HPC Market Update: HPC/AI Market Results, and High Growth Areas

SC23

www.HyperionResearch.com
www.hpcuserforum.com

Earl Joseph
(ejoseph@hyperionres.com)

About Hyperion Research

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

Hyperion Research mission:

- *Hyperion Research helps organizations make effective decisions and seize growth opportunities*
  - *By providing research and recommendations in high performance computing and emerging technology areas*

HPC User Forum mission:

- *To improve the health of the HPC/AI/QC industry*
  - *Through open discussions, information sharing and initiatives involving HPC users in industry, government and academia along with HPC vendors and other interested parties*

The Hyperion Research Team

Analysts

- Earl Joseph, CEO
- Bob Sorensen, SVP Research
- Mark Nossokoff, Research Director
- Jaclyn Ludema, Analyst
- Melissa Riddle, Associate Analyst
- Thomas Sorensen, Associate Analyst
- Cary Sudan, Principal Survey Specialist

Operations

- Jean Sorensen, COO

International Consultants

- Katsuya Nishi, Japan and Asia
- Jie Wu, China & Technology Trends

Global Accounts

- Mike Thorp, Sr. Global Sales Executive
- Kurt Gantrish, Sr. Account Executive *

Data Collection

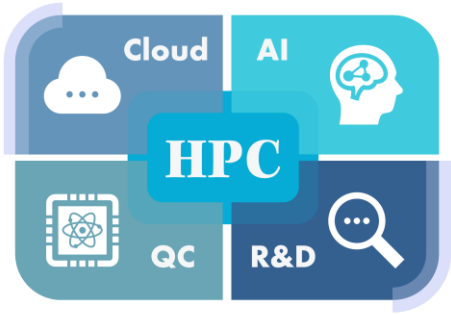
- Andrew Rugg, Certus Insights
- Kirsten Chapman, KC Associates

* primary individuals supporting this project

Our Research Areas

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

- **Traditional HPC**
- **AI: ML, DL, LLM & Large Scale AI**
- **Cloud Computing**
- **Quantum Computing**
- **Storage & Big Data**
- **Interconnects**
- **Software & Applications**
- **Power & Cooling**
- **The ROI and ROR from Using HPC**
- **Tracking all Processor Types & Growth Rates**
- **R&D and Engineering -- All Types of High Tech**
- **Edge Computing**
- **Staffing & Supply Chain Issues**



Today's Agenda

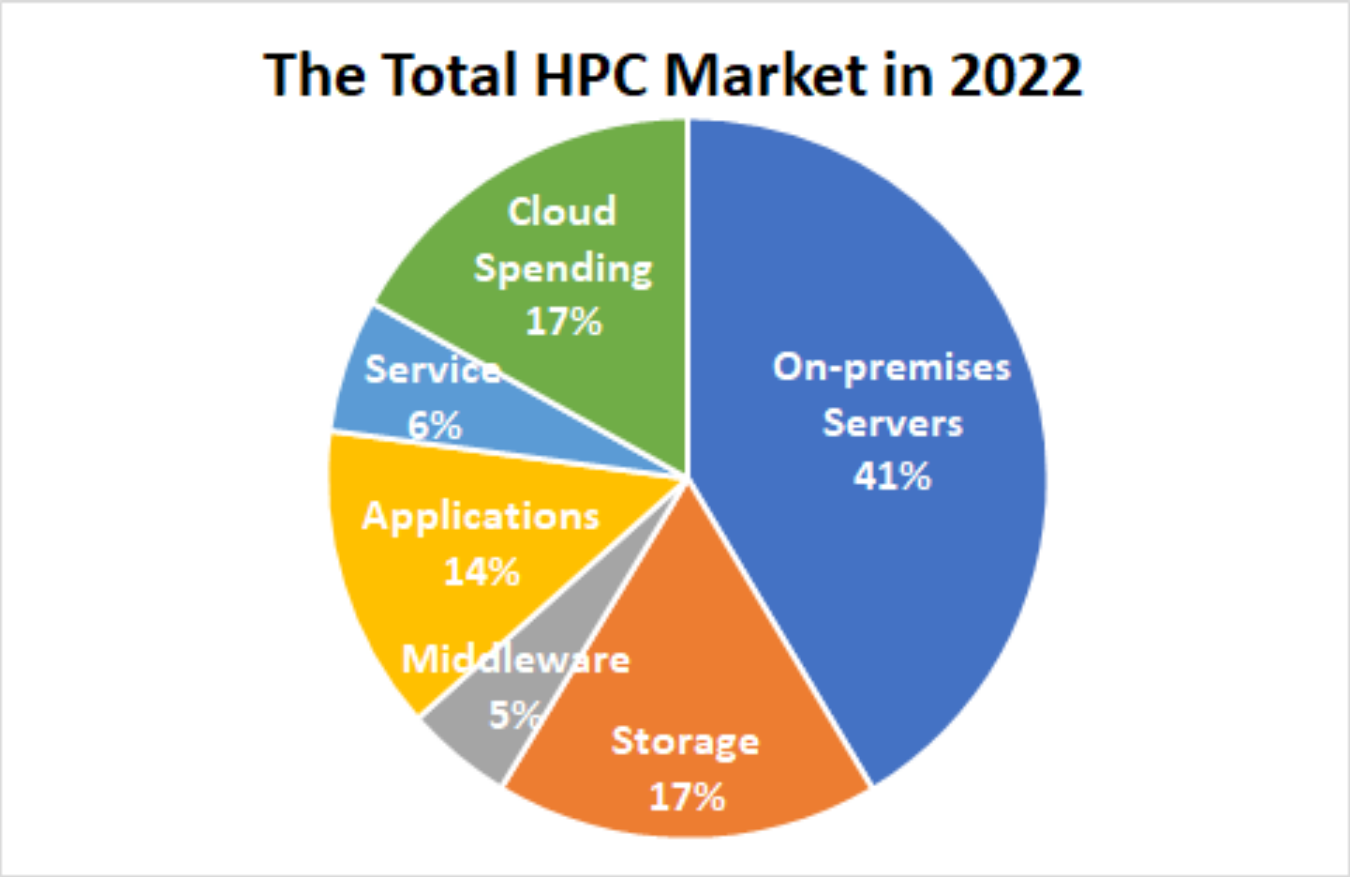
- Earl Joseph: Market Update
- Tom Sorensen: Update on the Intersection of AI and HPC
- Mark Nossokoff: State of HPC Cloud
- Jaclyn Ludema: Perspective on Sustainability in HPC
- Bob Sorensen: Exascale + Neo Exascale: What's Next?
- Mark Nossokoff: State of HPC Storage and Interconnects
- Bob Sorensen: The Global QC Market: Realistic and Steady Growth Ahead
- Melissa Riddle: HPC Applications and Verticals
- Earl Joseph: Conclusions

HYPERION RESEARCH

HPC Market Update

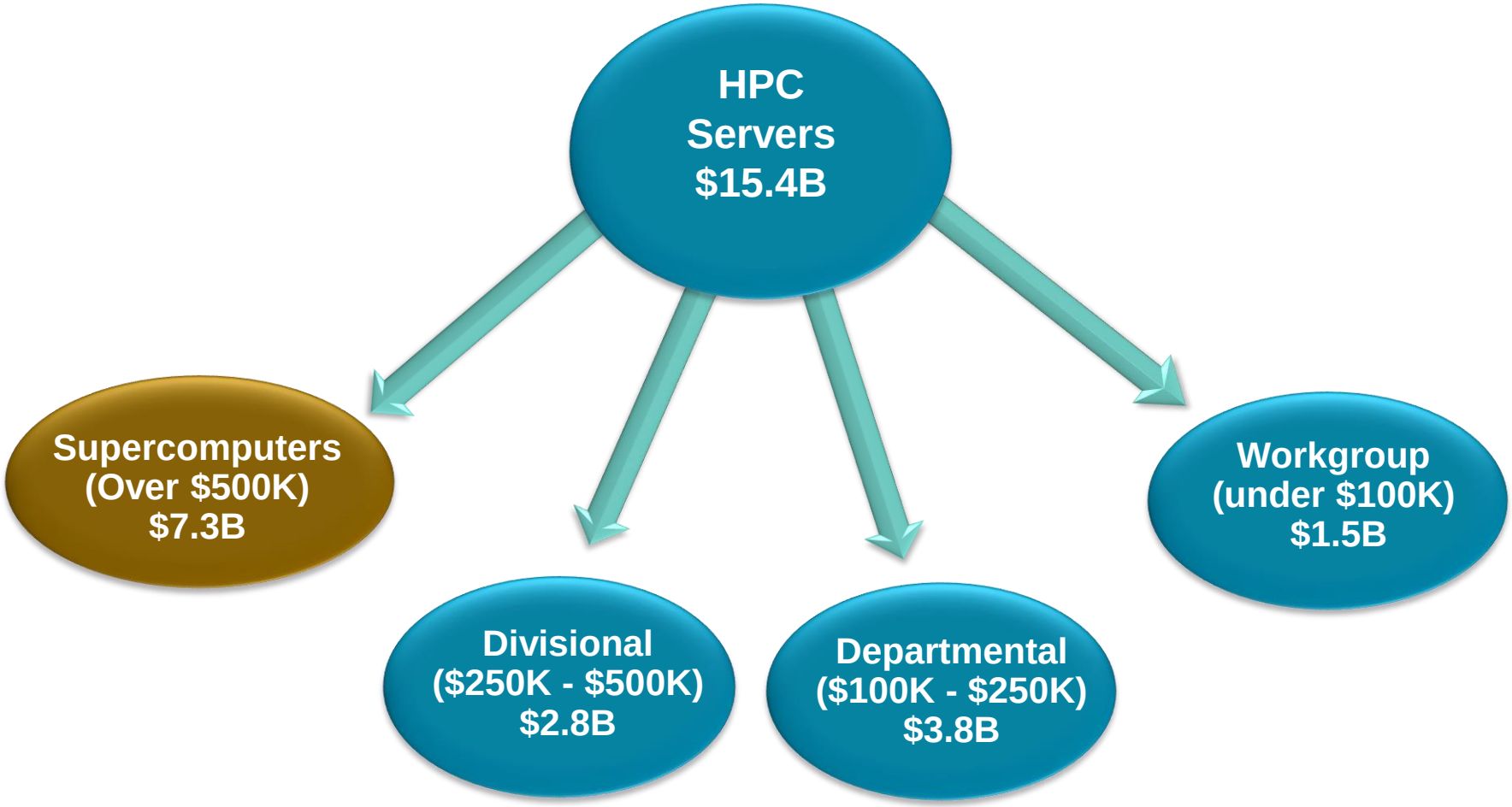
The Overall HPC Market in 2022

Looking at the overall HPC market, including servers, cloud usage, storage, software and repair services = $37.3 billion US dollars



The 2022 Worldwide On-Prem HPC Server Market: $15.4 Billion (up 4.3%)

2023 is projected to be around $16.5 Billion



2022 WW HPC On-Prem Market by Vendor and Sector ($ Millions)

HPC On-premises Server Market ($M)	
Vendor	**2022**
HPE	$5,137
Dell Technologies	$3,575
Lenovo	$1,201
Inspur	$1,073
Sugon	$603
IBM	$505
Atos	$480
Fujitsu	$230
NEC	$207
Penguin	$442
Other	$1,988
Total	**$15,441**
Source: Hyperion Research, 2023	

HPC On-premises Server Market ($M)	
Sector/Vertical	**2022**
Bio-Sciences	$1,449
CAE	$1,768
Chemical Engineering	$173
DCC & Distribution	$826
Economics/Financial	$757
EDA / IT / ISV	$873
Geosciences	$998
Mechanical Design	$57
Defense	$1,602
Government Lab	$3,342
University/Academic	$2,677
Weather	$700
Other	$221
Total	**$15,441**
Source: Hyperion Research, 2023	

First Half of 2023 WW HPC On-Prem
($ Millions)

9.5% growth for the first six months of 2023

First Half 2023 Growth Rate									
	2022				2023		First Half	First Half	Growth
$ millions	Q122	Q222	Q322	Q422	Q123	Q223	2022	2023	Rate
Supercomputer	$1,348	$1,533	$2,071	$2,267	$1,600	$1,729	$2,881	$3,330	15.6%
Divisional	$608	$656	$725	$815	$656	$673	$1,264	$1,329	5.1%
Departmental	$852	$900	$1,000	$1,073	$925	$886	$1,752	$1,811	3.4%
Workgroup	$337	$347	$412	$423	$355	$384	$684	$739	8.0%
Total Revenue	**$3,144**	**$3,437**	**$4,209**	**$4,578**	**$3,536**	**$3,673**	**$6,582**	**$7,209**	**9.5%**
Source: Hyperion Research, September 2023									

5-Year On-Prem HPC Server Forecast

8.0% yearly average growth over the next 5 years
LLMs and other AI are driving growth increases

On-Prem HPC Server Revenue Forecast by Competitive Segment								
$ millions	2021	2022	2023	2024	2025	2026	2027	CAGR 22-27
Supercomputer	$6,971	$7,219	$7,958	$8,839	$9,577	$10,390	$11,360	9.5%
Divisional	$2,783	$2,805	$3,067	$3,371	$3,580	$3,897	$4,182	8.3%
Departmental	$3,614	$3,826	$4,020	$4,397	$4,660	$4,981	$5,336	6.9%
Workgroup	$1,412	$1,519	$1,441	$1,507	$1,553	$1,626	$1,707	2.4%
Total	**$14,781**	**$15,369**	**$16,486**	**$18,113**	**$19,369**	**$20,894**	**$22,586**	**8.0%**
Source: Hyperion Research, September 2023								

