Analysis of Field Data for Repairable Systems

David Trindade, Ph.D.
Distinguished Engineer
Sun Microsystems, Inc.

Photo by Dave Trindade

7x24 Exchange 2006 Fall Conference *Scottsdale, Arizona*

Objectives

To highlight some key considerations for the analysis of reliability data from **repairable systems**

To clarify some aspects and **limitations** of MTBF as a reliability measure for repairable systems

To show some simple but powerful **techniques** for the analysis and modeling of repairable system data

To illustrate - using actual field examples - how such techniques offer valuable **insight** into important reliability issues.

Definition of a Repairable System

A system is ***repairable*** if, following a failure at time ***t***, the system can be **restored to satisfactory operation** by any action.

Examples include servers, computers, automobiles, airplanes, locomotives, major appliances, utilities, air-conditioners, and networks. Besides ***failure times***, other measures of interest may involve cost, downtime, resources used, etc.

Reliability of Repairable Systems

Function of many factors:

Hardware

Basic system design

Operating conditions

Environment

Type and quality of repairs

Materials used

Supplier component quality

Software

Software compatibility

Software robustness

Applications and load

Patches and upgrades

Maintenance practices

Human behavior

A Key Factor: System Age

System age is the elapsed time starting at installation turn-on. Age measures the ***total running hours*** from time zero. Also called power-on hours (POH) or operating hours.

We carefully distinguish *age* from *times between failures*, which are called the *interarrival times*.

Important Property of Repairable Systems

Failures occur sequentially in time.

The *sequence order of interarrival times* provides information that is very important for correct analysis.

Sequence of Times between Failures Provides Valuable Information!

If the times between successive failures are getting **longer**, then the system reliability is **improving**.

Conversely, if the times between failures are becoming **shorter**, the reliability of the system is **degrading**.

If the times show no trend (relatively **stable**), the system is neither improving or degrading, a characteristic of what is called a *renewal process*.

Renewal Process

In a renewal process, the times between failures are observations from a single failure distribution, that is, the times are independent and identically distributed (*i.i.d.*).

In a renewal process, there is *no trend*, since we're dealing with a single population of interarrival times with a *fixed mean,* called the *MTBF*. After repair, the system is *as good as new*.

Example of a Renewal Process: Simple Replacement

Single component system: light bulb.

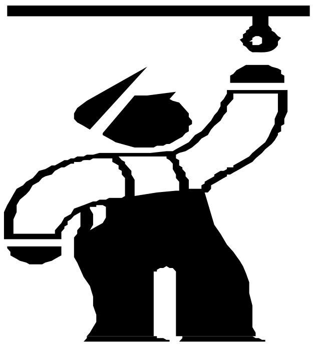
Light bulb is replaced upon failure with a light bulb from the same population as the one replaced.

Stock of spare parts all basically identical.

Stable environment and use.

Single distribution of failure times

- Independent
- Identically distributed



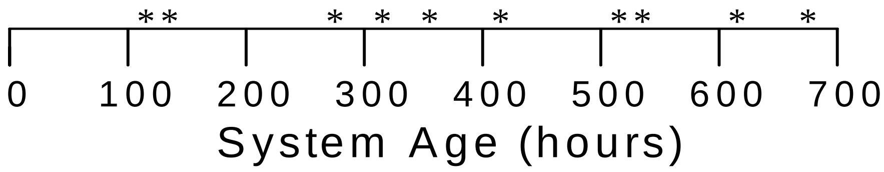
Analysis of Renewal Process

Consider a **single** system for which the **times to make repairs are negligible** compared to the failure times.

Ten failures are reported at the system ages (in hours):

106, 132, 289, 309, 352, 407, 523, 544, 611, 660.

The occurrence of repairs is

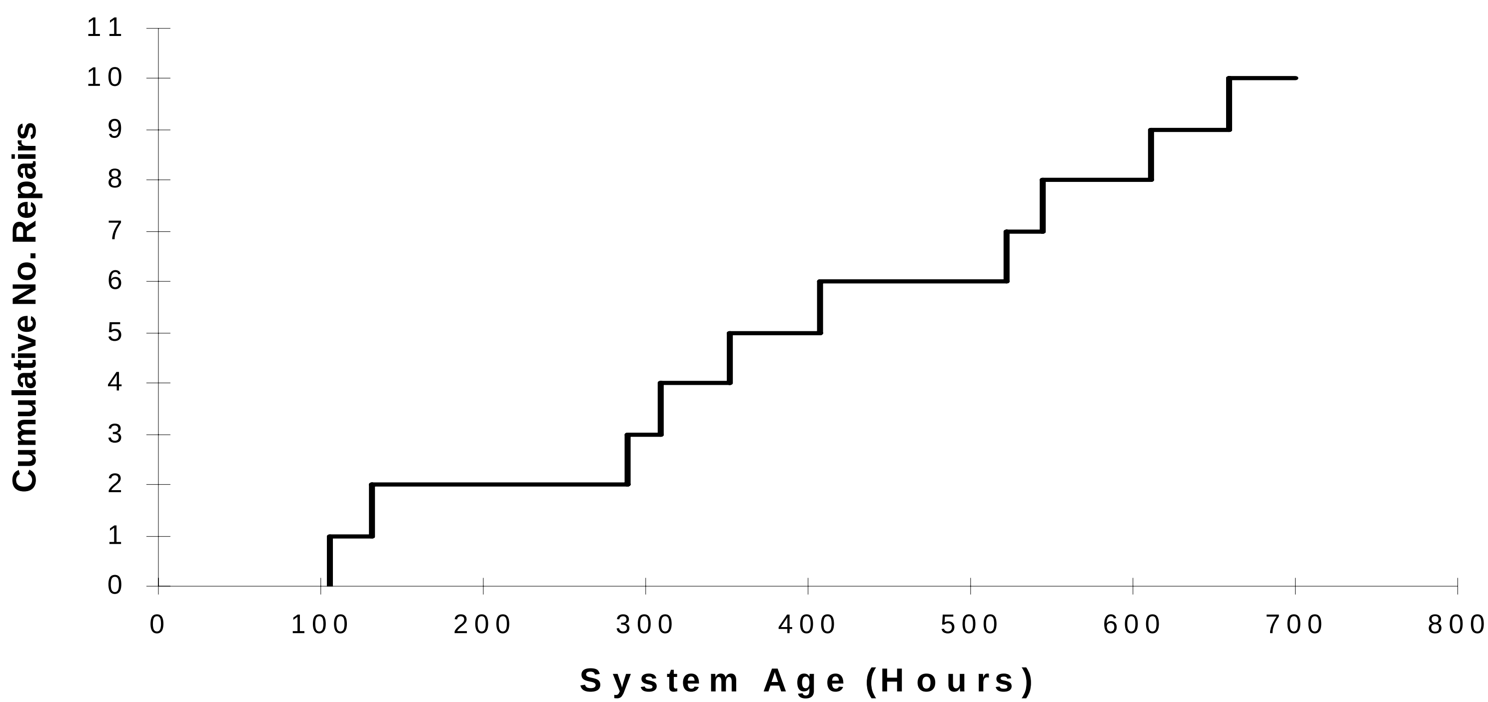


Do you see any pattern?

Cumulative Plot for Single System

A very revealing and useful graph is the **cumulative plot**: the **cumulative number** of repairs, *N(t)*, is plotted against the system **age**, ***t***, at repair.

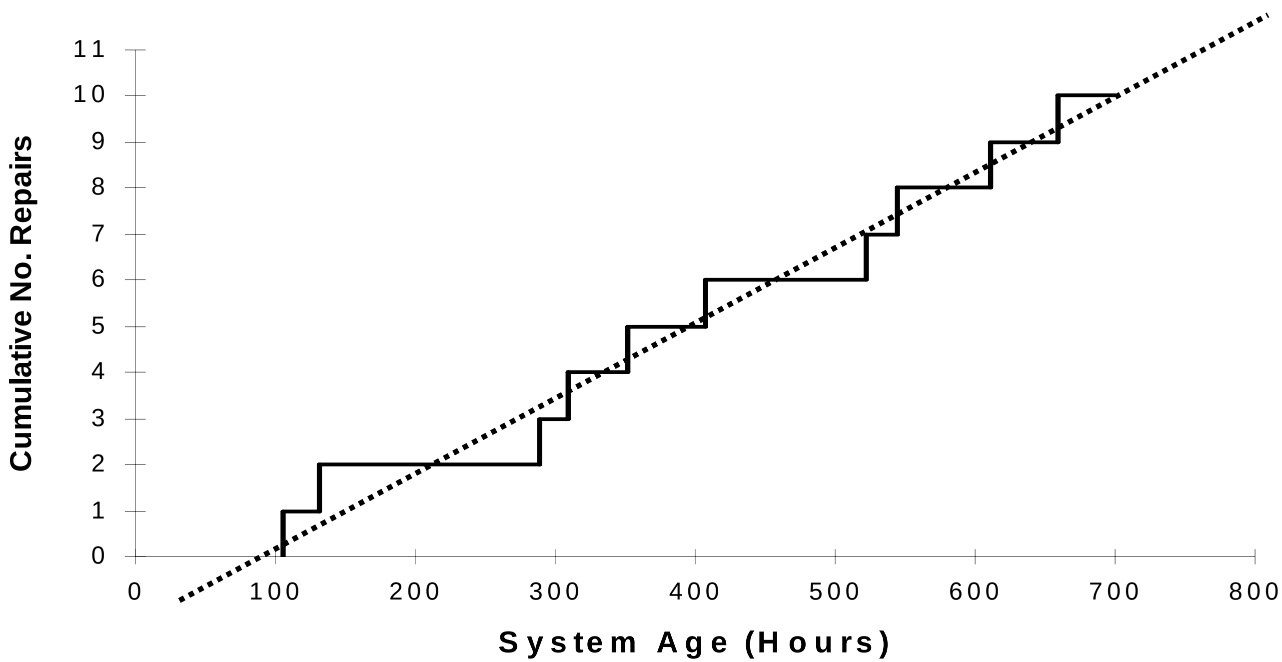
For the renewal data, the cumulative plot is:



N(t) Follows a Straight Line

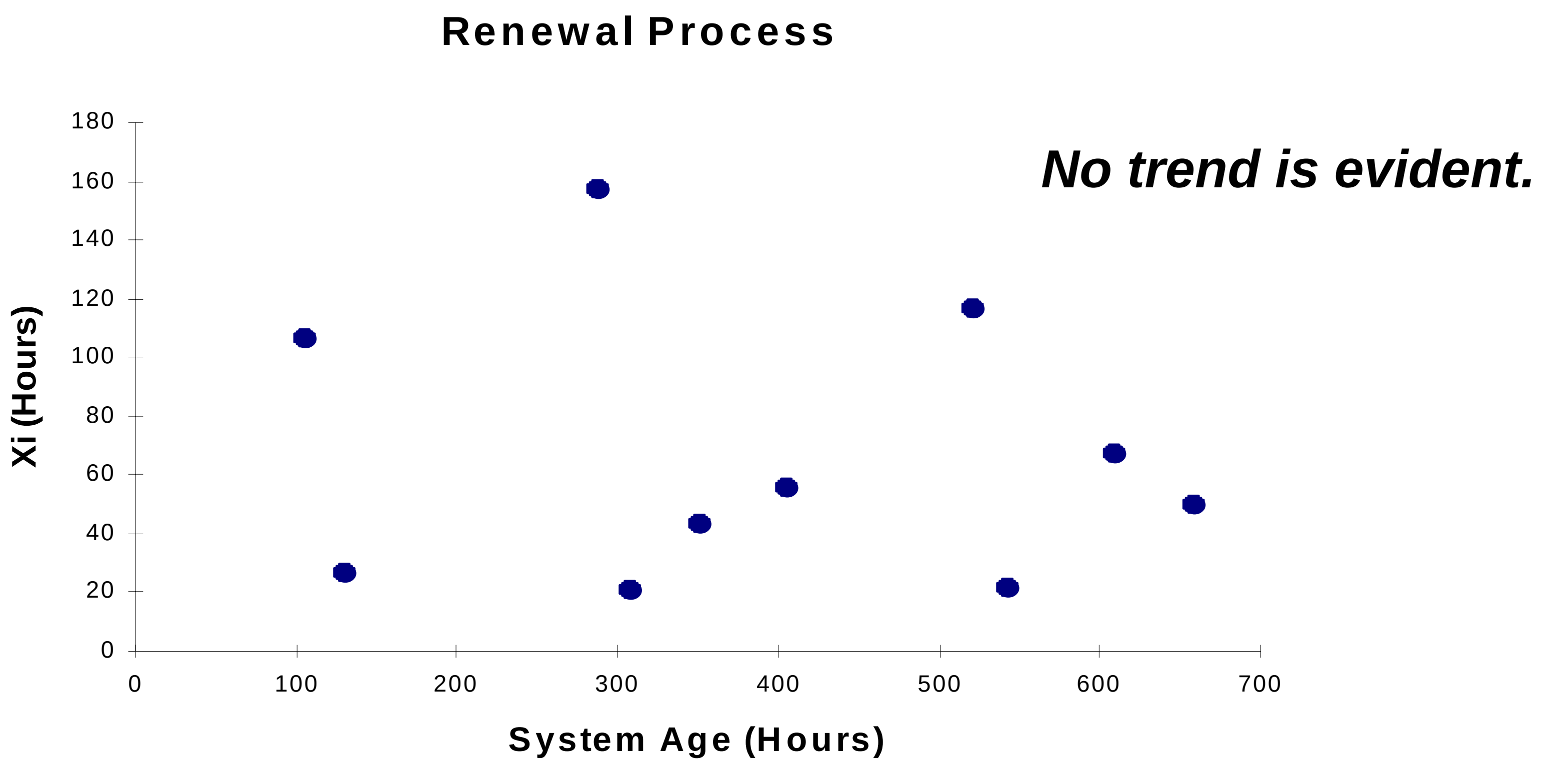
For a renewal process, the times between failures are i.i.d., that is, from a single population having a ***constant*** mean time between failures (MTBF).

Consequently, the cumulative plot of *N(t)* versus *t* should appear to follow a **straight line.**



Times Between Failures Versus System Age

A plot of the sequential ***interarrival times*** versus the system age is also a very useful chart for revealing any patterns.



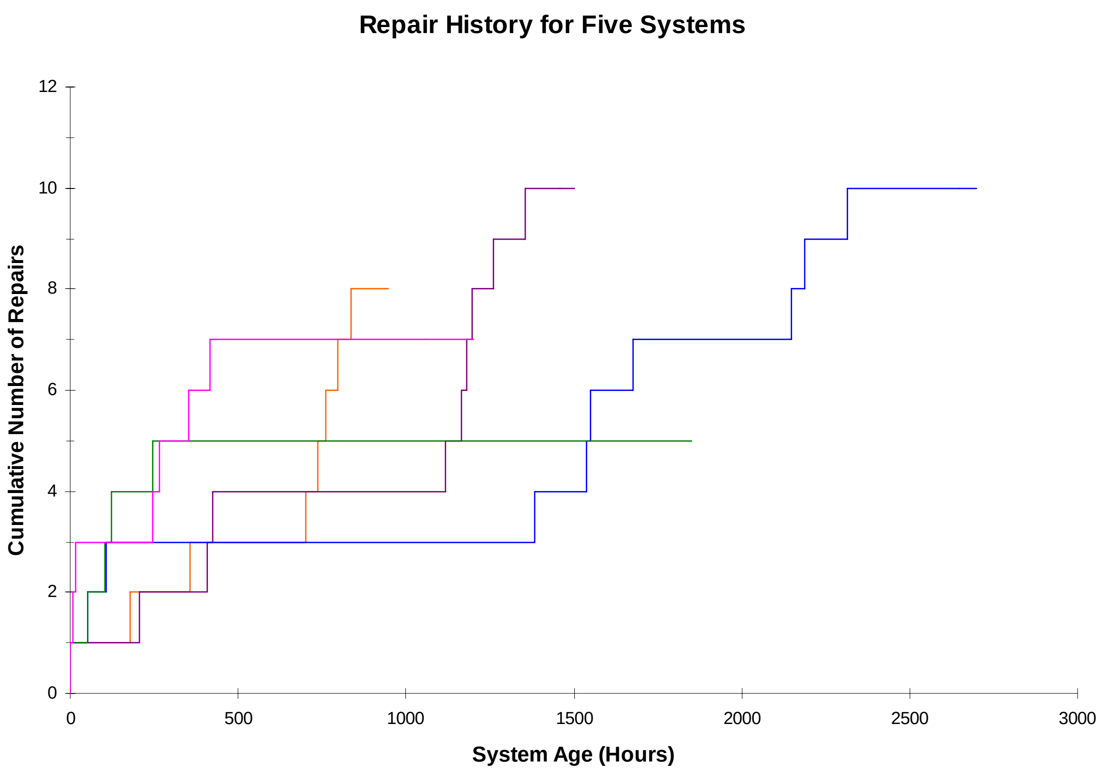
Analysis of a Group of Repairable Systems

Often we want to analyze the reliability behavior of **many identical or similar systems.** Because the systems are most likely installed on different dates, the system ages will vary, resulting in what we call *multicensored data.*

Example: servers installed in a datacenter at different dates throughout the year will have various field ages.

Graphical Approach to Multi-System Analysis

Consider a group of five systems, installed on different dates. Individual repair histories *N(t)* are shown as steps at each repair. All starting times are referenced to zero.

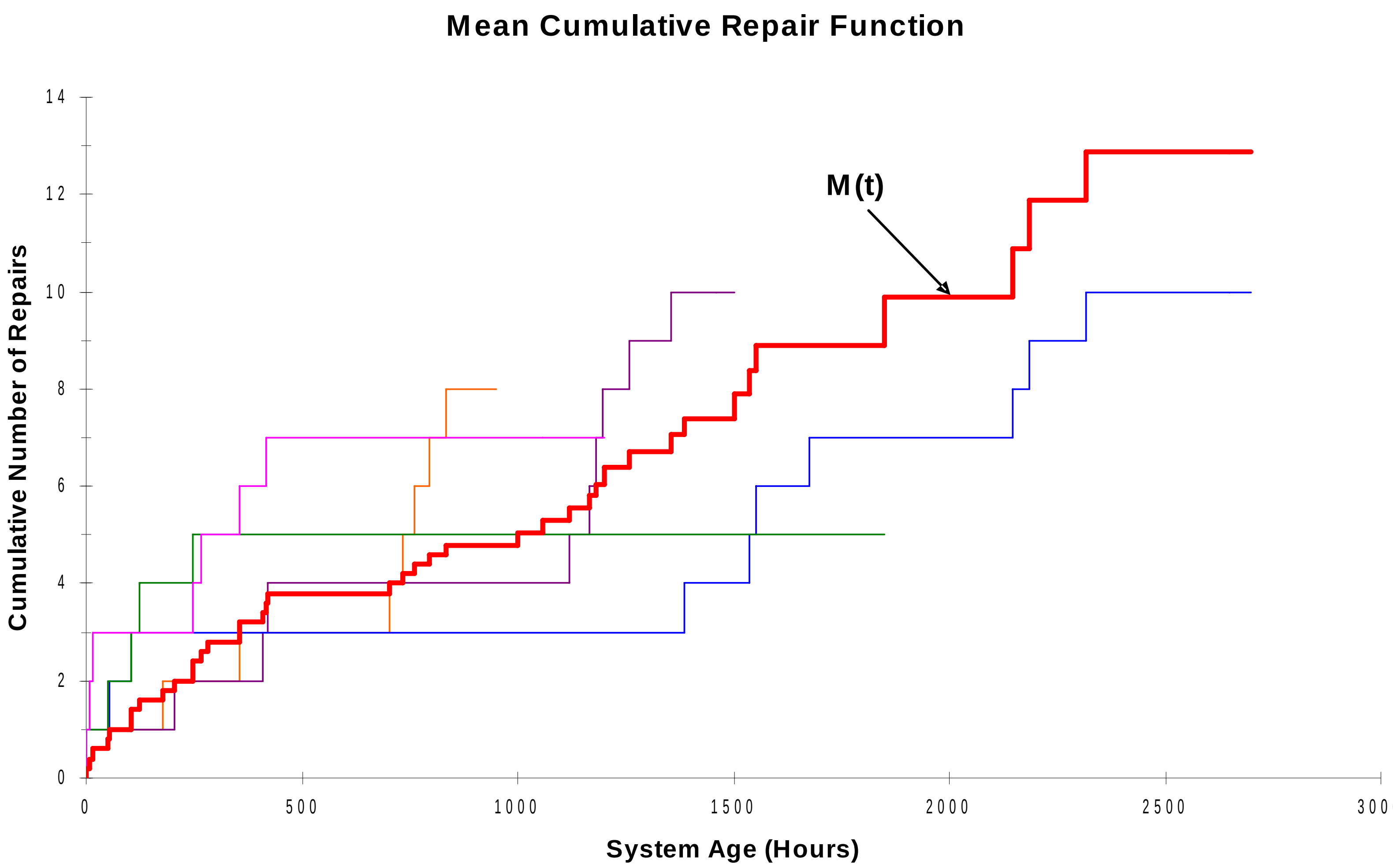


Mean Cumulative Function: MCF

We envision a **single** curve denoted by *M(t)* that gives the **average or mean number of repairs per system** at time *t*. Consider a vertical slice across the individual histories. Such a curve is called the *mean cumulative function* or *MCF*.

Estimation of *M(t)* must account for the **number of systems operational (at risk) at any system age.** Simple statistical procedures can easily handle multicensored data.

MCF for Five Systems



Considerations in Analysis

How **precise** is the estimate of *M(t)*? Confidence intervals can be applied.

