draft-francois-limited-scope-specifics-00

Pierre Francois
Bruno Quoitin

Goal(1)

- end-to-end control-plane/data-plane inconsistency can occur when your neighbors play with limited scope more specific prefixes
- Leads to BGP policy violations

Goal(2)

- Document solutions to the problem
  - Detection
  - Re-action
  - Anticipation (?)

Observation 1

- Data plane is often disregarded when thinking about BGP
- *"A BGP router will pick a path towards a given destination by applying the following rules"*

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

Observation 1

- Data plane is often disregarded when thinking about BGP
- *“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
  ...

Think FIB

- Traffic follows **data-plane** state

- 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 II

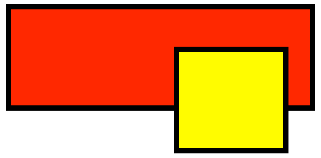
Typical recognized BGP community values

- If you are my customer or a customer of my customers, you can tag
- 65000:XXX : Do not advertise to ASXXX

Legend

🟥 A BGP Prefix advertisement for p/P

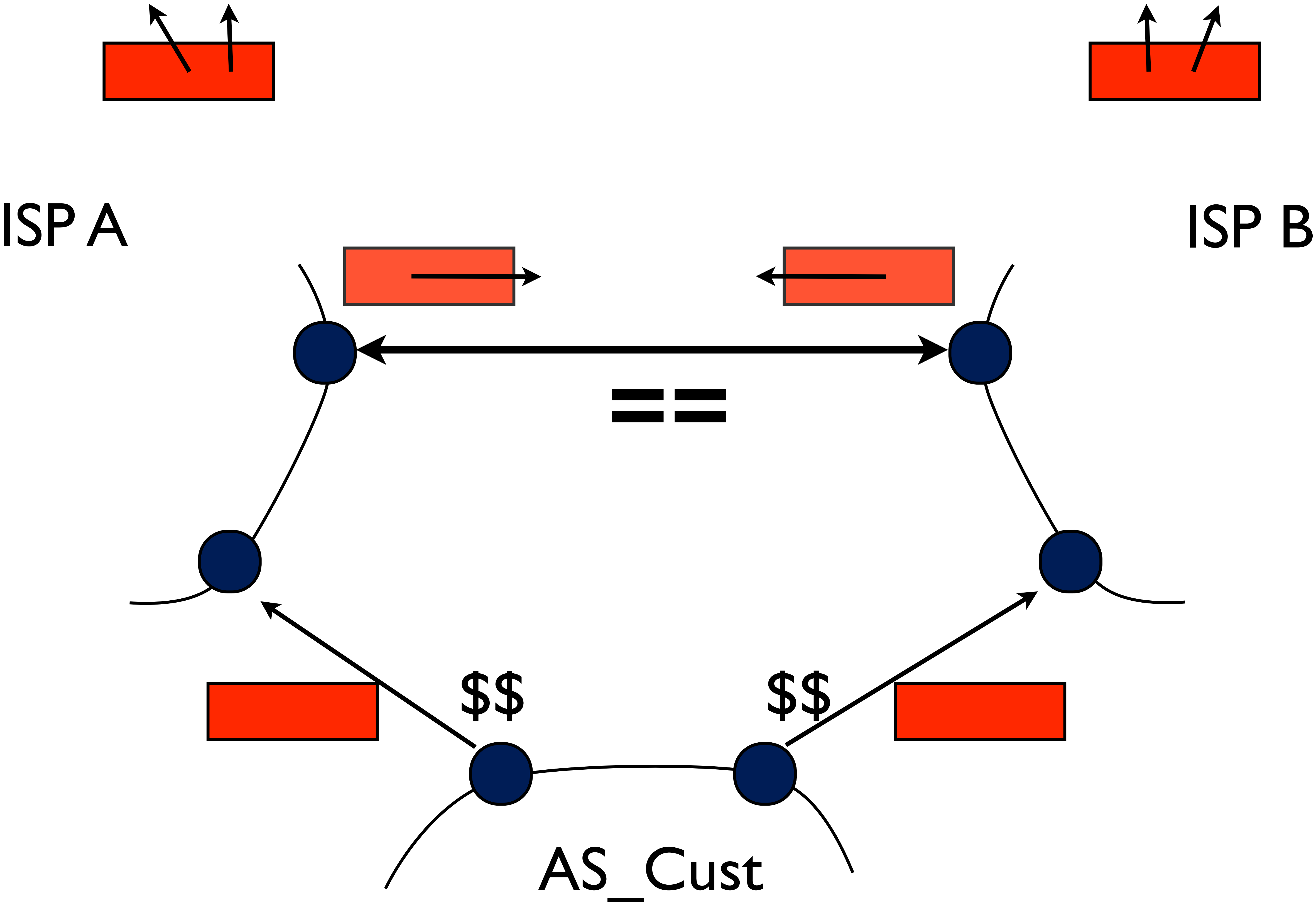
🟨 An advertisement of a prefix more specific than p/P, say p/P+2



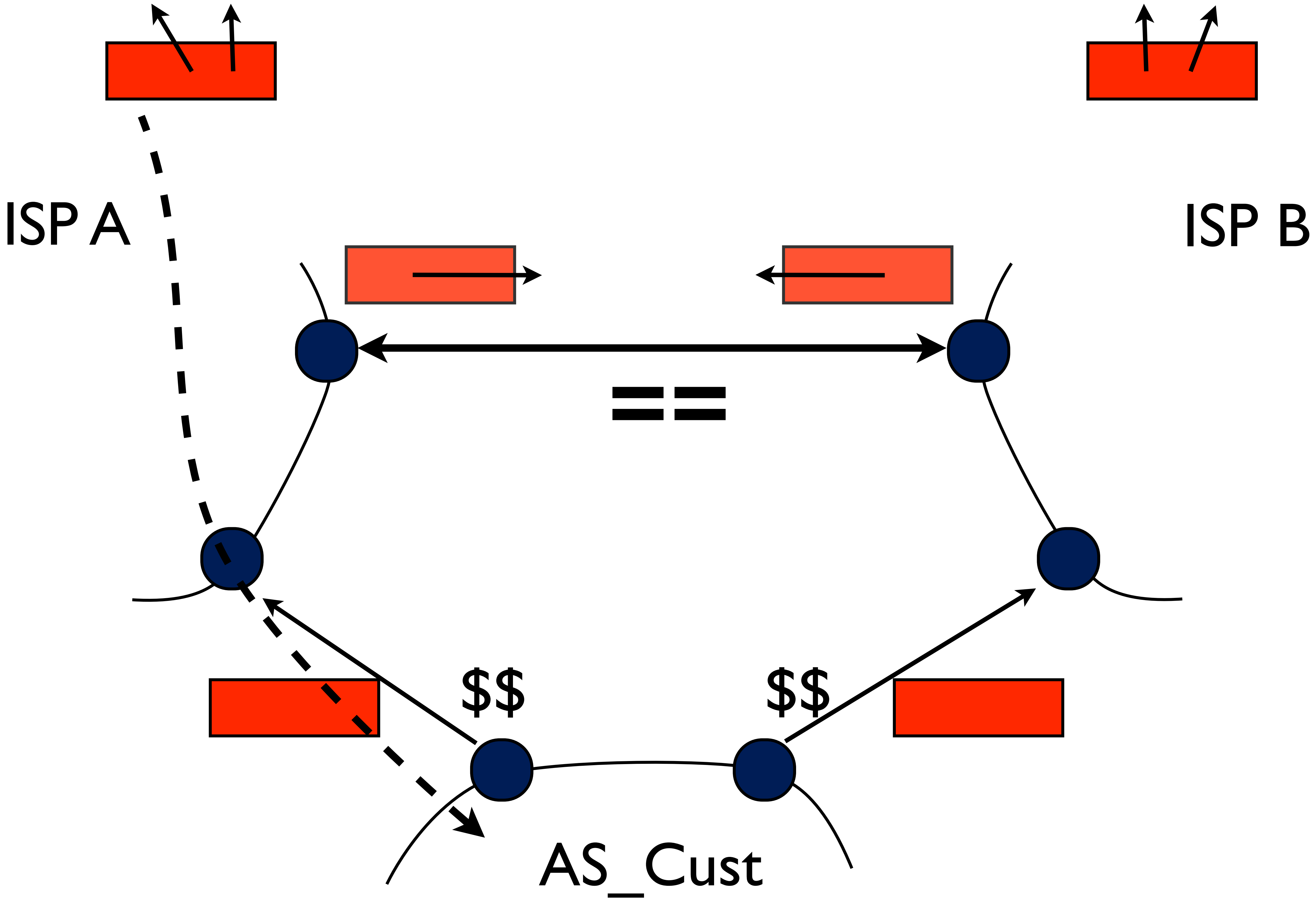
What can you do with these communities ?

- Play with 🟨
- Assume A and B are providers of AS_Cust
- B allows such community tagging
- A and B are peers
- AS_Cust turns “don’t advertise to AS X” values into a only “advertise to A”
  Just put them all but A

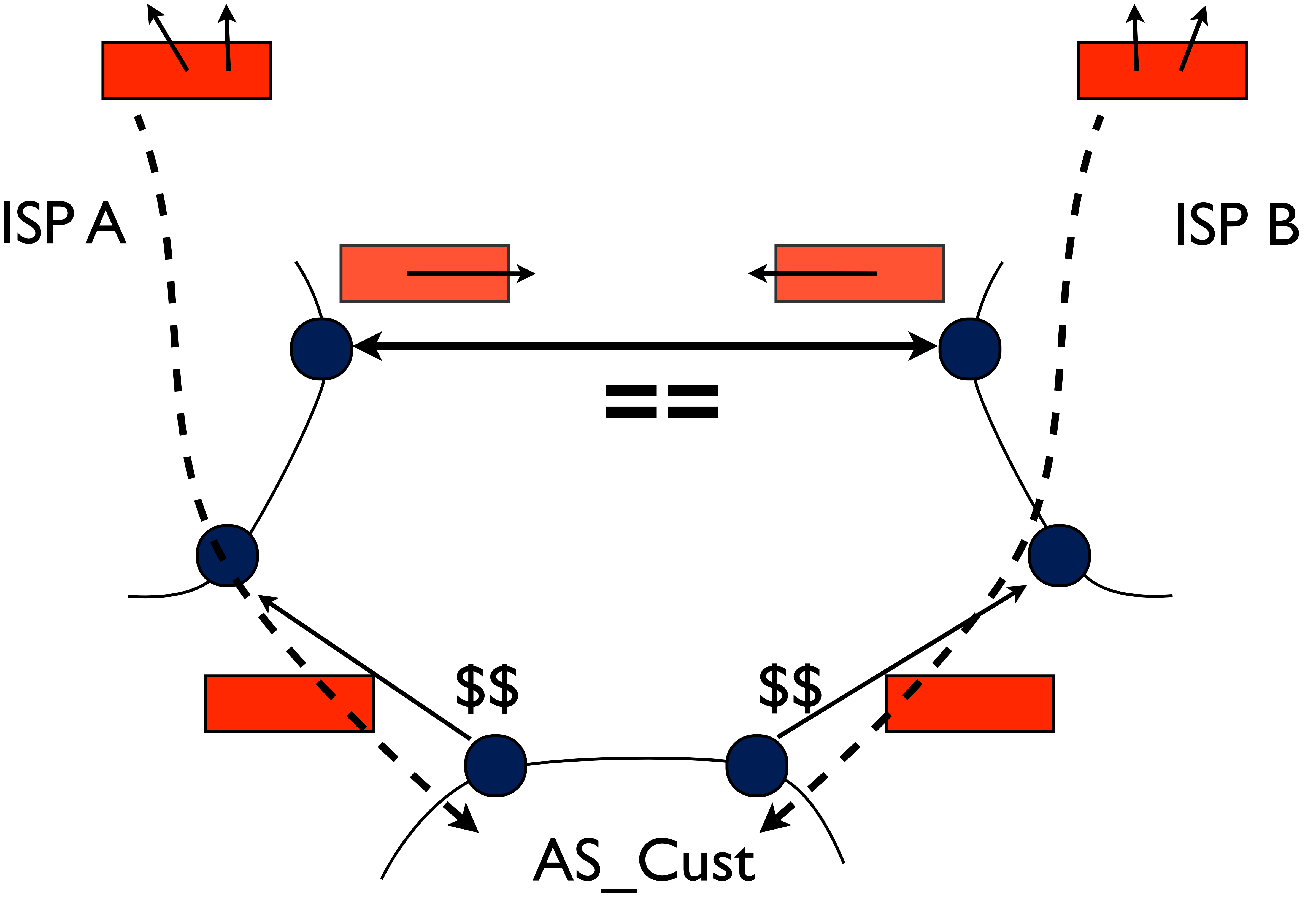
Initial routing status
control-plane (only) driven forwarding



