intel.

Google Cloud's Anthos on Bare Metal with Intel® Architecture

Get the simplicity and flexibility of Google Cloud's Anthos along with the cost efficiency, control, and predictable performance of bare metal—all running on high-performance Intel® technology in the data center or at the edge

Anthos

Solution Benefits

- Modernize existing applications and build cloud-native apps anywhere to promote **agility and cost savings**.

- Eliminate the hypervisor layer to help **reduce costs and lower latency**.

- Achieve **predictable performance** with a bare metal environment that eliminates noisy neighbors.

- Use verified Intel reference architectures to confidently **deploy Anthos on Bare Metal in the data center and at the edge**.

Executive Summary

Google Cloud customers want to modernize existing applications and build cloud-native apps anywhere to promote agility and cost savings—and that's exactly what Google Cloud's Anthos delivers. However, because some workloads—especially those sensitive to latency—are not suited to a virtualized environment, Google now offers Google Cloud's Anthos on Bare Metal. This solution combines the application platform benefits of Anthos with the hardware/software stack control, high-performance, and cost efficiency of a bare metal environment.

Google and Intel collaborated to create verified Intel® architecture-based reference configurations for Anthos on Bare Metal that can provide fast insights, low latency, and high throughput. Powerful Intel® hardware features drive the predictable performance that latency-sensitive workloads require in the data center, as well as edge workloads such as retail inventory applications or analytics at the edge.

Without the expense of virtualization, Anthos on Bare Metal can help lower costs, potentially increase performance, and help protect against noisy neighbors often found in a multi-tenant environment. Running Anthos on Bare Metal, combined with Intel® hardware, can help enterprises meet their cost and performance goals.

Author

Todd Christ
Senior Solutions Architect,
Data Platforms Group

Intel and Google teamed up to offer verified reference architectures for data center and edge deployments of Google Cloud's Anthos on Bare Metal.

Business Challenge: Virtualization Isn't for Everyone

Typical cloud environments provide users with a virtual machine as the building block for their infrastructure. While this is an incredibly powerful and flexible resource, it abstracts users from directly accessing and making decisions about the hardware—diminishing control over the solution stack components. In addition, a virtualized environment poses several other potential problems:

- Virtualization can increase costs because customers must pay for the hypervisor layer.
- Shared CPU, memory and storage, and network resources can lead to unpredictable performance that can significantly impact large-scale or latency-sensitive workloads like financial transactions.
- Security concerns can arise in a multitenant environment, especially for workloads that must meet stringent data protection and privacy regulations such as those in the healthcare industry.

Bare metal environments can solve many of these issues but can introduce problems of their own, such as having to manually perform lifecycle management, diagnostics, health checks, logging, and monitoring. Enterprises are seeking a way to gain the benefits of bare metal without the management drawbacks.

Solution Value: Cost Efficiency, Control, Ease of Deployment—and Performance

Using Anthos, customers can begin—or accelerate—their journey to a true hybrid cloud, and eventually to a multicloud solution. Both in the cloud and on-premises, Anthos lets customers take advantage of modern cloud-native technologies like containers, serverless computing, service meshes, and consistent policy management.

By removing the costs and latency associated with the hypervisor layer, Anthos on Bare Metal lets customers cost-effectively run even computationally intensive applications like big data analytics and machine learning. Customers can take advantage of Anthos benefits—centralized management, increased flexibility, and developer agility—even for demanding applications. Customers can also choose the OS and a hardware configuration that suits their workloads.

Anthos on Bare Metal ships with built-in networking, lifecycle management, diagnostics, health checks, logging, and monitoring to simplify management and deployment. Anthos enables up to a 55 percent increase in platform operations efficiency.[1] Use cases for Anthos on Bare Metal include data center workloads like machine learning and in-memory databases, as well as edge workloads such as analytics at the edge. Even for edge deployments, Anthos on Bare Metal lets customers use the Google Cloud Console to manage all applications deployed to Anthos edge locations. And whether in the data center or at the edge, the Intel reference architectures for Anthos on Bare Metal help ensure workloads run optimally.

Anthos (whether on bare metal or not), provides these additional benefits:

- **Flexibility.** Enables unified cloud-resource management, seamless application portability, and data mobility across cloud environments, on premises, and with public cloud providers (including Google Cloud, Amazon Web Services, and Microsoft Azure), all enabled by Intel architecture.
- **Scalability.** Helps customers securely and reliably scale, manage, and extend workloads from enterprise to cloud, without the need for reconfiguration, application changes, or testing.

