



The RDBMS for Next-Gen Networks and Services

The Clustra Zero-Downtime Database for the 24x7 World

Your name

Your title

Your e-mail

Your phone number

- Introductions
- The Clustra Database — how it works
- Next steps

Meeting Objectives

- Present the Clustra Database
- Explain our architecture — how we do it
- Resolve key issues
- Set up product evaluation

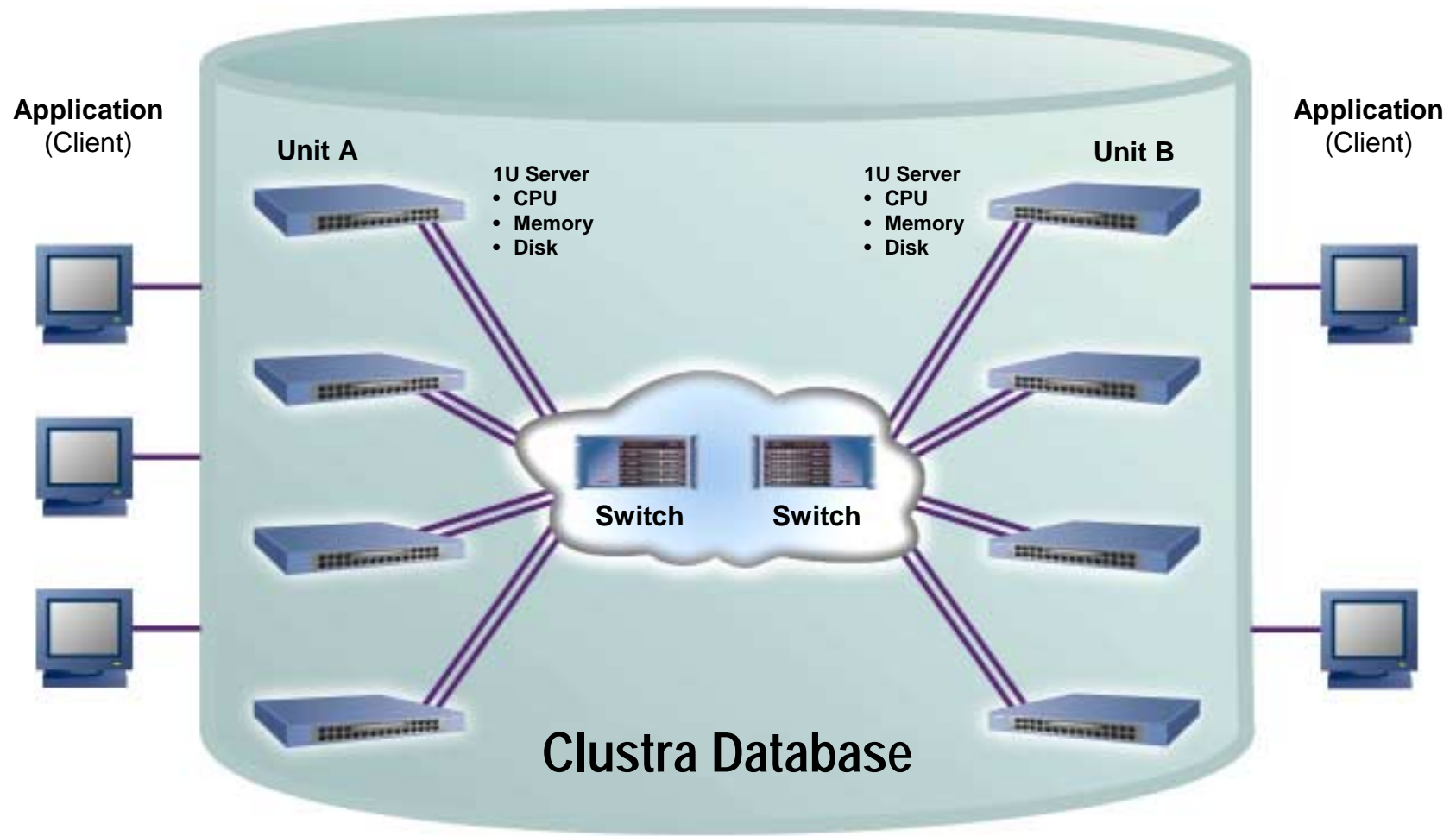


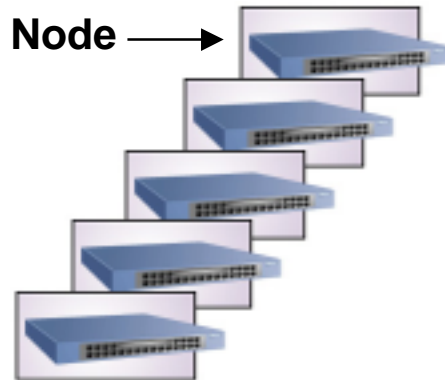
Technical Overview

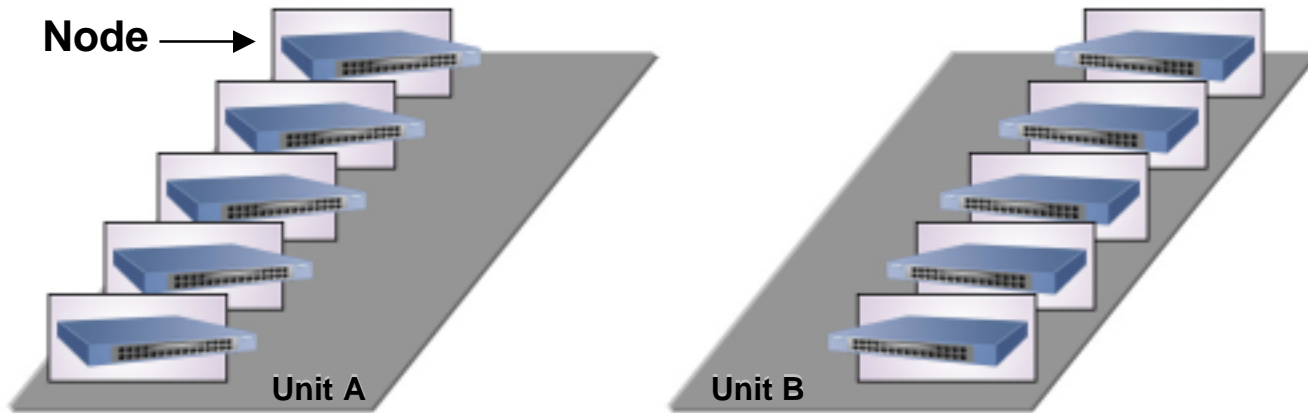
Product Architecture

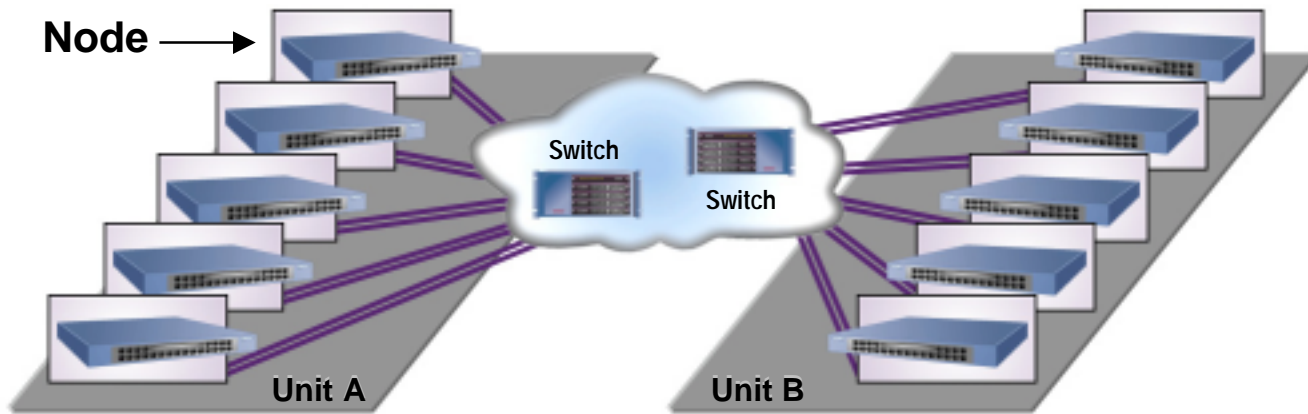
- Industry-standard interfaces
SQL92, ODBC, JDBC and real-time API
- Continuous availability
Max 2 min unavailability per year (better than Class 5)
- Linear scaling transaction rates
- Real-time response times
Max 5 ms for 95% of the transactions (TCP-B like)
- Commodity hardware and open software
Workstations / PC, UNIX, Ethernet / ATM / FDDI

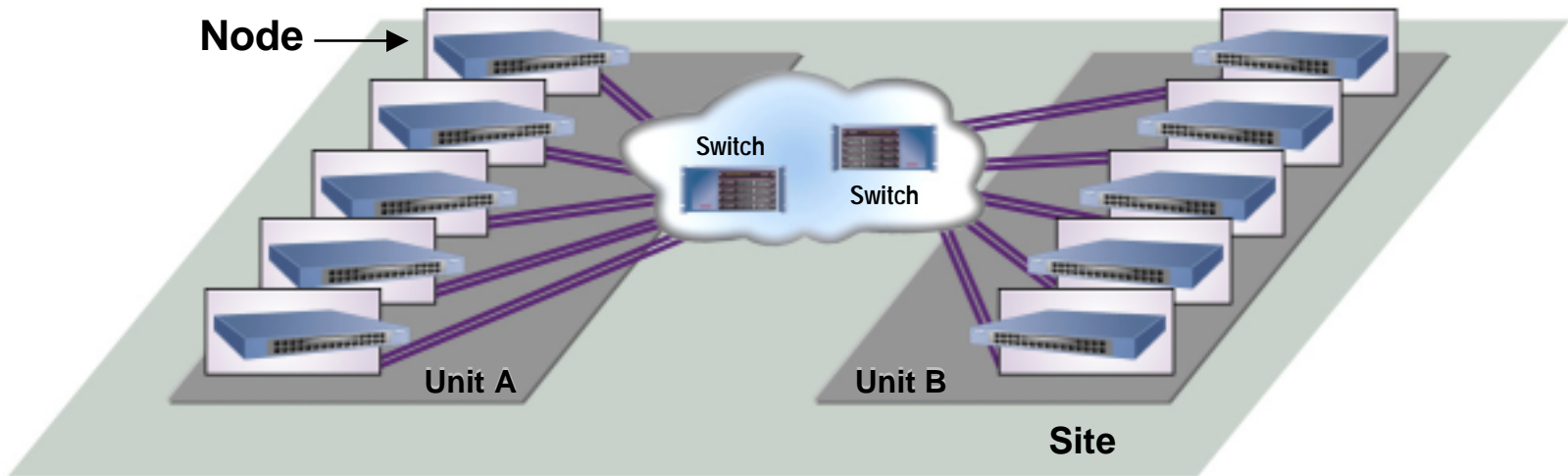
How it Works





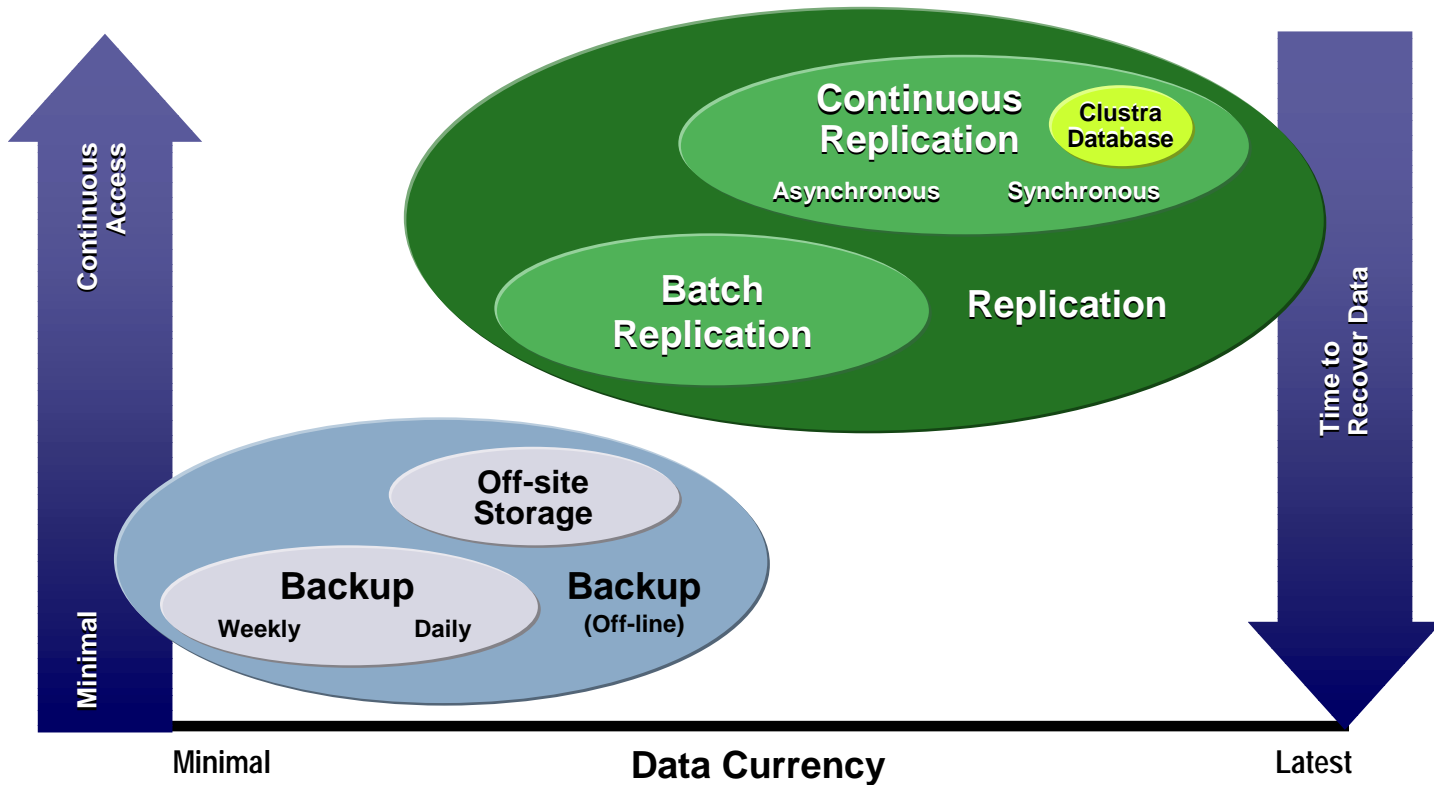




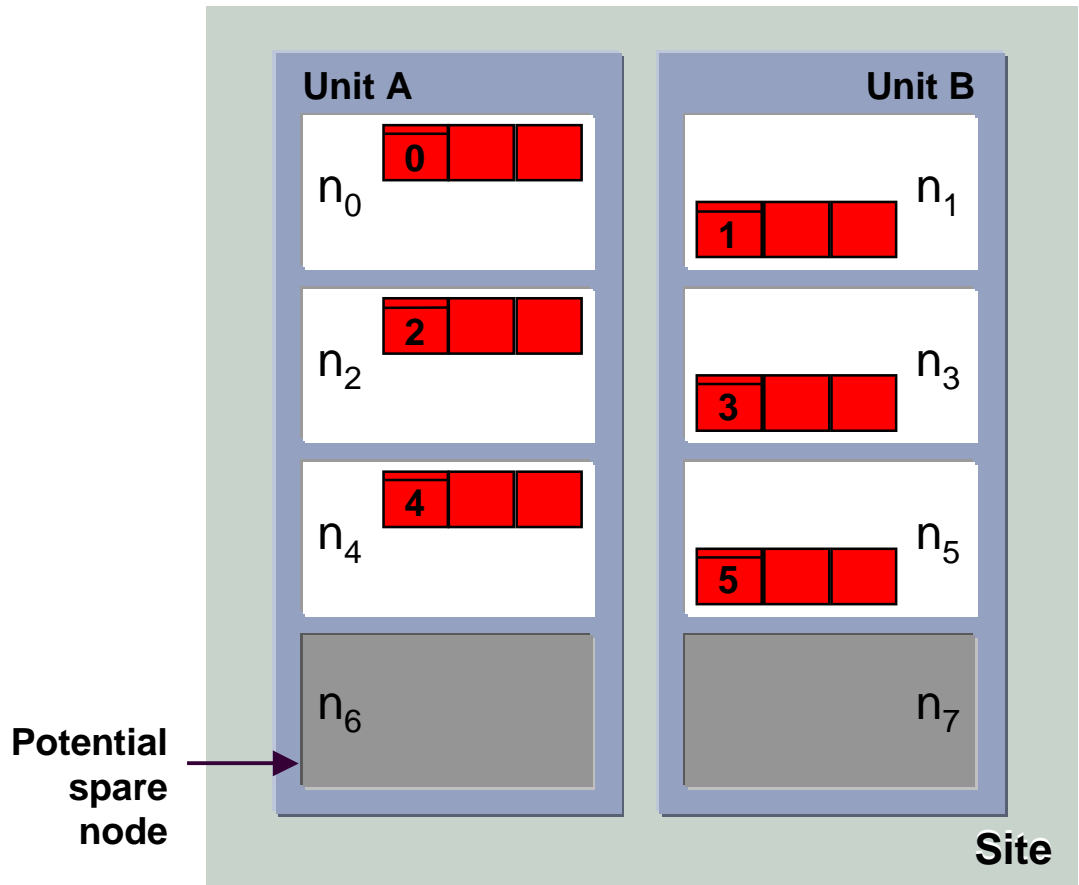
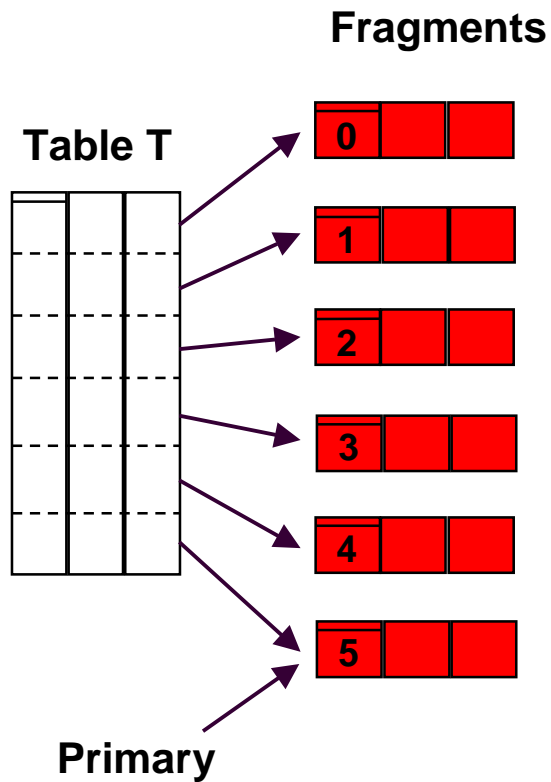