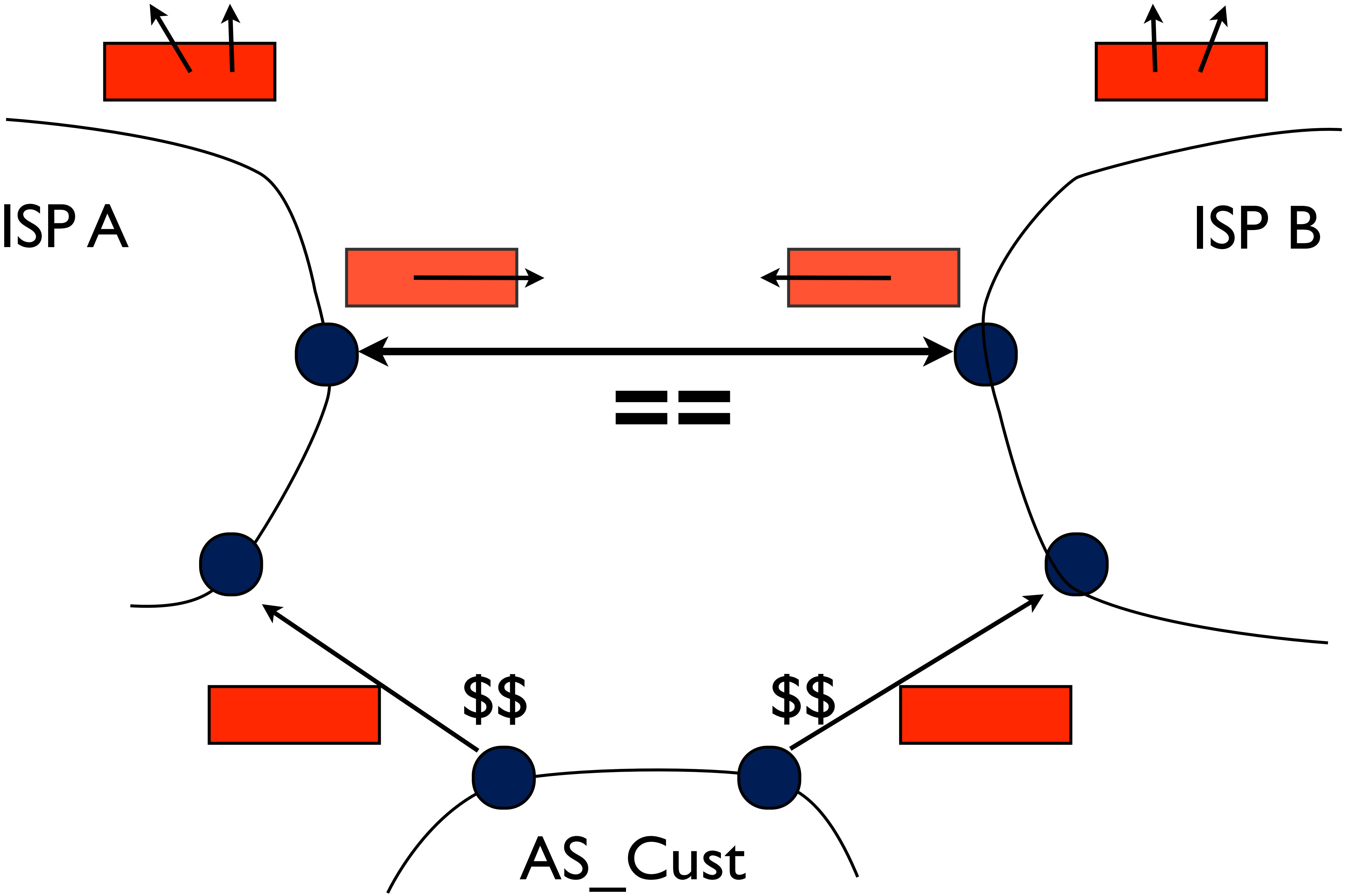
Initial routing status
control-plane (only) driven forwarding



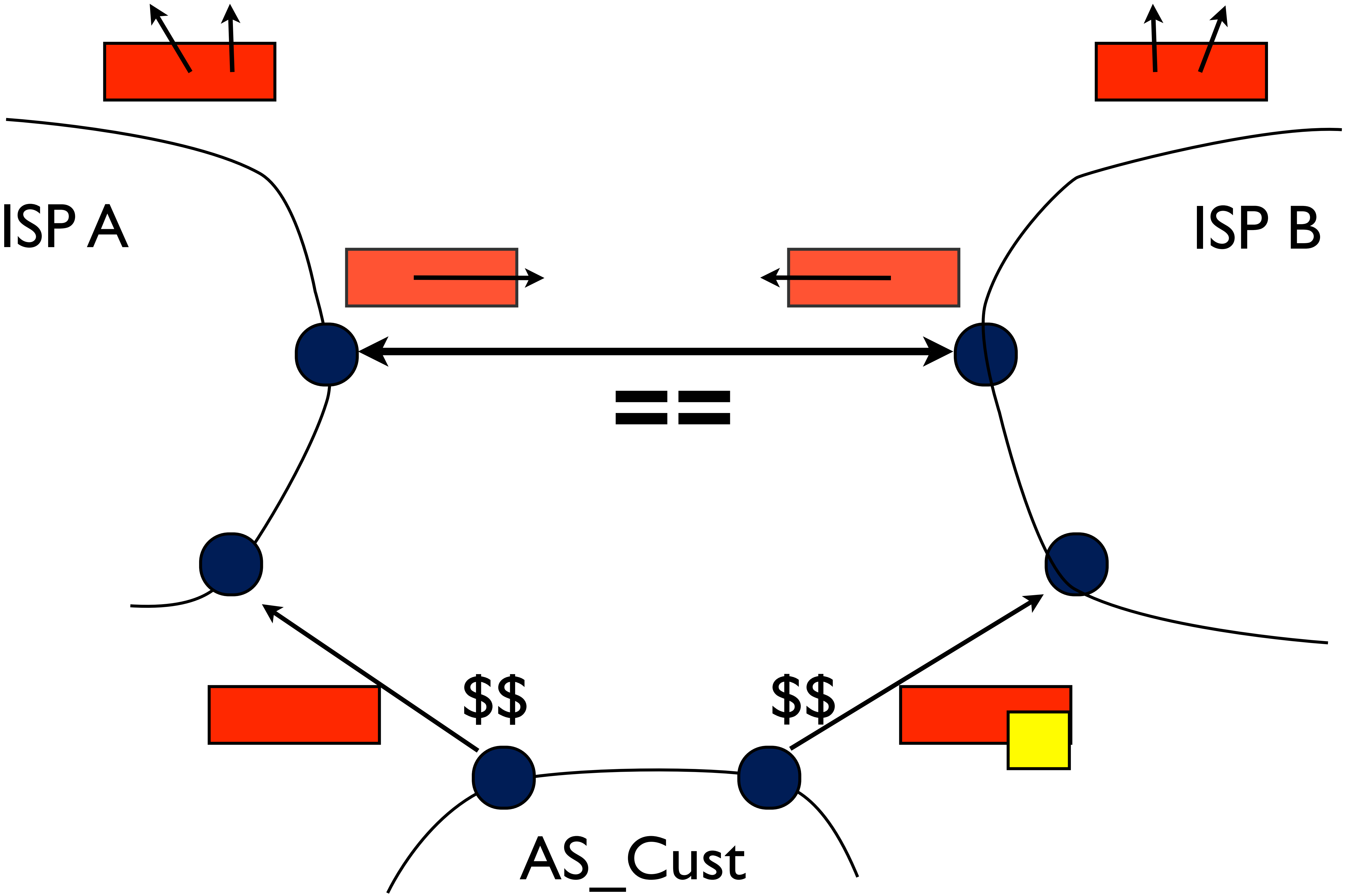
Initial routing status control-plane (only) driven forwarding



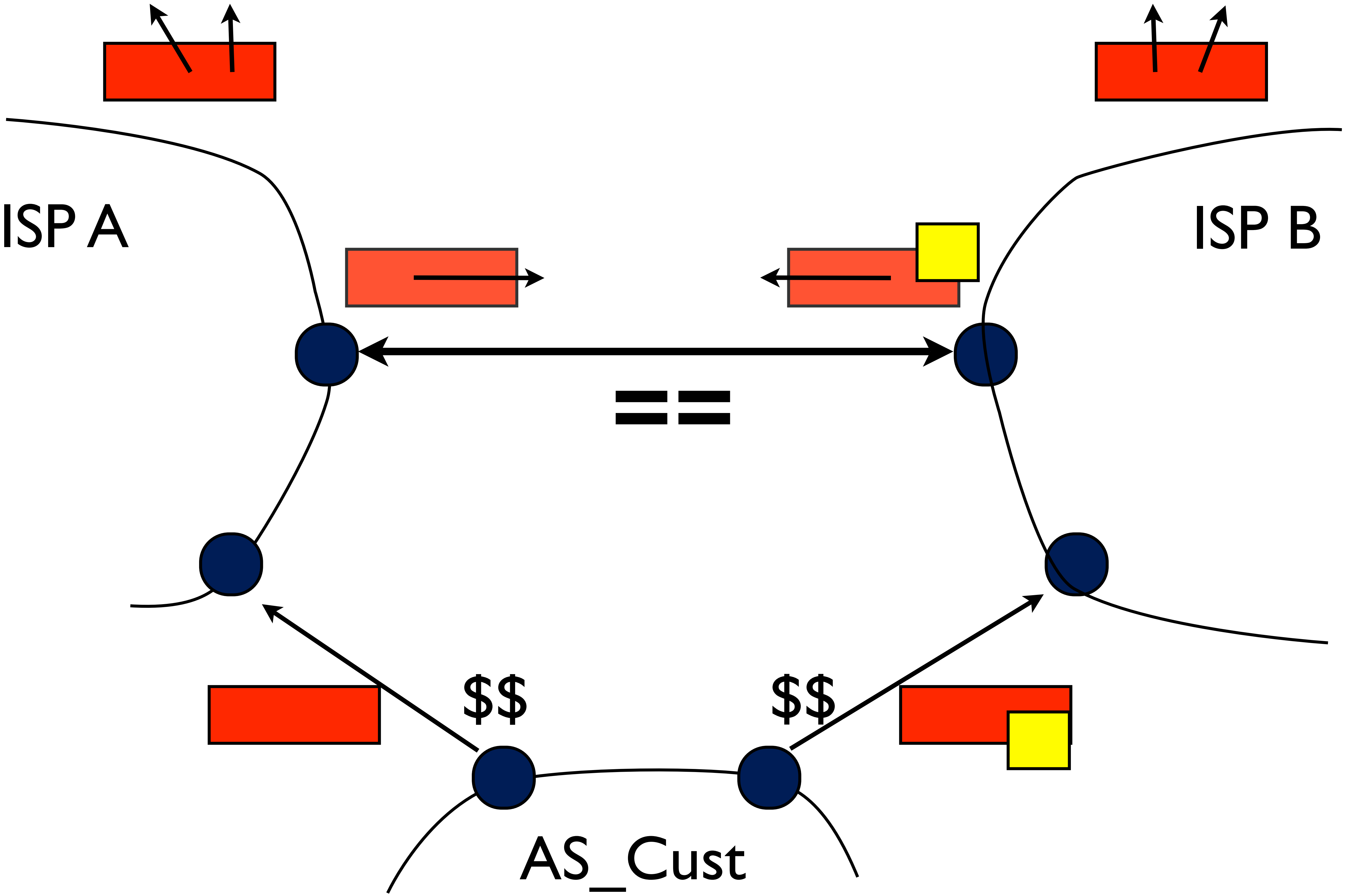
Let’s start playing : Inbound TE, increase RIB/FIB of everyone



