Replacing External Configuration Service with Chord Peer-to-Peer Protocol

Cloud Data Bases WS2021-22

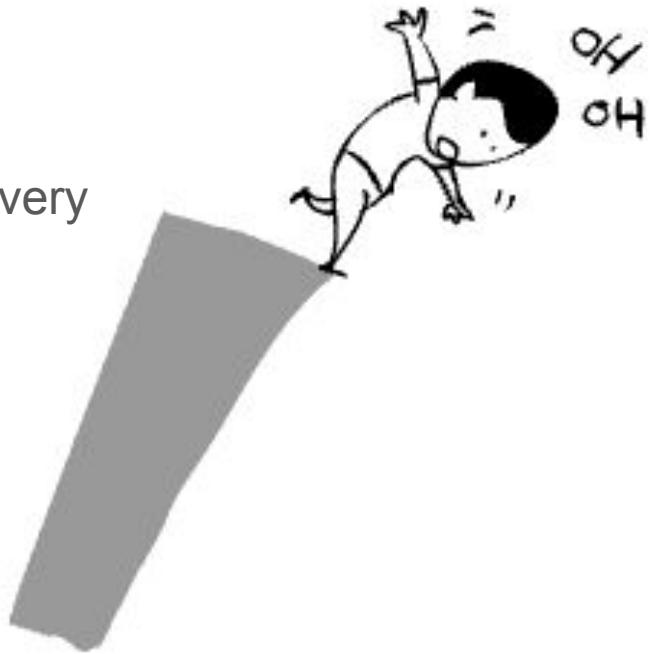
Group 13 - Final Project

Authors:
David Silva
Krisela Skenderi
Lukas Bernwald

Motivation

ECS - single point of failure of the database.

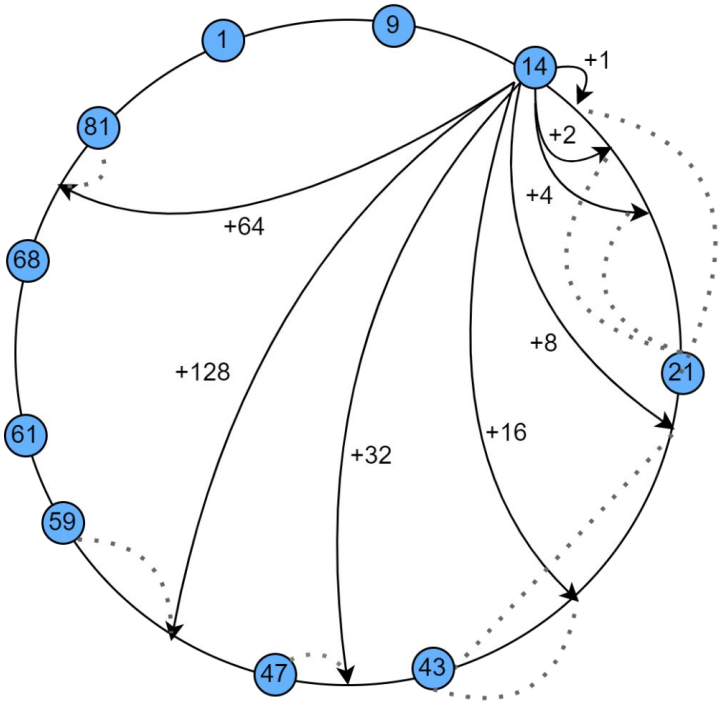
- Server Join
- Failure Detection (Heartbeat) and Recovery
- Metadata Update
- Replication
- Handoff



Why Chord?

- **Decentralized**, scalable Peer-to-Peer system
- **Load-Balancing**: Based on Consistent Hashing
- **Efficient Lookups**: O(log(n)) lookup complexity
- **Low maintenance**: Only a few nodes are affected on server join/departure
- **Easy adaptation** from the previous system

Chord overview - structure



node identifier

i^{th} entry corrsponds to successor($n + 2^{i-1} \mod 2^m$)

