Flash Memory Summit

PCIe® 7.0 Specification: 128 GT/s bandwidth for Future Data-Intensive Markets

Presenter: Al Yanes, PCI-SIG® President and Board Chair

PCI Express® 7.0 Specification & Status

Flash Memory Summit

PCIe 7.0 specification, version 0.3 is now live for PCI-SIG® members; The full PCIe® 7.0 specification is targeted for release in 2025

- **What does Version 0.3 mean?**
  - The first review draft of the specification is complete and has received work group approval

- **Feature Goals:**
  - Delivering 128 GT/s data rate and up to 512 GB/s bi-directionally via x16 configuration
  - Utilizing PAM4 signaling
  - Defining the channel parameters
  - Continuing to deliver the low-latency and high-reliability targets
  - Improving power efficiency
  - Maintaining backwards compatibility with all previous generations of PCIe technology

Revision	Max Data Rate	Encoding	Signaling
PCIe 7.0 (2025)	128.0 GT/s	1b/1b (Flit Mode*)	PAM4
PCIe 6.0 (2022)	64.0 GT/s	1b/1b (Flit Mode*)	PAM4
PCIe 5.0 (2019)	32.0 GT/s	128b/130b	NRZ
PCIe 4.0 (2017)	16.0 GT/s	128b/130b	NRZ
PCIe 3.0 (2010)	8.0 GT/s	128b/130b	NRZ
PCIe 2.0 (2007)	5.0 GT/s	8b/10b	NRZ
PCIe 1.0 (2003)	2.5 GT/s	8b/10b	NRZ

(*Flit Mode also enabled in other Data Rate with their respective encoding)

PCI Express® 6.0 Specification & Status

PCIe® 6.0 specification was released to members on January 11, 2022

- **Key Features:**
  - 64 GT/s raw data rate and up to 256 GB/s via x16 configuration; doubles the bandwidth and power efficiency from PCIe 5.0 specification (32GT/s)
  - Pulse Amplitude Modulation with 4 levels (PAM4) signaling and leverages existing PAM4 already available in the industry
  - Lightweight Forward Error Correct (FEC) and Cyclic Redundancy Check (CRC) mitigate the bit error rate increase associated with PAM4 signaling
  - Flit (flow control unit) based encoding supports PAM4 modulation and enables more than double the bandwidth gain
  - Updated Packet layout used in Flit Mode to provide additional functionality and simplify processing
  - Maintains backward compatibility with all previous generations of PCIe architecture

- **Adoption:**
  - Various PCIe 6.0 specification enabled products available in the industry; visit our sponsors for more information

- **Compliance Program Progress:**
  - FYI Testing targeted for Q1 2024

PCI-SIG® Specification Progress: At a Glance

- PCIe® specification release and compliance program tracking to 3 years
- Early products are available prior to compliance program dates, but product availability generally tracks with compliance program dates

Specification Revision	Version 1.0 of Specification Completed	Compliance Program Live
PCIe 4.0	October 2017	August 2019
PCIe 5.0	March 2019	April 2022
PCIe 6.0	January 2022	Est. March 2024
PCIe 7.0	Est. for 2025	2027

