How in the H*ll Do You Lose A Layer?!!

John Day

Dublin 2014

OSI only had a Network Layer, but
the Internet has an Internet Layer!

- Noel Chiappa, 1999

Preamble

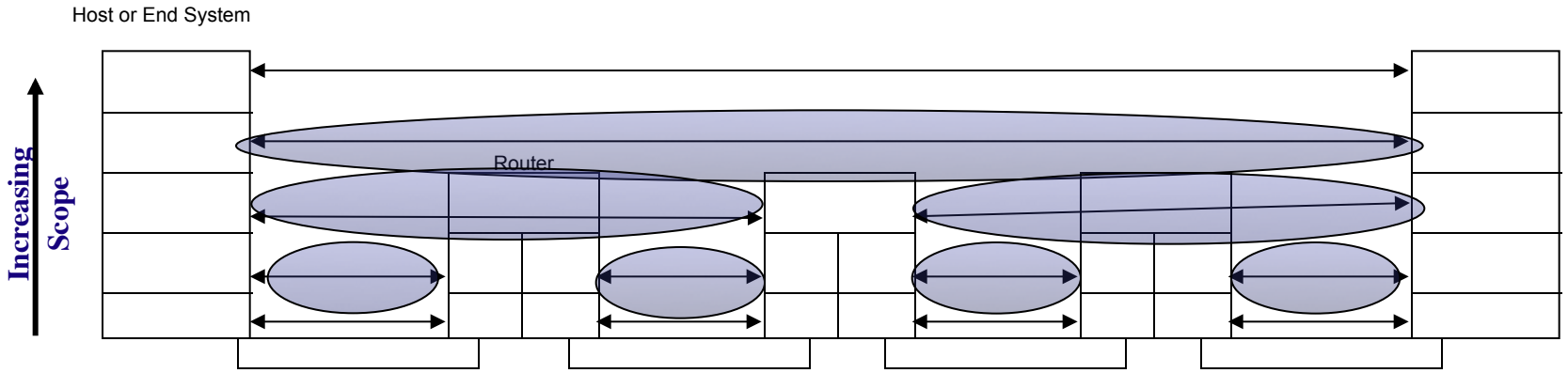
- A Couple of Remarks on the Nature of Layering and a Quiz:
- The advent of packet switching required much more complex software than heretofore, and so the concept of layers was brought in from operating systems.
- In operating systems, layers were seemingly a convenience, one design choice.
- Why Do We Use Layers in Network Architecture?

- In networks, they are a *necessity*.

The (really) Important Thing about Layers

(From first lecture of my intro to networks course)

- Networks have loci of distributed shared state with different scopes
- At the very least, differences of scope require different layers.
- It is *this* property that makes the earlier telephony or datacomm "beads-on-a-string" model *untenable*.
  - – Or any other proposal that does not accommodate scope.
- This has always been understood.



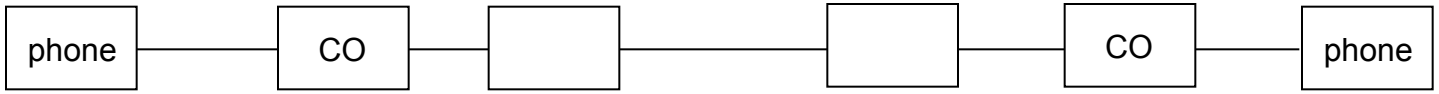
Refining the Concept of Layer

- The Necessary and (usually) Sufficient Condition for a Layer is that there are loci of shared state of different scope.
    - For Operating Systems and Networks, layers are ranges of resource allocation.
- If there are layers of the same scope, their functions must be completely independent.
- Dykstra wasn't wrong: Functions do not repeat

  . . . in layers of the same scope.
    - The hardware at the time was so constrained he could only see one scope.
- If there is partitioning within the layer, it will generally be orthogonal to the attributes that impose layers.

  – Do All Layered Models Follow These Rules? Probably not. At least Resource Allocation models. Perhaps all those that exhibit scope.

The Beads on A String Model



- The Nature of their early technology led the Phone Companies to Adopt what could be called, a "Beads-on-a-String" architecture.
  - Deterministic, Hierarchical, master/slave.
- Perfectly reasonable for what they had.
- The model not only organized the work,
  - But was also used to define markets: Who got sell what.
  - This was what was taught in most data comm courses prior to the 1980s.
    - And for some, in a fundamental sense, never left.

Packet Switching

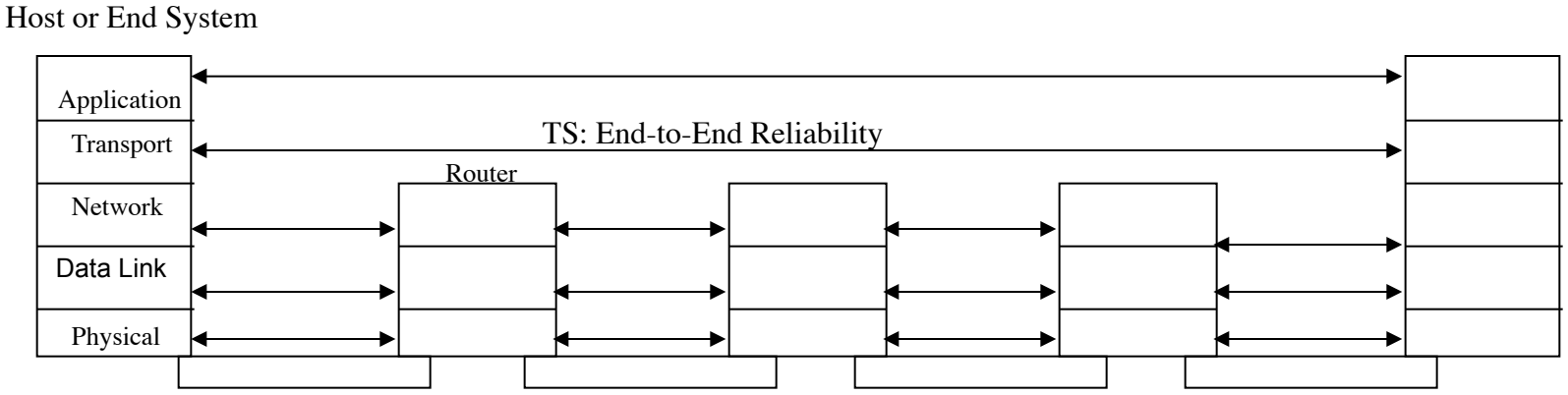
- In the early 1960s, Paul Baran at The Rand Corporation writes a series of reports investigating the networking requirements for the DoD.
  - Donald Davies at NPL in the UK had the same idea
- He finds that the requirements for data are very different than those for voice.
  - Data is bursty. Voice is continuous.
  - Data connections are short. Voice connections have long durations.
- Data would be sent in individual packets, rather than as continuous stream, on a path through the network.
- Packet switching is born and
- By the late 1960s, the Advance Research Projects Agency decides building one would reduce the cost of research and so we have the ARPANET.

But was Packet Switching a Major Breakthrough?

- Strange as it may seem, it depends.
    - During this period many things are age dependent.
- If your formative years had occurred prior to the mid-60s (pre-boomer), your model of communication was defined by telephony.
  - Then this is revolutionary.
- If you are younger (boomer), your model is determined by computers.
  - Data is in buffers, How do you do communications:
    - Pick up a buffer and send it.
  - What could be more obvious!
    - That it was independently invented (and probably more than twice) supports that.
- But there was a more radical idea coming!

