Outline

- Introduction
- Background
- Distributed DBMS Architecture
- Distributed Database Design
- Distributed Query Processing
- Distributed Transaction Management
  - Transaction Concepts and Models
  - Distributed Concurrency Control
  - Distributed Reliability
- Building Distributed Database Systems (RAID)
- Mobile Database Systems
- Privacy, Trust, and Authentication
- Peer to Peer Systems

Useful References

- Textbook *Principles of Distributed Database Systems*,

  Chapter 12.1, 12.2
- J. Gray and A. Reuter. *Transaction Processing - Concepts and Techniques*. Morgan Kaufmann, 1993. (Copy on reserve in MATH library)
- Bharat Bhargava (Ed.), *Concurrency Control and Reliability in Distributed Systems*, Van Nostrand and Reinhold Publishers, 1987. (Copy on reserve in LWSN reception office book shelf)

Reliability

In case of a crash, recover to a consistent (or correct state) and continue processing.

Types of Failures

1. Node failure
2. Communication line of failure
3. Loss of a message **(or delay of a message)**
4. Network partition
5. Any combination of above

Approaches to Reliability

1. Audit trails (or logs)
2. Two phase commit protocol
3. Retry based on timing mechanism
4. Reconfigure
5. Allow enough concurrency which permits definite recovery (avoid certain types of conflicting parallelism)
6. Crash resistance design

Recovery Controller

Types of failures:

* transaction failure
* site failure (local or remote)
* communication system failure

Transaction failure

UNDO/REDO Logs

transparent transaction

(effects of execution in private workspace)

⇒ Failure does not affect the rest of the system

Site failure

volatile storage lost

stable storage lost

processing capability lost

(no new transactions accepted)

System Restart

Types of transactions:

1. In commitment phase
2. Committed actions reflected in real/stable
3. Have not yet begun
4. In prelude (have done only undoable actions)

We need:

stable undo log; stable redo log (at commit);

perform redo log (after commit)

Problem:

entry into undo log; performing the action

Solution:

undo actions $\neg$ < T, A, E >

must be restartable (or idempotent)

DO – UNDO

≡ UNDO

≡ DO – UNDO – UNDO – UNDO --- UNDO

Site Failures (simple ideas)

Local site failure

- Transaction committed ⇒ do nothing
- Transaction semi-committed ⇒ abort
- Transaction computing/validating ⇒ abort

AVOIDS BLOCKING

Remote site failure

- Assume failed site will accept transaction
- Send abort/commit messages to failed site via spoolers

Initialization of failed site

- Update for globally committed transaction before validating other transactions
- If spooler crashed, request other sites to send list of committed transactions

Communication Failures (simple ideas)

Communication system failure

- Network partition
- Lost message
- Message order messed up

Network partition solutions

- Semi-commit in all partitions and commit on reconnection (updates available to user with warning)
- Commit transactions if primary copy token for all entities within the partition
- Consider commutative actions
- Compensating transactions

Compensation

Compensating transactions

- Commit transactions in all partitions
- Break cycle by removing semi-committed transactions
- Otherwise abort transactions that are invisible to the environment

  (no incident edges)
- Pay the price of committing such transactions and issue compensating transactions

Recomputing cost

- Size of readset/writeset
- Computation complexity

Reliability and Fault-tolerate Parameters

Problem:

How to maintain

atomicity

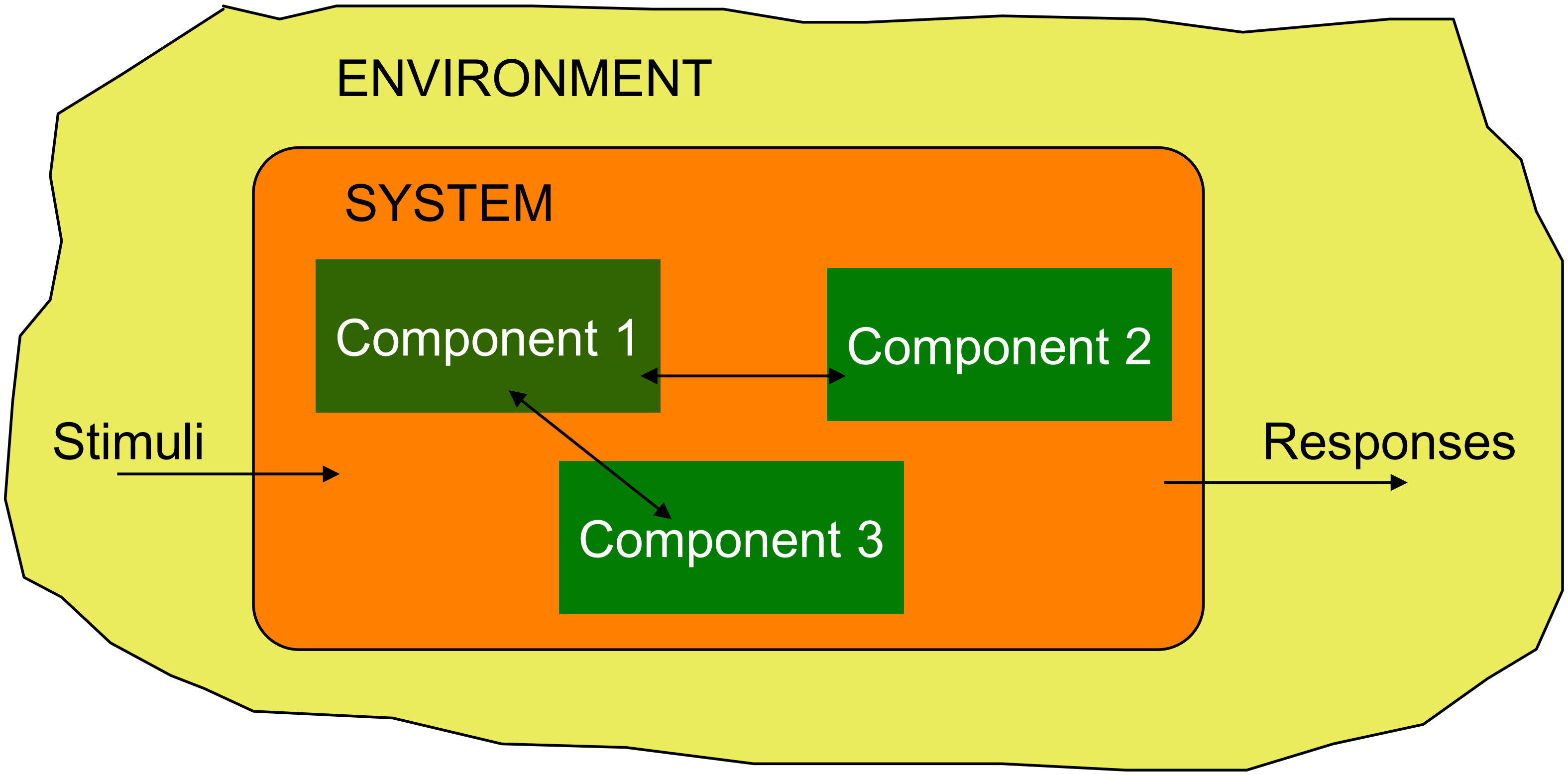
durability

properties of transactions

Fundamental Definitions

- Reliability
  - A measure of success with which a system conforms to some authoritative specification of its behavior.
  - Probability that the system has not experienced any failures within a given time period.
  - Typically used to describe systems that cannot be repaired or where the continuous operation of the system is critical.
- Availability
  - The fraction of the time that a system meets its specification.
  - The probability that the system is operational at a given time t.

