An Internet Outlook

Geoff Huston

October 2001

So far, the Internet has made an arbitrary number of good and bad decisions in the design of networking components.

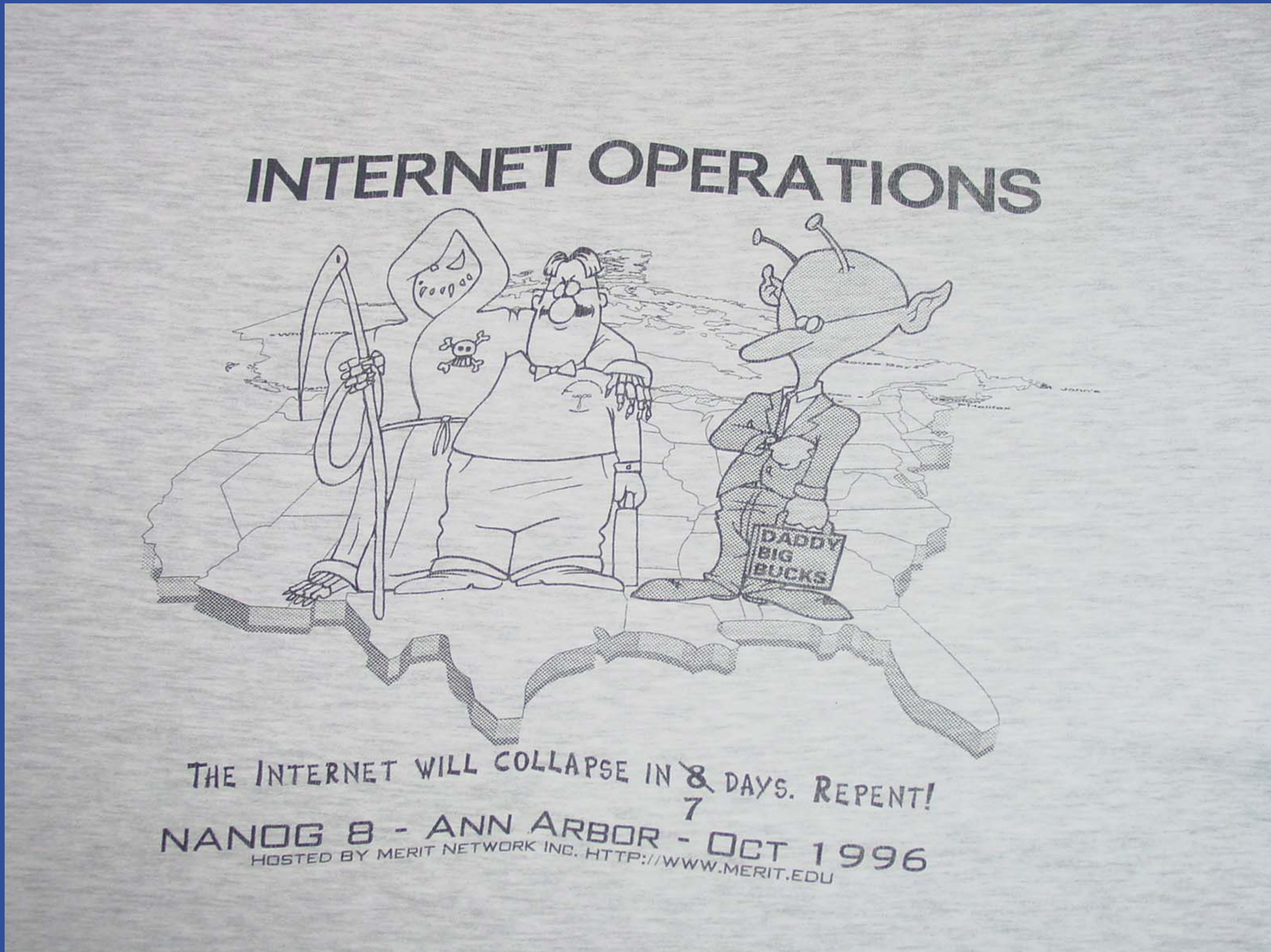
The good decisions were, of course, triumphs of a rational process at work.

In the case of the bad decisions, Moore's law has come to the rescue every time.

This may not continue to be the case...

The Internet Today

- Still in the mode of rapid uptake with disruptive external effects on related activities
- No visible sign of market saturation
- Continual expansion into new services and markets
- No fixed service model
- Changing supply models and supplier industries
- Any change to this model will be for economic, not technical reasons


INTERNET OPERATIONS
DADDY BIG BUCKS
THE INTERNET WILL COLLAPSE IN ~~8~~ 7 DAYS. REPENT!
NANOG 8 - ANN ARBOR - OCT 1996
HOSTED BY MERIT NETWORK INC. HTTP://WWW.MERIT.EDU

Collapse of the Internet Predicted – gifs at 11

- The Internet has been the subject of extraordinary scaling pressure for over a decade
- The continual concern is that with the increased pressures of commercial use the network will overload in a number of traffic concentration black spots and collapse under the pressure
- The reality so far is that the network has managed to continue to scale to meet evolving business needs without drama or disruption

- Will this continue?

Lets look at:

- Backbone Engineering
- End System Requirements
- Performance Issues
- Scaling Trust

