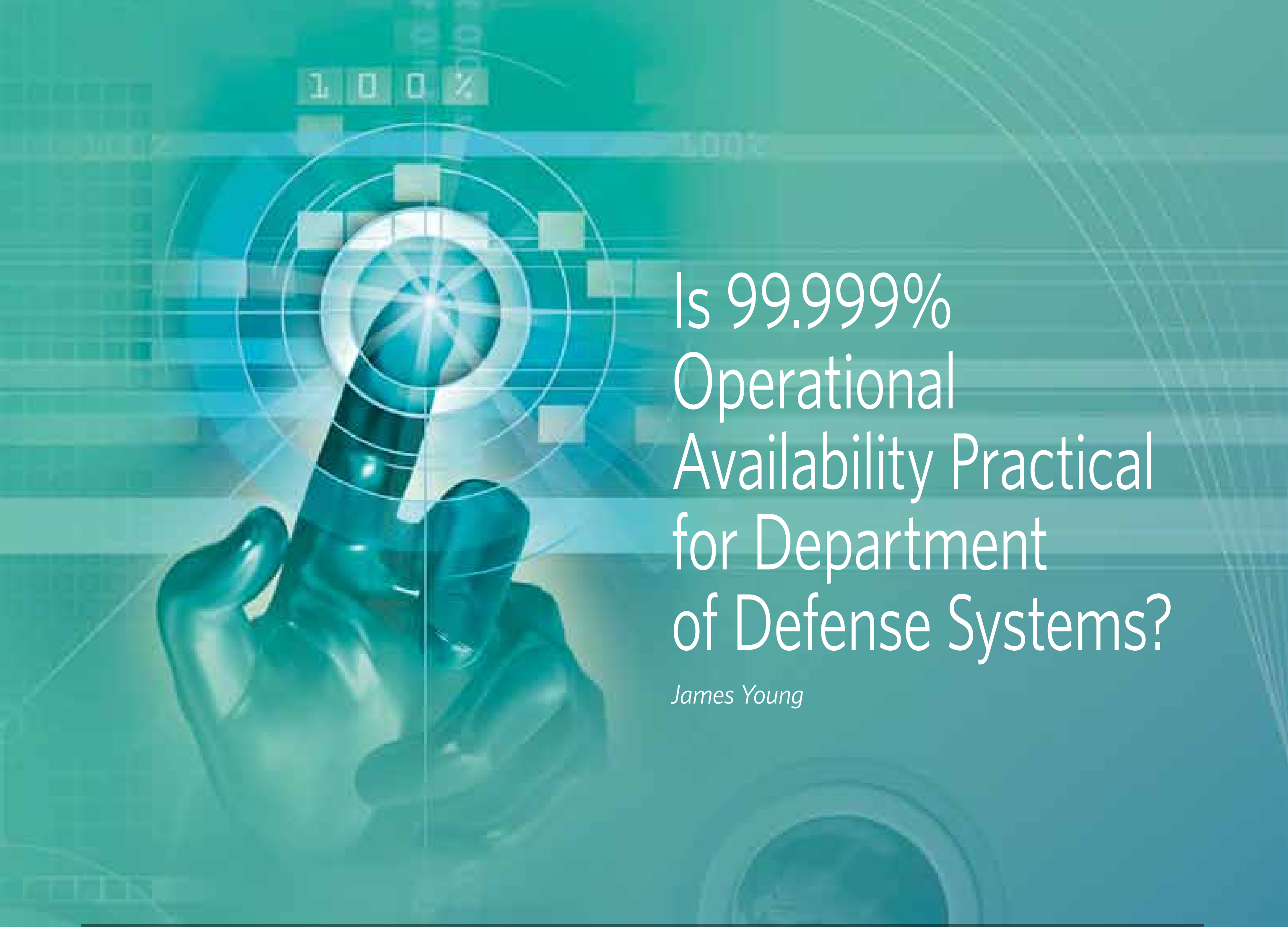


Is 99.999% Operational Availability Practical for Department of Defense Systems?

