



# **DSC PROJECT**

# **CASSANDRA DB**

**TEAM 12**

**MOHAMMED ANSAR AHMED**

**TAY JIA JIUN JOSHUA**

**KOH JIN RUI VINNY**

**VINCENTIUS ROGER**

**NATHAN ALDRICH WIRYAWAN**

# CONTENT

- 01** INTRODUCTION OF PROBLEM
- 02** OVERALL SYSTEM ARCHITECTURE
- 03** FEATURES IMPLEMENTED
- 04** SCALABILITY
- 05** CORRECTNESS
- 06** FAULT TOLERANCE
- 07** EXPERIMENTS / TESTING
- 08** LIMITATIONS

# INTRODUCTION

APACHE CASSANDRA IS A HIGHLY SCALABLE, DISTRIBUTED DATABASE DESIGNED FOR APPLICATIONS REQUIRING HIGH WRITE THROUGHPUT, FAULT TOLERANCE, AND LOW-LATENCY ACCESS ACROSS LARGE DATASETS.

MAKES IT IDEAL FOR TIME-SERIES DATA, SUCH AS WEB ANALYTICS.

STRATEGY N°1

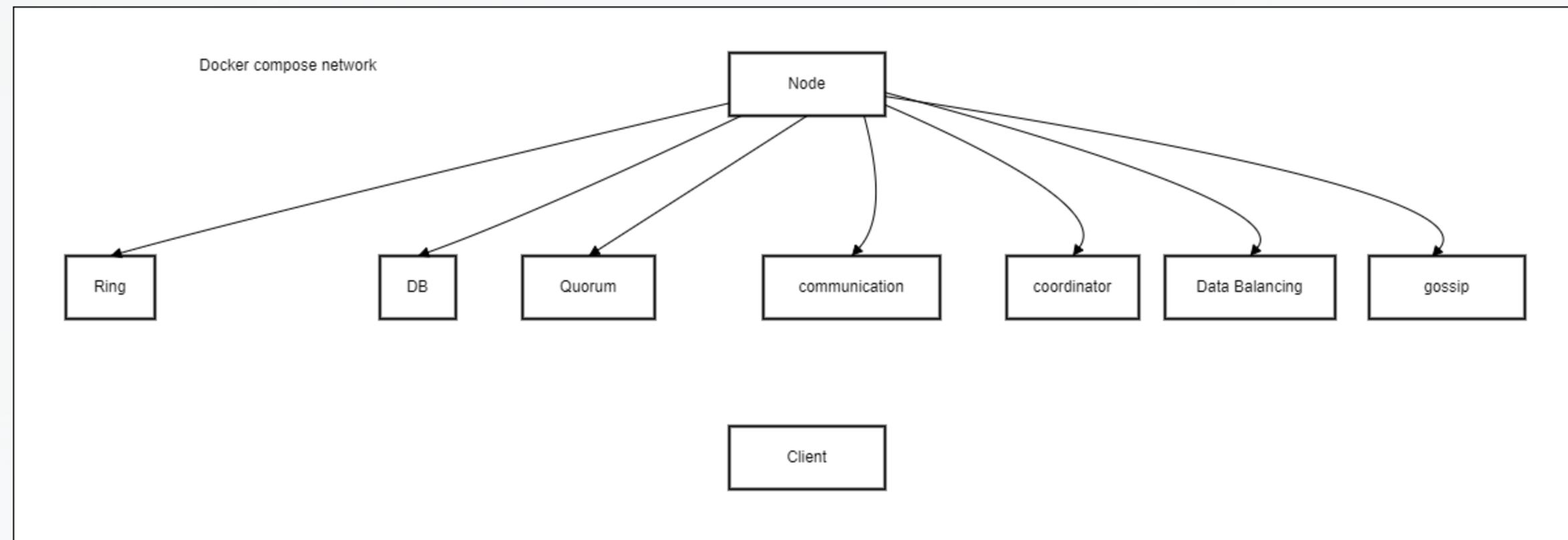
STRATEGY N°2

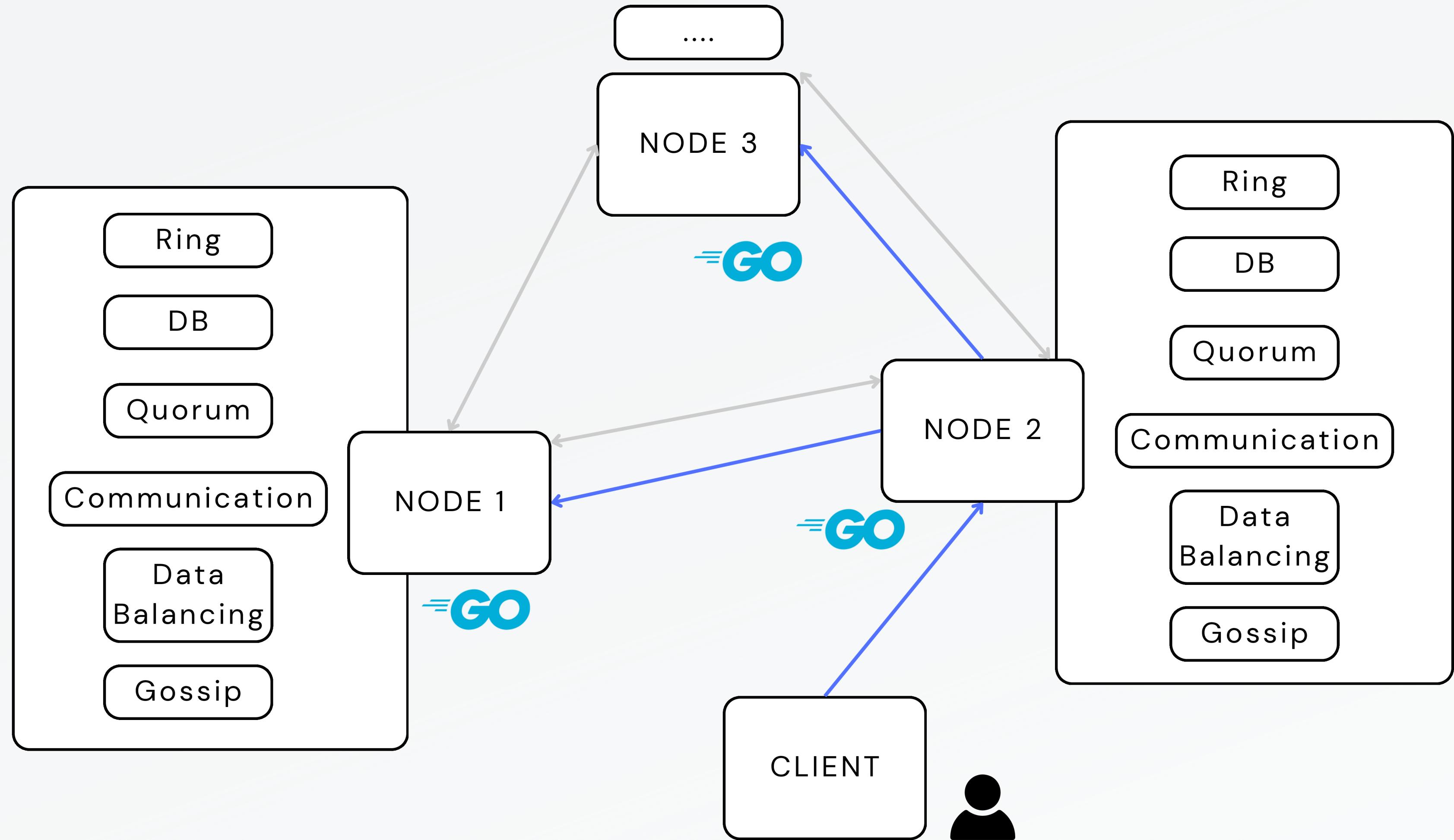
STRATEGY N°3

# SYSTEM ARCHITECTURE

OUR PROJECT HAS 2 TYPES OF NODES:

- 1.CLIENT/WEB ->SEND READ/WRITE TO NODES
- 2.CASSANDRA NODES->THE NODE CONTACTED WILL ACT AS COORDINATOR AND COMMUNICATE WITH OTHER NODES





# FEATURES IMPLEMENTED

- READ/WRITE WORKFLOW
- CONSISTENT HASHING
- QUORUM
- GOSSIP PROTOCOL
- DATA REBALANCING
- READ REPAIR
- FRONT END VISUALIZATION FOR RING
- FRONTEND CLIENT (SIMULATE THE WEB)

