



HYPERION RESEARCH

Hyperion Research Market Update

May 2026

www.HyperionResearch.com
www.hpcuserforum.com

**Earl Joseph, Bob Sorensen, Mark Nossokoff,
Tom Sorensen, and Jaclyn Ludema**

Hyperion Research Team

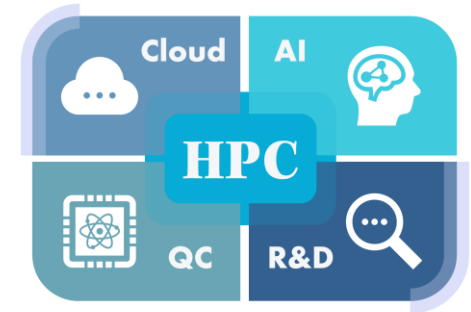
Earl Joseph	Bob Sorensen	Mark Nossokoff	Jaclyn Ludema	Thomas Sorensen	Mike Thorp	Jean Sorensen
<ul style="list-style-type: none">•CEO Hyperion Research•Executive Director HPC/AI User Forum	<ul style="list-style-type: none">•Senior Vice President of Research•Chief Quantum Computing Analyst•Lead AI Analyst	<ul style="list-style-type: none">•Research Director•Chief Storage Analyst•Lead Cloud Analyst	<ul style="list-style-type: none">•Lead Sustainability Analyst•Cloud Analyst	<ul style="list-style-type: none">•Principal Analyst AI/HPC	<ul style="list-style-type: none">•Senior Global Account Executive	<ul style="list-style-type: none">•Chief Operating Officer•HR Director



Example Research Areas

(www.HyperionResearch.com & www.HPCUserForum.com)

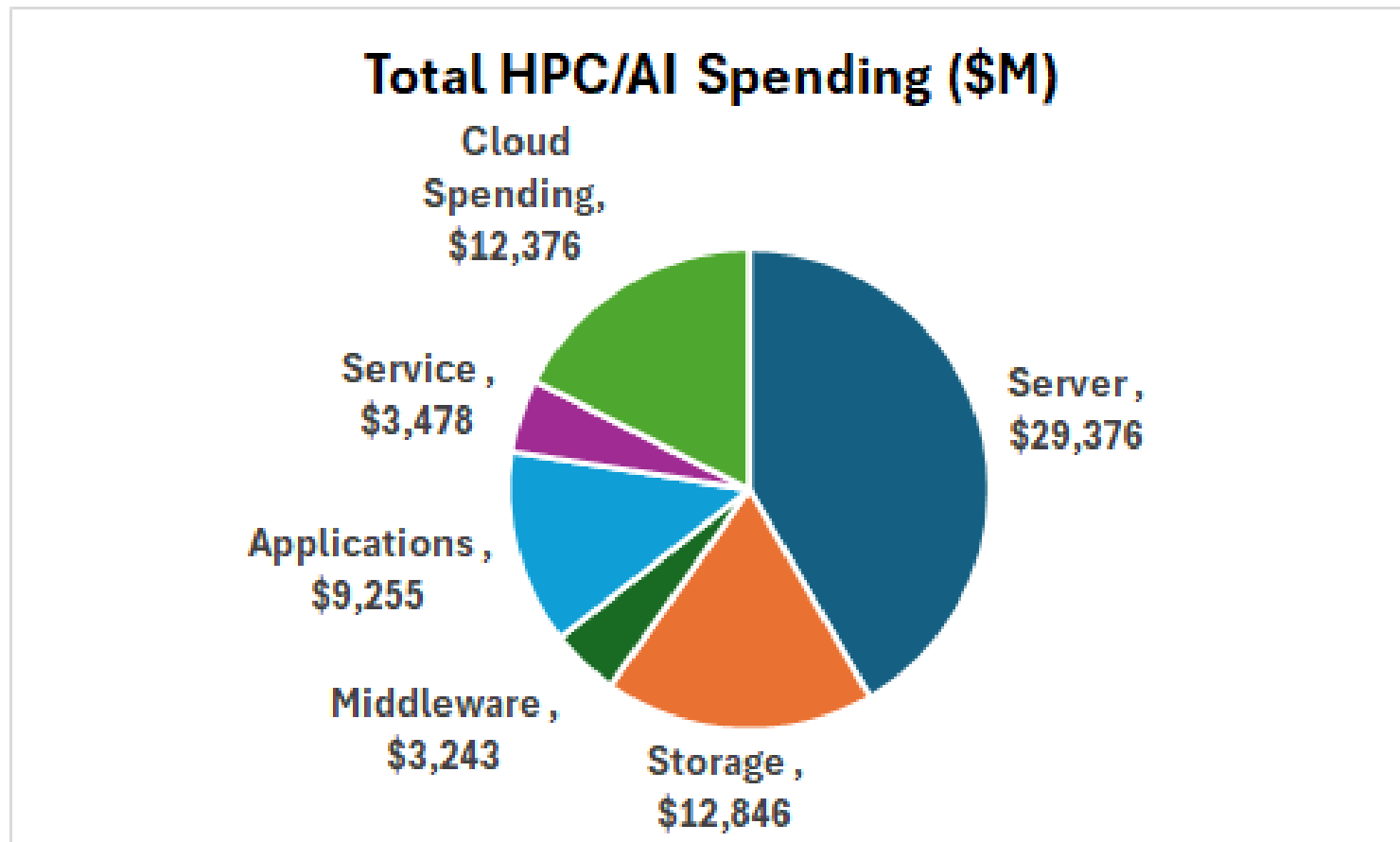
- **Traditional HPC**
- **AI for Science**
- **Cloud Computing**
- **Quantum Computing**
- **Storage & Data**
- **Interconnects**
- **Software & Applications**
- **The ROI and Scientific Returns from HPC**
- **Power & Cooling**
- **Tracking Processor Types & Growth Rates**
- **R&D and Engineering Success Stories**
- **Supply Chain Issues**
- **Sustainability**



