WHITE PAPER

Managing life cycle and network interoperability challenges on Navy platforms

Introduction

Navy combat, C4ISR [command, control, communications, computers, intelligence, surveillance, and reconnaissance], and machinery control systems are characterized by a wide range of programs, all of which have their own system architecture, configuration, and composition, mostly program-specific and irrespective of adjacent systems. This disconnected environment creates roadblocks to cost-effective technology refresh and generates institutional inefficiencies, resulting in increased program expense in the long term and less effective warfighting capabilities overall. However, open computing strategies – such as those initiated by social media giants Facebook and LinkedIn – as well as manufacturing practices common in the auto and freight transport industry can be leveraged by the Navy embedded computing industry to solve the interoperability and long-term cost challenges created by these disparate systems.

"*Once a new technology rolls over you, if you're not part of the steamroller, you're part of the road.*"

Stewart Brand

RICK STUDLEY, CHIEF TECHNOLOGIST, MERCURY SYSTEMS
TRUSTED MISSION SOLUTIONS
JULY 2018

mercury systems™

Problem domain/landscape

In our view, the plethora of disparate computing elements within Navy systems expands the number of combinations or permutations by the factorial of the number of those individual unique elements. Worse yet, each system has its own logistics tail and support requirements. Each undergoes expensive certifications to virtually the same extended operational environment – each system independent of the others.

With this approach, the Navy loses the ability to leverage a truly common processing or common display system, not just across the ship or platform, but also across the entire fleet. Economies of scale are not achieved and significant costs are hidden and locked in during initial procurement of the system.

Solution

To that end, we propose a Navy standard Hyper Infrastructure – a tactical system architecture based on a common modular equipment rack; HyperComposable compute, storage, networking, graphics, and special function modules; and HyperScale Leaf and Spine Interconnect Fabric, as the foundation for any number of mission-critical weapons, combat, C4ISR, and machinery control systems. These systems will be composed of a relatively small set (m) of common compute, storage, and I/O modules (as building blocks) for (n) systems of virtually any size and scope – where (m) is potentially orders of magnitude smaller than (n).

For existing legacy network topologies, these HyperComposable systems can be deployed just like any contemporary commercial rack-mount or bladed server system. For new construction and extensive technology insertion cycles, the adoption of a high-performance (100Gb+), low latency (<3µS end-to-end), highly resilient Infiniband or Ethernet fabrics as the core network will further enable the concept of the disaggregated system, and may remain viable for half (or more) of the life of the ship or platform.

Beyond total cost of ownership

Total cost of ownership calculations are often applied to technology platforms, but they seldom factor in the parallel paths of environmental qualifications, software certifications, shipboard industrial work, spares requirements (relative to every other system on the platform), training and support costs. Moreover, calculations often do not consider the difficulty of implementing and performing unforecasted technology refresh and regular periods of technology insertion that may take the ships' systems offline for extended periods of time.

When compared with large-scale commercial enterprises (Amazon Web Services, Microsoft Azure Cloud, Netflix, Google, Facebook, et-al) one can see a radically different strategy.

Our goal is to adapt these commercial-style strategies to define and deliver an enduring and flexible shipboard combat systems compute, I/O, and network architecture based on an open, scalable, and extensible modular designed for reusability, composability, and synchronous technology insertion. The Navy would then be more able to avoid high costs on future systems by abstracting contemporary technologies from shipboard infrastructures, maximizing commonality across the ship's various networks and data centers, while minimizing non-value-add activities and redundancies.

Leveraging commercial market success

For Navy, and indeed DoD-wide, embedded computing systems we propose the adoption of strategies that have had profound and positive effects within three totally different commercial markets – social media, shipping containers, and automotive manufacturing.

Social media: HyperScale Cloud, Open Compute Project®
and Open 19 Foundation

> ***According to Facebook*** "*Loading a user's home page typically requires accessing hundreds of servers, processing tens of thousands of individual pieces of data, and delivering the information selected in less than one second.*" What is astounding is the fact that there are more than 2 billion Facebook members and more than 1.1 billion daily active users, 84.5 percent of whom are outside of North America.

