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
Example Research Areas


- Traditional HPC
- AI, ML, DL, Graph
- Cloud Computing
- Storage & Data
- Interconnects
- Software & Applications
- Power & Cooling
- Tracking all Processor Types & Growth rates
- Quantum Computing
- R&D and Engineering -- all types
- Edge Computing
- Supply Chain Issues

© Hyperion Research 2022
Agenda

• HPC Market Update
• A Quick Update on Exascale Systems
• HPC in the Cloud Update
• Storage and Interconnects Update
• Multi-Client Study Results
• HPC Applications Update
• Global AI Update
• The Quantum Computing Sector
• HPC Innovation Awards
• Conclusions
HPC Market Update
Top Trends in HPC

The first half of 2022 was soft due to supply chain issues – $6.6 billion (US$) in revenues
  • We still expect the second half of the year to do better
  • 2021 grew by a strong 9%
  • Supercomputers are growing better at 5.3% (1H2022)

There are many high growth areas:
  • AI, ML and DL
  • Using clouds to run HPC workloads
  • GPUs
  • Storage

Software continues to be a major area that needs improvement
The lack of technical experts is creating a major constraint
The Overall HPC Market in 2021

Looking at the overall HPC market, including servers, cloud usage, storage, software and repair services = $34.8 billion USD

2021 HPC Market
($34.8 Billion)

- Servers: 42.4%
- Cloud: 14.7%
- Applications: 14.2%
- Middleware: 5.0%
- Storage: 17.2%
- Service: 6.5%
The 2021 Worldwide On-Prem HPC Server Market: $14.8 Billion (up 9%)

2022 is projected to reach $16 to $17 Billion

- Supercomputers (Over $500K): $6.9B
- Divisional ($250K - $500K): $2.8B
- Departmental ($100K - $250K): $3.7B
- Workgroup (under $100K): $1.4B

© Hyperion Research 2022
## WW HPC On-Prem Market By Vendor ($ Millions)

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Server Revenues ($M)</th>
<th>Market Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPE</td>
<td>5,050</td>
<td>34.2%</td>
</tr>
<tr>
<td>Dell Technologies</td>
<td>3,213</td>
<td>21.8%</td>
</tr>
<tr>
<td>Lenovo</td>
<td>1,174</td>
<td>8.0%</td>
</tr>
<tr>
<td>Inspur</td>
<td>993</td>
<td>6.7%</td>
</tr>
<tr>
<td>Atos</td>
<td>542</td>
<td>3.7%</td>
</tr>
<tr>
<td>Sugon</td>
<td>525</td>
<td>3.6%</td>
</tr>
<tr>
<td>IBM</td>
<td>463</td>
<td>3.1%</td>
</tr>
<tr>
<td>Penguin</td>
<td>378</td>
<td>2.6%</td>
</tr>
<tr>
<td>Fujitsu</td>
<td>176</td>
<td>1.2%</td>
</tr>
<tr>
<td>NEC</td>
<td>173</td>
<td>1.2%</td>
</tr>
<tr>
<td>Other</td>
<td>2,076</td>
<td>14.1%</td>
</tr>
<tr>
<td><strong>Total On-Prem HPC</strong></td>
<td><strong>14,763</strong></td>
<td><strong>100.0%</strong></td>
</tr>
</tbody>
</table>

*Source: Hyperion Research, 2022*
## WW HPC Market By Vertical ($ Millions)

<table>
<thead>
<tr>
<th>2021 WW On-Prem High-Performance Systems Revenue by Applications ($M)</th>
<th>2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bio-Sciences</td>
<td>1,455</td>
</tr>
<tr>
<td>CAE</td>
<td>1,767</td>
</tr>
<tr>
<td>Chemical Engineering</td>
<td>177</td>
</tr>
<tr>
<td>DCC &amp; Distribution</td>
<td>807</td>
</tr>
<tr>
<td>Economics/Financial</td>
<td>703</td>
</tr>
<tr>
<td>EDA / IT / ISV</td>
<td>849</td>
</tr>
<tr>
<td>Geosciences</td>
<td>1,010</td>
</tr>
<tr>
<td>Mechanical Design</td>
<td>59</td>
</tr>
<tr>
<td>Defense</td>
<td>1,552</td>
</tr>
<tr>
<td>Government Lab</td>
<td>2,866</td>
</tr>
<tr>
<td>University/Academic</td>
<td>2,637</td>
</tr>
<tr>
<td>Weather</td>
<td>681</td>
</tr>
<tr>
<td>Other</td>
<td>199</td>
</tr>
<tr>
<td>Total Revenue</td>
<td>14,763</td>
</tr>
</tbody>
</table>

*Source: Hyperion Research, 2022*
The Broader On-premise Market Areas ($ Millions)

2021 total on-prem HPC spending reached $29.7B (excluding cloud spending, which brings it to $34.8B)

<table>
<thead>
<tr>
<th>Revenues by the Broader HPC Market Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>2019</strong></td>
</tr>
<tr>
<td>Server</td>
</tr>
<tr>
<td>13,368</td>
</tr>
<tr>
<td>Storage</td>
</tr>
<tr>
<td>5,288</td>
</tr>
<tr>
<td>Middleware</td>
</tr>
<tr>
<td>1,572</td>
</tr>
<tr>
<td>Applications</td>
</tr>
<tr>
<td>4,569</td>
</tr>
<tr>
<td>Service</td>
</tr>
<tr>
<td>2,181</td>
</tr>
<tr>
<td>Total Revenue</td>
</tr>
<tr>
<td>26,979</td>
</tr>
<tr>
<td><strong>2020</strong></td>
</tr>
<tr>
<td>Server</td>
</tr>
<tr>
<td>13,523</td>
</tr>
<tr>
<td>Storage</td>
</tr>
<tr>
<td>5,079</td>
</tr>
<tr>
<td>Middleware</td>
</tr>
<tr>
<td>1,491</td>
</tr>
<tr>
<td>Applications</td>
</tr>
<tr>
<td>4,315</td>
</tr>
<tr>
<td>Service</td>
</tr>
<tr>
<td>2,015</td>
</tr>
<tr>
<td>Total Revenue</td>
</tr>
<tr>
<td>26,423</td>
</tr>
<tr>
<td><strong>2021</strong></td>
</tr>
<tr>
<td>Server</td>
</tr>
<tr>
<td>14,763</td>
</tr>
<tr>
<td>Storage</td>
</tr>
<tr>
<td>5,984</td>
</tr>
<tr>
<td>Middleware</td>
</tr>
<tr>
<td>1,731</td>
</tr>
<tr>
<td>Applications</td>
</tr>
<tr>
<td>4,952</td>
</tr>
<tr>
<td>Service</td>
</tr>
<tr>
<td>2,267</td>
</tr>
<tr>
<td>Total Revenue</td>
</tr>
<tr>
<td>29,697</td>
</tr>
</tbody>
</table>

