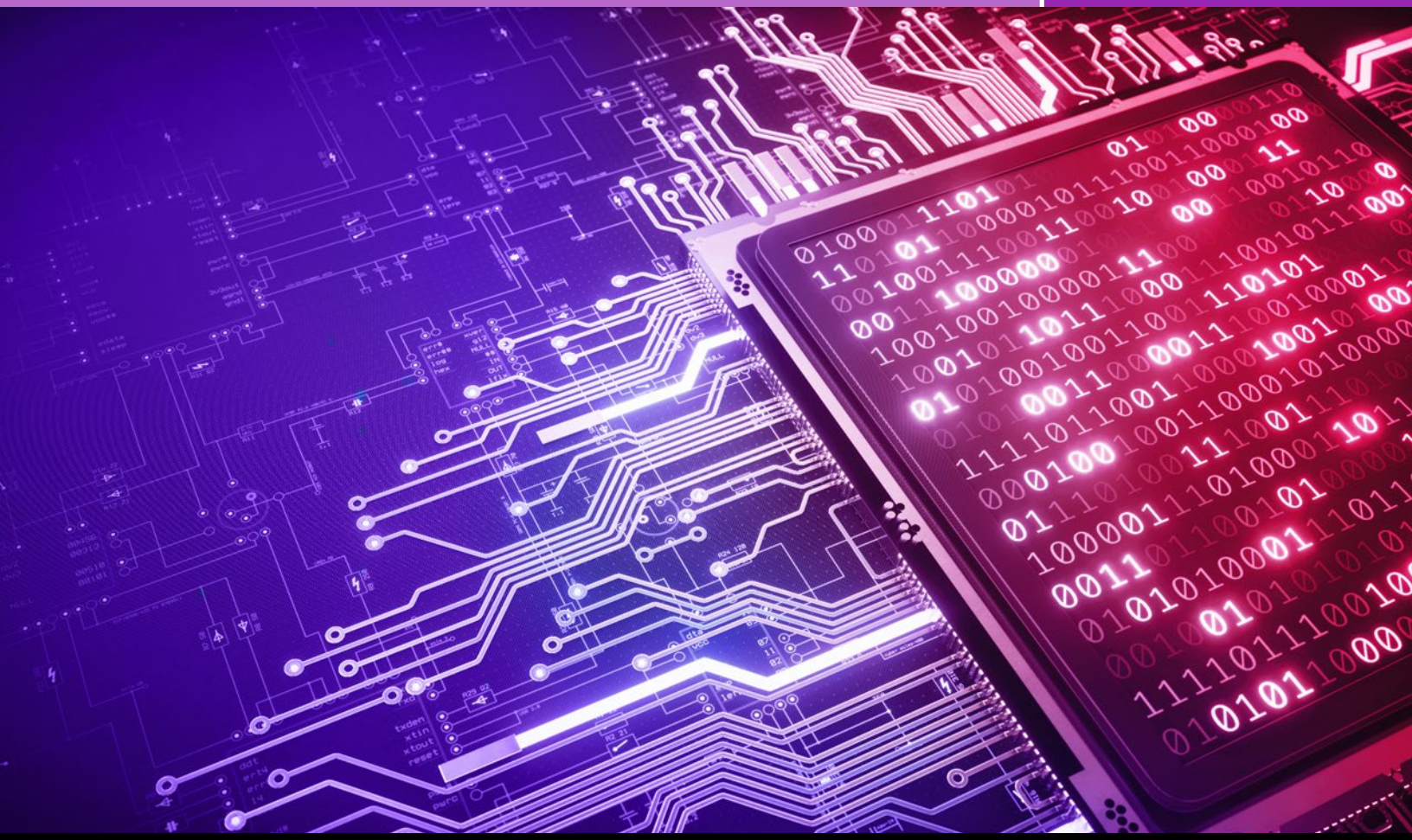
WHITE PAPER



Scale AI: Engineering the Next Leap in LPDDR6 Low-Power Memory

A Practical Look at How LPDDR6 Increases Bandwidth, with Less Energy, Better Margins for AI Systems

KEYSIGHT

Why LPDDR6 Matters for Scaling AI

Scaling AI often gets framed as more GPUs and bigger clusters. In practice, scaling is constrained by the whole system, and each gain in compute or throughput shifts pressure onto other layers, including bandwidth, power, and thermal margins. In other words, scaling is not a single upgrade, it is a system-wide engineering exercise.

