parc®
A Xerox Company

PIT AGGREGATION

Marc Mosko, Nacho Solis, J.J. Garcia-Luna-Aceves

CCNxCon 2015 (Palo Alto, CA)

INTEREST AGGREGATION

- When a ‘similar’ Interest arrives at a forwarder, the forwarder may aggregate the Interest with a previous Interest, reducing upstream traffic.
- A ‘similar’ interest is one that could be satisfied by any response to a previous Interest -- difficult to calculate.
- ‘Similar’ is usually simplified to mean ‘asks exactly the same question’.
- Common problem for CCNx and NDN.

WHEN DOES AGGREGATION HAPPEN

- Aggregation can only happen during the RTT of the first Interest/Content Object exchange, otherwise its cached.
- It's a pretty small window for lively data.
- Other situations: *Interest loss*, *Probing for 'new' content*
- We target Interest loss as the main use for PIT aggregation.

NFD APPROACH [1]

- Default InterestLiftetime is 4 seconds
- Aggregate similar Interests up to the Interest Lifetime (e.g. seconds). *unsatisfyTimer* based on InterstLifetime.
- Some applications use RTT estimates (e.g. milli-seconds) for InterstLifetime, some use subscription (e.g seconds) period.
- *Best Route Strategy*: suppress duplicates within MIN_RETRANSMISSION_INTERVAL (100 msec), with possible other methods (RTT estimates or exponential backoff).
- Some strategies retransmit Interests and possibly change the nonce (the wantNewNonce flag).

1. http://named-data.net/publications/techreports/ndn-0021-4-nfd-developer-guide/

EDGE AGGREGATION [2]

- Aggregate at edge, forward everything in the middle keeping colliding hashes longer (e.g. *Best Route Strategy*).
- PIT in core only stores fingerprint
- Operates on 16 msec time quanta (up to 16 sec) for expiration
- Uses heuristic of 9xRTT as PIT lifetime if there's a collision

2. Haowei Yuan; Crowley, P., "Scalable Pending Interest Table design: From principles to practice," INFOCOM, 2014 Proceedings IEEE , vol., no., pp.2049,2057, April 27 2014-May 2 2014.

DHT APPROACH [3]

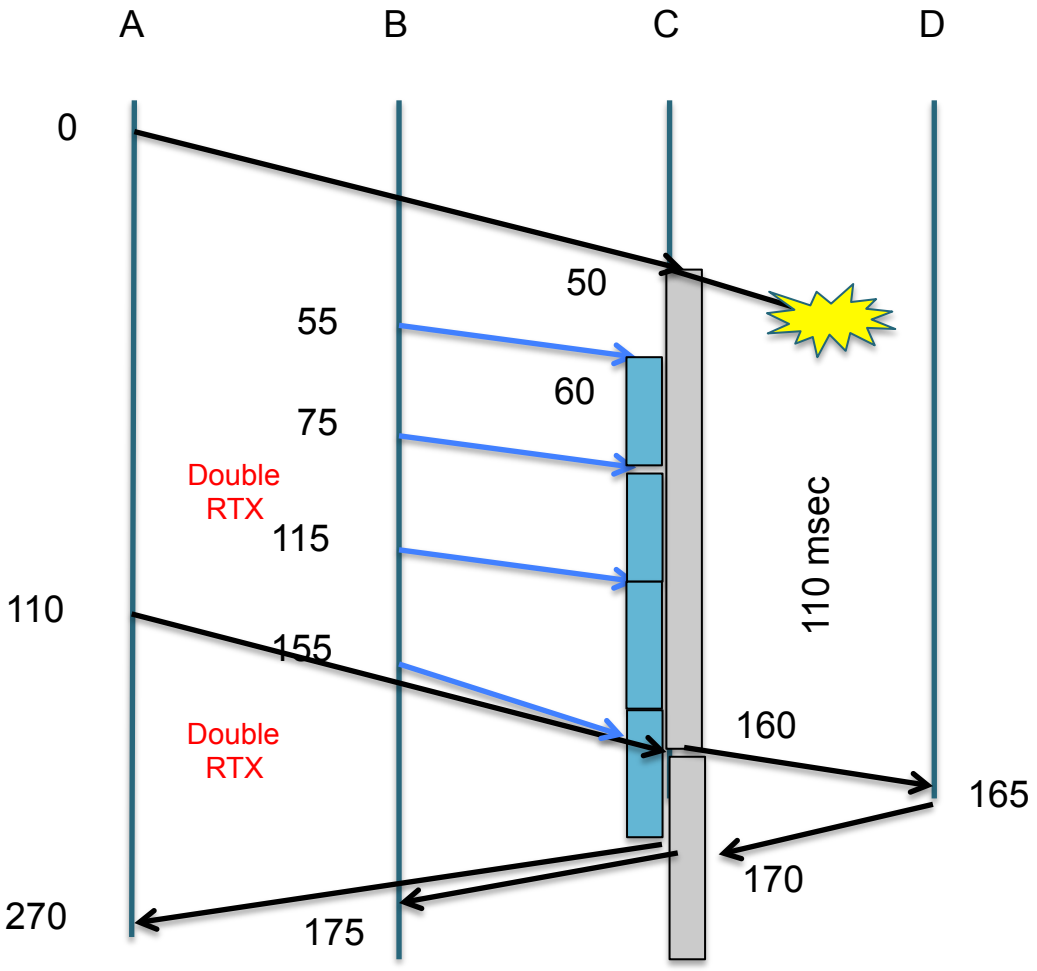
- PIT is a DHT between line cards.
- Aggregate all similar Interests.
- Interests have a timeout, estimates average around 80 msec, possibly up to 500 msec to 1 sec. Does not propose a specific method to pick one.
- Does not specifically address retransmissions.