Source: Hyperion Research, 2022
HPC Market Forecasts
## 5-Year On-Prem HPC Server Forecast

6.9% growth over the next 5 years

<table>
<thead>
<tr>
<th>$US Millions</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
<th>CAGR 21-26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supercomputer</td>
<td>6,926</td>
<td>8,125</td>
<td>9,023</td>
<td>9,846</td>
<td>9,661</td>
<td>9,543</td>
<td>6.6%</td>
</tr>
<tr>
<td>Divisional</td>
<td>2,803</td>
<td>3,077</td>
<td>3,397</td>
<td>3,681</td>
<td>3,677</td>
<td>4,397</td>
<td>9.4%</td>
</tr>
<tr>
<td>Departmental</td>
<td>3,648</td>
<td>3,909</td>
<td>4,316</td>
<td>4,645</td>
<td>4,704</td>
<td>5,137</td>
<td>7.1%</td>
</tr>
<tr>
<td>Workgroup</td>
<td>1,373</td>
<td>1,392</td>
<td>1,472</td>
<td>1,525</td>
<td>1,450</td>
<td>1,471</td>
<td>1.4%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>14,763</td>
<td>16,503</td>
<td>18,208</td>
<td>19,697</td>
<td>19,492</td>
<td>20,549</td>
<td>6.9%</td>
</tr>
</tbody>
</table>
5-year HPC Server Forecast Changes

Quarterly adjustments due to covid

Worldwide HPC On-Premises Server Market Revenues

- Early covid concerns
- Extended supply chain issues
- Exascale prices drop

© Hyperion Research 2022
## New Supercomputer Subsegments

<table>
<thead>
<tr>
<th>New Supercomputer Subsegments</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
<th>CAGR 21-26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership/Exascale-class Systems</td>
<td>0</td>
<td>1,065</td>
<td>1,000</td>
<td>2,050</td>
<td>2,300</td>
<td>2,450</td>
<td>2,000</td>
<td>2,150</td>
<td>16.5%</td>
</tr>
<tr>
<td>Supercomputers-Large ($3M and up)</td>
<td>3,858</td>
<td>3,808</td>
<td>4,477</td>
<td>4,516</td>
<td>4,995</td>
<td>5,420</td>
<td>5,615</td>
<td>5,419</td>
<td>3.9%</td>
</tr>
<tr>
<td>Supercomputers-Entry Level ($500k to $3M)</td>
<td>1,260</td>
<td>1,250</td>
<td>1,450</td>
<td>1,559</td>
<td>1,728</td>
<td>1,975</td>
<td>2,046</td>
<td>1,975</td>
<td>6.4%</td>
</tr>
<tr>
<td><strong>Total Supercomputers ($500K and up)</strong></td>
<td>5,118</td>
<td>6,123</td>
<td>6,926</td>
<td>8,125</td>
<td>9,023</td>
<td>9,846</td>
<td>9,661</td>
<td>9,543</td>
<td>6.6%</td>
</tr>
</tbody>
</table>

*Source: Hyperion Research, May 2022*
High Growth Areas

*These are redefining the HPC sector*

- The use of external clouds for running HPC workloads
- AI, ML and DL
- New processor types and accelerators/GPUs
- Storage
- Exascale-class systems
- Quantum technologies
The HPC Cloud Market Will See Strong Growth in 2022

The growth will build on the fundamental changes in buying behavior seen in 2021

- HPC & AI buyers around the world revealed for the first time that HPC buyers are planning to shift some of their on-premises budgets to spending in the cloud
  - The shift is fundamental because up to 2021 very few sites were taking money from the on-premises budgets for cloud computing
- End user spending on public cloud resources to run HPC workloads is projected to grow substantially in 2022, at a rate greater than 23%, and will exceed US $6.2 billion
- This major shift in buying behavior doesn’t mean that on-premises HPC systems are going away
  - The on-premises HPC server market is anticipated to exhibit healthy growth, 7%-8% a year, over the forecast period
HPC Cloud Usage Forecast

17.6% growth over the next 5 years

Using Clouds For HPC ($M)
HPC-enabled AI Server Forecast

5-year CAGR expected to exceed 22% growth
## Use Of Different AI/ML/DL Approaches

*From our end-user MCS study*

<table>
<thead>
<tr>
<th>Category</th>
<th>Responses</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine learning</td>
<td>99</td>
<td>70.2%</td>
</tr>
<tr>
<td>Deep learning</td>
<td>86</td>
<td>61.0%</td>
</tr>
<tr>
<td>Graph analysis</td>
<td>25</td>
<td>17.7%</td>
</tr>
<tr>
<td>Cognitive computing</td>
<td>24</td>
<td>17.0%</td>
</tr>
<tr>
<td>Semantic analysis</td>
<td>22</td>
<td>15.6%</td>
</tr>
<tr>
<td>Other big data/analytics</td>
<td>41</td>
<td>29.1%</td>
</tr>
<tr>
<td>We don't plan to run applications of these types</td>
<td>9</td>
<td>6.4%</td>
</tr>
</tbody>
</table>

*n = 141

Source: Hyperion Research, 2021
83% Of Sites Have Accelerators Or Co-processors Today