One of the constraints that shows up early is memory: bandwidth, latency, and energy efficiency increasingly define system performance as AI workloads scale. Against that backdrop, JEDEC LPDDR6 is positioned as a next-generation low-power memory standard designed to raise per-pin data rates beyond 10.6 Gbps while targeting meaningful reductions in active and standby power versus by up to ~30% the prior generation.

Because of the improved performance at lower power, LPDDR6 impact is expected to extend well beyond mobile: the combination of higher throughput and improved power efficiency makes LPDDR6 a strong fit for AI/ML workloads, high-speed digital systems, automotive platforms, data centers, and edge deployments, where maximizing performance per watt is critical.

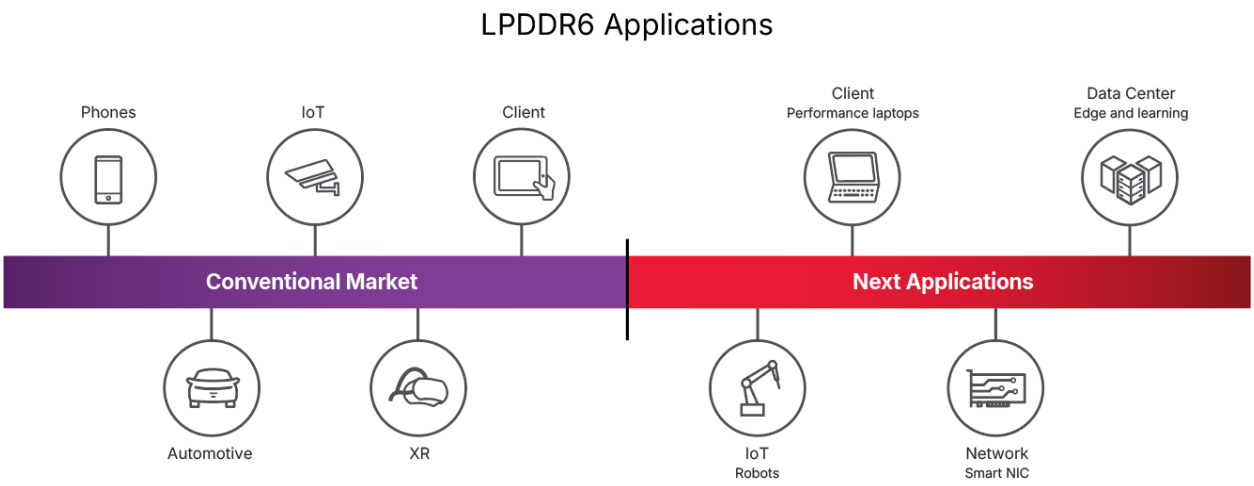


Figure 1. LPDDR6 expands beyond mobile. Conventional LPDDR apps (phone, client, IoT, automotive, XR) extending into AI data center, Smart NIC, robotics, and high-performance laptops.

What Are the Differences Between LPDDR6 and LPDDR5?

To support AI applications and high-performance computing (HPC), and building on the legacy of LPDDR5 and LPDDR5X, LPDDR6 introduces a suite of architectural and electrical innovations that redefine the balance between performance, power, and reliability.