The individual *N(t)* plots give us some idea about the **distribution** of the **number of repairs** across the systems at time *t*, but what *model* describes the *number of systems* with zero, one, two…, failures at time *t*?

Can we identify *anomalous* systems?

Graphical analysis is important, but we need to supplement with analytical and modeling tools.

Renewal Process for Single System

One useful model is to assume the interarrival times X_i come from a single *exponential* distribution with constant failure rate λ.

The *MTBF* equals the reciprocal of the failure rate = $1/\lambda$.

The actual *number of failures N(t) in time t* has a *Poisson distribution.* The expected number is λt.

Called a *homogeneous Poisson process (HPP).*

MTBF Estimation for a Single System

$$MTBF = \frac{\text{System Operating Hours During Time Period}}{\text{Number of Failures in Time Period}}$$

Recurrence Rate (ROCOF): $RR = 1/MTBF$

E.g., a system operating for 3000 hours with three failures to date has an *MTBF* of 1000 hours or an annualized *RR* of 8.8 failures/year.

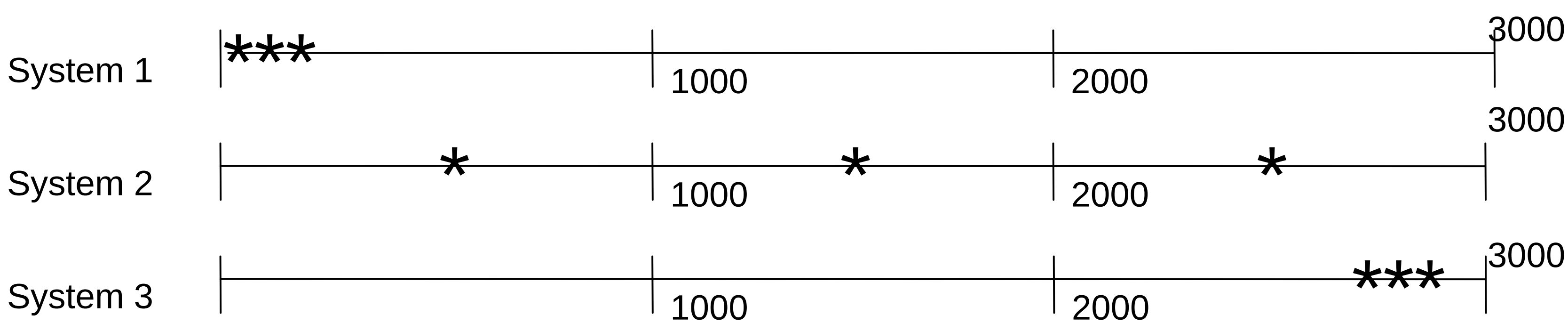
Caution: *RR* and *MTBF* are Summary Statistics – Hide Information

Consider three systems operating for 3000 hours, each with *MTBF* of 1000 hours

System 1 had three failures at 30, 70, 120 hours and no further failures

System 2 had three failures at 720, 1580, and 2550 hours

System 3 had three failures at 2780, 2850, and 2920 hours



Same *MTBF* but are these systems behaving the same?

MTBF for a Group of Systems Over Some Time Period

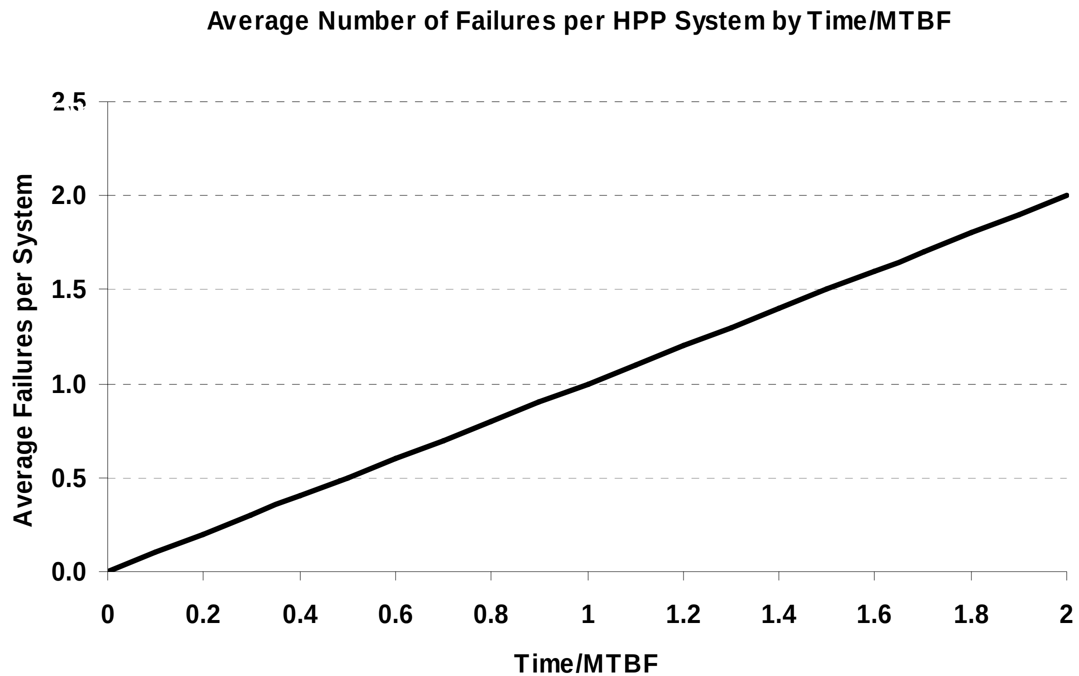
Simple to calculate for any period of time, say monthly:

$$MTBF = \frac{\text{Total Operating Hours on Systems}}{\text{Total Number of Failures}}$$

Problem: Formula assumes **all systems, system hours,** and **all failures are equivalent.** There is no distinction among systems, early life, constant random, and wearout modes of failures during the time period of interest. As a result, such an MTBF may be meaningless.

MTBF (*HPP* Repairable Systems)

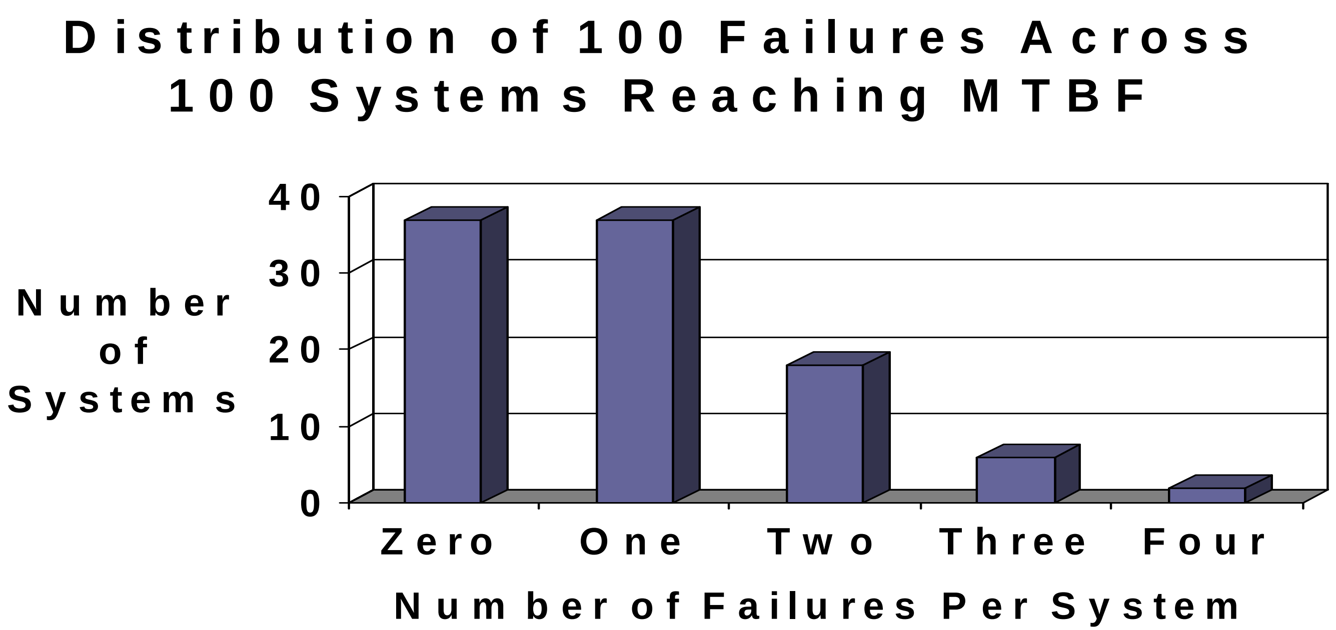
Consider a collection of *similar HPP* systems all with the *same MTBF*. How does the average number of failures per system grow with time?



When all systems reach their MTBF, there will be an average of *one failure per system*, but how are the failures actually distributed across the systems ?

HPP Model Example

Consider 100 systems each with MTBF = 1,000 hours. On the average there will be a total of 100 failures when all reach *t* = 1,000 hours. However, there will *NOT* be exactly one failure on every system.