The Anthos ecosystem includes Google Cloud's app modernization platform, Cloud Run. This platform provides custom machines types, virtual private cloud networking, and integration with existing Kubernetes-based solutions. Cloud Run for Anthos provides a flexible serverless development platform on Google Kubernetes Engine. Cloud Run for Anthos is powered by Knative, an open source project that supports serverless workloads on Kubernetes. Cloud Run is also available as a fully managed serverless platform, without Kubernetes. Customers can take advantage of Cloud Run for Anthos to further improve the application management lifecycle.

Solution Architecture: Anthos on Bare Metal in the Data Center

Data volumes are growing. Machine learning is becoming a part of many enterprise processes. Faster analytics insights help drive competitive edge. Stringent security requirements demand fast encryption and decryption. These trends mean that many data center workloads can benefit from an infrastructure characterized by predictable, low-latency performance. Anthos on Bare Metal, enabled by the latest Intel® technologies, can provide what these data center workloads need. The data center Intel reference architecture for Anthos on Bare Metal is shown in Figure 1, while Table 1 provides the full bill of materials. This reference architecture is suitable for real-time analytics, machine- and deep-learning, video streaming, and other latency- and performance-sensitive workloads.
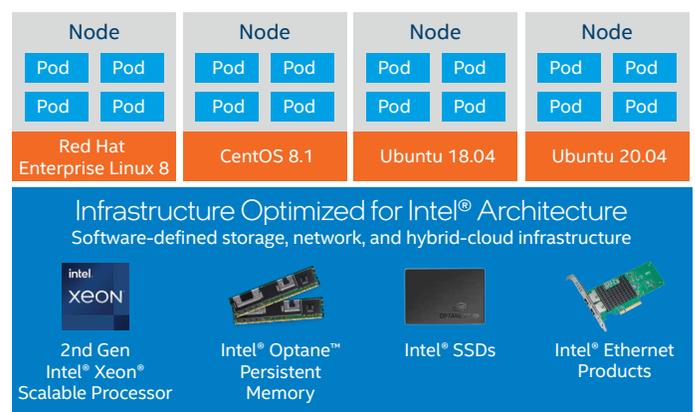

Figure 1. Anthos on Bare Metal in the data center runs on 2nd Gen Intel® Xeon® Scalable processors, Intel® Optane™ persistent memory, and other hardware innovations from Intel.

For the data center reference architecture, the Intel® Xeon® Gold 6240Y processor (2.6 GHz, 18 cores) provides an optimized balance of price and performance in a mainstream data center configuration. This processor supports several Intel technologies that can boost performance, including:

- **More memory, fast storage** with Intel® Optane™ persistent memory (PMem) in Storage over App Direct Mode. Intel Optane PMem offers high density—up to 512 GB per module—and persistence. Organizations can use PMem to cost effectively expand the capacity of available memory or higher quantities of "hot" data available for in-memory databases, analytics, and other demanding workloads. PMem, along with Intel® Optane™ solid state drives (Intel® Optane™ SSDs), represents a new class of memory and storage technology that helps fill the gaps between DRAM and traditional storage technologies (see Figure 2).

- **Fast deep-learning inference** with Intel® Deep Learning Boost (Intel® DL Boost) and Intel® AVX-512. Intel DL Boost uses Vector Neural Network Instructions to help speed up deep-learning inference, by combining three instructions into a single instruction.
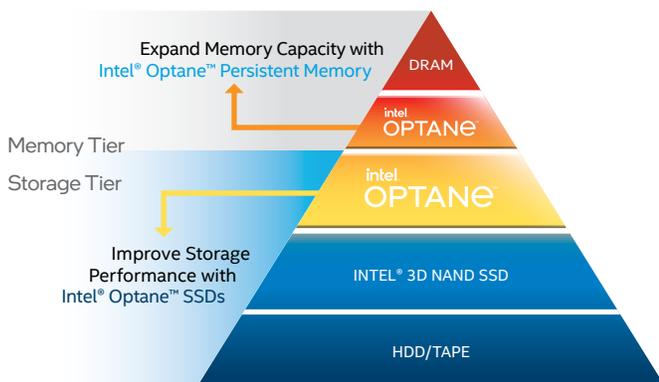
Figure 2. Intel® Optane™ technology helps fill the gap between DRAM and traditional storage technologies such as 3D NAND and hard disk drives.

