Chapter 8
Application Drivers and System Requirements

8.1. Introduction
Architects and end-users alike have enjoyed increased system performance and increased energy efficiency over the past many decades as a result of transistor scaling according to Moore’s law. Additionally, as application domains continue to expand beyond traditional computing and towards more bio-inspired and bio-aware solutions, systems and the elements that comprise them must better enable these use cases. As the industry transitions to a new era of computing systems and applications, we must advance computing systems with a full-stack approach. We consider the impacts of a wide variety of application domains to drive the impact of future applications, as well as to influence the direction of key enabling technologies covered by the MAPT roadmap: data center and HPC, mobile communications and infrastructure, edge and IoT, automotive, bio-applications and health, security and privacy, and defense and harsh environments.

Each of these application areas will evolve in distinct ways and will require domain-specific systems to achieve next-level performance. It is critical that we understand these workloads, the systems that will be built to run them, and their implications. Systems and applications can change in either an evolutionary or a revolutionary manner:

- Evolutionary: Transistor performance, interconnect performance, communications performance, etc.
- Revolutionary: New domain-specific architectures, compilers, frameworks, etc. that innovate atop new building blocks like 3D stacking, analog accelerators, compute in- or near-memory, etc.

Each application class will likely evolve in a different manner, as they possess diverse requirements and follow different use cases. We detail the important technical areas for each class of applications and systems below.

8.2. Data Center and High-Performance Computing (HPC)
Computing technologies are main drivers for many of the world’s technological advances: In many cases, these application areas require broad advances across the entire compute stack, including compute engines and processors, high-bandwidth and efficient memories, high-speed and efficient network infrastructure, and high-density storage technologies paired with a software and compiler ecosystem that coordinates parallelization across individual resources.
Future HPC use cases will require advances in scalability, RAS, application latency, and application throughput. Focus areas for this class of workloads include:

1. **Data analytics**, both structured and unstructured
2. **Biological, genome-based, and other similar analysis**, including an understanding how these function in humans and beyond
3. **Simulation techniques**: Molecular dynamics, computational fluid dynamics, aerodynamics, and structural analysis
4. **Large-scale inference and training**, including natural-language processing, image recognition, recommendation systems, etc.
5. **Applications enabled by quantum computing**

Improvements in these areas will influence and be influenced by technology enabling aspects across a wide variety of vectors:

1. Advanced CMOS technology: increased energy efficiency, increased transistor performance, and additional transistor scaling, in many cases driven by heterogeneous integration, MEMS, and analog circuitry.
2. More efficient chip-to-chip interfaces, including optical interconnects and heterogeneous integration.
3. Improvements in memory systems and memory devices through heterogenous integration and advanced packaging, increased memory density, and proximity of memory to compute through techniques like near-memory processing or processing-in-memory while maintaining system-level RAS guarantees.
4. Advanced cooling solutions, including liquid and immersion cooling, better thermal interface materials, etc.

### 8.3. Automotive

Between 2022 and 2035, automobiles will be transformed in almost every way, to further improve efficiency, safety, driver experience and sustainability, as we drive toward a world in which vehicles are **perfectly safe with zero emissions**. Many of these changes will impact other transportation modes including commercial and industrial vehicles, motorized scooters, and aviation. In short, vehicles will become more Autonomous, Sustainable, Networked and Connected, driving special requirements for semiconductor devices and packaging. Figure 1 summarizes some of the upcoming trends seen by many in the industry.
Table 8.1 summarizes the manner in which these automotive trends drive the microelectronic technology and package requirements. Although other applications also drive the need for high compute performance, cost-effective heterogeneous integration and integrated high voltage capabilities, these requirements will be pushed in unique ways as vehicles evolve over the next 10-20 years.

Table 8.1: Automotive Trends Driving Unique Requirements
Autonomous Vehicles will contain enormous processing power in the hundreds of TFlop range for sensor fusion, specialized domain specific processing of sensor data to model the environment around the car and assist or even direct all aspects of vehicle control with low latency. New Domain-based and Zonal architectures will require intra-vehicle networks which can manage data traffic approaching 100 Gbit/sec, and together with wireless connection to the infrastructure and internet, will enable key new capabilities such as SW-defined vehicles, immersive infotainment and new ownership models. Sustainability will drive adoption of carbon neutral propulsion including battery and fuel-cell EVs, also requiring significant processing power. Overall, these capabilities will require High Performance Compute power and performance to be supported inside the car, while meeting specific auto constraints including wide temperature range (-40C to 150C Tj), quality and reliability targets, functional safety and security, and harsh thermal management limits. See figure 2 for some quantitative performance predictions.

