Multicast Address Allocation

Deborah Estrin, Mark Handley,
Satish Kumar, Pavlin Radoslavov,
Dave Thaler

(along with Cengiz Alaettinoglu,
Steve Deering, Dino Farinacci,
Ramesh Govindan, Van Jacobson,
Dave Meyer)

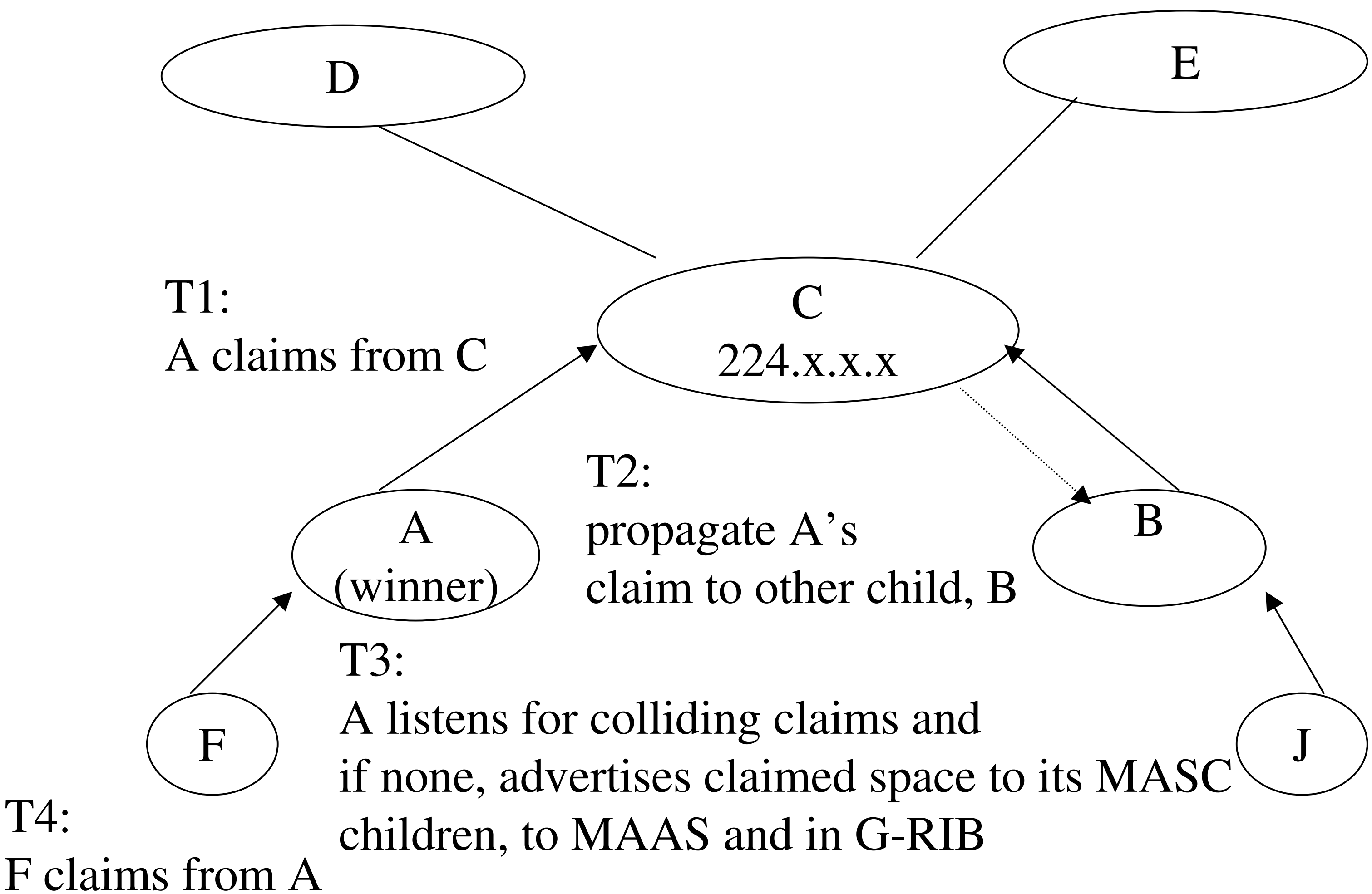
Overview

- “Multicast-Address-Set Claim” (MASC) protocol results in association of group ranges with AS’s
- BR’s announce local ranges internally on locally-scoped group; Address allocators (MAAS) should allocate out of these ranges
- BR’s inject local ranges into G-RIB of BGP to propagate information that will pull BGMP Joins and packets toward the AS

