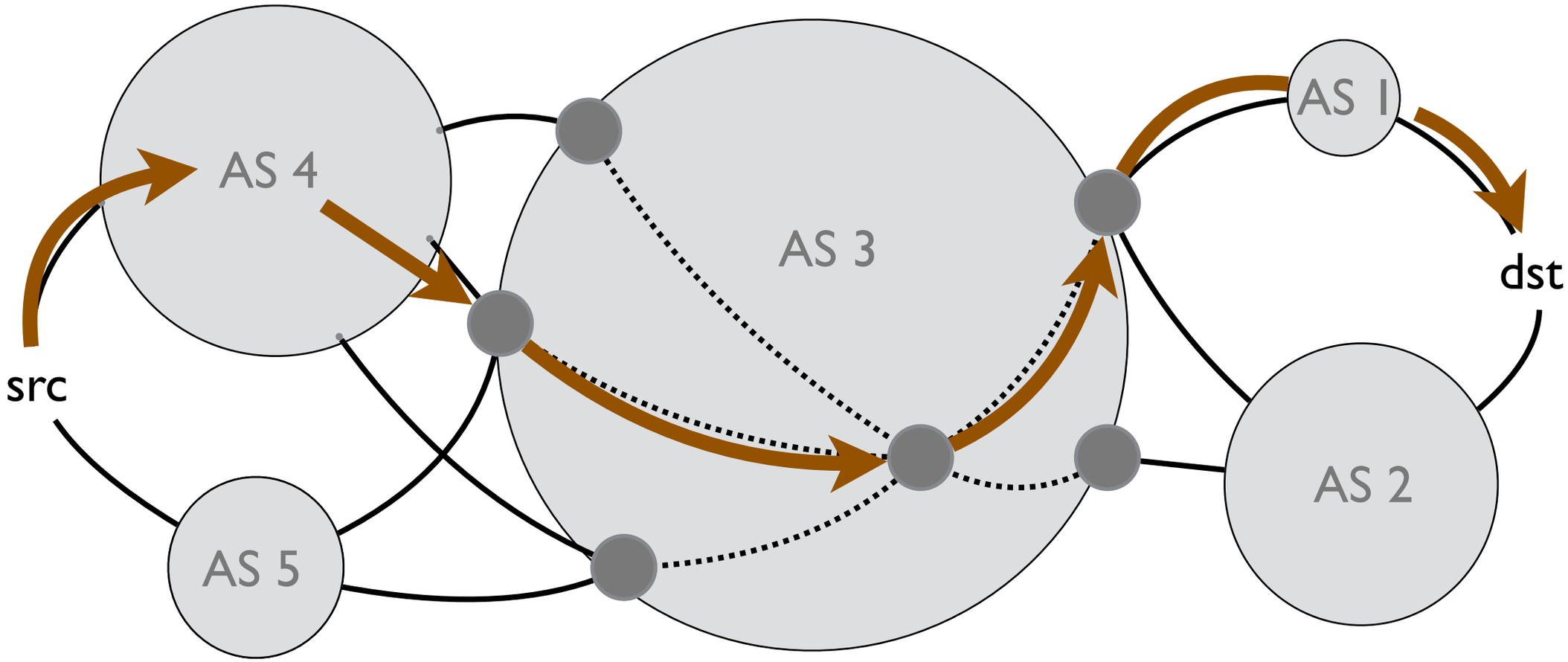
Enabling Policy Innovation in Interdomain Routing: A Software-Defined Approach

Anduo Wang[*] Zhijia Chen[*] Tony Yang[†] Minlan Yu[‡]
[*]Temple University [†]Johns Hopkins University [‡]Harvard University

April 4, 2019 SOSR'19

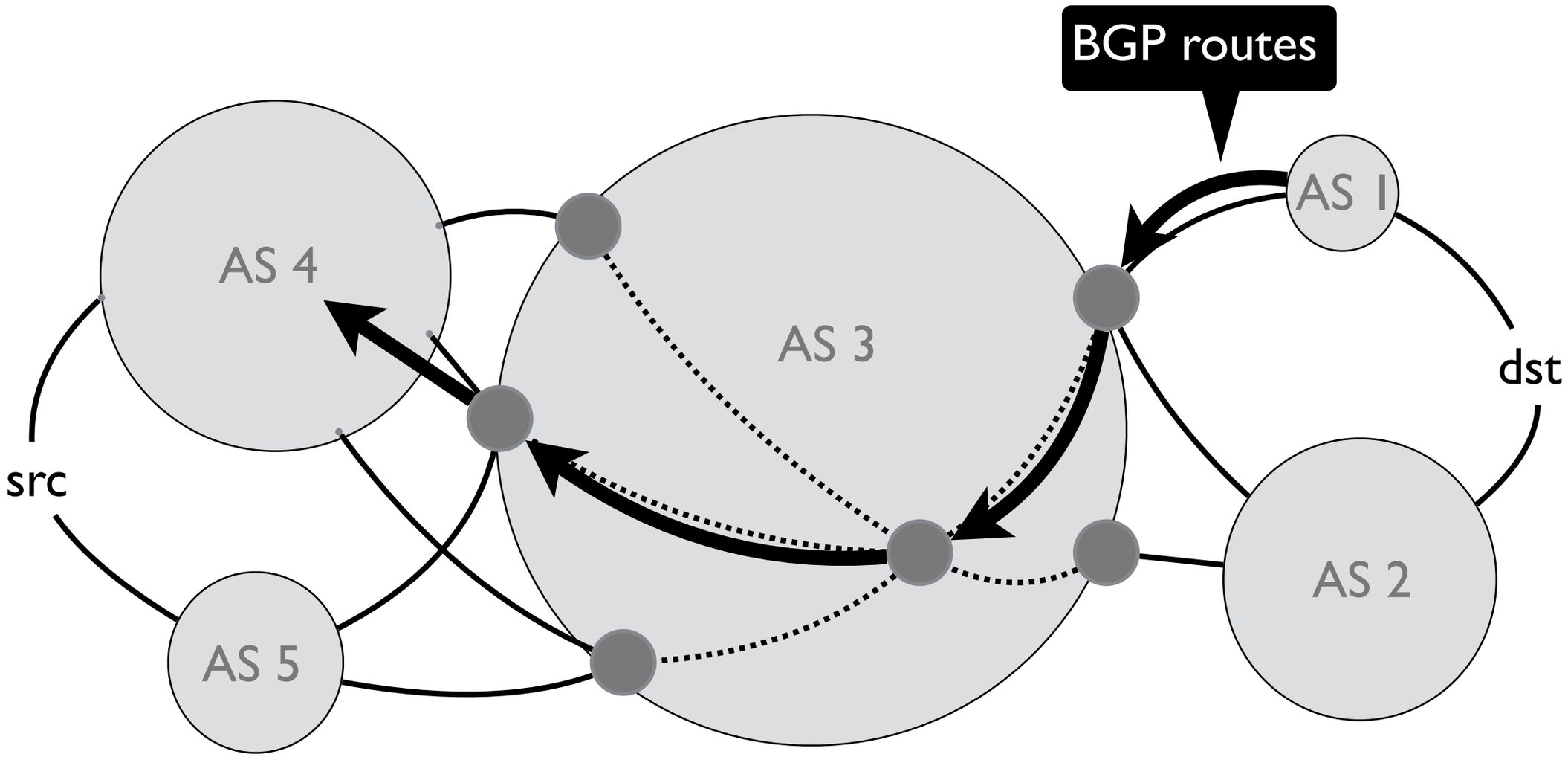
interdomain routing

determining data path connecting communicating hosts



interdomain routing

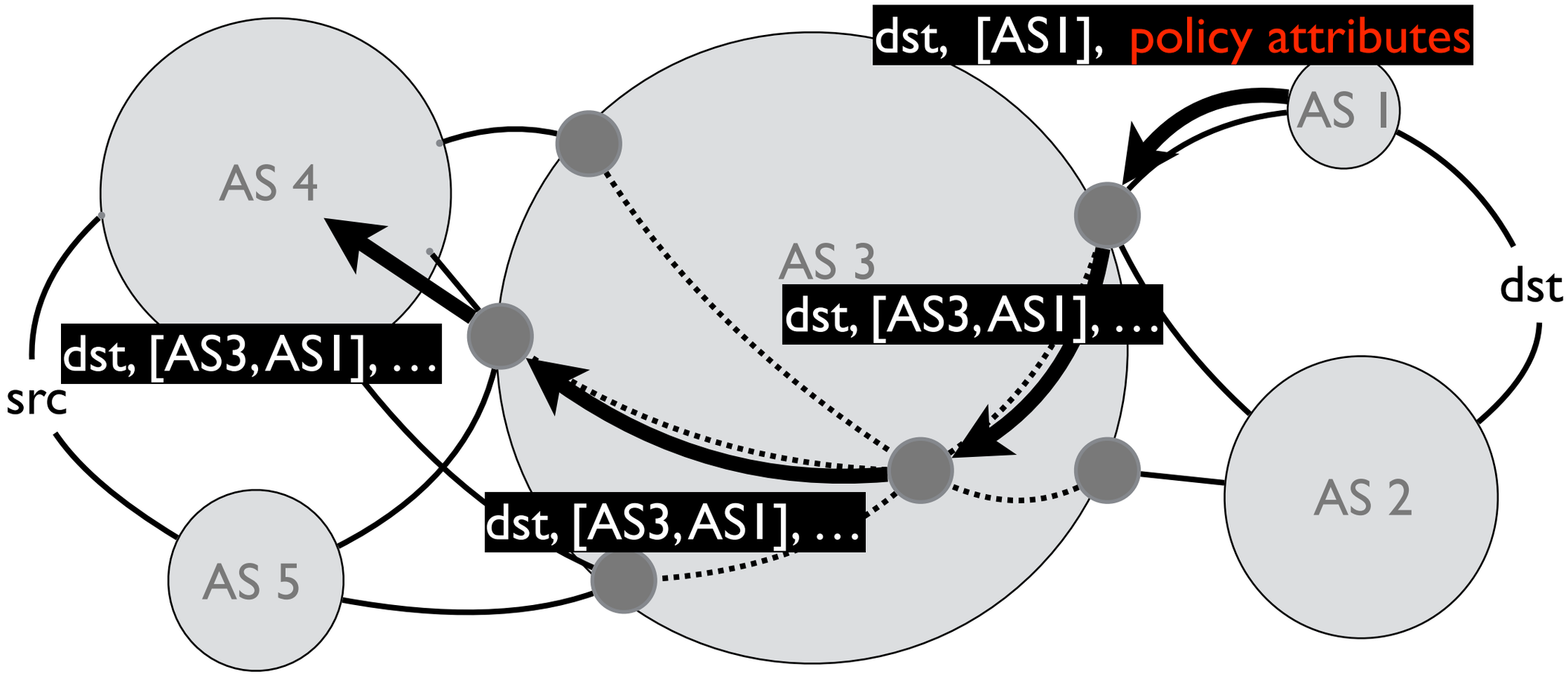
border gateway protocol (BGP) — the only de-facto interdomain routing system



BGP and AS policy

BGP supports AS policies

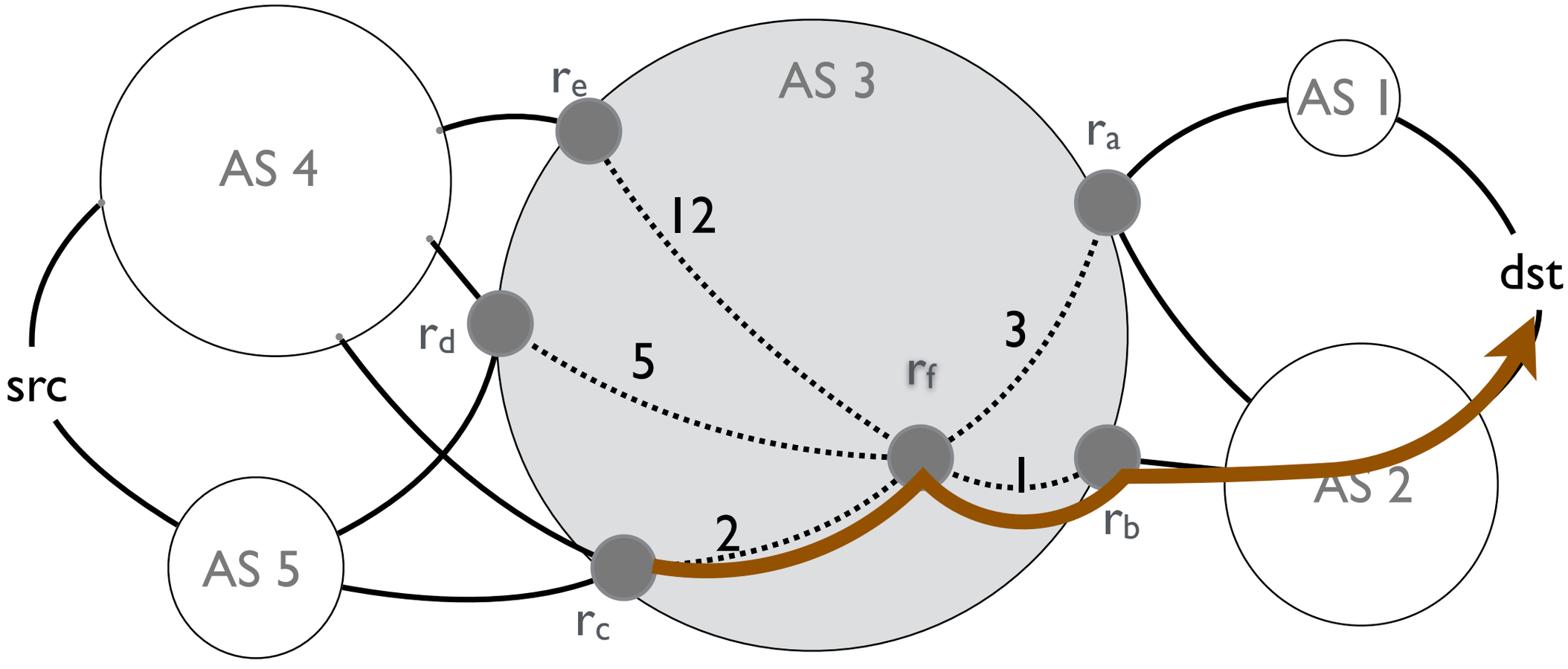
- extending path-vector system
- overriding the shortest AS-path behavior



BGP and AS policy

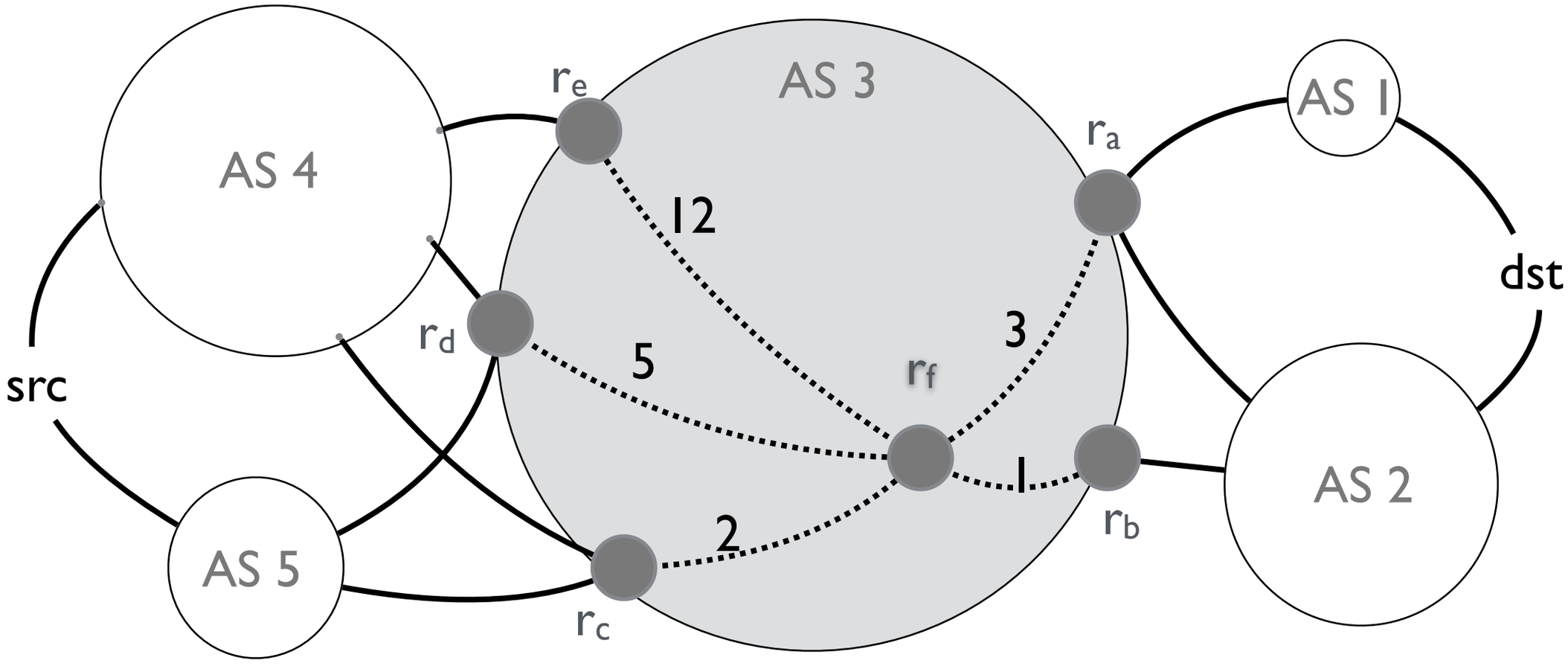
(example) hot potato policy of AS 3

- select a path that minimizes internal cost



influence policy in the downstream

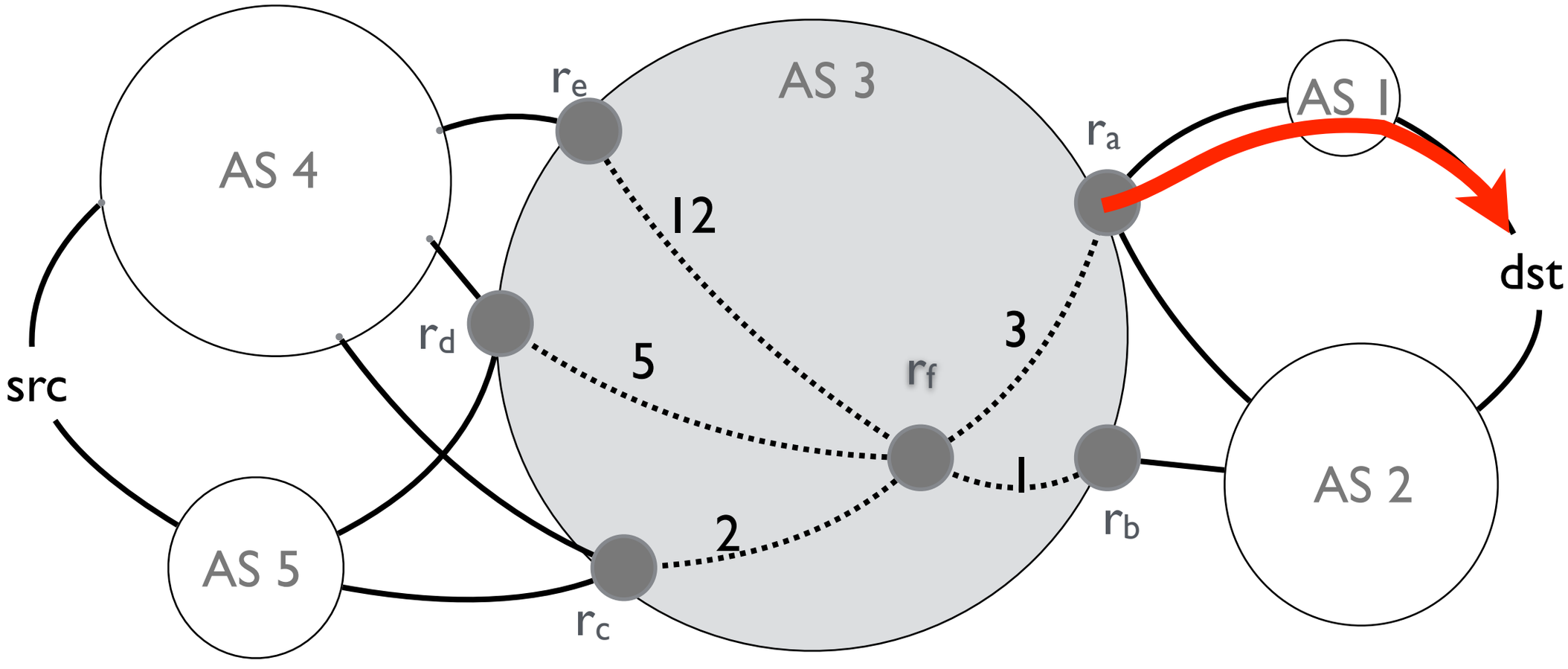
can AS 4 / AS 5 influence routes in AS3?



influence the downstream

can AS 4 / AS 5 influence routes in AS3?