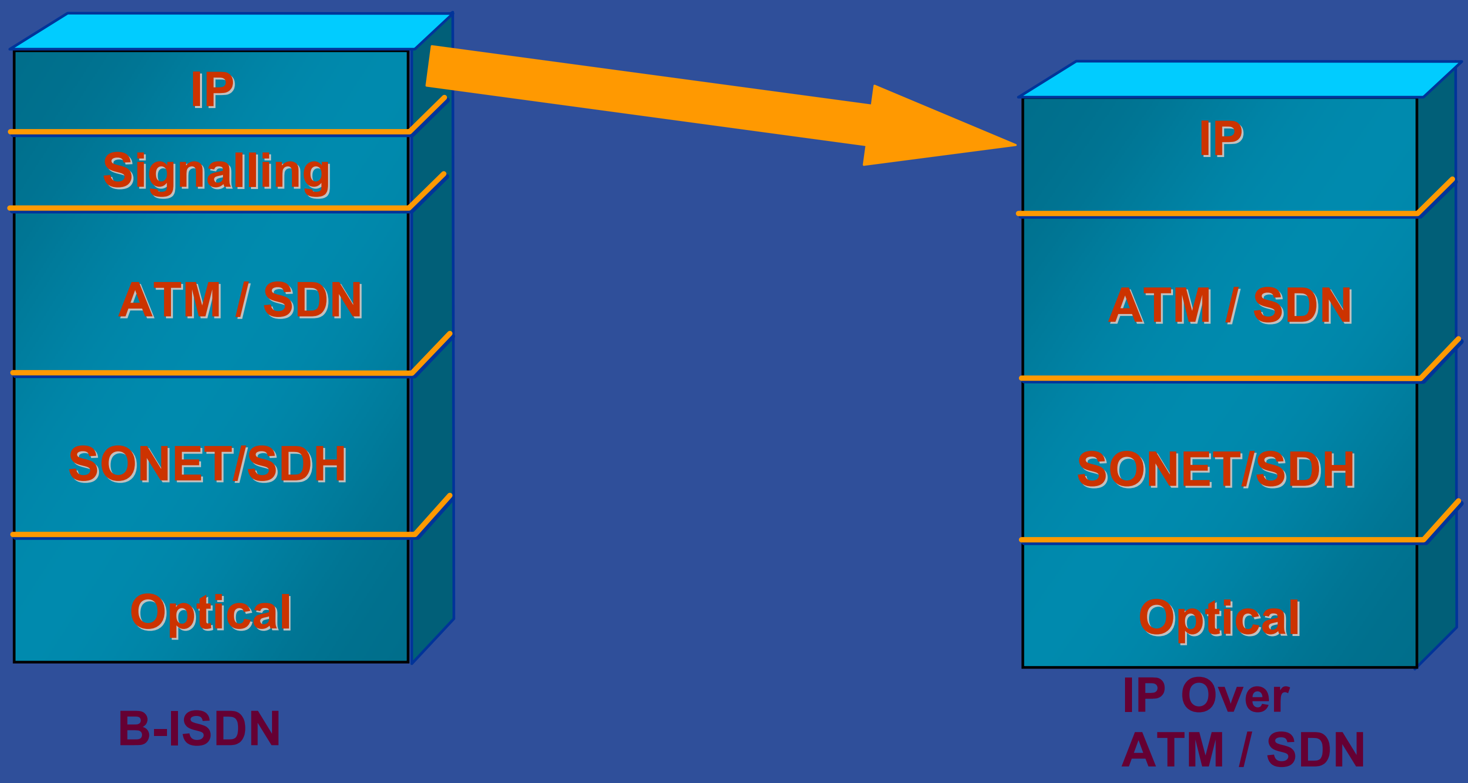
The Bandwidth Challenge

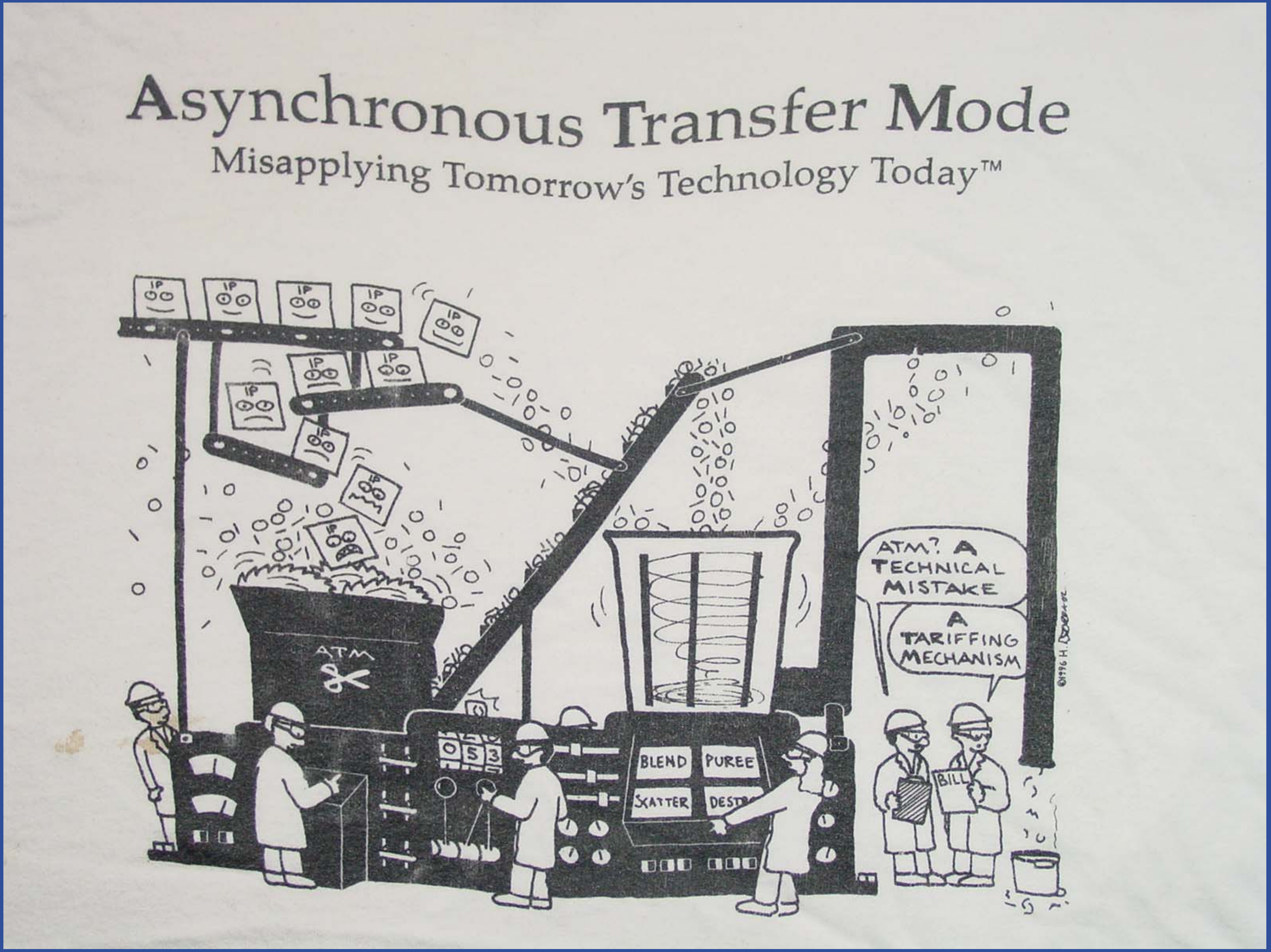
- On the Internet demand is highly **elastic**
  - Edge devices use TCP, a rate adaptive transmission protocol. Individual edge devices can sustain multi-megabit per second data flows
  - Network capacity requirement is the product of the number of edge devices multiplied by the user's performance expectation
      - Both values are increasing
  - Internet Bandwidth is exponentially increasing number
    - Rate of bandwidth demand is a doubling each 12 months
    - Moore's Law doubles processing capacity every 18 months

Backbone Technologies

- PSTN Carrier Hierarchy
  - Low speed, high complexity, high unit cost
  - 10^6 bits per second carriage speeds
- ATM