From our end-user MCS study

<table>
<thead>
<tr>
<th>How many co-processors or accelerators are in your largest HPC technical server?</th>
<th>Responses</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>23</td>
<td>17.3%</td>
</tr>
<tr>
<td>Less than 32</td>
<td>28</td>
<td>21.1%</td>
</tr>
<tr>
<td>32 to less than 64</td>
<td>18</td>
<td>13.5%</td>
</tr>
<tr>
<td>64 to less than 100</td>
<td>19</td>
<td>14.3%</td>
</tr>
<tr>
<td>100 to less than 500</td>
<td>18</td>
<td>13.5%</td>
</tr>
<tr>
<td>500 to less than 1,000</td>
<td>11</td>
<td>8.3%</td>
</tr>
<tr>
<td>1,000 to less than 5,000</td>
<td>10</td>
<td>7.5%</td>
</tr>
<tr>
<td>5,000 to less than 10,000</td>
<td>4</td>
<td>3.0%</td>
</tr>
<tr>
<td>10,000 or more</td>
<td>2</td>
<td>1.5%</td>
</tr>
</tbody>
</table>

Source: Hyperion Research, 2021

© Hyperion Research 2022
GPU/Accelerator Forecast

Anticipated high growth for accelerators over next 5 years
Storage Growth Rates

HPC storage is growing quickly, driven by AI, big data and growing modeling/simulation model sizes
Agenda

• HPC Market Update
• A Quick Update on Exascale Systems
• HPC in the Cloud Update
• Storage and Interconnects Update
• Multi-Client Study Results
• HPC Applications Update
• Global AI Update
• The Quantum Computing Sector
• HPC Innovation Awards
• Conclusions
Questions?

We welcome questions, comments and suggestions

Please contact us at: info@hyperionres.com

© Hyperion Research 2022
### The Exascale Market (System Acceptances)

*Over 30 systems and over $10 billion in value*

<table>
<thead>
<tr>
<th>Year Accepted</th>
<th>China</th>
<th>Europe</th>
<th>Japan</th>
<th>US</th>
<th>Other Countries*</th>
<th>Total Systems</th>
<th>Total Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020</td>
<td>2 exascale ~$350M each</td>
<td>1 pre-exascale system ~$180M</td>
<td>1 near-exascale system ~$1 B</td>
<td>1 pre-exascale system ~$200M</td>
<td>--</td>
<td>1</td>
<td>$1.0B</td>
</tr>
<tr>
<td>2021</td>
<td>2 exascale ~$350M each</td>
<td>1 pre-exascale system ~$180M</td>
<td>?</td>
<td>1 near-exascale system ~$200M</td>
<td>--</td>
<td>4</td>
<td>$1.1B</td>
</tr>
<tr>
<td>2022</td>
<td>1 exascale ~$350M each</td>
<td>2 pre-exascale systems ~$190 each</td>
<td>1 near-exascale system ~$150M</td>
<td>1 exascale systems ~$600M</td>
<td>--</td>
<td>5</td>
<td>$1.5B</td>
</tr>
<tr>
<td>2023</td>
<td>1 exascale system ~$350M</td>
<td>1 or 2 pre-exascale systems ~$150M each</td>
<td>1 near-exascale system ~$150M</td>
<td>1 or 2 exascale systems ~$600M each</td>
<td>--</td>
<td>4-5</td>
<td>$1.8B - $2.4B</td>
</tr>
<tr>
<td>2024</td>
<td>1 exascale system ~$350M</td>
<td>1 exascale ~$500M, plus 1 exascale (or pre) systems ~$200 M</td>
<td>?</td>
<td>1 or 2 exascale systems ~$400M each</td>
<td>1 exascale system ~$200M</td>
<td>4-6</td>
<td>$1.2B - $1.9B</td>
</tr>
<tr>
<td>2025</td>
<td>1 or 2 exascale system ~$300M each</td>
<td>1 or 2 exascale systems ~$350M each</td>
<td>1 near-exascale system ~$150M</td>
<td>1 or 2 exascale systems ~$350M each</td>
<td>1 exascale system ~$150M</td>
<td>5-8</td>
<td>$1.3B - $2.3B</td>
</tr>
<tr>
<td>2026</td>
<td>1 or 2 exascale system ~$300M each</td>
<td>1 or 2 exascale systems ~$325M each</td>
<td>?</td>
<td>1 or 2 exascale systems ~$350M each</td>
<td>1 or 2 exascale systems ~$150M each</td>
<td>4-8</td>
<td>$1.1B - $2.2B</td>
</tr>
<tr>
<td>2027</td>
<td>1 or 2 exascale systems ~$250M each</td>
<td>1 or 2 exascale systems ~$300M each</td>
<td>?</td>
<td>1 or 2 exascale systems ~$300M each</td>
<td>1 or 2 exascale systems ~$150M each</td>
<td>4-8</td>
<td>$1.0B - $2.0B</td>
</tr>
</tbody>
</table>

| Total         | 8-11 | 8-12 | 4 | 7-12 | 4-6 | 31-45 | $10B - $14B |

*Includes S. Korea, Singapore, Australia, Russia, Canada, India, Israel, Saudi Arabia, etc.*

*Source: Hyperion Research, July 2022*
A Quick Update on Exascale Systems

SC 22

Bob Sorensen

www.HyperionResearch.com
www.hpcuserforum.com
Near-Term US Exascale Plans

*Three systems over two 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: $R_{peak} = 1.68$ EFlops, $R_{max} = 1.1$ EFlops
  - 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**, ~ 1EF DP sustained
  - 08/21: Polaris testbed system (44PF DP and 1.5EF AI)
  - Cray Shasta architecture with Intel Xeons and Intel Xe GPU
  - Delivery in late 2022, acceptance in 2023 (delayed at least 12 months)

- **El Capitan:** DOE NNSA’s LLNL
  - ~ 2EF
  - Cray Shasta architecture AMD EPYC processors, next generation Radeon Instinct GPUs
  - Fully deployed in 2023
China Exascale Status

The official situation

- **Sunway Pro OceanLight**
  - ~1.3 EFlops Rpeak, ~1.05 EFlops Rmax
  - 35 MW, 38 million cores
  - ShenWei post-Alpha CPU
  - National Supercomputing Center-Wuxi

- **Tianhe-3**
  - Dual-chip FeiTeng ARM and matrix accelerator nodes
  - ~ 1.7 EFlops Rpeak, 1.3 EFlops Rmax
  - NSCC-Tianjin

- **Sugon**
  - Hygon processors (low confidence), may go AMD Zen4
  - NSCC-Shenzhen
China Exascale Status

The unofficial reality?