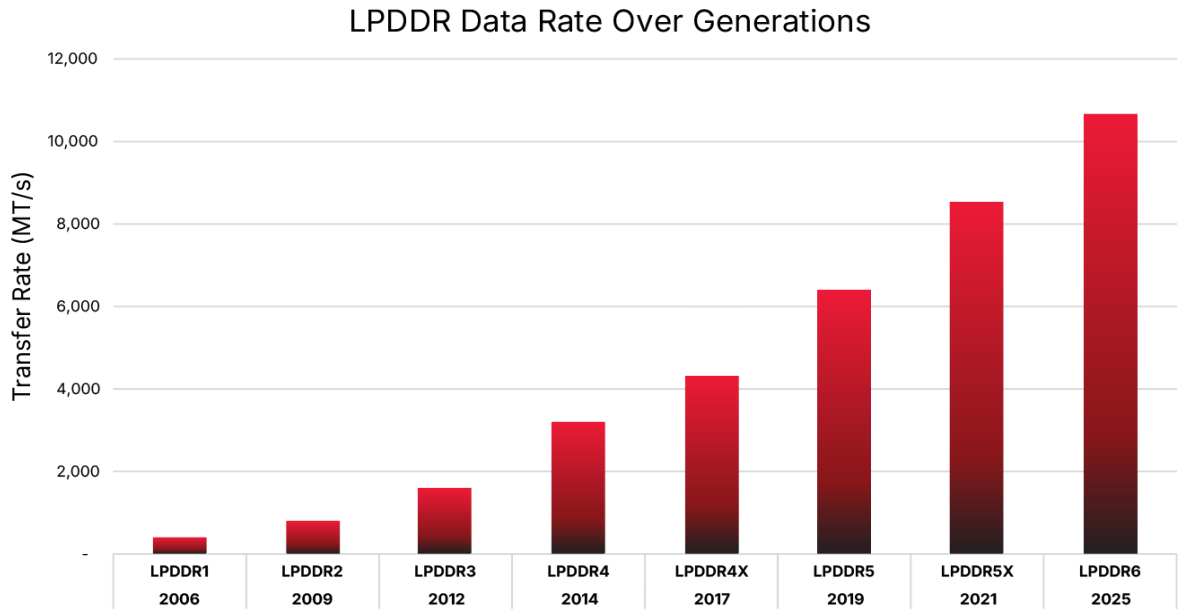


Figure 2. JEDEC LPDDR timeline. Generational rise in MT/s (mega-transfers/sec) per pin (data rate per DQ).

Performance (Increasing Bandwidth)

LPDDR6 significantly boosts bandwidth through a combination of higher signaling speeds, a refined dual sub-channel architecture, and increased data transfer efficiency.

Speed and Signaling

LPDDR6 is designed to outperform its predecessor. According to JEDEC standards, LPDDR6 speeds will start at 10.667 Gbps and scale up to 14.4 Gbps and beyond. This represents a leap over LPDDR5X, which typically caps out near 9.6 Gbps.

Architectural Shift: 24-bit Dual Sub-Channels

Beyond raw speed, LPDDR6 introduces a fundamental change to its channel architecture. While LPDDR5 utilized a 16-bit channel width, LPDDR6 moves to a 24-bit width, composed of two 12-bit sub-channels.

This shift also optimizes physical pin usage. In LPDDR5/5X, interface pins like Clock (CK) and Chip Select (CS) were duplicated for each channel. In LPDDR6, a single CK pair and non-command signals (like RESET) can be shared across sub-channels, effectively reducing the overall pin count and routing complexity.

Burst Length and Effective Bandwidth

To maximize these speed gains, LPDDR6 introduces a new Burst Length of 24 (BL24), a step up from the BL16 standard in LPDDR5. By combining these higher data rates with the wider 24-bit channel, LPDDR6 is expected to deliver:

- **Initial targets:** 28.5 GBps effective bandwidth.
- **Defined:** Up to 38.4 GBps (at 14.4 Gbps), doubling the performance of LPDDR5X-9600.

Wide Non-Return-to-Zero (NRZ) (Staying with 2-Level Signaling)

Rather than moving to multi-level signaling such as PAM4 (Pulse-Amplitude Modulation, 4-levels), LPDDR6 scales bandwidth while staying with the 2-Level NRZ signals. This preserves larger voltage margin and simpler receiver behavior, helping keep I/O power and signal integrity risk under control at higher per-pin rates.

LPDDR6 then gains throughput by going wider: two 12-bit sub-channels per die (24 DQ total) and a BL24 minimum transfer. Together, NRZ + wider parallelism delivers higher effective bandwidth without relying on modulation complexity.