The Cyclades Architecture

(1972)



Cigale Subnet

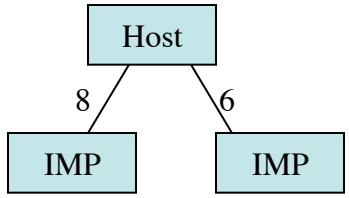
- Transport Service provides end-to-end reliability.
- In that case, hop-by-hop reliability does not have to be perfect.
  - Need only be sufficiently reliable to make end-to-end cost effective.
- Over a connectionless datagram network, Cigale
  - Yields a simpler, more effective and robust data network.
- CYCLADES brings in the traditional layering from operating systems.
- This represents a new model, in fact, a new paradigm completely at odds with the beads-on-a-string model.

The New Model Had 4 Characteristics

- It was a *peer network* of communicating equals not a hierarchical network connecting a mainframe master with terminal slaves.
- The approach required coordinating *distributed shared state at different scopes*, which were treated as black boxes. This lead to the concept of layers being adopted from operating systems and
- There was a shift from largely deterministic to *non-deterministic* approaches, not just with datagrams in networks, but also with interrupt driven, as opposed to polled, operating systems, and physical media like Ethernet, and last but far from least,
- This was a computing model, *a distributed computing model*, not a Telecom or Data comm model. The network was the infrastructure of a computing facility.
- These sound innocuous enough. They weren't. Not by a long shot!

1972 Was an Interesting Year

- Tinker AFB joined the ‘Net exposing the multihoming problem.



- The ARPANET had its coming out at ICCC ‘72.
- As fallout from ICCC 72, the research networks decided it would be good to form an International Network Working Group.
  - ARPANET, NPL, CYCLADES, and other researchers
  - Chartered as IFIP WG6.1 very soon after

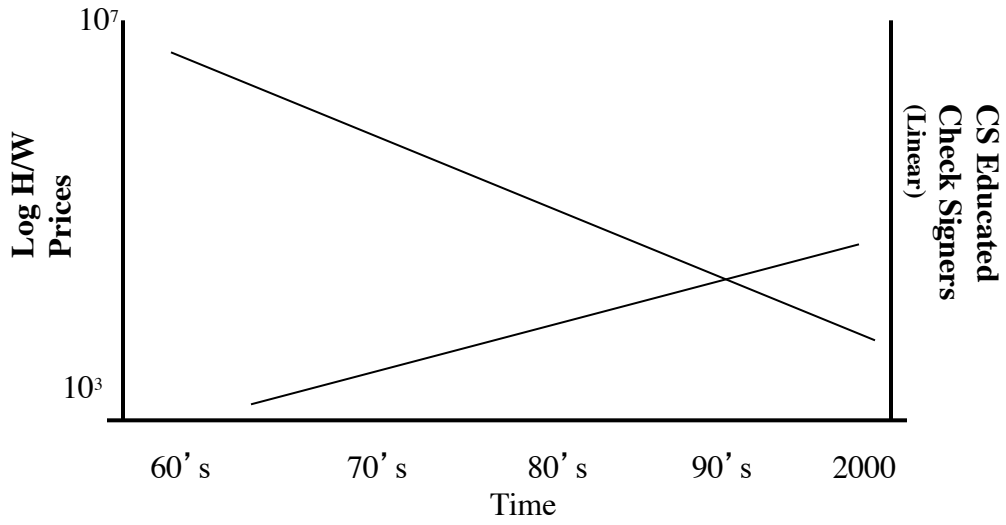
- Major project was an Internetwork Transport Protocol.
  - Also a virtual terminal protocol
  - And work on formal description techniques

A Nasty Brawl Was Shaping Up

The Phone Companies
Against
the Computer Companies
and
Everyone against IBM

IBM had Two Problems

Computing and Memory Prices were headed South . . . Fast.
Computing Power and Capacity were headed North . . . Fast.
By the late 70s, it was clear that IBM's days as the dominant computer maker were numbered

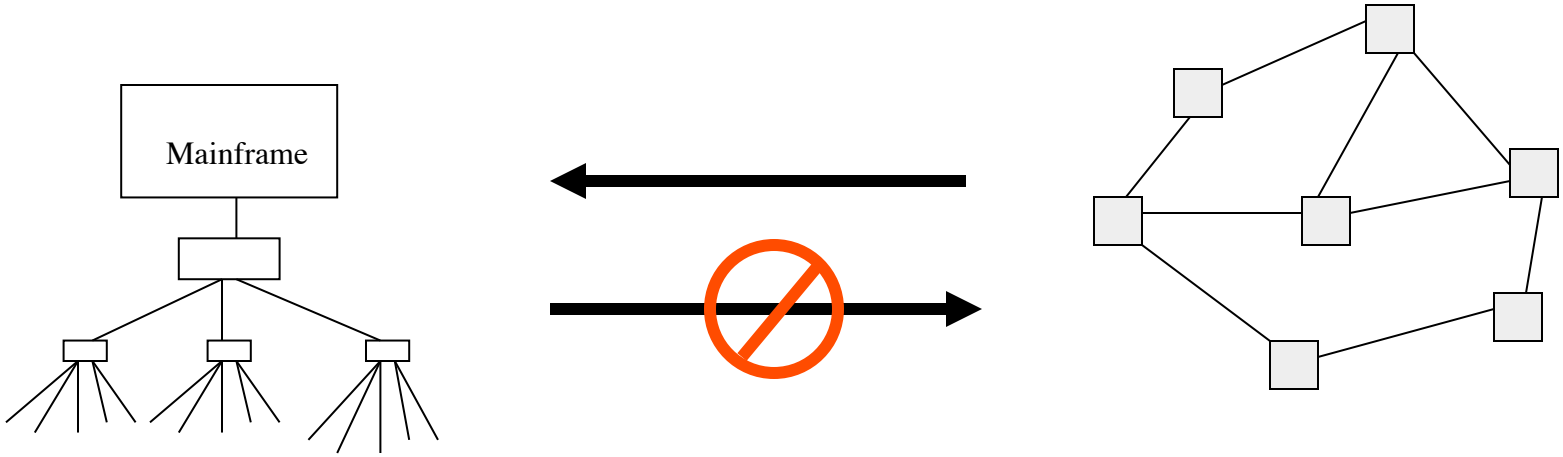


And if that weren't enough.

In Networking

IBM Found Itself at a Dead-End

You can always make a peer architecture hierarchical
But you can't go the other way.