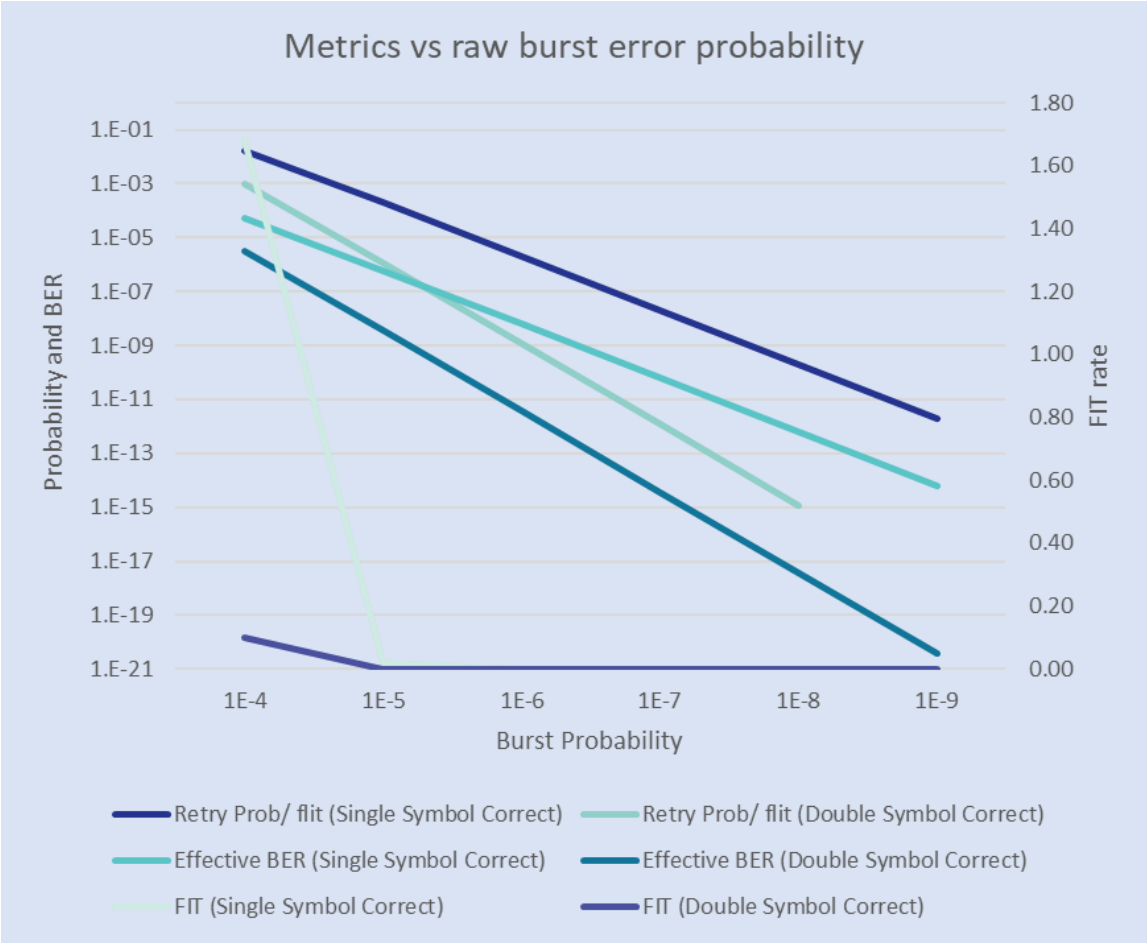
Handling Errors and Metrics Used for Evaluation

- Two mechanisms to correct errors
- Correction through FEC (Forward Error Correction)
- Latency and complexity increases exponentially with the number of Symbols corrected
- Detection of errors by CRC => Link Level Retry (a strength of PCIe)
- Detection is linear: latency, complexity and bandwidth overheads
- Need a robust CRC to keep FIT << 1 (FIT: Failure in Time – No of failures in 109 hours)

- Metrics: Prob of Retry (or b/w loss due to retry) and FIT
- Need to use both means of correction to achieve:
  - Low latency and complexity
  - Retry probability at acceptable level (no noticeable performance impact)
  - Low Bandwidth overhead due to FEC, CRC, and retry
- **Need to keep FEC correction latency low (2ns) to meet the performance needs of Load/Store/IO**

Our Approach: Light-weight FEC and Retry

- Light-weight FEC, strong CRC, and keep the overall latency (including retry) really low so that the Ld/St applications do not suffer latency penalty
- We are better off retrying a packet with 10-6 (or 10-5) probability with a retry latency of 100ns vs having a FEC latency impact of 100ns with a much lower retry probability
- **Low latency mechanism w/ FBER of 1E-6 to meet the metrics (latency, area, power, bandwidth**