*The Grand Challenge is to enable the vehicle of 2035 to become a data center on wheels.*

![Path Towards Full Autonomy](image)

*Figure 8.2  Extended Performance Requirements [2]*

Similarly, increasing Autonomy and Driver Assistance will continue to require integration of special functions such as large RAM and/or NVM, mmWave and other RF communication and ranging technologies. Networked vehicles will drive the need for high speed PHYs and related hardware.
EVs and other sustainable transport will put a strong emphasis on precision analog for battery management, NVM, sensors etc. Infotainment and other Connected Vehicle functions will drive the total memory space (both DRAM and Flash) requirements.

All of these vehicle advances will demand increased integration of heterogeneous functions including NVM, sensors, analog/RF, and processing power. Critically, low latency is a requirement for the real-time functions across these automotive trends, requiring close proximity and low parasitics, high data speeds and reduced power consumption/dissipation. While continued monolithic integration will offer part of the answer, advance multi-die packaging will play a major role. A crucial point is that these multi-die solutions need to scale to meet the power and cost requirements for automotive.

*The Grand Challenge in this regime is to integrate all required functions at the package level, with performance, time-to-market and cost better than today’s monolithic integrations within acceptable power constraints.*

Finally, all the trends mentioned here will drive higher voltage analog requirements, from increased IO voltages on advanced logic, to 3.3V/5V analog for sensor data processing, to the 20V/48V and up to 800V handling needed for EV drivetrains and battery management.

Automotive applications will also lead to unique requirements for some of the MAPT roadmap’s cross-cut groups. Specifically, Security & Privacy will be key issues to be addressed in future automotive electronics, and improved Sustainability & Energy Efficiency will be among the more important outcomes of next-gen automotive development.

Automotive security is critical for widespread adoption of autonomous and connected driving. Security ranging from trusted supply chain, secure hardware, malware, and virus detection will be critical as well as prevention attacks on cloud services and infrastructure as well as robustness in the presence of radar and image jamming and other targeted attacks that can disrupt traffic and threaten national security by causing widespread gridlock. A comprehensive approach is needed to address these security issues as well as to prevent vehicle theft and ensure personal safety and national security.

The sustainability of automobiles includes manufacture, operation, and recycling during the vehicle lifecycle. Automobile manufacturing requires assembly of a multitude of materials/components including metals, plastics, glass, chemicals, electronics, batteries, etc. all of which require varying amounts of carbon to produce. The operation of automobiles produces a carbon footprint which depends upon the vehicle size, weight, performance, energy source (gas, diesel, hydrogen, electric) and vehicle utilization which may be revolutionized by autonomous vehicles. Finally, the recycling of the automobile components is a key part of the vehicle ecosystem which can be optimized by design choices.

In summary, automotive applications for future cars and trucks will continue to leverage developments in other semiconductor markets, such as high-performance computing, but will drive unique requirements which must be met while taking maximum advantage of the performance, power, cost, security, and efficiency gains. These advances will also benefit other vehicle types (scooters, e-bikes) and transportation modes (long haul trucks, trains, aviation, aerospace and maritime).
8.4. IoT and Edge

The next several decades will see electronics integrated into nearly every location where there are humans and industry, and even in natural locations that we want to monitor to protect from humans and industry. In short, electronics will be truly ubiquitous. We cannot predict every application for electronics across every industry, but we can understand the broad expectation that electronics enables enhanced observations of environments and materials, automation of processes, automation of decision making, and an overall amplification of human effort. In the electronics manufacturing industry, our focus here is on IoT (internet of things) and Edge computing.

IoT and Edge computing occupy emerging market spaces adjacent to many established markets such as industrial automation and personal electronics. IoT and Edge computing is a complex system of technological and other components closely coupled to user behaviors. This area is typified by markets and applications that sit at the edge of what is possible and economical in close integration of unconventional semiconductor components with those that are currently conventional. As time goes on, that leading edge of unconventional integration will expand, leaving behind new markets and new conventions for semiconductors.