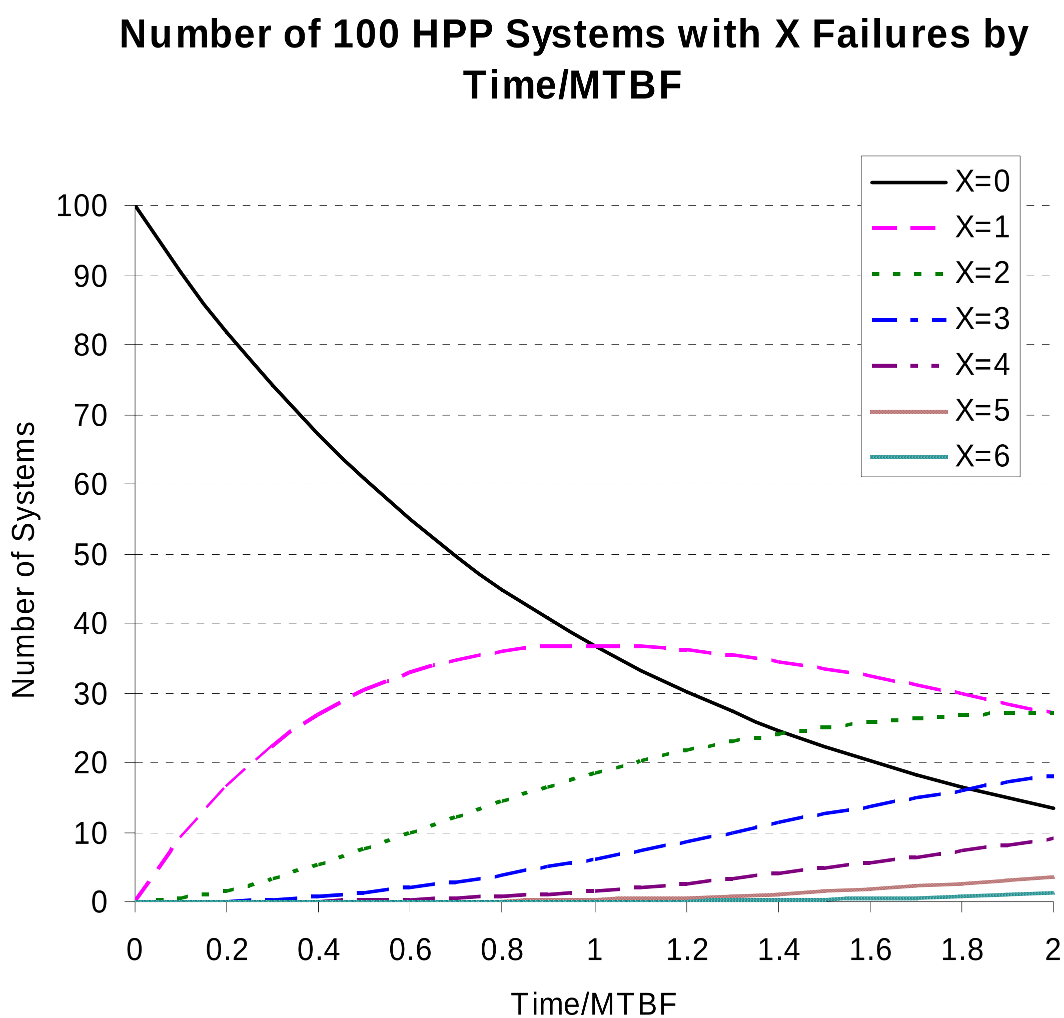
37 systems will have *no* failures

63 systems will have at least one failure: 37(1), 18(2), 6(3), 2(4).

Only those 37 systems with one failure will match the expected *MTBF*.

Customers with multiple failures will see lower *MTBF*s.

HPP Model Implications



We have a basis here for identifying anomalous systems.

Example of a Non-Renewal Process

Consider a light bulb which is replaced upon failure but the cooling fan inside the equipment is degrading, causing a gradually rising temperature.

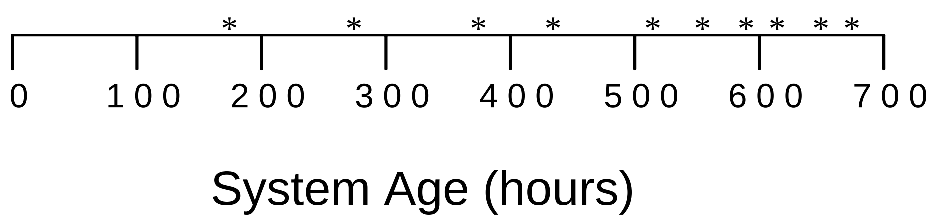
Times of replacement bulb failures are getting shorter.

There is not a single distribution of independent failure times (no constant *MTBF*).

To analyze system behavior properly, we must look at the occurrence order of failures .

Graphical Analysis of Non-Renewal Processes

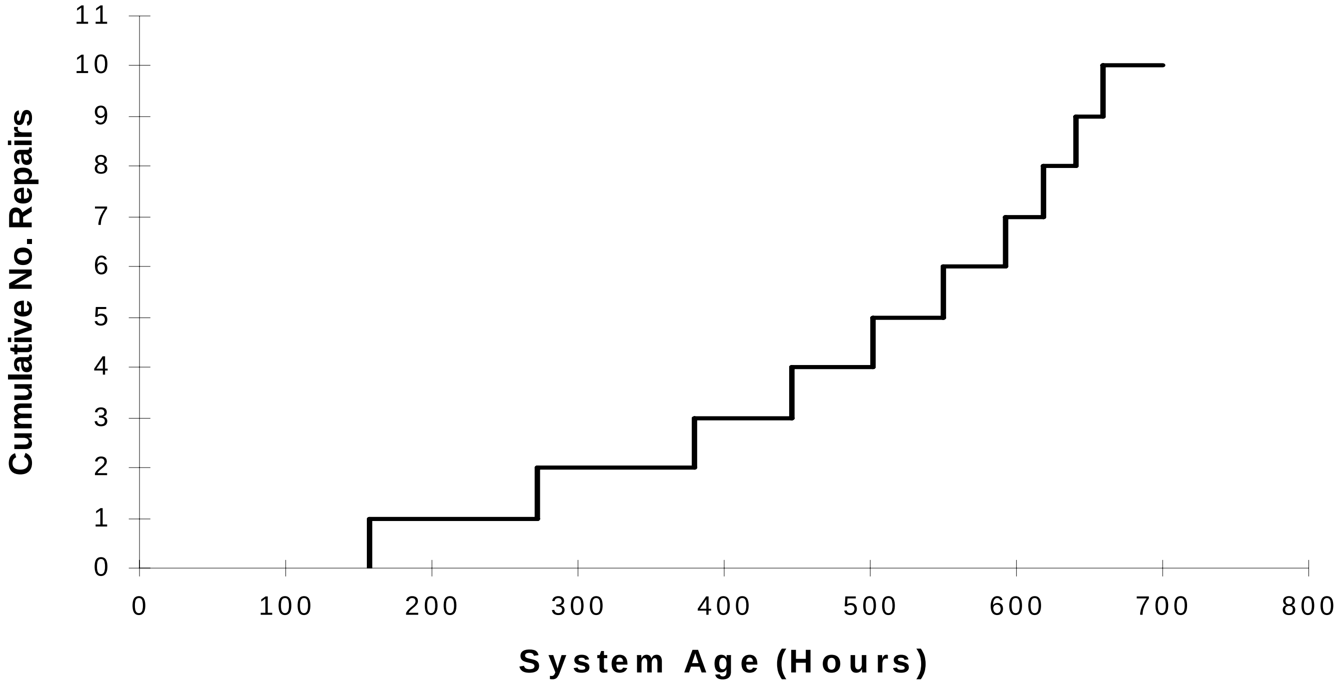
Suppose the repairs occurred at the following times
157, 273, 379, 446, 501, 550, 593, 619, 640, 660.
The line sketch is



Do you see a pattern?

Cumulative Plot

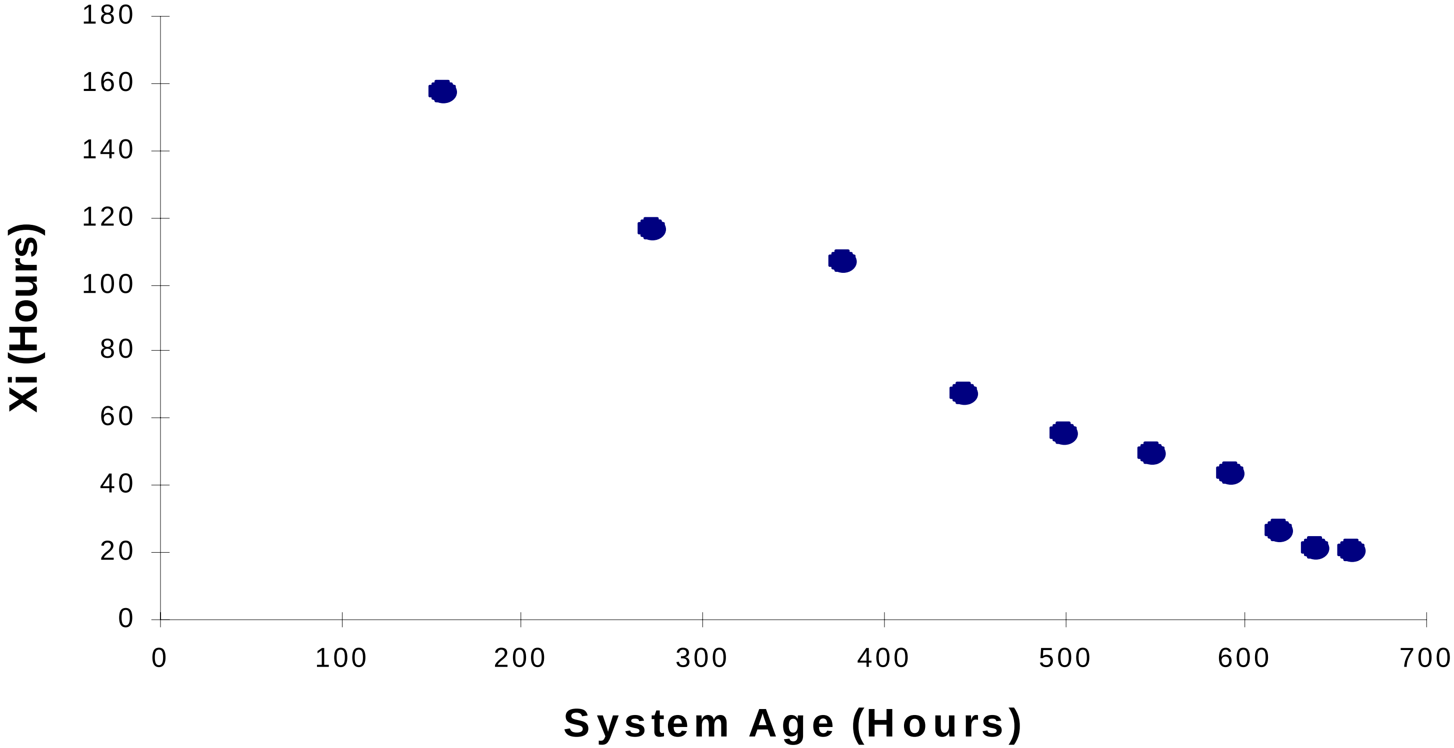
The cumulative plot is shown below.



The curvature shows the frequency of repairs increasing in time, indicating system **degradation.**

Interarrival Times Versus System Age, Degrading

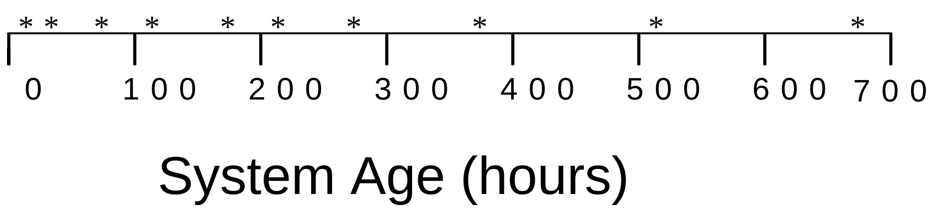
Larger is better.



Another Repairable System History

Suppose the observed consecutive repairs times were
20, 41, 67, 110, 159, 214, 281, 397, 503, 660.