To achieve this quality of service, Facebook, LinkedIn, and virtually every contemporary data center relies on HyperScale Spine and Leaf Interconnect fabrics for low-latency, high multipath bandwidth and extreme resilience to hardware failures. The ability to provide continuous and consistent service level agreements (SLAs) for each customer – irrespective of the millions of other simultaneous users – fuels these organizations' success.

Along these lines, Facebook adopted a data center strategy facilitated by HyperScale fabrics and the Open Compute Project (OCP).

> ***OCP is*** *"a collaborative community focused on redesigning hardware technology to efficiently support the growing demands on [large scale] compute infrastructure."* According to Jay Parikh, VP Infrastructure Engineering at Facebook, between 2011 and 2014, *"Facebook saved more than $1.2 billion by using Open Compute designs to streamline its datacenter and servers… marginal gains, compound dramatically over time..."*

LinkedIn adopted a similar approach for small to medium-sized data centers called Open19 Foundation. While LinkedIn currently operates more than 150,000 servers, their smallest data center instantiation is 16 servers.

As with Open19 Foundation, our strategies are based on a set of key infrastructure visions:

- Unlimited bandwidth
- Zero latency
- Compute on demand
- Disaggregation
- Programmable data center
- Self-healing

While no technology has "unlimited bandwidth" or "zero latency," we want to enable network infrastructure performance where, for all intents and purposes, the applications (and users) effectively experience unlimited bandwidth and zero latency. Once achieved, on-demand compute resources can be made available anywhere. Every server on an ECMP [equal cost multipath routing] nonblocking fabric is essentially "equidistant" from every other server, and any server can potentially host any application. Application modules, functioning as Virtual Machines (VMs) and/or containers, are now free to run anywhere with equal access to the network, cooperating servers, and storage.

The OCP and the Open19 Foundation have demonstrated this capability and enabled great flexibility for more than 250 participating entities, including major banks, internet, cloud, and telecommunications providers.

> ***According to Frank Frankovsky*** former Facebook, VP Hardware Design & Supply Chain: *Bringing the hottest new CPUs into our environment can have a big impact on performance and efficiency. Why not just swap out the CPUs and leave the rest of the server and rack elements in place, rather than rolling in pre-packaged rack-loads of new servers?"*

By preinstalling fully cabled, bare racks to their data centers, LinkedIn technicians can install, provision, and bring online 96 servers in less than 90 minutes. Cloud providers like Equinix provide customers with the option to lease a single-server module as part of a bare metal cloud.

Shipping industry: The intermodal shipping container

Prior to the ISO shipping container's introduction, products were manually handled in inefficient ways. The standardized ISO container solved these issues by abstracting the various products from the entire end-to-end shipping, storage, and transportation infrastructure.

A Navy standard Hyper Infrastructure will define a set of standard, HyperComposable module configurations – micro containers – that provide this benefit in the datacenter in the same way the Intermodal Shipping Container does for transcontinental shipping. HyperComposable micro containers normalize size/volume, mechanical mounting, power, environmental resilience, and cooling. By abstracting physically encapsulated technologies from the infrastructure, standard modules can be easily removed and replaced by newer technology with zero associated shipboard industrial work. Older modules can be repurposed as spares or redeployed on other platforms that may not require or have a budget for the latest technologies.

Auto manufacturing

The current best practice in the auto industry is to develop a set of common extensible platforms that can be tailored to build specific vehicle types by adding modular assemblies: engines, bodies, drive trains, cockpits, etc. By the year 2020, 95 percent of ~33M units manufactured by the top 12 OEMs will be based on an average of three OEM core platforms each. Common platforms in the auto industry clearly have a proven track record of reducing cost and time to market.