HPC/AI Market Update

2025 Was a Strong Growth Year

16.9% overall growth in spending!



- **15.0% growth in on-premises servers**
- **29.7% growth in the use of clouds**
- **Over \$70 billion in total spending**

2025 HPC/AI Market By Vendor

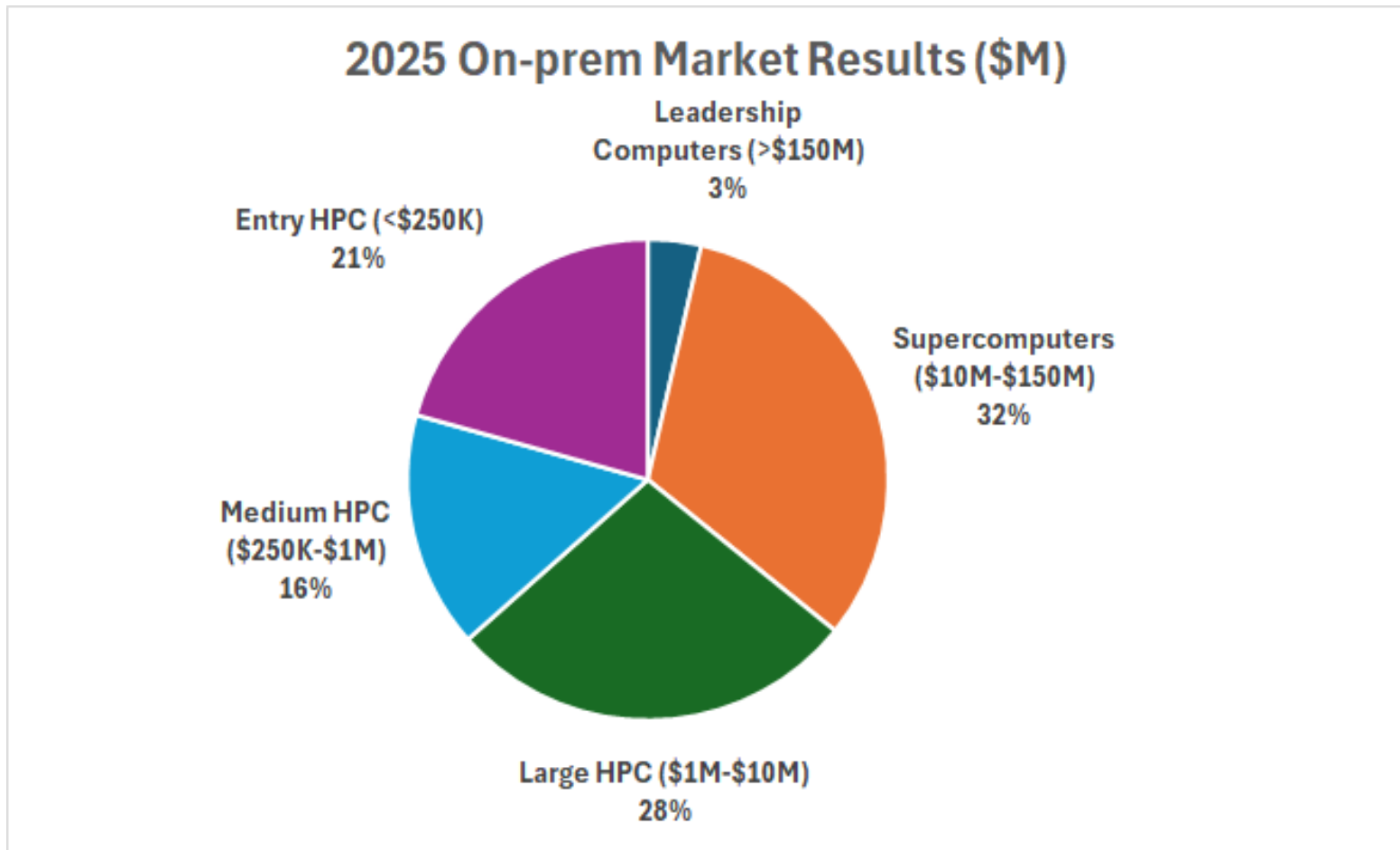
On-prem sever sales were just under \$30 billion in 2026