On-Prem Broader Market Forecast

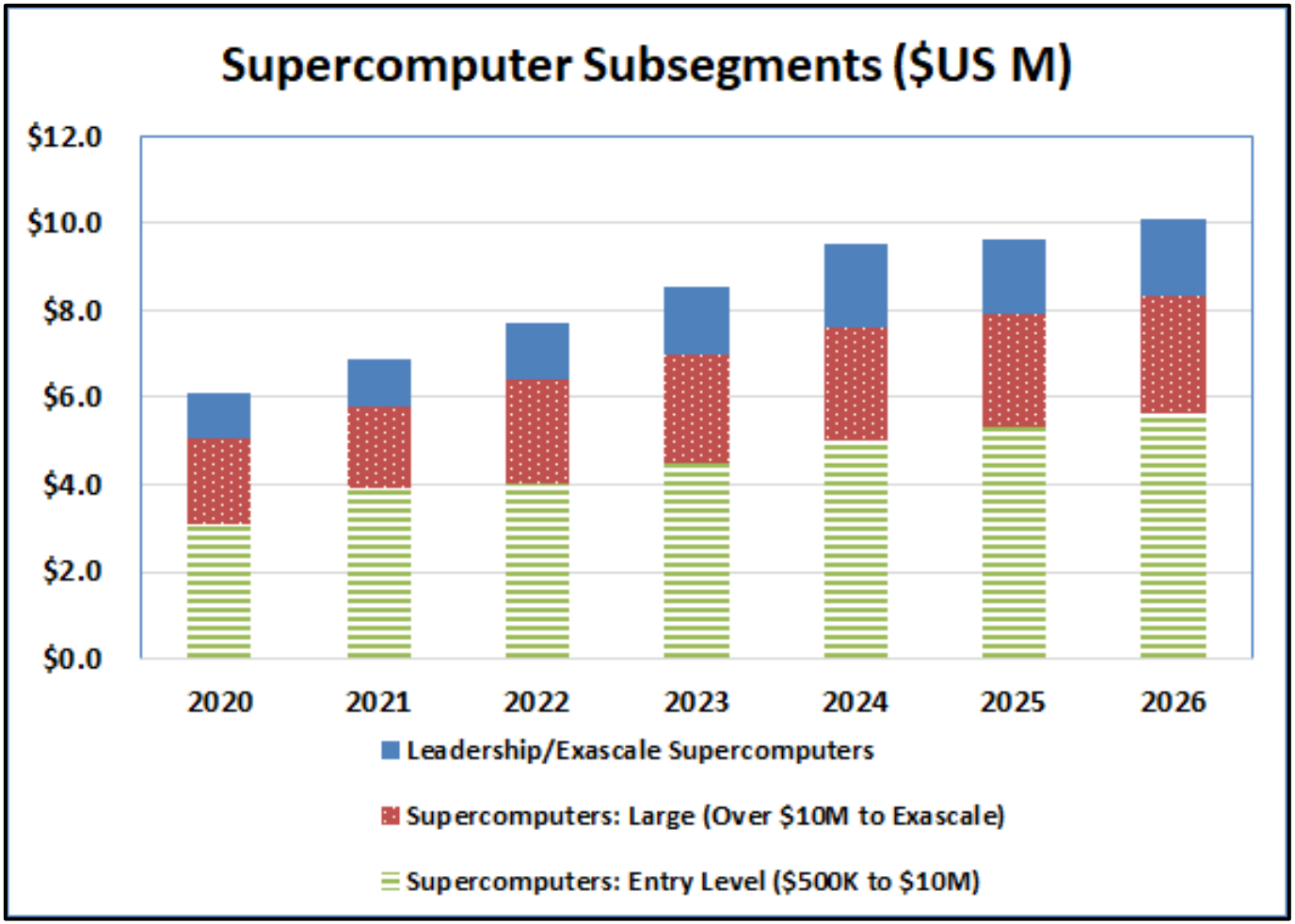
Overall market to exceed $44 billion by 2027

On-Prem Revenues by the Broader HPC Market Areas								
$ millions	2021	2022	2023	2024	2025	2026	2027	CAGR 22-27
Server	$14,781	$15,369	$16,486	$18,113	$19,369	$20,894	$22,586	8.0%
Storage	$5,985	$6,380	$6,924	$7,759	$8,434	$9,161	$10,007	9.4%
Middleware	$1,733	$1,781	$1,887	$2,049	$2,160	$2,316	$2,503	7.0%
Applications	$4,960	$5,069	$5,320	$5,729	$6,045	$6,446	$6,935	6.5%
Service	$2,272	$2,214	$2,220	$2,286	$2,323	$2,344	$2,508	2.5%
Total Revenue	**$29,731**	**$30,813**	**$32,836**	**$35,936**	**$38,331**	**$41,161**	**$44,539**	**7.6%**

Source: Hyperion Research, September 2023

Supercomputer Subsegments

The market for systems under $10M US is very large



High Growth Areas

Relative Growth Rates

The use of LLMs and other AI will likely increase these rates

- **8.0% On-premises HPC/AI servers**
  - 7.6% HPC/AI broader on-premises market
- **17.2 % GPU boards/accelerators**
  - With increased prices, GPU revenues are growing at a considerably higher rate
- **17.9% Running HPC workloads in the cloud**
- **22.7% AI systems**
  - 32.2% for DL

The Exascale Market (System Acceptances)

Exascale and Near-Exascale Leadership Systems (2020 to 2028)							
Year Accepted	**China**	**Europe**	**Japan**	**US**	**Other Countries***	**Total Systems**	**Total Value**
2020			1 near-exascale system ~$1.1B			1	$1.1B
2021	2 exascale ~$350M each	1 pre-exascale system ~$180M	--	1 pre-exascale system ~$200M	--	4	$1.1B
2022	1 exascale ~$350M	2 pre-exascale systems ~$390M **total**	--	1 exascale system ~$600M (2/3 accepted 2022)	--	4	$1.1B
2023	1 exascale system ~$350M	1 or 2 pre-exascale systems ~$150M each	1 near-exascale system ~$150M	1 exascale system ~$600M + remaining 1/3 of Frontier system	--	5-6	$1.5B - $1.6B
2024	1 exascale system ~$350M	1 exascale ~$350M, plus 1 exascale (or pre) system ~$200M	?	1 exascale system ~$600M	1 pre-exascale system ~$125M	5	~$1.6B
2025	1 or 2 exascale systems ~$300M each	2 or 3 exascale systems ~$350M each	1 exascale system ~$200M	1 or 2 exascale systems ~$350M each	1 near-exascale system ~$125M	6-9	$1.7B - $2.7B
2026	2 exascale systems ~$300M each	2 or 3 exascale systems ~$325M each	?	1 or 2 exascale systems ~$325M each	1 or 2 exascale systems ~$150M each	6-9	$1.7B - $2.5B
2027	2 exascale systems ~$275M each	2 or 3 exascale systems ~$300M	1 exascale system ~$150M	1 or 2 exascale systems ~$275M each	2 or 3 exascale systems ~$130M each	8-11	$1.8B - $2.5B
2028	2 exascale systems ~$250M each	2 or 3 exascale systems ~$275M	1 or 2 exascale systems ~$150M each	1 or 2 exascale systems ~$275M each	2 or 3 exascale systems ~$125M each	8-12	$1.7B - $2.6B
Total	**12-13**	**14-19**	**5-6**	**8-12**	**7-10**	**47-61**	**$13.4B - $16.8B**

94.3% of Sites Have Accelerators in Their Largest System Today

Up from 82.7% having accelerators in 2021

In Mid 2021

How many co-processors or accelerators are in your largest HPC technical server?		
	Responses	Percent
None	23	17.3%
Less than 32	28	21.1%
32 to less than 64	18	13.5%
64 to less than 100	19	14.3%
100 to less than 500	18	13.5%
500 to less than 1,000	11	8.3%
1,000 to less than 5,000	10	7.5%
5,000 to less than 10,000	4	3.0%
10,000 or more	2	1.5%
n = 133		
Source: Hyperion Research, 2021		

In Late 2022

Largest System Accelerator Count		
Q: How many compute-oriented accelerators/co-processors are in your largest on-premises HPC technical server?		
		Overall Percent
None		5.7%
Less than 32		24.4%
32 to less than 64		15.3%
64 to less than 100		12.5%
100 to less than 500		13.1%
500 to less than 1,000		7.4%
1,000 to less than 5,000		7.4%
5,000 to less than 10,000		2.8%
10,000 to less than 50,000		2.3%
50,000 to less than 100,000		4.0%
100,000 to less than 250,000		3.4%
250,000 to less than 500,000		0.6%
750,000 to less than 1,000,000		0.6%
1,000,000 to less than 5,000,000		0.6%
n = 176; 104; 20; 52		
Source: Hyperion Research, 2023		

Accelerator Plans for Next Purchases

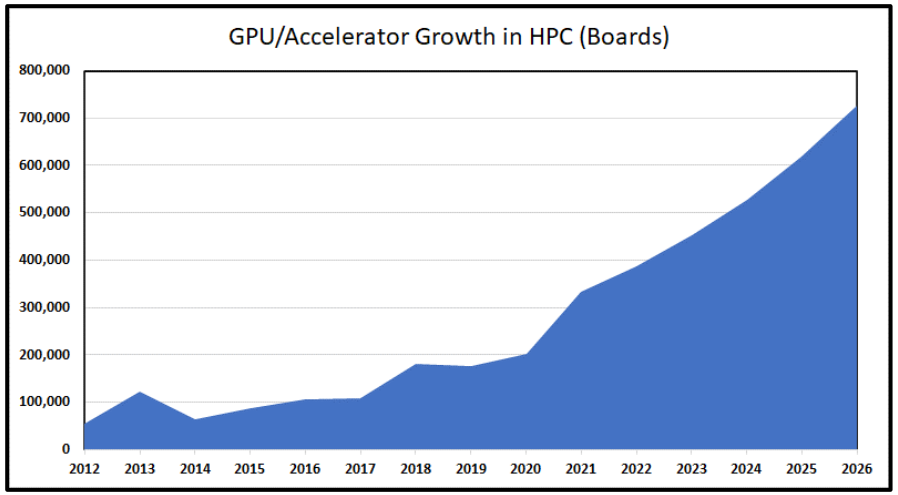
From our recent end-user MCS study

Planned Processing Elements by Sector				
Q: In the next 12 – 18 months, which of these processing elements do you expect will be incorporated into your HPC/AI/HPDA compute resources? Select all that apply:				
	Overall Percent	**Industry Percent**	**Government Percent**	**Academia Percent**
GPUs	74.0%	67.9%	85.0%	82.7%
TPUs (tensor processing units)	24.3%	27.5%	25.0%	17.3%
FPGAs	22.7%	28.4%	15.0%	13.5%
Single-purpose AI processors	11.0%	12.8%	5.0%	9.6%
ASICs	8.3%	11.9%	0.0%	3.8%
Neuromorphic processors	7.7%	9.2%	10.0%	3.8%
eASICs	2.2%	3.7%	0.0%	0.0%
Other	2.8%	2.8%	0.0%	3.8%
None	5.5%	7.3%	5.0%	1.9%
n = 181; 109; 20; 52				
Source: Hyperion Research, 2023				

GPU/Accelerator Forecast

Anticipated high growth for accelerators over next 5 years

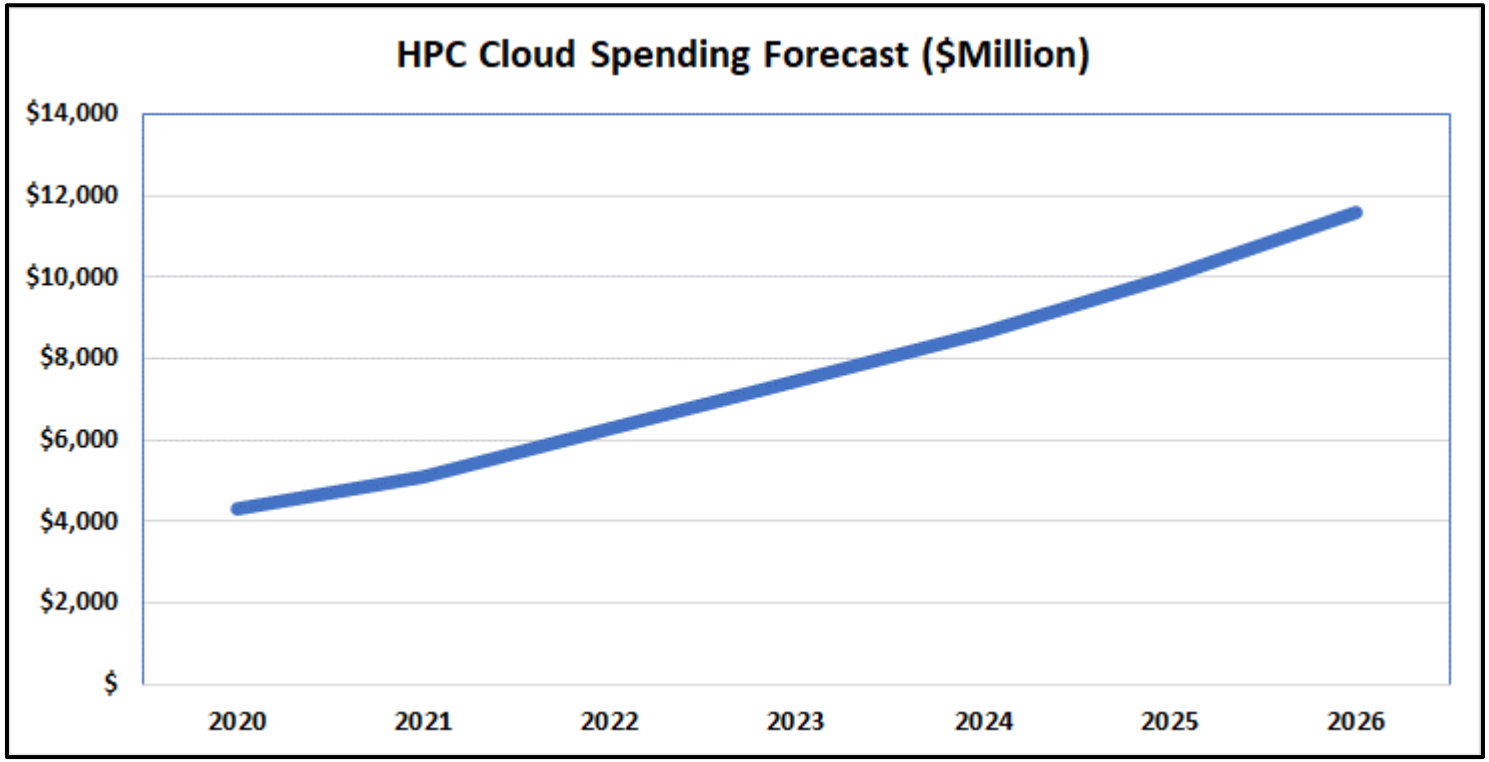
GPU & Accelerators Forecast								
	2021	2022	2023	2024	2025	2026	2027	CAGR 22-27
GPC & Accelerator Boards	334,037	386,035	476,371	546,438	639,405	742,853	854,280	17.2%
Source: Hyperion Research, September 2023								