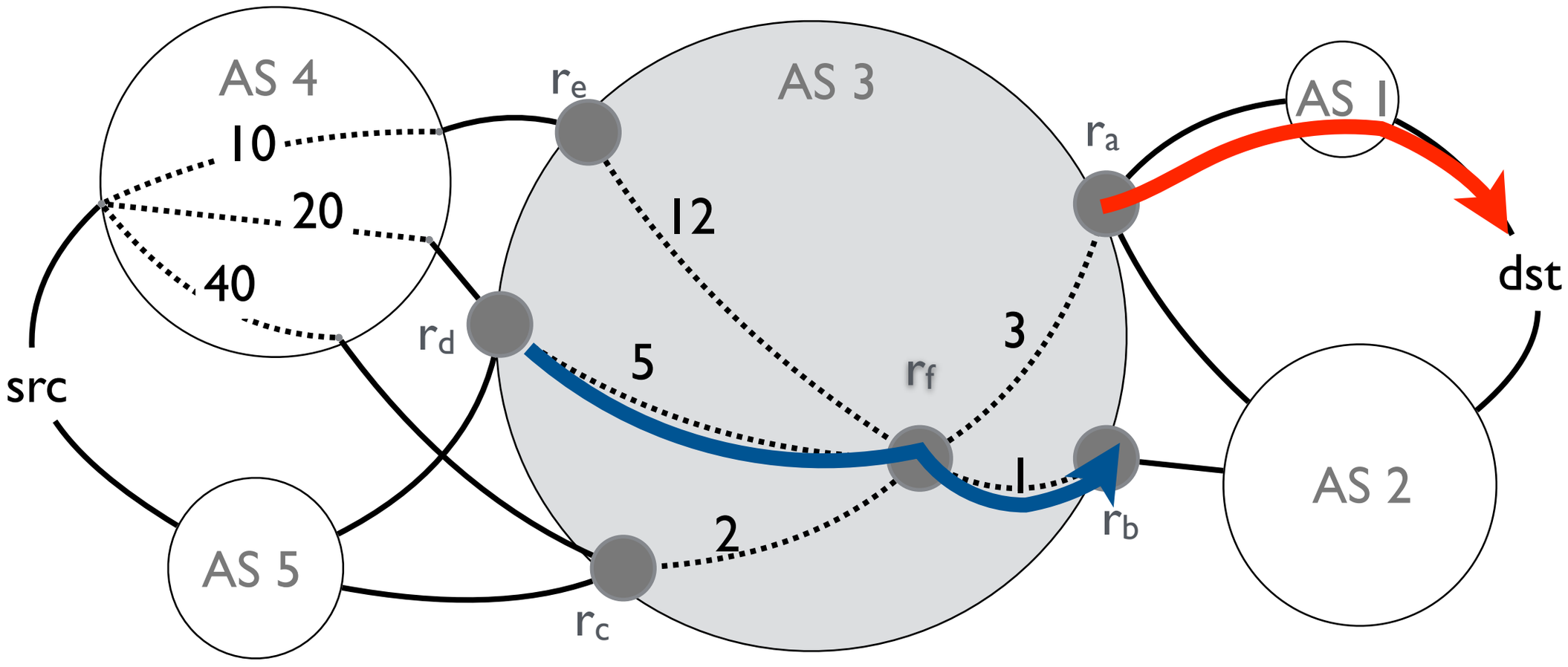
- AS 5 requests AS 3 to bypass AS 2?



influence the downstream

can AS 4 / AS 5 influence routes in AS3?

- AS 5 requests AS 3 to bypass AS 2?
- AS 4 demands AS 3 for joint traffic engineering?



influence the downstream

but the flow of policy attribute in BGP is unidirectional

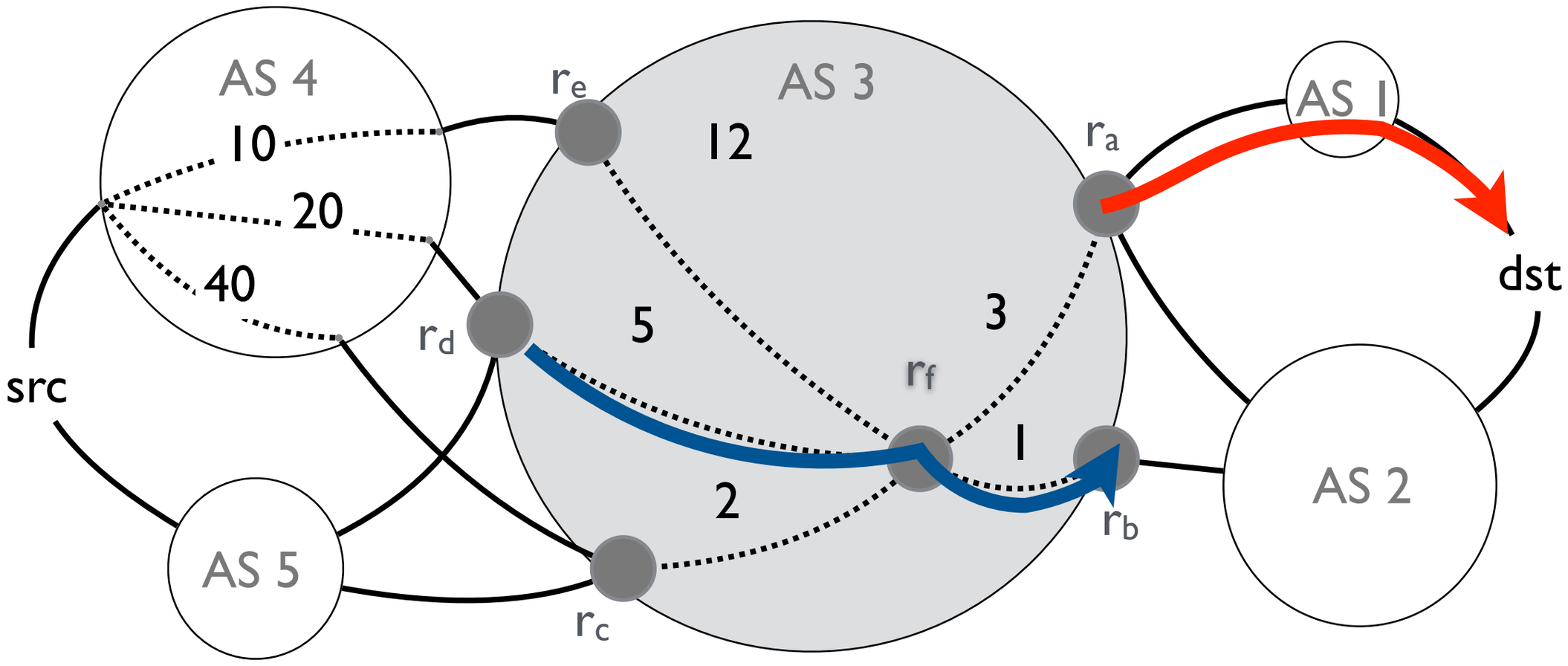
- policy carried by the routes

BGP policy is restricted
new extensions to path-vector?

influence the downstream

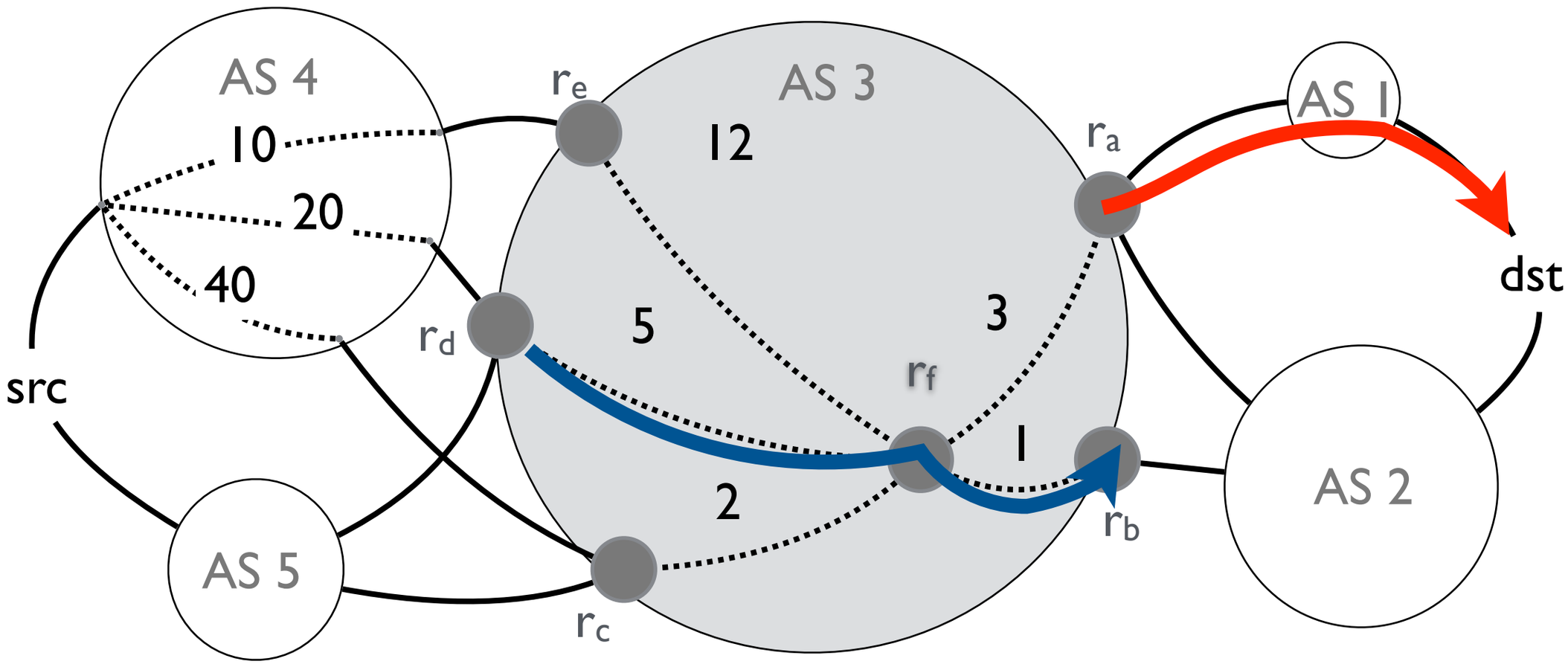
can AS 4 / AS 5 influence routes in AS3?

- AS 5 … — BGP + negotiation [MIRO, sigcomm'08]
- AS 4 … — BGP + new attribute [Wiser, sigcomm'07]



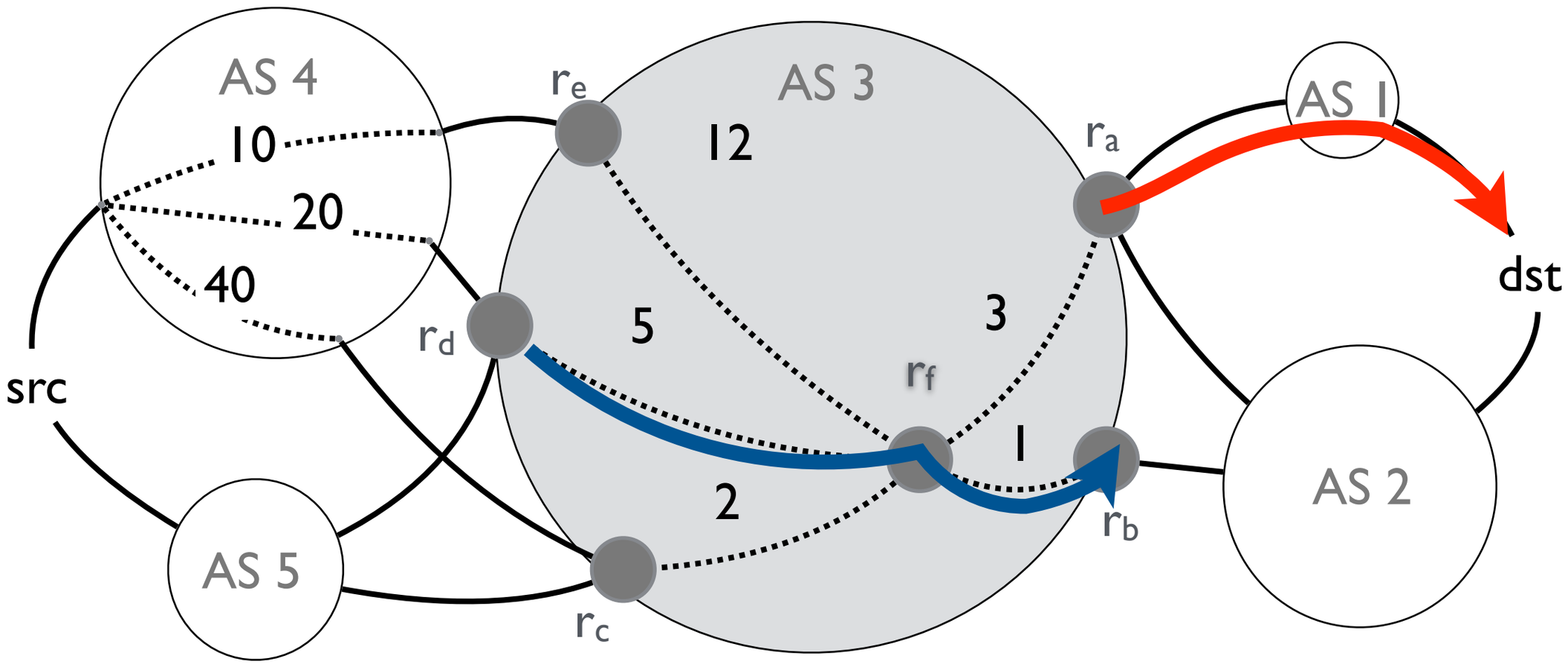
coordination

simultaneously, AS 5 wants to avoid AS 2, AS 4 demands joint TE?



coordination

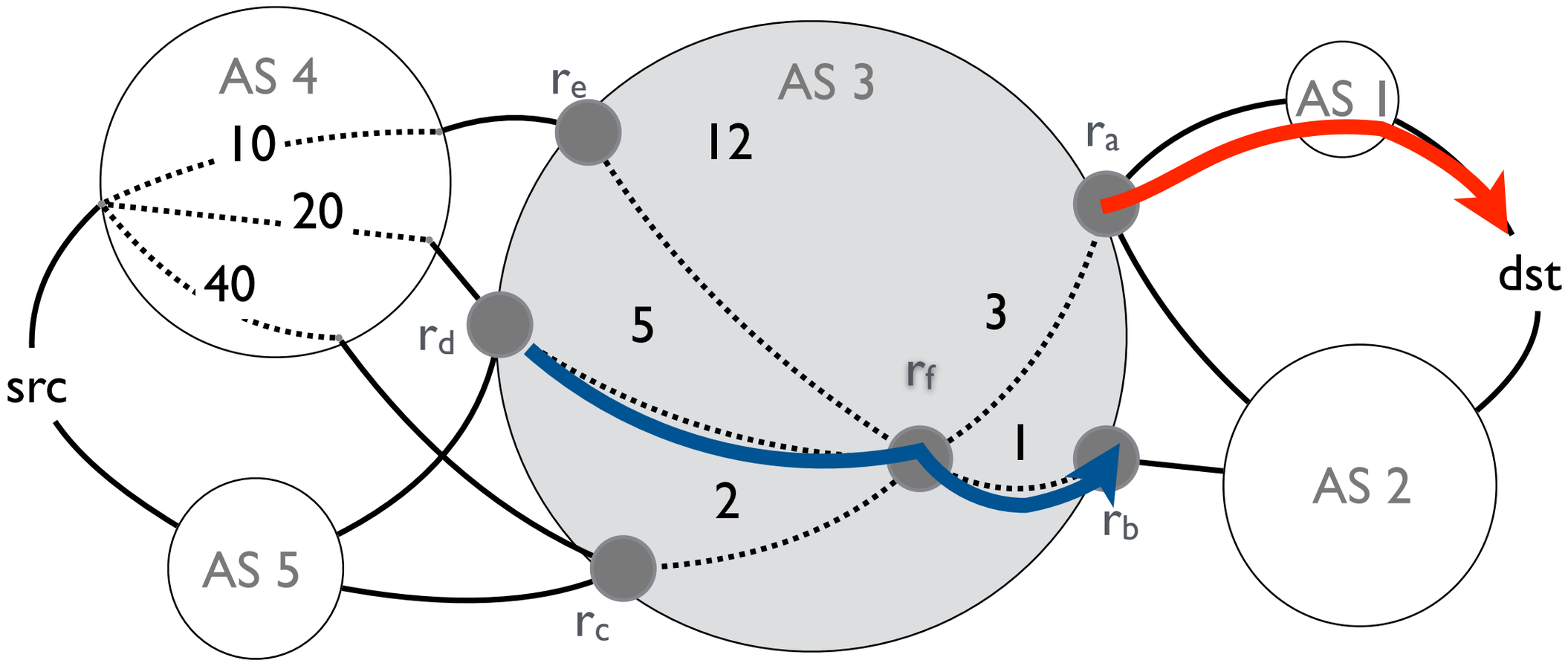
still impossible with (BGP + MIRO + Wiser)



coordination

(AS 5 avoids AS 2)+(AS 4 demands joint TE)

