BGP Policy violations in the data-plane

Pierre Francois, Institute IMDEA Networks
Paolo Lucente, PMACCT

pierre.francois@imdea.org
paolo@pmacct.net

Agenda

- Two well-known facts about routing...
- leading to policy violations...
- watch your network !

Observation 1

- Policy-constrained path selection in BGP...
  Flexible, per-prefix granularity

- *"A BGP-router's* ***route processor*** *will pick a path towards a given destination* ***prefix*** *by applying the following rules"*

  Weight
  Local-pref
  As Path Length
  IGP/Med
  ...

Observation 1

- ... dominated in the data-plane
- A **FIB** will pick a path towards a given destination **address** by applying the following rules

  Longest prefix match to get the prefix

  (
  Best path towards that prefix was picked based on
  Weight
  Local-pref
  As Path Length
  IGP/Med
  ...)

Observation 11

- Common to provide a lot of routing flexibility
- Route propagation control offered by Sprint
  - Have to be a customer of Sprint
  - 65000:XXX : Do not advertise to ASXXX
    can be AOL, NTT, BT, Level3, GBLX, Verizon, AT&T, ...

Powerful complementary means to limit path knowledge

- Selective advertisement, performed locally
- Selective advertisement, triggered remotely

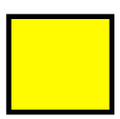
Control-plane/Data-plane can mismatch

- Paths for **overlapping** prefixes are controlled independently
  - By yourself
  - By your BGP neighborhood
- Forwarding plane dominated by the longest prefix match rule

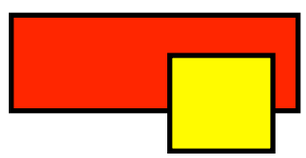
- What if your policy differs for overlapping prefixes ?

Toy case study

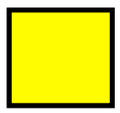
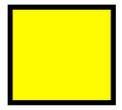
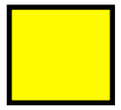
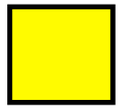
A BGP advertisement for NLRI P/p



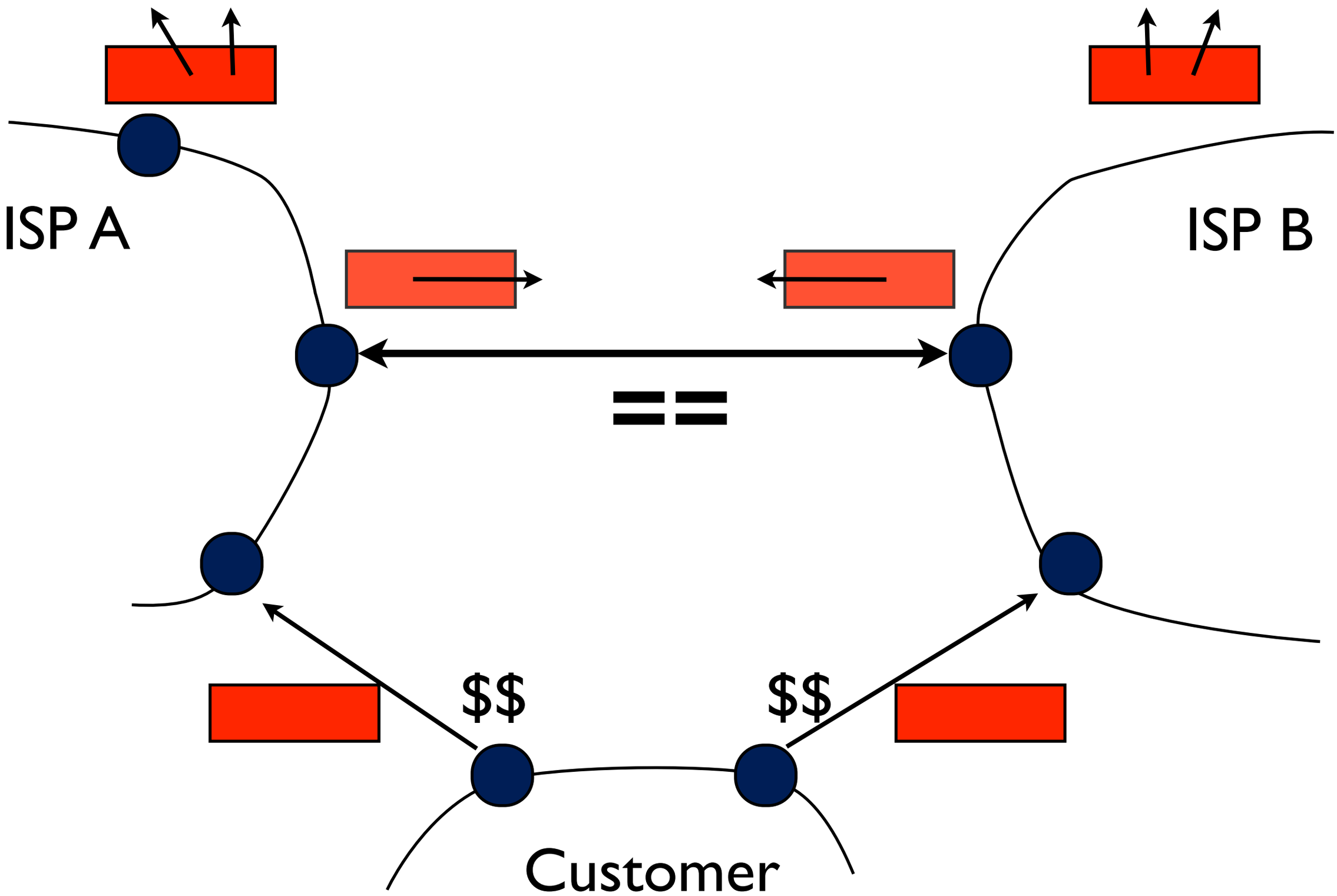
A BGP advertisement of a prefix more specific than P/p, say P/p+1



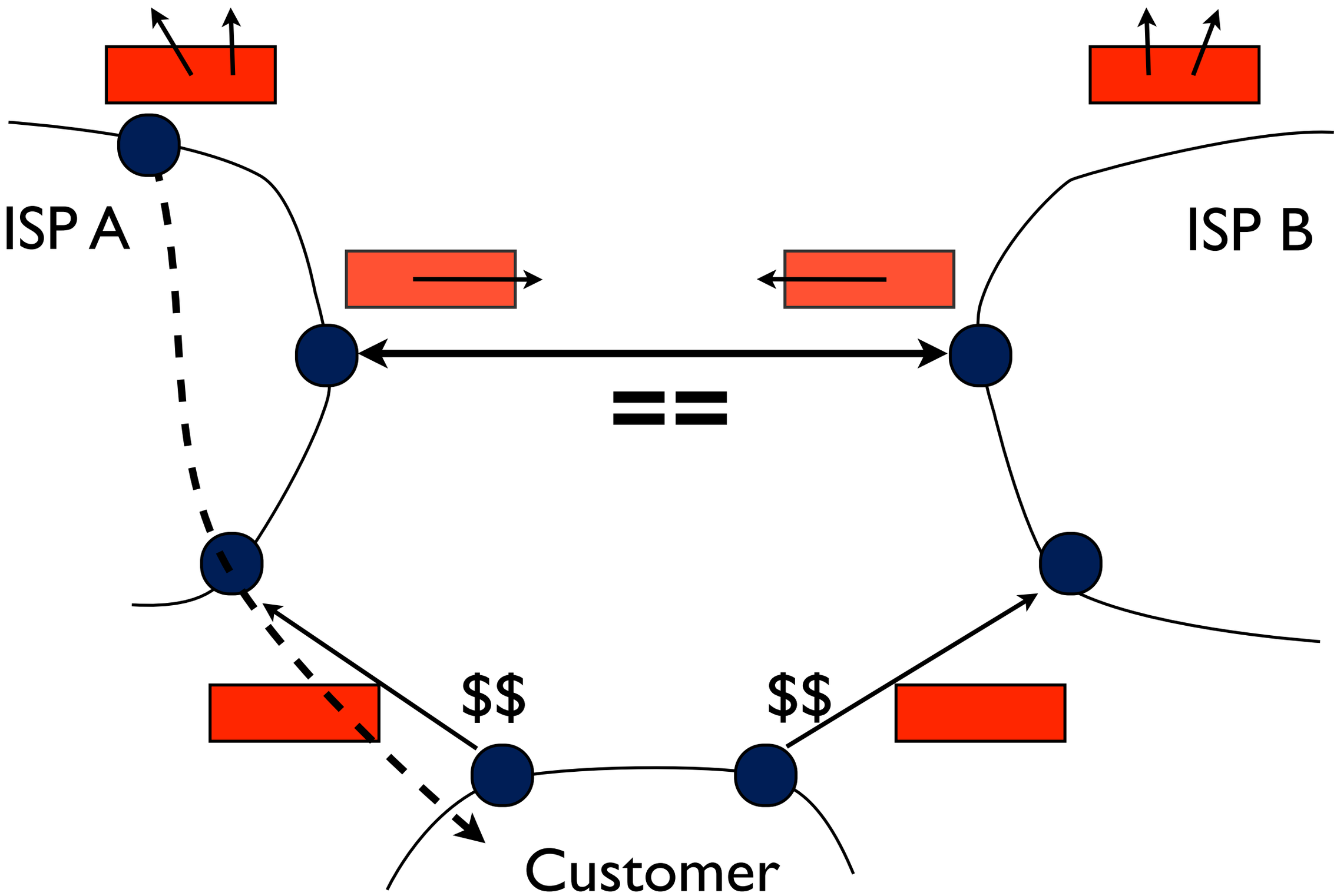
The BGP policy violation trick

- Play with  and communities
- Make  reach only a subset of the ASes
  - Some ASes forward  according to 
  - Until packet reaches an AS knowing 
  - Resulting data-plane not necessarily fitting everyone's policy...