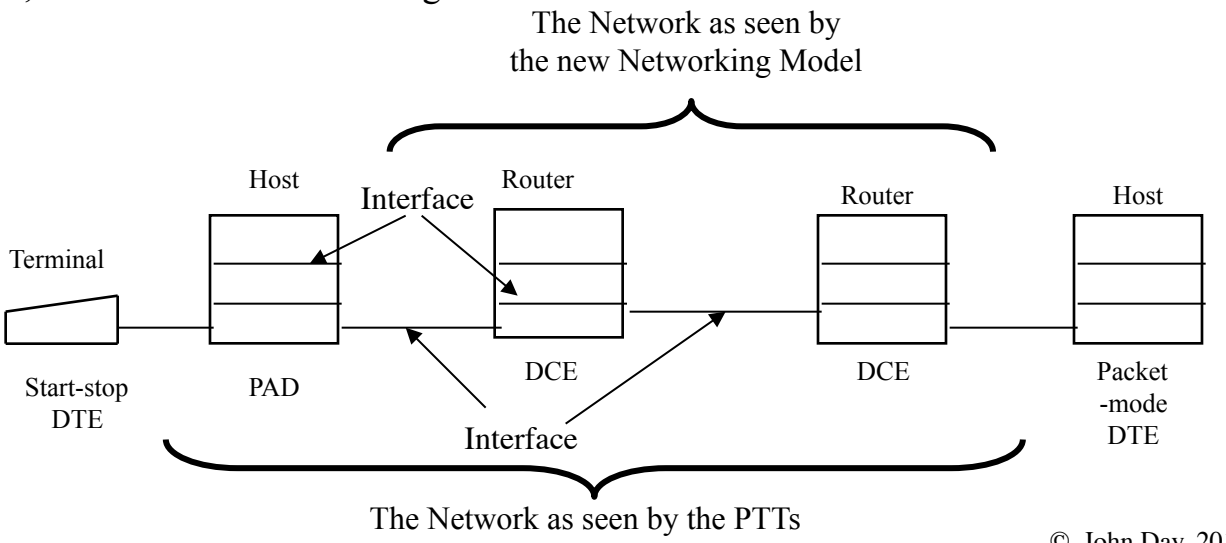


But IBM and the PTTs had carefully stayed out of each other's turf.

Had IBM made SNA a peer network and subset it for the 70s hierarchical market, the Internet would have been nothing but an interesting research project.

TPC:* The Beads-on-a-String Model

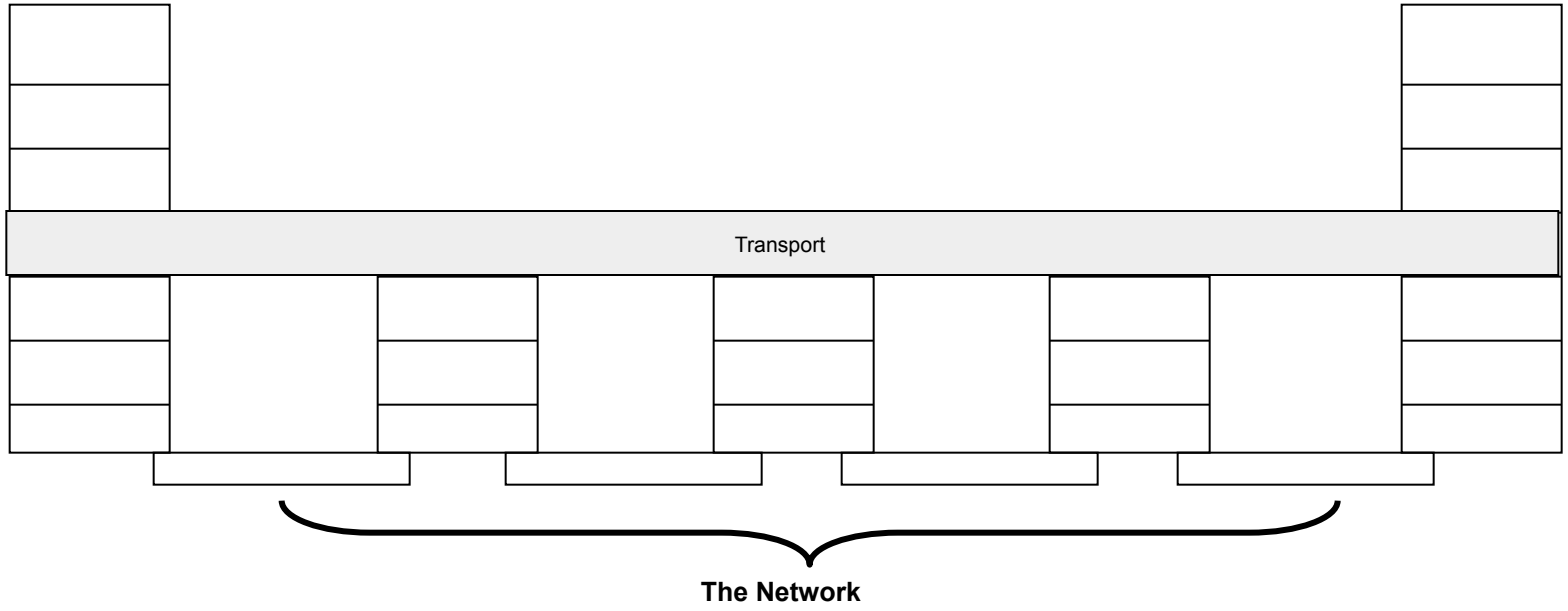
- Meanwhile TPC continues with what it is familiar with.
  - Emulating the phone system in computers
  - Who cares about this academic connectionless stuff? We have *real* networks to build.
  - How do you charge for usage in a best-effort service?
- Asymmetrical/Connections/Deterministic
  - And a tendency toward hierarchy
- This Model Can not Represent Scope.
- Purpose of the architecture is to define who owns what boxes (protect a monopoly).
  - If you hear, X is *in* the network and X isn't involved with moving bits or managing moving bits, then it is beads-on-a-string.



* "The Phone Company" with a nod to "**The President's Analyst**"

While the New Model Made Perfect Sense to Computing, It Was a Threat to Phone Companies.

- Transport Seals Off the Lower Layers from Applications.
  - Making the Network a Commodity, with very little possibility for value-add.
- TPC counters that Transport Layers are unnecessary, *their* networks are reliable.



And they have their head in the sand, "Data will never exceed voice traffic"

Meanwhile Back at INWG

There Were Two Proposals

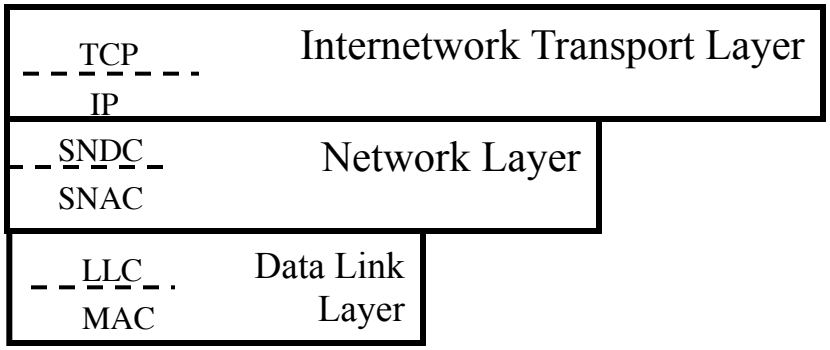
- INWG 39 based on the early TCP and
- INWG 61 based on CYCLADES TS.
- And a healthy debate, see Alex McKenzie, "INWG and the Conception of the Internet: An Eyewitness Account" IEEE Annals of the History of Computing, 2011.
- Two sticking points
  - How fragmentation should work
  - Whether the data flow was an undifferentiated stream or maintained the integrity of the units sent (letter).
- These were not major differences compared to the forces bearing down on them.

After a Hot Debate

- A Synthesis was proposed: INWG 96
- There was a vote in 1976, which approved INWG 96.
- As Alex says, NPL and CYCLADES immediately said they would convert to INWG 96; EIN said it would deploy only INWG 96.
- "But we were all shocked and amazed when Bob Kahn announced that DARPA researchers were too close to completing implementation of the updated INWG 39 protocol to incur the expense of switching to another design. As events proved, Kahn was wrong (or had other motives); the final TCP/IP specification was written in 1980 after at least four revisions."
  - Neither was right. The real breakthrough came two years later.
- But the differences weren't the most interesting thing about this event.

