State of the Art on BGP Economics

Costas Kalogiros – AUEB

8/1/2008

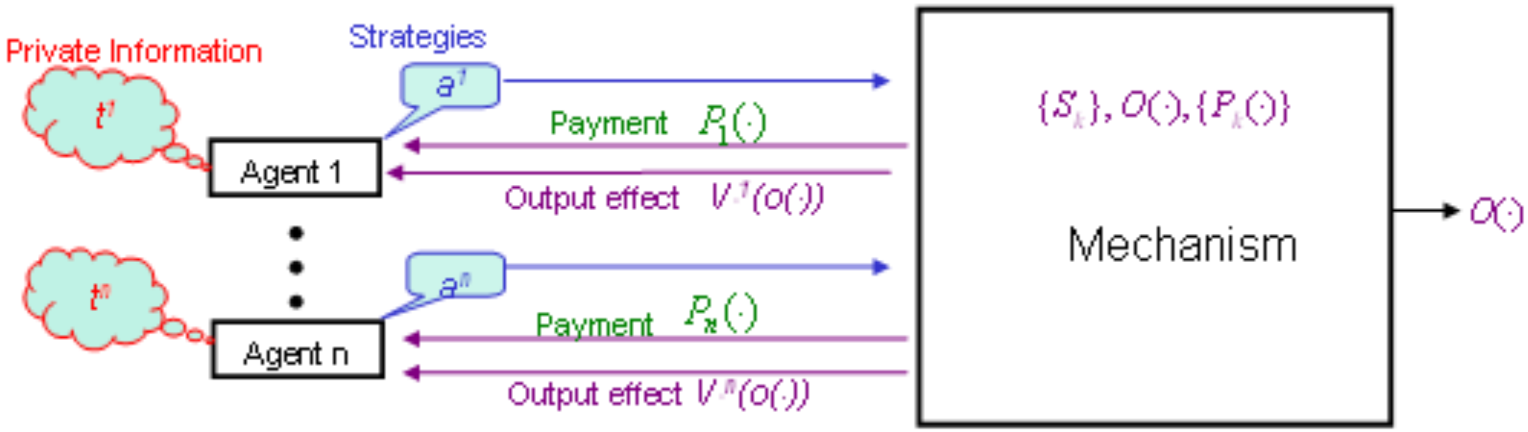
Outline

- Intro to Mechanism Design
- Incentive Compatibility & Convergence of BGP
- Mechanism Design for BGP
- Architectures for Cooperative routing

- You can find the report here:

Mechanism Design vs. Game Theory

- In a scenario of multiple selfish agents, each with private information
  - Game Theory: What are the possible equilibrium outcomes for a specific game?
  - Mechanism Design = "Inverse Game Theory": Given globally desired goals, what game should we design so that agents end up achieving those goals?
    - Game rules= Strategy set + output function + payments



Revelation principle

"Any game that does not elicit truthful declaration of private information can be substituted by a game that is *incentive compatible*".

Desired properties of Mechanism Design

a) Incentive Compatibility: Truthful revelation of a participant's private information, given that all other agents also tell the truth. In equilibrium, lying doesn't increase participant's benefit.
   – BGP: per packet transit cost, utility of a path, …

b) Individual Rationality: Agents will have positive Net Benefit from participating

Strategy-Proofness

- Strategy-proofness: Telling the truth regardless of what other agents declare.
  - Incentive Compatibility + Dominant Strategy Equilibrium

- Dominant strategy equilibrium: an agent keeps the same strategy regardless of other agents' strategies

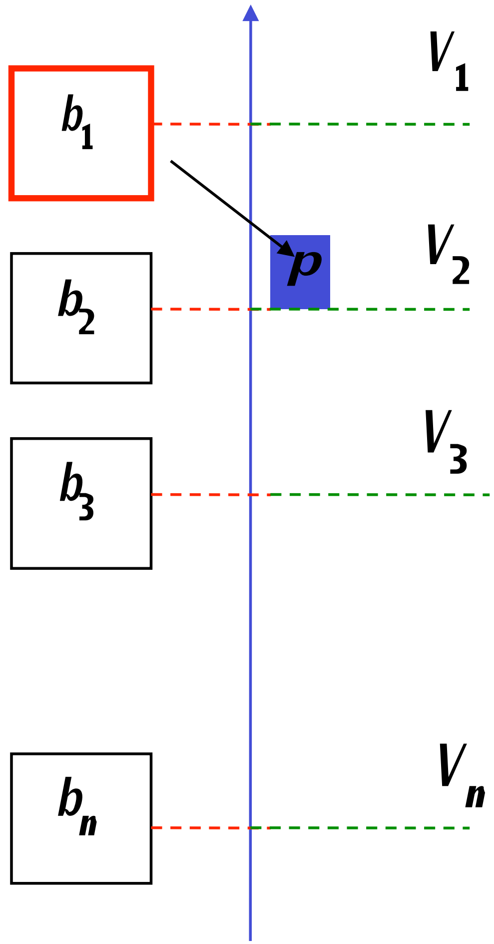
How Strategy-Proofness can be achieved?

- Use VCG class of payments: Payments towards/from an agent are based on her positive/negative effect to the rest agents, not on what she declares.
  - An agent's effect is computed by estimating what would happen if she was absent (all others being the same).
  - Of course, the outcome depends on what all agents declare.

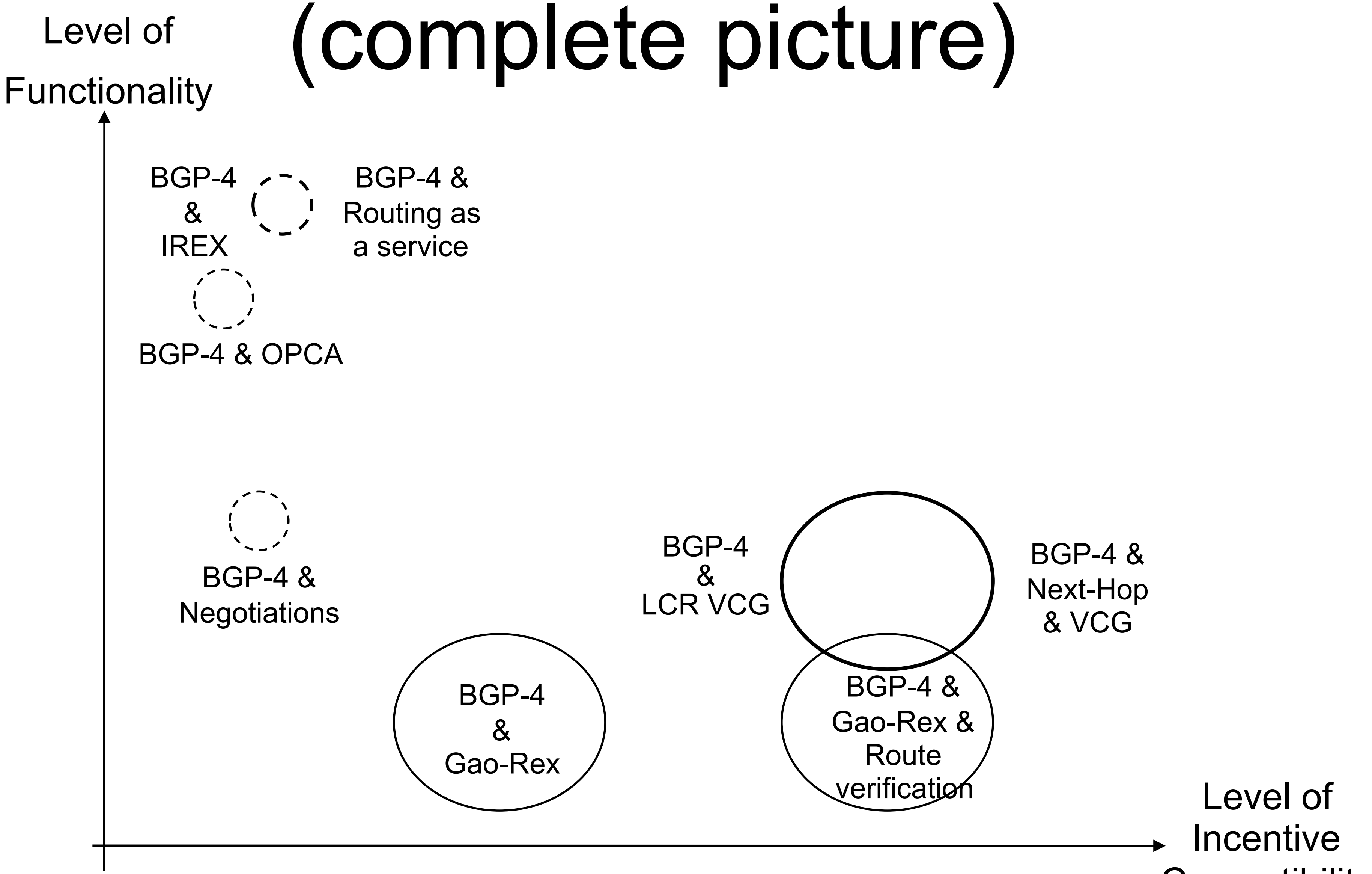
- But, payments to each agent can be arbitrarily larger than her cost
  - The mechanism executor must find a way to finance it.