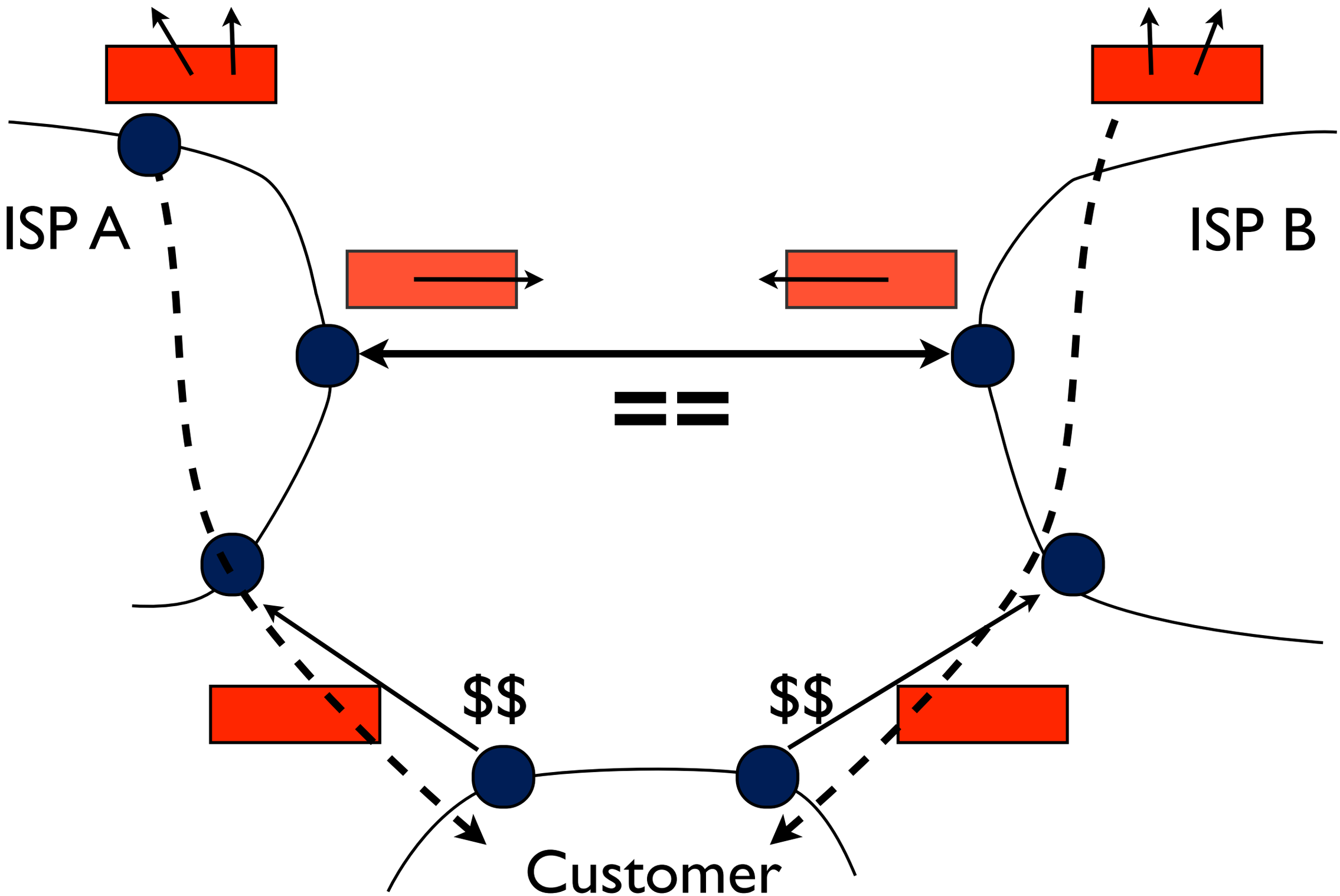
Initial routing status



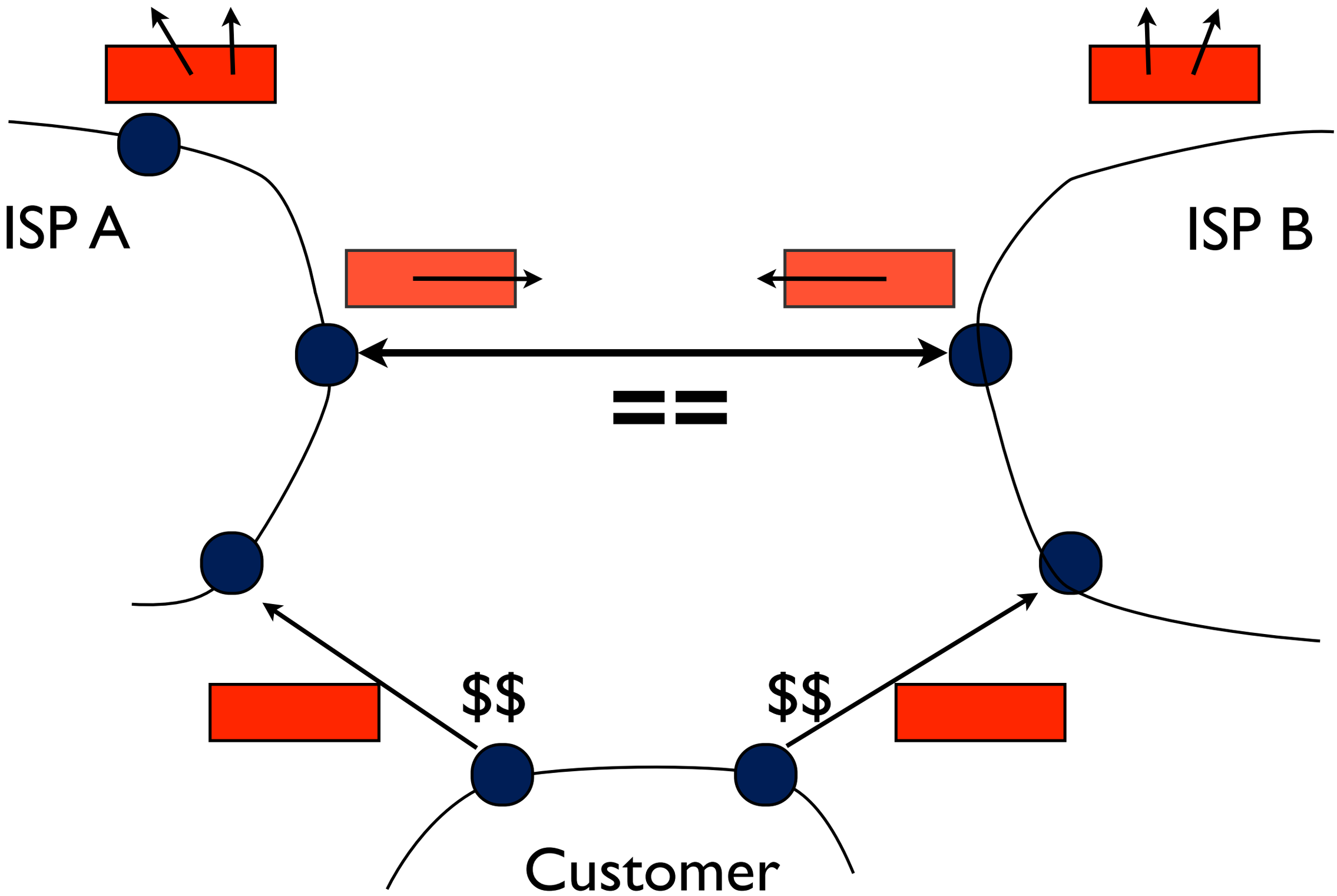
Initial routing status



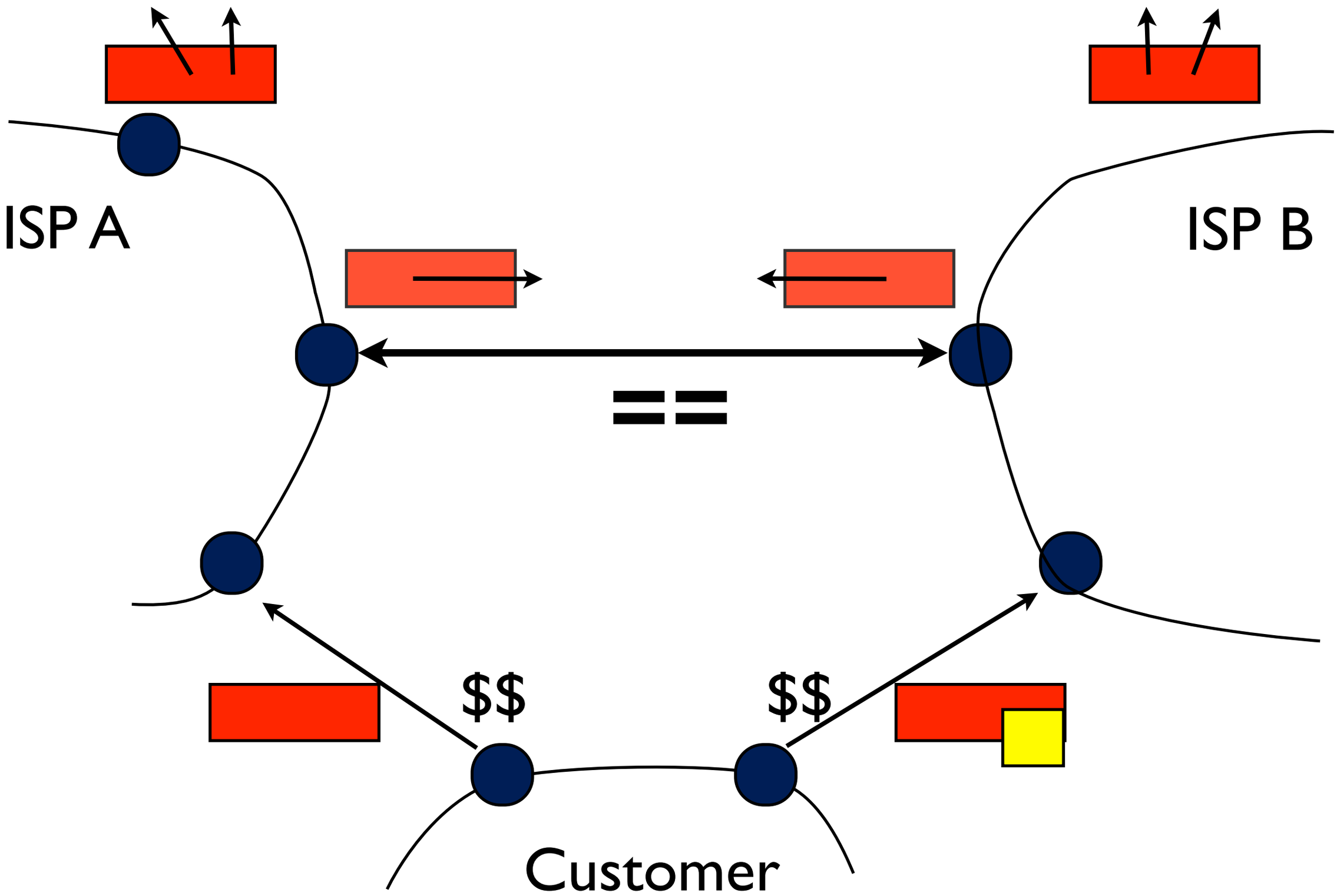
Initial routing status