HPC Cloud Usage Forecast

17.9% growth over the next 5 years

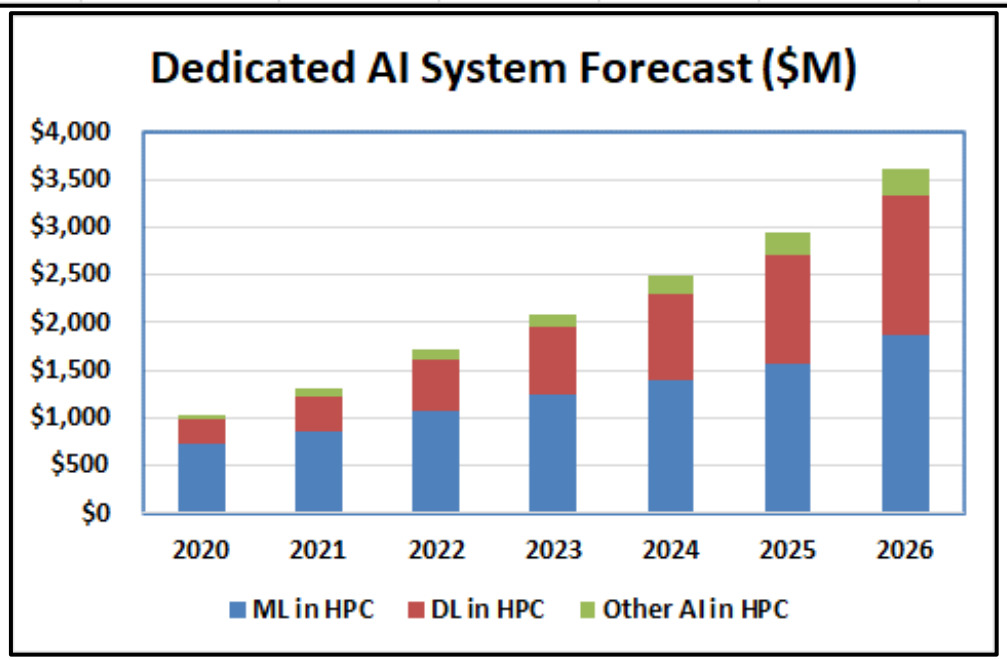
HPC Cloud Spending ($ Million)								
	2020	2021	2022	2023	2024	2025	2026	CAGR 21 to 26
HPC Cloud Spending	$4,300	$5,100	$6,304	$7,472	$8,630	$10,011	$11,613	17.9%
Source: Hyperion Research, 2023								



AI Forecast

22.7% growth over the next 5 years

Worldwide HPC-Enabled AI Forecast (ML, DL, & Other AI) Server Revenue ($M)								
	2020	2021	2022	2023	2024	2025	2026	CAGR 21-26
ML in HPC	$719	$861	$1,081	$1,243	$1,391	$1,568	$1,859	16.6%
DL in HPC	$263	$364	$532	$708	$919	$1,147	$1,468	32.2%
Other AI in HPC	$57	$75	$104	$132	$173	$226	$292	31.3%
Total AI Server Revenue	$1,039	$1,300	$1,718	$2,083	$2,484	$2,941	$3,619	22.7%
Source: Hyperion Research, 2023								



Today's Agenda

- Earl Joseph: Market Update
- Tom Sorensen: Update on the Intersection of AI and HPC
- Mark Nossokoff: State of HPC Cloud
- Jaclyn Ludema: Perspective on Sustainability in HPC
- Bob Sorensen: Exascale + Neo Exascale: What's Next?
- Mark Nossokoff: State of HPC Storage and Interconnects
- Bob Sorensen: The Global QC Market: Realistic and Steady Growth Ahead
- Melissa Riddle: HPC Applications and Verticals
- Earl Joseph: Conclusions

Updates on the Intersection of AI and HPC

Tom Sorensen

November 2023

A New Category of HPC Workload

- HPC users adopting/integrating AI at high rates
- Many methods and models, LLMs draw attention
- ~90% of HPC users surveyed currently or plan to use AI methods for workloads
- AI methods introduce new demands on sites:
  - Hardware (processors, interconnect, data access)
  - Software (data management, queueing, dev tools)
  - Expertise (procurement strategy, maintenance, troubleshooting)
  - Regulatory (data provenance, privacy, legal)

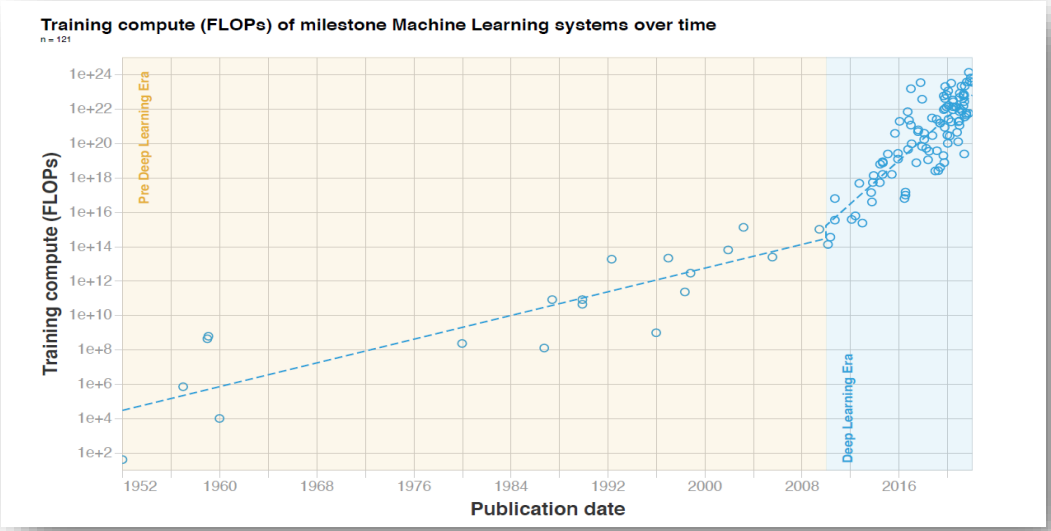
Framing LLM/HPC Requirements

Three elements dominate scaling of LLMs on HPCs

- **Compute: the absolute number of floating-point operations needed to train an LLM to a desired degree of accuracy**
- **Dataset size: input dataset used for training the LLM**
- **Model size: number of tokens or parameters**
  - The larger the number of parameters, the more nuance in the model's understanding of each word's meaning and context
- **This scaling heuristic been called the ideal gas law of machine learning**
    - PV= nRT encompasses a range of complex action
    - Scaling moves here as a f(C,D,M)
- **LLM requirements ultimately define necessary HPC specifications**

LLMs Consume Significant Flops

LLM flops growth eclipses Top 500 growth



- Pre 2010:
  - On the order of 2 X 10^{12} (200 Tflops)
  - Flops requirements doubling every 21.3 month
  - But not a lot of data points
- Post 2010 to Current:
  - Currently on the order of 6 X 10^{22} flops (60 Zettaflops)
  - Flops requirements doubling every 5.6 months
  - Roughly 11X faster than HPC Top 1 Linpack performance growth rate

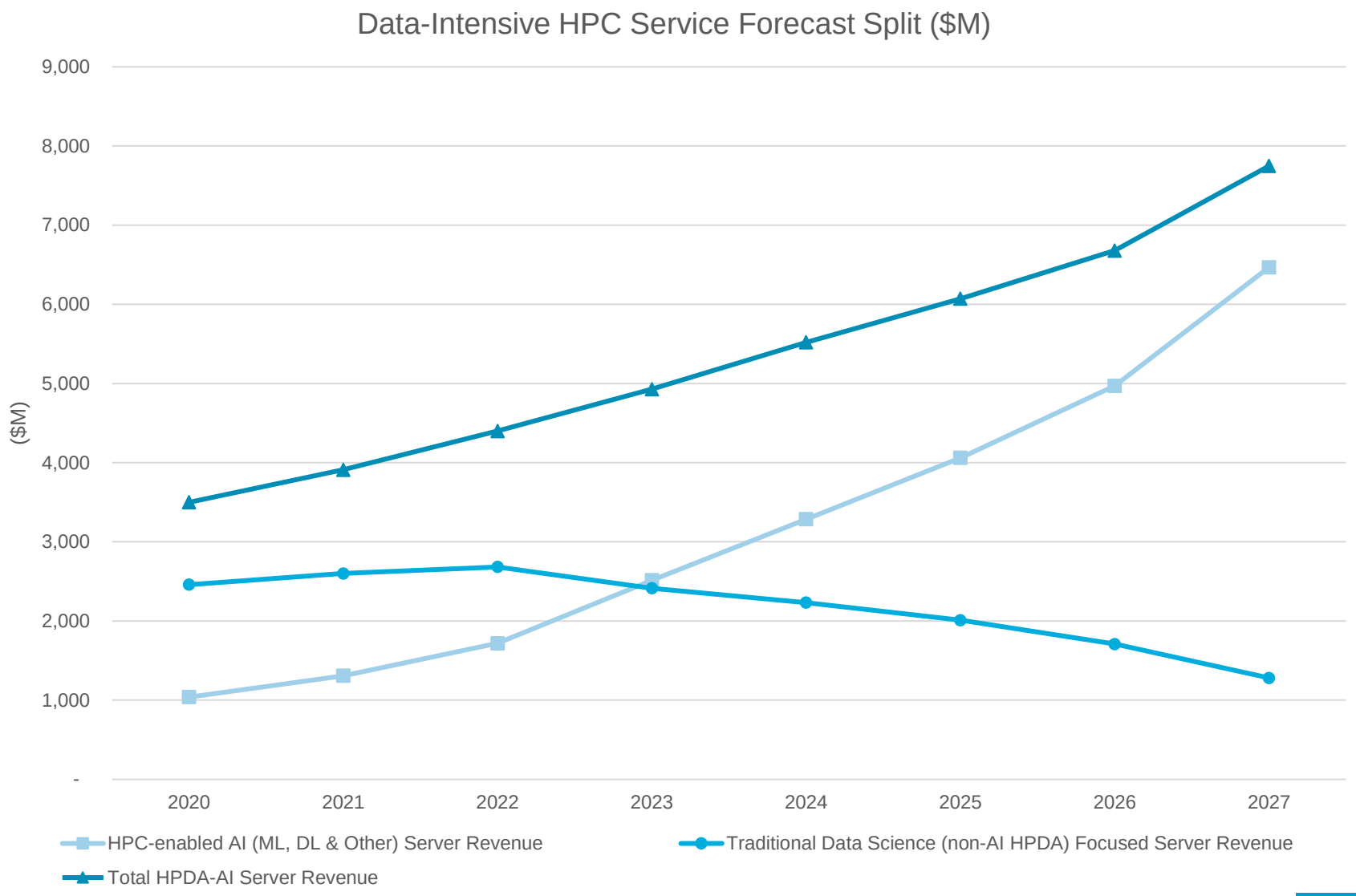
See Compute Trends Across Three Eras of Machine Learning, arXiv:2202.05924

Select Key Findings from AI Use Study

From a survey of HPC users leveraging LLMs

- **1:** LLMs are considered to be an important emerging asset for both current and planned HPC-related activity.
- **3:** Respondent organizations are looking at a broad set of LLM-related end uses.
- **5:** Many different HPC-related science and engineering algorithms were seen as viable for LLM enhancements.
- **9**: Open source is currently the most preferred option for accessing LLM software.
- **10:** Survey respondents looked to a wide range of LLM expertise to support the various stages of LLM development spanning foundation model construction, fine-tuning procedures, LLM integration into existing workloads, and supporting inference operations.

Data-Intensive HPC Forecast ($M)



Putting This All Together

Is this (another) new HPC architectural paradigm in the works?

- **Based on a recent LLM analysis by Riken**
- **GTP variant flops requirements**
  - GPT-3.5 (ChatGPT): $3X10^{24}$ flops (estimated)
  - GPT-4.0: $3X10^{25}$ flops (estimated)
- **OpenAI System: Microsoft/Open AI collaboration**
  - Top 5 system when stood up
  - GPU-based BF16 312 Tflop/s x 25,000 = 7.8 Eflop/s TPP
  - GPT3.5 (ChatGPT): 4.5 days X 2
  - GPT-4.0 45 days X 2
- **Fugaku:**
  - FP32 6.76 Tflop/s X 158,976 = 1.07 Eflop/s (TPP)
  - GPT3.5 (ChatGPT): 32 days X 10
  - GPT-4.0 45 days X 2: 328 days X 10 ~= 8.9 years

Distributed Training of Large Language Models on Fugaku, https://t.co/idofa7Tjyu

HYPERION RESEARCH

QUESTIONS?