Example: Vickrey (second-price sealed-bid) auction

- Sealed bids
- Item is awarded to the highest bidder
  - at a price equal to the second highest bid
- Dominant strategy: submit a bid equal to true valuation
  → incentive compatibility;
  guarantees efficiency

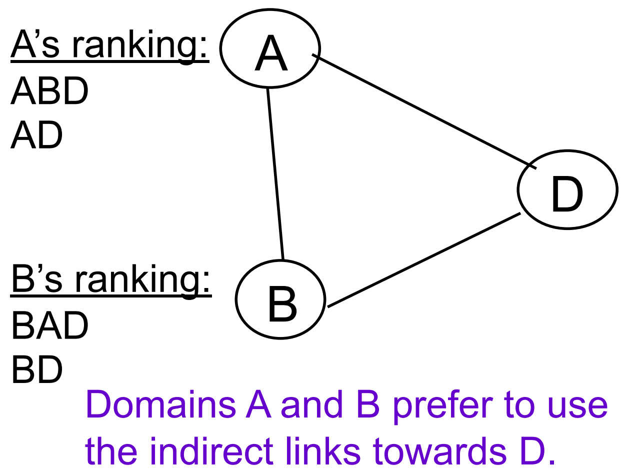


3D State-of-the-art map (complete picture)



BGP Convergence – outline & example

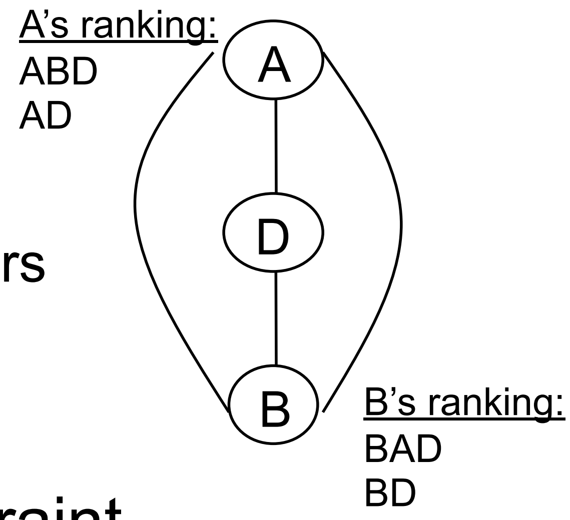
- **Convergence of BGP**: the state where no domain would change to a different route
- A group of AS'es can configure mutually incompatible policies → when does BGP converge?
- Stability may depend on the exact timing of reception and processing of the advertisements



A Stable Internet Routing without Global Coordination [L. Gao and J. Rexford]

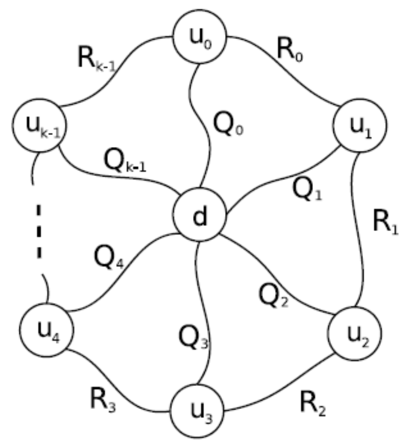
- Proved that BGP can converge given that ASes follow 3 “straight-forward” routing policies (Gao-Rex conditions)
  - No customer-provider cycles
  - Prefer customers to peers and peers to providers
  - Provide transit services only to customers

Simple example of a dispute wheel



- A special case of “no dispute wheel” constraint [Griffin, Sheperd, Wilfong]
  - A sufficient but not necessary condition for stability

The general case of a dispute wheel : $v_i(Q_i) \leq v_i(R_iQ_{i+1})$

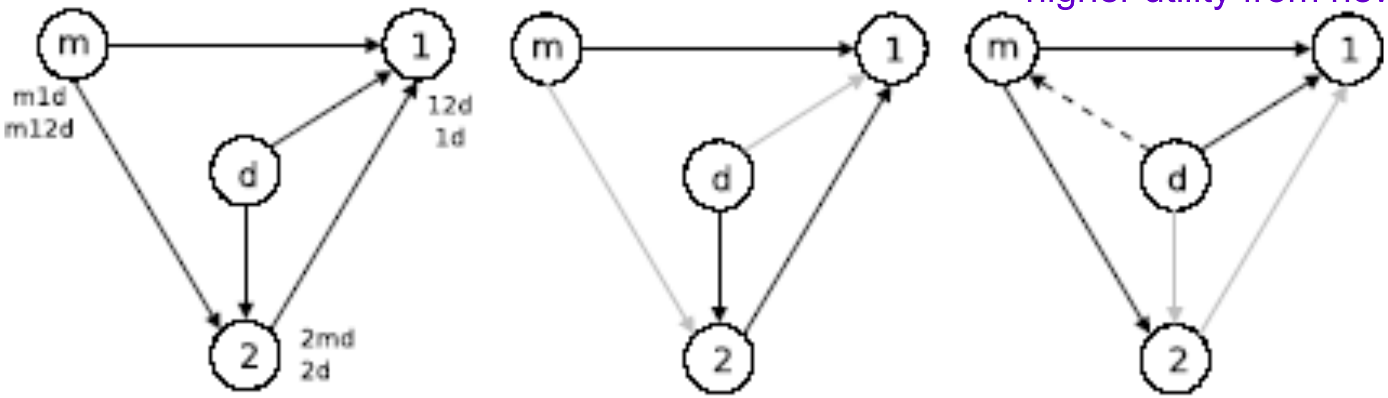


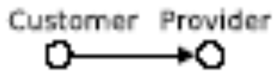
Is BGP incentive compatible?

- **Incentive compatibility of BGP**: Do ASes have the incentive to announce their true routes?
  - **The Strategic Justification for BGP** [Levin, Schapira, Zohar] (working paper)
- … No! However, BGP, with Gao-Rexford conditions, is immune to all forms of rational manipulation but one (lying about the availability of routes).
- BGP with *Route verification* is incentive-compatible in *ex-post Nash equilibrium*, if the routing policies do not induce a dispute wheel
  - *Route verification*: existence of route can be somehow proved

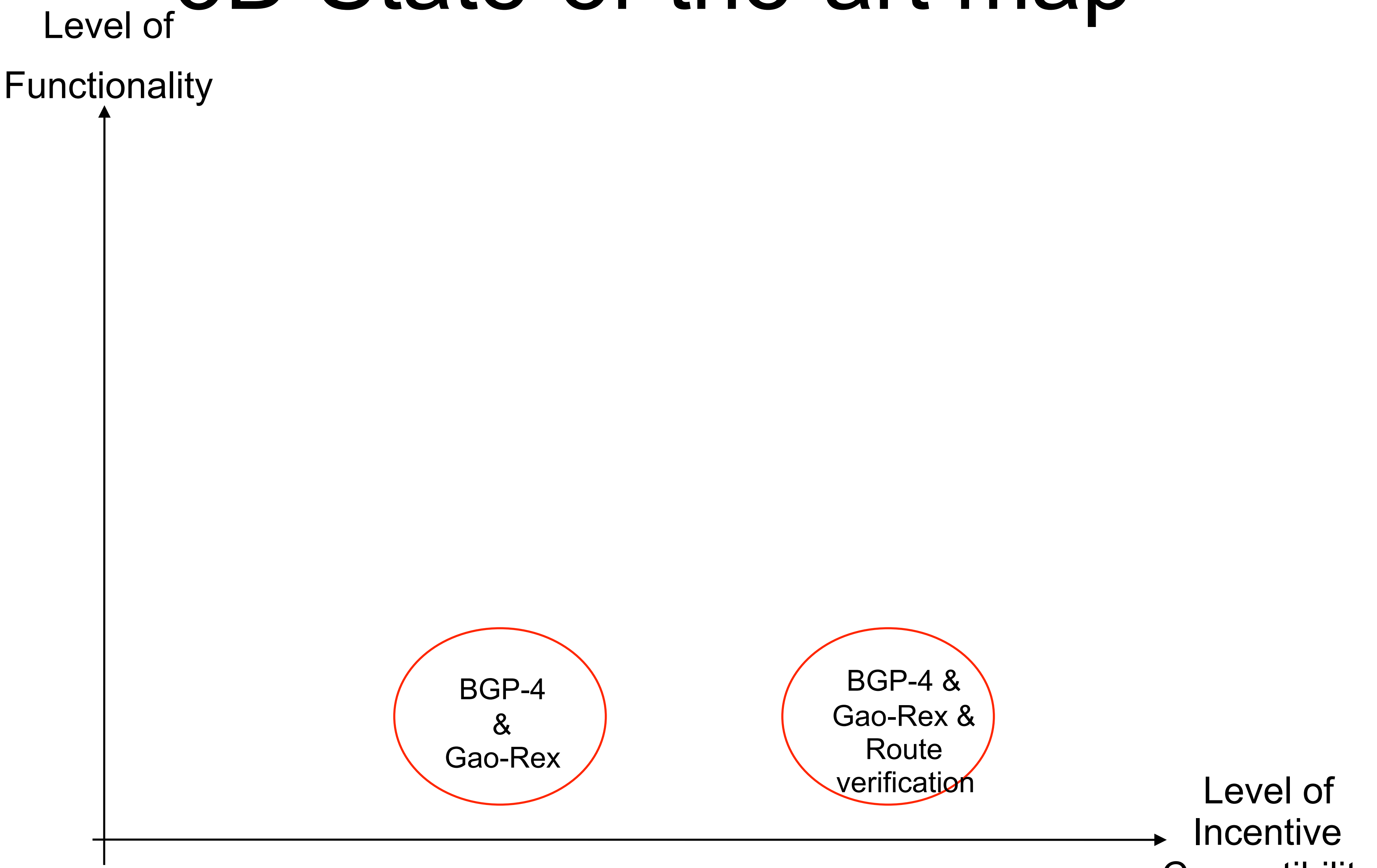
Stable routes conforming to Gao-Rex conditions

m lies about existence of md and achieves higher utility from new equilibrium outcome