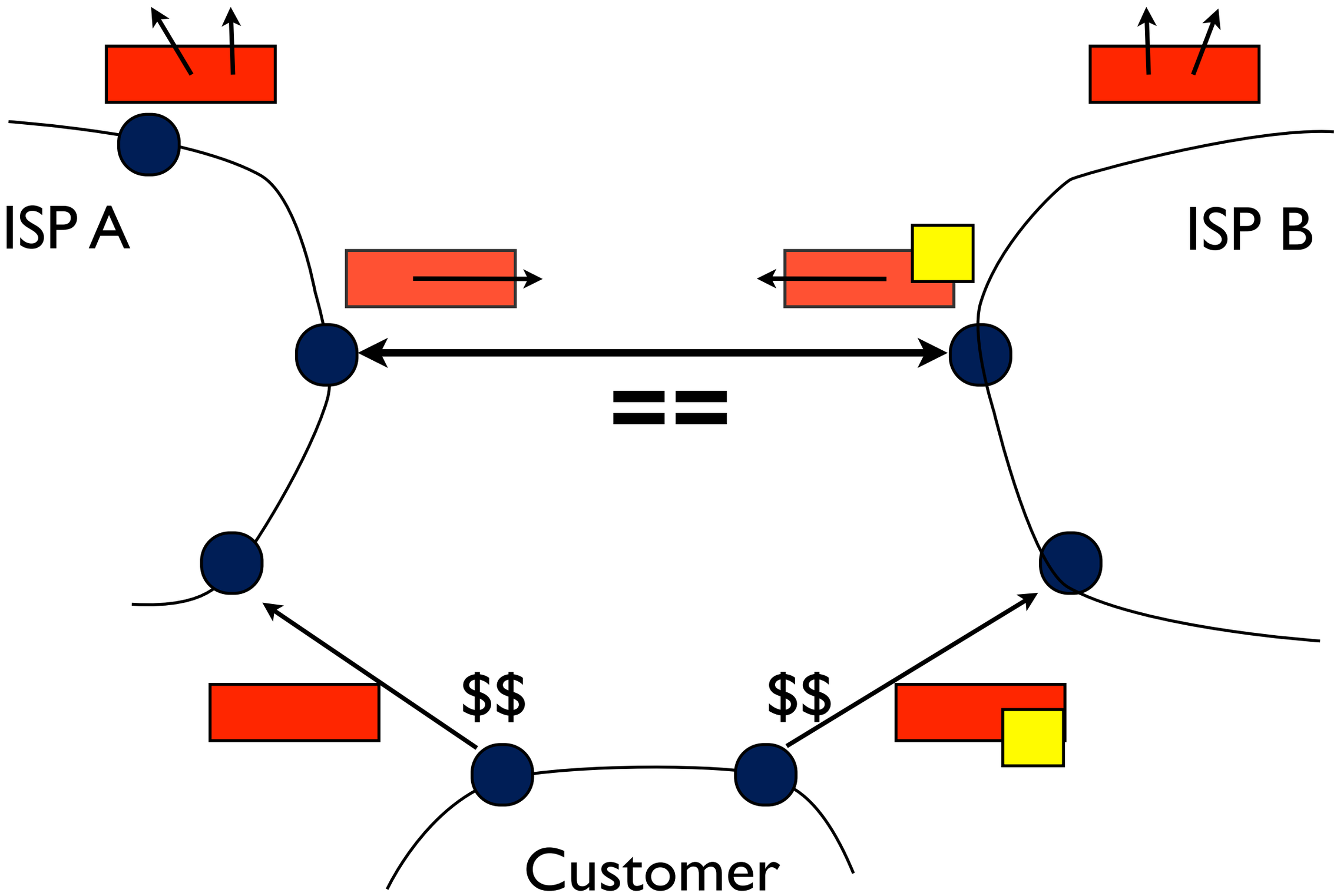
Inbound TE, selective advertisement of a more specific prefix



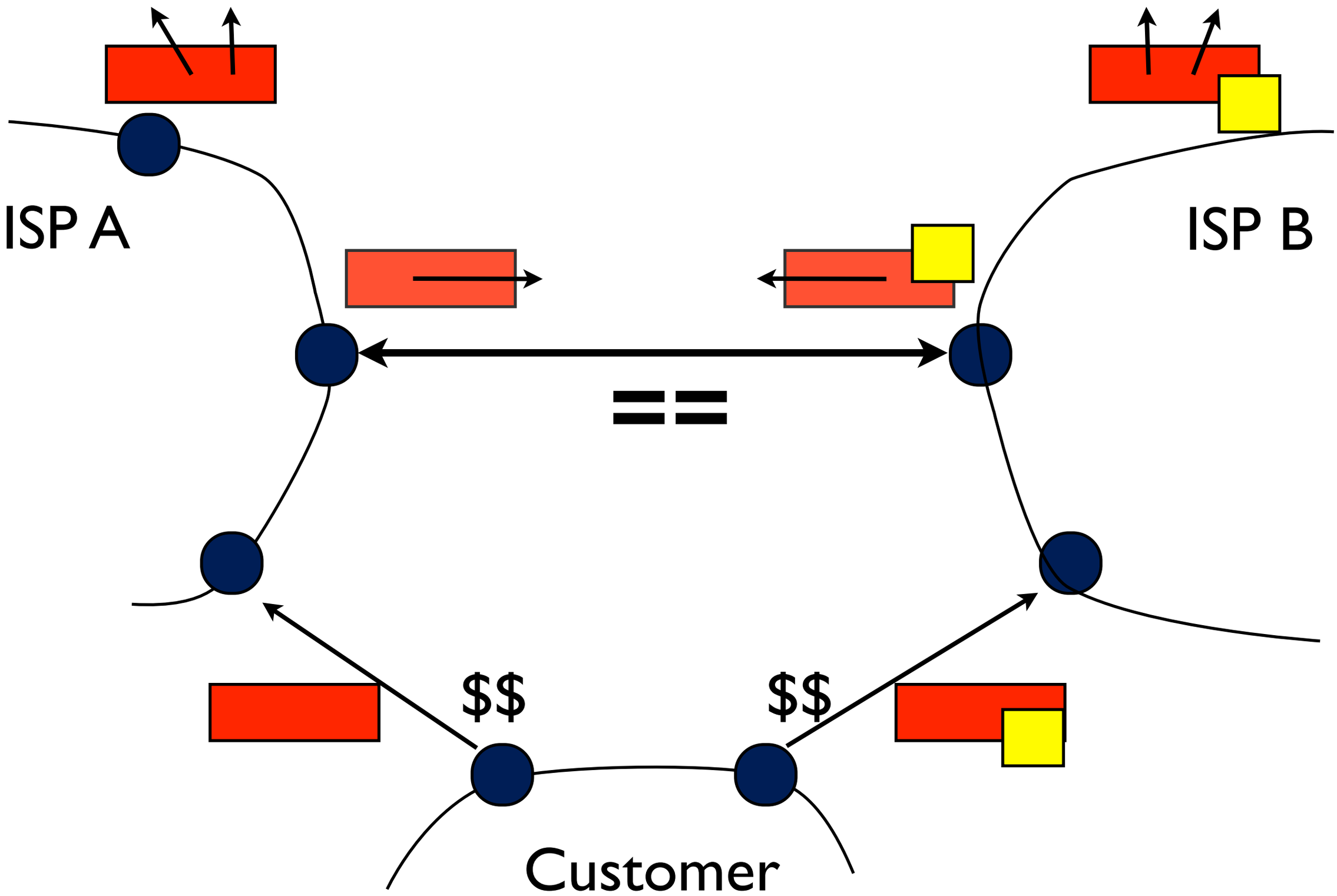
Inbound TE, selective advertisement of a more specific prefix