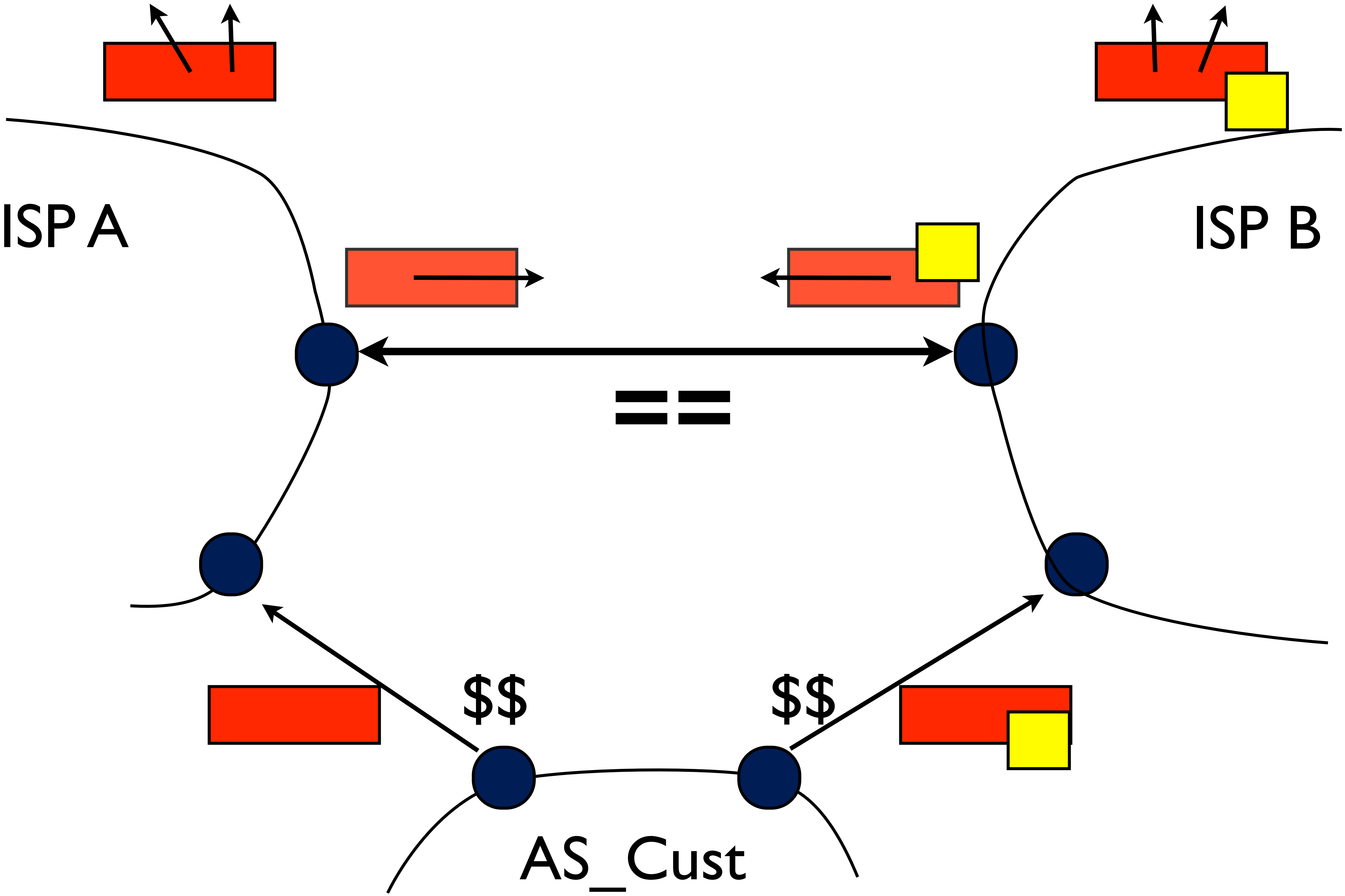
Let’s start playing : Inbound TE, increase RIB/FIB of everyone



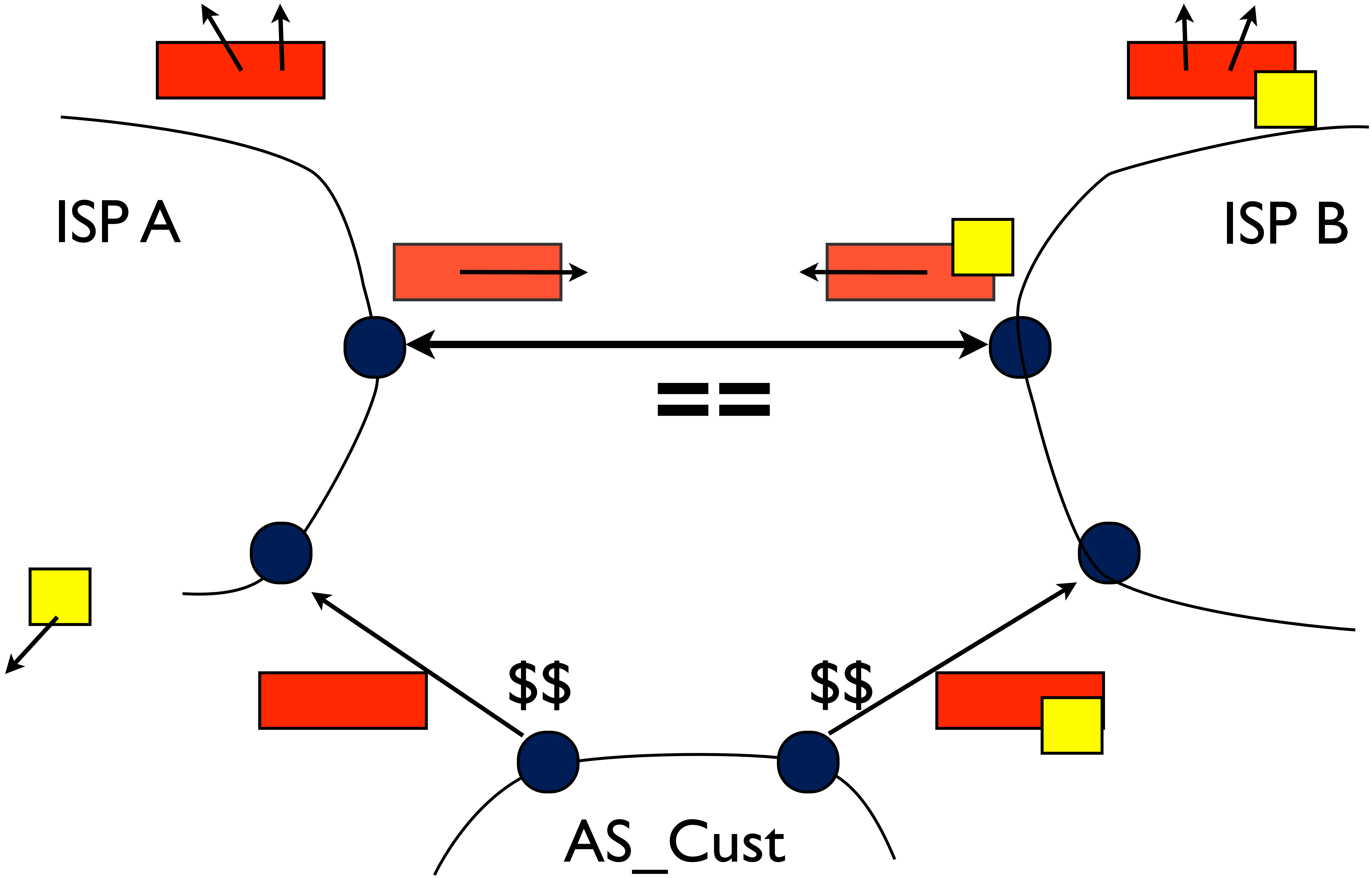
Let’s start playing : Inbound TE, increase RIB/FIB of everyone



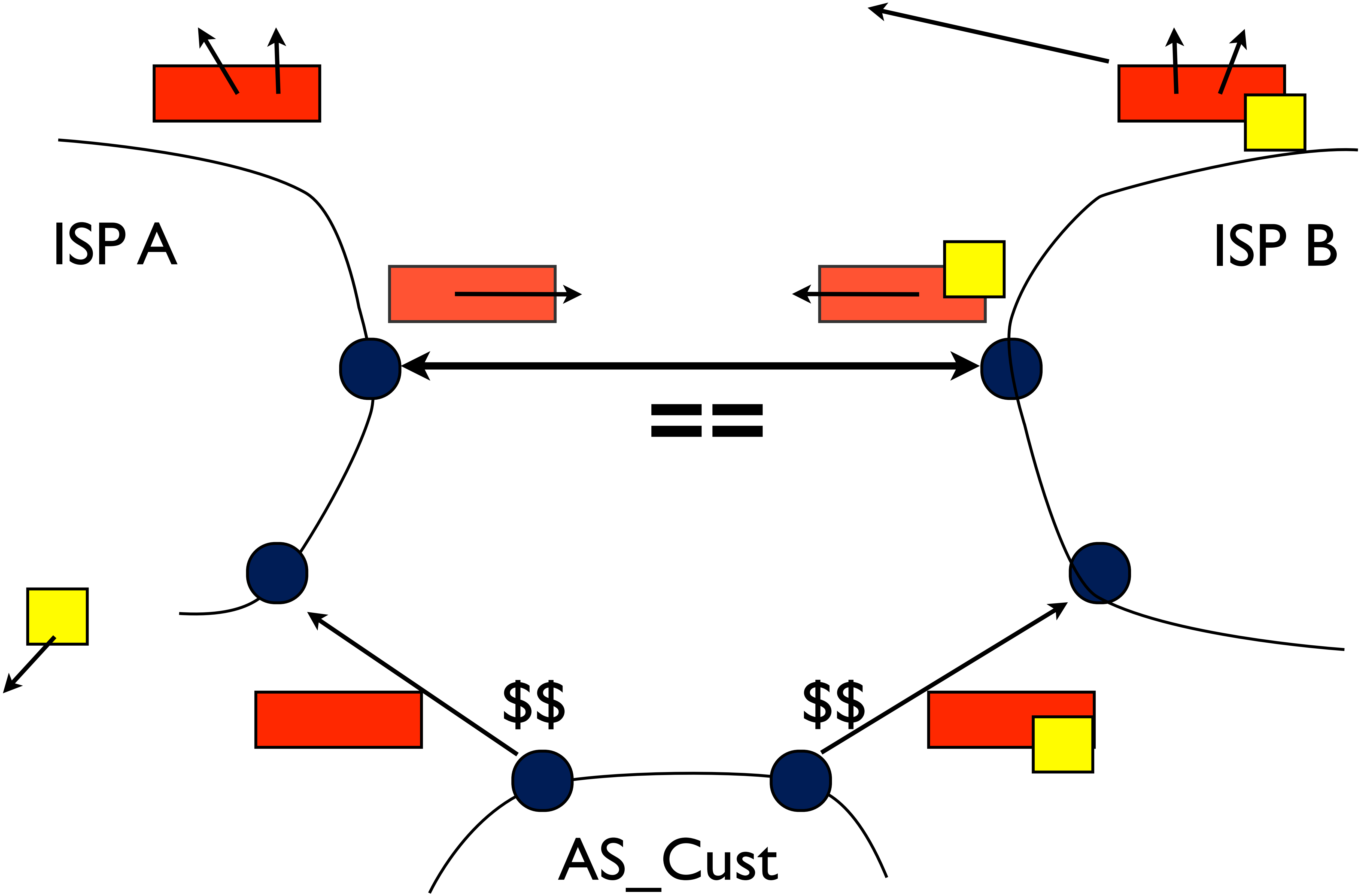
Let’s start playing : Inbound TE, increase RIB/FIB of everyone