m = number of bits in key/node identifier

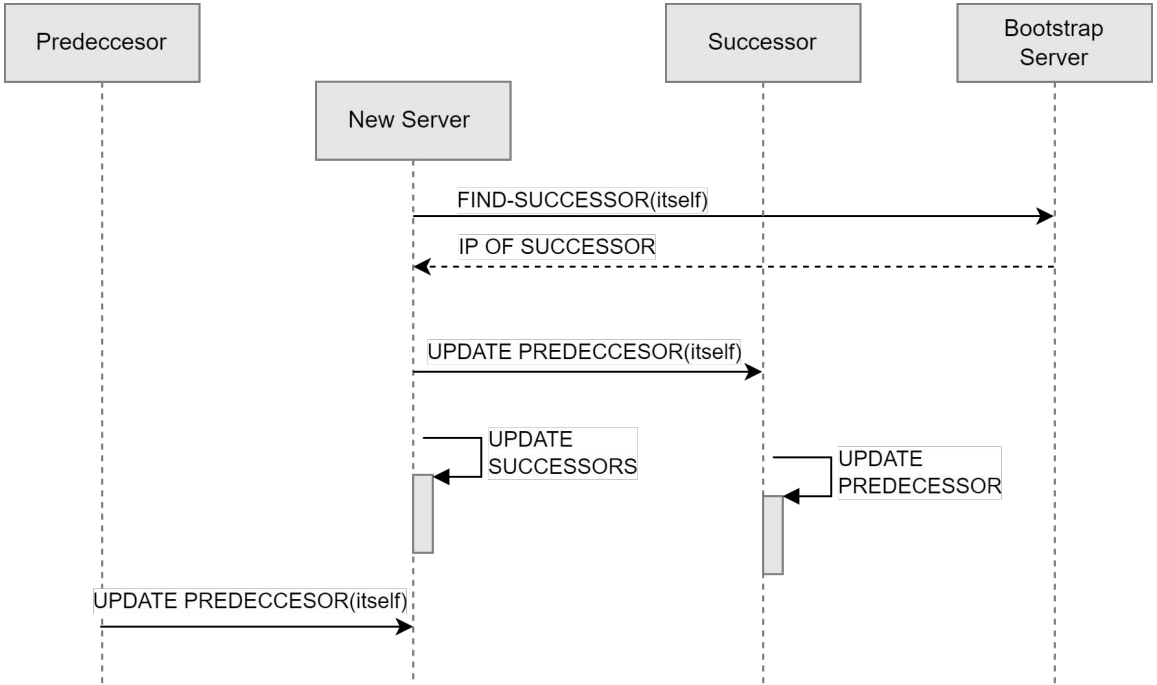
FINGER TABLE OF N14

N14 + 1	N21
N14 + 2	N21
N14 + 4	N21
N14 + 8	N43
N14 + 16	N43
N14 + 32	N47
N14 + 64	N81
N14 + 128	N59

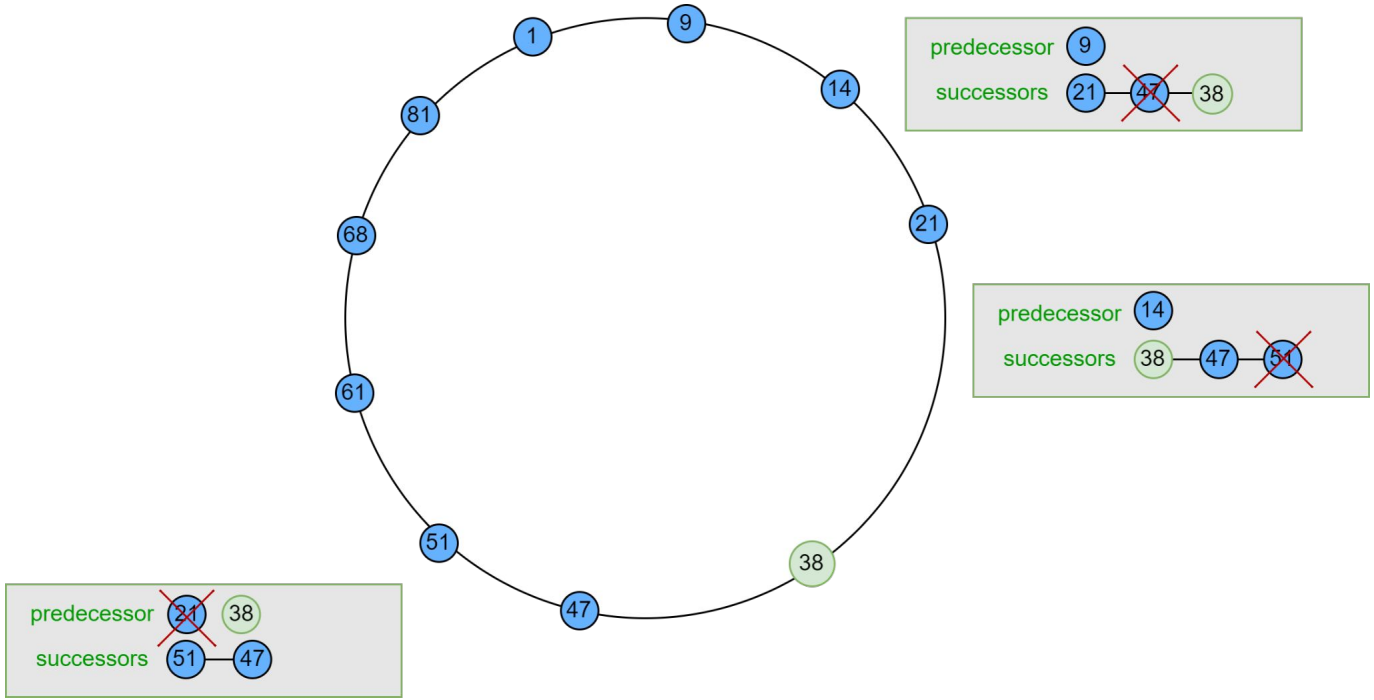
predecessor 9

successors 21 43 47

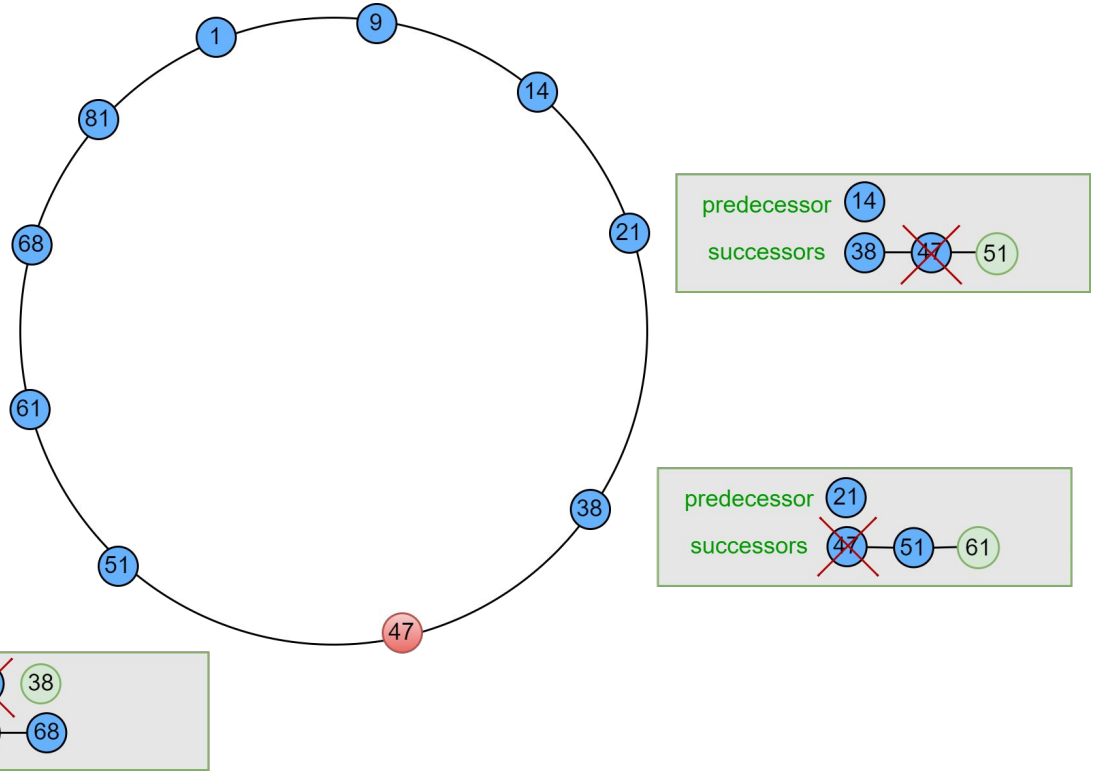