Goals

- Allow aggregation of Class-D routes in G-RIB
- Allow efficient use of address space via dynamic associations, not static allocations
- Avoid third-party dependencies (e.g., no registry)

MASC Example



Step 1: Each AS picks one or more "parent" AS's

- A "parent" AS is one whose address space will be used
- Customer/regional AS's which have a default route can use the AS(s) to which they point default
- Default-free AS's may use static configuration of parent AS(s), or be configured with "No parent" = "I'm a top-level BB provider"

Step 2: Select subrange based on MASC advertisements received

- Listen for parents’ MASC advertisements
- Claim unused subrange of parent’s space and advertise to parent
- Child and parent announce subrange to others without injecting it into BGP; subrange is only a potential association
- MASC advertisements sent separately from BGP to avoid interaction with BGP filtering; independent TCP connections
- More details later

Step 3: Wait for 48 hours(?) to detect collisions

- Need to wait longer than network partitions last
- If collision occurs, pick another subrange

Step 4: If no collisions, then advertise route

- Mark successfully-claimed subrange as "in-use" in MASC advertisements
- Inject group route into BGP in G-RIB
- BR's send successfully-claimed ranges to a locally-scoped group of Address allocators
  - Address allocators SHOULD then allocate from that range

Step 5: BR's monitor address space utilization

- If space gets tight, request new/larger range
- If space empties out, can give back space by:
  - First, decrease space in Prefix-Advertisements sent internally
  - Later, stop defending claim in MASC and withdraw from BGP G-RIB

Claim-Collide vs. Query-Response

- Rationale #1:
  - At the highest level want to avoid third party dependency of single top level server
  - Stick with one mechanism since what is on top is not configured or fixed
- Rationale #2:
  - Q-R approach requires coordination among redundant servers
  - Added complexity and/or added fragmentation of address space and/or still need to deal with collisions
- Rationale #3: Gut intuition :}