The Evolution of the IP Transport Stack

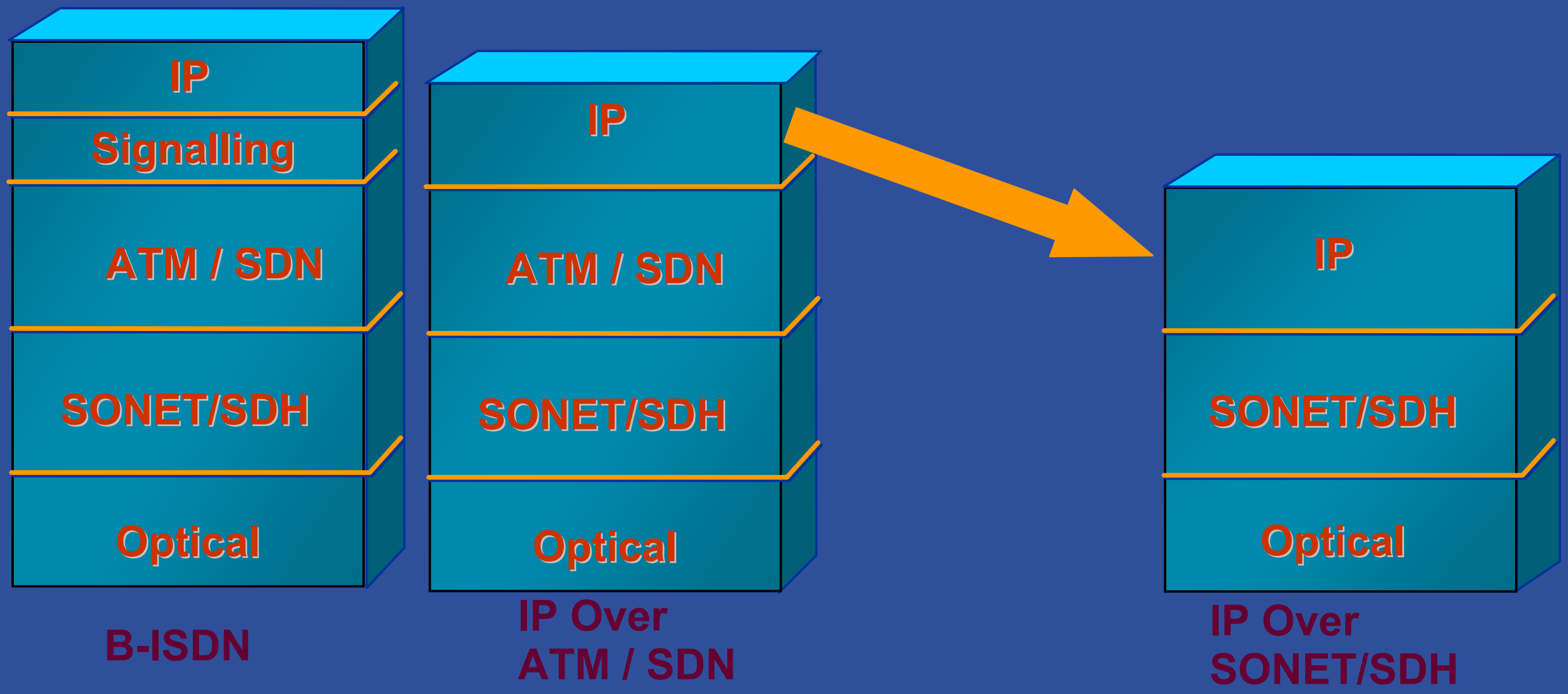



Asynchronous Transfer Mode
Misapplying Tomorrow's Technology Today™
IP
ATM
053
BLEND
PUREE
SCATTER
BILL
ATM? A TECHNICAL MISTAKE
A TARIFFING MECHANISM

Backbone Technologies

- PSTN Carrier Hierarchy
- ATM
  - Issues of IP performance,and complexity, and the need for a clear future path for increased speed at lower cost
  - 10^8 bits per second carriage speeds
- SDH / SONET