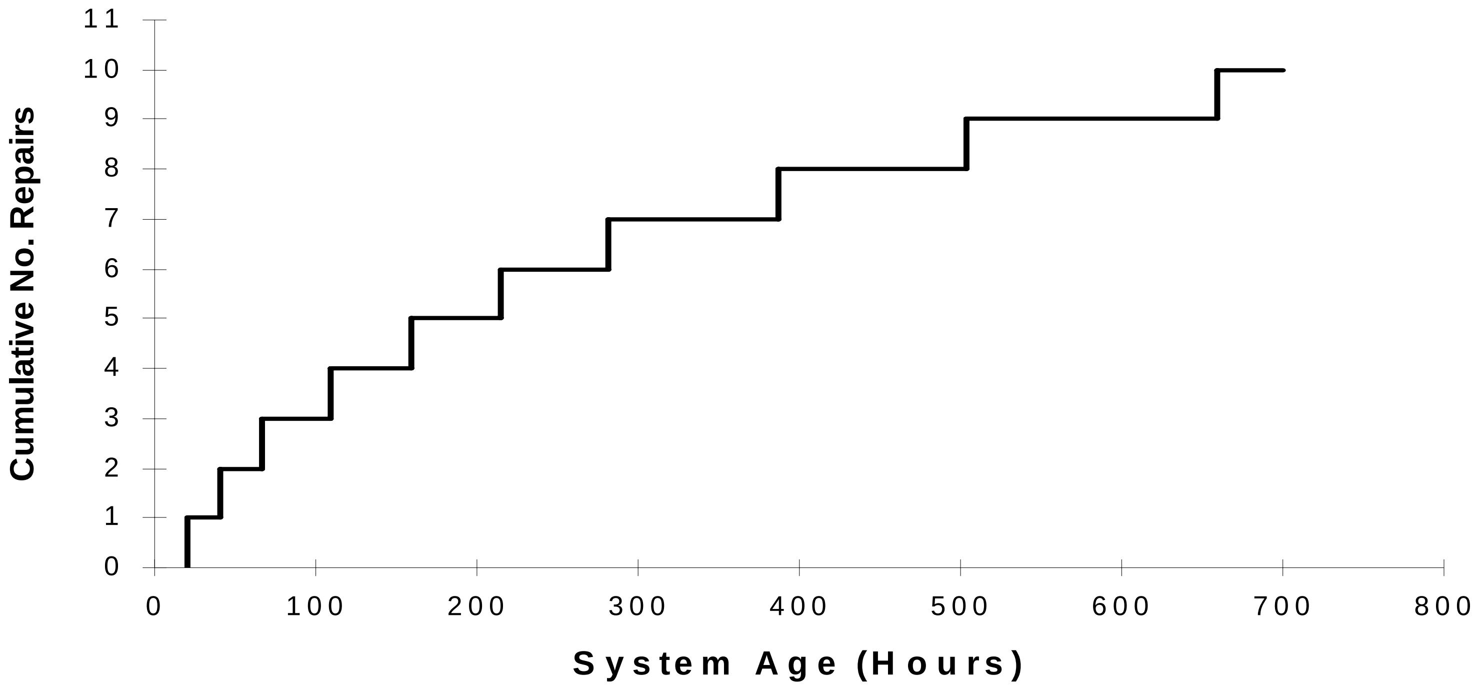
A line sketch of the pattern of repairs shows:



Do you see a pattern?

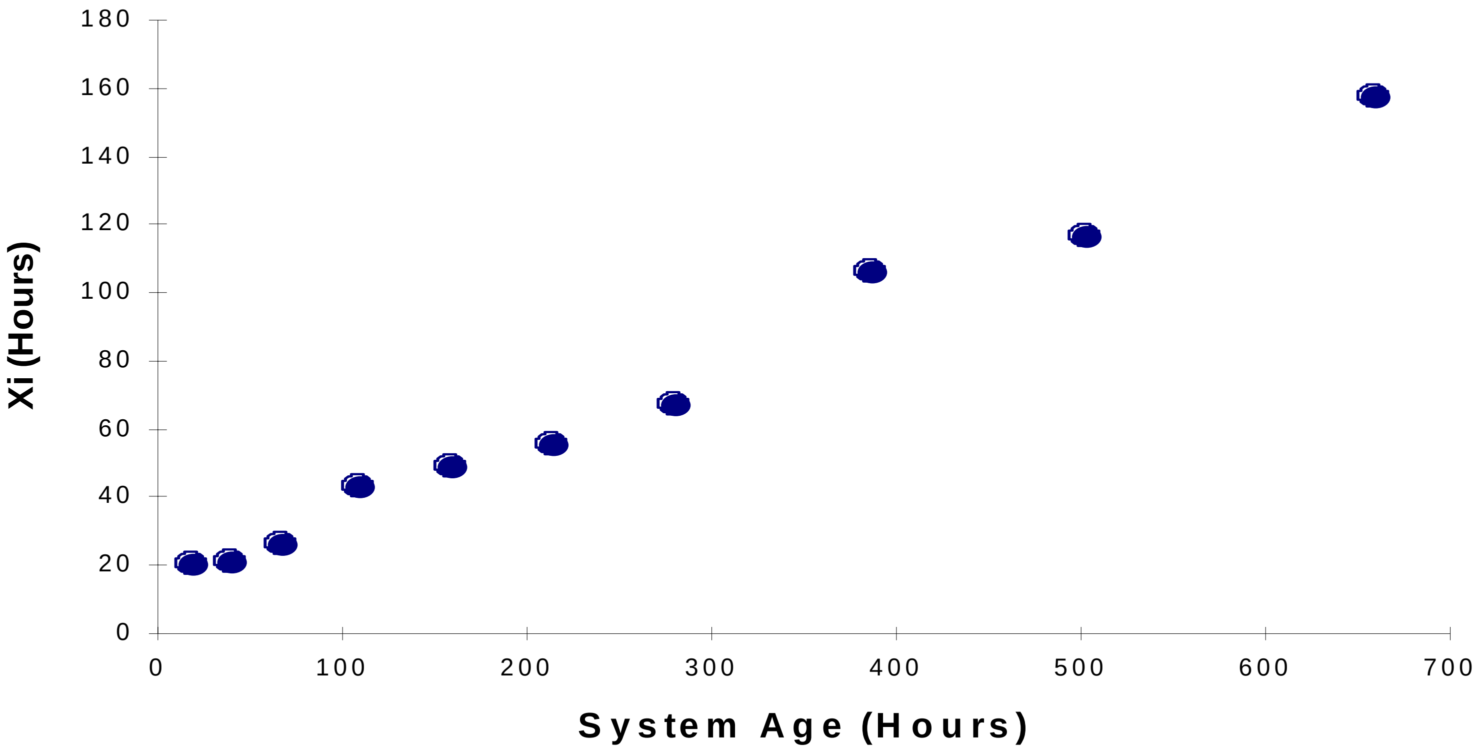
Cumulative Plot

The cumulative plot for this set of data is shown below.



The curvature suggests a decreasing frequency of repairs, that is, an improving recurrence rate.

Interarrival Times Versus System Age, Improving



Larger is better.

MTBF Comparisons (Interarrival Times)

Stable Renewal Process

106, 26, 157, 20, 43, 55, 116, 21, 67, 49

MTBF = 66

Degrading Process

157, 116, 106, 67, 55, 49, 43, 26, 21, 20

MTBF = 66

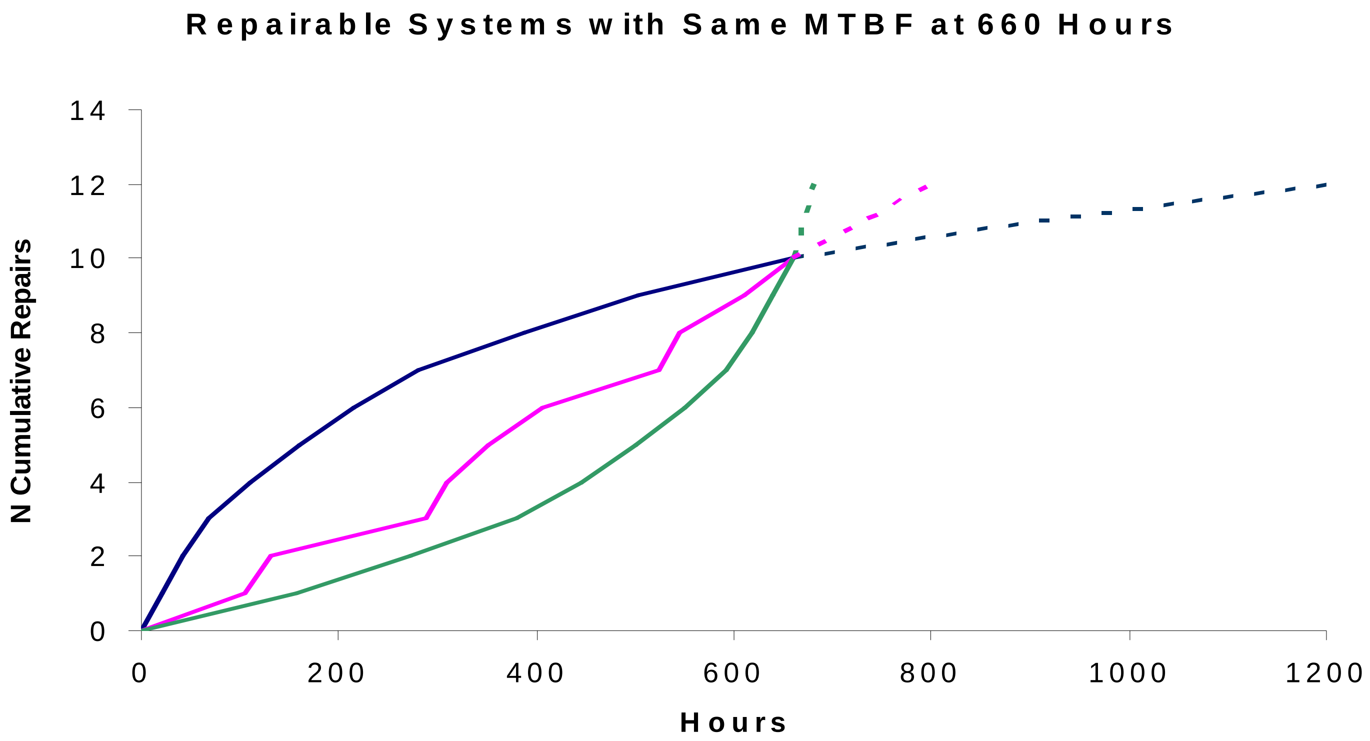
Improving Process

20, 21, 26, 43, 49, 55, 67, 106, 116, 157

MTBF = 66

The data are the same! Only the *order* has changed. *MTBF*s and *RR*s are identical! Yet, the behavior is vastly different!

MTBF Comparisons



MTBF may tell us on the average where we are at some time, but *MTBF* may not reveal how we got there or where we're headed.

MTBF - Misinterpretations

Expected or Typical Lifetime of a System

During the years 1996-1998, the average annual death rate in the US for children ages 5-14 was 20.8 per 100,000 resident population.

The average failure rate is thus 0.02%/yr

The MTBF is 4,800 years!