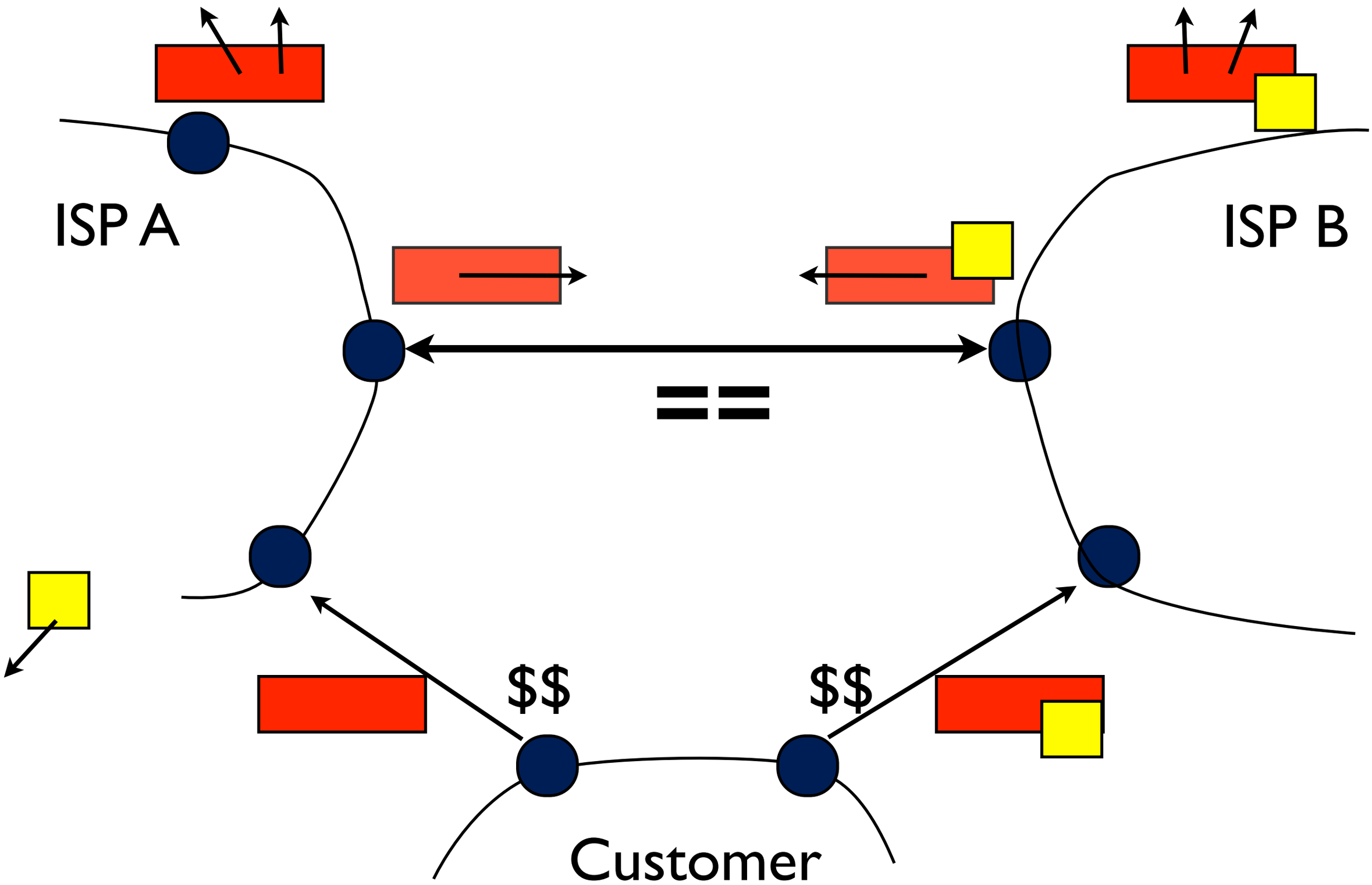
Inbound TE, selective advertisement of a more specific prefix



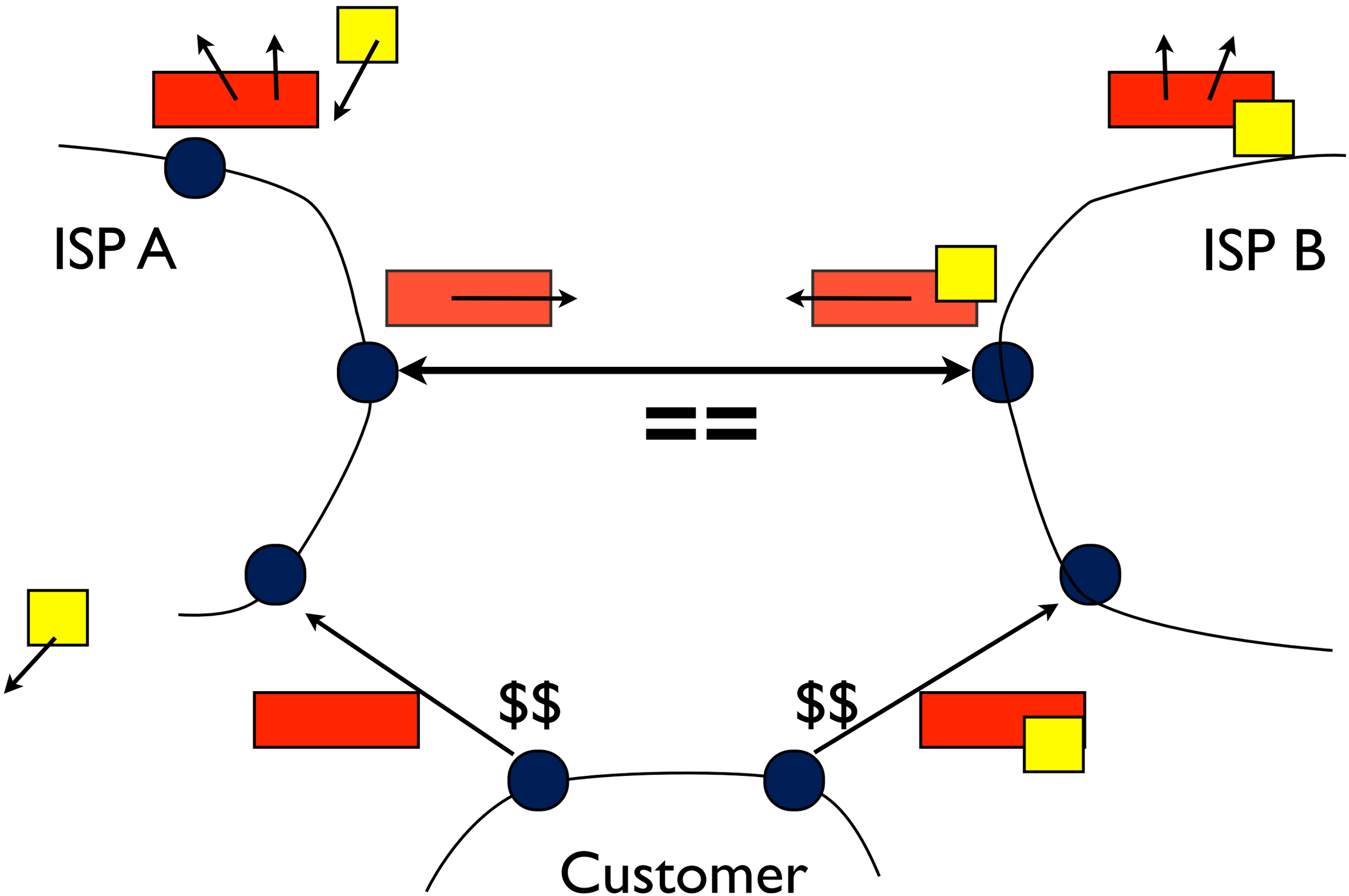
Inbound TE, selective advertisement of a more specific prefix



