Dependable Computing: Basic Concepts and Definitions

Lecture 2

Learning Objectives

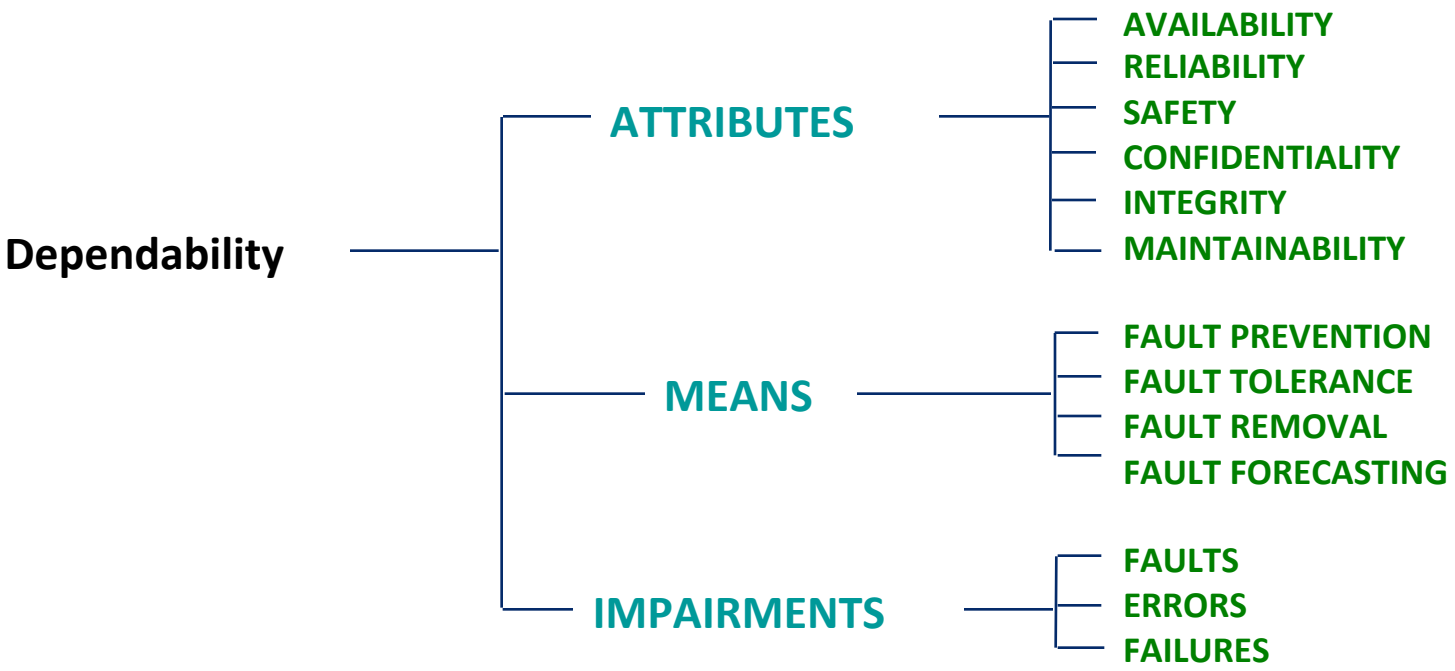
- At the end of this lecture, you should be able to:
  - **Define** fault-tolerance terms such as reliability, availability, safety and distinguish between them
  - **Identify** faults, errors and failures based on system descriptions and scenarios
  - **Classify** faults, errors and failures into various types
  - **Categorize** fault-tolerance techniques based on which phase they are applied
  - **Apply** common fault-tolerance strategies to problem scenarios and systems

What is dependability ?

- IFIP WG 10.4 on dependability
  - *[..] the trustworthiness of a computing system which allows reliance to be **justifiably** placed on the service it delivers*

- Incorporates the following notions:
  - Availability, Reliability, Maintainability (traditional)
  - Safety, security and Integrity (modern)

Dependability: Attributes, Means and Impairments



Dependability Attributes

- **Availability:** Readiness for correct service
- **Reliability**: Continuity for correct service
- **Safety:** Absence of catastrophic consequences
- **Integrity:** Absence of improper modifications
- **Maintainability:** ability to undergo modifications and repairs

