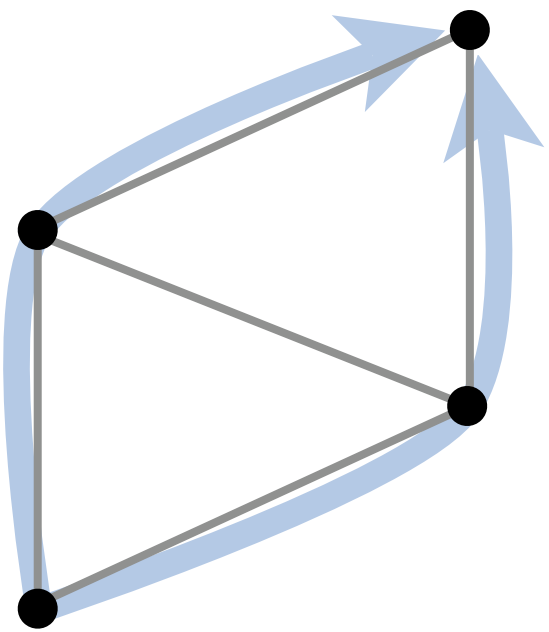
PATHLET ROUTING

Brighten Godfrey
Scott Shenker
Ion Stoica
{pbg,shenker,istoica}@cs.berkeley.edu

UC Berkeley

Hotnets 2008

multipath internet routing



good for everyone!

reliability source observes directly, reacts quickly

path quality source observes directly, knows what it wants

money network providers can sell new service

good for everyone

"The Route Not Taken" [Frost 1920]

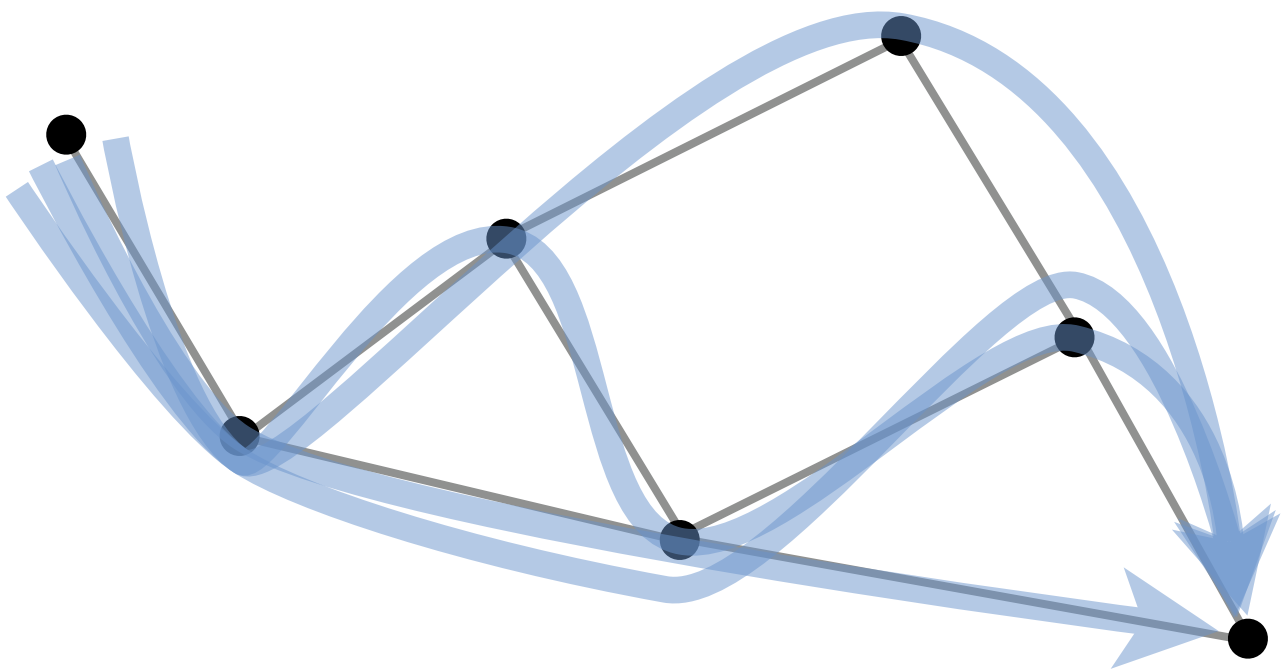
Two routes diverged in a network, and I --
I took the one less transited,
And that has reduced latency by up to 41%.

why don't we have it?

