

Bringing AI to the data centre

 By [Jim Holland](#)

17 Apr 2019

It's time to bring AI to the data centre. SA Businesses especially those involved in data centre management should start adopting and embracing AI now if they still want to continue to be relevant.



Jim Holland is Country Manager at Lenovo Data Center Group Southern Africa

Artificial intelligence (AI) is front and centre for many enterprises because of its potential to deliver significant benefits to both top and bottom lines of businesses. It can help enterprises make better decisions by leveraging their data to reduce costs by improving process efficiencies and to increase revenues and gain competitive advantage by bringing new offerings to market more swiftly.

Further, AI has enormous potential to impact human lives. AI can help deliver better healthcare by reducing human errors in diagnosis, enhancing understanding and treatment of diseases and dramatically accelerating drug discovery. It can also be utilised to enhance personal security by identifying potential threats from humans or nature before they become an issue, such as more accurate weather forecasts, or enhanced, in-car traffic monitoring.

Many of the intelligent applications delivering these transformational benefits in enterprises need huge amounts of data, particularly using new methods like deep learning, which enables AI today. Much of this data still resides in on-premise data centres due to security restrictions, privacy laws or even the economics of moving sheer amounts of data. These issues are much more prominent in certain verticals like healthcare, manufacturing, finance and research organizations. When developing and deploying AI applications in these scenarios, it makes more sense to bring the compute power closer to the data than the other way around.



How is AI changing business in 2019?

Tarun Gurjal 10 Apr 2019



To be successful with AI efforts, enterprises need more than just a hardware provider for their computing needs. They need a partner who will work hand-in-hand to help reduce business uncertainty and technical complexity associated with this emerging technology. Lenovo is taking a unique approach to meet customer needs through several offerings, which goes beyond infrastructure. Together with Launch:AI Workshops, Lenovo's global AI Innovation Centres help customers identify use cases that can deliver business value and execute proofs-of-concept by providing AI expertise and optimized infrastructure to reduce business risk and prove technical viability.

The next phase

The next phase of executing AI projects even bring many more challenges, particularly in procuring hardware and software tools to support data scientists and developers while optimizing total cost of ownership (TCO). Currently, much of the AI development using deep learning method makes use of open source software frameworks such as TensorFlow, Caffe, MxNet, etc., whereas enterprise IT experience is mostly with packaged software applications, which are easy to manage. Adding further complexity is the need to support multiple developers using various frameworks and versions to accomplish the same task. In this context, enterprises need supporting development tools to leverage open source effectively. Our award-winning Lenovo intelligent Computing Orchestration (LiCO) simplifies AI development by efficiently managing cluster resources, open source frameworks and typical AI workflows.

Procuring infrastructure for AI is a balancing act for many data centre IT admins because of the need to satisfy multiple requirements. These typically include meeting the diverse performance demands from data scientists, AI engineers and software engineers, while optimizing the TCO. A difficult needle to thread!



Cutting through the AI hype: Explainable AI, flexible workforce

Fadi Kanafani 4 Apr 2019



The optimal solution would have the performance of a purpose-built system, yet have the flexibility to run multiple applications, even non-AI workloads. With this blueprint in mind, the Lenovo ThinkSystem SR670 was designed primarily for scale-out AI workloads, but with flexibility to handle traditional high-performance computing (HPC), virtual desktop infrastructure (VDI), video processing, etc. Further, it efficiently scales from experimentation with a couple of nodes to large-scale deployments in distributed training environments, which LiCO also supports as an added benefit.

The ThinkSystem SR670 has two models, each based on the different AI workloads in data centre, supporting either NVIDIA T4 or NVIDIA V100 Tensor Core GPUs. The newest ThinkSystem SR670 offering (announcing at NVIDIA GTC'19) supports eight T4 GPUs in a 2U server. With NVIDIA NGC-Ready validation, T4 servers can excel across the full range of accelerated workloads— machine learning, deep learning, virtual desktops, and high-performance computing (HPC). And, at 70 watts each, the T4 GPUs alleviate the power and cooling requirements to populate these servers at scale in an existing enterprise data centre.

For deep learning training or running HPC applications exclusively, the SR670 currently supports four NVIDIA V100 GPUs, and delivers the best total compute capacity with 448 deep learning teraflops or 28 double-precision teraflops. Further, the 32GB onboard HBM2 memory is suitable for training large neural networks and HPC computational models with in-memory requirements. Together, with the dedicated PCIe communication channels between CPU and GPUs in ThinkSystem SR670, this configuration becomes the ideal choice for IT to get the best performance for these two workloads.

By using the SR670 and LiCO, IT admins will be able to balance their performance requirements and TCO by designing their data center infrastructure for a better return on investment (ROI). Now is the time to bring AI closer to the 'data' center for building production-grade intelligent applications.

ABOUT JIM HOLLAND

Jim Holland is Country Manager at Lenovo Data Center Group Southern Africa

- Bringing AI to the data centre - 17 Apr 2019
- Cultivating a connected world with emerging technology - 21 Feb 2019
- AI: The driver for mind-blowing business growth - 31 Jan 2019
- Trends empowering intelligent transformation - 3 Dec 2018
- Encouraging a new breed of storage for the digital age - 23 Nov 2018

[View my profile and articles...](#)

For more, visit: <https://www.bizcommunity.com>