

And You Thought the Wall Street Meltdown was Fun!

John Day
The Pouzin Society
May 2009

‘They’re English,’ said old Fairford, ‘and you can get the English to do anything if you put it to them right. The trouble with the English is they try all the wrong ways first.’

- John Masfield,
poet laureate of the Great Britain
The Bird of Dawning, 1932

Houston, We Have a Problem

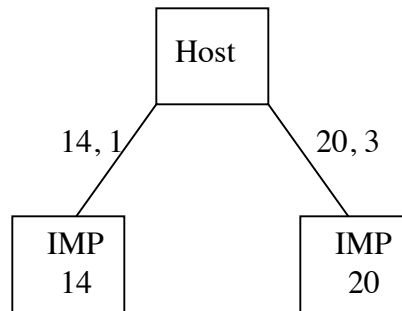
- Two years ago, a major presentation at IEPG.
 - Router table size was on the increase again,
 - And at a rate that was either quadratic or exponential
 - Due to multihoming.
 - Moore's Law won't bail us out this time.
 - This is a big time crisis. We are in big trouble.
 - See Vince Fuller presentation to the IPEG October 2006.
- If not fixed, it is the end of the Internet as we know it.
 - Net will fragment. Costs in the core will skyrocket.

- This problem isn't new.

We Have Known about the Multihoming Problem for 35 Years

- And we have known the answer that long.

What is Multihoming?



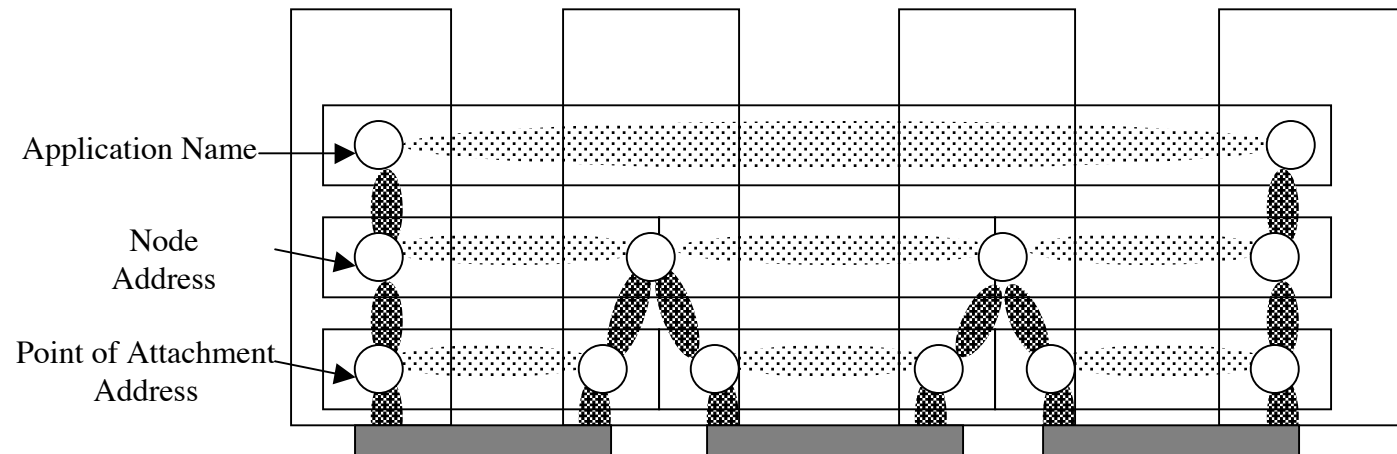
First encountered in 1972 when Tinker AFB joined the ARPANET and wanted redundant connections.

If routing is based on the interface addresses and a system has multiple interfaces then routing can't know that the interfaces go to the same place. It looks to routing like different systems.

It has never been fixed. Kludged, yes; fixed, no.

