The Death of Transit and the Future Internet



Dave Crosby, https://www.flickr.com/photos/wikidave/4044498586/

Geoff Huston

Chief Scientist, APNIC

This presentation is not about any specific network details:

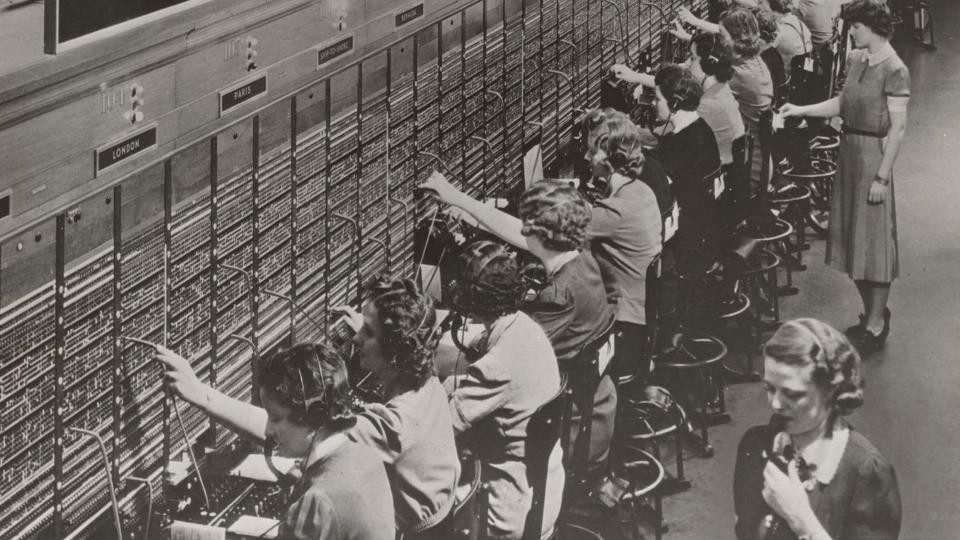
Or specific network plans Or particular services Or any particular technology Or anything like that

It's about architecture

And, in particular, about the evolution of network architecture in the internet

Our Heritage

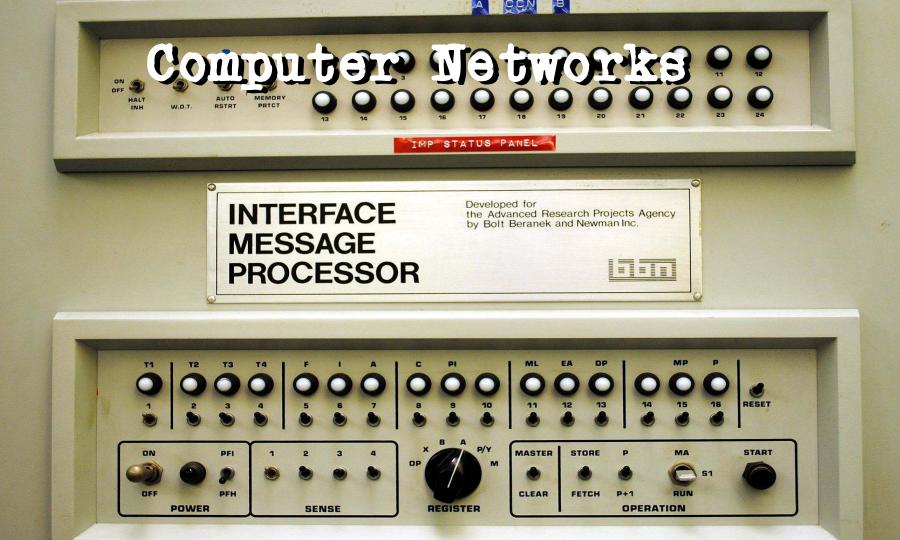
The Telephone Network



The Telephone Network

The major technology achievement of the twentieth century

- Connected handsets to handsets
- The network was intentionally transparent
- Real time virtual circuit support between connected edge devices
- Network-centric architecture with minimal functionality in the edge devices



