GENERAL DYNAMICS C4 Systems

Six Common Mistakes Engineers Make When Dealing with Availability

Dr. Jeffrey Harris General Dynamics C4 Systems Scottsdale, Arizona Jeff.Harris@GDC4S.com

October 23, 2008

What is Availability?

- Availability is the statistical probability that a system will be ready for your use when you want to use it
- Often confused with a different but related statistical topic – reliability
- Availability is important because
 - Users build expectations about the system around it
 - Engineers design to it
 - The cost of the system is driven by it
 - Operations people get paid by it



Design for High Availability is Risk Mitigation

- There is always a risk that some components in a system will fail
- Specifying that a system will have a high availability transfers that risk from the owner of the system to the designer
- The designer must take steps to mitigate that risk, and will charge the user for everything he does
- The more the risk is reduced, the more is costs the owner
- The progression is not necessarily linear

Common Mistakes

- Problems arise when the customer does not understand
 - What availability means
 - How the system is architected
- Communication can help avoid these common pitfalls
- Six Common Mistakes
 - Expressing availability as time
 - Treating availability as a single dimension
 - Failing to limit the scope of availability
 - Failing to identify partial availability
 - Failing to account for preventative maintenance
 - Measuring performance short term

Mistake #1 – Expressing Availability as Time

- This is a common chart shown for availability
- Unfortunately it is wrong, and conveys inappropriate expectations

Availability	Unavailability (Minutes/Year)
99.9% 99.999% 99.9999% 99.99999%	525.6 52.56 5.256 0.5256 0.05256

Availability is not time, it is a probability

A Better Slide

Availability	Unavailability
99.9%	0.1%
99.99%	0.01%
99.999%	0.001%
99.9999%	0.0001%
99.99999%	0.00001%

- This chart introduces no erroneous information
- Often the unavailability of a system is what we really want to know - outage

Mistake #2 – Treating Availability as a one Dimensional Point

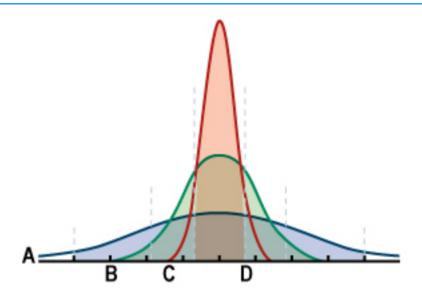
- Albert Einstein is quoted as having said
 - "Everything should be as simple as possible, but no simpler."
- Why?
 - As a university professor, I always strove to make things as simple as possible, in order to communicate understanding of them
 - But when you make things simpler than they are, you lose essential information.
 - Design a box for me
 - The height is 2 feet
 - The depth is 1 foot
 - Go
- This is what we do with Availability all the time
- Availability is not one dimensional, its two dimensional

Another Look at the Misleading Chart

Availability	Unavailability (Minutes/Year)	
 99.9% 99.99%	525.6 52.56	
99.999%	5.256	
99.9999% 99.99999%	0.5256 0.05256	

- Is a 5-9's system always down for 5.256 minutes each year?
- When in the year? At the beginning? At the end?

Availability is a Probability, and Probabilities have Distributions



- All four systems described by curves A,B,C, and D have the same statistical mean
- Yet they describe systems that behave differently

Consider a System with 50% Availability

- Sometimes to understand something you have to take it to a ridiculous extreme
- Lets consider a system with 50% availability
- All of the following systems have 50% availability
 - One year on One year off
 - Six months on Six months off
 - One day on One day off
 - One minute on One minute off

Mistake #3 – Measuring Actual Performance in a Short Time Horizon

- Statistical probabilities are meant to represent behavior across long time horizons such as 30 years
- A mistake is made when compensation is based on short term measurements
- Consider the 50% availability system and a six month measuring period
 - If the system is in a mode of 1 year up and 1 year down, you will believe that the system is 100% available
 - If the system is 1 year down and 1 year up, you will believe the system has 0% availability
 - If the system is 1 day up and 1 day down, you will believe that the system is 50% available

How Should Performance be Measured?

