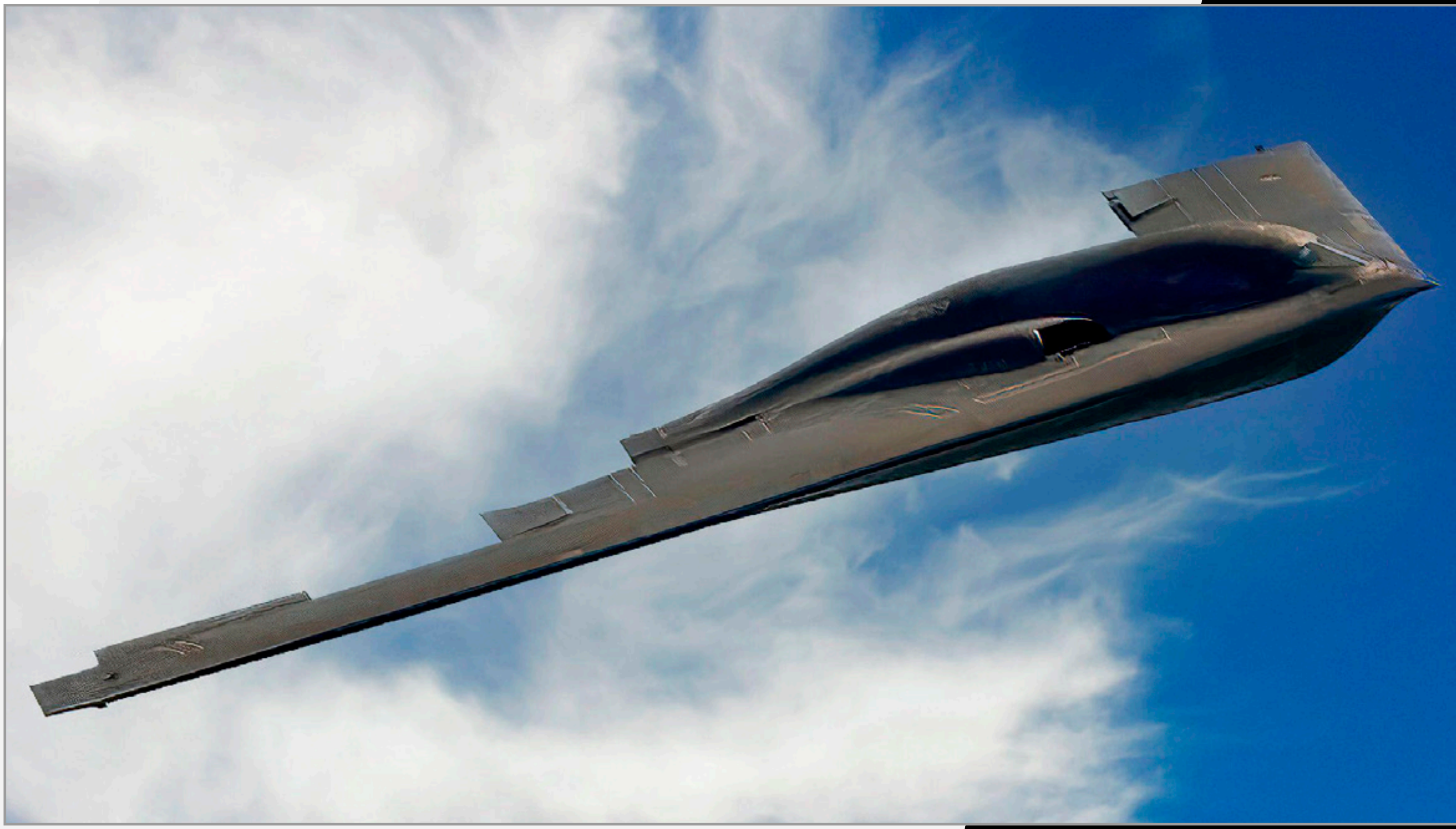
CURTISS-WRIGHT

Removing 10 GbE Network Attached Storage Bottlenecks

Network attached storage (NAS) devices have the ability to protect and store critical data at high-performance speeds.