tsorensen@hyperionres.com
ejoseph@hyperionres.com

State of HPC Cloud

SC23 Virtual Breakfast Briefing
November 2023

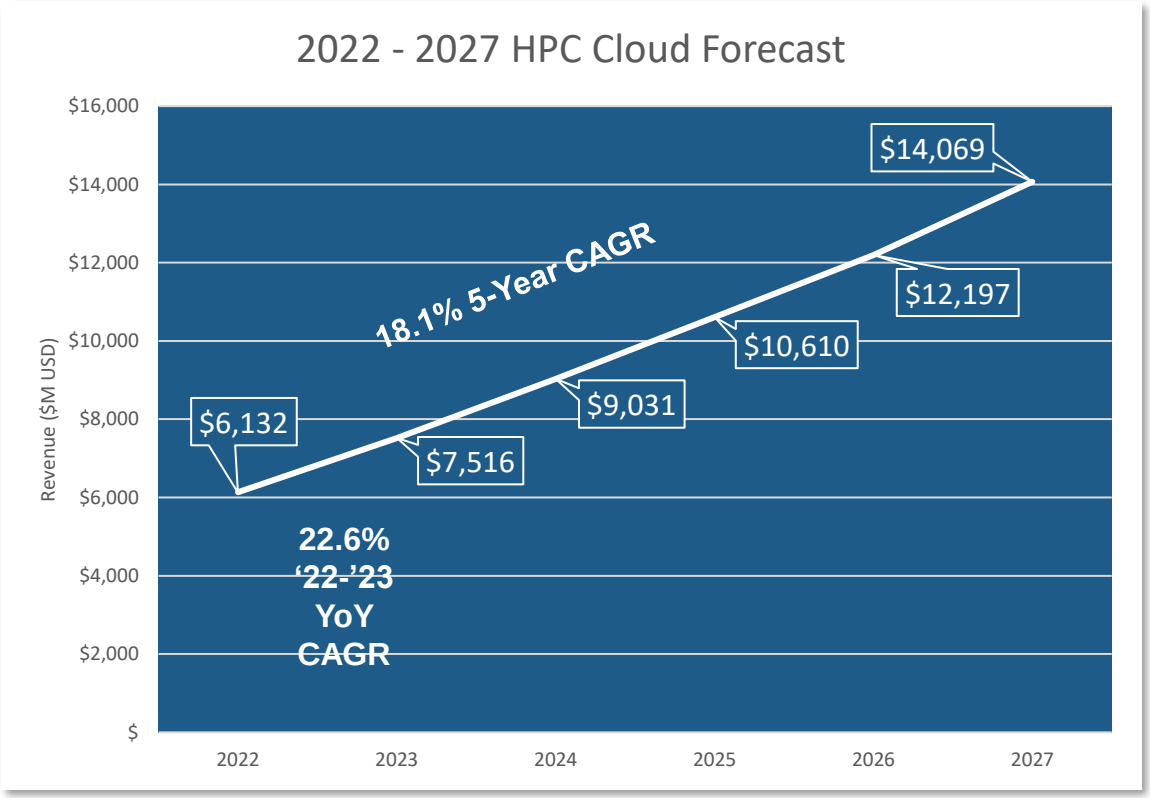
Mark Nossokoff

www.HyperionResearch.com
www.hpcuserforum.com

HPC Cloud Forecast

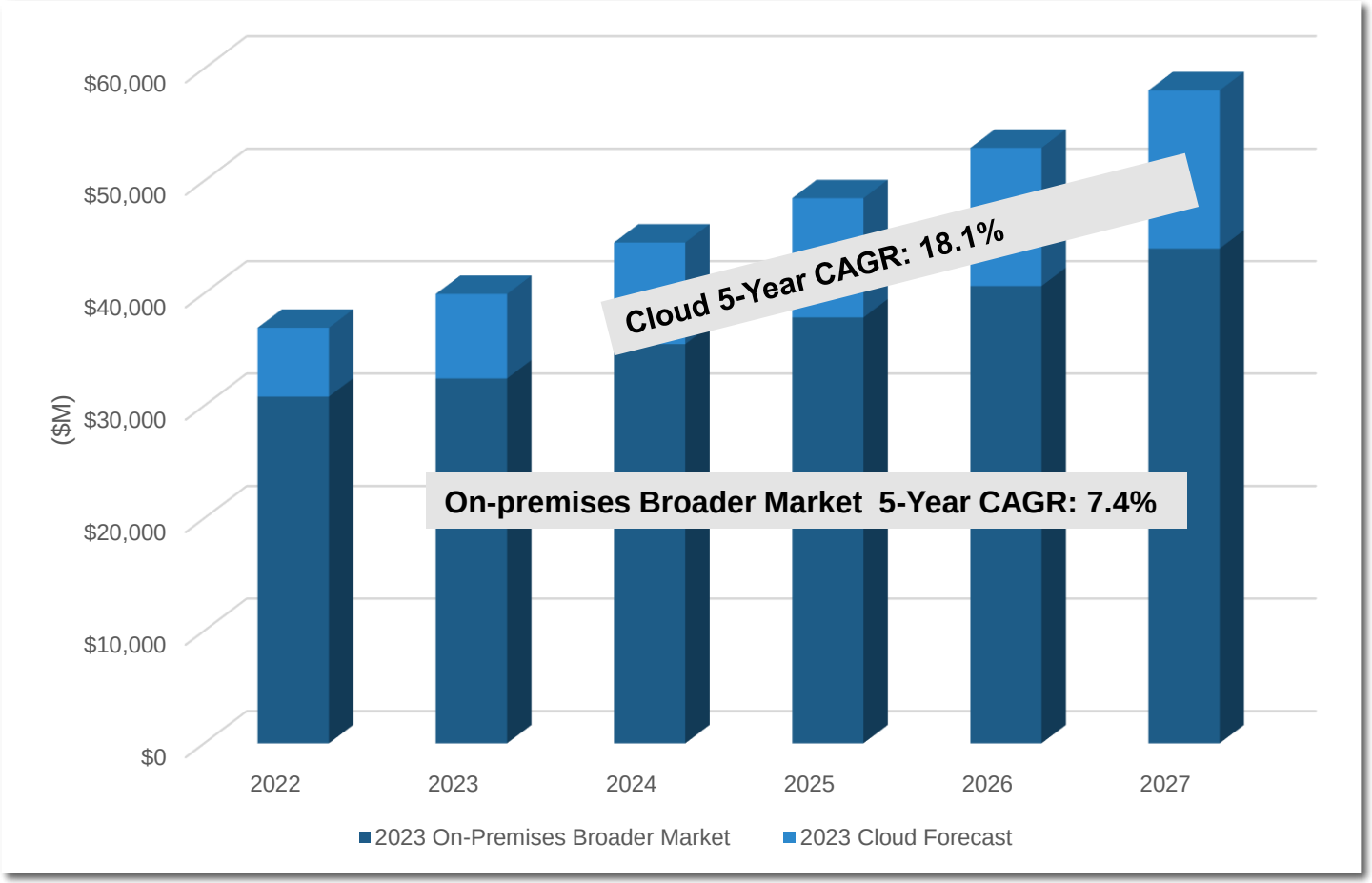
HPC cloud revenue expected to exceed $14 billion by 2027

- **Global buyers accelerating shift of on-premises HPC budgets to the cloud**



The Total HPC Market: On-Premises and Cloud Computing

The cloud market continues to outpace on-premises growth

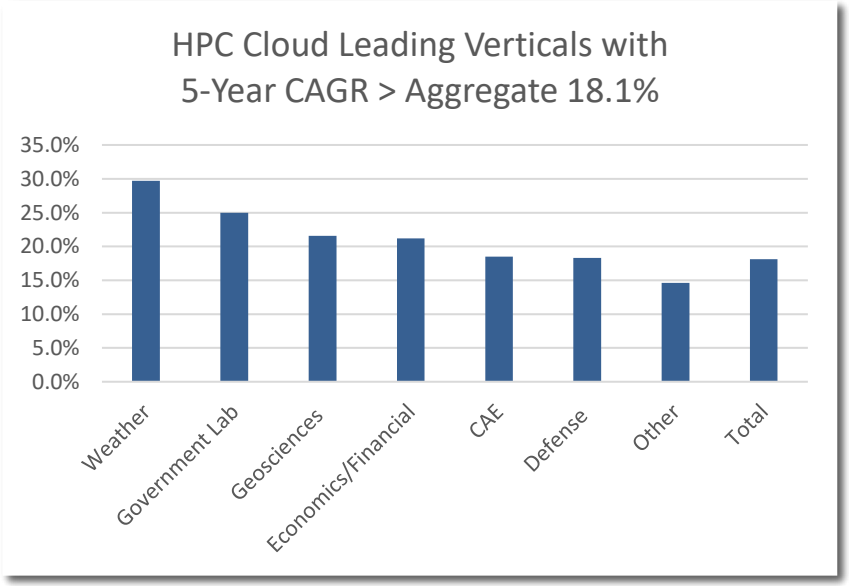


HPC Cloud Forecast by Verticals

Changes emerging in vertical adoption of cloud resources

HPC Cloud Forecast by Vertical 2022-2027 – All Verticals			
	2022 ($M)	2027 ($M)	CAGR 22-27 (%)
Bio-Sciences	$1,557	$2,821	12.6%
CAE	$1,289	$3,017	18.5%
Defense	$471	$1,091	18.3%
Government Lab	$443	$1,351	25.0%
EDA	$443	$975	17.1%
Geosciences	$403	$1,069	21.6%
Economics/Financial	$398	$1,041	21.2%
DCC & Distribution	$347	$732	16.1%
University/Academic	$276	$549	14.7%
Weather	$184	$675	29.7%
Chemical Engineering	$154	$267	11.7%
Mechanical Design	$27	$45	10.6%
Other	$141	$436	25.3%
Total	$6,132	$14,069	18.1%

Source: Hyperion Research, 2023



Other: EDA, DCC, Univ/Academic, BioSciences, Chem Engr, Mech Engr

Source: Hyperion Research, 2023

Highest

2nd highest

3rd highest

HPC Cloud Market Drivers

Rapid adoption of LLMs not fully reflected in market numbers…yet

- **AI & LLMs, including availability of GPUs**
- **Investments by cloud service providers (CSPs) to ease migrations to and integration with the cloud**
- **Users' maturing understanding of an expanding number of cloud-appropriate workloads**
- **Other recurring drivers with shifting priority order**
  - Cost-effectiveness relative to same job on-premises
  - Flexibility with surge workloads
  - Scale of available resources relative to on-premises infrastructure
  - Access to new technologies

Supply-Side Cloud Service Trends

Increasing CSP internal investments across multiple fronts are facing competition from focused entrants

- **CSP investments**
  - Integration of latest merchant CPU and GPU vendors
    - Intel, AMD, NVIDIA
  - Development of captive application-specific accelerators
    - AWS: Graviton, Inferentia, Trainium
    - Google: TPUs
  - Networking
    - Google: Apollo OCS
  - Solution-focused
    - Vertical market integrations
    - Range of price and performance capabilities
- **GPU Cloud**
  - Providing access to scarce supply of GPUs
  - Coreweave
- **Previously vertically focused cloud service providers are expanding their market coverage**
  - CGG
- **System vendors offering on-premises OPEX business models and cloud hosting services**
  - Dell: APEX, HPCaaS
  - HPE: HPE Greenlake, HPE Greenlake for LLMs
  - IBM: IBM Cloud HPC
  - Lenovo: TruScale

HYPERION RESEARCH

Questions?

mnossokoff@hyperionres.com

Trends in HPC Sustainability

SC23 Virtual Breakfast Briefing
October 2023

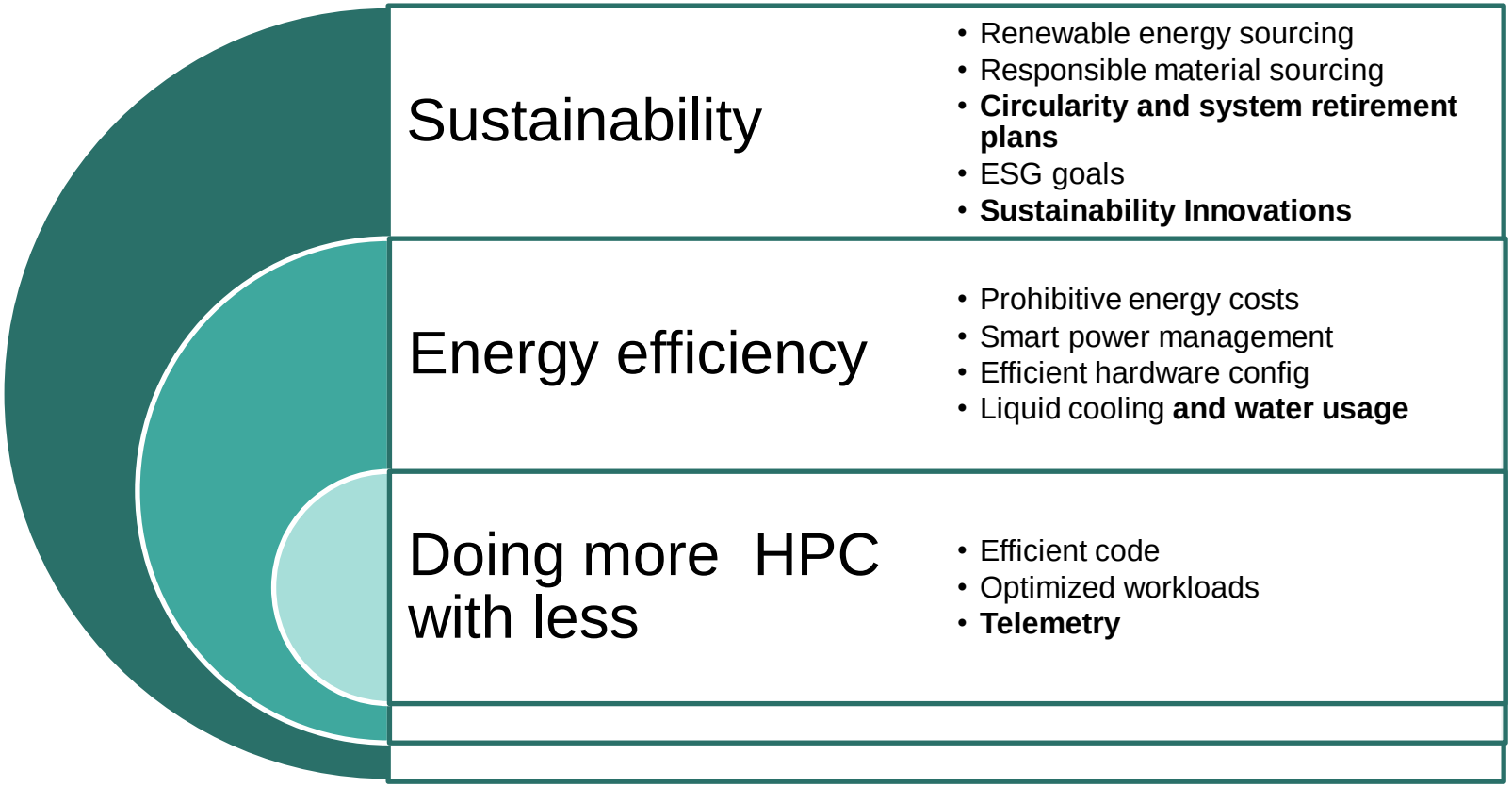
Jaclyn Ludema

www.HyperionResearch.com
www.hpcuserforum.com

Trends in Sustainability

Sites are sharing their sustainability priorities

- **Sustainability innovations, telemetry, water usage, and circularity are gaining priority at many HPC sites**



Sustainability Innovations

Applying HPC to worldwide environmental concerns

- **Digital twins of the earth for sustainability research**
  - Mapping earth systems
  - Better understanding of natural and human impacts on the environment
- **Using HPC enabled mod/sim allows for better, more accurate designs. This translates to less materials wasted in research and development stages**
- **Reducing environmental impact of oil and gas exploration and extraction**
- **Many datacenters would like for the environmental benefit of their application/innovation to “count” in their assessment of overall site sustainability**

Telemetry

Benefits to measuring the issue at hand

- **Incentivizing Energy Efficient Code**
  - When power consumption of a workload is not measured, the natural motivation of users is to get their answers as fast as possible
  - By measuring energy usage per workload, and making that data available to users, users can be incentivized to get a slower answer that uses less power
  - TACC- considering switching to charging for total energy consumed rather than wall clock hours used
- **Knowledge to inform new purchases**
  - Is it worth porting code to GPUs for energy efficiency?
  - Informed decisions on when to retire equipment
  - Energy use comparisons between on-prem and cloud

Water Usage

Liquid cooling brings with it a new sustainability concern

- **With liquid cooling gaining popularity at datacenters worldwide, water sourcing and efficient water usage are becoming a sustainability area of consideration**
- **Grey Water Systems**
  - Argonne National Lab uses grey water system that is filtered on site, reducing potable water intake by $1.2 million annually
- **Warm Water Systems**
  - Evaporative cooling systems reduce water usage
  - NASA NAS- evaporative cooling system, at a warm water temperature of 90 degrees Fahrenheit, saves 5.5 million gallons of water per year, and 6 million kWh

Circularity

Plans for equipment from procurement to retirement

- **A new focus on datacenter equipment retirement programs and recycling plans has emerged amongst industry leading sites**
- **Retirement programs are being considered during procurement stages**
- **Extending equipment life has also become an area of sustainability consideration**

HYPERION RESEARCH

Questions?