MTBF in Qualification Activity

Manufacturers often provide MTBF estimates obtained by stressing *many units* for *short periods* of qualification times.

Bending paper clips example.

What's the Point?

Using a summary statistic like the *MTBF* can be misleading and potentially ***risky*** if we do not distinguish between stable or trending processes.

We need to analyze the **ordered** times between failures versus the system ***age*** to determine system reliability behavior.

MTBF can be an Inadequate Measure of System Reliability

Valid only for a renewal process.

We cannot ignore potential and likely real age effects.

We need to check for validity of *HPP*.

A better and less assuming approach to measure reliability is to *analyze the data versus system age*, that is, apply *time dependent reliability (TDR) analysis*.

Customers Want More than an *MTBF* Measure for Reliability

What is the reliability of servers? What should it be?

What are the causes of downtime?

What can we expect going forward?

The answers can be provided by time dependent reliability (*TDR*) analysis.

Case Study Example: E10K Performance at Customer's Sites

- Customer has five E10K systems installed in early 2001.
- Customer is concerned by failures over last year that have caused downtime and impacted production.
- How did we typically react in the past?

Typical Response to Customer

Your measured *MTBF* is *X* hours.

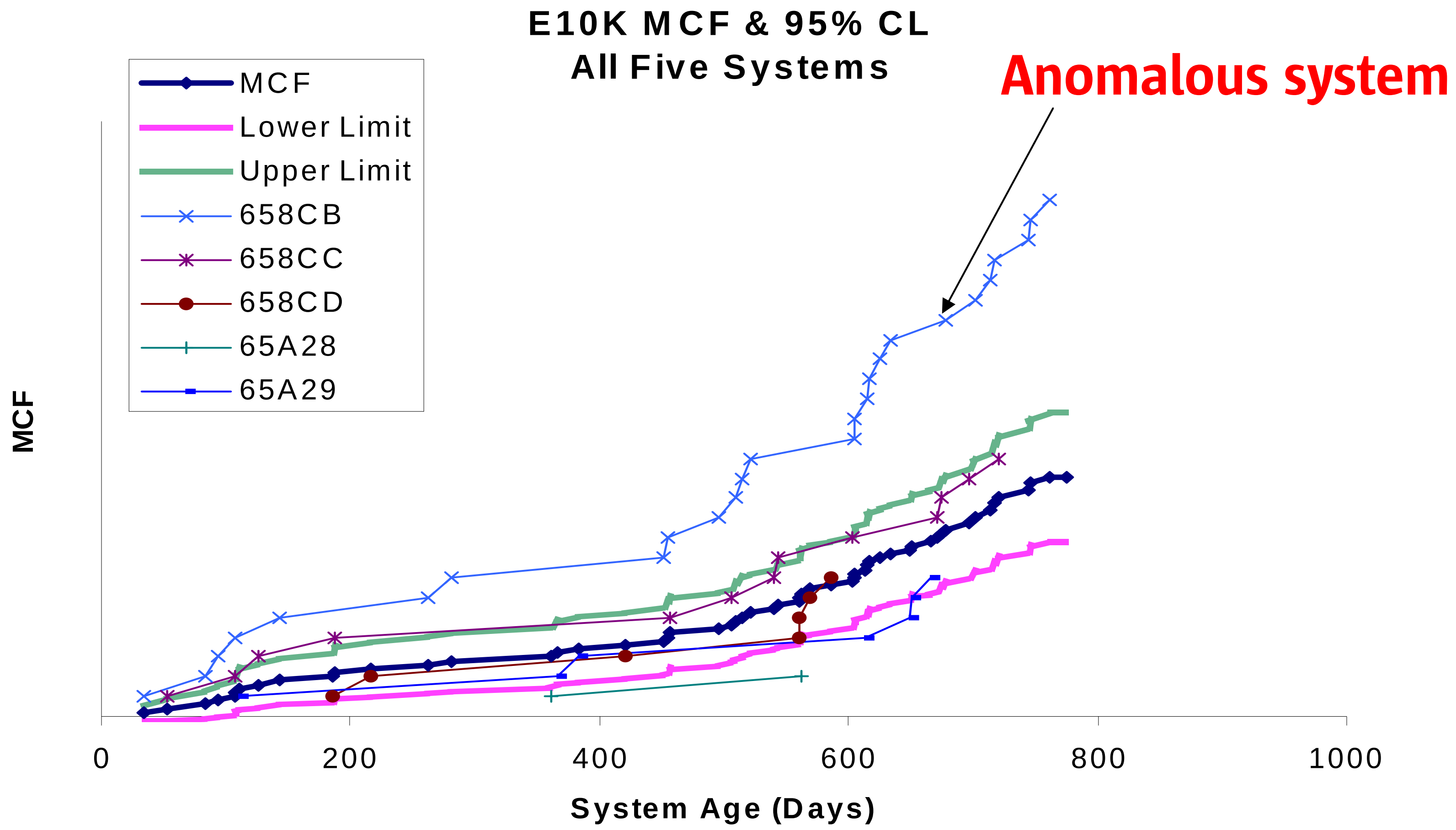
We're very concerned about reliability.

We'll work with you to improve.

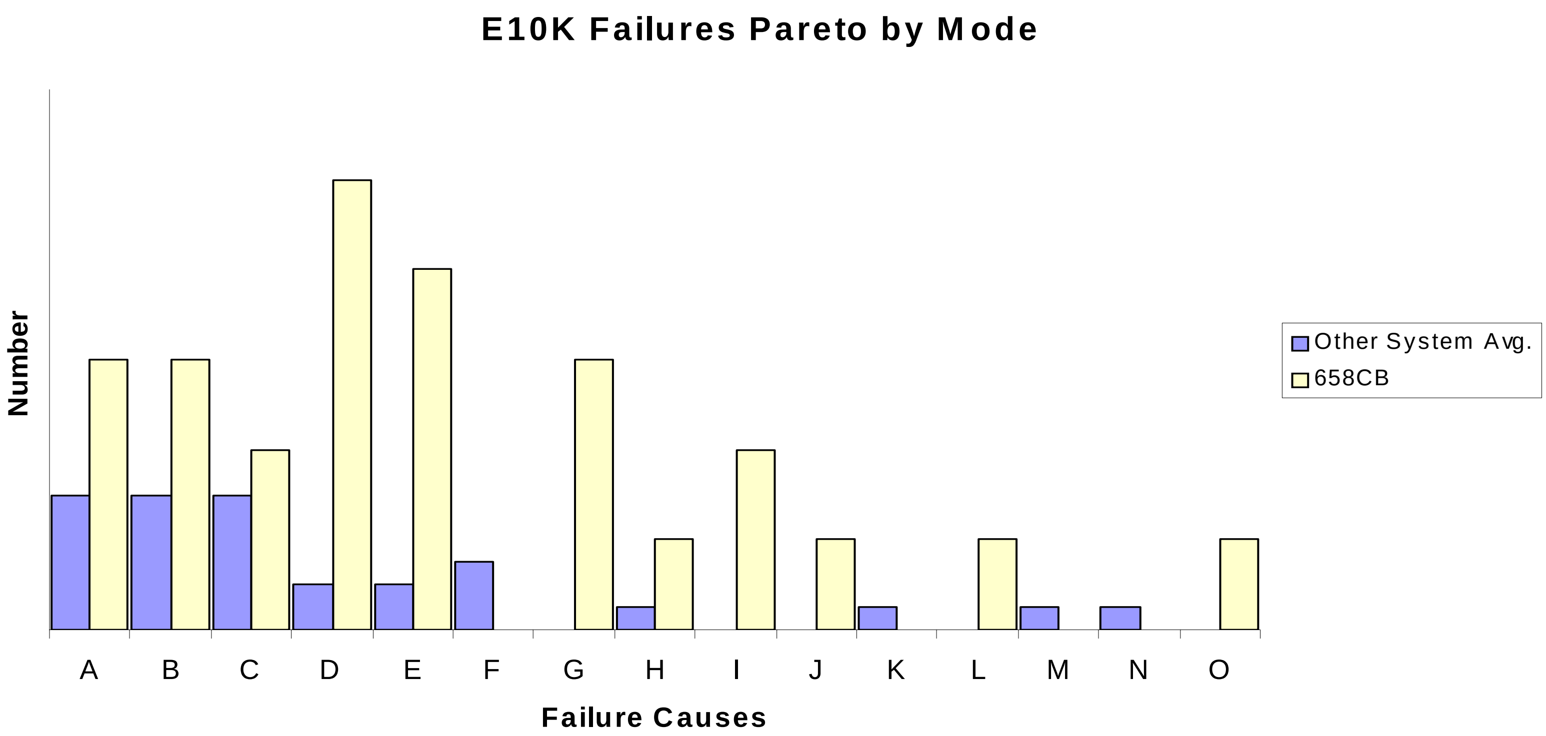
Result: Customer dissatisfied. Doesn't know what to expect going forward.

How do we react now?

TDR Plots for E10K Systems (with Confidence Limits on *MCF*)



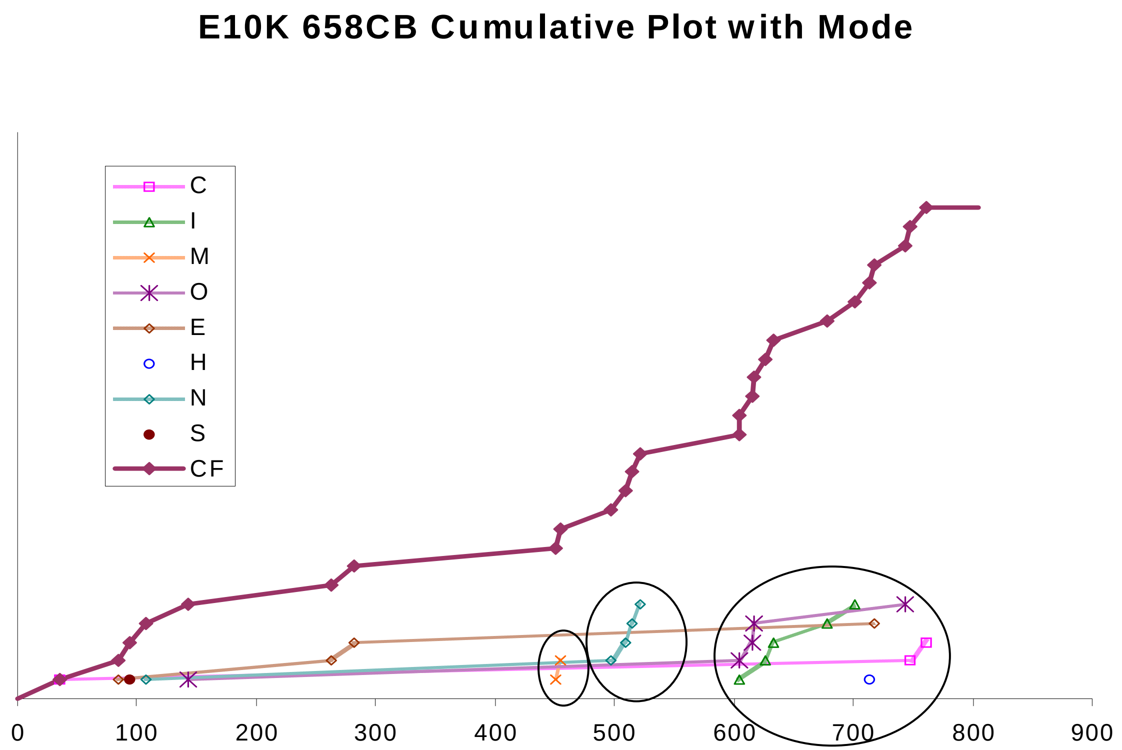
E10K Cause Code Comparison
(Pareto Chart of Count)



Pareto chart shows different distribution of cause codes for anomalous system compared to other four servers.