Chord overview - server join



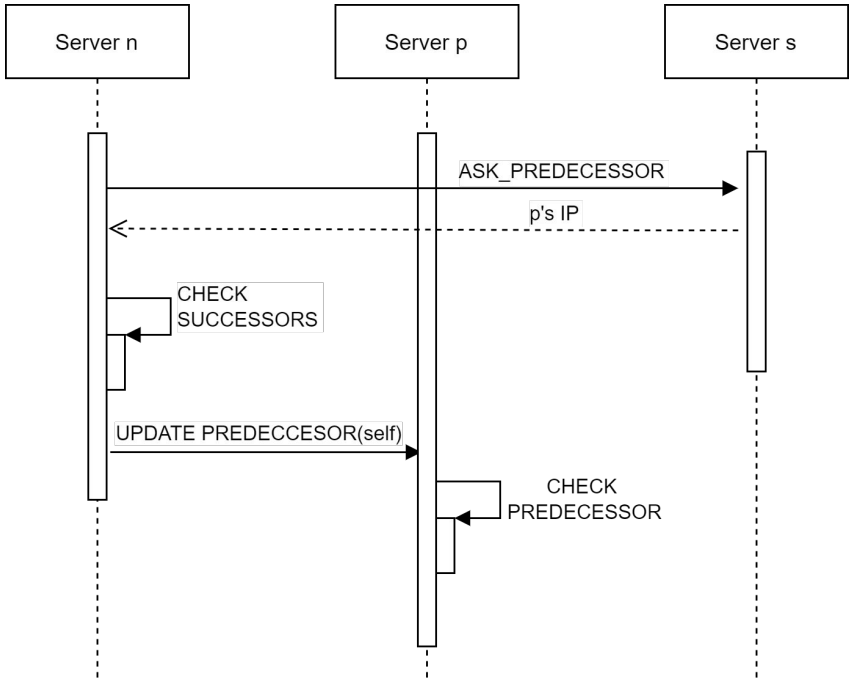
Maintaining the ring - server join



Maintaining the ring - server departure

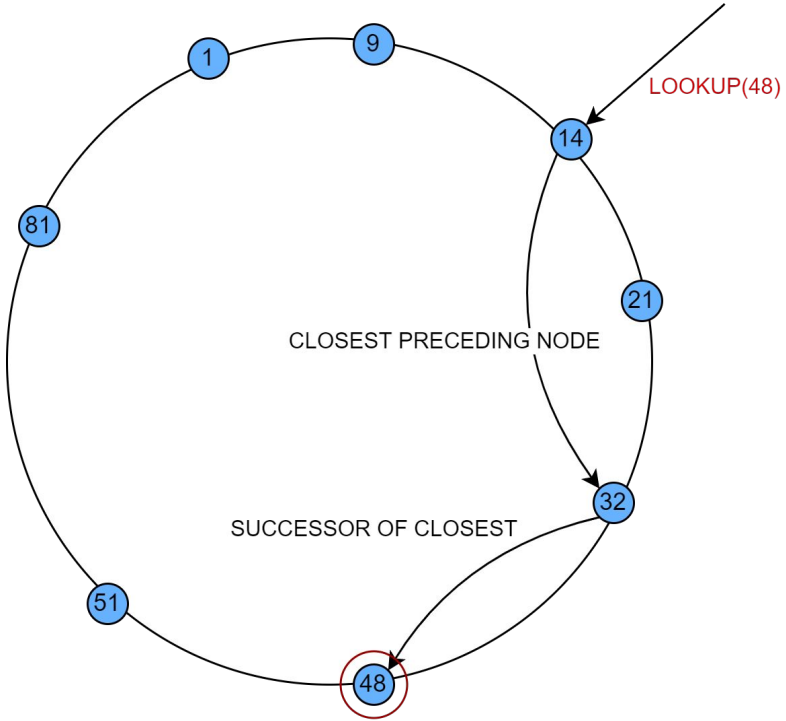


Maintaining the ring - stabilization



- ***p*** has joined between ***n*** and ***s***

Maintaining the ring - “fix fingers”



