

# Reconciling Zero-config with Efficiency in Enterprises

Chang Kim and Jennifer Rexford  
Computer Science Department  
Princeton University  
{chkim, jrex}@cs.princeton.edu

## 1. Problem Statement

Zero-configuration networking has long been yearned for especially by enterprise network administrators because their resources for control and management are more limited than in service providers. At the same time, enterprise networks are facing two challenges: i) increase of volatility by incorporating portable hosts, and ii) increase of scale by natural growth and by combining multiple remote sites via VPNs. Both of these factors intensify configuration overhead and sub-optimal resource usage. Although self-learning Ethernet bridging [1] can partly solve these problems, it poses some non-negligible side effects as well. From a management perspective, Ethernet bridges have to avoid employing back-up paths (a.k.a., loops) or, when back-up paths are inevitable for survivability, a spanning tree protocol must be configured [2]. In terms of performance, forwarding along a spanning tree does not scale because it concentrates loads and increases path lengths [3]. Spanning tree based flooding, combined with missing TTL support, also requires a conservative approach to failover, resulting in a poor convergence rate [4]. IP, on the other hand, is more efficient in utilizing redundant resources and rapid in failover, but it falls short on “plug-and-play” capability. Its hierarchical addressing and routing requires subnet configurations<sup>1</sup>, although these are not critical parts for enhancing overall networking performance. Hierarchical addressing is also inefficient because address blocks are not optimally utilized. Poor support for mobility is yet another matter.

Given the poor state of the art, we argue that enterprise networking should be re-designed to minimize the dependence on configurations for basic functions, such as routing and addressing, yet to be efficient and scalable. Meanwhile, the new architecture should be incrementally deployable. Specific requirements that such a networking framework should meet are as follows:

- A network must work effectively with minimal or zero configuration on end hosts and network nodes (bridges or routers)
- Network nodes must make optimal use of topological richness by capitalizing on pair-wise shortest paths.
- Flooding for unknown unicast destinations should be avoided because it is unscalable and dangerous. Alternative discovery and dissemination methods must be substantially (i.e., at least an order of magnitude) more scalable than the existing technologies. Such a level of scalability should be ensured even with unstable network nodes and a large number of mobile hosts.
- Networking schemes must be backward compatible with both end hosts and conventional networking equipment. The former means retaining packet formats and service models, such as ARP, DHCP, multicast/broadcast, etc. The latter encompasses support for partial deployment, reuse of the existing hardware and software, etc.
- Networking functions should be able to support key operational tasks, such as traffic engineering, reachability control (access-lists).

## 2. SEIZE (Scalable and Efficient Zero-conf Enterprise) Architecture

We expect the following straw man architecture to serve as one possible solution for a better control plane for enterprises networks.

- Ethernet addressing and frame format  
Ethernet addresses (IEEE 802 MAC-48 addresses) are used as unique identifiers of interfaces across

---

<sup>1</sup>Although DHCP can automate host configurations, operators still have to configure subnets on interfaces and routing instances.

an entire network. Since Ethernet addresses are flat and unique, no subnet configurations are required on network nodes. Intra-enterprise mobility also does not entail host reconfiguration. IP addresses are given to end-hosts only for external reachability and application-layer compatibility. Conceptually, an entire enterprise appears as a large single IP subnet carrying data end-to-end in the Ethernet format. This guarantees backward compatibility to end-hosts because they can use the same Ethernet interfaces and protocol implementations.

- Link-state protocol for distributing topology information

A link-state protocol allows network nodes<sup>2</sup> to unanimously share a complete view of the connectivity among themselves, except in transient periods. Using this topology information, each network node maintains shortest paths to all other nodes. Note that the link-state protocol does not disseminate end-hosts' information for the sake of scalability.

- Hash-based end-host location and address dissemination

To maintain (i.e., to register, deregister, and look-up) end-hosts' locations and addresses (MAC and layer-3 addresses), network nodes use *consistent hash* [5] with the end-host's address as the key. That is, each network node maintains only a small portion of the entire end-hosts' information, and the mapping is dictated by the consistent hash. This dissemination scheme ensures that, when an end-host's information needs update, only a constant number of network nodes are involved in the process. Additionally, the overhead to deal with unstable network nodes is also minimized.

- Backward compatible vs. scalable end-host discovery

For backward compatibility, one can support the conventional "discovery-from-data" mechanism used by Ethernet. This mechanism intrinsically requires flooding to deliver data to an unknown destination, which is unscalable and dangerous<sup>3</sup> with a large number of hosts. As a supplement, upon discovering a new host, one can store the end-host's information using the hash-based dissemination scheme. This efficiently reduces redundant flooding attempts for unknown or forgotten hosts. On the other hand, when scalability and safety is more demanding a concern than backward compatibility, an active host management scheme which tightly monitors hosts' availability can obviate the need for flooding, significantly enhancing the entire network's efficiency and scalability.

- Optimal vs. scalable delivery

Since each network node possesses partial knowledge about end-hosts, when a network node needs to deliver a packet to an end-host whose location is not locally available, it transmits the packet to a "relay" network node that is in charge of maintaining the destination host's location as per the consistent hash. This relay can be accomplished by tunneling or address swapping. For performance's sake, one can optimize this detour path by letting the initiating network node keep the destination host's location in its cache and use a direct tunnel to the destination. This exercise of trading scalability for optimality can be dynamically adjusted by controlling the number of cacheable paths at each node according to administrative goals.

- Securing flooding

For the sake of service-level compatibility, an enterprise network must support broadcast/multicast as well. Flooding, which is a conventional method to support broadcast, proliferates packets when it coincides with a transient loop. As the Ethernet frame format does not carry a TTL, packet proliferation can easily bog down involved network nodes. Pseudo-flooding systematically replicates a unicast packet along a pre-determined cycle-free graph and can replace the conventional physical flooding. Because pseudo-flooding is built on top of unicast, a loop does not proliferate packets. With an intelligent loop detection scheme that does not resort on TTL, network nodes can aggressively remove packets trapped in a transient loop, providing better loop-evasion performance than IP does.

---

<sup>2</sup>We intentionally use a generic term, "network node", because the packet delivery entity in our architecture is different both from the conventional bridges and routers.

<sup>3</sup>Some malicious attacks, such as MAC spoofing and ARP flooding, become more devastating as more end-hosts get involved.

### 3. Summary

Conventional enterprise networking architectures require unnecessary configuration overhead, yet unscalable and sub-optimal. We argue that there must be a better control framework for enterprises that works effectively with minimal configurations, and efficiently with unstable network nodes and a large number of volatile hosts. Our solution provides an initial draft for further research.

### References

- [1] IEEE Std 802.1D 2004, IEEE Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges, IEEE Computer Society and ANSI.
- [2] R. Perlman. “Rbridges: Transparent Routing,” in *Proceedings of Infocom 2004*, Hong Kong, March 2004.
- [3] T. Rodeheffer, C. Thekkath, and D. Anderson, “SmartBridge: A Scalable Bridge Architecture,” in *Proceedings of ACM SIGCOMM*, pp. 205-216, August 2000.
- [4] A. Myers, T. S. Eugene Ng, and H. Zhang, “Rethinking the Service Model: Scaling Ethernet to a Million Nodes,” in *Proceedings of HotNets III*, November, 2004.
- [5] D. Karger, E. Lehman, T. Leighton, M. Levine, D. Lewin, and R. Panigrahy, “Consistent Hashing and Random Trees: Tools for Relieving Hot Spots on the World Wide Web,” in *Proceedings of ACM Symposium on Theory of Computing*, pp. 654-663, 1997.