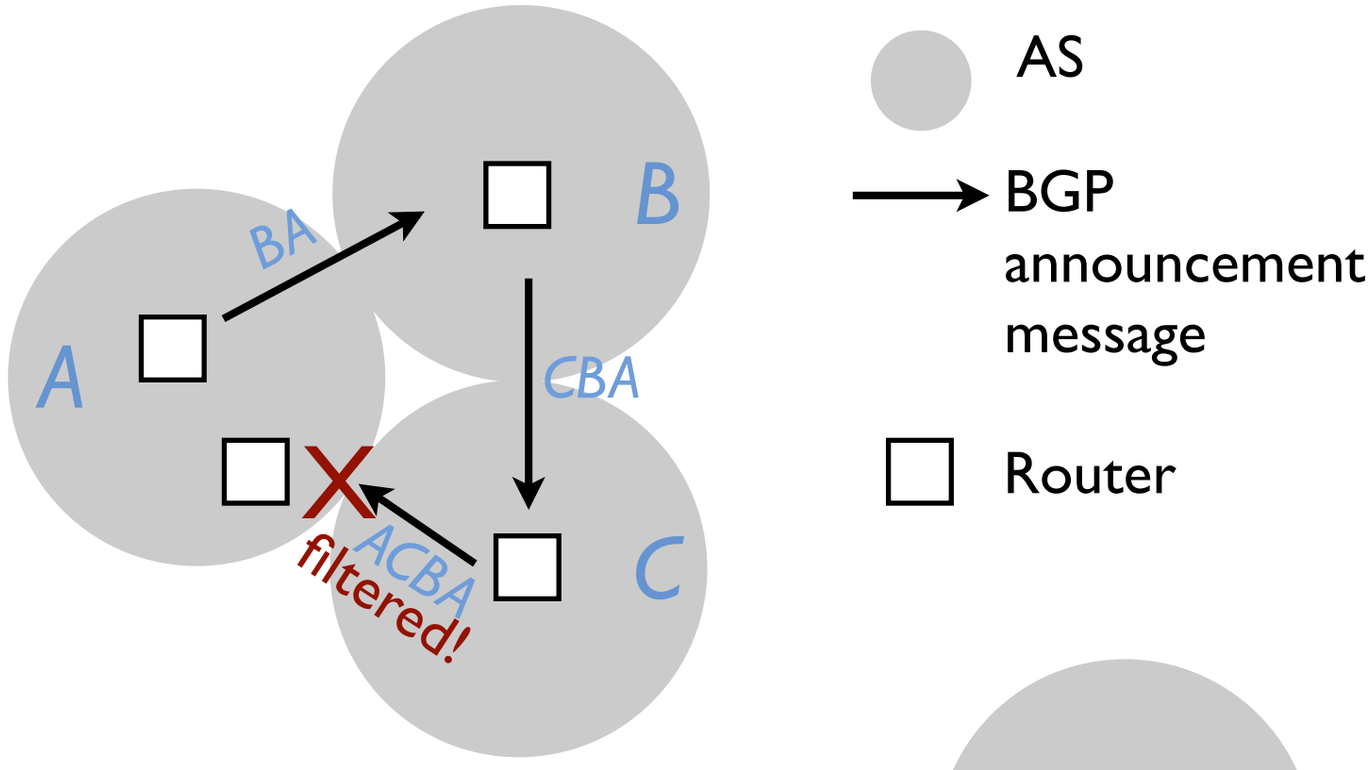
The AS Partitioning Problem in Pathlet Routing

Brighten Godfrey
pbg@illinois.edu
July 2, 2009

AS partitioning problem

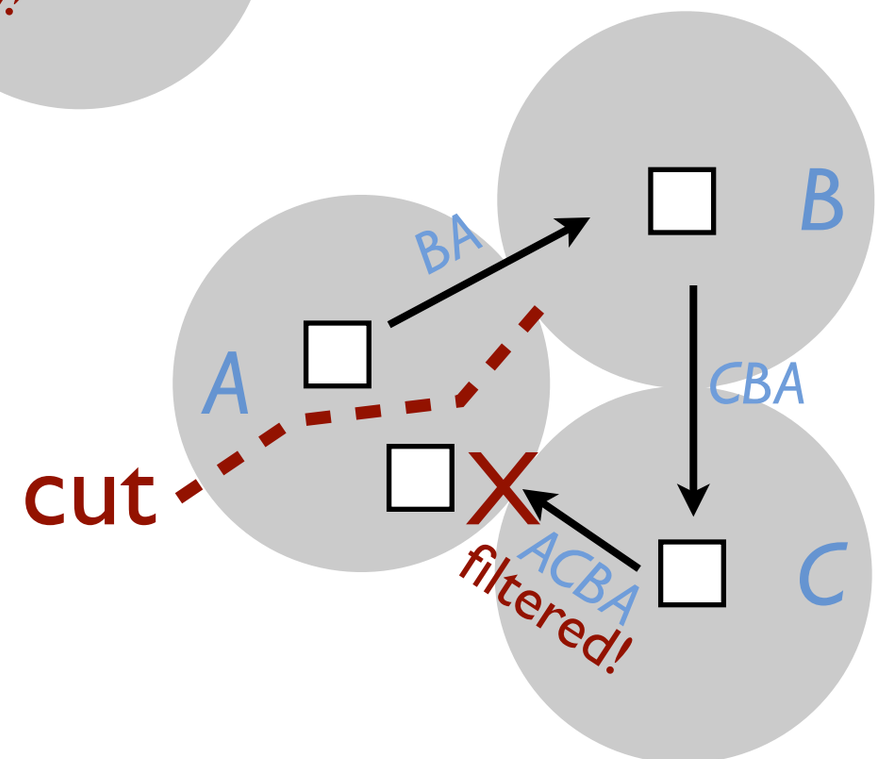
BGP operates very roughly at the level of ASes: an AS is a node in the graph.