We welcome questions, comments, and suggestions

Please contact us at:

jludema@hyperionres.com

mnossokoff@hyperionres.com

Exascale + Neo Exascale: What's Next?

SC23
November 2023

www.HyperionResearch.com
www.hpcuserforum.com

Bob Sorensen

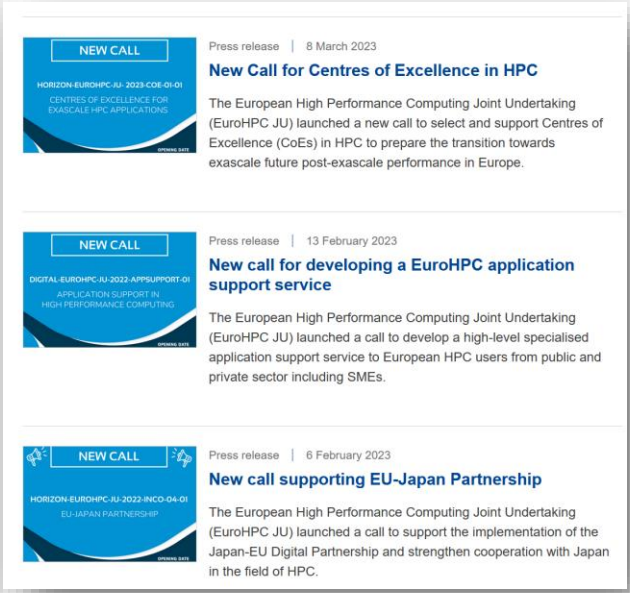
No Exascale Talk Complete Without the Top 500

On June 2023 list, US-5, EU-2, China-2, Japan-1

Rank	System/Site	# Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	Frontier - HPE Cray, DOE/SC/Oak Ridge National Laboratory, United States	8,699,904	1,194.00	1,679.82	22,703
2	Supercomputer Fugaku - RIKEN Center for Computational Science, Japan	7,630,848	442.01	537.21	29,899
3	LUMI - EuroHPC/CSC, Finland	2,220,288	309.1	428.7	6,016
4	Leonardo - EuroHPC/CINECA, Italy	1,824,768	238.7	304.47	7,404
5	Summit - DOE/SC/Oak Ridge National Laboratory, United States	2,414,592	148.6	200.79	10,096
6	Sierra - DOE/NNSA/LLNL, United States	1,572,480	94.64	125.71	7,438
7	Sunway TaihuLight - National Supercomputing Center in Wuxi, China	10,649,600	93.01	125.44	15,371
8	Perlmutter - DOE/SC/LBNL/NERSC, United States	761,856	70.87	93.75	2,589
9	Selene - NVIDIA Corporation, United States	555,520	63.46	79.22	2,646
10	Tianhe-2A - National Super Computer Center in Guangzhou, China	4,981,760	61.44	100.68	18,482

EuroHPC: Driving EU HPC Progress

The central focus of advanced computing in Europe



- Jointly funded by its 33 members with a budget of around EUR 7 billion between 2021-2027
- Develop, deploy, extend, and maintain in the EU a world-leading federated, secure and hyper-connected supercomputing, quantum computing, service and data infrastructure ecosystem

EuroHPC JU Activities and Plans

Varying workloads: varying architectures

- **An array of HPCs accessible to EU membership:**
  - **LUMI** in Finland (#3 on Top 500 list)
    - 375 Pflop/s sustained, 550Pflop/s TPP
    - Lumi-G (GPU centric),
    - Lumi-C (CPU only),
    - Lumi-D (Data Analytics: large memory),
    - Lumi-K partitions (containers, cloud services)
  - **LEONARDO** in Italy (#4 on Top 500 list)
    - 249 Pflop/s sustained, 323 Plop/s TPP
    - CPU partition (9 Pflop/s), GPU partition (240 Pflop/s)
  - **Vega** in Slovenia
    - 6.3 Pflop/s sustained, 10 Pflop/s TPP
    - CPU partition (960 nodes), GPU partition (60 nodes)
  - **MeluXina** in Luxembourg
    - 12.8 Plop/s, sustained, 18.3 Plop/s TPP
    - CPU partition (570 nodes), Accelerator Module (220 nodes)
  - **Discoverer** in Bulgaria
    - 4.5 Pflop/s sustained, 5.9 Pflop/s TPP
    - CPU only (1128 nodes)

EuroHPC JU Activities and Plans

Varying workloads: varying architectures

- **An array of HPCs in play:**
  - **Karolina** in the Czech Republic
    - 9.6 Pflop/s sustained 15.7 Pflop/s TPP
    - Four processing main partitions:
      - Standard numerical simulations: 720 servers
      - Accelerator: 70 servers with 8 GPUs per server
      - Large data set processing partition: 24 TB shared memory
      - Cloud service provider: 36 servers
  - **Deucalion** in Portugal
    - 7.22 Pflops sustained, 10 Pflops TPP
    - ARM partition (1632 nodes), x86 partition (500 nodes), accelerated partition (33 nodes)
  - **MareNostrum5** in Spain (still TBD)
    - 205 Pflops sustained, 314 Pflops TPP
    - GPP (general purpose partition), ACC (Accelerated partition,) NGT GPP (Next Gen Tech GPP partition), NGT ACC (Next Gen Tech ACC)

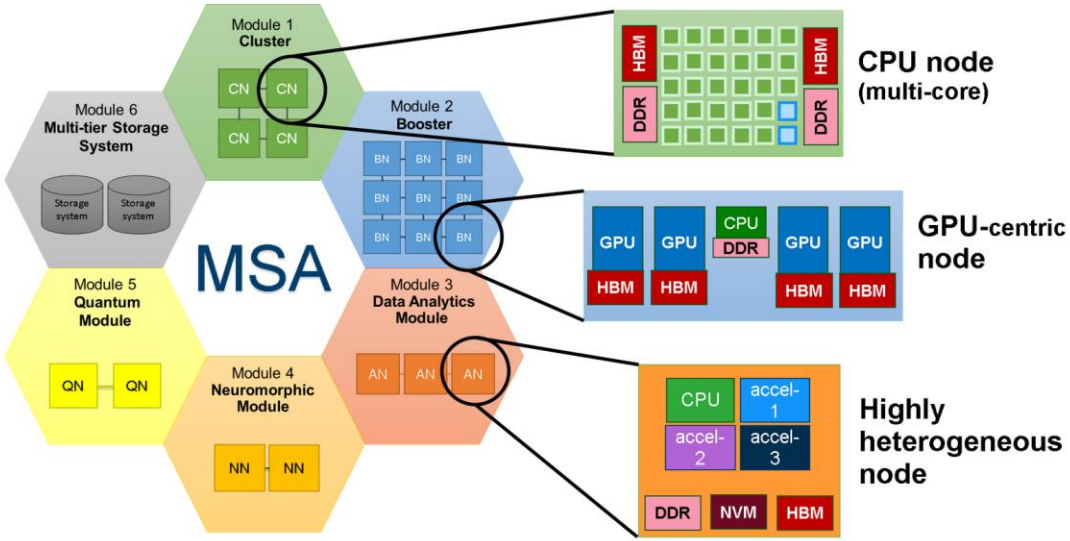
EU Plans for First Exascale System

JUPITER (Joint Undertaking Pioneer for Innovative and Transformative Exascale Research

- **Planned for 2024**
- **Installed at the Julich Supercomputing Center Germany**
  - Already hosts:
    - JUWELS Booster (#12), Module 1 (#93)
    - JURECA Data Centric Module (#61)
- **Average power is anticipated to be up to 15 megawatts**
- **Overall system and operation costs: 500 million euros**
    - 250 million euros: EuroHPC JU
    - 250 million euros in equal parts by the German Federal Ministry of Education and Research (BMBF) and the Ministry of Culture and Science of the State of North Rhine-Westphalia (MKW NRW)

EU Plans for First Exascale System

Is Jupiter the prototype for neo exascale systems?



https://www.fz-juelich.de/en/ias/jsc/about-us/structure/divisions/technology-division/next-gen-arch-proto/msa

- **Based on Eviden's Bull Sequana XH3000 technology**
- **Partnered with ParTec AG**
  - Specializes in the development and manufacture of modular supercomputers and quantum computers
- **SiPearl Arm-based Rhea Processors (fab TSMC))**
- **NVIDIA GPU (fab TSMC)**

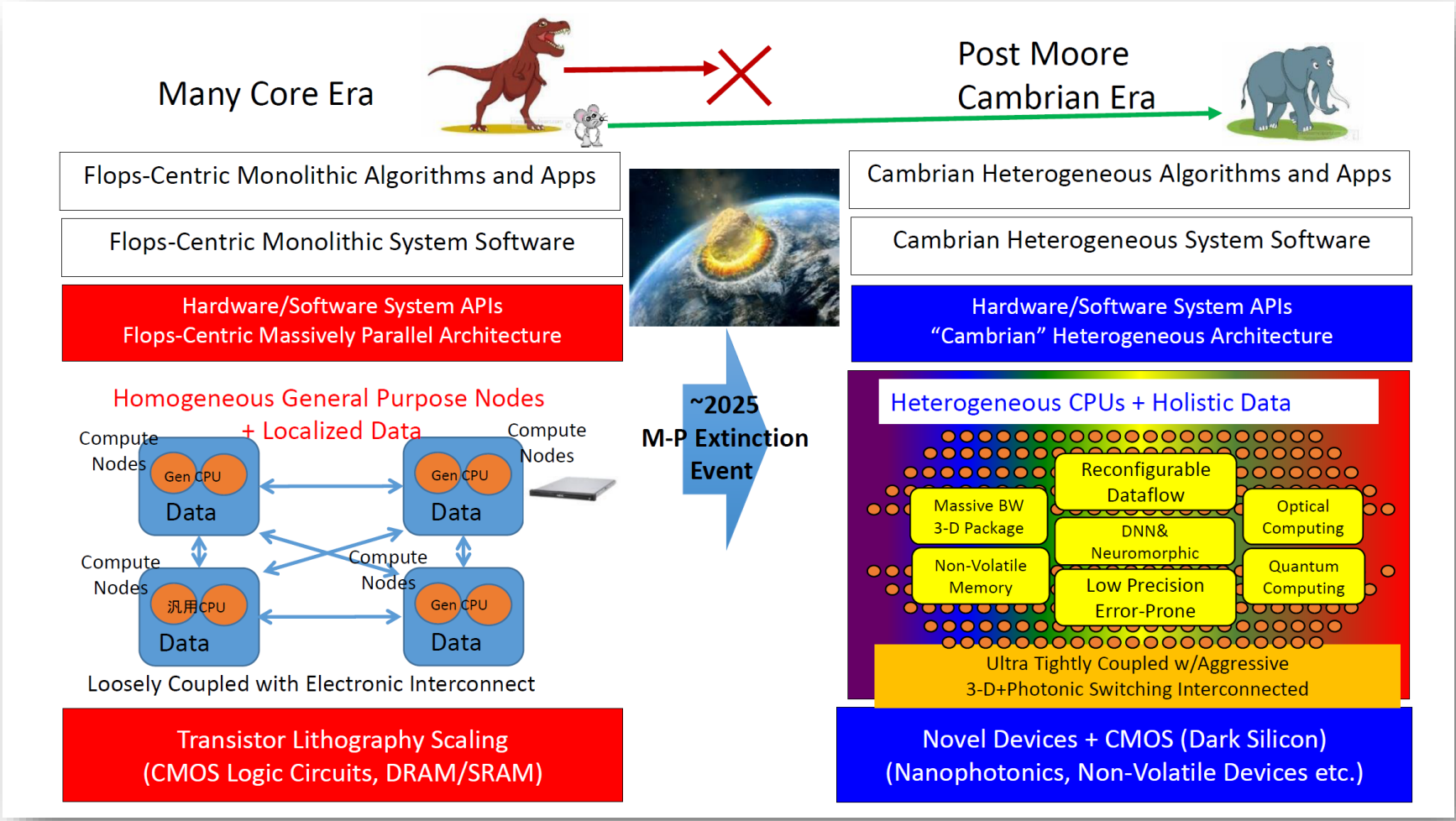
UK Plans for First Exascale System

50X performance of ARCHER2, UK's fastest system

- **Systems will be installed at EPCC's Advanced Computing Facility in Edinburgh**
- **Installation of the first phase is due to begin in 2025**
- **Part of a £900 million investment to upgrade the UK's next-generation compute capacity**
  - Speculations call for a £500 million budget for the HPC
- **More details to follow….**

Japan's Plan for 'Fugaku NEXT'

Not a straight-line projection



Satoshi Matsuoka, Director Riken RCCS ASCAC Presentation

China Exascale Status

- **No official announcements**
- **Last new Chinese appearance in Top 10 list was June 2018**
  - #7 Sunway TaihuLight in Wuxi (2016)
  - #10 Tianhe-2A at NUDT (2018)
  - Probably no new announcements this time
  - Most likely politically, not technologically, motivated
- **Strong evidence of at least five or more other Chinese systems that could make Top 10 list today**
- **Work, however, is being done**
  - Presentation on Wednesday: 5 ExaFlop/s HPL-MxP Benchmark with Linear Scalability on the 40-Million-Core Sunway Supercomputer

Rank	Site	Computer	Cores	HPL-AI (Eflop/s)	TOP500 Rank	HPL Rmax (Eflop/s)	Speedup
1	DOE/SC/ORNL	Frontier	8,730,112	9.9507	1	1.1940	8.3
2	EuroHPC/CSC	LUMI	2,174,976	2.168	3	0.3091	7.0
3	RIKEN	Fugaku	7,630,848	2.000	2	0.4420	4.5

Near-Term US Exascale Status

Three systems over two (or more) years with budget of ~ $1.8 billion

- **Frontier: DOE Office of Science: Oak Ridge National Laboratory**
  - First US exascale system in US
  - June Top 500 List: Rpeak = 1.68 Eflop/s, Rmax = 1.1 Eflop/s
  - 21 MW to run LINPAC
  - Cray Shasta with AMD EPYC CPU and AMD Radeon Instinct GPUs
  - Full user operations January 2023 (some delay)
- **Aurora: DOE Office of Science, Argonne National Laboratory**
  - **60MW**, ~ 1Eflop/s DP sustained, 2Eflop/s TPP
  - Cray Shasta architecture with Intel Xeons and Intel Xe GPU
  - On the list this week?
  - Delivery in late 2023, acceptance in 2024 (delayed at least X months)
- **El Capitan: DOE NNSA's LLNL**
  - ~ 2 Eflop/s
  - Cray Shasta architecture with AMD EPYC processors, next generation Radeon Instinct GPUs
  - Fully deployed in 2024

US Exascale Plans Going Forward

A new US Government procurement paradigm?