For years, defense and aerospace organizations installed direct attached storage (DAS) in deployed vehicles. However, as storage needs increase and Ethernet has become commonplace in mission vehicles, NAS solutions have become an essential option for data storage. When a NAS device is used, the stored data can be made available to all devices and computers on the Ethernet network, not just a single computer.

Contents



NAS Components

A NAS device will have three basic sections or subsystems – input/output (I/O), a central processor unit (CPU), and memory (MEM). By definition, a NAS device connects to an Ethernet network providing remote data storage to the network clients.

That Ethernet network may operate at one gigabit per second (1 GbE)[i], ten gigabits per second (10 GbE)[ii], or higher (40 GbE, 100 GbE). The selection of the Ethernet speed of the network depends on the clients and how much data they are expected to store on and retrieve from the NAS. In a deployed vehicle, network clients may be mission computers, sensor management computers, streaming cameras, radars, sonars, or other devices that collect or manage data.

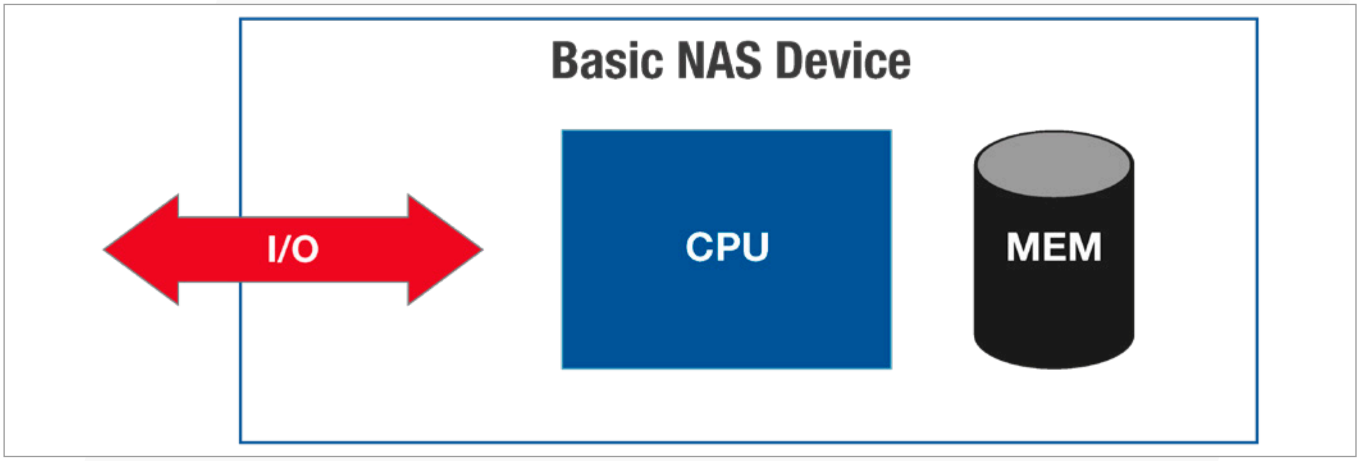


Figure 1 - Basic NAS Block Diagram

1 GbE Example NAS

If the clients are all 1 GbE based, then a NAS like that shown in Figure 2 may be used. The example 1 GbE NAS has two 1 GbE ports and is designed to fit into very small spaces on deployed vehicles. Its throughput speed is limited by the combination of the I/O (2 x 1 GbE ports), the CPU (1 ATOM processor), and the MEM (1 SATA solid-state drive or SSD). Also, encryption mechanisms inside the example 1GbE NAS affect the throughput.

Gigabit Ethernet has an actual line rate of 1,250 megabits per second (Mb/s). Because it uses 8b/10b encoding, the effective line rate is 1,000 Mb/s. Dividing 1,000 Mb/s by 8 bits per byte gives a theoretical byte rate of 125 megabytes per second (MB/s). Other factors like switch efficiency and network congestion may affect actual data transfer.

Figure 2 - 1 GbE NAS Example (DTS1)

10 GbE Example NAS

If some or all the clients are 10 GbE based, then a NAS as fast as the one shown in Figure 3 will be required. The example 10 GbE NAS is designed for high-speed data storage and can be matched with modern high-speed sensors on deployed vehicles. It has two 10 GbE optical ports, a multi-core CPU, and four high-speed SSDs.