FINGER TABLE OF N14

N14 + 1	N21
N14 + 2	N21
N14 + 4	N21
N14 + 8	N32
N14 + 16	N32
N14 + 32	N51
N14 + 64	N81

N48

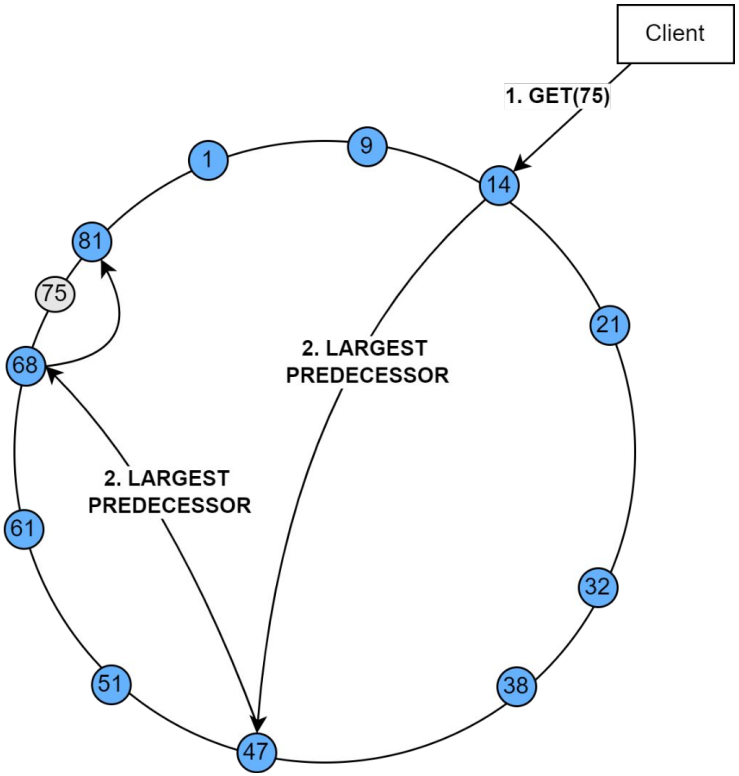
Client GET - lookup

FINGER TABLE OF N14

N14 + 1	N21
N14 + 2	N21
N14 + 4	N21
N14 + 8	N32
N14 + 16	N32
N14 + 32	**N47**
N14 + 64	N81

FINGER TABLE OF N47

N47 + 1	N51
N47 + 2	N51
N47 + 4	N51
N47 + 8	N61
N47+ 16	**N68**
N47 + 32	N81
N47 + 64	N21



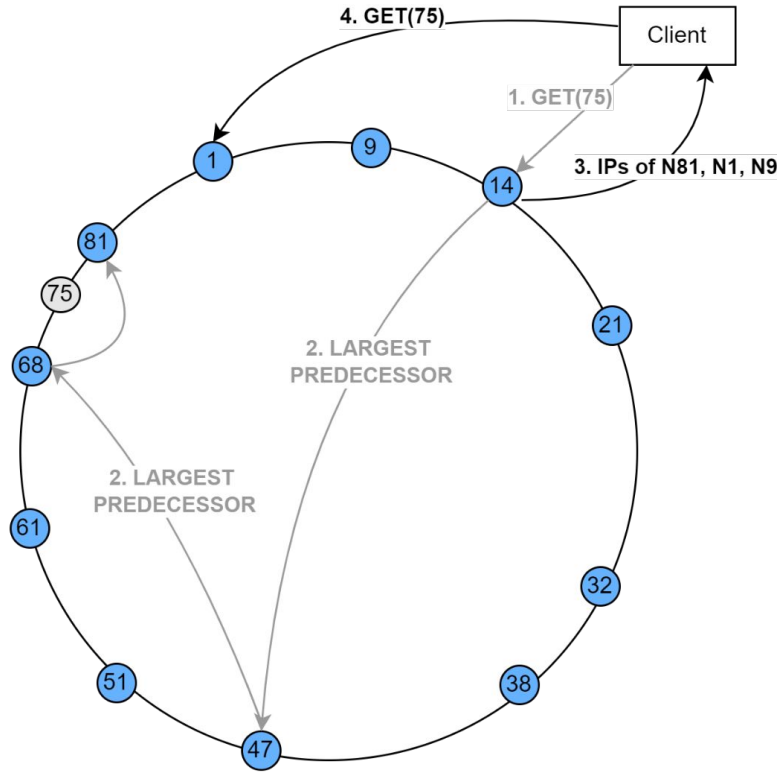