- **Sunway Pro OceanLight**
  - Up and running since March 2021

- **Tianhe-3**
  - Up and running in last year (?)

- **Sugon**
  - Potentially delayed

- **No official announcements**

- **No entries for June 2021, November 2021, June 2022 Top 500 list**
  - Maybe this time around: likely not
  - Political decision

- **Strong evidence of at least five other Chinese systems that could make top 10 list today**
**EU HPC Plans**

*Exascale plans going forward*

- **EU plan calls for acquisition of two exascale systems in the 2021-2024 timeframe**
  - At least one to use European technology: specifically using an EPI-developed processor
  - Additional procurements in Germany in 2024, 2025
  - EU may include 2 additional ES systems in 2023-2026

- **Post Exascale System around 2027**
  - Plans call for integration and deployment of the first hybrid HPC/quantum infrastructure in Europe

---DRAFT---

<table>
<thead>
<tr>
<th>HPC Infrastructure</th>
<th>2019 &amp; 2020</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
<th>2027</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 pre-exascale + 5 petascale systems</td>
<td>Several mid-range, pre-exascale and 2 exascale systems</td>
<td>exascale and post-exascale HPC systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quantum infrastructure</th>
<th>2019 &amp; 2020</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
<th>2027</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot Quantum simulators interfacing with HPC systems (100+ Quantum units)</td>
<td>QComputer/ QSimulators (NISQ) with Basic HPC integration</td>
<td>QComputer/QSimulators (NISQ) with Full HPC integration - HPC Accelerators</td>
<td>Prototype QComputers fitted with Error Correction and robust Qbits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Leonardo Flores Añover, Senior Expert
DG CNECT, HPC & Quantum Technology Unit - European Commission 2021

© Hyperion Research 2022
SC22: What I’ll Be Looking At

- Mixed precision: not just AI, in traditional mod/sim
- Hybrid quantum/classical efforts
- Post exascale architectures
- Liquid cooling
- End of ‘end of’
QUESTIONS?

bsorensen@hyperionres.com

*Float to the top or sink to the bottom. Everything in the middle is the churn.* - Amos Burton, The Expanse
HPC Cloud Forecast

*HPC cloud revenue is expected to exceed $11 billion by 2026*

- Storage-specific components comprise roughly 1/3 of cloud revenue for HPC
- AI and other data-intensive applications are a high growth segment for cloud adoption in HPC
The Total HPC Market: On-Prem and Cloud Computing

The cloud market is smaller, but growing faster

© Hyperion Research 2022
Cloud Spend Distribution

*Compute comprises more than half of cloud spend*

- **57% of cloud spend focused on compute instances:**
  - Roughly 50/50 split between accelerated and non-accelerated compute instances in the cloud
  - Slightly more spend on reserved instances over preemptive

---

Source: Hyperion Research Multi-Client End User Study 2022
Impact of HPC Cloud

*Organizations are increasingly factoring cloud into future on-premises deployment plans*

- **Public cloud resources have historically been seen as complementary to on-premises**
  - Many longitudinal studies show that cloud is used primarily for burst capabilities by many HPC users
  - This perception is shifting:
    - A recent study showed that almost 50% of the users are altering on-premises deployments due to cloud

- **Migrating HPC workloads to cloud platforms requires new skills for center managers & researchers**
  - Much of the education and training on using the cloud focuses on which workloads can and should be run in the cloud versus remain on-premises
  - IT departments are factoring in data movement and security as they expand their resource pools to consist of cloud resources
No One-Size-Fits-All

Cloud and on-prem decisions depend on variety of factors

- Choosing where to run HPC applications depends on a wide set of factors:
  - Internal skillsets and knowledge
  - Budget and cost
  - Performance characteristics
  - Hardware availability
  - Time constraints
  - Other

- Understanding which applications can be run cost-effectively and performant in the cloud can aid in on-premises system design and optimization
  - What technologies to invest in?
  - What scale to deploy on-premises?
  - Which middleware and software tools are necessary to optimally run HPC applications across platforms?

- There is no single solution for all HPC user sites
Models for Accessing HPC Cloud

- **On-premises Infrastructure**
  - Private
  - Cloud-like

- **Hosted Cloud**
  - Hybrid

- **Enablement Services**
  - Multi-cloud
  - ISV

- **Hosted Infrastructure**
  - Co-lo
  - Managed Service Providers
  - Industry Vertical Providers
  - Hosted HPC Providers

- **Direct to Cloud** (Includes full service and DIY)

© Hyperion Research 2022
Continue the conversation

anorton@hyperionres.com
mnossokoff@hyperionres.com
HPC Storage Growth Continues

Demand increasing across all sectors and verticals

- Storage historically the highest growth HPC element
- Storage represents ~20% of on-premises HPC spending and growing
- Almost half of sites surveyed expect their storage budgets to increase more than 5%

On premises HPC Spend - 2021
Total 2021 HPC Spend: ~ $29.7B

<table>
<thead>
<tr>
<th>Area ($M)</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
<th>CAGR 21-'26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server</td>
<td>$14,748</td>
<td>$16,077</td>
<td>$17,738</td>
<td>$19,565</td>
<td>$19,495</td>
<td>$20,481</td>
<td>6.8%</td>
</tr>
<tr>
<td>Add-on Storage</td>
<td>$5,971</td>
<td>$6,677</td>
<td>$7,457</td>
<td>$8,388</td>
<td>$8,491</td>
<td>$9,027</td>
<td>8.6%</td>
</tr>
<tr>
<td>Middleware</td>
<td>$1,729</td>
<td>$1,863</td>
<td>$2,030</td>
<td>$2,212</td>
<td>$2,172</td>
<td>$2,268</td>
<td>5.6%</td>
</tr>
<tr>
<td>Applications</td>
<td>$4,948</td>
<td>$5,302</td>
<td>$5,731</td>
<td>$6,195</td>
<td>$6,089</td>
<td>$6,326</td>
<td>5.0%</td>
</tr>
<tr>
<td>Service</td>
<td>$2,266</td>
<td>$2,316</td>
<td>$2,389</td>
<td>$2,468</td>
<td>$2,336</td>
<td>$2,296</td>
<td>0.3%</td>
</tr>
<tr>
<td>Total Revenue</td>
<td>$29,662</td>
<td>$32,236</td>
<td>$35,345</td>
<td>$38,828</td>
<td>$38,584</td>
<td>$40,398</td>
<td>6.4%</td>
</tr>
</tbody>
</table>