Similarly, a Navy-standard Hyper Infrastructure Common Modular Equipment Rack (CMER) will function as a universal platform for any number of applications. The strategy is to deploy CMERs at any number of end points of the platform HyperScale fabric. Applications developed in the lab on a specific combination of HyperComposable Modules can be deployed as the exact same set of HyperComposable modules, VMs, and/or Linux containers. An immutable binary file, in the form of a Linux container, can physically configure the CMSR module interconnect switch system. (See Figure 1: Disaggregation)

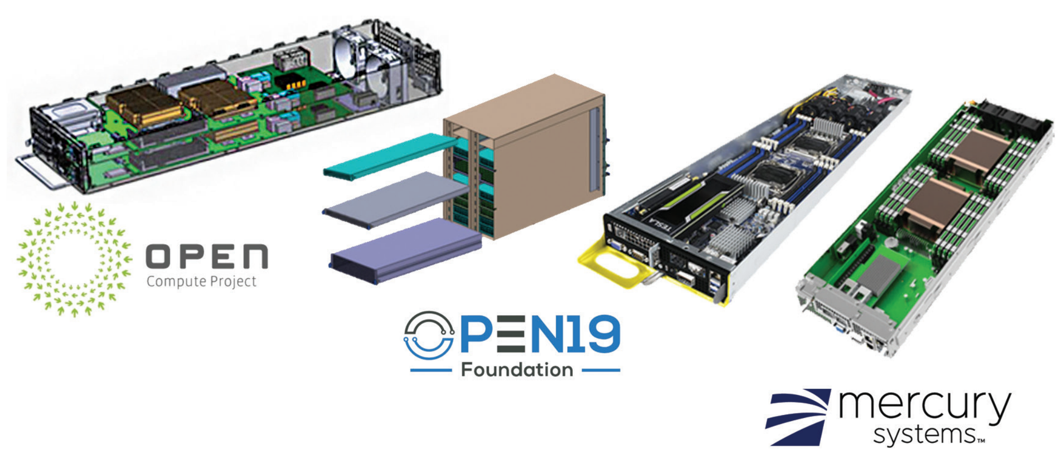


Image Source: The Stack, Adlink Technologies, Open Compute Project, Open19 Foundation

Figure 1. Disaggregation: The trend is disaggregation of storage, servers and networks through open, modular, and software defined everything (SDx)

Defining elements of a common modular infrastructure for the Navy

To apply these strategies to naval systems, first we need to define a common modular infrastructure. We propose a Navy-standard tactical system infrastructure based on:

1. "HyperScale" Leaf and Spine Interconnect Fabric;
2. Common Modular Equipment Rack (CMER);
3. Common Modular Subrack (CMSR);
4. "HyperComposable" compute, storage, networking, graphics, and special function modules; and
5. Software-defined everything; i.e., virtualization for converged and hyperconverged system-level building blocks.

Together, these elements form an enduring foundation for any number of mission-critical weapons, combat, C4ISR, and machinery control systems. These systems will comprise a relatively small set (m) of common compute, storage, and I/O modules (as building blocks) for (n) systems of virtually any size and scope – where (m) is potentially orders of magnitude smaller than (n). See Figure 2: Strategic Hierarchy

For existing or legacy network topologies, HyperComposable systems can be deployed in the same way as any contemporary commercial rackmount or bladed server system. For new construction and extensive technology insertion cycles, the adoption of a high-performance, low latency, highly resilient – versus redundant – Infiniband or Ethernet fabric as the core network will further enable the concept of the disaggregated system, and may remain viable for half (or more) of the life of the ship or platform.

This strategy is firmly aligned with Modular Open Systems Approach (MOSA) guidelines within both business and technical areas, with particular focus on open modular design; technology insertion; extensibility; reusability; and most importantly, composability.

Composability enables configuration of a virtually limitless range of systems from the smallest possible number and type of common compute, I/O, and graphics modules. Any combination of a small set of common modules should be deployable anywhere on a network in virtually any system configuration. (See Figure 3: Common Modular Rack.)

Composable modules will include:

1. CMSRs;
2. Common processing modules (e.g., Intel, AMD, Power, ARM);
3. Common I/O modules;
4. Common display modules;
5. Common attached processor module – GPGPU and FPGA;
6. Common network switch modules;
7. Common storage modules – HDD/SSD (SAS, SATA); U.2 NVME; RSSD NVME; and
8. Common special function modules – e.g., WAN fabric extender, system resource manager.