Client GET - execute

FINGER TABLE OF N14

N14 + 1	N21
N14 + 2	N21
N14 + 4	N21
N14 + 8	N32
N14 + 16	N32
N14 + 32	**N47**
N14 + 64	N81

FINGER TABLE OF N47

N47 + 1	N51
N47 + 2	N51
N47 + 4	N51
N47 + 8	N61
N47+ 16	**N68**
N47 + 32	N81
N47 + 64	N21



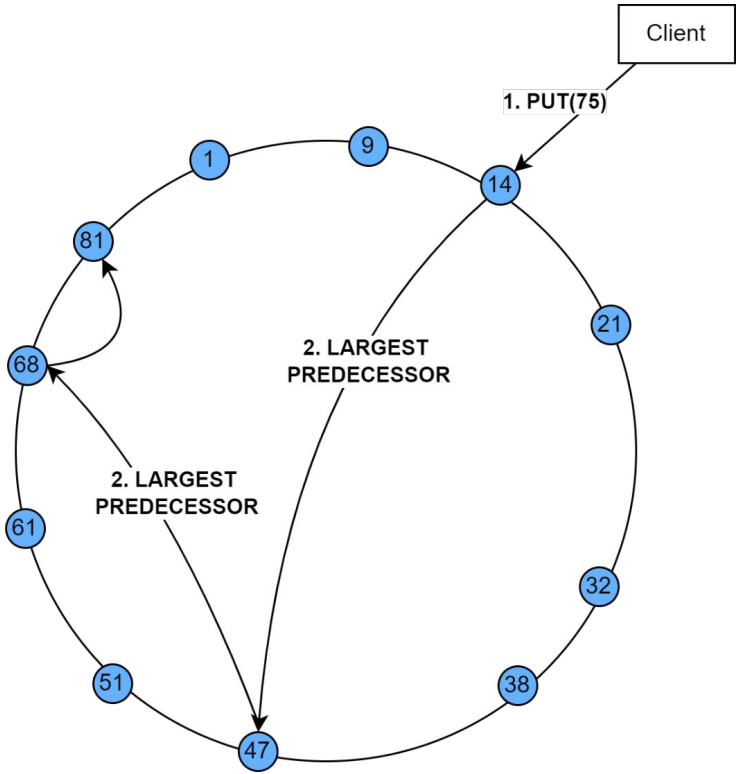
Client PUT - lookup

FINGER TABLE OF N14

N14 + 1	N21
N14 + 2	N21
N14 + 4	N21
N14 + 8	N32
N14 + 16	N32
N14 + 32	**N47**

FINGER TABLE OF N47