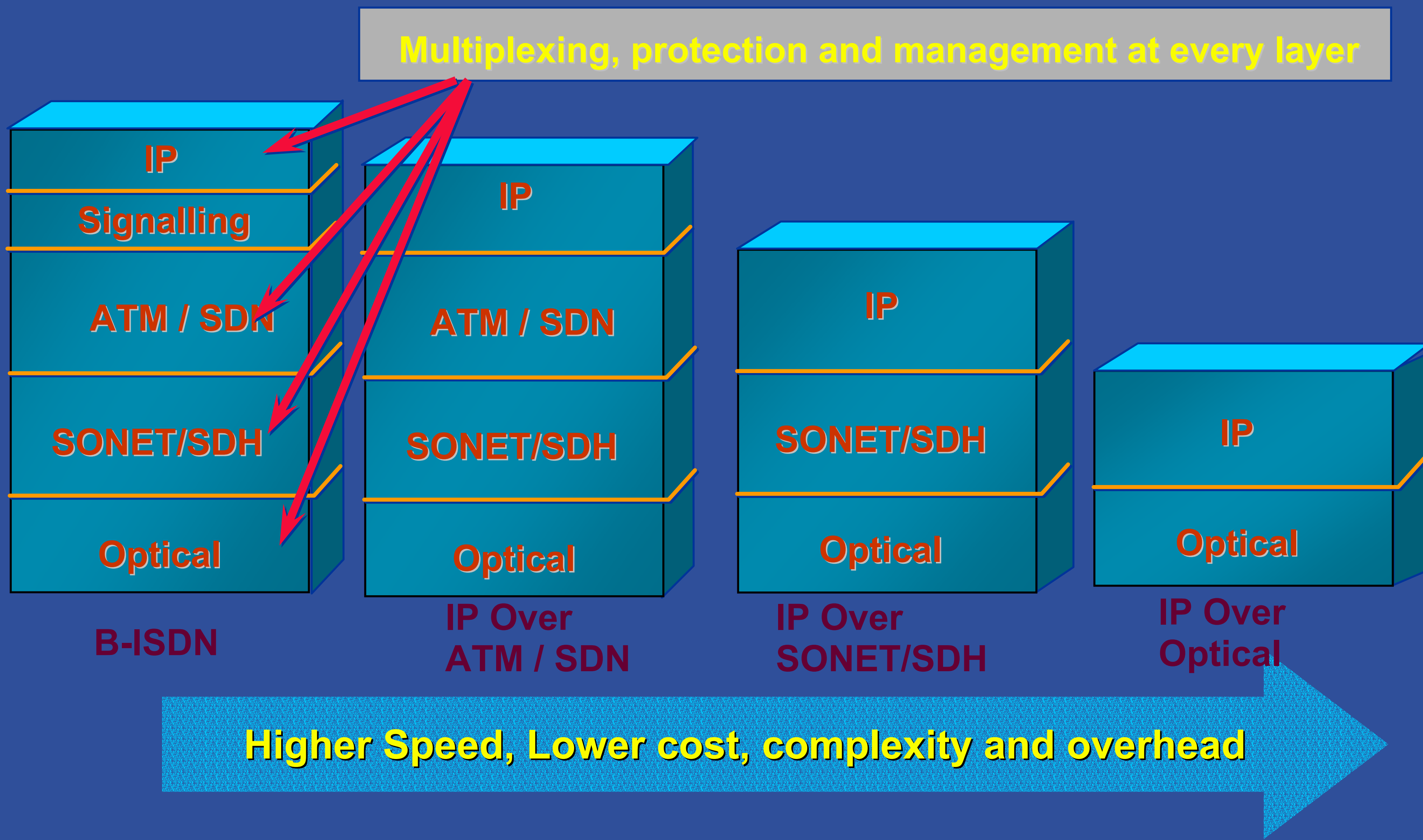
The Evolution of the IP Transport Stack



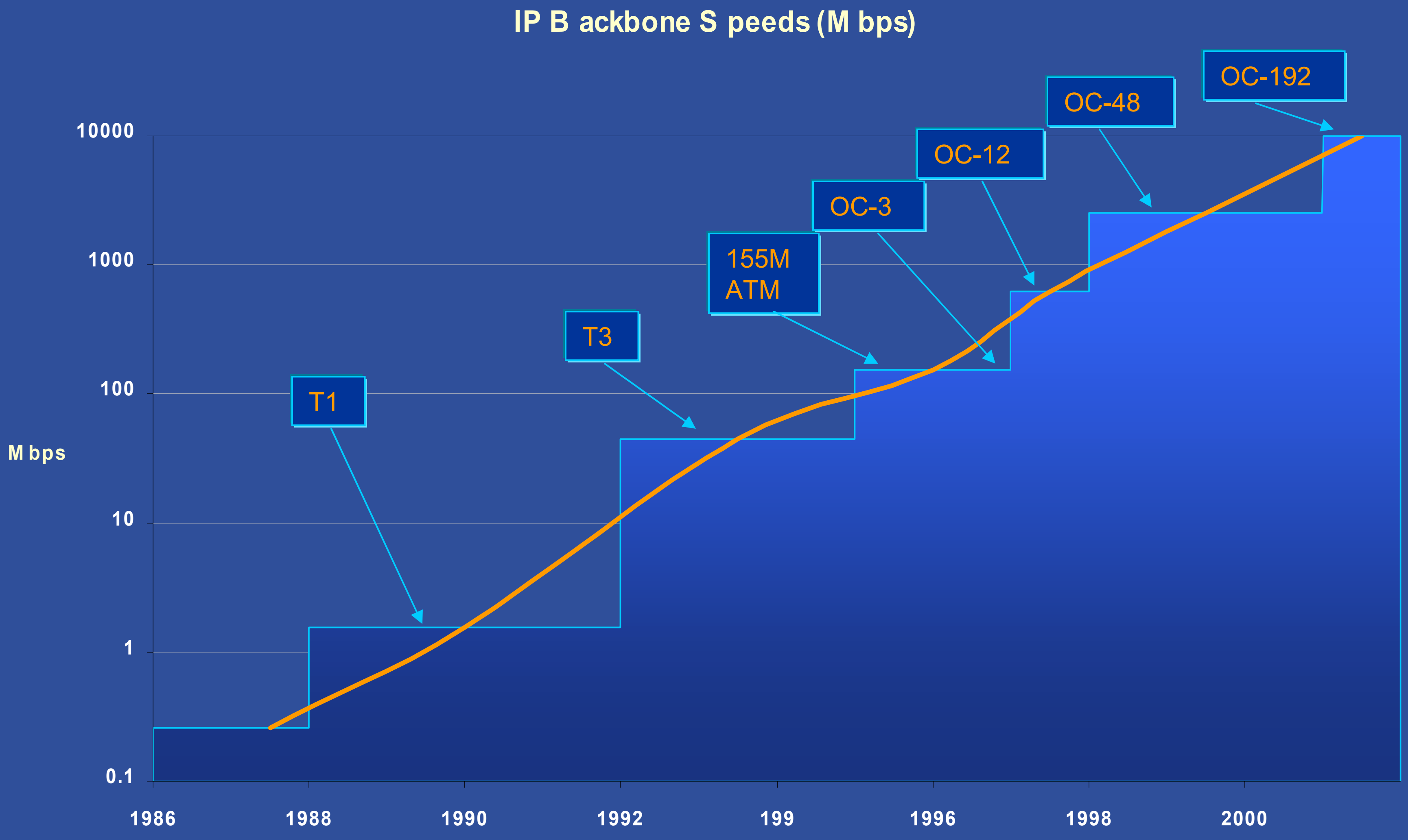
Backbone Technologies

- PSTN Carrier Hierarchy
- ATM
- SDH / SONET
  - 10^9 bits per second carriage speeds
- Unclocked packet over fibre?
  - 10 / 40 / 100 GigE?

