The Evolution of Internet Architecture

Geoff Huston AM
Chief Scientist, APNIC
April 2025

The Victorian Internet - the Telegraph

- In 1800 **Alessandro Volta** invented the battery that allowed electricity to be used in a controlled manner
- In 1820 **Hans Christian Oersted** demonstrated the connection between electric current and magnetism
- In the 1830's in the UK **William Cooke** and **Charles Wheatstone** used a system of five pointers to send text - the first use was railway signalling in the UK
- In the 1840's **Samuel Morse** developed a simpler system using a keypad to complete a circuit.
- By 1861 telegraph lines spanned the US
- By 1870 an undersea cables spanned the Atlantic

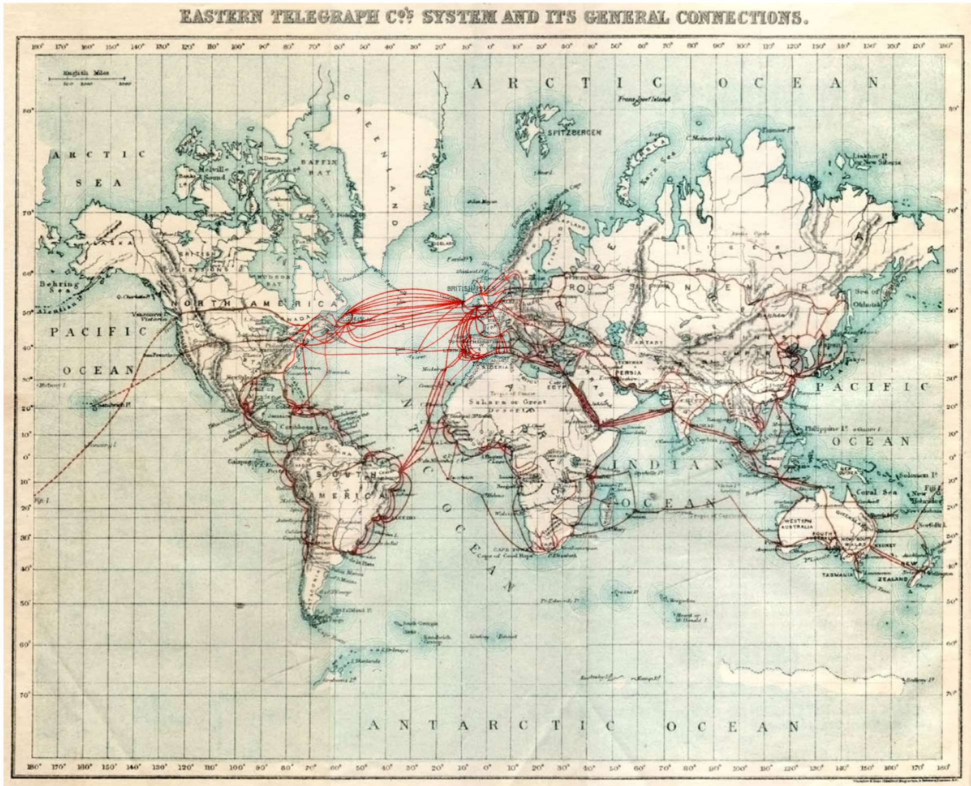
The Great Telegraph Boom

The period from the 1850's to the 1900 saw major investments in national and international telegraph cable systems

Most of the initial international investment activity was from the UK – by 1982 British companies owned and operated two thirds of the world's telegraph cables.

In the US newspapers expanded 5-fold in the period 1840 – 1860 as 50,000 miles of telegraph cable were installed

When combined with the railway this became an effective means for the projection of power and control – enterprises saw opportunities in extensive reach, creating private monopolies to complement the older state-sponsored monopolies



The Next Wave: the Telephone

- First shown to the world at the 1876 World Exposition at Philadelphia, its invention triggered a struggle to the death between Western Union's telegraph and Bell's telephone
  - Although Western Union never fully appreciated that the telephone was an existential threat to the telegraph until it was simply too late.
- Thousands of regional telephone companies appeared all over the world in the following years

The Formation of the Telephone Cartel

- Theodore Vail - President of American Telephone and Telegraph (twice!) - oversaw the construction of a national monopoly masquerading as a public utility through the Kingsbury Commitment with US Congress in 1913
- AT&T divested itself of Western Union Telegraph and in return created a substantial private monopoly under the catch cry of "one policy, one system and one universal service"
- Other countries emulated this transformation from competition to national monopoly in just a few years, using existing telegraph monopoly to subsume telephone operators into public utility structures

The Telephone Network

This was the major technology achievement of the twentieth century. This network:

- Connected handsets to handsets
- Was intentionally transparent
- Was implemented by real time virtual circuit support between connected edge devices
- And was designed as a network-centric architecture with minimal functionality in the edge devices

The rise of Computer Networks

The original concept for computer networks was based entirely on the telephone network