10 GbE has an actual line rate of 10,320 megabits per second (Mb/s). Because it uses 64b/66b encoding, the effective line rate is 10,000 Mb/s. Dividing 10,000 by 8 bits per byte gives a theoretical byte rate of 1,250 megabytes per second (MB/s) which is ten times (10x) faster than 1 GbE. As noted previously, other factors may affect actual data transfer rates.

So, the 10 GbE NAS should absorb 10x more data than the 1 GbE NAS simply due to the faster I/O ports. What factors are required to support that throughput speed increase?

Figure 3 - 10 GbE NAS Example (HSR10)

Critical Factors for High-Speed NAS

As noted, any NAS device has three essential components – I/O, CPU, and MEM. Each factor or element must be matched to the other to support consistent data flow to the NAS storage and back out again when required.

NOTE: In this brief review, only write speeds will be considered. That is, writing to the solid-state disk(s). In a thorough analysis for a program, the read speeds from the disks must also be considered. The methodology will be the same.

Input/Output (I/O)

The example 1 GbE NAS in Figure 2 has two 1 GbE ports. Each port can transfer a maximum of 125 MB/s for a total of 250 MB/s. The example 1GbE NAS system does not actually store data at that maximum possible port speed. The primary reason for this is the two encryption layers – one hardware and one software. One integrated circuit (IC) performs the hardware full disk encryption (HWFDE) prior to storage on the internal SSD, limiting throughput to 150 MB/s. The software full disk encryption (SWFDE) further limits the effective throughput to slightly more than 100 MB/s.

The example 10 GbE NAS in Figure 3 uses a COTS I/O card from a 3rd party company. The two 10 GbE ports can theoretically transport 2,500 MB/s of data into (or out of) the NAS. These two 10 GbE ports are ten times (10x) faster than the two 1GbE ports in the example 1 GbE NAS. To match that speed properly, the I/O card must have an appropriate interface to the NAS.

In the example 10 GbE NAS, the 10 GbE I/O card interfaces with the NAS system via a PCI Express (PCIe) bus. The PCIe bus is generation 3 (Gen 3), with eight lanes, dramatically increasing the throughput. PCIe Gen 3 x 8 can transfer data at a rate of 7,880 MB/s. That is more than enough to handle the maximum write speed of 2,500 MB/s of the two 10 GbE ports. It can handle the read speeds as well.

As shown in Figure 4, the 10 GbE I/O board throughput is limited to the lowest rate, which is 2,500 MB/s. So, the whole NAS system can work no faster than 2,500 MB/s. Next, let us look at the memory on the backend, and last, let's look at the CPU to determine any other potential bottlenecks.

The 10 GbE I/O card has features that allow it to maximize the throughput, including a TCP/IP offload engine (TOE).

NOTE: This analysis only discusses data-at-rest (DAR), which is internal to the NAS. When data is in transit to or from the NAS device, it is known as data-in-transit (DIT).

