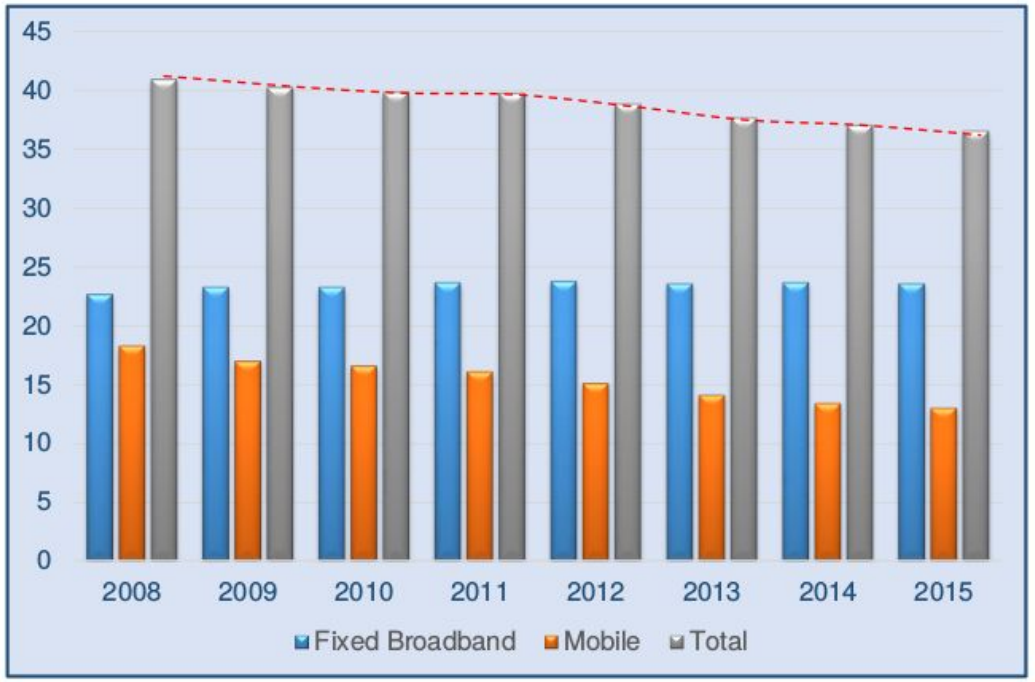
A datacenter-based Internet

UNIVERSIDAD DE LA REPUBLICA URUGUAY

- Data is the most important component
  - Networking follows data
- There is a whole world of “virtuality” inside a datacenter, for instance:
  - Virtual Machines (VMs)
  - **Virtual networks** to connect such VMs (inside and across datacenters)
- Which means the network must provide (at least):
  - (Fast) content replication
  - (Fast) Virtual Machine migration
- Many application are based on L2 (ethernet) connectivity
  - Therefore, **good virtual networking solutions for L2 are required**

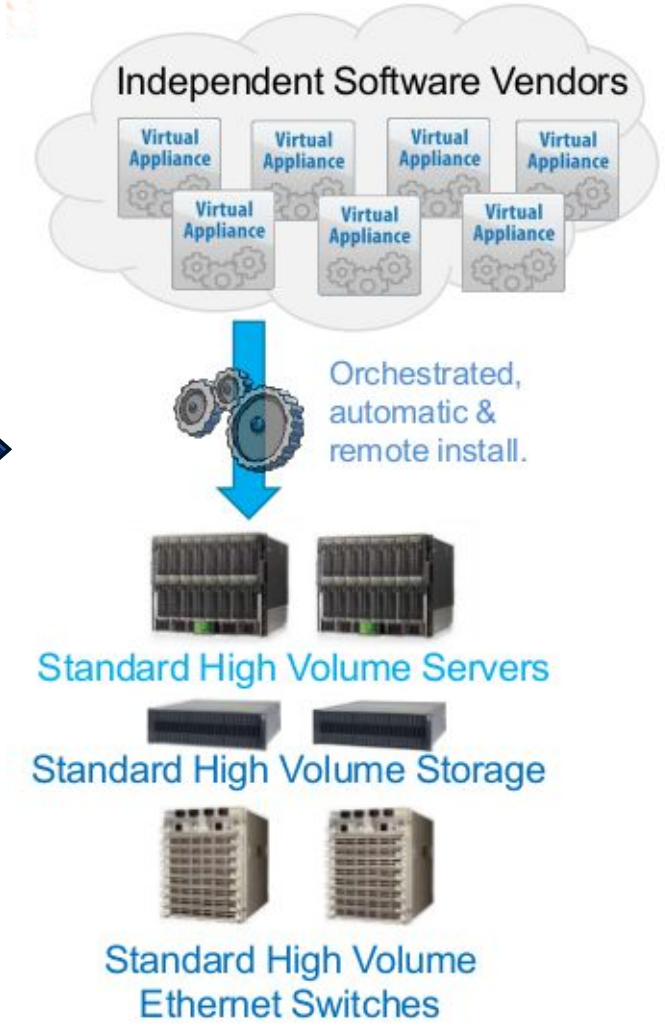
What about network service providers?





