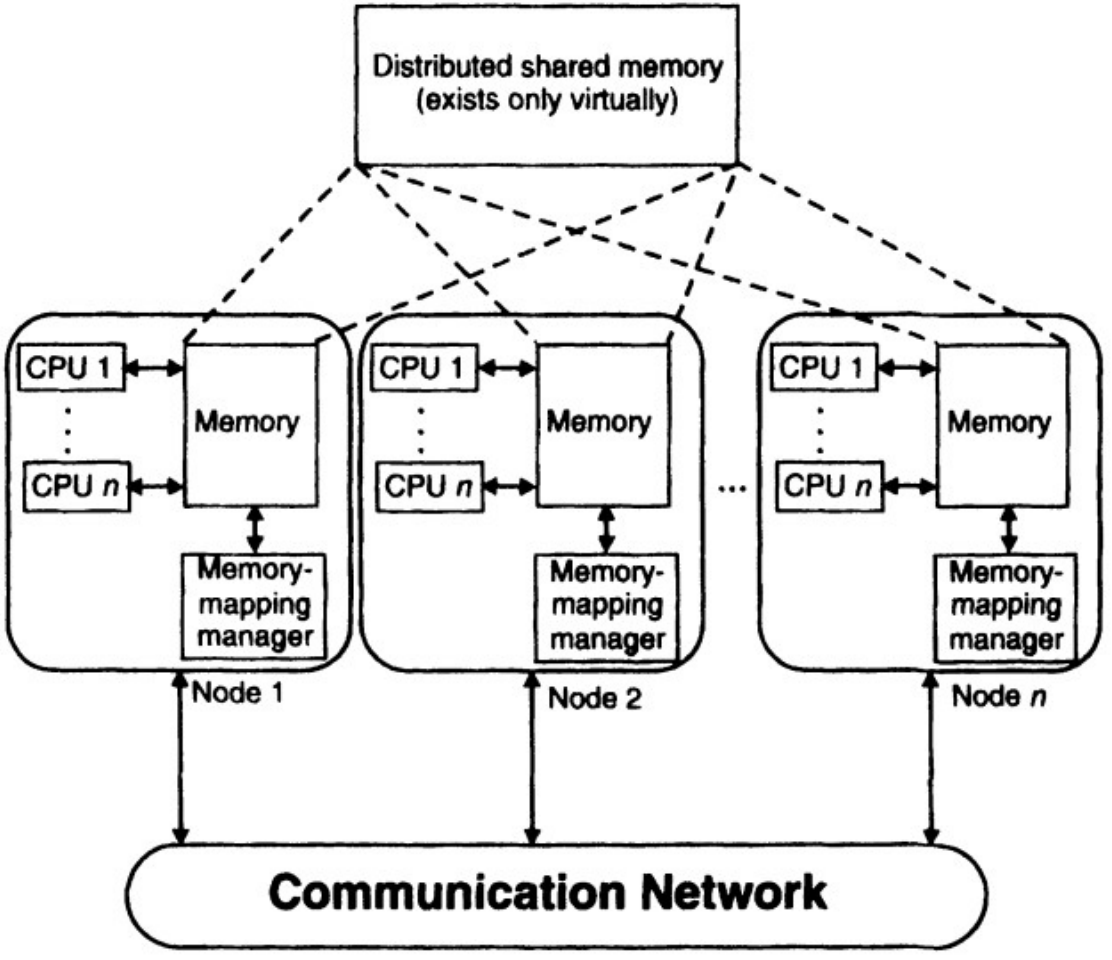
Distributed Shared Memory

2 basic paradigms for IPC:-
1. shared memory paradigm
2. message-passing paradigm -RPC


Distributed shared memory
(exists only virtually)
CPU 1
CPU n
Memory
Memory-
mapping
manager
Node 1
CPU 1
CPU n
Memory
Memory-
mapping
manager
Node 2
...
CPU 1
CPU n
Memory
Memory-
mapping
manager
Node n
Communication Network

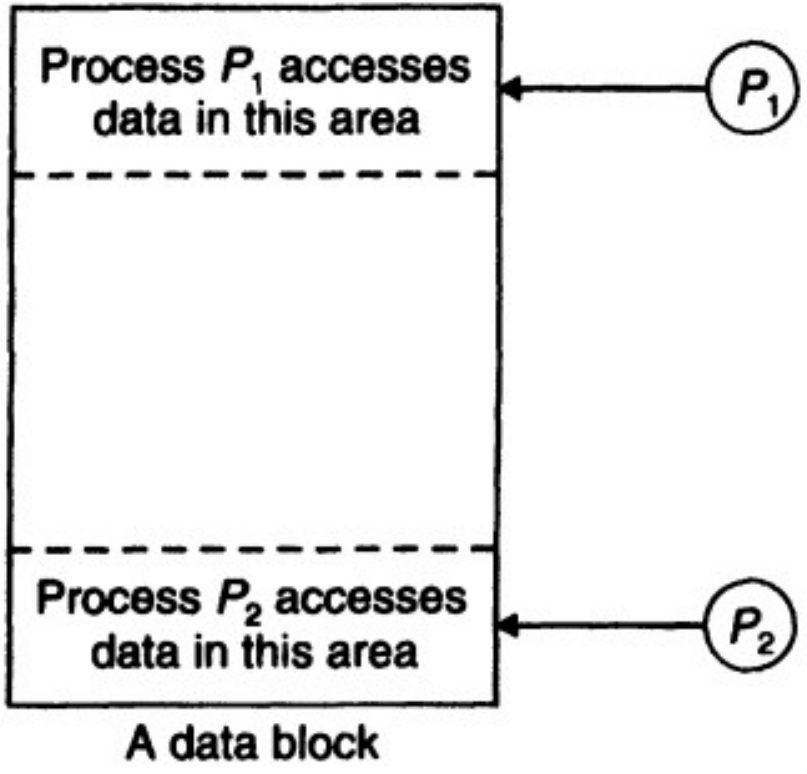
DESIGN AND IMPLEMENTATION ISSUES OF DSM

- Granularity
- Structure of shared-memory space
- Memory coherence & access synchronization
- Data location & access
- Replacement strategy
- Thrashing
- Heterogeneity

Granularity

- Usually large block size is favoured over small ones.
- Factors affecting the choice of block size:-
- Paging overhead: is less for large block size compared to small block size
- Directory size: larger the block size, smaller the directory, as directory information about the block has to be maintained.
- Thrashing: may occur with any block size but more likely with larger block size as different regions of same block can be updated by processes on different node, causing data block transfers that are not necessary with smaller block size.
- False sharing: The larger is the block size, the higher is the probability of false sharing, due to the fact that the same data block may contain different data structures that are used independently.

False sharing



CONSISTENCY MODELS

- Degree of consistency that has to be maintained for shared memory data to work correctly for applications.
- Set of rules that the applications has to obey

Cntd...