Continuous Replication

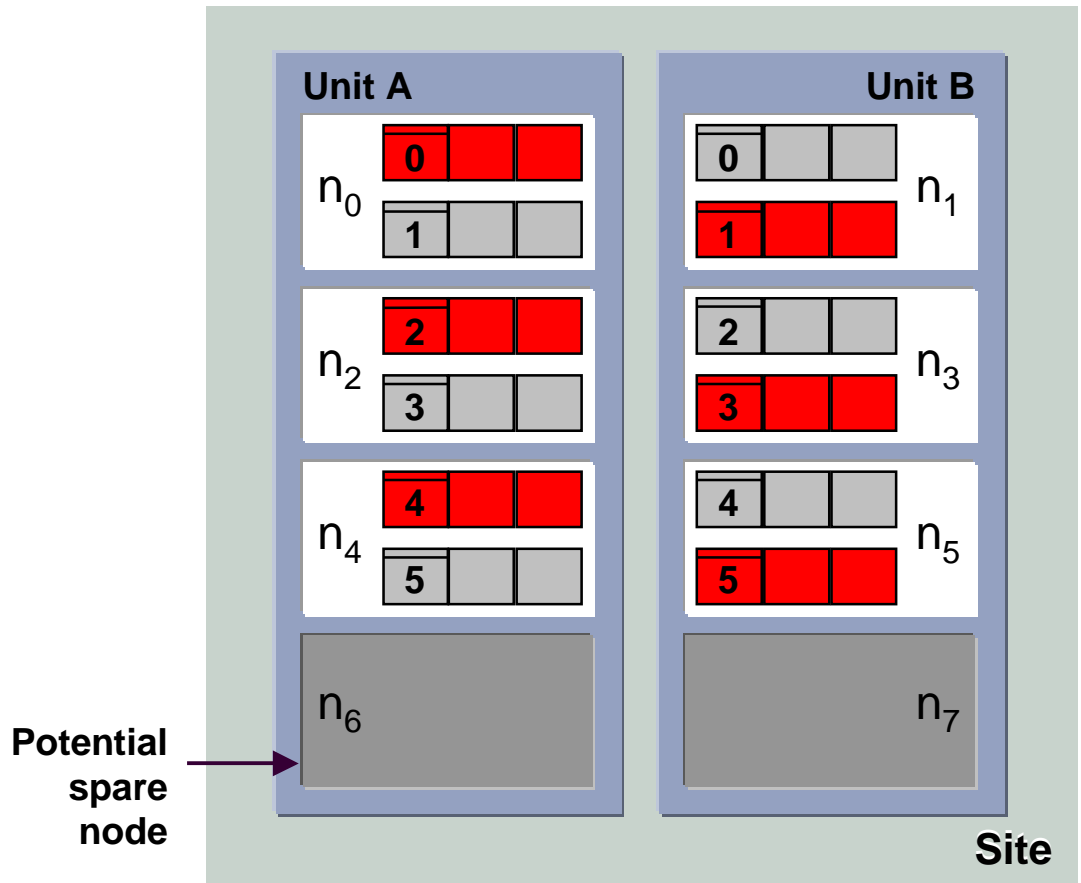
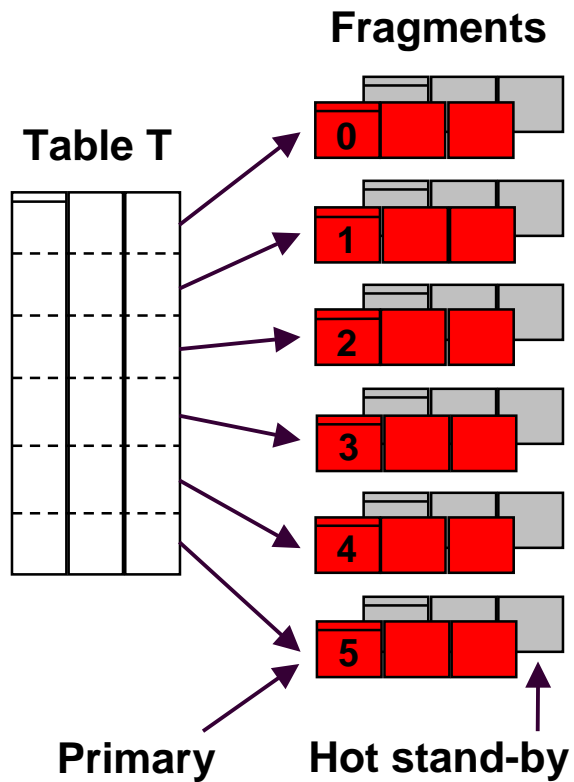
Databases



Data Distribution & Replication Scheme

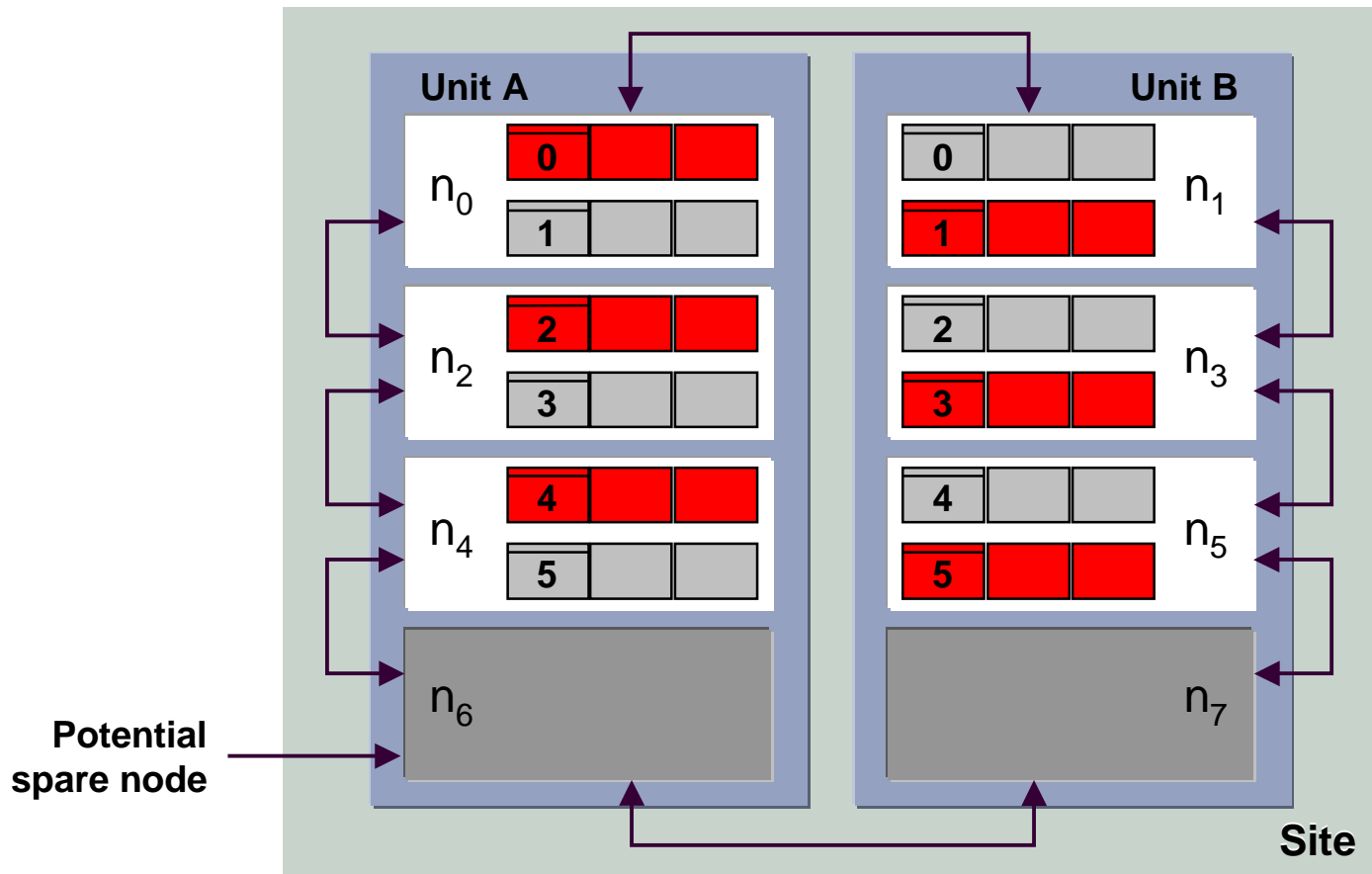


Data Distribution & Replication Scheme



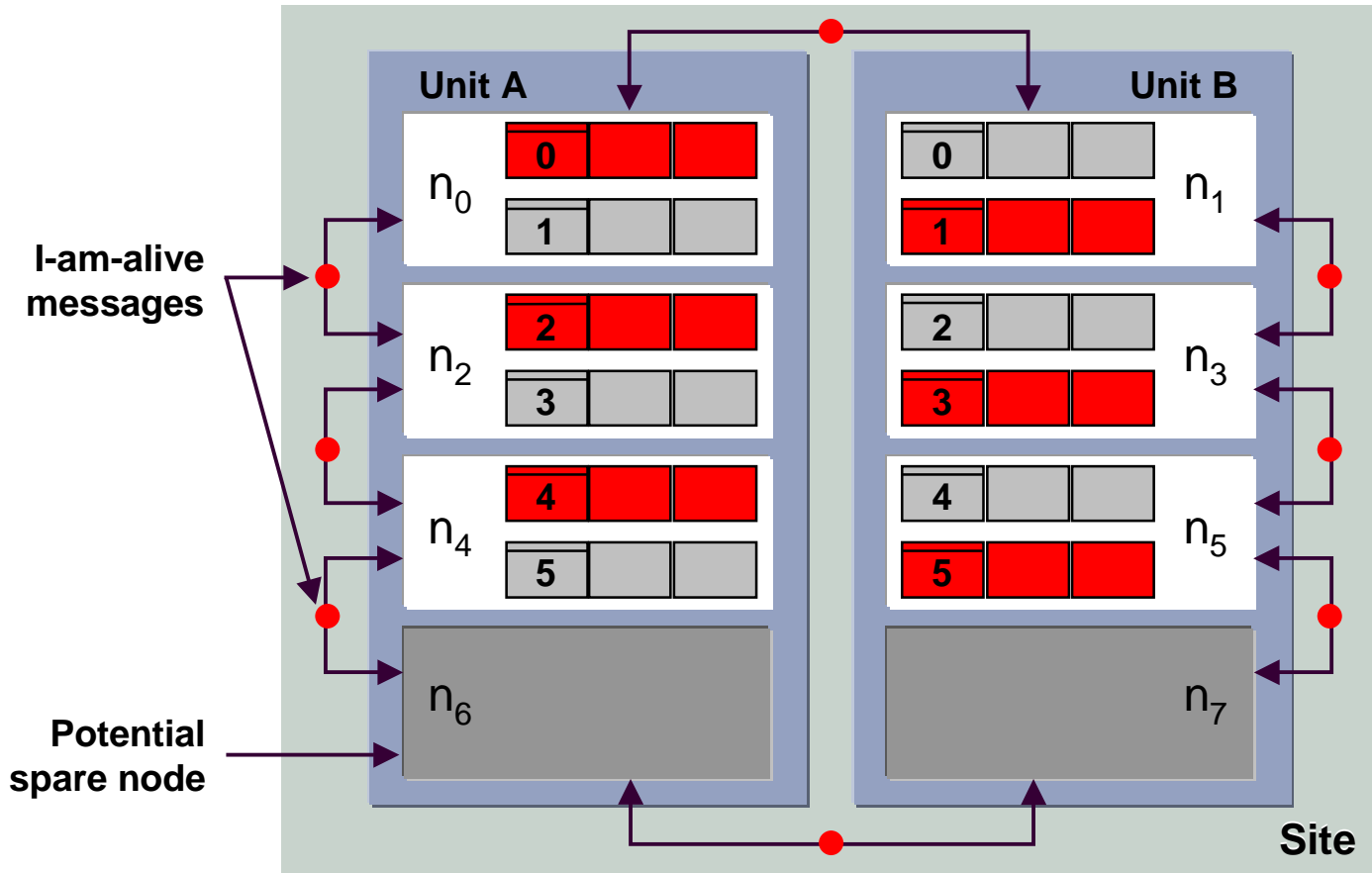
Failure Handling:

Logical Neighbors



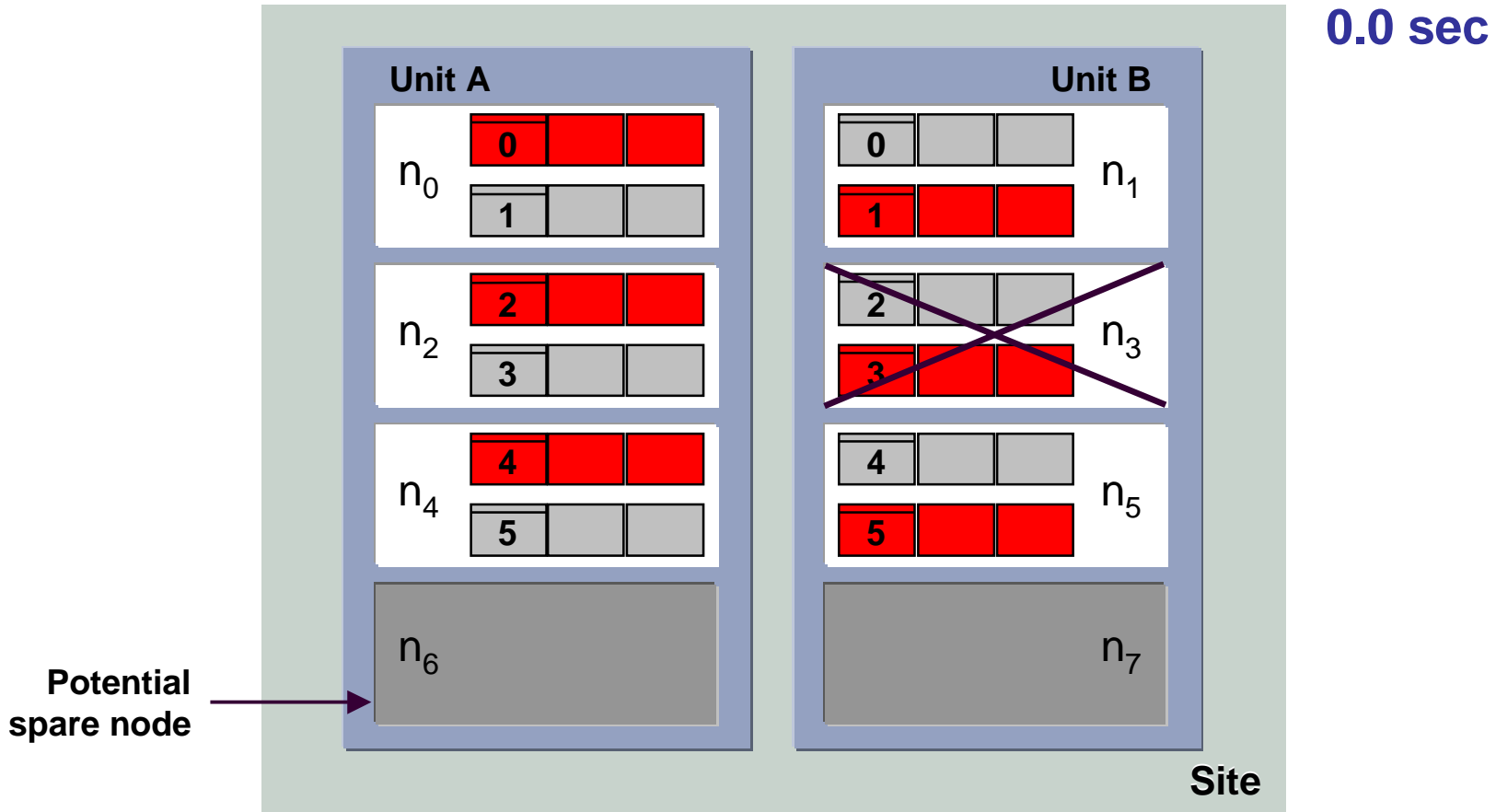
Failure Handling:

Heart Beat Monitoring



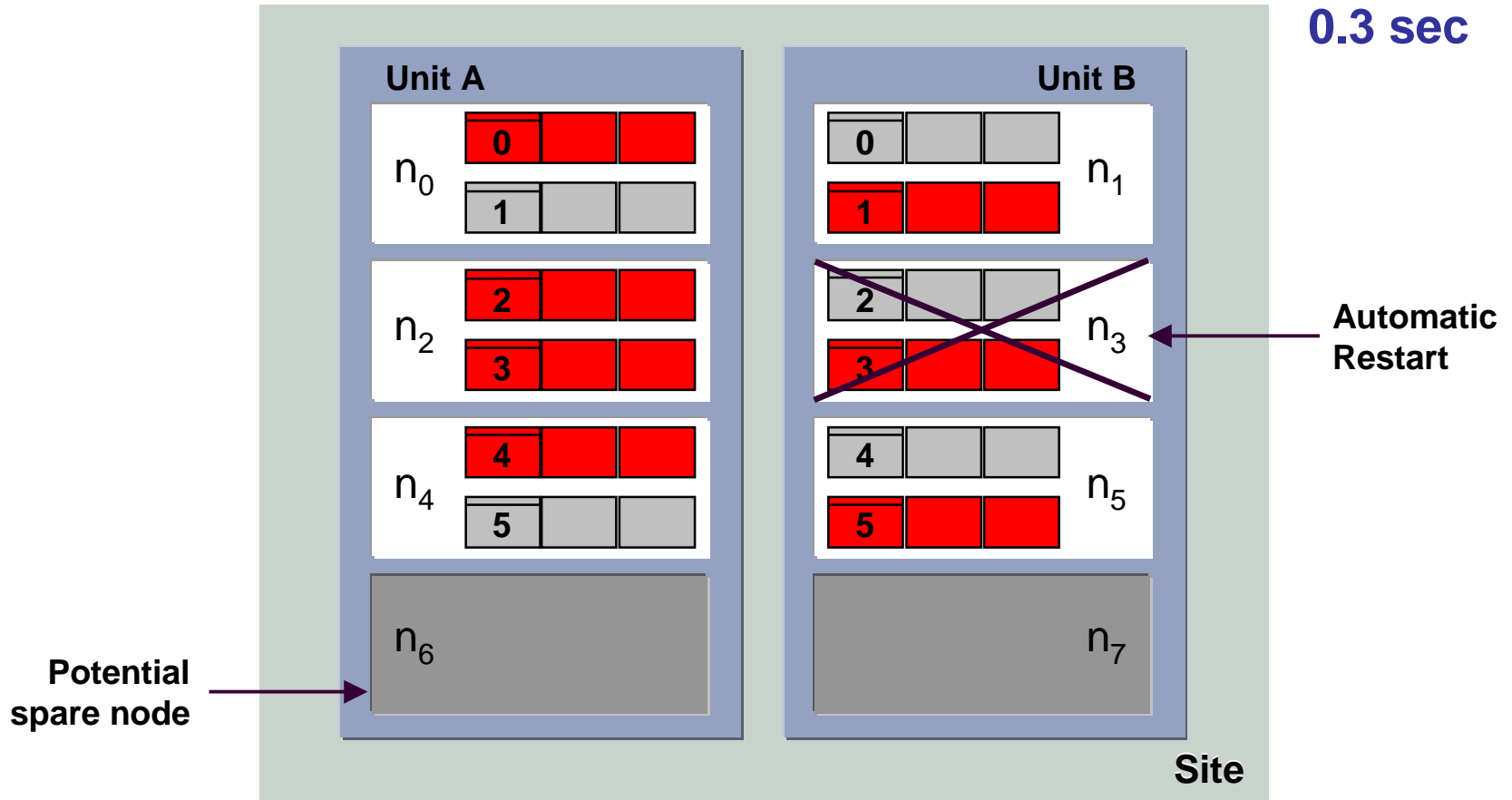
Failure Handling:

Failure at Node n_3



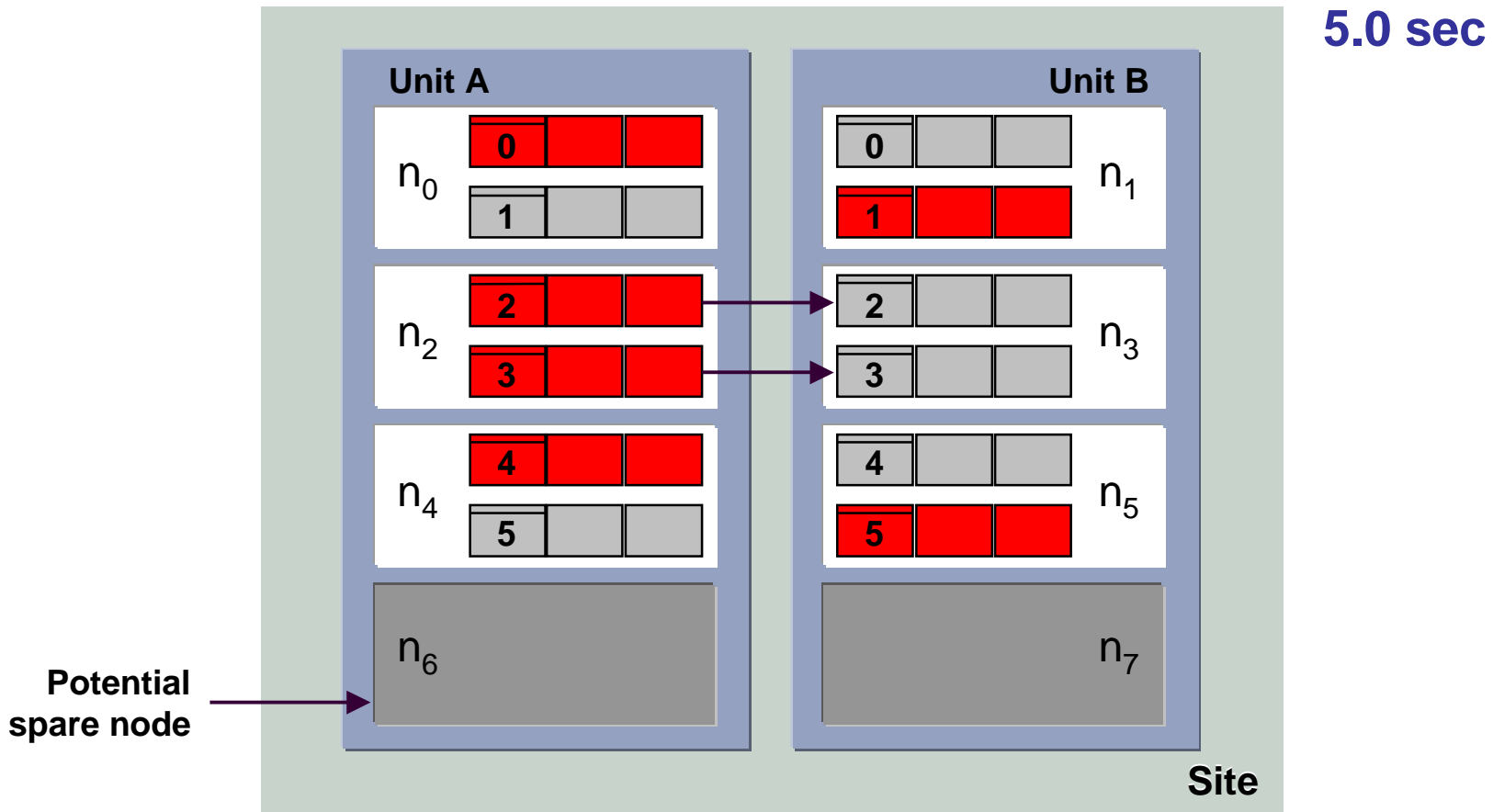
Failure Handling:

Takeover and Restart



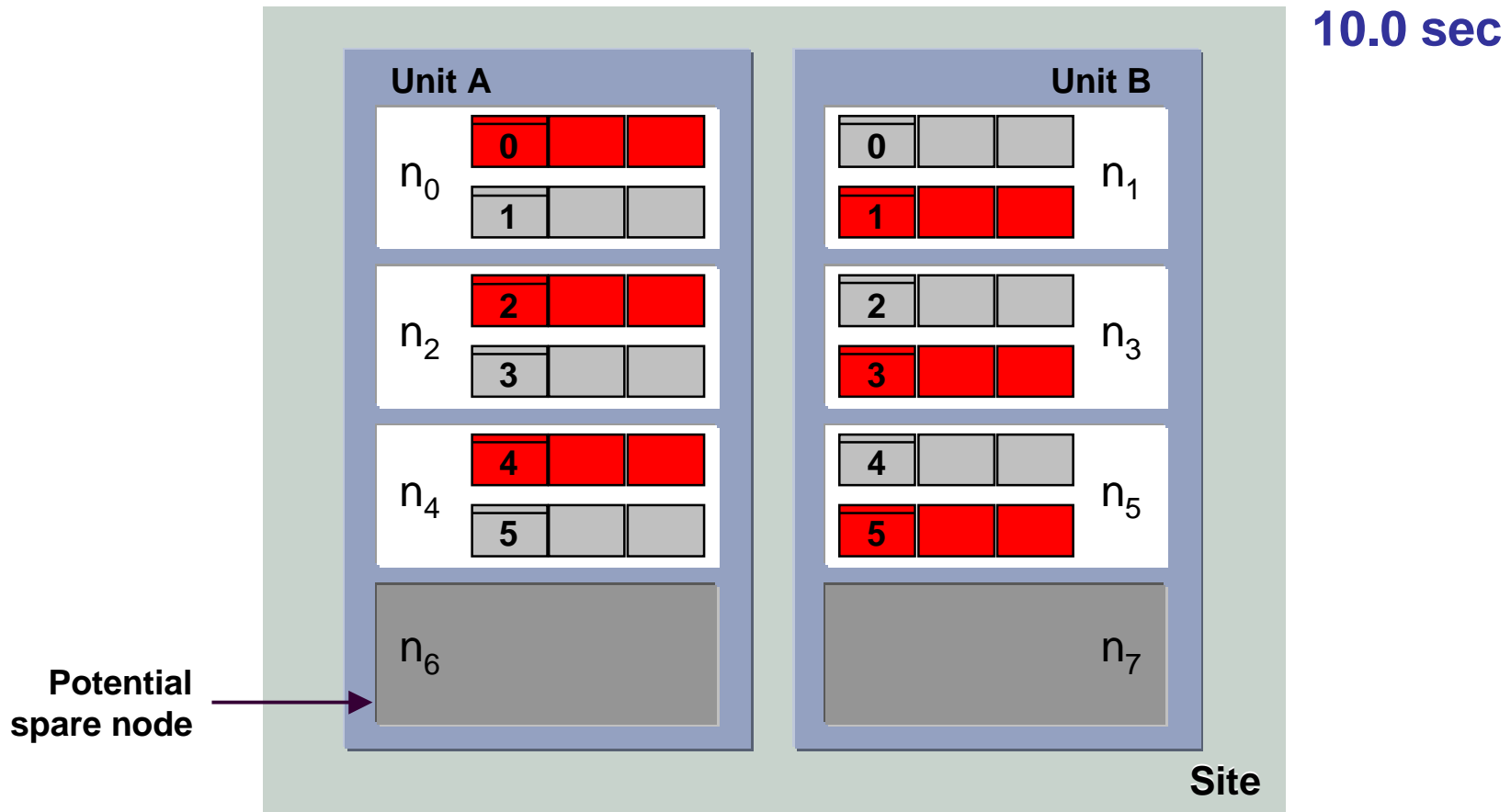
Failure Handling:

Take Back by n_3



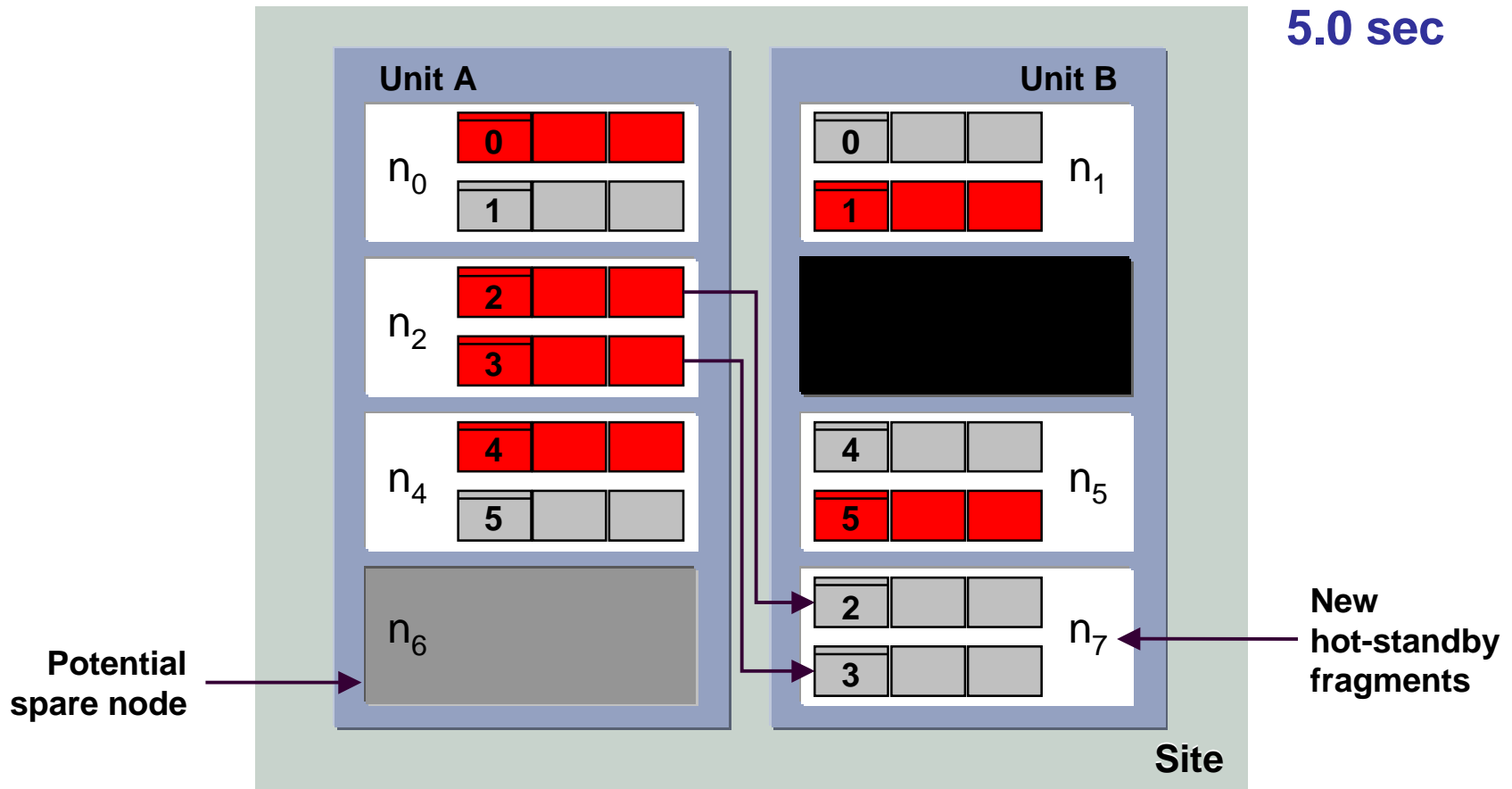
Failure Handling:

Back to Original Configuration



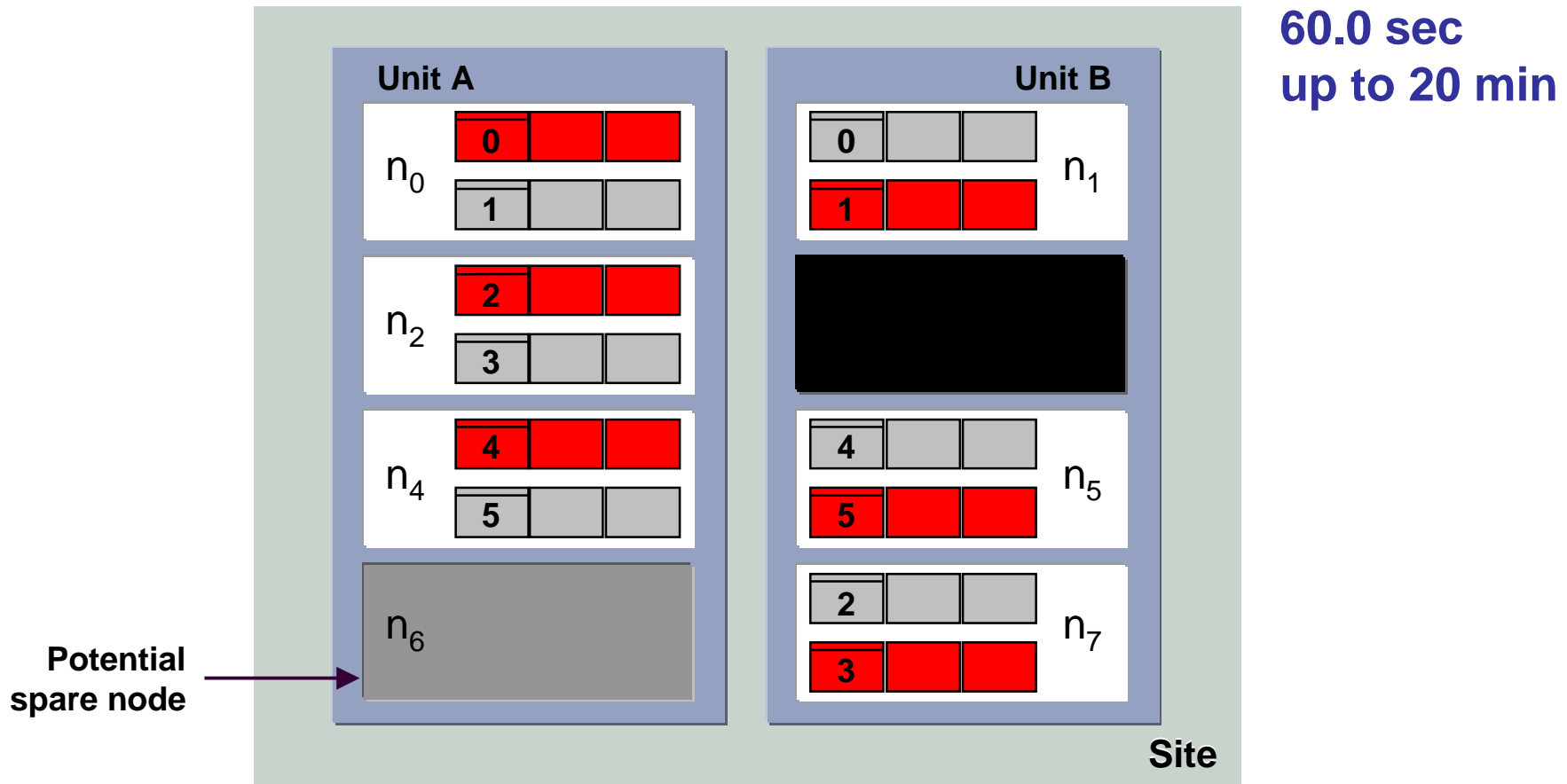
Failure Handling:

Repair to Spare Node



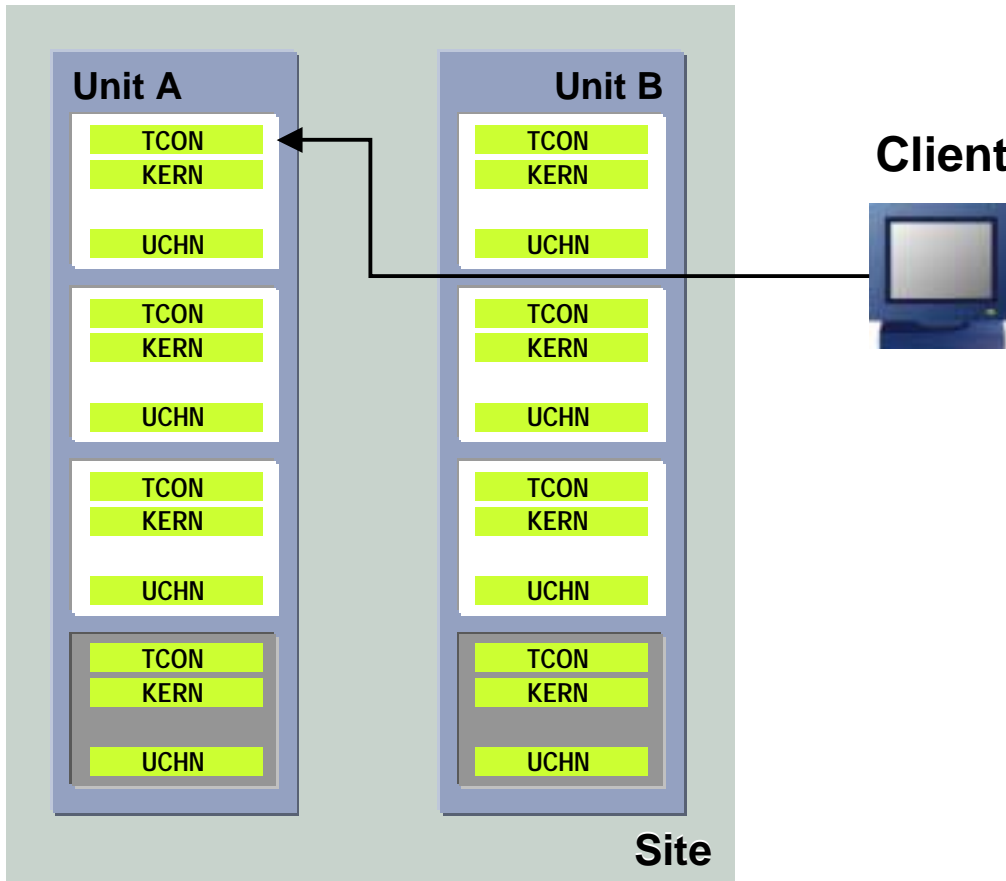
Failure Handling:

Repair Complete, n_7 Replacing n_3



Transaction Execution

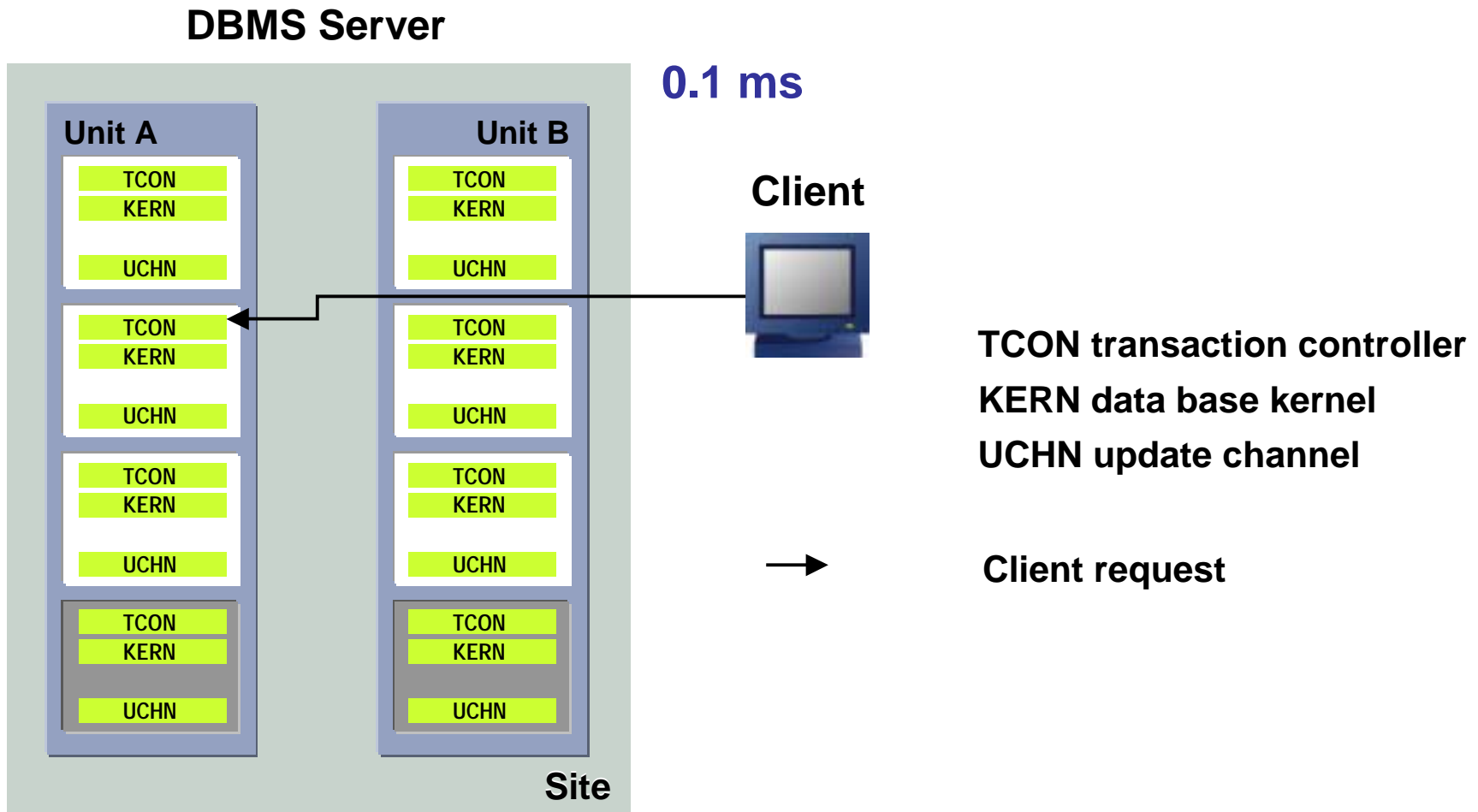
DBMS Server



TCON transaction controller
KERN data base kernel
UCHN update channel

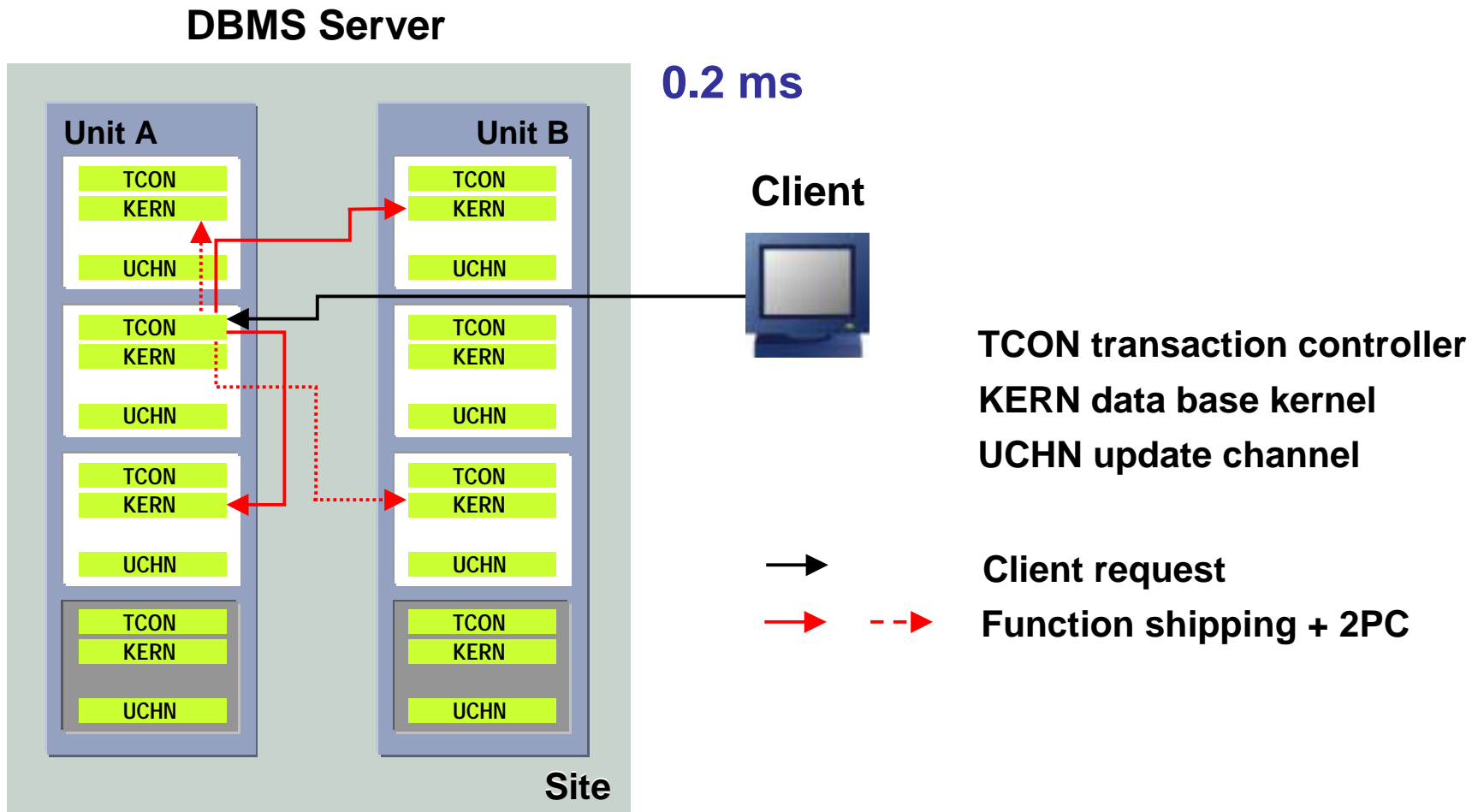
Transaction Execution:

Primary and Hot Standby Session



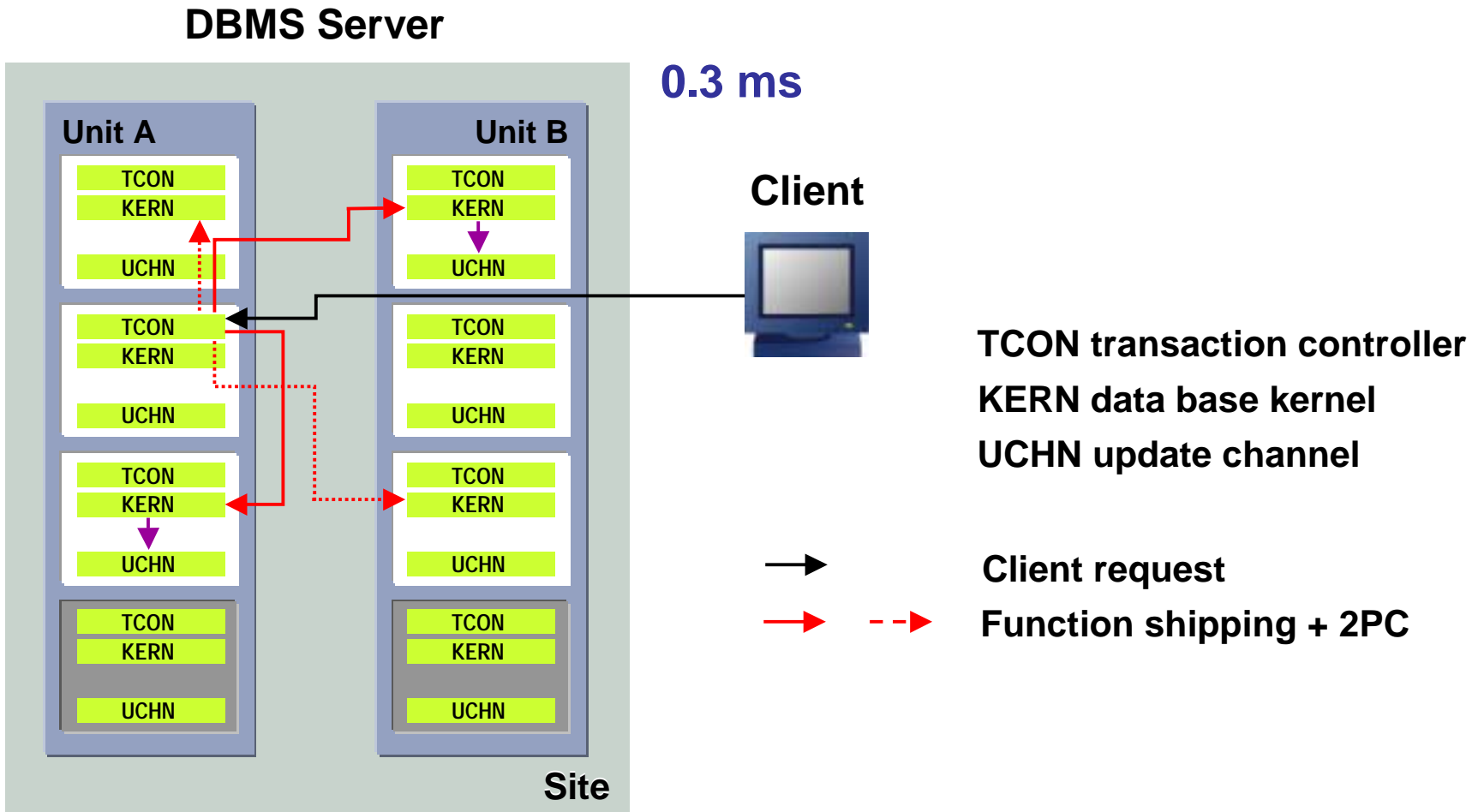
Transaction Execution:

Function Shipping to Primary and Hot Standby



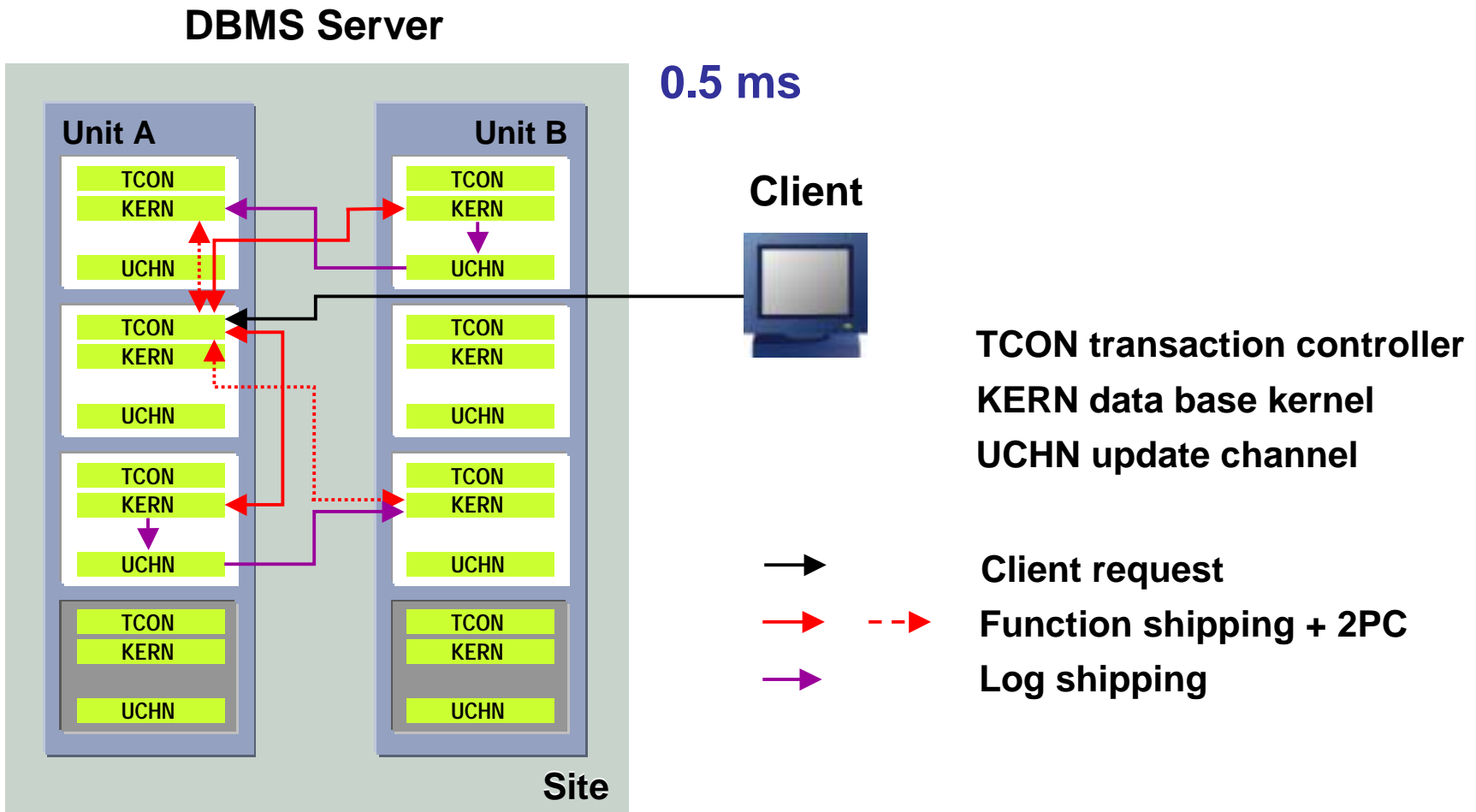
Transaction Execution:

Primary Update



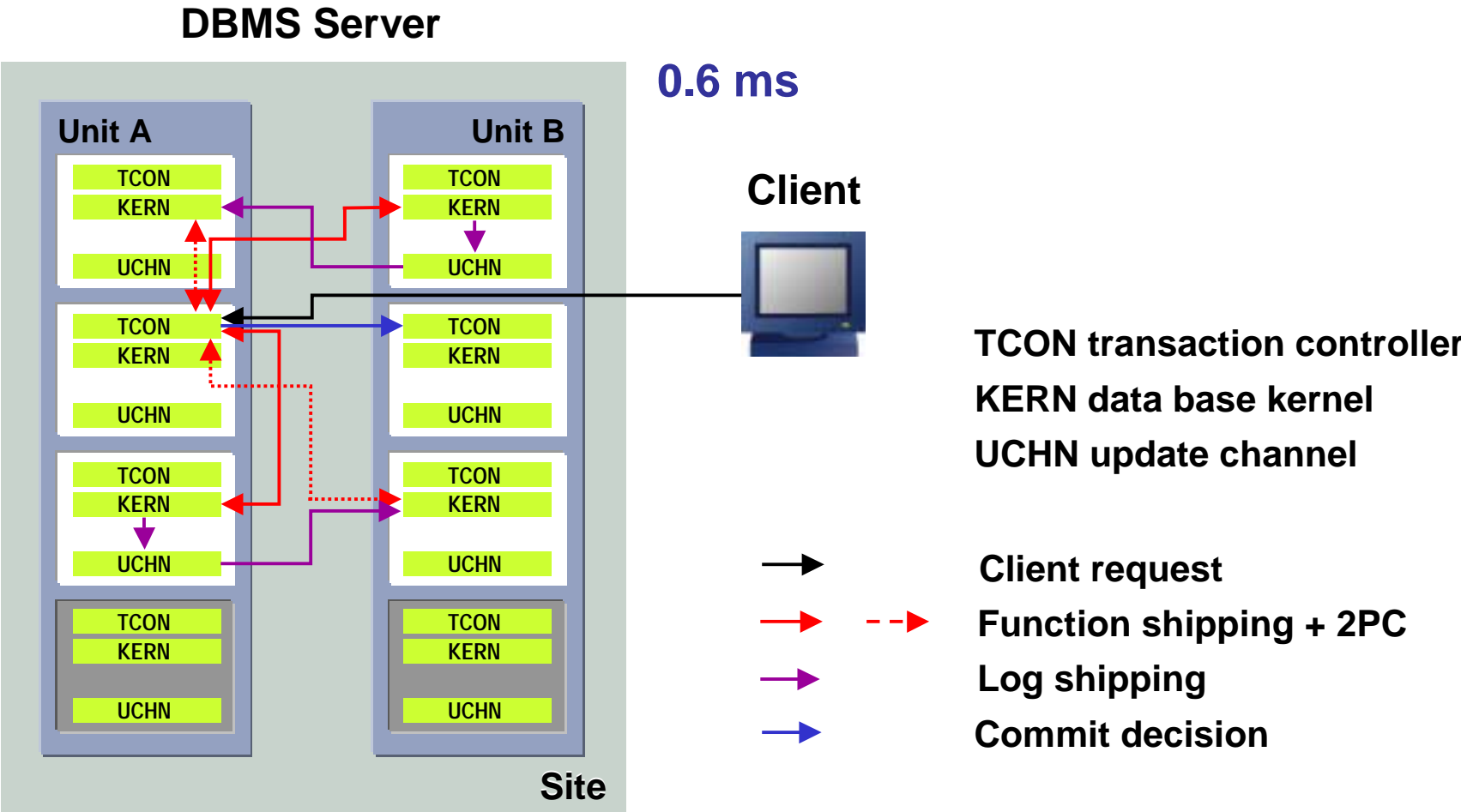
Transaction Execution:

Ready to Commit



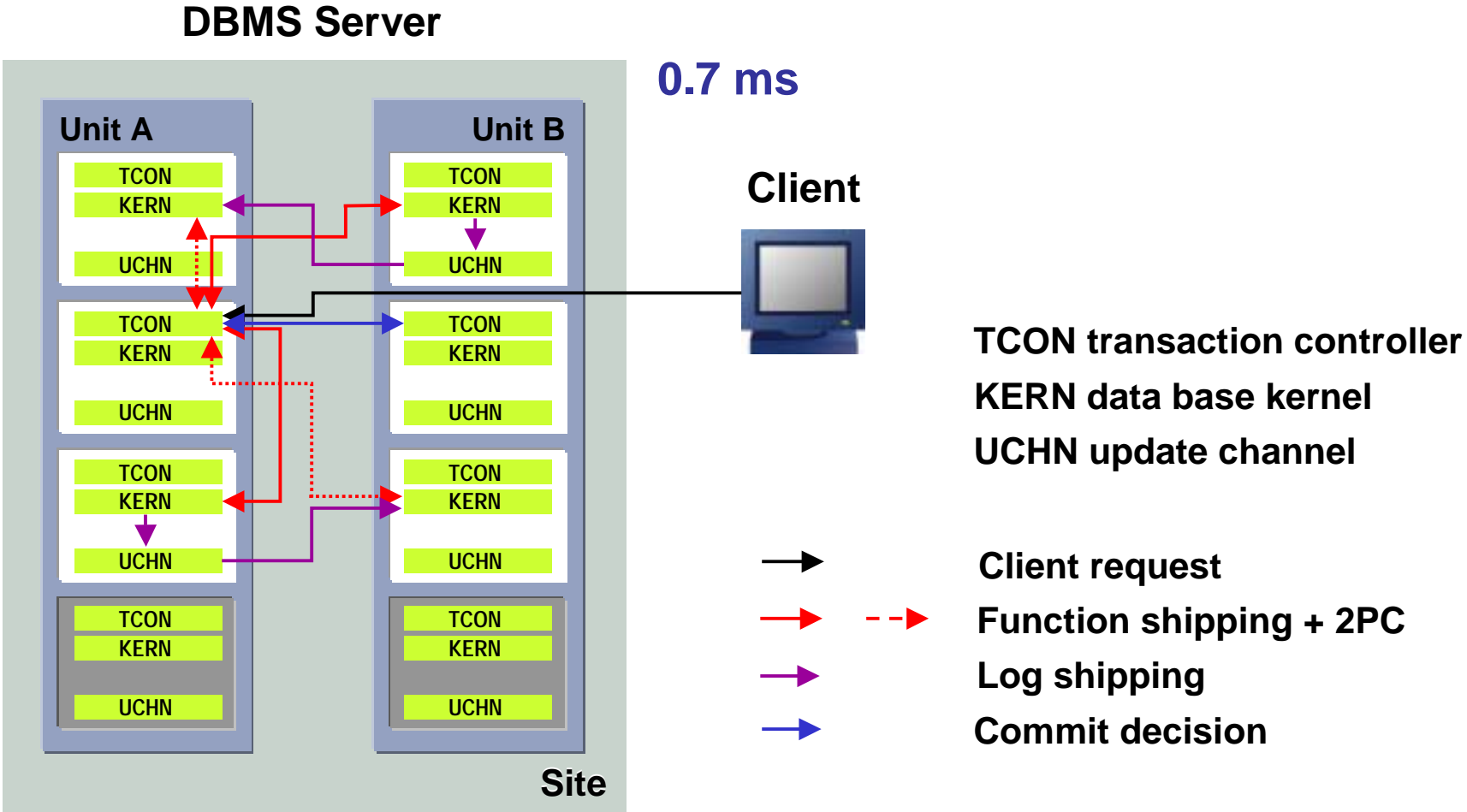
Transaction Execution:

Commit Decision



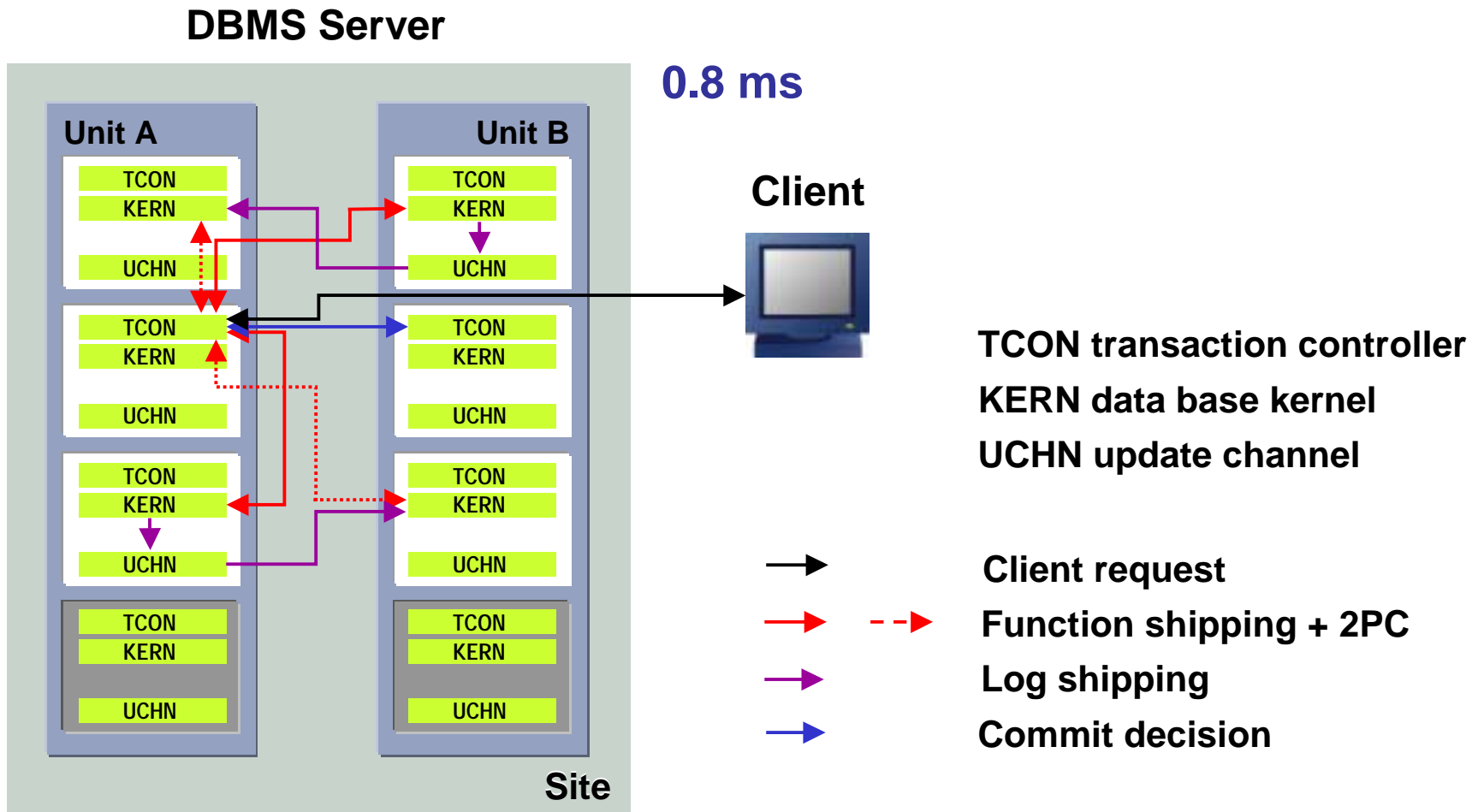
Transaction Execution:

Commit Decision, Ready to Commit



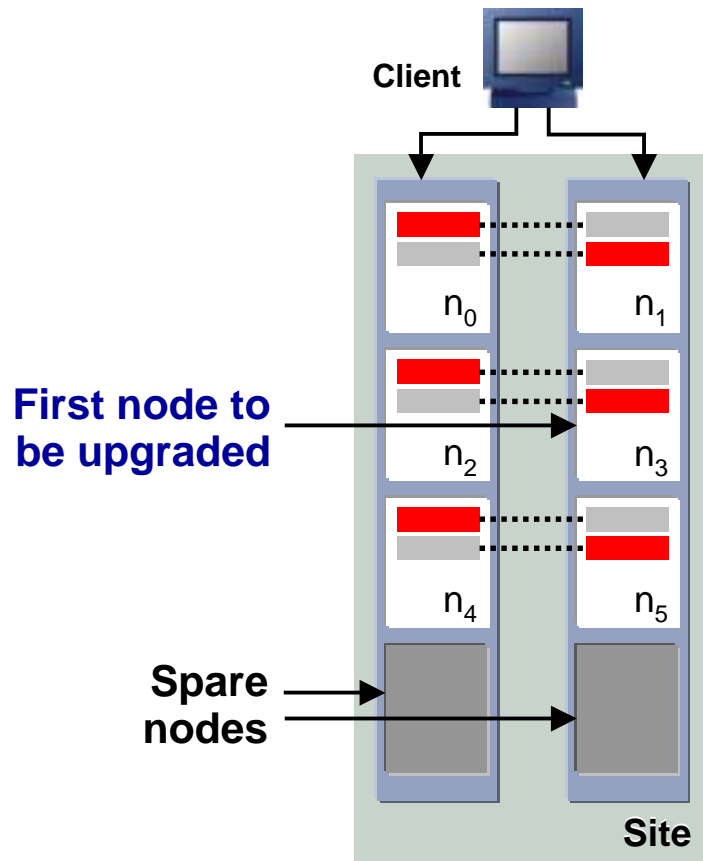
Transaction Execution:

Response to Client



Rolling Upgrades

Original Configuration

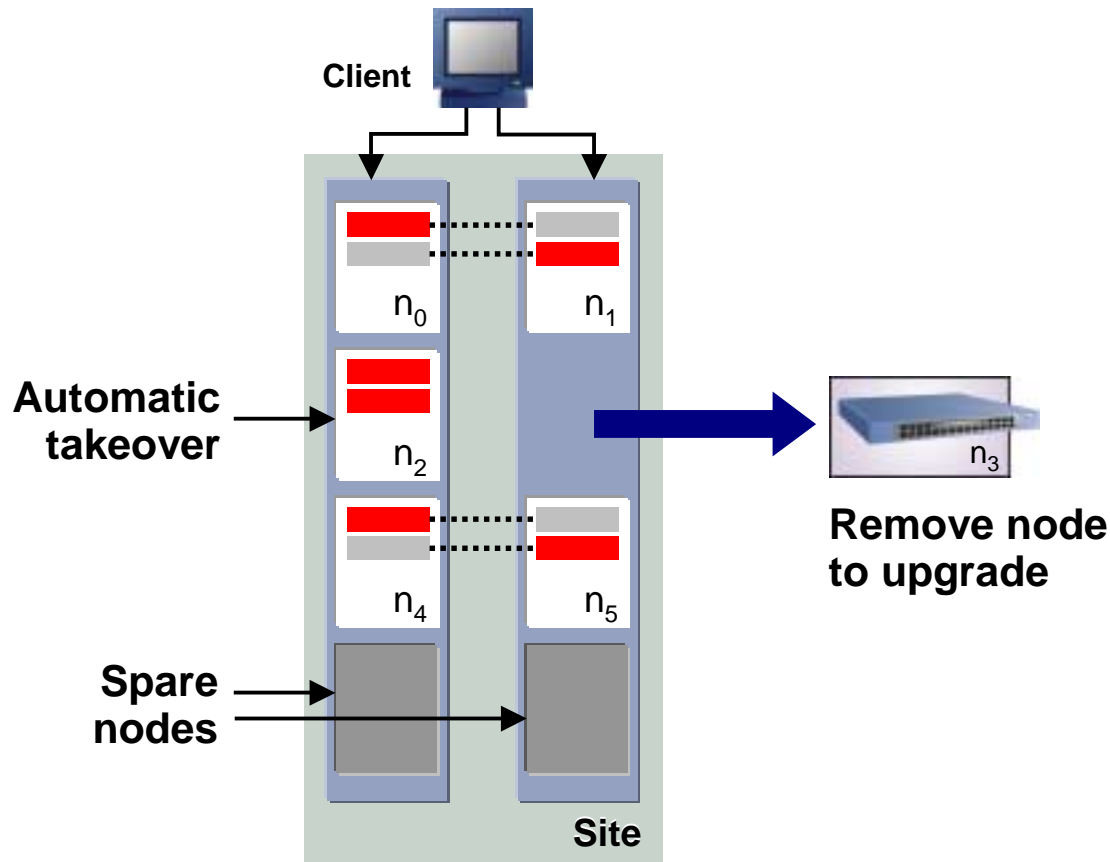


- Online
- Non-blocking
- Hardware and system software
- Clustra Database backward compatibility