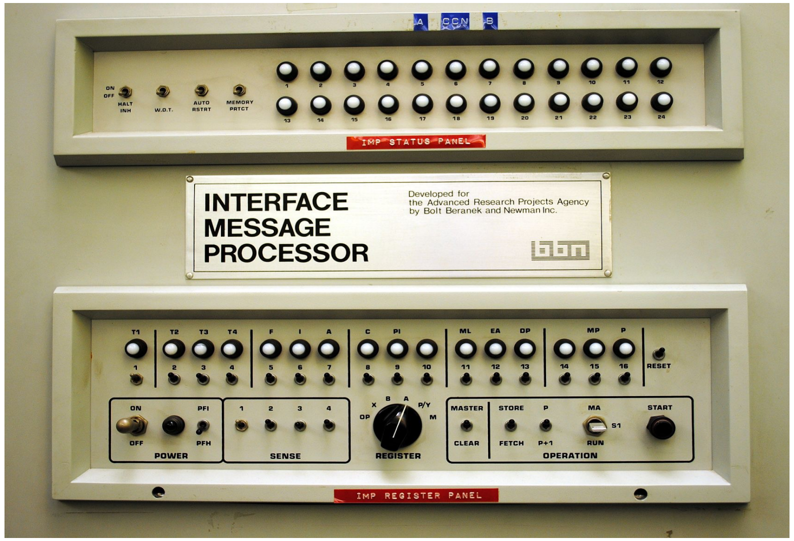
The network was there to enable connected computers to exchange data within a context of dynamic point-to-point two-party connections:

- All connected computers were able to initiate or receive “calls”
- A connected computer could not call ”the network” – the network was an invisible common substrate
- It made no difference if the network had active or passive internal elements

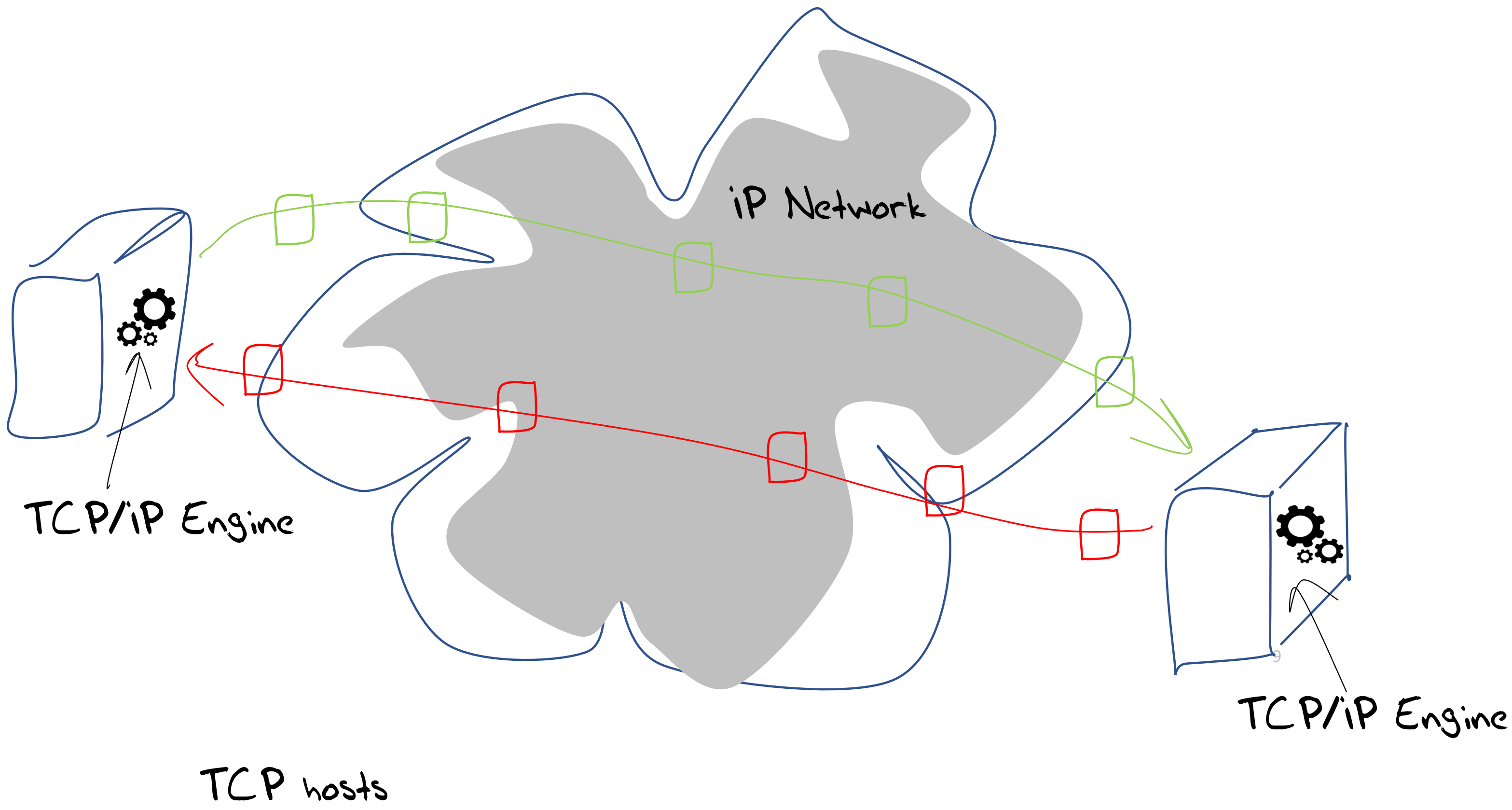
Internet Architecture (c1980's)

"End-to-End" Datagram Packet design:

- Connected computer to computer
- All data is segmented into independent packets
- The network switching function was stateless
  - No virtual circuits, no dynamic state for packets to follow
- All endpoints are uniquely addressed
- 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 (c1980's)



The Result was Revolutionary!