Computer Networks

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

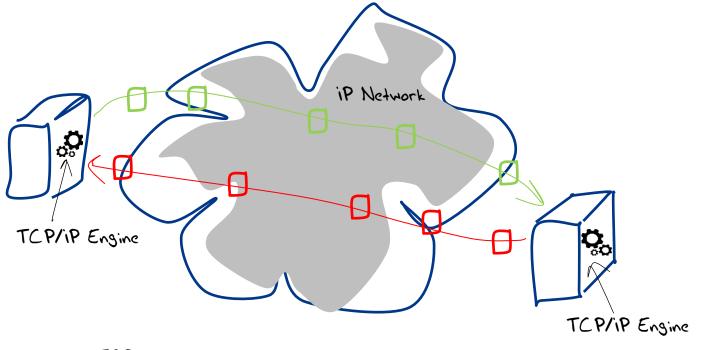
- The network was there to enable connected computers to exchange data
 - All connected computers were able to initiate or receive "calls"
 - A connected computer could not call "the network" the network was an invisible common substrate
 - It made no difference if the network had active or passive internal elements

Internet Architecture (c1980's)

"End-to-End" 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
- 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)



TCP hosts

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

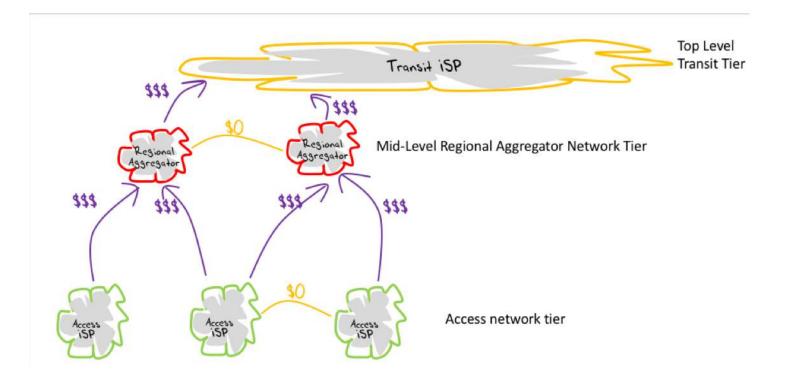
Role Specialization

In the regulated world of national telephone operators every telephone network was "equal"

Markets do not normally support such outcomes, and we see 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 and differentiating on the flow of revenues between networks

The 1990's Internet



Transit Networks were "special"

- These were the so-called "Tier One" networks
- These networks collectively managed the "default-free zone" and arbitrated reachability on the network
- These networks were at the apex of the money flow within the Internet ecosystem
- They effectively formed a cartel that defined the Internet as we knew it at the time

Enter Content

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

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 Carriage

Who pays whom?

- The only reason why access networks have clients is because there are content services that clients want to access
 - Carriage access providers should directly pay for content for their users

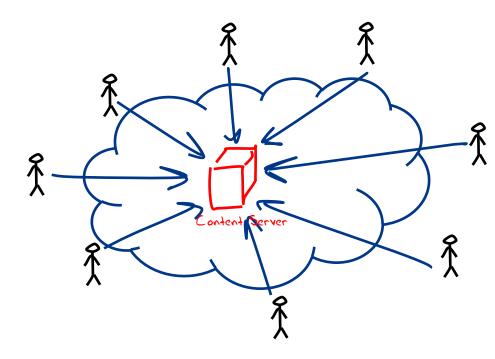
- There is no "end-to-end" financial settlement model in the Internet both "ends" pay for access and network providers settle between themselves. To a carriage network, content is just another client
 - Content providers should pay for carriage, just like any other client

Content vs Carriage

Who pays whom?