Source: Hyperion Research, 2021
HPC Storage and the Cloud

Cloud adoption for storage remains strong and growing

- **Storage ~ 1/3 total HPC spending in the cloud**
- **Spending on persistent, durable storage 2x greater than ephemeral, temporal storage**

**2022 Cloud Budget Distribution**

- **Compute Instances 57%**
- **Persistent Storage 23%**
- **Ephemeral Storage 11%**
- **Software Licenses 7%**
- **Other 2%**

**n = 94**

Source: Hyperion Research, 2022

**HPC Cloud and Storage Forecast**

- **$1.7B cloud storage spend in 2021**
- **Cloud storage growth ~ 2.3x on-premises storage growth**
- **Total cloud spending projected to overtake on-premises storage spending in 2024**

© Hyperion Research 2022
2022 HPC Storage Vendor Preferences

*Dell Technologies continues as top preferred storage vendor*

- Dell Technologies first overall; preferred by Industry & Academia
- IBM second preferred overall, tied with HPE/Cray in Government and Academia
- DDN 4th continues as preferred independent storage vendor
- Users continue to prefer sourcing storage from their system providers

© Hyperion Research 2022
Storage Technology Ecosystem
Large % of sites adopt both NAS/scaleout and parallel file systems on their largest systems

2022 NAS/Scaleout File System Adoption - Largest System

- NFS, 80.3%
- OneFS, 5.3%
- QumuloFS, 8.3%
- VAST, 5.3%
- Other, 0.8%

n = 132
% represents adoption at sites that indicated they deploy NAS/Scaleout on their largest system
Source: Hyperion Research, 2022

2022 Parallel File System Adoption - Largest System

- Lustre 46%
- Spectrum Scale / GPFS 22%
- PNFS 8%
- BeeGFS 6%
- Other 14%
- CEPH 4%

n = 139
% represents adoption at sites that indicated they deploy parallel file systems on their largest system
Source: Hyperion Research, 2022
Interconnect Architecture

Subtle shift expected in interconnect implementation

<table>
<thead>
<tr>
<th>Network Architecture</th>
<th>Current</th>
<th>Next Procurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent Networks</td>
<td>54.1%</td>
<td>47%</td>
</tr>
<tr>
<td>Converged Networks</td>
<td>45.9%</td>
<td>53%</td>
</tr>
</tbody>
</table>

• Site’s largest systems architecture preference today skew towards independent system-system and system-storage
• Site’s expressed preference to skew toward converged system-system and system-storage network in next procurement
• What’s happening:
  • Networking rates are increasing to the point of being able to support the demands of converged networks
  • Features are being implemented to optimize overall performance
  • Requirements to support the mixes workloads of compute-intensive and data-intensive workloads are being addressed to alleviate the need and expense of separate networks

© Hyperion Research 2022
Independent Networks: System-System

Ethernet preferred over Infiniband for system-system

- **Ethernet**
  - 44.9% of sites surveyed with independent networks
  - 100Gb overtakes 10Gbit as most widely deployed within Ethernet sites

- **InfiniBand**
  - 35.9% of sites surveyed with independent networks
  - HDR 200 catches up with EDR 100 Gbit as most widely deployed at Infiniband sites

- **Other**
  - Omni-Path = 4.5%
  - Other = 4.5%

* % of site adoption Ethernet and Infiniband, respectively

© Hyperion Research 2021
Independent Networks: **System-Storage**

*Ethernet preferred over Infiniband for system-storage*

- **Ethernet**
  - 47.8% of sites surveyed with independent networks
  - 100Gb most widely deployed at almost half of Ethernet sites

- **InfiniBand**
  - 34.9% of sites surveyed with independent networks
  - EDR 100 Gbit most widely deployed at Infiniband sites

- **Other**
  - Omni-Path = 3.5%
  - Other = 1.2%

* % of site adoption Ethernet and Infiniband, respectively

© Hyperion Research 2021
Converged Networks

*Infiniband preferred over Ethernet for converged networks*

- **Ethernet**
  - 42.5% of sites surveyed with independent networks
  - 100Gb most widely deployed within Ethernet sites

- **InfiniBand**
  - 45.2% of sites surveyed with independent networks
  - EDR 100 Gbit most widely deployed at InfiniBand sites

- **Other**
  - Omni-Path = 4.1%
  - Other = 8.2%

* % of site adoption Ethernet and InfiniBand, respectively

© Hyperion Research 2021
Future Research Direction
Future Research Direction

Broad range of topics across diverse storage ecosystem

- Grow on-premises census data including HDD, solid state and tape
  - Use cases
    - Temporal, durable
    - File, block, object
    - Scratch, user, home directory, project, campaign, archive
  - Capacity
  - Internal to servers and compute nodes

- Expand cloud storage coverage
  - Storage landscape model
  - Workload usage and requirements
  - Impact of containers

- Edge computing implications on storage
  - Capacities
  - Architectures
  - Computational storage

- Intranode interconnects and protocols
  - NVMeoF
  - CXL
  - Memory pooling
  - Chiplet standards and adoption
    - UCIe
    - OCP Bunch of Wires (BoW)

- Memory topics
  - Big memory
  - In-and-near memory computing
PLEASE SHARE YOUR THOUGHTS!

info@hyperionres.com
mnossokoff@hyperionres.com

© Hyperion Research 2022
2022 Multi-Client Study Results
HPC Users Are Optimistic About Budgets, Prioritizing Hardware and Storage