3D State-of-the-art map



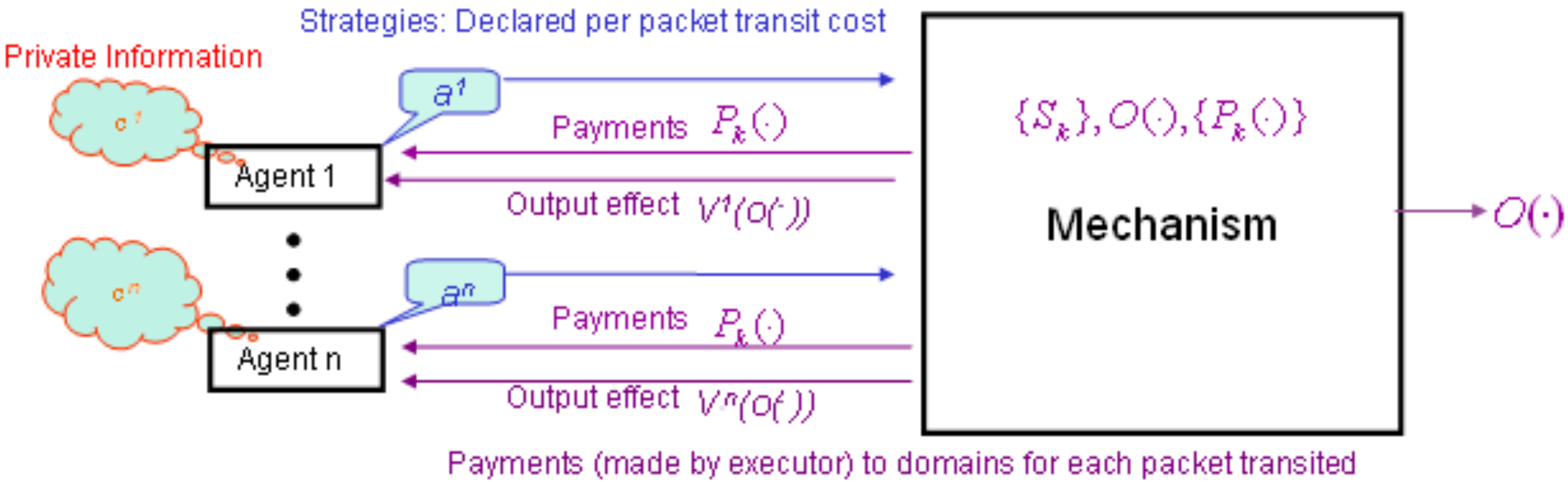
A BGP-based mechanism for lowest-cost routing

J. Feigenbaum, C. Papadimitriou, R. Sami, and S. Shenker

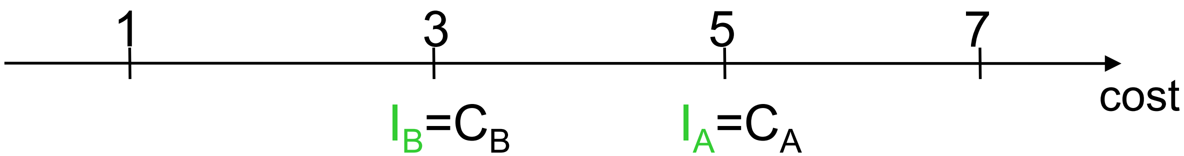
- A strategy-proof mechanism for truthful revelation of private information (cost per packet) by each AS, that outputs lowest-cost paths (LCP) and computes payments only to participating ASes on the LCP to a destination
  - Cost c_k is independent of i and j

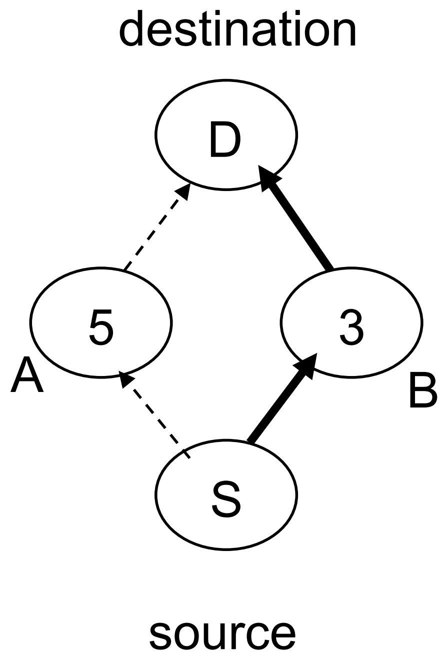
- Payment to agent K for path (i,j) according to VCG (note: graph is assumed to be biconnected)
  - If K on LCP path: $p_k^i(j)$ = (K-avoiding least cost path) – (current least cost path excluding K's cost)
  - Otherwise: $p_k^i(j)$ = 0

A BGP-based mechanism for lowest-cost routing



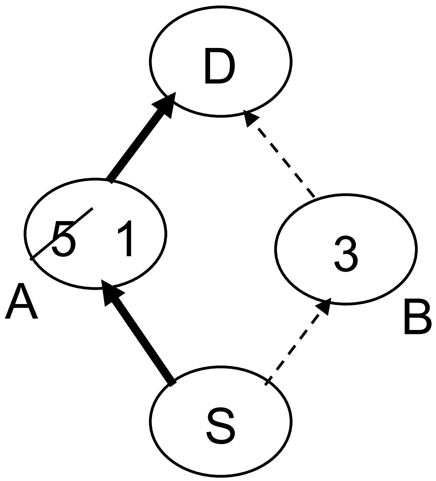
A BGP-based mechanism for lowest-cost routing: Examples & Incentive Compatibility (1/2)

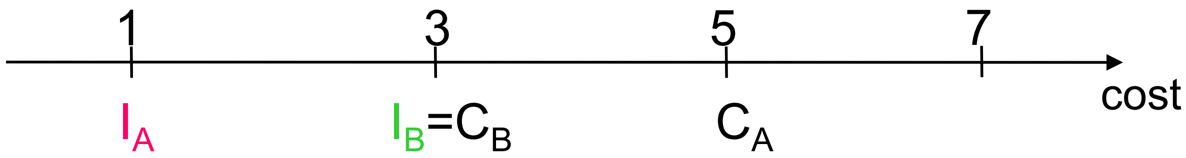




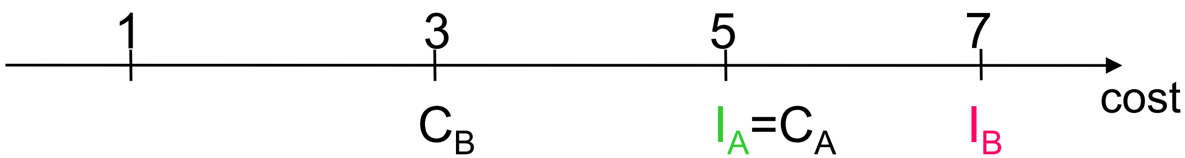
A and B report their true costs. B is on the LCP and she is paid:
P_B= LCP(C_B=∞) + LCP –I_B = 5 + 3 – 3 = 5

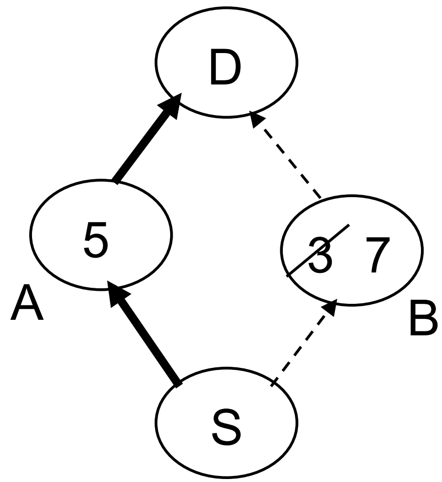
A BGP-based mechanism for lowest-cost routing: Examples & Incentive Compatibility (2/2)





A is understating her cost, while B is truth telling.
A is chosen as the LCP to D but she can't recover her cost!
$P_A = LCP(C_A=\infty) + LCP - I_A = 3 + 1 - 1 = 3 < C_A$