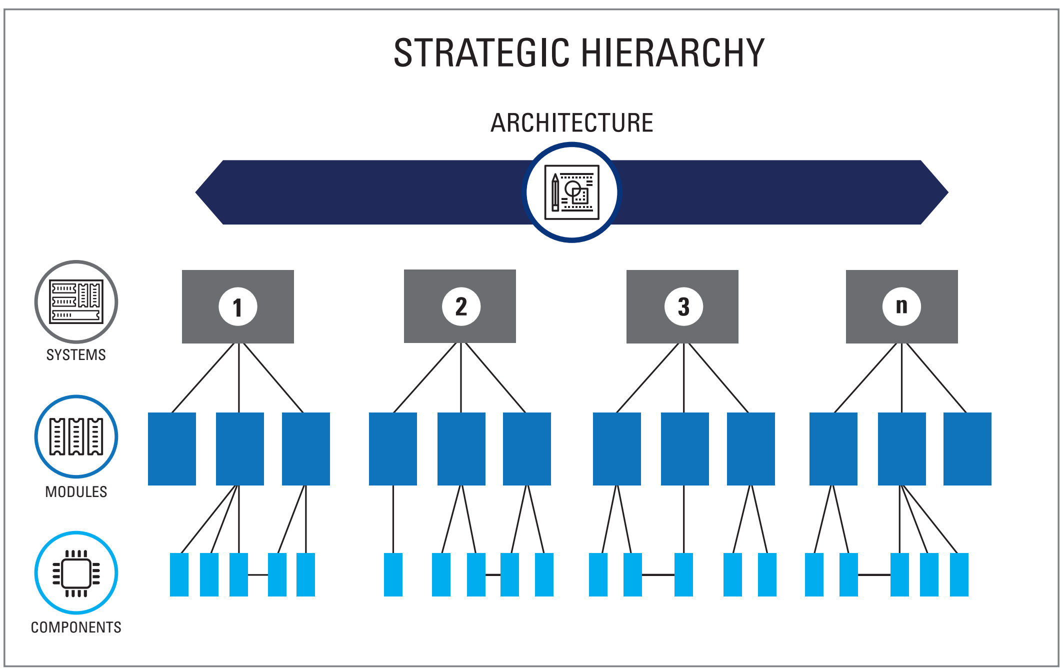


Figure 2. Strategic Hierarchy. Mercury High Density Modular Systems is in alignment with contemporary commercial enterprise data center strategies: Every element is a server, Appliance-free systems, Multi-supplier for every component, Software defined functions where possible

