# VERSAL™ HBM SERIES

**6X memory bandwidth at 65% lower power per bit vs. four LPDDR4[1]**
**2X adaptable compute engines[2] for evolving algorithms and protocols**
**1.8X faster connectivity[3] to adapt for emerging networks**

## OVERVIEW

The Versal™ HBM series enables the convergence of fast memory, adaptable compute, and secure connectivity in a single platform. The Versal HBM series is architected to keep up with the higher memory needs of the most compute intensive, memory bound applications, providing adaptable acceleration for data center, wired networking, test and measurement, and aerospace and defense applications.

Based on the fourth generation of stacked silicon interconnect technology (SSI technology), the Versal HBM adaptive SoCs integrate advanced HBM2e DRAM, providing 819 GBps of memory bandwidth and 2 GB of capacity for 6X more memory bandwidth and 65% lower power per bit than LPDDR4 solutions[1].

Built on the foundation of the Versal Premium series, the Versal HBM adaptive SoC integrates an extensive set of multi-terabit, power-optimized connectivity cores and 112G PAM4 transceivers to adapt to emerging network protocols. While nearly doubling the transceiver speed, the Versal HBM adaptive SoC also enables you to secure every layer of the network infrastructure with built-in encryption engines.

As an adaptive, heterogeneous compute platform, the Versal HBM series is engineered to accelerate a wide range of workloads with large data sets. The adaptable compute engines, coupled with the customizable memory hierarchy, enable massive parallelism and adaptability for evolving algorithms and emerging protocols by matching workloads to the appropriate engines and interfaces.

Providing a design-entry point for any developer, including Vivado™ design suite and Vitis™ unified software platform, the Versal HBM adaptive SoC unlocks a new class of acceleration for artificial intelligence (AI), machine learning (ML), database acceleration, next-gen firewall, advanced data processing system, and a breadth of other compute-intensive, memory-bound applications.

## HIGHLIGHTS

**Alleviating Network and Compute Bottlenecks for Large Data Sets**
- 819 GBps bandwidth for faster runtime
- 32 GB capacity to process bigger data sets
- Lowest power-per-bit solution, easing power and thermal constraints and costs
- Global memory access with built-in switch for minimized design size and power

**Scalable, Secure Connectivity for Next-Gen Compute and Networking Infrastructure**
- 112G PAM4 transceivers adaptable to emerging network modules and protocols
- 100G and 600G Ethernet cores enabling a wide variety of data rates and standards
- 600G Interlaken cores with FEC for chip-to-chip interconnect
- 400G High-Speed Cryptography Engines for inline network security

**Adaptable Acceleration for Evolving Algorithms and Protocols**
- Adaptable Engines with programmable memory hierarchy for higher compute density
- Enhanced DSP Engines to support a wide range of data types for diverse workloads
- Programmable network on chip (NoC) for high bandwidth IP interconnect

## TARGET APPLICATIONS

### DATA CENTER
- Machine Learning Acceleration
- Compute Pre-Processing and Buffering
- Database Acceleration and Analytics

### WIRED NETWORKING
- Network Security Acceleration
- Search and Look-up System
- 800G Switch / Router

### TEST AND MEASUREMENT
- Network Testers
- Packet Capturing System
- Data Capturing System

### AEROSPACE AND DEFENSE
- Signal Processing
- Secure Communication Equipment

## FEATURE HIGHLIGHTS

| | |
|---|---|
| **Scalar Engines** | • Dual-core Arm® Cortex®-A72 application processing unit for Linux-class operating systems<br>• Dual-core Arm Cortex-R5F real-time processing unit for low latency and determinism<br>• Platform management for quick boot, power & thermal management, and safety & security enclave |
| **Adaptable Hardware Engines** | • Adaptable for any workload, including packet processing, ML models, security algorithms<br>• High bandwidth, low latency data movement between engines and I/Os<br>• Up to 93 TBps of programmable memory hierarchy for optimal compute efficiency |
| **Intelligent Engines** | • Enhanced DSP Engines (DSP58) for high-precision floating point & low latency<br>• Up to 75 TOPs with INT8 and 17.5 TFLOPs of DSP compute bandwidth for acceleration |
| **High Bandwidth Memory** | • 819 GBps bandwidth eliminates network and compute bottlenecks<br>• Up to 32 GB capacity for higher performance operation on larger data sets |
| **Programmable Network on Chip** | • Multi-terabit network on chip (NoC) with built-in arbitration and Quality of Service<br>• Programmable memory-mapped access and interconnect to all engines and custom logic<br>• Built-in switch for global access to any HBM memory location to minimize complexity and power<br>• Easy IP and kernel placement |
| **112G PAM4 and 32G NRZ Transceivers** | • Supports the latest optical and electrical communication standards<br>• Up to 5.6 Tbps of serial bandwidth for high density communication interfaces |
| **Integrated PCIe® Gen5 with DMA and CCIX** | • CPU-to-accelerator communication for next-generation data center applications<br>• Hardened, queue-based DMA Engines for efficient memory access<br>• Symmetric/asymmetric access to memory with cache coherent protocol support |
| **Integrated 600G Ethernet and 100G Multirate Ethernet Cores** | • Up to 2.4 Tbps of scalable Ethernet throughput<br>• Multirate: 400/200/100/50/40/25/10G<br>• Multi-standard: FlexE, Flex-O, eCPRI, FCoE, and OTN |
| **Integrated 600G Interlaken Cores with FEC** | • Scalable chip-to-chip interconnect from 12.5 Gbps to 600 Gbps<br>• Integrated FEC for power-optimized error correction |
| **400G High-Speed Crypto Engines** | • AES-GCM-256/128 Cryptography Engines<br>• Up to 1.2 Tbps line-rate encryption throughput<br>• 400G of line-rate crypto per core<br>• MACsec, IPsec supported with soft logic wrapper |

# NEXT STEPS

• For more information about the AMD Versal HBM series, visit www.amd.com/versal-hbm.

1. Based on AMD internal analysis in May 2023, comparing a system implementation of a single Versal HBM VH1542 device with in-package HBM2E to a Versal Premium VP1502 device implementation with four LPDDR4-4266 components. Assuming sequential memory accesses with 40% read/write transactions. Power calculation generated using AMD Power Design Manager and a third-party system power calculator. Configurations may vary, yielding different results. (VER-013)
2. Logic density vs. Virtex™ UltraScale+™ HBM FPGA
3. Based on AMD internal analysis in June 2023, comparing the Versal HBM VH1782 device with GTYP and GTM transceivers, to a Virtex UltraScale+ HBM VU47P device with GTY transceivers. Configurations may vary, yielding different results. (VER-017)