Even if everyone involved wants multipath,
no way to do it in BGP!

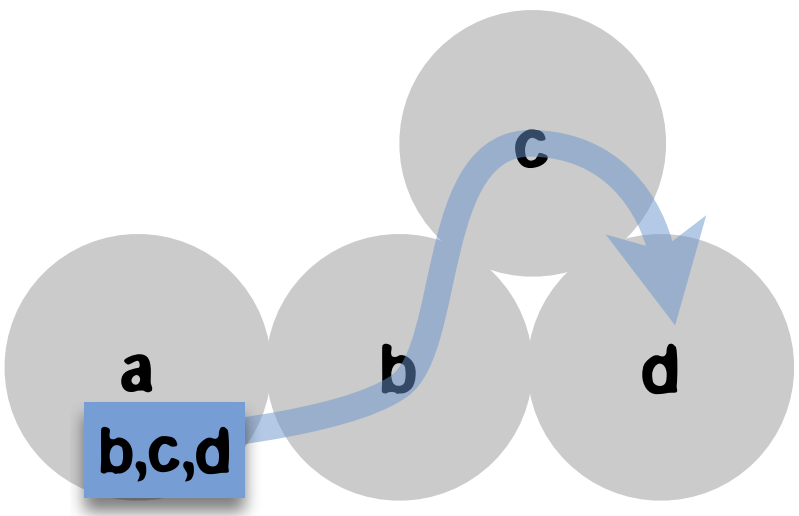
All paths
blocked
except one.

Offer more?
State explosion!



why don't we have it?

AS-level source routing...



...gives network owners
no control.

why don't we have it?

path vector (BGP)

great at blocking paths!
bad at allowing them.

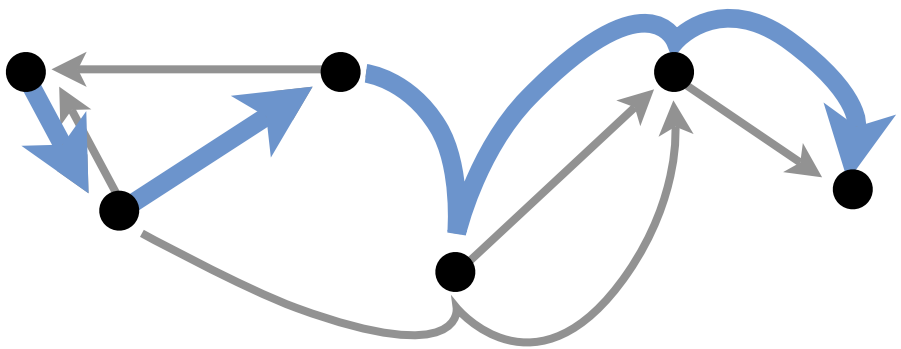
AS-level source routing

great at allowing paths!
bad at blocking them.

Highly constrained routing policies.

goal: flexible policy control → many paths

pathlet routing



fragments of paths (pathlets)

