Mobile and Heterogeneous databases

Heterogeneous Distributed Databases

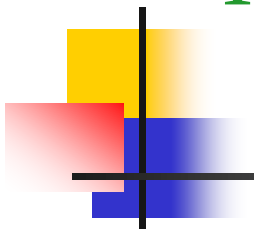
Transaction Processing

A.R. Hurson

Computer Science

Missouri Science & Technology

Heterogeneous Distributed Databases

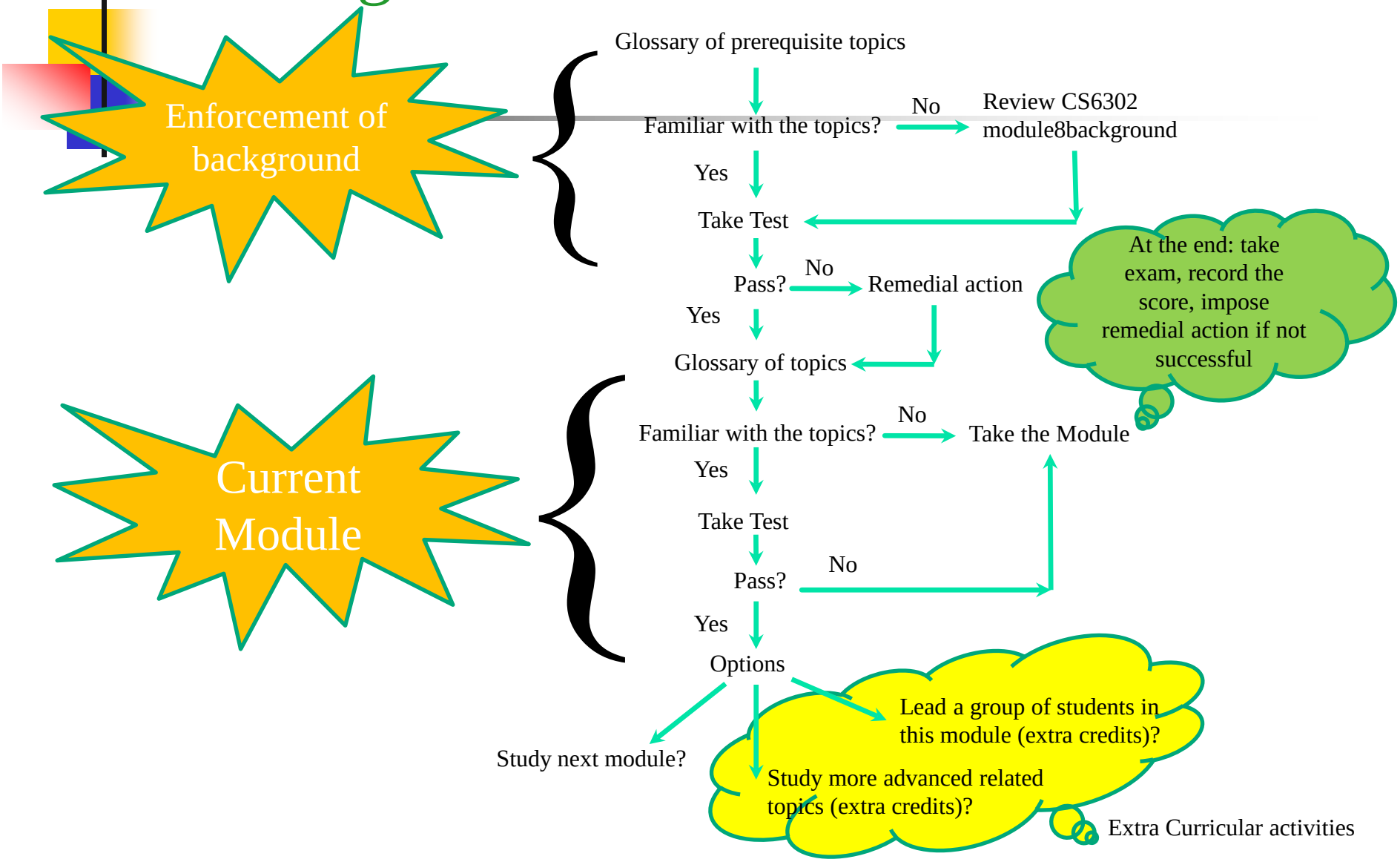


Note, this unit will be covered in two lectures. In case you finish it earlier, then you have the following options:

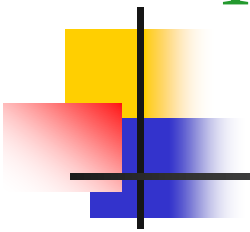
1) Take the early test and start CS6302.module9
2) Study the supplement module (supplement CS6302.module8)
3) Act as a helper to help other students in studying CS6302.module8

Note, options 2 and 3 have extra credits as noted in course outline.