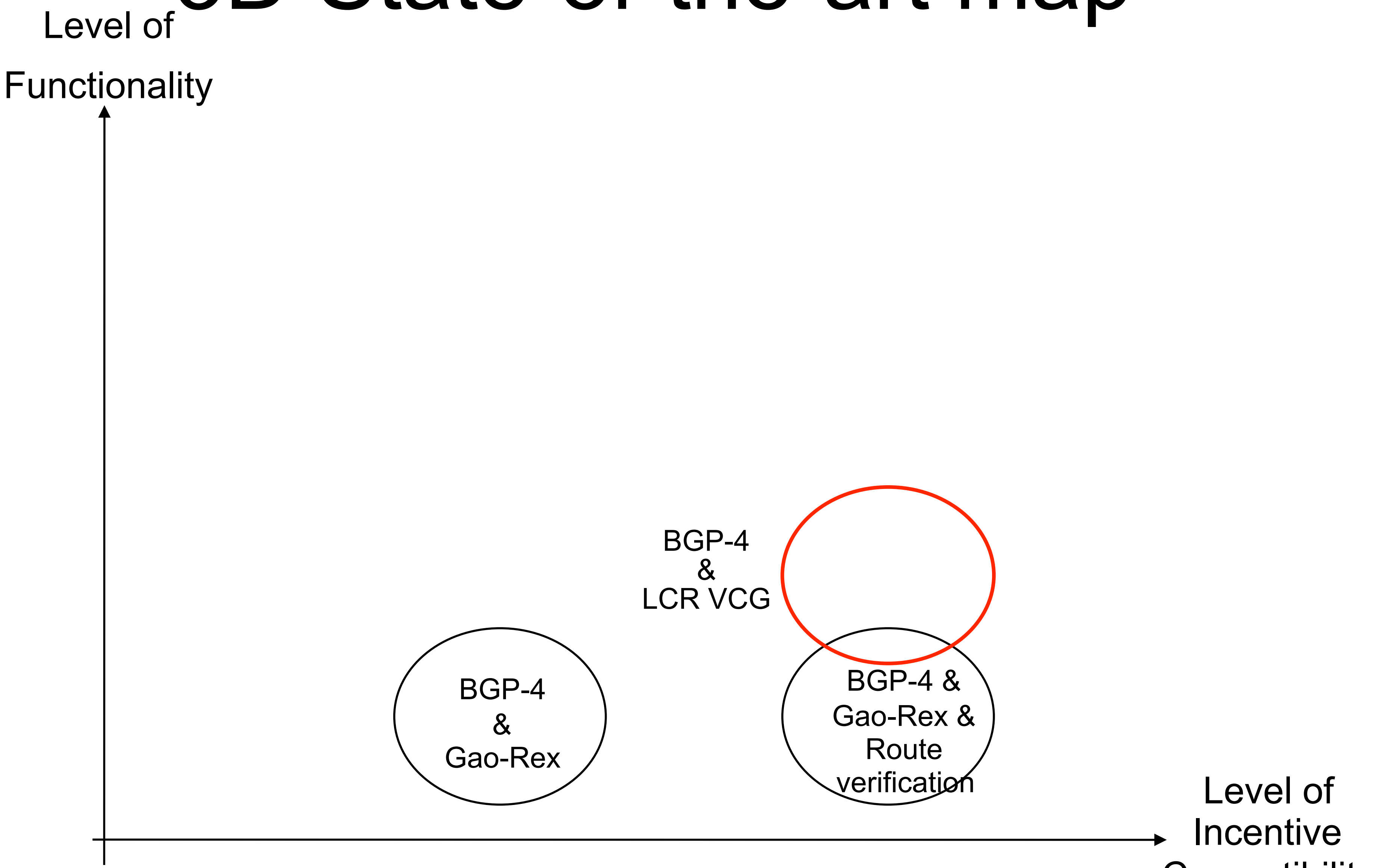


B is overstating her cost, while A is truth telling.
$P_B=0$ since B is not on the LCP to D!

A BGP-based mechanism for lowest-cost routing

- Issues that affect mechanism's applicability:
  - Mechanism must be financed by central planner
    - If central planner would try recover payments by charging end-users for their traffic, then users' demand could be lower than a perfectly competitive market
  - Restricts policies to least-cost paths
  - Collusion of some ASes can sustain higher prices by overstating costs
  - One stage game (but route advertising is a repeating process)

3D State-of-the-art map



Incentive-compatible interdomain routing

J. Feigenbaum, V. Ramachandran, and M. Schapira

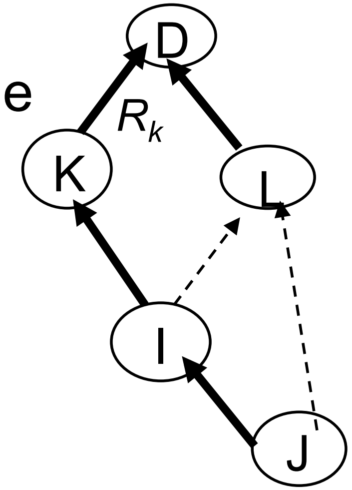
- Identifies three conditions that together form a sufficient constraint on policies to permit the computation of semantically-rich, stable and welfare-maximizing routes by any path-vector protocol.
  a. Dispute wheel freeness (sufficient condition for stability)
  b. Policy consistency (valuation based on the same and undeniable criterion, e.g. path length)
  c. Consistent filtering (no selective advertisements → advertise the best route to all legitimate neighors)

- "**next-hop**" policy: An AS decides among available routes to a destination based solely on the routes' next hop. The above 3 conditions are satisfied (a., c. by Gao-Rex conditions and b. by default)

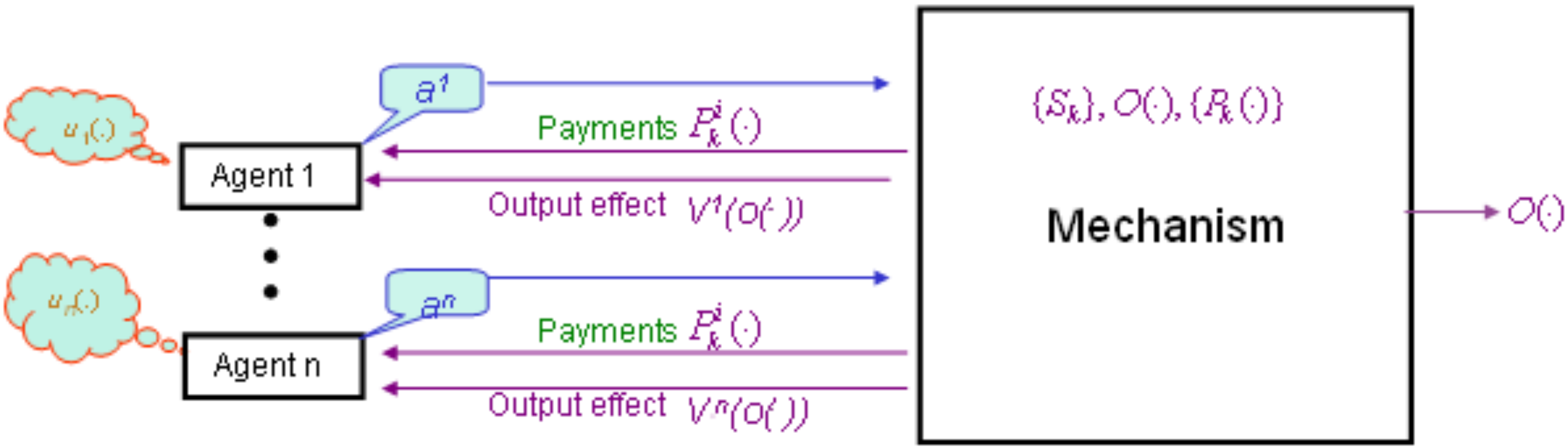
Incentive-compatible interdomain routing

- The payment to domain K on behalf of J depends only on the valuations at some neighbor *I* (not affected directly by K):
  - If *I* selects K as next-hop: $p_k^i(T_D) = u_i(T_D) - u_i(T_D^{-k})$;
  - Otherwise: $p_k^i(T_D) = 0$

  } VCG



- An agent's utility from the routing tree T_D to destination d is: $u_k(T_D) = u_k(R_k) + p_k(T_D)$



Incentive-compatible interdomain routing: Examples & Incentive Compatibility (1/2)

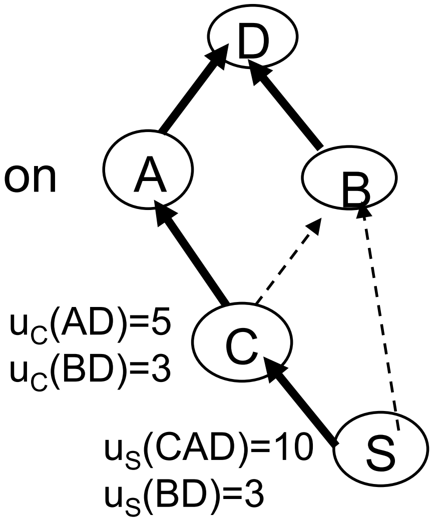
Agent's C utility with truthful revelation of route valuation (according to next-hop):

$$u_C(T_D) = u_C(R_C) + p_C(T_D)$$

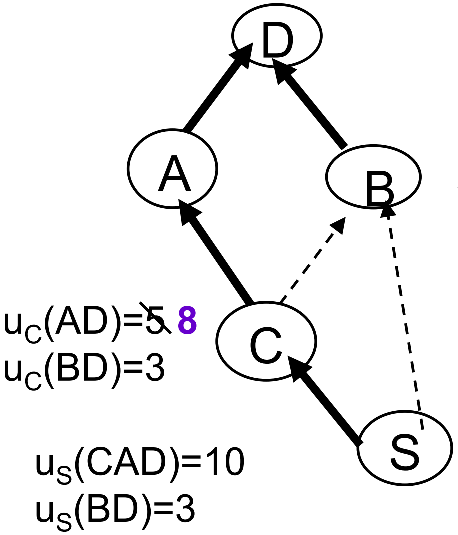
$$= u_C(R_C) + \overset{s}{p}_C(T_D)$$

$$= 5 + (10-3)$$

$$= 12$$

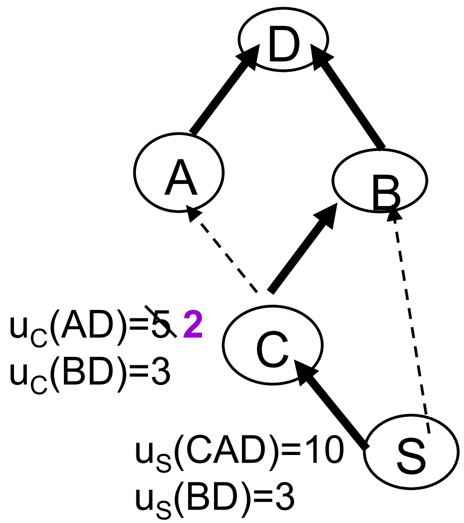


Incentive-compatible interdomain routing: Examples & Incentive Compatibility (2/2)



Agent's C utility with overstating the value of the best route (according to next-hop):

$u_C(T_D) = u_C(R_C) + p^{S}_C(T_D) = 8 + (10-3) = 15$
but, it's an artificial increase!