For example, if an AS appears twice in the ASPATH, it is assumed there is a loop and the route is filtered.



Problem: this topological aggregation breaks if an AS is internally partitioned.

Here, although the overall internetwork is connected, the two halves of *A* cannot reach each other.



AS partitioning problem

We ran into the same problem in pathlet routing (to appear in SIGCOMM'09).

This wasn't the main focus of the paper but it turns out that in pathlet routing, there is a clean solution.

The key idea is that pathlet routing works at the level of vnodes (virtual nodes) which have essentially arbitrary meaning. If an AS partitions, new vnodes can be created to represent the new topology which is not aggregatable in the same way as the old one.

Next we'll look at the pathlet routing protocol, and how it can solve the AS partitioning problem.

Pathlet routing

Pathlet routing has two building blocks:

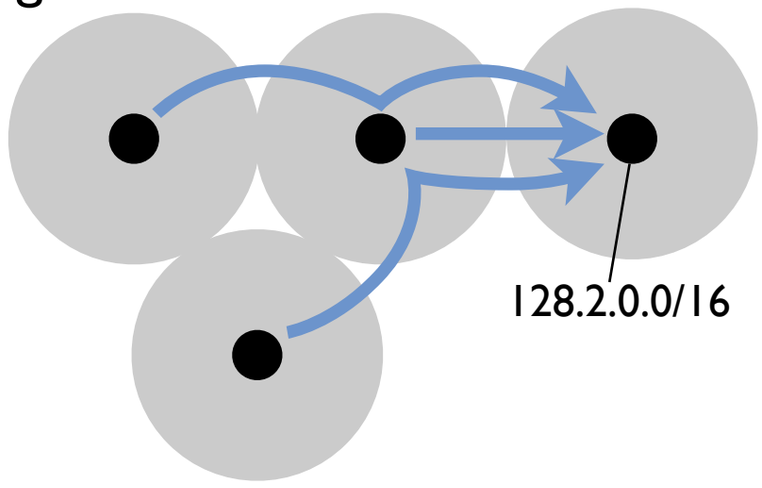
- a vnode is a virtual node within an AS.
- a pathlet is a fragment of an end-to-end-path: a sequence of vnodes.

Senders do source routing over pathlets.

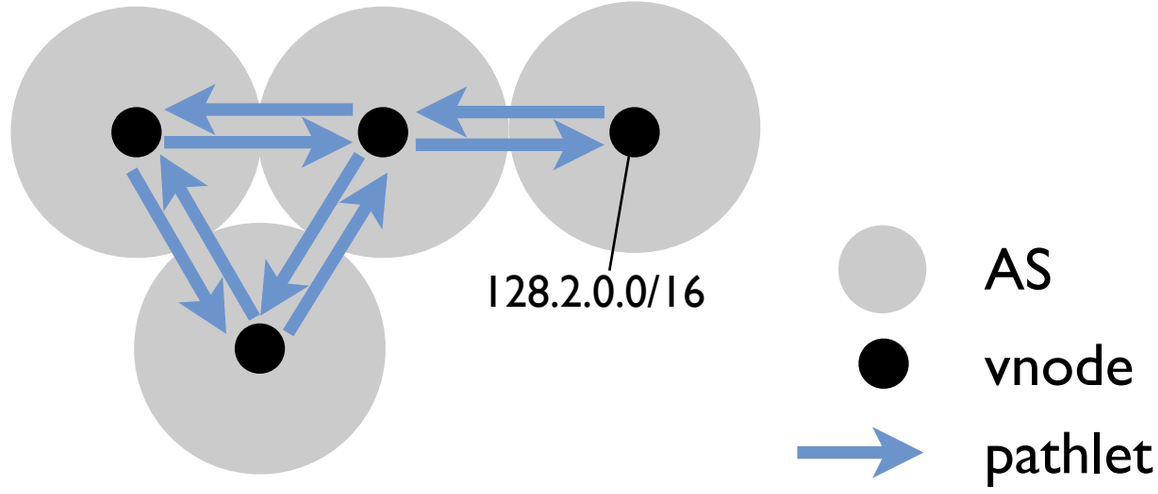
Our paper discusses several benefits of this architecture. The common theme is flexibility:

- pathlet routing can emulate the policies expressible in BGP, loose and strict source routing, MIRO, LISP, and NIRA.
- it enables "local transit policies" which have tiny forwarding tables and give sources choice over paths, while preserving some policy control.
- variation in style of policies used in different parts of the Internet is seamlessly possible, in the spirit of Clark et al.'s recommendations in their "Tussle in Cyberspace" paper.