+

source routing

goal: flexible policy control → many paths

outline

- the protocol
- emulating other protocols
- local transit (LT) policies

pathlet routing

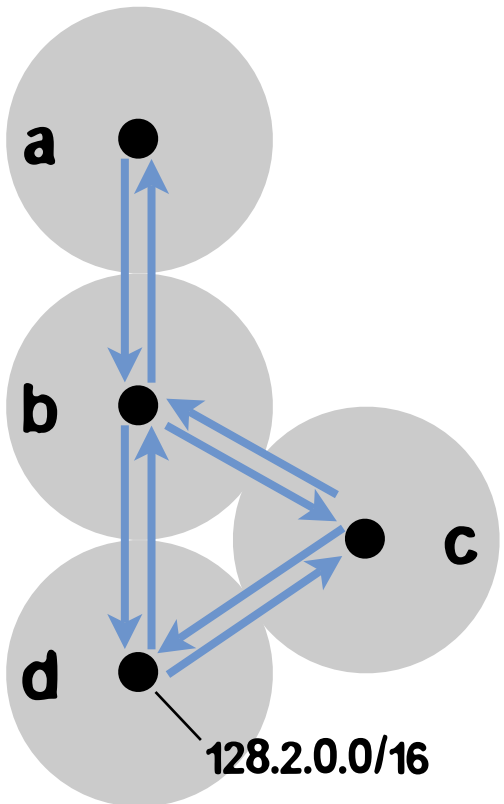
1. **vnode:** virtual node within an AS
2. **pathlet:** sequence of vnodes
3. **announce** pathlets
4. source lists pathlets in packet