By stripping out network-centric virtual circuit states and removing time synchronicity the resultant packet carriage network was minimal in design and cost and maximized flexibility and efficiency

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

But

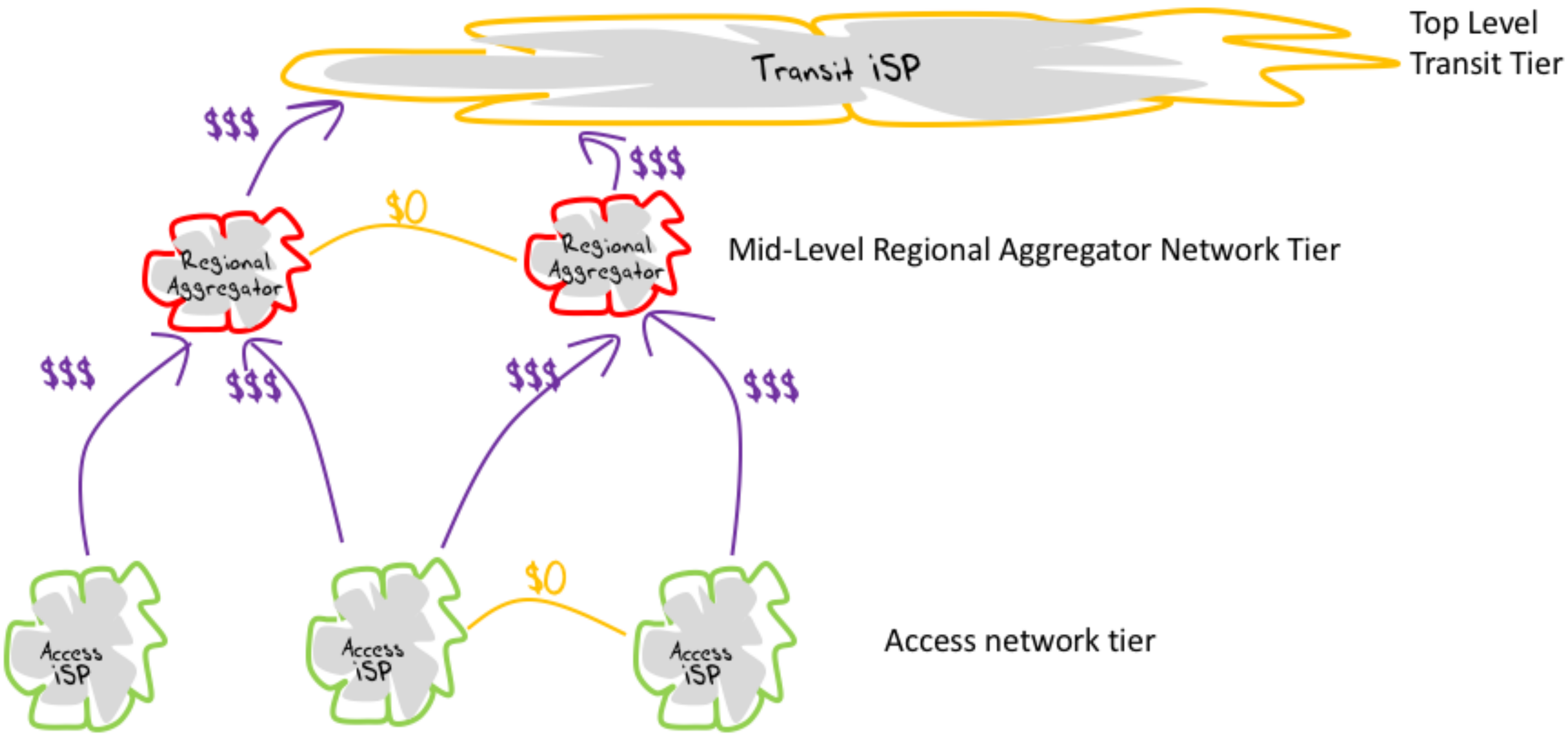
The price of runaway success is uncontrolled growth!

