Future Network Needs

Geoff Huston AM

Chief Scientist, APNIC

The Internet used to be simple…

1980's:

- The network was the transmission fabric for computers
- It was just a packet transmission facility
- Every other function was performed by attached mainframe computers

"dumb" network, "smart" devices

Then we went client/server

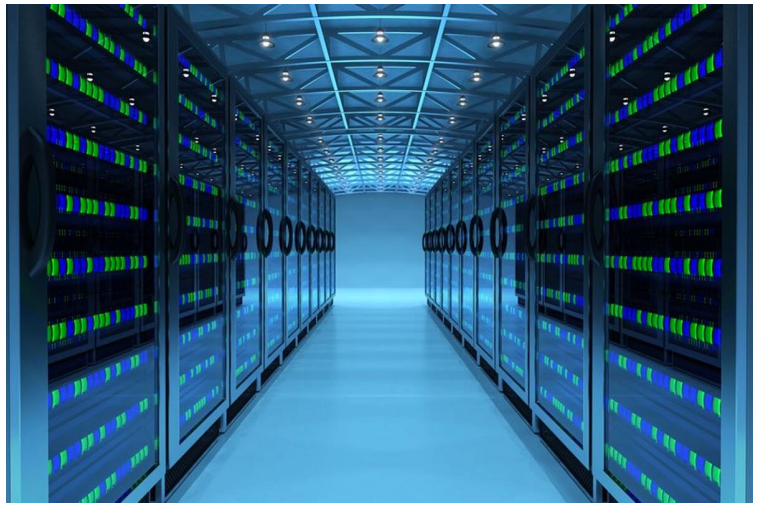
1990's:

- The rise of the Personal Computer as the "customer's computer"
- We started to make a distinction between "customers" and "network"
  - The naming system was pulled into the network
  - The routing system was pulled into the network
  - Messaging, content and services were pulled into the network
- We created the asymmetric client/server network architecture for the Internet

Internet Infrastructure of 2000

Rapid expansion of infrastructure in many directions:

- Exchanges, Peering Points and Gateways
- Transit and Traffic Engineering
- Data Centres and Service “Farms”
- Quality of Service Engineering
- MPLS, VPNs and related network segmentation approaches
- Mobility Support
- Customer Access Networks
- Content Distribution Networks



Aren't these all "different" networks?

- Well, yes they are
- The true genius of the Internet was to separate the service environment from the link technology
  - Each time we invented a new comms technology we could just “map” the Internet onto it
  - This preserved the value of the investment in “the Internet” across successive generations of comms technologies

What about the coming decades?