2025 HPC/AI On-Prem Sales		
Vendor	2025 Sales	2025 Share
HPE	\$6,485	22.1%
Dell Technologies	\$5,342	18.2%
Lenovo	\$1,783	6.1%
Inspur	\$1,236	4.2%
Sugon	\$693	2.4%
Atos	\$464	1.6%
IBM	\$453	1.5%
Penguin	\$508	1.7%
Fujitsu	\$259	0.9%
NEC	\$238	0.8%
Other HPC Suppliers	\$2,735	9.3%
Non-Traditional Suppliers	\$9,180	31.2%
Total	\$29,376	100.0%

Source: Hyperion Research, April 2026

2025 HPC/AI Market By Segment

Supercomputers (\$10M-\$150M) is the largest segment at \$9.5 billion



2025 HPC/AI Market By Vertical

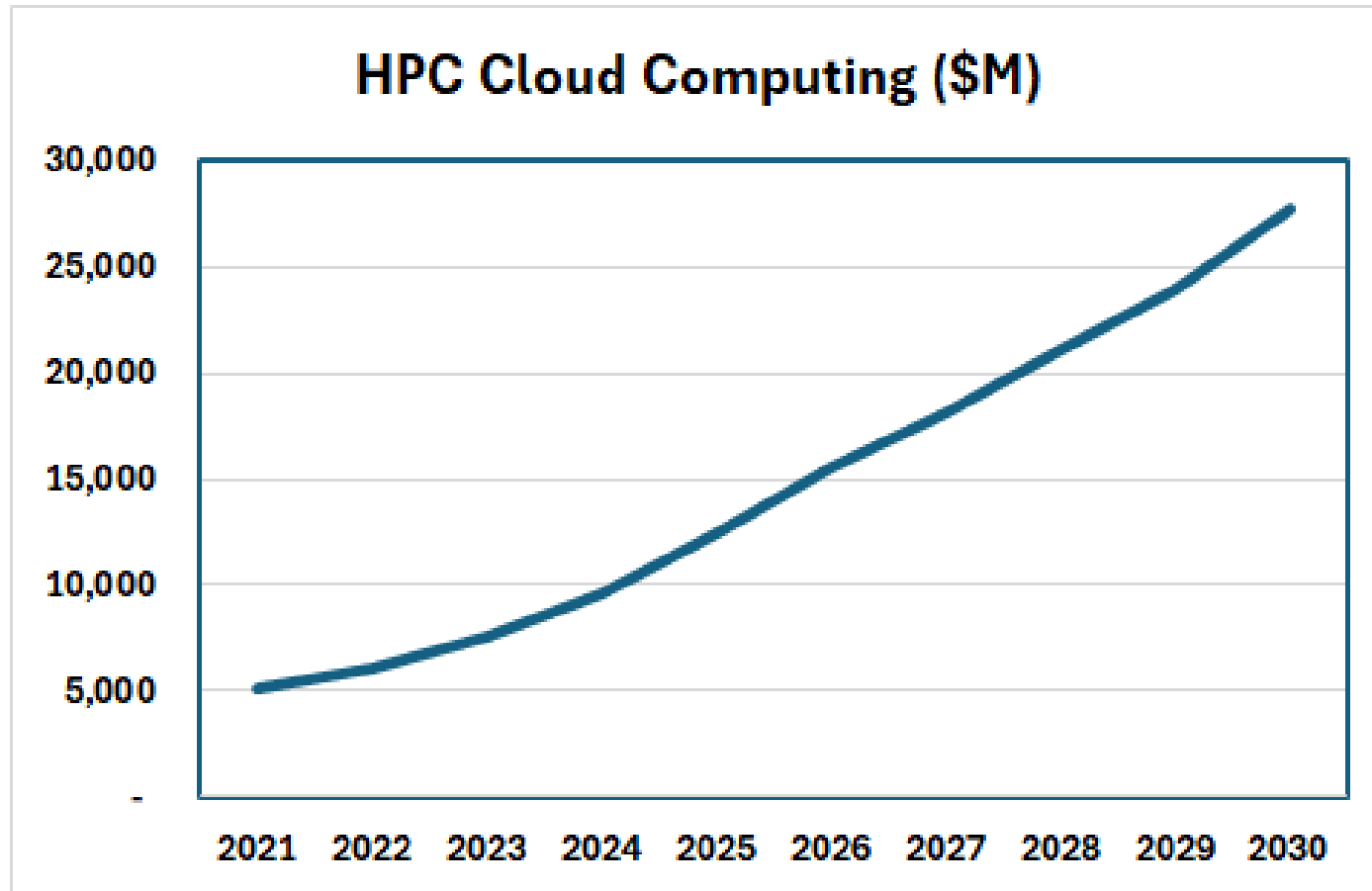
Five verticals are now over \$2 billion, and two are over \$4 billion

