Fault Tolerance

Techniques for creating reliable systems out of unreliable components.

- Error Detection
- Error Masking / Correction
- Redundancy

Fault: Defect with the potential to cause a failure.

- active / latent
- persistent / intermittent / transient

Error: Incorrect behavior which could lead to failure if it is not masked.

- undetectable / detectable / maskable
- untolerated / reported / tolerated

Failure: When a component or module does not meet its specifications

An active fault causes an error, if the error is not corrected it could lead to a failure.

Fail-fast: Component reports the error

Fail-safe: Bad values are transformed to “safe” values (i.e. blinking red stop-light)

Fail-soft: Operate correctly but with decreased performance or reduced features

Error-masking: Component meets the specification despite the error.

Hamming distance: The number of bits that are different between two valid representations

For a code with a hamming distance d:

(d-1) bit errors are detectable.

(d-1)/2 bit errors are correctable.

d=1: all errors are undetectable

d=2: single-bit errors are detectable

d=3: single-bit errors are correctable

d=4: double-bit errors are detectable and single-bit errors are correctable.

Forward Error Correction:

Always send extra information to correct errors

useful when it is it hard to retransmit

(i.e. broadcast, one-way, real-time)

Backward Error Correction:

Detect errors and retransmit lost data

$$\text{Availability} = \frac{\text{time the system was up}}{\text{time it should have been}}$$

$$= \frac{\text{MTTF}}{\text{MTTF + MTTR}}$$

MTTF: mean time to failure

MTTR: mean time to recovery

MTBF: mean time between failures = MTTF +MTTR

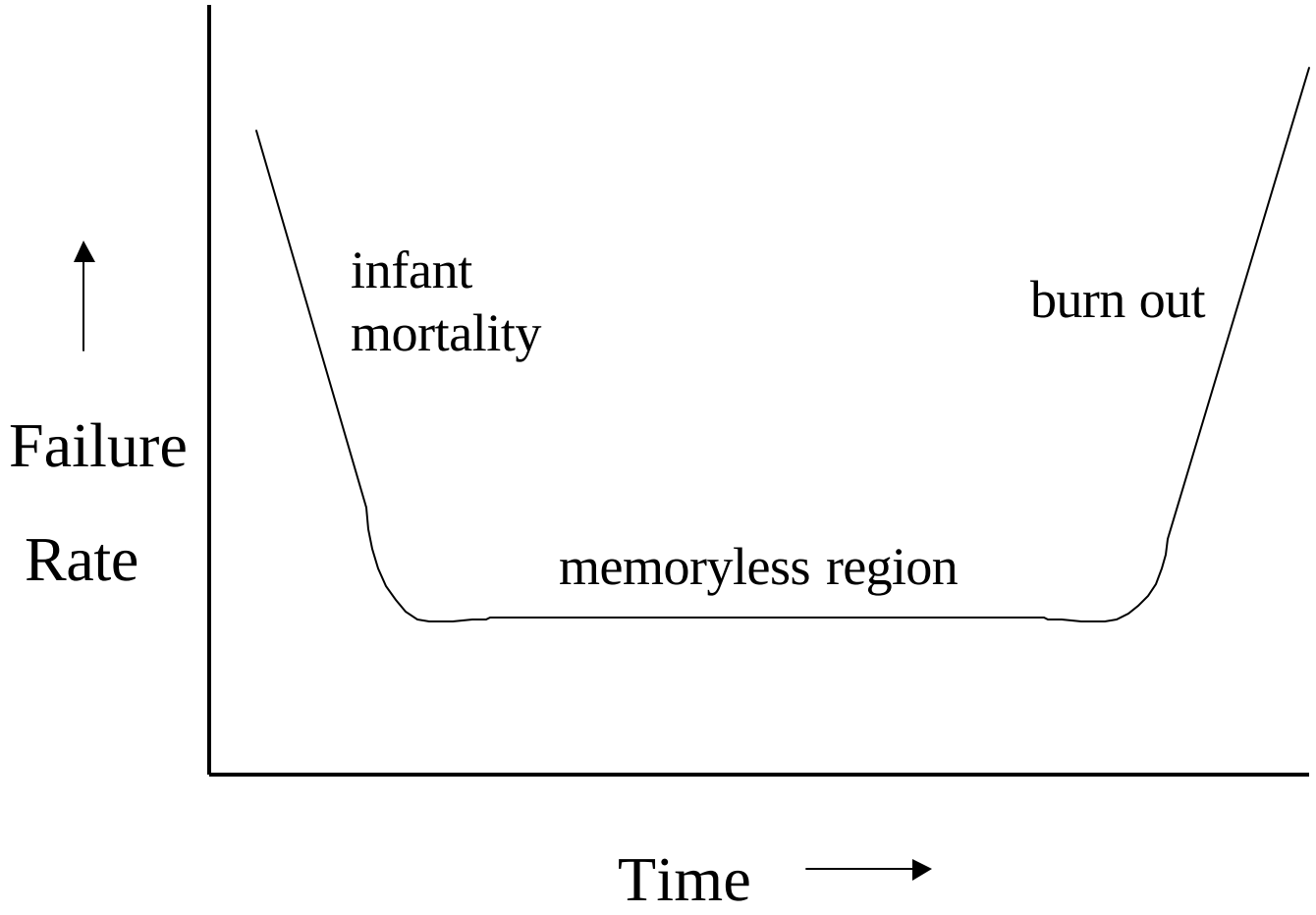
Often expressed in nines: three-nines = 99.9%

Common Assumptions:

memoryless - failure rate in independent of time

independent – failures are unrelated

“Bathtub curve”



Redundancy:

N-Modular Redundancy (NMR):

Provide identical inputs to N systems and connect the outputs to a voting system. More than N/2 must agree for the system to function properly

Fail-Fast NMR:

Build system out of fail-fast components and exclude components reporting errors from voting.

MTTF of the system decreases!

$$MTTF_{system} = \frac{MTTF}{N} + \frac{MTTF}{N-1} + \ldots + \frac{MTTF}{1} = MTTF * \ln(N)$$

NMR with Repair: Dramatically increases availability

Other lessons:

- Consider the application – identify each potential failure and the associated risk
- Avoid rarely used components
- Monitor error rates
- Design for iteration and use feedback