- The underlying connectivity fabric and IP routing systems were under scaling pressure
- We needed a way to respond to continual growth in a way that allowed aggregation of individual requirements
- The response was aggregation in addressing, routing and connectivity

Scaling Routing and Connectivity by Role Specialization

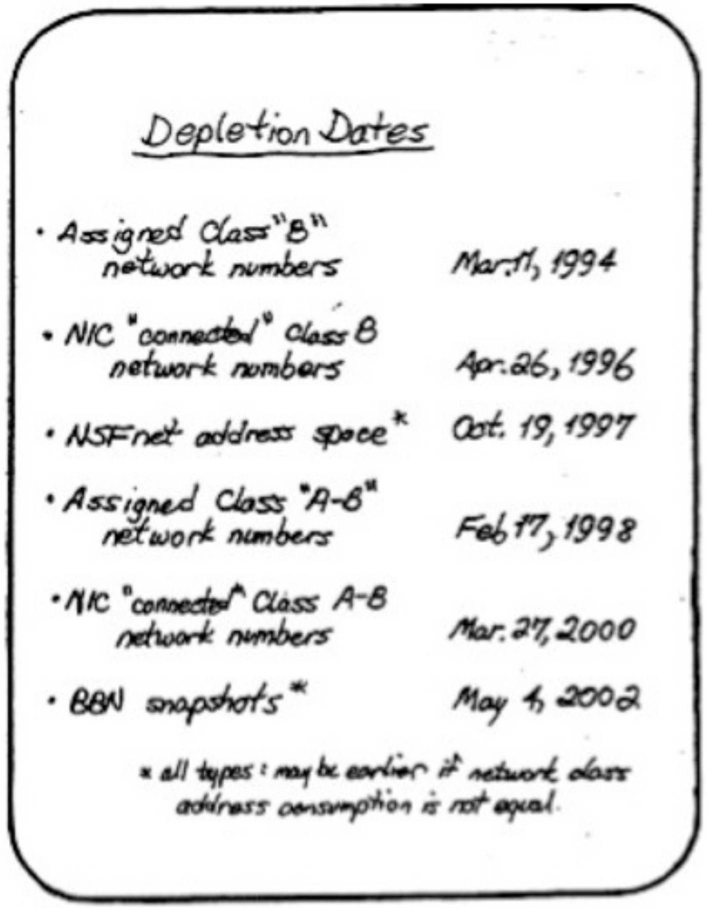
- In the regulated world of national telephone operators every telephone network was "equal"
- Markets do not normally support such outcomes, and we saw role specialization as a way of sustaining efficient distribution chains to support public services
- We rapidly started differentiating between Internet networks differentiating on roles and services to allow effective aggregation and differentiating on the flow of revenues between networks

The 1990's Internet



But

- That wasn’t the only scaling pressure point
  - By 1990 it was clear that the Internet address plan was failing to match the demands of growth of connected hosts
  - Our response was IPv6 – an expansion of the address field in IP – which changed little else



Frank Solensky – IETF August 1990

Some changes are easy…

But some are hard!

- It was relatively easy to change the underlying connection framework and the routing system
  - The scale of the change was a few thousand entities, and the changes were able to be performed in a piecemeal fashion
- It was extremely hard to the change the IP protocol in connected devices
  - The scale was far greater, the protocol platform was immature and the economic incentives to change did not exist

The cost of peer-to-peer

- Uniquely addressing each connected device has been very resistant to scaling pressures
- The pragmatic response has been to dispense with unique addresses for most connected devices, and instead of further refining a peer-to-peer network we've turned our attention to moving into client/server networks

Client/Server

Breaking the edge into **clients** and **servers**

- Access networks service the needs of “clients”
- Clients are not directly reachable by other clients
- Clients only connect to services
- Clients don’t need persistent external addresses

Services need persistence in terms of identity to facilitate client connection, while clients don’t