Heterogeneous Distributed Databases

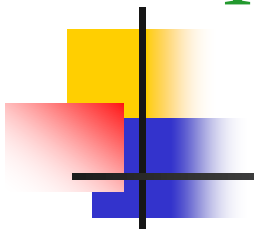


Heterogeneous Distributed Databases



- You are expected to be familiar with:
  - Heterogeneous Distributed Databases,
- If not, you need to study CS6302.module3, and CS6302.module7

Heterogeneous Distributed Databases



Access and manipulate data across multiple heterogeneous pre-existing data sources in sometimes, somewhere and anytime anywhere, access environment:

- Multidatabase System (MDBS)
- Mobile DataAccess System (MDAS)

Need an efficient global transaction management scheme with an eye on issues such as:

- Heterogeneity, autonomy, network bandwidth, frequent disconnections, and limited processing power and resources.

Heterogeneous Distributed Databases

The Summary Schemas Model has been developed as an extension to multidatabase systems to provide linguistic support to automatically identify semantically similar entities with different access terms.

Heterogeneous Distributed Databases

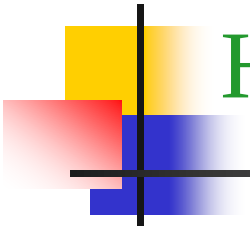
Summary Schemas Model – Advantages

- The meta-data is by orders of magnitude smaller than the meta-data generated by the Global-schema approach.
- Preserves local DBMS autonomy.
- Provides good system scalability.
- Reduces average search time.
- Resolves imprecise queries.

Heterogeneous Distributed Databases

Summary Schemas Model

- Prototyped a Client/Server based SSM
- Access Control and Security in SSM
- Transaction management in SSM
- Query processing in SSM



Heterogeneous Distributed Databases

Summary Schemas Model

- Lessons learned
  - Lack of portability,
  - Lack of stability,
  - Relying on network connectivity.

Heterogeneous Distributed Databases

Summary Schemas Model

- Current distributed and multidatabase systems are designed to allow timely and reliable access to large amounts of autonomous and heterogeneous data sources
  — Sometimes, somewhere access environment.

Heterogeneous Distributed Databases

Summary Schemas Model

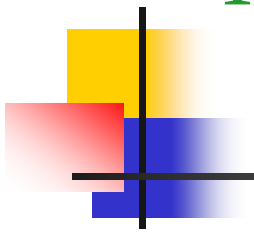
- Adding mobility and wireless connection to the traditional multidatabase environments allow anytime, anywhere access to the information sources.
- However, this advantage comes at the expense of additional complexities due to the network bandwidth, frequent disconnections, and limited processing power and resources.

Heterogeneous Distributed Databases

Extended Summary Schemas Model

- Add mobility to the Summary Schemas Model (MDAS)
- Transaction management in MDAS
- Power management
- Application of Software Agent
- Multimedia support
- Community of local ad-hoc nodes/sensor networks