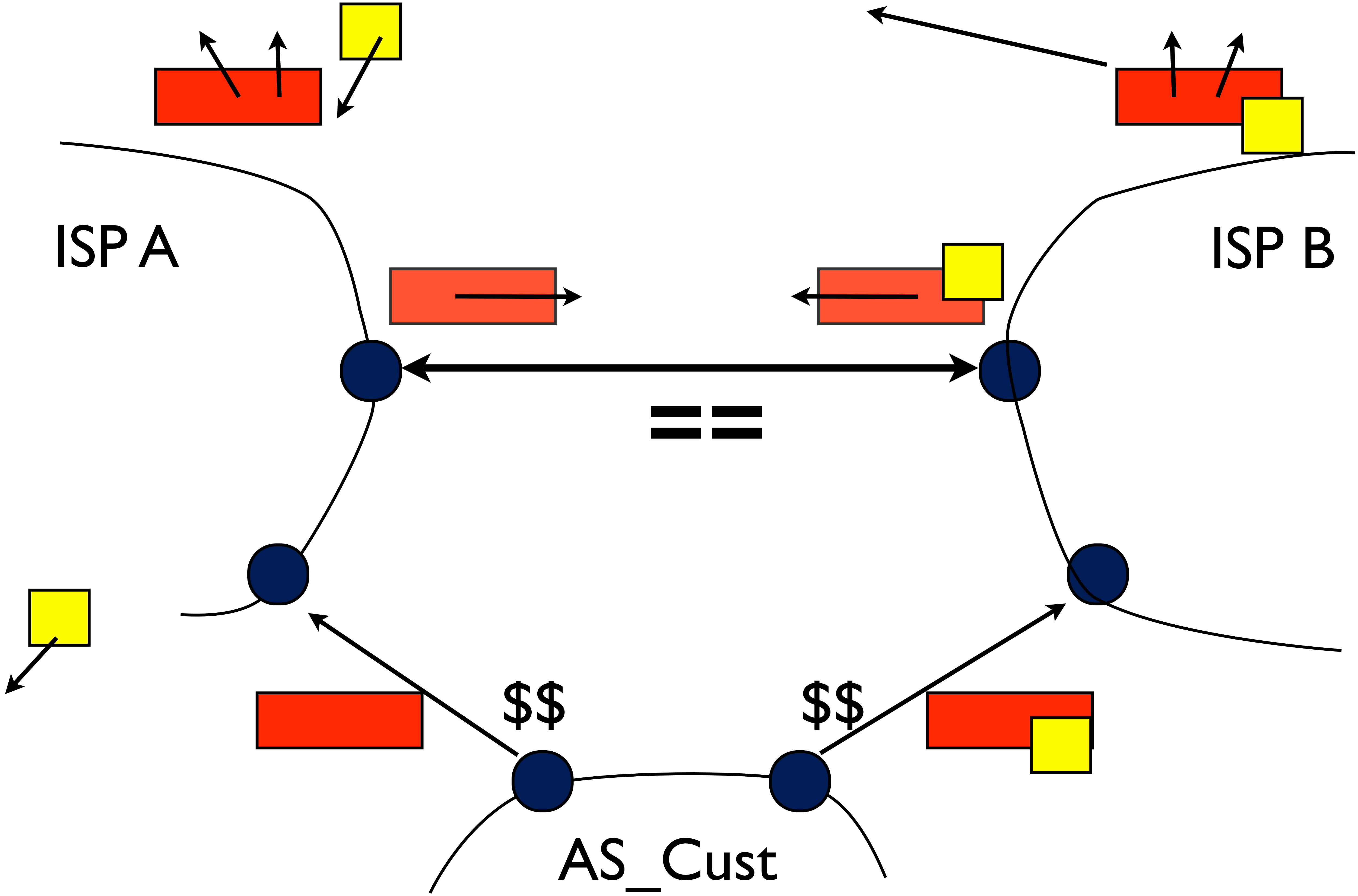
Let’s start playing : Inbound TE, increase RIB/FIB of everyone



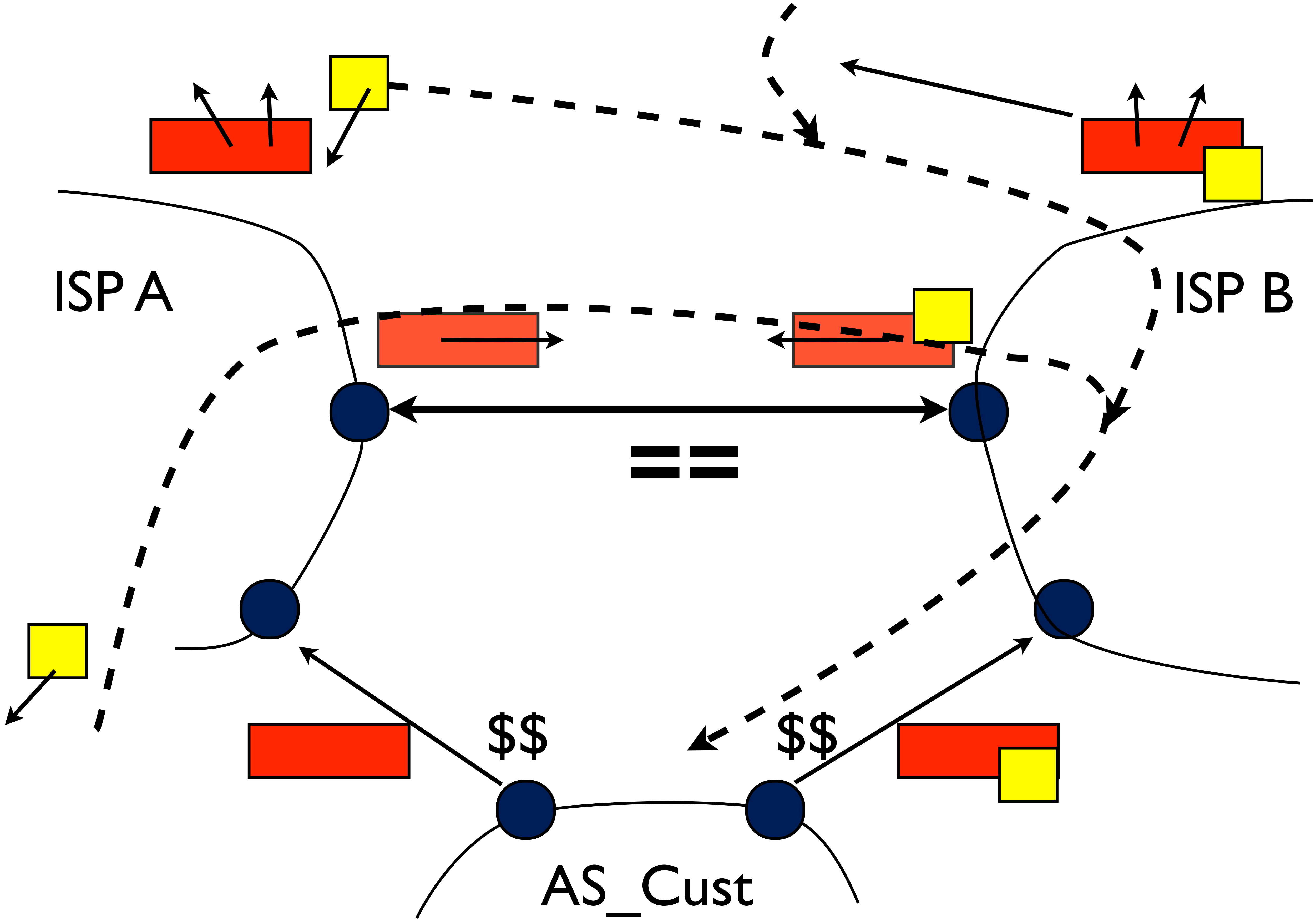
Let’s start playing : Inbound TE, increase RIB/FIB of everyone



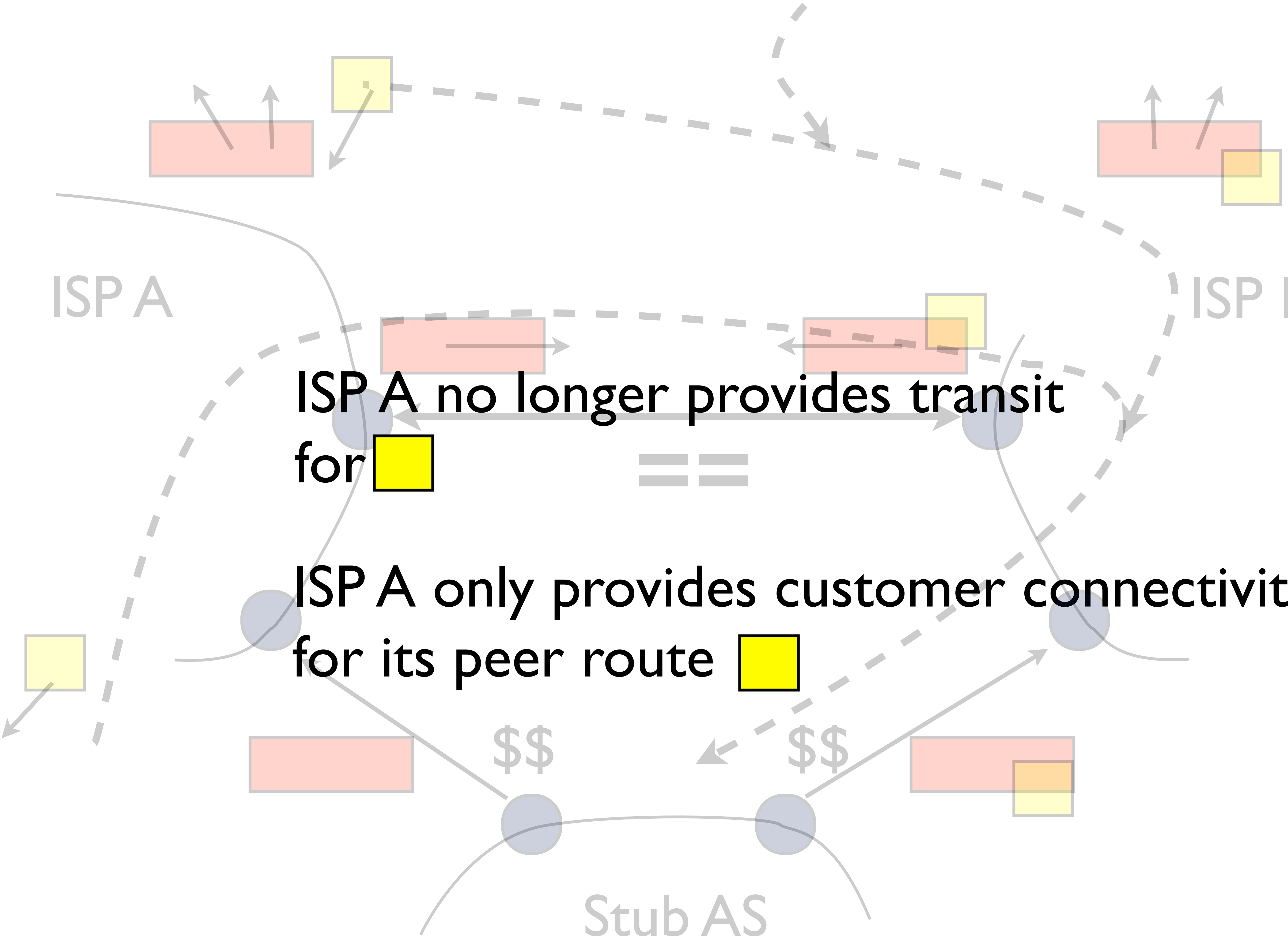
Let’s start playing : Inbound TE, increase RIB/FIB of everyone