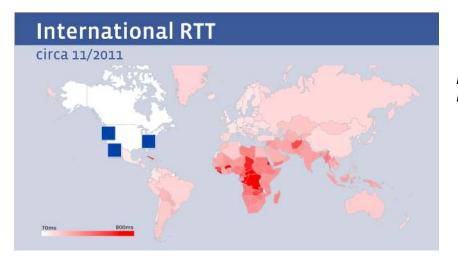
- The only reason why access networks have clients is because there
- are content services that clients want to fight by
 Carriage access provider resolved this fight users
 The content folk resolved the top and created users
 The content folk the top and created users
 The going over the top with end users
 The going bips directly with end users
 The going bips directly as settlement model in the Internet both relationships and network providers settle between access and network providers settle between theruseives. To a carriage network, content is just another client
 - Content providers should pay for carriage, just like any other client

The Evolution of Content Service



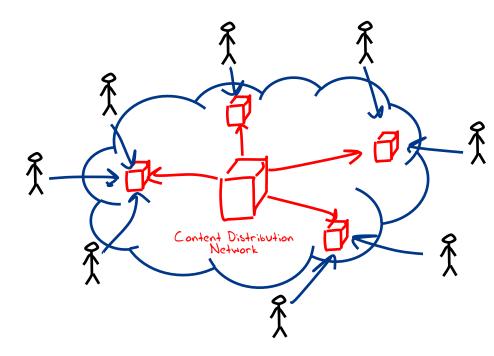
The Tyranny of Distance

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



Facebook presentation at NANOG 68

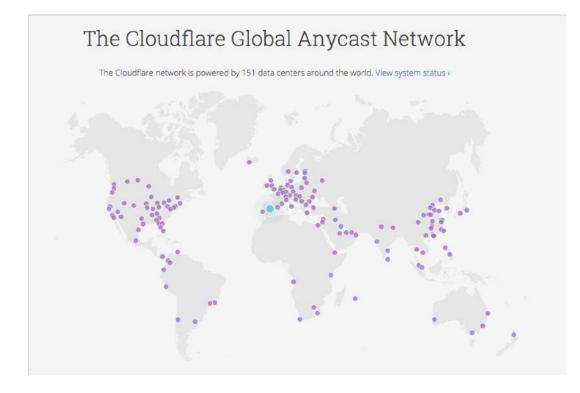
Enter Content Distribution



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!





fastly Products Customers Pricing Support

Blog Contact Login Search

Sign up

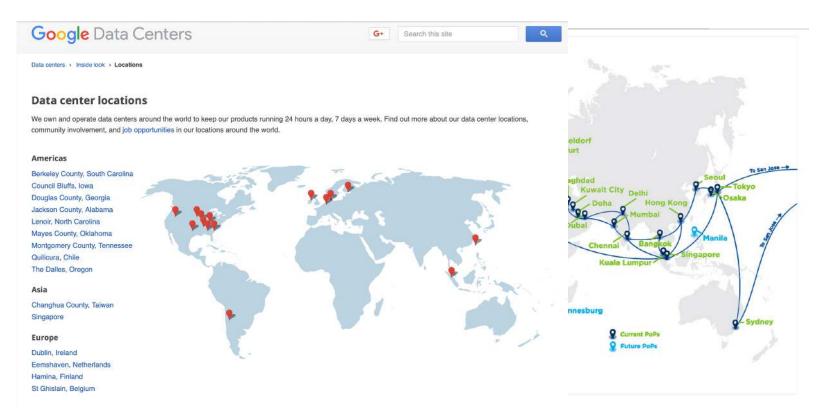
A new architecture for the modern internet