HPC/AI On-prem System Installations (\$M)		
	2024	2025
Bio-Sciences	2,302	2,622
CAE	2,751	3,134
Chemical Engineering	305	341
DCC & Distribution	1,384	1,610
Economics/Financial	1,336	1,526
EDA / IT / ISV	1,497	1,701
Geosciences	1,550	1,782
Mechanical Design	28	27
Defense	2,695	3,213
Government Lab	6,115	7,129
University/Academic	4,039	4,542
Weather	1,137	1,288
Other	415	462
Total Revenue	25,554	29,376
<i>Source: Hyperion Research, April 2026</i>		

Cloud Computing for HPC/AI

Projected to reach close to \$30 billion by 2030

Grew close to 30% in 2025



The HPC/AI Market Should See Growth in 2026

... but there are some major concerns

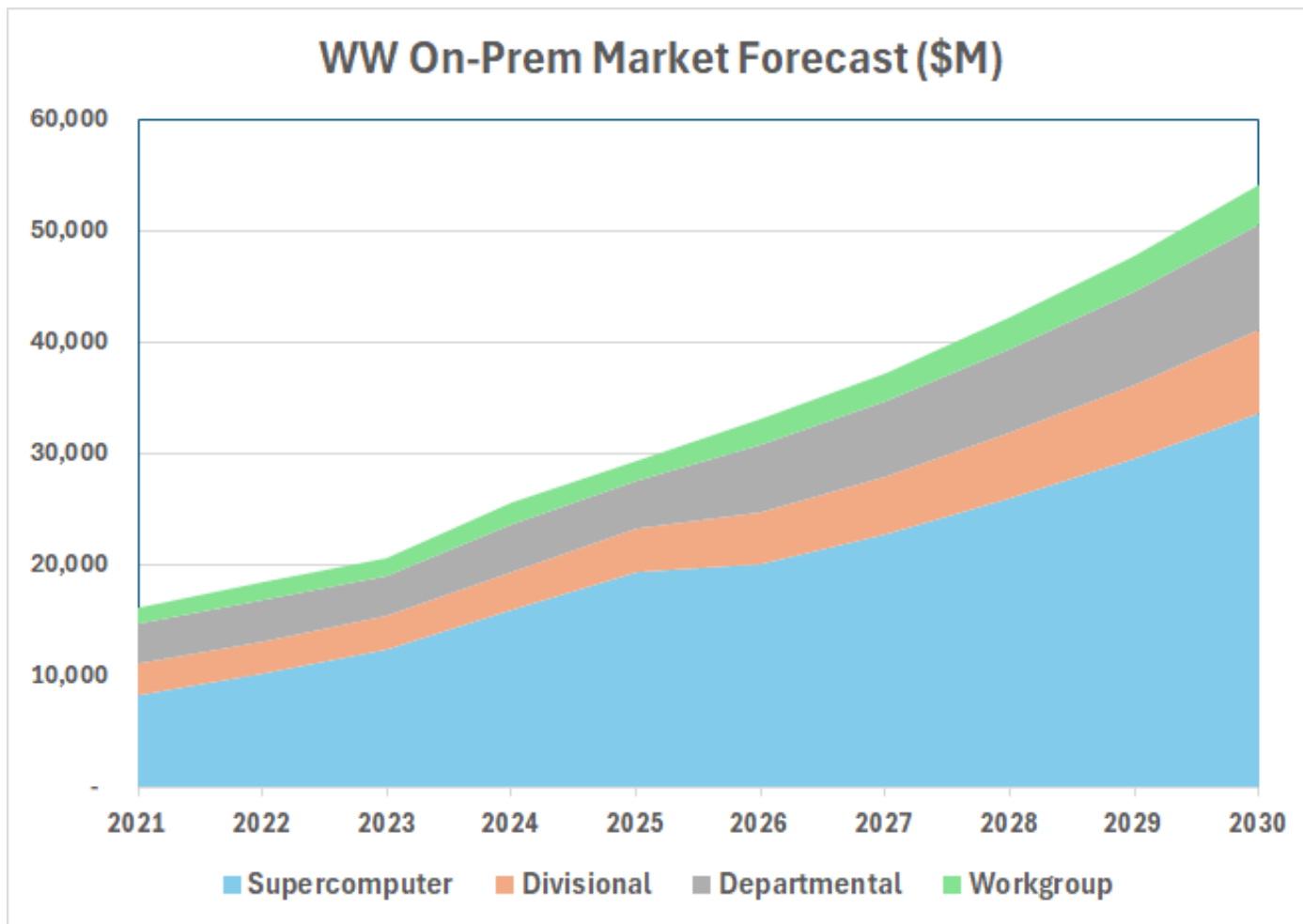
- **The global economic situation and changing trade rules could have a major impact to IT build outs in 2026**
- **Non-science AI systems are getting major attention and \$\$**
- **Supply chain issues are still impacting installations (e.g., GPUs, memory & SSDs)**
- **Exascale system acceptances are seeing delays**
- **The lower end of the on-premises market continues to struggle**