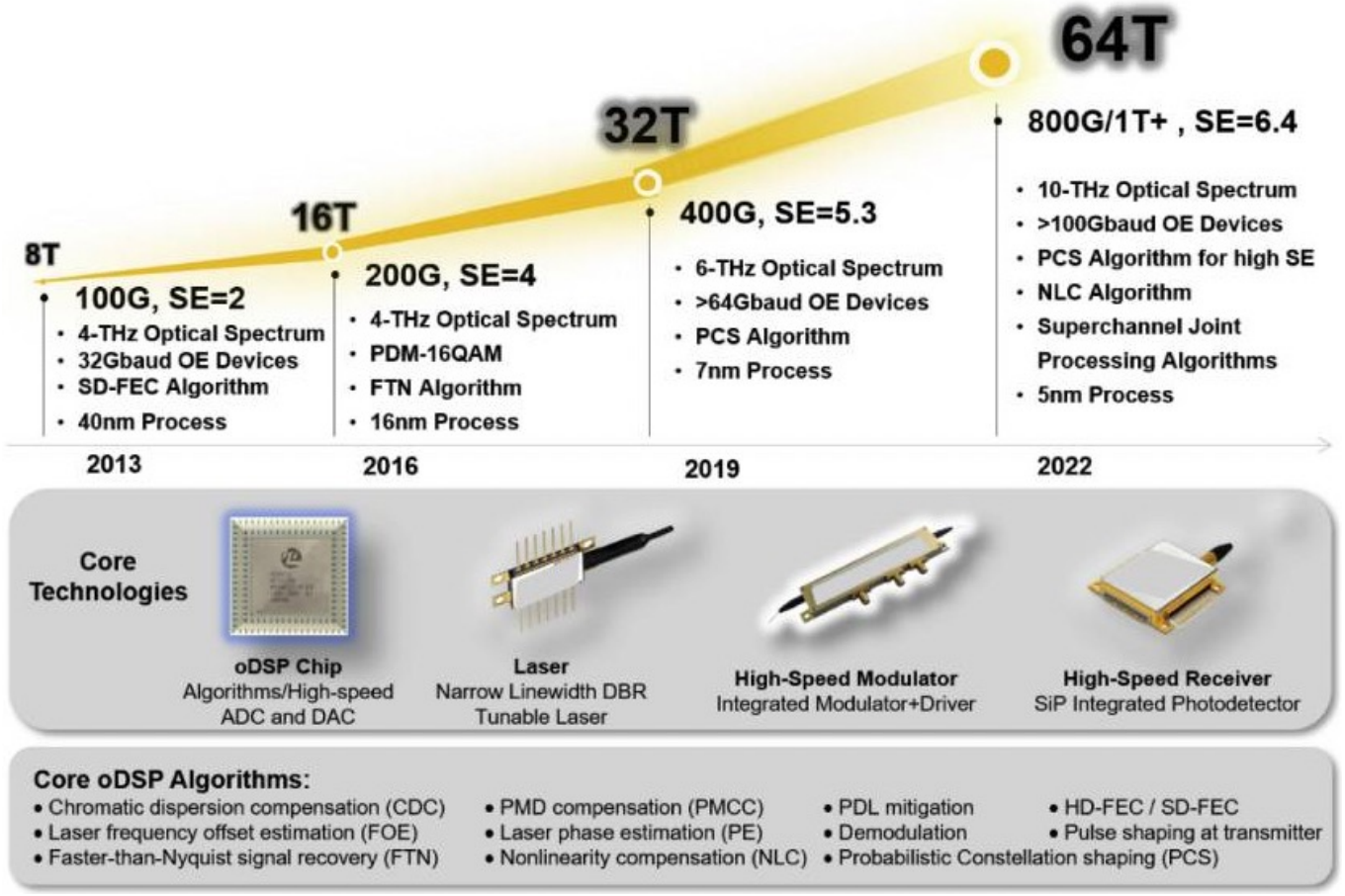
- The seeds of the dominant factors of the future environment are probably with us today
- The problem is that a lot of other seeds are here as well, and sifting out the significant from the merely distracting is the challenge
- So with that in mind lets work out the big drivers in today's environment…

What's driving change today?

- From scarcity to abundance!
- For many years the demand for communications services outstripped available capacity
- We used price as distribution function to moderate demand to match available capacity
- But this is no longer the case – available capacity in the communications domain far outpaces demand