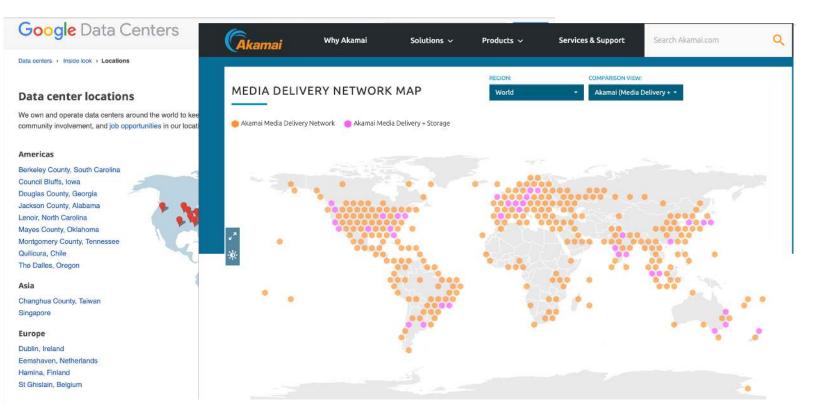
Deploying thousands of small, scattered points of presence (POPs) may have worked for legacy CDNs in the clail up era, but the internet has become increasingly dynamic, and spinning disks no longer get the job done. Fastly has taken a fundamentally different approach, we've focused our efforts on placing fewer, more powerful POPs at strategic locations around the world. With Tier 1 transit, solid-state or ve (SSD) powered servers, and an engineering team that lives to optimize for speed, we ve built a blazing fast network that requires less hardware to deliver comprehensive global reach. Fastly's high-density POPs enable us to serve more from cache, including static and event-driven content. This improves your cache hit ratio, resulting in better user experiences.











Today's Internet Architecture

We've split the network into **clients** and **servers**

- Web servers
- Streaming servers
- 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

Role Reversal

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

Role Reversal

And that means changes to the network. - Public Networks no long ound implications for the internet via ISP it has some profound implications to/from service portals

Private Networks carry content to service portals via CDN services

Who's building now?

Almost all new submarine international cable projects are heavily underwritten by content providers, not carriers

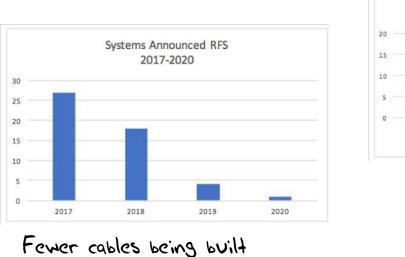
Large content providers have huge and often unpredictable traffic requirements, especially among their own data centers. Their capacity needs are at such a scale that it makes sense for them, on their biggest routes, to build rather than to buy. Owning subsea fibre pairs also gives them the flexibility to upgrade when they see fit, rather than being beholden to a third-party submarine cable operator."

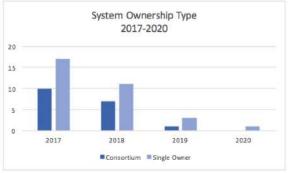
Tim Stronge of Telegeography, January 2017

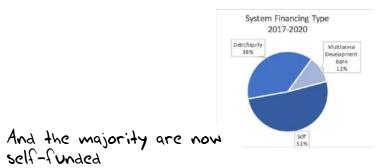


Submarine Cables

And those that are being built are now single owner cables





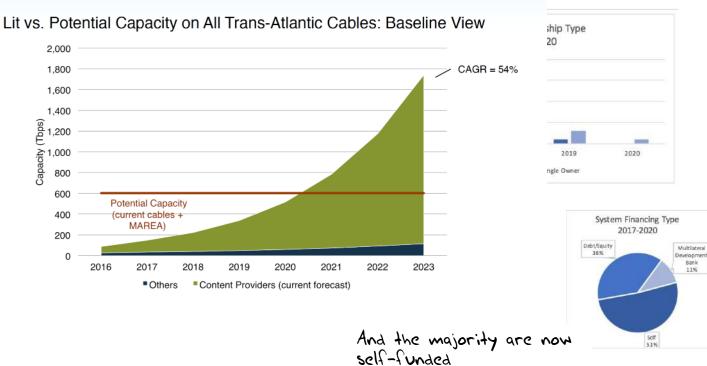


Tim Stronge, Telegeography, Jan 2017

Submarine Cables

being built are now

Growth depends on content



Tim Stronge, Telegeography, Sept 2017

Submarine

Growth depends on con

Lit vs. Potential Capacity on All Trans-Atlantic



Submarine cables now interconnect content data centers

To address the growing reporting and analysis needs of the submarine fiber industry. STF Analytics continues its Market Sector Report series - designed to provide the industry with the information it needs to make

marine Telecoms Market Sector Re port is a bi-monthly product covering a specific sector of the submarine fiber industry, coinciding with the theme of each issue of the SubTel Forum Magazine. The second edition of this report addresses the data center and Over-The-Top (OTT) provider aspect of the submarine fiber industry.