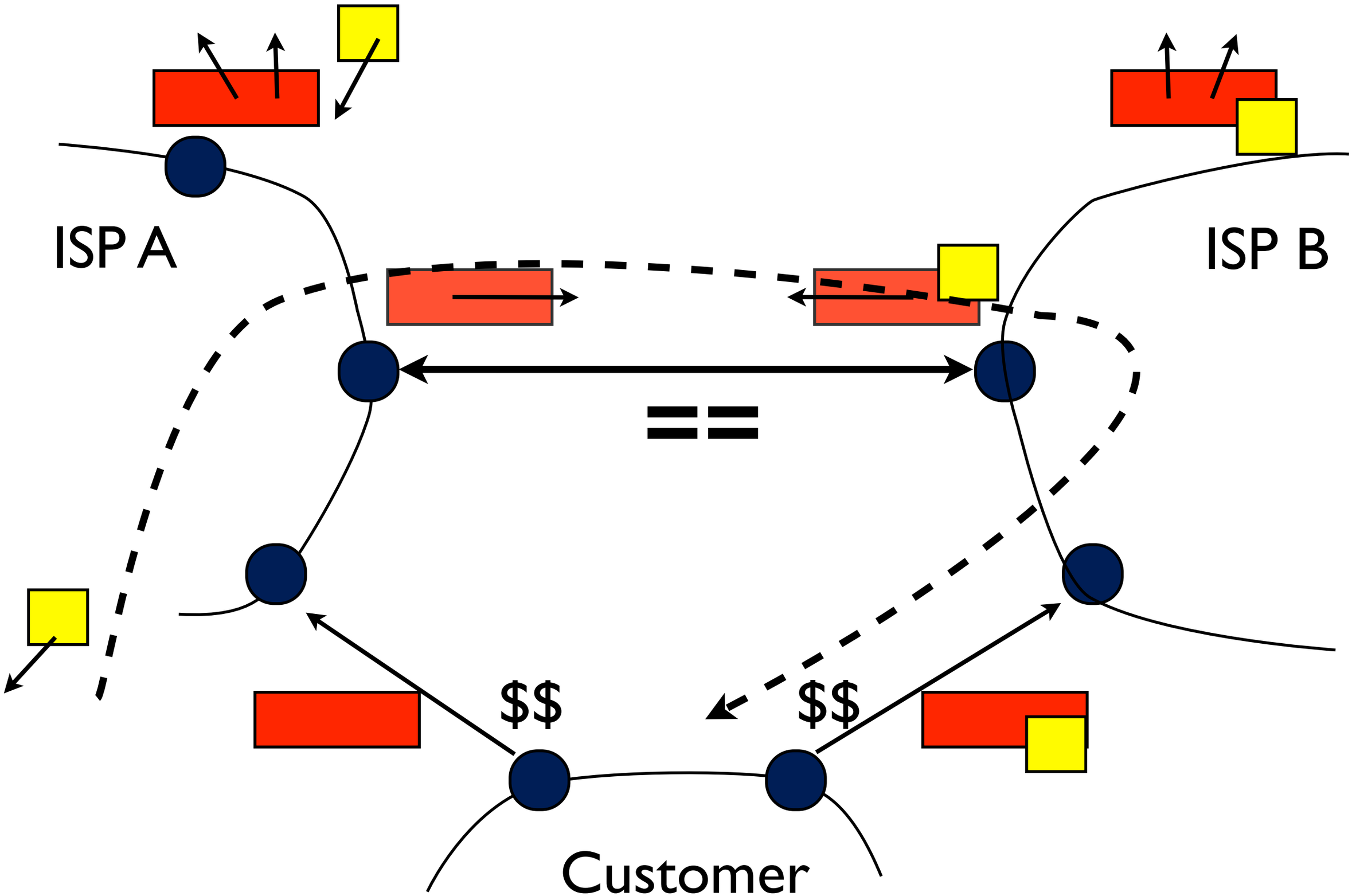
Inbound TE, selective advertisement of a more specific prefix



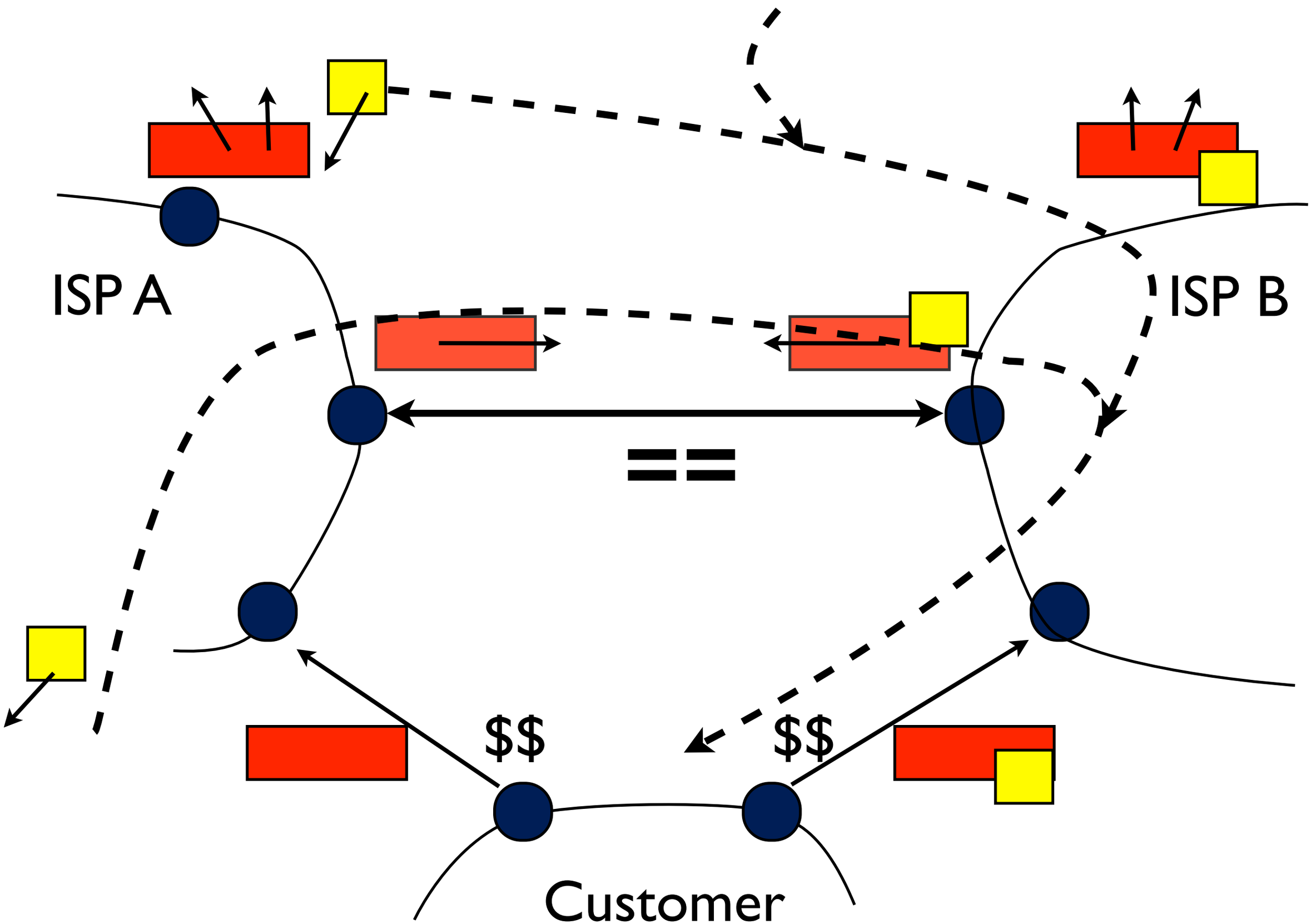
Inbound TE, selective advertisement of a more specific prefix



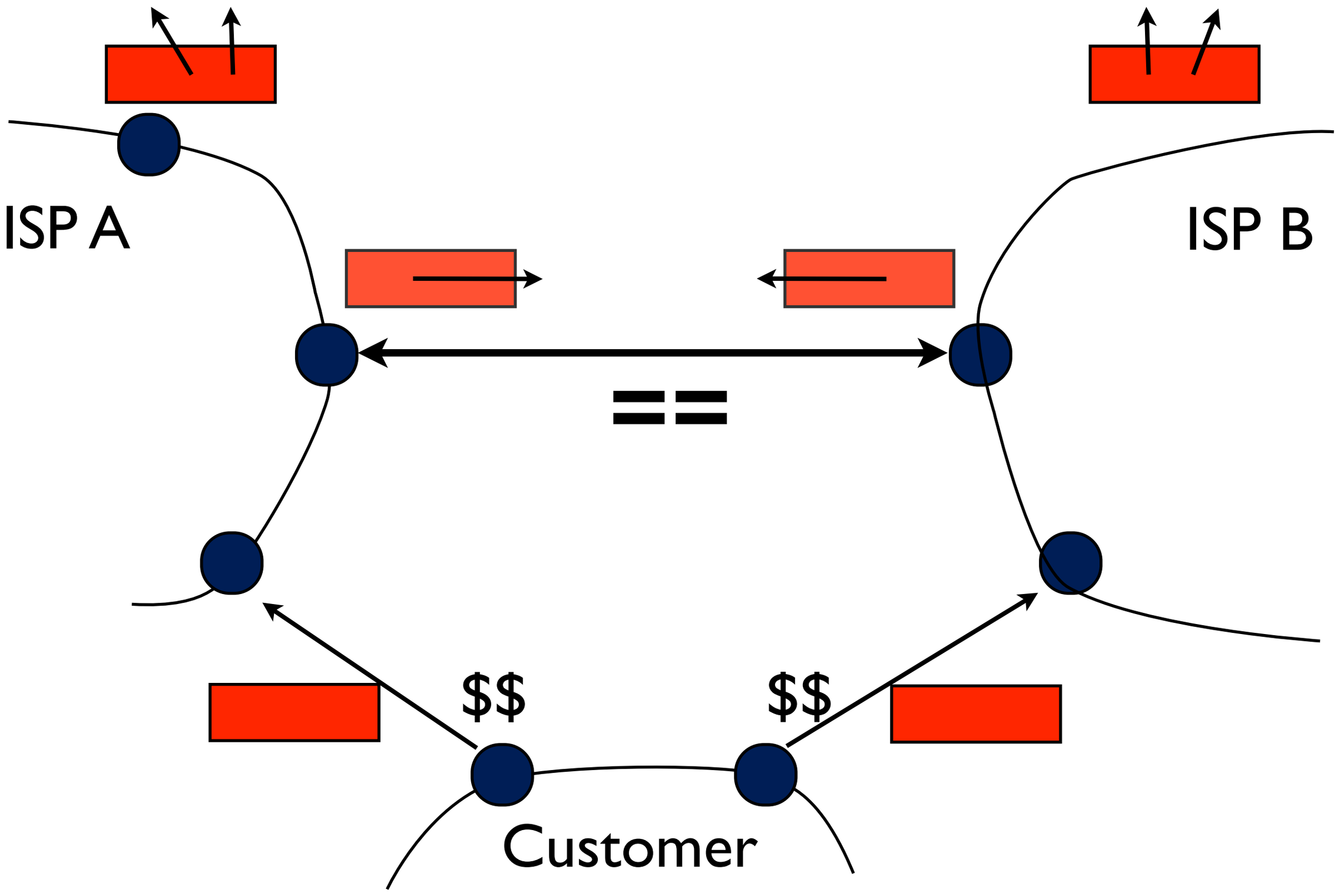
Inbound TE, selective advertisement of a more specific prefix