- **Growth drivers include:**
 - **New use cases especially in AI are providing new areas for users to advance their research**
 - **Countries and companies around the world continue to recognize the value of being innovative and investing in R&D to advance society, grow revenues, reduce costs, and become more competitive**

HPC/AI On-prem Server Forecast

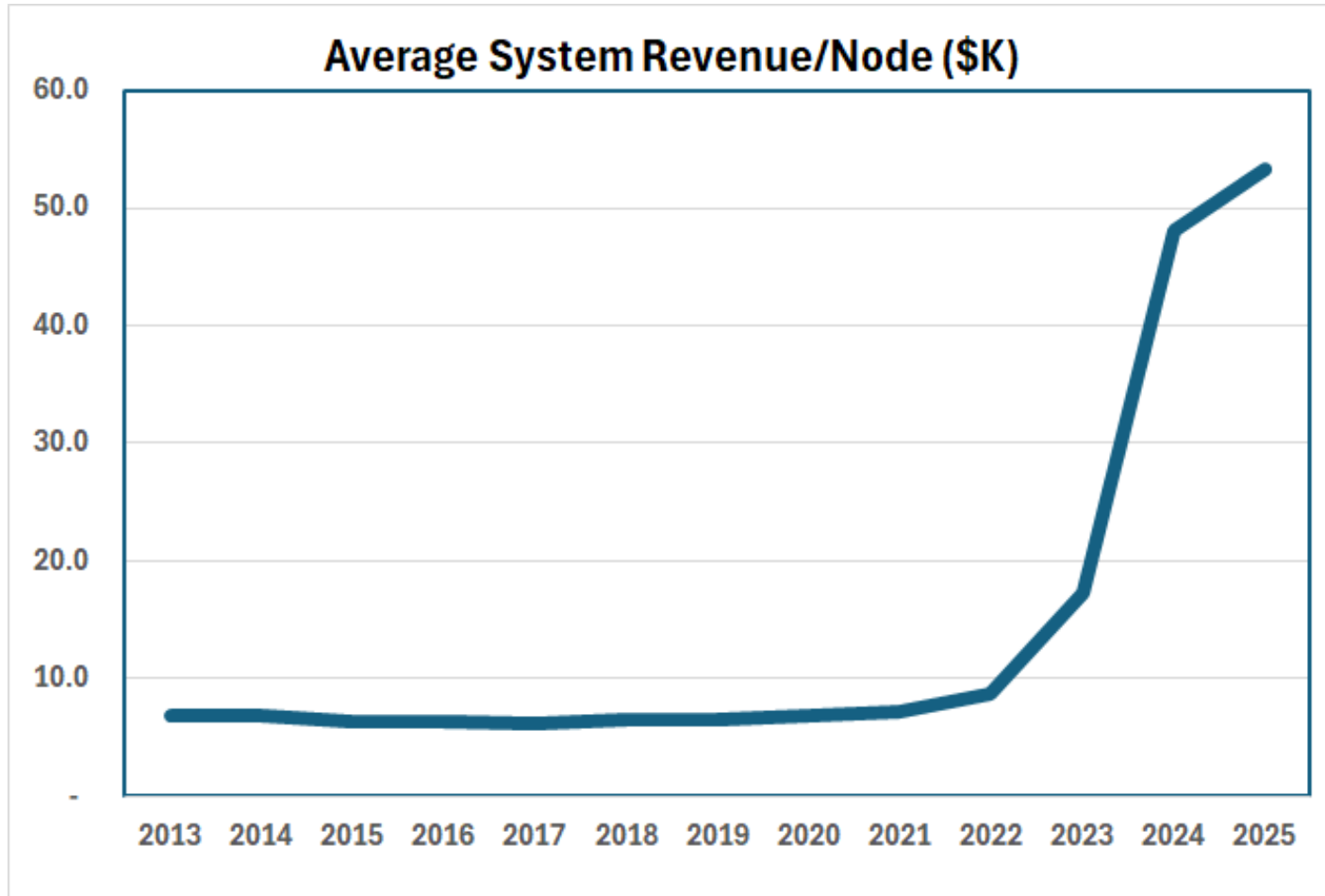
On-prem servers are projected to reach \$54 billion by 2030