STF Analytics collected and analyzed data derived from a variety of public, commercial and scientific sources

6 SUBMARINE TELECOMS MAGAZINE

to best analyze and project market conditions. While every care is taken in preparing this report, these are our best estimates based on information provis ed and discussed in this industry. The following Executive Summ provides an overview of the topics

EXECUTIVE SUMMARY

The world continues to consume more and more bandwidth as digital activity for both enterprise and consumer applications move to "the cloud The companies behind these services - including the likes of Facebook, Google and Microsoft - continue to grow at nearly astronomical rates in effort to keep up with demand - pre

this month's re

BY KIERAN CLARK

SUBMAR

MARKET SECTOR REPORT

DATA CENTER

PROVIDER EDITION

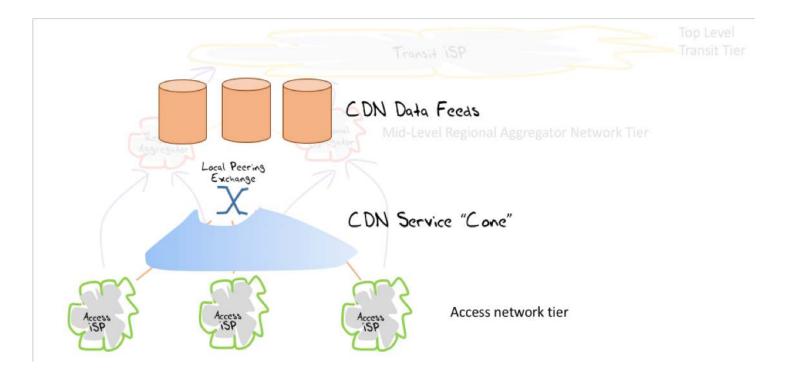
ST

& OTT

senting arine fiber industry to pro w infrastructure for the growing digital economy.Data Center and OTI providers are an increasingly integral part of the submarine cable system development process. OTT providers. like Facebook, Google, Microsoft - and Amazon in the near future - are moving from capacity purchasers to cable owners. Meanwhile, data center providers work to bridge the gap between the cable landing station and backhaul or interconnection services in an attempt to maximize network efficiency and throughput for their customers by attempting to bring once disparate infrastructure into a single facility.

Submarine Telecoms Magazine, November 2018

Today's Internet Architecture



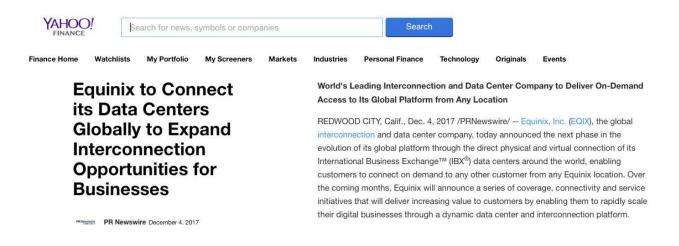
Who needs Transit?

- 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?

Closed Transit?

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



Transit?

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

Internet Names and Addresses?

If the Internet is (or maybe soon will be) a collection of discrete CDN service 'cones' then why do we expect end users to pay for the maintenance of:

- A global address plan?
- A global name system?
- A single global network?

It's not just Death of Transit

It's the re-purposing of the entire network

- Service provisioning sits within cloud providers and distributed data centres
- Edge computers are now acting as televisions into the clouded world of data
- The distinction between personal and public data realms is disappearing into the realm of corporately owned private data empires

Exactly where are we?

