

Tangram

A robust and reliable directory system based on DHT algorithms

Index

- Introduction
 - Collaboration, Problem, Existing solutions
- State of the Art
 - P2P, 2nd Generation P2P
- Building a Distributed database
 - Tangram nodes and external, queries, optimized PDU
- Replication
 - Local and geographical, Finite-state machines
- Performance
 - Load balancing, topology, performance figures
- Conclusions and future work

Collaboration Project

- Universidad Politécnica de Madrid



Departamento de Ingeniería de Sistemas Telemáticos
ETS de Ingenieros de Telecomunicación
Universidad Politécnica de Madrid



POLITÉCNICA

- Ericsson España



TAKING YOU FORWARD



Problem domain

- Telecom Operators directory databases.
- High Availability
 - 99.9999% or better figures
- Real time demands
 - You don't want delays on your phone calls.

Existing solutions

- Complex scalability.
- Manual configuration.
- Provisioning vs traffic phase.

P2P

- Mainly used for file sharing.
- Technology should be neutral
- Are there other possible uses?

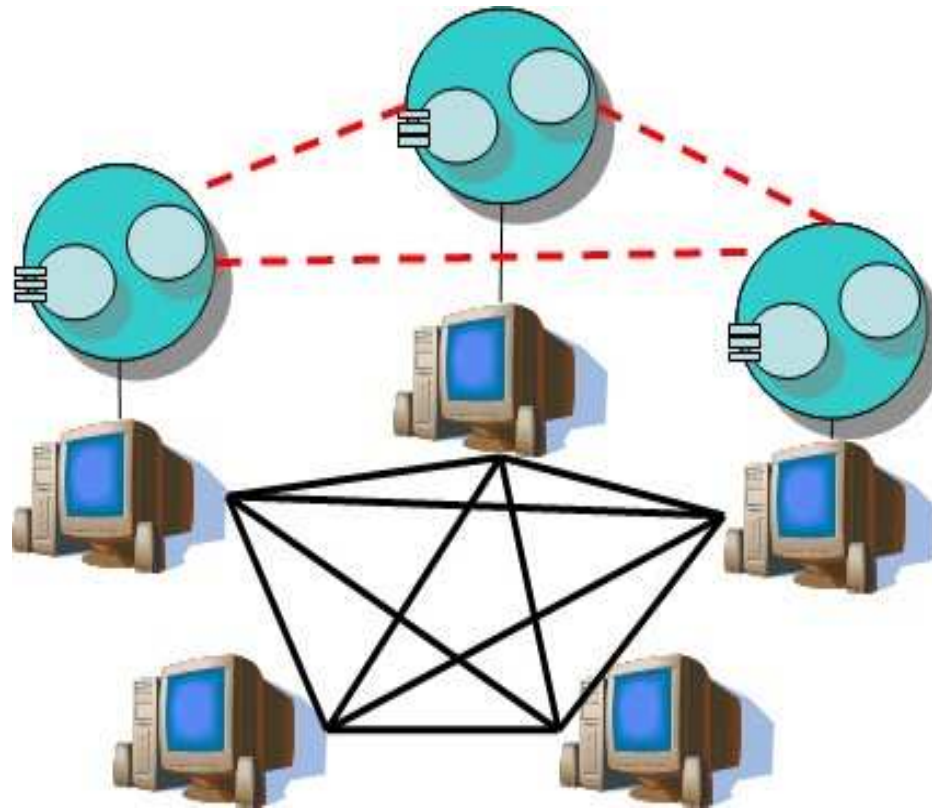
Second Generation P2P

- DHT based P2P.
- Distributed Hash Table:
 - Hash Table index with routing algorithm instead of local pointers to data.

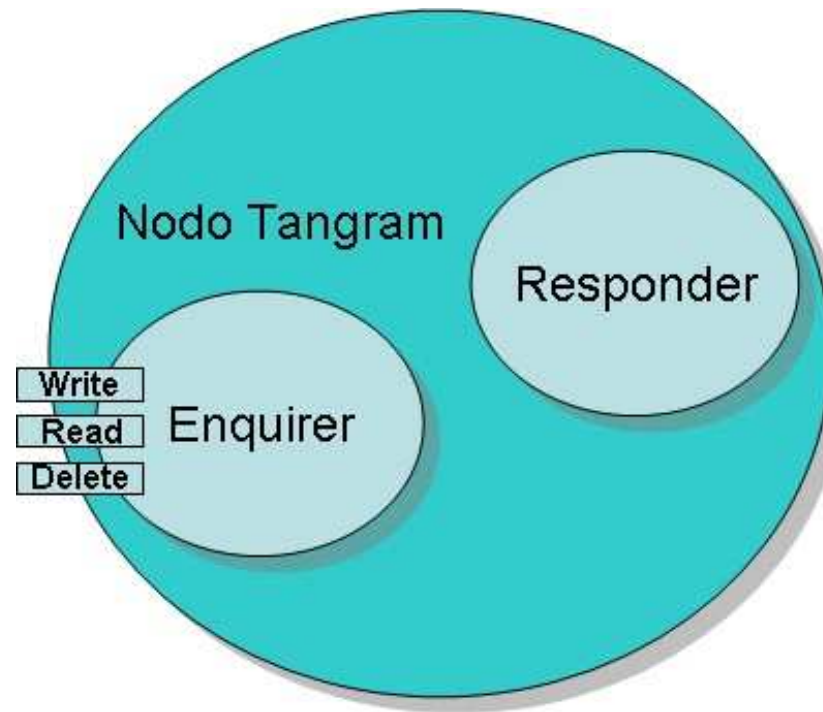
Building a Distributed database

- Advantages of P2P:
 - Self-configuration.
 - Fast response time.
 - Not single point of failure.
 - No bottleneck in front end.
- In memory database for faster response.