• Half of surveyed users (52%) expect HPC budgets to increase by at least 5% over the next year
  • Top categories included HPC hardware, add-on storage, and public cloud
• Most users reported willingness to pay a 10-15% premium for their desired system attributes
  • Top categories included better processors, larger/faster memory, higher performance external I/O and storage, and better density/power/cooling
• Average site storage capacity increased significantly over the past year among Industry sites
• A quarter of overall public cloud spending in HPC (23%) is persistent storage
• Ethernet continues taking share from InfiniBand

© Hyperion Research 2022
Other Priorities: Acceleration, Data Locality and HPC/AI Expertise

• Many sites report using accelerators on their most used or most important application
• Most cloud compute instances (64%) are reserved and many (43%) are accelerated
• Top limitations for increasing public cloud use are budgets and data locality/speed. Among Industry sites, data security is #1 concern
• Expertise is a top concern for both HPC and AI, outranked only by budget concerns
• Following Red Hat’s change to the CentOS business and support model, Ubuntu Linux has risen in popularity as a free alternative CentOS usage declines
Emerging Technologies Are Gaining Traction: AI, Edge Computing, And Composable Infrastructures

- **AI popularity continues to rise**
  - Virtually all users report plans to use AI methodologies
  - In Industry, AI is overtaking traditional mod/sim as % of workload
  - AI-specific software licenses such as Databricks and Anaconda are almost exclusively being purchased by Industry sites

- **A quarter of users (28%) expect to employ edge computing within 2 years**
  - Top motivators include improving real-time data collection/processing, accelerating HPC applications, access to IoT devices for data collection, and a wider range of sensor data

- **Composable infrastructures are gaining interest and attention, as are DPUs**
HPC Applications Market View

- $4.95 billion (14.2% of the broader market)
- For every dollar spent on servers, applications add 33 cents to the price tag
- The applications market exceeded growth expectations from 2020 to 2021. Last year it was expected to grow by 6.8%, and we’re seeing 14.8% growth so far

### Revenues by the Broader HPC Market Areas

<table>
<thead>
<tr>
<th></th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server</td>
<td>13,368</td>
<td>13,523</td>
<td>14,763</td>
</tr>
<tr>
<td>Storage</td>
<td>5,288</td>
<td>5,079</td>
<td>5,984</td>
</tr>
<tr>
<td>Middleware</td>
<td>1,572</td>
<td>1,491</td>
<td>1,731</td>
</tr>
<tr>
<td>Applications</td>
<td>4,569</td>
<td>4,315</td>
<td>4,952</td>
</tr>
<tr>
<td>Service</td>
<td>2,181</td>
<td>2,015</td>
<td>2,267</td>
</tr>
<tr>
<td><strong>Total Revenue</strong></td>
<td>26,979</td>
<td>26,423</td>
<td>29,697</td>
</tr>
</tbody>
</table>

Source: Hyperion Research, 2022
Migration of CentOS Users

Supporting numbers of the trend to migrate away from CentOS

- Only 37.6% of MCS 2022 respondents report using CentOS this year, with 72.1% of those users reporting plans to migrate from CentOS to another operating system
- Ubuntu Linux is gaining popularity

Migration from CentOS, what OS are you migrating to?

<table>
<thead>
<tr>
<th></th>
<th>Overall Percent</th>
<th>Industry Percent</th>
<th>Government Percent</th>
<th>Academia Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Hat Linux</td>
<td>23.5%</td>
<td>31.6%</td>
<td>12.5%</td>
<td>13.6%</td>
</tr>
<tr>
<td>Rocky Linux</td>
<td>20.6%</td>
<td>15.8%</td>
<td>37.5%</td>
<td>22.7%</td>
</tr>
<tr>
<td>Ubuntu Linux</td>
<td>19.1%</td>
<td>21.1%</td>
<td>25.0%</td>
<td>13.6%</td>
</tr>
<tr>
<td>Alma Linux</td>
<td>4.4%</td>
<td>2.6%</td>
<td>-</td>
<td>9.1%</td>
</tr>
<tr>
<td>SUSE Linux</td>
<td>2.9%</td>
<td>5.3%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Windows</td>
<td>1.5%</td>
<td>2.6%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>I will continue to use CentOS and not migrate</td>
<td>27.9%</td>
<td>21.1%</td>
<td>25.0%</td>
<td>40.9%</td>
</tr>
</tbody>
</table>

Source: Hyperion Research, 2022
Accelerated Applications

Finding meaningful uses for GPUs and accelerators is a priority

• Use of GPUs and accelerators are on the rise among HPC users
• On average, MCS 2022 respondents spend a third (32%) of their HPC cycles on accelerated applications
  • 87.8% of respondents have at least one accelerated critical application
  • The average site has 7.5 accelerated applications
  • 74.6% of respondents use coprocessors/accelerators on their most important or most used application (at least some of the time)
    • Industry respondents lead accelerator/coprocessor usage at 82%
Application Run Time

Government is speeding up

• From the MCS 2022, 29% of the most important or most used applications have a typical runtime of 24 hours or more
  • This is a remarkable shift from last year’s study, where 40% of respondents reported typical runtimes of 24 hours or more for their most important/most used applications
  • The government sector has the most notable change, where last year 61% of most important/most used applications were 24 hours or more, and this year that number has shrunk to 35%

• Growing run time was a key finding of last year’s MCS
QUESTIONS?

info@hyperionres.com
jludema@hyperionres.com
Global AI Update

SC22
November 2022
Alex Norton and Tom Sorensen
On-Prem HPC Market Segmentation

Server classification based on end-user application

HPC Servers

Data-Intensive

Traditional Data Science
- Monte Carlo Apps
- Optimization Apps
- Pattern Recognition

HPC-Enabled AI
- Model Training
- Surrogate Models
- Graph Analysis

Machine Learning

Deep Learning

Compute-Intensive

- CFD
- Reservoir Modelling

Other AI
- Graph Analysis
- Semantic Analysis

Note: Definitions can be found in the appendix of the slide deck.
HPC-enabled AI Forecast