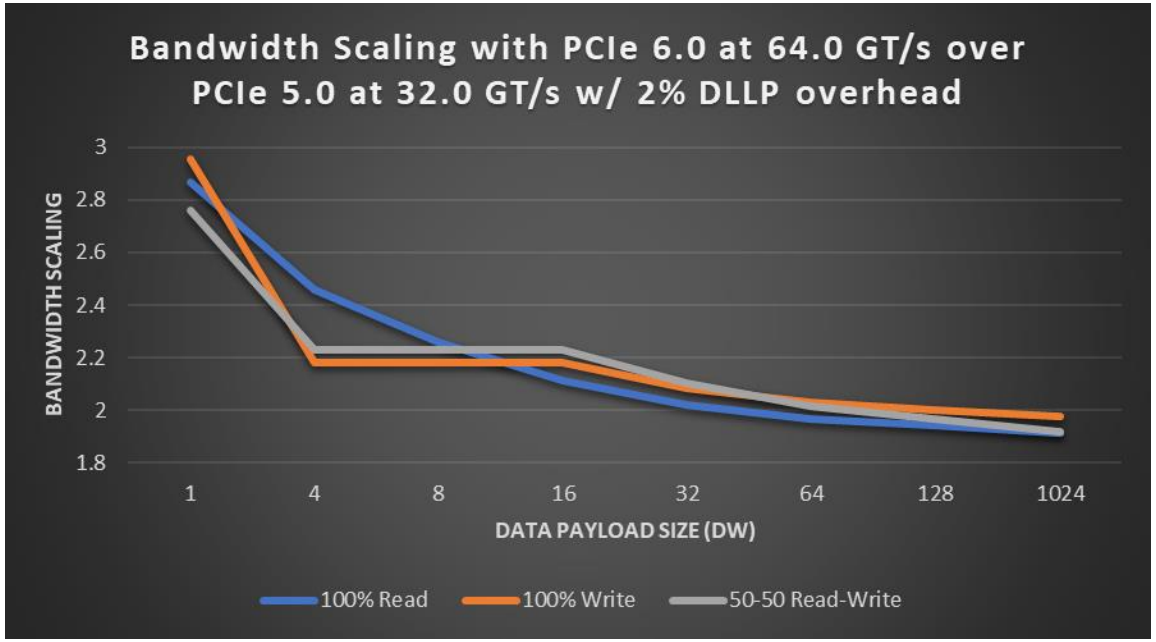
Flit Encoding PCIe® 7.0 and 6.0 Specification: Low-latency w/ High Efficiency

- Flit (flow control unit) based: FEC needs fixed set of bytes
- Correction in flit => CRC (detection) in flits => Retry at flit level
- Lower data rates will also use the same flit once enabled
- Flit size: 256B
  - 236B TLP, 6B DLP, 8B CRC, 6B FEC
  - No Sync hdr, no Framing Token (TLP reformat), no T(DL)LP CRC
  - Improved bandwidth utilization due to overhead amortization
  - Flit Latency: 2ns x16, 4ns x8, 8 ns x4, 16 ns x2, 32 ns x1
  - Guaranteed Ack and credit exchange => low Latency, low storage
- Optimization: Retry error flit only with existing Go-Back-N retry
- **Low latency improves performance and reduces area**