# SCALABILITY

- HIGH WRITE AVAILABILITY
- SCALING NUMBER OF NODES
  - CONSISTENT HASHING
  - GOSSIP
  - DATA REPPLICATION

# CORRECTNESS

- EVENTUAL CONSISTENCY
- READ QUORUM - ENSURES THAT A MAJORITY OF NODES RETURN MATCHING RESPONSES
- WRITE QUORUM - ENSURES THAT A MAJORITY OF NODES HAVE COMMITTED THE WRITE
- READ REPAIR - IF A NODE RETURNS A STALE READ DURING READ QUORUM, IT WILL BE SENT THE UPDATED DATA

READ REPAIR ALSO ENSURES THAT TEMPORARY FAULTS THAT RECOVER WITHIN THE GOSSIP TIME OUT MIGHT CAUSE SOME DATA TO NOT BE REPLICATED CORRECTLY.

# FAULT TOLERANCE

- GOSSIP PROTOCOL TO DETECT NODES ARE DOWN
- DATA REBALANCING -> ADD NEW NODE WHEN GOSSIP DETECTS A NODE IS DOWN
- REPLICAS
- MAXIMUM CONCURRENT SYSTEM FAULT TOLERANCE:  $R - 1$  CONCURRENT FAILURES
- MAXIMUM NON-CONCURRENT FAILURES:  $N - R$  FAILURES

## SOLUTION:

- TYPICAL CONVENTION OF HOSTING NODES IN DIFFERENT DATA CENTER'S OR RACKS WITHIN DATA CENTERS CAN REDUCE THE POSSIBILITY OF HITTING THE MAX CONCURRENT FAILURE LIMITS