CHARTING A PATH IN A SHIFTING TECHNICAL AND GEOPOLITICAL LANDSCAPE:
POST-EXASCALE COMPUTING FOR THE NATIONAL NUCLEAR SECURITY ADMINISTRATION

FINDING 2.1: Semiconductor manufacturing is now largely in the hands of offshore vendors who may experience supply-chain risk; U.S. sources are lagging.

FINDING 2.2: All U.S. exascale systems are being produced by a single integrator, introducing both a technical and an economic risk.

FINDING 2.3: The joint Exascale Computing Project created a software stack for moving systems software and applications to exascale platforms, but although DOE has issued an initial call for proposals in 2023, there is not yet a plan to sustain it.

FINDING 2.4: Cloud providers are engaged in hardware and software innovations and will have more market influence in both technology and talent but are not aligned with NNSA requirements.

National Academies of Sciences, Engineering, and Medicine. 2023. *Charting a Path in a Shifting Technical and Geopolitical Landscape: Post-Exascale Computing for the National Nuclear Security Administration*. Washington, DC: The National Academies Press. https://doi.org/10.17226/26916.

HYPERION RESEARCH

QUESTIONS?

bsorensen@hyperionres.com

Insufficient facts always invite danger.

- Spock, *Stardate: 3141.9.*

State of HPC Storage and Interconnects

SC23 Virtual Breakfast Briefing November 2023

Mark Nossokoff

www.HyperionResearch.com
www.hpcuserforum.com

Hyperion Research's 2023 Predictions

1. Strong growth in the leadership-class segment will support modest growth across the global on-premises HPC market.
2. The advanced computing sector and its associated supply chain will become increasingly driven by national and regional government policies that stress domestic capabilities.
3. Sustainability and energy efficiency considerations will become a dominant factor in many procurements.
4. Cloud utilization will shift towards production workloads leading to initial erosion of on-premises spending in low end of the market.
5. 2023 will be the year of AI regulation.
6. AI will become more pervasive in production tier deployments due to users' higher confidence in its abilities and ease of use.
7. HPC system architectures will bifurcate between systems optimized for one set of applications and those designed to address many.
8. Divergent requirements of traditional and modern workloads will move architectural focal points from compute to interconnects and storage systems.
9. Interest in edge computing for HPC will rise in 2023, especially in the industry sector, but spending will be muted.
10. Growth at many HPC sites will be stunted due to the continued difficulty in acquiring and retaining talent.

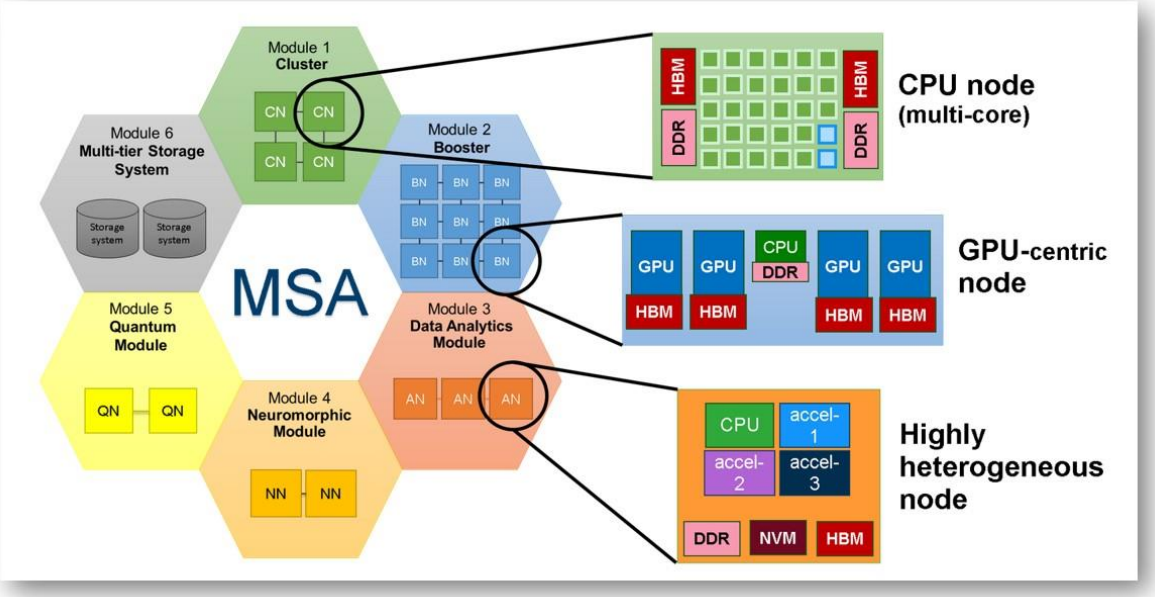
HPC System Architecture Changes

HPC system architectures will bifurcate between systems optimized for one set of applications and those designed fora myriad of applications

- **Future system designs for HPC users will factor in new requirements:**
  - New workloads, like AI and big data
  - New areas of research
  - New anticipated scale of data and computation
- **Major system decisions will split between:**
  - Support of a much larger and diverse set of building blocks
  - Single, heterogeneous system to address wide set of applications
  - Multiple, smaller systems with specific applications in mind
  - Public clouds for specific sets of applications
- **Heterogeneous systems will incorporate:**
  - Data intensive vs. processing intensive designs
  - Node configurations to multiple accelerators and expanded memory profiles
  - Infrastructure accelerators (e.g., DPUs), to processes from CPUs/GPUs
  - Complex storage infrastructure to address different I/O profiles
- **Smaller systems will be designed to target applications like AI, Big Data, or traditional modeling/simulation**
  - AI systems will most likely have more accelerated nodes
  - This scenario requires data centers to be knowledgeable of the requirements of novel and established applications

EU Plans for First Exascale System

Large HPC systems will be built with very diverse hardware building blocks



Source: https://www.fz-juelich.de/en/ias/jsc/about-us/structure/divisions/technology-division/next-gen-arch-proto/msa

- **Potential RFP benchmarks (24 in total) hinting at anticipated workloads**
  - Traditional Computing: Graph 500, HPCG, HPL, Stream
  - Accelerated: NekRS, a GPU Navier-Stokes Solver
  - Quantum
    - JUQCS,JUQCS--G
  - AI
    - AI-NLP (GPT)
    - AI-CNN (ResNet)

Storage and Interconnects: A New Architectural Focal Point

The divergent requirements of traditional HPC modeling/simulation and AI workloads will move HPC architectural focal points from compute to system interconnects and storage systems

- **Internode system interconnects will be critical for performance and scalability of composable system elements**
  - InfiniBand and Ethernet dominance is expected to continue
  - Shift from independent node-node and storage networks to converged networks
- **Intranode interconnects such as CXL are emerging to address composable memory**
- **Storage architectures are evolving to address broad challenges across the entire ecosystem**
  - Compute-intensive vs. data-intensive
  - IO profiles (large block sequential vs. small block random)
  - Access methods (file vs. block vs. object)
  - Access frequency (hot vs. archive vs. cold)
  - Locality (centralized datacenter vs. cloud vs. edge)
  - Enforced consistency (strict POSIX vs. relaxed POSIX)

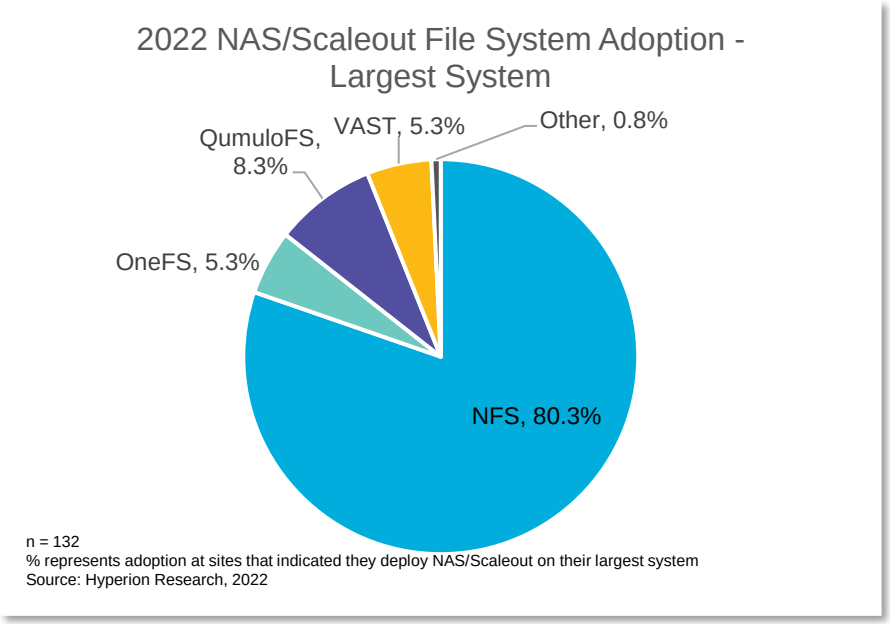
Large Implications for File Systems

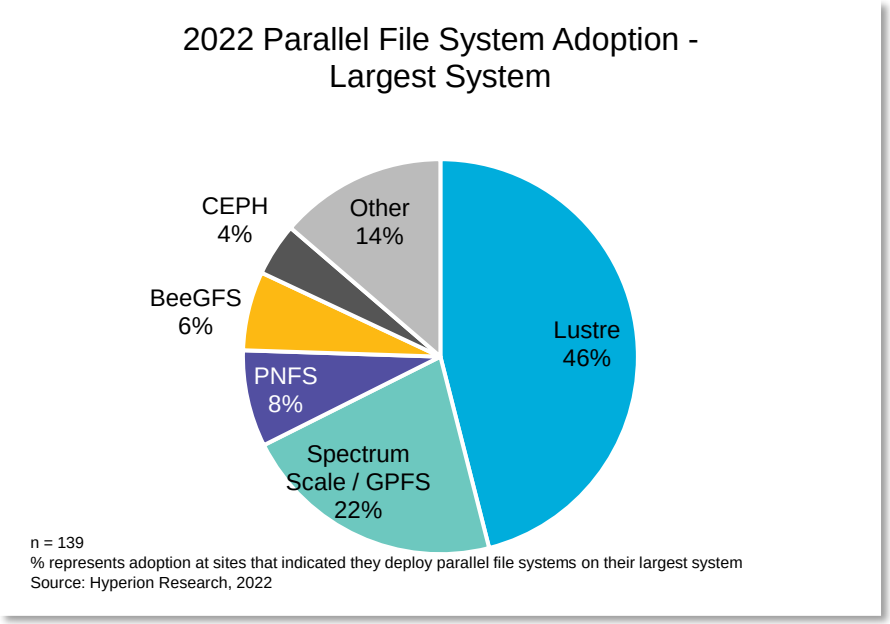
Lustre, Spectrum Scale, NFS currently dominate

- **Systems typically require multiple file systems to address the variety of I/O profiles**
  - Compute-intensive vs. data-intensive
  - IO profiles
  - Access methods
  - Access frequency
  - Locality
  - Enforced consistency
- **CSPs providing support for parallel file systems**
  - AWS: FSx for Lustre
  - Google: Parallelstore (based on DAOS)
  - Microsoft Azure: Managed Lustre Service
- **Advancements are occurring in each file systems but not equally across each area**
- **In search of the elusive global parallel file system**
  - Captive storage system investments
  - Independent start-ups

File System Preferences – Largest System

Large % of sites adopt both NAS/scaleout and parallel file systems on their largest systems





- ***Note: Data is from 2022 Global Site Survey. 2023 Global Site Survey will launch in December.***

Interconnect Snapshot Since ISC23

- **Heightened Regional Focus**
  - EuroHPC call for *Innovation Action in Low Latency and High Bandwidth Interconnects*
- **Many moving parts within the interconnect ecosystem**
  - NVIDIA IB, Ethernet, DPU,
  - Broadcom Jericho3 AI switch and PCIe Gen6&7 roadmap
  - Rockport Networks re-branding as Cerio and releases new PCIe platform
  - Cornelis Networks continuing OPA roadmap investment
  - CSPs and hyperscalars making internal investments
  - HPE Slingshot maturing (Frontier, Aurora)
  - Strong ecosystem and broad emergence of products for CXL
- **Ethernet gaining momentum for increased adoption in HPC/AI**
  - Evolution driven by AI impacts to networks
  - Ultra Ethernet Consortium
  - ETH Zurich research
- **Chiplet interconnects**
  - UCIe
  - OCP BoW
  - Start-up innovations
- **Further investment in and deployment of optical I/O**
  - Ayar Labs funding
  - Lightmatter funding
  - Lightelligence products and solutions
  - Google Apollo OCS (Optical Circuit Switching)

EuroHPC call for *Innovation Action in Low Latency and High Bandwidth Interconnects**

Support the R&I technology development of innovative and competitive European HPC inter-node interconnect technology

- **Call open from August 1, 2023 – January 31, 2024**
- **Proposals should outline ability to:**
  - Develop a roadmap for European scalable inter-node interconnects targeting HPC exascale and post-exascale systems. The roadmap should take into account the EuroHPC supported work in this area such as the components being developed in the EuroHPC RED-SEA project as well as in the area of processors and accelerators.
  - Develop the inter-node interconnect hardware addressing design, development, testing and tape-out as well as integration in test-beds. The work should foster synergies with the EuroHPC supported work in the area of processors and accelerators.
  - Develop the software, installation, configuration and management tools for the developed interconnect, driven by the needs of relevant HPC workflows and application requirements.
  - Address issues like high bandwidth, low latency, power efficiency, virtualisation, scalability, reliability, security, etc.

