

Multipath Transport, Resource Pooling, and implications for Routing

Mark Handley, UCL and XORP, Inc

Also:

Damon Wischik, UCL

Marcelo Bagnulo Braun, UC3M

The members of **Trilogy** project: www.trilogy-project.org



A few things that will stress routing ...

- Scalability (natural growth).
- Ubiquitous multihoming for robustness.
- Increased demand for very fast recovery from failures.
 - VoIP, IPTV, Games - need sub-second recovery times.
 - Net becoming critical for business.
- Resilience to flash crowds, DDoS attacks, earthquakes, etc.
- 4 billion IP-connected cellphones with multiple radios that can be used simultaneously.

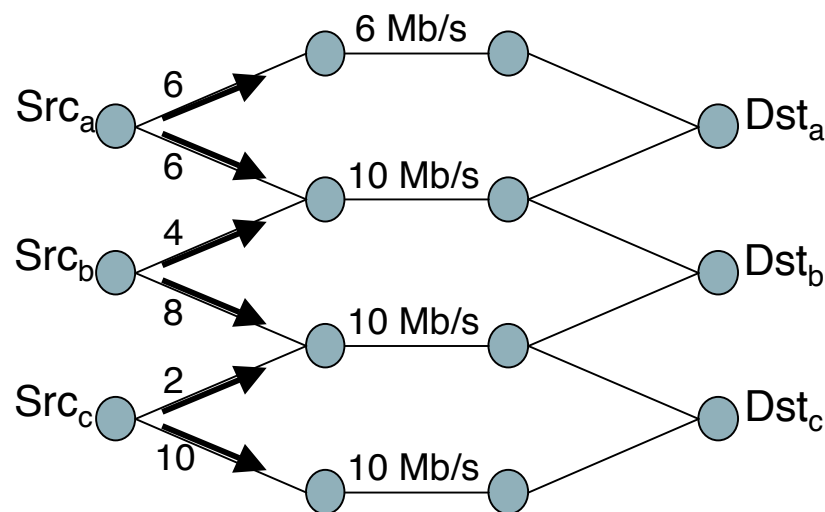
Assertion

- We can't make routing scale, and simultaneously use routing to solve:
 - Multihoming.
 - Fast recovery.
 - Short-timescale traffic engineering.
 - Mobility (eg Boeing Connexion).

Resource Pooling?

Make a network's resources behave like a single pooled resource.

- Aim is to increase reliability, flexibility and efficiency.
- Method is to build mechanisms for shifting load between the various parts of the network.

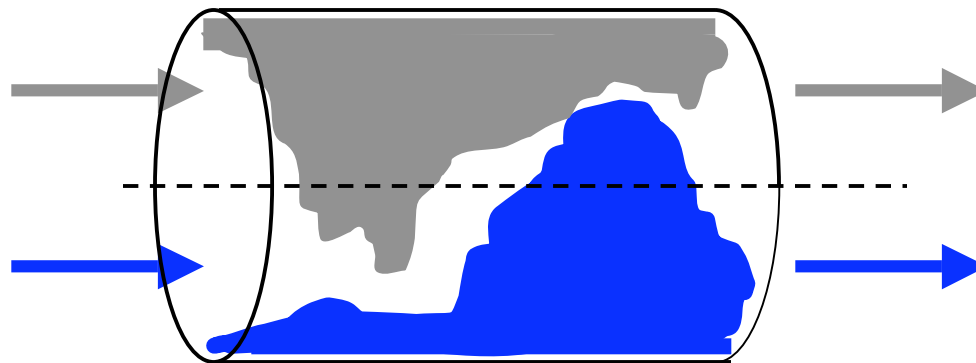


Everyone keeps reinventing resource pooling to solve their own local problems.

Resource Pooling is not new...

Computer communication is bursty, so a virtual circuit-based model with rate allocations gives poor utilization.

- A packet-switched network pools the capacity of a single link.
 - Goal: high utilization
- Router queues pool capacity from one time interval to the next
 - Goal: high utilization, robustness to arrival patterns



We're doing resource pooling in routing

- **BGP traffic engineering:**

- Slow manual process to pool resources across peering links.
- Financial goal - match revenues to costs.
- Robustness goal - prevent overload.