N47 + 1	N51
N47 + 2	N51
N47 + 4	N51
N47 + 8	N61
N47+ 16	**N68**
N47 + 32	N81



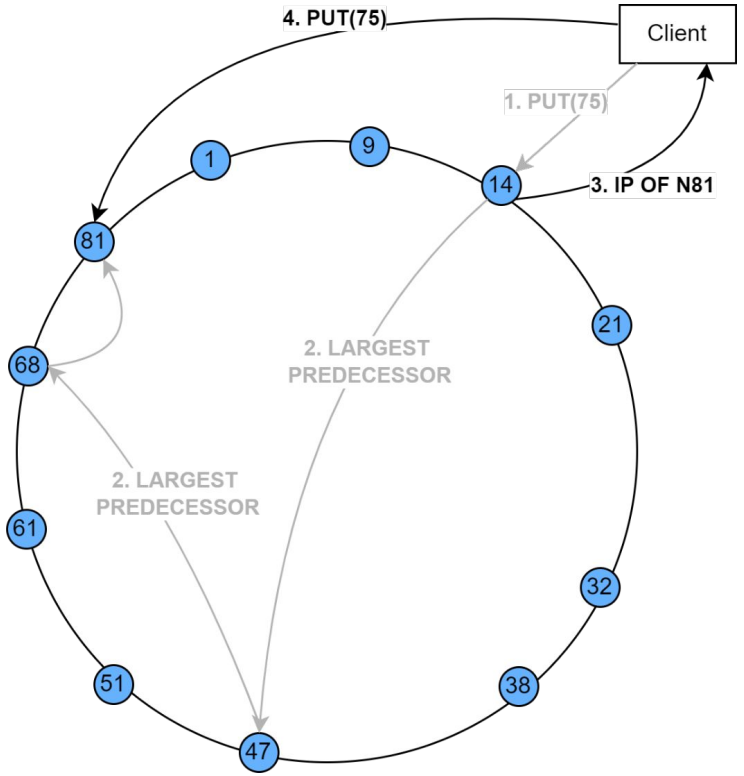
Client PUT - execute

FINGER TABLE OF N14

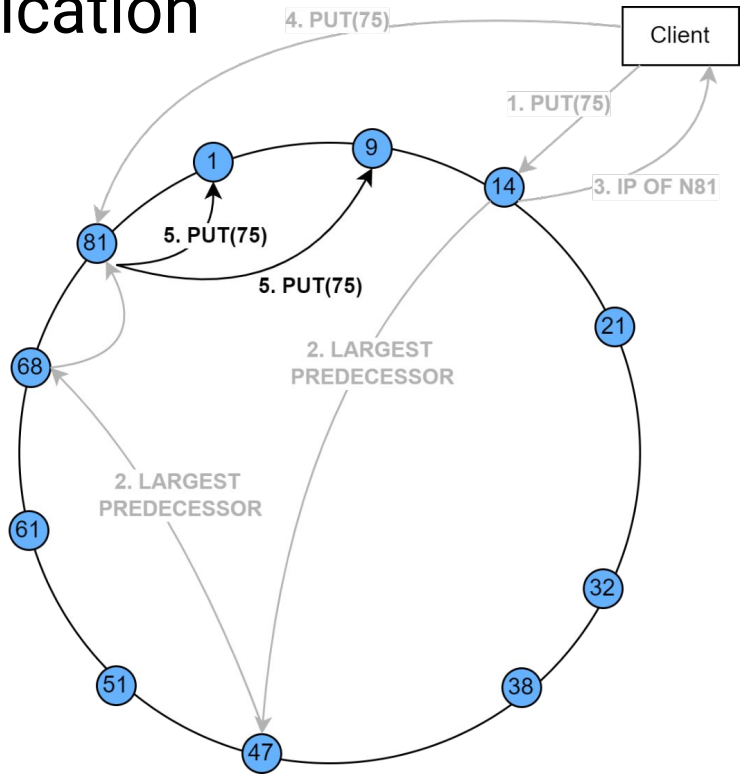
N14 + 1	N21
N14 + 2	N21
N14 + 4	N21
N14 + 8	N32
N14 + 16	N32
N14 + 32	**N47**

FINGER TABLE OF N47

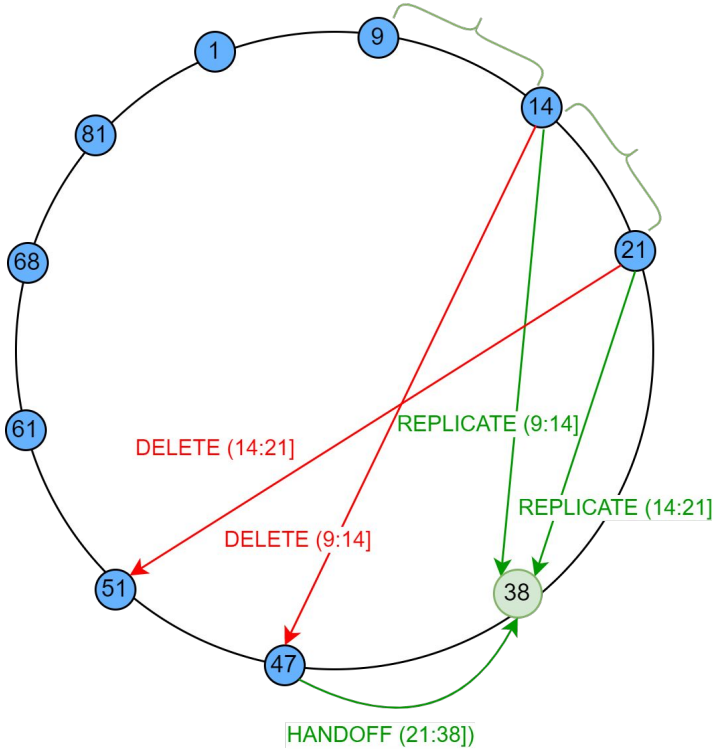
N47 + 1	N51
N47 + 2	N51
N47 + 4	N51
N47 + 8	N61
N47+ 16	**N68**
N47 + 32	N81



