A Route Deflection Approach to Minimize Routing Disruptions for Inter-AS Traffic Engineering

Kin-Hon Ho, George Pavlou, Stylianos Georgoulas and Mina Amin
Centre for Communication Systems Research,
University of Surrey, Guildford, Surrey, GU2 7XH, UK

Email: { K.Ho, G.Pavlou, S.Georgoulas, M.Amin }@surrey.ac.uk

Poster@IEEE INFOCOM 2006, Barcelona, Spain

BGP-based Traffic Engineering

- **Traffic Engineering (TE) using BGP**
  - To control traffic entering and exiting an Autonomous System (AS), through optimal ingress/egress point selection
  - Common TE objectives are, e.g., to satisfy inter-AS link capacity constraints, to achieve inter-AS load balancing, and/or to minimize peering cost, etc.
  - TE solutions are realized by adjusting BGP route attributes, e.g. local preferences and even IGP link weights (if necessary) for hot-potato routing
  - A recent survey has indicated an increasing usage of BGP-based TE
  - However, BGP-based TE may cause routing disruptions …

Routing Disruptions

- **Any transient or persistent perturbation of network performance caused by routing changes**
- **Several reasons causing routing disruptions**
  - Slow BGP convergence. During this time, service is disrupted
  - Routing table re-computation increases the computation load on the router processor, thus affecting packet forwarding performance
  - Re-advertising route updates to upstream ASes, which may subsequently change their best routes, possibly bypassing the advertising AS
  - If IGP link weight is changed in order to achieve TE objectives, the routing of many traffic flows will be affected within the network

Motivation

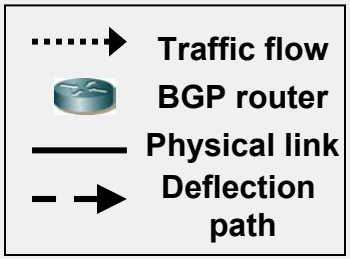
- **Looking for stable TE**
  - It is desirable not only to achieve TE objectives, but also to minimize routing disruptions
  - Therefore, it is important to maintain stable BGP routing while providing alternate routes to forward traffic around the congestion points
  - We propose using inter-AS deflection routing

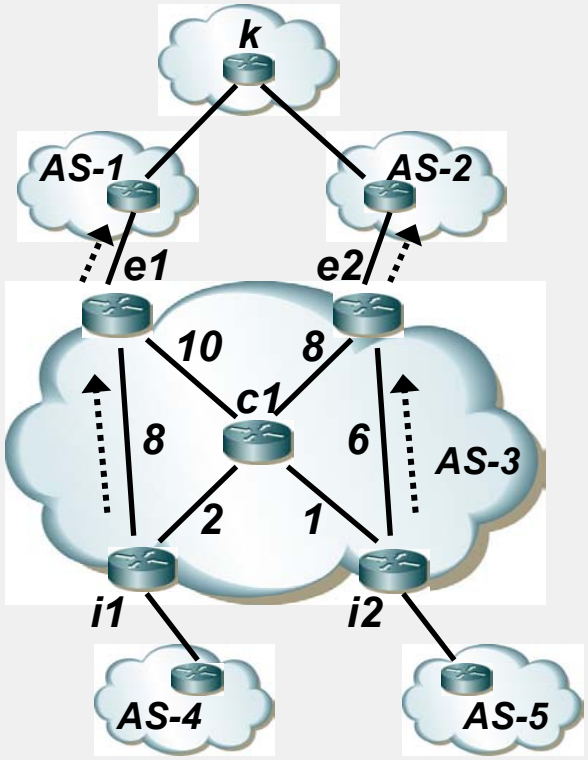
Inter-AS Deflection Routing (1/4)

- **Overview of inter-AS deflection routing**
  - When overloading is detected over an inter-AS link, the incident border (egress) router makes a local traffic forwarding decision to divert some traffic from the overloaded link to another underloaded inter-AS link
  - Since the traffic diversion is done locally, the incident egress router does not need to initiate any new route update because no BGP route attribute will be re-configured

Inter-AS Deflection Routing (2/4)