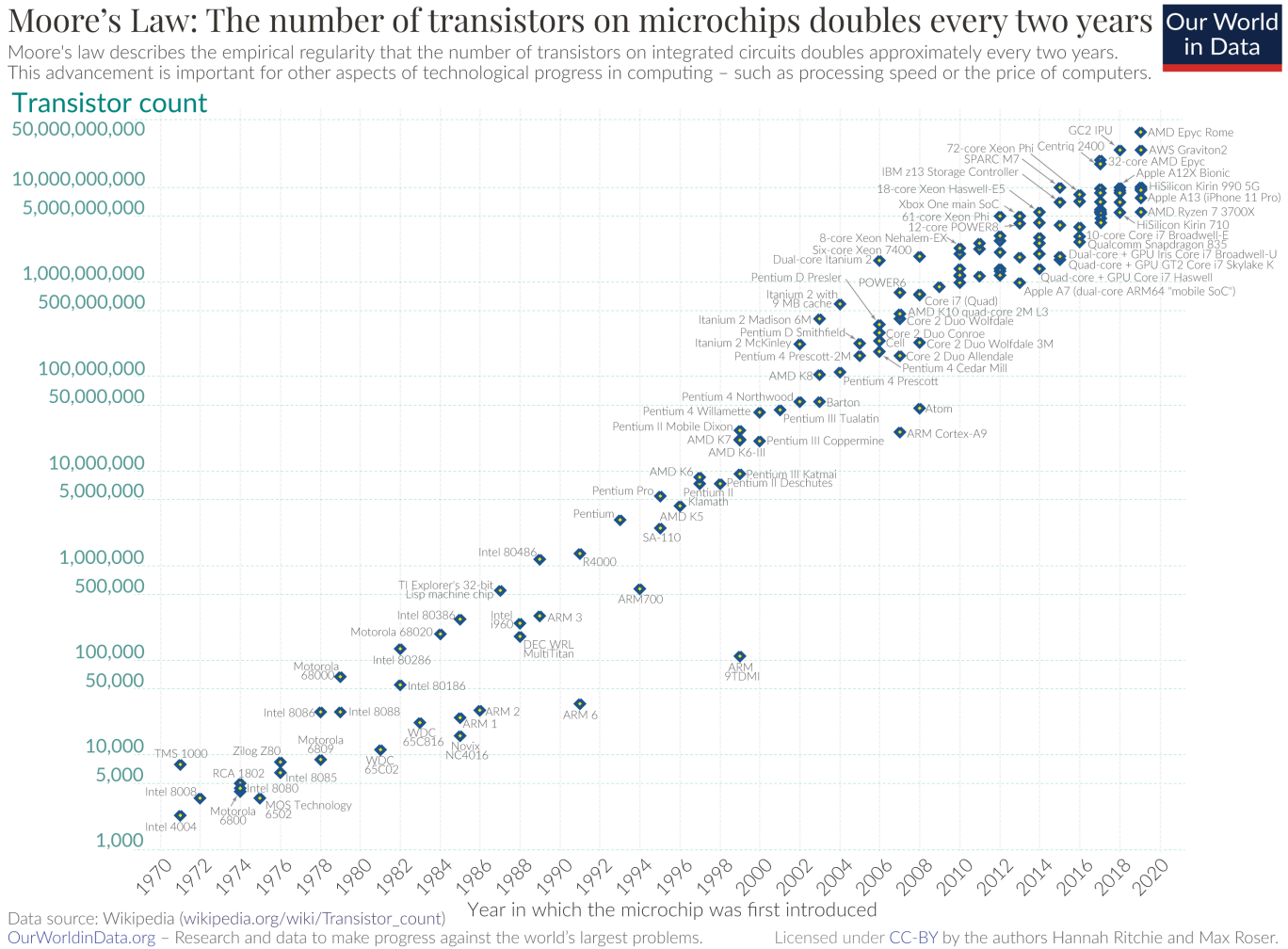
Abundant Capacity

Fibre cables continue to deliver massive capacity increases within relatively constant unit cost of deployment