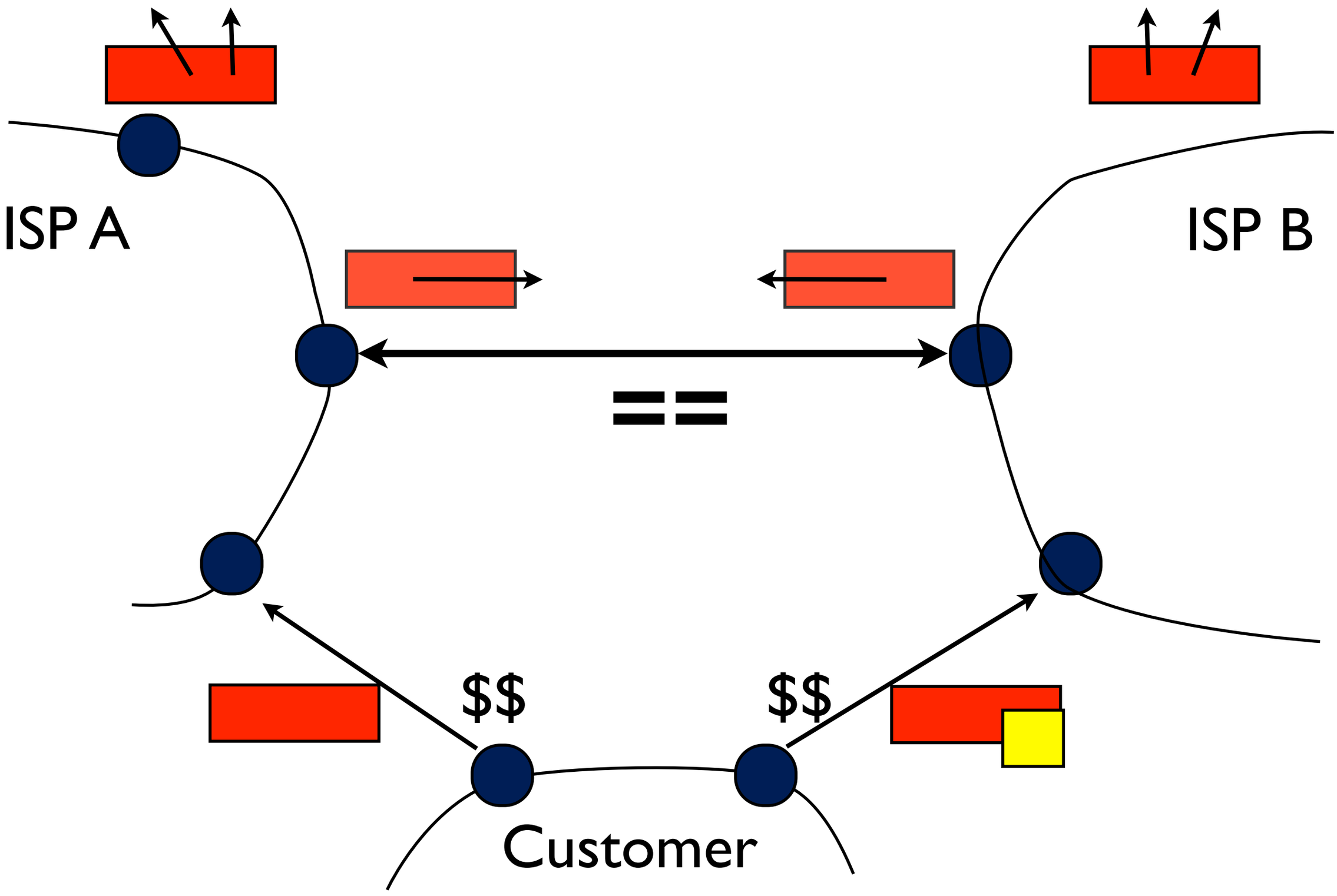
Inbound TE, selective advertisement of a more specific prefix



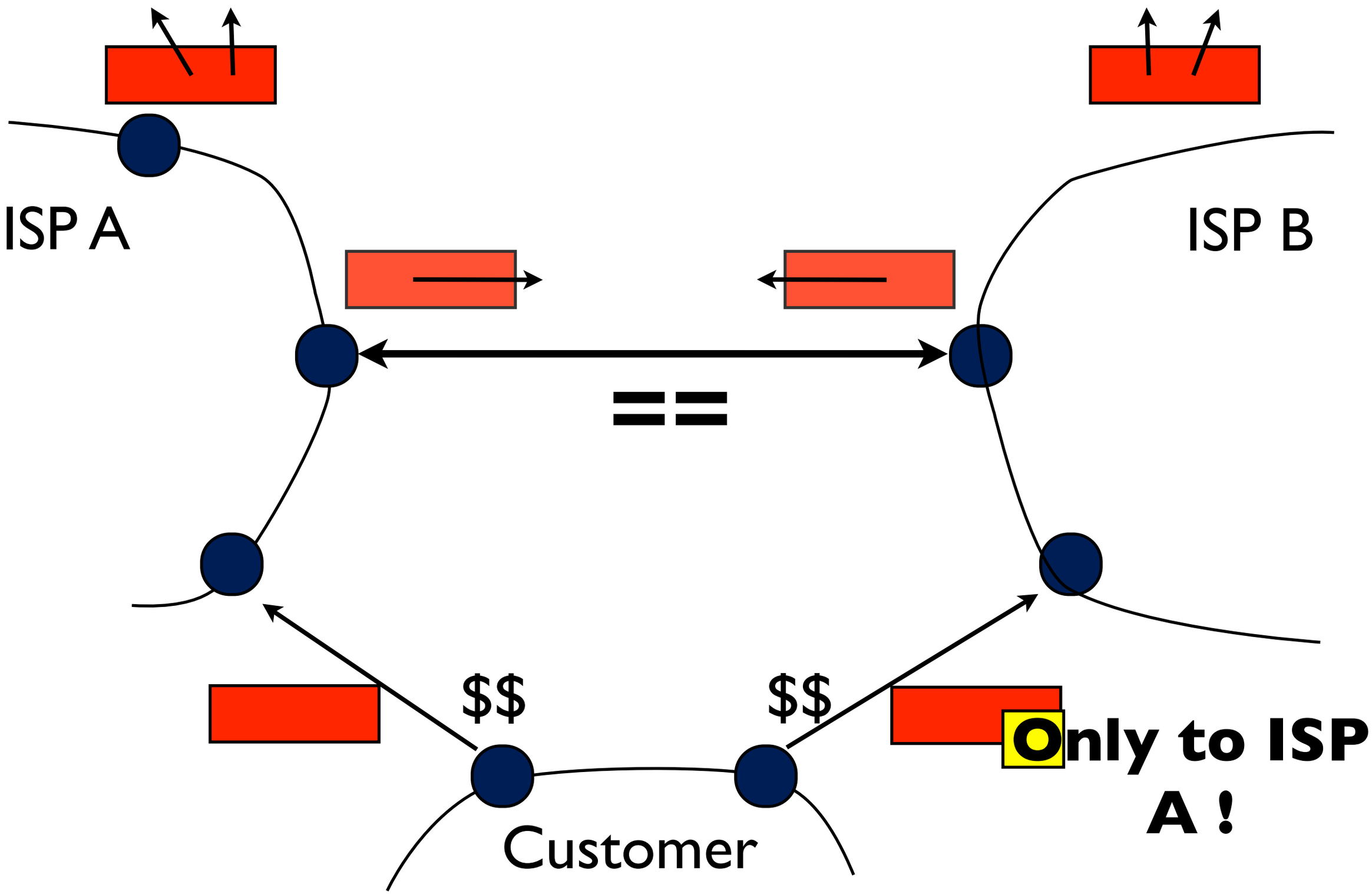
Scope the advertisement of the more specific