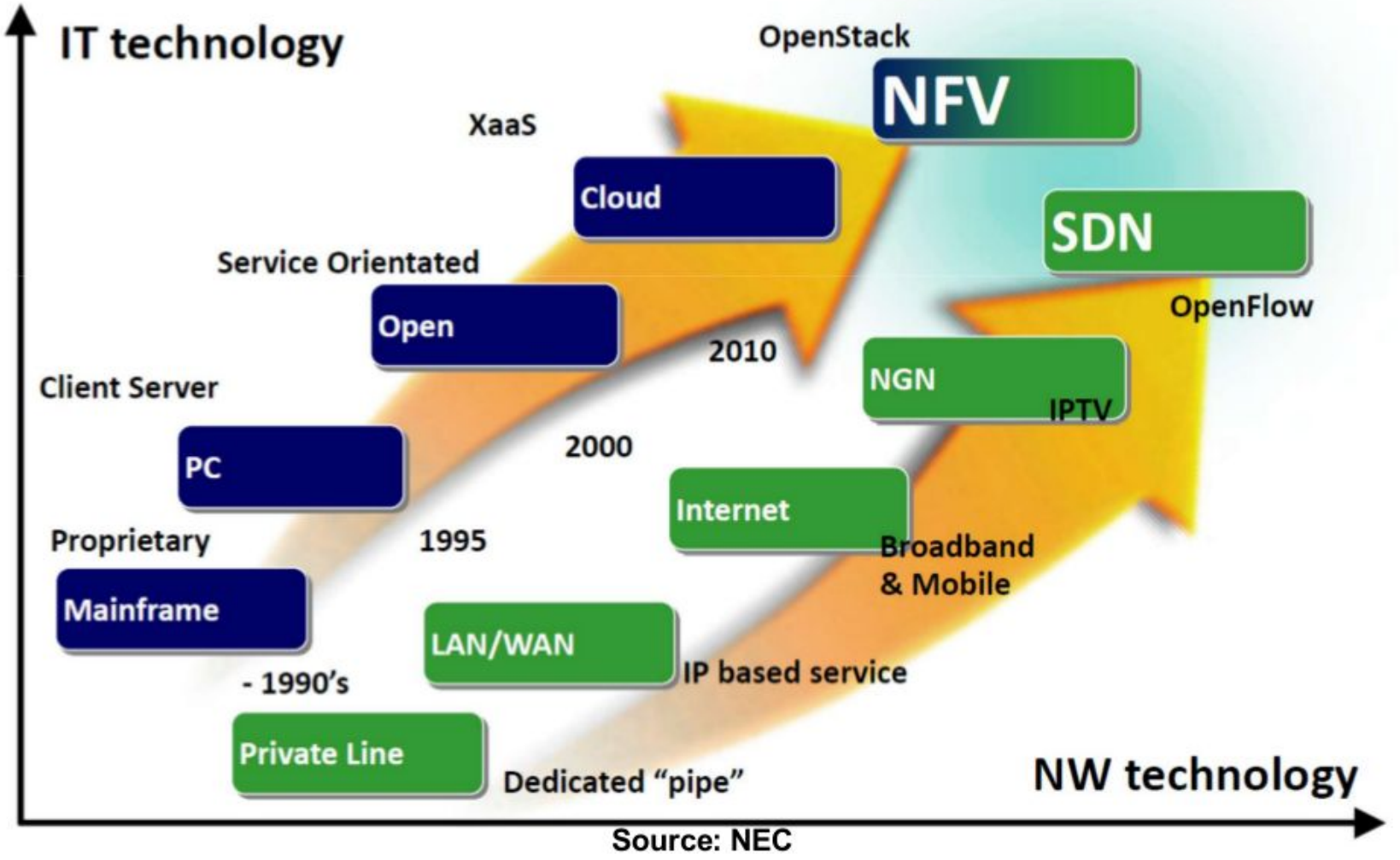
Operational costs increase (CAPEX, OPEX), benefit decrease (ARPU)

What about network service providers?



Network Function Virtualization: from propritary, expensive appliances to standard hardware (and/ or VMs running on standard servers!)

What about network service providers?



SDN + NFV

UNIVERSIDAD DE LA REPUBLICA URUGUAY

- NFV: device and network function architecture redefinition
  - Seeking to fulfill Service Providers (SP) needs:
    - Lower CAPEX removing proprietary hardware
    - Consolidate multiple network functions over standard industrial platforms

- SDN: network architecture redefinition
  - Joint academic and IETF proposal:
    - Separate forwarding and control plane
    - Add the capacity for network behaviour programming using well definded interfaces

SDN + NFV

UNIVERSIDAD DE LA REPUBLICA URUGUAY

- SDN: forward traffic in a flexible (programmable) way, both in physical and virtual environments

- NFV: flexible location of functional blocks (Virtualized Network Functions – NFVs) in the network/cloud

- SDN & NFV are complementary tools to achieve complete network programmability

Conclusion, upcoming subjects

UNIVERSIDAD DE LA REPUBLICA URUGUAY

- Virtual networks in the datacenter era
  - Particularly, L2 VPNs
  - What’s the role of BGP here?
- SDN&NFV

References

[Ghuston2018] Geoff Huston, The Death of Transitand Beyond. Online: https://labs.apnic.net/presentations/store/2018-03-02-death-of-transit.pdf

[GoogleB4] Jain, S.; Kumar, A.; Mandal, S.; Ong, J.; Poutievski, L.; Singh, A.; Venkata, S.; Wanderer, J.; Zhou, J.; Zhu, M.; Zolla, J.; Hölzle, U.; Stuart, S. & Vahdat, A. B4: Experience with a Globally-deployed Software Defined Wan Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM, ACM, 2013, 3-14