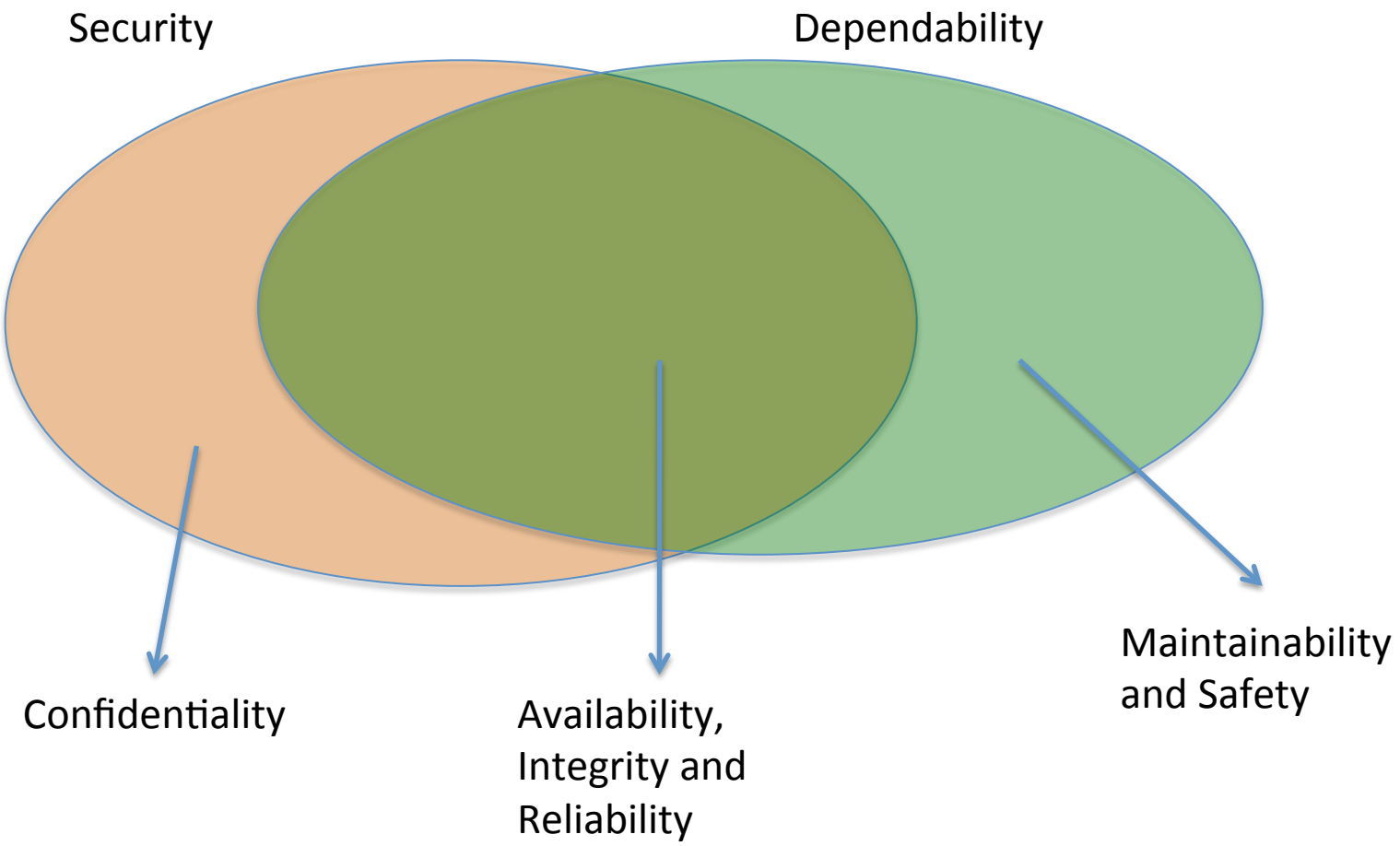
Which of the following do the systems satisfy ?

- A database system fails every 1 minute, but recovers in 0.1 micro-seconds. The recovery is guaranteed to occur after every failure.
- A web server has downtime of 1 month a year, but it goes down at the same month every year and is down for the entire month
- A missile system has an expected downtime of 1 minute a year, and will hit its target with 99.999% certainity. However, occasionally it may backfire and hit an object close to the missile launcher itself.
- A nuclear power plant system will lock down whenever any improper change is made to it. Once locked down, it needs the sysadmin to initiate a complex control sequence to bring it up again. This operation may take anywhere from few hours to a few days.
- A computer system on the stock trading floor has only 3 minutes of downtime a year, and is always up during critical operations. It also prevents any modifications to it (including code updates) unless three different operators coordinate to simultaneously apply them.