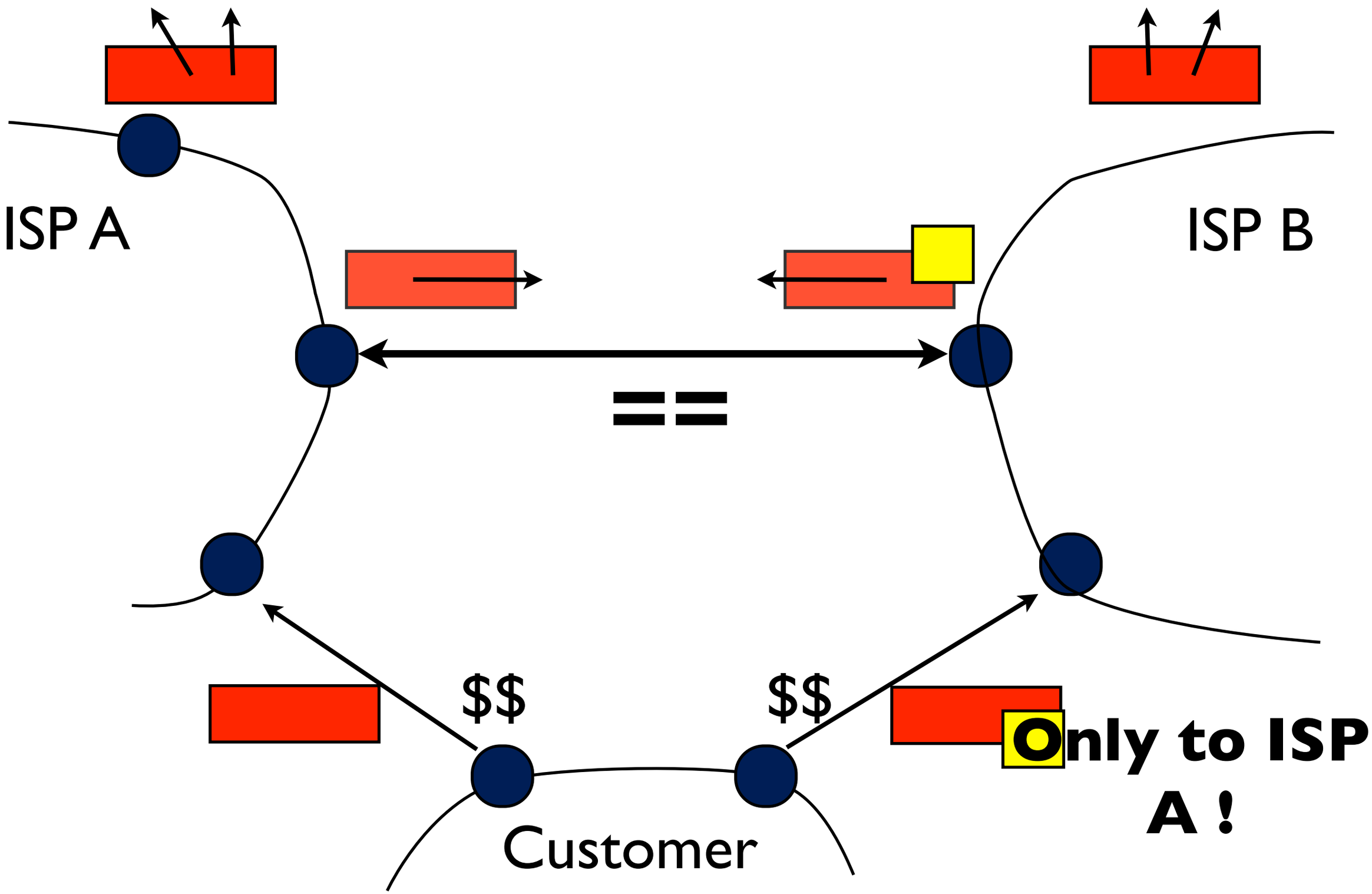
Scope the advertisement of the more specific



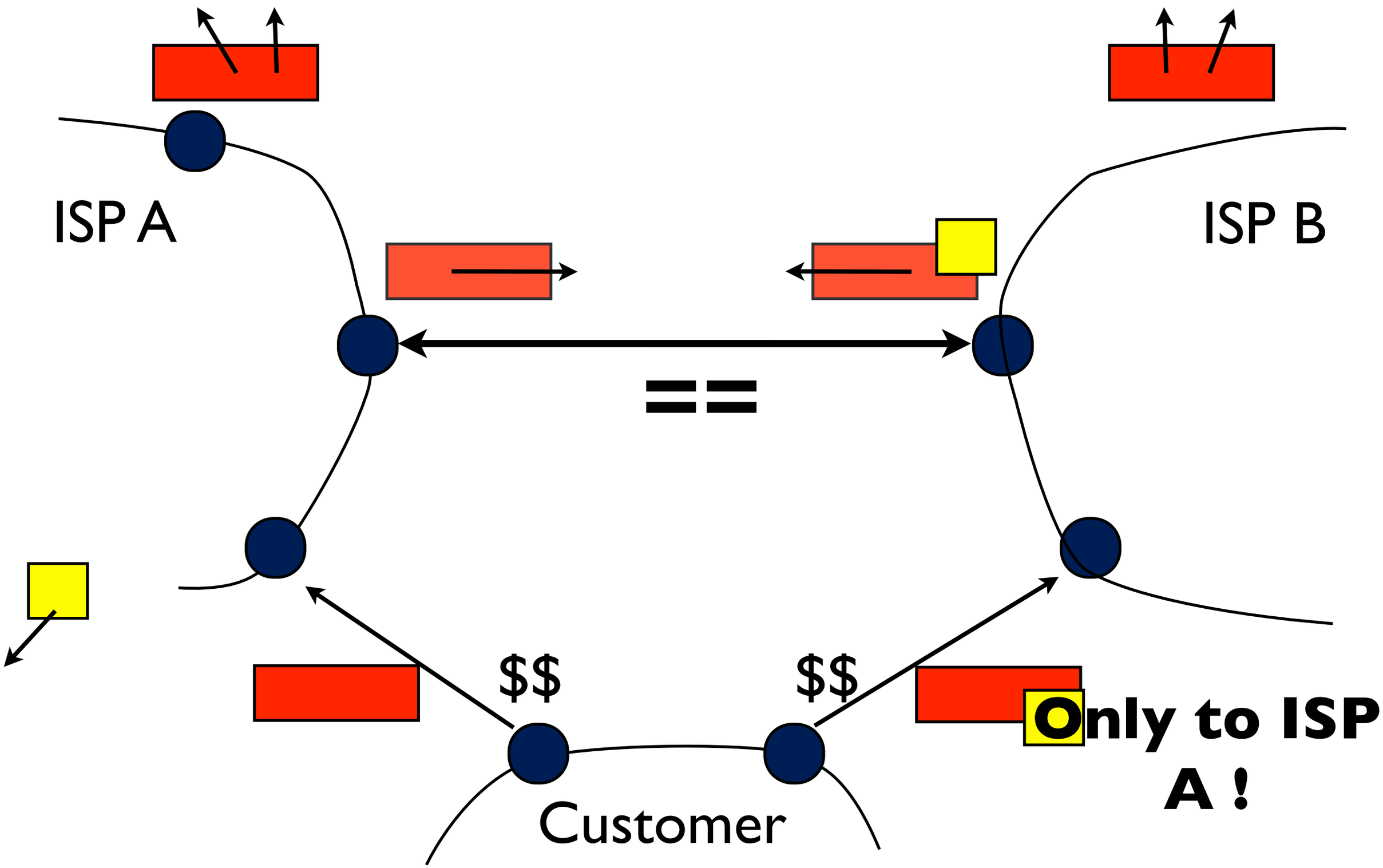
Scope the advertisement of the more specific



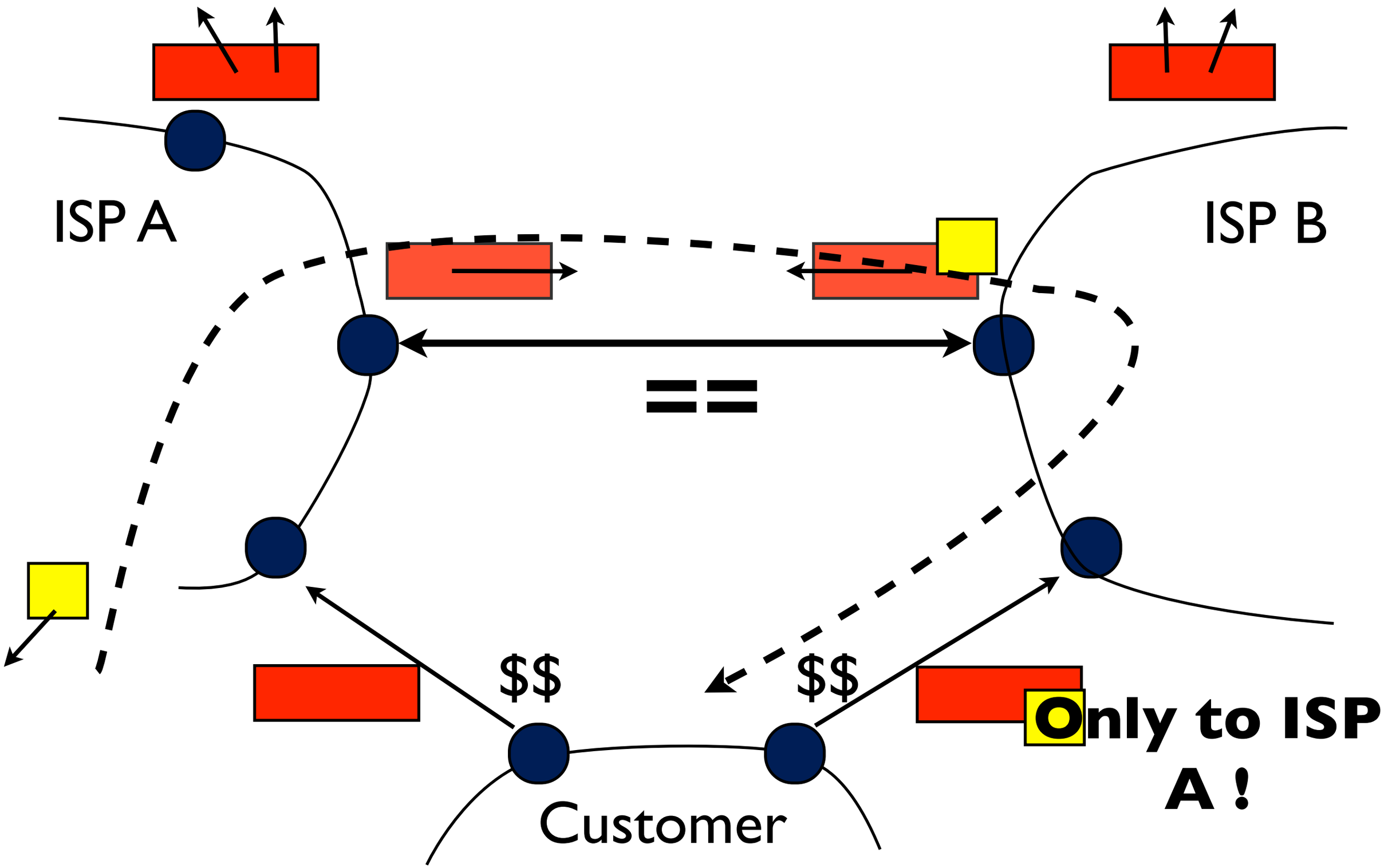
Scope the advertisement of the more specific