- We started this journey building a telephone network for computers to communicate between each other
- But now one-way content distribution lies at the core of today's Internet
- This content distribution role is an enterprise service framework rather than a public carriage service
- The internal parts of the carriage network are now being privatized and removed from public regulatory oversight

Policy?

If CDN feeder networks are private networks, and there is little residual public carriage other than last mile access networks, then what do we really mean by "public communications policy"?

In the regulatory world 'content' is *commerce*, not *carriage*!

Policy?

In today's Internet what do we mean in a policy sense by concepts such as:

"universal service obligation"

"network neutrality"

"rights of access" or even

"market dominance"

when we are talking about diverse CDNs as the dominant actors in the Internet?

The Large and the Largest

	Company	Market Cap (B)	
é	Apple	1,091	
amazon	Amazon	976	
Microsoft	Microsoft	877	
Alphabet	Alphabet	839	
	Berkshire Hathaway	523	
facebook	Facebook	473	
E2. Alibaba Group encent 腾讯	Alibaba	423	
	Tencent	388	
	JPMorgan Chase	379	
	Johnson & Johnson	370	

The world's 10 largest publicly traded companies, as ranked by their market capitalization, Q3, 2018

Content Really is King!

- None of these seven technology companies are a telephone company, or even a transit ISP, or even an ISP at all!
- All of them have pushed aside carriage networks in order to maintain direct relationships with billions of consumers
- These valuable consumer relationships are based on content services, not carriage

Content Consolidation

- There are not thousands of content service platforms
 any more
 - There are just a few left
- And the space is dominated by a small number of dominant actors who set the rules of engagement for all others

Content Consolidation

"The size and scale of the attacks that can now easily be launched online make it such that if you don't have a network like Cloudflare in front of your content, and you upset anyone, you will be knocked offline.

•••

In a not-so-distant future, if we're not there already, it may be that if you're going to put content on the Internet you'll need to use a company with a giant network like Cloudflare, Google, Microsoft, Facebook, Amazon, or Alibaba.

•••

Without a clear framework as a guide for content regulation, a small number of companies will largely determine what can and cannot be online.

Consolidation?

Alphabet is primarily an advertising company that dabbles in blue-sky technology projects. ever in the history of the world has a single company had so much control over what people know and think. Yet Washington has been slow to recognize that Google's power is a problem, much less embrace the obvious solution: breaking the company up.

Google accounts for about 90 percent of all Internet searches; by any honest assessment, it holds a monopoly at the very gateway to information in the modern world. From there, the company's power radiates outward, dominating everything from maps to smartphone operating systems to video distribution — vacuuming up huge quantities of highly specific data about users along the way.

Boston Globe, June 14 2018

Competition or Cartel?

With a small number of truly massive enterprises at the heart of the area of digital content and service is this still a space that is shaped by competitive pressures?

Or do these dominant incumbents get to set their own terms of engagement with each other, with users, and even with the public sector?

Competition or Cartel?

With a small number of truly massive enterprises d' heart of the area of digital content and service of the area of digital content and service of the area of digital content and service of the sound, it's not area that is shaped by compositive might sound, it's not area of the area of digital content and service of the sound, it's not area of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not are of the area of digital content and service of the sound, it's not area of the sound of the

We've been here before...



American Art: The Gilded Age

M ark Twain coined the phrase "the Gilded Age" in 1873. This nerm, with its connotations of superficiality and estentiations wealth, During that period of rapid industrialization, the contrast between the lifestyles of so called robber barons and average workers was enormous. The metaphor of gilded surfaces resonates in the richly decorated possessions of the roling class, from domestic firmiture to picture frames.