Broader HPC/AI on-prem market is projected to exceed \$100 billion



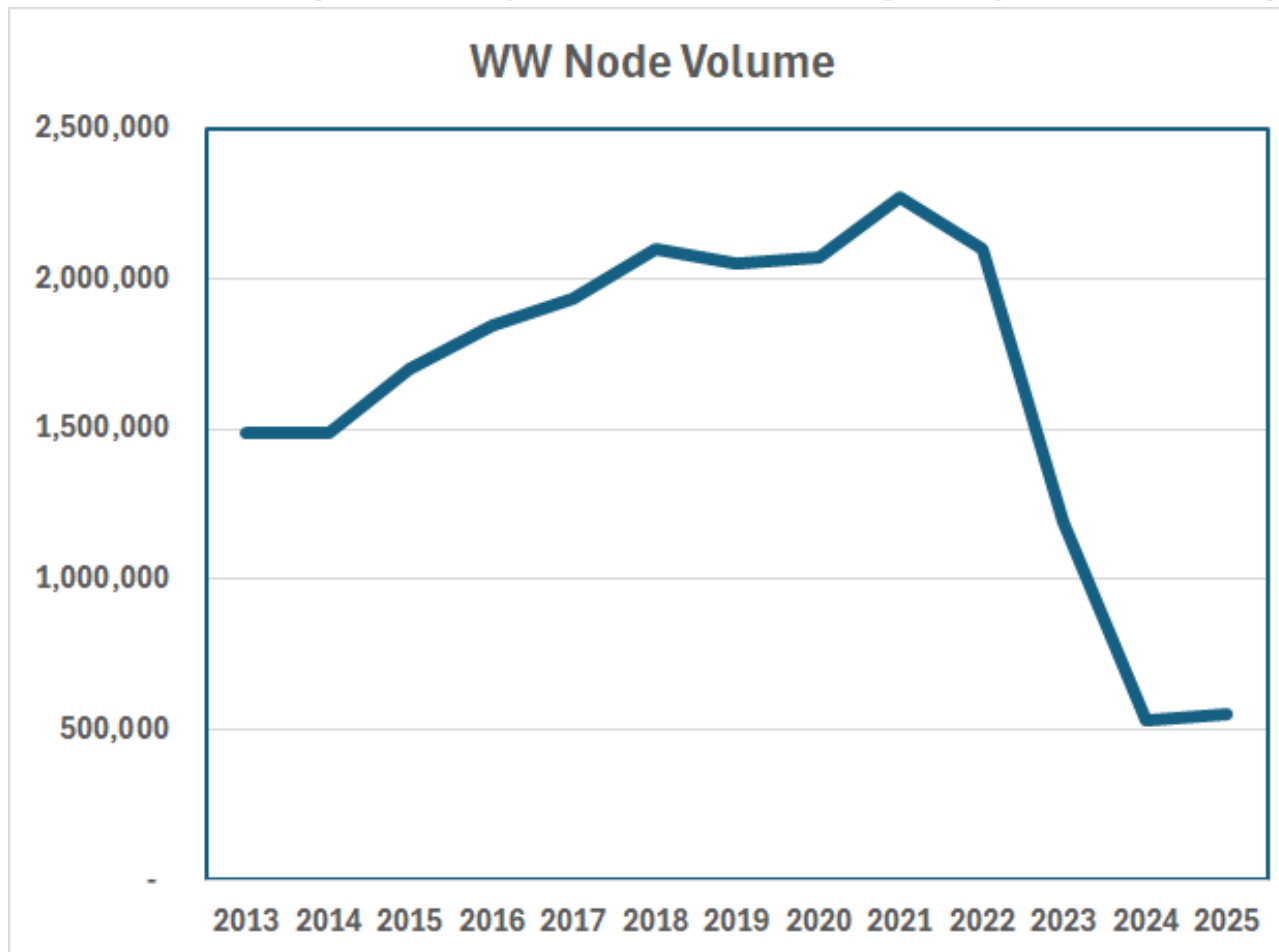
A Major Market Shift: High Node Costs

Driven by more systems having GPUs, higher GPU costs and higher node costs (memory, liquid cooling, higher power)



A Major Market Shift: Fewer Node Purchases

Driven by more systems having GPUs, higher GPU costs and higher node costs (memory, liquid cooling, higher power)



Some Soon To Be Published AI Findings From Our MCS Study

2026 MCS Study Findings: AI Strategies

*94.8% are using or will soon be using AI/LLMs
But 8.2% plan to stop using them*

Characterizing AI/LLM Strategy	
Q. Which of the following best describes your generative AI/LLM strategy?	
Strategy Characterization	Percentage of Sites
We use generative AI/LLMs today and plan to continue using it over the next 12-18 months	71.1%
We do not use generative AI/LLMs today but plan to start in the next 12-18 months	15.5%
We use generative AI/LLMs today but plan to stop using it within the next 12-18 months	8.2%
We do not use generative AI/LLMs today and do not plan to within the next 12-18 months	5.2%
n = 97	
Source: Hyperion Research, 2026	



2026 MCS Study Findings: Processors Planned for AI

69.2% are using or will soon be using GPUs for AI/LLMs workloads

42.4% are using CPUs

Note: most sites use multiple types of processors

Processing Elements to Support AI/HPC	
Q. In the next 12 - 18 months, which of these processing elements do you expect will be incorporated into your HPC/AI/HPDA computing? Please check all that apply:	
Processing Element	Percentage of Sites
GPUs	69.2%
CPUs	42.5%
AI-specific processors	25.0%
TPUs (tensor processing units)	23.3%
FPGAs	10.8%
Neuromorphic processors	5.0%
VPUs (vision processing unit)	3.3%
In-house processors	3.3%
n = 120	
Notes: Respondents could select multiple options	
Source: Hyperion Research, 2026	