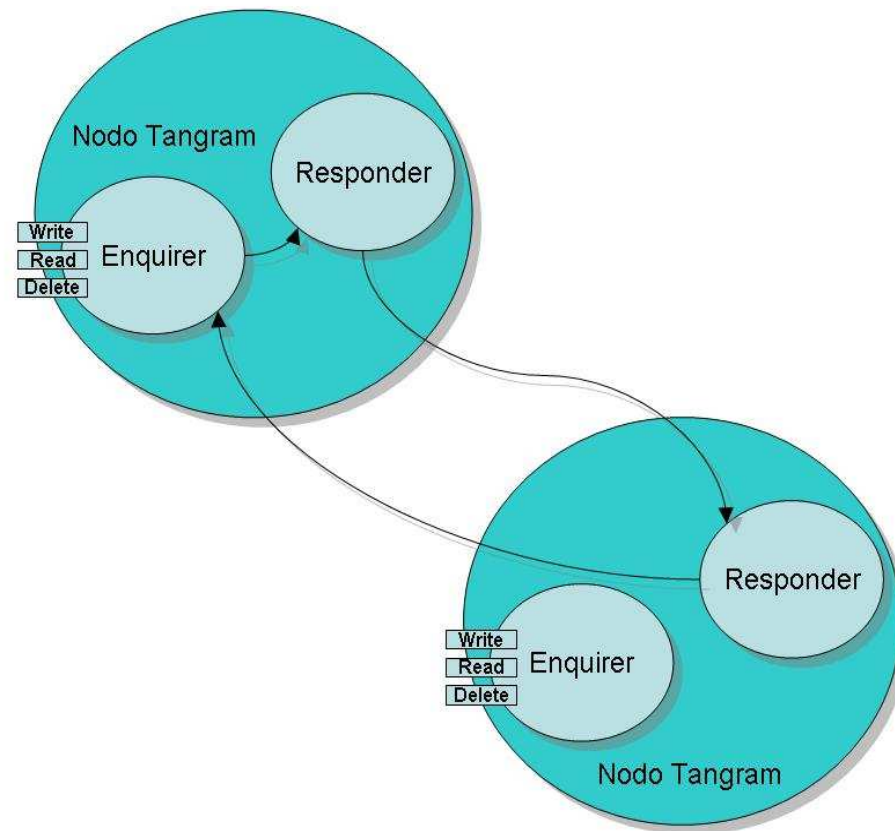
Tangram is an overlay network



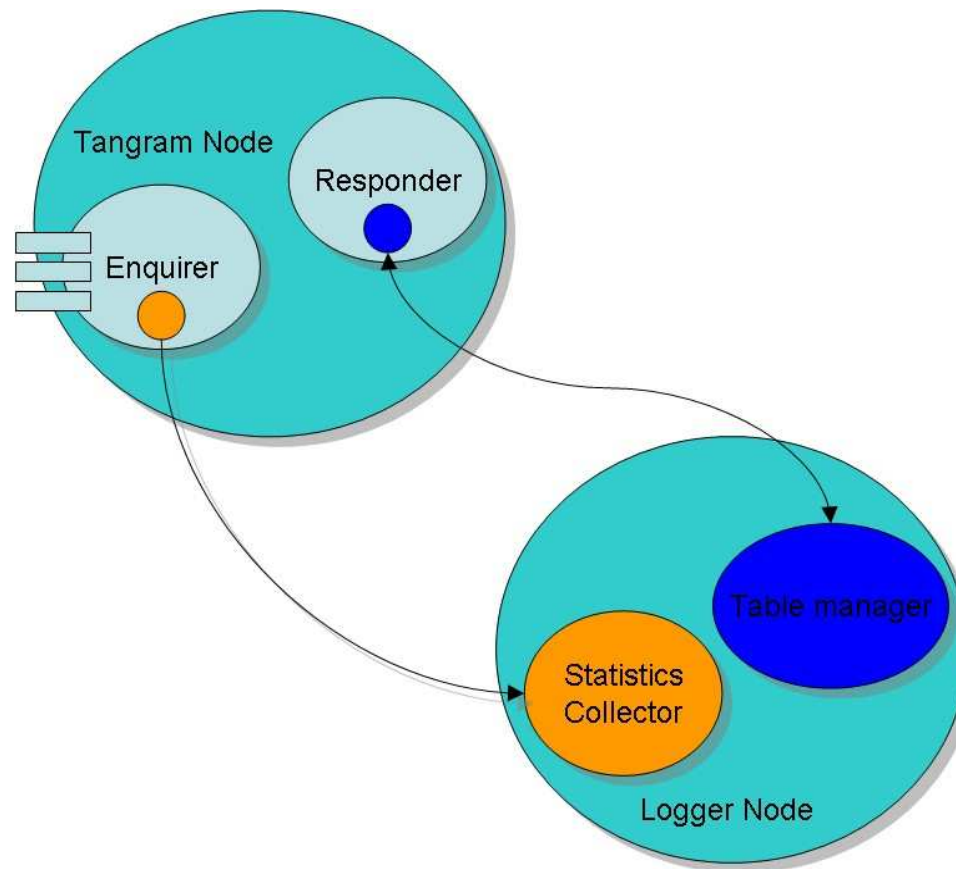
Minimalist Node structure



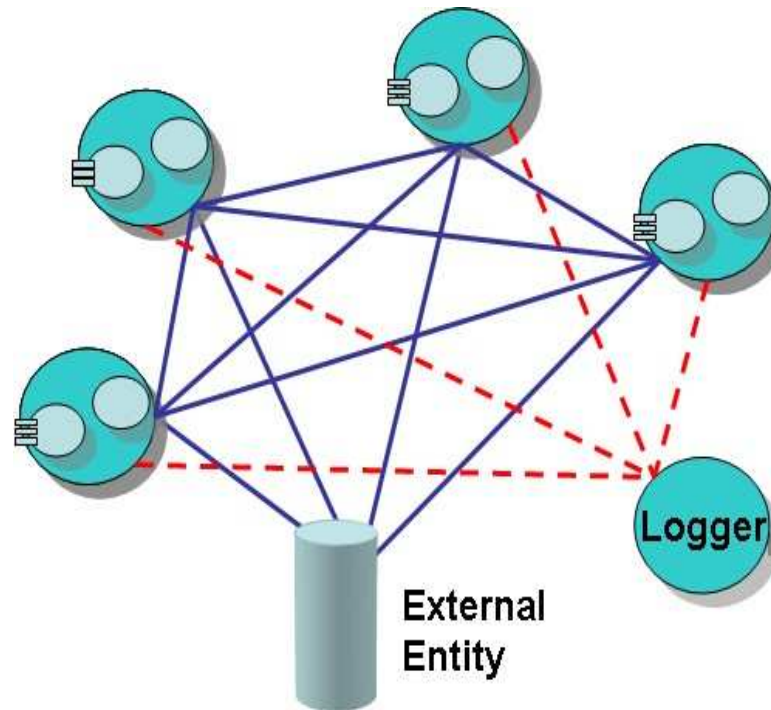
Every node is born equal