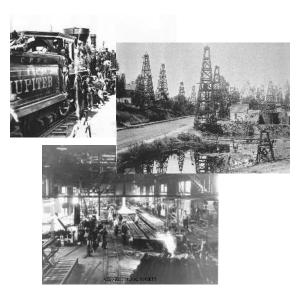
This gallery examines the leading cultural phenomenon of the 1870s and 1880s, the American Aesthetic movement, through a range of objects produced for affluent consumers. Aestheticism, rooted in the English philosophies of John Ruskin and William Morris, advanced the notion that a beautiful environment could promote moral and social reform. In the process, the Aesthetic movement helped to liberate American art and design from the confines of historicism by admitting fresh influences from foreign lands.

High Museum of Art, Atlanta

The Gilded Age

A term applied to America in the 1870 – 1890's about the building of industrial and commercial corporate giants on platforms that were a mix of industrial innovation and enterprise with elements of greed, corruption and labor exploitation

Andrew Carnegie - US Steel John Rockfeller - Standard Oil Theodore Vail - AT&T George Westinghouse – Rail Brakes Thomas Edison – General Electric J P Morgan - Banking



The Gilded Age

During this period in the United States the dominant position within industry and commerce was occupied by a very small number of players who were moving far faster than the regulatory measures of the day.

The resulting monopolies took the US decades to dismember, and even today many of these gilded age companies remain dominant in their field



At some point in the past decade or so the dominant position across the entire Internet has been occupied by a very small number of players who are moving far faster than the regulatory measures that were intended to curb the worst excesses of market dominance by a small clique of actors.



At some point in the past decade or so the dominant position across the entire Internet has been occupied by a very

small number of moving far faste measures that w the worst excess dominance by a

Alphabet				\$18.1 million
Amazon				
			\$13 million	
Facebook				
			\$11.5 million	
Microsoft				
		\$8.6 million		
Apple				
	\$7.1	million		

PUCK

small number of Total 2017 federal lobbying spends by tech giants

These actors have enough market influence to set their own rules of engagement with:

- Users,
- Each other,
- Third party suppliers,
- Regulators and Governments

By taking a leading position with these emergent technologies, these players are able to amass vast fortunes, with little in the way of accountability to a broader common public good

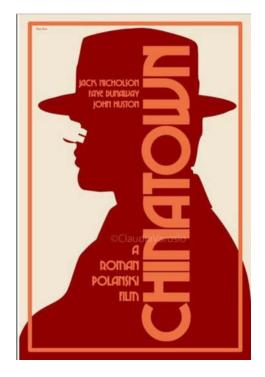
These actors have enough market influence to set their own rules of engagement with:

- Third parties this the internet we were - is this the internet we were to position with these emergent By ta technologies, these players are able to amass vast fortunes, with little in the way of accountability to a broader common public good

The Internet's 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)



What is this all about?

This is no longer just a conversation about incremental changes in carriage and communications within the Internet.

For me, the essential topic of this conversation is how we can strike a sustainable balance between an energetic private sector that has rapidly amassed overarching control of the digital service and content space, and the needs of the larger society in which we all would like some equity of opportunity to thrive and benefit from the outcomes of this new digital age.



Is it that these enterprises are:

- so big?
- exploitative of their workers?
- distorting markets?
- extracting monopoly rentals from consumers?
- not providing consumers what they want?

• In the quest for ever-faster service delivery we are seeing the return of proprietary solutions in applications and service delivery platforms that expose as little as possible to the underlying network platform

- Perhaps the problem is the looming demise of open technologies and open technology standards
 - Akamai uses Fast to improve content delivery
 - Google uses QUIC and BBR
 - Facebook and WhatsApp use strong encryption to hide the application from the network and the platform
 - Applications are no longer constructed on a platform of common libraries provided by the platform
 - Applications are now paranoid and avoid exposing their behaviour wherever and whenever possible
 - Applications are increasingly reluctant to use standard open technologies in standard and open ways

- In its place we are seeing a resurgence of various closed technologies that create a set of datacentre-to-application bindings that are impervious to all third parties
- These closed architectures make minimal assumptions about a common network protocol, a common name space or even a common name space
- What happens to the efforts that support open technologies, open standards and open networking in such a world?

Where does all this head?

Where does all this head?

i just don't know!

But i'm not sure that it's all good!