Heterogeneous Distributed Databases

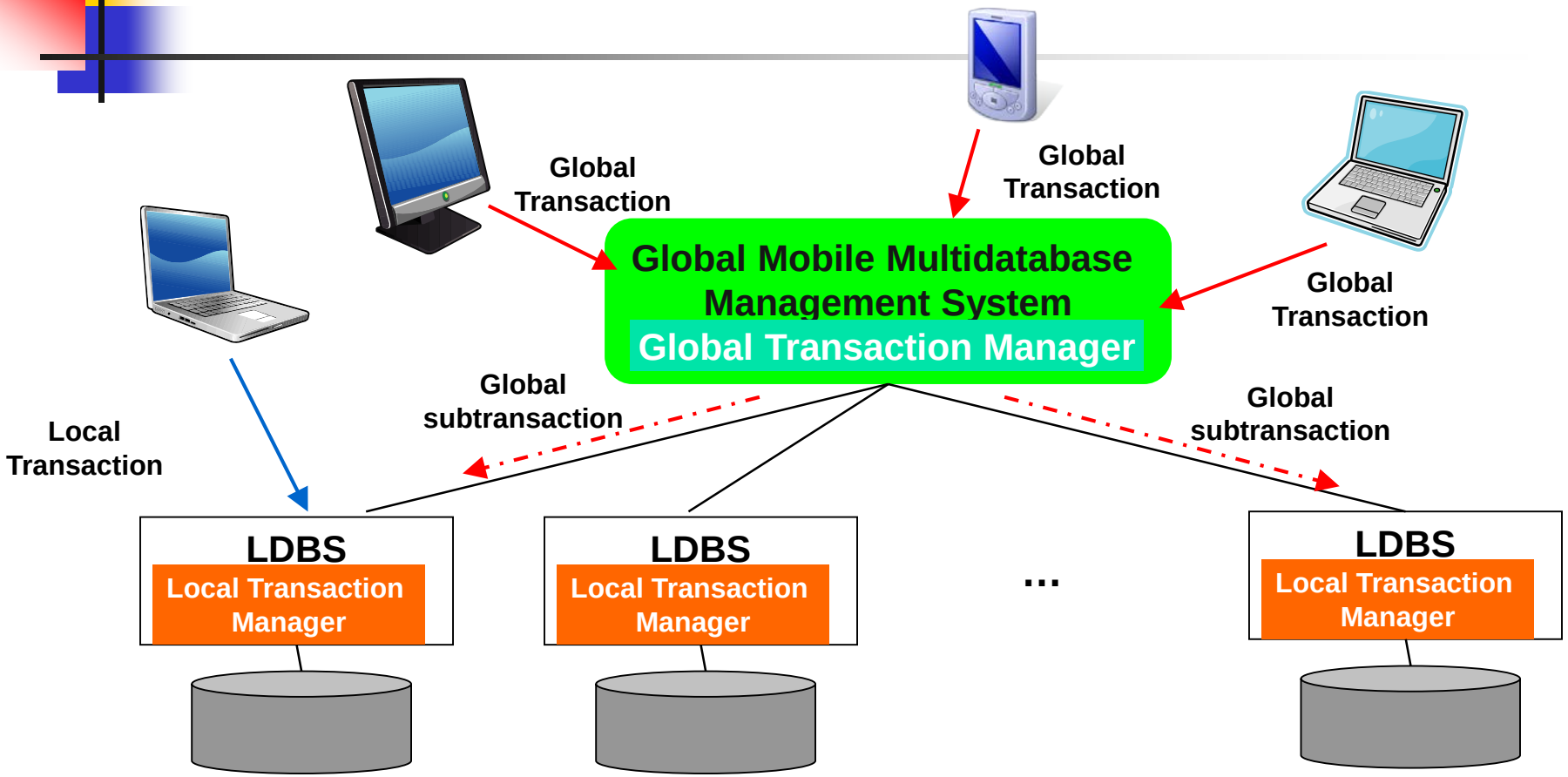


It was decided to use agent technology to overcome technological constraints.

Mobile agent-based computation paradigm responds to the:

- Limited resources of mobile devices by migrating tasks to more powerful servers on the network,
- Intermittent connectivity by supporting disconnected operations, and
- Reduced Bandwidth by moving computation to the data.

Heterogeneous Distributed Databases

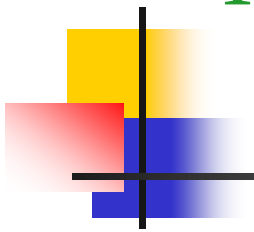


Local databases residing on wired network receiving transactions from both fixed and mobile clients.

Heterogeneous Distributed Databases

- Autonomy and heterogeneity
  - Direct and Indirect Conflicts
- Intermittent and unreliable network connectivity
  - Excessive transaction aborts
  - Long-lived Transactions (LLT)
  - Overcommitted resources
- Power and processing constraints
- Lower bandwidth

Heterogeneous Distributed Databases



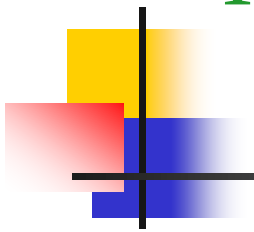
Efficient access and manipulation of data across multiple heterogeneous and autonomous databases via wireless connections should:

Preserve ACID properties.

Preserve autonomy of the local databases as much as possible.

Accommodate frequent disconnections.

Heterogeneous Distributed Databases



Non-blocking scheme at the global level.

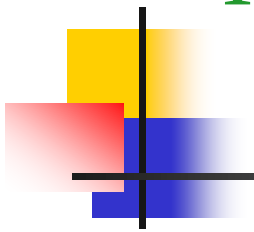
Parallel processing of global sub-transactions.

Reduced Communication (Agents can make local decision, avoiding acknowledgements).

No cascading Abort.