- Strict consistency model : strongest
- Sequential consistency model
- Causal consistency model
- Pipelined RAM consistency model
- Processor consistency model
- Weak consistency model
- Release consistency model

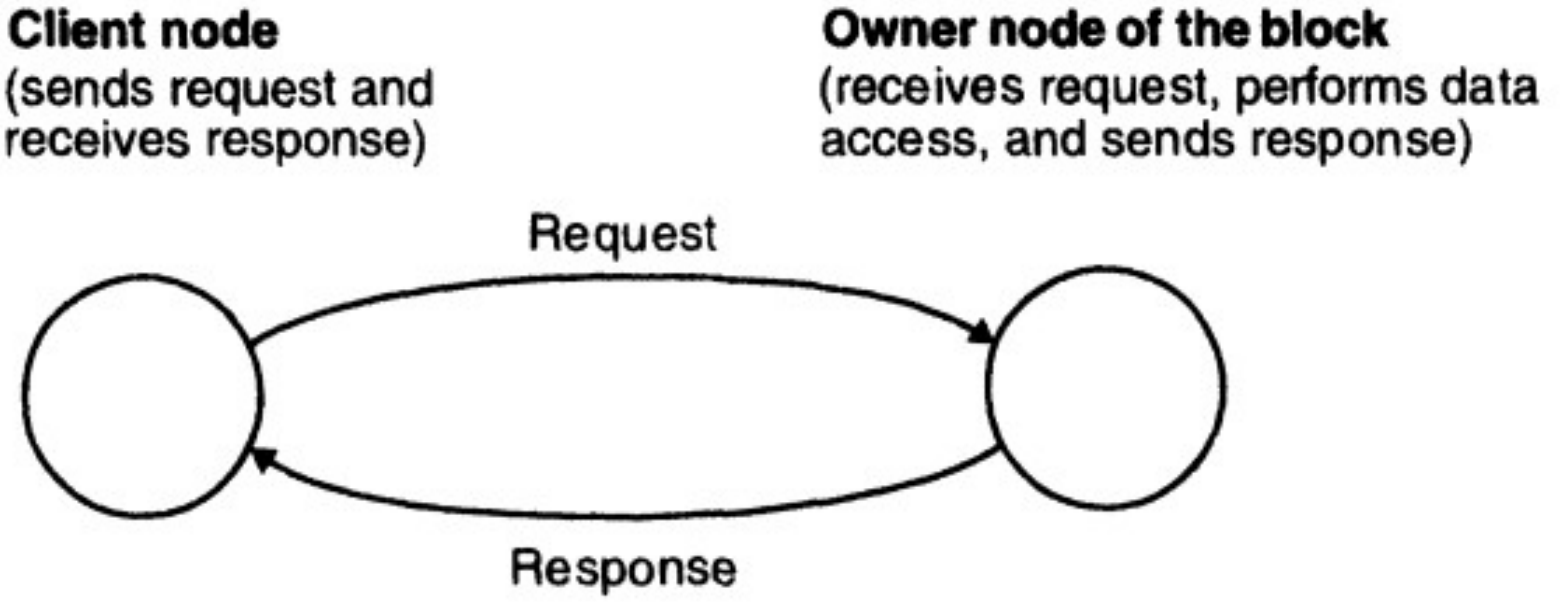
Implementing Sequential Consistency Model

- Protocols for implementing the sequential consistency model in a DSM system depend to a great extent on whether the DSM system allows replication and/or migration of shared-memory data blocks.
- The designer of a DSM system may choose from among the following replication and migration strategies [Stumm and Zhou 1990]:
- 1. Nonreplicated, nonmigrating blocks (NRNMBs)-The Central-Server Algorithm
- 2. Nonreplicated, migrating blocks (NRMBs)- The Migration Algorithm
- 3. Replicated, migrating blocks (RMBs)-The Full-Replication Algorithm
- 4. Replicated, nonmigrating blocks (RNMBs)-The Read-Replication Algorithm

Nonreplicated, Nonmigrating Blocks

- This is the simplest strategy for implementing a sequentially consistent DSM system.
- In this strategy, each block of the shared memory has a single copy whose location is always fixed.
- All access requests to a block from any node are sent to the owner node of the block, which has the only copy of the block.

Nonreplicated, Nonmigrating Blocks



NRNMB Cntd...

- Although the method is simple and easy to implement, it suffers from the following drawbacks:
  - Serializing data access creates a bottleneck.
  - Parallelism, which is a major advantage of DSM, is not possible with this method.

NRNMB Cntd...

- The NRNMB strategy has the following characteristics:
- 1. There is a single copy of each block in the entire system.
- 2. The location of a block never changes.

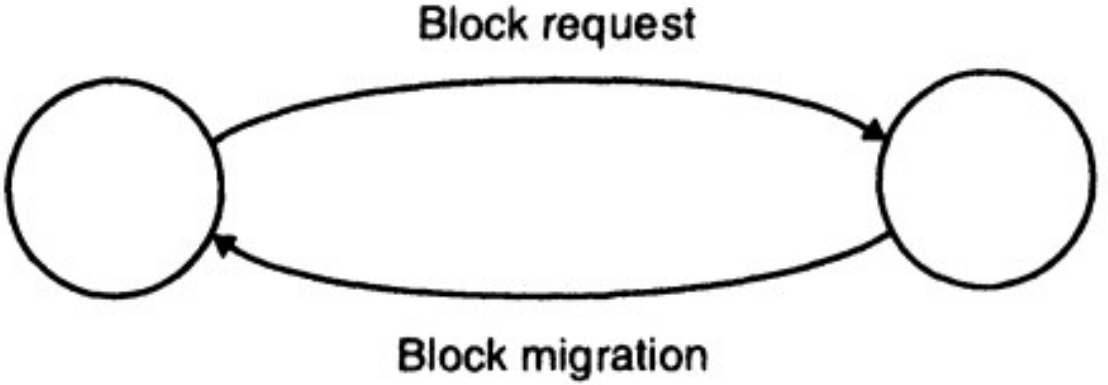
Nonreplicated, Migrating Blocks

- In this strategy each block of the shared memory has a single copy in the entire system.
- However, each access to a block causes the block to migrate from its current node to the node from where it is accessed.
- Therefore, unlike the previous strategy in which the owner node of a block always remains fixed,
- in this strategy the owner node of a block changes as soon as the block is migrated to a new node

NRMB Cntd...

Client node
(becomes new owner node of block after its migration)

Owner node
(owns the block before its migration)



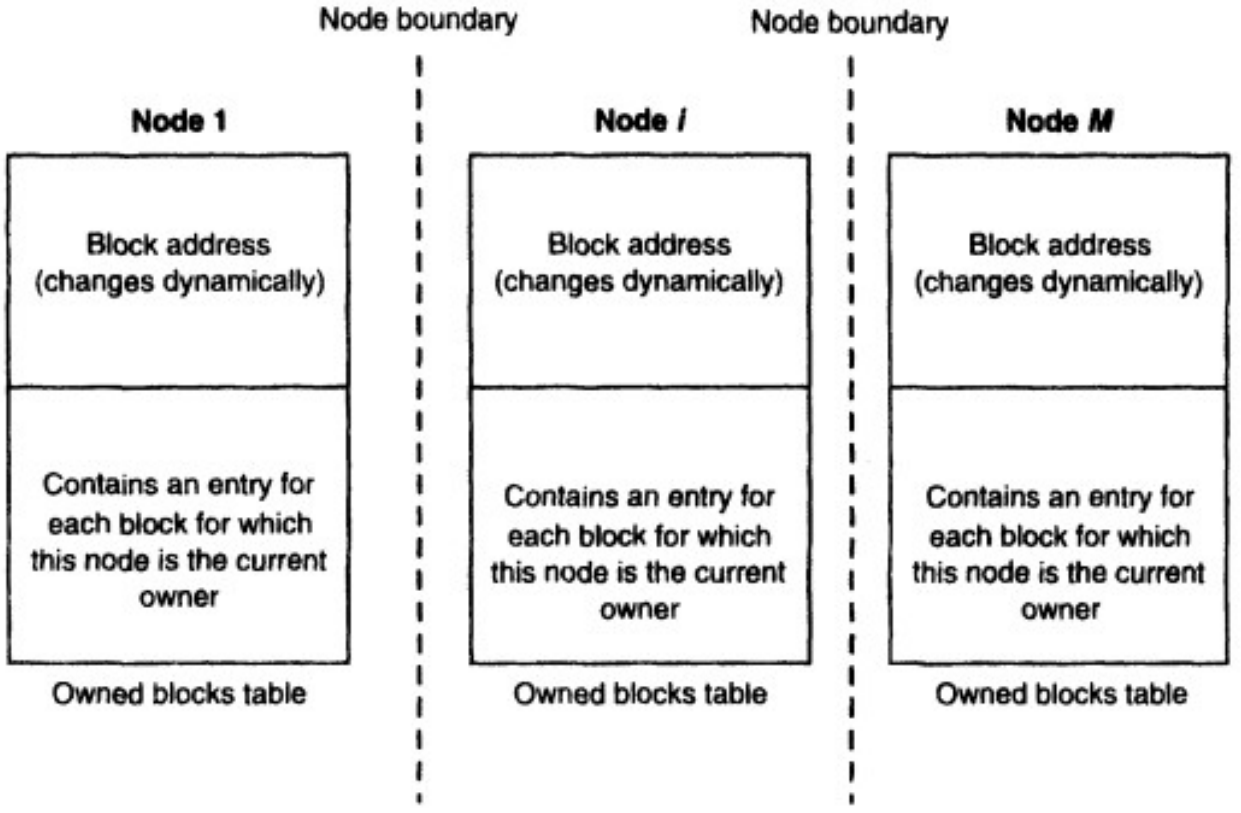
NRMB cntd...

