Towards Cooperative Self-Organized Replica Management

Work in Progress

David Hales, Andrea Marcozzi (University of Bologna)

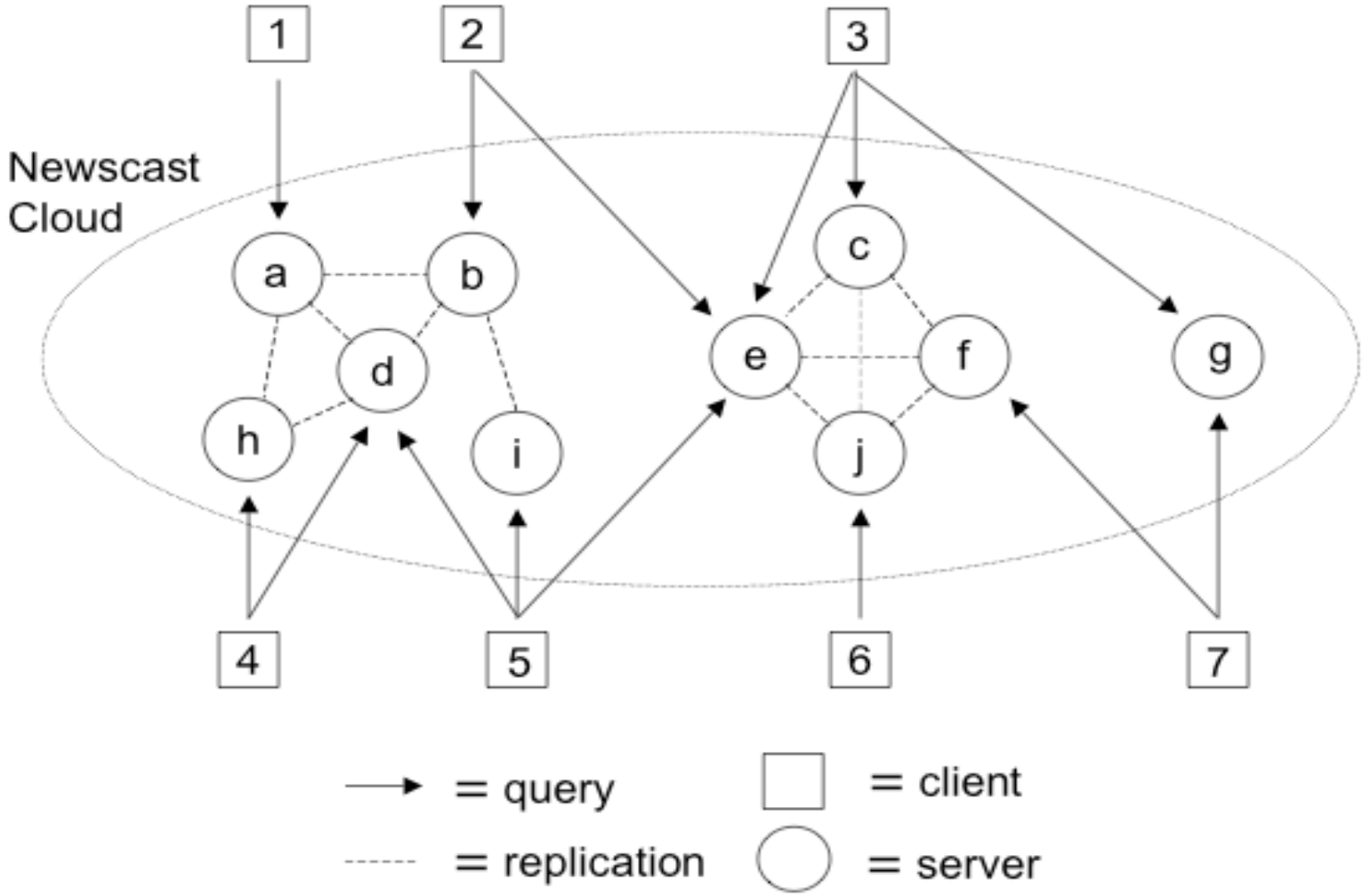
Giovanni Cortese (University of Rome, RadioLabs)

SASO 2007, Cambridge. Mass.

