A New Ecosystem for A New Era In Compute

Introduction - The Changing Ecosystem and its Importance

We live in a distributed world where APIs and microservices rule the datacenter, and the datacenter is no longer a physical construct; rather, it's a combination of clouds, edges and remote computing centers that drive the modern business. Cloud computing and edge computing require applications to be more mobile, resilient, manageable and secure, and this need, in turn, led to the emergence of distributed computing. Applications are no longer monolithic; rather, they are comprised of microservices. These applications are deployed across infrastructure - a range of servers and other compute platforms in multiple locations - to achieve those design goals of resiliency, mobility, manageability and security.

The software ecosystem – from the operating system to the toolsets to workloads and applications − has evolved in response to new compute paradigms and deployment models. This impacts the infrastructure choices to support this new, distributed environment. Just as open source software rules in the distributed computing environment, underlying platform architectures are being evaluated to meet the unique needs of these new workloads.

In this paper, we dive deeper into the forces and business drivers impacting change and how companies like Arm are responding through IP and a rich ecosystem of partners.

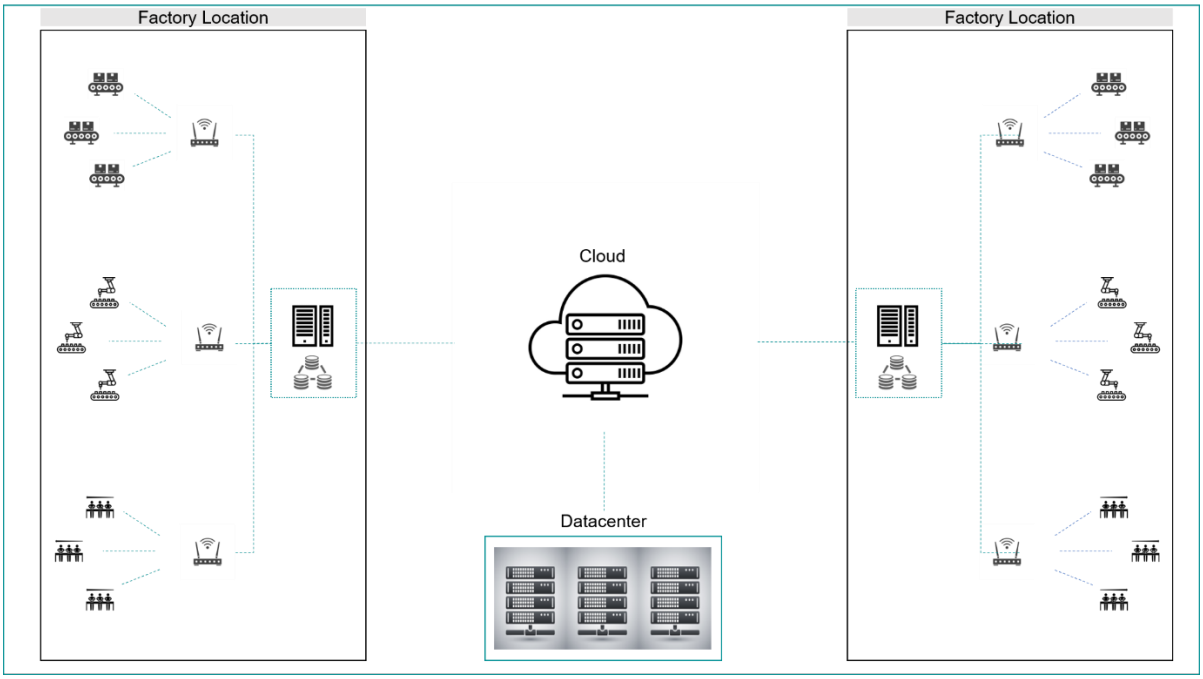
Distributed Architectures Crave Flexibility