- The method has the following **advantages** [Stumm and Zhou 1990]:
- 1. No communication costs are incurred when a process accesses data currently held locally.
- 2. It allows the applications to take advantage of data access locality. If an application exhibits high locality of reference, the cost of data migration is amortized over multiple accesses.
- However, the method suffers from the following **drawbacks:**
- 1. It is prone to thrashing problem. That is, a block may keep migrating frequently from one node to another, resulting in few memory accesses between migrations and thereby poor performance.
- 2. The advantage of parallelism cannot be availed in this method also.

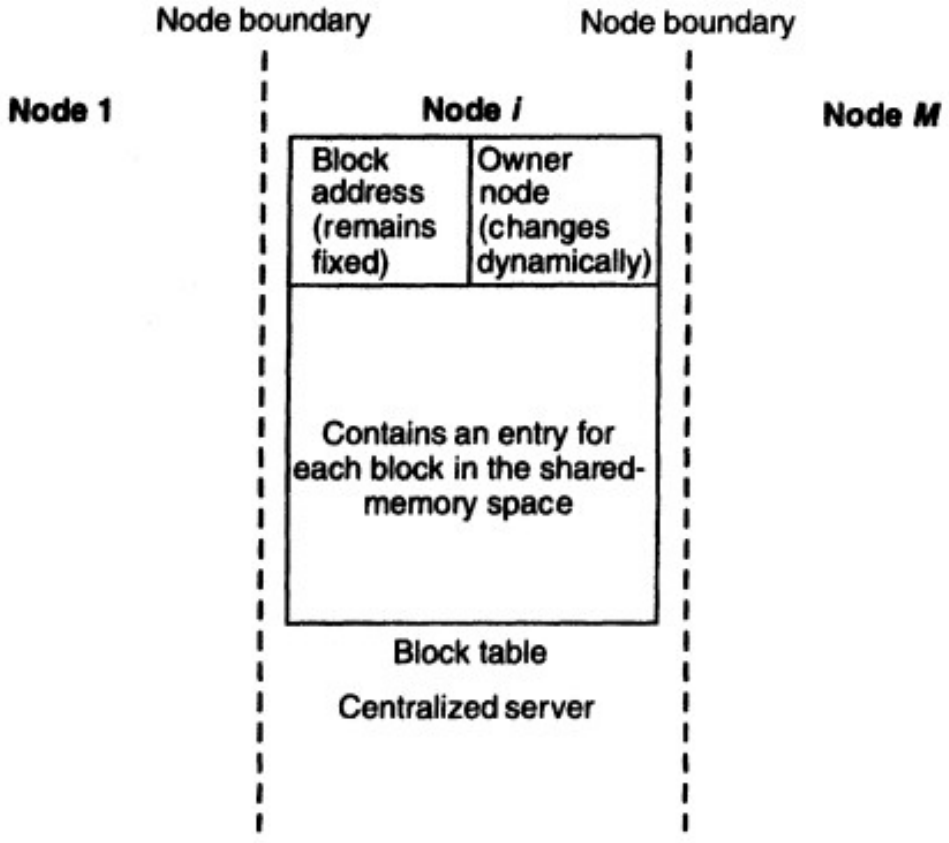
Data Locating in the NRMB Strategy.

- 1. Broadcasting.
- 2. Centralized-server algorithm.
- 3. Fixed distributed-server algorithm.
- 4. Dynamic distributed-server algorithm.

