Position Papers of the 20th Conference on Computer Science and Intelligence Systems (FedCSIS) pp. 107–107 DOI: 10.15439/2025F9634 ISSN 2300-5963 ACSIS, Vol. 44

Paths to Zero Emission Computing—Reducing Energy Consumption, and carbon emissions in HPC and AI environments

Tikiri Wanduragala
Lenovo, UK
twanduragala@lenovo.com

***Abstract*—In this position note, core issues involved in creation of zero emission data centers are summarized.**

***Index Terms*—data center, zero emission, HPC.**

DIGITAL transformation projects combined with rebalancing workloads between public and private clouds for reasons of sovereignty have given rise to an increase in demand for compute capacity globally. Conventional systems and data center designs have been able to accommodate the projected growth. However, the rise of AI and more demanding HPC environments using large numbers of GPU's and associated networking, high bandwidth memory and storage systems have radically changed systems design and energy requirements. In recent years power consumption has grown by at factor of 3 up to 500W for CPU's and 1000W+ for GPU's. This has resulted rack power consumption increasing from about 15KW to100KW+.

In essence HPC and AI environments consume significantly more energy which can result in increased CO2 emissions and water usage at the data center level. The International Energy Agency (IEA) Energy and AI report (April 2025) projects global Data Centre (DC) electricity consumption to double to 945TWh globally by 2030. A Mckinsey article on "The Cost of Compute – a $7 trillion race to scale data centers" (April 2025) put the "global demand for data center capacity could almost triple by 2030, with about 70 percent of that demand coming from AI workloads".

Conventional data centres and systems are struggling to cope with the demands being placed on them and place limits on the capabilities of HPC and AI unless these issues are addressed. With traditional air-cooled systems as much as 40% of the energy provided for compute can be lost by the cooling systems.

Lenovo is a leader in providing HPC and AI factory solutions globally, experience in designing energy efficient systems. From SW to gather data from the underlying infrastructure to optimise for performance or energy usage to a range of HW solutions branded as Lenovo Neptune offer both air- and water-cooled solutions to ensure energy is utilised as efficiently as possible.

As data centers take several forms and are specific to the workloads that are designed to run on the systems inside them. Classic air-cooled systems offer the most flexibility and in Hot / Cold Aisle configuration can achieve a PUE of 1.5 to 1.6. Air cooled systems combined with rear door heat exchangers using chilled water can improve the PUE to about 1.2. Direct warm water-cooled system can improve the PUE to 1.1 to 1.06 range.

Moving from air to direct warm water-cooled systems can result in several significant benefits:

- Density more compute power in a compact footprint
- Optimal performance by keeping components within thermal design power envelope
- Possibility of reducing the carbon footprint of the installation
- Possibility of re-using waste heat in other campus location
- Higher performance per watt

In the recent past the use of water-cooled technologies was in the realm of HPC and AI installations. As HPC and AI workloads are being deployed within enterprise computing environments a hybrid approach is sometimes taken with less demanding computational workloads running on air-cooled systems and the more demanding running on direct water-cooled systems.

To realise the full potential of HPC and AI in a cost-effective sustainable manner efficient use of energy is essential. Tried and tested technologies like Lenovo Neptune™ direct warm water cooling not only offers compute efficiencies but capturing the waste energy in water offers the possibility of heat re-use but heating campus buildings making carbon neutral computing a possibility.