NIRA: A NEW INTER-DOMAIN ROUTING ARCHITECTURE

Xiaowei Yang, David Clark and Arthur Berger

Presented by Sachin Kadloor
BGP ROUTING

• Link weights represent costs

• Default routes picked by BGP

  • A→B, A→B→C, A→B→C→D

• Need to tune attributes in order to use A→C and B→D links

• Manual tuning can lead to instabilities
OUTLINE OF THE TALK

• Routing
  • challenges and existing solutions (OSPF-TE, TeXCP)

• NIRA: New Internet Routing Architecture
  • design philosophy
  • a new addressing scheme
  • route discovery
  • forwarding
TRAFFIC ENGINEERING (TE)

• Mathematical framework
  • Given: directed graph $G = (V,E)$
  • vector $C = [C_e], e \in E$, of edge capacities
  • demand matrix $D = [D(s,t)], s,t \in V$
  • design $L = [l_e], l_e$ load along edge $e$, to satisfy demand
  • Optimization: minimize max. utilization ($l_e/C_e$)
OPEN SPF-TE (OSPF-TE)

- Offline (static) TE
- Uses long term average data to tune weights
- Avoids the risk of instability caused by real-time fluctuations
- Pre-compute the best re-routes in case of failures
- Chooses the best re-route which works for most failures
- Can lead to over-provisioning of network
TRAFFIC ENGINEERING WITH XCP (TEXCP)

- Online (dynamic) TE
- Recall: XCP gives explicit feedback
  - Efficiency and fairness controller
- TeXCP: routers give explicit feedback on utilization
  - load balancer
  - per-path XCP controller
NIRA

Let users decide how they want to route
NIRA: DESIGN PHILOSOPHY

• Users pick ISPs, but cannot control their packets’ routes
  - currently, each domain makes a local decision on what the next hop will be (according to BGP protocol)

• Alternative: users choose a path with better QoS
  - Stimulate competition among providers
  - Different applications have different needs
    - e.g. gaming, peer-to-peer downloads, support multi-path routing
CHALLENGES

• How does a user discover routes?
• How do you represent routes?
• How should the providers be compensated?
NIRA - OVERVIEW

• Organize the routers into hierarchy
• Use hierarchical addressing
• Nodes in tier-1 get allocated a chunk of address space
  • They allocate a fraction of it to their children
  • Children get an address from each parent
    • B→b, C→c, A→b.1 and c.1, B→b.2 and c.2
  • A source and destination address uniquely specifies the route
• with hierarchical addressing, an address represents a route to the core
ROUTE DISCOVERY AND REPRESENTATION

• TIPP (Topology Information Propagation Protocol) used to discover routes

  • information propagated only along hierarchy (not globally)

• With the new addressing scheme, a source address and a destination address efficiently represents a route

  • switch routes by switching address

• Note: route specified by source and destination address
NAME-TO-ROUTE MAPPING

• Bootstrapping a communication is directory lookup based

• Each user registers his name, addresses and preferences in a NRLS (Network to Route Lookup Server)

• Needs to be updated when domain level topology changes
PACKET FORWARDING

• Each router maintains 3 or 4 forwarding tables
  • uphill table: domain’s providers’ addresses
  • downhill table: domain’s address and its customers’ address
  • bridge table: domain’s neighbors with whom it has peering relationship
  • BGP table: if the router also participates in BGP routing
FORWARDING ALGORITHM

• A router uses longest prefix match to look up the destination address in its downhill table
  
  • if match is found, destination address used to route the packet

• If no match is found, use the uphill table, and the source address to forward the packet towards the core
  
  • if already in the core, or if a peering link supports forwarding, use bridge table to forward the packet
CONCLUSION

• NIRA allows users to pick routes
  • uphill routing based on source address, and downhill routing based on destination address

• Other benefits of NIRA
  • scalable: TIPP information propagates only along hierarchy
  • memory requirement of BGP scales linearly in IP prefixes advertised
  • security: limits source spoofing
THANK YOU.
ROUTEING

• Design goals:
  • optimality
  • low overhead
  • robustness
  • fast convergence
  • flexibility

• NIRA: route selection by user
ROUTE DISCOVERY

• TIPP (Topology Information Propagation Protocol)
  • a path-vector component to inform a user of the domain-level routes
  • a policy based link state component to inform a user of the dynamic network conditions
ROUTING

Inter-Domain routing (BGP)
NIRA
Intra-Domain routing (IGP)
RIP, OSPF, TeXCP
TIPP

- Link-state messages only propagated downward the hierarchy
  - ensures scalability

- Topology updated based on message heard from neighbors

- Inconsistencies resolved by believing the neighbor that is on the shortest failure-free path to the link that triggered update
DISTANCE-VECTOR ROUTING

• Distance Vector

  • each node knows the distance (cost) from itself to its neighbors

  • each node advertises a vector containing distances from itself to all other nodes in the network (initialize unknown distance with infinity)

• Uses Dijkstra’s algorithm to find shortest path

• Routing Information Protocol (RIP)
LINK-STATE ROUTING

• Each node learns the complete topology of the network

• Creates shortest paths to all destinations with itself as a root node

• Requires more overhead than Distance-Vector routing, but more robust and scalable

• Open Shortest Path First (OSPF)
TRAFFIC ENGINEERING WITH XCP (TEXCP)

- Problem: routing within domain
  - load balancing? reaction to failures?
- Solution: make routing adaptive
- Like XCP: Introduce load balancer and a feedback controller
- Formulate the routing problem as a min-max utilization optimization problem