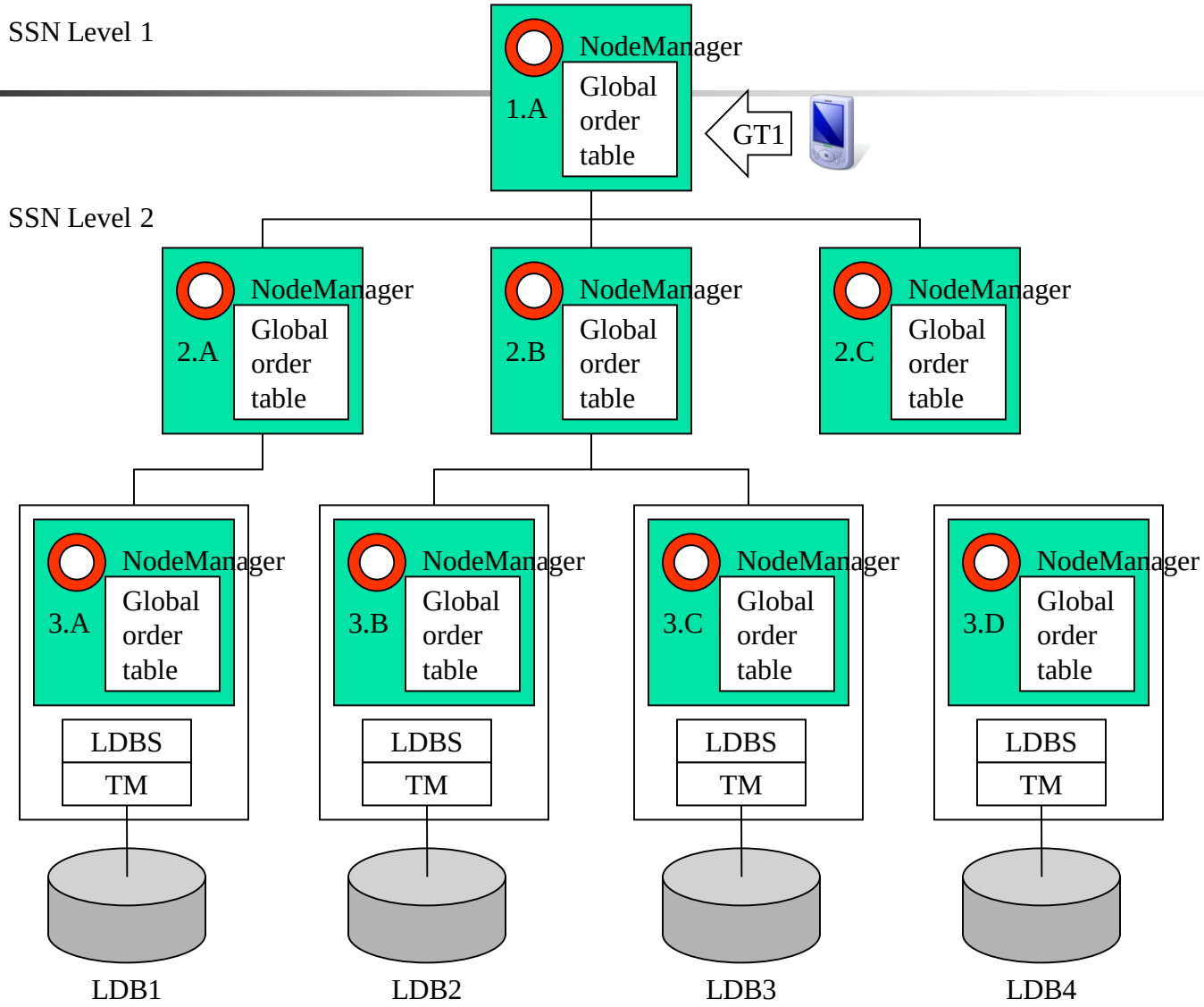
Mobility Support.

Heterogeneous Distributed Databases

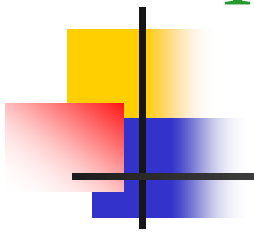


- Summary Schema Model (SSM)
- Global Transaction Agent
- Global Sub-Transaction Agent
- NodeManager
- Global Order Table