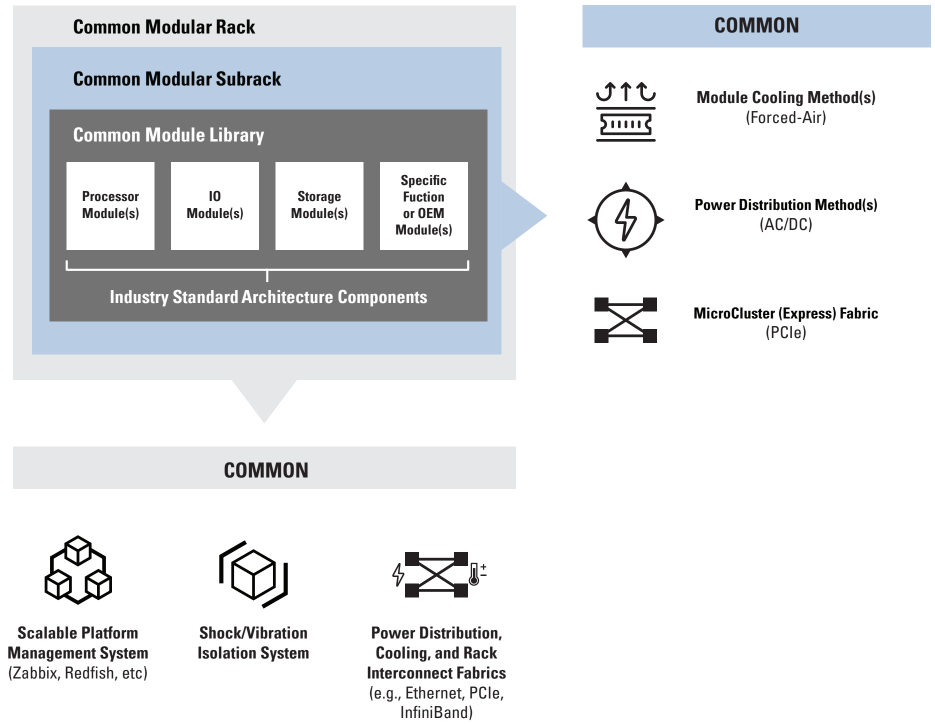


Figure 3: Common Modular Rack. Modular Commonality minimizes the need for redundancy in infrastructure design and component selection, qualification, environmental certification and non value-add shipboard industrial work

On any platform, a shipwide fabric based on InfiniBand and RDMA represents an excellent flexible organic shipboard interconnect "utility." Ship systems integration engineers will have the option to configure protected subnets to build any or all of the aforementioned systems on this utility grid.

An Equal Cost Multi-Path (ECMP) or Open Shortest Path First (OSPF) fabric based on 200 Gb Infiniband would likely last through several generations of technology insertion and will be the right choice for future systems, such as CESARS-SPECS, wherein large, simultaneous, and continuous super high-resolution video and infrared imagery will need to be captured, stored, and processed in real time to cue defensive weapons systems against stealthy, optically guided antiship missiles, unmanned aerial vehicles (UAVs), and even smaller threats like jet skis and rubber rafts

Disaggregation

Once deployed to the platform, a HyperScale fabric enables disaggregation of shipboard computing plants, including, disaggregation of: servers storage networks, and even equipment racks. A disaggregation strategy minimizes, or eliminates physical appliances via extreme modularity, composability, and software-defined everything (SDx). For example, servers are no longer part of a monolithic chassis and storage no longer exists local to the servers as islands of storage or within expensive appliances. Equipment racks together with modular subracks effectively represent "hotel" spaces for disaggregated modules, software-defined networks support virtualization of fewer higher-bandwidth pipes, and servers routinely connect to 100 GbE networks.

Creating an organic shipboard architecture via the Hyper Infrastructure

Navy platforms can leverage these commercial strategies to create an organic shipboard architecture that enables multiple options for hosting mission-critical applications from dedicated bare-metal machines to SDx. We call this the Hyper Infrastructure and its hierarchy is

1. HyperScale
2. HyperCompose
3. HyperConverge

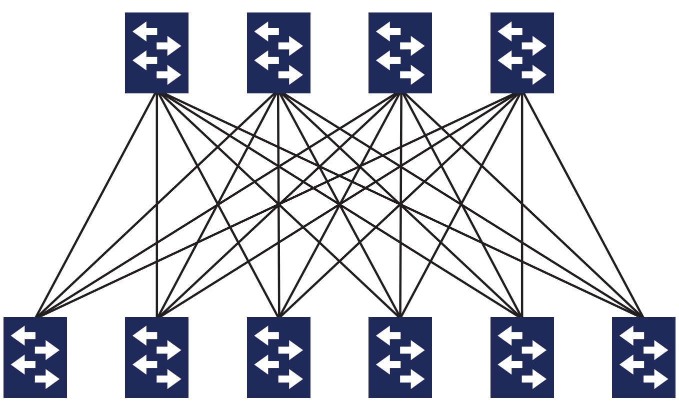
Instead of large naval platforms (ship, aircraft, etc.) hosting a plethora of stovepipe or dedicated systems on one or more traditional networks, the larger platform can comprise an interconnect utility with CMERs pre-installed. This organic platform or "utility" will comprise a high-bandwidth, low-latency, and deterministic shipwide data communications fabric (essentially a large, nonblocking crossbar switch) populated with a number of common processing, storage, I/O, and network function modules. We envision these modules being installed virtually anywhere and everywhere on the fabric, which we call the Hyperscale interconnect fabric.