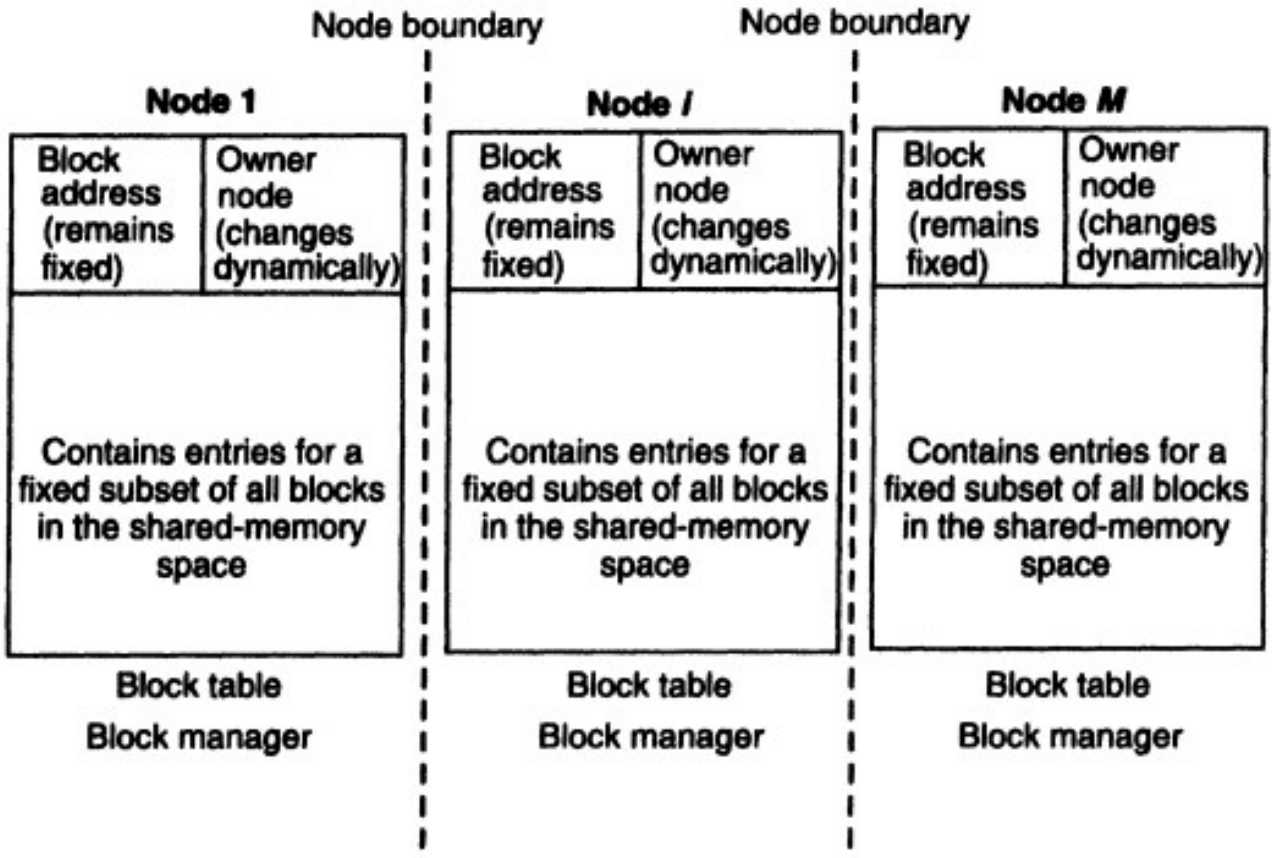
NRMB.....Broadcasting



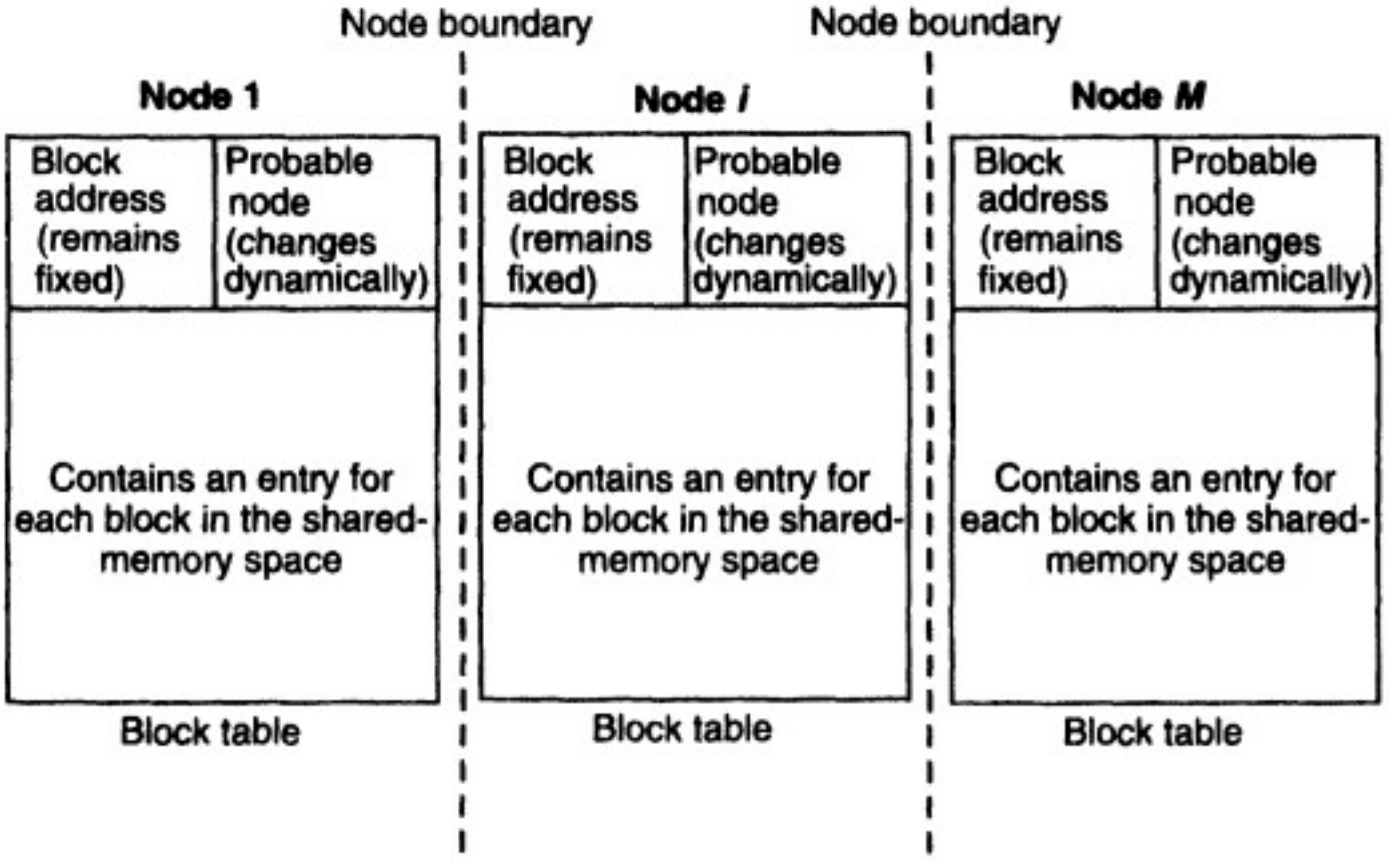
NRMB.....Centralized-server algorithm



NRMB....Fixed distributed-server algorithm



NRMB....Dynamic distributed-server algorithm.



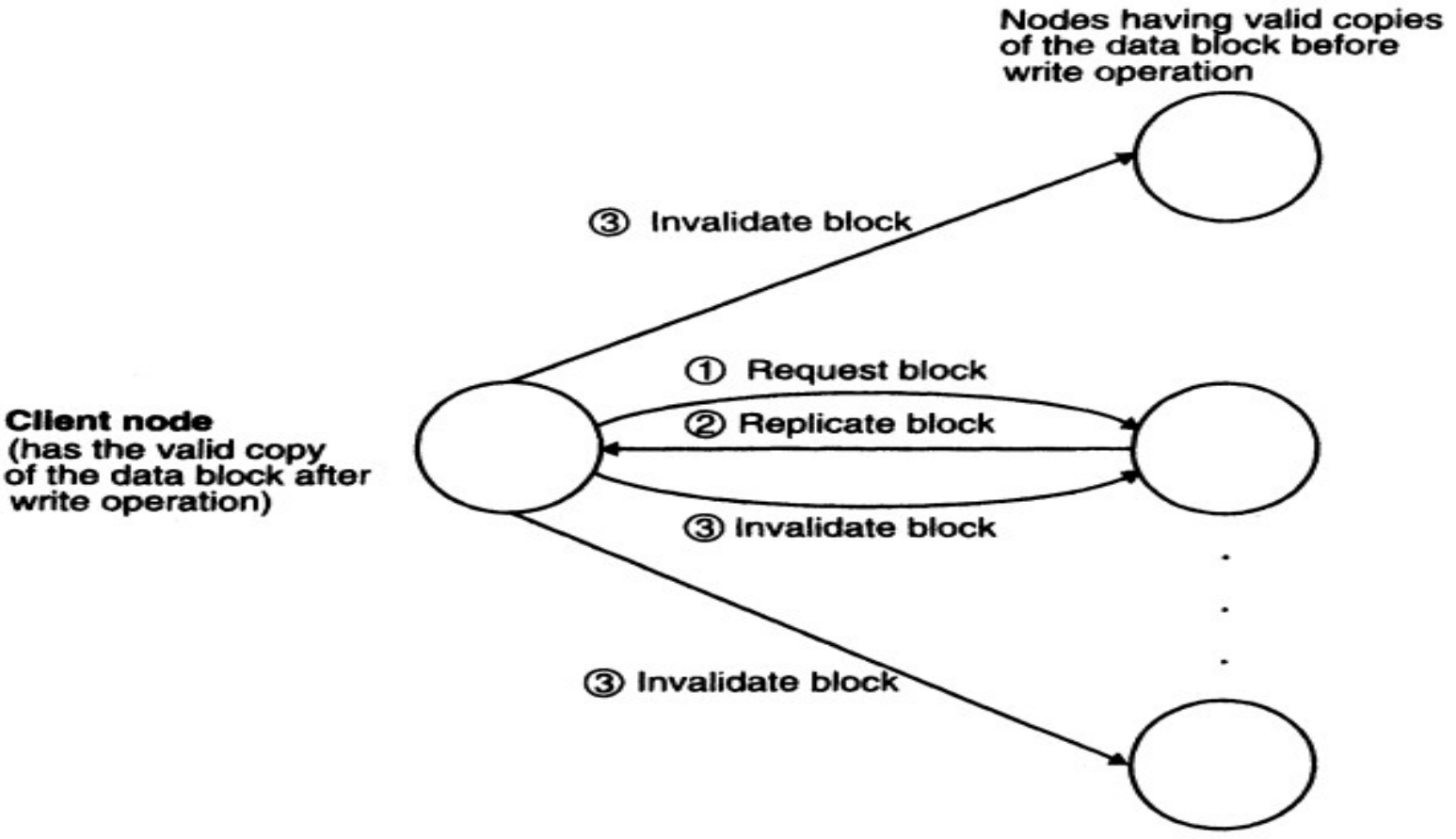
Replicated, Migrating Blocks (RMB)