* Source: https://eurohpc-ju.europa.eu/innovation-action-low-latency-and-high-bandwidth-interconnects_en

Report Card

High marks for both predictions

- **Taking the following together…**
  - Jupiter's architecture
  - Industry and ecosystem happenings since ISC23
  - Identification of interconnect technology of key strategic importance to EuroHPC

- **…supports pretty high grades for both the system architecture prediction and the storage and interconnect prediction.**
- **Looking towards 2024**
  - Continued and expanded storage and interconnect research
  - A new slate of predictions in January

HYPERION RESEARCH

Questions?

mnossokoff@hyperionres.com

The Global QC Market: Realistic and Steady Growth Ahead

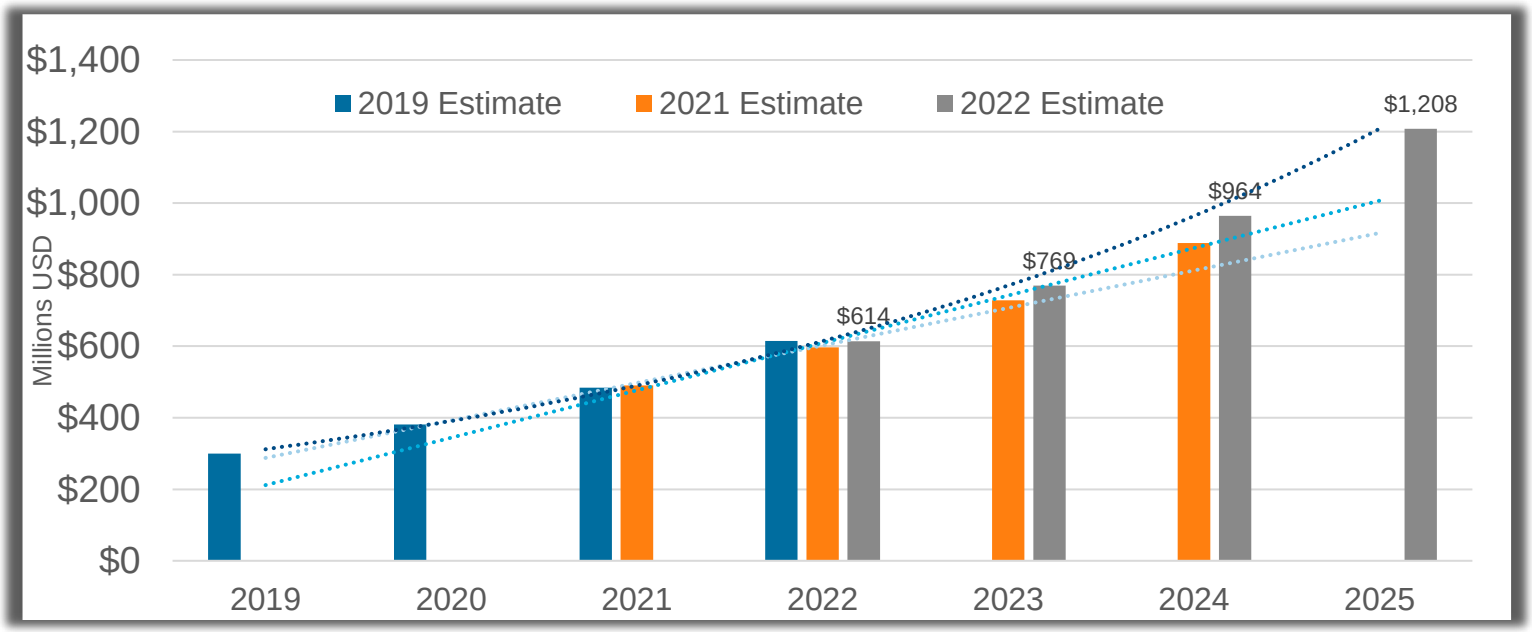
Bob Sorensen
Chief Analyst for Quantum Computing
Hyperion Research, LLC

www.HyperionResearch.com
www.hpcuserforum.com

QC Market Highlights

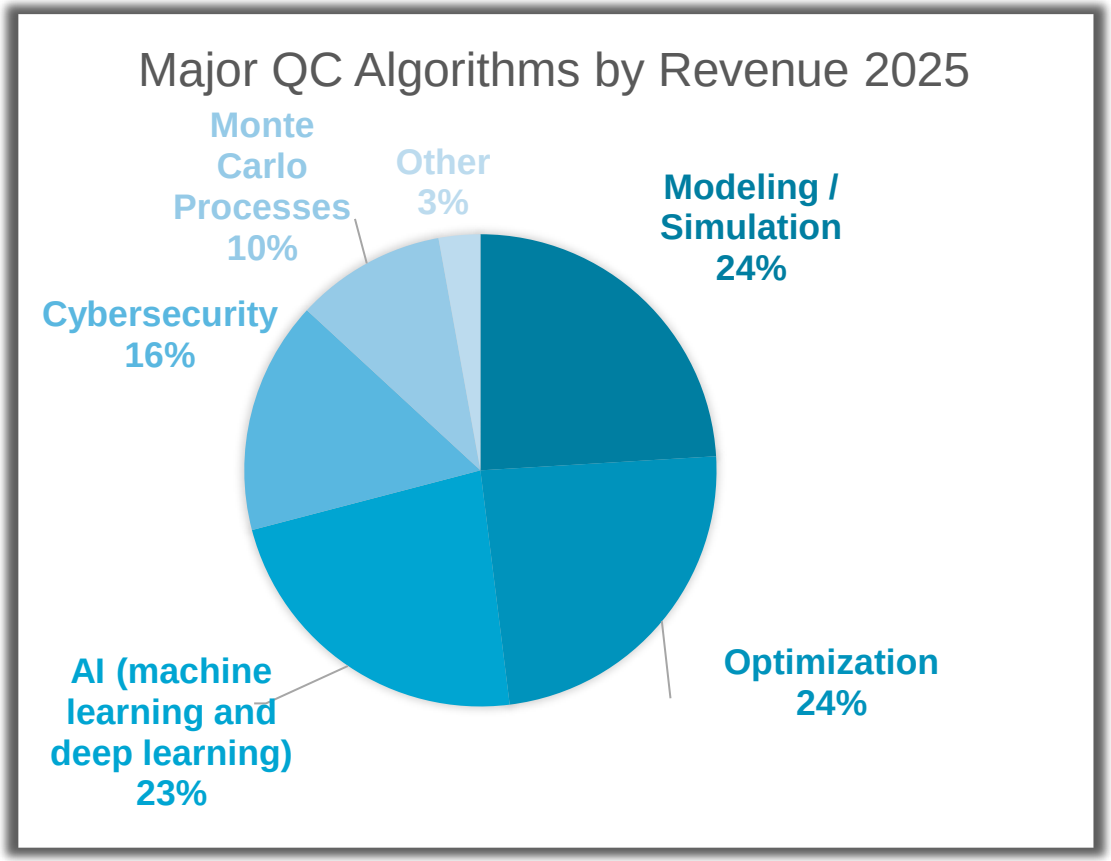
Building a data driven market forecast

- **The global QC market estimated to be worth $770 million USD in 2023**
  - Based on a survey of 145 respondents, 18 different countries, 108 QC companies,
- **Projected to grow at an annual rate of 25.3% out to 2025**
  - Driving the global QC market to approximately $1.2 billion USD in 2025



QC Market 2025: Major Algorithms by Revenue

Mod/sim, optimization, and AI share significant and near equal presence

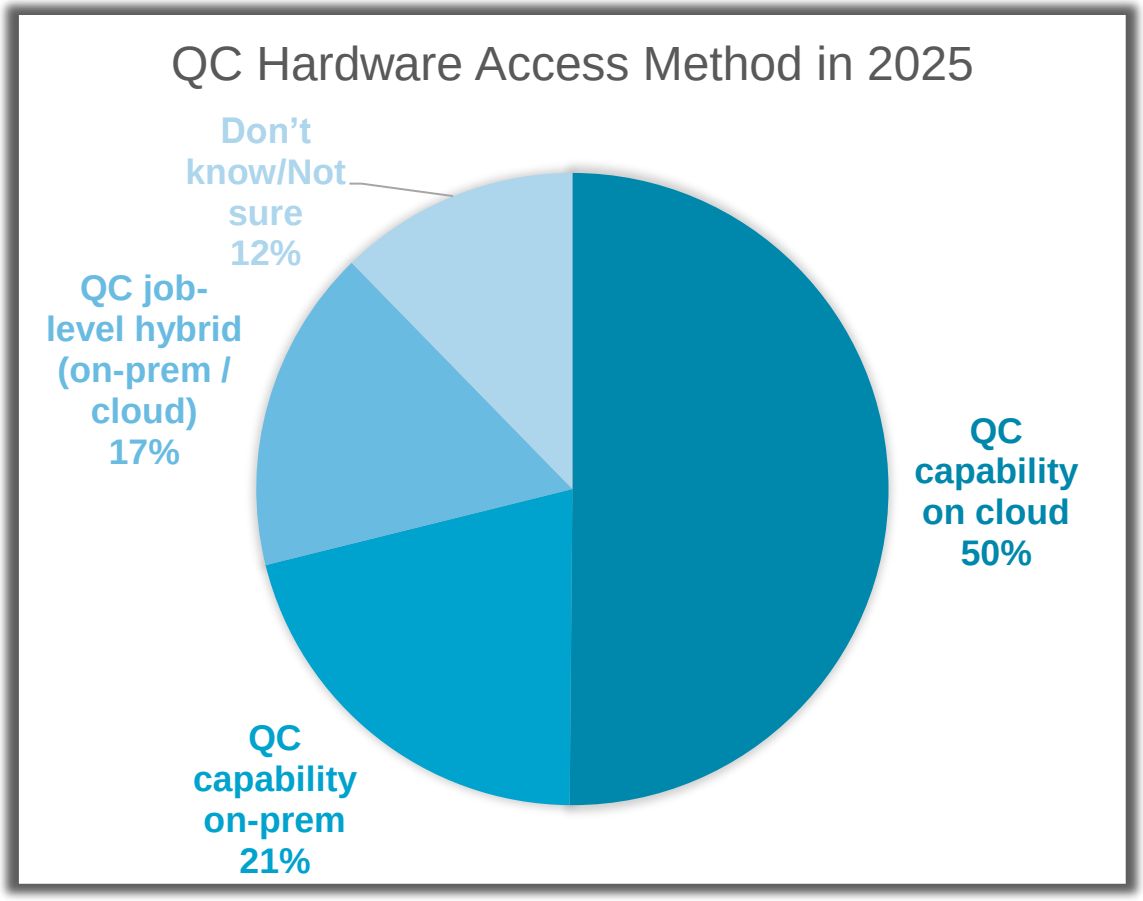


- **No major changes from last year's study**
  - **Slight downturn in machine learning**
  - **Slight upticks in mod/sim and optimization**
- **No specifics on "Other"**
- **Is this list really the complete set?**

N = 113

QC Market 2025: QC Access Method

Cloud continues to dominate as preferred QC hardware access method

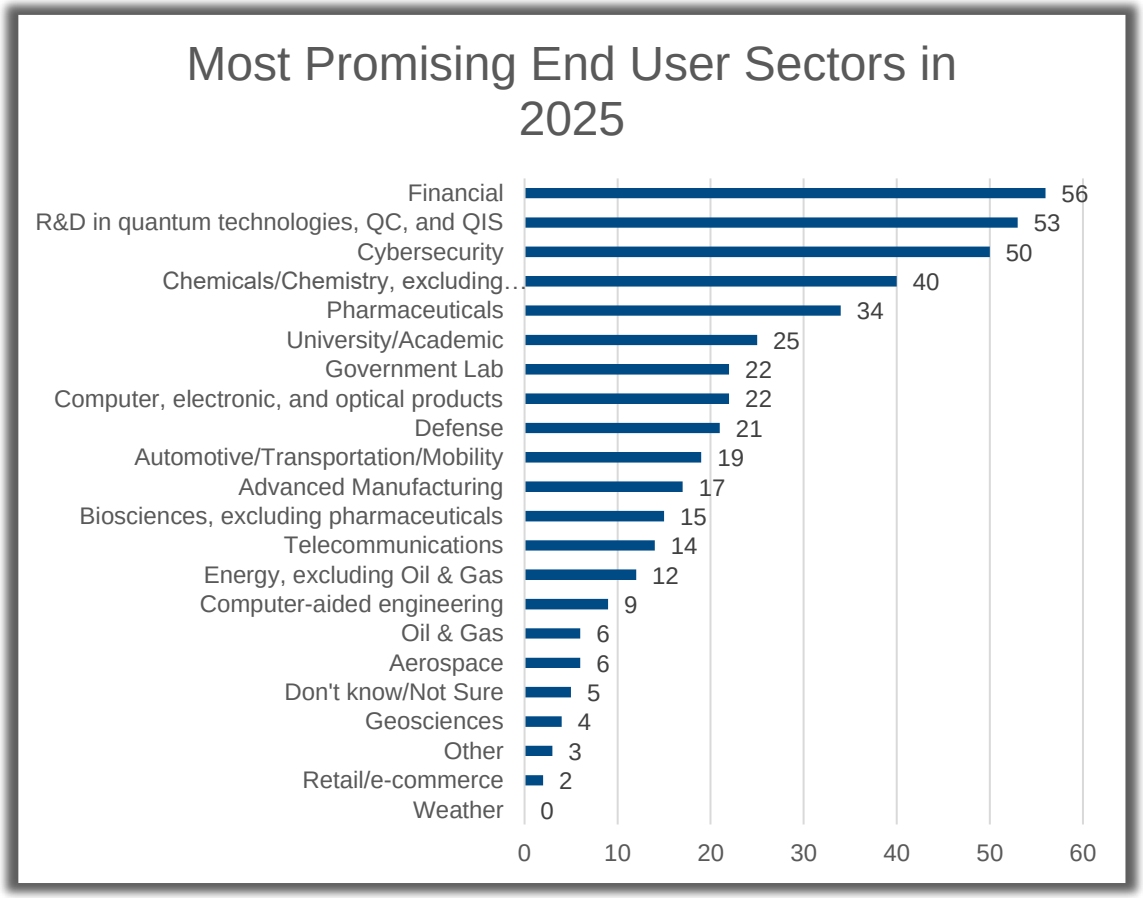


- **Cloud supports half of total QC hardware access**
  - **Combined with hybrid, cloud involved in 67% of total QC hardware market access**
- **No major changes from last year's**

