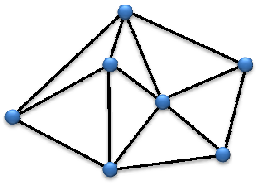
Floodless in SEATTLE:
A Scalable Ethernet Architecture for Large Enterprises

Changhoon Kim, Matthew Caesar,
and Jennifer Rexford

Ethernet in Enterprise Nets?

- Ethernet has substantial benefits
  - Simplifies network management, greatly reducing operational expense
  - Naturally supports host mobility
  - Enhances network flexibility
- Why do we still use IP routing inside a single network?

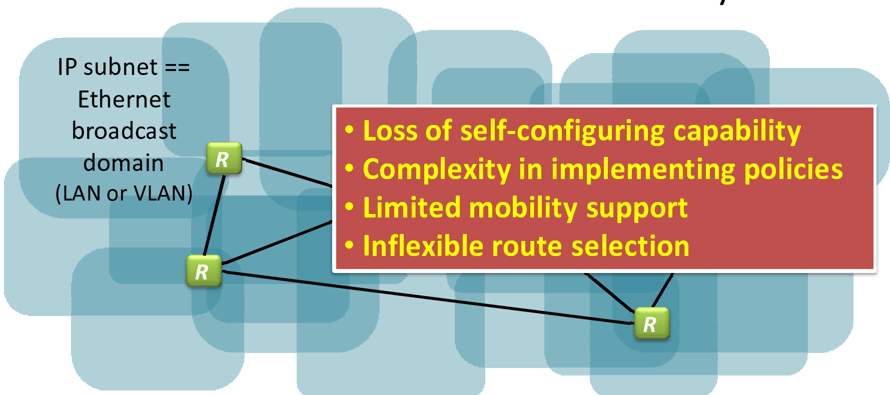
Ethernet Doesn't Scale!



- Reasons for poor scalability
  - Network-wide flooding
  - Frequent broadcasting
  - Unbalanced link utilization, low availability and throughput due to tree-based forwarding
- Limitations quickly growing with network size
- Scalability requirement is growing very fast
  - 50K ~ 1M hosts

Current Practice

A hybrid architecture comprised of several small Ethernet-based IP subnets interconnected by routers



Sacrifices Ethernet's simplicity and IP's efficiency only for scalability

Key Question and Contribution

- Can we maintain the same properties as Ethernet, yet **scales** to large networks?
- SEATTLE: The best of IP and Ethernet
  - Two orders of magnitude more scalable than Ethernet
  - Broadcast domains in **any size**
  - Vastly simpler network management, with host mobility and network flexibility
  - Shortest path forwarding

5

Objectives and Solutions

Objective	Approach	Solution
1. Avoiding flooding	Never broadcast unicast traffic	**Network-layer one-hop DHT**
2. Restraining broadcasting	Bootstrap hosts via unicast	
3. Reducing routing state	Populate host info only when and where it is needed	**Traffic-driven resolution with caching**
4. Shortest-path forwarding	Allow switches to learn topology	**L2 link-state routing maintaining only switch-level topology**