Economically, the impact of these technologies is referred to as the fourth industrial revolution. Such a grand statement clearly indicates the expected commercial impact that the combination of sensors, automation, digitization, and intelligence will have. The Smart Industry Readiness Index (SIRI) is a framework for understanding the process through which industry moves from a completely manual process to one which takes full advantage of automation and intelligence enabled by electronics. Figure 8.3 shows how facility intelligence is evaluated in this framework.
From a hardware and packaging point of view, the devices necessary to broadly enable this industrial automation require great flexibility in sensor fusion, including optics, sound, magnetic fields, temperature, pressure, and chemical detection. These are the practical problems for electronics. The challenge over the next 15 years is to bring the cost of deployment and use down through chiplet interface standardization with aggressive modernization and use of standards. Communication and interface standards are particularly important, which is why the Open Platform Communications Foundation recently published an open standard for communication between sensors and cloud applications. Broad implementations of standards such as these will be critical to achieving the economic success expected from Industry 4.0.

Home automation, as another example, consists of devices such as smart thermostats, network connected lightbulbs, and kitchen appliances. In the future power balancing and conservation will be controlled by a home “PMU” equivalent and network security and privacy in the home will be centralized via a home “NSP”. Energy efficiency and integrated cyber security are two defining characteristics of IoT and Edge development with significant improvements expected within 5 years. This prioritizes standardization and integration of low-power local networking, sensing, and computing.

Considering agricultural automation brings in the need for increased energy harvesting and packaging to support sampling of liquids and biological materials for chemistry, biochemistry, and genetics. In 15 years, new Smart Farm devices will track pests and will be used to track and validate the genetics and health of agricultural products at the farm and through the logistics chain. These systems will need to be low cost and high volume. This application will serve as the starting point for commercialization of standardized microfluidic integration at the packaging level that will find later use in the Synbio application areas.

Infrastructure monitoring is one of the most exciting areas of IoT and Edge development, in large part driven by advances during the Covid pandemic. In 15 years, monitoring wastewater for disease or unusual contamination will become common place of local government infrastructure. This requires
packaging and sensor integration as in agriculture, standards development as in industrial automation will, and will require applying evolving best practices for cyber security.

Table 8.2

<table>
<thead>
<tr>
<th>Low Power Operation, Standby, and Energy Harvesting</th>
<th>Home Automation</th>
<th>Agriculture</th>
<th>Industrial Automation</th>
<th>Infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td>High power. Access to power infrastructure. 3 Volt.</td>
<td>Low power. Battery operation and drone integration. Allow for multiple energy inputs. 1 Volt.</td>
<td>High power. Access to power infrastructure. 3 Volt.</td>
<td>Mixture of power sources and levels. Need for both low power and high power components in the same network.</td>
<td></td>
</tr>
<tr>
<td>&lt; 1W standby power 10 W operation range</td>
<td>&lt; 10 uW standby power 1 mW low end operation range 100 mW high end operation range</td>
<td>&lt; 1W standby power 10 W operation range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Packaging and Environmental Integration</td>
<td>Integration for consumer applications. PMU and HSP for the home.</td>
<td>Integration and packaging for continuous exposure of liquids or biological samples to sensor in integrated package.</td>
<td>Integration and packaging for wide variety of sensors including optics, electromagnetics, temperature, pressure, and chemicals.</td>
<td>Integration and packaging for wide variety of sensors including optics, electromagnetics, temperature, pressure, and chemicals.</td>
</tr>
<tr>
<td>Cost, volume, and timing</td>
<td>Potential &quot;premium market&quot; for IoT and Edge systems. 2023-2027</td>
<td>Extremely low cost, high volume systems required in purchase, deployment, and use. 2028-2031</td>
<td>Low cost, high volume systems required in purchase, deployment, and use. 2028-2031</td>
<td>Extremely low cost, high volume systems required in purchase, deployment, and use. 2032-2036</td>
</tr>
</tbody>
</table>

8.5. Bioapplications and Health

Synthetic biology (or ‘synbio’) harnesses the power of biological systems for useful purposes. Semiconductor-based applications of semiconductor synthetic biology1 (‘SemiSynBio’) are rapidly growing but have unique packaging considerations and constraints. These applications generally require the semiconductor device and packaging to directly contact or contain a liquid interface. They may also be required to operate in a narrow range of biologically-relevant temperatures, at ultra-low power.