As previously mentioned, the new datacenter is comprised of (hybrid) cloud, edge and IoT devices generating data every moment across multiple locations. Mobility rules the world. To fully take advantage of the power of cloud and edge environments and easier management, developers now architect "cloud-native" applications. Cloud-native applications are comprised of microservices that deliver unique functions. Microservices are wrapped in containers and may reside on the same server or can be distributed across multiple servers to reduce resource contention or leverage local resources, such as large memory footprints or (local) accelerators. Use cases drive application development and consider user interaction, data usage requirements and the contextual needs of the larger organization, such as interfacing to other applications and use and

integration of data. For example, an application architected to manage a large-scale manufacturing operation has to consider the collection, aggregation, cleansing and analysis of data on a specific assembly line or factory location. That data must also be hauled to a separate aggregation point where it can be merged with data from other factories to manage the optimal cost-production output.

Consider the above usage scenario. Multiple sensors are required on machinery to track up/down status as well as performance data. Each assembly line on a factory floor may require compute, networking and storage to collect machine data and perform the necessary aggregation, transformation and analysis. A micro-datacenter is used to aggregate data at the local factory, and that data is hauled to a more central location for the sake of optimizing production across multiple factory locations. Finally, deeper data analysis, warehousing and archiving are performed in the cloud or core datacenter.

FIGURE 1 – EDGE TO CLOUD TO DATACENTER



Source: Moor Insights & Strategy

The unique and shifting needs of the modern datacenter require software architects to reconsider the tools used to develop the applications that drive the business. The software ecosystem that can intelligently connect a range of IoT devices to the edge to the cloud and core datacenter is rooted in open source. For the sake of performance

and simplicity, this range of devices – from sensors to gateways to networking to servers – would share a common architecture.

Datacenter architects also feel the tension of the cloud and edge impact. Cloud-native applications crave a server architecture that can be tailored based on workload and function, and this server, in turn, requires an underlying CPU architecture that allows for design customizations for right-sized compute, tight integration and comprehensive security – from far edge to the core datacenter.

IT consumers require choice in deploying the right hardware platform for the right application in the right usage scenario. For example, Cloud Service Providers (CSPs) require choice to deploy the right infrastructure for the right service – whether those services are Infrastructure as a Service (IaaS), Platform as a Service (PaaS) or Functions as a Service (FaaS). Telecommunications companies need choice in deploying 5G infrastructure that delivers the best performance in the lowest power envelope. CIOs looking to bridge operational technology (OT) with traditional IT organizations require consistency of architecture to better manage application, data and platforms.

To meet these needs, technology providers require choice in the underlying architectures that power the solutions brought to market. Though x86 is a rich architecture, the ability to deliver tailored CPUs through a variety of silicon partners is non-existent due, in large part, to the business model of such CPU providers. As the sole provider of IP and product, tailoring product designs per customer or usage scenario is an unattainable business model.

The market requires an ecosystem of solutions providers to deliver the custom design tools and silicon and platform support required to enable the diversity of usage demanded by these new compute models. Moor Insights & Strategy sees Arm as being well-positioned to deliver the platform for this new class of infrastructure allowing the ecosystem to rapidly design, build and deliver solutions that enable software solutions providers to innovate and deliver at a faster pace, and enterprise IT organizations can deliver services that enable first-mover status in an ever-competitive business climate.

CLOUD-NATIVE RULES THE MODERN DATACENTER

Software innovation has long outpaced the innovation of hardware - specifically, CPU innovation, and it can be argued that software innovation has been stifled by the longer design and manufacturing process of hardware.

Cloud-native is a manifestation of the cloud deployment model. Hardware becomes pools of resources to enable workloads to run most efficiently, and these resources (compute, memory, storage and networking) are assembled and utilized in a composable fashion. A deployment model based on container technology abstracts hardware from the application. Services run encapsulated in containers that can run in a variety of compute environments.

Cloud-native exists because of the requirements (outlined earlier) brought on by cloud and edge computing. What makes cloud-native *thrive* is an open ecosystem of software and hardware vendors. While open source initially found success in supporting edge cases and "skunkworks," type projects in the datacenter, the open source movement matured in part due to the adoption of cloud. Today, roughly 70% of the world's web servers run on Linux or Unix.