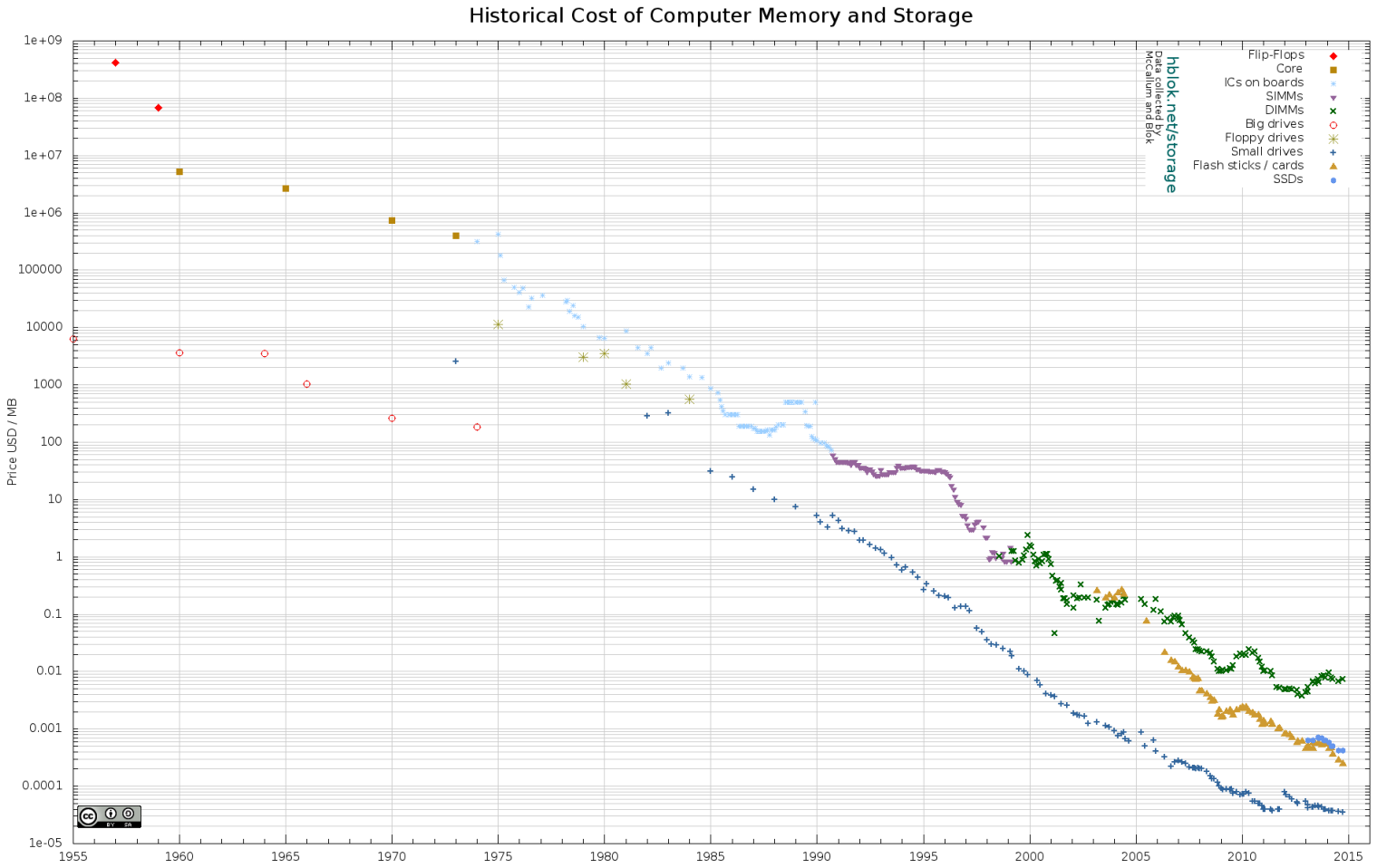
https://www.ncbi.nlm.nih.gov/

Abundant Compute Power



By Max Roser, Hannah Ritchie - https://ourworldindata.org/uploads/2020/11/Transistor-Count-over-time.png, CC BY 4.0, https://commons.wikimedia.org/w/index.php?curid=98219918

Abundant Storage



http://aiimpacts.org/wp-content/uploads/2015/07/storage_memory_prices_large-_hblok.net_.png

How can we use this abundance?

- By changing the communications provisioning model from *on demand* to *just in case*
- Instead of using the network to respond to users by delivering services *on demand* we've changed the service model to provision services close to the edge jusdt in case the user requests the service
- With this change we've been able to eliminate the factors of *distance* from the network and most network transactions occur over shorter network spans
- What does a *shorter* network enable?