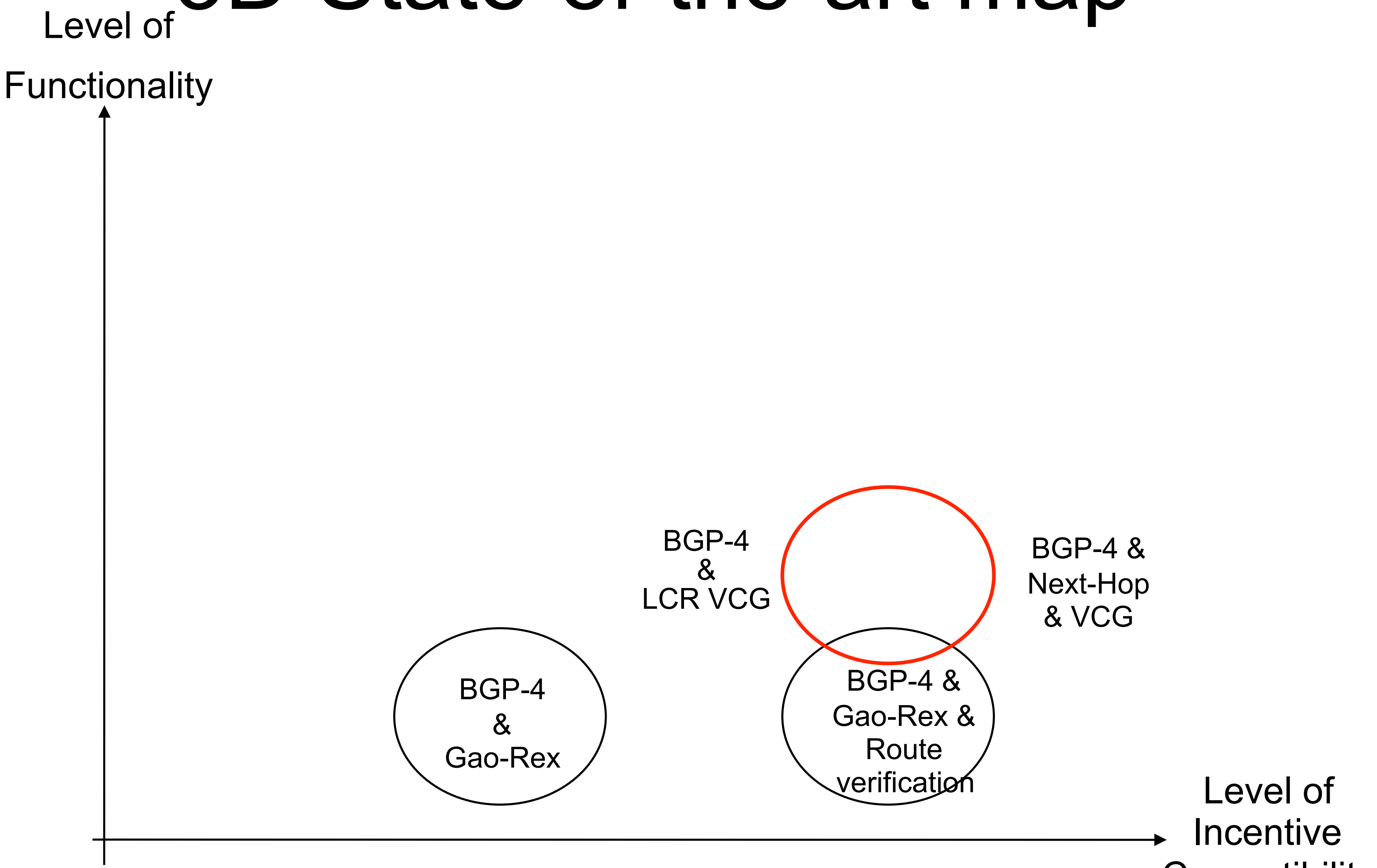
Agent's C utility with lying about the value of the best route (according to next-hop):

$u_C(T_D) = u_C(R_C) + p^{S}_C(T_D) = 3 + (10-3) = 10$
which is (artificially) lower than truthtelling!

Mechanism Design in BGP-outline

Agent = Domain	Domains' private information	Designer's target	Mechanism's Outcome	Incentive for truth telling
A BGP-based mechanism for lowest-cost routing. Joan Feigenbaum, Christos Papadimitriou, Rahul Sami, and Scott Shenker ,Principles of distributed computing, pages 173--182. ACM Press, 2002.	Intra domain per packet transit cost (same for all next-hop ASes)	Packets must be routed on the lowest cost path towards each destination (network efficiency)	•A path ∀ source destination AS pair (i,d) •Payments for each packet transited	$p_k^i(j)$ the mechanism payment on behalf of i, for a packet sent to j through k
Incentive-compatible interdomain routing Joan Feigenbaum, Vijay Ramachandran, and Michael Schapira, 7th Conference on Electronic Commerce, ACM Press, New York, 2006, pp. 130-139.	Domain's utility for each path it has learned	Routing tables that maximize utility of all domains	•a path ∀ source destination AS pair (i,d) •One-time payments	Domains do not pay for their preferences, but get paid according to others' declarations

3D State-of-the-art map



Cooperative routing - selected papers outline

- The Case for Separating Routing From Routers. N. Feamster, H. Balakrishnan, J. Rexford, A. Shaikh, and J. van der Merwe. In Proceedings of FDNA-04, 2004.
  - Routing tables should be managed by PCs; routers should forward packets by consulting the routing table

- Routing As a Service, K. Lakshminarayanan, I. Stoica, S. Shenker, 2004
  - A system architecture that allows third parties to propose routes based on users' preferences that the networking infrastructure will later use

Cooperative routing - selected papers outline

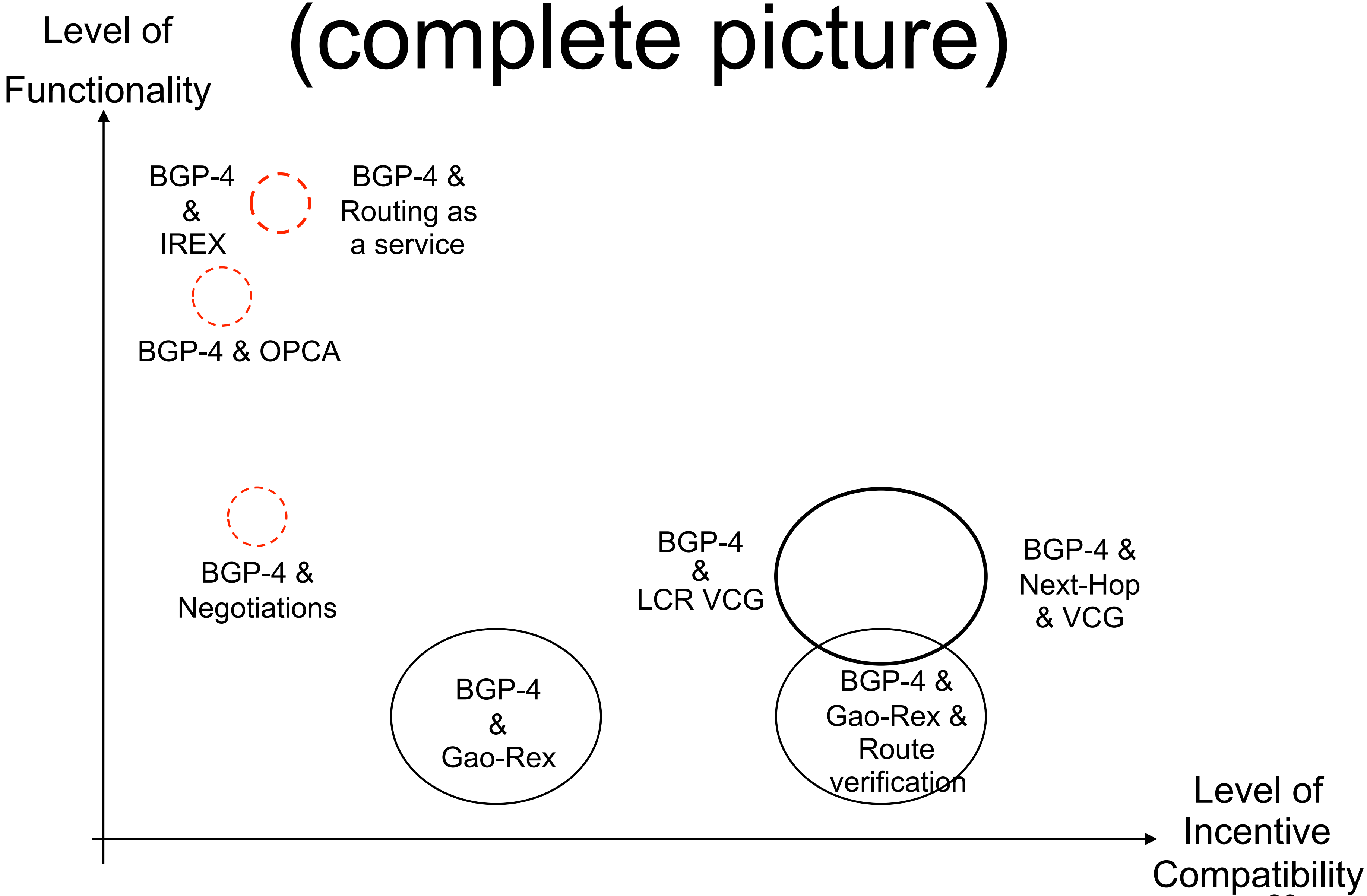
- Negotiation-based routing between neighboring ISPs, R. Mahajan, D. Wetherall, and T. Anderson, in Proc. USENIX Symposium on Networked Systems Design and Implementation, May 2005
  - Domains inform their neighbors about routing preferences in order to negotiate changes to current routes