- A major disadvantage of the nonreplication strategies is lack of parallelism because only the processes on one node can access data contained in a block at any given time.
- To increase parallelism, virtually all DSM systems replicate blocks.
- With replicated blocks, read operations can be carried out in parallel at multiple nodes by accessing the local copy of the data.
- Therefore, the average cost of read operations is reduced because no communication overhead is involved if a replica of the data exists at the local node.
- However, replication tends to increase the cost of write operations because for a write to a block all its replicas must be invalidated or updated to maintain consistency.

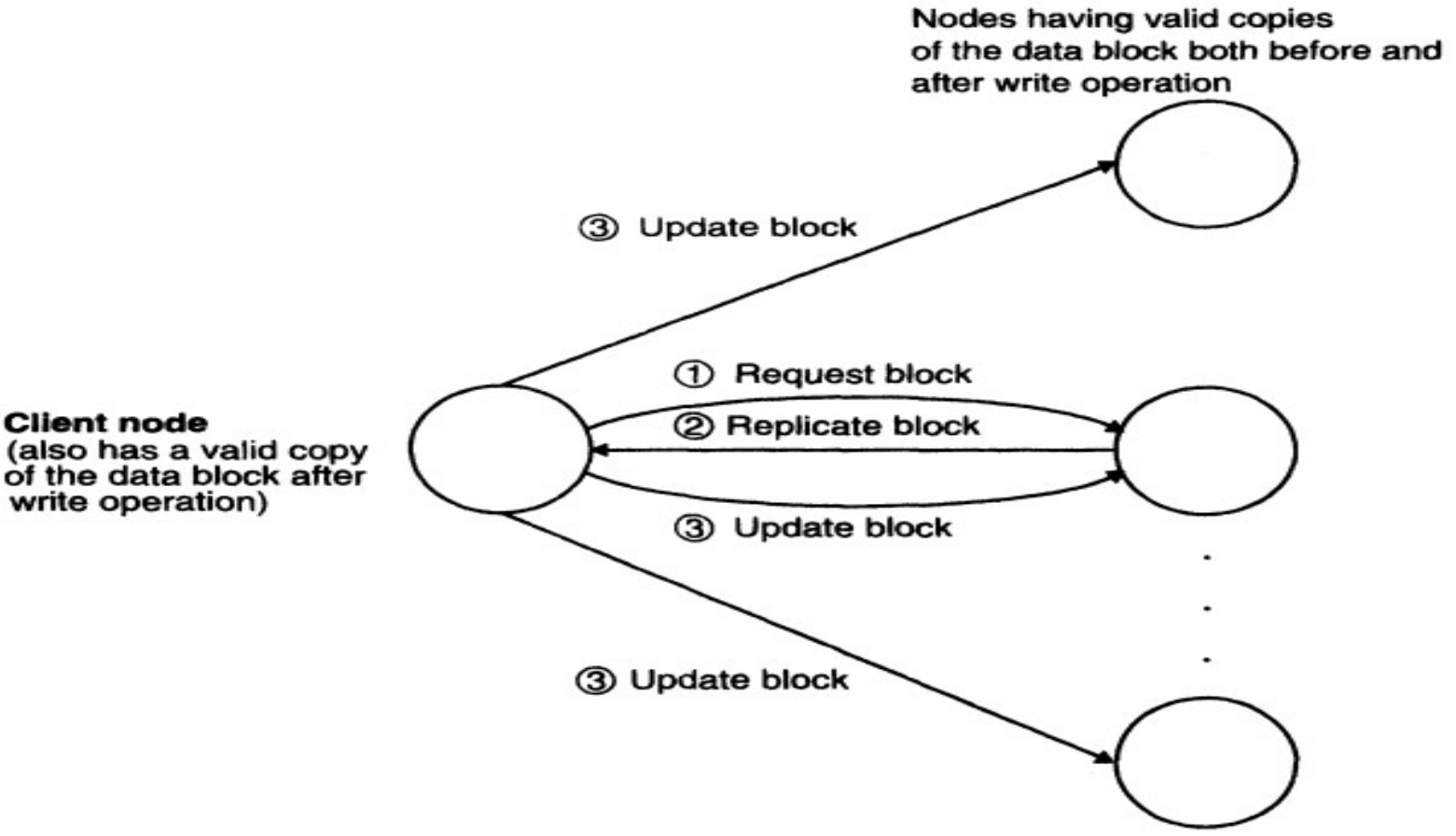
RMB cntd...

- The two basic protocols that may be used for ensuring sequential consistency in this case are as follows:
- 1. Write-invalidate.
- 2. Write-update.