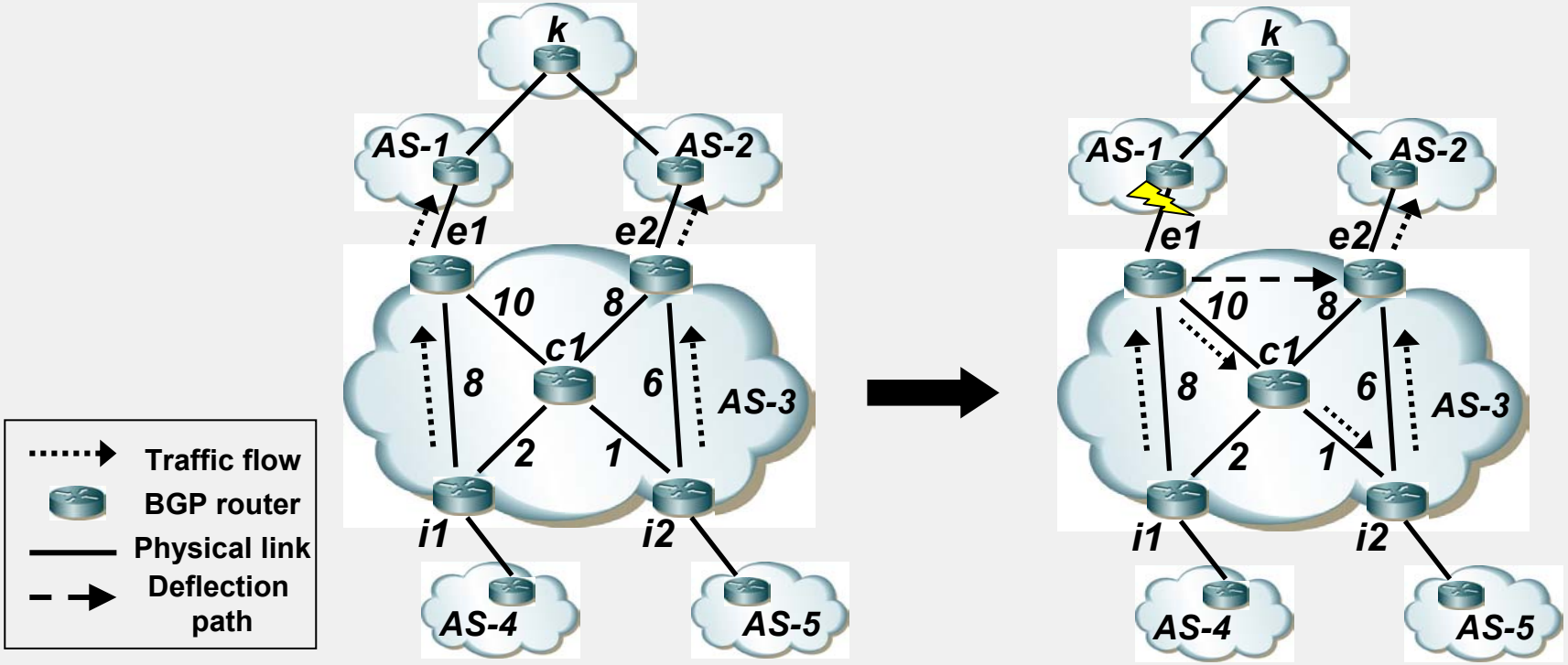
- **Example**





- Two BGP paths with equal route attributes to prefix *k* through *e1* and *e2*
- The number on each link represents IGP cost
- According to hot-potato routing, *e1* and *e2* carry the traffic for *k* from *i1* and *i2* respectively

Inter-AS Deflection Routing (3/4)



- When overloading is detected over the link between *e1* and *AS-1*, *e1* diverts the traffic for *k* to *e2* whose route towards *k* has been selected by *e1* as the alternate BGP route
- *e1* does not re-configure any BGP route attribute because the traffic is diverted locally
- *e1* & *e2* are *deflection* and *relay* routers respectively

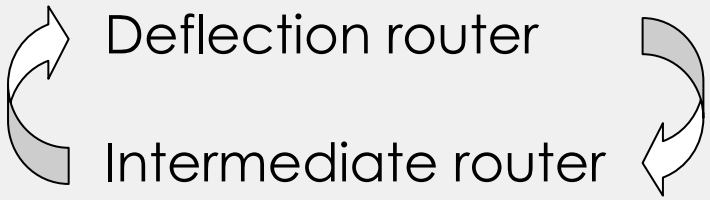
Inter-AS Deflection Routing (4/4)

- **Two possible ways of implementation**
  - Modify the next-hop or outgoing interface in the BGP Forwarding Information Base (FIB) table
    - According to hot-potato routing, *c1* selects as the best route to *k* the one through *e2*
    - *e1* could therefore divert the traffic to *c1*, which then forwards the traffic to *e2* towards *k*
  - Use explicit route to forward the traffic from the deflection router to the relay router
    - E.g. establish a MPLS LSP from *e1* to *e2* as the deflection path. This LSP may not be the shortest path

- **There are several important issues should be carefully addressed when using inter-AS deflection routing …**

Issues to be addressed (1/3)

- **Routing loop avoidance**
  - Loop may be formed if some intermediate routers on the IGP path to the relay router have the best route to the prefix through the deflection router



  - Two potential solutions may be considered
    - *Next-hop routing*: carefully select the alternate BGP route so that the deflection router has at least one adjacent router selected for which the alternate BGP route is the best router to the prefix
    - *Explicit routing*: establish an explicit route from the deflection router to the relay router

Issues to be addressed (2/3)

- **Sub-optimal intra-AS resource utilization**
  - The whole path for the deflection routing may be longer than the usual shortest path from the ingress router to the relay router, thus consuming extra network resources
  - The resource utilization depends on
    - Location of the relay router: the closer the deflection router to the relay router, the shorter the deflection path
    - Selection of the deflection path: the path does not have to be the shortest to the relay router but could be a long one for traffic engineering purposes
  - An intelligent algorithm can determine the optimal assignment of relay router and deflection path

Issues to be addressed (3/3)

- **Granularity of deflection routing**
  - Where should the deflection routing be performed?
    - At ingress, core or egress routers?
    - Different network resource utilization may be achieved as different deflection paths are used

~ The End ~

Thank you!