Basic System Concepts



External state

Internal state

Fundamental Definitions

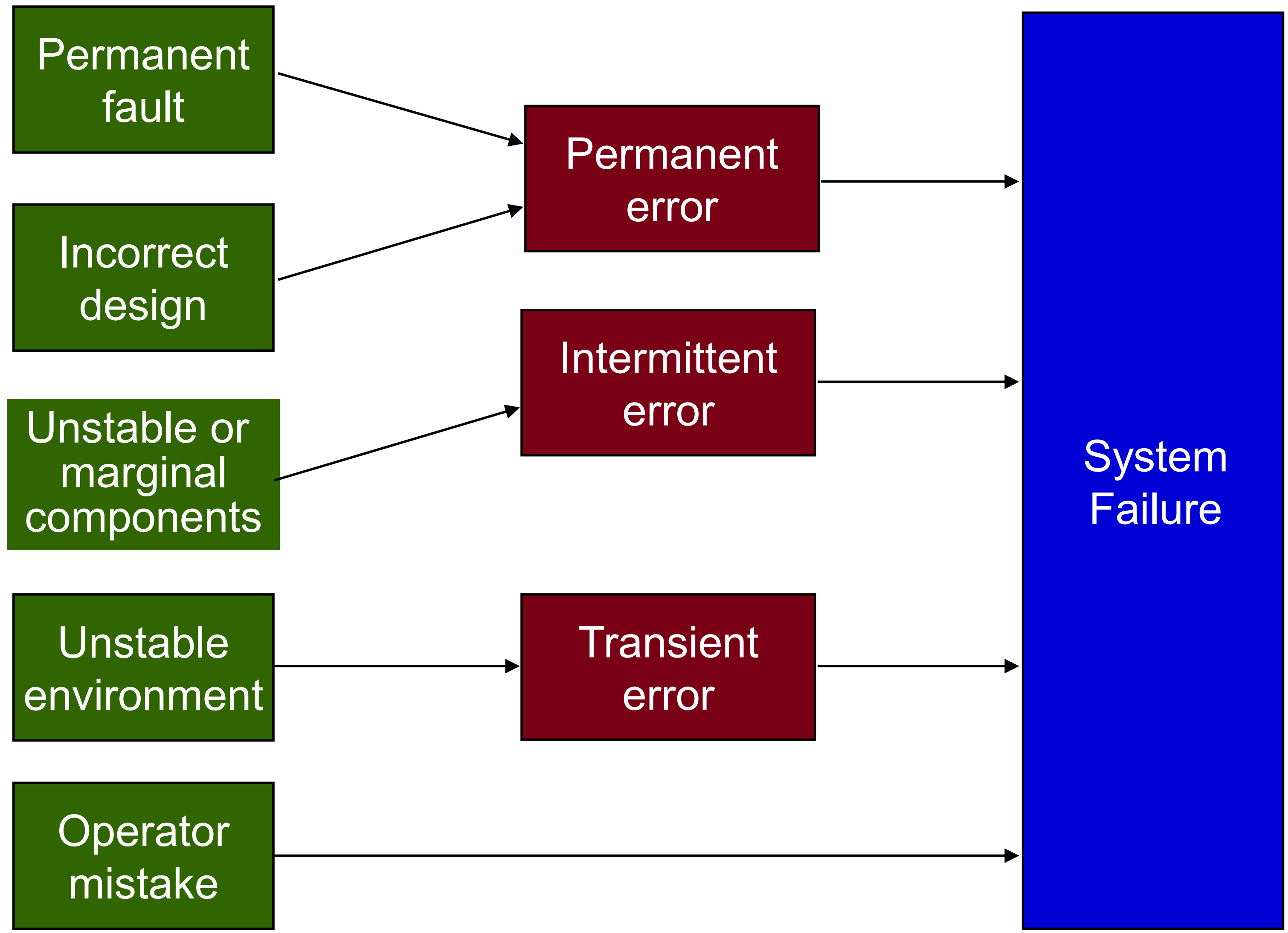
- Failure
  - The deviation of a system from the behavior that is described in its specification.
- Erroneous state
  - The internal state of a system such that there exist circumstances in which further processing, by the normal algorithms of the system, will lead to a failure which is not attributed to a subsequent fault.
- Error
  - The part of the state which is incorrect.
- Fault
  - An error in the internal states of the components of a system or in the design of a system.