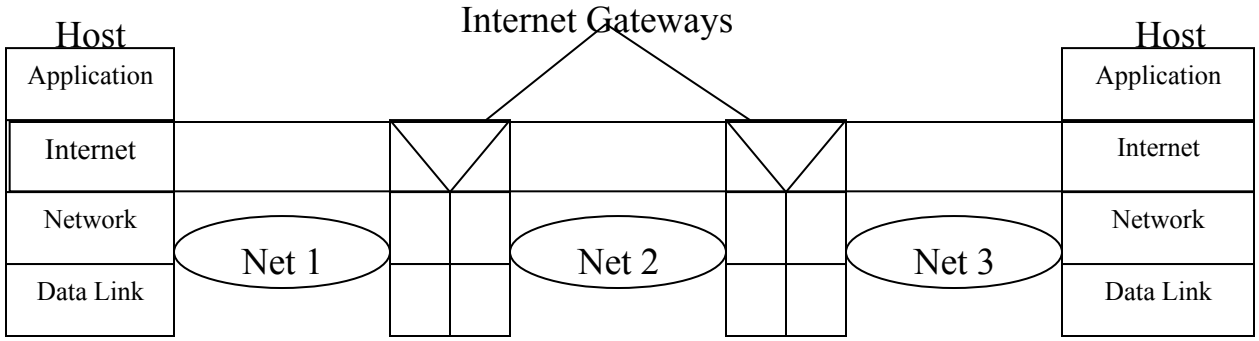
The *Similarity* Among all 3 Is Much More Interesting

- This is *before* IP was separated from TCP. All 3 of the Proposed *Transport* Protocols carried addresses.
- This means that the Architecture that INWG was working to was:



- Three Layers of Different Scope each with Addresses.
- If this does not hit you like a ton of bricks, you haven't been paying attention.
- This is *NOT* the architecture we have.

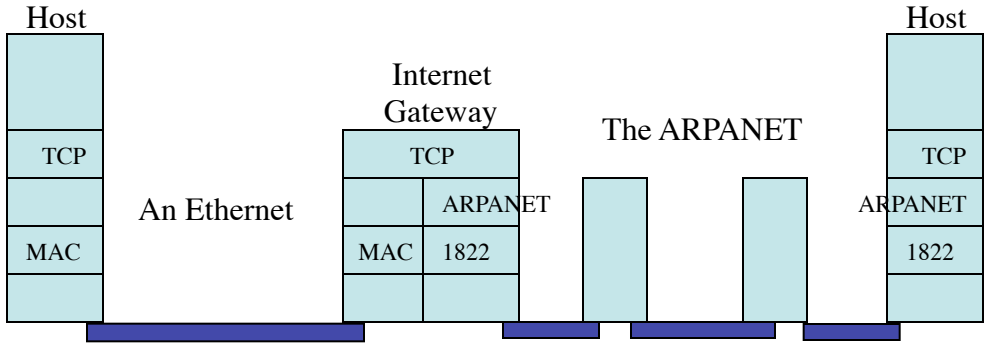
INWG's Internet Model



- An Internet Layer addressed Hosts and Internet Gateways.
- Several Network Layers of different scope, possibly different technology, addressing hosts on that network and that network's routers and its gateways.
  - Inter-domain routing at the Internet Layer; Intra-Domain routing at the Network Layer.
- Data Link Layer smallest scope with addresses for the devices (hosts or routers) on segment it connects
- The Internet LOST A LAYER!!

How Did They Lose a Layer?

- To Hazard a Guess: (This is subtle so pay close attention)
  - A Case of Sorcerer's Apprentices (Thought they knew more than they did)
  - The Internet was a DoD project with the ARPANET at its center
    - Built and operated by BBN. Only BBN made IMPs
      - In a sense, BBN was their phone company, e.g. provider.
  - The initial growth was LANs at the Edge connected by
    - Internet Gateways: Ethernet on one side; BBN 1822 or X.25 on the other.
    - The ARPANET had no "peers" in this environment.



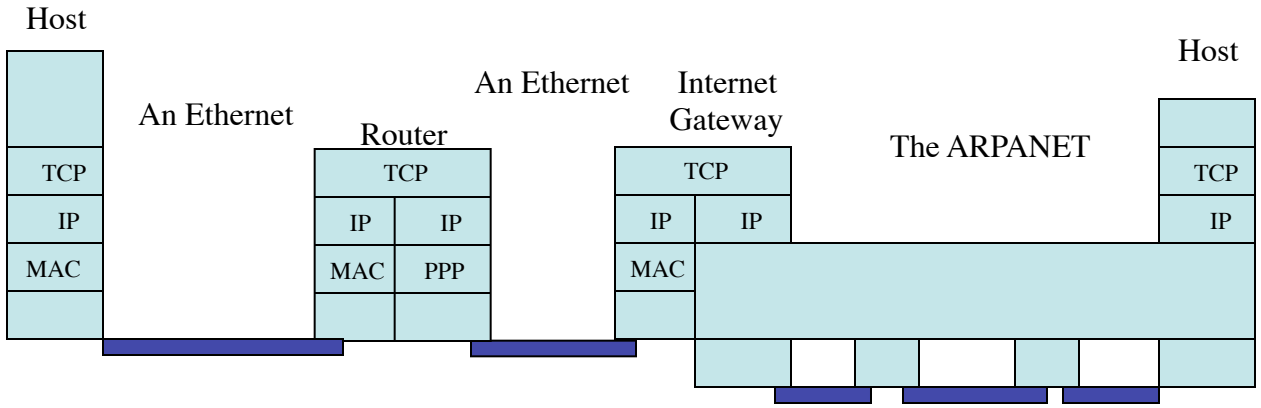
The View About 1976 Before IP is Split from TCP

Now we split IP from TCP

Remember, only one or two people involved
in this were also involved in INWG

How Did They Lose A Layer? II

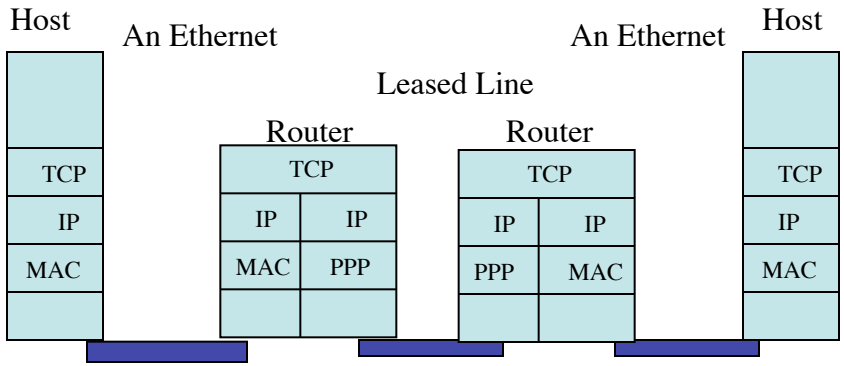
After IP is Split Off



The View After 1976 Now IP is Split from TCP

- But the ARPANET is a black box. Only BBN can see inside it.
- So to everyone else it looks like just another LAN.
  - They start to think that way.
- Most of the new entries are workstations on LANs being connected together over short and long distances (with leased lines).
- Which leads to a picture that looks like:

How Did They Lose a Layer? III



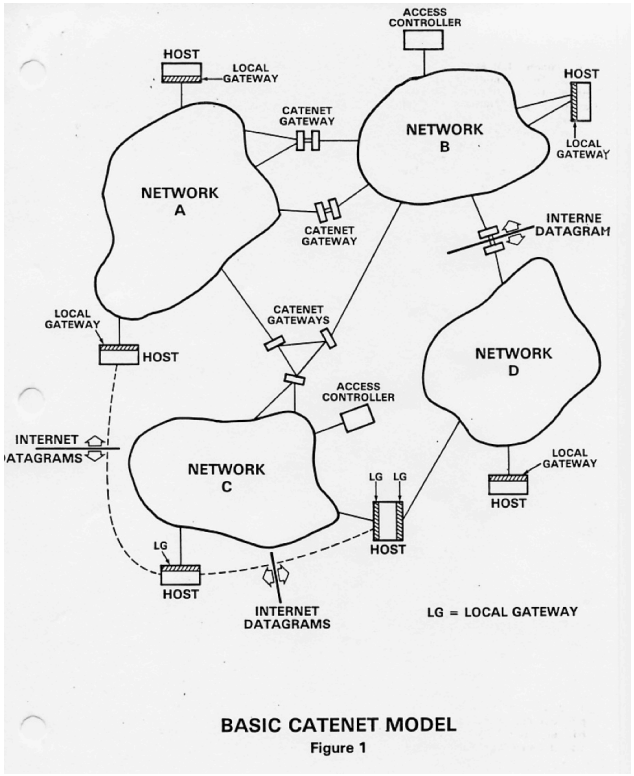
- And there are lots of them connecting to each other!
  - The ARPANET is becoming less and less important.
- Voilà! Did you see it disappear?
  - This is not an Internet! It is a beads-on-a-string *Network*!
    - Just like the ITU!!
    - We have Met the Enemy and He is Us! - Walt Kelly 1970
  - No Internet Gateways, only Routers. The term disappears in the early 80s
- Separating IP from TCP; not understanding the importance of scope; the misconception of one protocol, one layer; just doing the next thing with no plan; all contributed to being an Internet in name only.

But They Had Help: IEN #48

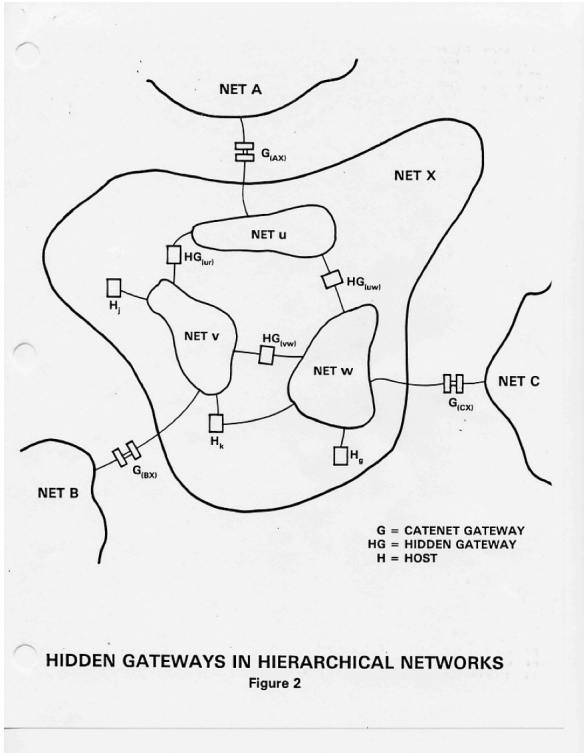
- "The Catenet Model of for Internetworking" by Vint Cerf July 1978
    - An important document for the Internet
- Which Says:
    - "The same host may have several addresses depending on how many nets it is connected to or how many lines connect it to a particular network."
        - This is 6 years after Tinker AFB, and addresses name interfaces, not nodes.
- As well as,
    - "To simplify the translation from internet address to local address, it is advantageous, if possible, to simply concatenate a network identifier with the local "host: addresses to create an internet address."
        - Thus making the addresses path dependent.
- And the Figures are Interesting!

Thanks to Vint Cerf for pointing me at this.

Absence of a Picture is Worth 1000 Words



BASIC CATENET MODEL
Figure 1



HIDDEN GATEWAYS IN HIERARCHICAL NETWORKS
Figure 2

– The diagrams are beads-on-a-string. No diagram of the layers. Almost no mention of layers at all. Layers are inferred. The elements are there, but no indication the implications are understood. Lot of talk of "gateway-halves." It does mention flow (congestion) control in the *networks*, i.e. the network layer.

So What Layer Did They Lose?

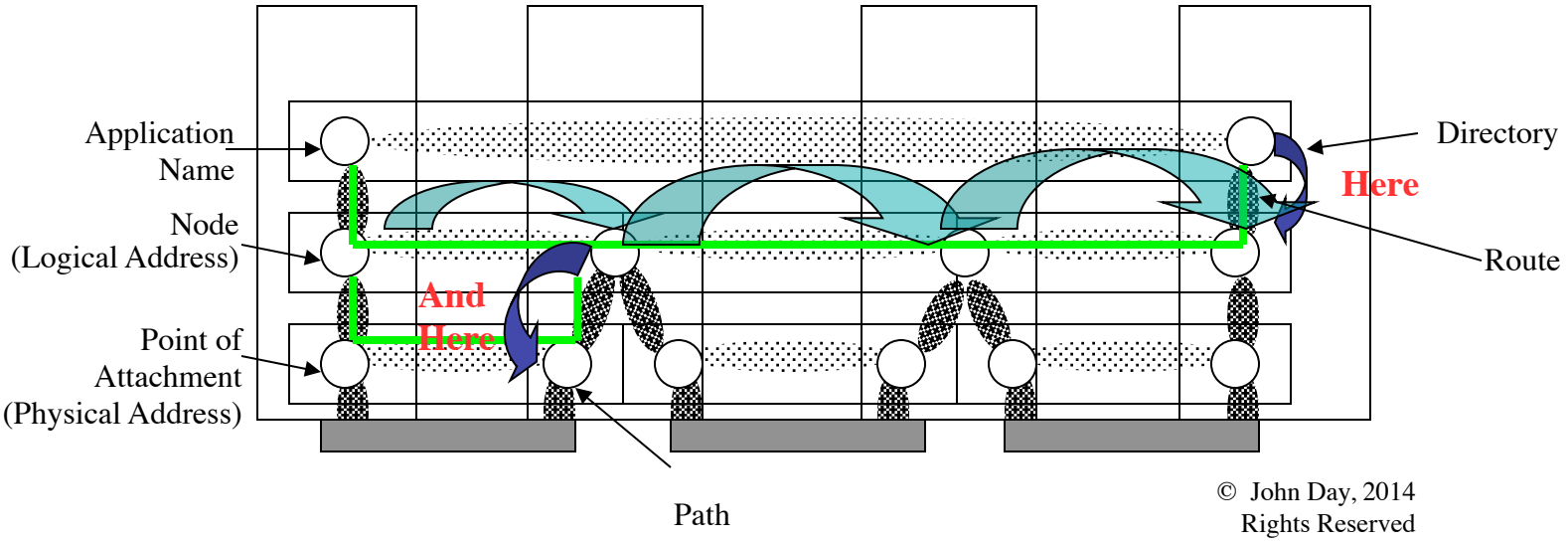
- It is not obvious.
- At first glance, one might say the Network Layer.
  - The Protocol is called IP after all!
  - Removing the ARPANET, "removed" the Network Layer,
  - Everything just dropped down.
- But the IP Address names the Interface, something in the layer below, just like ARPANET addresses did!
  - At best, IP names a network entity of some sort, at worst, a data link entity
  - Actions speak louder than words
- We must conclude that, . . .