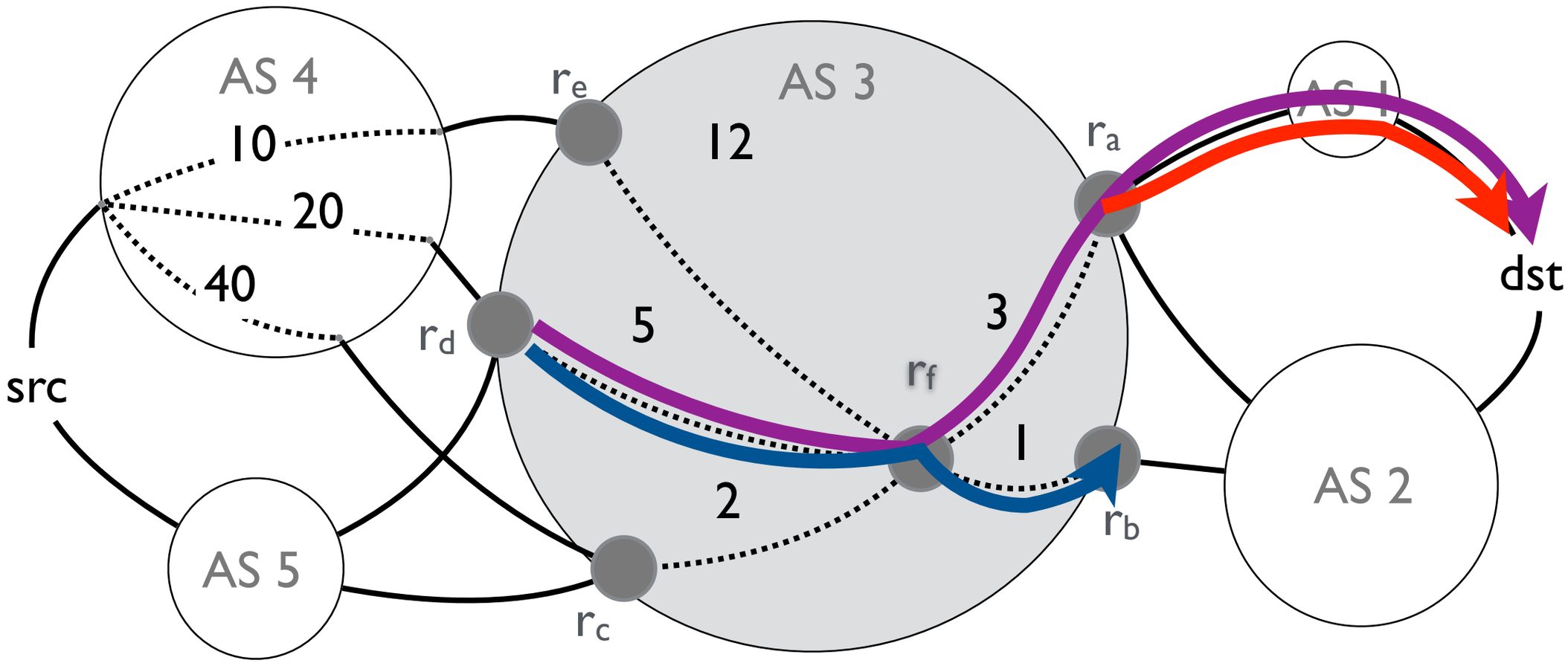
- AS 3 needs to *properly* combine the sub-routes: simple concatenation does not work!



coordination

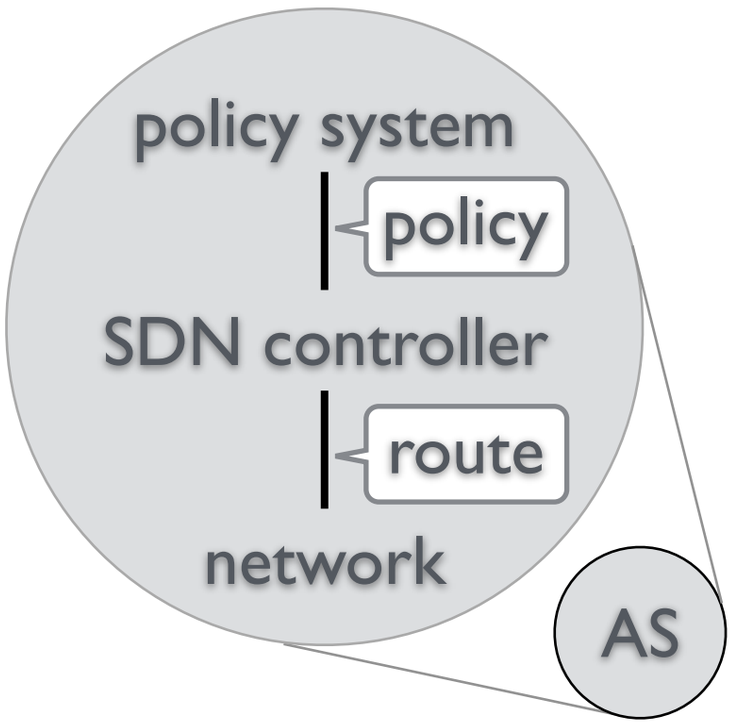
(AS 5 avoids AS 2)+(AS 4 demands joint TE)

- AS 3 needs to resolve conflicts — modify the subpath by the less important policy (AS 4 demands joint TE)



coupling routes and policies — including any path vector based policy (BGP, MIRO, Wiser …) — is inherently flawed

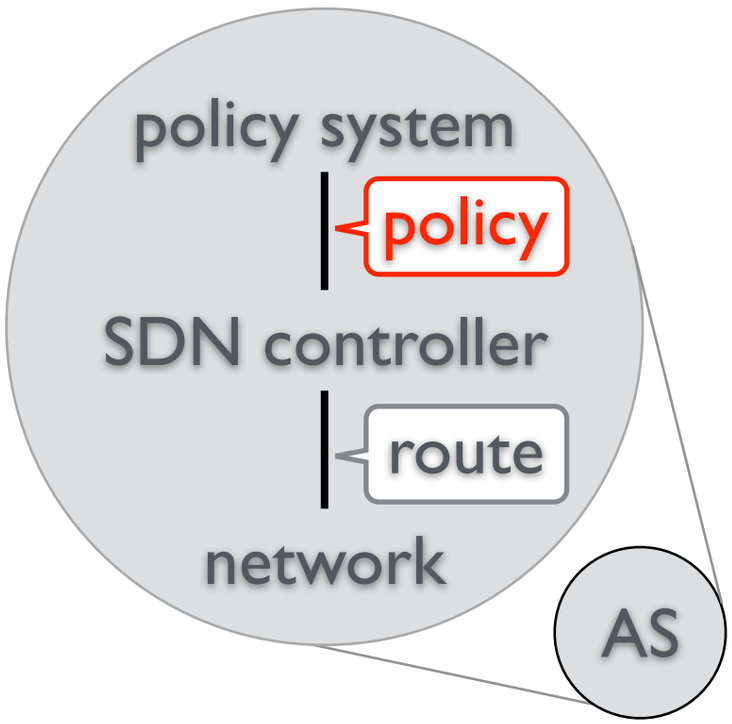
separate policies from routes



a new policy system with SDN

- SDN controller: route discovery and dissemination
- policy system: express and process high-level intention

separate policies from routes

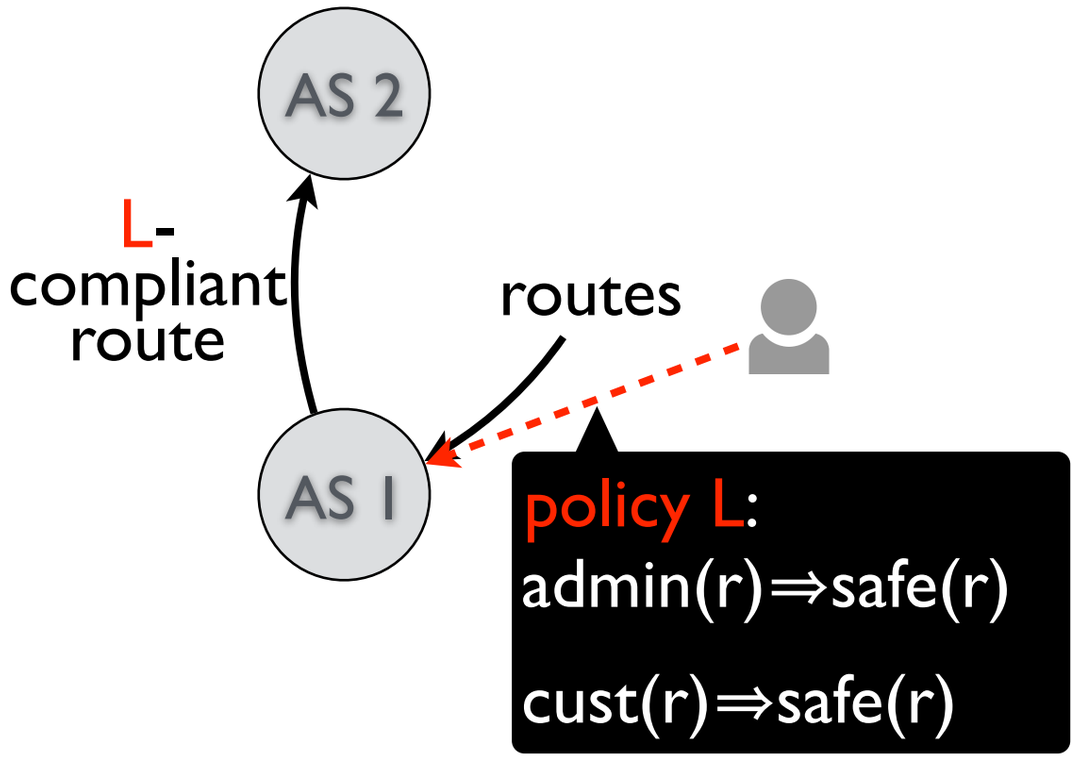


a new policy system with SDN

- SDN controller: route discovery and dissemination
- policy system: express and process high-level intention

key idea: making policies logic statements that — like routes — freely flow and interact

exchangeable logic policy

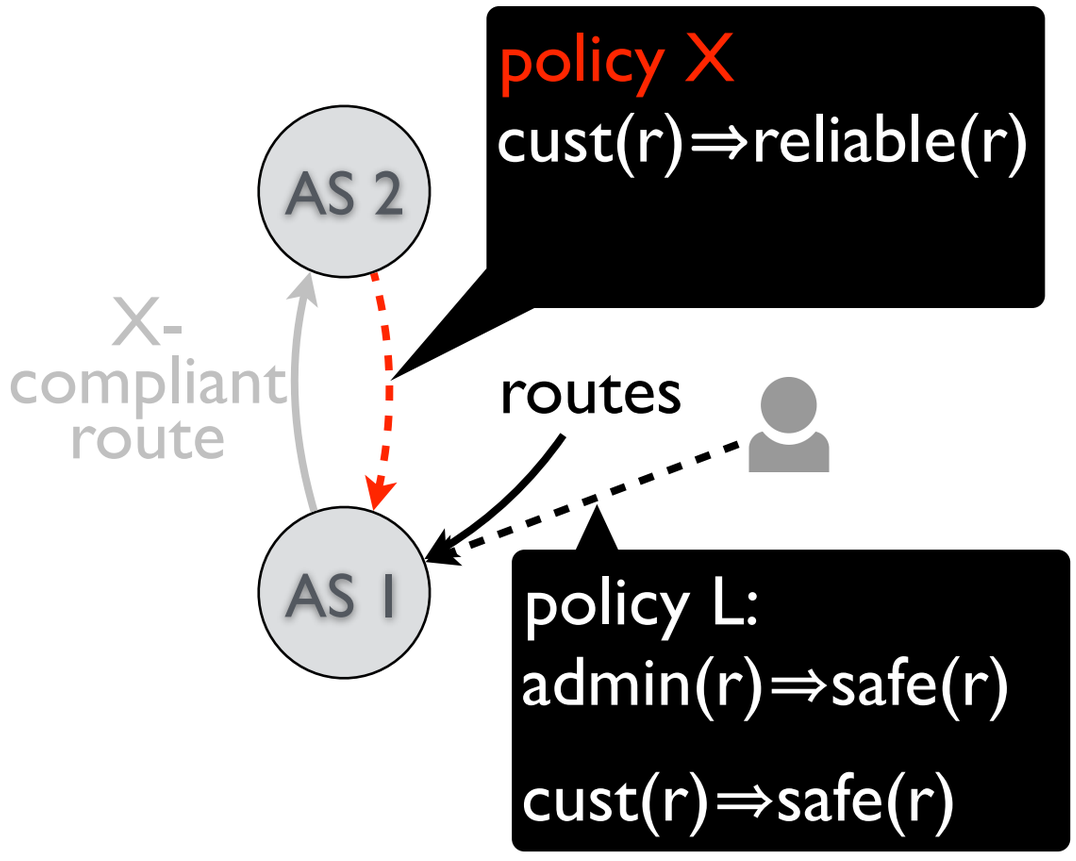


more flexible

- logic unifies disparate policies

route flow ⟶ policy flow (L,X,Y,…)⇢

exchangeable logic policy

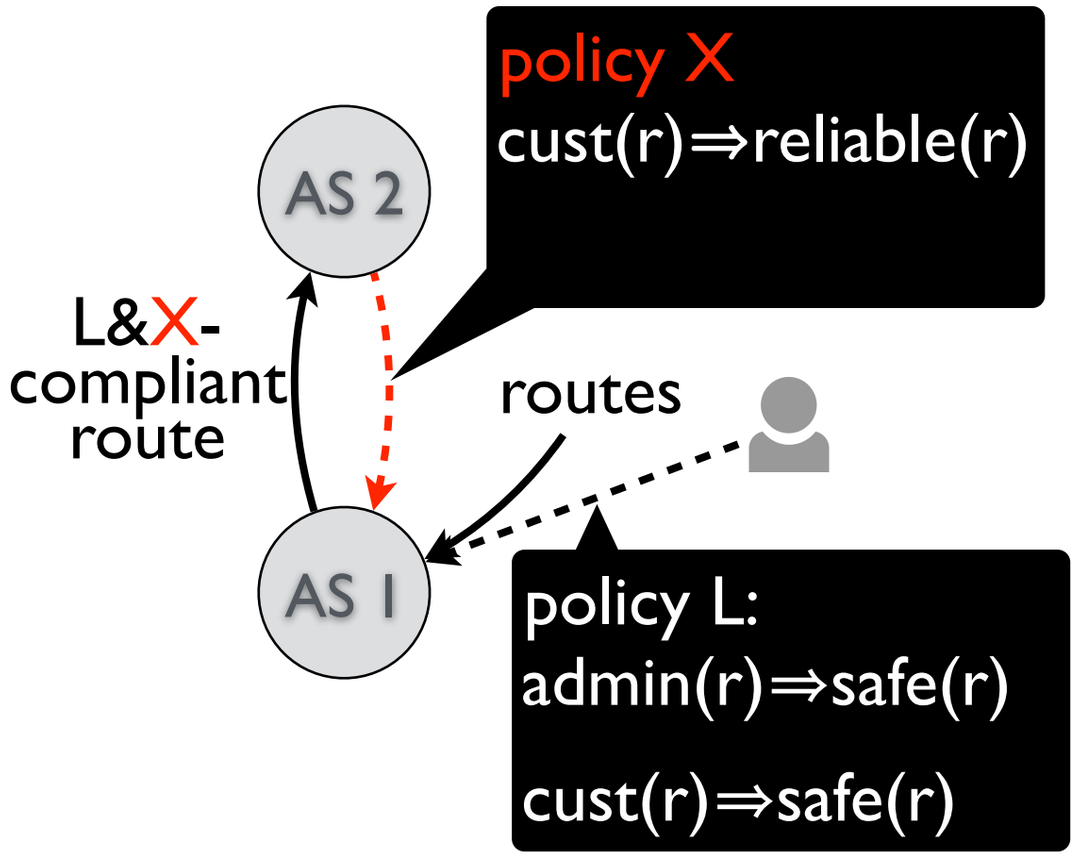


more flexible

- logic unifies disparate policies
- immediately allows control of the downstream

route flow ⟶ policy flow (L,X,Y,…)⇢

exchangeable logic policy

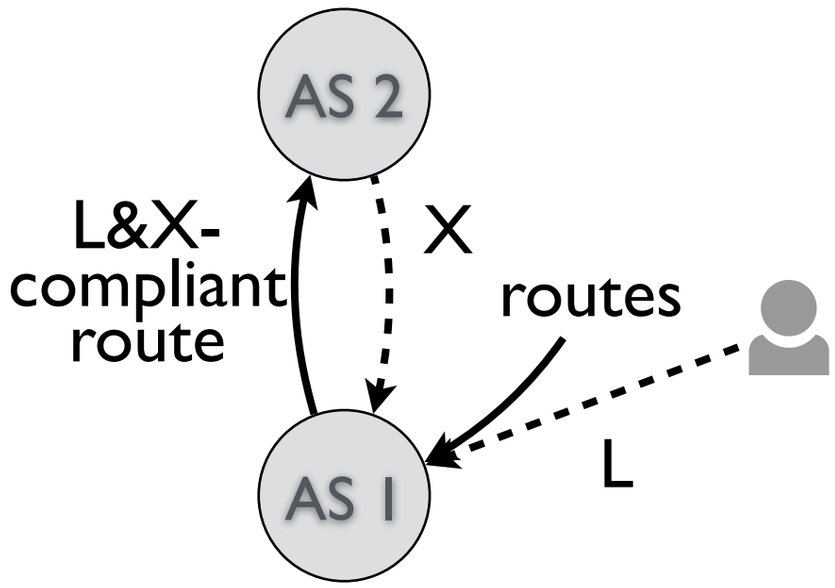


more flexible

- logic unifies disparate policies
- immediately allows control of the downstream