*** Meanwhile, avoid modifying end hosts**

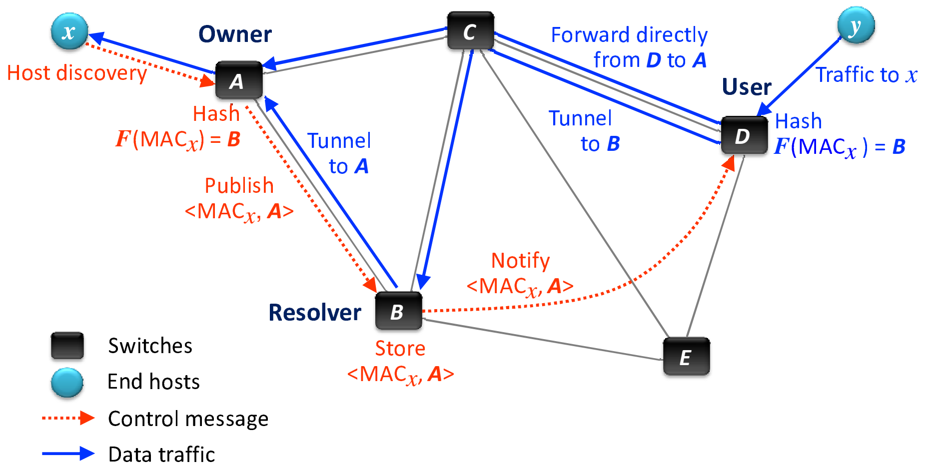
6

Network-layer One-hop DHT

- Switches maintain <*key*, *value*> pairs by commonly using a hash function *F*
  - *F*: Consistent hash mapping a key to a switch
  - *F* is defined over the live set of switches
  - LS routing ensures each switch knows about all the other live switches, enabling one-hop DHT operations
- Benefits
  - Fast and efficient reaction to changes
  - Reliability and capacity naturally growing with network size

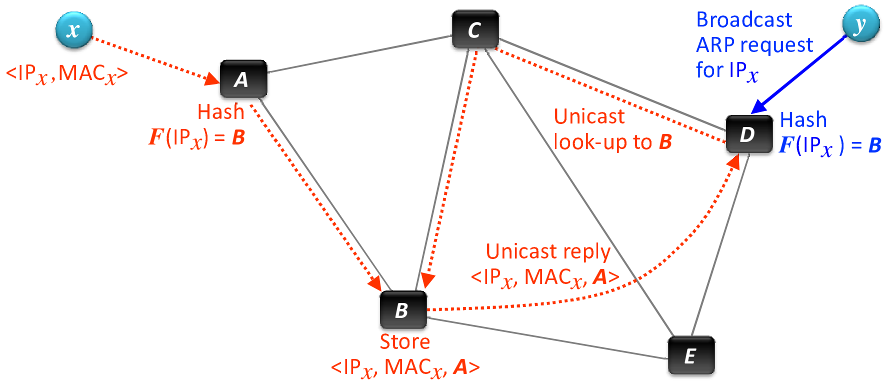
Location Resolution

<key, val> = <MAC addr, location>



Address Resolution

<key, val> = <IP addr, MAC addr>



Traffic following ARP takes a shortest path without separate location resolution

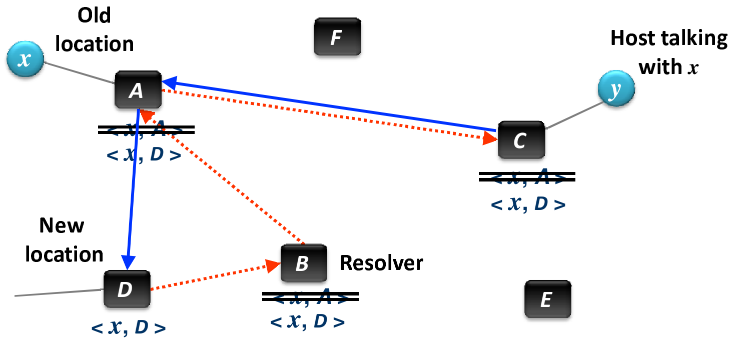
Handling Network Dynamics

- Events not modifying the set of live switches
  - E.g., most link failure/recovery
  - LS routing simply finds new shortest paths
- Events modifying the live set of switches
  - E.g., switch failure/recovery
  - F works differently after a change
  - Two simple operations ensure correctness
    - If $F_{new}(k)$!= $F_{old}(k)$, owner re-publishes to $F_{new}(k)$
    - Remove any <k,v> published by non-existing owners

Handling Host Dynamics

- Host location, MAC-addr, or IP-addr can change

Dealing with host mobility



MAC- or IP-address change can be handled similarly

11

Further Enhancements

- **Goal**: Dealing with switch-level heterogeneity
- **Solution**: Virtual switches