example

one per AS

one per link

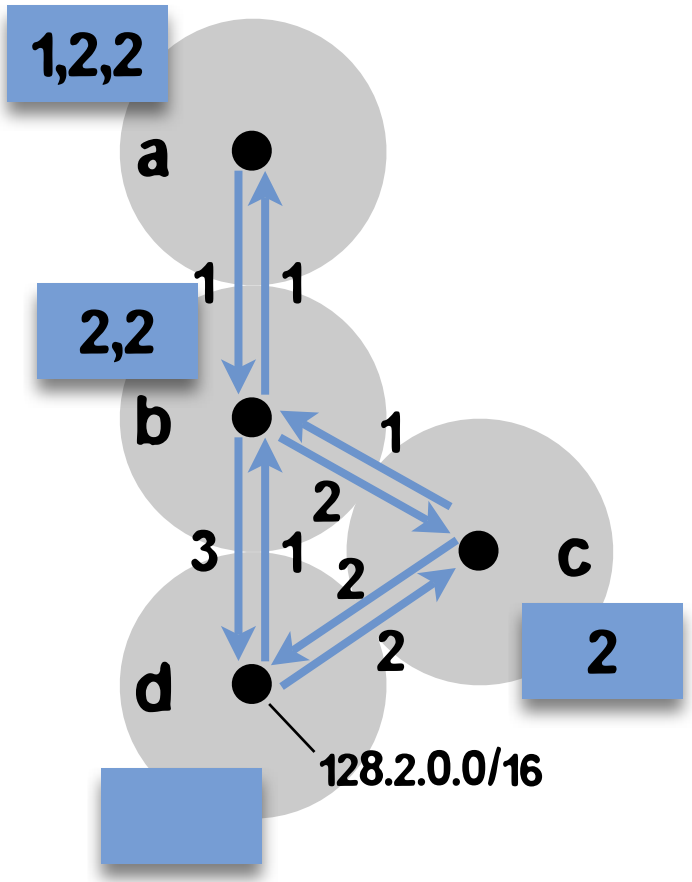
gossip all known pathlets to neighbors



= AS level source routing

forwarding plane

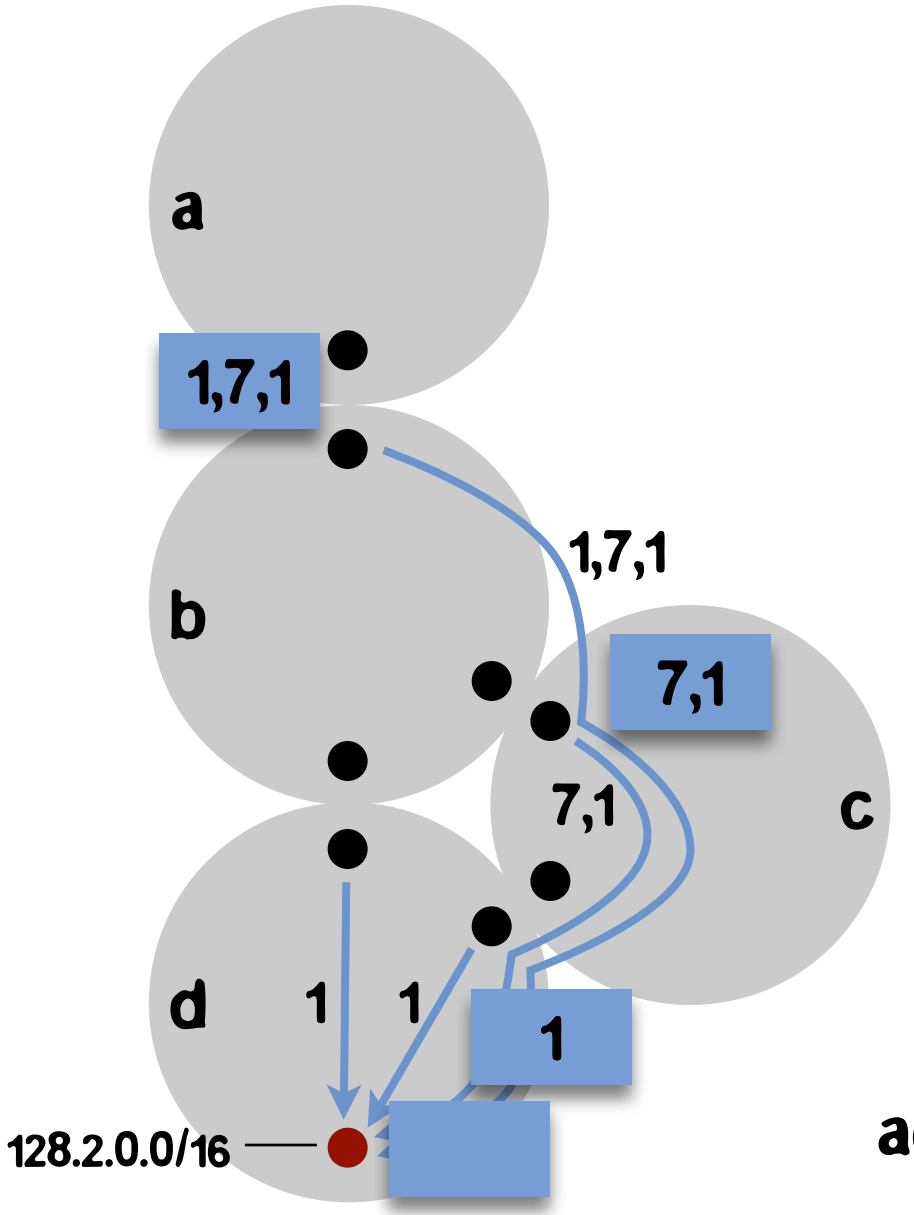
- pathlets tagged with Forwarding ID sequence
- packet contains list of FIDs
- forwarding table maps FID to, e.g., outgoing interface