- OPCA: Robust interdomain policy routing and traffic control. Sharad Agarwal, Chen-Nee Chuah, Randy H. Katz , In IEEE Openarch (New York, NY, April 2003).
  - A system where domains can ask providers on a problematic path to select another route

Cooperative routing - selected papers outline

- iREX: Inter-domain Resource Exchange Architecture, Ariffin Datuk Yahaya, Tobias Harks, Tatsuya Suda
  - A system where ASes advertise a selected path to destination based on price info and reputation. End-users contact each AS on the path to establish the path.

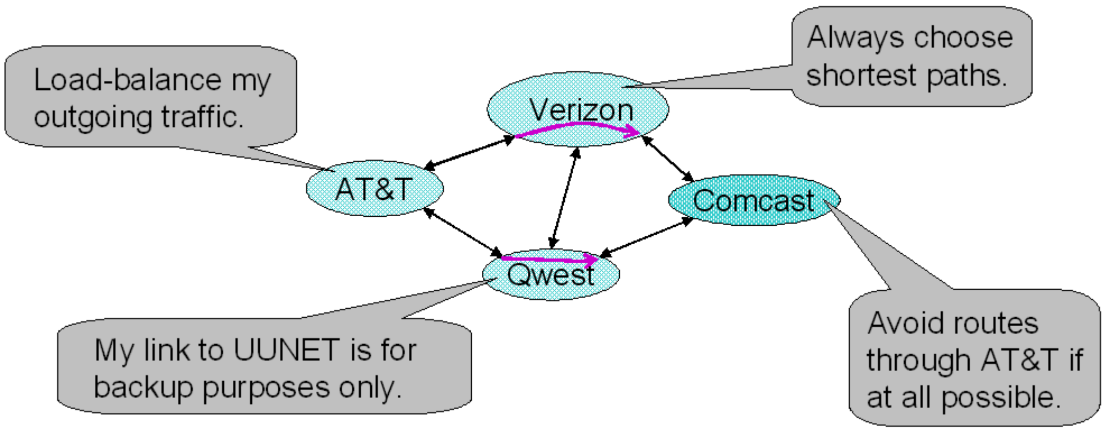
3D State-of-the-art map (complete picture)



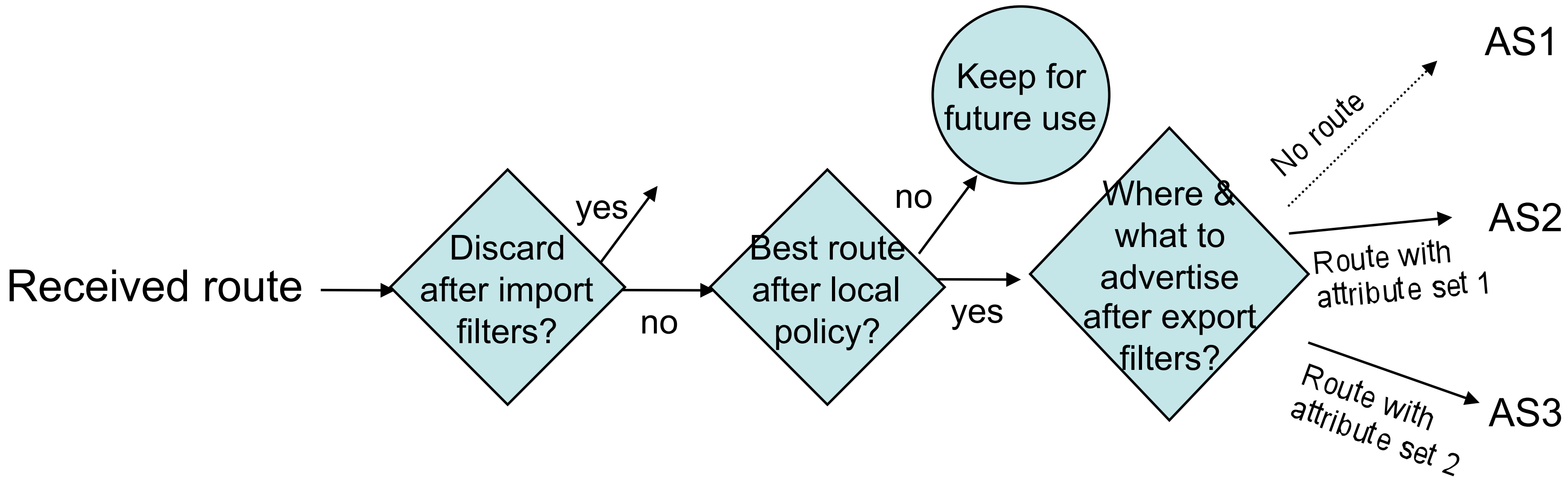
Backup slides

Inter-domain Routing with BGP

- Brings any-to-any reachability on the Internet
- Interesting property: BGP allows local routing policies but only one path to destination must be selected
- An AS has very limited influence on the routing policies of other domains

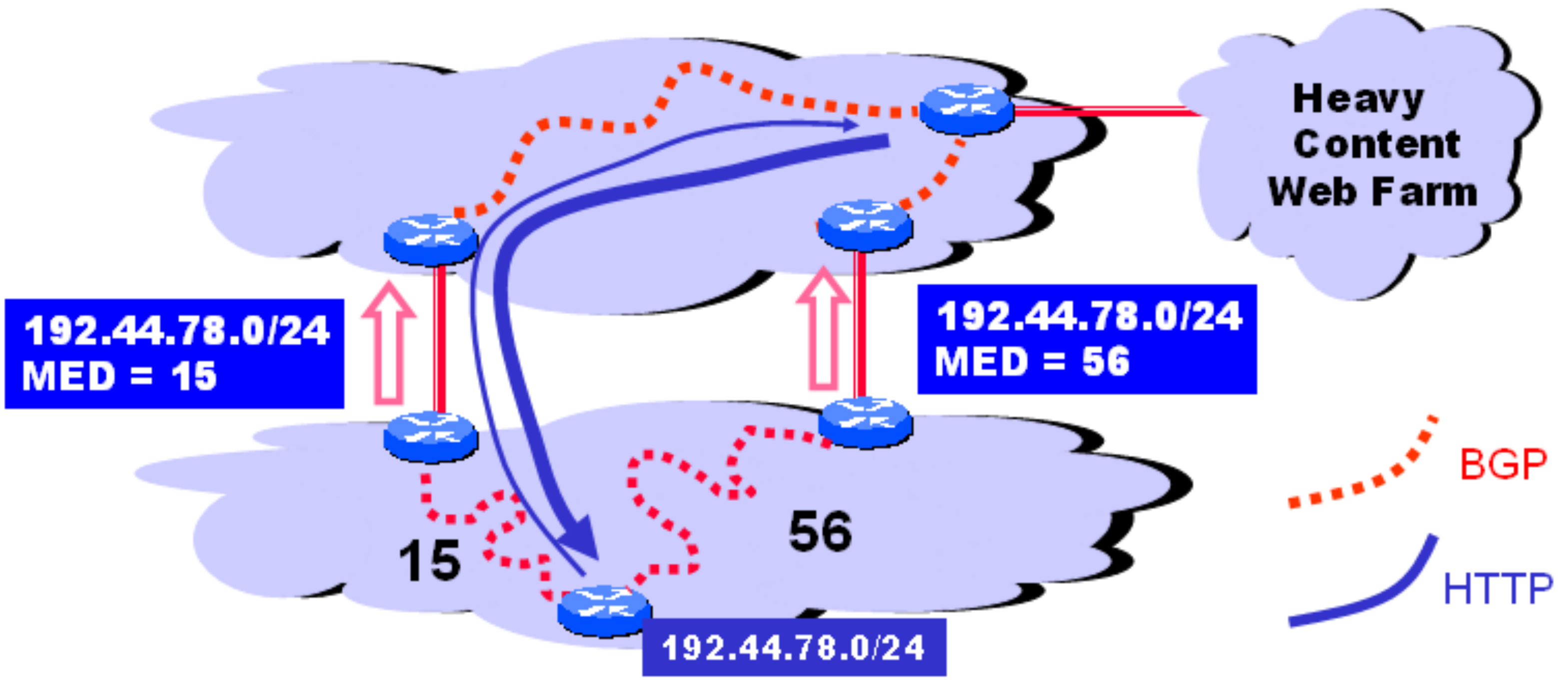


BGP Route Processing



- Each AS is supposed to advertise only one AS path for every set of destinations.
  - But some attributes can differ across recipients (even to the same neighbour, when interconnection happens at >=2 PoPs, e.g. MED)
- If an AS wants to have less inbound traffic it can adapt its export rules
  - However, he must always advertise her customers and her customer’s customers