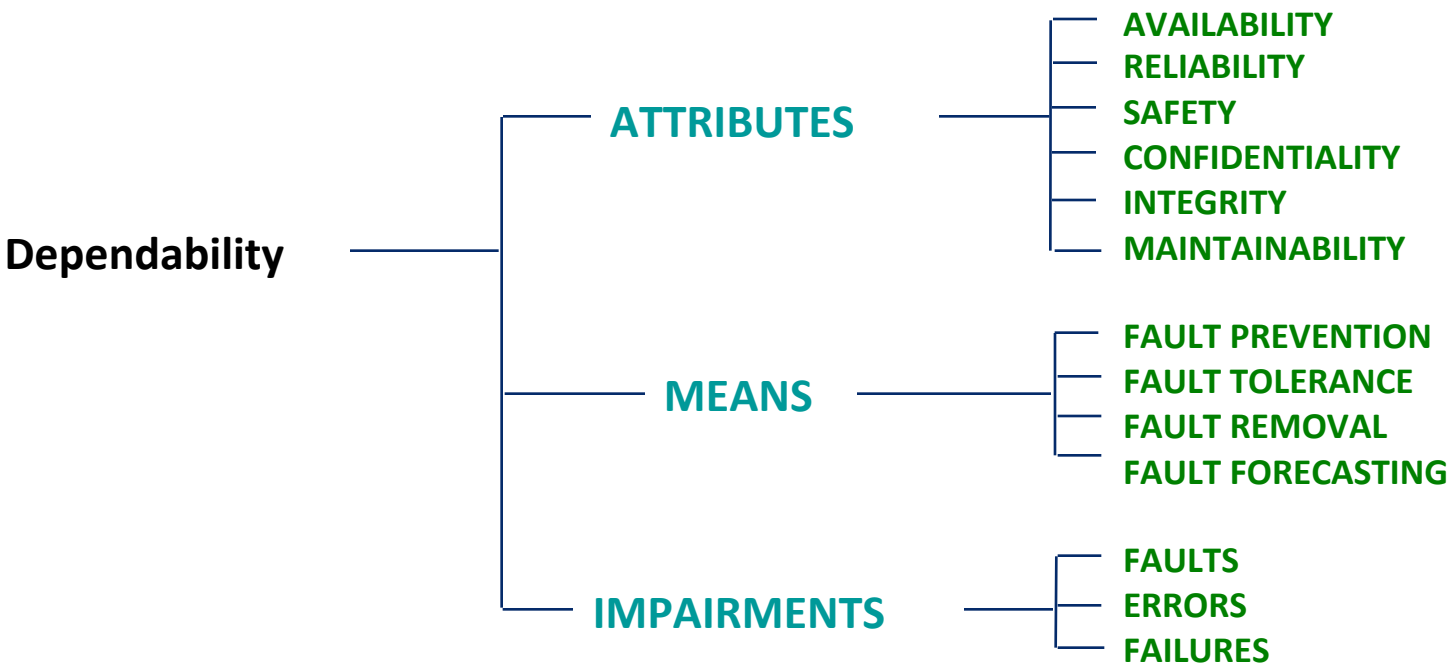
Examples of Attributes

Type of system	Availability	Reliability	Safety	Integrity	Maintainability
Aircraft systems	X	X	X	X	X
Nuclear reactor	X	X	X		X
Internet news page	X	X		X	X
Smart-phone	X		X	X	X
Pace-maker	X	X	X	X	
Automobile system		X	X	X	X