James Young

Changes are needed to make significant improvements to operational availability and must be considered as early as possible during the design cycle; however, after initial system development, design changes are typically cost-prohibitive. The Department of Defense needs to ensure maintenance and supportability are considered during all phases of the system development cycle, particularly during initial design. That becomes evident when one considers

***Young** is an integrated logistics manager at the Naval Surface Warfare Center, Crane Division, and he supports the Office of Naval Research in the C4ISR Department. A former Navy officer, he has a Bachelor of Science degree in occupational education and is currently pursuing a Master of Science degree in systems engineering at Johns Hopkins University Applied Physics Laboratory.*

that the largest cost of a system's life is consumed during the operating and support phase, and by the time a system reaches the production and deployment phase, at least 70 to 80 percent of the operating and support phase costs of the system are already set (see Figure 1).

Changes after the concept exploration/definition phase are cost-prohibitive and would require a substantial investment in redesign, remanufacturing, and production; as well as installation and fielding of the improved hardware/software, among other tasks. Supportability experts must be involved and be considered principal stakeholders during the early design phase of a system, allowing cost-effective supportability to be designed into the system. Even though some programs state that supportability and affordability are very important in the development of a new system, they are not provided the same importance as technical specifications or per-unit production costs. DoD is missing an opportunity to save significant money by ensuring life cycle costs and associated supportability are fully considered during early stages of system design.

Consider mean logistics delay time and the fact that it has a significant effect on operational availability. This article demonstrates that reducing mean logistics delay time and mean time to recovery—the average time that a device will take to recover from any failure—while increasing the value of the mean time between failures can easily be done.

Commercial Versus Government

Let's consider some initiatives that have worked for the commercial sector and consider applying them to government systems. Commercial satellite systems and commercial computer servers for financial institutions often reveal operational availability values approaching five nines, which indicate 99.999 percent availability. Satellite television and servers are important to a large number of people, as they will notice and be inconvenienced if their service is disrupted. They are also important to business. A loss of service means a loss of dollars. In some cases, millions of dollars per minute are lost in the event of a complete server or satellite failure.

Typical weapons system operational availability values are very good if the system achieves an operational availability of 90 percent. Keep in mind that with a critical weapon system, a loss of service at an inopportune time may cost a great deal more than millions of dollars per minute—we may lose hundreds, if not thousands, of American lives. Personnel loss is capability lost. So when we consider loss of service of a critical weapon system, we must also consider the importance of the system to safety as well as the effects on the defense of the United States.

What makes the commercial sector able to achieve 99.999 percent availability while DoD systems are lucky if they achieve 90 percent? Why can't DoD weapon systems be as reliable as commercial systems? Hot swapping and redundancy are two items reflected in the commercial world that can benefit DoD systems and help them achieve higher availability.

Let's look at a computer server and how it achieves very high availability. One method large financial institutions use is to choose highly reliable assemblies or modules for computer servers. For example, computer hard disk drives typically have a five-year warranty and a stated mean time between failures of approximately 1.2 million hours. If those commercial enterprise computer hard disk drives were like government weapon systems, government employees would need to replace the hard disk drive at least every six months and spend a great deal of time reloading their operating systems and applications software. Imagine the loss of productivity and capability to do our everyday jobs with hard disk drives like that.

Hot Swapping

Another aspect of commercial servers is the ability to hot swap assemblies or modules in the event of a failure. (Hot swap refers to the ability to swap or remove a module or circuit card assembly and replace it with power on. Normally, one must power the system off, remove the faulty module, install a new module, power the system back up, then use the system.) Virtually all high-end servers now have the ability to hot swap, and those servers usually only cost thousands of dollars. Typical weapon systems are in the millions or tens or hundreds of millions of dollars range yet have availability values much lower than the typical high-end server and do not have the ability to hot swap assemblies or modules.

One method commercial enterprise computers use to achieve near-100-percent availability is to write the software so that upon a hardware failure, the computer will deallocate the faulty assembly from the resource pool and task other assemblies to do the tasks required. Is it possible to do this with the computers/processors, memory, etc., in our critical weapon systems? Yes, it is! Hot-swappable technology has matured significantly over the past several years and is now at the point where cost-effective system designs can readily use the technology. In addition, the costs for hot-swappable modules are very close to non-hot-swappable modules. Hot swapping in computer servers is so common today that costs have dramatically reduced.