MED: Influencing incoming traffic

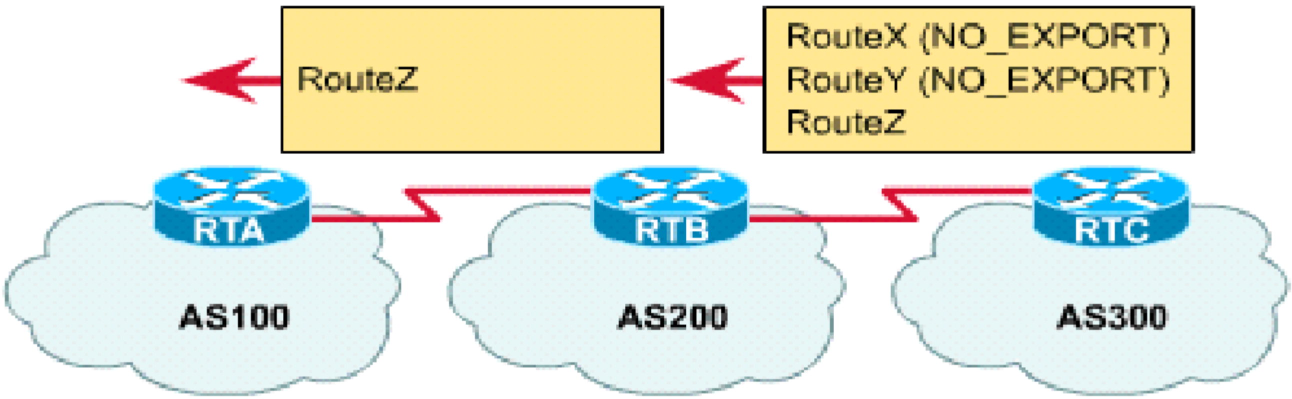


If multiple links exist between A & B,
MED determines which one B prefers that A uses (multi-exit discriminator)
Prefer lower MED values
This case is known as **Cold Potato Routing**
However some providers do not take into consideration MED attribute

MED values are comparable only between the same AS pair

Community: Influencing outgoing advertisements

- **Community attribute** is an optional transitive attribute
  - ASes can be organized into “communities”
- Well-known Community values include:
  - **NO_ADVERTISE** - A route carrying this community value, when received, should not be advertised to any BGP peer.
  - **Local-AS** - A route carrying this community value, when received, should be advertised to peers within the AS, but not advertised to peers in an external system.
  - **NO_EXPORT** - A route carrying this community value should not be advertised to peers outside a confederation
  - **Internet** - A route carrying this community value, when received, should be advertised to all other routers.

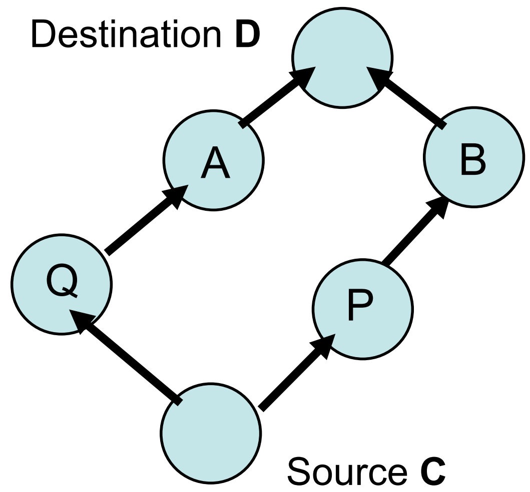


“Missing” functionality of BGP

Some thoughts…

Multi-homing

- With multi-homing, a single network has more than one connection to the Internet.
- Improves reliability and performance:
  - Can accommodate link failure
  - Bandwidth is sum of links to Internet
- Challenges
  - Multi-homing to different ISPs, while keeping the same address

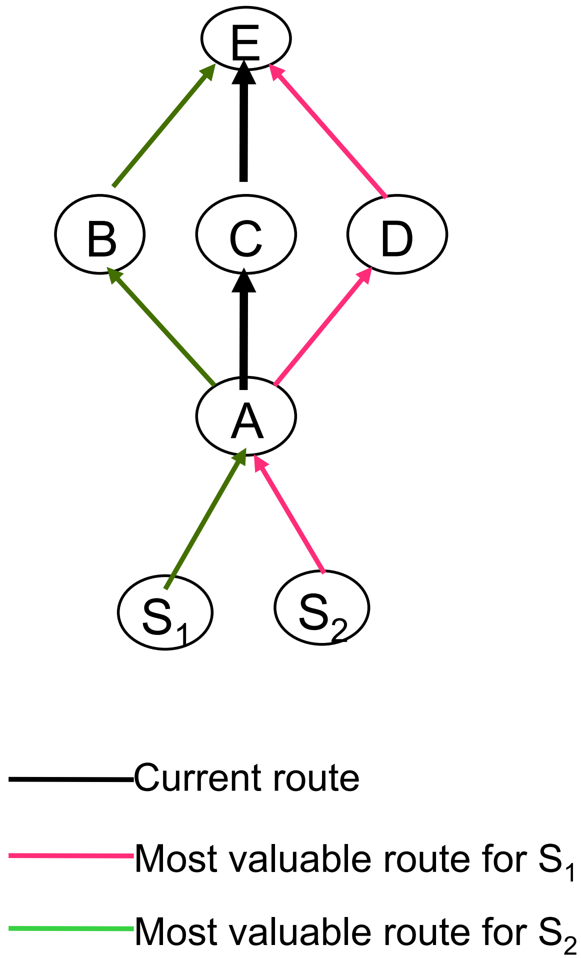


Influenced routing

- An AS "must" advertise to neighbors only the selected route towards each destination
  - e.g., AS1 cannot tell AS2 to route to other ASes in a manner different than what AS2 has chosen
- Neighbors (customer or providers) could somehow influence "best" route selection for incoming and outgoing traffic
  - Overlay networks is currently the way for "user-directed routing"
- How neighbors will learn about alternative routes and state their preferences?
- How does it affect convergence?
  - longer oscillations?
- How does it affect routing table size and router overhead?
  - if personal routing preferences are stored...

Influencing "best" route selection

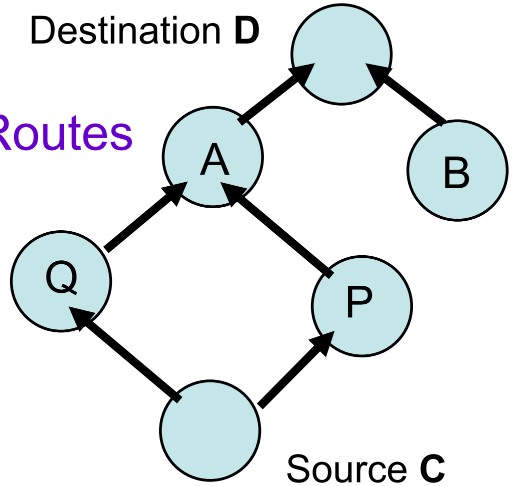
- A domain could be interested in influencing the route selected by previous domains on this path and propose changes (explicit paths or criteria used during evaluation of alternative routes).
- Issues:
  - Oscillations due to conflicting incentives for performing changes
    - If a mechanism for negotiation is absent (e.g. auctions, …)



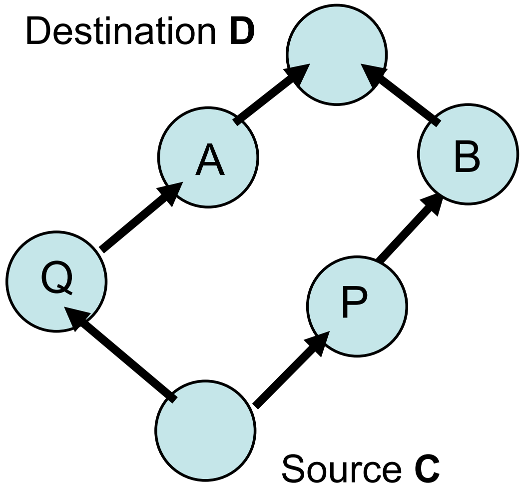
Asking to avoid an AS

- Source **C** wants redundant paths to destination and wants provider **P** to choose **B**'s route instead of **A**
- Issues:
  - How ASes will learn about alternatives routes and state their preferences?
  - Again, oscillations due to conflicting incentives for performing changes
    - If a mechanism for negotiation is absent (e.g. auctions, …)

Initial Routes



Routes after C's "complaint"



Solution Concepts

- The permitted *strategy set* affects the outcome
  - “solution concepts”
    - *Dominant strategy equilibrium*: an agent keeps the same strategy regardless of other agents' strategies
    - *Nash equilibrium*: given that all other agents keep their strategies, the agent has no incentive to change strategy
    - …