---

1 https://www.src.org/library/publication/p095387/p095387.pdf
levels or to interface with biological processes that take place at much longer time scales than other semiconductor applications.

SemiSynBio devices make excellent biosensors, allowing for sampling and detecting specific analytes against complex backgrounds. Sensing applications include environmental monitoring (e.g. detecting toxins in public water systems or viruses in ambient air samples) as well as medical applications for diagnostics (e.g. detecting pathogens or evidence of cancer from blood samples). Such sensing devices can range in complexity from simple, chemically functionalized silicon surfaces to entire communities of cells (a ‘biological front-end’) living atop of, and engaging in two-way communication with, a CMOS device (a ‘silicon back-end’).

![Fig. 8.4 Illustration of communities of cells communicating with CMOS devices](image)

Semiconductor devices have also recently been used to manufacture DNA for binary data storage. DNA is an ideal storage medium given its chemical stability and its nearly universal use in nature for reliable, long-term information storage. This universality is an additional strength: there will always be a large economic incentive to build capability to read DNA, ensuring that binary data stored in DNA will always be directly readable. Synthesis of DNA for binary data storage is an application requiring densely packed semiconductors, advanced electrochemistry, microfluidics, and enzymatics that could increase the density for binary data storage by orders of magnitude over existing data storage technologies.

SemiSynBio devices can also be used in personalized medicine both to diagnose patient conditions (in a point-of-care setting) as well as to manufacture treatments in real-time at individual-patient scales. Current generation micro-bioelectronic devices can be ingested and pass through the digestive tract, collecting data via onboard sensors such as Medtronic’s PillCam. Future generation devices may be small enough to travel through, or hold position within, blood vessels. These devices will monitor blood over time and transmit results outside of the body for collection and analysis. They may even use the turbulent passage of blood itself to generate power for ongoing operation. Such devices could also be used to locally manufacture therapeutic compounds at low doses in real time, responding to and addressing biological risks to the host including infection or cancer.

SemiSynBio devices also have applications in building brain/computer interfaces (BCIs). BCIs are an exquisitely complex marriage of semiconductors in direct contact with neurons and other nervous-system tissues. Such devices will allow for two-way communication between the brain and traditional
computing systems. Ensuring these devices have a low immune profile while embedding deep enough into the brain to communicate with all requisite brain structures remains a significant challenge.

Lastly, SemiSynBio-based devices can harness thermodynamics directly in molecular computation-based applications to efficiently solve otherwise intractable (or simply expensive) problems. These applications may make use of CMOS devices to carry out synthesis of the components that will carry out the computation (e.g. DNA or other polymers) or to read back the results of the computation from the fluid interface, or both.

Advancement in these applications over the next 5 years will require:

- Improved control over electrochemical and electro-optical interfaces with semiconductor devices and packaging strategies
- Integrating microfluidics for reagent distribution and waste or product removal into semiconductor packaging up to and including distribution of such hybrid devices into traditional data centers.
- Synthesis of ~1 billion unique DNA sequences on a single device in a single run
- Improving the ability of semiconductor devices to operate at lower temperatures required by many enzymatic processes
- High performance computing approaches to design- and model molecular computational approaches

On a longer time horizon of 10-15 years, to enable the broad promise of SemiSynBio technologies, advancements must:

- Advance microfluidics capabilities to enable not just reagent delivery and waste removal but direct, on-device ‘downstream processing’ to isolate and purify output compounds
- Lower the immune profile of SemiSynBio devices to allow their direct integration into- and long-term contact with living tissues and cell communities
- Enable creation- and maintenance of cell populations in artificial ecosystems for years in stand-alone SemiSynBio devices, using these populations to manufacture materials, directly communicate with other biological systems, or to sense changes in their environment.

Given the challenges facing binary data storage and the need for improved areal density called out in the SRC Decadal Plan, storing and recovering 1 EB of binary data in DNA in a standard data center rack should be considered a grand challenge for the advancement of SemiSynBio devices.