*5 year CAGR expected to reach over 22% growth*

<table>
<thead>
<tr>
<th>Forecast: Worldwide HPC server revenue breakout by compute-intensive and data-intensive focuses ($M)</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
<th>2024</th>
<th>2025</th>
<th>2026</th>
<th>CAGR 2021-2026</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worldwide HPC Server Revenue Forecast</td>
<td>13,519</td>
<td>14,750</td>
<td>16,503</td>
<td>18,208</td>
<td>19,697</td>
<td>19,492</td>
<td>20,549</td>
<td>6.9%</td>
</tr>
<tr>
<td><strong>Compute-Intensive</strong> Server Revenue</td>
<td>10,020</td>
<td>10,848</td>
<td>12,103</td>
<td>13,280</td>
<td>14,177</td>
<td>13,586</td>
<td>13,993</td>
<td>5.2%</td>
</tr>
<tr>
<td><strong>Data-Intensive</strong> Server Revenue</td>
<td>3,499</td>
<td>3,901</td>
<td>4,400</td>
<td>4,928</td>
<td>5,519</td>
<td>5,906</td>
<td>6,555</td>
<td>10.9%</td>
</tr>
<tr>
<td><strong>HPC-enabled AI (ML, DL &amp; Other) Server Revenue</strong></td>
<td>1,039</td>
<td>1,300</td>
<td>1,718</td>
<td>2,083</td>
<td>2,484</td>
<td>2,941</td>
<td>3,619</td>
<td>22.7%</td>
</tr>
<tr>
<td><strong>Traditional Data Science (non-AI HPDA) Focused Server Revenue</strong></td>
<td>2,460</td>
<td>2,601</td>
<td>2,682</td>
<td>2,845</td>
<td>3,036</td>
<td>2,965</td>
<td>2,937</td>
<td>2.5%</td>
</tr>
</tbody>
</table>
Future HPC System Design

AI and HPDA workloads pushing sites to consider new system architectures

- As workloads become more diverse, system designs have shifted:
  - Some sites are building single, large, heterogeneous systems to address a wide variety of applications
  - Some sites are building out multiple, smaller systems to handle different workloads specifically
  - Cloud resources are growing in utilization to address data-intensive workloads

- Technology options have diversified as well
  - New accelerator options, including AI-specific ASICs
  - Various memory, interconnect, and storage solutions

- Compute resource allocation should be treated as an optimization problem:
  - Find a balance among diverse technology options
  - Optimize for key workloads
Intersection of HPC and AI

Modeling and simulation workloads working in harmony with AI techniques

- **AI applications growing in the HPC space:**
  - Stand-alone AI models
  - AI incorporated into traditional simulation workloads:
    - Surrogate models
    - Data preparation and cleansing
    - Simulation steering with trained AI models

- **Mod/sim workloads benefiting from AI**
  - Acceleration of time to solution
  - Exploring new solution spaces
  - Parsing sparse matrices of data

- **AI benefiting from mod/sim workloads**
  - Generation of large synthetic datasets for training
  - Verification and testing of trained models in simulation
The Role of Explainability

For optimization, engagement, and compliance

- **Reproducibility and transparency as optimizers**
  - Automated monitoring hastens and ameliorates training
  - Timesaving capabilities (automated reporting, bias or drift detection) ease developer load and free valuable time
    - More models are developed overall
    - More models make it to production

- **Explainability drives engagement**
  - Many application spaces highly value auditability
  - Reassurance for previously hesitant domains
  - Contributes positively to development of AI workforce

- **Growing efforts to regulate and standardize**
  - Bolstered public knowledge and trust
  - Auditability as a legal obligation
  - Explainability tools mitigate regulatory fines
What did we miss?

tsorensen@hyperionres.com
anorton@hyperionres.com
QC Market Summary

Continued strong growth seen for global QC sector

• Based on a study of 112 QC suppliers from around the world, the estimated global QC market was worth about US $490 Million in 2021
  • The anticipated CAGR for the sector between 2021 and 2024 is 21.9%
  • The global QC market in 2024 is estimated to be approximately US $900 Million
• Current QC supplier base dominated by a few players
  • 49% of companies < US $500K, 7% > US $10 Million
• Cloud access model dominates for next three years
  • All cloud (43%) + hybrid (21%) = 64%
• Most Promising Market Segments
  • QC, Cybersecurity, Financial, Academic, and Chemical/Chemistry
In next 36 months, QC hardware/software near parity

N = 112
QC Algorithm by Revenues

ML & optimization, M/S and cyber close behind

QC Algorithm by Revenue (36 months)

- Machine learning: 25%
- Modeling / Simulation: 19%
- Optimization: 19%
- Cyber security: 14%
- Monte Carlo Processes: 9%
- Don’t know/Not sure: 12%
- Other: 2%

N = 112

© Hyperion Research 2022
Current Field of Commercial QC Hardware Suppliers

*Tracking the visible players from a global perspective*

- 44 identified QC hardware developers
- 12 quantum modalities under consideration
Commercial end users' interest in QC is high

- Almost 70% of 415 companies surveyed worldwide have some in-house QC program
  - Additional 20% plan to do so in the next few years
- QC technology is seen as offering a range of corporate-level benefits including improved research capabilities and increased revenue
- End users are looking for solutions in optimization, factory processes, scheduling, etc.
- Every vertical surveyed had a significant number of organizations currently involved in some level of QC activity
- Most interested companies already involved in data analysis, ML/DL, optimization, mod/sim, material science—some of the most promising areas of QC today
The QC Sector Matures

Transitioning from an R&D to market sector

D-Wave to Go Public with SPAC Deal; Expects ~$1.6B Market Valuation

Rigetti Computing, Inc. (RGTI)
NasdaqCM - NasdaqCM Real Time Price. Currency in USD

4.2100 -0.3200 (-7.06%)
As of 12:32PM EDT. Market open.