route flow ⟶ policy flow (L,X,Y,…) ⇢

exchangeable logic policy

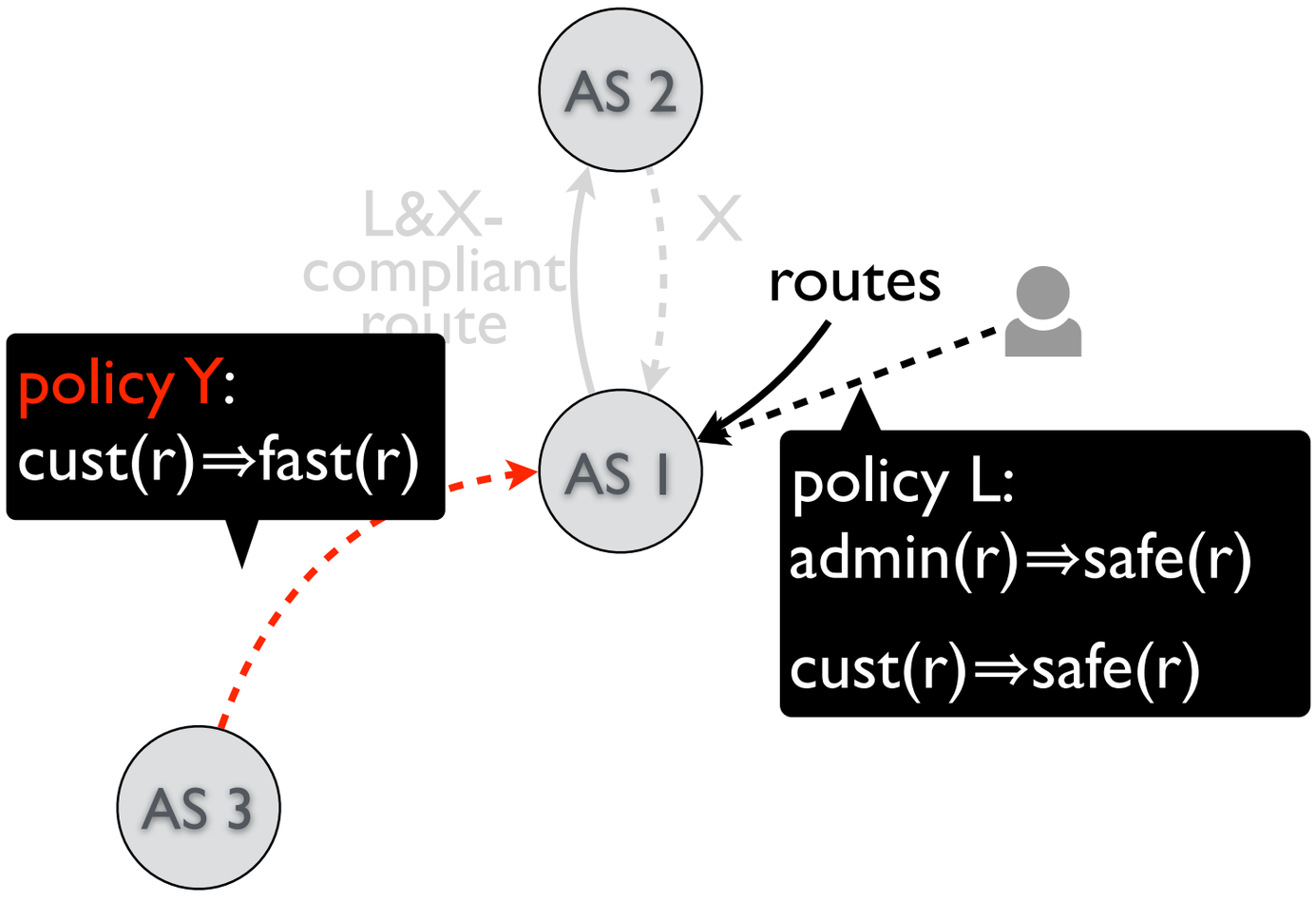


more flexible

- logic unifies disparate policies
- immediately allows control of the downstream

route flow ⟶ policy flow (L,X,Y,…) ⇢

exchangeable logic policy

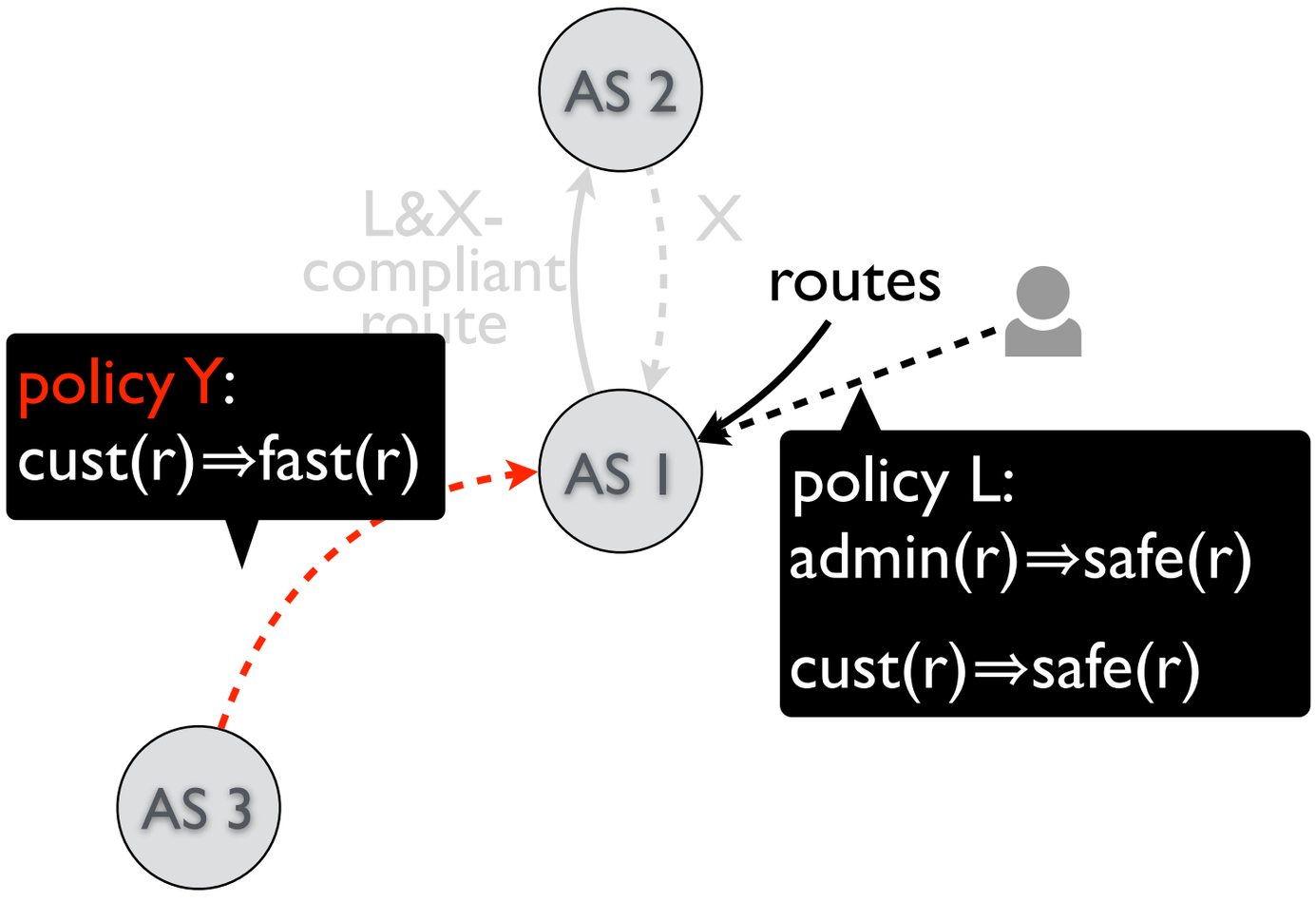


route flow ⟶ policy flow (L,X,Y,…)⇢

coordination

- compute the "impact" of a policy on another policy

exchangeable logic policy



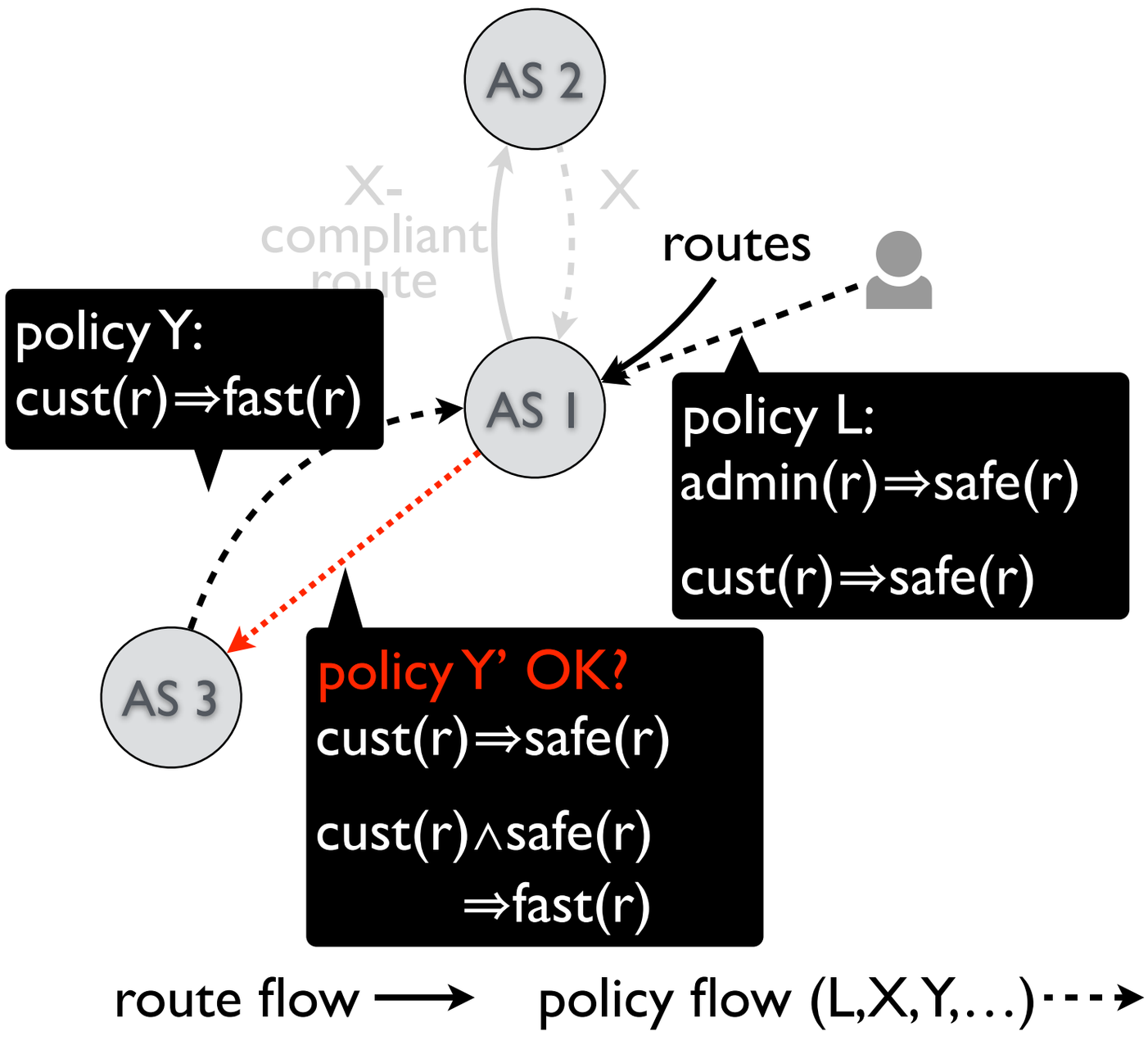
but some routes cannot be simultaneously safe and fast