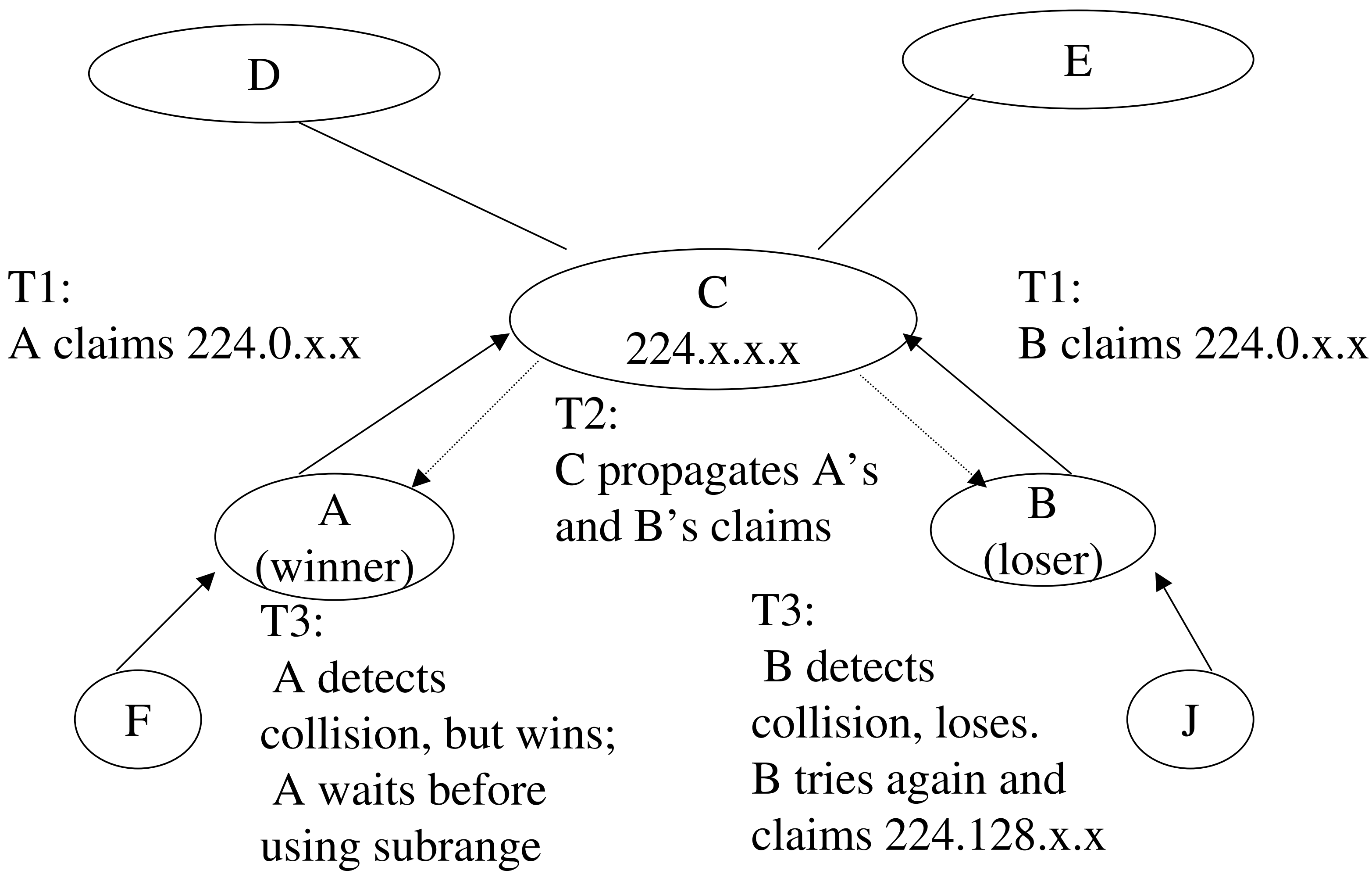
Group addresses have lifetimes

- Every group address has finite lifetime
- If address range still in use/needed, AS continues to include it in advertised claims
- When address range not needed, AS does not reclaim; becomes available to others
- Small set of very long lived addresses for quasi-permanent sessions

Select subrange

- Goals: efficient address space utilization and aggregatability of G-RIB
- Contiguous masks (Thaler's method)
  - prefixes handed out in reverse binary order (e.g., 000, 100, 010, 110 …)
  - to double space include first available bit from right (e.g., 00x owner gets 0xx, if available)
- Non-contiguous masks (Francis' Kampai)
  - Prefixes handed out in increasing order (e.g., 000, 001, 010, 011, …)
  - to double space include first available bit from right