Heterogeneous Distributed Databases



Heterogeneous Distributed Databases



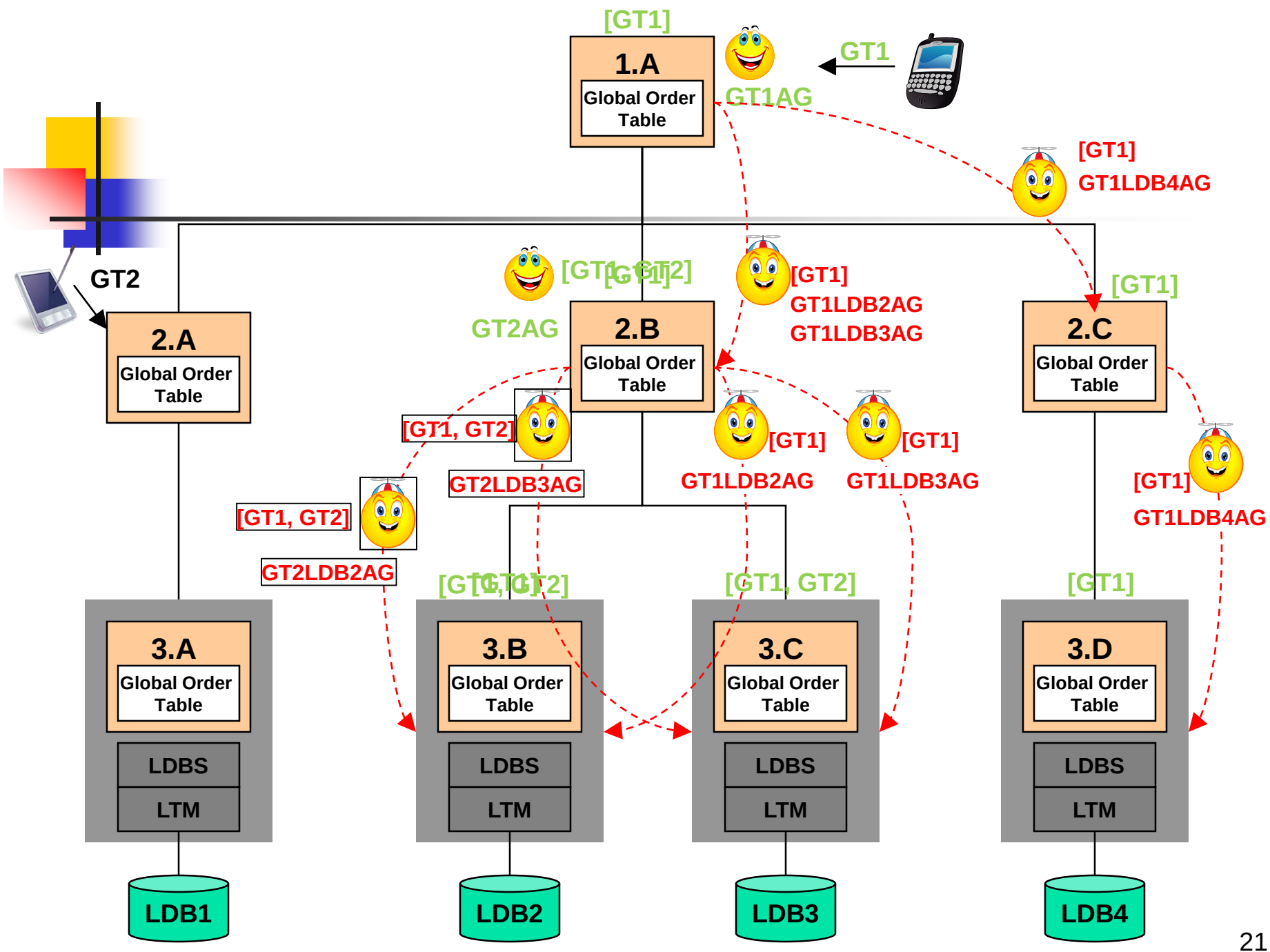
Phase1: Transaction Resolution and selection of the Global Transaction Coordinator

Phase2: Global Order

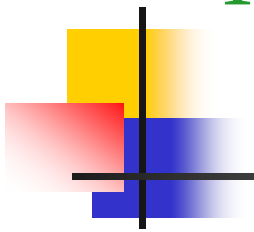
AT3M uses pessimistic approach.

Collaboration between NodeManagers and GSTAgent to perform a timestamp-based algorithm and determine a global execution order before the execution of the global sub-transactions.

Global order is enforced through out the hierarchy.


[GT1]
1.A
Global Order Table
GT1
GT1AG
[GT1]
GT1LDB4AG
GT2
[GT1]
[GT1, GT2]
GT2AG
2.A
Global Order Table
2.B
Global Order Table
[GT1]
GT1LDB2AG
GT1LDB3AG
[GT1]
2.C
Global Order Table
[GT1, GT2]
GT2LDB3AG
[GT1]
GT1LDB2AG
[GT1]
GT1LDB3AG
[GT1]
GT1LDB4AG
[GT1, GT2]
GT2LDB2AG
[GT1, GT2]
[GT1, GT2]
[GT1]
3.A
Global Order Table
3.B
Global Order Table
3.C
Global Order Table
3.D
Global Order Table
LDBS
LTM
LDBS
LTM
LDBS
LTM
LDBS
LTM
LDB1
LDB2
LDB3
LDB4

Heterogeneous Distributed Databases



Phase3: Global Order Enforcement

NodeManagers at leaf node preserve global serialization.

Phase4: Global Commit

Each GSTAgent after execution enters in the prepare-to-commit stage and reports back to the GTAgent.

The GTAgent after receiving all responses initiates global commit command.

The GSTAgent receiving global commit command commits the global subtransaction and terminates.

GTAgent receiving any abort request initiates global abort.