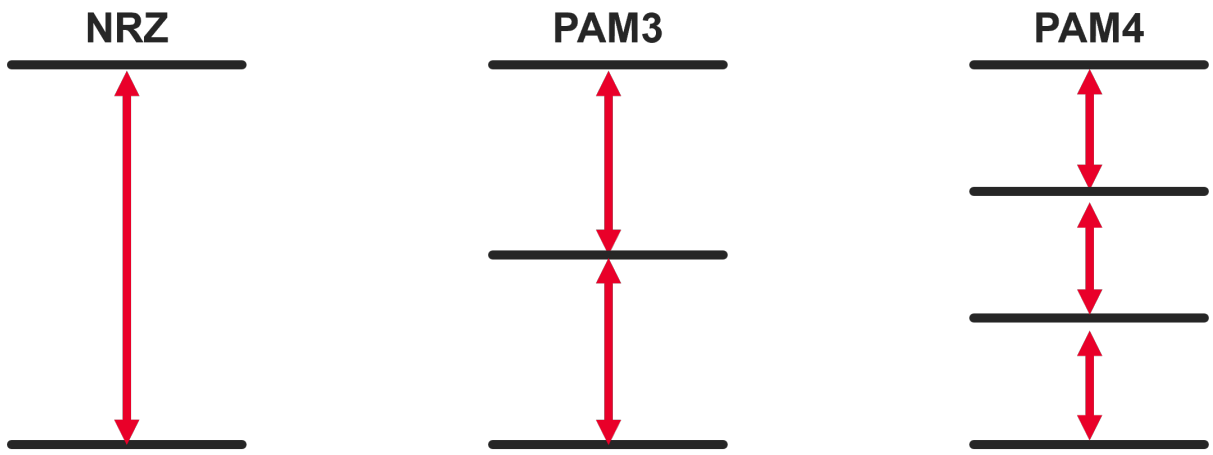


Figure 3. NRZ, PAM3, and PAM4 signaling compared by amplitude levels and eye spacing.

Reduced Sideband Pins and Metadata Integration

In previous generations, these functions required separate physical pins (often called Data mask inversion, DMI, pins). LPDDR6 eliminates these dedicated pins by embedding DBI, masking, and metadata directly into the data transfer itself.

While the elimination of DMI pins simplifies PCB routing and reduces pin count, it introduces a small amount of overhead. For a BL24 access, the system transfers 288 total bits to deliver 256 bits of user data and 32 bits of overhead (metadata, masking, ECC/EDC, etc.). This results in a payload efficiency of 88.9% (256/288), a factor that must be accounted for when calculating effective bandwidth from raw data rates.

Key Specification	LPDDR5X	LPDDR6
Channel width (DQ)	x16 total per channel (per device die)	2 sub-channels × 12 DQ = x24 total per channel (per device die)
Burst length (BL)	16 / 32	24 / 48
BT (bits transferred on DQ per minimum burst)	256 (BL16) / 512 (BL32)	288 (12-bit sub-ch × BL24) (256 DQ*BL + 32 Meta/DBI/ECC/EDC)
User data bits per minimum burst	256 (BL16) / 512 (BL32)	256 (per BL24 access)
Payload efficiency (user/BT)	256/256 = 100% (for BL16) 32 bits of overhead travel over dedicated physical pins (DMI)	256/288 = 88.9% (for BL24) 32 bits of overhead (meta, masking, ECC/EDC, etc.) are integrated directly into the data packets
Bandwidth per channel (GBps)	• Data rate 8,533 Mbps * 2 Bytes = 17.1 GBps • Data rate 9,600 Mbps * 2 Bytes = 19.2 GBps (for BL16)	• Data rate 10,667 Mbps * 3 Bytes * 89% = 28.5 GBps • Data rate 14,400 Mbps * 3 Bytes * 89% = 38.4 GBps (for BL24)

Table 1. LPDDR6 performance targets. >10 Gbps per pin, 24-bit channels (2×12), ~28.5 GB/s initially, up to ~38.4 GB/s at 14.4 Gbps.

Power (Lowering Power Usage)

Over the past two decades, LPDDR has steadily reduced operating voltage and evolved its power architecture. Across generations, the nominal core supply has moved from roughly 1.8 V toward ~1.0 V, and LPDDR5/5X introduced split VDD2 rails (VDD2H/VDD2L) to enable lower-voltage operation at lower speeds. LPDDR6 extends this approach with a dual-rail scheme (often described as VDD2C ~1.0 V and VDD2D ~0.875 V) to target ~20–30% lower operating power versus LPDDR5X.