They Lost the Internet Layer!!!

Wait A Minute!

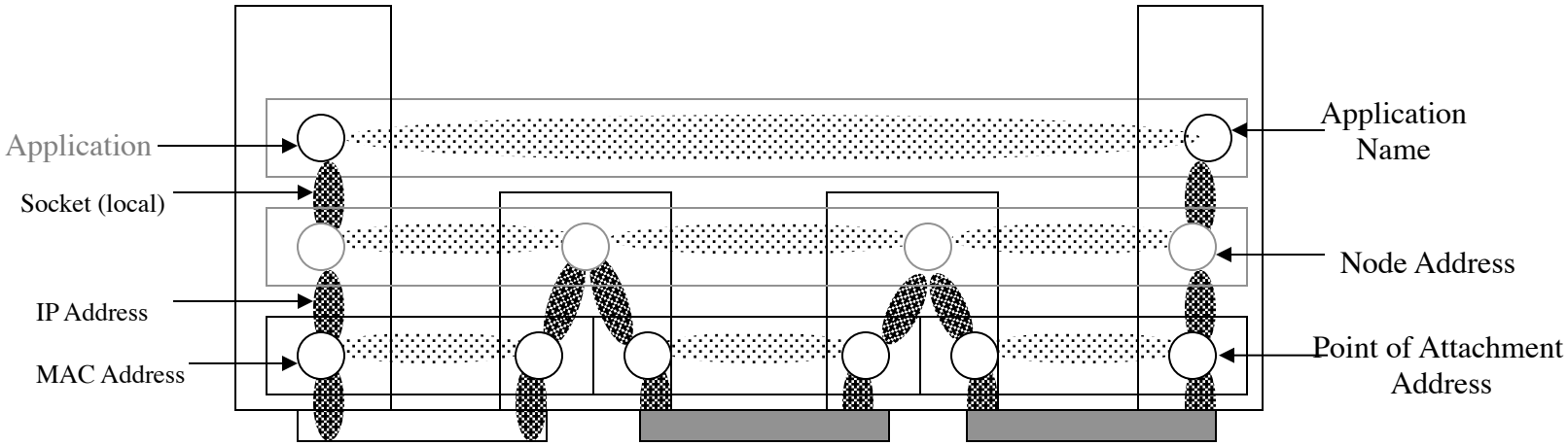
Names the Interface?

- Remember Tinker AFB? The answer was obvious. Just like OSs!
- Directory provides the mapping between Application-Names and the node addresses of all Applications reachable without an application relay.
- Routes are sequences of node addresses used to compute the next hop.
- Node to point of attachment mapping for all nearest neighbors to choose path to next hop. (Because there can be multiple paths to the next hop.)
- This last mapping and the Directory are the same:
  - Mapping of a name in the layer above to a name in the layer below of all nearest neighbors.



Not in the Internet

- The Internet only has a Point of Attachment Address, an interface.
  - Which is named twice!
  - No wonder there are addressing problems
- There are no node addresses or application names.
  - Domain names are macros for IP addresses
  - Sockets are Jump points in low memory
  - URLs name a path to an application



As if your computer worked only with absolute memory addresses.

(kinda like DOS, isn't it?)

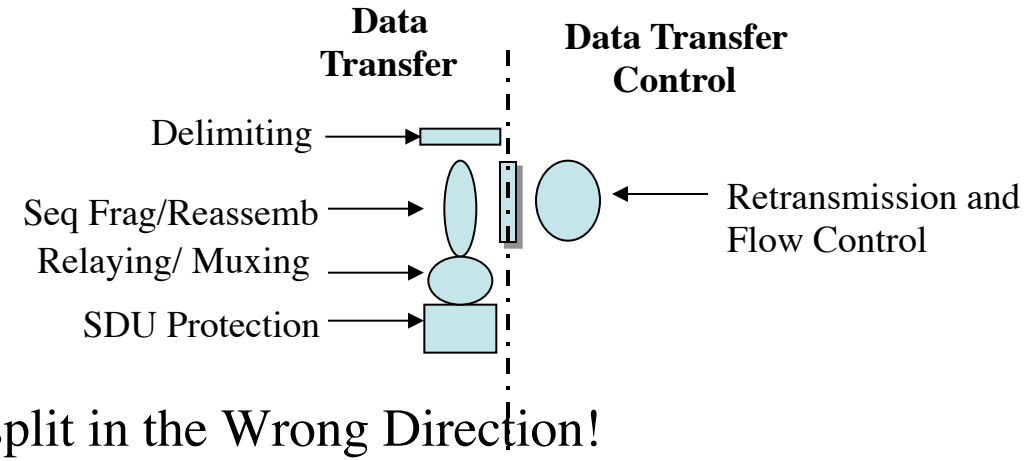
The Big Mistake:
Splitting IP from TCP

- The Rules say if there are two layers of the same scope, the functions must be completely independent.
- Are Separating Error and Flow Control from Relaying and Multiplexing independent? No!
  - IP also handles fragmentation across networks.
  - Remember, Don't repeat functions in different layers of the same scope.
- Problem: IP fragmentation doesn't work.
  - IP has to receive all of the fragments of the same packet to reassemble.
  - Retransmissions by TCP are distinct and not recognized by IP.
    - Must be held for MPL (5 secs!).
    - There can be considerable buffer space occupied.
- There is a fix: MTU Discovery.
  - The equivalent of "Doc, it hurts when I do this!" "Then don't do it."
    - Actually the fix doesn't work either: many nets filter ICMP.
  - Not a "big" problem, but big enough to be suspicious.

But it is the *Nature* of the Problem

That is Interesting

- The problem arises because there is a dependency between IP and TCP. The rule is broken.
  - It tries to make it a beads-on-a string solution.
- A Careful Analysis of this Class of Protocols shows that the Functions naturally cleave (orthogonally) along lines of Control and Data.



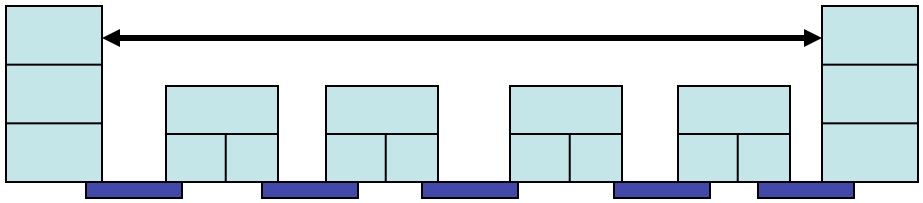
- TCP was split in the Wrong Direction!
    - It is one layer, not two.
  - IP was a bad idea.
- Are There Other Implications?

A Chance to Get Things on Track

- We knew in 1972, that we needed Application Names and some kind of Directory.
- Downloading the Host file from the NIC was clearly temporary.
- When the time came to automate it, it would be a good time to introduce Application Names!