Investing in Quantum Computing Stocks
An in-depth look at the best quantum computing stocks in the U.S stock market this year.
QUESTIONS?

bsorensen@hyperionres.com

Insufficient facts always invite danger.
- Spock, *Stardate: 3141.9.*
The HPC Innovation Award and Winners
Examples of Previous Winners
The Trophy for Winners

The Innovation Excellence Award

For the Outstanding Application of HPC

Global 2017

Presented to:

For the Outstanding Application of HPC for Business and Scientific Achievements

© Hyperion Research 2022
HPC Award Program Goals

• #1 Help to expand the use of HPC by showing real ROI examples:
  1. Expand the “Missing Middle” – SMBs, SMSs, etc. by providing examples of what can be done with HPC
  2. Show mainstream and leading edge HPC success stories

• #2 Create a large database of success stories across many industries/verticals/disciplines
  ▪ To help justify investments and show non-users ideas on how to adopt HPC in their environment
  ▪ Create many examples for funding bodies and politicians to use and better understand the value of HPC → to help grow public interest in expanding HPC investments
  ▪ For OEMs to demonstrate success stories using their products
Users Submit the Value of the Accomplishment

- Users are required to submit the value achieved with their HPC system, using 3 broad categories, following a very specific set of guidelines:
  
a) Dollar value of the HPC usage
   - e.g., made $$$ in new revenues, saved $$$ in costs, made $$$ in profits, etc.
  
b) Scientific or engineering accomplishment
   - e.g., discovered how xyz really works, developed a new drug that does xyz, etc.
  
c) Value to society as a whole
   - e.g., ended nuclear testing, made something safer, provided protection against xyz, etc.

... and the investment in HPC that was required
The Judgment Process – Clear, Fair And Transparent

- The ranking of the accomplishments are done by only HPC USERS, following very specific rules
- A three-step process is used:
  1. First the submission has to be complete with a clear “value” shown
     - Some submissions were good, but needed a little more information
  2. Secondly, the HPC User Forum Steering Committee evaluates if the submission presents a realistic assessment of the value/returns
  3. Then, in cases where the value isn’t clear or a deeper technical depth is required, the final evaluation is by experts in the specific area/discipline
SC22 Winners: HPC User Innovation Awards
HPC-driven design of innovative functional materials for catalysis, energy conversion, and storage

Andrey Lyalin, Hokkaido University ICReDD

- Successfully predicted a variety of unique materials for catalysis, energy conversion, and storage applications which were then experimentally confirmed
- Developed novel computational methods in quantum chemistry, materials informatics, and machine learning that can be used to discover further materials for energy and environmental applications

© Hyperion Research 2022
Fast-track computationally-driven discovery of new SARS-CoV-2 Main Protease (Mpro) inhibitors: from HPC to experimental drug candidates
Jean-Philip Piquemal, Sorbonne Université and Qubit Pharmaceuticals

• Synthesized 2 new families of compounds in less than 3 months that were experimentally validated
• Produced a high-resolution molecular dynamics simulation of the SARS-CoV2 Main protease protein
  • Modeled interaction with water
  • Unraveled potential druggable sites
  • Identified potential drug candidates
  • Predicted absolute free energies of binding with potential active molecules
• Opens routes for further drug discovery and design

© Hyperion Research 2022
An automated agent-based simulation method for the prediction of intensive care unit (ICU) occupancy due to COVID-19

Ralf Schneider, High Performance Computing Center Stuttgart

- Weekly production of 4-week forecasts of ICU occupancy in Germany due to COVID-19
  - This is the first time that a federal HPC-center in Germany delivers such a service in production
  - This proved for the first time that HPC for Global System Science can support decision makers in critical situations and can make a difference for society.

- Forecasts are simulated by a model initially developed by the Federal Institute for Population Research

High-Performance Computing Center Stuttgart

© Hyperion Research 2022
AI-based emulator of Computational Fluid Dynamic simulations for urban wind flow and gas dispersion
Francisco Ramirez Javega, Bettair Cities SL

- Air pollution is the single largest environmental health risk in Europe and a major cause of premature death and disease, but most urban air quality problems are hyper-local (e.g., land use, traffic)
- Ran CFD simulations of wind flow and dispersion patterns on a large data set and trained a DNN to simulate the results at a fraction of the time and cost
  - This research empowers local communities, municipalities, and regional governments with accessible, actionable information about local emissions in real-time
- This work has been funded within the FF4EuroHPC European project, Experiment 1012: Improving Bettair Air Quality Maps
- Part of the project was performed by a team from the Barcelona Supercomputing center using ALYA
Thank you!

For questions: info@hyperionres.com
Conclusions

• 2021 was a strong year with a 9% increase
  • 2022 is also expected to be a growth year
    • Exascale systems will help drive growth in 2022 to 2024
    • AI, HPDA/big data are hot areas
    • But supply chain issues continue

• New technologies are showing up large numbers:
  • Processors, AI hardware & software, memories, new storage approaches, etc.
  • Quantum technologies

• The cloud has become a viable option for many sites
  • HPC in the cloud is lifting the sector writ large

• Storage will likely see major growth driven by AI, big data and the need for much larger data sets

• There are growing concerns about the workforce

©Hyperion Research 2022
A Concern: HPC Expertise Shortage

The growing scarcity of HPC experts to implement new technologies is the number one roadblock for many HPC sites

• Two major trends:
  1) A shrinking HPC workforce
  2) A massive increase in system complexity

• HPC experts are an aging workforce
  • The pipeline of new HPC staff entering the workforce does not adequately match the outflow of retirees
  • Competition for HPC staff will intensify

• Increasingly complex workloads are more difficult to manage
  • Increasing HPC systems per site
  • Augmenting traditional modeling/simulation with AI and big data
  • Incorporating multiple processor types, co-processors, accelerators, and other specialized hardware
  • Balancing on-prem and cloud
  • And Enterprise IT users are entering HPC space, and need HPC expertise

• HPC users need major improvements in ease-of-use, ease-of-selection, & ease-of-optimization

© Hyperion Research 2022
Future Meetings

• April 18 - 19, 2023 – Princeton, NJ
  • In person HPC User Forum meeting

• May 23, 2023 – ISC23 Breakfast Briefing
  • In person at the Grand Elysée Hamburg

• September 6 - 7, 2023 – Tucson, AZ
  • In person HPC User Forum meeting

• October HPC User Forum in Europe
Questions?

We welcome questions, comments and suggestions

Please contact us at: info@hyperionres.com