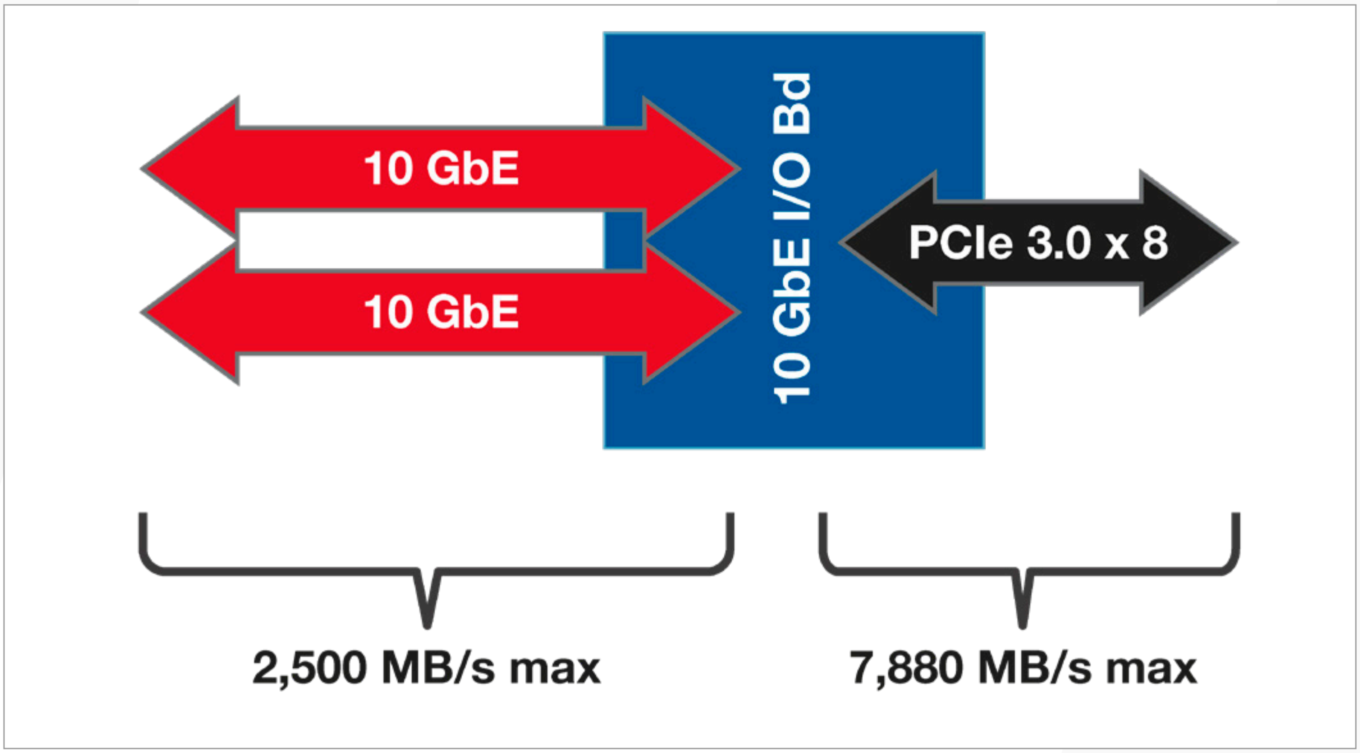


Figure 4 - I/O Board Throughput

Memory (MEM)

The purpose of a NAS device is to provide permanent data storage for the network clients. In deployed systems, solid-state drives are used instead of less reliable mechanical rotating drives. The SSDs can be either SATA or NVMe. Both SSD types use NAND flash for the actual storage. When data is stored on the MEM device, it is known as data-at-rest.

The 1 GbE NAS example in Figure 2 uses one 2.5” SATA SSD, which minimizes size, weight, and power (SWaP). The SSD is housed in a removable memory cartridge (RMC) for easy transport. The SATA SSD in the RMC can vary depending on the storage capacity selected and the NAND flash type (SLC, MLC, TLC[iii]). The eight terabyte (8 TB) SSD will be used for comparison, which is an MLC type. It is a SATA 3.0 device, meaning the SATA lane can transfer data up to 600 MB/s. The 1 GbE NAS system in Figure 2 operates slower than that maximum speed, but let’s use 600 MB/s for comparison. One SATA 3.0 SSD could not absorb the maximum data from two (or even one) 10 GbE ports. You would need either more SATA SSDs or faster different SSDs.

New NVMe SSDs are much faster than the SATA drives. In the example 10 GbE NAS in Figure 3, four NVMe SSDs are used in a removable memory module (RMM). Each NVMe SSD has a 4 TB capacity and a PCIe Gen 3 interface with four lanes with a maximum throughput rate of 3,900 MB/s each. With four of these NVMe SSDs, the maximum theoretical throughput for all four SSDs is 15,600 MB/s.

