Special Report

HPC Profiles in Leadership: DUG Technology

Mark Nossokoff and Jaclyn Ludema
October 2023

EXECUTIVE SUMMARY

Innovations are constantly occurring across all aspects of HPC infrastructure. Much of that innovation has been focused on the computing infrastructure:

- CPU performance once driven by faster clock speeds have evolved to integrating more cores.
- GPUs have emerged to advance performance of AI-related workloads.
- Application-specific accelerators are now being introduced to address unique elements of AI applications.

Advancements in these areas have created challenges in other areas, primarily related to the increasing power required to run each new generation of processing element. Innovations in equipment cooling and datacenter design and operations are evolving to address the power and cooling challenges. Specifically, immersion cooling has emerged as an option to mitigate increasing power demands and aims to provide substantial benefits to datacenter facilities. DUG Technology (DUG) has fully embraced immersion cooling for its computing infrastructure and provides a proof point for others considering taking the immersion cooling approach.

DUG TECHNOLOGY OVERVIEW

DUG, headquartered in Australia with datacenters in Australia (Perth), UK (London), U.S. (Houston, Texas), and Malaysia (Kuala Lumpur), specializes in analytical software development, big-data services, and reliable, green, high-performance computing (HPC). DUG solutions are designed to offer innovative software products and cost-effective, cloud-based HPC as a service backed with dedicated support for technology onboarding.

Leveraging its expertise in algorithm development, DUG aims to provide a broad range of geoscience products comprised of seismic data with customized services, software and HPC solutions. Industries currently served by DUG's solutions include radio-astronomy, biomedicine and meteorology, as well as the government and education sectors. Key to DUG's strategy and approach to the market is a focus on green technology for supercomputer-class infrastructure.

THE DUG HOUSTON DATACENTER

Located north of the west end of Houston's Energy Corridor on I-10, the DUG Houston datacenter is housed in a facility managed by Skybox Datacenters, the largest datacenter facility provider in Texas. DUG was the first tenant at Skybox's Houston site and the DUG compute data hall was custom-built to its specifications, explicitly designed for immersion cooling.

The DUG computing environment is composed of separate containers. Each container houses the equivalent of a 26U rack laid on its back. A typical container can have 13 2U servers, each with 2x dual-core nodes, providing 52 CPU cores in each container. They also include a DUG-developed heat pump and exchanger that with the current iteration of CPU technology exhausts roughly 5 degrees of heat from the immersion coolant. Each generation of CPU (and coming GPUs) will be tested to determine the appropriate heat transfer parameters. There are 220 uniform containers currently in service with an expansion capability of up to 720 containers.

While slowly changing, many of the servers available today are not yet immersion cooling ready. This necessitates on-site modifications to standard configurations, such as removing unnecessary fans and eliminating thermal epoxy to prevent dissolving into the coolant.

Storage is mixed between immersion containers and standard racks that are deployed along with networking in a co-located data hall at the Skybox Houston facility. Redundant high-speed networks provide connectivity to the compute data hall.

ANALYST OBSERVATIONS

Immersion Cooling Benefits

While initial datacenter design costs for immersion cooling may be higher than traditional air-cooled or direct-liquid-cooled datacenters, the benefits of immersion cooling relative to air-cooling may be achieved across a wide range of areas and can provide a payback on the investment:

- Reduced air conditioning expense
- Lower equipment cooling operating expense (e.g., elimination of hot/cold aisle designs and segmented air handling)
- Better reliability with more consistent temperature control
- Longer equipment life due to elimination of hardware hot spots and oxidation

There are also multiple non-financial benefits:

- Substantially lower acoustic decibel levels
- A reduced carbon footprint
- Reduced fire risk – high-flash point of the fluids means they act as a fire suppressant

Changes Required to Use Immersion Cooling Today

The current state of server vendor roadmaps provides a limited choice of options for CPU technology immediately ready for immersion cooling. This requires the datacenter facility to open each server and remove anything related to air cooling (e.g., fans) and any material that may breakdown in the immersion coolant (e.g., thermal epoxy). The act of cracking open a server to make these modifications likely voids the warranty on those systems. This may drive the need to factor in extra servers during the initial procurement to use as full replacements or spare parts.

While more immersion-ready systems are expected to be offered by multiple system vendors, this added up-front activity and expense may be a cause for the current limited industry adoption of immersion cooling.

FUTURE OUTLOOK

The power demands of CPUs and GPUs used in HPC infrastructure are expected to increase for the foreseeable future. Current high-end compute elements draw 700W each with certain SKUs approaching 1KW each. Keeping these components and their servers cooled is proving challenging for both server vendors and datacenter facility operators.

Air-cooling will continue to be employed for storage systems and networking infrastructure. Server architects and datacenter designers have a choice of liquid cooling options between direct liquid cooled and immersion, and there are further choices available within each cooling technology. There is no one-size-fits-all approach to determine which technology is best suited for a given system or user requirements. Criteria to consider when making the decision between the cooling technologies include availability of capital funding, operating expense targets, facility capabilities (e.g., retrofitting existing datacenter vs. greenfield facility availability), server availability and attitudes towards sustainability, energy efficiency, and carbon footprint reduction.

DUG has made the decision on immersion cooling as the cooling technology for their datacenters. The design point and IP they've developed are intended to provide a reliable infrastructure for uses and financial returns for DUG, and current indications are that these benefits are being realized.

About Hyperion Research, LLC

Hyperion Research provides data-driven research, analysis and recommendations for technologies, applications, and markets in high performance computing and emerging technology areas to help organizations worldwide make effective decisions and seize growth opportunities. Research includes market sizing and forecasting, share tracking, segmentation, technology, and related trend analysis, and both user and vendor analysis for multi-user technical server technology used for HPC and HPDA (high performance data analysis). We provide thought leadership and practical guidance for users, vendors, and other members of the HPC community by focusing on key market and technology trends across government, industry, commerce, and academia.

Headquarters

365 Summit Avenue
St. Paul, MN 55102
USA
612.812.5798
www.hpcuserforum.com and www.HyperionResearch.com