Solution Architecture: Anthos on Bare Metal at the Edge

Customers with edge applications can also benefit from a verified, Anthos-ready reference architecture for Anthos on Bare Metal. For example, a retail store might use an affordable, small-footprint Intel® NUC to run its store inventory application. Industrial use cases are also possible, such as gathering Internet of Things data and analyzing it at the edge to reduce the amount of data sent to Anthos deployments in the data center or cloud.

Anthos on Bare Metal provides the deployment and management platform for these edge applications, just as it does for workloads hosted in the data center. The edge Intel reference architecture for Anthos on Bare Metal is shown in Figure 3, while Table 2 provides the full bill of materials.

At the Edge
10th Gen Intel® Core™ i7-10710U Processor
Up to 4.7 GHz, 6 cores
Intel® UHD Graphics for 10th Gen Intel® Processors

Storage 1x Intel® SSD Pro 7600p Series 256 GB, M.2 80 mm
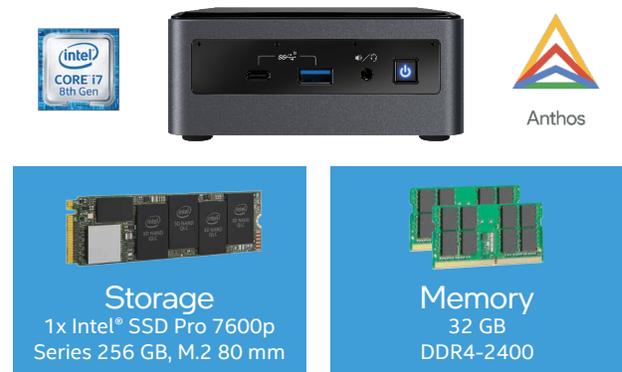
Memory 32 GB DDR4-2400

Figure 3. Anthos on Bare Metal at the edge runs on 10th Generation Intel® Core™ processors and includes other Intel technologies that boost performance and manageability.

For the edge reference architecture, the 10th Generation Intel® Core™ i7-10710U processor (up to 4.7 GHz, 6 cores) provides excellent performance in an edge environment. This processor supports several advanced Intel technologies that can improve performance and manageability, including the following:

- **High-speed graphics** with Intel® UHD Graphics for 10th Generation Intel® processors. This integrated graphics processor is a GPU built into the processor. Integrated graphics hardware doesn't use a separate memory bank for graphics/video. Instead, the GPU uses system memory that is shared with the CPU. Since integrated graphics is built into the processor, it typically uses less power and therefore creates less heat, which can result in a longer battery life.

- **Further capacity and performance** with the optional addition of Intel® Optane™ Memory H10 with Solid State Storage. This technology combines Intel® Optane™ memory and Intel® QLC 3D NAND storage to deliver a personalized computing experience with a new level of performance and large storage capacity.

Sizing and Scaling Considerations

Building an Anthos on Bare Metal solution has some base requirements, and we offer these recommendations for growth:

- **Data center.** In the data center, start with six servers as the baseline, and then grow this deployment up to 18 servers. Once the application requirements have exceeded this size, we suggest a new deployment of six (or more) servers in a separate rack to extend bare metal capacity across the data center.

- **Edge.** Edge solutions with Intel NUCs should start with a minimum of four systems and can grow up to 12 NUCs in design. Beyond 12 NUCs, consider moving to an Intel® Xeon® processor platform because that will simplify deployment and offer the scalability beyond what a NUC solution can provide.