Let’s start playing : Inbound TE, increase RIB/FIB of everyone



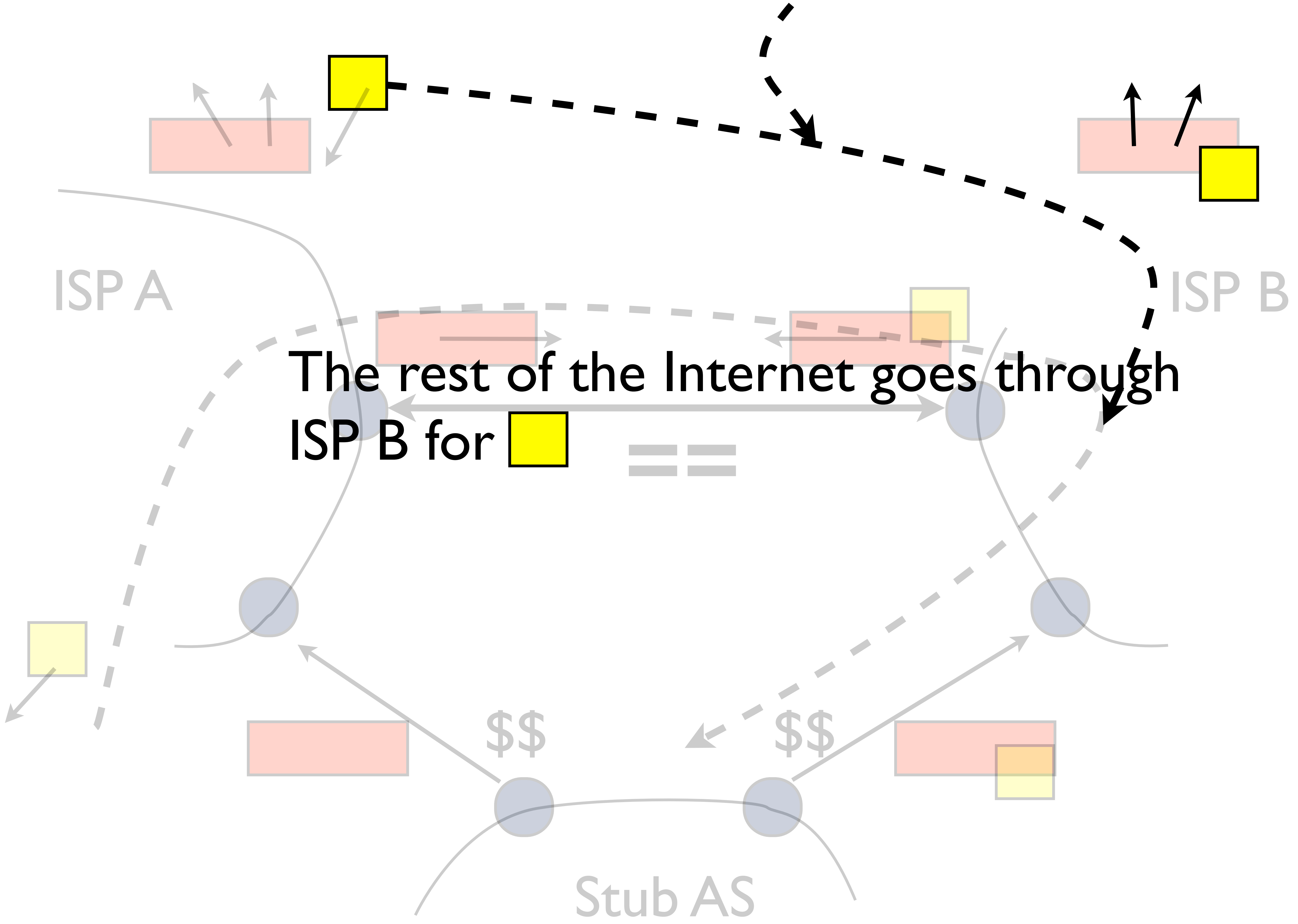
Let’s start playing : Inbound TE, increase RIB/FIB of everyone



ISP A no longer provides transit for 🟨

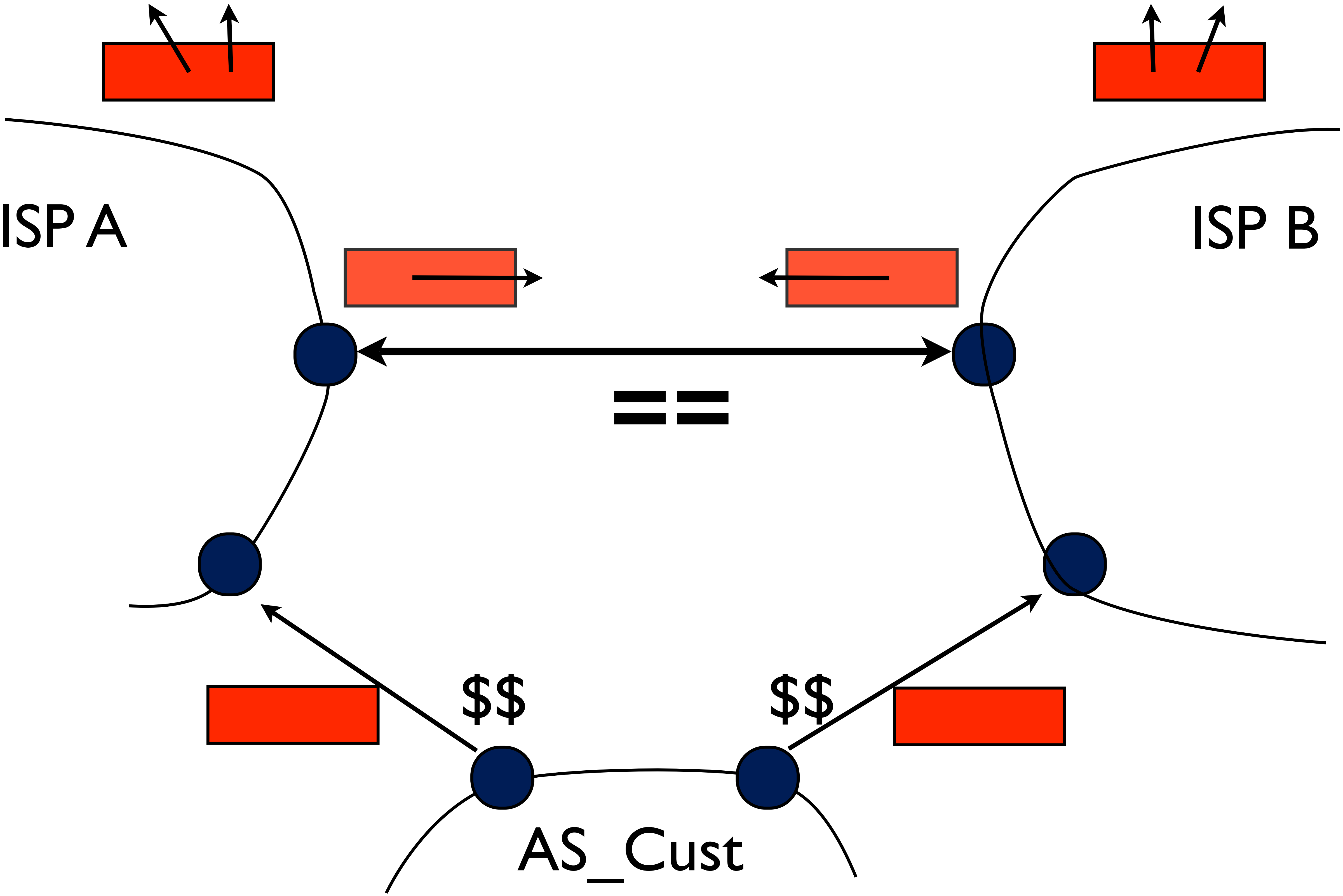
ISP A only provides customer connectivity for its peer route 🟨