route flow ⟶ policy flow (L,X,Y,…)⇢

coordination

- compute the "impact" of a policy on another policy

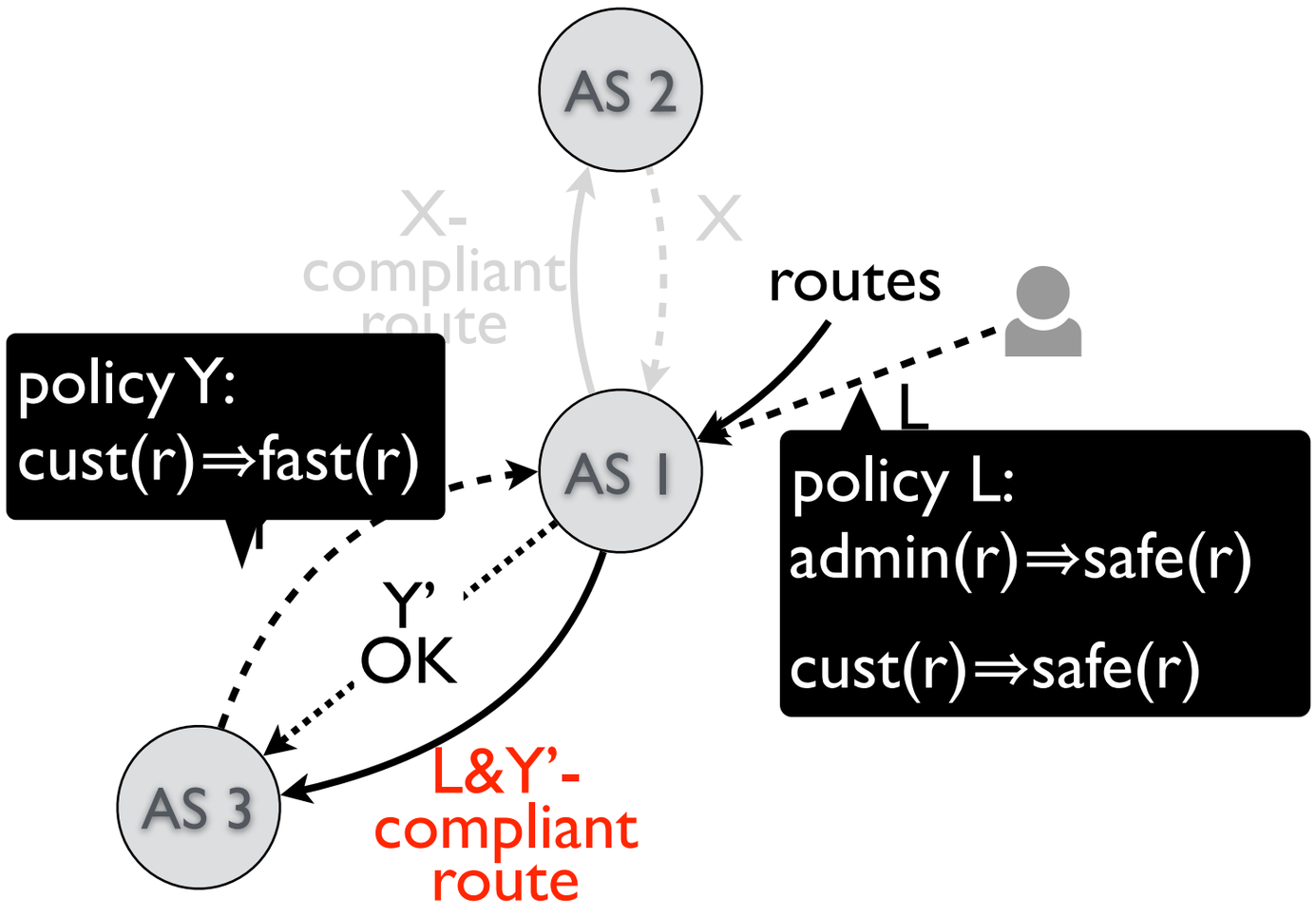
exchangeable logic policy



coordination

- compute the "impact" of a policy on another policy

exchangeable logic policy

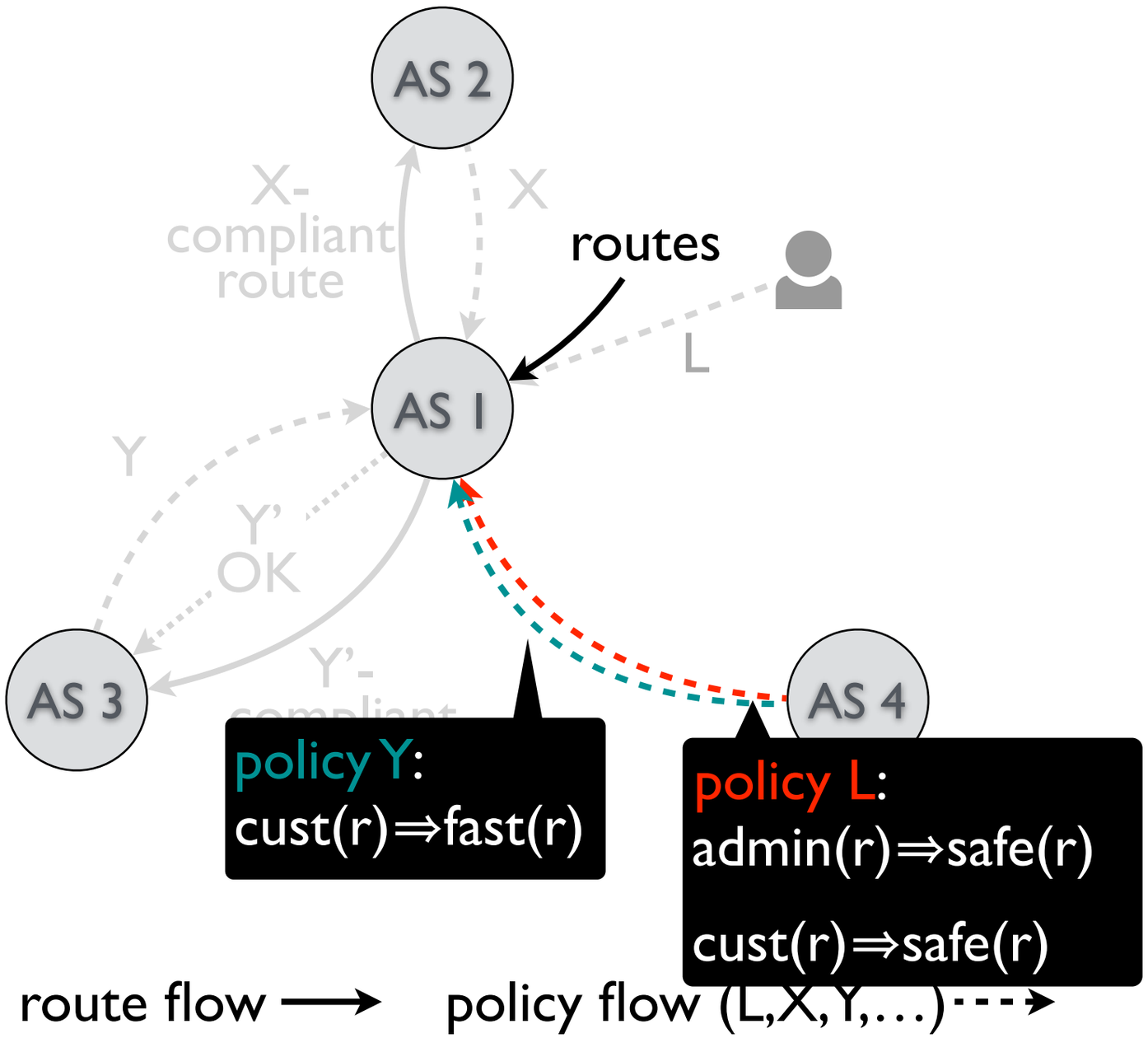


route flow ⟶ policy flow (L,X,Y,…)⇢

coordination

- compute the "impact" of a policy on another policy

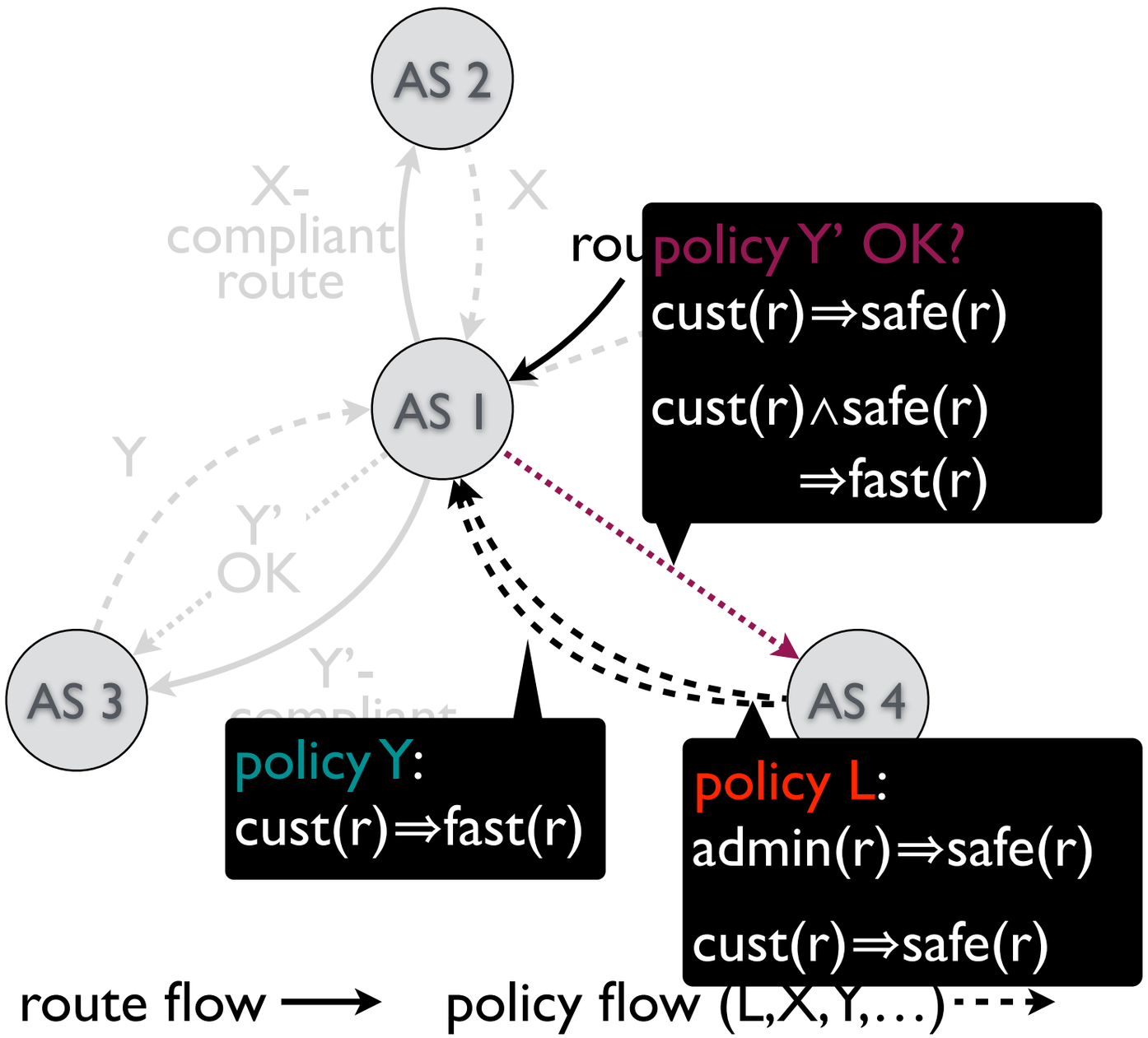
exchangeable logic policy



coordination

- compute the "impact" of a policy on another policy

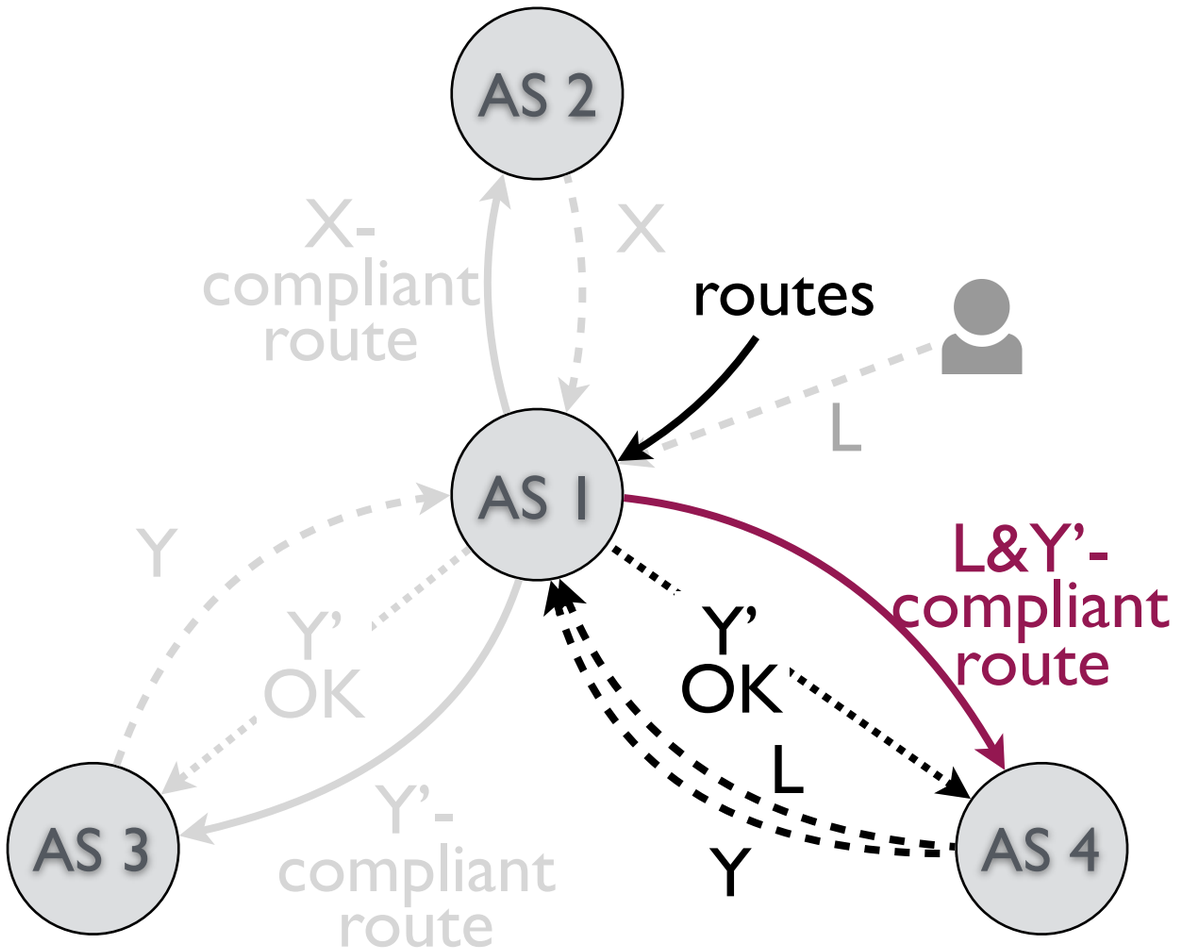
exchangeable logic policy



coordination

- compute the "impact" of a policy on another policy

exchangeable logic policy

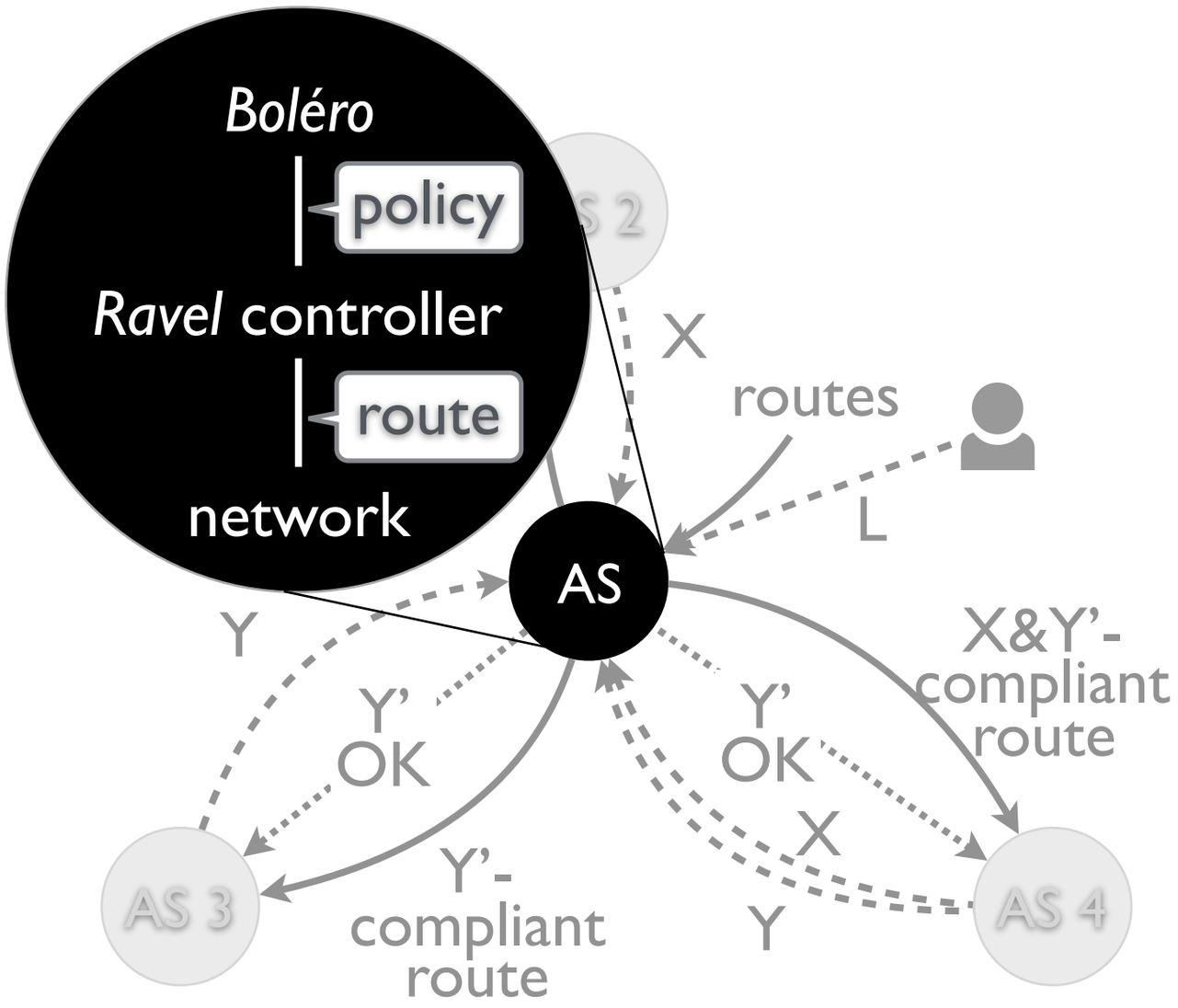


route flow ⟶ policy flow (L,X,Y,…) ⇢

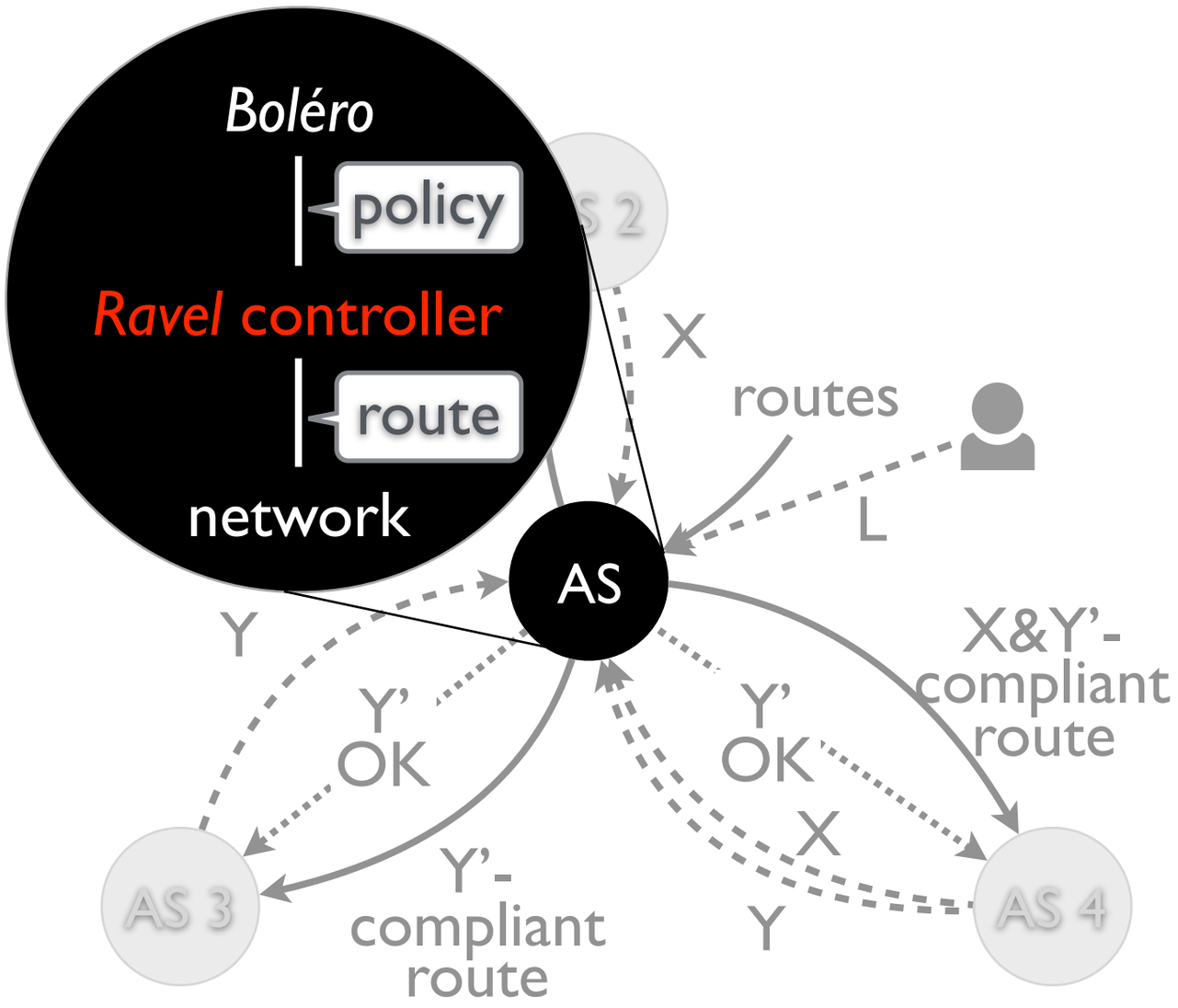
coordination

- compute the "impact" of a policy on another policy

Boléro — a realization with *Ravel*



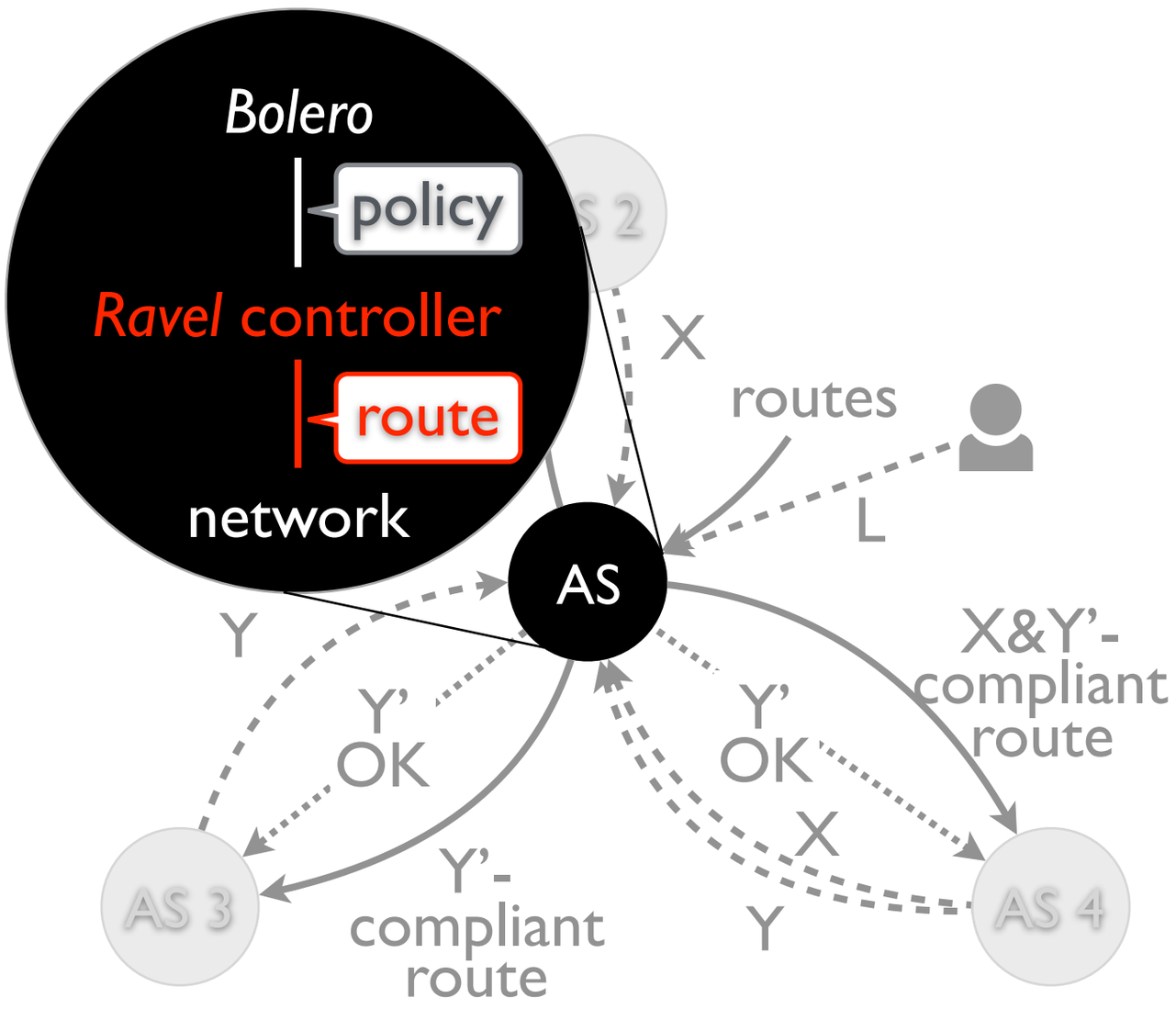
Boléro — a realization with *Ravel*



Ravel

- uses a high-performance database as the controller

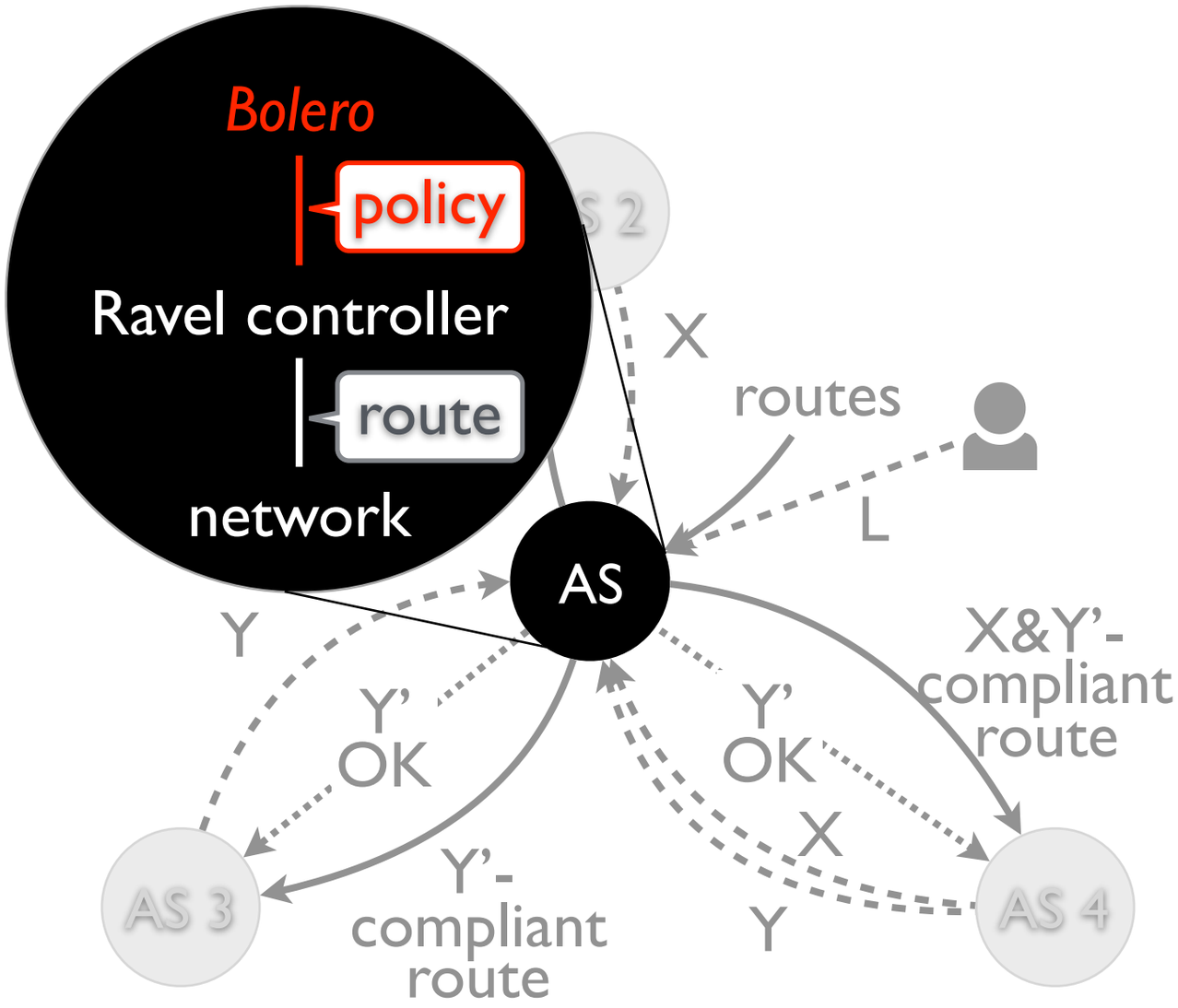
Boléro — a realization with *Ravel*