Client PUT - replication

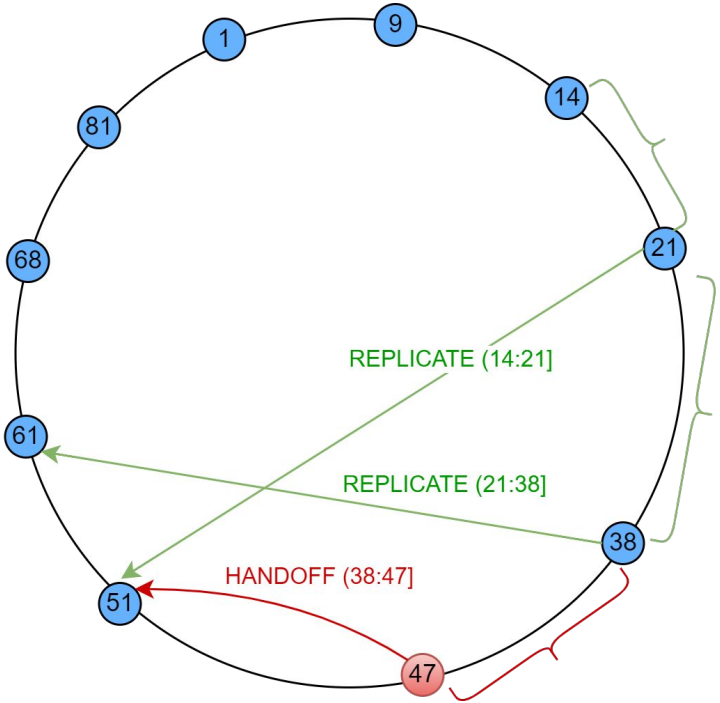


Server join - replication and handoff



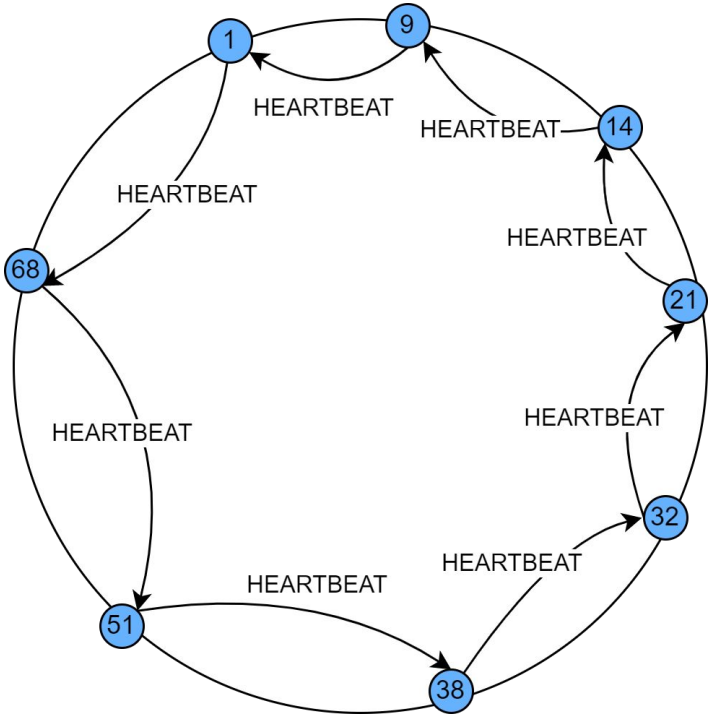
- Replication factor 2
- Server 38 joins
- Handoff keys in range (21,38]
- Replicate to new node
- Delete stale replication data

Server departure - replication and handoff



- Replication factor 2
- Server 47 leaves
- Handoff from [$pred_{47,}$ 47] to 51
- Update replication

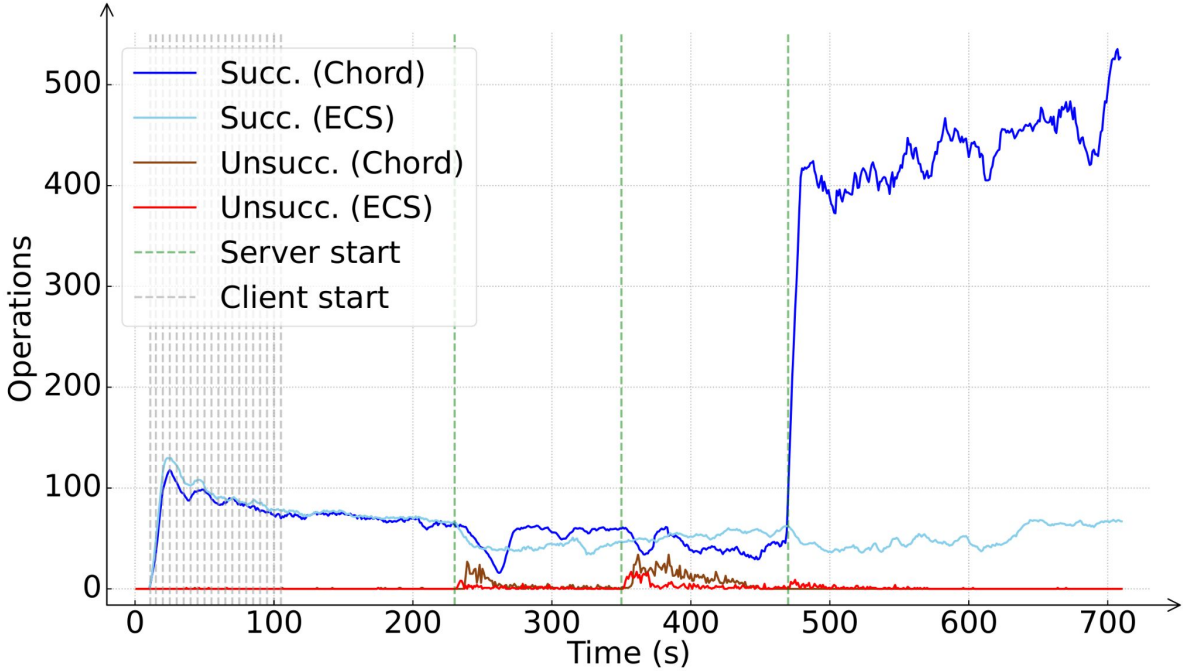
Failure detection - check predecessor



Experimental evaluation - overview

- Tool schedules timed events
  - Starts clients and servers
- Clients put, get and delete
- Dataset: Enron email
- 2 Experiments with ECS and Chord
  - LFU cache (500 size)
  - B-Tree minimum degree = 200
- Measure throughput and latency