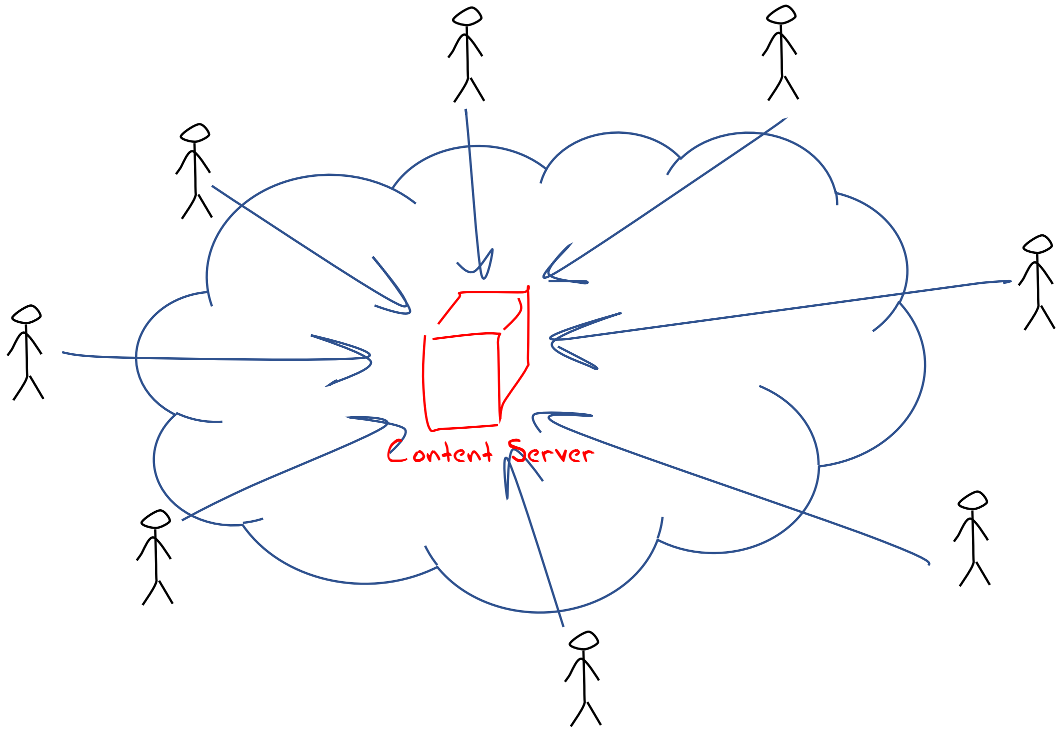
Addressing for Clients

- Clients do not need a persistent external address as they are never the target of service binding operation
  - Clients can “borrow” a session token from a common location-based address/port pool – NAT
- NATs are incrementally deployable
- Provide a network-level response to address scaling issues
- And are the mainstay of today’s IP-level Internet

Addressing for Servers

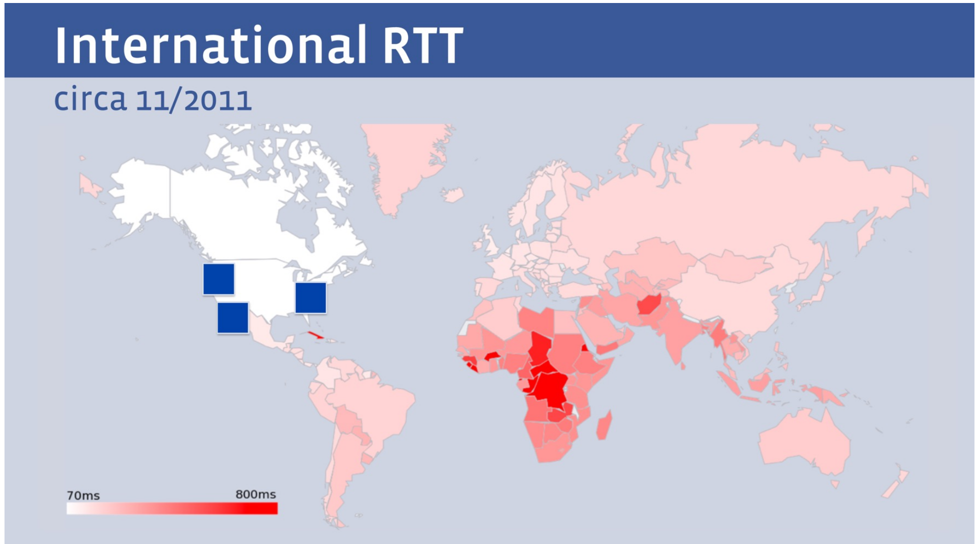
- Servers require a persistent service identity
- But that's not quite the same as a need for a persistent IP address
- The model of service provision has evolved over time…

The Evolution of Service Models



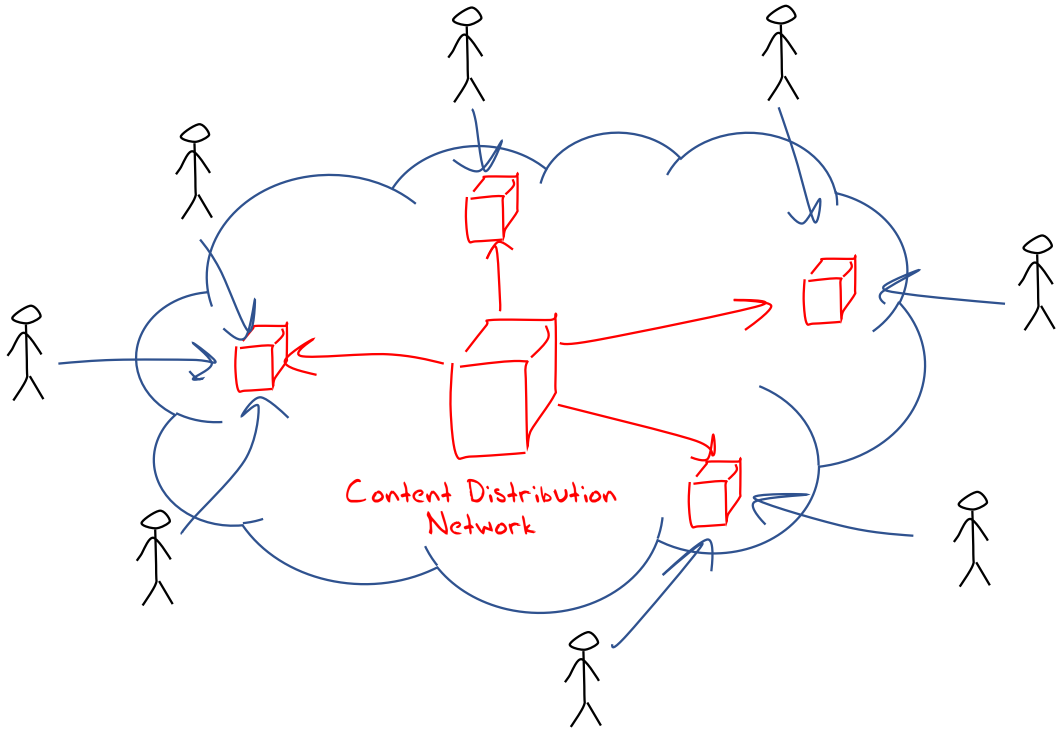
The Tyranny of Distance

But not all clients enjoy the same experience from a single service



Facebook presentation at NANOG 68 - 2011

Content Distribution Networks



Let them eat data!