Faults to Failures

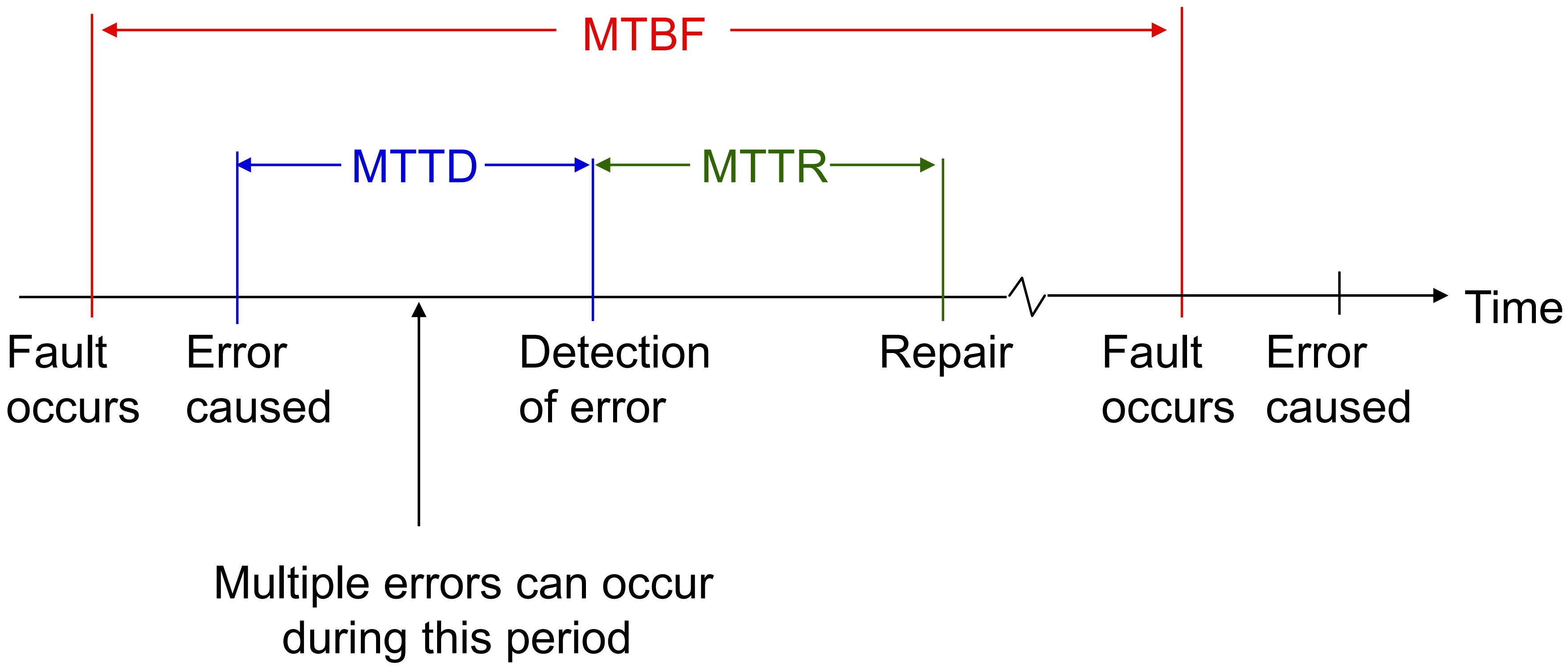
Types of Faults

- Hard faults
  - Permanent
  - Resulting failures are called hard failures
- Soft faults
  - Transient or intermittent
  - Account for more than 90% of all failures
  - Resulting failures are called soft failures

Fault Classification



Failures


MTBF
MTTD
MTTR
Time
Fault occurs
Error caused
Detection of error
Repair
Fault occurs
Error caused
Multiple errors can occur during this period

Fault Tolerance Measures

Reliability

$$R(t) = \Pr\{0 \text{ failures in time } [0,t] \mid \text{no failures at } t=0\}$$

If occurrence of failures is Poisson

$$R(t) = \Pr\{0 \text{ failures in time } [0,t]\}$$

Then

$$\Pr(k \text{ failures in time } [0,t] = \frac{e^{-m(t)}[m(t)]^k}{k!}$$

where $m(t)$ is known as the *hazard function* which gives the time-dependent failure rate of the component and is defined as

$$m(t) = \int_0^t z(x)dx$$

Fault-Tolerance Measures

Reliability

The mean number of failures in time [0, t] can be computed as

$$E[k] = \sum_{k=0}^{\infty} k \frac{e^{-m(t)}[m(t)]^k}{k!} = m(t)$$

and the variance can be be computed as

$$Var[k] = E[k^2] - (E[k])^2 = m(t)$$

Thus, reliability of a single component is

$$R(t) = e^{-m(t)}$$

and of a system consisting of n non-redundant components as

$$R_{sys}(t) = \prod_{i=1}^{n} R_i(t)$$

Fault-Tolerance Measures

Availability

$A(t) = \Pr\{\text{system is operational at time } t\}$

Assume

- Poisson failures with rate λ
- Repair time is exponentially distributed with mean $1/\mu$

Then, steady-state availability

$$A = \lim_{t \to \infty} A(t) = \frac{\mu}{\lambda + \mu}$$

Fault-Tolerance Measures

MTBF

Mean time between failures

$$\text{MTBF} = \int_0^\infty R(t)\,dt$$

MTTR

Mean time to repair

Availability

$$\frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}$$