Ravel

- uses a high-performance database as the controller
- discovers and disseminates routes (maintain database tables)

Boléro — a realization with *Ravel*



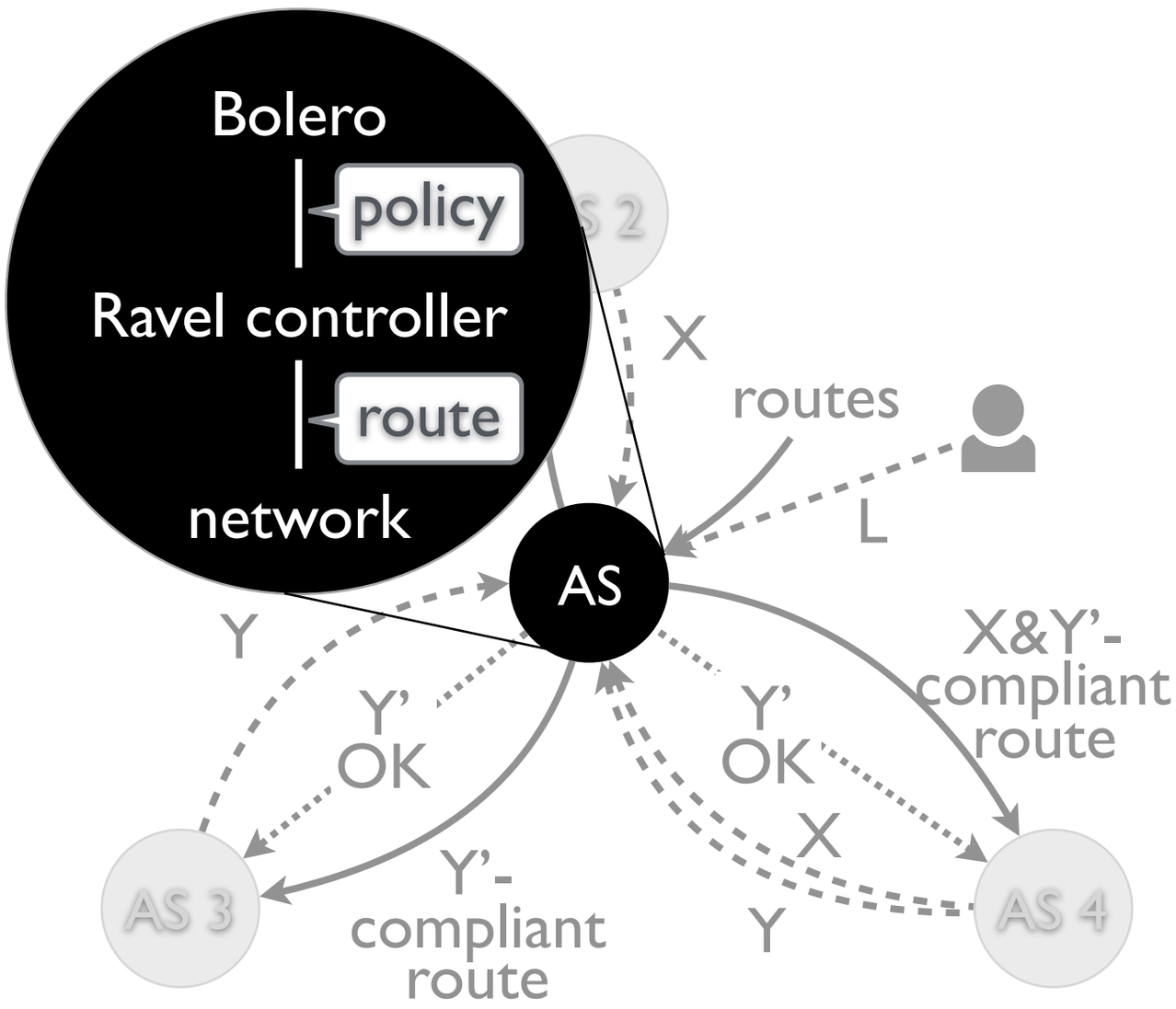
Ravel controller

- uses a high-performance database as the controller
- discovers and disseminates BGP routes

Boléro policy system

- exchanges and processes logic policies

Boléro — a realization with *Ravel*



Ravel controller

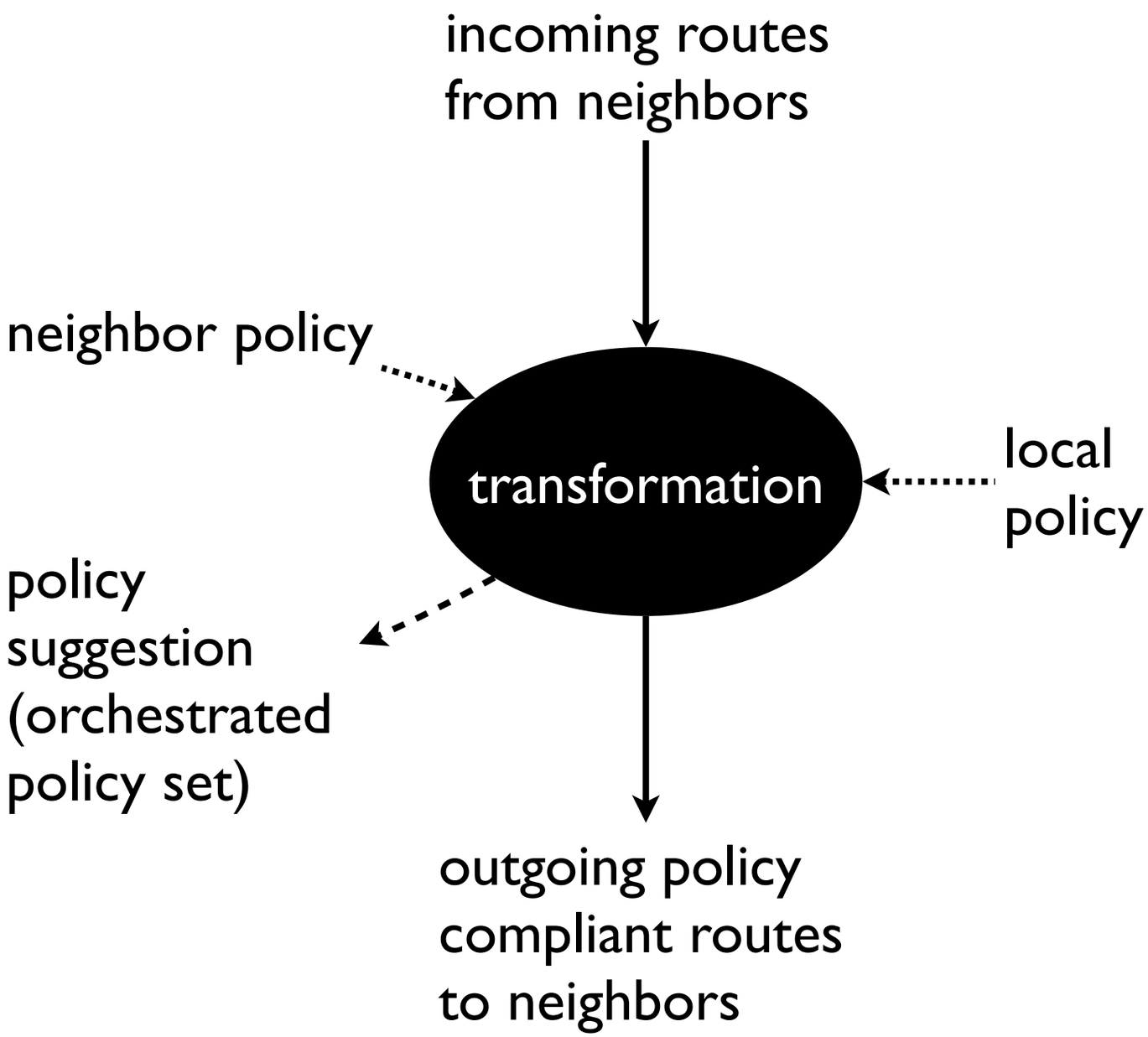
- uses a high-performance database as the controller
- discovers and disseminates BGP routes

Boléro policy system

- exchanges and processes logic policies

key idea: process policy — semantic knowledge — by semantic transformation

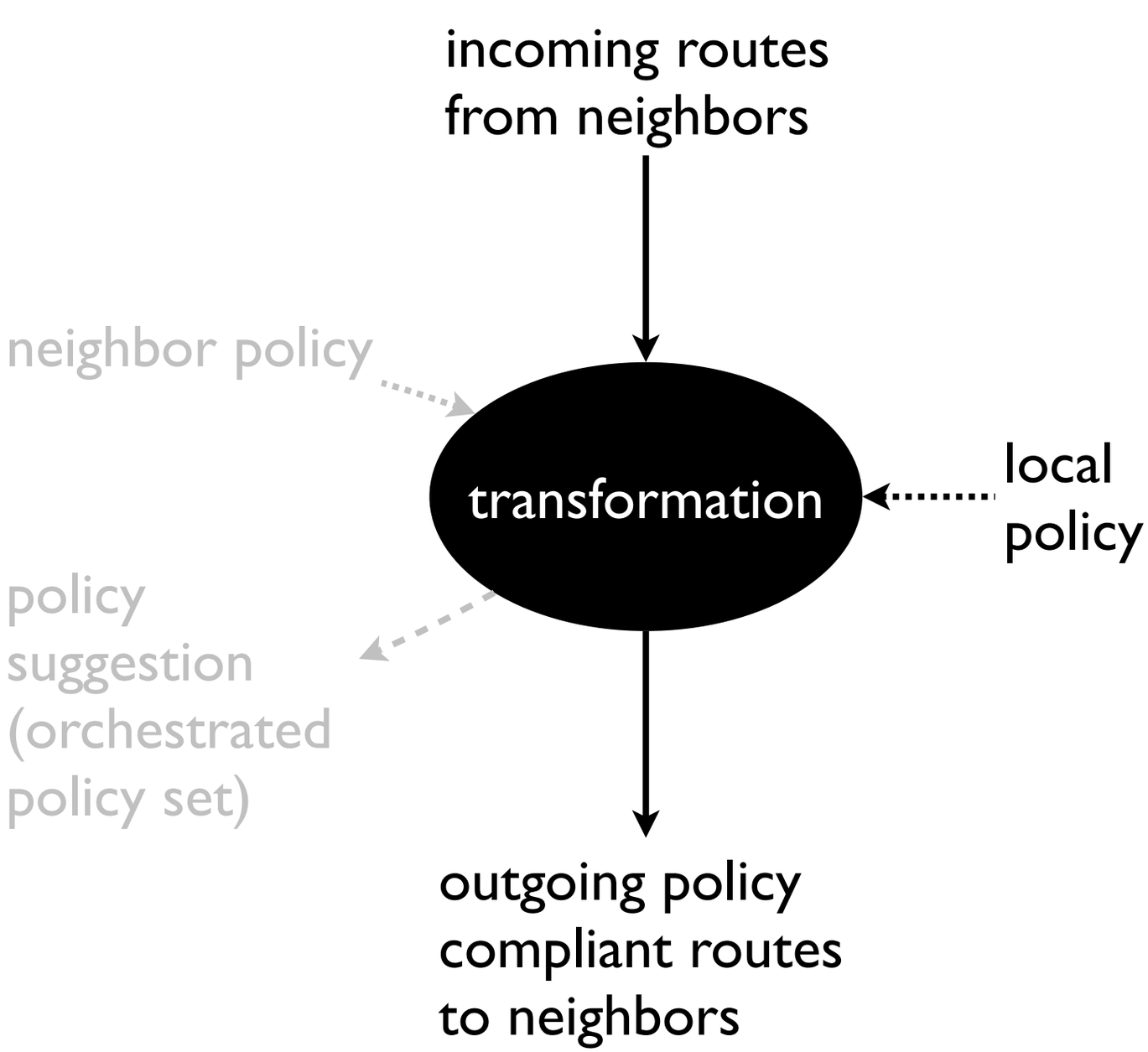
Boléro overview



transformation

- flexible policy routing
- policy coordination

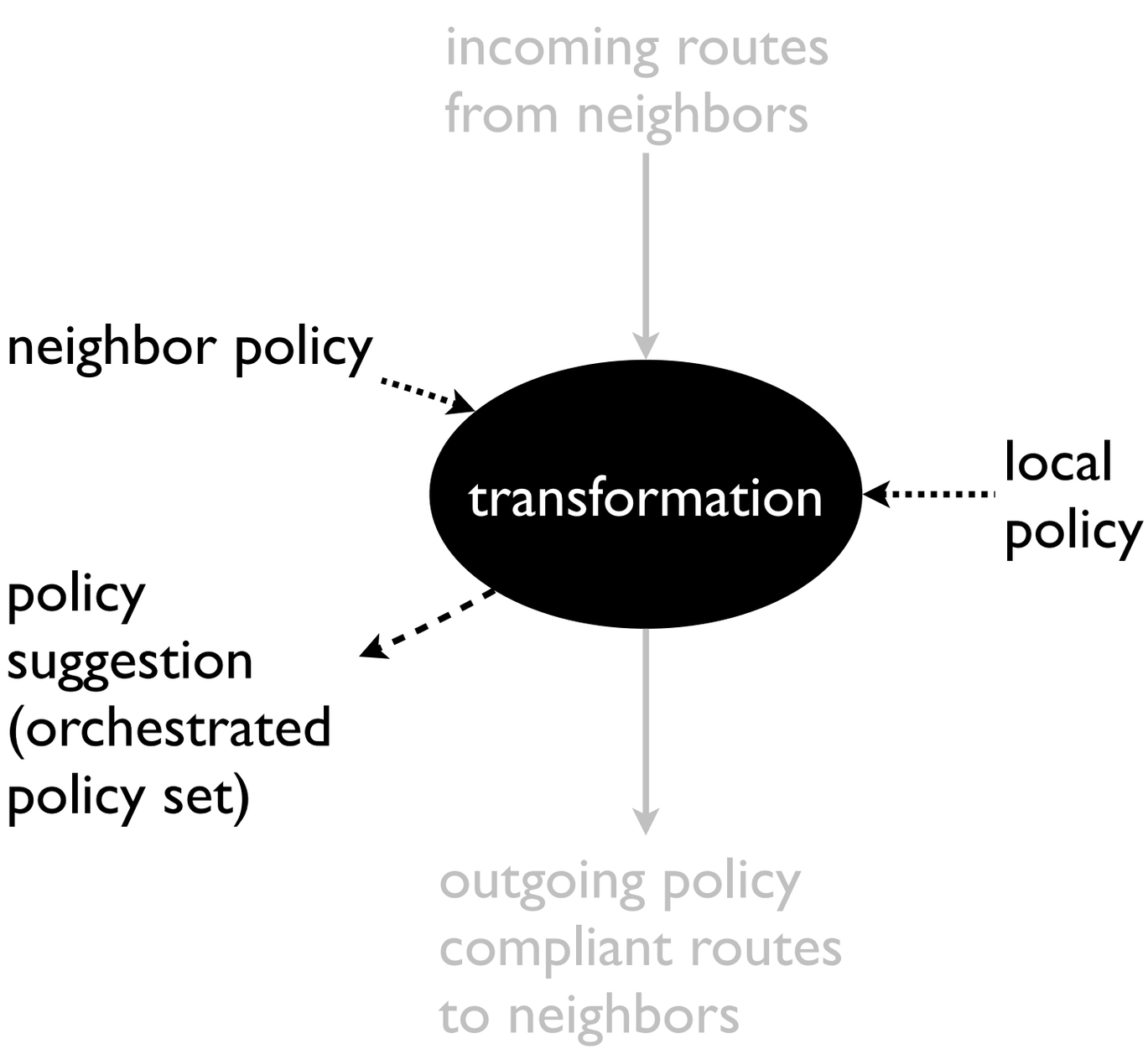
Boléro overview



transformation

- policy routing
  - transform routes using the policies

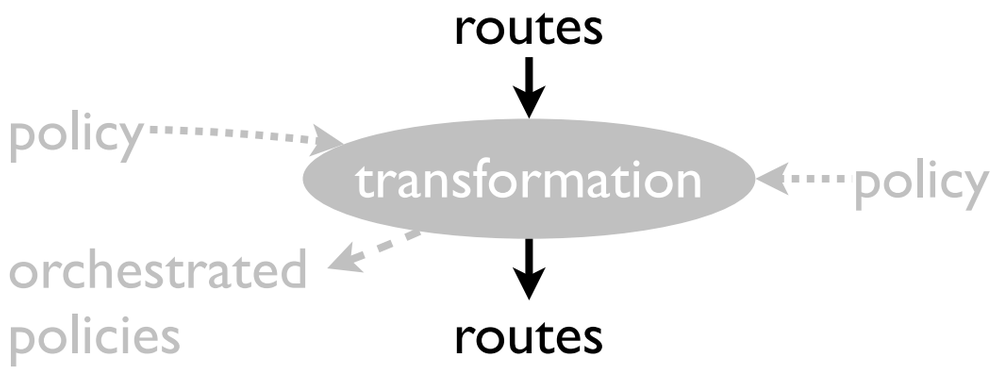
Boléro overview



transformation

- policy routing
  - transform routes using the policies
- policy coordination
  - transform one by another policy