Let’s start playing : Inbound TE, increase RIB/FIB of everyone

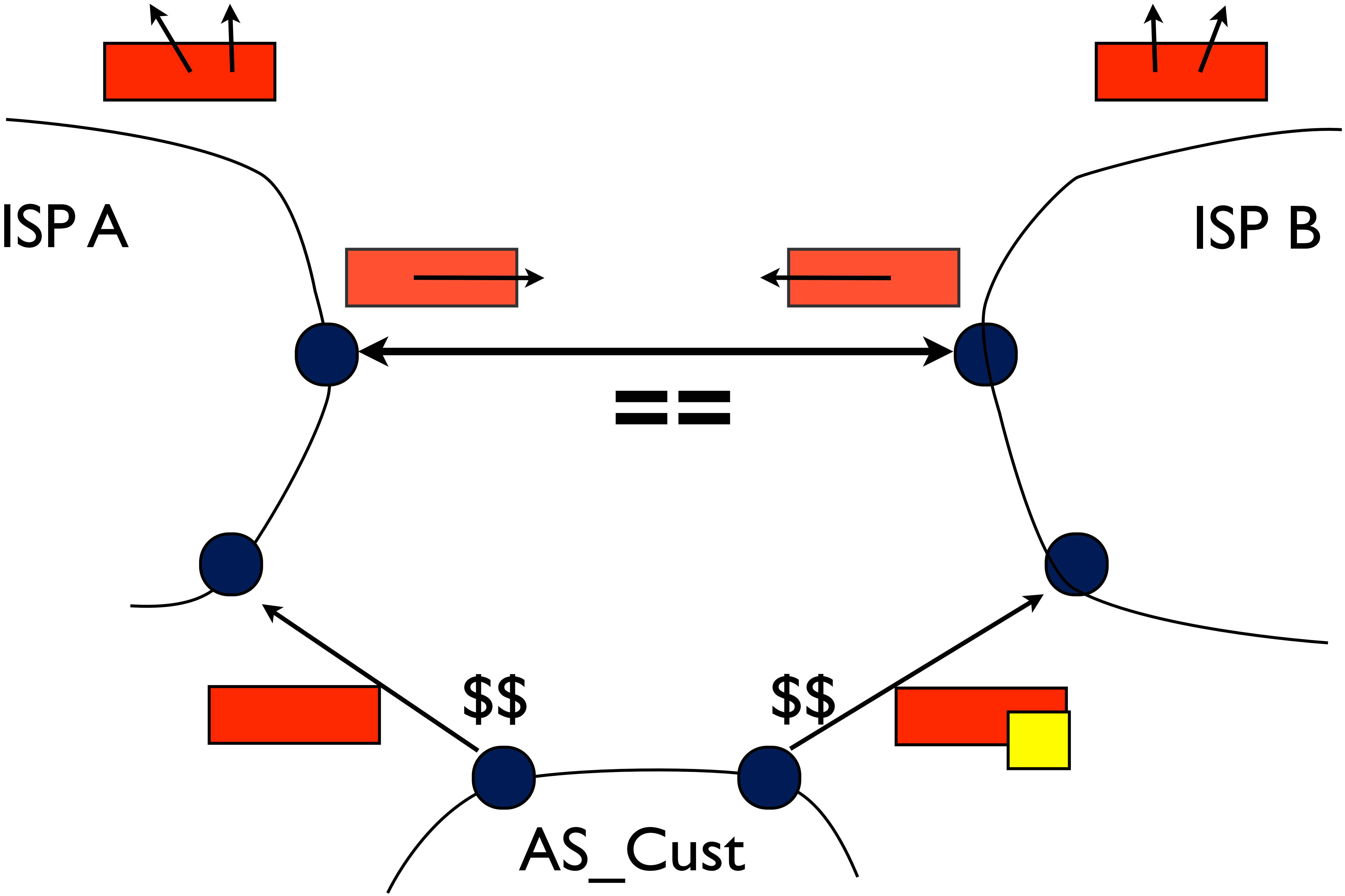


The rest of the Internet goes through ISP B for 🟨

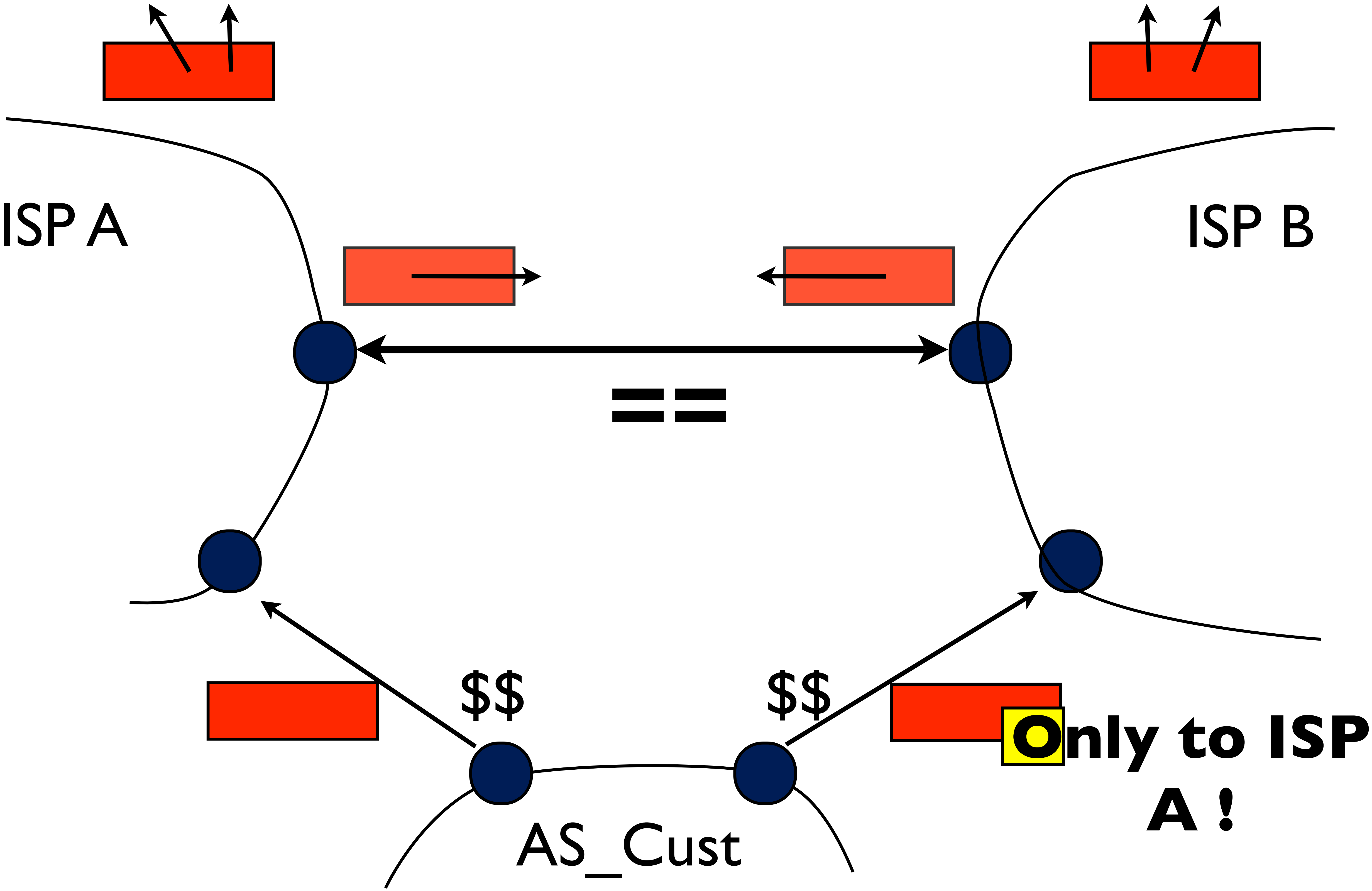
Let's start playing : Scope advertisement of the more specific



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

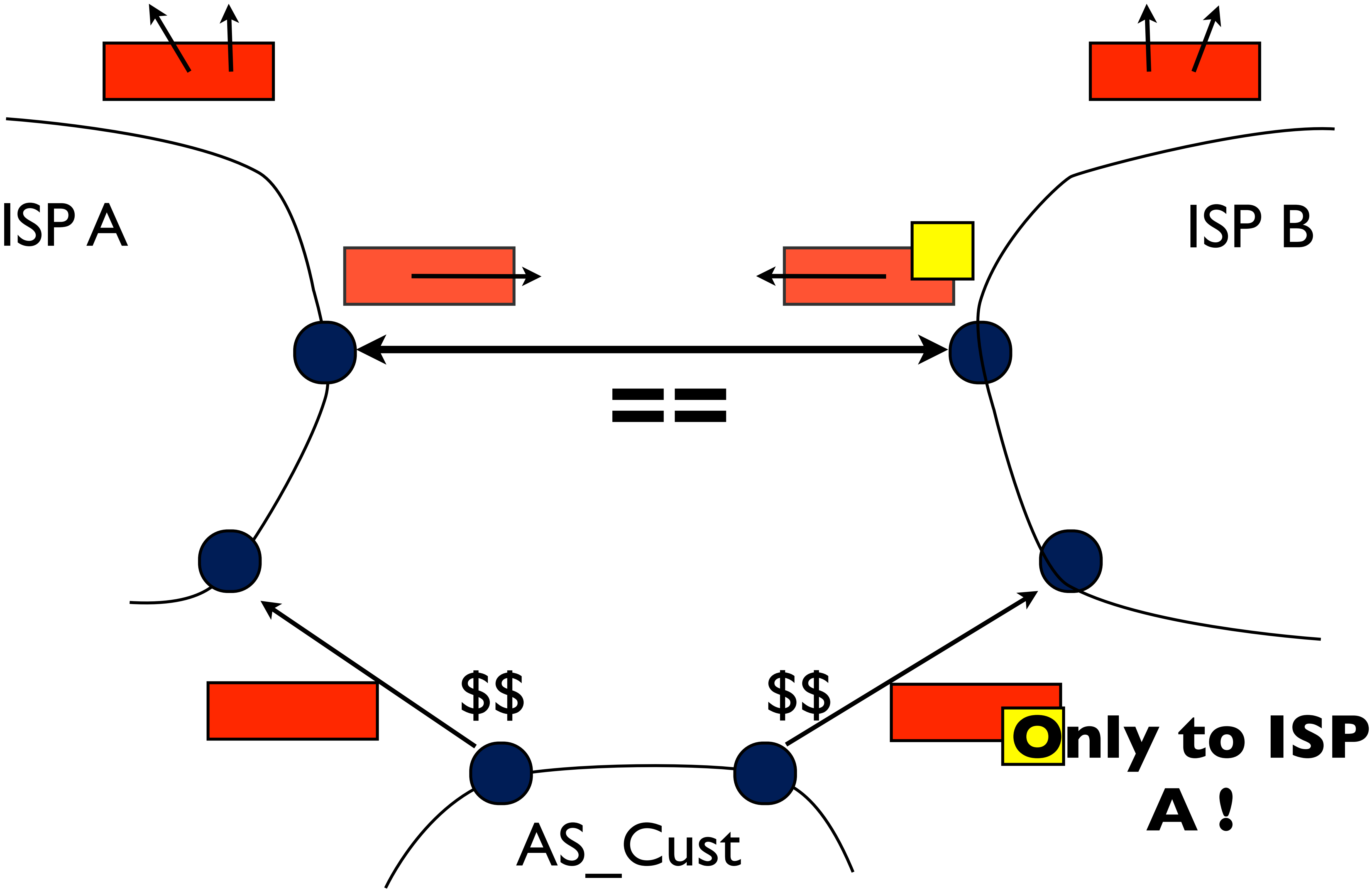


Let's start playing : Scope advertisement of the more specific

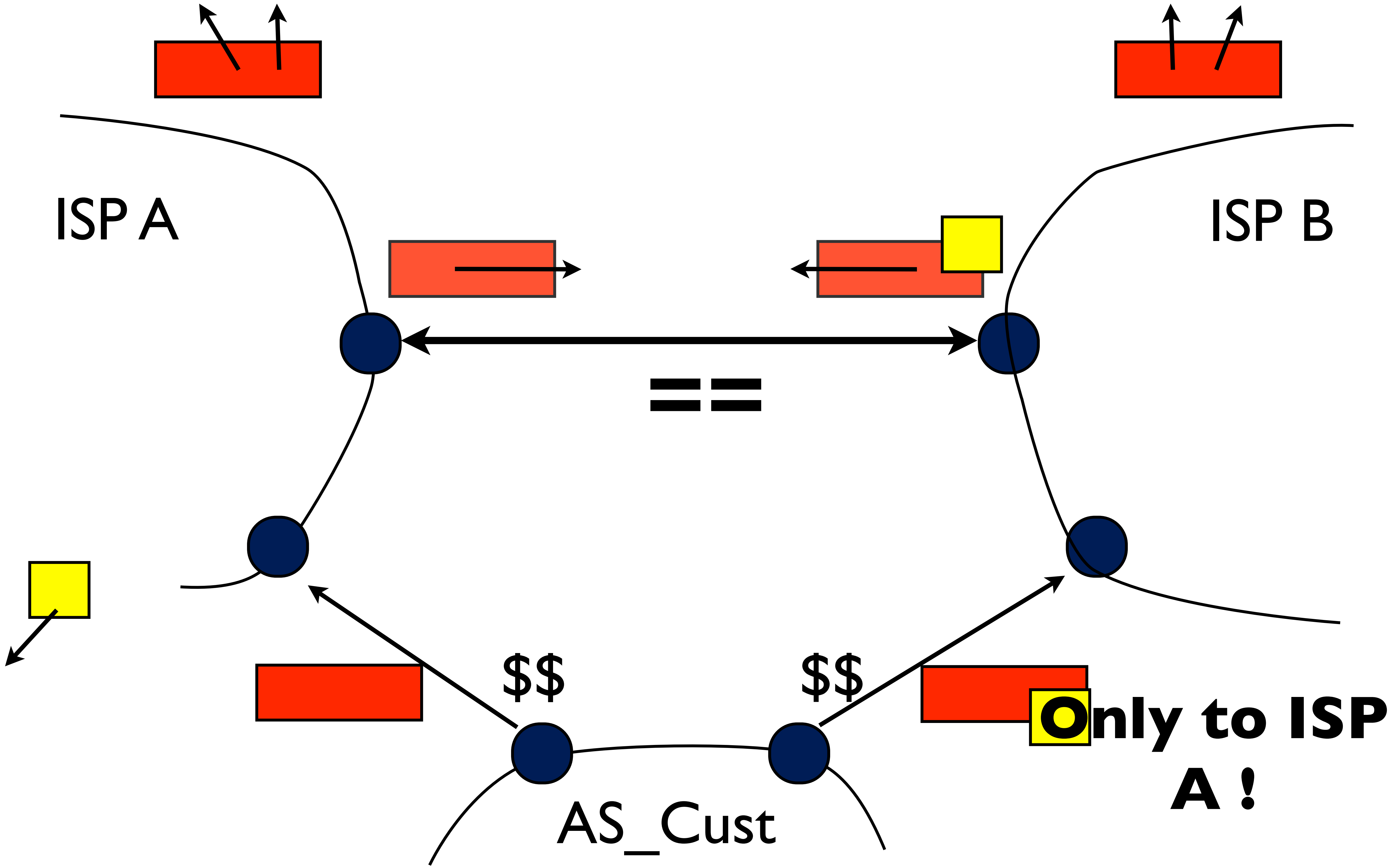


Only to ISP A !