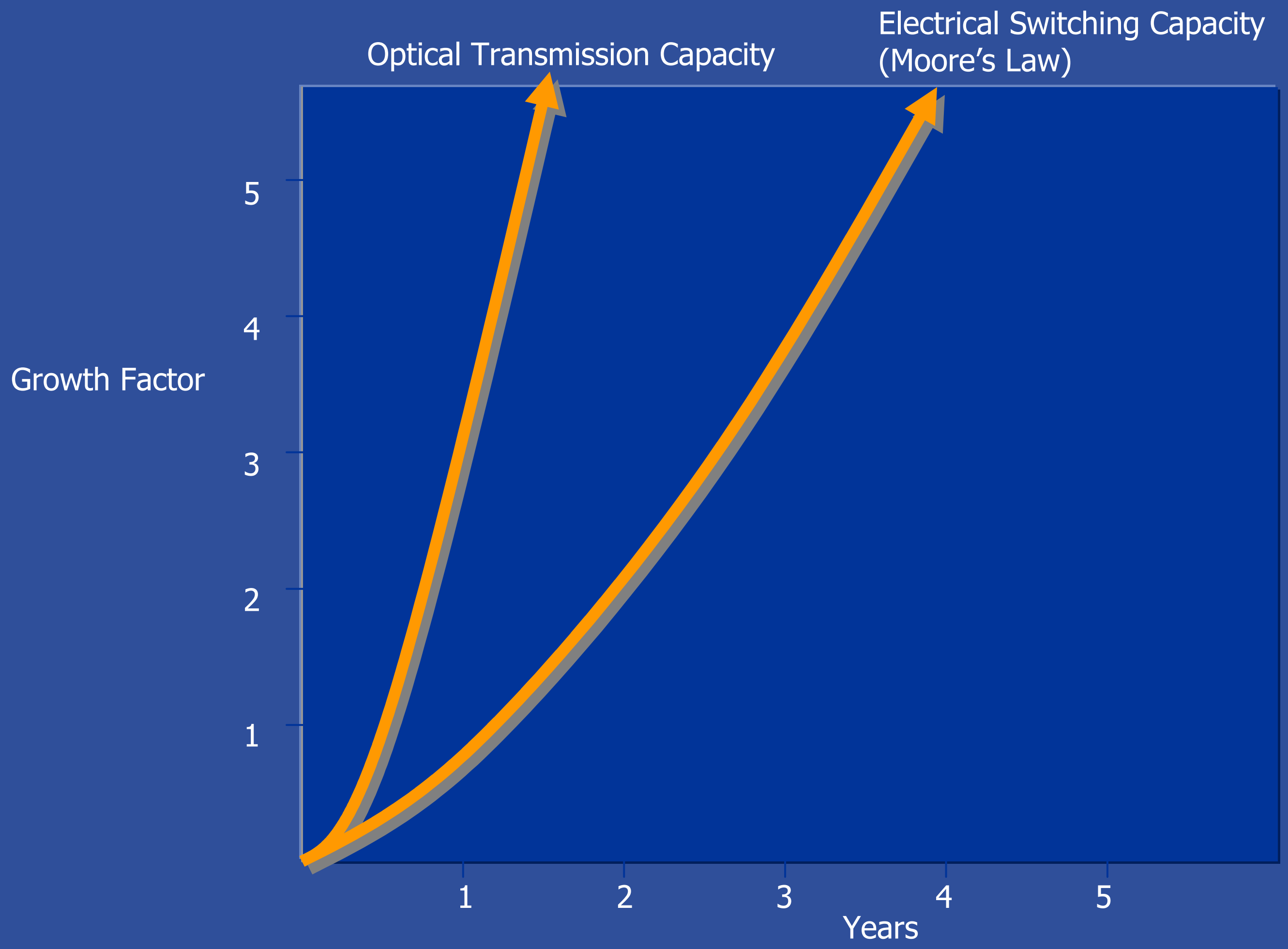
The Evolution of the IP Transport Stack



Internet Backbone Speeds



Recent Fibre Trends



Fibre speeds overwhelming Moore's law, implying that serial OEO switching architectures have a limited future

All-Optical switching systems appear to be necessary within 3 – 5 years

Physics Bites Back

- No confident expectation of cost effective 100G per-lambda equipment being deployed in the near future
- Current fibre capacity improvements being driven by increasing the number of coherent wavelengths per fibre system, not the bandwidth of each individual channel

IP Backbone Technology Directions

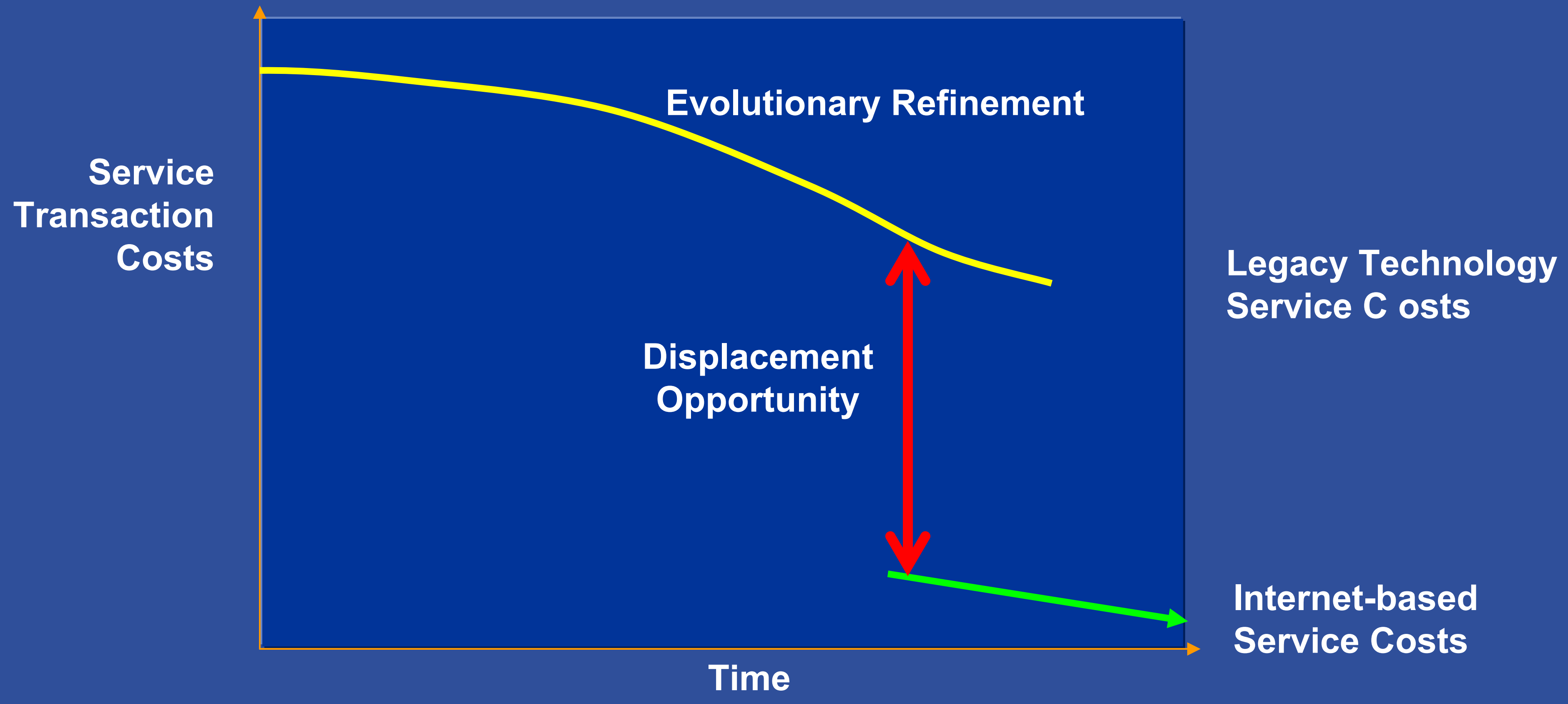
- POS / Ether Channel virtual circuit bonding
  - 10G – 40G concatenated systems
  - 3 – 4 year useful lifetime

- Lambda-agile optical switching systems
- GMPLS control systems
- MPLS-TE admission control systems
- Switching decisions pushed to the network edge (source routing, or virtual circuit models)
  - 100G – 10T systems
  - +3 years

IP Backbone Futures

- Assuming that we can make efficient use of all-IP abundant wavelength network:
  - The dramatic increases in fibre capacity are leading to long term sustained market oversupply in a number of long haul and last mile markets
  - Market oversupply typically leads to outcomes of price decline
- It appears this decline in basic transmission costs is already becoming apparent in the IP market

The Disruptive View of the Internet



Economics A01
as production costs decline…

- Implies a consequent drop in the retail market price
- The price drop exposes additional consumer markets through the inclusion of price-sensitive new services
- Rapidly exposed new market opportunities encourage agile high risk market entrants

Now lets relate this to the communications market…

- Local providers can substitute richer connectivity for parts of existing single upstream services
- Customers can multi-home across multiple providers to improve perceived resiliency
- Network hierarchies get replaced by network meshes interconnecting more entities

Is this evident today?