x8 Lanes	0	1	2	3	4	5	6	7
256 UI								
TLP Bytes	0	1	2	3	4	5	6	7
(0-299)	8	9	10	11	12	13	14	15
	16	17	18	19	20	21	22	23
	24	25	26	27	28	29	30	31
	32	33	34	35	36	37	38	39
	40	41	42	43	44	45	46	47
	48	49	50	51	52	53	54	55
	56	57	58	59	60	61	62	63
	64	65	66	67	68	69	70	71
	72	73	74	75	76	77	78	79
	80	81	82	83	84	85	86	87
	88	89	90	91	92	93	94	95
	96	97	98	99	100	101	102	103
	104	105	106	107	108	109	110	111
	112	113	114	115	116	117	118	119
	120	121	122	123	124	125	126	127
	128	129	130	131	132	133	134	135
	136	137	138	139	140	141	142	143
	144	145	146	147	148	149	150	151
	152	153	154	155	156	157	158	159
	160	161	162	163	164	165	166	167
	168	169	170	171	172	173	174	175
	176	177	178	179	180	181	182	183
	184	185	186	187	188	189	190	191
	192	193	194	195	196	197	198	199
	200	201	202	203	204	205	206	207
	208	209	210	211	212	213	214	215
	216	217	218	219	220	221	222	223
	224	225	226	227	228	229	230	231
	232	233	234	235	dlp0	dlp1	dlp2	dlp3
	dlp4	dlp5	crc0	crc1	crc2	crc3	crc4	crc5
	crc6	crc7	ecc0	ecc0	ecc0	ecc1	ecc1	ecc1

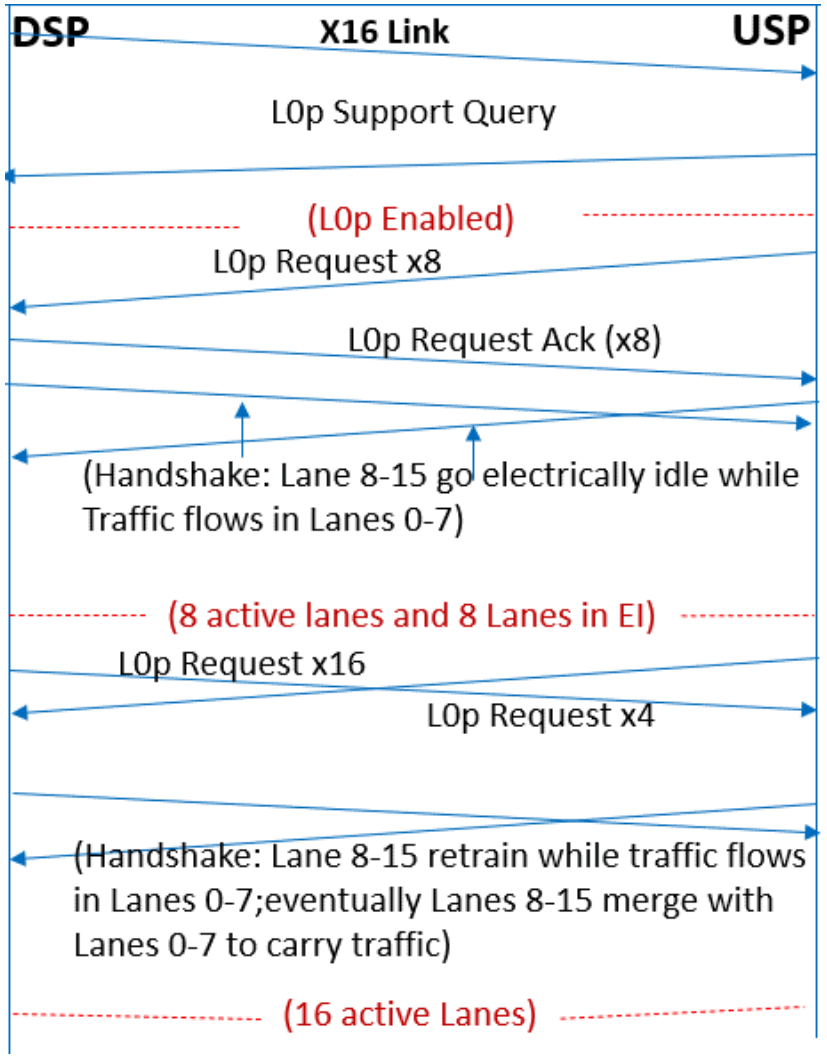
PCIe® 7.0 and 6.0 Specification Flit Mode Bandwidth at 64.0 GT/s