# TESTING

- UNIT AND SYSTEM TESTING

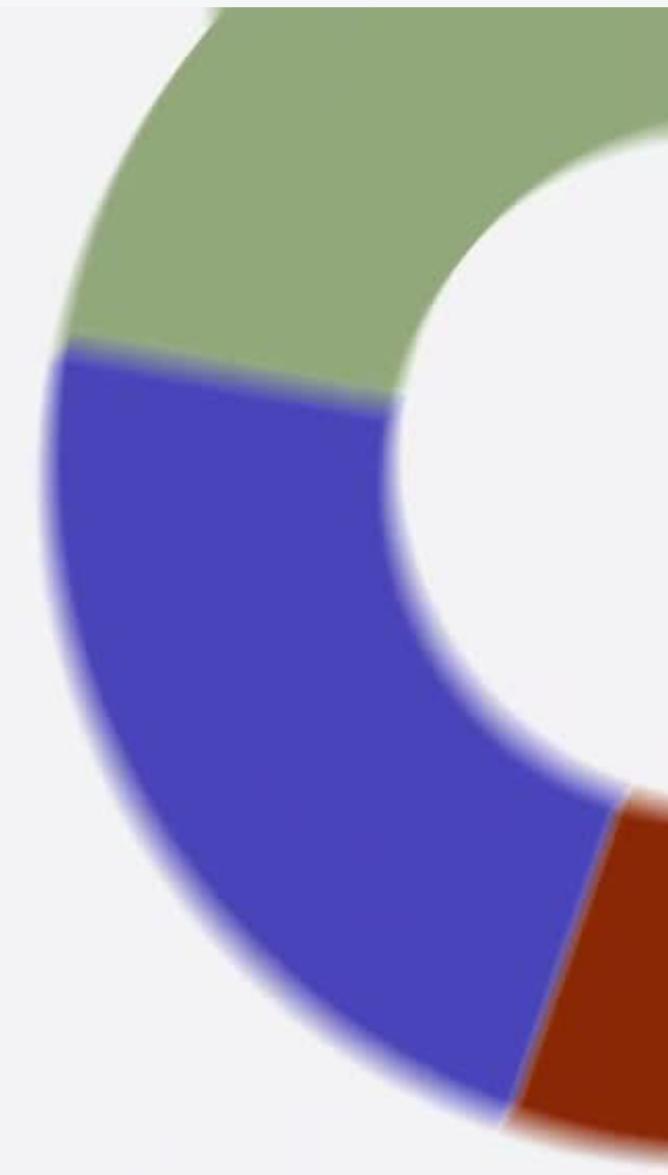
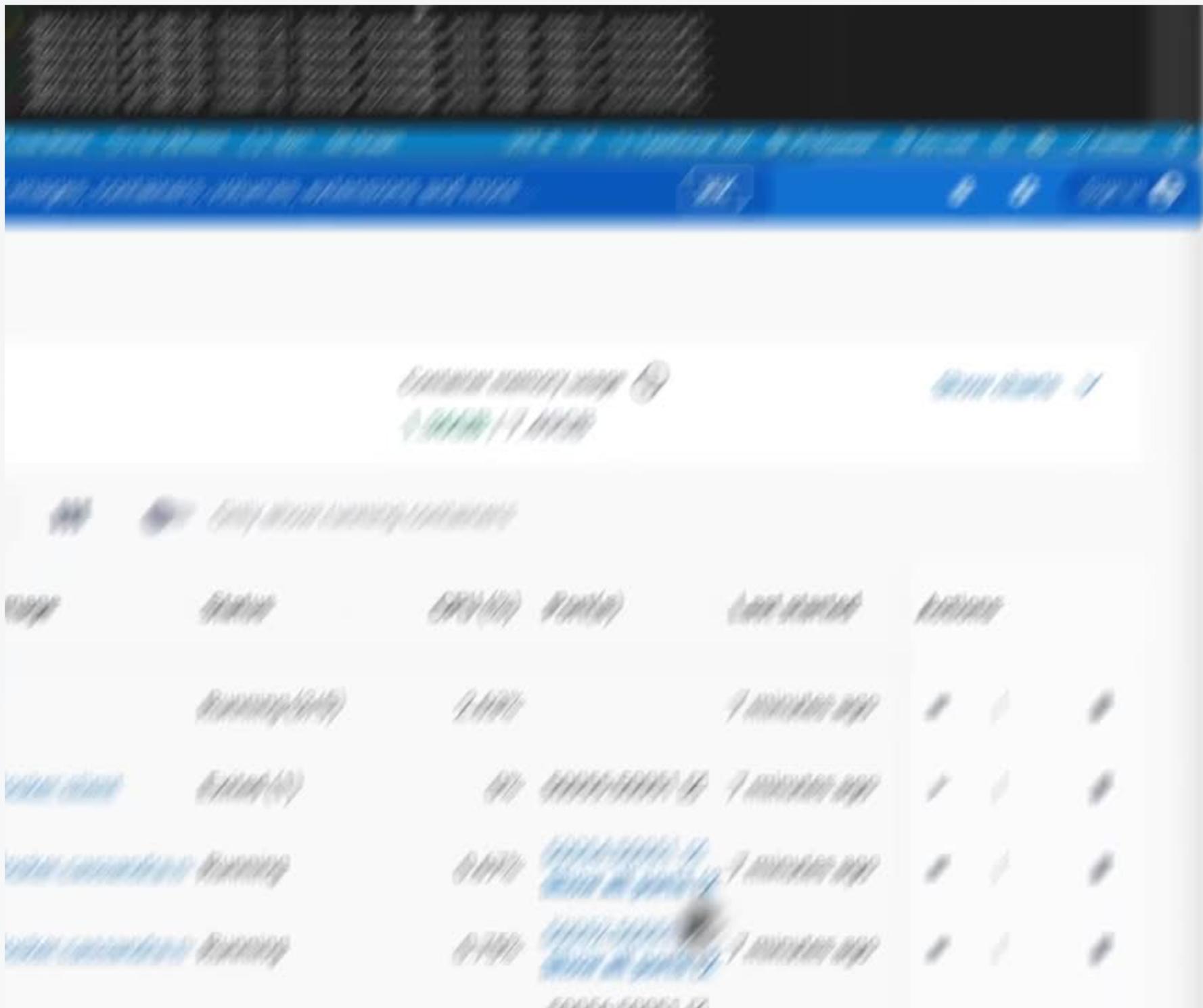
STRATEGY N°1