Heterogeneous Distributed Databases

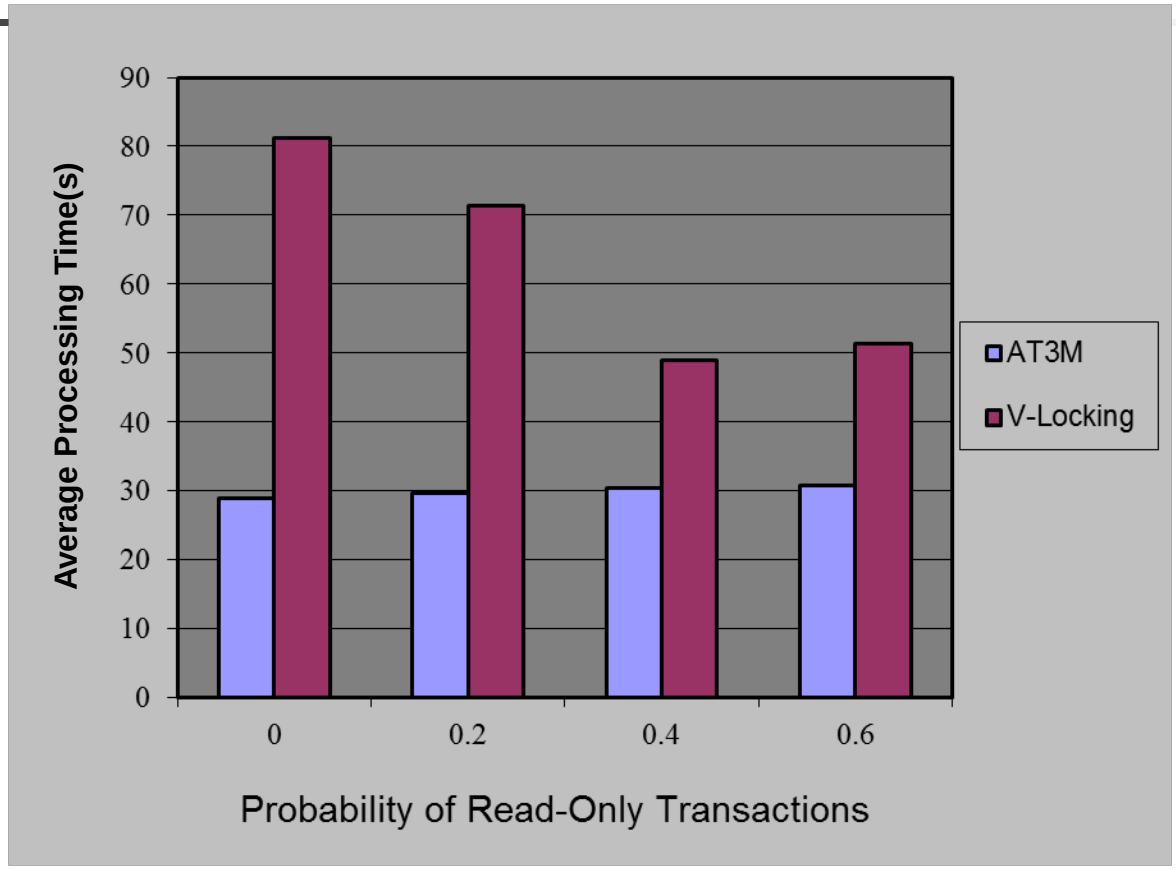
A simulator was developed using SimJava 2.0.

The mobile multidatabase system consists of several local databases.

Each local database contains 100 data items, 20 items are hot-spot (likely to be accessed).

AT3M is simulated and compared against V-Locking and Pre-Serialization algorithms.