- Bandwidth increase = 2X (BW efficiency of flit mode) / (BW efficiency in non-flit mode)
- Overall we see a >2X improvement in bandwidth (benefits most systems)
  - Efficiency gain reduces as TLP size increases
  - Beyond 512 B (128 DW) payload goes below 1
- Bandwidth efficiency improvement in flit mode due to the amortization of CRC, DLP, and ECC over a flit (8% overhead) – works out better than sync hdr, DLLP, Framing Token per TLP, and 4B CRC per TLP overheads in PCIe 5.0
- **Bandwidth Efficiency improvement causes > 2X bandwidth gain for up to 512B Payload in 64.0 GT/s flit mode**

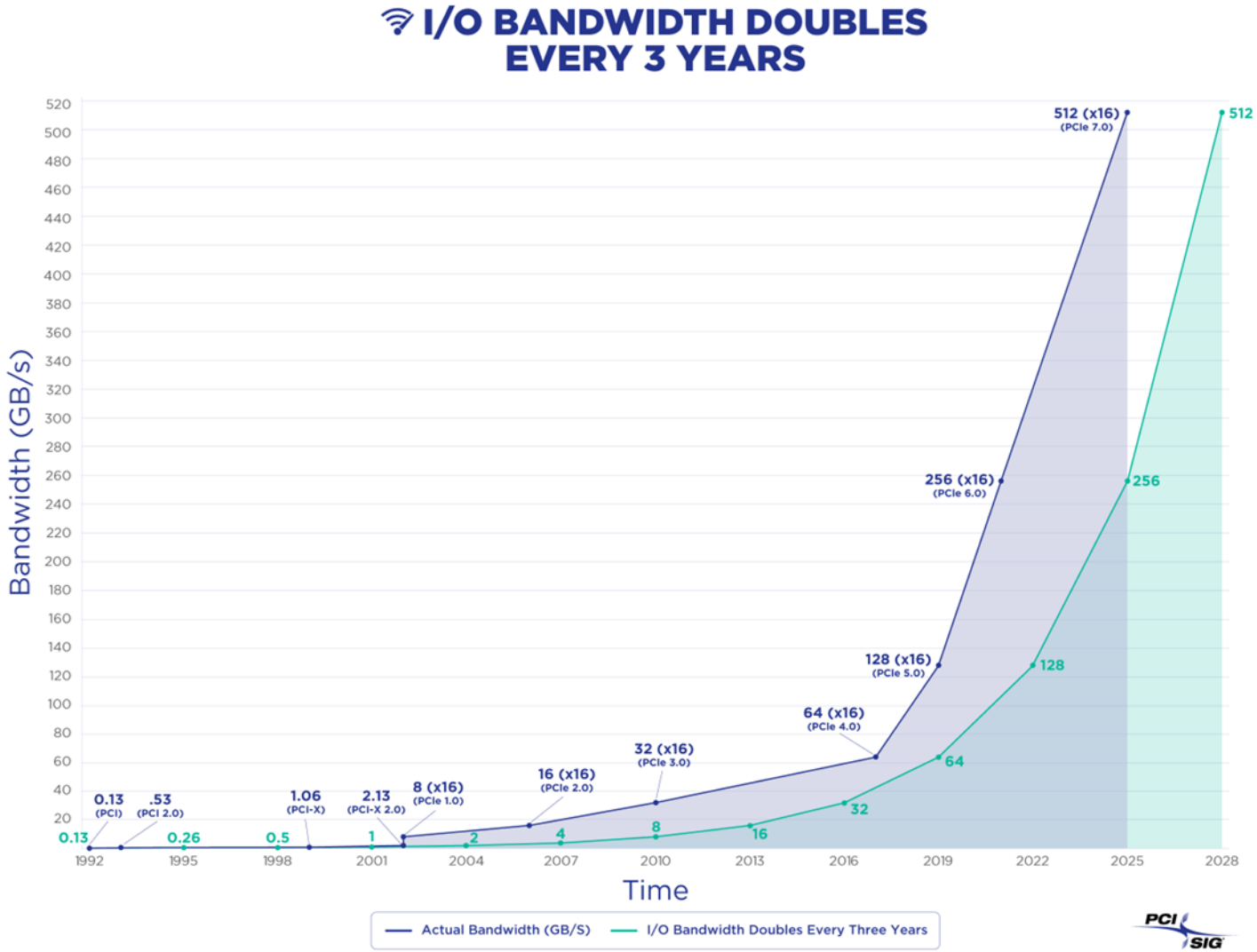


Motivation for a New Low Power State

- Existing low-power states: L0s, L1, Dynamic Link Width (DLW), Speed Change
  - Served well for the set of usages so far and will continue
- Increasingly there is demand for power consumption scaling with bandwidth usage without impacting traffic flow
- Solution: New state L0p – symmetric
  - Maintain at least one active Lane – they continue to carry traffic. Link still carries traffic during L0p width transition