outline

- the protocol
- emulating other protocols
- local transit (LT) policies

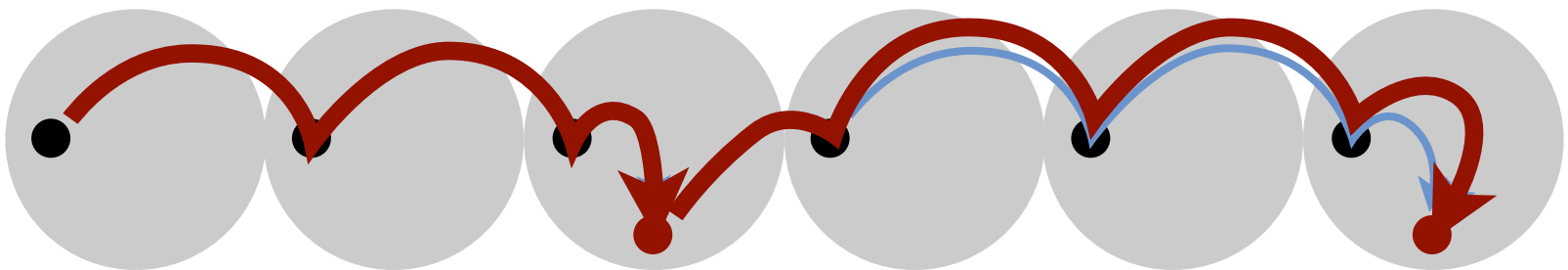
emulating BGP



(not discussed: how to verify route follows advertised policy)

emulating other protocols

- **MIRO** [Xu, Rexford, SIGCOMM'06]

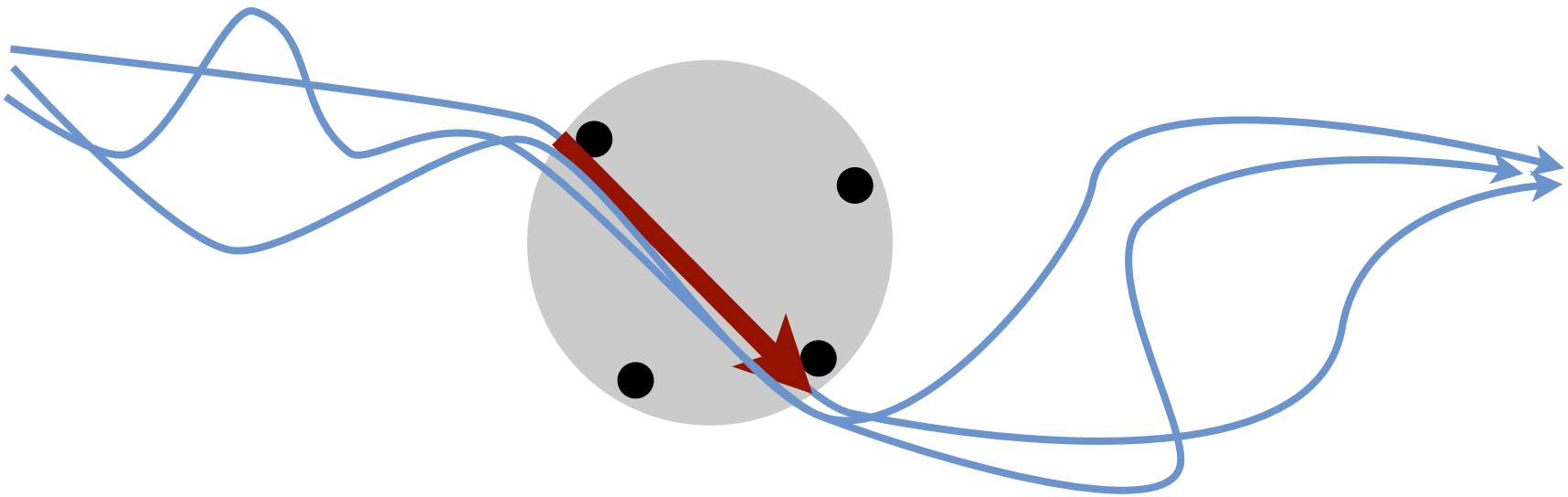


- **NIRA** [Yang, Clark, Berger, ToN'07]

outline

- the protocol
- emulating other protocols
- local transit (LT) policies

"local transit" policies



Each ingress --> egress pair
is either allowed or disallowed.

Subject to this, any path allowed!

LTP a common case?