EDGE INTRODUCES BOTH OPPORTUNITY AND CHALLENGES

Edge computing is comprised of many deployment models supporting countless use cases. Smart cities, manufacturing, healthcare and oil and gas are just a few examples. As organizations fully realize the benefits of the connected edge, Moor Insights believes the edge market will meet and exceed the forecast put forth by many industry pundits.

5G is the underlying technology that makes "the edge" work for organizations looking to capitalize on the amount of data generated in the wild. However, telecommunications providers deploying 5G require equipment of varying compute and power profiles everywhere and in every environmental condition. Ideally, the equipment deployed for the movement of data, video and voice across the 5G network would use an architecture that is also used by the connected devices and compute platforms generating this explosion of traffic.

THE IMPORTANCE OF "OPEN SILICON" IN AN OPEN SOURCE WORLD

To deliver tailored solutions that fully enable the use cases that are driving cloud computing and "the edge," software architects and developers require flexibility and choice. As previously mentioned, this requires a hardware ecosystem that can design tailored platforms. Hardware design firms must be able to rely on a network of CPU firms that can quickly design and deliver tailored silicon at a cost that makes delivering such hardware platforms financially viable.

This market dynamic should cause the silicon ecosystem to grow deeper and broader as these markets further mature. Companies such as Marvell, Ampere, Fujitsu, HiSilicon and others should see increased market opportunity from CSPs, telco operators and other hyperscalers, as well as the solutions providers emerging in the IoT/edge space. These companies will probably be joined by several new players that build practices around markets such as data analytics, high-performance computing (HPC), artificial intelligence (AI), etc.

However, for these companies to find success, they must have the ability to incorporate more agile design principles that stress time to market and architectural flexibility. For these reasons, we expect to see Arm IP continue to grow in popularity across the supplier to buyer continuum – from silicon to hardware to software to the datacenter.

For perhaps the best example of an organization relying on Arm IP to enable a better cost-performance compute environment, look no further than Amazon Web Services' recent launch of the A1 family of instances. This environment is powered by an in-house design around Arm's architecture. The result? AWS claims a 45% lower cost to customers running network intensive workloads on the "Graviton" processor.

AWS is not alone in its adoption of Arm-based servers. Microsoft has been public about its evaluation of Arm-based instances as an Azure offering, as has Alibaba cloud. Huawei recently announced a server CPU designed in-house to address both HPC and cloud instances.

Equally interesting is the adoption of Arm-based servers from the smaller, more agile cloud players like Packet, a bare metal as a service (BMaaS) cloud provider. This deployment demonstrates how smaller players can deliver differentiated services in an extremely competitive market populated by giants.

Much like open source software, MI&S believes the hyperscale community's adoption of "open source silicon" will subsequently lead to broader adoption in the mainstream enterprise IT market. Organizations that are not looking at the entire open source ecosystem (from silicon to software) run the risk of falling further behind the competition.