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!

How to Replicate Service?

IP Anycast

- Use the DNS to map the service name to an IP address
- Originate a route to this service address in BGP from many locations
- Use the BGP's preference for "shortest path" to offer the closet instance of each service platform to each user

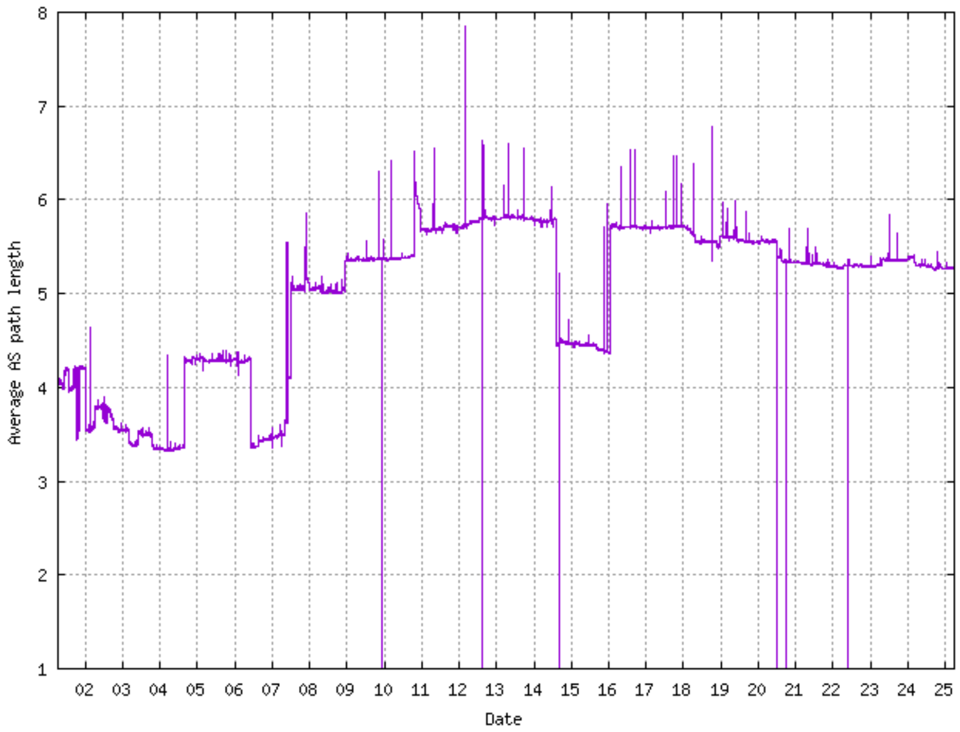
How to Replicate Service?

IP Anycast has some issues :

- BGP routes on minimizing AS Path length, not delay, load or responsiveness
- the BGP network is dense (heavily compressed) and the average AS Path length is too small
- Route instability throws off long-held sessions

This implies that Anycast service platforms are prone to poor service selection and low stability stability

Which means that anycast works best for short, simple service transactions (e.g. DNS) and not for longer sustained transactions (e.g. video streaming)



25 years of BGP Average AS Path Length – AS131072

How to Replicate Service?

- DNS Service Steering
  - Replicate the service on a number of platforms, each with their own IP address that is location-based.
  - Alter DNS authoritative servers to perform distance calculation between the DNS querier and the collection of available service platforms, and respond with the IP address of the "closest" service platform

```
$ dig www.sbs.com.au
;; ANSWER SECTION:
www.sbs.com.au.                179    IN    CNAME   www.sbs.com.au.edgekey.net.
www.sbs.com.au.edgekey.net.    179    IN    CNAME   e7065.b.akamaiedge.net.
e7065.b.akamaiedge.net.        3      IN    A       104.116.108.252
```

Client is located in Sydney, Australia

Server is located in Sydney, Australia

How to Replicate Service?

- DNS Service Steering
  - Replicate the service on a number of platforms, each with their own IP address that is location-based.
  - Alter DNS authoritative servers to perform distance calculation between the DNS querier and the collection of available service platforms, and respond with the IP address of the "closest" service platform

Client is located in Frankfurt, Germany

```
$ dig www.sbs.com.au
;; ANSWER SECTION:
www.sbs.com.au.                250    IN    CNAME   www.sbs.com.au.edgekey.net.
www.sbs.com.au.edgekey.net.    235    IN    CNAME   e7065.b.akamaiedge.net.
e7065.b.akamaiedge.net.        20     IN    A       23.35.228.216
```

Server is located in Frankfurt, Germany

How to Replicate Service?