STRATEGY N°2

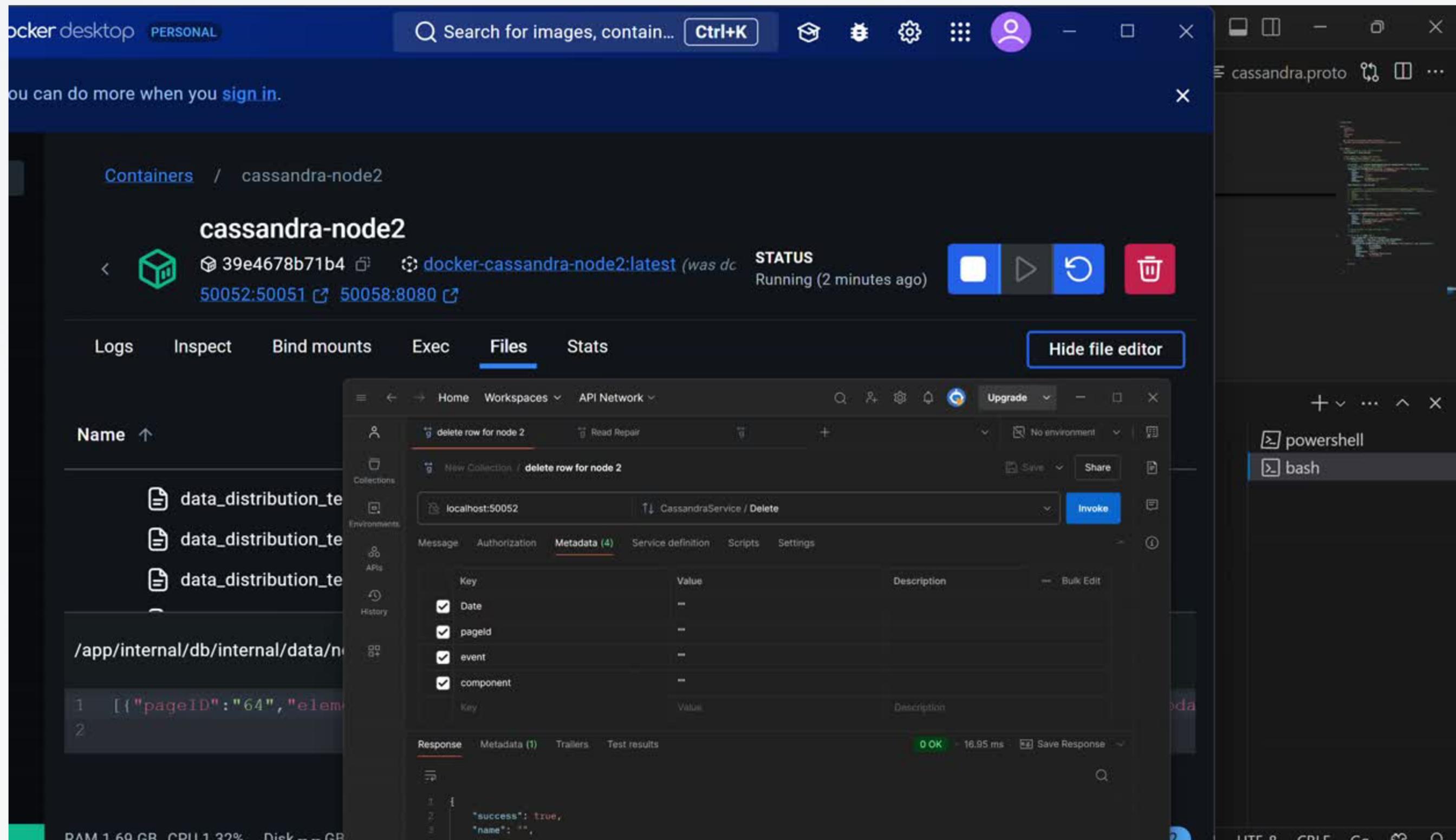
STRATEGY N°3

# **TESTING - DEMO**

# RING VISUALISER DEMO



# READ-REPAIR DEMO



DELETE ALL DATA IN NODE 2, UPON READING PG 1, NODE 2'S PG1 ROWS ARE RESTORED FROM READ REPAIR

# LIMITATIONS

- UNABLE TO MODIFY THE DATA STRUCTURE? CURRENTLY USES HARDCODED PARTITION KEY OF "PAGE ID" DUE TO TIME CONSTRAINTS
- INSUFFICIENT TIME TO IMPLEMENT HANDLING SOME TYPES OF BYZANTINE FAULTS (EG SENDING MALICIOUS MESSAGE)
- INSUFFICIENT TIME TO IMPLEMENT HINTED HANDOFF (WRITE REPAIR)

# THANK YOU

STRATEGY N°1

STRATEGY N°2

STRATEGY N°3