2026 MCS Study Findings: Barriers to AI Growth

Expertise, quality of training data, dealing with complexity and high costs are the top barriers

Barriers to AI Capability Growth	
Q. Are any of the following barriers to furthering your AI capabilities? Please select all that apply:	
Barrier	Percentage of Sites
Level of in-house AI expertise	36.7%
Quality of available training data	35.0%
Complexity with integrating AI models into existing HPC workloads	30.0%
High/uncertain development costs	28.3%
Concerns with technical issues (i.e., scaling, hallucinations, ease of maintenance)	21.7%
High/uncertain operational costs	20.8%
Scale of available training data	19.2%
Access to specialized hardware	17.5%
The technology is moving too fast for credible assessment of value	17.5%
Lack of demonstrated return on investment	16.7%
Level of AI vendor or 3rd party expertise	10.8%
Access to specialized software	10.0%
Confusion/uncertainty with vendor selection	8.3%
Other (please specify)	3.3%
n = 120	
Notes: Respondents could select multiple options	
Source: Hyperion Research, 2026	

In Summary

Conclusions

- **2025 was a strong growth year**
 - AI, GPUs, and cloud are high growth areas
 - New QC systems are being installed around the world
- **New technologies are showing up large numbers:**
 - Generative AI, smarter AI, LLMs and SLLs are fueling a new level of growth
 - Processors, AI hardware & software, memories, new storage approaches, etc.
 - The cloud has become a viable option for many HPC workloads
- **Storage will likely see major growth driven by AI, big data and the need for much larger data sets**
- **There are growing concerns around system costs, power, talent and political changes**