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

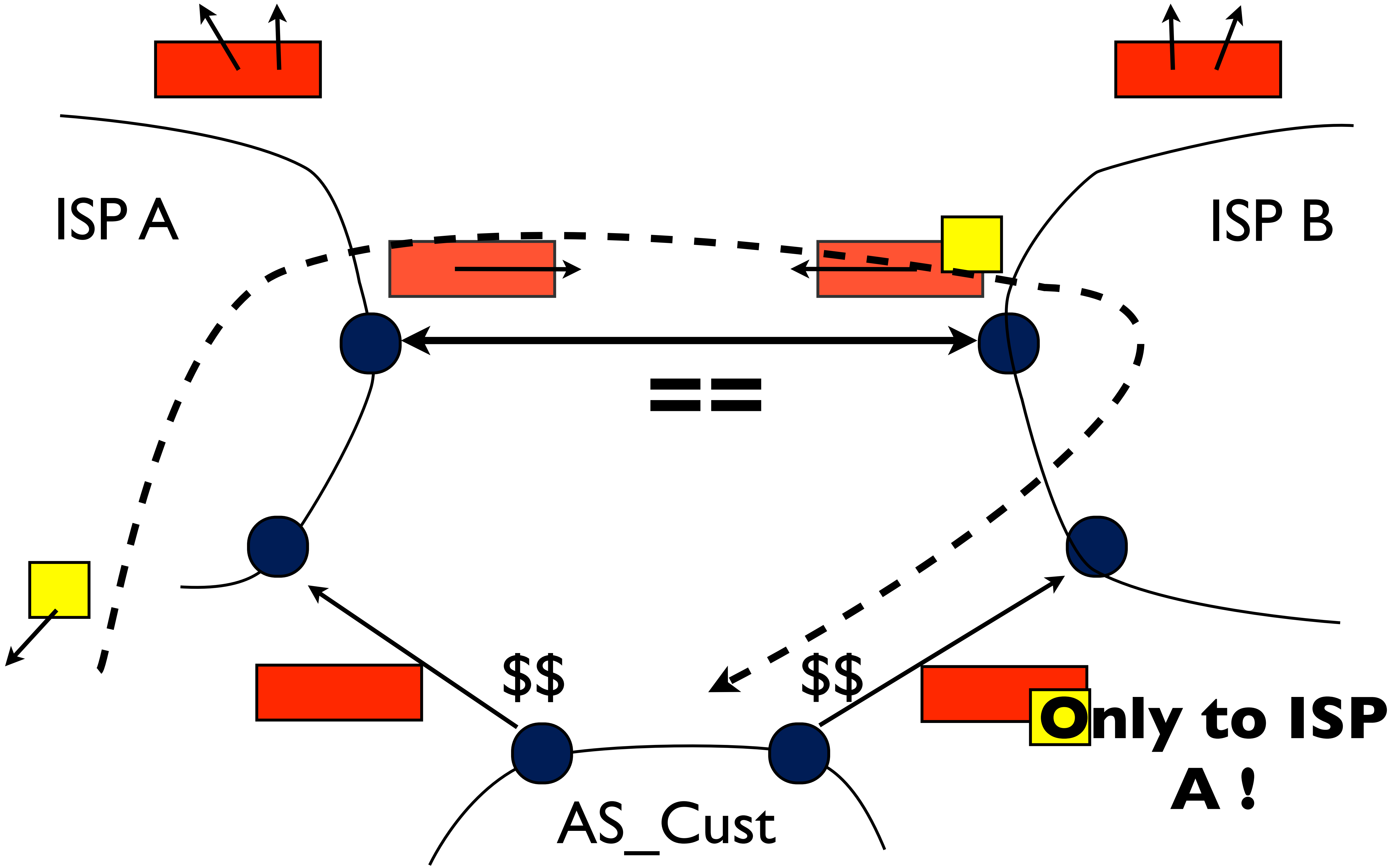


Let’s start playing : Scope advertisement of the more specific



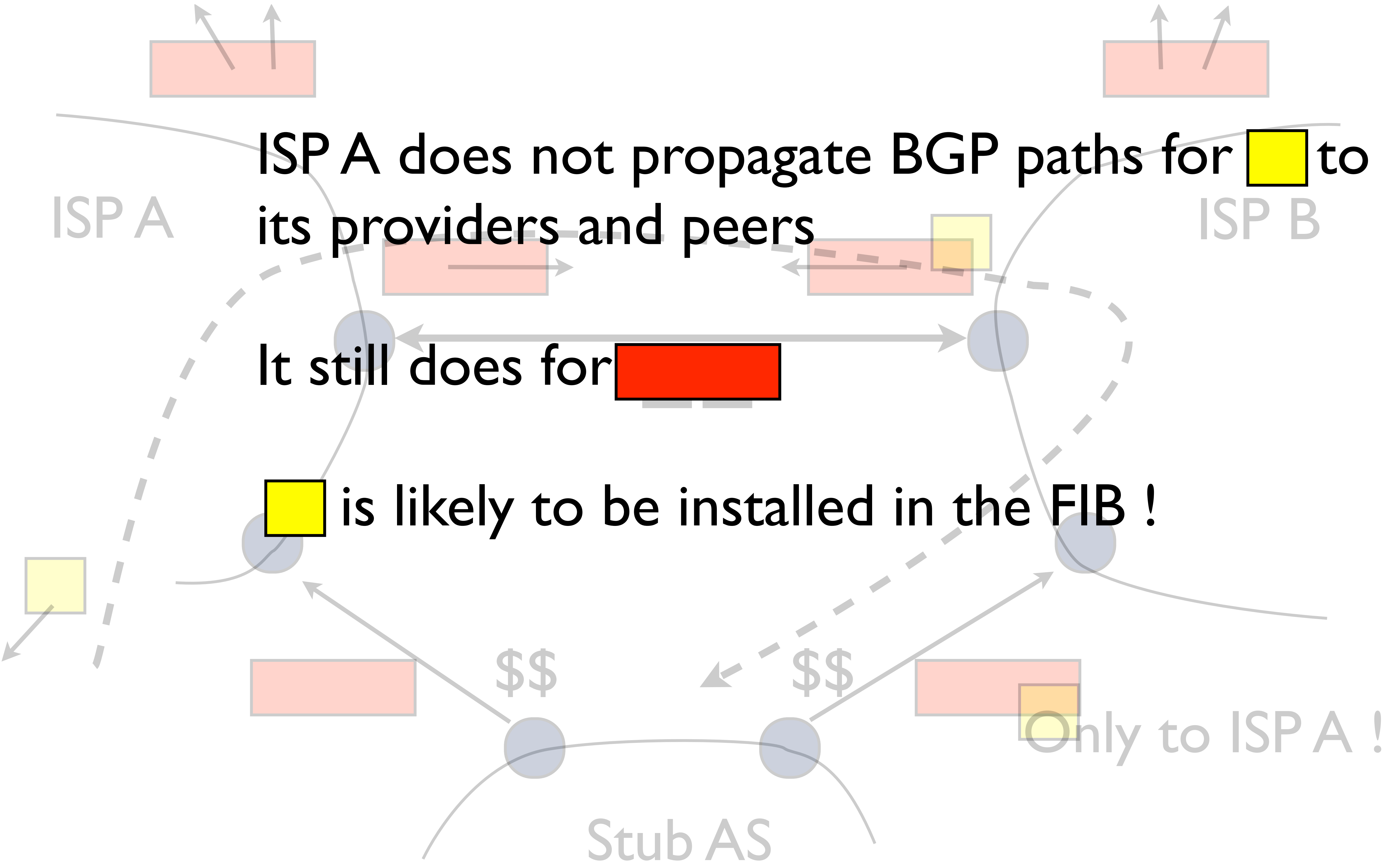
Only to ISP A !

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



Only to ISP A !