- DNS Steering has some issues:
  - This approach assumes that the DNS recursive resolver is located close to the end client - which breaks down with the use of remote DNS recursive resolvers (Google, Cloudflare, Quad 9)
    - The inclusion of the Client’s subnet into the DNS Query is somewhat of a privacy leak
  - Dynamic changes to the service constellation is hard to control due to DNS content caching.
  - Load Balancing across multiple service platforms can be challenging

How to Replicate Service?

- Application Steering
  - Break the content into separate service “chunks”
  - Replicate the service chunks on a number of platforms, each with their own IP address that is location-based.
  - Have the application direct the client to the “closest” chunk server
  - Reassign the client to a different server location if application performance falls below some threshold

Service Models

- **IP anycast** is highly effective for short transactions (e.g., DNS over UDP)
- **DNS steering** is effective for high-speed service transactions
- **Application steering** is effective for extended transactions using decomposable elements (e.g., video streaming)

Changes to the Internet

These changes are a reversal of the Internet's original service delivery model

- Instead of using a network to connect local users to remote services in a "just in time" delivery model we now are able pre-provision services to multiple locations that are local to users in a "just in case" service provisioning model
- When local users access locally provided services there is no reliance on the network to span a distance gap
- No distance implies that we can deliver services that are faster and cheaper

Change

Greater capacity in edge networks has enabled…

Greater use of high-volume streaming content, which has lead to …

Adoption of higher capacity technologies in edge networks, which

Generates economies of scale that enables …

Reductions in the unit cost of carriage in edge networks

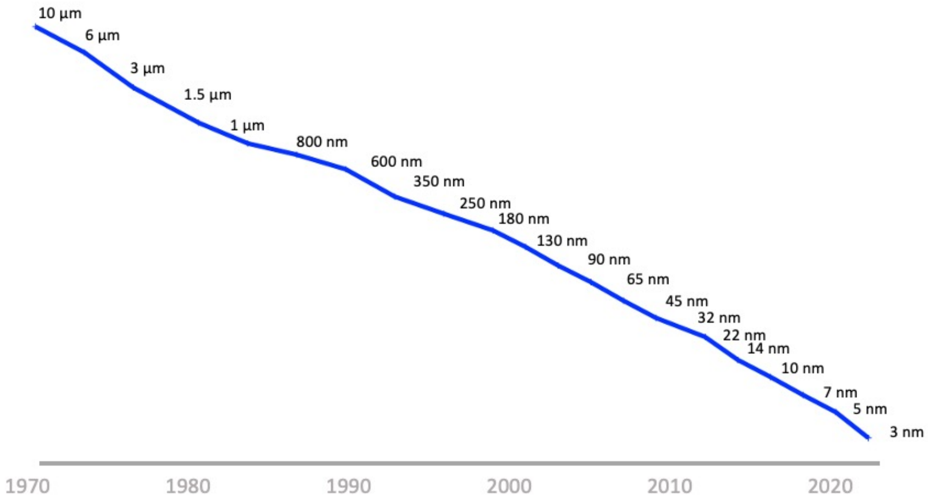
“Bigger” induces “Faster” and “Cheaper”!

How did this happen?

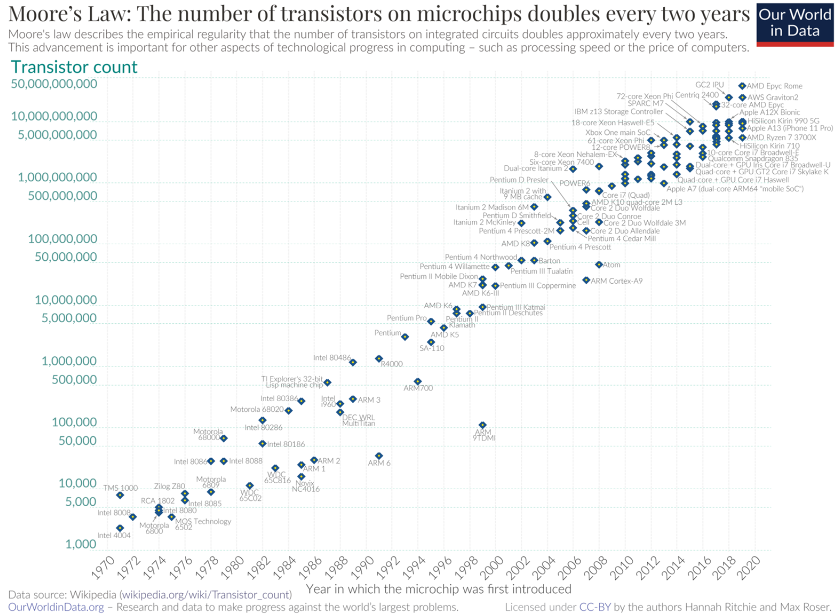
- The Internet is constructed on the foundation of a market-based economy, not a command-and-control economy
- It was assumed that an open market-based activity would generate efficient outcomes based on competitive pressures between providers
- Yet the Internet is not a highly competitive environment!