  - Expect L0p PHY power savings similar to turning off power for the idle Lanes
- L0p enables power consumption proportionate to bandwidth usage without interrupting traffic flow



PCI-SIG® Roadmap



PCIe® Speeds/Feeds - Pick Your Bandwidth

- **Flexible to meet needs from handheld/client to server/HPC**
- **~Max Total Bandwidth = Max RX bandwidth + Max TX bandwidth**
- **35 Permutations yielding 11 unique bandwidth profiles**
- **Encoding overhead and header efficiency not included**

Specifications	Lanes				
	x1	x2	x4	x8	x16
2.5 GT/s (PCIe 1.x +)	500 MB/S	1 GB/S	2 GB/S	4 GB/S	8 GB/S
5.0 GT/s (PCIe 2.x +)	1 GB/S	2 GB/S	4 GB/S	8 GB/S	16 GB/S
8.0 GT/s (PCIe 3.x +)	2 GB/S	4 GB/S	8 GB/S	16 GB/S	32 GB/S
16.0 GT/s (PCIe 4.x +)	4 GB/S	8 GB/S	16 GB/S	32 GB/S	64 GB/S
32.0 GT/s (PCIe 5.x +)	8 GB/S	16 GB/S	32 GB/S	64 GB/S	128 GB/S
64.0 GT/s (PCIe 6.x +)	16 GB/S	32 GB/S	64 GB/S	128 GB/S	256 GB/S
128.0 GT/s (PCIe 7.x +)	32 GB/S	64 GB/S	128 GB/S	256 GB/S	512 GB/S

+ = data rate supported by this and subsequent spec revisions.