■ Primary fragment ■ Hot-standby fragment

Rolling Upgrades

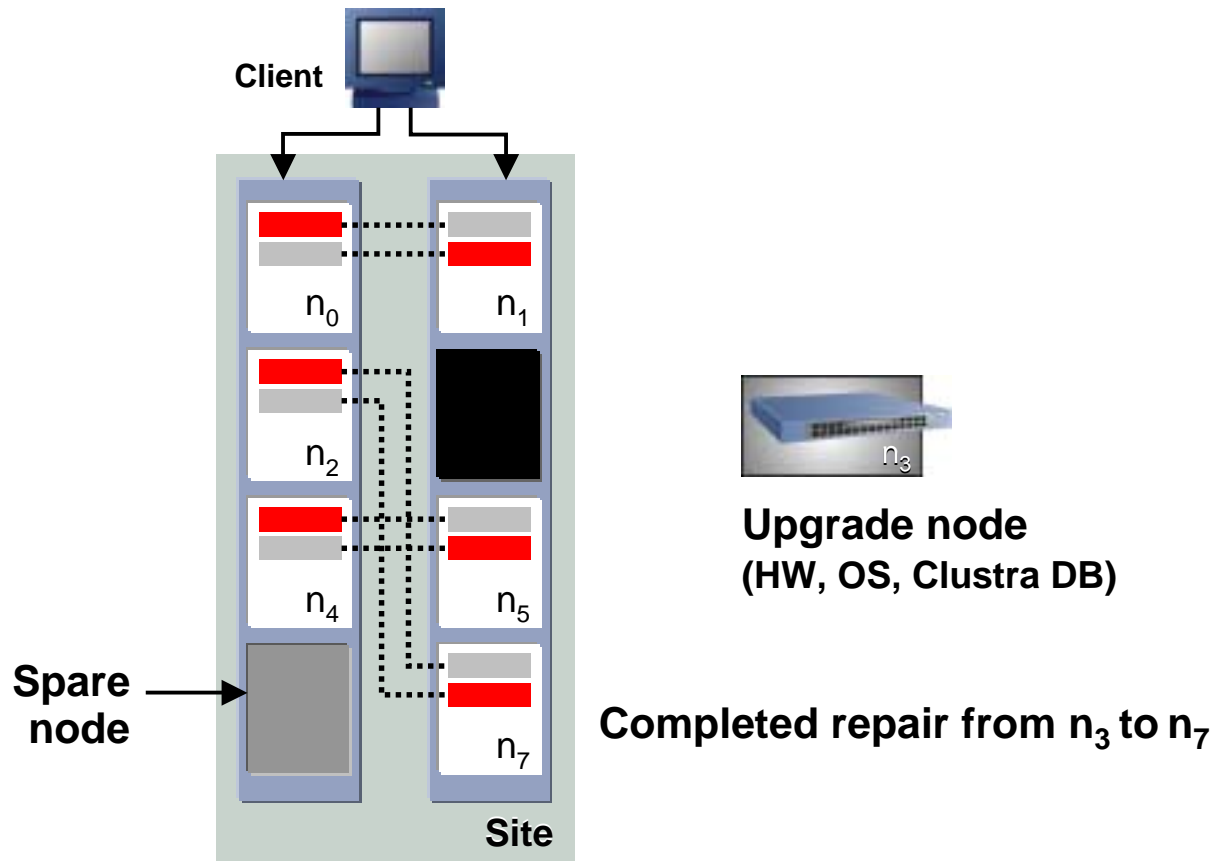
Pull Out Node



■ Primary fragment ■ Hot-standby fragment

Rolling Upgrades

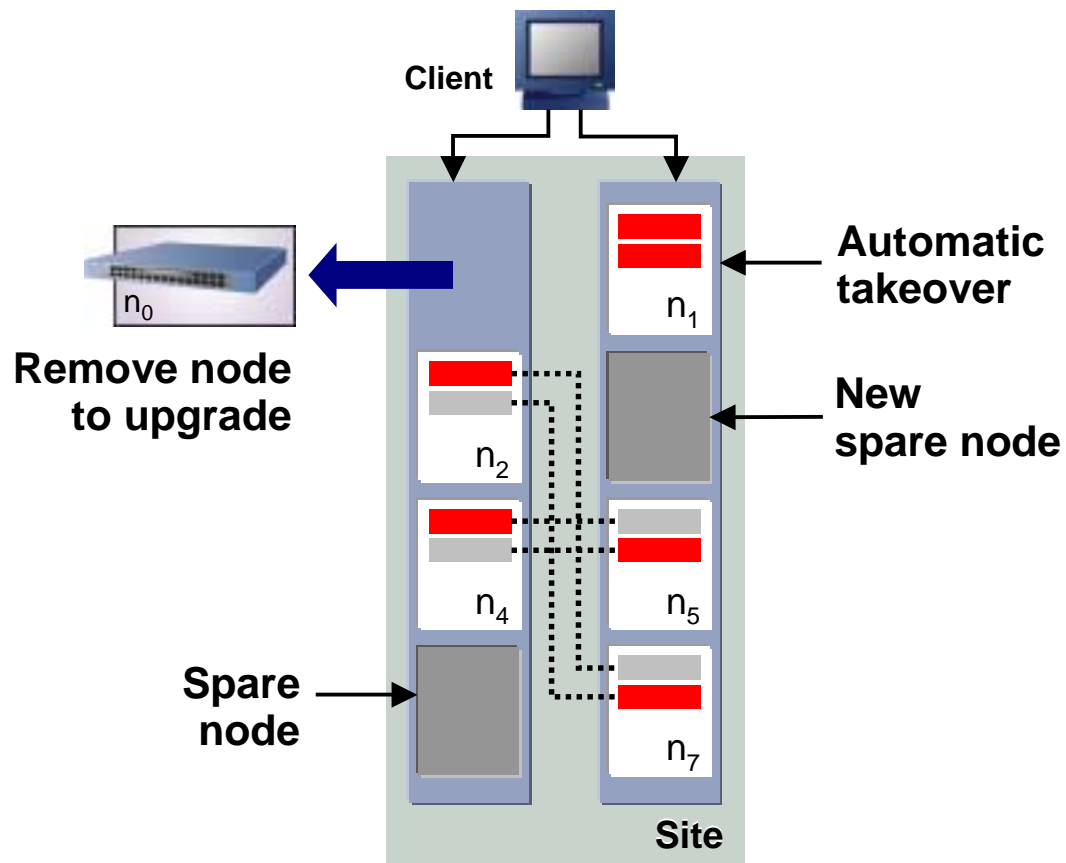
Upgrade and Repair to Spare Node



■ Primary fragment ■ Hot-standby fragment

Rolling Upgrades

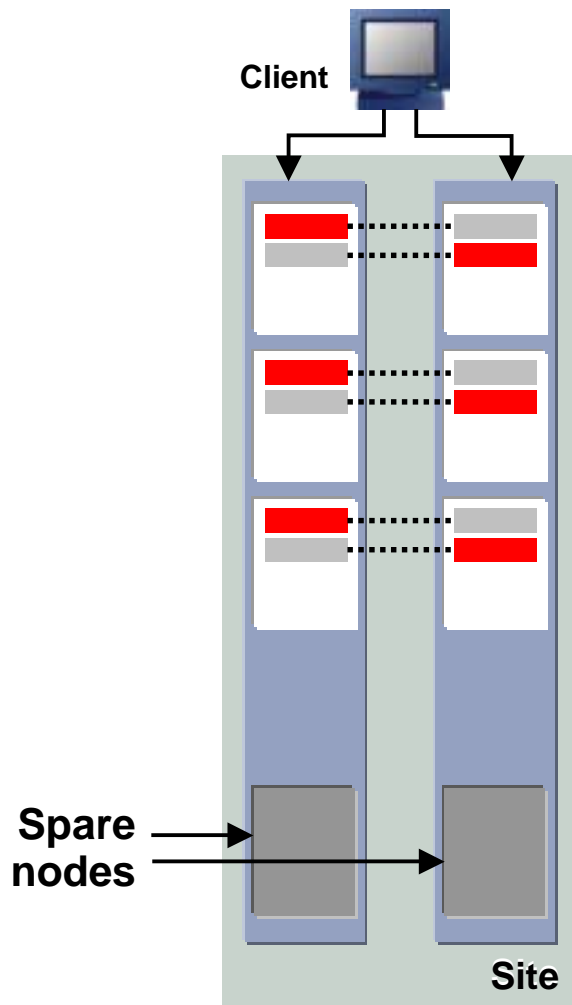
Replace New Spare Node, Upgrade Next



■ Primary fragment ■ Hot-standby fragment

Incremental Scaling of a Site

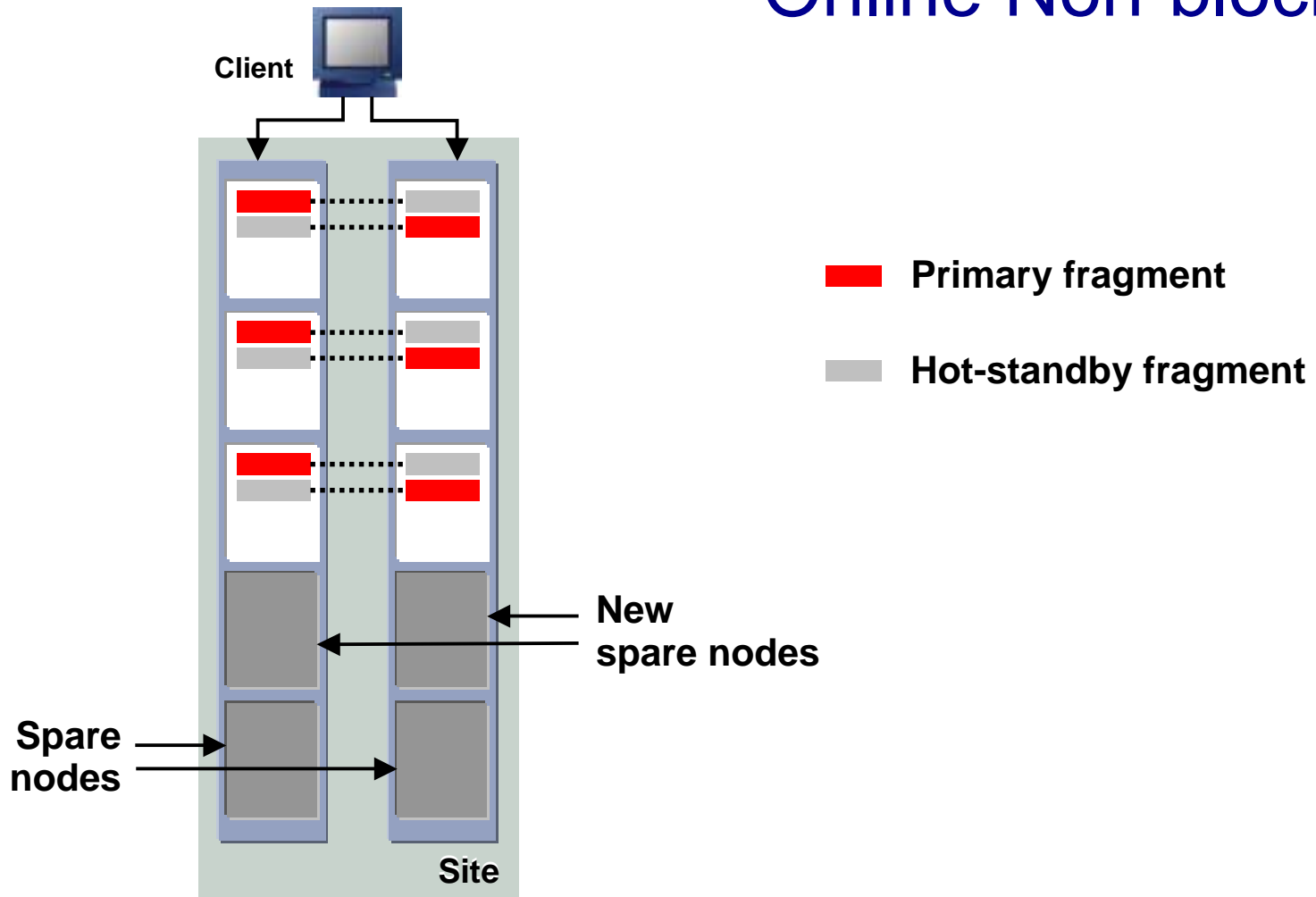
Online Non-blocking



-  Primary fragment
-  Hot-standby fragment

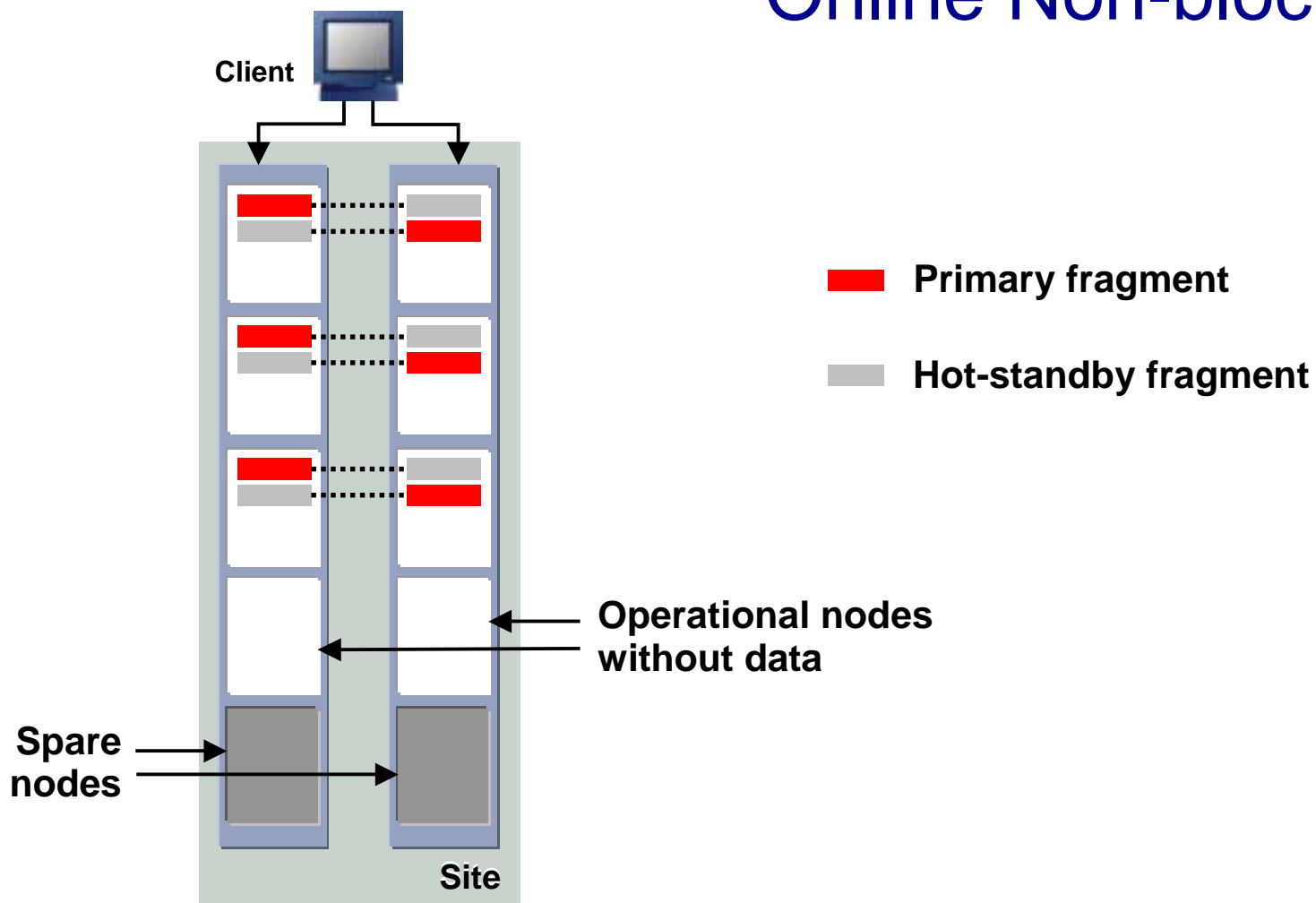
Incremental Scaling of a Site

Online Non-blocking



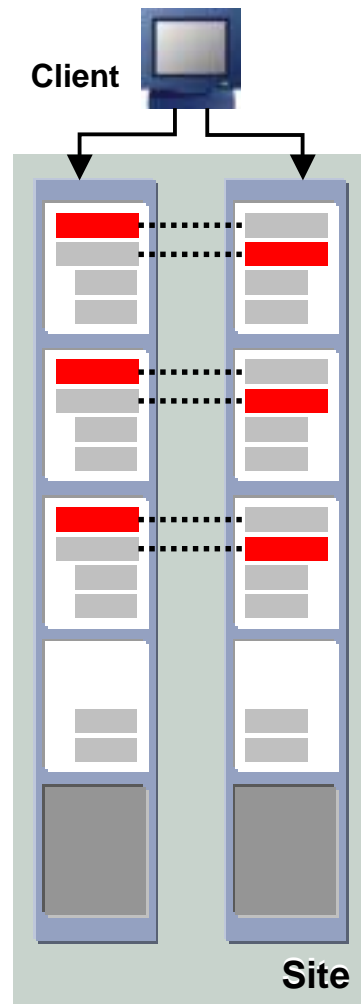
Incremental Scaling of a Site

Online Non-blocking



Incremental Scaling of a Site

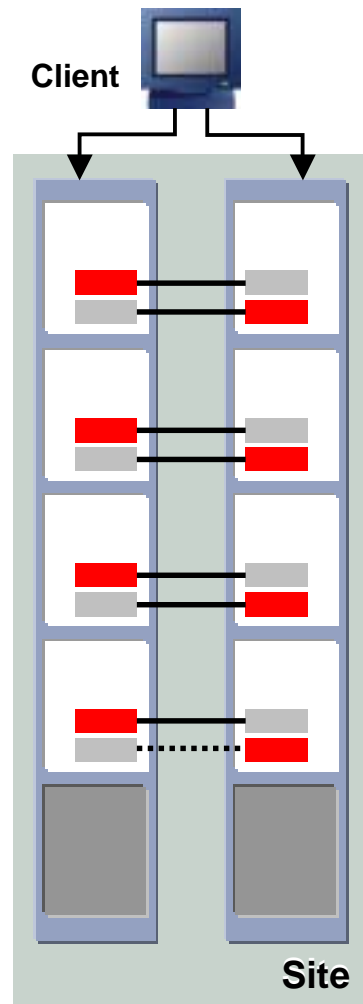
Online Non-blocking



-  Primary fragment
-  Hot-standby fragment

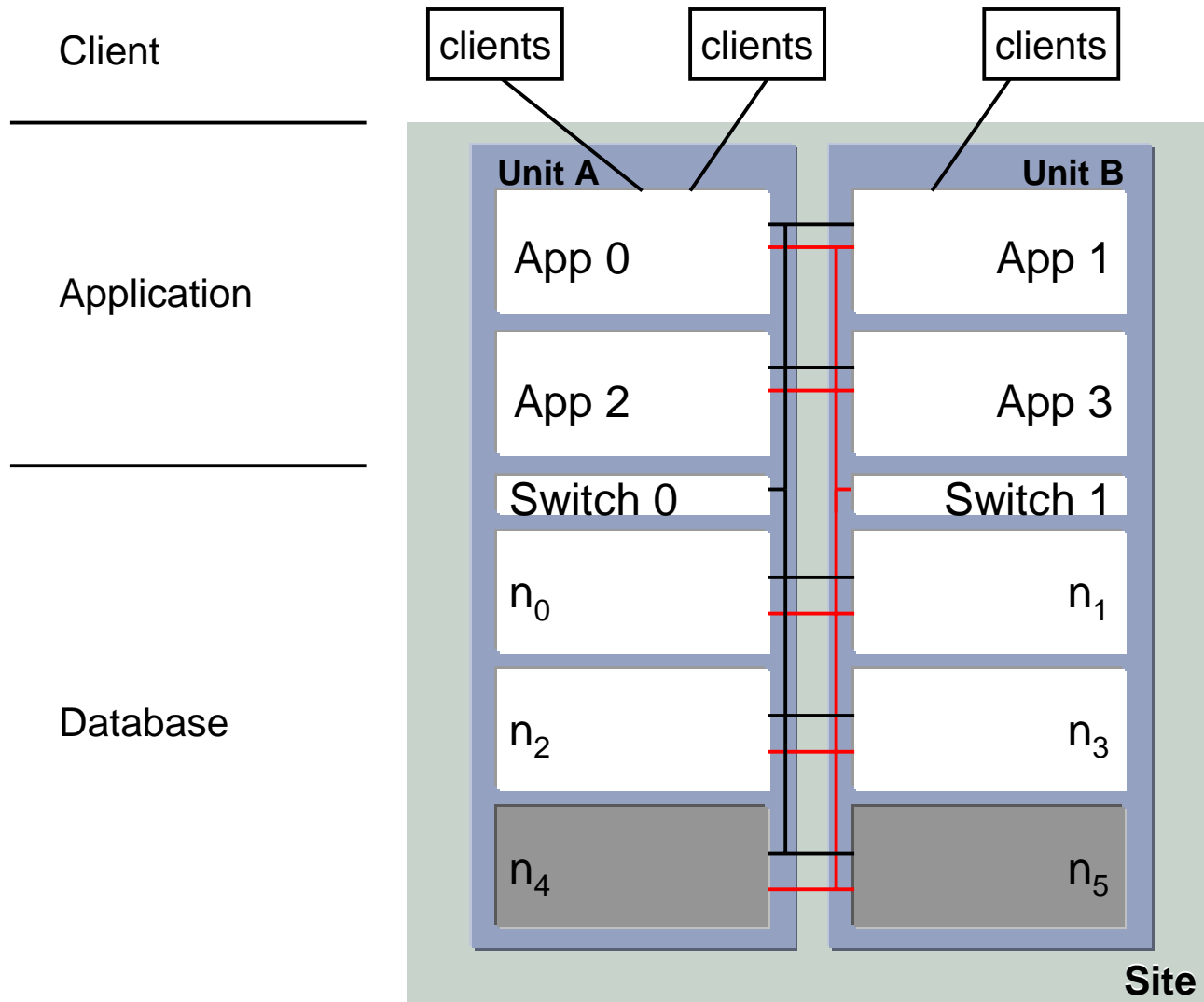
Incremental Scaling of a Site

Online Non-blocking



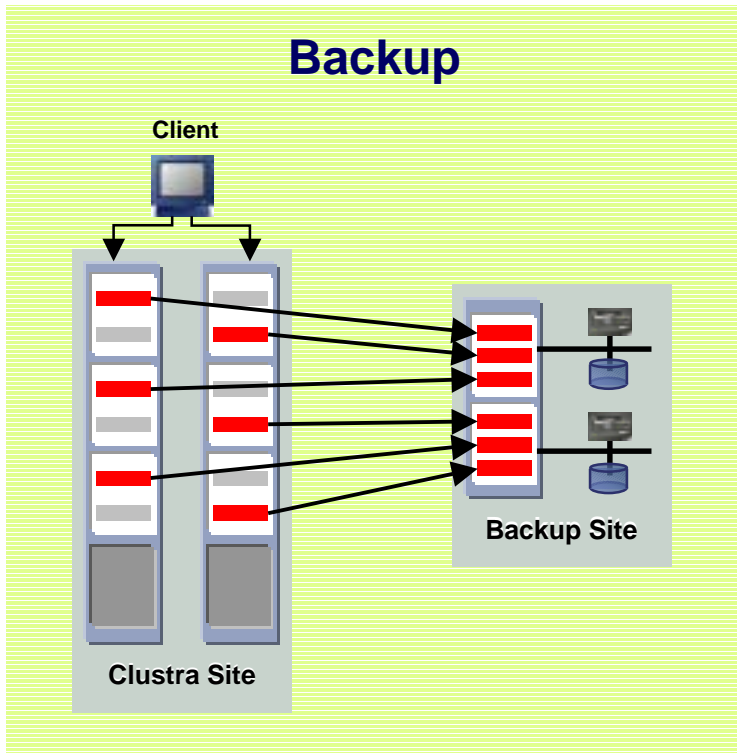
- Primary fragment
- Hot-standby fragment