- **OSPF/MPLS traffic engineering:**

- Slow mostly manual process to pool resources across internal ISP links.
- Primary: Robustness - prevent overload.
- Secondary: Higher utilization.

- **BT, AT&T (and others) dynamic alternative routing**

- Robustness to overload.
- Provide higher availability than the availability of the links/switches themselves (pool reliability)

Recent resource pooling trends

■ Multihoming

- Primary goal: pool reliability.
- Secondary goal: pool capacity

■ Google, Akamai, CDNs

- Pool reliability of servers, datacenters, ISPs.
- Pool bandwidth.
- Pool latency??

■ Bittorrent

- Overall: Pool upstream capacity (over space and time)
- Per-chunk: pool for reliability from unreliable servers.

Summary:

Motivations for Resource Pooling

- Robustness
- Increased capacity or utilization

Currently two main resource pooling mechanisms:

- **Routing-based traffic engineering.**
 - Inter-domain routing is too slow and doesn't scale well (especially if a human is in the control loop)
 - Intra-domain routing is better, but not fast enough with a human in the loop, not stable if automatic.
 - There are many examples where no network-based flow-based mechanism can achieve pooling.
- **Application-based load-balancing between multiple servers.**
 - Pretty effective, but strong tussle with what the network operators are doing.

So what might work?

- **Multipath.**
 - Only real way to get robustness is redundancy.
- **Multihoming**, via multiple addresses.
 - Can aggregate routing information.
- **Mobility**, via adding and removing addresses.
 - No need to involve the routing system, or use non-aggregatable addresses.

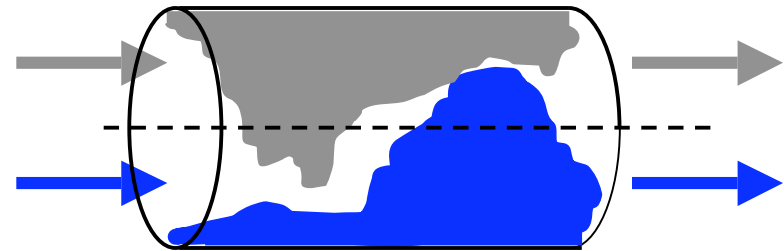
So what might work?

- **Multipath-capable transport layers.**
 - Use multiple subflows within transport connections.
 - Congestion control subflows independently
 - Traffic moves to the less congested paths.
- Note the involvement of congestion control is crucial.
 - Link the backoff parameters for stability and fairness (Kelly+Voice).
 - You can't solve this problem at the IP layer.

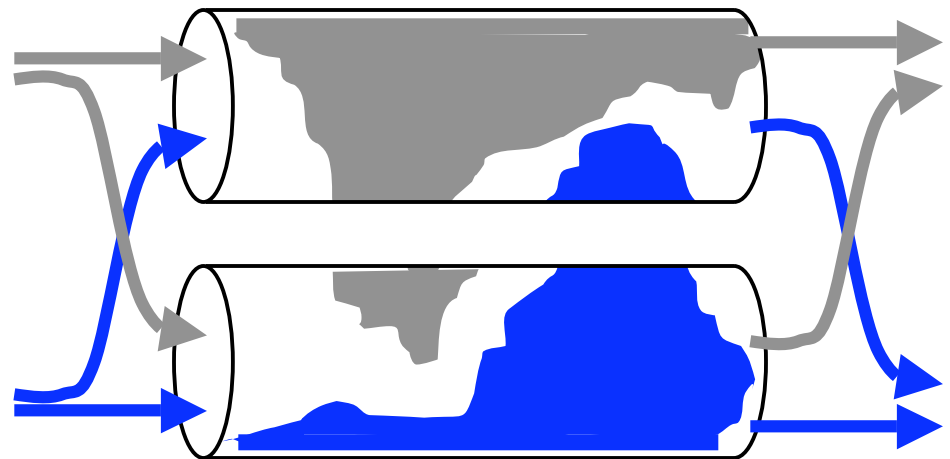
Multipath transport

- Multipath transport allows multiple links to be treated as a single pooled resource.
- Traffic moves away from congested links.
- Larger bursts can be accommodated.