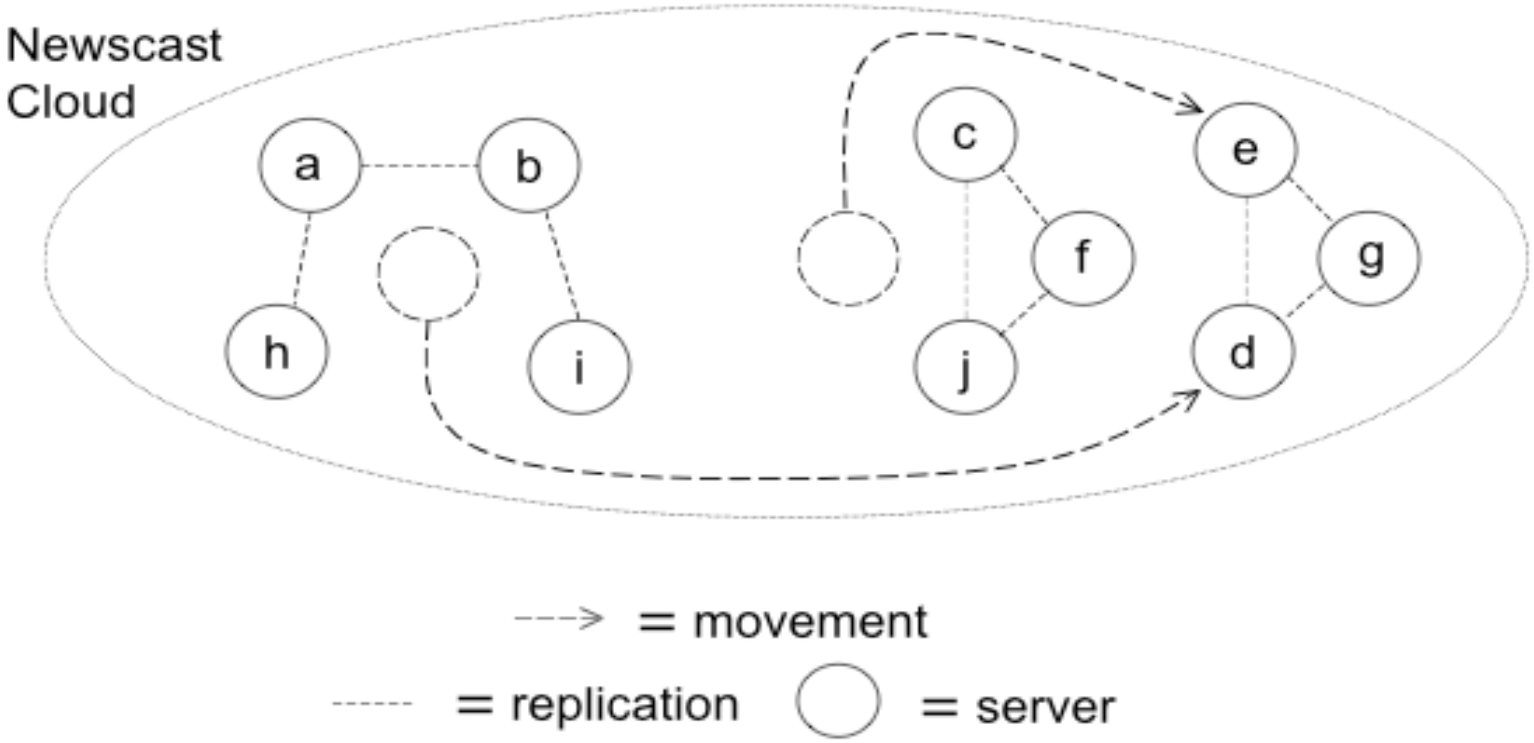
CacheWorld: Cooperative content replication

- Cooperative Self-Organized Content Replication
- Assuming protocols for
  - Replicating content between nodes
  - Redirecting queries (requests for content) between nodes
  - Peer sampling over a population of nodes
- We want:
  - Simple protocol for cooperatively coordinating these services to maximise system capacity
  - With incentives for nodes to cooperate
  - Dynamically adjusting to varying load and node entry and exit

CacheWorld outline



CacheWorld outline



CacheWorld Outline Algorithm

- capacity and load for each node specify different scenarios
- maximum number of neighbours (k) currently fixed
- nodes “satisfied” if all queries submitted to them are answered (over a given period - the load cycle)
- nodes associated with single unique content item replicated between linked neighbours
- nodes are “receptive” if they have spare capacity or are not satisfied