Recent 'extensions' to classic Mechanism Design

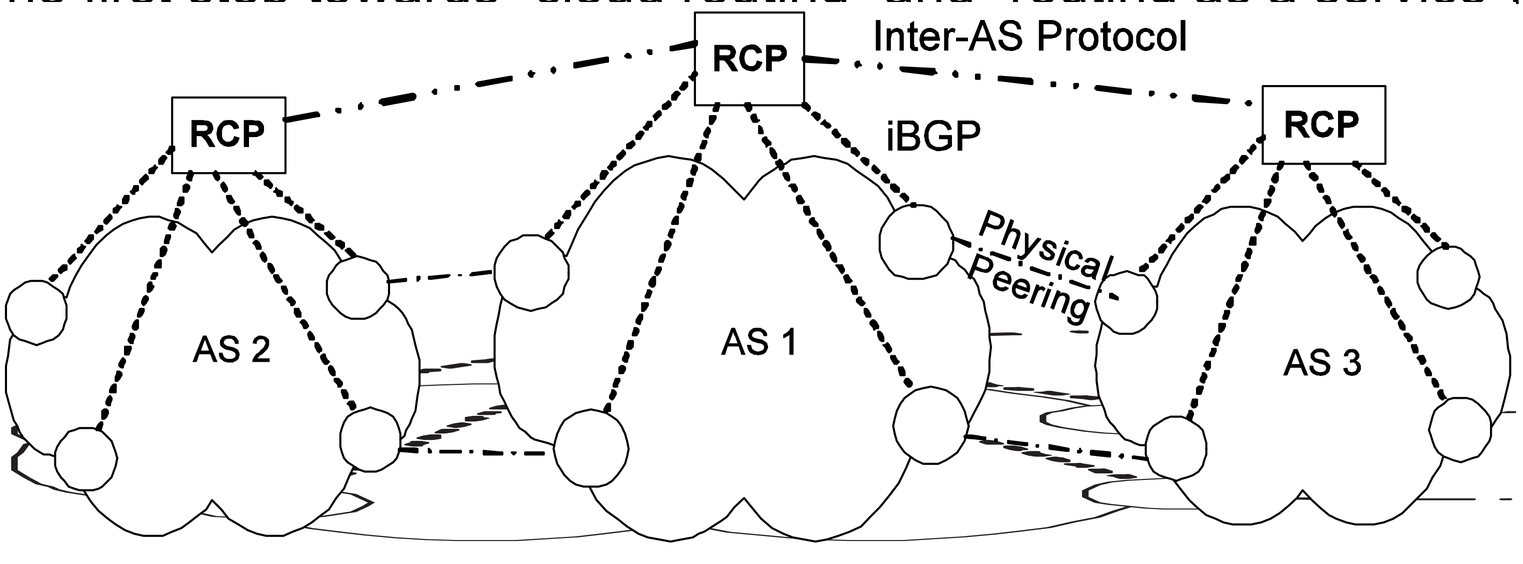
- In distributed Mechanism Design agents participate in the execution
  - Communication Compatibility
    - Truthful propagation of mechanism's messages.
  - Computation Compatibility
    - Truthful computation of mechanism's steps to reach an outcome (important for distributed mechanisms)
- If agents' Strategy set includes actions and computations Strategy-Proofness should include *Incentive Compatibility*, *Communication Compatibility* and *Computation Compatibility*
  - An intuition (not a theorem…)

Negotiation-based routing between neighboring ISPs

- ISP's make routing decision based on their own view of inter-network which leads to instability and inefficiency.
- Two-way information exchange model where neighboring ISPs provide their preferences on alternative routes and negotiate on changes.
  - Improving the path for one of the flows may hurt one of the ISPs but a set of improvements will cause a win - win situation
  - Initially, <u>for each alternative route</u>, ISPs disclose a rough measure (in the integral range [-P,P]) of the cost or benefit.
    - Opaque preferences provide a basis for negotiation between ISPs with different objectives.
    - Similar to MED…
  - Negotiation procedure (until they agree to stop):
    - Decide who will be the first (next) to propose an alternative
    - One who gets the turn proposes an alternative based on local and remote preferences
    - Other ISP decide whether to accept the proposal or not.

The Case for Separating Routing from Routers

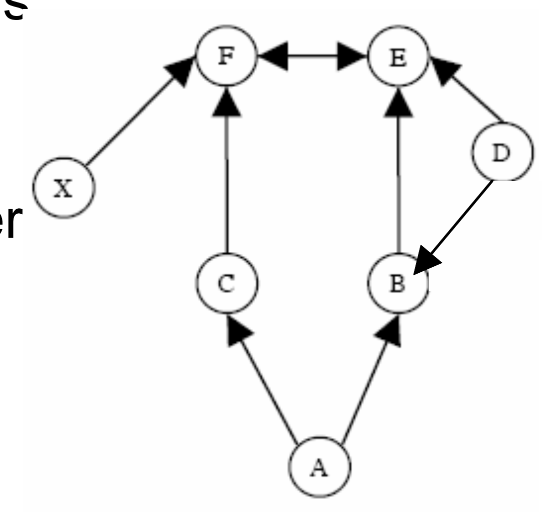
- Inter-domain routing protocol functionality should be separated from the routers
  - Routing is too important, but complicated also
- Routing Control Platform
  - Separates control plane logic from routers (data plane)
    - “Dumb”, but fast…
    - routing as a service (internal)
- Single entity (proxy) in each domain
- The first step towards “cloud routing” and “routing as a service”…



OPCA: Robust Inter-domain Policy Routing and Traffic Control

- Policy exchanges on top of BGP (overlay)
  - to support better management of incoming traffic across multiple incoming paths, and
  - to reduce the fail-over time of inter-domain paths
- A change on the routing between two distant ASes requires coordination
  - negotiations with the remote and, possibly, other intermediate policy agents
- Problems:
  - Requires knowledge of all proxies and overlay topology
    - Maybe difficult to infer (confidentiality reasons, …)

A→B indicates that
B is A's provider



X uses $X \rightarrow F \rightarrow C \rightarrow A$

$C \rightarrow A$ link fails

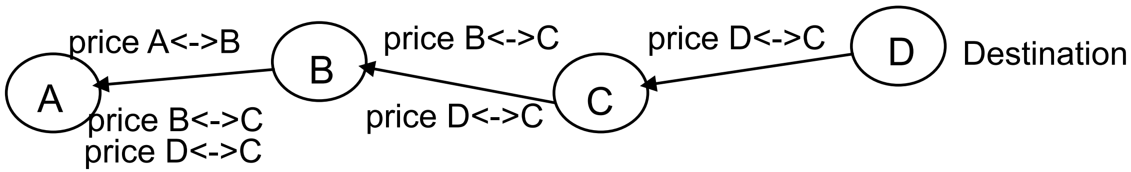
A wants X to failover to the $F \rightarrow E \rightarrow B \rightarrow A$

A's PA has to contact F's PA to make the routing change.

iREX: Inter-domain Resource Exchange Architecture

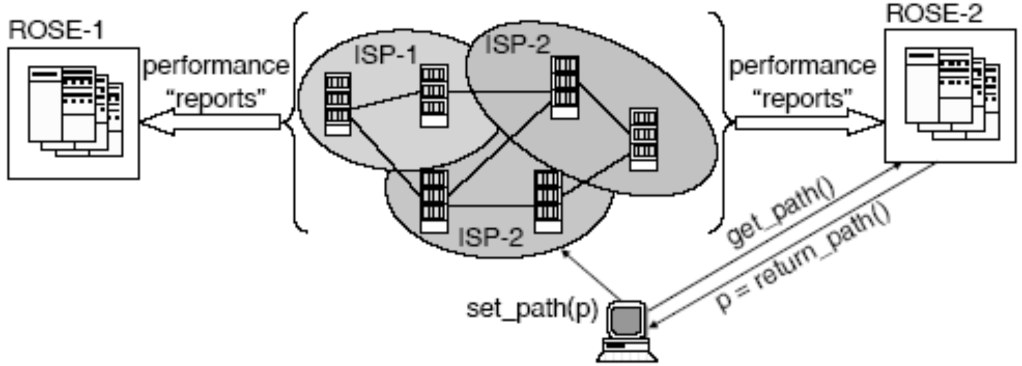
- Overlay architecture for automated deployment of flexible End-to-End (E2E) Inter-Domain (ID) Quality-of-Service (QoS) policy among resource user and resource provider ISPs
  - Economic-aware, distributed mechanisms (pricing and reputation mechanism)
  - Flavour of “source routing”, but packet routing is based on BGP
- Each *provider* ISP (seller) advertises the “real time” price per QoS level, per unit time (based on congestion) of the available inter-domain bandwidth
- Each neighbor (intermediate) advertises the cheapest reputable resources for the whole path
- Each *user* ISP (buyer) requests resources from each AS on the selected path
  - Each intermediate is able to propose a new path to buyer ('short circuit')
  - However, in general, intermediate ASes increase the complexity of the system
- Problems: central bank, fake paths (hide previous ASes), ...
  - In general: not incentive compatible

The IP address of each AS proxy is also propagated to A



Routing As a Service

- Goal: balance users' and ISPs' desire for traffic control, while allowing routing flexibility (instead of BGP-based) and less network overhead (instead of overlay)
- Forwarding infrastructure (ISPs)
  - Inter-domain overlay topology constructi
  - Routing functionality
    - Label-switching (for overlay traffic)
    - BGP-based (for non-overlay traffic)
  - Primitives for inserting routes
  - Provides information to *Route selectors*



- Route selector (Third parties)
  - Aggregates network information
  - Computes overlay paths on behalf of *End hosts*
  - Competes with other selectors for customers
- End host (Customers)
  - Queries route selector for "best" path
  - Setup the "best" path in the *Forwarding infrastructure*