representation

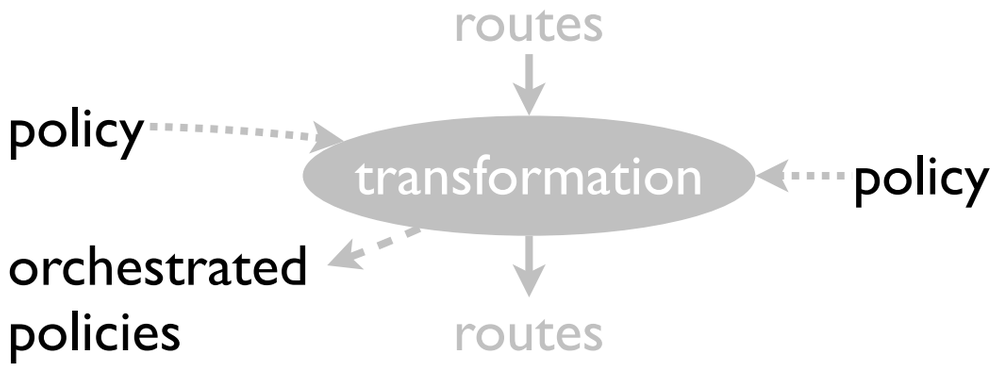


route

prefix	next hop	AS path	…

- routing state as queryable tables
  - *factual data*

representation



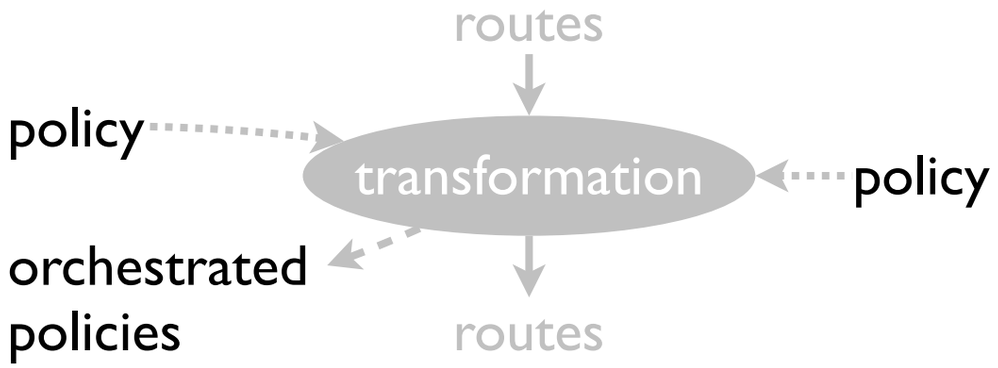
route

prefix	next hop	AS path	…

```
--- MIRO-like policy
:- route(D,N,P), ('AS2' in P).
```

- routing state as queryable tables
  - *factual data*
- policy as data integrity constraint (IC)
  - logical statement about what (must be avoided) are the valid route data
  - *semantic data*

representation



route

prefix	next hop	AS path	…

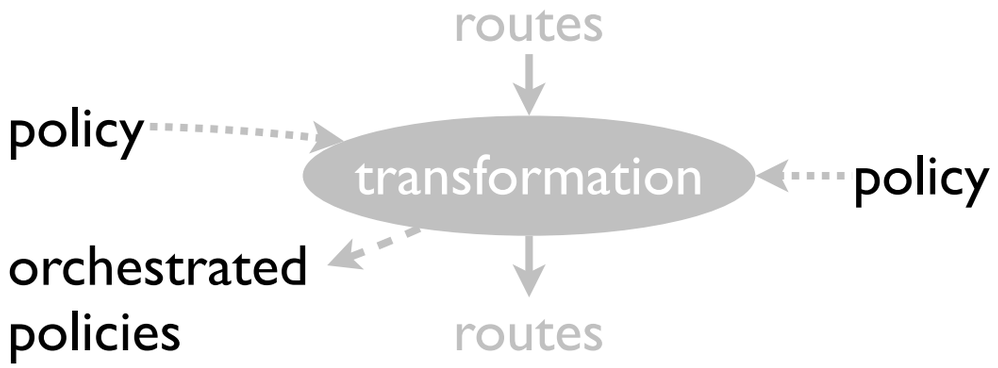
```
--- MIRO-like policy
:- route(D,N,P), ('AS2' in P).
```

meaning

```
False <—
route(I,R,P)/\ {(‘AS2’ in P)}
```

- routing state as queryable tables
  - *factual data*
- policy as data integrity constraint (IC)
  - logical statement about what (must be avoided) are the valid route data
  - *semantic data*

representation



route

prefix	next hop	AS path	…

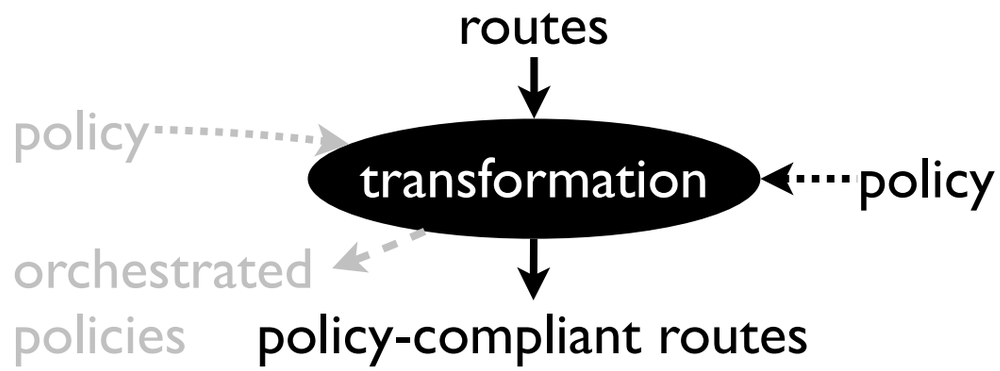
```
--- MIRO-like policy
:- route(D,N,P), ('AS2' in P).
```

```
--- Wiser-like policy
:- route(D,N,P),
   Wiser(D,R,C),Advertise(R,C2),
   Wiser(D,R',C'),Advertise(R',C2'),
   C+C2>C'+C2'.
```

- routing state as queryable tables
  - *factual data*
- policy as data integrity constraint (IC)
  - logical statement about what (must be avoided) are the valid route data

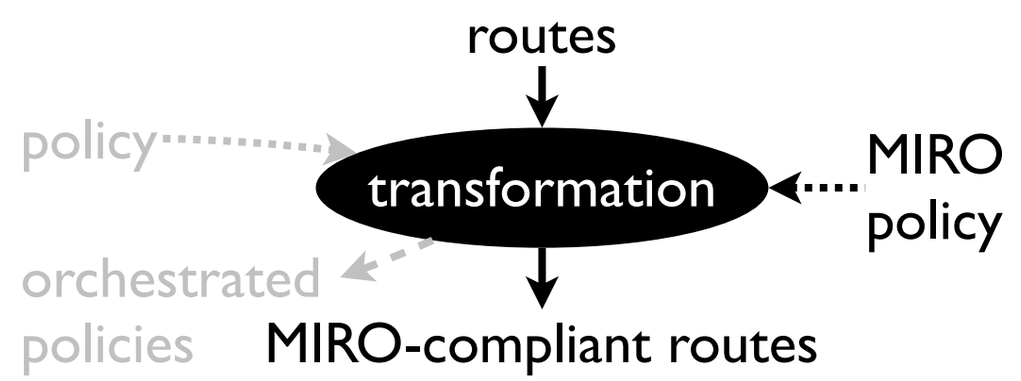
semantic data

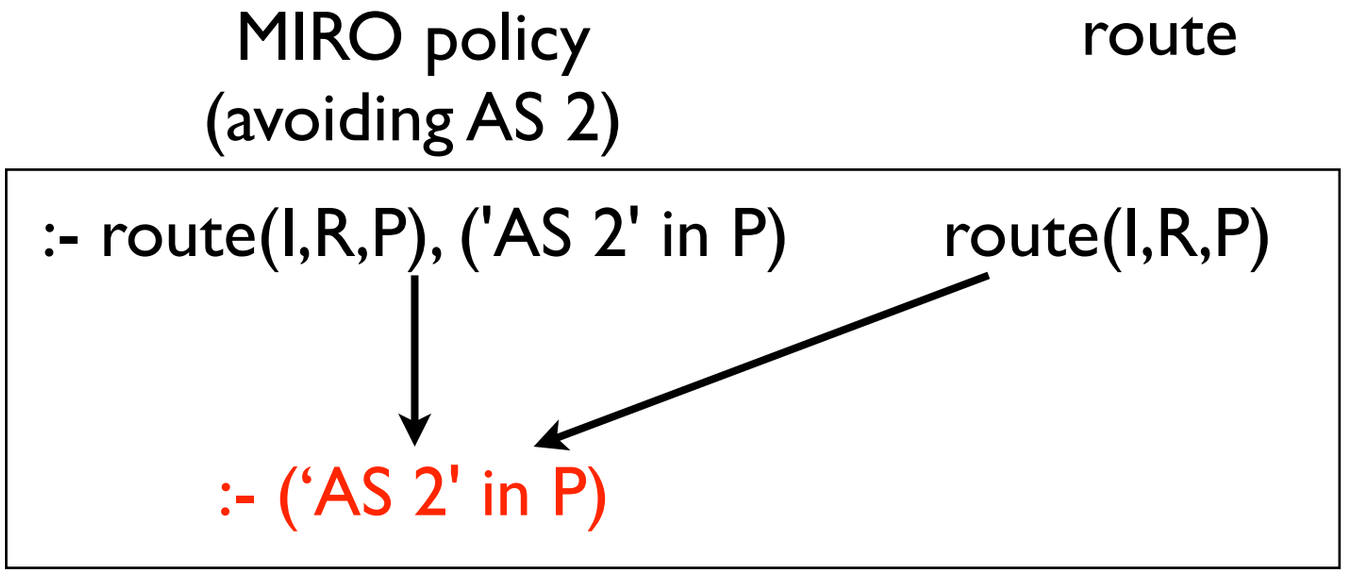
route transformation



- compute the "impact" of a policy

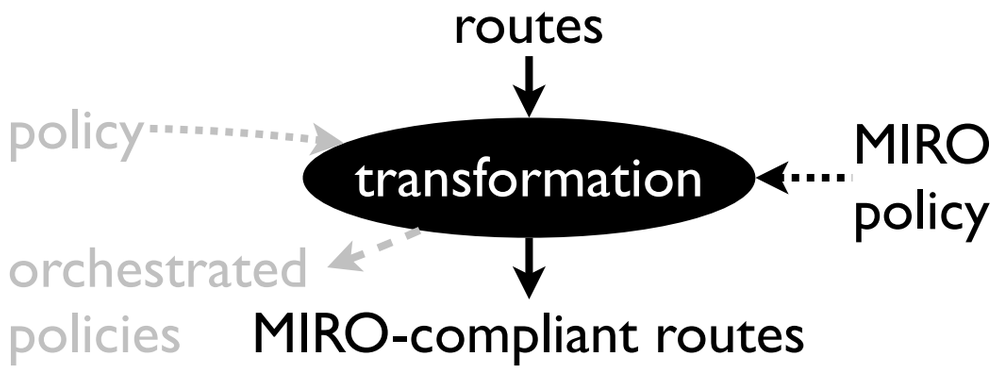
route transformation

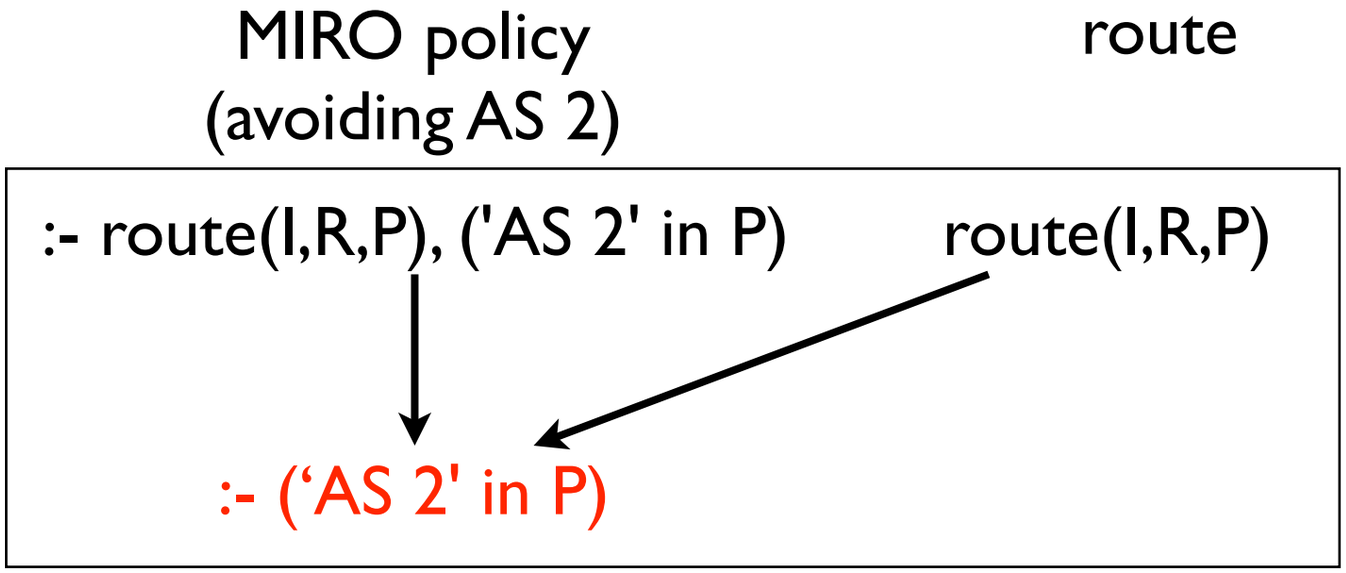




- compute the "impact" of a policy
  - obtain residue by partial subsumption

route transformation



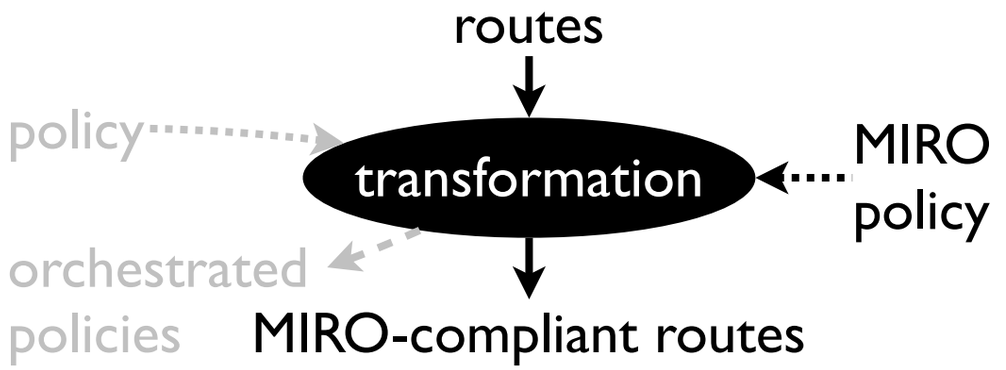


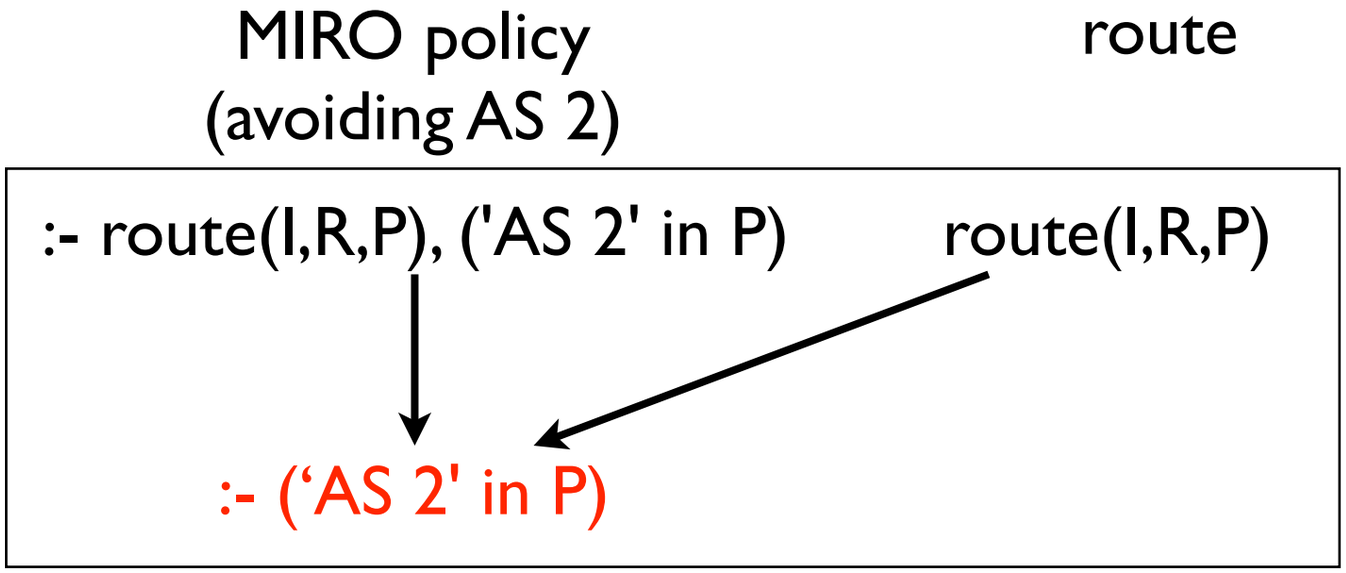
- compute the "impact" of a policy
  - obtain residue by partial subsumption
- attach the residue