  (indeed, it's strongly consolidated and not very competitive at all!)
- What's driving this evolution in the Internet's basic architecture?

The Driver of Change: Moore's Law



Silicon Chip Track Width over time



Silicon Chip transistor counts

Year	Mode	Baud	Capacity/Lambda	Cable Capacity	DSP
2010	PM-QPSK	32 GBd	100G	8T, C-Band	40nm
2015	PM-16QAM	32 GBd	200G	19.2T, Ext C	28nm
2017	PM-32QAM	56 GBd	400G	19.2T, Ext C	28nm
2019	PM-64QAM	68 GBd	600G	38T, Ext C	16nm
2020	PS-PM-64QAM	100 GBd	800G	42T, Ext C	7nm
2022	PCS-144QAM	190 GBd	2.2T	105T, Ext C	5nm

Year	Processor	Cores	Transistors	Clock	Cost	$/core
2019	Rome	64	40B	2.25GHz	$6,950	$109
2022	Milan	64	26B	2.20GHz	$8,800	$138
2022	Genoa	96	90B	2.40GHz	$10,625	$110
2023	Bergamo	128	82B	2.25GHz	$11,900	$92

What does this mean?

- The economics of silicon chip evolution have a profound impact on the computing space - few technologies has been able to survive more than 5 years in this sector!
  - What was too expensive, too slow, or just impossible to scale up becomes quickly viable when the currency of computation and storage changes so quickly
- The result is that no business plan has been able to survive more than 5 years in the computing/communications marketplace!
- From planning, to debut, to consolidation, maturity, and then to obsolescence, a market service offering has at best just 5 years to do it all!

Moore's Law is BRUTAL!

- What is driving the economics of digital delivery systems in today's networks is **not** the historical use of pricing as a means of rationing access to a scarce common resource
- This is an environment that has switched over to abundance of processing, storage and communications
- Consolidation and Centrality of goods and service provision is not a surprising outcome in this space – its INEVITABLE
  - What would be far more surprising would be if consolidation and centrality was NOT the outcome!

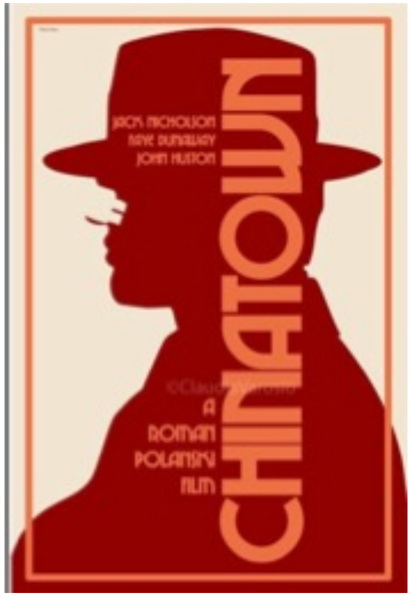
Moore's Law is BRUTAL!

This is a tough environment where:

- Smaller entities almost always fail
- Some larger entities may get sucked up by being acquired by yet larger entities
- And only the very largest of entities can afford to buy a future

Gittes: How much are you worth?
Cross: I've no idea. How much do you want?
Gittes: I just want to know what you're worth. Over ten million?
Cross: Oh my, yes!
Gittes: Why are you doing it? How much better can you eat? What can you buy that you can't already afford?
Cross: The future, Mr. Gittes - the future!

Chinatown (1974)



Change

Abundance and scale have driven radical changes across the Internet's basic architecture

- Networks are no longer share common transit services that connect users to services

  ("sharing" is so yesterday!)
- Content distributors are using abundance of computing, storage and communications capacity to bring content and service replicants to each user in advance of actual use (pre-provisioning just in case)

We pre-provision content and service at the edge of access networks and no longer rely on networks to carry user's traffic to remote service points.

What about network architecture?

- We’ve moved beyond address-based network architectures
  - Address uniqueness is a “relative” concept, not a “universal” requirement
  - Routing has largely been replaced by DNS service selection
- Service names are the basic distinguisher in the network
  - We use service names to establish a secured transaction context (TLS)
  - We use service names to provide authenticity of the service (Domain Name Certification)
  - We use the DNS to map a service description to a network rendezvous profile

Is this what "Named Data Networking" was all about?

- Yes and No

- The basic packet forwarding mechanisms are unchanged from IP – its just that the endpoint addresses used in a session have no enduring permanence beyond that session
- Services are named, but that's not quite the same as naming data
- What we have today in the Internet's architecture* is neither structured nor designed – its more of a hybrid situation that blends a named service environment with IP transport

* it's probably not even an "architecture" in anything but the loosest sense of that word!

Where is this heading?

- We are where we are as a result of the inexorable pressures of Moore's Law on the technology underpinnings of the Internet, combined with the more chaotic forces of market pressures and macro-economics
- This coupled with the scaling pressures bought about by the displacement of more traditional service delivery models by their digital analogues through ubiquitous low-cost digital capability
- But it be foolish to think that Moore's Law will continue to deliver improvements indefinitely – it can't and it won't
- But what happens then is entirely unclear!

Thanks!