3. Varvello, M.; Perino, D.; Linguaglossa, L., "On the design and implementation of a wire-speed pending interest table," Computer Communications Workshops (INFOCOM WKSHPS), 2013 IEEE Conference on , vol., no., pp.369,374, 14-19 April 2013

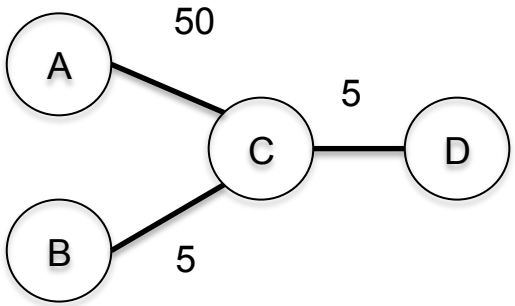
SHORTCOMINGS OF RTT STRATEGIES

- Removing PIT entries based on average RTT misses a lot of Content Objects. Makes RTT estimation difficult.
- If a client (or intermediate system) uses some σ or constant over the mean, that that will over-suppress retransmissions.
- For a forwarder to maintain RTT estimates for a name prefix, it must understand what is a good prefix (or keep everything) and that is long-term storage.
- RTT strategies usually need to decrement the lifetime in-flight to account for link delays on the milli-second level.

THROTTLING RETRANSMISSIONS



Not to scale



B sees this as 4 losses, doubles RTX timeout twice (20 to 40 to 80 msec) and now has a highly flawed RTX timeout.

B sees 120 msec delay, A sees 270 msec delay.