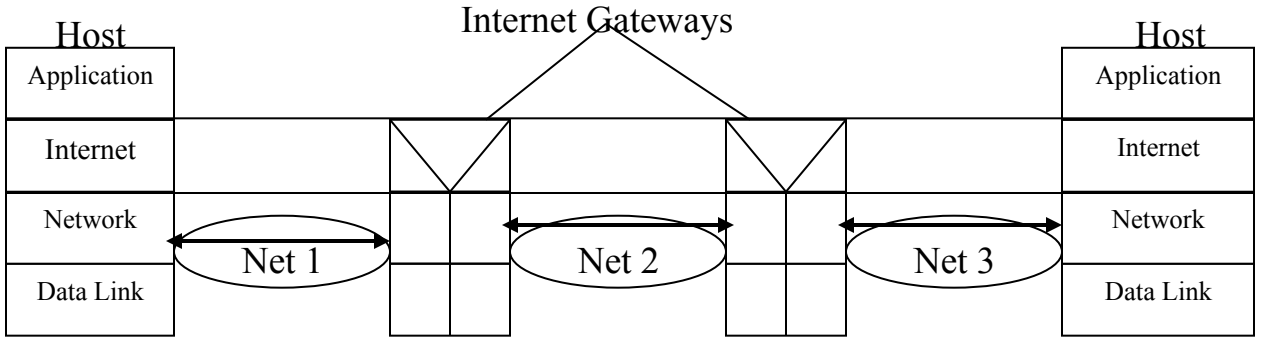
- Nope, Just Automate the Host File. Big step backwards with DNS.
- Now we have domain names
  - Macros for IP addresses
- And URLs
  - Macros for jump points in low memory
  - The path to the Application is named, but Nothing names the *Application*.

Then in ‘86: Congestion Collapse



- Caught Flat-footed. Why? Everyone knew about this?
  - Had been investigated for 15 years at that point
- With a *Network* Architecture they put it in Transport.
  - Worst place.
- Most important property of any congestion control scheme is minimizing *time to notify*. Internet maximizes it and its variance.
- And implicit detection makes it predatory.
  - Virtually impossible to fix
- Whereas,

Congestion Control in an Internet is Clearly a Network Problem



- With an Internet Architecture, it clearly goes in the Network Layer
  - Which was what everyone else thought.
- Time to Notify can be bounded and with less variance.
- Explicit Congestion Detection confines its effects to a specific network and to a specific layer.

Would be Nice to Manage the Network

- All Management is Overhead! We need to minimize it.
  - Then need Efficiency, Commonality, Minimize Uncertainty
- With a choice between a object-oriented protocol (HEMS) and a "simple" approach (SNMP), IETF goes with "simple" to maximize inefficiency
  - Must be simple, has Largest Implementation of the 3: SNMP, HEMS, CMIP.
  - Every thing about it contributes to inefficiency
    - UDP maximizes traffic and makes it hard to snapshot tables
    - No means to operate on multiple objects (scope and filter). Can be many orders of magnitude more requests
    - No attempt at commonality across MIBs.
    - Polls?! Assumes network is mostly failing!
    - Use BER, with no ability to use PER. Requests are 50% - 80% larger
- Router vendors played them for suckers and they fell for it.
  - Not secure, can't use for configuration.
    - (Isn't ASN.1 an encryption algorithm?)
  - Much better to send passwords in the clear.
  - **It is all about account control**

IPv6 Still Names the Interface?

Why on Earth?

- Known about this problem since 1972
  - No Multihoming, kludged mobility
  - Notice in an Internet Architecture this is straightforward.
  - Signs the Internet crowd didn't understand the Tinker AFB lesson.
  - DARPA never made them fix it.
- By the Time of IPng, Tradition trumps Everything.

- When they can't ignore it any longer, and given post-IPng trauma they look for a workaround.
- "Deep thought" yields

Voilà!

Loc/Id Split!

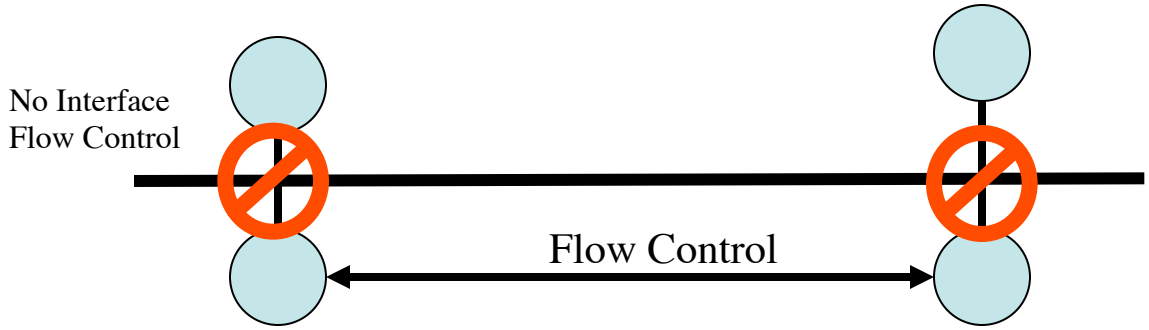
Loc/ID Split

(these are people who
lost a layer to begin with, right?)

- You've got to be kidding?! Right!?
- First off:
  - Saltzer [1977] defines "resolve" as in "resolving a name" as "to locate an object in a particular context, given its name."
  - All names in computing locate something.
    - So either nothing can be identified without locating it, nor located without identifying it, OR
    - anything that doesn't locate something is being used outside its context
    - Hence this is either a false distinction or it is meaningless.
- Second, one must route to the end of the path.
  - The locator is on the path to the *end*, it isn't the end.
    - Getting to the last box is *not* the end of the path. (beads-on-a-string again)
  - The "identifier" locates the end of the path but they aren't routing on it.
  - No wonder it doesn't scale
- There is no workaround. IP is fundamentally flawed.

Never Get a Busy Signal on the Internet!

2010 They Discovered Buffer Bloat!



- Golly Gee Whiz! What a Surprise!!
- With Plenty of Memory in NICs, Getting huge amounts of buffer space backing up behind flow control.
- Well, Duh! What did you think was going to happen?
  - If you push back, it has to go somewhere!
  - Now you can have local congestion collapse!
- If peer flow control in the protocol, pretty obvious one needs interface flow control as well.

But What About Security?

- Security?
- Don't you read the papers?!
  - It is terrible! And all signs are getting worse.
  - IPsec makes IP connection-oriented, so much for resiliency to failure.
  - Everything does their own, so very expensive.
- Privacy? Can't fix it, so same reaction as for QoS
  - You don't need it in the brave new world.
- They say the Reason is that Never Considered It at the Beginning.
  - Later we will see how ignoring security can lead to better security.
- There have been a lot of "after the fact" attempts to improve it.
  - With the usual results: greater complexity, overhead, new threats.

Taking Stock

- The Internet has:
  - Botched the protocol design
  - Botched the architecture
  - Botched the naming and addressing
  - When they had an opportunity move in the right direction with application names, they didn't. They did DNS.
  - When they had an opportunity to move in the right direction with node addresses, they didn't. They did IPv6.
  - More than Botched Network Management
  - Botched the Congestion Control twice
  - Once so bad it probably can not be fixed.
  - Botched Security!

- By my count this makes them 0 for 9!
- It defies reason! Do these guys have an anti-Midas touch or wha!?

But It is a Triumph!