- How is this richer connectivity and associated richer non-aggregated policy environment expressed today?
  - More finer grained prefixes injected into the BGP routing system
  - Continuing increase in the number of Autonomous Systems in the routing space
  - Greater levels of multi-homing
- These trends are visible today in the Internet’s routing system

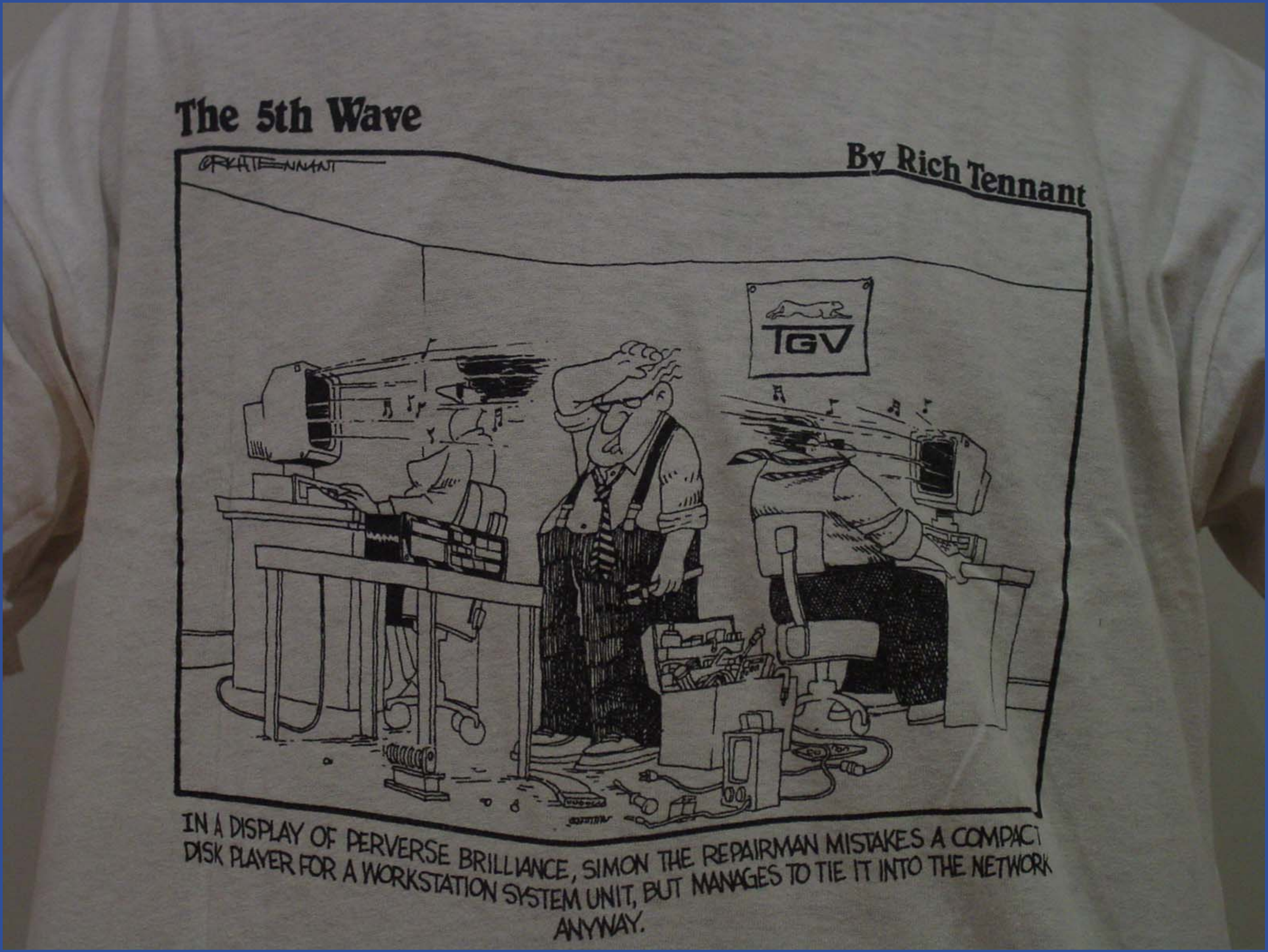
Backbone Futures

- Backbones transmission networks are getting faster
  - Not by 'larger channels'
  - By more available fibre channels
  - By a denser mesh of connectivity with more complex topologies
- This requires
  - More complex switches
  - Faster switching capacities
  - More capable routing protocols

Edge Systems

Edge Systems

- The Internet is moving beyond screens, keyboards and the web
- A world of devices that embed processing and communications capabilities inside the device


The 5th Wave
By Rich Tennant
TGV
IN A DISPLAY OF PERVERSE BRILLIANCE, SIMON THE REPAIRMAN MISTAKES A COMPACT DISK PLAYER FOR A WORKSTATION SYSTEM UNIT, BUT MANAGES TO TIE IT INTO THE NETWORK ANYWAY.

Edge Systems

- With the shift towards a device-based Internet, the next question is how can we place these billions of devices into a single coherent network?
- What changes in the network architecture are implied by this shift?

Scaling the Network

- Billions of devices calls for billions of network addresses
- Billions of mobile devices calls for a more sophisticated view of the difference between identity, location and path

Scaling the Network - The IPv4 View

- Use DHCP to undertake short term address recycling
- Use NATs to associate clients with temporary (32 + 16) bit aliases
- Use IP encapsulation to use the outer IP address for location and the inner IP address for identity
- And just add massive amounts of middleware
  - Use helper agents to support server-side initiated transactions behind NATS
  - Use application level gateways to drive applications across disparate network domains
  - Use walled gardens of functionality to isolate services to particular network sub-domains

Scaling the Network

- Or change the base protocol

Scaling the Network
- The IPv6 View

- Extend the address space so as to be able to uniquely address every connected device at the IP level
- Remove the distinction between clients and servers
- Use an internal 64/64 bit split to contain location and identity address components
- Remove middleware and use clear end-to-end application design principles
- Provide a simple base to support complex service-peer networking services
  - End-to-end security, mobility, service-based e2e QoS, zeroconf, etc

How big are we talking here?


FIRST ANNIVERSARY
TOASTER NET '91
FLYING TOASTERS Artwork © 1991 BERKELEY SYSTEMS, INC.
From AFTER DARK, the Ultimate Screen Saver. Reproduced under
under license agreement by EPILOGUE Technology Corp.

IP network requirements: Scaling by numbers

- Number of distinct devices
  - $O(10^{10})$
- Number of network transactions
  - $O(10^{11}/sec)$
- A range of transaction characteristics
  - $10 – 10^{10}$ bytes per transaction
- End-to-end available bandwidth
  - $10^3 – 10^{10}$ bits /sec
- End-to-end latency
  - $10^{-6} – 10^1$ sec

Scale Objectives

- Currently, the IP network with IPv4 encompasses a scale factor of 10^6
- Near-term scale factors of deployment of
  - Personal mobile services
  - Embedded utility services

  will lift this to a scale factor of around 10^{10}
- How can we scale the existing architecture by a factor of 10,000 and improve the cost efficiency of the base service by a unit cost factor of at least 1,000 at the same time?