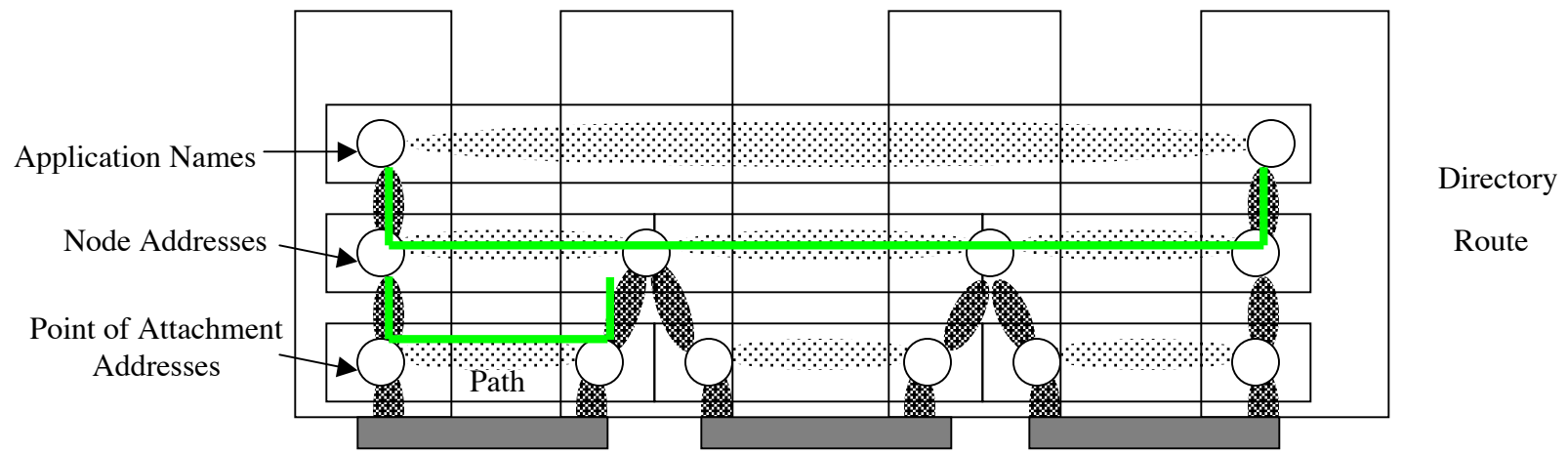
Saltzer Explained the Answer in 1982

- Application names map to node addresses.
- Node addresses map to points of attachment addresses.
- Routes are sequences of points of attachments.
 - Just as in an operating system.
 - But internetworks are different from networks.



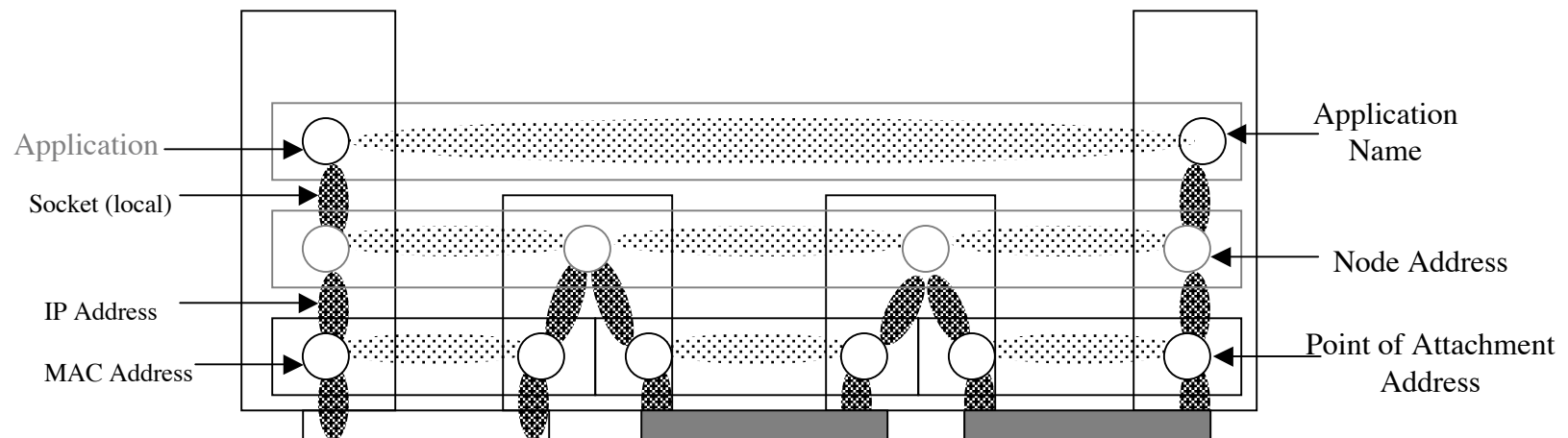
But Networks Are More General

- There can be more than one path between next hops, (they might even be networks)
- Directory maintains the mapping between Application-Names and the node addresses of all Applications.
- Routes are sequences of node addresses used to compute the next hop.
- Node to point of attachment mapping for all nearest neighbors to choose the path to next hop. (Saltzer missed this because it hadn't occurred yet.)
- But notice! 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.



Applying Saltzer to the Internet

- The most striking feature is that half of the addressing architecture is missing.
 - No wonder there are addressing problems.
 - The only identifier we have for anything is the IP address.
- There are no node addresses and no application names.
 - And the point of attachment is named twice!
- At best, IP is a Catenet Protocol, not an Internet Protocol.



As if your computer worked only with absolute memory addresses.