N = 145

QC Market 2025: Top Three End User Sectors

Financial, QC R&D and cybersecurity on top, but broad applicability envisioned



- **Financial chosen by one in three respondents as a top three most promising QC end user sector**
  - **But only narrowly ahead of QC R&D and cybersecurity**
- **Nearly every sector choice deemed important by some**

N = 145, Select top three

QC Early Adopter Study Summary

- **Wide range of QC-related activities currently underway in commercial sector**
  - **485** organizations contacted to locate 300 QC early adopters (62%)
    - **34%** exploring options and monitoring technology
    - **26%** conducting use case analysis
    - **14%** engaged in production use
- **Greatest hurdles to QC adoption are complexities with integrating QC into existing IT infrastructure and clearly demonstrating QC ROI**
- **Positive QC adopter plans for next 2-3 years**
  - More than half are looking at measured growth
  - One-third see aggressive efforts
  - Only a small number **(2%)** indicate disillusionment with QC

Some Closing Thoughts

Time permitting

- **National sovereignty concerns are on the rise and could become invasive to the progress of the sector writ large**
- **QC professional services could become a critical contributor to the prospects of the sector**
  - Composed of a mix of QC and non-QC-specific consulting entities
- **As in HPC, large government procurements could soon alter the trajectory of the sector**
  - A double-edged sword in the making?
- **Questions about continued VC investment (currently at an impressive magnitude and rate of growth) is causing no small degree of angst**
  - Many private firms will continue to rely on private funding instruments for the near future

HYPERION RESEARCH

QUESTIONS?

bsorensen@hyperionres.com

Too great a burden of knowledge can clog the wheels of imagination. When a distinguished but elderly scientist states that something is possible, he is almost certainly right. When he states that something is impossible, he is very probably wrong.

-Arthur C. Clark

HPC Applications and Verticals

SC23 Virtual Breakfast Briefing
November 2023

Melissa Riddle

www.HyperionResearch.com
www.hpcuserforum.com

HYPERION RESEARCH

Applications and Verticals Forecasts

HPC Applications and Middleware Revenue Forecast

HPC applications and middleware revenue expected to collectively exceed $9 billion by 2027

- **Today, applications (e.g., ISV software) and middleware (e.g., compilers, schedulers) make up 22% of the HPC on-premises broader market**
- **Middleware growth (7.0% CAGR) is expected to slightly outpace applications growth (6.5% CAGR)**
  - Both are expected to grow slightly slower than the market writ large (7.6% CAGR)

Revenues by the Broader HPC Market Areas

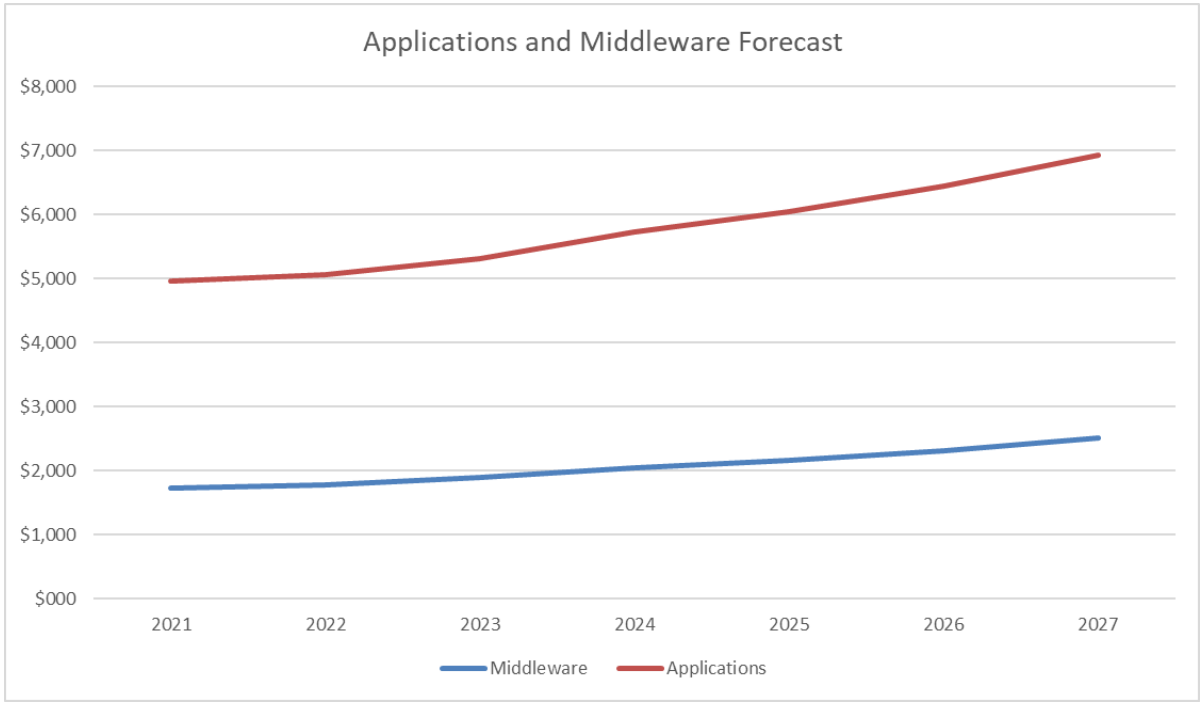
	2021	2022	2023	2024	2025	2026	2027	CAGR 22-27
Server	$14,781	$15,369	$16,486	$18,113	$19,369	$20,894	$22,586	8.0%
Storage	$5,985	$6,380	$6,924	$7,759	$8,434	$9,161	$10,007	9.4%
Middleware	$1,733	$1,781	$1,887	$2,049	$2,160	$2,316	$2,503	7.0%
Applications	$4,960	$5,069	$5,320	$5,729	$6,045	$6,446	$6,935	6.5%
Service	$2,272	$2,214	$2,220	$2,286	$2,323	$2,344	$2,508	2.5%
Total Revenue	$29,731	$30,813	$32,836	$35,936	$38,331	$41,161	$44,539	7.6%

Source: Hyperion Research, November 2023

HPC Applications and Middleware Revenue Forecast (cont'd.)

HPC applications and middleware revenue expected to collectively exceed $9 billion by 2027

- **Applications software represents the majority of these revenues at $6.9 billion in 2027**



HPC Server Forecast by Vertical

Most verticals expected to exceed $1B in servers each annually by 2027

- **Fastest growing verticals include Other, Weather, CAE, Geosciences, University/Academic, and Bio-Sciences**

2022-2027 HPC Server Forecast by Vertical ($M)			
	2022	2027	CAGR 22-27
Bio-Sciences	$1,443	$2,132	8.1%
CAE	$1,760	$2,830	10.0%
Chemical Engineering	$172	$247	7.5%
DCC & Distribution	$822	$1,187	7.6%
Economics/Financial	$755	$1,071	7.2%
EDA / IT / ISV	$870	$1,205	6.7%
Geosciences	$992	$1,490	8.5%
Mechanical Design	$57	$75	5.9%
Defense	$1,597	$2,259	7.2%
Government Lab	$3,324	$4,438	6.0%
University/Academic	$2,648	$3,972	8.4%
Weather	$697	$1,183	11.1%
Other	$232	$495	16.3%
Total Revenue	$15,369	$22,586	8.0%
Source: Hyperion Research, November 2023			

HPC Verticals Forecast

Looking at the total HPC broader market (including cloud), most verticals represent over $1B each today

- **The largest verticals are driven by on-premises revenues: Government Lab, University/Academic, CAE, Bio-Sciences, and Defense**

2022 HPC Revenues by Vertical ($M)

	On-Premises Servers	On-Premises Broader Market	Cloud	Total HPC (On-Premises + Cloud)
Bio-Sciences	$1,443	$2,892	$1,557	$4,450
CAE	$1,760	$3,530	$1,289	$4,818
Chemical Engineering	$172	$345	$154	$499
DCC & Distribution	$822	$1,648	$347	$1,995
Economics/Financial	$755	$1,513	$398	$1,911
EDA / IT / ISV	$870	$1,744	$443	$2,187
Geosciences	$992	$1,989	$403	$2,392
Mechanical Design	$57	$113	$27	$141
Defense	$1,597	$3,201	$471	$3,672
Government Lab	$3,324	$6,665	$443	$7,107
University/Academic	$2,648	$5,308	$276	$5,585
Weather	$697	$1,398	$184	$1,582
Other	$232	$466	$141	$607
Total Revenue	$15,369	$30,813	$6,132	$36,945

Source: Hyperion Research, November 2023

HPC Verticals Forecast (cont'd.)

All verticals now expect higher growth in the cloud than their respective on-premises growth rates

- **Fastest growing verticals include Other, Weather, CAE, Geosciences, and Economics/Financial**

2022-2027 Forecasted Growth Rates by Vertical

	On-Premises Servers	On-Premises Broader Market	Cloud	Total HPC (On-Premises + Cloud)
Bio-Sciences	8.1%	7.8%	12.6%	9.6%
CAE	10.0%	9.6%	18.5%	12.3%
Chemical Engineering	7.5%	7.2%	11.7%	8.7%
DCC & Distribution	7.6%	7.3%	16.1%	9.0%
Economics/Financial	7.2%	6.9%	21.2%	10.5%
EDA / IT / ISV	6.7%	6.4%	17.1%	8.9%
Geosciences	8.5%	8.1%	21.6%	10.9%
Mechanical Design	5.9%	5.6%	10.6%	6.6%
Defense	7.2%	6.8%	18.3%	8.6%
Government Lab	6.0%	5.6%	25.0%	7.3%
University/Academic	8.4%	8.1%	14.7%	8.5%
Weather	11.1%	10.8%	29.7%	13.7%
Other	16.3%	16.0%	25.3%	18.4%
Total Revenue	8.0%	7.6%	18.1%	9.7%

Source: Hyperion Research, November 2023

HYPERION RESEARCH

Applications and Verticals Trends

Applications Trends

Software opportunities and challenges

- **Increased cloud adoption increases concern about software licensing terms**
  - HPC users continue to report that application availability and/or pricing is a significant limitation on cloud growth
  - Users are likely to select schedulers with favorable licensing terms both on-premises and in the cloud (e.g., SLURM)
- **Operating system upheaval continues with rippling effects from Red Hat's CentOS evolution**
  - Some end users evaluating multiple operating systems
  - Has led to the development of open source partnerships, an outpouring of pledges, etc.

Verticals Trends: AI Adoption

Verticals with highest rates of AI adoption include DCC, Government Lab, Bio-Sciences, and Defense

- **Verticals with highest AI revenues include Government Lab, University/Academic, Defense, and Bio-Sciences**

2022 AI HPC Revenues by Vertical ($M)

	All HPC Servers	HPC AI Servers	Vertical as % of HPC	Vertical as % of AI	AI as % of HPC
Bio-Sciences	$1,443	$215	9.4%	12.5%	14.9%
CAE	$1,760	$110	11.5%	6.4%	6.2%
Chemical Engineering	$172	$17	1.1%	1.0%	10.0%
DCC & Distribution	$822	$178	5.3%	10.3%	21.6%
Economics/Financial	$755	$48	4.9%	2.8%	6.4%
EDA / IT / ISV	$870	$23	5.7%	1.4%	2.7%
Geosciences	$992	$28	6.5%	1.6%	2.8%
Mechanical Design	$57	$	0.4%	0.0%	0.0%
Defense	$1,597	$217	10.4%	12.6%	13.6%
Government Lab	$3,324	$580	21.6%	33.8%	17.4%
University/Academic	$2,648	$257	17.2%	15.0%	9.7%
Weather	$697	$30	4.5%	1.7%	4.3%
Other	$232	$15	1.5%	0.9%	6.4%
Total Revenue	$15,369	$1,718	100.0%	100.0%	11.2%

Source: Hyperion Research, November 2023

HYPERION RESEARCH

Questions?

mriddle@hyperionres.com

SC23 Update Conclusions

SC23

www.HyperionResearch.com
www.hpcuserforum.com

Earl Joseph

Overall Conclusions

- **2023 is expected to be a healthy growth year**
  - The first half of 2023 grew at 9.5%
  - Generative AI & LLMs are major growth drivers
  - Exascale systems will help drive growth in 2023-2025
  - GPUs, cloud, & other AI/ML/DL are high growth areas
- **New technologies are showing large numbers:**
  - Processors, AI hardware & software, memories, new storage approaches, Quantum, etc.
  - Composability may fit well in certain applications
- **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 still concerns about the supply chain and growing concerns around power/sustainability & talent**

Conclusions: Questions That We Are Exploring

- **How will new and aggressive computational demands for AI alter the overall HPC hardware and software market landscape?**
  - Where will LLMs have the greatest impact over the next few years?
- **Is this the year the quantum computing becomes a self-sustaining industry sector?**
  - What will be the QC use cases that attract the most end use attention?
- **How are the growing capabilities of HPC in the cloud changing the overall HPC sector?**
  - What are the key opportunities and challenges in operating an efficient effective hybrid on-prem/cloud environment?
  - When will the next tipping point happen in the use of clouds for HPC & AI?
- **What will be the path for post-exascale system evolution, and how will it impact mainstream HPC design and end use?**
- **What new or emerging technologies will drive future HPC & AI performance, power and/or sustainability gains?**
  - Composable computing? Optical I/O? New CPUs & accelerators? Cooling and/or packaging advances?

HYPERION RESEARCH

Thank You For Joining Us Today!

We Welcome Questions, Comments And Suggestions

You can contact us at:

info@hyperionres.com