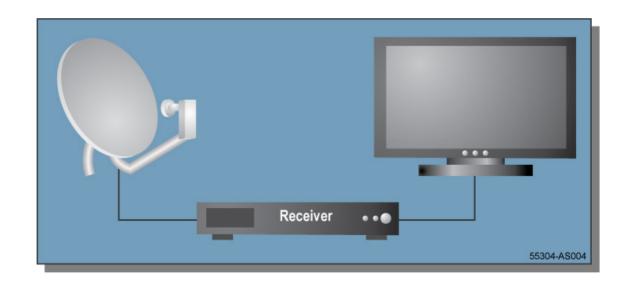
- The Availability formula is composed of Mean Time Between Failure (MTBF) and Mean Time To Repair (MTTR)
- Both MTBF and MTTR are distributions
- When you do the calculation by hand you are using the mean of each distribution
- Predicted availability should be described with a system model and then simulated

Availability (A) = MTBF/(MTBF + MTTR)

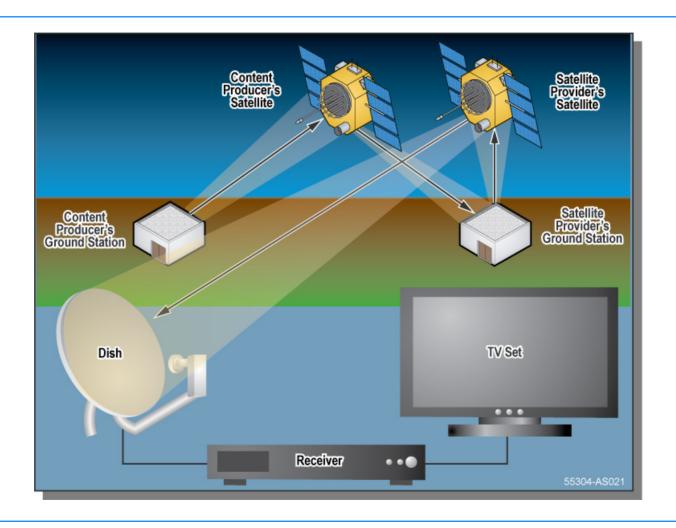
Unavailability (U) =
$$\frac{MTTR}{MTBF}$$

Mistake #4 – Failing to Limit the Scope of Availability

Always ensure that the scope of availability is within your scope of control



The Bigger Picture

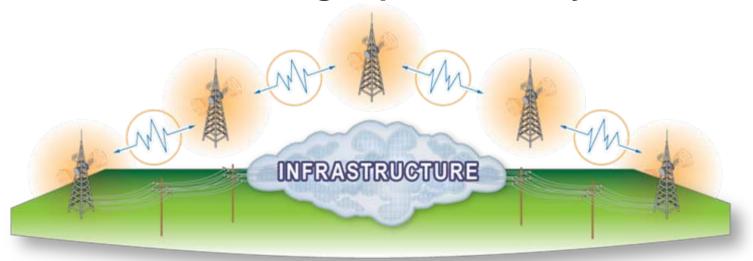


GENERAL DYNAMICS

C4 Systems

Mistake #5 – Failing to Identify Opportunities for "Partial Availability"

- Modern systems are extremely complex and often spread out over large geographical areas
- Availability should be expressed on a sub-system or geographical basis where applicable
- View this as a learning experience for your customer



Mistake #6 – Failing to Account for Preventative Maintenance

- You and your customer need to decide if preventative maintenance is inside or outside the scope of the availability calculations
- It drives radically different designs
- I advocate for full communication with the customer



Summary

- Problems arise when the customer does not understand
 - What availability means
 - How the system is architected
- Communication can help avoid these common pitfalls
- Six Common Mistakes
 - Expressing availability as time
 - Treating availability as a single dimension
 - Failing to limit the scope of availability
 - Failing to identify partial availability
 - Failing to account for preventative maintenance
 - Measuring performance short term

Thank You

Contact

Jeff Harris
Jeff.Harris@GDC4S.COM

GENERAL DYNAMICS

C4 Systems