Table 1. Bill of Materials for Data Center Reference Architecture for Anthos on Bare Metal

	Item	Model/Spec/Quantity
Control/Master 1 node	CPU	2x Intel® Xeon® Gold 6240Y processor (2.6 GHz 18 cores)
	Memory	192 GB memory (12x 16 GB)
	Persistent Memory	N/A
	Boot Drive	2x Intel® SSD DC S4500 240 GB RAID1 (SATA)
	Storage Drive	N/A
	Network	10/25 GbE Intel® Ethernet Adapter X710-DA2 dual port or above
	Management Network	10/25 GbE Intel Ethernet Adapter X710-DA2 dual port or above
Compute/Worker 3 nodes – Ceph and Rook	CPU	2x Intel Xeon Gold 6240Y processor (2.6 GHz 18 cores)
	Memory	384 GB memory (24x 16 GB)
	Persistent Memory	N/A
	Boot Drive	2x Intel SSD DC S4500 240 GB RAID1 (SATA)
	Storage Drive	3x 3.5 TB Intel® SSD D3-S4510 (SATA)
	Network	10/25 GbE Intel® Ethernet Adapter XXV710-DA2 dual port or above
	Management Network	10/25 GbE Intel Ethernet Adapter XXV710-DA2 dual port or above
Compute/Worker 2 nodes	CPU	2x Intel Xeon Gold 6240Y processor (2.6 GHz 18 cores)
	Memory	384 GB memory (12x 32 GB)
	Intel® Optane™ persistent memory	1 TB (8x 128 GB) in Storage over App Direct Mode
	Boot Drive	2x Intel SSD DC S4500 240 GB RAID1 (SATA)
	Storage Drive	N/A (storage on Intel Optane PMem)
	Network	10/25 GbE Intel Ethernet Adapter XXV710-DA2 dual port or above
	Management Network	10/25 GbE Intel Ethernet Adapter XXV710-DA2 dual port or above

Note: The data center-focused reference design showcases a single control node for simplicity in deployment, but for high-availability, customers can also deploy three control nodes.

Table 2. Bill of Materials for Edge Reference Architecture for Anthos on Bare Metal

	Item	Model/Spec/Quantity
Intel® NUC 10 Performance Kit NUC10i7FNK	CPU	10th Generation Intel® Core™ i7-10710U processor (1.1 GHz default, 4.7 GHz with Intel® Turbo Boost Technology, 6 cores)
	Memory (upgraded)	32 GB 2x DDR4-2400 SODIMM
	Storage and Boot Drive	1x Intel® SSD Pro 7600p 256 GB, M.2 80 mm (3D NAND, tri-level cell, PCIe 3.1 x4, NVMe)

Conclusion

Google Cloud's Anthos on Bare Metal customers value stellar and predictable performance, enhanced security and reliability, and access to recent technology innovations. They get all of this and more using the bare metal reference architectures for data center and edge, based on the latest Intel technology.

The global bare metal cloud market is expected to grow from USD 5.7 billion in 2020 to USD 26.2 billion by 2025, at a compound annual growth rate of 38.4 percent.[2] The collaboration between Intel and Google on the Anthos on Bare Metal launch will open a wide array of new use cases for industries that have not yet adopted a cloud first approach, but can now benefit from the same technologies available on Google Cloud. From the data center to the edge, Intel's reference architectures for Anthos on Bare Metal can help accelerate customer and partner adoption of Anthos on Bare Metal.

Software Innovation from Intel for Anthos on Bare Metal Users

Intel helped develop the PMEM-CSI operator, which facilitates deploying and managing the PMEM-CSI driver on a Kubernetes cluster. This operator is based on the tCoreOS operator-sdk tools and APIs.

PMEM-CSI is a CSI storage driver for container orchestrators like Kubernetes. It makes local persistent memory (such as Intel Optane PMem) available as a filesystem volume to container applications. It can currently utilize non-volatile memory devices that can be controlled via the libndctl utility library. All of the source code required to build PMEM-CSI is available under open source licenses. The source code files identify which external Go modules are used. Binaries are distributed as container images on DockerHub.

For more information, read Introduction to PMEM-CSI for Kubernetes, Application examples, and Persistent Memory Programming.

Learn More

You may find the following resources helpful:

- Intel Xeon Scalable processors
- Intel Optane persistent memory
- Intel AES-NI
- Intel Deep Learning Boost
- Intel NUCs
- Intel Optane Memory H10 with Solid State Storage
- Google Cloud's Anthos on Bare Metal

Find the solution that is right for your organization.
Contact your Intel representative or visit <URL>.

Solution Provided By:

1 The economic benefits of Anthos; https://cloud.google.com/anthos/forrester-tei-report

2 Bare Metal Cloud Market Size, Share & Trends Analysis Report; https://www.grandviewresearch.com/industry-analysis/bare-metal-cloud-market