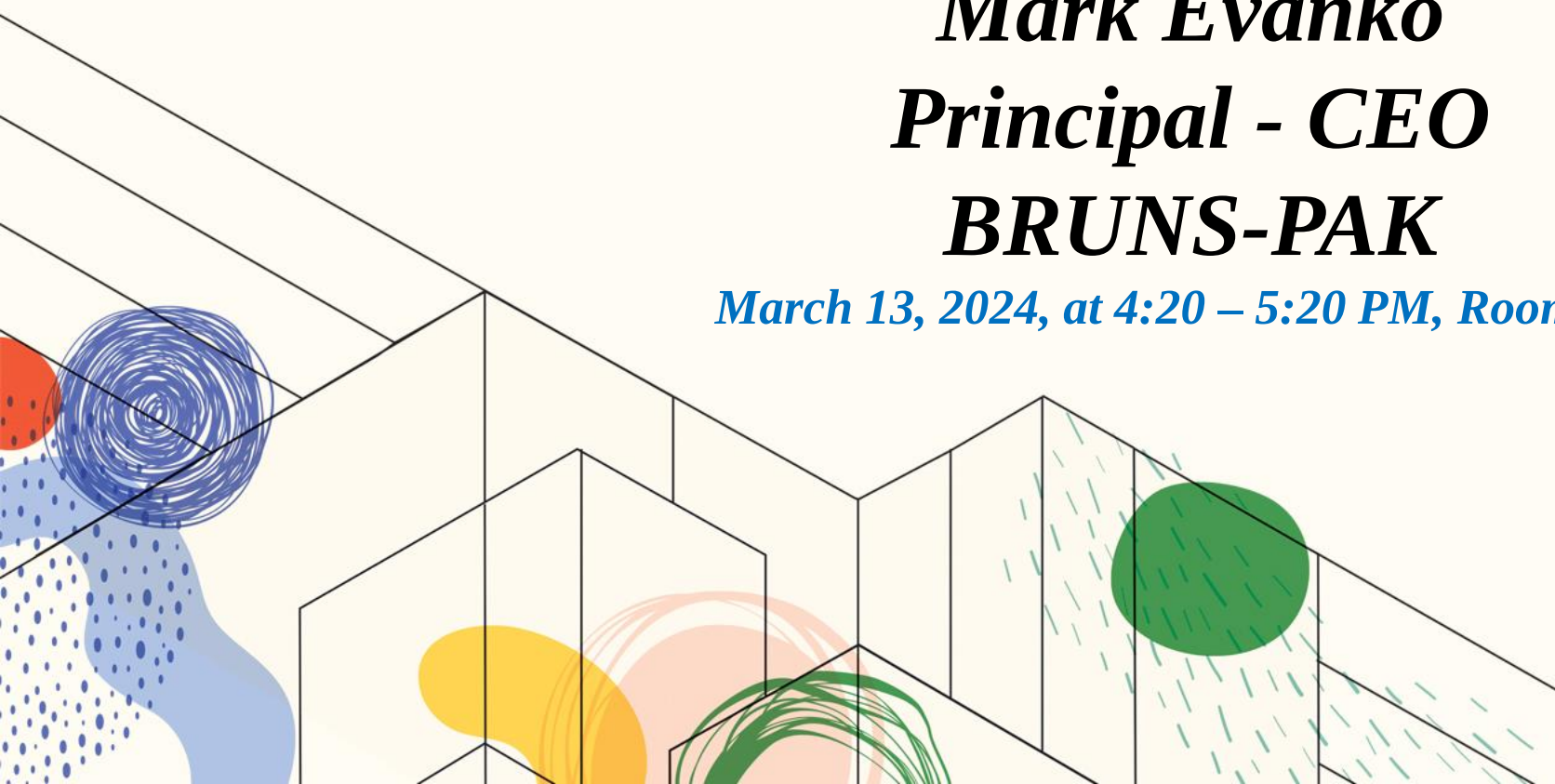
BRUNS-PAK
DATA CENTER "HYBRID" SOLUTIONS

Data Centers: High Performance Computing and Artificial Intelligence

Mark Evanko
Principal - CEO
BRUNS-PAK

March 13, 2024, at 4:20 – 5:20 PM, Room 330



NFMT 2024
National Facilities Management & Technology March 12-14, 2024 • Baltimore

Learning Objectives

1. What are the data center facility impacts (A/MEP) associated with HPC/AI?
2. How can I optimize the data center energy efficiency in deploying HPC/AI?
3. What are specific data center facility infrastructure solutions for HPC/AI?
4. What are the trends associated with planning and managing HPC/AI?

Agenda

Section I:	***The Data Center High Performance Compute (HPC) and Artificial Intelligence (AI) Challenges and Solutions – 2024***
Section II:	***The Overall Data Center "19 Elements" and the Compatibility of HPC/AI***
Section III:	***The Current Data Center Activities Surrounding HPC/AI***
Section IV:	***The 2024 Results of HPC/AI Deployments***
Section V:	***Conclusion***
Section IV:	***Questions and Answers***

Section I:

Data Center High Performance Compute (HPC) and Artificial Intelligence (AI) Challenges and Solutions - 2024

Section I: The Data Center High Performance Compute (HPC) and Artificial Intelligence (AI) Challenges and Solutions – 2024

1) The increased use of technologies including the internet of things (IoT), artificial intelligence (AI), and machine learning (ML) have created a large demand for processing massive amounts of data at a faster rate. The result is high performance compute clusters being deployed to solve the problem.

