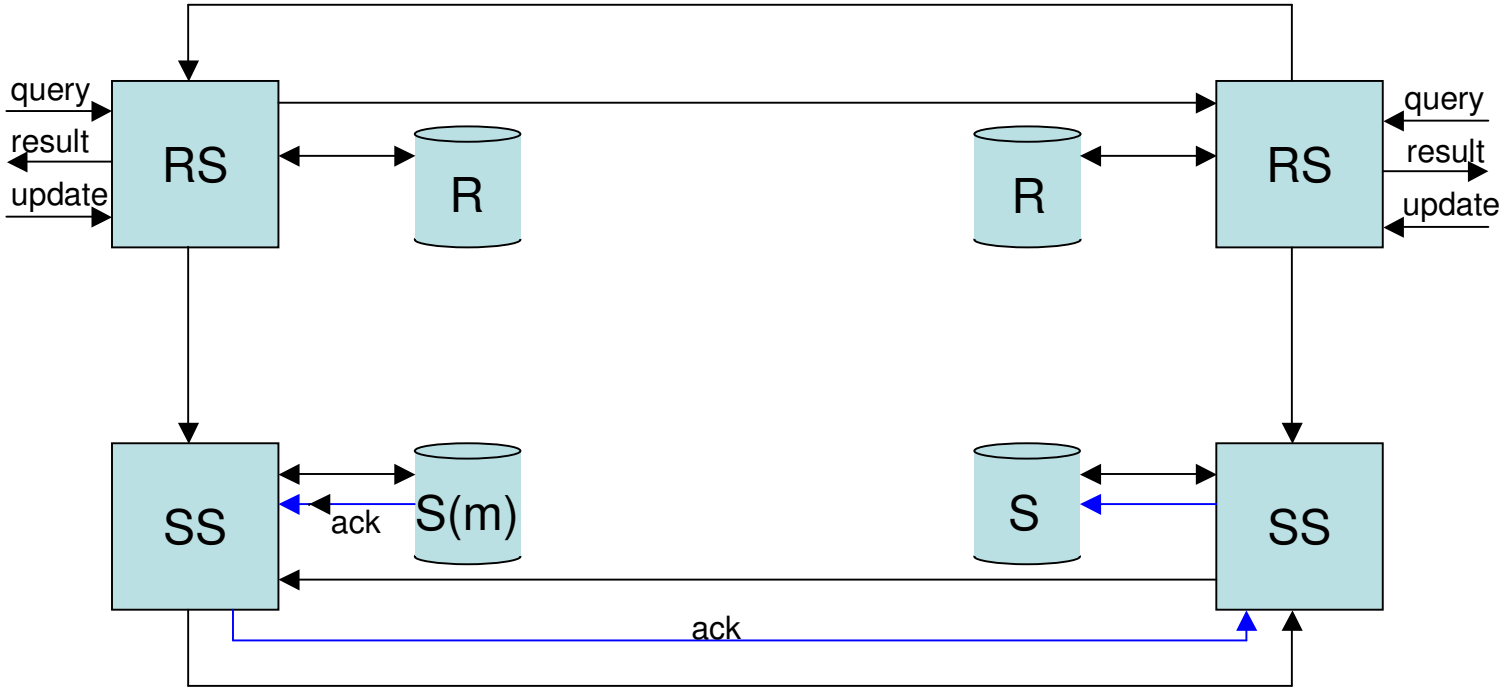
Schema Replication Feedback

Paul N. Taylor

Topology

- Distributed ‘Gossip’ Registry replication. Master/Slave Schema Replication
- 1 Schema per Registry model (only 1 Master schema per VO)
- Queries aimed at ‘local’ schemas
- Updates via local Schemas but aimed at ‘Master’ schema for authorisation



Summary of Feedback

- Move away from static view of DB tables, columns and types in ReplicatableObjects
  - Time pressure on how general this can be
  - Affects quite a lot of code throughout Registry/Schema methods
- Use the leader-election ‘bully’ algorithm to determine the Master schema
  - Based on hierarchy (e.g highest IP number) ordering of Schemas
  - Voting passes a query up the chain to find new Master
  - If Master dies, new election takes place with remaining schemas. Gets the next highest ordered Schema to be Master
  - If Master re-starts, new election initiated, so Master role is re-gained

Summary of Feedback (cont.)

- Queries from Schema are handled locally at remote schemas
- Updates to Schema first contact the Master.
- If authorised, Master stores new table first, send ack to requesting Schema, then post update to all remaining Schemas
- Adding a new Registry could be handled by telling new Registry about an already ‘connected’ Registry.
  - New Registry sends add-me message to ‘connected’ Registry, which updates its config file, and forwards add-me to all other known registries.
  - Would need an add-me (and forward) and add-me-only (no forward) methods.

Summary of Feedback (cont.)

- On Master re-start, as part of leader election, need a piggy-back message to send current state to other Schemas.
  - Inconsistent state is highlighted when compared
  - Merge states to bring ‘old’ Schema(s) up-to-date
- Election of leader when sub-nets merge which each had a leader
  - Decide the leader based on IP address method
    - IP address unique and works across entire VO
  - How to handle name clashes between sub-nets?
    - Could table names be appended with Producer id?
      - What to do with Producer trying to publish to ‘changed’ table name
- Security
  - Validation of registrys/schemas during replication
    - Do you know/trust who you’re sending to?
  - Adding a new registry/schema to the VO
    - Is this new registry/schema known/trusted?