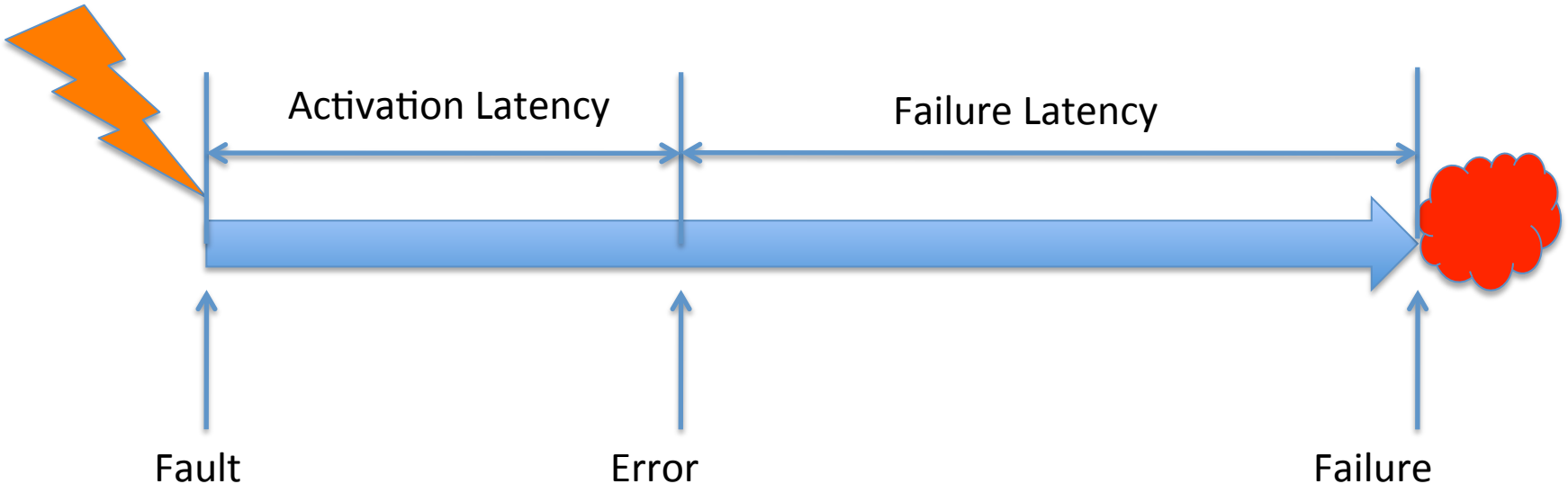
Dependability and Security



Dependability: Attributes, Means and Impairments



Dependability Impairments



Fault – Defect in the system (e.g., soft error)

Error – Deviation of system behavior from fault-free run

Failure – Violation of system's specification (e.g., crash)

