

## Versal HBM Series

- > 8X memory bandwidth<sup>1</sup> at 63% lower power vs. DDR5
- > 2X faster secure connectivity<sup>2</sup> to adapt for emerging networks
- > 2X adaptable compute engines<sup>3</sup> for evolving algorithms and protocols

### 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 & measurement, and aerospace & defense applications.

Based on the fourth generation of stacked silicon interconnect technology (SSI technology), the Versal HBM ACAPs series integrate the most advanced HBM2e DRAM, providing 820GB/s of memory bandwidth and 32GB of capacity for 8X more memory bandwidth and 63% lower power than DDR5 solutions<sup>1</sup>.

Built on the foundation of the Versal Premium series, the Versal HBM series integrates an extensive set of multi-terabit, power-optimized connectivity cores and 112G PAM4 transceivers to adapt to emerging network protocols and modules. While doubling the transceiver speed, the Versal HBM series secures 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 amount of 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 series unlocks a new class of acceleration for artificial intelligence (AI), machine learning (ML), database acceleration, next-generation firewall, advanced data processing, and a breadth of other compute-intensive, memory-bound applications.

### HIGHLIGHTS

#### Alleviating Network and Compute Bottlenecks for Large Data Sets

- > 820GB/s memory access bandwidth for faster runtimes
- > 32GB HBM 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 protocols
- > 600G Interlaken cores with FEC for chip-to-chip interconnect
- > 400G High-Speed Crypto 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 & Measurement

- > Network Testers
- > Packet Capturing System
- > Data Capturing System

#### AEROSPACE AND DEFENSE

- > Radar
- > Data and Signal Processing System
- > Secure Communication Equipment

## FEATURES

FEATURE HIGHLIGHTS	
Scalar Engines	<ul style="list-style-type: none"> <li>&gt; Dual-core Arm® Cortex®-A72 application processing unit for Linux-class operating systems</li> <li>&gt; Dual-core Arm Cortex-R5F real-time processing unit for low latency and determinism</li> <li>&gt; Platform management for quick boot, power &amp; thermal management, and safety &amp; security enclave</li> </ul>
Adaptable Hardware Engines	<ul style="list-style-type: none"> <li>&gt; Adaptable for any workload, including packet processing, ML models, security algorithms</li> <li>&gt; High bandwidth, low latency data movement between engines and I/Os</li> <li>&gt; Up to 93TB/s of programmable memory hierarchy for optimal compute efficiency</li> </ul>
Intelligent Engines	<ul style="list-style-type: none"> <li>&gt; Enhanced DSP Engines (DSP58) for high-precision floating point &amp; low latency</li> <li>&gt; Up to 75TOPs with INT8 and 17.5TFLOPs of DSP compute bandwidth for acceleration</li> </ul>
High Bandwidth Memory	<ul style="list-style-type: none"> <li>&gt; 820GB/s bandwidth eliminates network and compute bottlenecks</li> <li>&gt; Up to 32GB capacity for higher performance operation on larger data sets</li> </ul>
Programmable Network on Chip	<ul style="list-style-type: none"> <li>&gt; Multi-terabit network on chip (NoC) with built-in arbitration and Quality of Service</li> <li>&gt; Programmable memory-mapped access and interconnect to all engines and custom logic</li> <li>&gt; Built-in switch for global access to any HBM memory location to minimize complexity and power</li> <li>&gt; Easy IP and kernel placement</li> </ul>
112G PAM4 and 32G NRZ Transceivers	<ul style="list-style-type: none"> <li>&gt; Supports the latest optical and electrical communication standards</li> <li>&gt; Up to 5.6Tb/s of serial bandwidth for high density communication interfaces</li> </ul>
Integrated PCIe® Gen5 with DMA, CCIX, and CXL	<ul style="list-style-type: none"> <li>&gt; CPU-to-accelerator communication for next-generation data center applications</li> <li>&gt; Hardened, queue-based DMA Engines for efficient memory access</li> <li>&gt; Symmetric/asymmetric access to memory with cache coherent protocol support</li> </ul>
Integrated 600G Ethernet and 100G Multirate Ethernet Cores	<ul style="list-style-type: none"> <li>&gt; Up to 2.4Tb/s of scalable Ethernet throughput</li> <li>&gt; Multirate: 400/200/100/50/40/25/10G</li> <li>&gt; Multi-standard: FlexE, Flex-O, eCPRI, FCoE, and OTN</li> </ul>
Integrated 600G Interlaken Cores with FEC	<ul style="list-style-type: none"> <li>&gt; Scalable chip-to-chip interconnect from 12.5Gb/s to 600Gb/s</li> <li>&gt; Integrated FEC for power-optimized error correction</li> </ul>
400G High-Speed Crypto Engines	<ul style="list-style-type: none"> <li>&gt; AES-GCM-256/128 Cryptography Engines</li> <li>&gt; Up to 1.2Tb/s line-rate encryption throughput</li> <li>&gt; 400G of line-rate crypto per core</li> <li>&gt; MACsec, IPsec supported with soft logic wrapper</li> </ul>

## TAKE THE NEXT STEP

For more information about the Xilinx® Versal HBM series, visit [www.xilinx.com/versal-hbm](http://www.xilinx.com/versal-hbm)

## Corporate Headquarters

Xilinx, Inc.  
2100 Logic Drive  
San Jose, CA 95124  
USA  
Tel: 408-559-7778  
www.xilinx.com

## Xilinx Europe

Xilinx Europe  
Bianconi Avenue  
Citywest Business Campus  
Saggart, County Dublin  
Ireland  
Tel: +353-1-464-0311  
www.xilinx.com

## Japan

Xilinx K.K.  
Art Village Osaki Central Tower 4F  
1-2-2 Osaki, Shinagawa-ku  
Tokyo 141-0032 Japan  
Tel: +81-3-6744-7777  
japan.xilinx.com

## Asia Pacific Pte. Ltd.

Xilinx, Asia Pacific  
5 Changi Business Park  
Singapore 486040  
Tel: +65-6407-3000  
www.xilinx.com

## India

Xilinx India Technology Services Pvt. Ltd.  
Block A, B, C, 8th & 13th floors,  
Meenakshi Tech Park, Survey No. 39  
Gachibowli(V), Seri Lingampally (M),  
Hyderabad -500 084  
Tel: +91-40-6721-4747  
www.xilinx.com