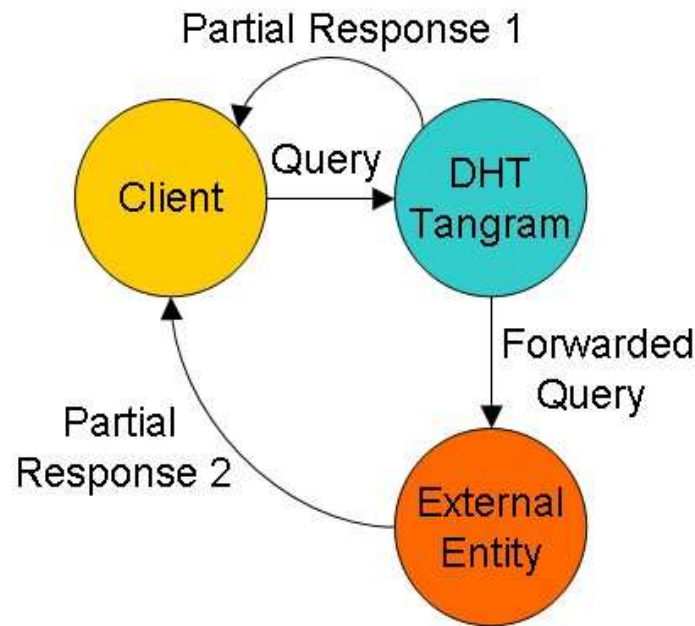
Monitoring the nodes



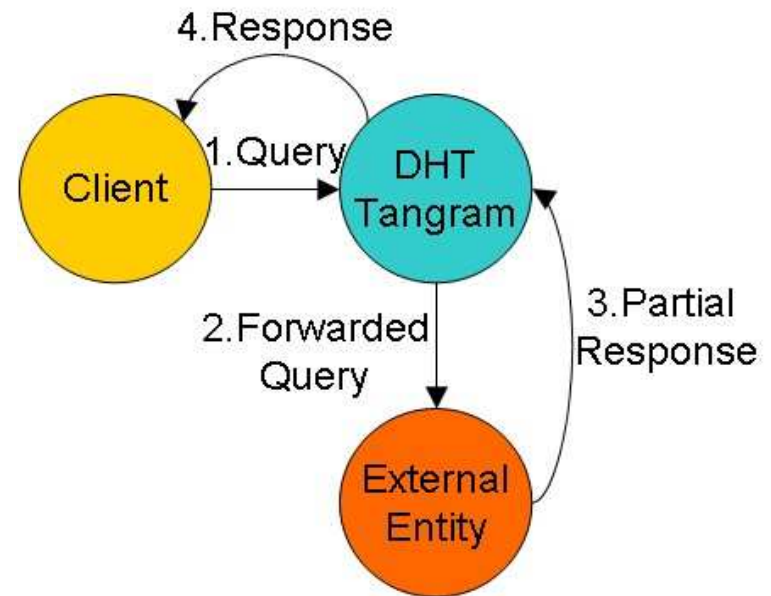
External Nodes can be added



External nodes 2-working ways

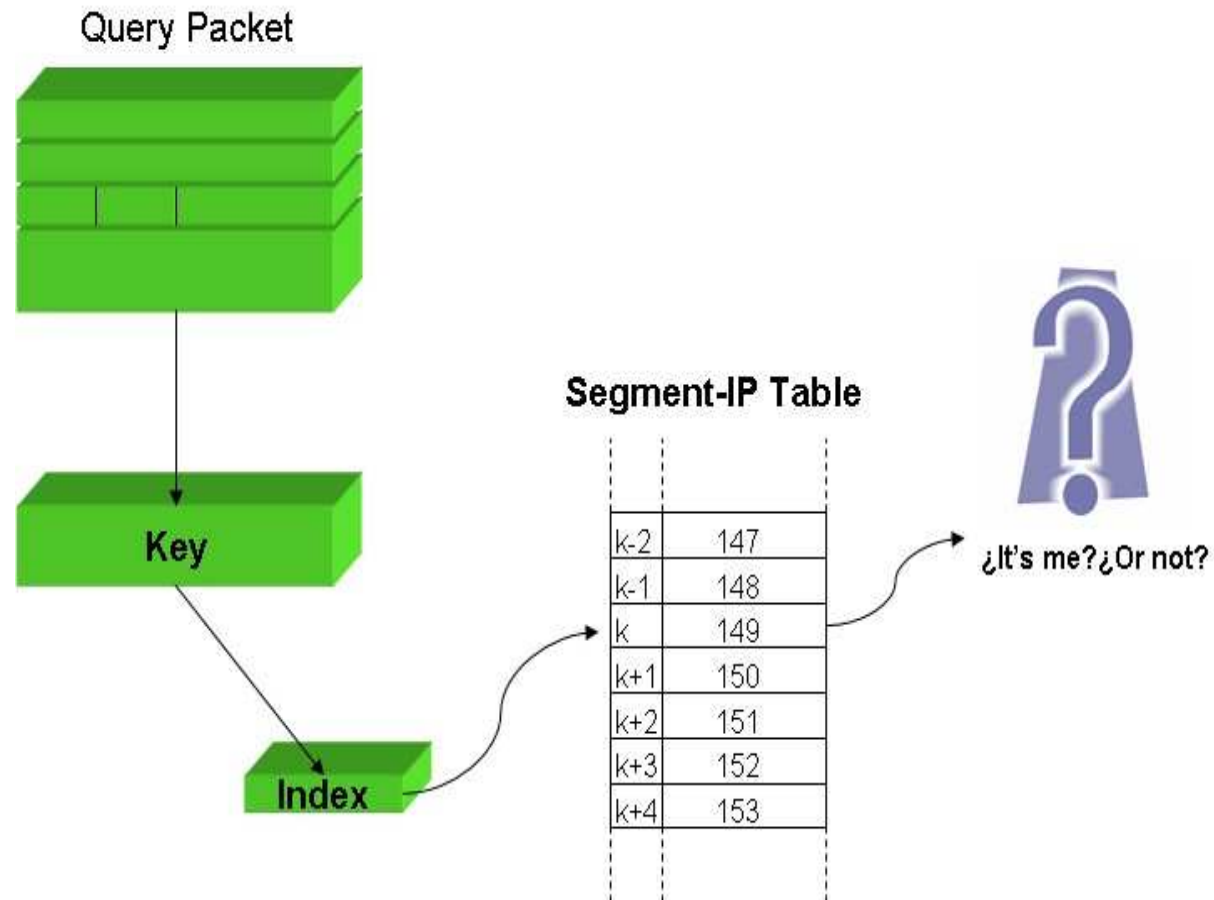


Direct response

