

Dr. Multicast: Harnessing IP Multicast in Data Centers

NSDI '09 Best Poster Award Recipient



Cornell University

Pros of IP Multicast

Simple communication paradigm.

Useful in data centers:

- data replication,
- service monitors,
- load balancers,
- publish-subscribe systems.

Widely supported.

Cons of IP Multicast

No Policy Control

- Any node can send to any group.

No Group Scalability

- NICs use small, imperfect filters.
- Switches flood all ports if state is exceeded.

No Traffic Rate Scalability

- Multicast *storms* overrun the network.

Wishlist

Acceptable Use Policy (AUP)

- Enable *control* of IP Multicast (IPMC).

Optimized Resource Use

- Use IPMC as far as it scales, then resort to another form of multicast.
- Collapse similar groups.

Rate-Limiting

- Limit IPMC traffic in a fair way.

Ymir Vigfusson



Hussam Abu-Libdeh



Mahesh Balakrishnan (Microsoft)



Ken Birman



Gregory Chockler (IBM)



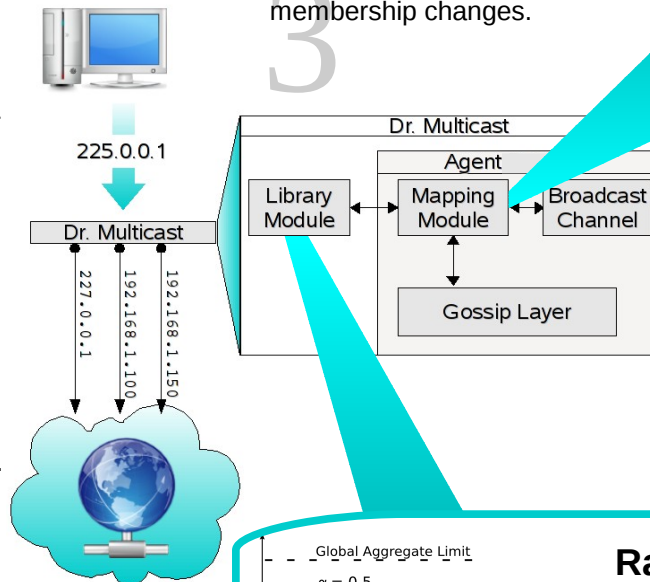
Yoav Tock (IBM)



1
Transparent library layer between application and network. Enforces AUP.

2
Logical groups are mapped to physical IPMC and/or unicast.

3
Gossip & broadcast channel ensure delivery of policy and membership changes.



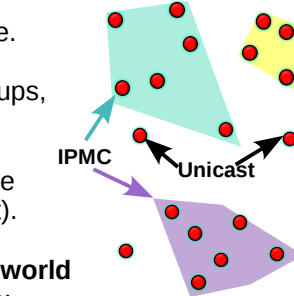
Optimizing Resource Use

Global **group membership** service.

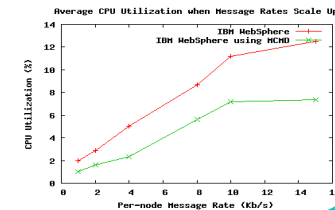
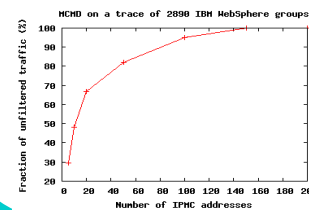
Maps **physical IPMC** to "best" groups, others use **unicast**.

Collapses **similar groups** using the *k*-means clustering algorithm (right).

Compression opportunities in **real-world** systems, such as **IBM Websphere**:



● = group in a metric space.



Rate-Limiting

Receivers **monitor** and **report** group rates.

Senders **slow down** if traffic exceeds limit.

A dynamic **subset** of senders slow down (left).

Experiments (right) indicate **fast reaction**.