8.6. Mobile, Communications, and Infrastructure

5G-Advanced starts the second phase of the 5G decade, bringing a new wave of wireless technology innovations. It is envisioned to push technology boundaries in two broad directions:

1. Strengthen the 5G system foundation through advanced downlink/uplink MIMO, enhanced mobility, mobile integrated access/backhaul (IAB) and smart repeaters, evolved duplexing, AI/ML data-driven designs, green networks, and more.
2. Proliferate 5G to virtually all devices, deployments, and use cases including boundless extended reality (XR), NR-Light (RedCap) evolutions, expanded sidelink, expanded positioning, drones and enhanced satellites and multicast enhancements.
Additional enhancements are anticipated in dynamic spectrum sharing (DSS), multi-SIM, in-device coexistence, small data transmission, quality of experience, carrier aggregation, self-organizing network/minimization of drive test (SON/MDT), and more.

Massive multiple in multiple out (mMIMO) systems are highly desired for dense, urban areas, with the largest commercial focus currently on TDD applications between 2.5 and 4.2 GHz. This is known in general terms as sub-6 GHz and has been the biggest area of cellular infrastructure growth in the last 5 years.

Looking forward there is the need for significant amount of emphasis in the follow areas:

- **Frequency allocations for increasing instantaneous bandwidth (IBW):** One of the biggest drivers for millimeter-wave (mmW) research, typically in the 28-49 GHz range, has been enough contiguous bandwidth to offer the opportunity at increasing IBW to 800 MHz this is a primary factor to improve data-rates, and lower latency. Sub-6 GHz systems are typically 200 MHz of IBW, moving to 400 MHz. The major issues for mmW systems are propagation loss and system efficiency (energy consumption).

- **Continued system efficiency improvement and power consumption:** Doherty power amplifiers are the workhorse architecture for both mMIMO, as well as more traditional high power, Macro base-stations (usually 4 – 8T), and for very good reasons—they are not overly complex, and they can work up to ~ 40% relative bandwidth with high efficiency in the 6 – 9 dB average power backoff from peak power. Note that most FDD and TDD systems operate with signal peak-average ratio in that same 6 – 9 dB range. Digital Predistortion (DPD) is the main linearization technique. Research areas are to continue to push wave-shaping to higher efficiency amplifier classes such as class F, inverse class F, class J, etc... and still allow linearity correction. Reducing the number of nonlinear coefficients in the DPD ASICs also allow improved overall system efficiency.

In general, the current and future mobile communication infrastructures are based on what is called a Service-Based Architecture (SBA), which implements IT network principles and a cloud-native design approach within the radio access networks (RAN) and core network (CN). In particular, the designs eliminate the centralized infrastructures in favor of cloud native distributed infrastructure. For example, 5G Core (5GC) implements the new 3GPP network architecture that will unleash the full power of 5G standalone (i.e., it is not dependent on any other network architecture) enabling faster connectivity speeds, ultra-low latency, and higher bit rates with high grade of network reliability. These capabilities, combined with network automation, network slicing and edge computing, are instrumental to address multiple verticals and enable an ecosystem for innovation with use cases such as: enhanced mobile broadband, ultra-reliable low latency communication, massive machine type communication, and time-critical communication. Figure 8.5 below is providing the end-to-end latency in mobile communication infrastructures divided into sub-domains which is given by:

\[ L_{\text{end-to-end}} = \text{Latency experienced at the RAN} \left( \frac{1}{2} L_{\text{radio}} + L_{\text{Transport Network}} \right) + \text{Latency experienced at the Core Network} \left( L_{\text{CN}} + \frac{1}{2} L_{\text{User Plane Functions-to-Application Server}} \right) + L_{\text{Application Server}} + L_{\text{Peering Point}}. \]
In current 5G deployment, there are several reports of uplink and downlink radio latencies varying from 1 ms to 5 ms depending on the RAN configuration. The latency introduced by the transport and core networks is usually modeled using fixed values. For example, some reports considered that the one-way core network latency is 200 μs or 100 μs for non-standalone and standalone networks. Also, latency introduced in the link between the core network and the location of the Application Server (AS) was introduced in the order of 5.4 μs. Finally, the core network latency is estimated as a single value of 20 ms considering the processing delay introduced by the core network nodes with neglecting the propagation or queuing delays. The final end-to-end latency can reach to 201 ms.

In terms of power consumption, the core network workloads are optimally placed on highly distributed compute platforms to deliver optimal performance as per use case requirements. This distributed cloud platforms are interlinked, and workloads are orchestrated centrally to achieve elasticity based on availability. These distributed mobile communication cloud platforms could be used for hosting web-scale player application platforms or web-scale players compute platforms can be used to host core network nodes.