Heterogeneous Distributed Databases

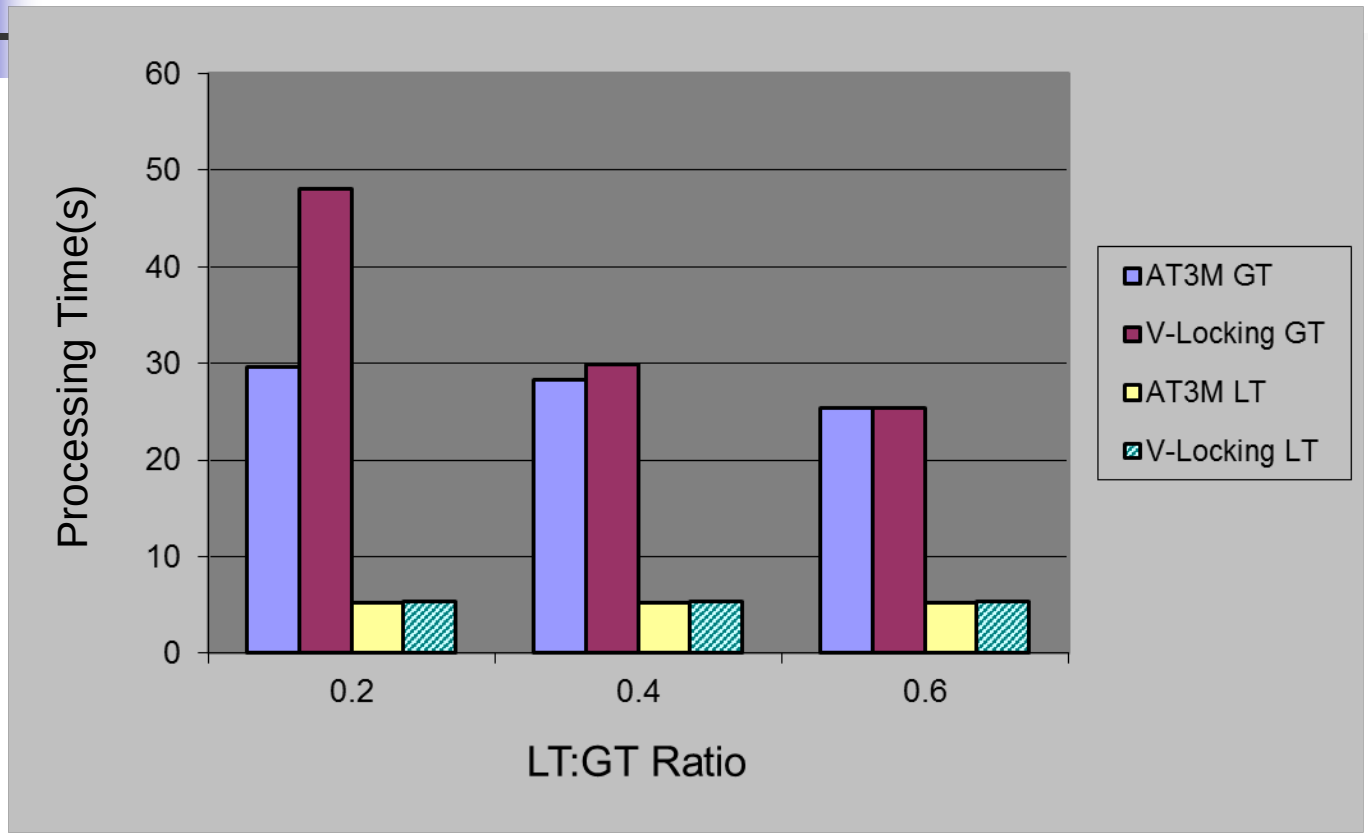


Average Processing Time *VS.* Probability of Read-Only Transactions

Heterogeneous Distributed Databases

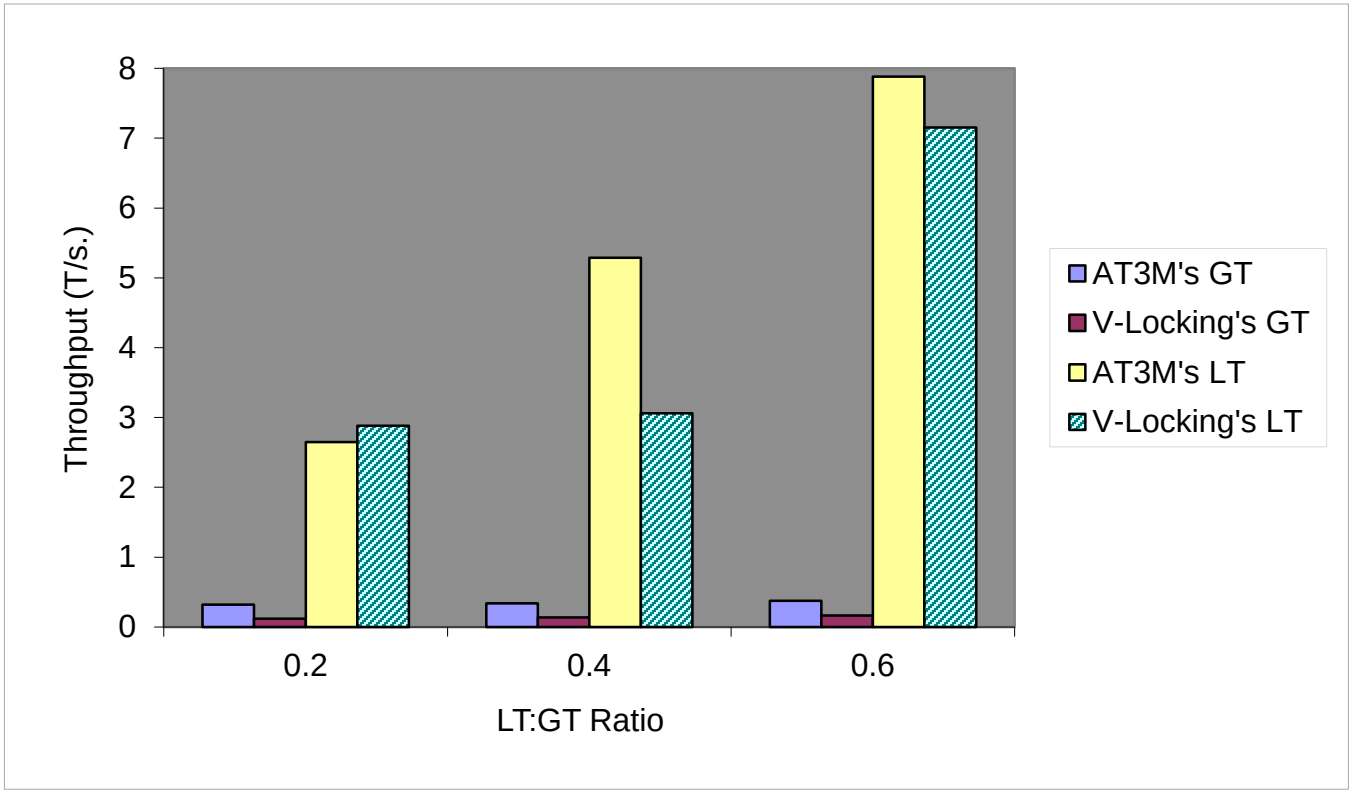
Number of Communication Messages per Transaction *VS.* Probability of Read-Only Transactions