(By that argument, so was DOS)

- But It Works!
- So did DOS. Still does.
- 'With Sufficient Thrust even Pigs Can Fly!' - RFC 1925
- As long as fiber and Moore's Law stayed ahead of Internet Growth, there was no need to confront the mistakes.
  - Or even notice that they were mistakes.

- Now it is catching up to us, is limiting, and it can't be fixed.
  - Fundamentally flawed from the start, a dead end.
  - Any further effort based on IP is a waste of time and effort.
    - Throwing good money after bad
  - Every patch (and that is all we are seeing) is taking us further from where we need to be.

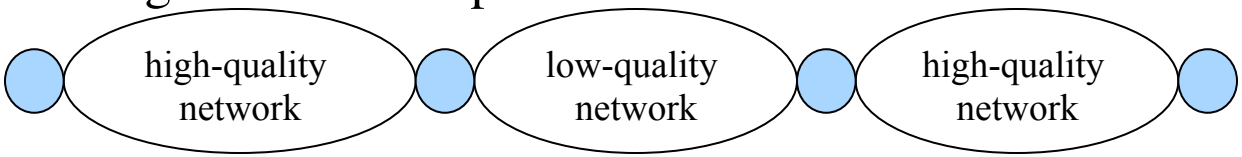
Want to Feel Really Bad?

- A New eBook* James Pelkey's "Entrepreneurial Capitalism and Innovation: A History of Computer Communications, 1968-1988" paints a different picture:
  - First companies were developing LAN products
    - Workstations coming in but SNA is still dominant
  - Then products to connect LANs together in the workplace.
    - Novell and others are making inroads.
  - Then connecting LANs over distances to create corporate networks.
    - Corporations were concerned about security and wanted their own networks
  - In the Middle of this is dumped free software and a subsidized ISP but with a flawed architecture and a lot of hype: The Internet!!

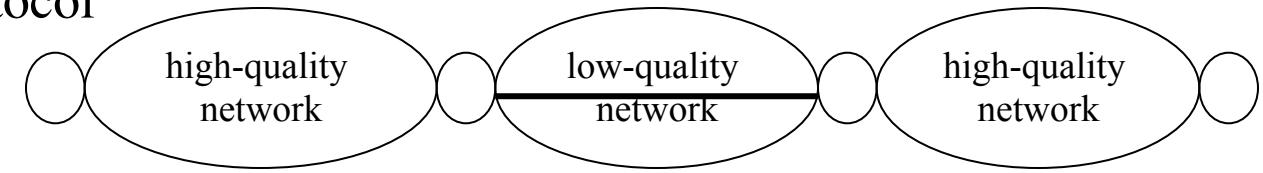
- The Meddling of Big Government Caused the Entire Mess
  - How Do I Know This is What Would Have Happened?
  - Because It Did.

It Was the Computer Companies Who Were Doing the OSI Network Layer

- The other major effort at the time.
- The well-known 7-layer model was adopted at the first meeting in March 1978 and frozen. After that, they had to work within that.
- They knew they would have to accommodate different networks of different quality and different technology.
- One of their concerns in the Network Layer deliberations was interworking over a less capable network:

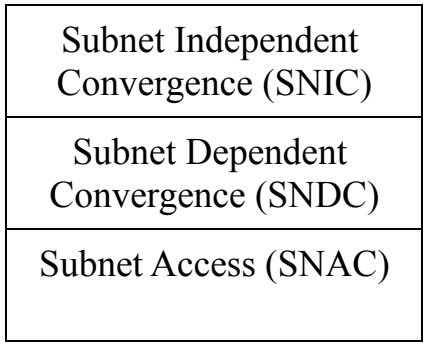


- Would need to enhance the less capable network with an additional protocol

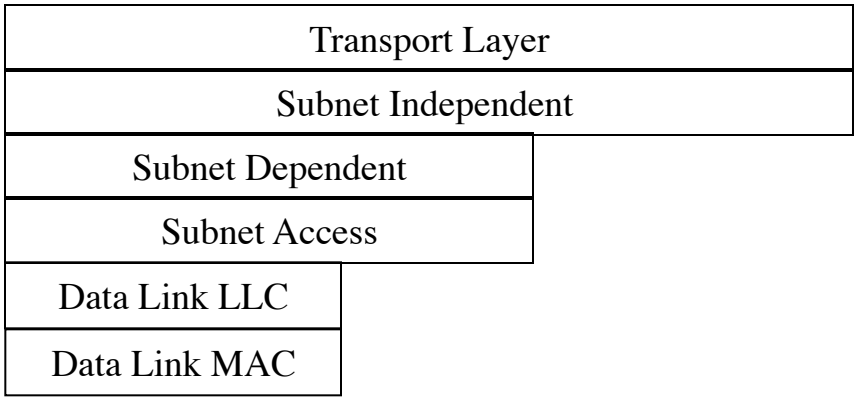


They Sub-Divided the Network Layer

- This concern and the recognition that there would be different networks interworking lead the computer companies to divide the Network Layer into three sublayers, which might be optional depending on configuration:

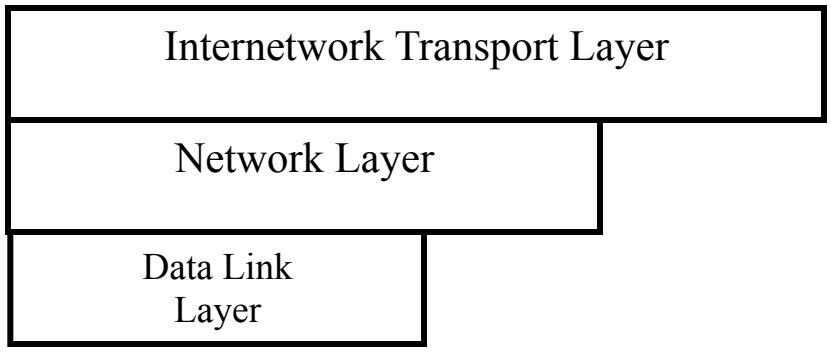


And Came to the INWG Model



- With the Transport Layer, this is the same as the INWG model.
- For different political reasons, OSI made a similar split to TCP/IP.
  - Remember PTTs didn't want a Transport Layer at all.
  - This is independent confirmation. None of the OSI Network Layer group had been involved in INWG.
- So OSI had an Internet Architecture and the Internet has an ITU-like Network Architecture.
  - You just can't make this stuff up!
- And signs of a repeating structure.

INWG Had Been on The Right Track



- They were Close to Seeing it was a Repeating Structure
  - A Structure We Will See Again.
- Had the Research Community Maintained a United Front. Had They Not Assumed They Had Final Answers.
- Had Politics Not Intervened in a Major Way. Had Business Addressed Markets as They Arose.
  - Internet boom and bust would probably have been much moderated
- We Could Have Avoided Many of the Current Problems
  - There Would Still be Security Threats, but far fewer.

Not the Results I Expected

- 20 Years Ago when I embarked on this effort to nail down what it was I knew about Networking, I assumed that the Internet and OSI weren't that different.
  - There were some things in the Internet I knew we hadn't fixed, but they weren't hard to fix. There were some advances that were in OSI we needed to include and a lot of junk from the PTTs to get rid of.
- But the more I pulled on threads sticking out here and there, the more the whole thing unraveled.
  - The more it became apparent that there was much more wrong than first suspected. Most "innovation" and "advance" in the Internet was a myth.
- But in its place emerged an incredibly simple model of extraordinary simplicity and beauty. And the more we push on it, the more it revealed.
- It has truly opened a New World for us to explore.

Questions?