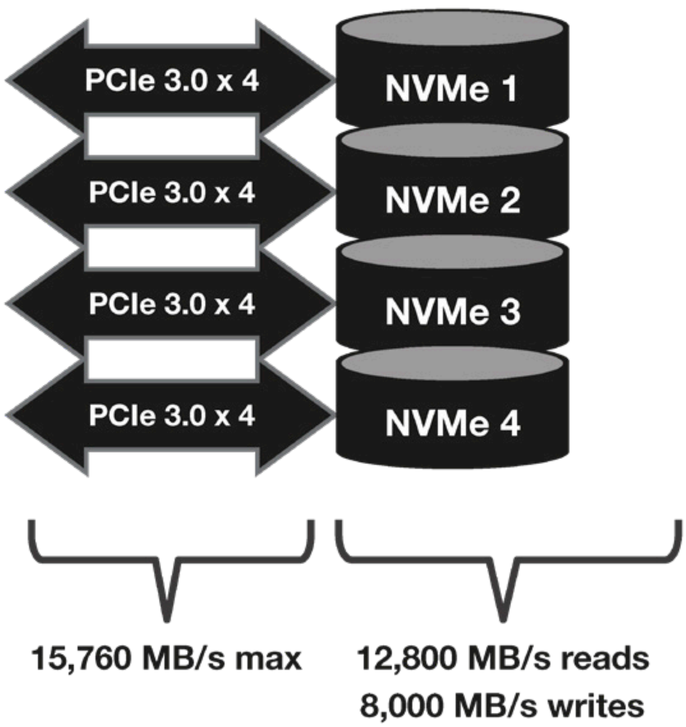


Figure 5 – NVMe Memory Throughputs

The PCIe interfaces are not the memory chain's limiting factor. The NAND flash behind the PCIe interface is the limiting factor. The resulting throughput for each NVMe SSD is 2,000 MB/s writes and 3,200 MB/s reads. With four SSDs, the total throughputs are 8,000 MB/s writes and 12,800 MB/s reads. Again, let's focus on the write speed.

So, you can see that the MEM system will not limit the NAS throughput. You are still limited by the two 10 GbE ports and their throughput of 2,500 MB/s. Now let's turn our attention to the CPU between the I/O and the MEM systems.

CPU

The CPU used in the 1 GbE NAS example is an Intel® Atom® processor with low power consumption. That selection of a low-power CPU matches the objective of that NAS, which is SWaP-optimization and is powerful enough to support the throughput rate of slightly more than 100 MB/s of the NAS system.

For the 10 GbE NAS example, the CPU is a generally available commercial off-the-shelf (COTS) module. At the heart of this CPU module is an Intel® processor with multiple cores, enhancing the number of simultaneous operations it can perform and thus enhancing the speed. The processor module consumes 35 watts (W) of power which exceeds the total power dissipation of the 1 GbE NAS example, which uses only 15W.

For the 10 GbE NAS example, the primary objective was to maximize throughput speed and remove potential bottlenecks. The CPU module supports two separate PCIe buses – one for the front-end I/O and the other for the backend memory.

The CPU supports the 10 GbE I/O card and has a PCIe 3 x 8 bus, which matches the front-end I/O card bus. Supporting the NVMe memory, the CPU has a PCIe 3 x 16 bus split into four groups, one 4-lane group for each NVMe SSD.