Identify fault, error and failure

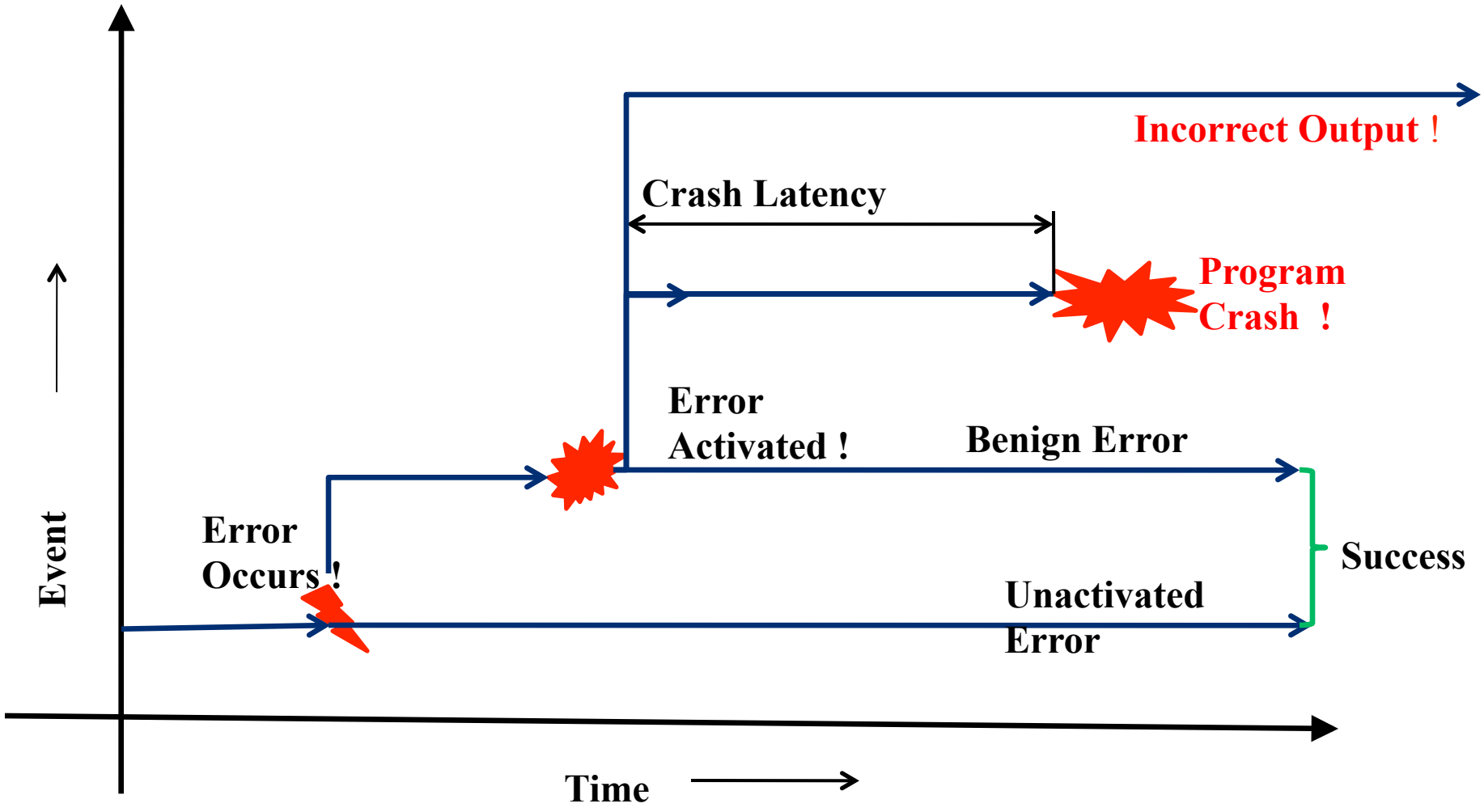
- A soft error in a processor causes corruption of a data word stored in the cache. This word is read and de-referenced in the program, which leads to an “out of bounds” exception (e.g., seg. fault).
- A program has a logical bug that is trigerred only by certain test cases (i.e., when they exercise it). When the bug is triggered, it causes the program to compute a value incorrectly and the wrong value is printed as part of the program’s output.