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

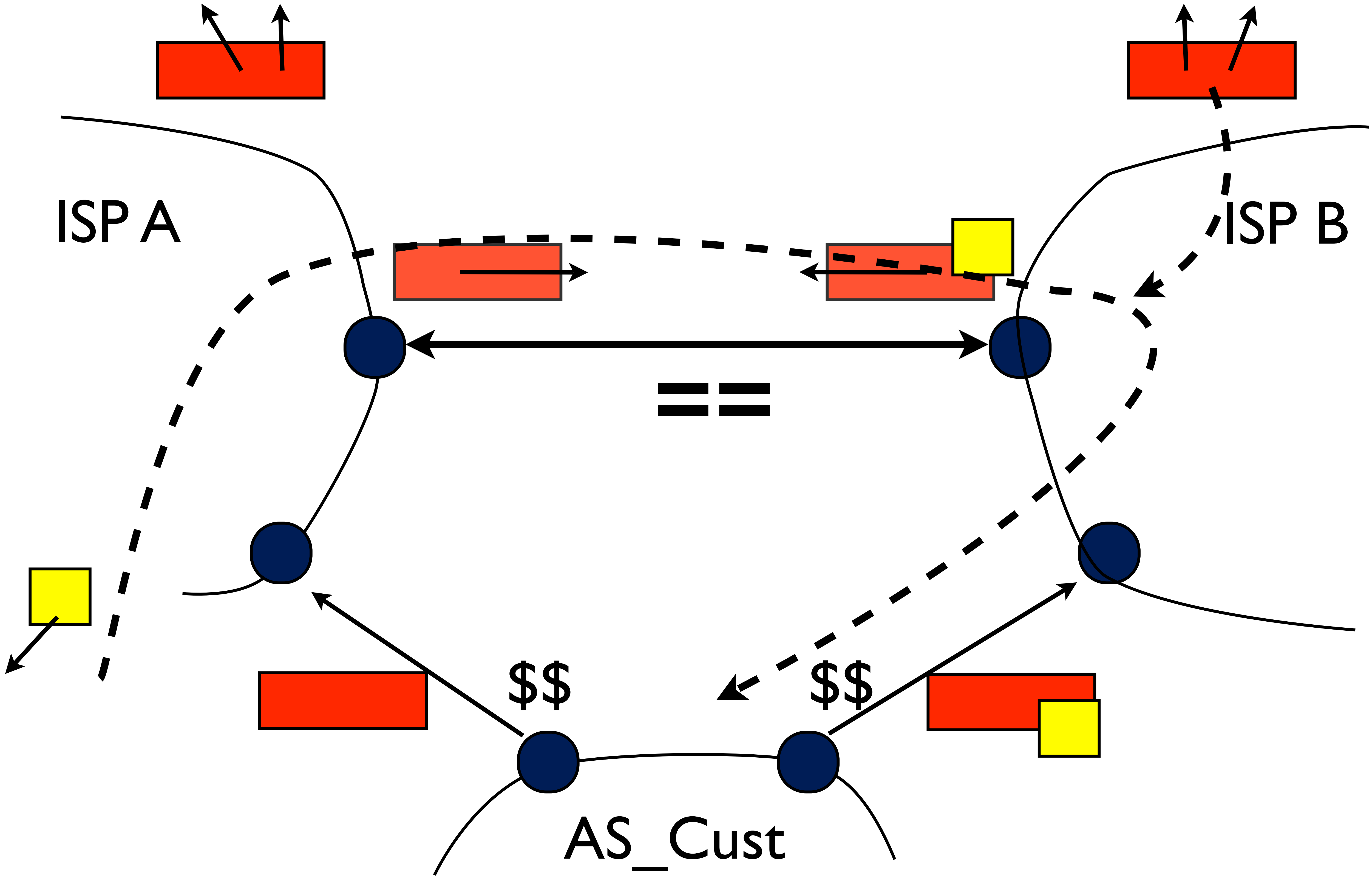


ISP A does not propagate BGP paths for 🟨 to its providers and peers

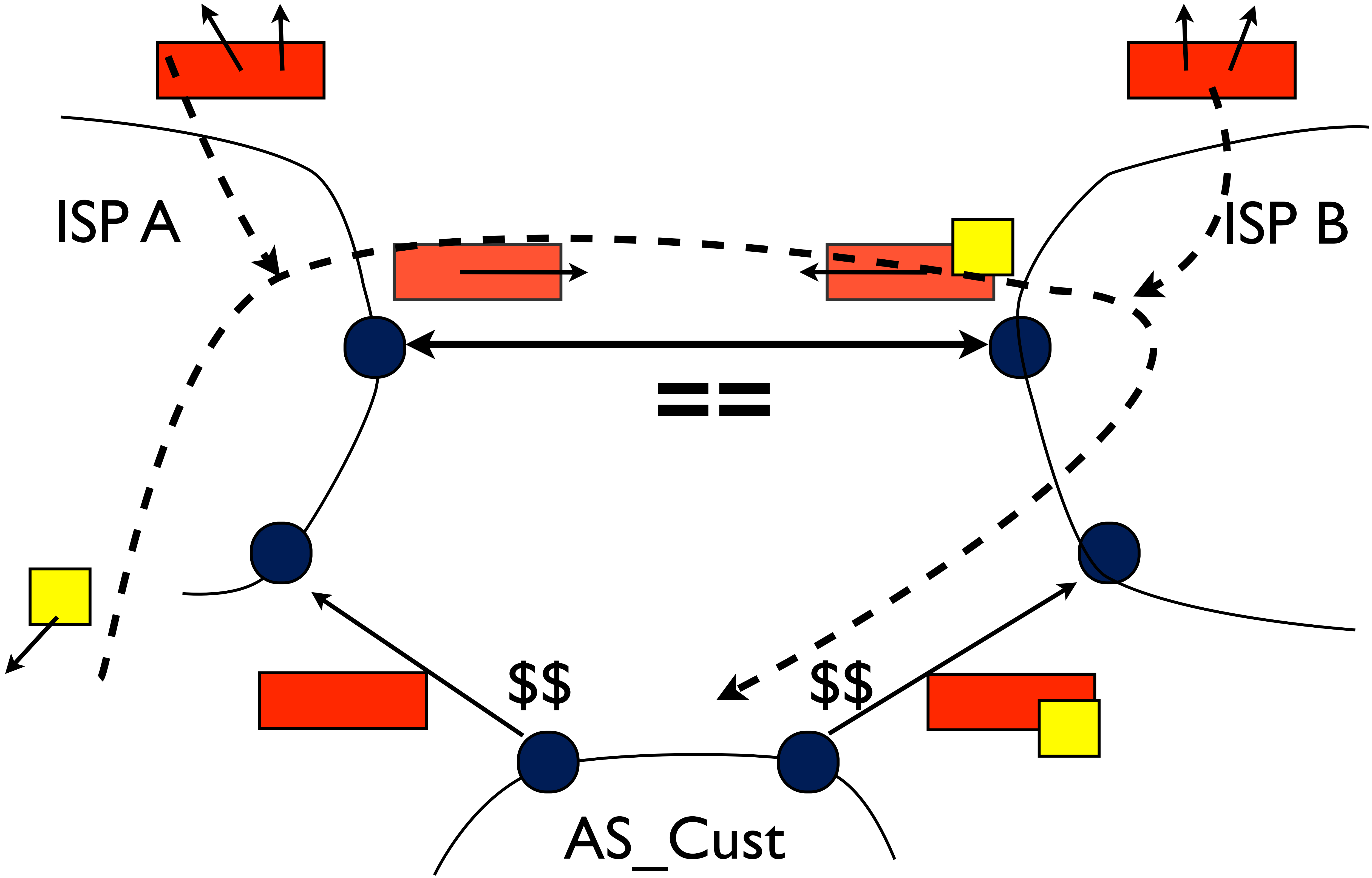
It still does for 🟥

🟨 is likely to be installed in the FIB !

New paths in the network



New paths in the network



How to detect

Data-plane

- Netflow
  - Am I transiting traffic from X to Y ?
  - Warning upon policy violating (X,Y)

How to detect
Control plane

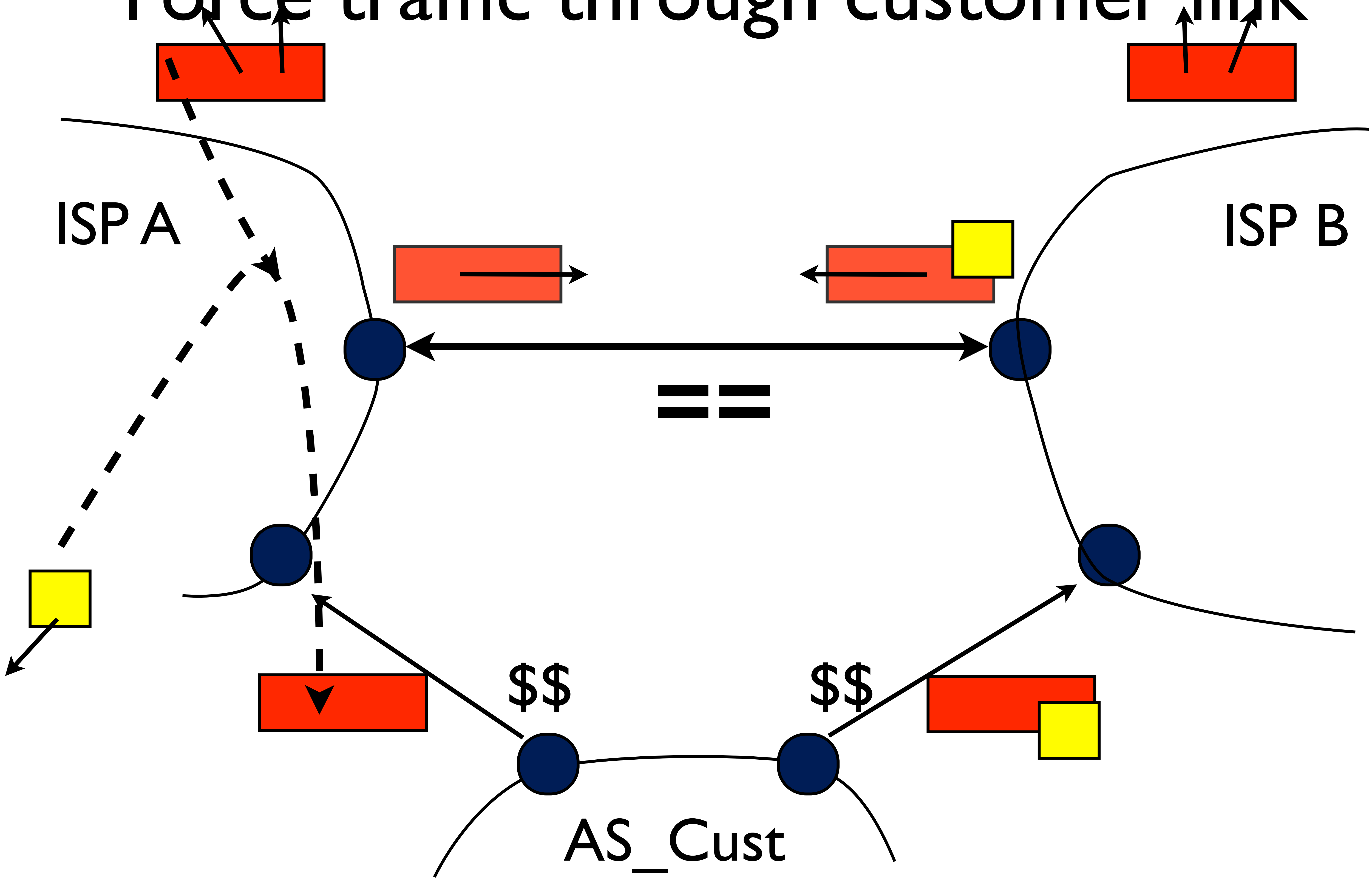
- Getting a more specific route of a customer from a peer and not from the customer is not a sufficient criterium
- Not receiving it from other providers is a good hint
  - Means that your provider does not have a route to p/P+2, and is likely routing according to p/P
  - ...
- As many RIB checks as there are ways to violate policies...
- Often required to "look elsewhere"...

How to react
DAC

- Deliver, Account, Charge
  - consider your "peer" as a customer for that share of traffic
  - negotiate

How to react

Force traffic through customer link



How ?

- Filter out the more specific ?
- Do “Neighbor-Specific” forwarding ?

How to react ?

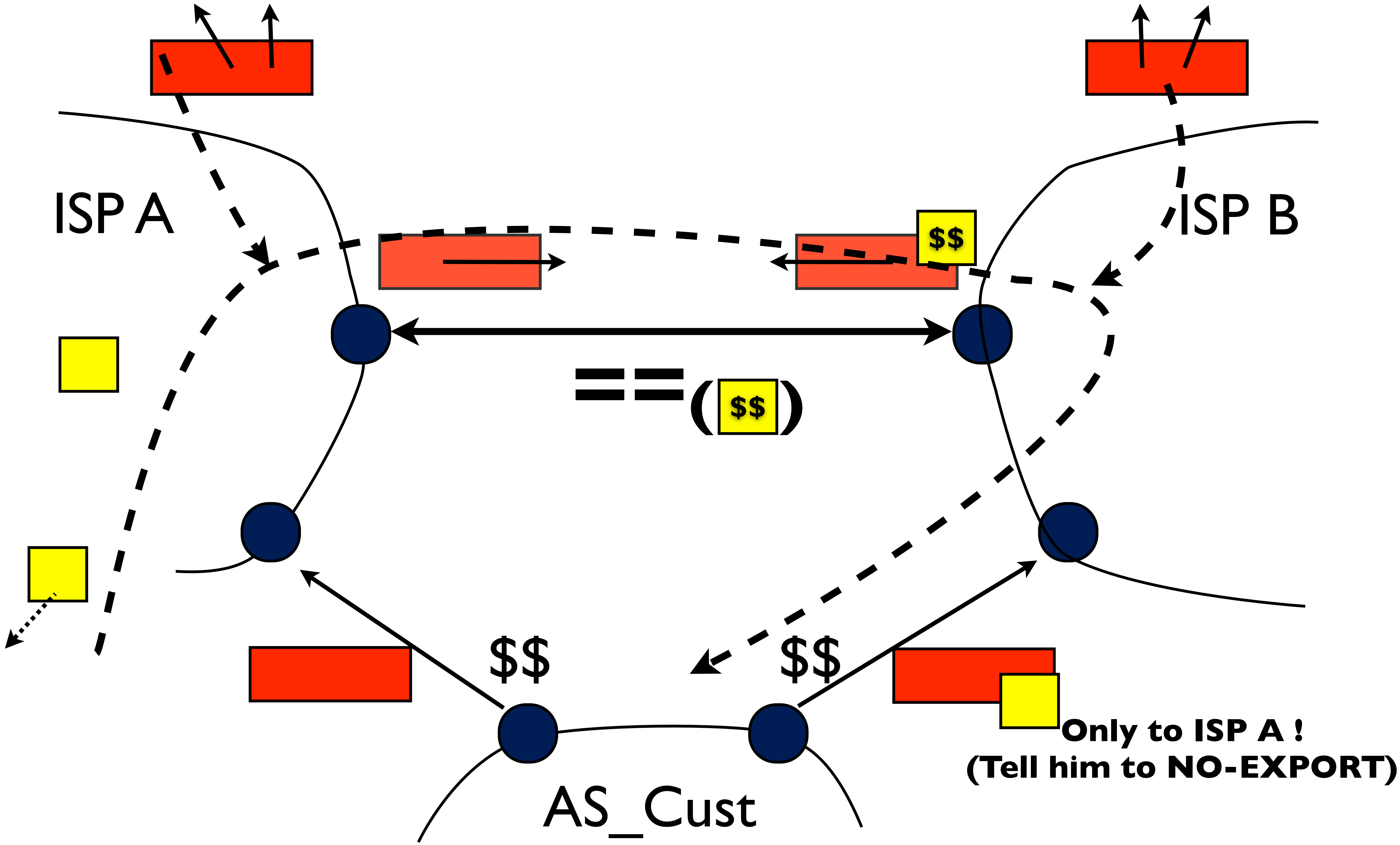
Drop

- Install ACLs or empty routes to p/P+2 at providers and peers entry points

How to anticipate ?

- Pretty hard to avoid false positives with anticipant solutions
- Neighbor-Specific BGP is kind of an anticipant solution
- Scripted ACL generation is kind of an anticipating **drop-based** solution

New paths in the network



Conclusions

- BGP Policies can be violated using
  - more specific prefixes with scope limitation
  - Lacks of documentation
- Automated solutions are not trivial, should be discussed
- Dropping maybe not **THE** solution
- Detection in the data-plane may be easy
- Neighbor-Specific BGP routing ?