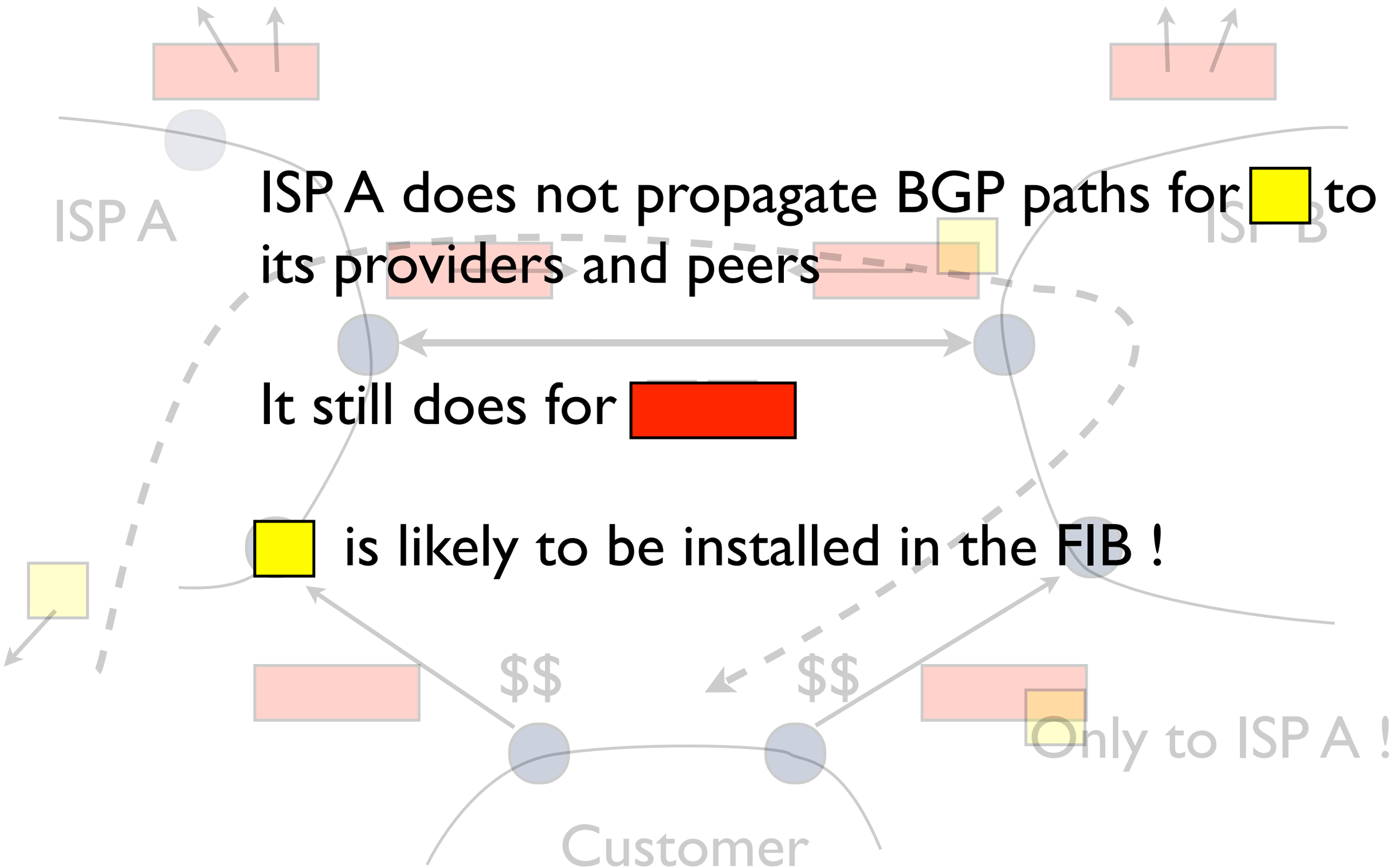
Scope the advertisement of the more specific



Scope the advertisement of the more specific



Let's start playing : Scope advertisement of the more specific

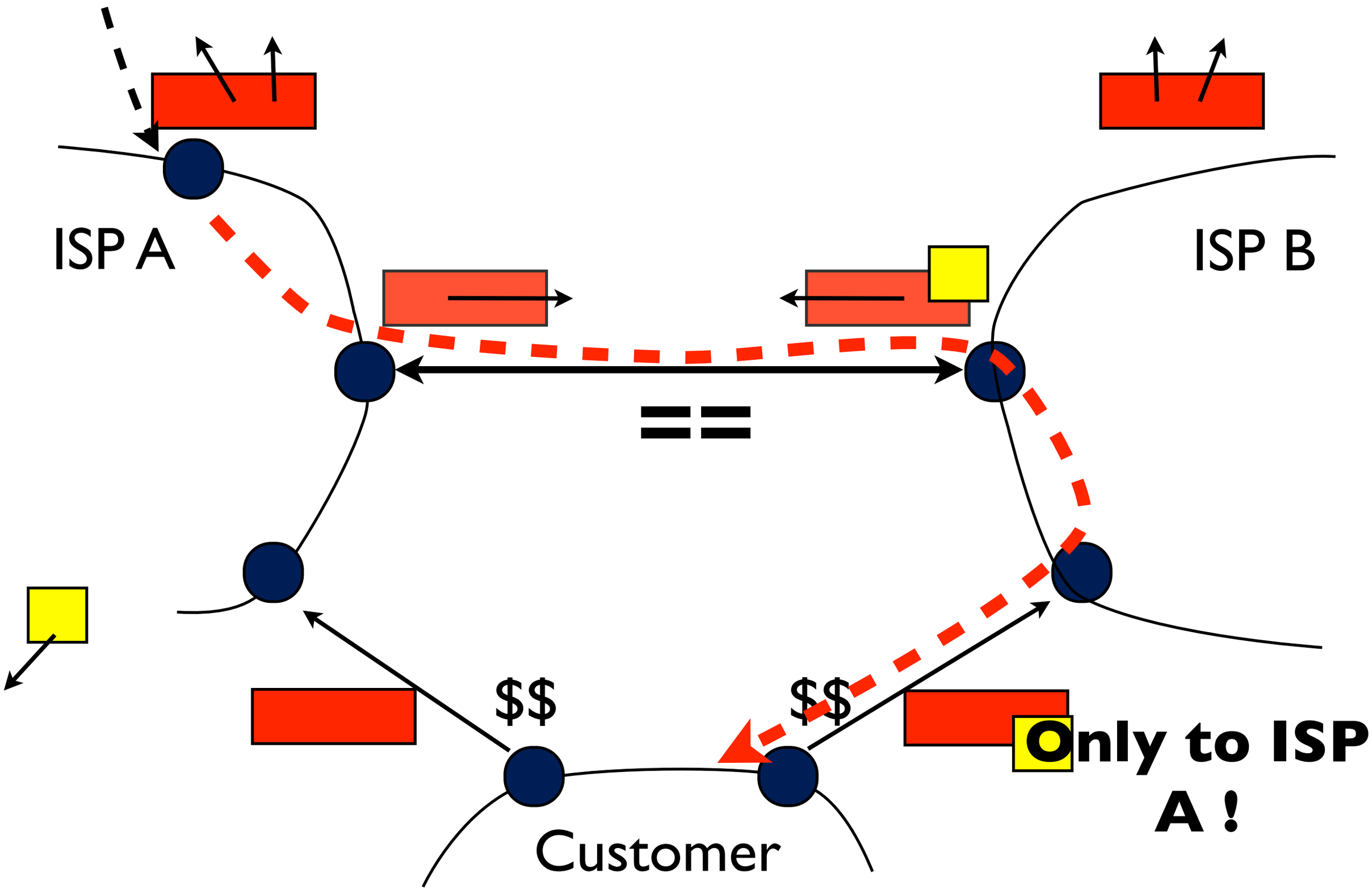


ISP A does not propagate BGP paths for to its providers and peers

It still does for

is likely to be installed in the FIB !

New path in the network



This is annoying

- Your policies can be violated
- Your flexible routing service can turn **you** into a transit thief when misused by **your** customers
- “Nothing breaks” when the violation takes place
- Ex. : Just consider the Tier-I clique...

So what can you do ?

- Forward differently
- Filter-out / Drop
- Monitor !

Forwarding differently

- Deploy BGP so as to have forwarding at an incoming interface solely based on policy fitting paths
  - Put the Internet in VRFs
  - Careful configuration of import rules
- Complex, Costly

Filtering out / Drop

- Drop packets, at ingress, for routes that are not supposed to be served there
  - Assume malicious behavior by default
  - Interrupts service from/to customers
- Filter out, at egress
  - Range served as if the msp did not exist

Monitor

- You got the means to monitor ingress-egress traffic demand to run your business, right ?
- “Just” check if counters for non-policy compliant transit
  - Pick the phone when counters are not at 0
  - Filter-out if the issue is not getting fixed early enough
- Seems like few operators run the check

PMACCT

- Tool developed by Paolo Lucente
  (See talk at RIPE 61 plenary)
- Policy violation check is a matter of a couple of lines

  http://wiki.pmacct.net/DetectingRoutingViolations
- Tools integrating with pmacct can benefit from this work
  (ie. Cariden)

Thanks !