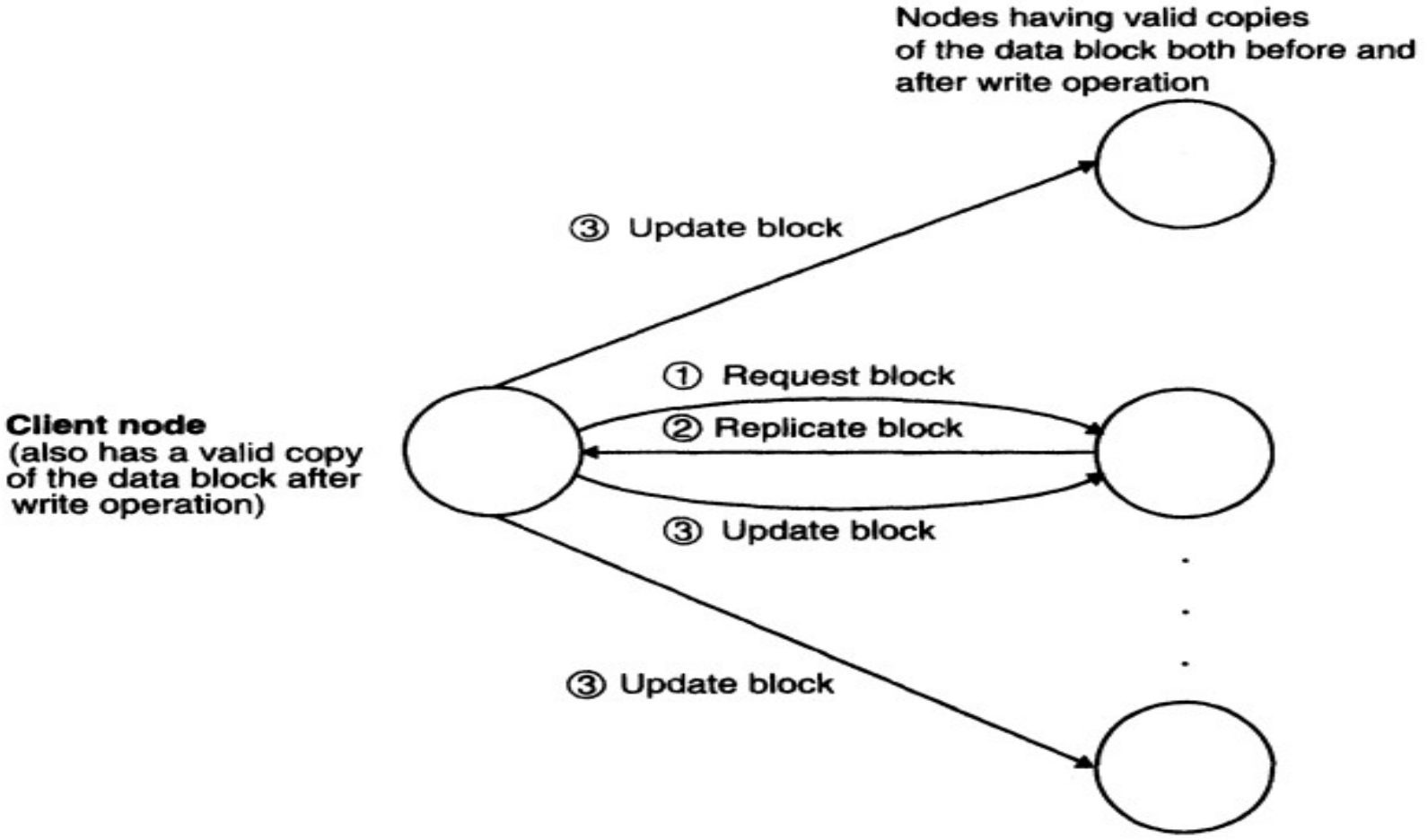
RMB....Write-invalidate



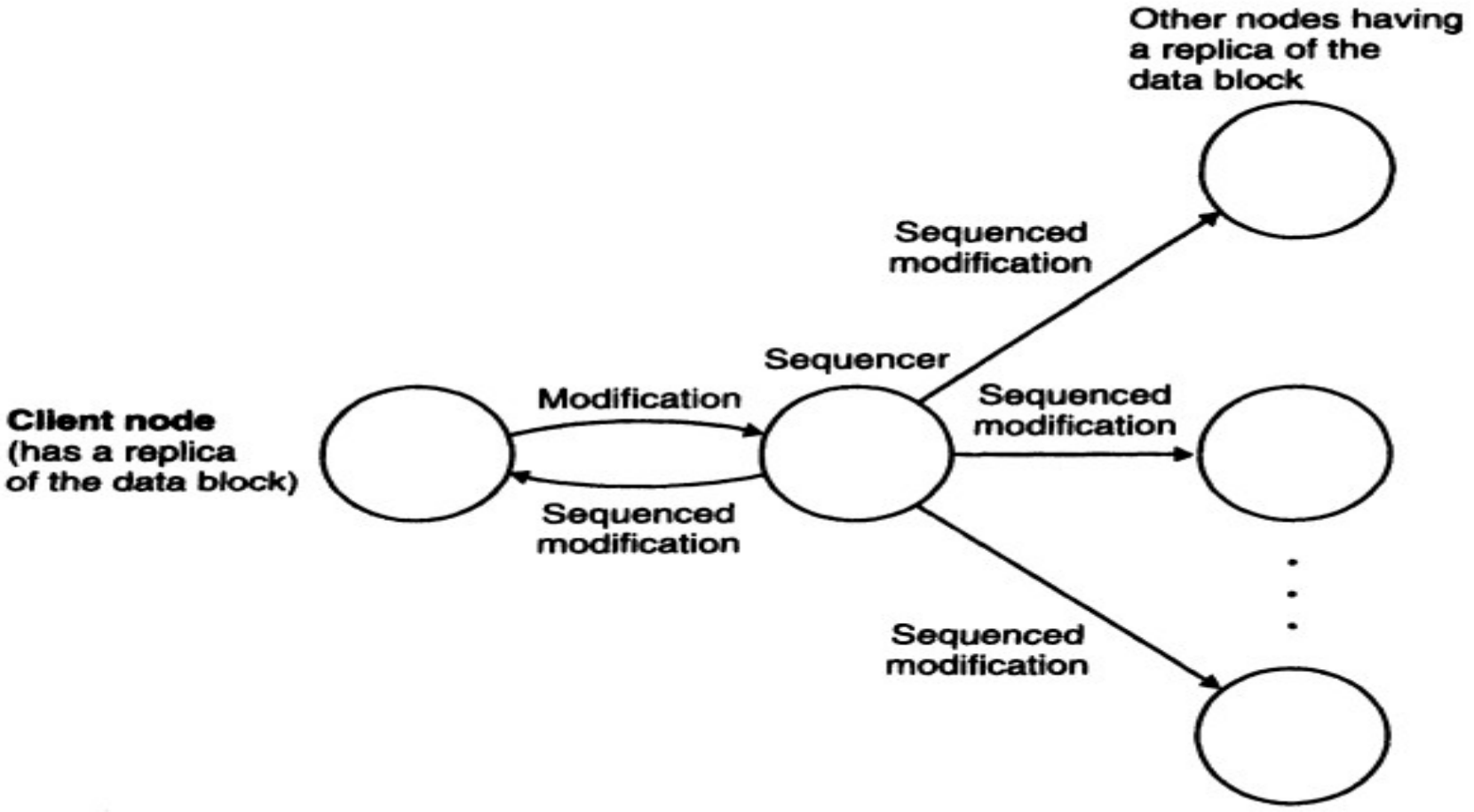
RMB....Write-update.



RMB....Write-update.