Bigger

- Increasing **transmission capacity** by using photonic amplifiers, wavelength multiplexing and phase/amplitude/polarisation modulation for fibre cables
- Serving content and service transactions by distributing the load across many individual platforms through **server and content aggregation**
- The rise of high capacity mobile edge networks and mobile platforms add massive volumes to content delivery

- To manage this massive load shift we've stopped pushing content and transactions across the network and instead **we serve from the edge**

Faster

- Reduce latency - stop pushing content and transactions across the network and instead **serve from the edge**
- The rise of CDNs serve (almost) all Internet content and services from massively scaled distributed delivery systems.

- The “Packet Miles” to deliver content to users has shrunk - that’s faster!
- The development of high frequency cellular data systems (4G/5G) has resulted in a highly capable last mile access network with Gigabit capacity
- Applications are being re-engineered to meet faster response criteria

- Compressed interactions across shorter distances using higher capacity circuitry results in a much faster Internet

Better

- If “better” means “more trustworthy” and “more privacy” then we are making progress at last!
  - Encryption is close to ubiquitous in the world of web services
  - TLS 1.3 is moving to seal up the last open TLS porthole, the SNI field
  - QUIC is sealing up the transport controls from the networks
  - Oblivious DNS and Oblivious HTTP is moving to isolate knowledge of the querier from the name being queried
  - The content, application, and platform sectors have all taken the privacy agenda up with enthusiasm, to the extent that whether networks are trustable or not doesn’t matter any more – **all network infrastructure is uniformly treated as untrustable!**

Cheaper

- We are living in a world of abundant comms and computing capacity
- And working in an industry when there are significant economies of scale
- And being largely funded by capitalising a collective asset that is infeasible to capitalise individually – the advertisement market
- The result is that a former luxury service accessible to just a few has been transformed into an affordable mass-market commodity service available to all

So it's all good!

Right?

Longer Term Trends?

Pushing EVERYTHING out of the network and over to applications

- Transmission infrastructure is becoming an abundant commodity
  - Network sharing technology (multiplexing) is decreasingly relevant
- We have so much network and computing that we no longer have to bring consumers to service delivery points - instead, we are bringing services towards consumers and using the content frameworks to replicate servers and services
- With so much computing and storage **the application is becoming the service**, rather than just a window to a remotely operated service

Do Networks matter any more?

- We have increasingly stripped out network-centric functionality in out search for lower cost, higher speed, and better agility
- We are pushing functions out to the edge and ultimately off “the network” altogether and what is left is just dumb pipes
- What defines “the Internet”?
  - A common shared transmission fabric, a common suite of protocols and a common protocol address pool?

  or
  - A disparate collection of services that share common referential mechanisms?

Some issues to think about

What matters today?

- **End Point Addressing** – IPv4 / IPv6 / IPv? Absolute? Relative?
  - Is universal unique end-point addressing a 1980's concept who's time has come and gone?
  - If network transactions are localised then what is the residual role of unique global end point addressing for clients or services?
  - And if we cannot find a role then why should we bother?
  - Who decides when to drop it?
    - Is this a market function, so that a network that uses local addressing can operate from an even lower cost base gains a competitive market edge?
    - Or are carriage services so cheap already that the relative benefit in discarding the last vestiges of unique global addresses are so small that its just not worth bothering about?

Some issues to think about

What matters today?

- **Naming and Name Spaces** – DNS evolution?
  - Are ”names” a common attribute of the network, or an attribute of a service environment or application realm?
  - Should names be persistent over time?
  - Is the resolution of a name absolute or relative to the content of the resolution query?
  - In a world of densely replicated service delivery points how does a client rendezvous with the “best” service point? Does the client work it out? Or the network? Or the service?
  - If names are an attribute of applications then why do we need a single name domain? Surely each application realm can define its own name space? How can we associate a referential name space with a given name?
  - If both names and addresses are ephemeral and unstable then what defines the Internet?

Some issues to think about

What matters today?

- **Referential Frameworks?**
  - Without a common referential space then how do we usefully communicate?
  - What do we mean by “common” when we think about referential frameeworks?
  - How can we join the ‘fuzzy’ human language spaces with the tightly constrained deterministic computer-based symbol spaces?

Thanks!