Static: Does not show time dependency of causes.

Dynamic Analysis of Anomalous System



System cumulative plot with cause codes shows the *dynamic* time dependent effects of different failure types on reliability. *TDR* reveals *clustering* of failure causes.

E10K Cause Code Summary

Problems appeared when customer put systems into full production mode.

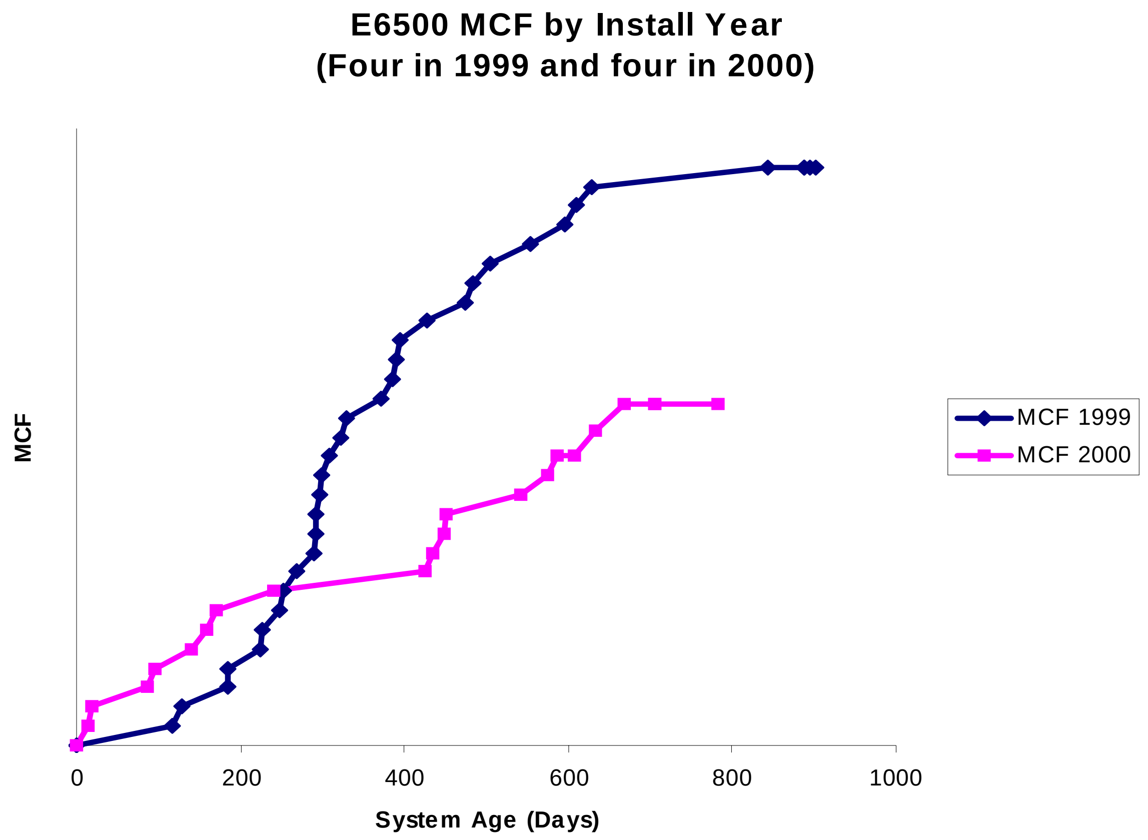
Multiple failures for same or different causes in a short time period revealed inability to diagnose and repair correctly the first time.

Failure rates consequently are made artificially higher and *MTBF*s lower by repeated repairs for same problem.

Possible *MCF* Comparisons

By platform
By customer
By vintage
By age (left and right censoring)
By calendar date
By location
By failure cause
By supplier
By technology
By payload or applications

TDR for Vintage Analysis *MCF* by Year Installed



Significant difference between years.

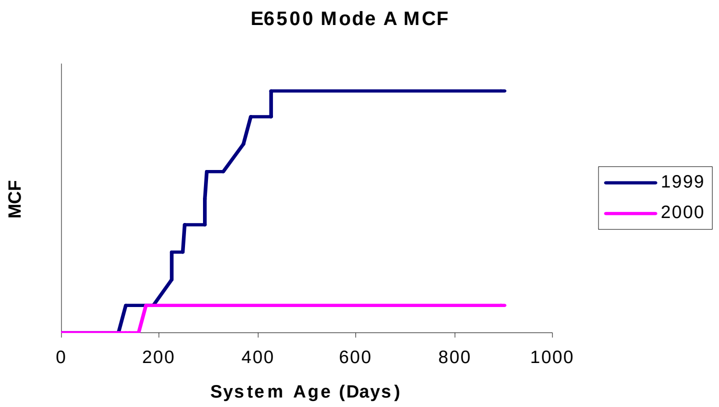
TDR Analysis of Individual Cause Codes

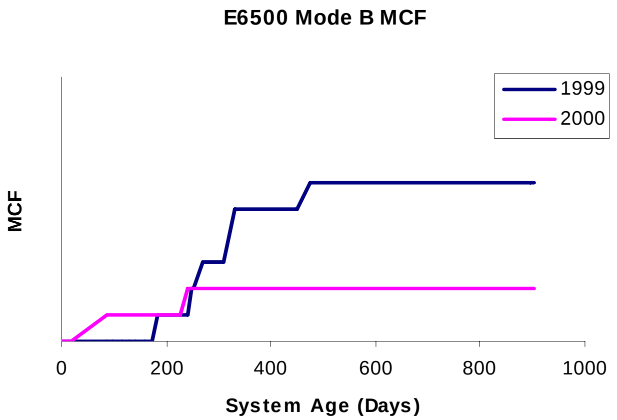
Individual *MCF* plots can be constructed for each cause code to reveal failure trends.

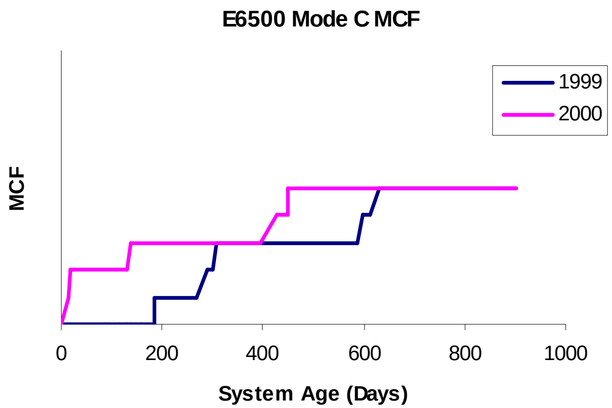
TDR analysis of each cause code provides time related information not available through static Pareto analysis.

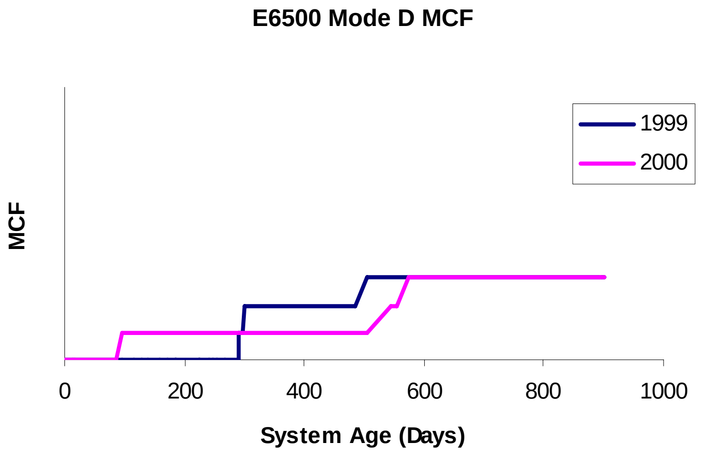
We will illustrate this concept with a group of E6500 systems.

E6500 Cause Analysis (*MCF* of Top Four Causes by System Age)



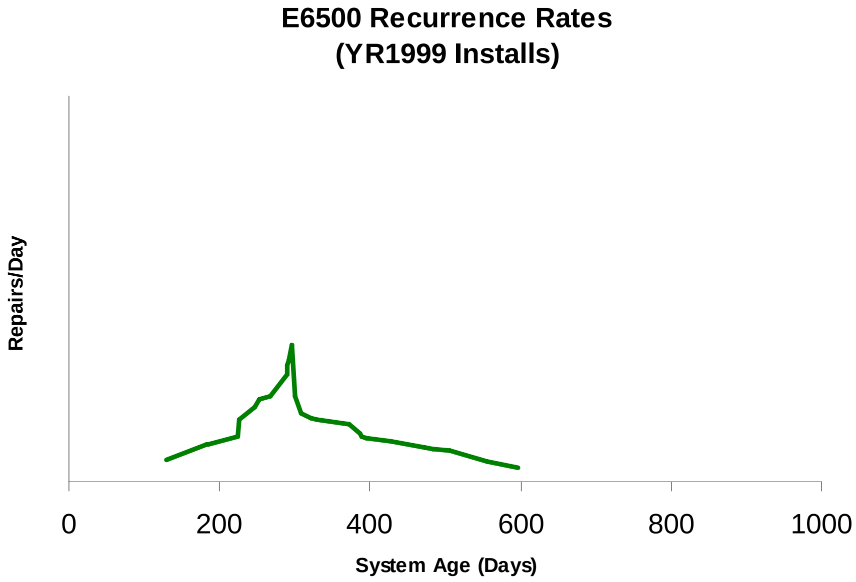


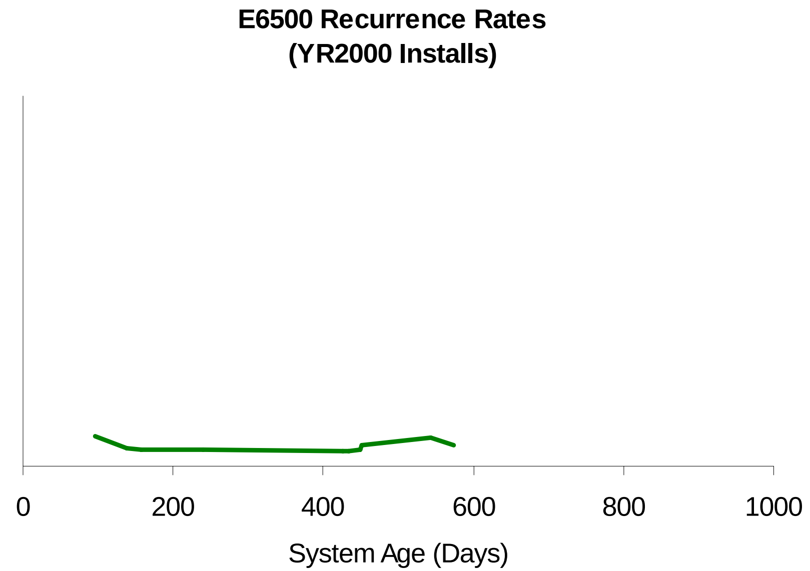




E6500 Recurrence Rates

By numerical differentiation of the *MCF*, it is possible to estimate repair or recurrence rates versus the system age.