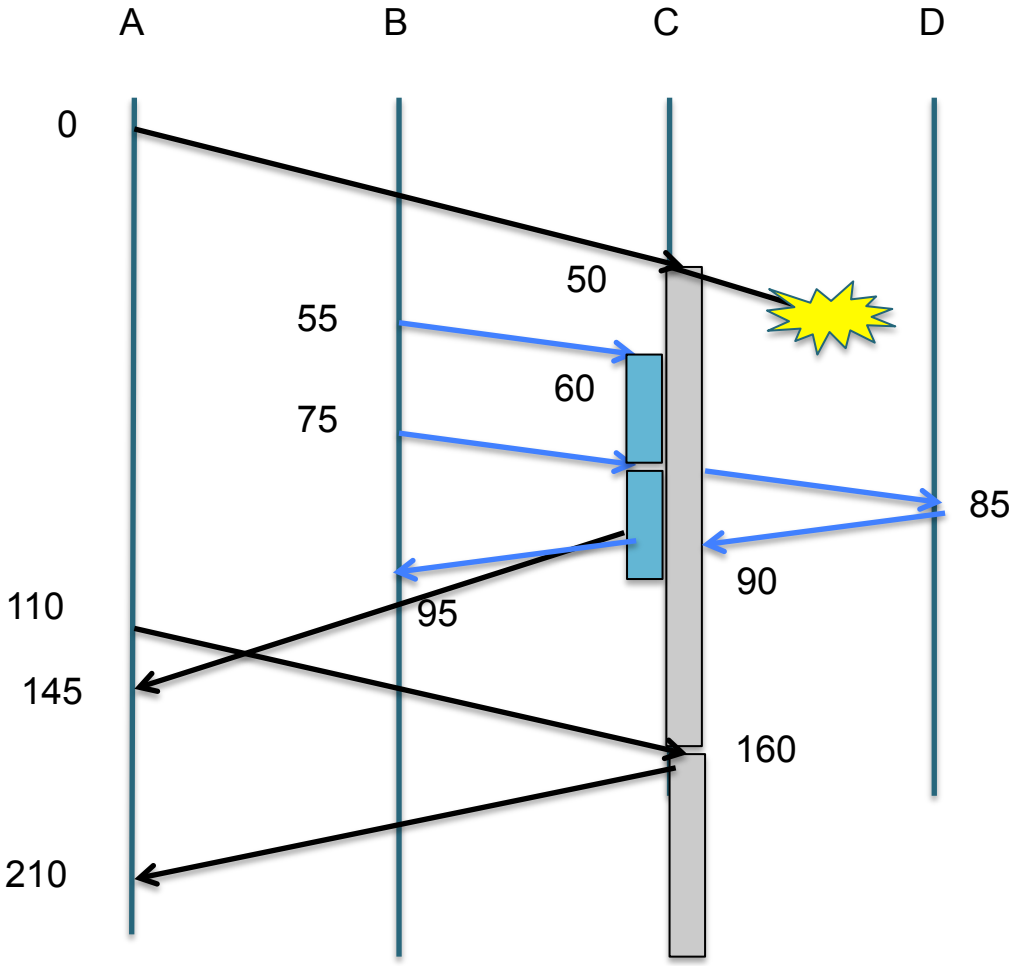
parc
A Xerox Company

PROPOSED ALGORITHM

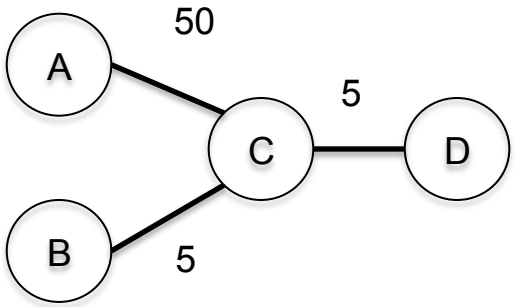
- The first Interest for a tuple {name, selectors†} is always forwarded. Create a PIT entry with expiry based on InterestLifetime (tracked per predecessor).
- A similar Interest *from a different predecessor* may be aggregated. Update PIT entry and possibly extend the Lifetime.
- A similar Interest *from an existing predecessor* is always forwarded, we classify it as a retransmission. Update the PIT entry and possibly extend the Lifetime.
- When a Content Object returns, forward to all predecessors with valid Lifetimes. Remove PIT entry.

† Selectors for CCNx 1.0 are the KeyIdRestriction and ContentObjectHash restriction.

ALLOWING RETRANSMISSIONS



Not to scale



B sees 1 loss, does not increase RTX estimate. Response to B RTX satisfies A too. Response to A crosses its first RTX.

B sees 45 msec delay (2.6x faster), A sees 145 msec delay (1.8x faster).

parc
A Xerox Company

OTHER ADVANTAGES

- Clients can use subscription times, not RTT estimates, for InterestLifetime (forwarders can still discard early). RTX timer independent of what goes in Interest.
- Forwarders do not need RTT estimates.
- Forwarders use LRU or other eviction strategy on PIT, not necessarily timers for each PIT entry.
- Much greater flow isolation that previous PIT aggregation schemes, each predecessor can use own RTX values.

CONCLUSION

- Now the default PIT aggregation method in the CCNx 1.0 forwarder.
- Partners doing video applications have reported significantly better performance.
- Rigorous studies to follow.


Change, disruption, innovation
www.ccnx.org
parc
A Xerox Company


parc®
A Xerox Company
THANK YOU.