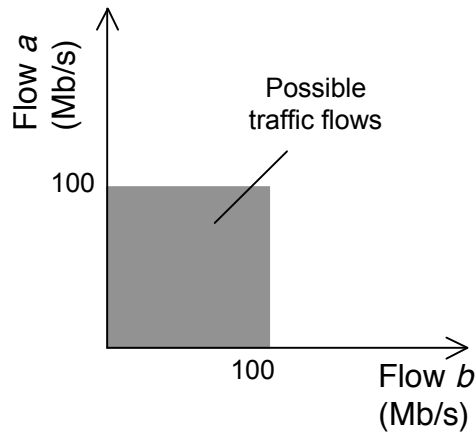
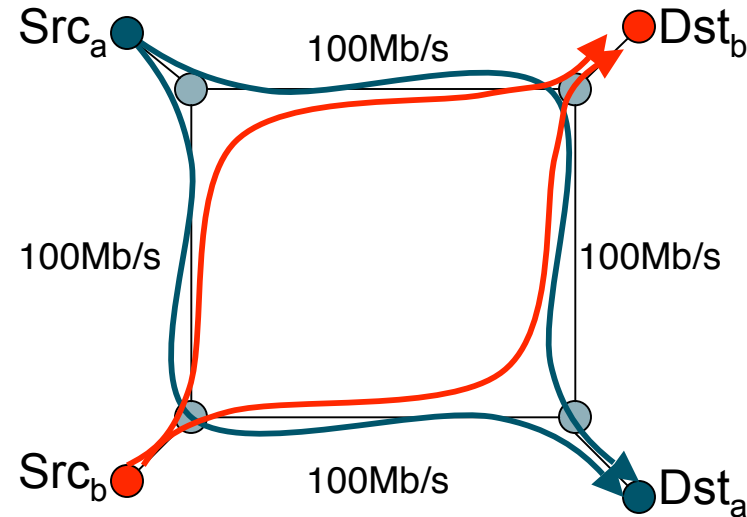
ARPAnet resource pooling:



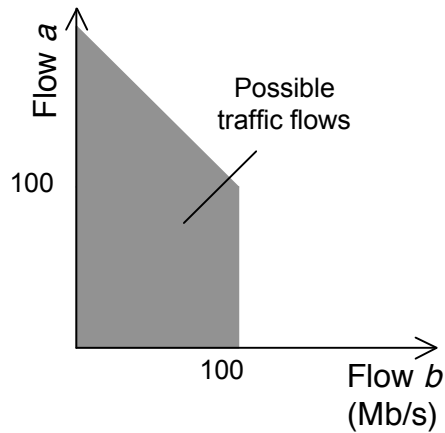
Multipath resource pooling:



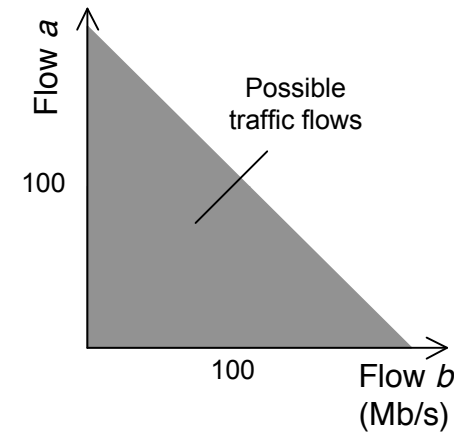
Resource pooling
allows a wider range
of traffic matrices



No multi-path flows

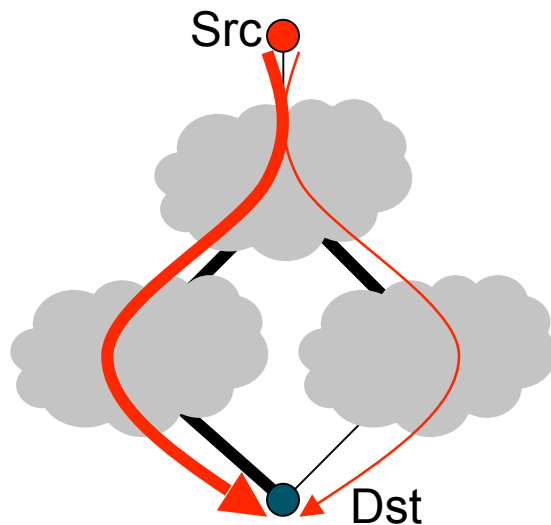


Only flow a is multi-path.