2) The HPC/AI "power density per cabinet" are ranging from 30 – 90+ kW and are scalable.

3) The HPC/AI solution of:
   - ✓ ***Compute***
   - ✓ ***Network***
   - ✓ ***Storage***

typically do not have the same data center facility infrastructure reliability (N+1 / 2N) as the Tier I production environment.

Section I: The Data Center High Performance Compute (HPC) and Artificial Intelligence (AI) Challenges and Solutions – 2024

4) HPC/AI users transcend most all industries:

Research	*Medical (i.e. Watson)*
Energy	*Government*
Aerospace	*Retail*
Weather	*Banking/Financial*
Automotive	*Traffic*
Pharmaceuticals	*Marketing/Predictive Analytics*
Real Estate	*Military*

Section II:

The Overall Data Center “19 Elements” and the Compatibility of HPC/AI

Section II: *The Overall Data Center "19 Elements" and the Compatibility of HPC/AI*

Elements of a Successful Data Center Project
Data Center "Hybrid" Design/Build Solutions

The Hybrid "2024 Transformation" Efficient Data Center Elements

Section II: *The Overall Data Center "19 Elements" and the Compatibility of HPC/AI*

- ✓ ***Disparate Facility Infrastructure Reliability of HPC /AI (N) vs. Production Environment (N+1 / 2N / Uptime Tier III/IV).***
- ✓ ***HPC/AI Modular/Scalable Rack Densities From 30 to 90+ kW***
- ✓ ***Separate HPC/AI Facility From Production Environment?***

Section III:

The Current Data Center Activities Surrounding HPC/AI

Section III: The Current Data Center Activities Surrounding HPC/AI

1) Understand the end user client short/long term high performance compute (HPC)/artificial intelligence (AI) mission and the corresponding rack/kW load densities.

2) Understand the end user client compute short/long term “uptime” requirements associated with the HPC/AI deployment. Specifically delineate between:
   - ***Front End Nodes***
   - ***General Processing***

3) Translate the HPC/AI short/long term “kW per rack vertical demand” to understand the immediate vs. long term electrical/mechanical impact to the data center raised floor space.

Section III: The Current Data Center Activities Surrounding HPC/AI

4) Identify the short/long term mechanical/thermodynamic impact options of HPC/AI including:
   a. ***Containment***
   b. ***Rear Door Heat Exchangers***
   c. ***In-Row Cooling***
   d. ***To the Chip Cooling/Immersion Cooling***

5) Confirm compute hardware manufacturer warranties associated with the chip cooling/immersion cooling.

6) Develop a short/long term data center facility A/MEP modular strategy to scale with HPC/AI.

7) The initial cost of deployment vs. return on investment. ***CRITICAL!!***

Section IV:

The 2024 Results of HPC/AI Deployments

Section IV: The 2024 Results of HPC/AI Deployments

1) Tremendous speed of processing results. What took weeks/months of processing application data can be decreased within hours/days.

2) Data center facility improvements are designed/deployed to accommodate a modular/scalable/flexible/reliable “HPC/AI” deployment.

3) Many end user HPC/AI applications do not require tier III/IV A/MEP superstructures. See end user support.

4) Large data center real estate consolidation results realized by deployment of “HPC/AI” loads from legacy less dense loads.

5) HPC/AI is delivering real time data to the client community results in increased profits, market share, revenue, and marketing visibility.

Section IV: The 2024 Results of HPC/AI Deployments

6) HPC/AI is being designed from a mechanical standpoint to operate in higher inlet (°f) temperatures.

7) Most all compute manufacturers continue to deploy HPC/AI now and in the future based on ROI.

8) 2024 HPC deployments at BRUNS-PAK up to 100kW per cabinet.

9) Survey of CIO industry leaders found that 75% believe if their business does not address AI/HPC by 2025, they face bankruptcy.

10) <u>OPTION</u>: Retrofit/renovate/upgrade legacy data centers to meet HPC/AI objectives vs. modular vs. new.

11) HPC/AI expected to double data center power usage in the next three years.

Section V:

Conclusion

Section V: Conclusion

- ✓ HPC/AI deployments are growing at an expediential level.
- ✓ The growth of HPC/AI is based on the increased use of technologies like internet of things (IoT), machine learning (ML), and now quantum computing.
- ✓ The initial rates of return of "properly deployed" HPC/AI is reporting less than three (3) years.
- ✓ The industries served by HPC/AI are exploding.
- ✓ Compute manufacturers including IBM, Dell, HPE, and Lenovo will continue to develop high powered compute racks.
- ✓ Cloud/colocation/EDGE/container/enterprise data centers continue to evolve to address the optimal solutions.

Section VI:

Questions and Answers

Section VI: Questions and Answers

Questions and Answer Session

THANK YOU!

Mark Evanko

Principal - CEO

BRUNS-PAK WORLDWIDE, INC.

999 New Durham Road

Edison, NJ 08817

PH: 732-248-4455

Email: mevanko@bruns-pak.com

Website: www.bruns-pak.com

BRUNS-PAK
DATA CENTER "HYBRID" SOLUTIONS

NFMT2024
National Facilities Management & Technology March 12-14, 2024 • Baltimore