We Welcome Questions, Comments and Suggestions

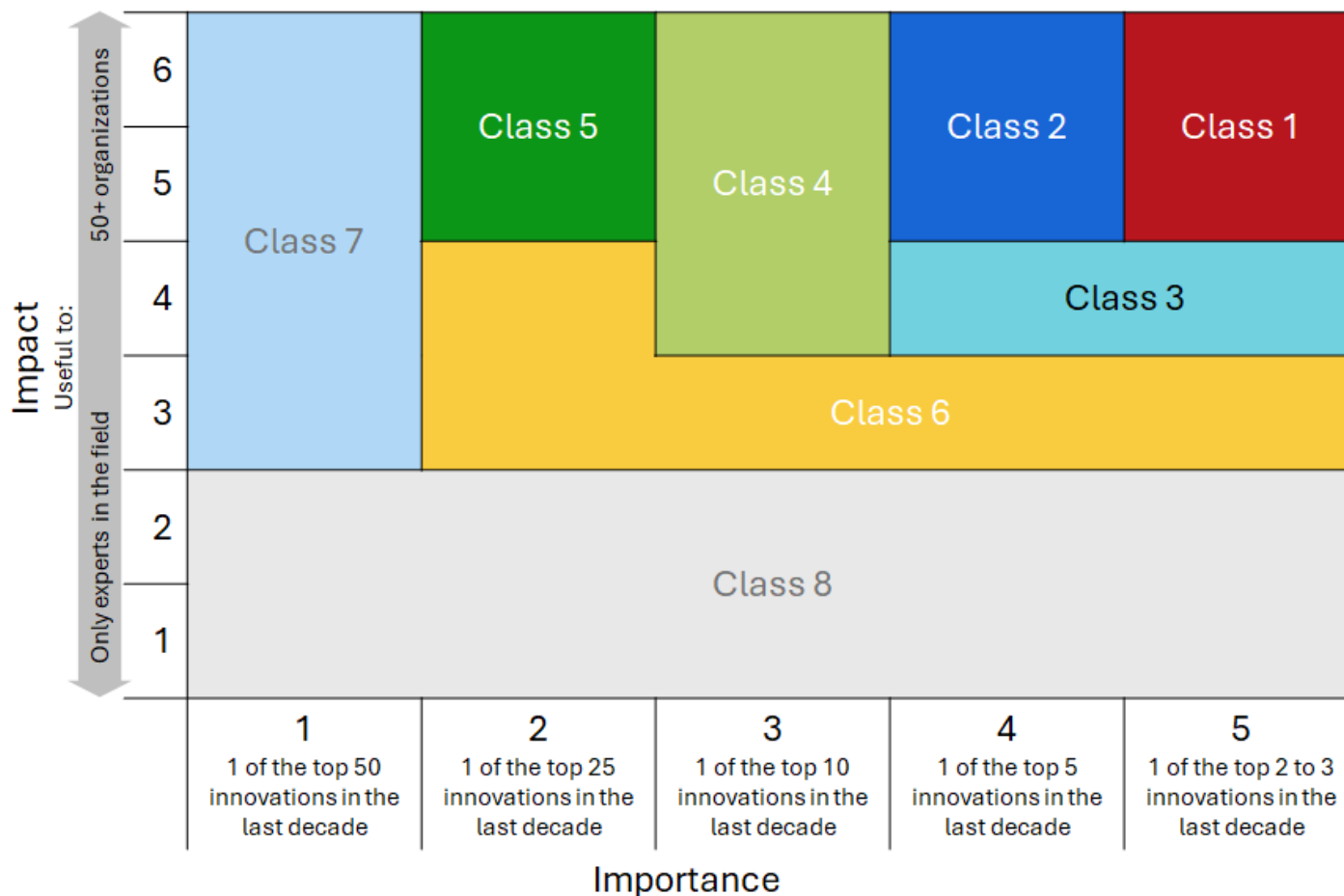


Please contact us at:
info@hyperionres.com

Measuring and Comparing Leadership Computing

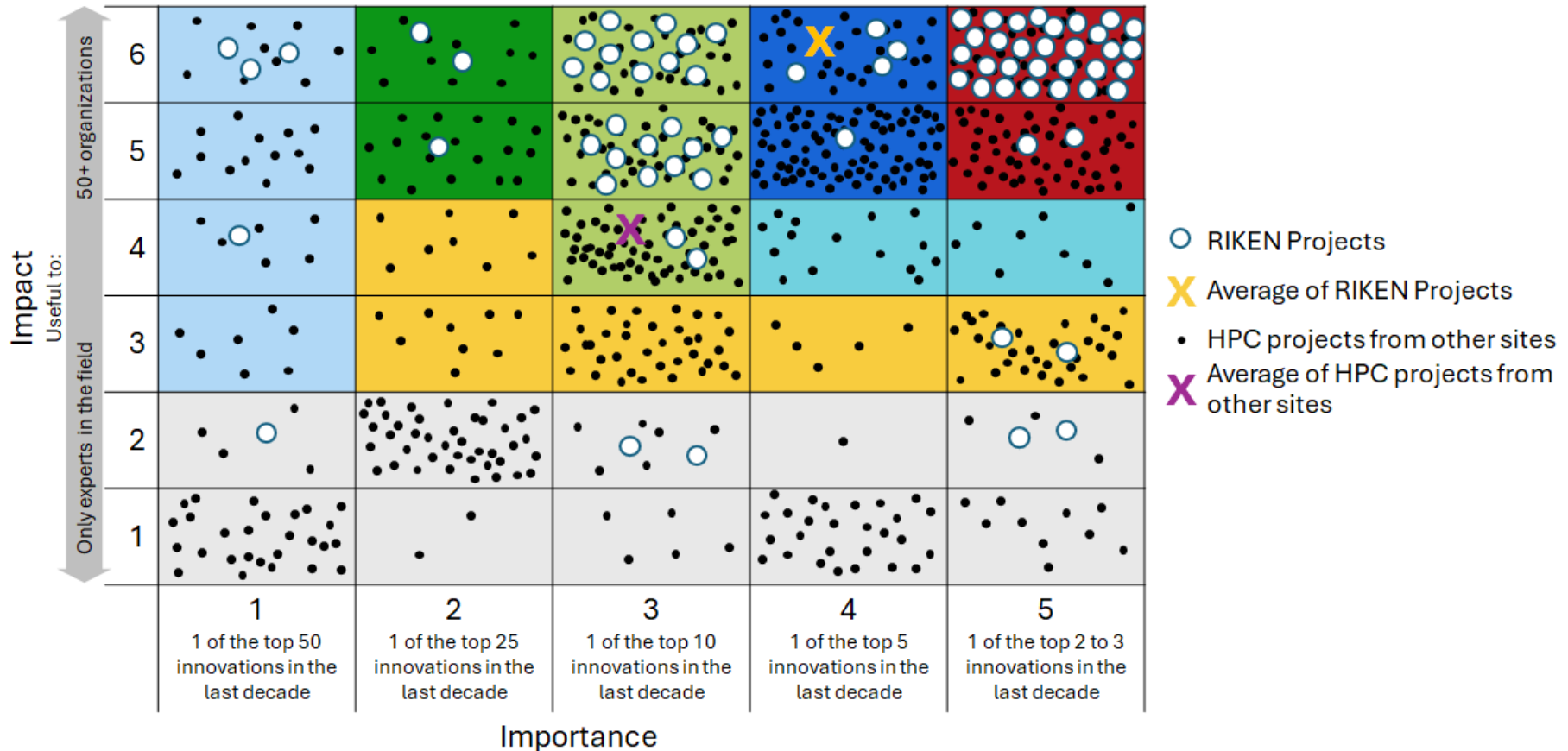
A New Way to Show the Value of Leadership Computing

Using two scales: innovation importance level, and how broadly impactful are the results



A New Way to Show the Value of Leadership Computing - RIKEN

An example from a 2024 study compared to 650 other projects



A New Way to Show the Value of Leadership Computing - NERSC

An example from a 2024 study compared to 650 other projects

Innovation Class Mapping: Showing Participating NERSC projects