Performance and Service Quality

Scaling Performance

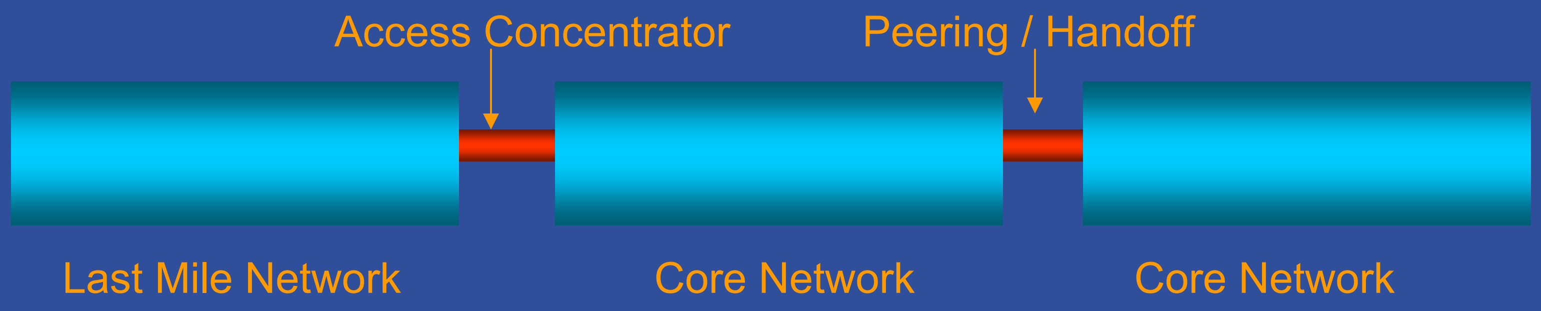
- Application performance lags other aspects of the network
- Network QoS is premature. The order of tasks appears to be:
  1. Correct poor last mile performance
  2. Address end-to-end capacity needs
  3. Correct poor TCP implementations
  4. Remove non-congestion packet loss events
  5. Then look at residual network QoS requirements

1. Poor Last Mile Performance

- Physical plant
  - Fibre last mile deployments
  - DSL last mile
- Access network deployment models
  - What’s the priority:
    - Low cost to the consumer?
    - Equal access for multiple providers?
    - Maximize per-user bandwidth?
    - Maximize contestable bandwidth?
    - Maximize financial return to the investor / operator?

2. End-to-End Capacity

- Network capacity is not uniformly provisioned
- Congestion is a localized condition



3. TCP Performance

- 95% of all traffic uses TCP transport
- 70% of all TCP volume is passed in a small number of long held high volume sessions (heavy tail distribution)
- Most TCP implementations are poorly implemented or poorly tuned
  - Correct tuning offer 300% performance improvements to high volume high speed transactions (web100's wizard margin)

4. Packet Loss

- TCP Performance
  BW = (MSS / latency) * (0.7 / SQRT(loss rate))
- Improve performance by
  - Decrease latency (speed of light issues)
  - Reduce loss rate
  - Increase packet size
- 75Mbps at 80ms with 1472 MSS requires 10^{-7} loss rate
  - That's a very challenging number
- Cellular Wireless has a 10^{-4} loss rate
  - High performance wireless systems may require application level gateways for sustained performance

5. Network QoS

- Current backbone networks exhibit low jitter and low packet loss levels due to low loading levels
  - Small margin of opportunity for QoS measures in the network
- Improved edge performance may increase pressure on backbone networks
  - Which in turn may provide for greater opportunity for network QoS
  - Or simply call for better engineered applications

Performance

- Is performance a case of better application engineering with more accurate adaptation to the operating characteristics of the network?
- Can active queue techniques in the switching interior of the network create defined service outcomes efficiently?
- How much of the characteristics of interaction between applications and network do we understand today?

Trust

Just unplug?

U.S. GOVERNMENT SEEKS INPUT TO BUILD ITS OWN NET

The federal government is considering creating its own Internet. Called GovNet, the proposed network would provide secure government communications. Spearheading the effort is Richard Clarke, special advisor to President Bush for cyberspace security. With the help of the General Services Administration (GSA), Clarke is collecting information from the U.S. telecom sector about creating an exclusive telecom network. The GSA Web site features a Request for Information (RFI) on the project. GovNet is intended to be a private, IP-based voice and data network with no outside commercial or public links, the GSA said. It must also be impenetrable to the existing Internet, viruses, and interruptions, according to the agency. GovNet should be able to support video as well as critical government functions, according to the RFI.

(InfoWorld.com, 11 October 2001)

Trust

Every network incorporates some degree of trust

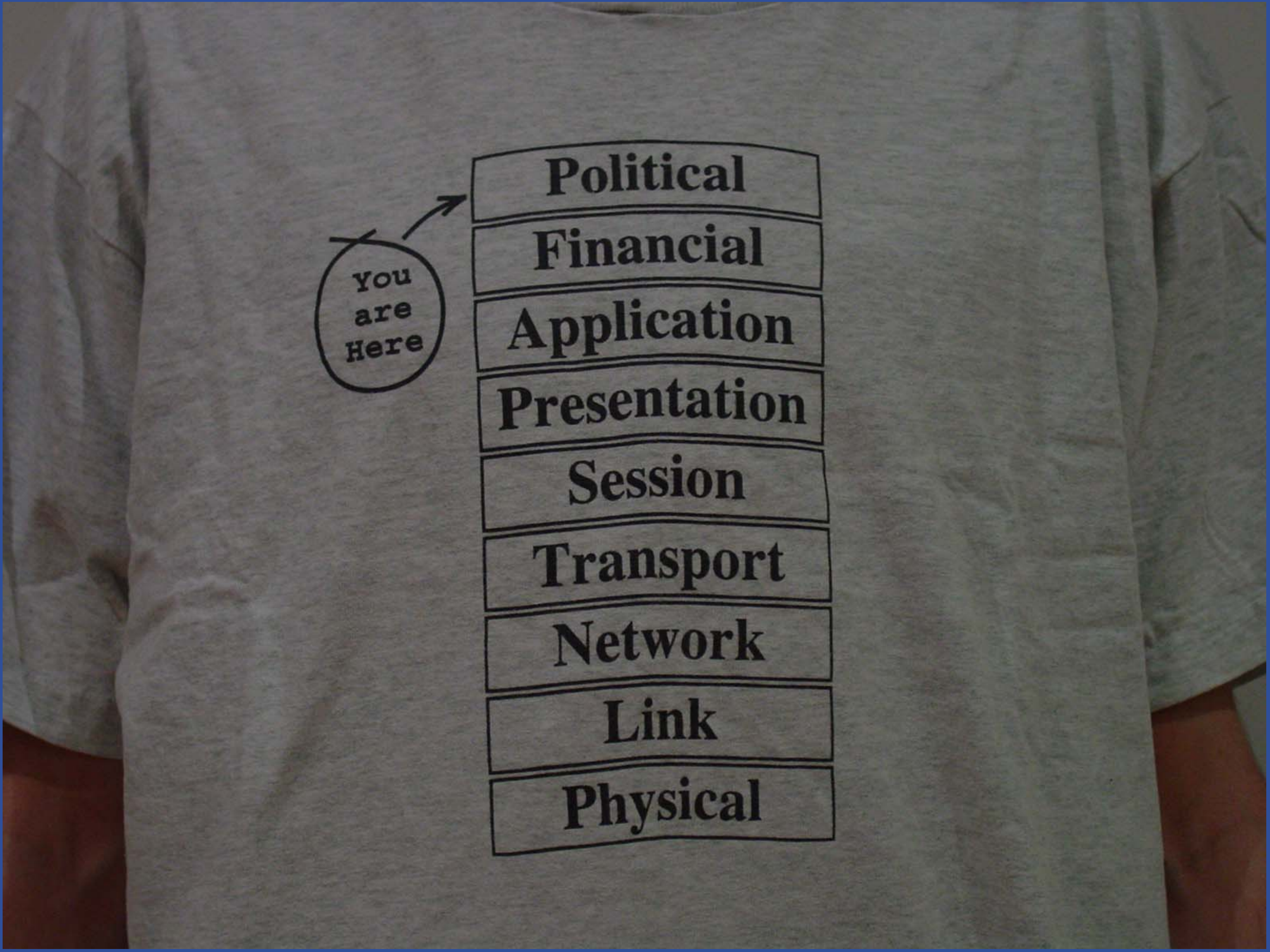
The thin-core thick-edge service model of the Internet places heavy reliance on *edge-to-edge* trust