- **Goal**: Attaining very high availability of resolution
- **Solution**: Replication via multiple hash functions

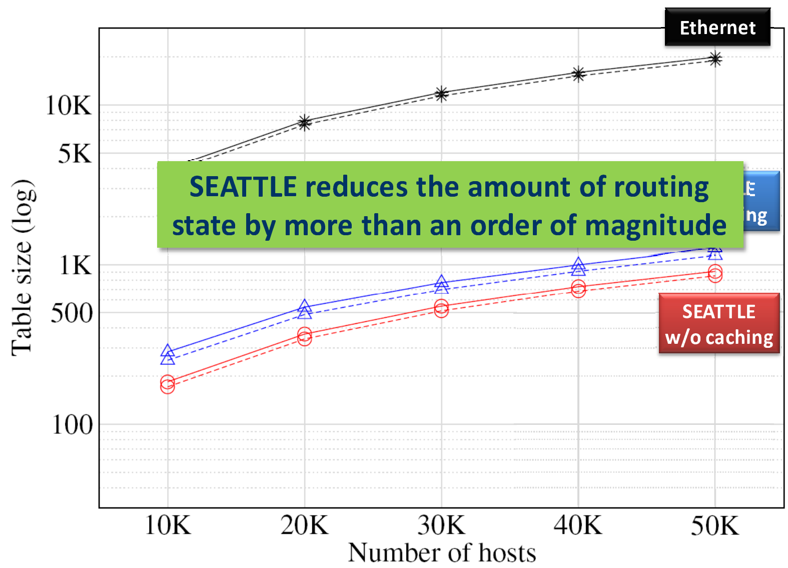
- **Goal**: Dividing administrative control to sub-units
- **Solution**: Multi-level one-hop DHT

12

Performance Evaluation

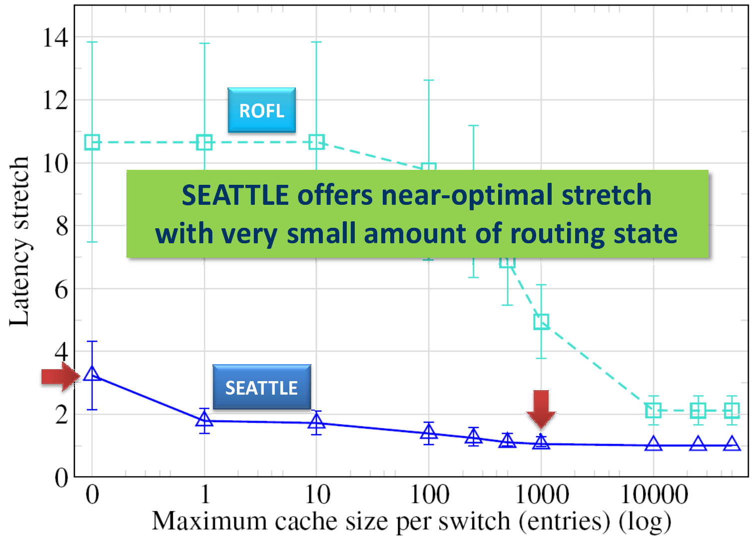
- Large-scale packet-level simulation
  - Event-driven simulator optimized for control-plane evaluation
  - Synthetic traffic based on real traces from LBNL
    - Inflated the trace while preserving original properties
  - Real topologies from campus, data centers, and ISPs
- Emulation with prototype switches
  - Click/XORP implementation

Amount of Routing State



Cache Size vs. Stretch

Stretch = actual path length / shortest path length (in latency)



15

Conclusion and Future Work

- SEATTLE is a plug-and-playable network architecture ensuring both scalability and efficiency
- Enabling design decisions
  - One-hop DHT tightly coupled with LS routing
  - Reactive location resolution and caching
  - Shortest-path forwarding
- Future work
  - Using SEATTLE to improve network security
  - Utilizing indirect delivery for load balancing
  - Optimizations when end hosts can be changed

16