We often hear the argument that hot swapping is much, much harder to do with radio frequency devices and circuits and other government technologies. But look at the commercial and government satellite industry. A quick Internet search will reveal thousands of vendors advertising their hot-swappable power supplies, processing boards, memory boards, storage devices, radio frequency and digital amplifiers, switches, and so on. If industry is doing it, why can't government? Why are we not performing hot swapping in critical weapon systems? We should be using hot-swappable assemblies as much as practically possible in our systems.

Redundancy

Another area of consideration as DoD seeks to achieve 99.999 percent availability is redundancy. Have you noticed how the phone system works fine the vast majority of the time? Have you also noticed that when a catastrophe happens (like the Sept. 11, 2001 terrorist attacks), suddenly you cannot call anywhere? That indicates there is excess capacity built into the phone system for typical usage, but in the event of a disaster, the system cannot handle the volume, and the excess capacity is all used up. If the phone system were more critical, then excess capacity would enable us to call whenever we wanted—even during catastrophic events.

DoD should build in some excess capacity for critical weapon systems during the early design phase so warfighters never experience the inability perform vital tasks. How much excess capacity to build in must be determined based on the criticality of the functions. We need to do some analysis and choose the optimal level of redundancy, highly reliable assemblies, hot-swappable assemblies, excess capacity, etc., in our critical weapon system design. Single-point-of-failure items are good candidates for built-in redundancy.

Redundancy is typically viewed as cost prohibitive, but it should be considered for most critical functions. If we have a system design and conduct some analyses to determine very critical functions, then we can do a cost-versus-capability analysis to determine if the operational importance of the

Figure 1. **Life Cycle Cost by Phase**

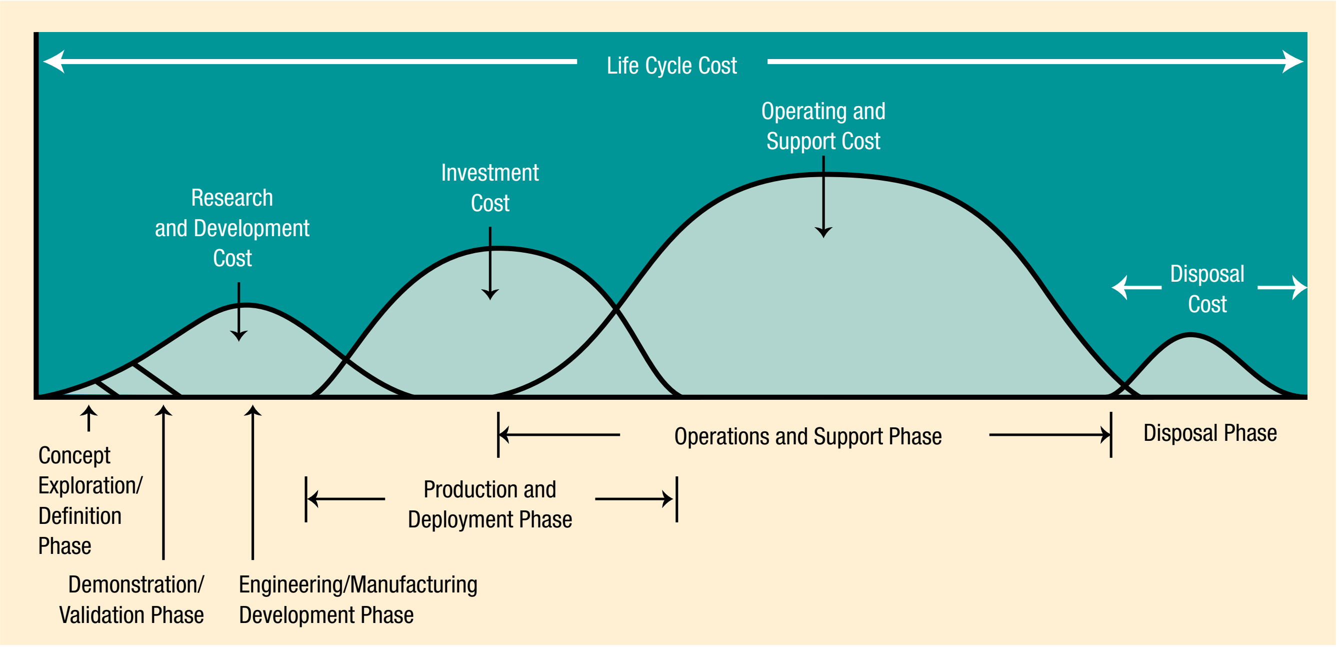
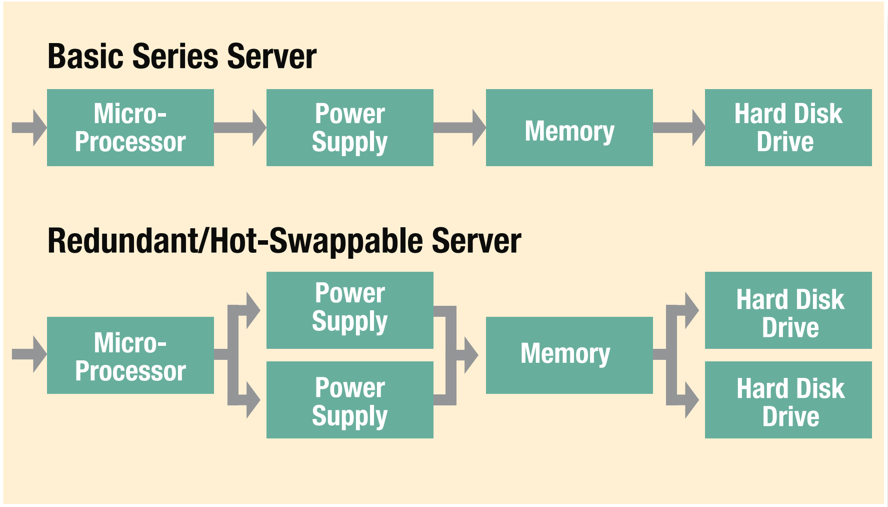