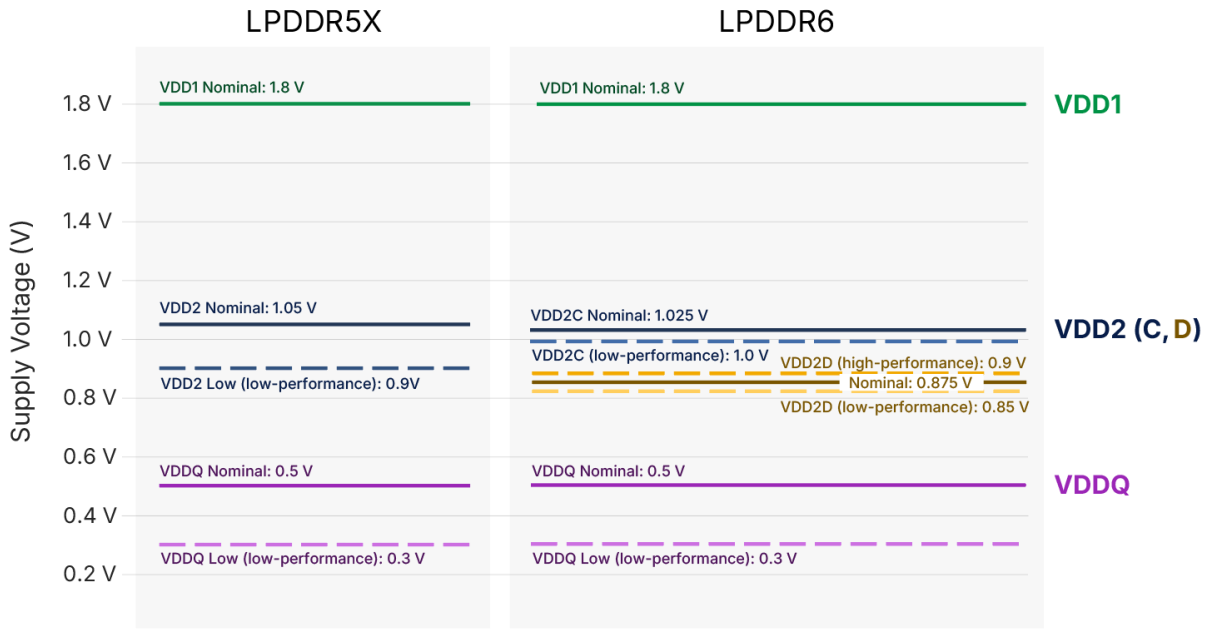


Figure 4. LPDDR5X → LPDDR6 rails & DVFS. Moves from VDD2H/VDD2L to dual rails (~1.0 V/~0.875 V) with more Dynamic Voltage Frequency Scaling (DVFS) operating performance points for lower operating power.

While the dual-rail DVFS scheme primarily targets core and interface supply power, LPDDR6 also reduces I/O energy per bit through signaling optimizations, most notably Enhanced Write Data Bus Inversion (DBI). DBI is an I/O signaling technique used in LPDDR memory to reduce power consumption by keeping more zeros than ones in the data stream, leveraging reduced termination power when a low level is driven in a terminated interface (Low-Voltage Swing Terminated Logic, LVSTL, introduced in earlier generations). In essence, it inverts the data being transmitted on the bus if there are more 1s than 0s (or vice versa) to minimize signal transitions and thus reduce power consumption. This technique is particularly effective because low-voltage signaling (as used in LPDDR) typically consumes less power when driving a '0' than a '1'. Enhanced Write DBI in LPDDR6 provides a mechanism for the host to select optimal data patterns in addition to LPDDR5's count-based approach. To achieve this, the host periodically computes optimum masks to perform logical operations (exclusive OR - XOR) with data bits based on observed data traffic and configures the mode registers accordingly.

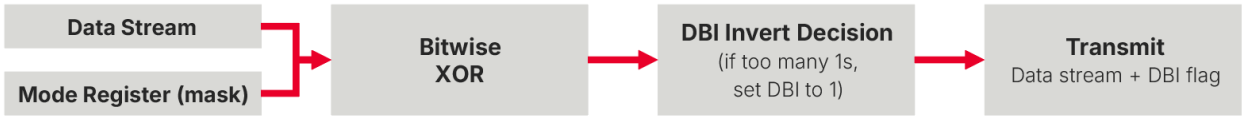


Figure 5. Enhanced Write DBI. Host computes optimal XOR masks, store on mode registers, and inverts/encodes write data to reduce switching/termination power.

Finally, LPDDR6 allows the host to enable Dynamic Efficiency Mode, reducing SDRAM power at the expense of up to half the available data bandwidth, depending on the system's performance-versus-power tradeoff. In this mode, the host can operate using only one sub-channel's I/O to access data across both sub-channels, as illustrated below. This reduces power primarily because fewer I/O lines are switching and being terminated, lowering both dynamic toggle power and driver/termination current, while also reducing associated internal interface activity.