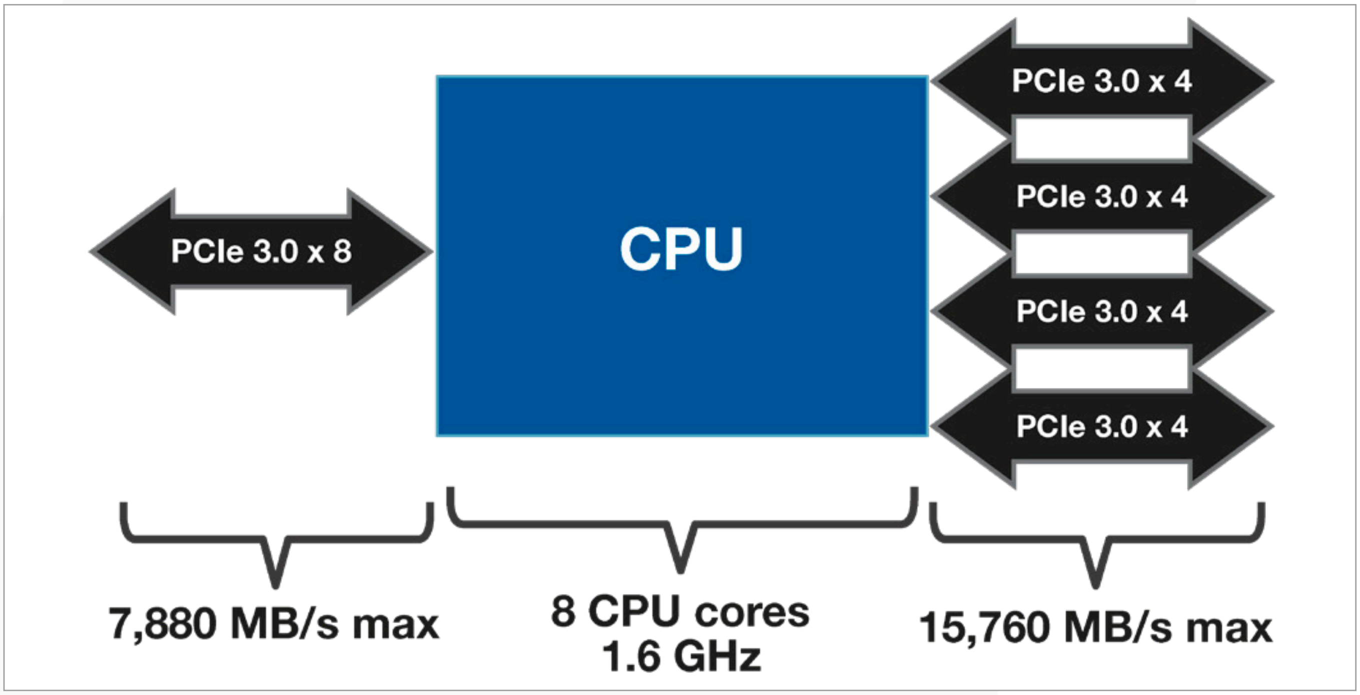


Figure 6 - CPU Throughput

Summary and Throughput Results

When evaluating a NAS device, determine your requirements and set priorities – SWaP, throughput, capacity, flexibility, etc.

If you minimize size, weight, or power, then that will usually mean reduced speed and throughput. As noted previously, the CPU alone in the 10 GbE NAS example consumes more power than the entire 1 GbE NAS example.

Regarding throughput, you must look at the entire system. If you have two 10 GbE ports, you must ensure that the CPU and MEM systems can handle the data. A fast PCIe bus can move the data from the I/O ports to the CPU. A fast processor can handle the data, and then a fast PCIe bus moves it along to the MEM system. Fast memory like NVMe is needed to absorb all the data without slowing down the system. Look at the entire system. Mis-matched components or systems can mean poor throughput performance.

The example 10 GbE NAS was designed with very high throughput in mind. The front-end I/O supports 10 GbE with two optical ports. The I/O performance is supported by a multi-core CPU which can absorb the incoming data and output the data efficiently to the MEM. The MEM system uses four NVMe SSDs, which have more than enough bandwidth.

As a result of careful design and planning, the 10 GbE NAS example can be written to and read from at very high rates. Throughput tests[iv] showed that the 10 GbE NAS example provides the following results:

- Store (write) data at 1,960 MB/s (1.96 GB/s) which represents 78.4% of the possible for two 10 GbE ports
- Read data at 2,130 MB/s (2.13 GB/s) which represents 85.2% of the possible for two 10 GbE ports

A summary of the throughput numbers for each section are shown in Table 1. The write and read test results are included in the last two rows.

NAS SECTION	INPUT	OUTPUT	COMMENT
10 GbE I/O Card	2,500	7,880	Two 10 GbE ports in PCIe 3.0 x 8 lanes out
CPU Module	7,880	15,760	PCIe 3.0 x 8 in PCIe 3.0 x 16 out
Removable Memory Module	15,760	12,880	PCIe 3.0 x 16 in Four NVMe memory in
Write Test Results	1,960		78.4% of Two 10 GbE Ports
Read Test Results		2,130	85.2% of Two 10 GbE Ports

Table 1 - Summary of Throughput for Each NAS Section (MB/s)

References

> [i] 1 Gigabit Ethernet may be known as 1 GbE, 1 GE, 1 GigE, or simply GbE. This paper uses 1 GbE.

> [ii] 10 Gigabit Ethernet may be known as 10 GbE, 10 GE, or 10 GigE. This paper uses 10 GbE.

> [iii] NAND flash types:
- SLC = single-level cell
- MLC = multi-level cell
- TLC = triple-level cell

> [iv] Available upon request to the factory.

Discover more

Products
> DTS1 & DTS1+: 1-slot Rugged Network Attached File Server
> HSR10 10GbE Network Attached Storage

Author

Paul Davis
Director, Product Management (ret.)
Curtiss-Wright Defense Solutions