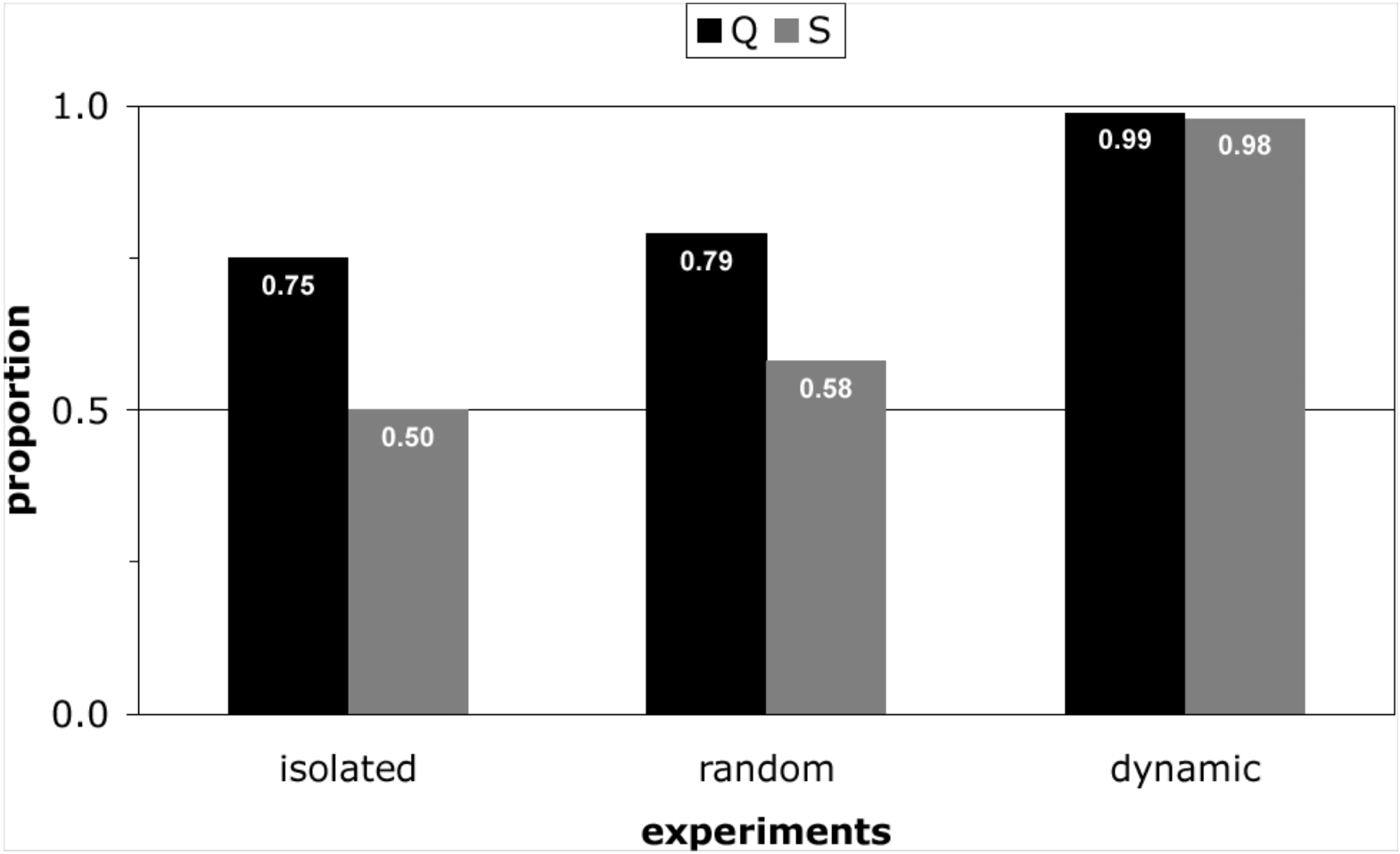
Passive thread

```
on receiving a query q, node i:
        if not overloaded, service q directly
        else if neighbors > 0 and q is not
        already a redirected query
                j ← selectRandomNeighbor()
                redirect q to j
        end if
```

Active thread

```
periodically each node i:
        if not satisfied
                drop all neighbor links
                j ← selectRandomPeer()
                if j is receptive then link to j
        end if
```