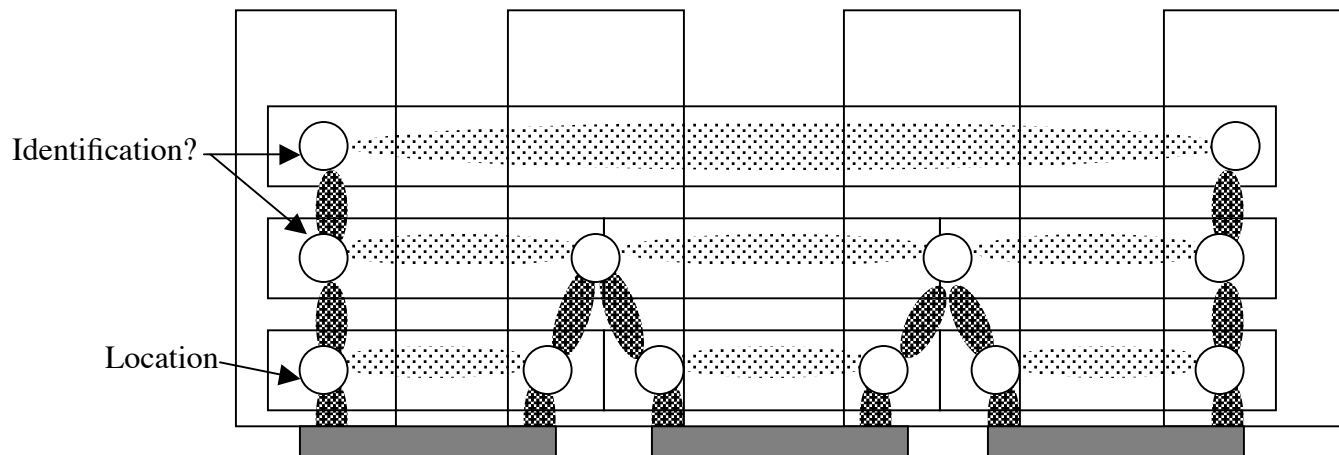
So Why Wasn't It Fixed?

- Odd that a DoD network touted to survive nuclear attack didn't support redundant links. Lots of good "reasons:"
 - Not that many hosts need to be multi-homed.
 - Not then, but the ones that did were the ones everyone wanted to get to.
 - Not everyone should have to bear the cost for a few.
 - Classic committee politics: Put a condition on the solution that guarantees any proposal will be rejected (asymmetry in this case)
 - Also assumes there is a cost.
 - Multihoming will be to different providers, so no point.
 - Assumption is wrong and even if right assumes a static network.
 - Actually, we did try to fix it but it was rejected by the IETF.
 - Finally, they deemed to look at the problem.

Around the turn of the century

The Answer to Multihoming is loc/id split

- The reasoning went: the semantics of the IP address are overloaded as both locator and identifier. (Just one problem)
- It is impossible to locate something without also identifying it.
 - Especially in computing, where virtually all identifiers are “locators.”
 - Interestingly Saltzer [1977] defines “resolving a name” as ‘*locating* an object given its name’ (emphasis added)
 - Hmmmmm, already indications of a false distinction



That Brings Us Back to Here

Houston, We Have a Problem

- Two years ago, major presentation at IETF.
 - Router table size is on the increase, due to multihoming.
 - Moore's Law won't bail us out this time.
 - This is a big time crisis. We are in big trouble.
- If not fixed, it is the end of the Internet as we know it.
 - Net will fragment. Costs in the core will skyrocket (appear unbounded).
 - NetworkWorld sits on the story for a year.
 - Tons of papers written on loc/id split!
- Finally, Cisco and others start proposing solutions.
 - Mostly requiring patches involving NATs and a bevy of new protocols.
 - This makes everyone nervous, but what else to do?
 - This is a crisis!

Then Last October

- Dave Meyer, Cisco, gives a talk at IEEE CCW, “The End of the Internet as We Know It”:
 - *Assertion: The lack of a reasoned approach to both the IPv4 run-out problem (data plane) and the growth of routing state (control plane) are life-threatening to the (end-to-end) Internet we all know and love.*
 - The situation is worse than I thought. The IPv6 group has totally blown it. v6 group didn't realize there were router table issues that needed to be solved.
 - An Address 4 times bigger creates much much larger scaling problems,
... well duh.