Hyperscale Interconnect Fabric

HyperScale, Leaf-and-Spine (i.e., ECMP) fabrics (Figure 4) are highly modular designs with a well-defined, regularly repeating topology. Within an ECMP or OSPF fabric, traffic is split across many available paths rather than pushed onto a smaller number of higher speed paths. HyperScale is capability rather than size – it has the advantage of using small open switches in an architecture that scales to any size without changing the building blocks.

Critically, fabric performance is quantifiable in regular mathematical terms - even if you don't know why a particular fabric performs a certain way (theoretically), you can still know how it will perform under specific conditions. The most common case of this capacity is calculating the oversubscription rate on the fabric; e.g., the total amount of traffic the network can switch without contention.

Spine (Fabric Switches)



Leaf Switches

Figure 4. High Performance Enterprise Class Switched Fabric: Features include 1. All End-Points Equidistant from One Another, 2. Low Latency, Zero Jitter, Non-Blocking, 3. Zero Packet Loss, Wire Rate Performance at All Packet Sizes and Port Combinations, 4. Predictable Performance, Fairly Dividing Traffic in All Scenarios, 5. Better Buffering with predictable Buffer Allocation to Any Port & Packet Size

When properly designed, minimum and maximum delays (i.e., jitter) across the fabric can effectively be determined. It then becomes straightforward to determine at what level a fabric is going to introduce buffering (latency) as a result of link contention. From a combat system network design perspective, this is a fabric's crucial defining characteristic.

To paraphrase economist and technology visionary George Gilder, this key attribute enables the disintegration of monolithic machines across the fabric into a set of special-purpose appliances. The resulting appliances (or- modules) can be recombined as building blocks to form the various functions of each combat, C2, or machinery control system.

Leveraging the fabrics to build an organic computing utility grid

Hyperscale fabrics will function as an organic shipboard interconnect utility grid with known and predictable properties. This grid will enable distribution of a common modular infrastructure to equipment rooms across the ship and facilitate easy deployment of modules to bays, within these subracks, without the need to tune the geolocation of hardware or applications. Module specifications will define a common electromechanical configuration, with predefined kinetic and thermal resilience properties. Lastly, the grid supports the composition of (n) systems from a common hardware library of (m) appliance modules (server, storage, I/O).

A shipwide, flexible fabric (hyperscale) infrastructure scales simply by adding paths with equal performance that can scale in any dimension. This arrangement represents a structured, uniform, future-proof topology and a simple path for growth. Power and cooling are optimized as well by enabling mixed loads for more efficient use of spaces (e.g., colocating server, storage, console controllers).

Hypercompositibility and the CMER

Hypercomposability begins with a CMER. Once the interconnect fabric is in place, a set of common equipment racks can be predeployed to any equipment room or space with fabric end points. These racks might comprise two or more identical (unpopulated) common modular subracks and redundant 100 Gb leaf switches.

Such a proposed shipboard data processing utility will include:

1. CMERs providing identical power distribution, cooling (power in and power out), and kinetics mitigation systems (shock/vibration).
2. Within the CMER, two CMSRs can provide a 14U "hotel" space for 20 common processing, I/O, storage, graphics, attached coprocessors, and specialized network-based functions (e.g. WAN fabric extender), and dual interconnect switch planes.
3. Within the CMSR, each module slot or bay will be connected via dual 128 Gbps PCI Express links to two programmable ExpressFabric interconnect planes within the CMSR.
4. Switch planes connect all bays in one or two nonblocking PCI Express "Clos Fabrics" for a guaranteed, available bisection bandwidth in excess of 2 Tb/s. It's expected this architecture will be completely compatible with Gen4 PCI Express and/or future high-speed interconnects.
5. CMSR switch planes facilitate a resilient, nonblocking PCIe-speed (128 Gb/s) IP network between processing modules within the common modular subrack without the need for an external switch.
6. Each CMSR slot will accept any common module in any configuration required for the larger system function – the switch planes are programmable such that each 7U subrack can host (10) high-performance server modules, or a combination of server modules and storage, I/O, or attached coprocessors. Switch fabrics within two subracks may also be directly interconnected for added flexibility.