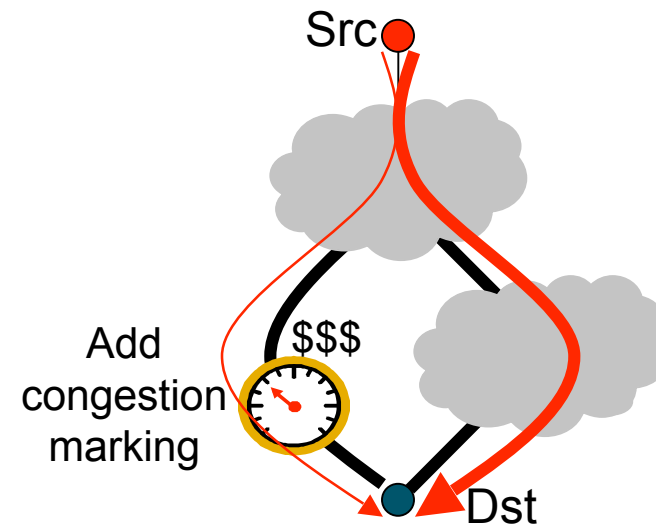


Both flows are multi-path

Multipath Traffic Engineering

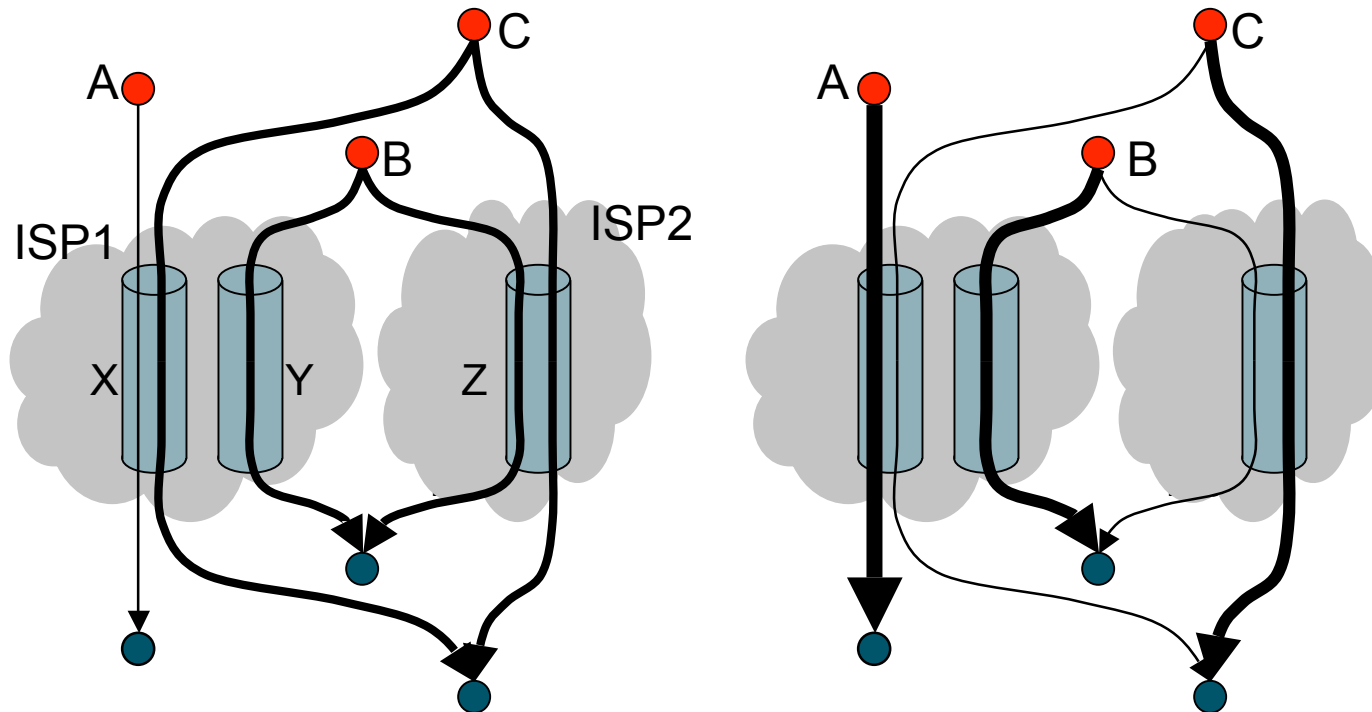


- Balancing across dissimilar speed links



- balancing across dissimilar cost links

End-systems can optimize globally (often ISPs cannot)