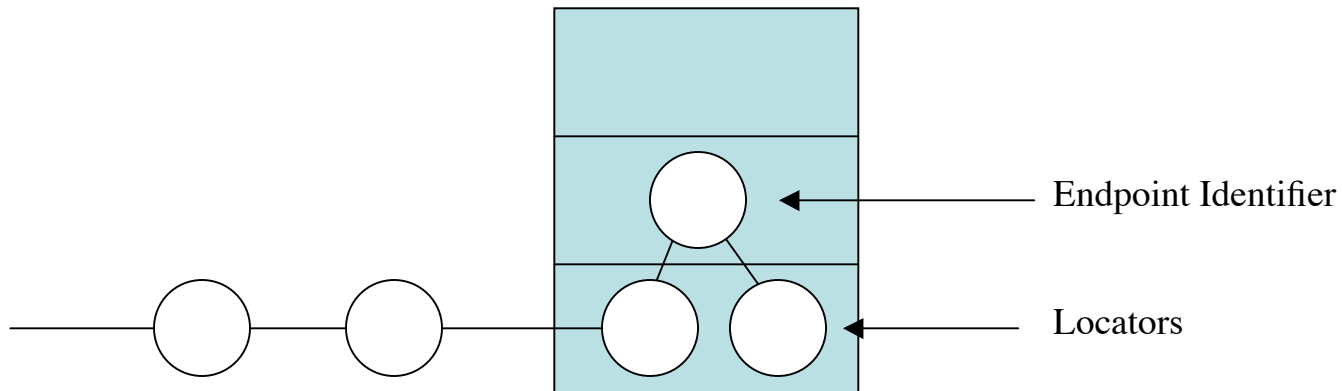
In November

Houston, We have a Bigger Problem

- Dave Meyer says, ‘I have an “architectural issue” to discuss.’
 - He has come across two problems in implementing LISP.
 - Both are $O(n^2)$ with the number of destinations.
 - Both require doing path discovery.
 - Path discovery doesn’t scale. LISP won’t scale. QED
 - He suspects that any loc/id approach will have the same problem.
 - draft-meyer-loc-id-implications-01.txt
 - In case you didn’t notice, we just went to DefCon5
 - If Dave’s suspicions are right, then we are up sh. . . up a creek without a paddle.
- Dave: Why hasn’t anyone noticed this in the last 15 years?

Dave is Right

- All proposals based on loc/id split will have the same flaw.
- Once put in the context of the model, it is obvious.
- Let us look at it carefully. What do the locator and the identifier name?



The Locator Locates the Wrong Thing!

The locator is part of the path, not the final destination.

No wonder Dave ran into *path* discovery issues.

Apply the e2e principle!

Solve the Problem in the Hosts

- There are poor misguided souls claiming this.
 - Mostly done by changing the definition.
 - Not worrying that it might take seconds if not minutes to do the failover.
- Remember the fundamental problem is that the network doesn't know that two paths go to the same place.
 - This is a problem of *delivering*, not sending.
- There is no solution as long as one routes only on the interface address. Which means . . .

If Loc/ID Doesn't Scale

- Then no protocol that routes only on the interface scales.
- In other words,

IP is Fundamentally Flawed

(v4 or v6)

Wanna Hear the Real P*sser?

- We had the right answer in 1992.
 - It was rejected by the IETF.
 - Why? Hubris and the Failure of University Education
- And to add insult to injury, **it was deployed in the routers.**
 - We could have spent the last 15 years working on transition
 - Rather than 100s of millions on a small incremental step that provides no benefit to your bottom line and doesn't scale.
- You just can't make stuff like this up!

So What Is the IETF Plan B?

- Plan A
 - What are the major Universities working on?
 - Plan A
 - What are the vendors working on?
 - Plan A
-
- Talk about scary, there was a meeting of luminaries in March to discuss the addressing crisis
 - The agenda: to determine what form of loc/id split to use!
 - It is important to get those deck chairs lined up very neatly.

But Everything Works Fine!

- Yes and it will for a few more years.
- When you notice it means we are already over the cliff.

“Has he vertigo?

No, only about

ten stories more!”

- Ogden Nash

If We Stay the Course?

- Not entirely clear.
 - Increasing router table size (and desire to control costs) will probably cause carriers to arbitrarily limit the number of routes they handle.
 - Exhaustion of v4 address space is predicted to create a market in v4 addresses. This will exacerbate the problem.
 - Current proposals greatly increase the cost in the core.
- Conjecture:
 - The Internet will begin to fragment
 - Costs will skyrocket. (Fuller predicts unbounded cost.)
 - Undoubtedly unforeseen consequences

So Now What?

- Back to IPv7 and CLNP?
 - Much better to go forward, than backward.
- A Crisis is a Terrible Thing to Waste.
 - Some of us have learned a lot about architecture since 1992.
- We can kill several birds with one stone: PNA
 - Multihoming, mobility, and multicast for free.
 - Security and congestion control
 - Reduce capex and opex costs by factors of 10.
 - And scales indefinitely.
- That's my next meeting.