PCIe® Architecture: One Interconnect – Infinite Opportunities

Flash Memory Summit

HPC / Cloud

Data Center / Enterprise Servers

Artificial Intelligence / Machine Learning

Automotive

Internet of Things

Military / Aerospace

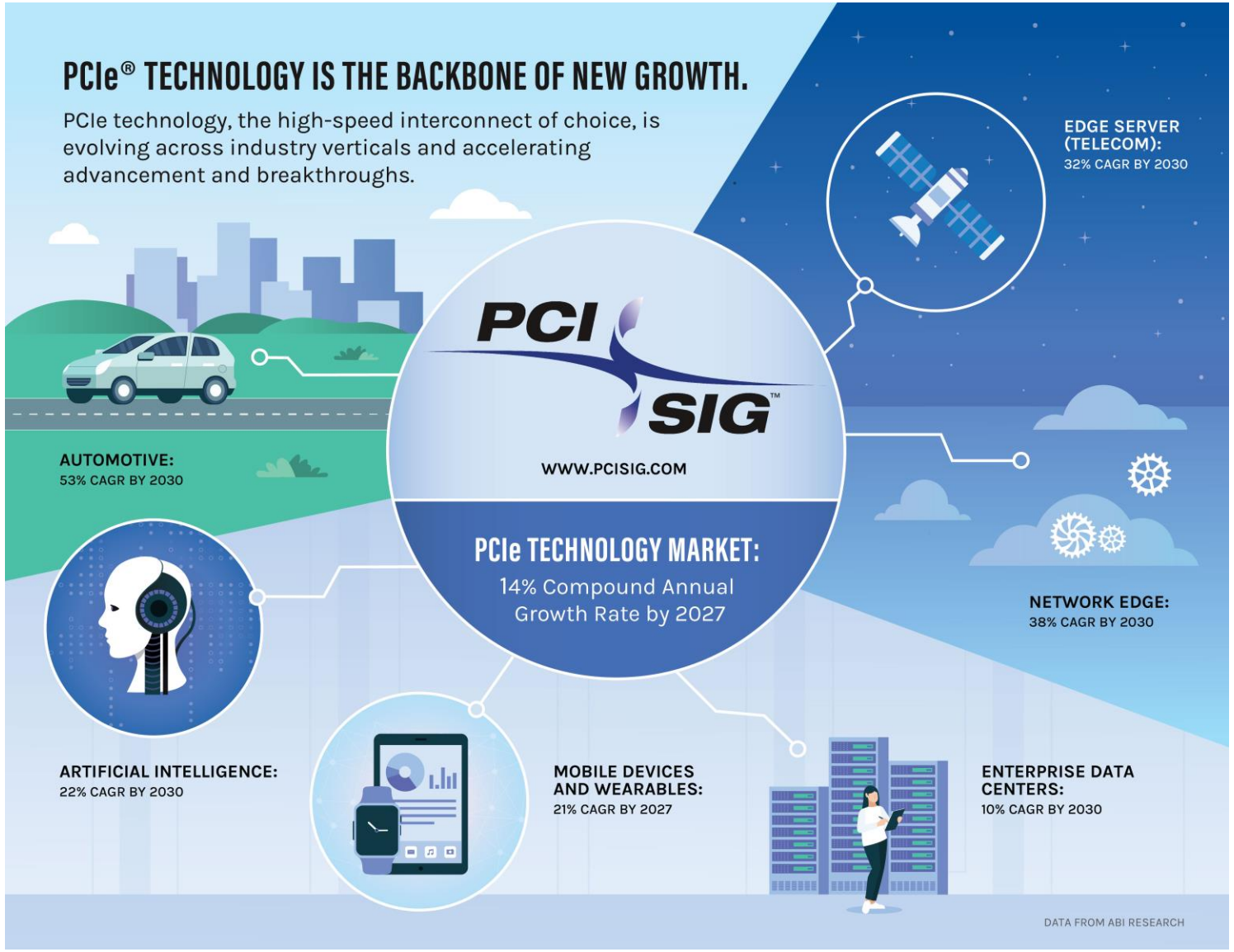
Storage

PCIe® Technology Market Data Forecasted Through 2027

Key Highlights from the Report Developed by ABI Research:

- **Automotive and network edge** verticals offer the highest growth opportunities for PCIe technology, with total addressable market (TAM) compound **annual growth rates (CAGRs) of 53% and 38**
- High-performance applications such as **data centers will contribute to sustained long-run demand** for new PCIe generation deployment. Performance is not the only driver of PCIe technology adoption, as verticals increasingly look at **power efficiency, security and "time-to-value" as crucial issues**
- **AI industry adoption will be high**, as PCIe technology offers decision makers agility through forwards and backwards compatibility, **improving time-to-value and lowering deployment risk**
- PCIe technology will **perform well in the mobile devices vertical**, as the quick pace of market innovation will **necessitate a discrete component interconnect**
- **The PCIe 6.0 specification low power feature (L0p) will be a major driver of deployment**, as power efficiency becomes a central strategy for adopters with a closer **focus on sustainability and lowering operational costs**

PCIe® Technology Market Forecast



Q & A