Heterogeneous Distributed Databases



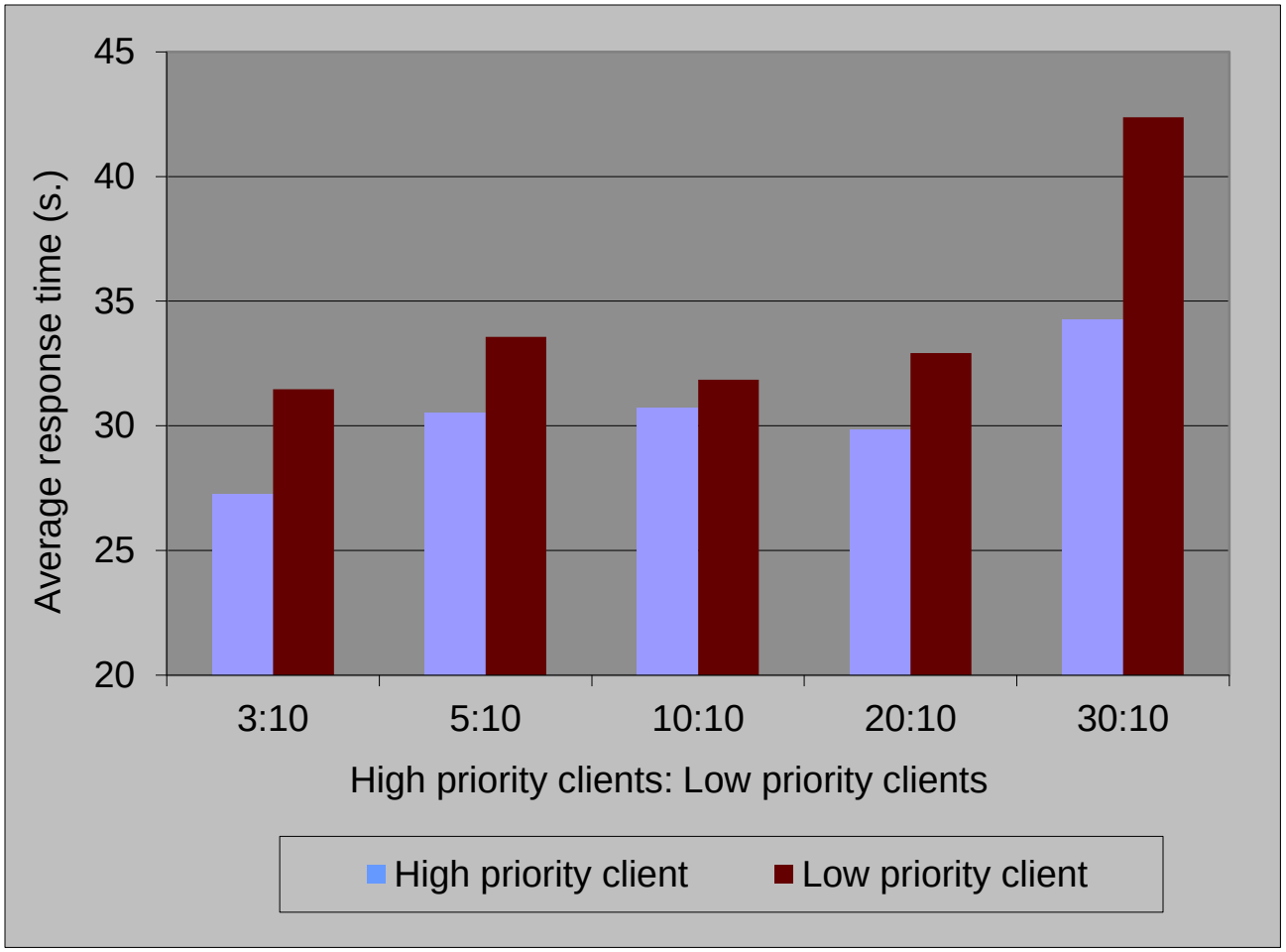
Average Processing Time for Different Levels of Local And Global Transactions Ratio

Heterogeneous Distributed Databases

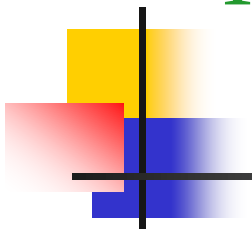


Throughput for Different Levels of Local And Global Transactions Ratio

Heterogeneous Distributed Databases



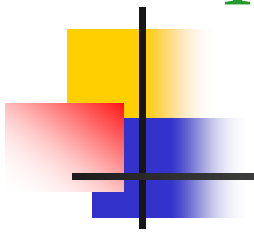
Heterogeneous Distributed Databases



An agent-based pessimistic non-locking approach that preserves the local autonomy and technological constraints.

It is a promising approach as it provides better average processing time and lower number of communication messages.

Heterogeneous Distributed Databases



Extension of AT3M to operate in pervasive computing environment.

Extension of AT3M to operate in environment with replicated databases.