This reliance on a distributed *edge-to-edge* trust is visible in

- IP address assignments
- IP routing system
- DNS integrity
- End-to-End packet delivery
- Application integrity
- Mobility
- Network resource management

Scaling Trust

- Are the solutions to a high distributed trust model a case of more widespread use of various encryption and authentication technologies?
- Is deployment of such technologies accepted by <u>all</u> interested parties?


You are Here
Political
Financial
Application
Presentation
Session
Transport
Network
Link
Physical

Improving Trust Models

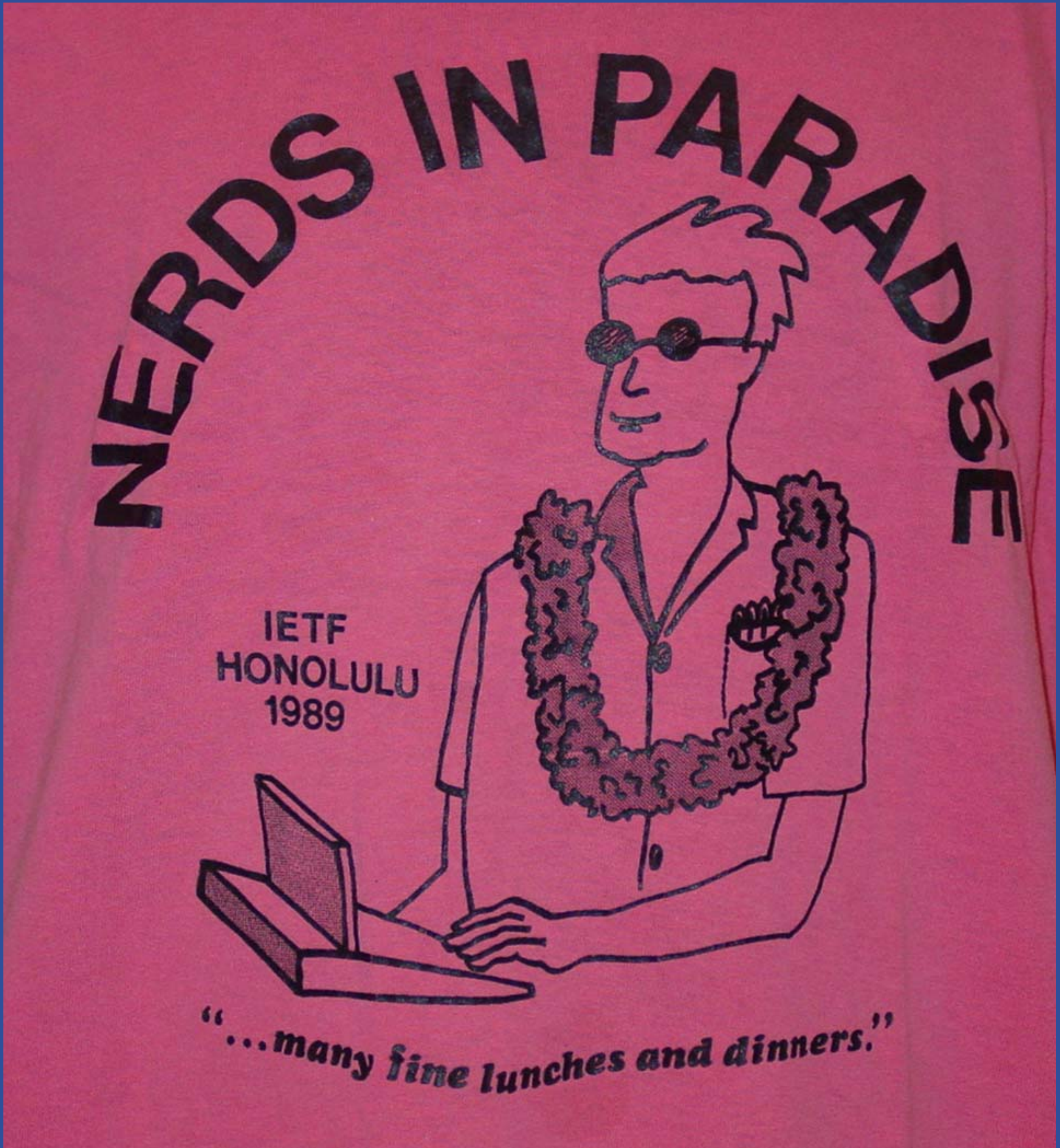
- Many of the component technologies are available for use today
- But a comprehensive supporting framework of trusted third parties and reference points remains elusive

The Outlook

- The Internet's major contribution has been cheap services:
  - Strong assumption set about cooperative behavior and mutual trust
  - Strong assumption set regarding simple networks and edge-based 'value added' services
- Scaling the Internet must also continue to reduce the cost of the Internet
  - Its likely that simple, short term evolutionary steps will continue be favoured over more complex large-scale changes to the network or application models

There is much to do

- And it sounds like fun!


NERDS IN PARADISE
IETF
HONOLULU
1989
"...many fine lunches and dinners."