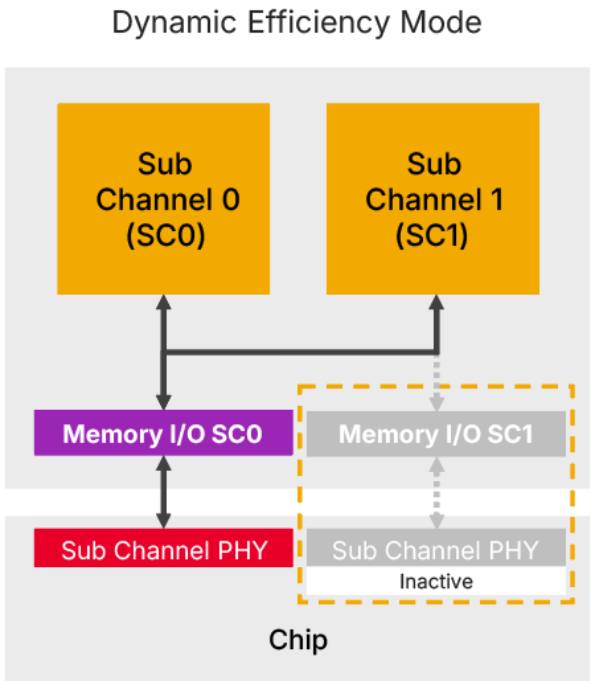


Figure 6. Dynamic Efficiency Mode. One sub-channel I/O can access both, cutting I/O activity and power, with up to ~½ bandwidth tradeoff.

Reliability (Enhanced RAS)

LPDDR6 introduces a new signal, ALERT, which provides a mechanism for LPDDR6 devices to signal faults to the host. The ALERT pin will be High-Z when inactive, so that multiple device ALERT outputs can be connected to a single host input. The device can register up to 15 different faults via four Fault Registers. Faults such as Write ECC/EDC errors can be masked or reported in real time.

Advancing LPDDR6 Validation (Going Beyond Physical Pins)

As LPDDR6 pushes per-pin data rates toward 14.4 Gbps, the transition from theoretical simulation to hardware validation presents new engineering challenges. In addition, with dedicated DMI pins removed, LPDDR6 carries additional data within DQ — simplifying routing and reducing pin count but increasing the complexity of signal integrity analysis.

To address this, the industry has evolved its validation workflow to bridge the gap between pre-silicon design and physical testing:

- **Pre-Silicon Design:** Engineers increasingly model LPDDR6 channel and package configurations. This allows for early evaluation of link performance and potential signal loss before physical hardware is fabricated.
- **Extrapolated Eye-Mask Requirements:** Another evolution in the LPDDR6 specification is the transition to extrapolated eye-mask testing. Unlike previous generations that focused on timing and signal-quality measurements like slew rate and setup/hold timing, LPDDR6 requires an eye-mask analysis extrapolated to a Bit Error Rate (BER) of 1E-16. This ensures long-term reliability in high-speed AI and edge computing environments.
- **Read/Write Data Timing - WCK & RDQS vs DQ Precision:** Increased focus on Validation of the critical timing relationship between the Read/Write Clocks and the Data (DQ) signals. Using advanced compliance software and high-bandwidth oscilloscopes, designers can verify margins and ensure signal integrity at the extreme speeds required for LPDDR6.