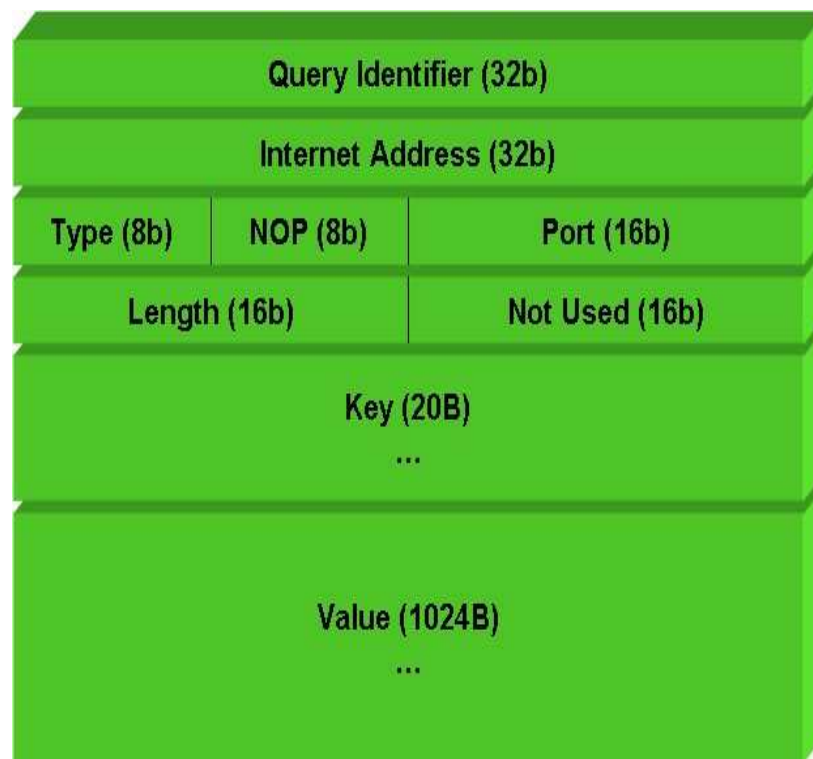


Forwarded response

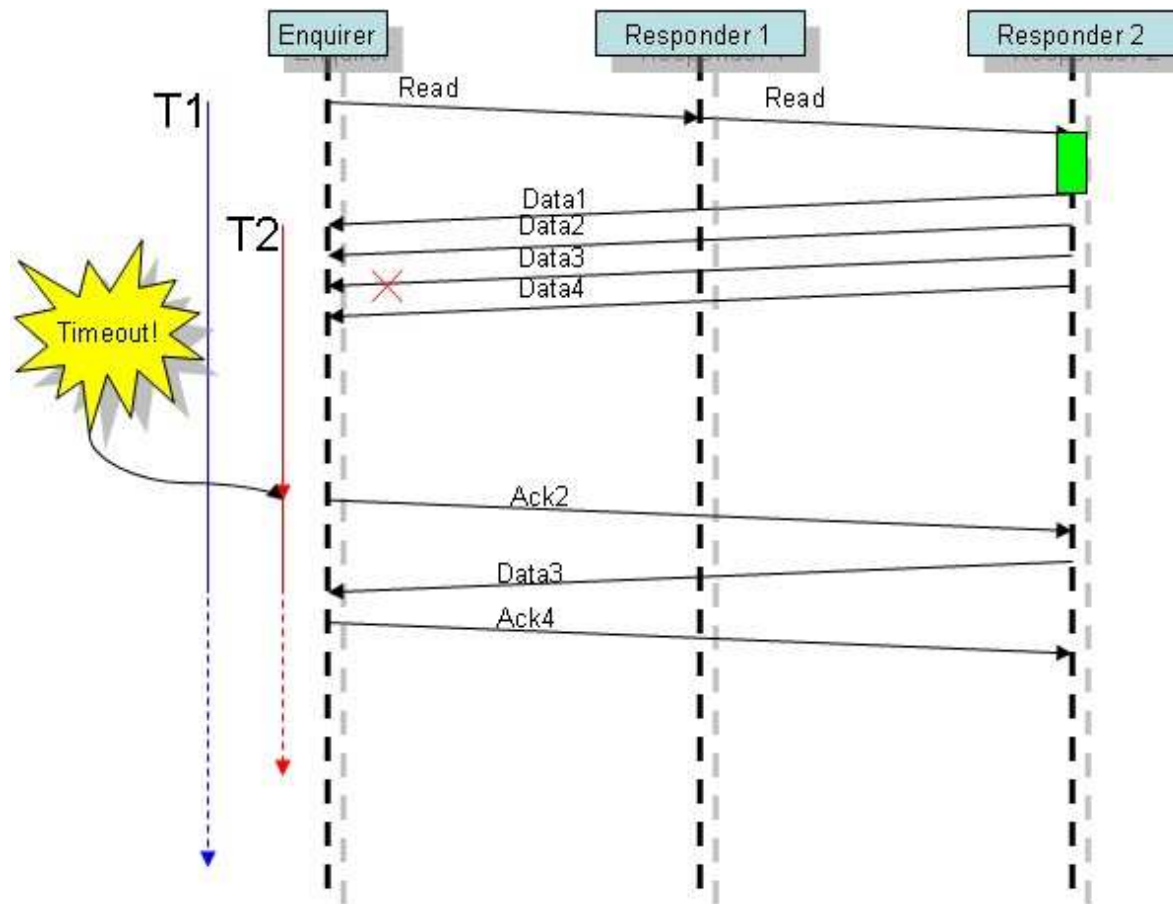
Tangram Query in Action



Tangram optimized PDU



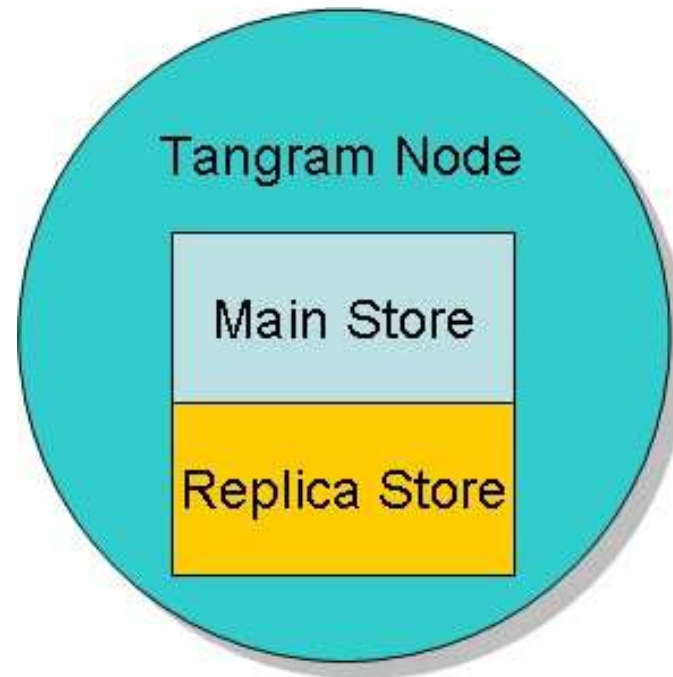
Retransmission of lost packets