The “Resource Pooling Principle”

- **Observation 1:** Resource pooling is often the only practical way to achieve resilience at acceptable cost.
- **Observation 2:** Resource pooling is also a cost-effective way to achieve flexibility and high utilization.
- **Consequence:** At every place in a network architecture where sufficient diversity of resources is available, designers will attempt to design their own resource pooling mechanisms.
- **Principle:** A network architecture is effective overall, only if the resource pooling mechanisms used by its components are both *effective* and *do not conflict with each other*.

Corollary of the “Resource Pooling Principle”

- **Principle:** A network architecture is effective overall, only if the resource pooling mechanisms used by its components are both *effective* and *do not conflict with each other*.
- **Corollary:** The most effective way to do resource pooling in the Internet is to harness the responsiveness of the end systems in the most generic way possible, as this maximizes the benefits while minimizing the conflicts.

Multipath Transport Design Space

Multipath TCP:

- So obvious it's been proposed at least ~~four~~ *five* times (originally by Huitema?).
 - SCTP is already going there.
- We now understand that multipath TCP, if done appropriately, can go a long way towards solving network-wide traffic engineering problems.
- We're starting to understand the consequences of not solving the issue in a general way.

Multi-server HTTP:

- Request chunks of file, each from a different server.
- Better pooling, but less general.

Peer-to-peer?

What about Multipath Routing?

- You can achieve resource pooling using the routing system if:
 - Routers can make a choice (at a flow granularity) of more than one path for traffic forwarding to a destination.
 - The load-balancing between the paths is done based on the measured congestion on those paths to that destination.
- This has the same effect of moving traffic away from congested paths.
 - It's harder to make stable.
 - It doesn't provide resilience for individual flows (still need to re-route very quickly).

What if most flows are multi-path?

- Greatly reduced need to do traffic engineering by tuning routing.
 - Eg. incoming traffic to a multi-homed site naturally balances between both links.
- Can use aggregated PA addresses for routing of multi-homed edge sites.
 - Reduces the prefixes advertised.
 - Reduces the churn, as failures in edge links no longer trigger global routing updates.

Aggregated PA addresses for multihoming?

- For multipath-capable end-systems, failure detection is best done at the transport level (much faster).
 - Need to bootstrap connections - need multiple addresses in DNS.
- For legacy end systems, failure recovery is more problematic.
 - Can restart a connection using a different address from DNS (unsatisfying).
 - Tunnelling from one ISP to the other is feasible, but ugly.
 - 8+8 would make this easy.
 - Advertise a more specific via working path only when other path has failed. (removes some of the benefits, but at least is only an interim solution).
 - Directed BGP updates?

PA addresses?

- IPv4 is probably a lost cause for PA addresses.
 - Other benefits of multipath transport still apply though.
- Everyone wants PI addresses anyway.
 - No-one wants to renumber.
- Mobile hosts will have to renumber anyway.
- For non-mobile hosts, if PI addresses are really needed, one-to-one address re-writing to a PA address is probably the best scalable solution.
 - Six/One?

Summary

- Multipath transport can deliver resource pooling.
 - This is the closest thing to load-dependent routing that is likely to scale globally and be stable.
- People will attempt to build resource pooling solutions anyway.
 - Such solutions will conflict with each other.
 - Multipath transport minimizes the bad interactions between such solutions.
- We need to think carefully about the division of control functionality.
 - What is the role of routing and network-based traffic engineering?

Key Theory Papers

- *Path selection and multipath congestion control*, P. Key and L. Massoulié and D. Towsley, Proc. IEEE Infocom'', 2007.
- *Stability of end-to-end algorithms for joint routing and rate control*, F. Kelly and T. Voice, Computer Communications Review, Vol 35, Number 2, April 2005.

Position Paper

- The Resource Pooling Principle, M. Handley, M. Bagnulo-Braun and D. Wischik,
- <http://www.cs.ucl.ac.uk/staff/m.handley/papers/resource-pooling-principle.pdf>