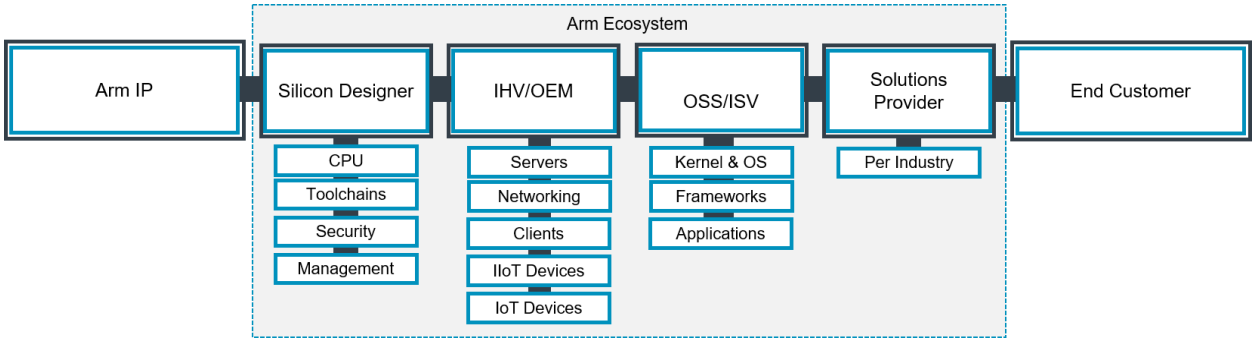
ARM – THE CORNERSTONE OF THE OPEN SOURCE HARDWARE AND SOFTWARE ECOSYSTEM

MI&S sees silicon (the CPU) as the foundational element of the open source ecosystem. As such, Arm is well-positioned as the provider of open silicon intellectual

property and is worthy of serious consideration across the hardware and software ecosystem, as well as by software architects and datacenter architects. With Arm:

- ***Silicon designers*** can leverage an architecture designed for enterprise-class server compute, networking and storage and pair that with silicon IP designed down to the devices that make up the IoT.
- ***Systems designers*** can quickly design and build systems that meet the very specific needs of workloads, industries and environments.
- ***Software developers*** have the flexibility to develop solutions that span the range of use (device to server) with a richness of performance and security capabilities from a vibrant open source software community that has long embraced the Arm architecture.
- ***Solutions providers*** can bundle systems and software to deliver industry-specific solutions with a range of functionality, power and cost.

FIGURE 2 – THE ARM ECOSYSTEM



Source: Moor Insights & Strategy

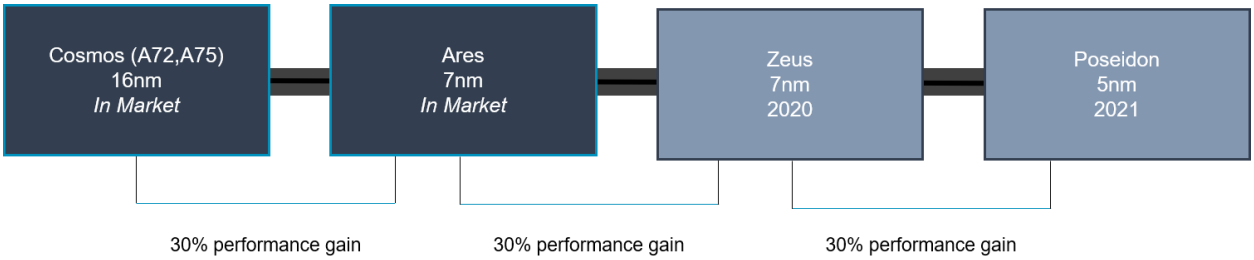
IT infrastructure providers of all types can leverage any of these ecosystem players and realize real benefits around performance and cost for CSPs and hyperscalers, cost and customization of software stacks.

NEOVERSE – THE MANIFESTATION OF ARM'S DATACENTER VISION

Arm has long been viewed as the IP provider for phone and device chips. In fact, because of the amount of its IP that resides in a variety of datacenter infrastructure (such as routers and security appliances), Arm estimates its footprint across the space is approximately 30%.

The recent launch of Arm Neoverse signaled a significant investment to the datacenter infrastructure space. Neoverse is a family of datacenter class infrastructure IP designed to compete with x86 and other architectures on performance in a lower power envelope. The Neoverse roadmap shows not only Arm's commitment to the infrastructure market but its commitment to delivering performance gains that should make it a viable alternative to its x86 competitors.

FIGURE 3 – ARM NEOVERSE IP SCHEDULE



Source: Moor Insights & Strategy

The first platforms launched under Arm's Neoverse brand were its N1 and E1 platforms. N1 is Arm's platform supporting higher-end server compute where the E1 platform would be considered the platform for silicon designers looking to develop solutions for network and edge computing use cases.

WHAT THE FUTURE HOLDS

The use of technology is what drives its evolution. Fifteen years ago, the average person owned a mobile phone, a media player, a laptop, a portable GPS and a personal reader. Before the mainstream deployment of virtualization, dedicated server platforms designed to support workloads sat highly underutilized in racks populating the datacenter.