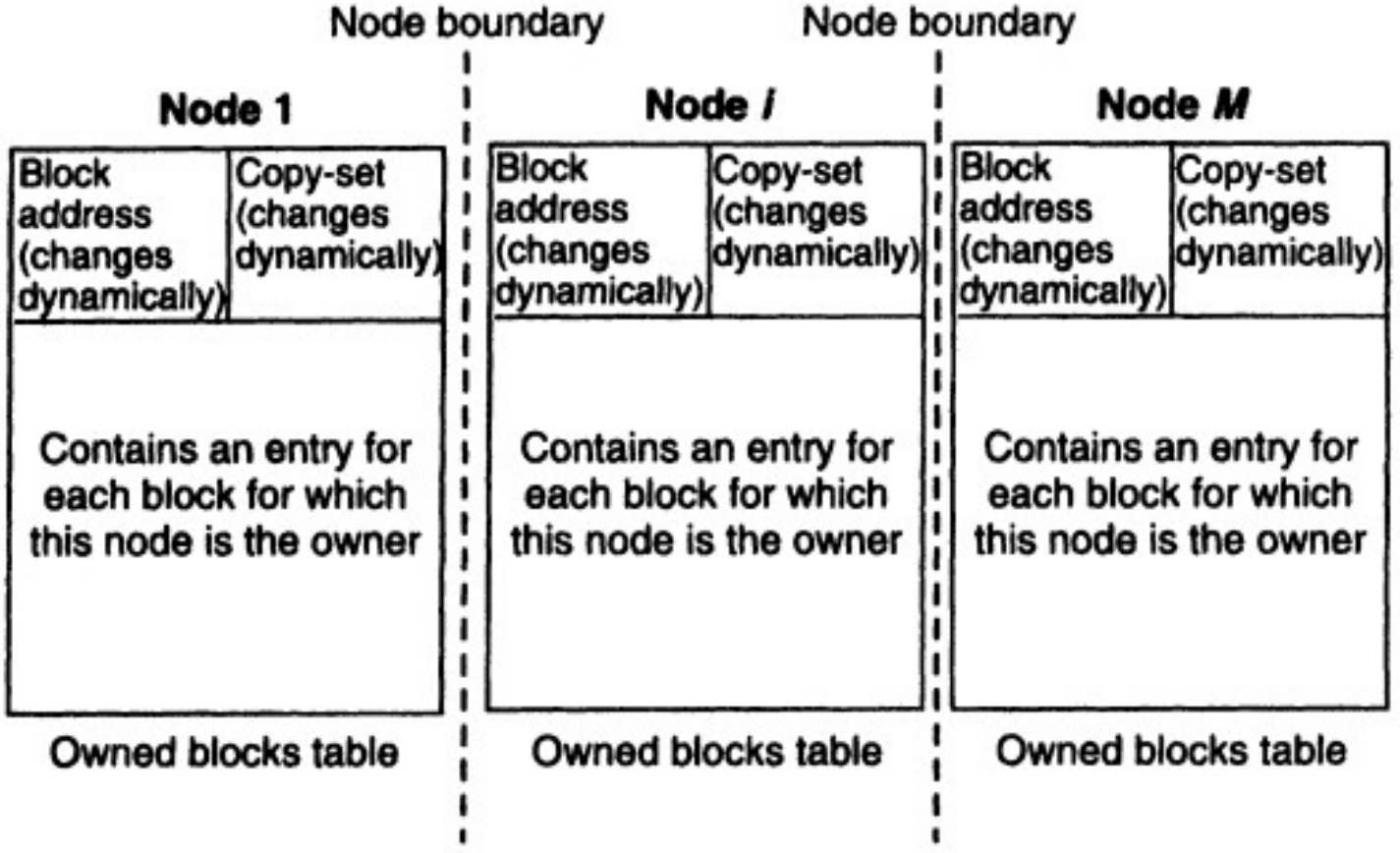
RMB....Write-update.



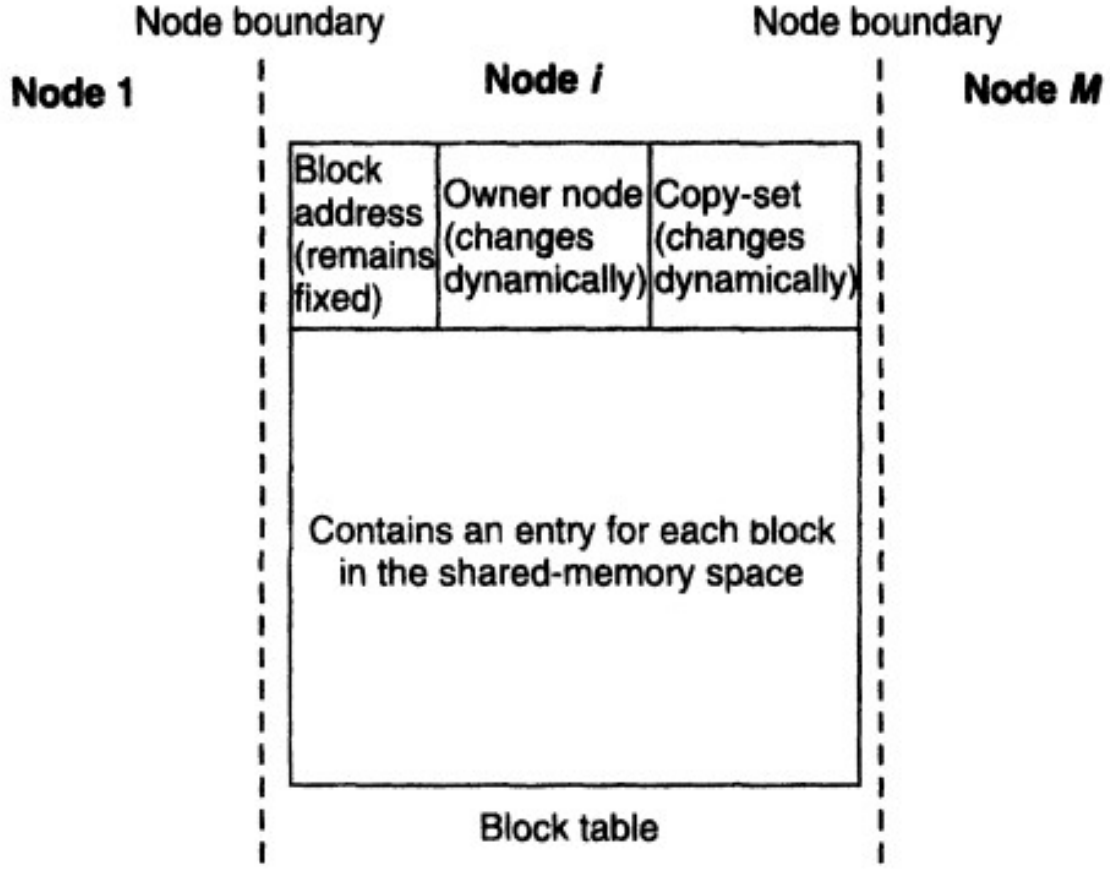
RMB...Data Locating issues

- The following data-locating issues are involved in the write-invalidate protocol used with the RMB strategy:
- 1. Locating the owner of a block. An owner of a block is the node that owns the block, namely, the most recent node to have write access to it.
- 2. Keeping track of the nodes that currently have a valid copy of the block.
- One of the following algorithms may be used to address these two issues:
- 1. Broadcasting.
- 2. Centralized-server algorithm.
- 3. Fixed distributed-server algorithm.
- 4. Dynamic distributed-server algorithm.