Life Cycle Cost
Operating and Support Cost
Investment Cost
Research and Development Cost
Disposal Cost
Operations and Support Phase
Disposal Phase
Concept Exploration/ Definition Phase
Production and Deployment Phase
Demonstration/ Validation Phase
Engineering/Manufacturing Development Phase

Figure 2. **Basic Server Configuration Versus a Redundant/Hot-Swappable Server**



functions is worth spending more money to have redundancy and/or excess capacity.

Hot Swapping and Redundancy Examples

For the greater operational availability techniques I've discussed to be fully realized, new system hardware and software designs must periodically and automatically check the status of all assemblies in the background without affecting normal operation; electrically remove or disconnect faulty modules from the resource pool; provide seamless operation to the operator; automatically notify maintenance personnel of fault conditions with full descriptors for action required; enable hot swap capability; and reallocate the new assembly to the resource pool.

To illustrate those tasks fully, let's consider a very basic example of a typical server and the effects of redundancy and hot-swappable assemblies on the overall cost and availability of the system and plot this as a representation of cost versus availability over the life cycle of the system. Let's consider a basic cost analysis of each of these systems. Figure 2 compares a basic server with a server with redundancy and hot swapping.

If we were to consider the support cost of the basic and high-end servers, we would discover an increase in costs for the modules to support the redundant and hot-swappable system. A simple example of that is illustrated in Figure 3. You'll notice that the cost of each module that is hot-swappable is higher than the basic server. Also, you'll notice we will be paying for more failures. You might ask, "Is paying approximately 50 percent more in parts costs per year a viable option?" At first glance, it doesn't appear to be wise thing to do; however, with the addition of redundant modules, as well as the ability to hot swap in the event of a failure, the mean time to recovery will be much less than if we had to power the system down.