Fault masking and benign errors

- **Not all faults lead to errors**
  - Faults can be masked because they are not activated (e.g., faults in unread locations)
  - Faults can also be corrected before they lead to errors (e.g., memory scrubbing in ECC)

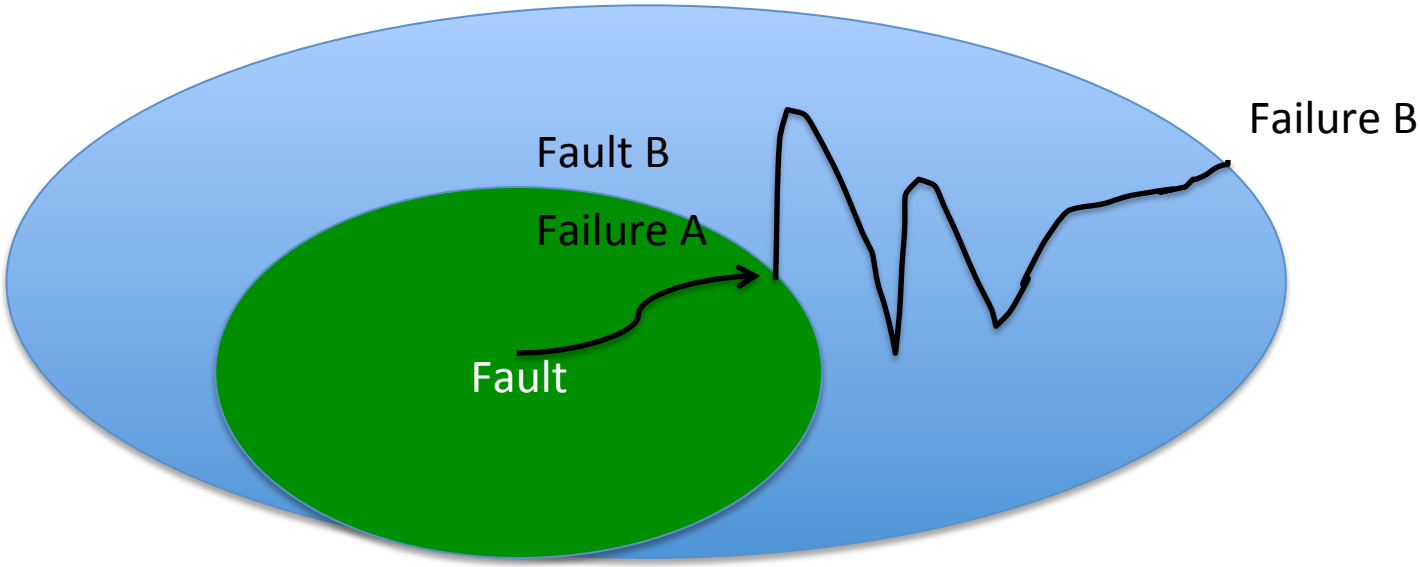
- **Not all errors lead to failures (benign errors)**
  - Error may affect inconsequential system state
  - System may have redundancy to correct error