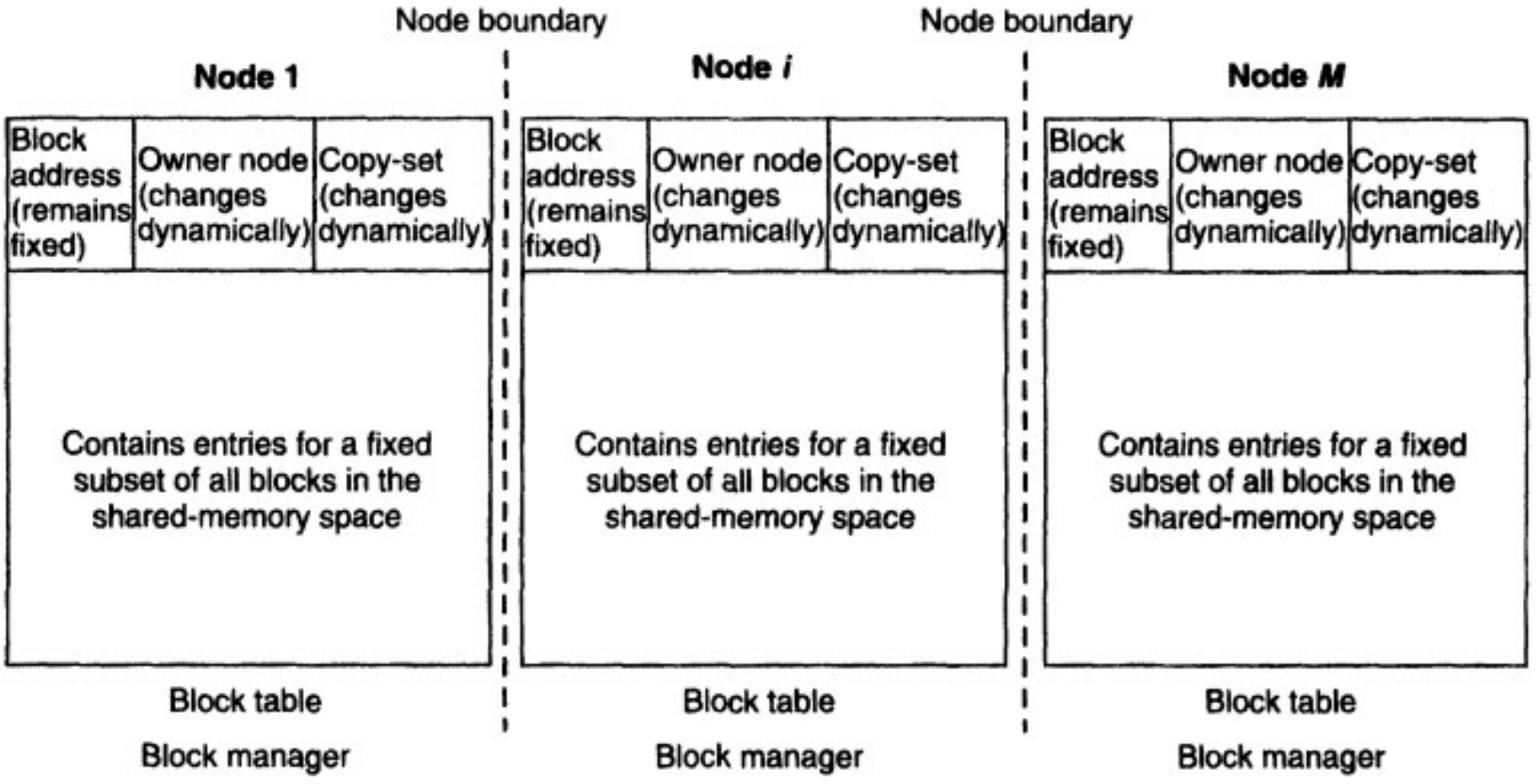
RMB.....Broadcasting



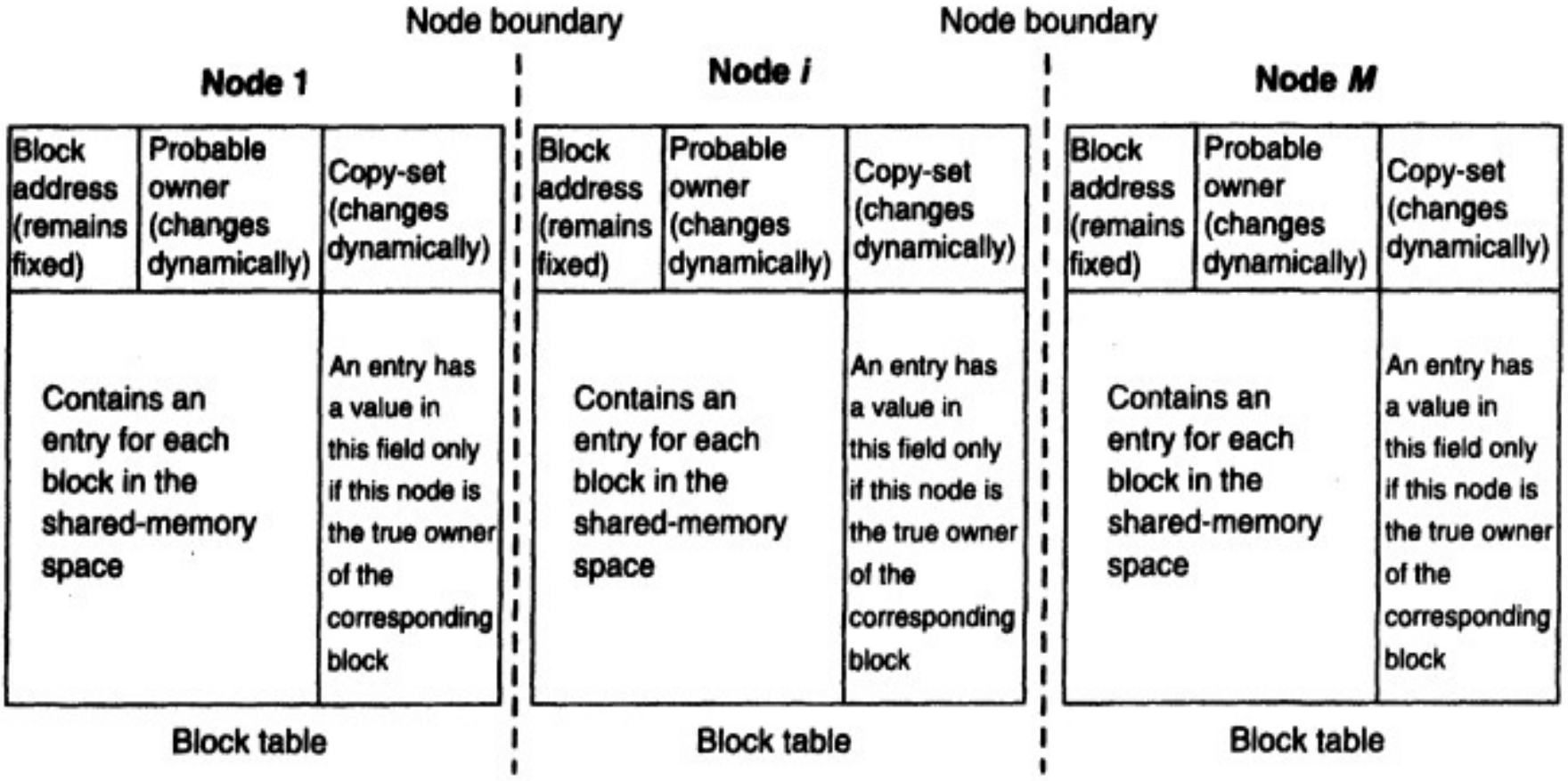
RMB.....Centralized-server algorithm



RMB....Fixed distributed-server algorithm



RMB....Dynamic distributed-server algorithm.



Replicated, Non Migrating Blocks (RNMB)

- In this strategy, a shared-memory block may be replicated at multiple nodes of the system, but the location of each replica is fixed.
- A read or write access to a memory address is carried out by sending the access request to one of the nodes having a replica of the block containing the memory address.
- All replicas of a block are kept consistent by updating them all in case of a write access.
- A protocol similar to the write-update protocol is used for this purpose. Sequential consistency is ensured by using a global sequencer to sequence the write operations of all nodes

Data Locating in the RNMB Strategy.

- The RNMB strategy has the following characteristics:
- 1. The replica locations of a block never change.
- 2. All replicas of a data block are kept consistent.
- 3. Only a read request can be directly sent to one of the nodes having a replica of the block containing the memory address on which the read request is performed and all write requests have to be first sent to the sequencer.

Data Locating in the RNMB Strategy.

- the best approach of data locating for handling read/ write operations in this case is to have
- a block table at each node and
- a sequence table with the sequencer

Structure and locations of block table and sequence table in the centralized sequencer data-locating mechanism for RNMB strategy