The mobile communication infrastructures are expected to support critical communications with ultra-reliable and low latency features. In particular, extremely high reliability (e.g., 99.9999% control of process automation) is expected with an end-user experience of data rate up to 100 Mbps and an end-to-end latency of 50 ms.

In Table 8.3 below, the hardware layer of the core network, radio access network, and user equipment are considered. For the current values in the core network, the calculations are provided based on several
available chips and the forecast is driven by low latency, high throughput, and deterministic performance. In general, the core hardware is usually racked in traditional data centers, with associated cooling and power infrastructure available. However, this should evolve to meet the future requirements of compute and power efficiency imposed by the telecom standards body (e.g., 3GPP). For instance, as an example, in the current 5G core network, Arm based Graviton2, 64-core processor is considered (~1W per core @2.5GHz) based on Arm’s Neoverse N1 design, 4-8 physical CPU cores per core network node, or 3rd Generation Intel Xeon Scalable 6338N processor 8–40 CPU cores, with network processors 20–36 cores, or AMD’s EPYC 7571 processor. For the RAN, the current values are considering Arm-based v8 cores at 3.3 GHz, or 4th Gen Intel Xeon Scalable processor, or AMD EPYC 7003 Series Processors.

Table 8.3: A system-level roadmap for the mobile communications infrastructure

<table>
<thead>
<tr>
<th>Attribute/Parameter</th>
<th>Unit</th>
<th>Current</th>
<th>5-Years</th>
<th>10-15 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Network (CN)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The hardware/infrastructure layer includes server blade and sled, Network adapter, Internal controller, and storages.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latency</td>
<td>ms</td>
<td>~2</td>
<td>~1</td>
<td>~0.5</td>
</tr>
<tr>
<td>Power Consumption (CPU*)</td>
<td>Watt per socket</td>
<td>180</td>
<td>111</td>
<td>70</td>
</tr>
<tr>
<td>Memory bandwidth</td>
<td>Mega Transfers per Second</td>
<td>3200</td>
<td>5120</td>
<td>8192</td>
</tr>
<tr>
<td>Memory capacity</td>
<td>GB DDR4</td>
<td>256</td>
<td>512</td>
<td>1024</td>
</tr>
<tr>
<td>Radio Access Network (RAN)</td>
<td>The hardware/infrastructure layer includes server blade and sled, network switches, and storages. This layer may have sub-processors such as SoCs, GPUs, and FPGAs to enhance complex computations required in the RAN to support near real time intelligent controller.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latency</td>
<td>ms</td>
<td>~5</td>
<td>~1</td>
<td>~0.5</td>
</tr>
<tr>
<td>Power Consumption (CPU*)</td>
<td>Watt per socket</td>
<td>140</td>
<td>111</td>
<td>70</td>
</tr>
<tr>
<td>Memory bandwidth</td>
<td>Mega Transfers per Second</td>
<td>4800</td>
<td>7750</td>
<td>12500</td>
</tr>
<tr>
<td>Memory capacity</td>
<td>GB DDR5</td>
<td>512</td>
<td>1024</td>
<td>2048</td>
</tr>
</tbody>
</table>

*At idle and around double under full load

Green – Process developed and ready for manufacturing,
Yellow - Additional development work needed, and
Red – Major development efforted needed for HVM.
Table 8.4: A system-level modeling that includes some aspects of link-level modeling.