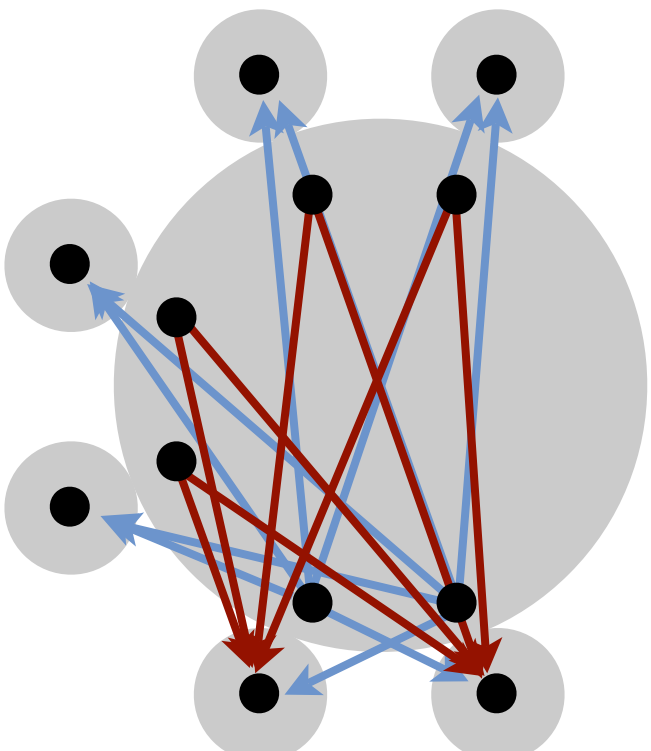
- capture a network's direct costs
- valley freeness is a LT policy, and the common case in BGP export policies today

valley free routing as an LTP

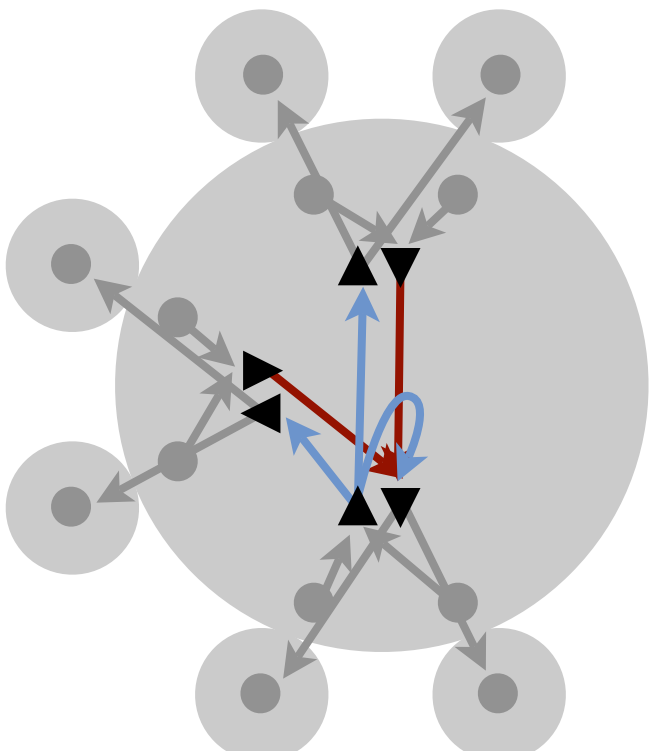
"customers can route to anyone;
anyone can route to customers"

providers

peers



customers



scalability

	BGP	Pathlet routing, class-based LT policies
forwarding table entries	$O(kn)$	$O(d+k)$
control plane entries	$O(dkn)$	$O((d+k)n)$
control plane messaging	$> O(Lkn/d)$	$O(dn)$

n = # ASes d = mean # neighbors

L = mean path len k = prefixes per AS

forwarding table entries

		current Internet (CAIDA/APNIC):
BGP	one per destination (IP prefix)	266,073 entries
pathlet routing, LT policies	one for each pathlet starting at the router	2,317 entries, max 6 entries, mean

conclusion

- **pathlet routing:** flexible policies --> multipath with many choices, better scalability
- can't emulate everything, e.g. FBR [Zhu, Gritter, Cheriton '03]
- emulate others? path splicing [Motiwala, Elmore, Feamster, Vempala 2008], Routing Deflections [Yang, Wetherall 2006]
- challenge for all multipath protocols: different payment for different paths?

thanks: fonts by tom7