Typical Application Architecture

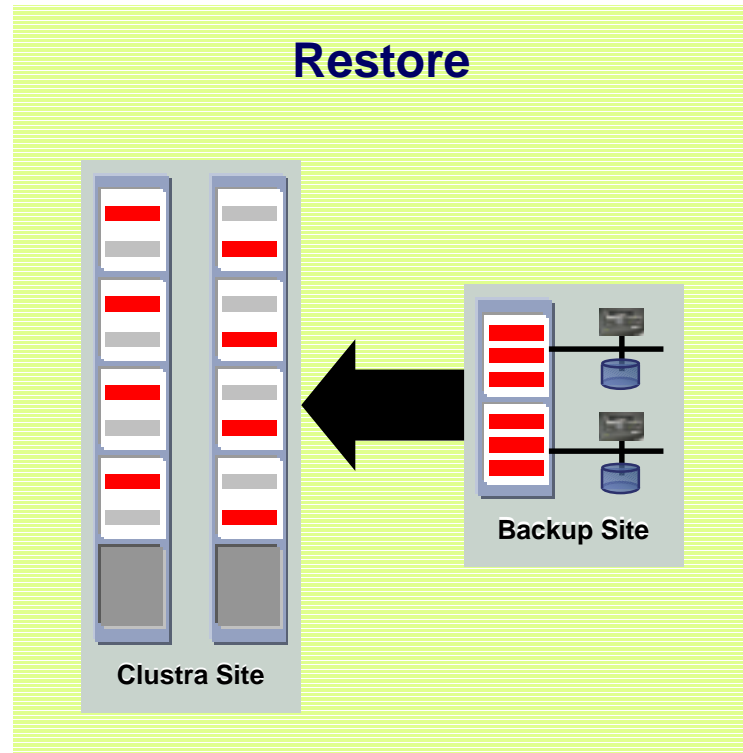


Backup and Restore

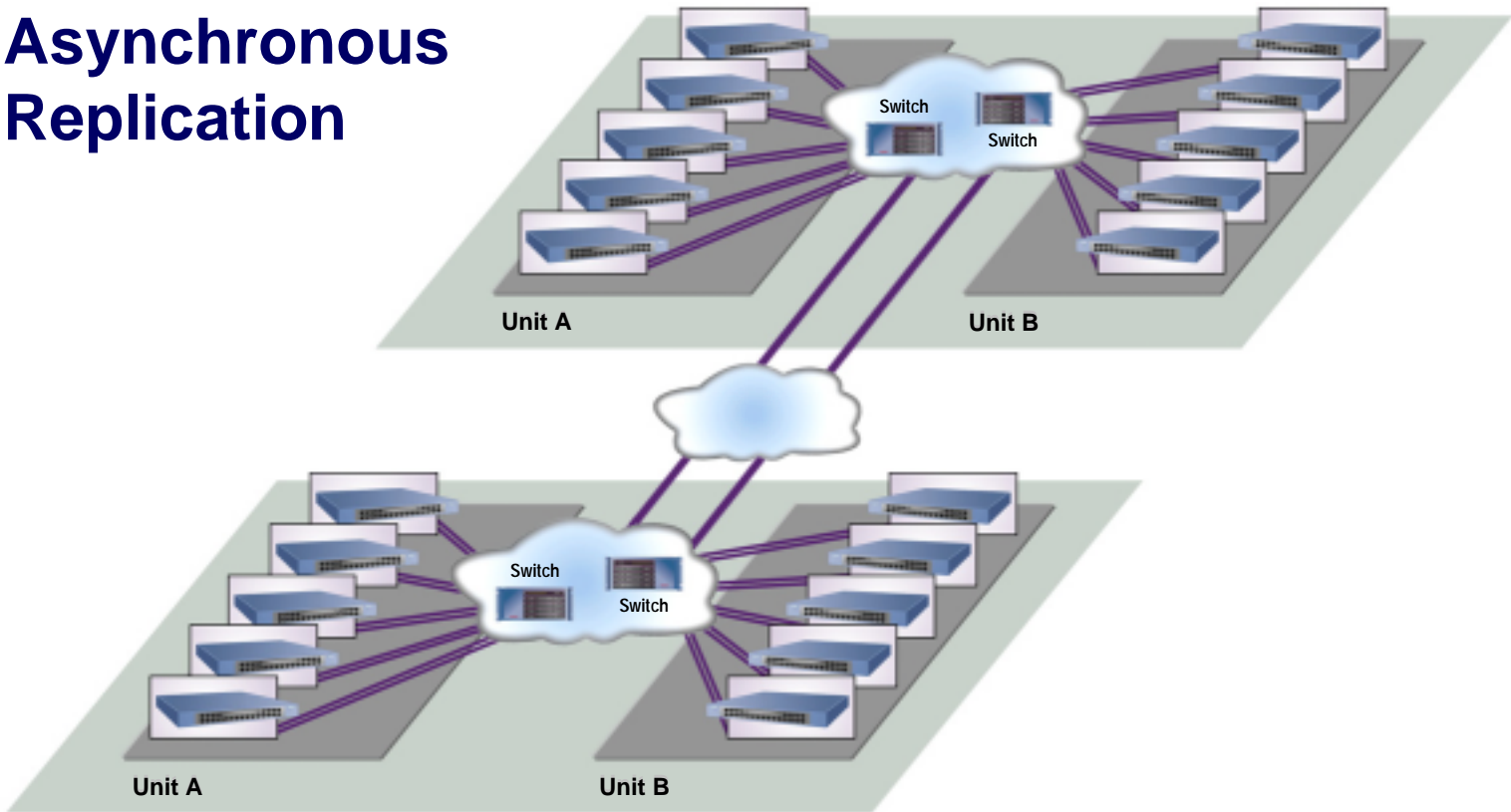
Parallel Non-Blocking



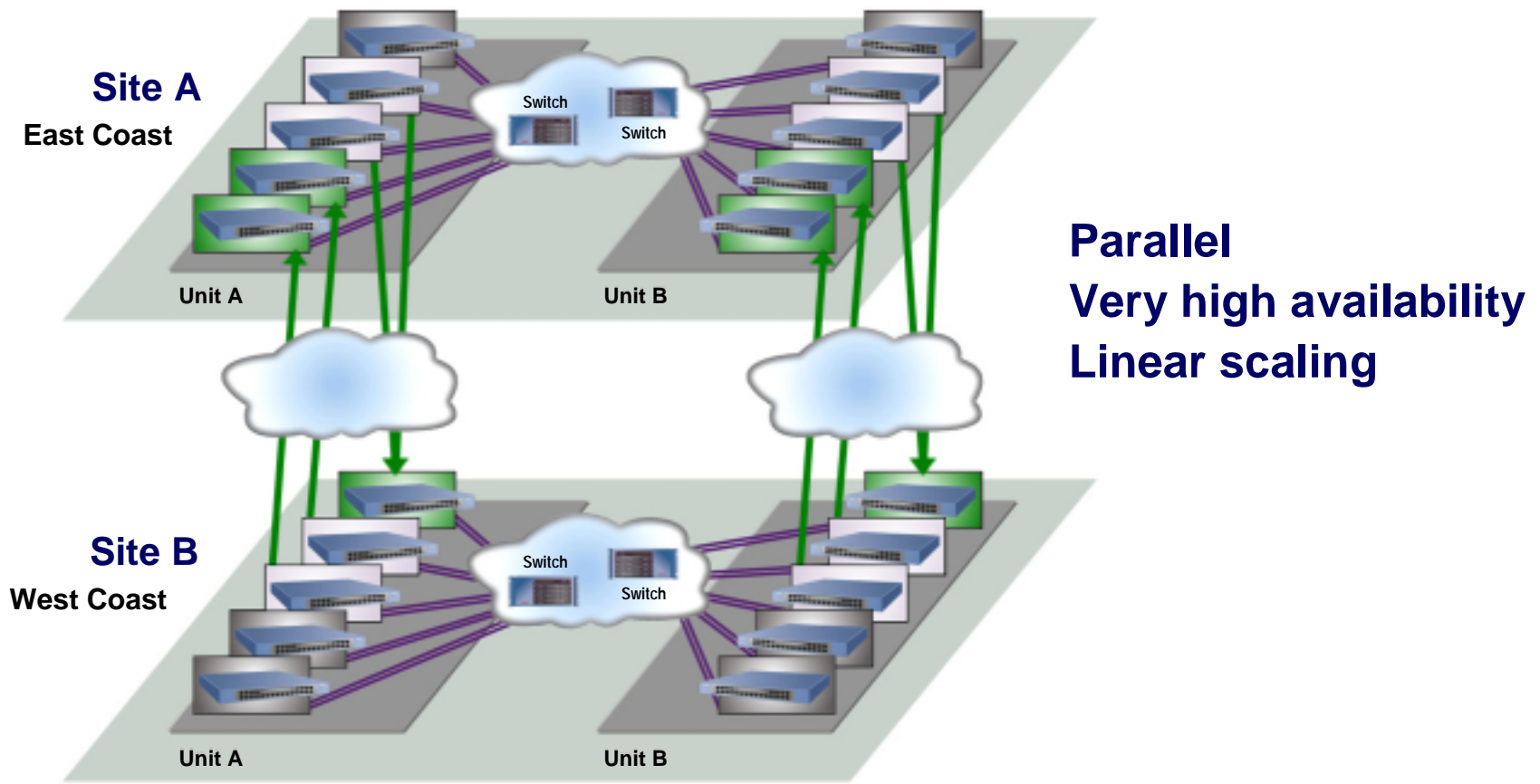
Refragmentation



Asynchronous Replication

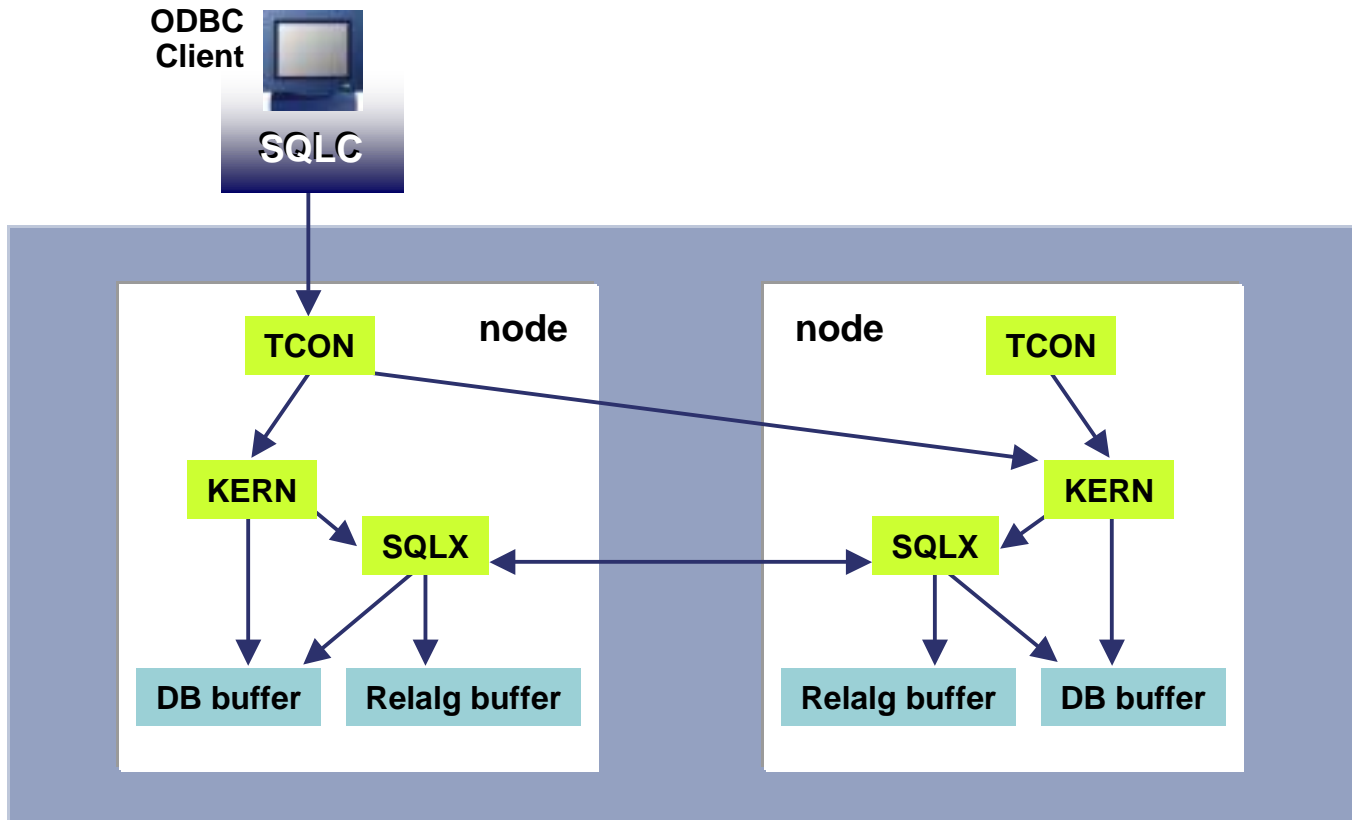


Asynchronous Replication

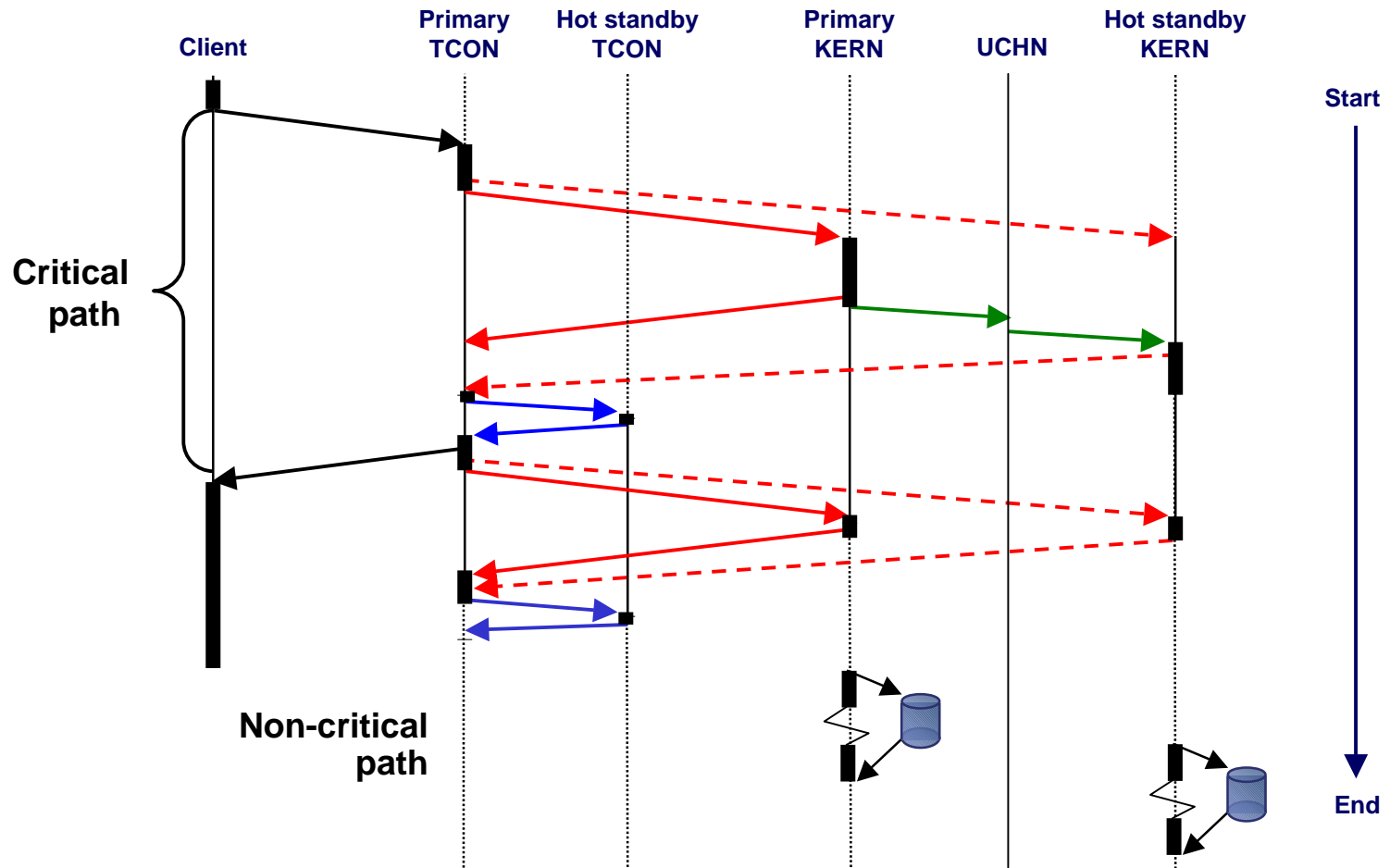


SQL: Relational Algebra

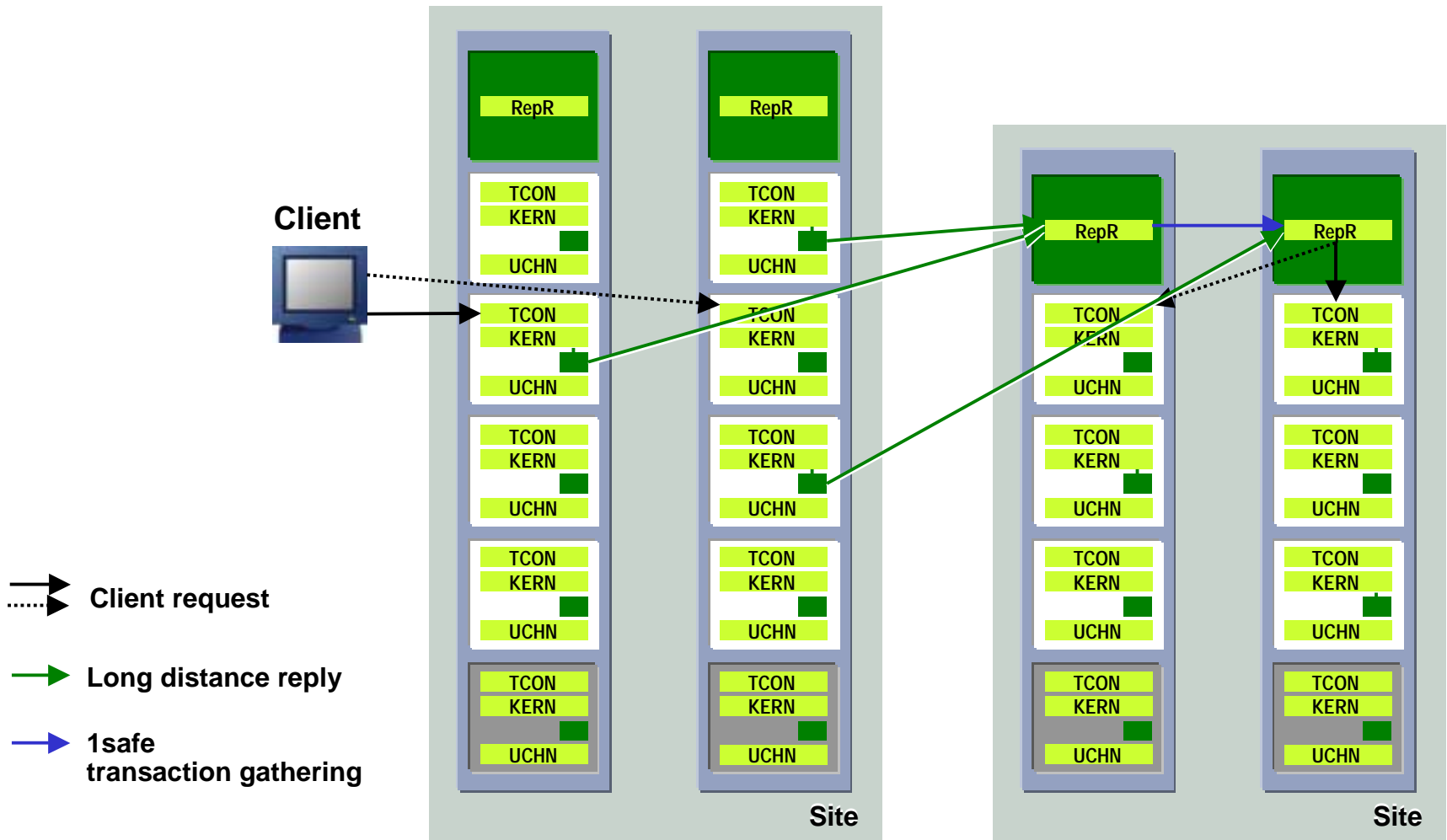
Parallel Execution in Kernel



Transaction Execution



Asynchronous Replication



Next Steps



Thank You

Rolf Rander Næss

rolfn@clustra.com

+47 8800 4243

www.clustra.com