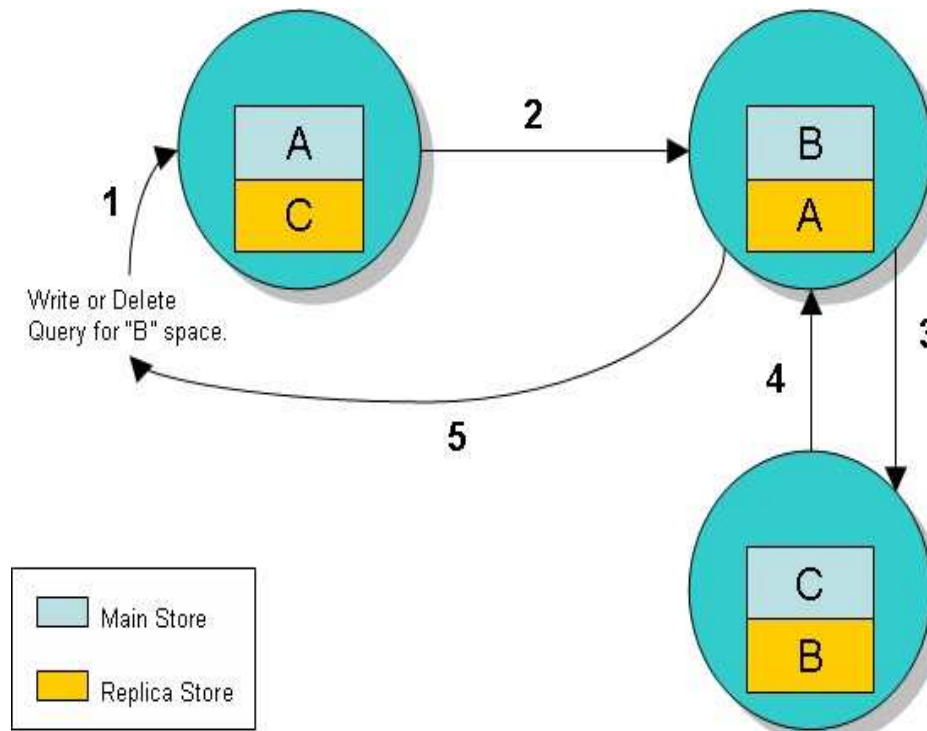
Replication

- 2 levels:
 - Local redundancy (in case of node failure).
 - Geographical redundancy (whole site failure).
- The overall system must always be on.