Pathlet routing example with the simple case of a single vnode per AS, and pathlets set up to emulate BGP-style policies (i.e., allowing only a single route to the destination IP prefix):



Example which allows any route:



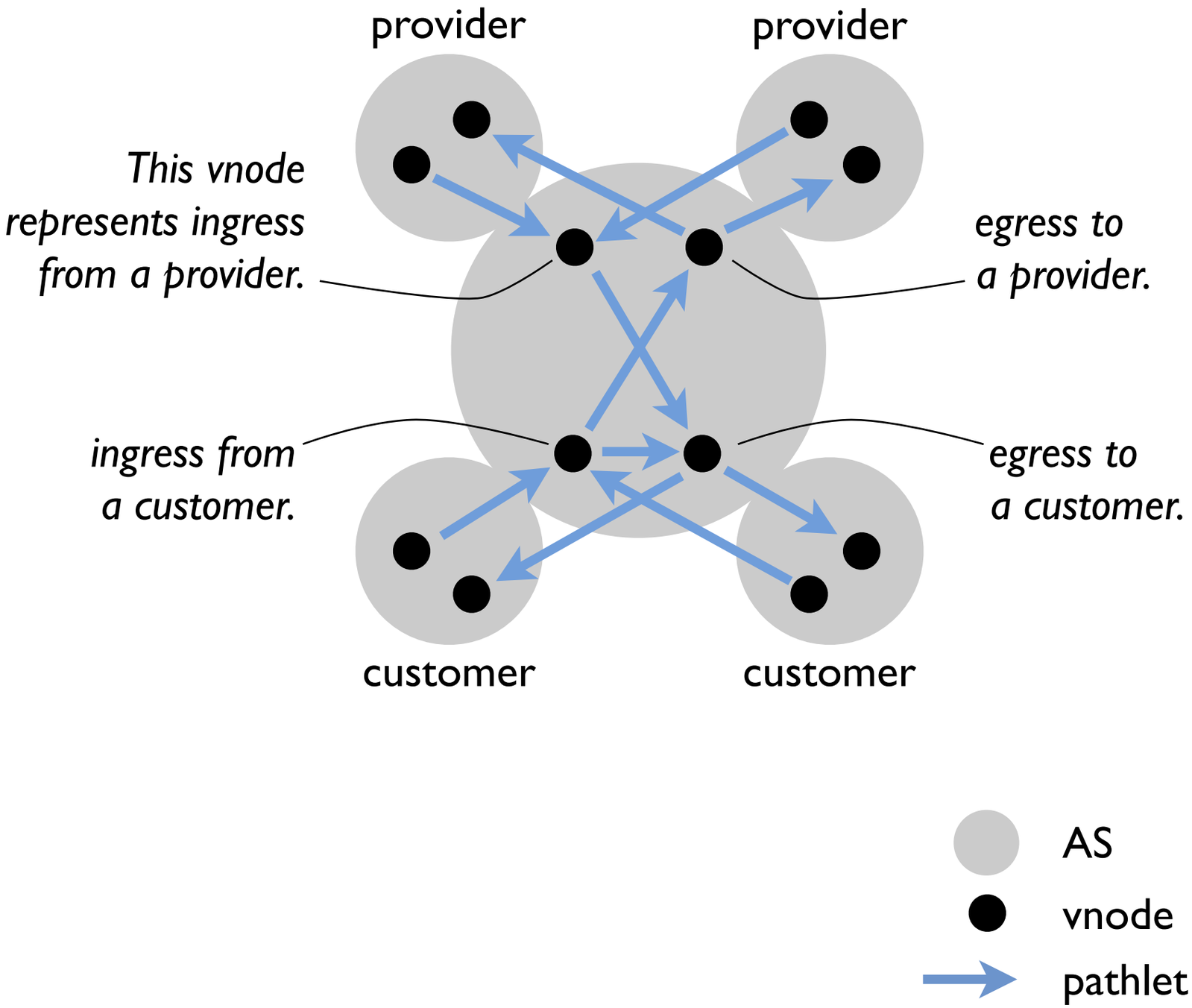
vnodes

For the AS partitioning problem, vnodes are the important concept.

Vnodes are powerful because a vnode can represent arbitrary granularities, such as an entire AS; a geographical region; or part of a VPN.

Another example: at right, we show how to represent a valley-free routing policy. Customers can route anywhere, but the AS does not provide transit service between its two providers. Subject to this, any end-to-end path is allowed.

But what if the AS partitions?



Handling partitioning

If the AS partitions, the abstraction of the previous slide no longer matches the physical topology.

Fortunately, since vnodes are entirely virtual, we can fix the problem:

1. Run a consensus algorithm to determine which routers are connected within a partition in the AS.
2. Each partition creates its own set of vnodes (with separate identifiers).

To put it another way, vnodes allow us to do on-the-fly de-aggregation when necessary.