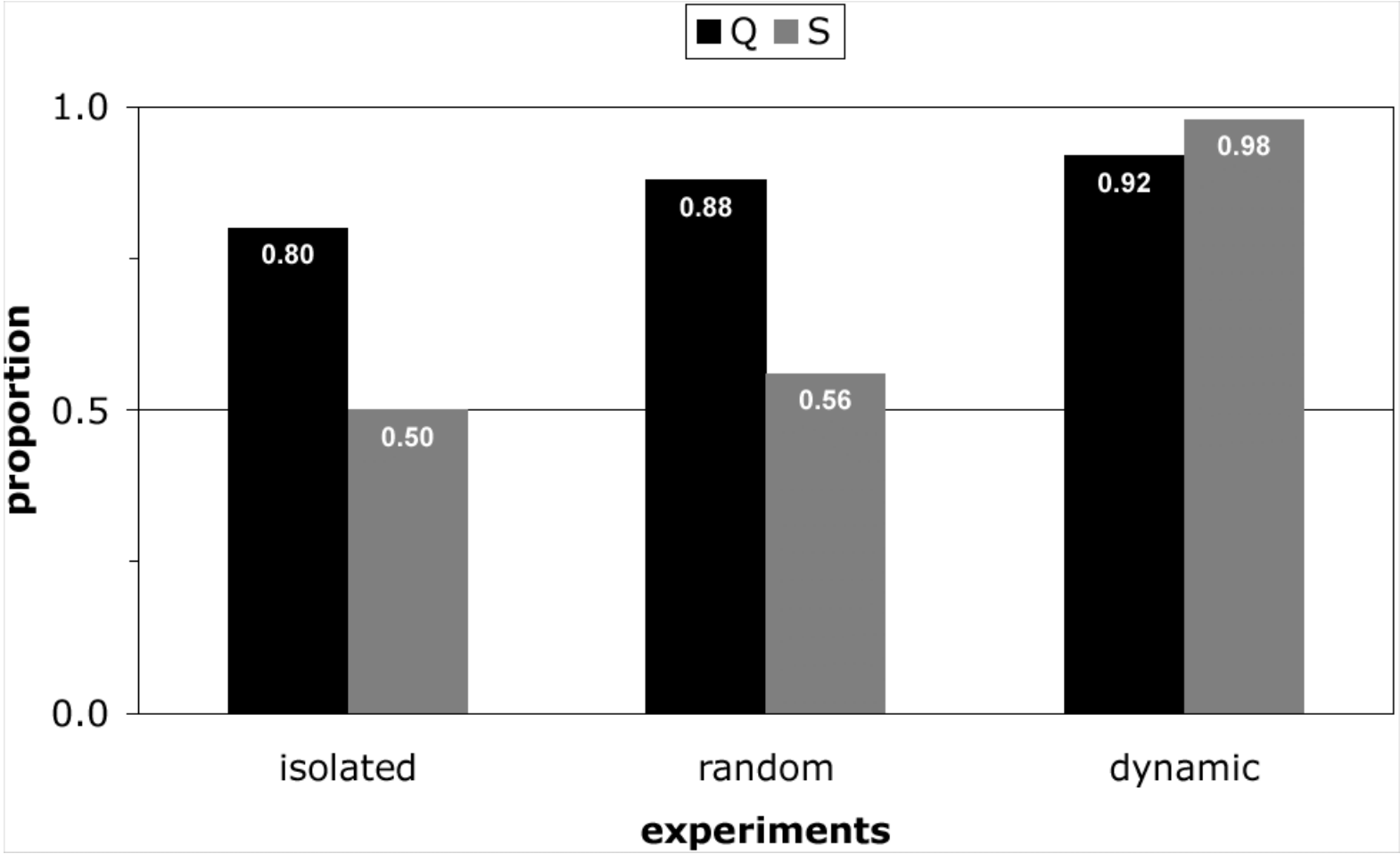
Some initial results



Q = queries answered, S = satisfied nodes

(very simple scenario, half nodes underloaded, half overloaded, k = 1)

Some initial results



Q = queries answered, S = satisfied nodes

(less simple scenario, half nodes under-loaded, half overloaded, k = 4)

CacheWorld in progress

- Very initial results, with simple load / capacity scenarios
  - Nodes replicate and serve a single “content” item
  - Not modelling cost of replication process
  - Fixed loads and capacities
- Hence more realistic scenarios needed and comparison with existing protocols (Squirrel, Globule)
- Still not tested with malicious nodes, pure free-riders and churn. But reasonably confident will degrade gracefully
- Specially factoring in replication cost could help prevent bad guys
- Varying of satisfaction threshold
- simple “loyalty” approach (where preference is given to older links) could lead to much better results but this is on-going. Interesting this could link to a lot of work from “evolutionary economics” (Kirman’s Marseille Fish Market studies / models)

WP5.3 Summary

- Related Publications:
  - Hales, D. and Arteconi, S. (2006) SLACER: A Self-Organizing Protocol for Coordination in P2P Networks. IEEE Intelligent Systems 21(2):29-35
  - Hales, D. (2006) Emergent Group-Level Selection in a Peer-to-Peer Network. Complexus 2006:3.
  - Hales, D. and Babaoglu, O. (2006) Towards Automatic Social Bootstrapping of Peer-to-Peer Protocols. ACM SIGOPS Operating Systems Review 40(3)
  - Arteconi, S., Hales, D., Babaoglu, O. (2007) Greedy Cheating Liars and the Fools Who Believe Them. Proc. Workshop on Engineering Self-Organising Applications (ESOA2006), Springer

Get these from www.davidhales.com