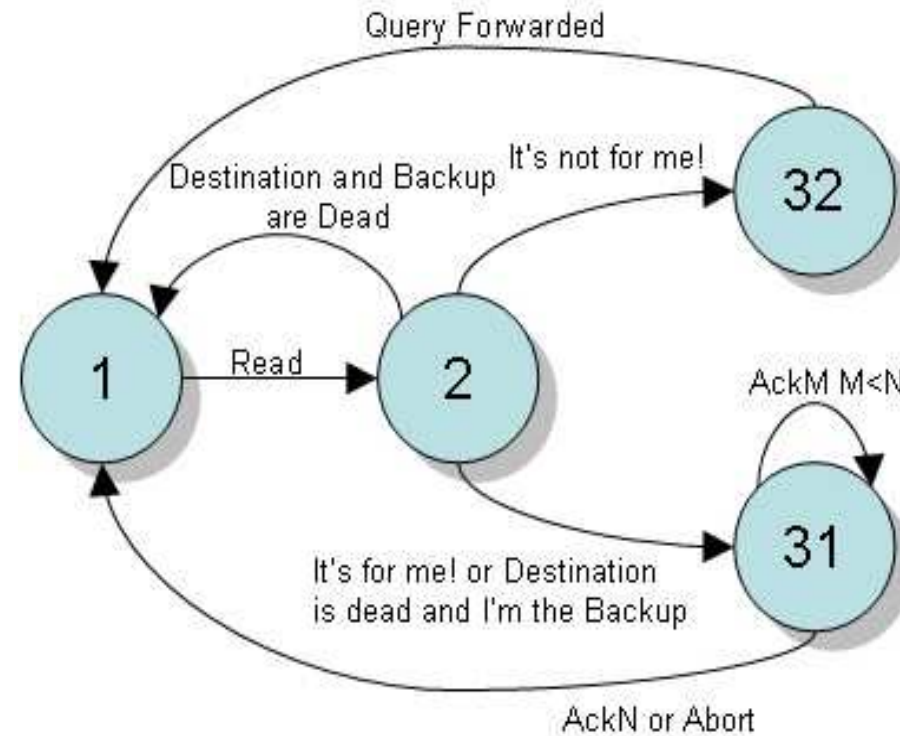
Node Replication



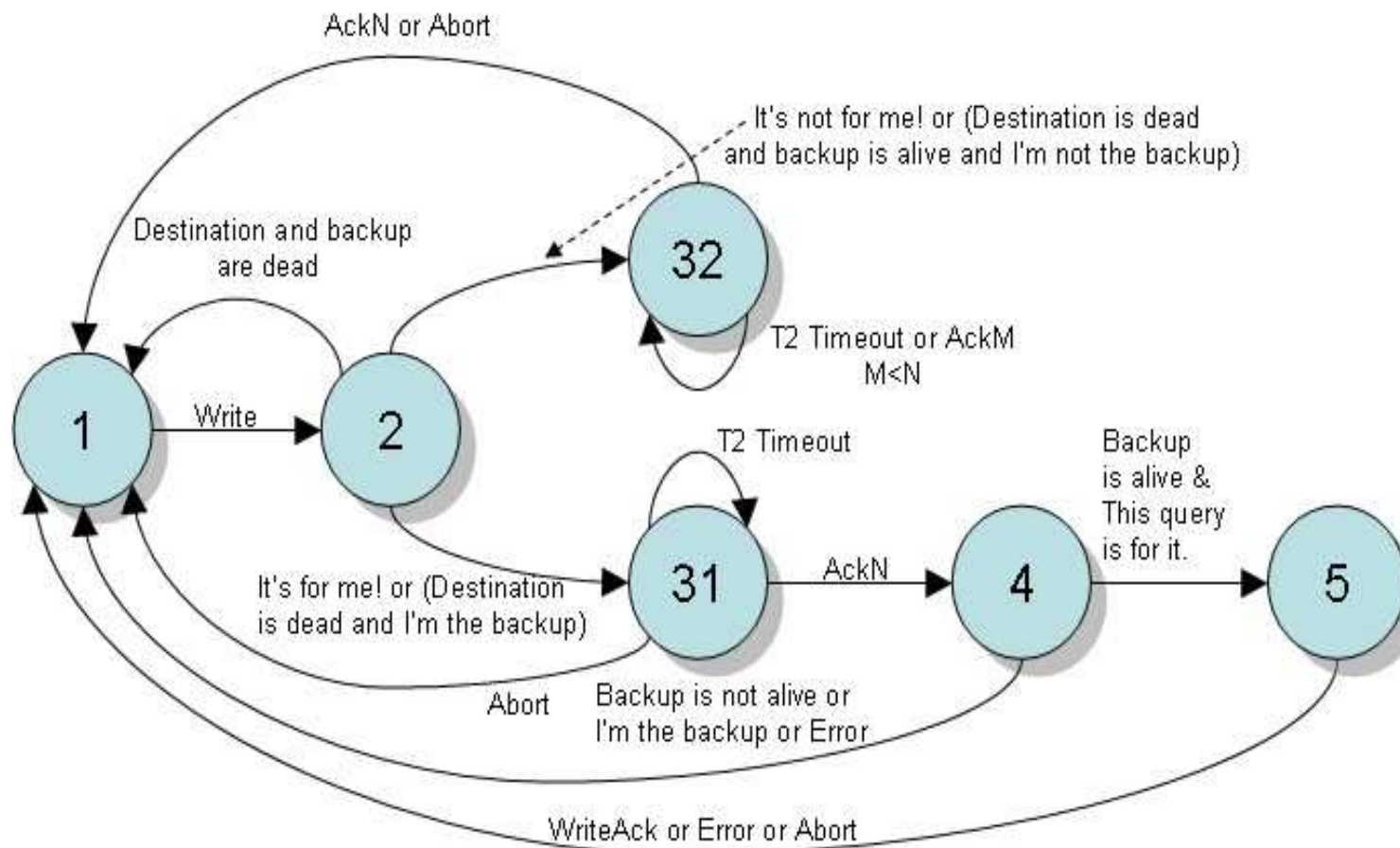
Several nodes replicated



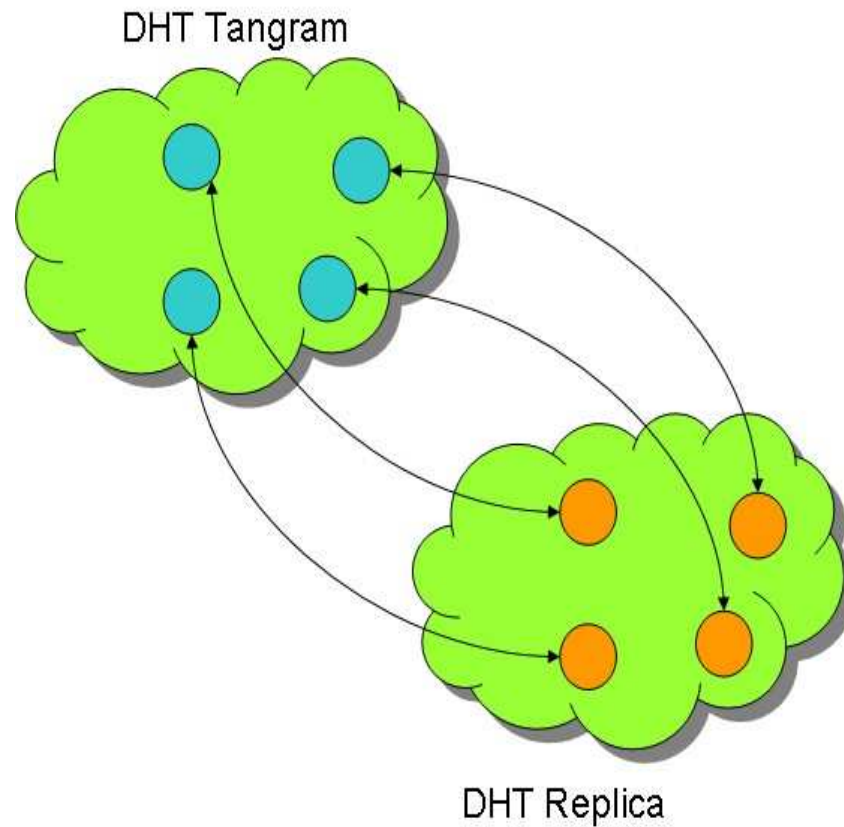
Tangram Replicated-Read FSM



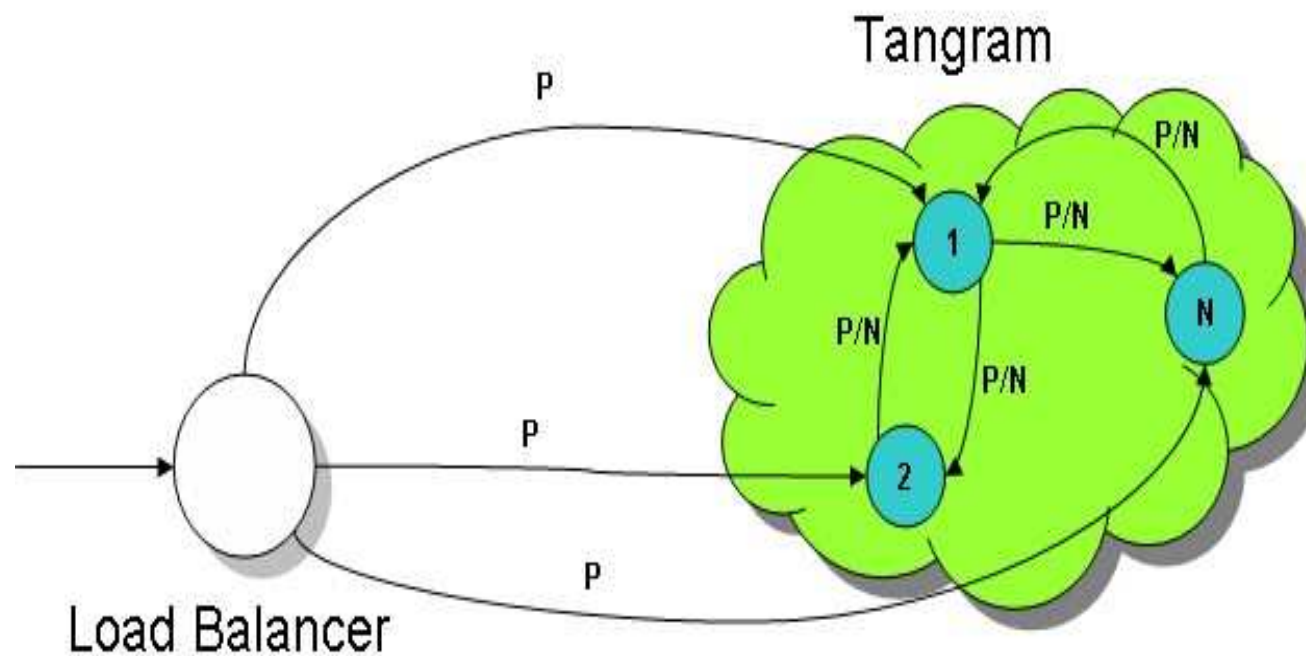
Tangram Replicated-Write FSM



Geographical Replication



Sharing the load

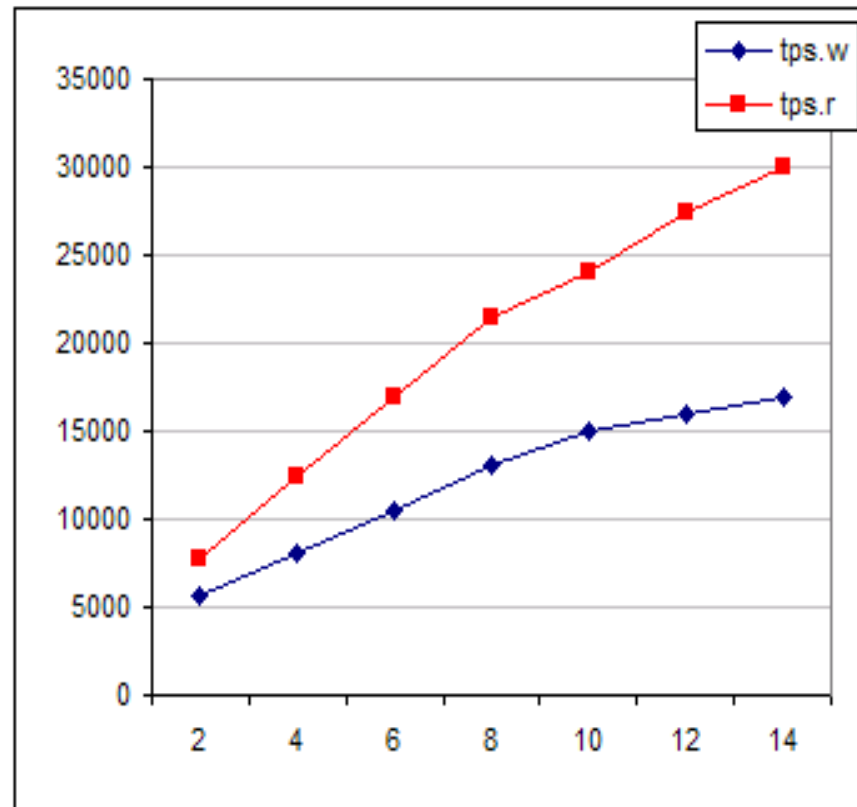


From Grid to Cluster

- Started with Grid / P2P principles
 - Distributed processing and storage.
 - Self-organizing nodes.
- Telecom-specific issues added in
 - Operation fully controlled for high availability.
 - Nodes are typically in cluster-like topology.

Performance Read & Write

TPS



Number of Nodes

Conclusions

- DHT used for real-time databases
 - Performance figures are excellent.
 - Self-organization could reduce costs.
- Adaptable to telecom environment:
 - No single point of failure
 - High availability, redundancy

Future works

- Open source distribution
 - Open for collaboration with the community.
- Experiments on Planet-Lab
 - Real grid like adaptation.
- Contact us: tangram@agws.dit.upm.es

Questions?