<table>
<thead>
<tr>
<th>Modeling Attribute</th>
<th>State of the Art</th>
<th>Future Need</th>
</tr>
</thead>
<tbody>
<tr>
<td>specification of the network building blocks</td>
<td>Interconnect Building Block (consists of switch, arbitrator, and buffer blocks for creating standard pipelined routers) with set of devices that are characterized using the basic element device model.</td>
<td>extend with novel photonic building blocks deriving detailed values for delay and energy dissipation. The model shall include parameters for clock rate, buffer size, channel width, and number of virtual channels. Also, shall include additional methods for interfacing with photonic devices.</td>
</tr>
<tr>
<td>specification of the target application</td>
<td>specifying the target application that supports the use of both synthetically generated traffic patterns (e.g., random, hotspot, nearest neighbor, and tornado) and communication traces (monitoring the network traffic during the execution of a real application).</td>
<td>integration with a cycle-accurate microarchitecture modeling.</td>
</tr>
<tr>
<td>modeling of the network architecture</td>
<td>The devices from the Interconnect Building Block Library can be combined to create higher-order networking components and entire interconnection network topologies. Nonblocking switch can be derived by connecting various devices from the Building Block.</td>
<td>Network architect shall optimize the topology design to target specific requirements such as packet size, latency, and/or throughput and account for the target applications.</td>
</tr>
<tr>
<td>physical-layer characterization</td>
<td>involves the characterization of the network architecture at the physical layer including metrics such as power budget, crosstalk, and power dissipation.</td>
<td>Shall be determined from the aggregate performance of the individual photonic devices. Also, enables an accurate first-order physical characterization of an entire network through the</td>
</tr>
</tbody>
</table>
characterization of a small number of foundational components.

<table>
<thead>
<tr>
<th>System-level performance analysis</th>
<th>measures the system-level performance characteristics of the network architecture in terms of data throughput and latency.</th>
<th>Includes physical layer properties to measure the impact on network functionality and scalability to determine the overall system performance.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iterative refinement of parameters and design</td>
<td>performance results and analysis of the modeled network are used iteratively to refine the topology design and device parameters to optimize the overall performance.</td>
<td>Integration of ML prediction methods (e.g., LSTM, SVM, KNN) to the physical-layer characterizations. This shall impact subsequent analysis of a system with improved crossings.</td>
</tr>
</tbody>
</table>

Green – Methods known,  
Yellow – Methods need to be improved,  
Red – Methods need to be developed.

8.7. Defense and Harsh Environments

While defense applications have driven unique requirements for a low volume and harsh environment microelectronics, commercial sectors that operate in harsh environments should see a growth in symbiotic relationships between the defense and commercial sectors. There are three primary application environments in which the sustainability and robustness of electronics are intrinsically required: land/terrestrial, sea/air, and space. Each of these environments inherently dictate requirements for full operation (temperature, radiation, shock/vibe, security, and lifetime). Either for defense or commercial applications, there are a few key technical attributes listed below that must be assessed based on the intended operational environment.

**Thermal management:** Higher power operation ultimately drives junction temperature beyond operational functionality and impacting reliability. Advanced thermal management techniques (active and passive) within a 3D package need to be developed that minimize operational complexity and power consumption, particularly as advanced packaging can collocate more components in smaller volumes. It is expected that advanced packaging will minimize the ability to utilize current techniques such as passive heat sinks from surrounding architecture.

**Lifetime & reliability:** Application environments are one main driver for rigorous lifetime requirements, such as space. Without the ability to repair or upgrade, in addition to high cost for insertion/deployment, decade+ lifetimes paired with high reliability are necessary for systems such as satellites. For defense applications, long lifecycles for qualification of hardware for deployment ultimately drives longer lifetime requirements for investments. Across all sectors, high consequence systems (e.g., automotive, aerospace) drive a necessity for high reliability. Schemas for heterogeneous
integration will need to push towards high reliability and decade+ lifetimes for insertion for high consequence and harsh environment applications.

**Verification & Validation (Security):** Defense applications have a high threshold for ensuring secure operation and performance. Additionally, as COTS components are utilized across many applications, coupled with complexities for supply chain risk and subversion, advanced packaging will push the level of complication for verification and validation of sub-components. The industry will need to significantly advance robust processes, metrology, and testing techniques within a 3D package, not only for verification and validation aspects, but to ensure performance, lifetime, and reliability metrics.

**8.9. State of the Art / Product Examples**

*This section still under development*

**8.10. Limitations with Current Technology**

- Future architectures across all applications and systems domains will become power constrained in the coming years, which will limit the application-level advances possible.
- Intercommunication bandwidth limits the scalability of applications and systems. Memory systems, for example, are limited in their host-consumable bandwidth given the packaging options available to systems architects.
- Cross-stack technologists in high demand; however, very few engineers possess the ability to reason across the entire technical stack.
- Lack of infrastructure to prototype application and workload tradeoffs on core, foundational technologies similar to those explored in the MAPT roadmap.

**8.11. Challenges, Future Needs and Possible solutions**

*This section still under development*