Other Concepts

Some other concepts DoD should consider during the design phase:

Fault-Tolerant/Switching

Many systems use fault-tolerant designs that switch over to other devices or reroute signals when faults occur, thereby increasing overall availability. If automatic fault switching is included in the early design phase, it becomes a viable option to achieve high levels of availability. Fault-tolerant designs and switching can be leveraged and applied to an entire system rack. In the event of a failure, the operator receives a fault message/indication. The system continues normal operations while maintenance personnel removes and replaces the faulty module. The repair is accomplished without shutting the software down, powering the server down, or loading/initializing software.

Cost-Based Selection/Optimization

Cost must be one of the major determinants when architecting a system-level design. Operations and support costs play a major role in overall system costs, while development and production are mere fractions of the overall costs of the system life cycle. Designs that leverage cost as an independent variable and influence the design will result in significant savings over the life cycle of the system.

Critical Functions Analysis

A critical function analysis is required to determine if redundancy, fault tolerance, very-high-reliability parts, or ready spares, etc., are needed and are appropriate for the design, or at least for the most critical functions the system performs. In order to determine which components, modules and/or assemblies are critical, an analysis must be performed. If the critical functions analysis reveals

DoD needs to ensure maintenance and supportability are considered during all phases of the system development cycle.

Figure 3. **Basic Cost of Server Hardware**

Basic Series Server

Module	Quantity	Cost Each	Mean Time Between Failure	#Fails/Year	Cost Sub-Total	Hot Swap
Micro-Proc	1	$1,500	28,000	0.31	$469.29	No
Power Supply	1	$500	12,000	0.73	$365	No
Memory	1	$700	50,000	0.18	$122.64	No
Hard Disk Drive	1	$300	70,000	0.13	$37.54	No

Redundant/Parallel Server

Module	Quantity	Cost Each	Mean Time Between Failure	#Fails/Year	Cost Sub-Total	Hot Swap
Micro-Proc	1	$1,500	28,000	0.31	$469.29	No
Power Supply	2	$600	12,000	1.46	$876	Yes
Memory	1	$700	50,000	0.18	$122.64	No
Hard Disk Drive	2	$360	70,000	0.25	$90.10	Yes

single-point failures in the design, those failures should be dealt with by selecting highly reliable parts and applying redundancy, fault-tolerant design via switching to other devices, etc.

Ready Spares

The methodologies I've discussed will keep a system running in the event of failure, but eventually, a replacement part will be needed. Currently, in many cases, two weeks is a reasonable time to wait for a replacement part; however, that is not an acceptable length of time if we're to aim for greater operational availability. The spare must be readily available and easily installed for us to realize the maximum benefits of the methods I've discussed. An inventory of ready spares of the most critical assemblies should be stocked in equipment spaces in order to enable rapid removal and replacement upon failure.

If we apply the concepts previously described, particularly redundancy, or have excess capacity for critical functions, then the system can provide near-perfect operational capability even upon failure of critical modules or assemblies—giving us time to replace the part with a spare. For example, if a system has an optimal response time of 10 microseconds and, in a degraded mode, the response time is 15 microseconds, then a slightly degraded response time can easily be tolerated for the relatively small amount of time it will take to hot swap the faulty assembly with a ready spare. Ready spares of critical assemblies must be on hand for trained technicians to quickly and efficiently hot swap the faulty assembly and go from degraded operation to full capability within minutes.

Weighing Costs

We must weigh costs versus operational availability. A constant argument with system design is how much operational availability can we afford? I think we should apply more resources and money during system design to the methodologies I've mentioned. If we do that, we can make cost-effective improvements to the system and improve operational availability; and in the event of a failure, the system can still operate in a satisfactory manner. The excess capacity and/or redundancy will enable the system-level performance to stay practically constant, and the operator may not even notice a change in performance. But we must conduct analyses to determine what the effects on performance would be versus how much we are willing to spend for more operational capability. In most cases, paying a little additional procurement and support cost is justified if significant improvements in operational availability are achieved.

By studying the initiatives mentioned in this article, we can obtain near-perfect availability for DoD systems at very reasonable costs. We should all strive to provide our service personnel with systems that are as reliable as practically possible, are relatively easy to repair, and have near perfect operational availability. The technology to accomplish this is available now and is affordable. What are we waiting for?

The author welcomes comments and questions and can be contacted at james.m.young@navy.mil.