Any configuration assembled, tested, and certified in the lab can be taken to the ship as a set of common modules and potentially installed by a ship's company. The subrack switch plane fabric manager is programmed via a containerized configuration file and remains locked down unless specifically otherwise authorized. The fabric manager will interrogate each installed module and allow it to be used only if the unique ID of each module matches that which is stored within the immutable software container (binary file) used to configure the fabric connections between each module in the system. In the future, this structure will allow an automatic (at boot time) audit of all installed modules and configuration of those modules no matter what their location within each of the CMSRs.

Module interconnect configurations will be determined at the time of hardware-software integration. The resulting interconnect topology, together with the specific module load out, will be converted to a container (immutable binary file), loaded, and run whenever any or all modules in the system are booted. Whenever a module is removed and replaced or is rebooted, the event will be recorded as a "critical event," and may be prevented under predefined conditions.

Long-term cost savings and performance benefits

The common module library we are proposing is analogous to the NAVSEA Common Source Library. The CML will consist of environmentally robust microcontainers that may be precertified for the full, extended operational environment for all systems simultaneously. The extreme modularity and "composability" of the CMER and CMSR will facilitate multiple technology-insertion cycles simply by replacing modules on the ship. Whatever has been integrated, tested, and certified in the lab will run with equal performance and results on the target platform. See Figure 5: Common Module Library;

The total cost of ownership for these platforms (as mentioned at the beginning of this paper) can be subjective, but the Navy and the Department of Defense as a whole needs to look at ways of reducing long-term life cycle costs while improving performance. And, as we have demonstrated, both the technologies and methods exist so that they can start today.

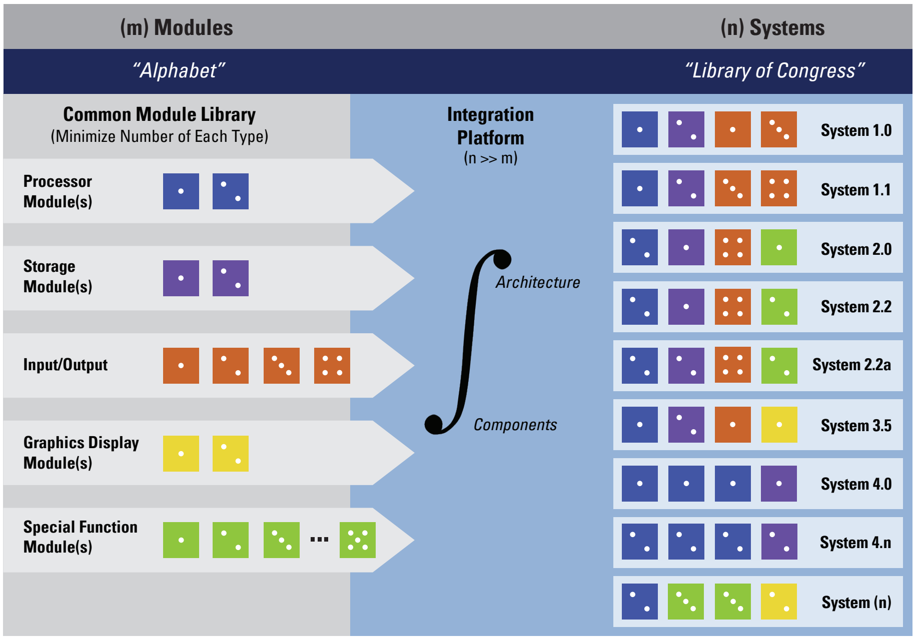


Figure 5: Common Module Library: Analogous to the AEGIS Common Source library, utilizing common modules simplifies logistics, streamlines deployments, and presents an opportunity for module reuse across any number of different programs, not only within the same ship class, but across the entire fleet.

To learn more, visit us at tms.mrcy.com.

3450.00E-0718-wp-Navy-platforms

INNOVATION THAT MATTERS ™

CORPORATE HEADQUARTERS
50 Minuteman Road • Andover, MA 01810 USA
(978) 967-1401 • (866) 627-6951 • Fax (978) 256-3599

TRUSTED MISSION SOLUTIONS
47200 Bayside Parkway Fremont • CA 94538
(510) 252-0870 • Fax: (510) 490-5529