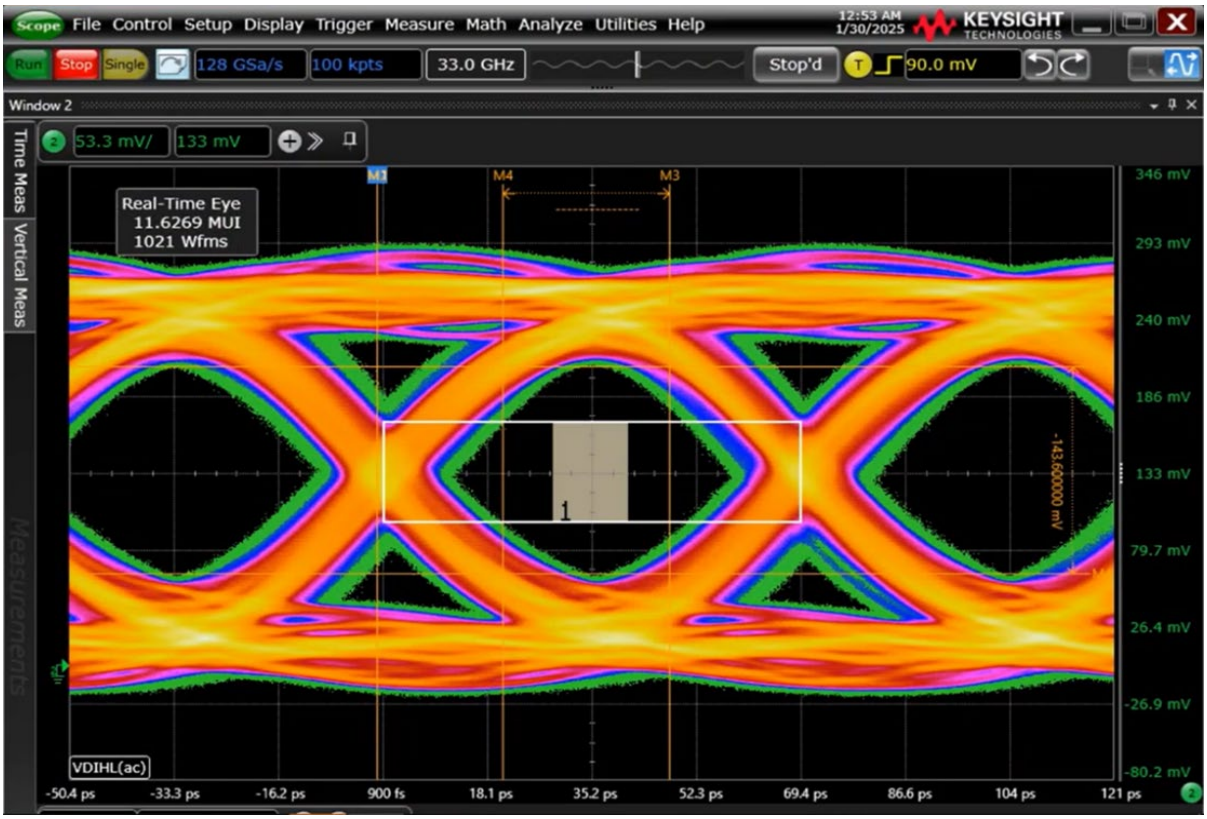


Figure 7. LPDDR6 transmitter compliance results. Real-time eye diagram captured on a Keysight UXR oscilloscope showing eye opening and compliance mask.

Keysight's Role in LPDDR6 Enablement

Test complexity has grown with the adoption of next-generation memory devices such as LPDDR6, HBM4, and GDDR7. These technologies demand advanced test methods to ensure reliability and performance and reducing test times while maintaining accuracy is a constant challenge.

At Keysight, we've been actively involved in shaping and supporting LPDDR6 through our compliance and validation solutions. Keysight offers a complete LPDDR6 design and test solution to support the next technology wave for memory systems. The solution significantly improves device and system validation (both timing and signal integrity), providing new test automation tools necessary for advancing AI, especially in mobile and edge devices.

Keysight's complete workflow solution consists of transmitter and receiver test applications and the Advanced Design System (ADS) Memory Designer workflow solution.

- **Physical Layer - Transmitter Testing:** Keysight D9360LDDC LPDDR6 Tx Compliance Test Software
  - Fully automated compliance testing with extrapolated eye-mask analysis.
  - Reduce validation time with fully automated compliance testing and characterization
  - Capture precise measurements quickly using industry-leading low-noise technology
  - Debug design issues faster with streamlined data analysis tools

- Analyze device BER performance with extrapolated eye mask margin testing
- Achieve accurate signal measurements directly from BGA packages with specialized de-embedding capabilities

- **Physical Layer - Receiver Testing:** Keysight M80896RCA LPDDR6 Receiver Conformance and Characterization Test Application
  - BER testing under stressed eye conditions, including jitter, crosstalk, and voltage swing sweeps
  - Validate designs confidently using proven Bit Error Ratio testing methodology
  - Pinpoint performance issues early by testing against multiple jitter, crosstalk, and noise scenarios
  - Maximize signal integrity through detailed BER analysis and receiver equalization optimization
  - Ensure high interoperability with both device and host controller validation

- **Physical Layer - Probing:** A wide range of probe heads (RC and RCRC) optimized for LPDDR6's fine-pitch, high-speed signaling

- **Physical Layer – Memory Interface System Simulation:** Keysight EDA Memory Designer (W3622B)
  - ADS Memory Designer enables memory system simulation for eye margin, BER, timing, skew, and more, with highest fidelity.
  - ADS Memory Designer LPDDR4/5/6 and other standards, including HBM, DDR, GDDR, and flash memory