Even though 1999 vintage spiked around 300 days, recurrence rate has since decreased to be the same as 2000 vintage systems.

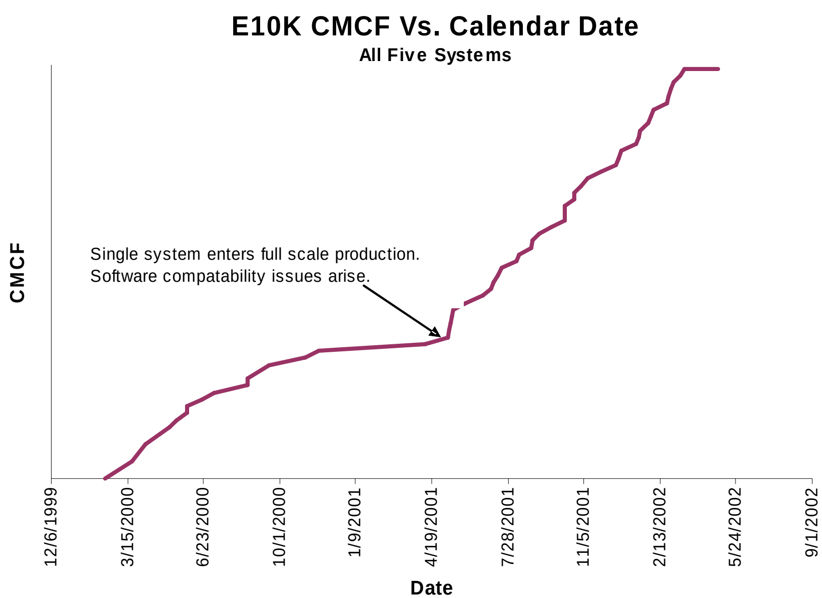
MCF: System Age Versus Calendar Date

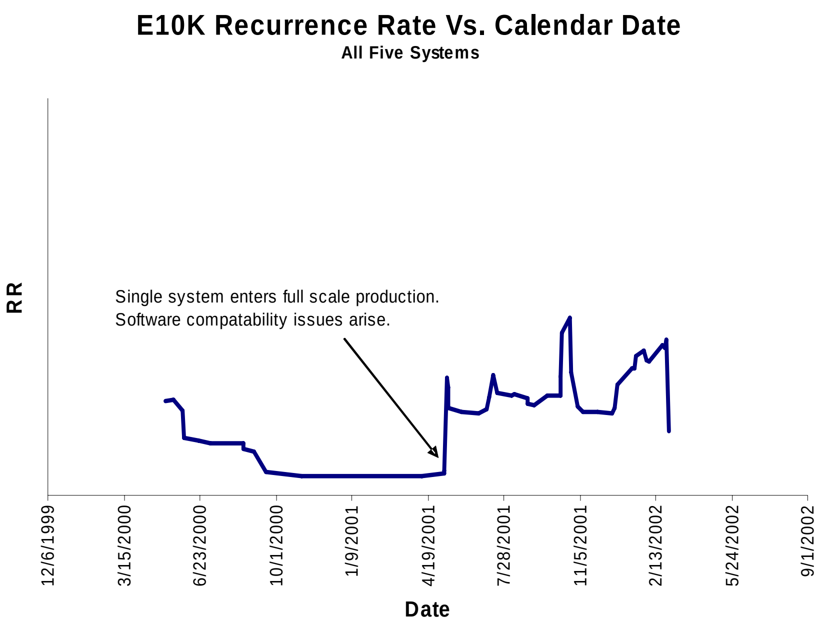
MCF versus *system age* shows cumulative repairs per system that depend on the system operating hours.

MCF versus *calendar date* may reveal repairs across systems associated with actions (physical relocation, software patches, upgrades, etc.) during specific time periods.

Both types of plots are useful for cause analysis.

E10K Calendar Date Analysis





On 5/1/2001, system entered full production mode. Anomalous behavior caused by several software compatibility issues.

TDR Data Needs

For *each and every* system type (*full inventory by serial number*) at a specific customer site:

Date installed (Install)

Date data capture began (*Begin*)

Date of each failure, *if any (Failure)*

Failure cause for each failure

Description of repairs

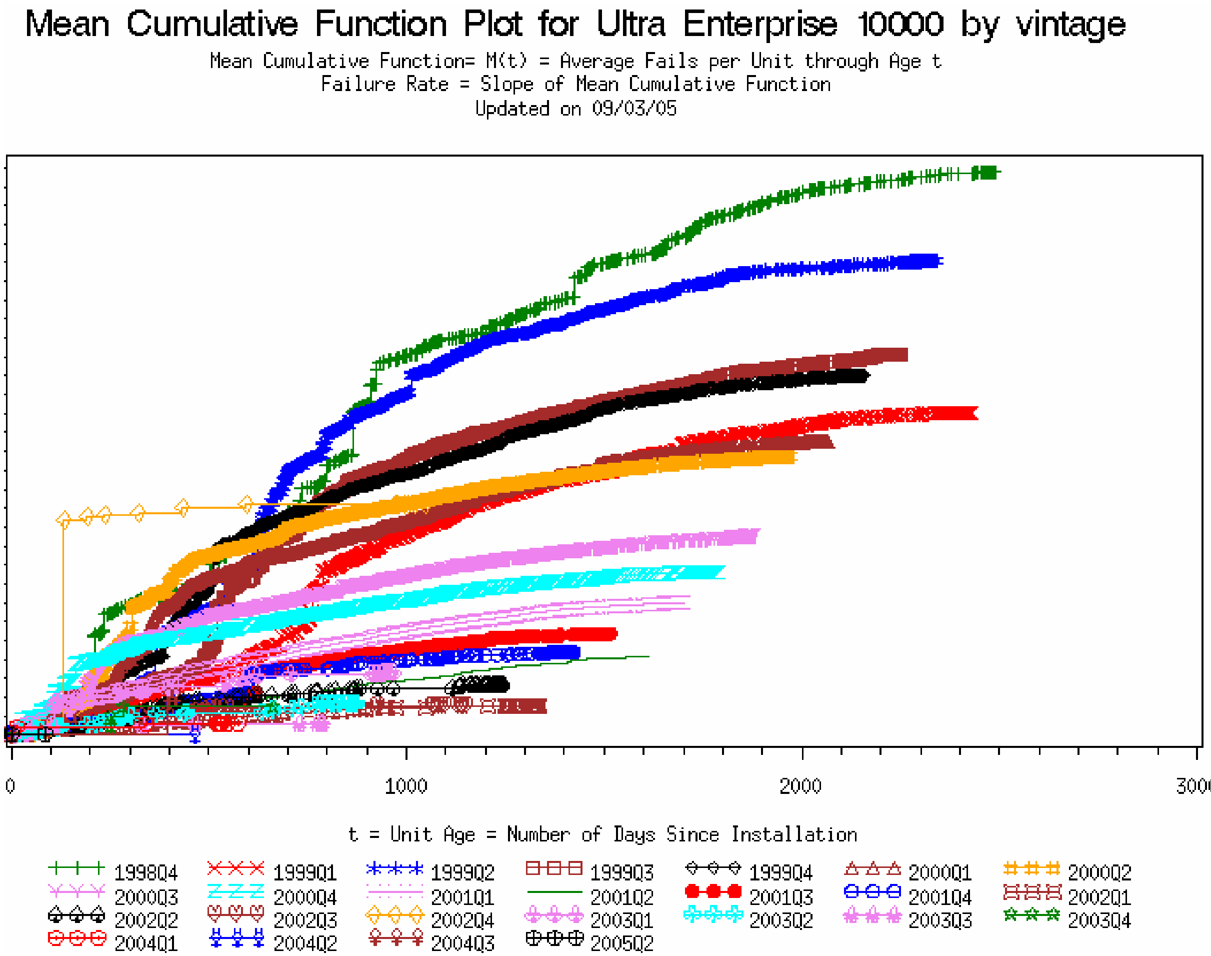
Current or removal dates (*End*)

Configuration information

Special comments (applications, load, etc.)

Sun Product Quality Portal Example

Note reliability improvement across vintages.
Longest MCF curves represent oldest vintages.



Brief Recap for Analysis of Repairable Systems

Study the *pattern of times between failures* of systems to show reliability performance.

Use graphical and statistical tools for analysis.

Understand the limitations of summary statistics like *MTBF*.

Recognize the *power of TDR analysis* to reveal what's happening with systems in the field.

Remember...

Reliability is *time dependent*.

TDR analysis can reveal trends.

Think **cumulative plots**, *MCF*, and **recurrence rates.**

Display results **graphically**.

Track failures and downtime **by system** versus **age and calendar dates.**

Identify **anomalous** behavior.

Exploit TDR **analysis** to drive appropriate preventive and corrective actions.

References

H. Ascher, “A set of number is NOT a data-set”, *IEEE Transactions on Reliability,* Vol 48 No. 2 pp 135-140, June 1999.

J. Usher, “ Case Study : Reliability models and misconceptions”, *Quality Engineering*, 6(2), pp 261-271, 1993

W. Nelson, *Recurrence Events Data Analysis for Product Repairs, Disease Recurrences and Other Applications*, ASA-SIAM Series in Statistics and Applied Probability, 2003.

P.A. Tobias, D. C. Trindade, *Applied Reliability*, 2nd ed., Chapman and Hall/CRC, 1995.

D. C. Trindade, Swami Nathan, Simple Plots for Monitoring the Field Reliability of Repairable Systems, *Proceedings of the Annual Reliability and Maintainability Symposium (RAMS)*, Alexandria, Virginia, 2005.

Analysis of Field Data for Repairable Systems

David C. Trindade, Ph.D.

david.trindade@sun.com

Contributions by Swami Nathan, Bill Chen, Doug Ray, and Jeff Glosup.

Photo by Dave Trindade