Experiment 1 - Behavior comparison



- 3 → 6 servers
- Replication factor = 1
- 0 → 20 clients

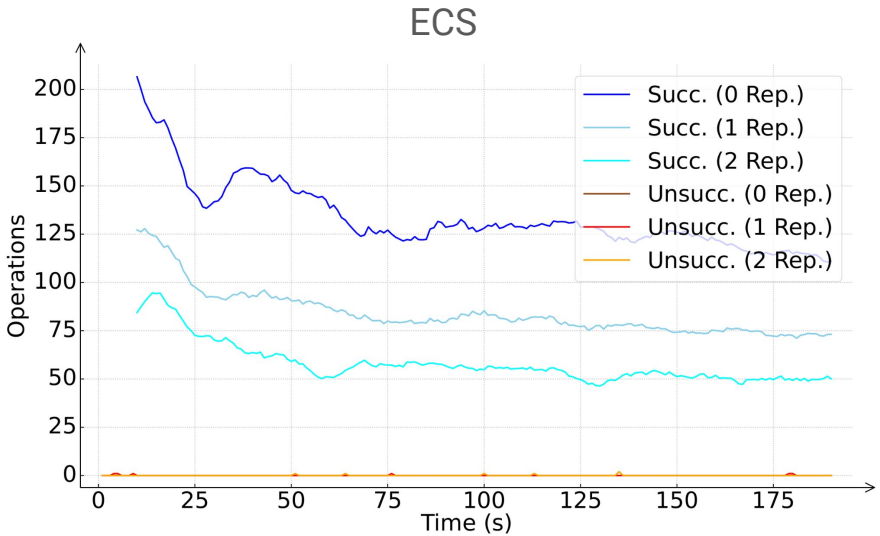
Experiment 1 - Behavior comparison

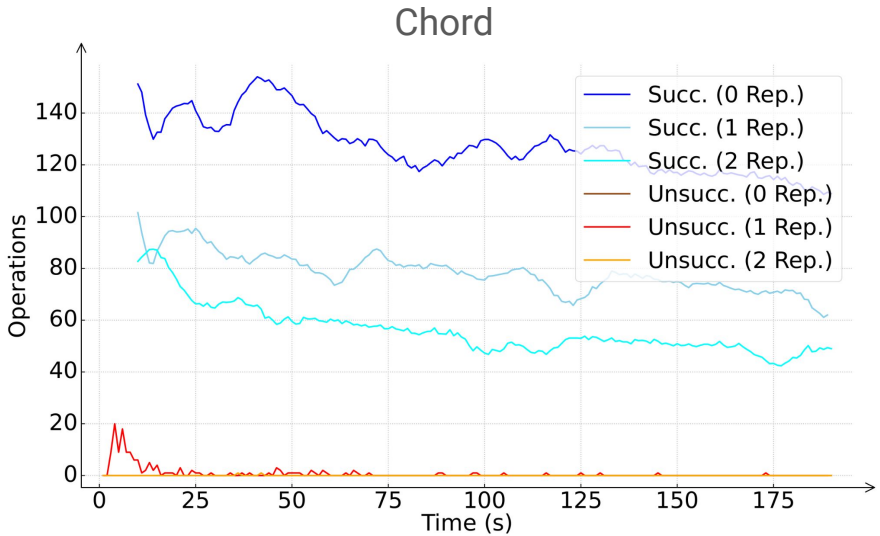
- Request latency (in *ms*)

ECS			Chord		
GET	PUT	DEL	GET	PUT	DEL
272.2	329.9	359.8	30.7	266.3	268.8

Experiment 2 - Replication performance

- 6 Servers (same settings), 20 Clients, variable replication factor





Experiment 2 - Replication performance

- Request latency (in *ms*)

	ECS			Chord		
Repl. factor	GET	PUT	DEL	GET	PUT	DEL
0	133.3	159.4	167.9	143.8	166.2	174.8
1	205.9	260.5	275.2	217.4	285.9	281.2
2	326.2	351.8	379.5	332.5	369.2	388.6

Experimental evaluation - remarks

- Network latency and data distribution
- Eventual consistency
- Complexity of lookup operation: O(log(n)) compared to O(1)
- Hardware limitations
- Non-deterministic behavior

Conclusion

- More complex than ECS
  - Increased network communication
  - Eventual consistency
- Performs similarly to potentially better
- No Single Point of Failure
- Potentially higher scalability
- Further research needed

Replacing External Configuration Service with Chord Peer-to-Peer Protocol

Cloud Data Bases WS2021-22

Final Project

Authors:
David Silva
Krisela Skenderi
Lukas Bernwald

Questions?