Fifteen years ago, the idea of microservices delivered as containers in a distributed model seemed impossible. The idea of using the data produced by sensors on machines to gain efficiencies and competitive advantages seemed like science fiction, never mind the actual deployment of AI/ML.

Looking ahead 15 years, the future state of IT is largely a mystery, but we know that the world will be ever more connected. The 8 billion or so people inhabiting this planet will be generating more data than ever, as will be the technologies and tools that are

designed to make our lives simpler and safer, and the market will see a hardware and software ecosystem filled with new players challenging the norm with disruptive technologies and solutions.

Because of this, the heterogeneous compute won't just be a requirement; it will be table stakes. An open ecosystem that can drive performance, security and efficiency across the device-to-cloud continuum will be the norm. This open ecosystem will start with an open silicon design.

Because of this, companies like Arm should be well-positioned as they allow their ecosystem partners the greatest flexibility in terms of performance and performance across a range of devices.

CALL TO ACTION

IT Infrastructure is undergoing a major change. Though proprietary software once dominated the datacenter landscape, open source software has become more pervasive and is the likely distribution for some of the most advanced applications. In other words, open source is driving business.

MI&S believes successful IT organizations will rationalize deployments of open source software. The wise IT organization will understand where open source is the right fit in the right situation. Likewise, the smart IT shop will know when proprietary solutions are best suited to meet the needs of the business.

As IT organizations continue to strategically embrace and extend open source software, the larger open source movement follows in support. The movement extends from software to hardware and from hardware to the CPUs that power datacenter infrastructure. This larger open source movement further abstracts software from the underlying hardware architecture, which can remove performance and optimization disparities that may have previously existed between competing architectures.

Because of these factors, Arm is well-positioned in the market. It is, in many ways, the cornerstone of the open source community with a depth and breadth of partners that (Arm would argue) is unmatched in the industry. One could argue that Arm is the open source silicon IP for the open source community.

For several reasons, MI&S recommends IT organizations of all sizes to consider Arm-based solutions for the workloads deployed in support of the business:

- Breadth and depth of the hardware and software ecosystem;
- Faster time to market, faster time to value, the potential for first-mover status in the market;
- Range of solutions that can meet the power, cost and performance budgets of organizations;
- Consistency of architecture from the datacenter to the cloud to the edge; and
- Simplicity of management and strengthening of security.

Finally, pay close attention to Arm's progress with Neoverse. The products that partners will release from the N1 and E1 platforms may have major impacts on the server market in the years to come.

Check out more at https://www.arm.com/products/silicon-ip-cpu

IMPORTANT INFORMATION ABOUT THIS PAPER

CONTRIBUTOR
Matt Kimball, Senior Analyst at Moor Insights & Strategy

PUBLISHER
Patrick Moorhead, Founder, President, & Principal Analyst at Moor Insights & Strategy

INQUIRIES
Contact us if you would like to discuss this report, and Moor Insights & Strategy will respond promptly.

CITATIONS
This paper can be cited by accredited press and analysts but must be cited in-context, displaying the author's name, author's title, and "Moor Insights & Strategy". Non-press and non-analysts must receive prior written permission by Moor Insights & Strategy for any citations.

LICENSING

DISCLOSURES
This paper was commissioned by Arm. Moor Insights & Strategy provides research, analysis, advising, and consulting to many high-tech companies mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

DISCLAIMER
The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. Moor Insights & Strategy disclaims all warranties as to the accuracy, completeness, or adequacy of such information and shall have no liability for errors, omissions, or inadequacies in such information. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators and not as precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on what the future holds, they are subject to risks and uncertainties that could cause actual results to differ materially. You are cautioned not to place undue reliance on these forecasts and forward-looking statements, which reflect our opinions only as of the date of publication for this document. Please keep in mind that we are not obligating ourselves to revise or publicly release the results of any revision to these forecasts and forward-looking statements in light of new information or future events.