MIRO-compliant route

```
route(I,R,P) {:-('AS2' in P)}
```

route transformation





- compute the "impact" of a policy
  - obtain residue by partial subsumption
- attach the residue

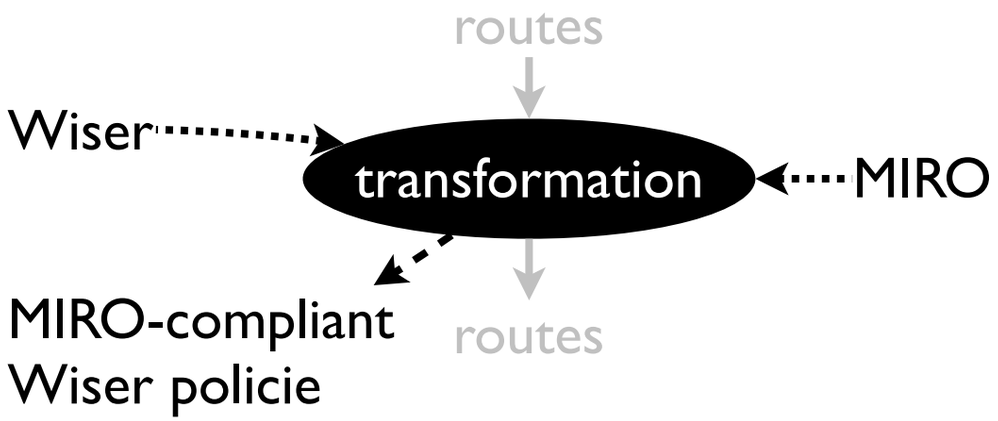
MIRO-compliant route

```
route(I,R,P) {:-('AS2' in P)}
```

meaning

{¬('AS2' in P)}

policy transformation



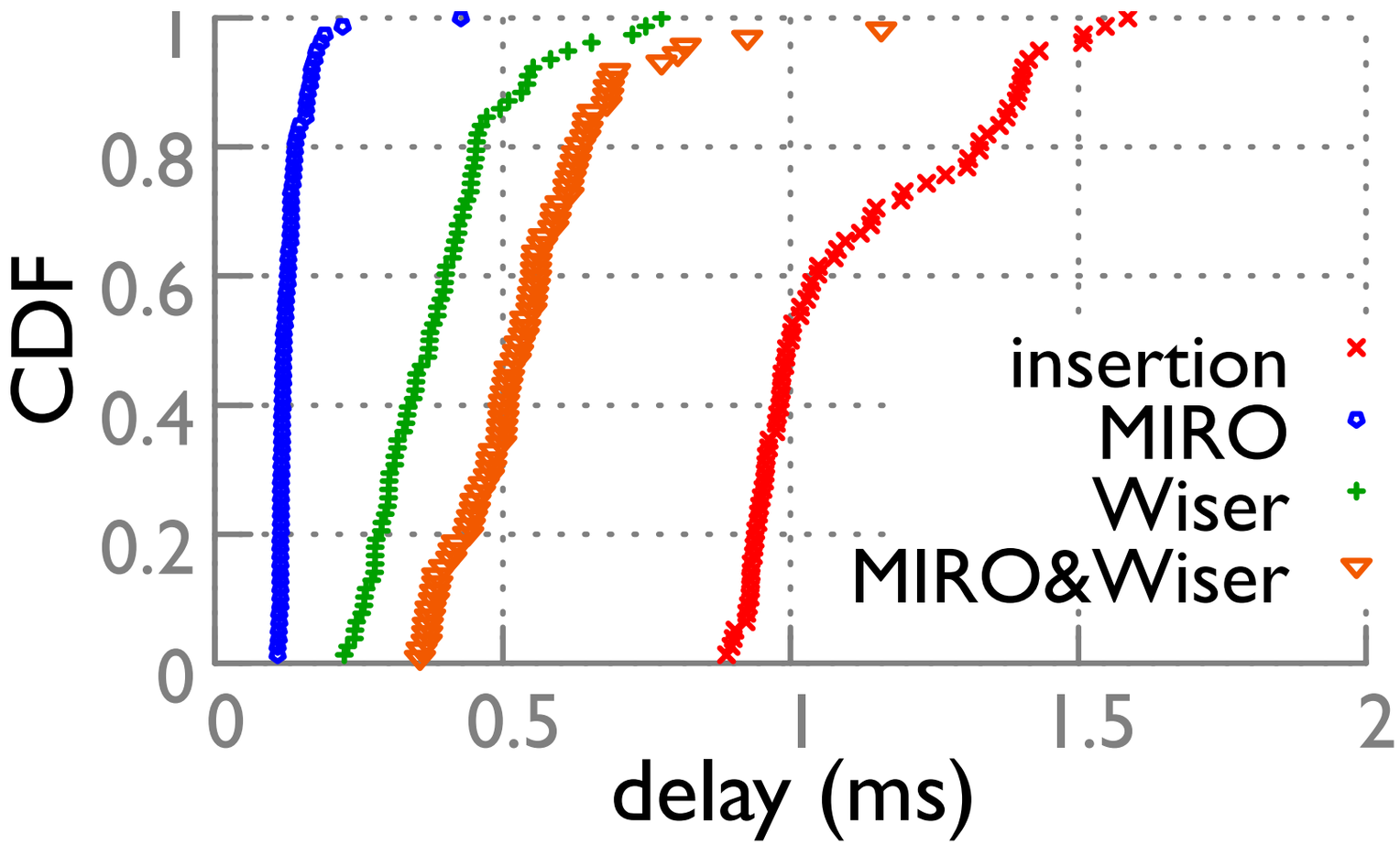

```
--- Wiser policy
:- route(D,N,P),
   Wiser(D,R,C),Advertise(R,C₂),
   Wiser(D,R’,C’),Advertise(R’,C₂'),
   C+C2>C'+C2'.
```

```
--- MIRO-compliant Wiser policy
:- route(D,N,P),
   Wiser(D,R,C),Advertise(R,C₂),
   Wiser(D,R’,C’),Advertise(R’,C₂'),
   C+C2>C’+C2’, {¬(‘AS2' in P)}.
```

- compute the “impact” of a policy
- attach the residue

preliminary result

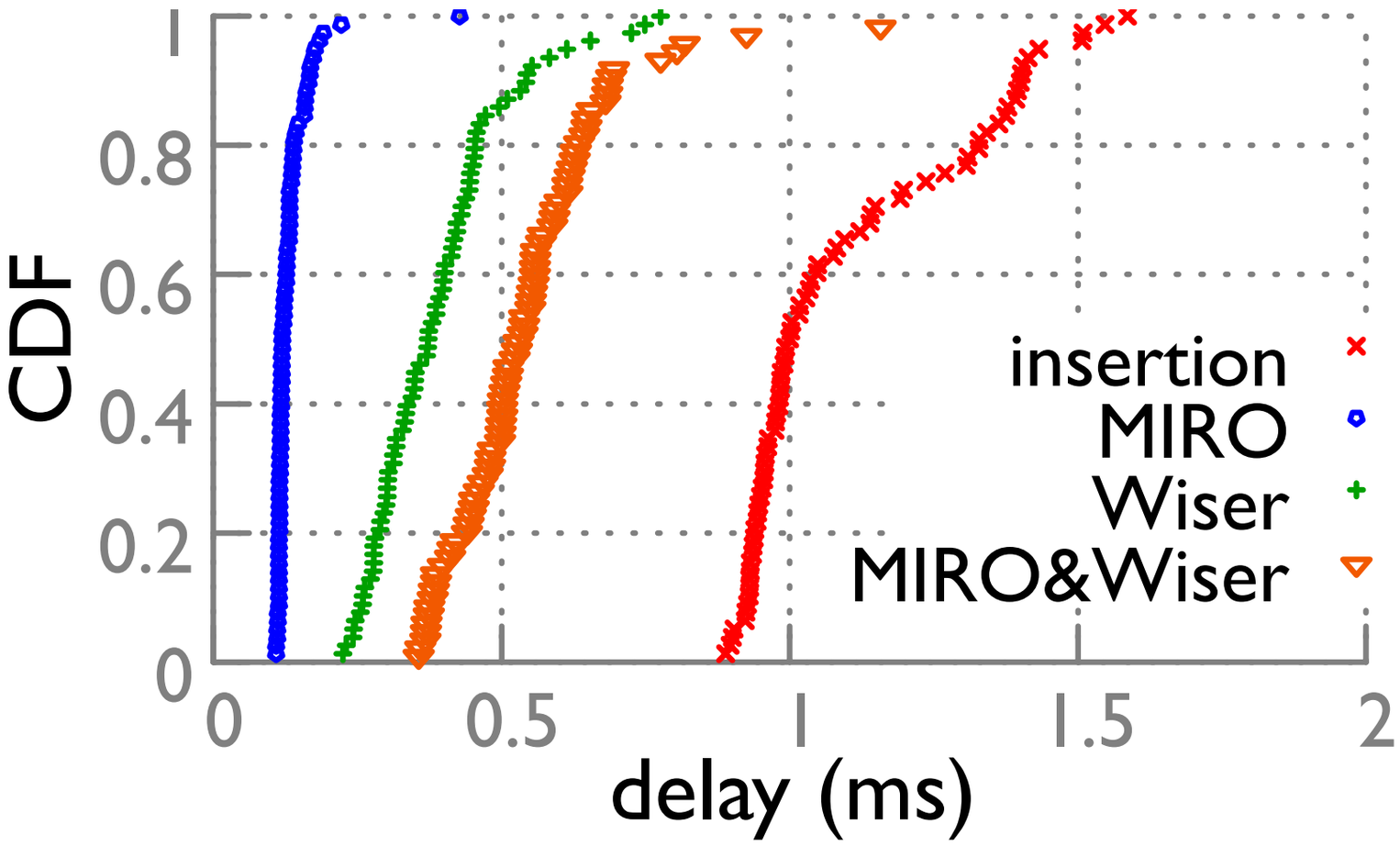
time the database delay for computing policy-compliant routes



- policies
  - MIRO
  - Wiser
  - MIRO&Wiser
- topology: Pocketful ISP topology embedded in Skitter AS-level topology
- incoming routes
  - Routeview BGP feeds

preliminary result

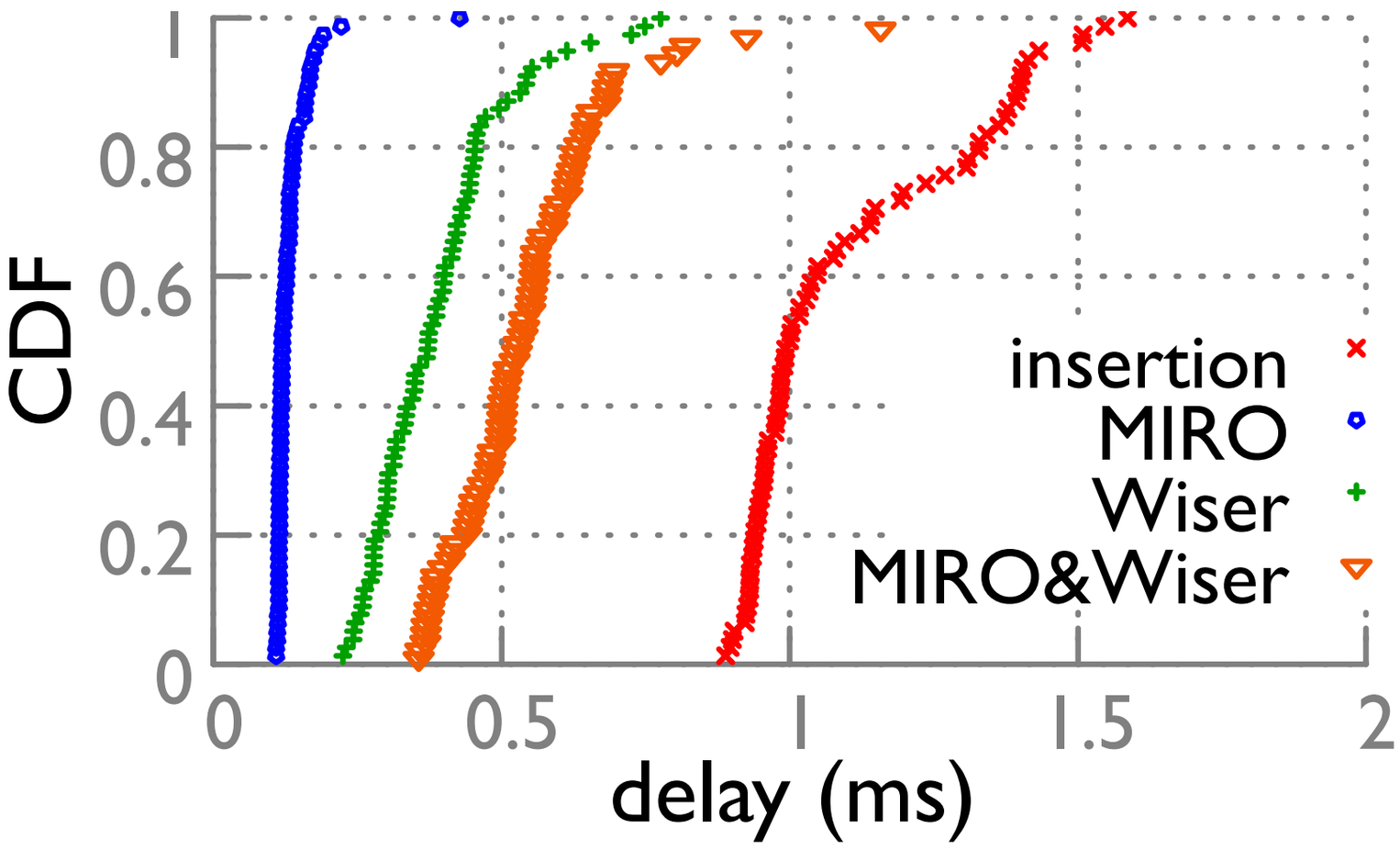
time the database delay for 10,000 BGP feeds



- delay is small and scales well
  - 95% route insertion < 1.424ms
  - 95% MIRO < .174ms
  - 95% Wiser < .64ms
  - 95% Wiser&MIRO < .844ms

preliminary result

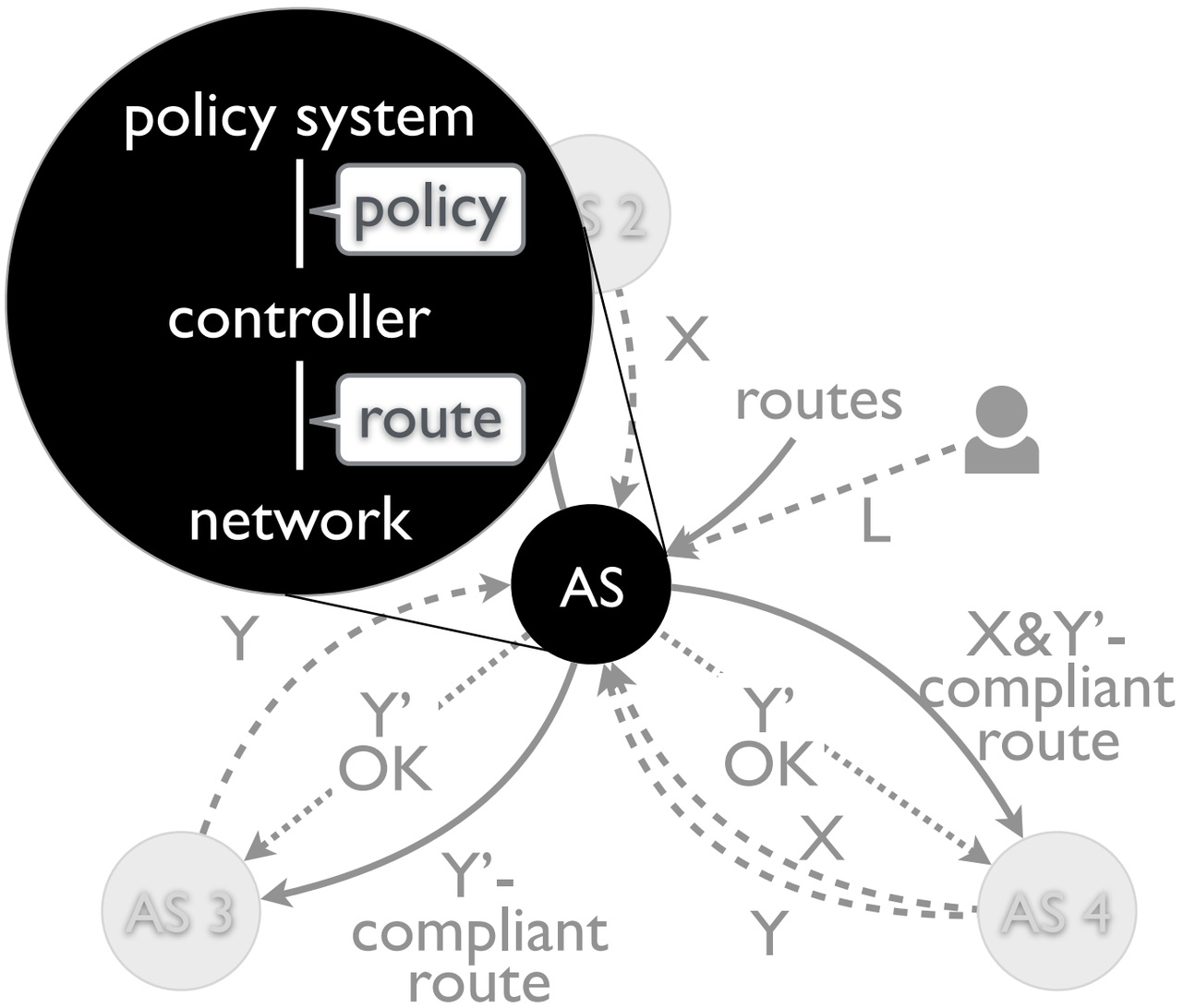
time the database delay for 10,000 BGP feeds



- delay is small and scales well
- delay grows as policy becomes more complex

conclusion

coupling routes and policies is inherently flawed



this talk

- decouple policies from routes by a new policy system with SDN

benefits

- flexible policies
- automated coordination

future work

- implementation
- anonymize policies