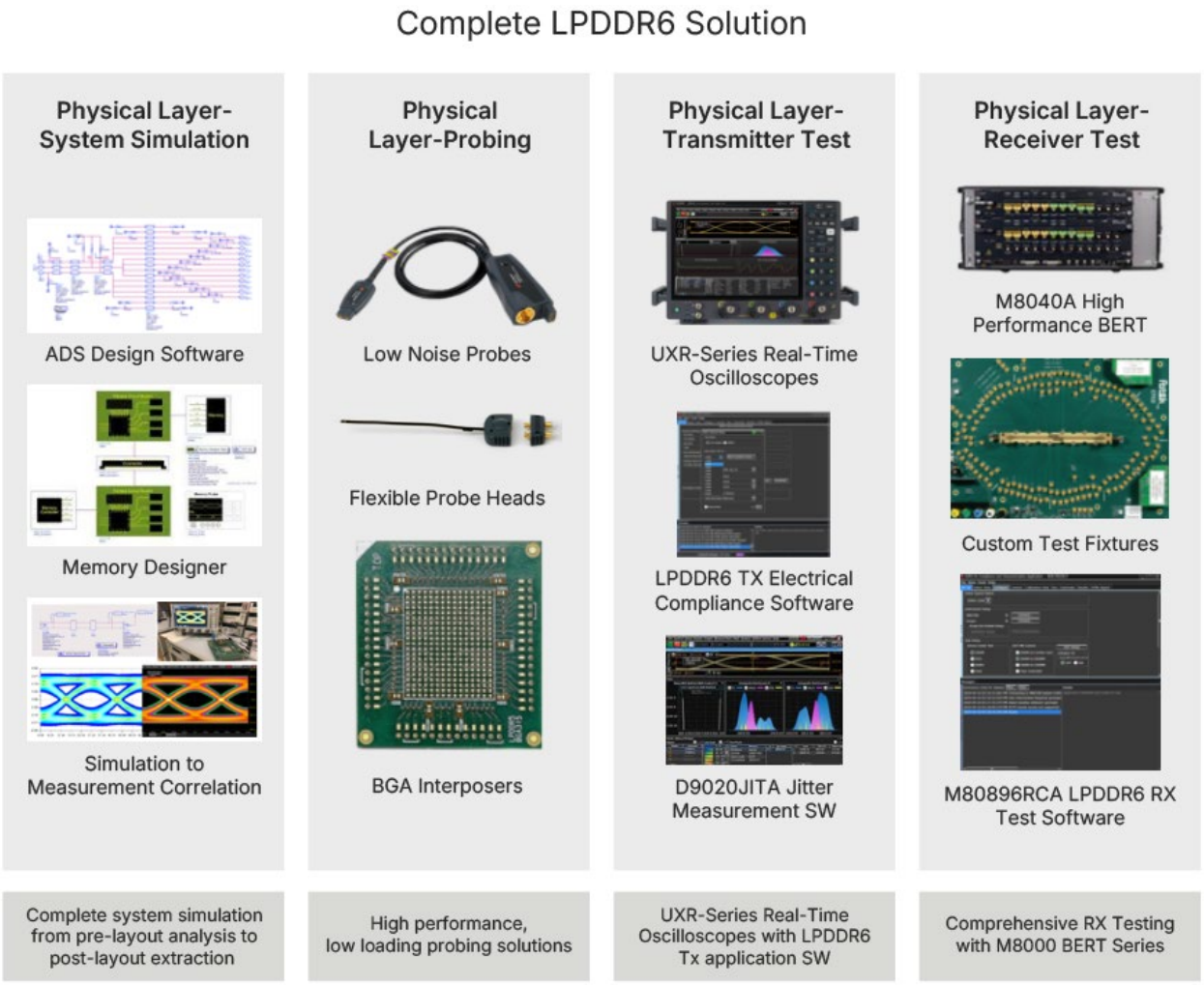


Figure 8. Keysight LPDDR6 workflow. Simulation (ADS), probing, TX compliance (scope + SW), and RX/BER validation (BERT + fixtures + app).

Conclusion

LPDDR6 is not just a faster LPDDR, it is a system-level shift aimed at the exact bottlenecks that show up as AI scales: bandwidth per watt, tighter signal margins, and reliability under real workloads. By pushing per-pin rates beyond 10.6 Gbps, widening the interface with dual 12-bit sub-channels, integrating sideband functions into the data stream, and extending DVFS with dual-rail operation, LPDDR6 changes the performance and power math while raising the bar on validation, including BER-focused eye-mask requirements. The takeaway is simple: if scaling AI is an engineering exercise across the whole stack, LPDDR6 becomes a key lever for improving throughput without paying an additional power and thermal tax, but it demands a modern validation workflow to prove it holds up in silicon and in systems.

Learn more about Keysight's Memory solutions