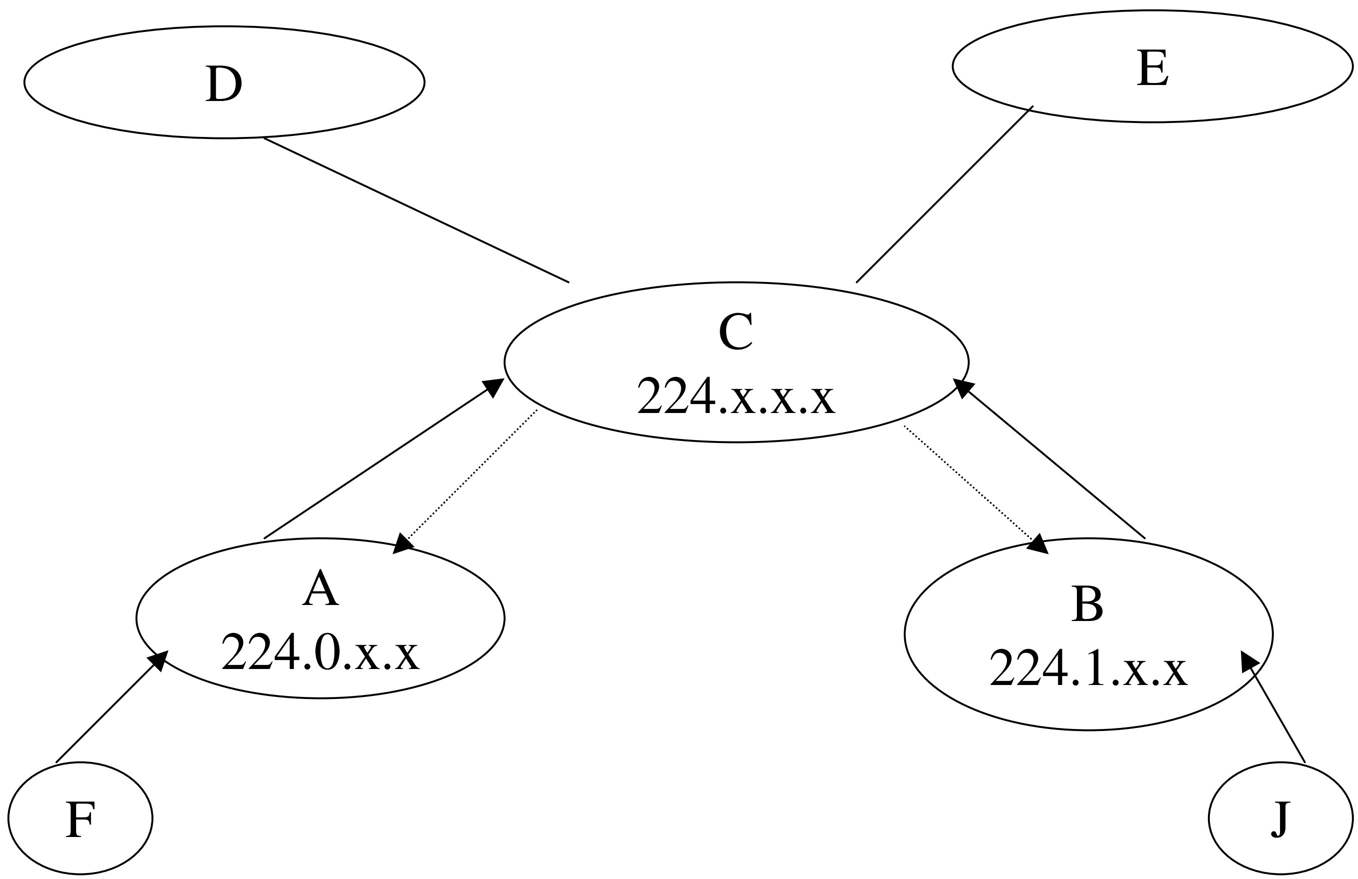
Selecting which new address space to claim: Example



Expanding address space: Example

•A knows 224.1.x.x in use and collision will occur if it claims:
Contig: 224.0.x.x/15, or
Kampai: 224.0.x.x w/mask
(11111111)(**11111110**)(00000000)(0000000)

•So, to expand, A claims:
Contiguous: new prefix 224.128.x.x/16, or
Kampai: 224.0.x.x, with mask
(11111111)(**11111101**)(00000000)(00000000)



Evaluation: open issue

- How much worse is non-contiguous? Does it warrant the manageability concerns regarding discontiguous masks?
- Metrics:
  - Number of prefixes (G-RIB size)
  - Maximum address space utilization
- Study effect of lifetimes, demand distributions, and size estimation

Preliminary findings

- As address space utilization increases, number of prefixes for contiguous is many times higher, when number of children per parent is uniform
- If number of children per parent is non-uniform, number of prefixes for contiguous is much higher
- Contiguous performance can be improved greatly by better subrange-size estimation techniques; static estimation leads to unacceptable results as demand changes

Simulation Results

- See http://netweb.usc.edu/pavlin/masc-results/