Anatomy of an Error



Inter-connected Systems

- One system's failure (system A) may be another one's fault/error (system B)



Examples

- **Example 1**
  - A short in an integrated circuit is a failure (with respect to the function of the circuit)
  - The consequence (e.g., stuck at a Boolean value) is a fault that stays dormant until activated
  - Upon activation (invoking the faulty component by applying an input) the fault becomes active and produces an error
  - If and when the propagated error affects the delivered service (e.g., information content), a failure occurs
- **Example 2**
  - The result of an error by a programmer leads to a failure to write the correct instruction or data.
  - This results in a dormant fault in the written software (e.g., faulty instruction)
  - Upon activation the fault become active and produces an error
  - When the error affects the delivered service , a failure occurs

Fault Classification

Temporal

- Transient
  - Occur only once at a location
  - Eg., cosmic ray strikes
- Intermittent
  - Periodically recur at a location
  - E.g., timing violations
- Permanent
  - Continuous and stable occurrence at a location
  - E.g., stuck-at-faults

Origin

- Physical faults:
  - Occur due to physical phenomena such as EM effects, threshold, effects etc. or due to environmental conditions such as temperature or workload
- Human-made faults:
  - Programming errors, mis-configuration, human operator errors etc.

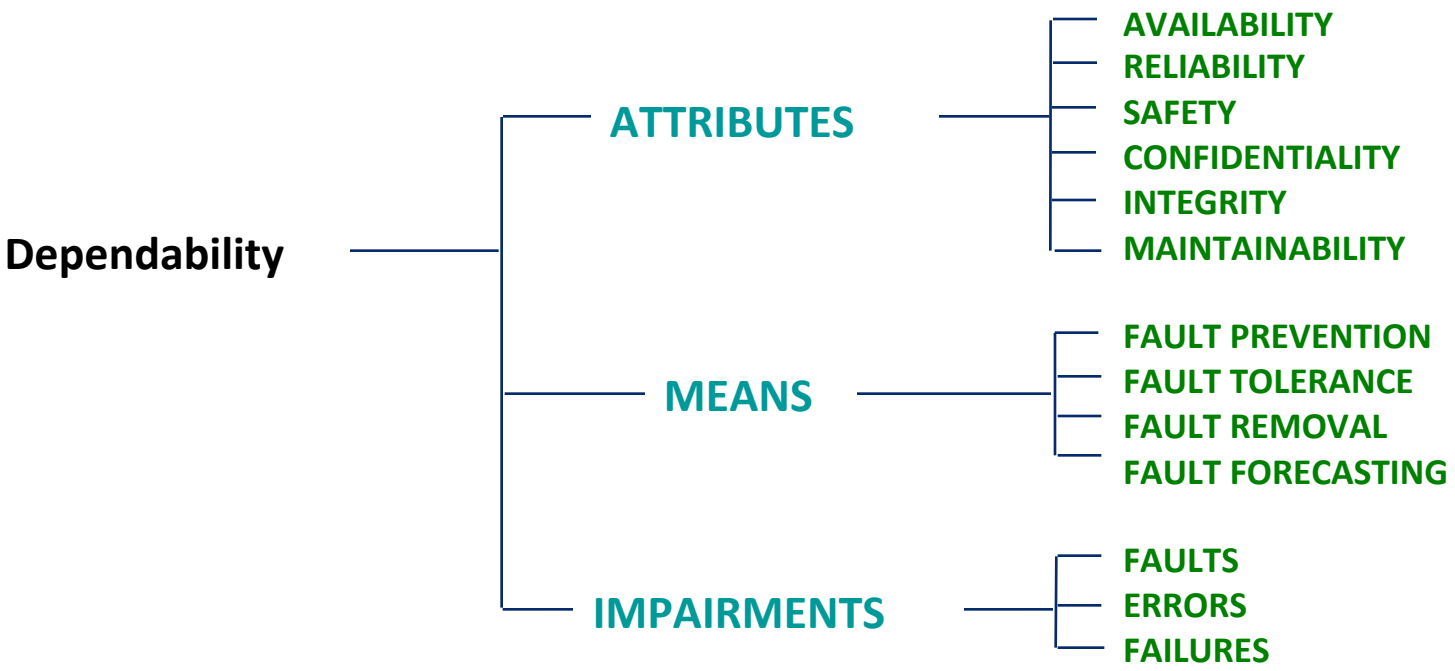
Classification of Failures

Criteria	Classification 1	Classification 2
Nature	Halt (fail-stop)	Erratic (Babbling)
Detection	Signaled	Un-signaled (fail silent)
Consistency	Consistent	Inconsistent
Severity	Minor	Catastrophic

Classification of Errors

Criteria	Classification 1	Classification 2
Domain	Timing	Content
Detection	Detected	Latent
Consistency	Consistent	Inconsistent
Severity	Minor	Catastrophic

Dependability: Attributes, Means and Impairments



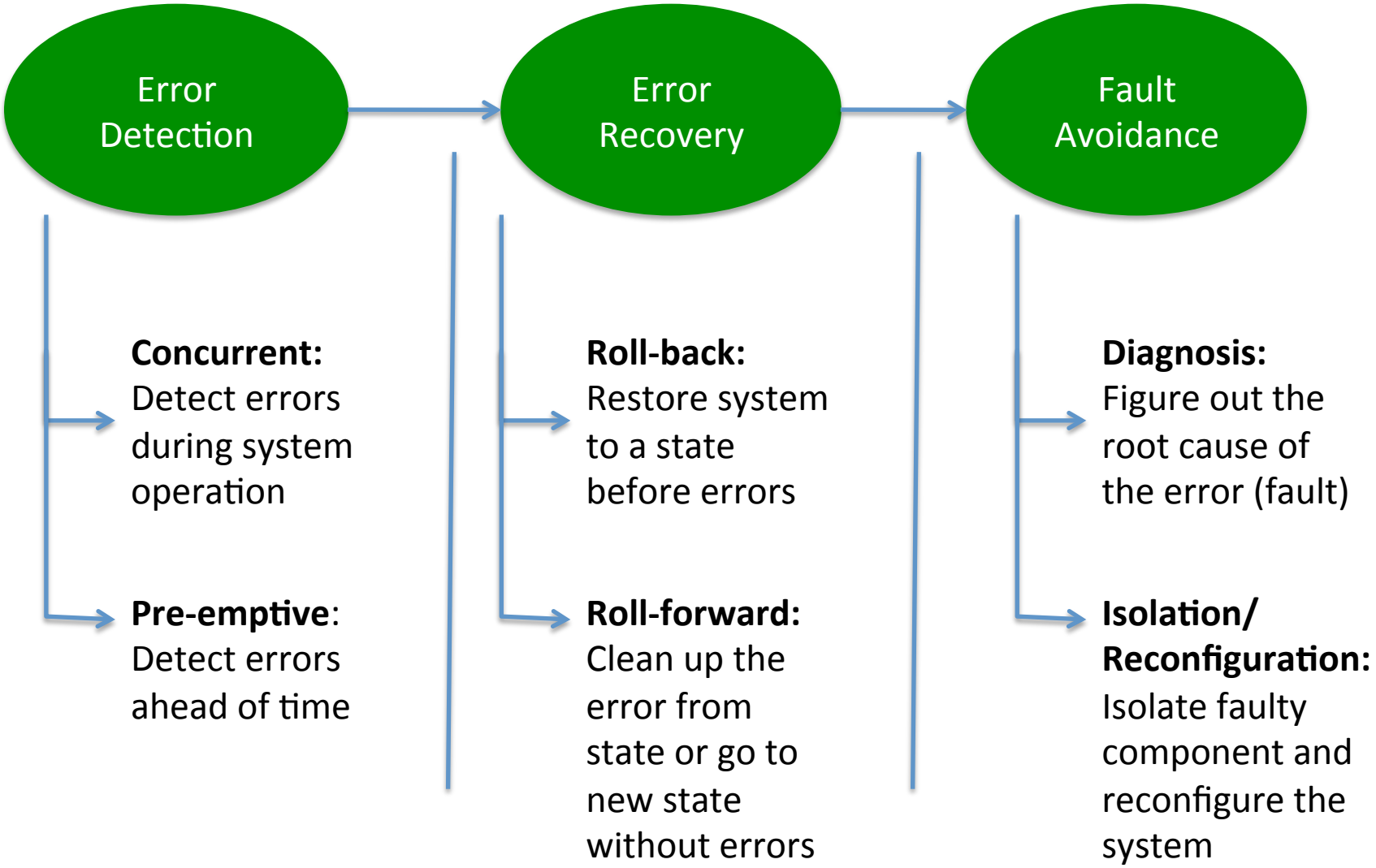
Dependability Means

- **Fault prevention**
  - Prevent occurrence of faults
- **Fault tolerance**
  - Avoid service failures in the presence of faults
- **Fault removal**
  - Reduce the number or occurrence of faults
- **Fault forecasting**
  - Predict how many faults remain and their likely consequences (increase assurance in system)

Examples

- Consider a space mission that has to be deployed for long periods of time (say a few years)
  - **Fault prevention**: Use of reliable coding practices, (e.g., defensive programming) and safe languages
  - **Fault-tolerance**: Provide ability to detect and repair failed components during missions
  - **Fault-removal**: Remove the faults at runtime or record the commonly observed faults for later removal
  - **Fault-forecasting:** Predict how likely a future mission is to succeed based on the lessons learned

Fault-tolerance techniques



Fault Tolerance

- **The ability to provide continued correct operation despite the presence of faults**
  - Encompasses a broad rage of techniques ranging from low-level devices to application software
  - Important to ensure that the service behaves as expected provided fault belongs to fault-model
  - There is no such thing as perfect fault-tolerance Every fault-tolerance technique has a coverage and a fault-model over which it is evaluated.

Error Detection

- **Concurrent detection during normal operation**
  - Watchdog timer
  - Software assertions
  - Process pairs
- **Preemptive detection preempts the failure**
  - Spare checking
  - Memory scrubbing
  - Software rejuvenation

Error Recovery

- **Rollback:** Restores the state of the system to one before the fault. Useful for transient and intermittent faults.

- **Rollforward:** Corrects the fault and allows system to make forward progress, if possible.

- **Compensation:** Leverage natural redundancy of the system to mask the error.

Fault Avoidance

- **Diagnosis**
  - Identify the root cause of the fault (location, type)
- **Isolation**
  - Physically/Logically excludes faulty component
- **Reconfiguration**
  - Switches in non-faulty components and remaps the tasks to the non-faulty components
- **Reinitialization**
  - Record the new system state and updates the external entities that interface with the system (if necessary)

Example: Hardware Fault Tolerance

- Multi-core system experiences a fault in one of the cores. Error is **detected** through a concurrent check. What are the possible recovery actions that can be taken ?

Example: Software Fault Tolerance

- Two identical copies of a software program are run concurrently to check each other. If the output of one differs from the output of the other, what recovery actions can be taken ?

Coverage

- Every fault-tolerance technique has associated coverage, i.e., the probability that it detects/ corrects a fault, given that a fault has occurred
  - Coverage can be reduced by defects in the fault-tolerance mechanism's implementation ("who will check the checker ?")
  - Coverage can reduced because of common mode failures or because assumptions made during the FTM's design are violated in practice

Summary

- **Dependability concepts:** definitions, means, attributes and impairments, systems view

- Recommended reading:

Avizienis, A., Laprie, J., Randell, B., and Landwehr, C.

Basic Concepts and Taxonomy of Dependable and Secure Computing. *IEEE Trans. Dependable and Secure Computing,* Vol 1, Issue 1, 2004, 11-33.