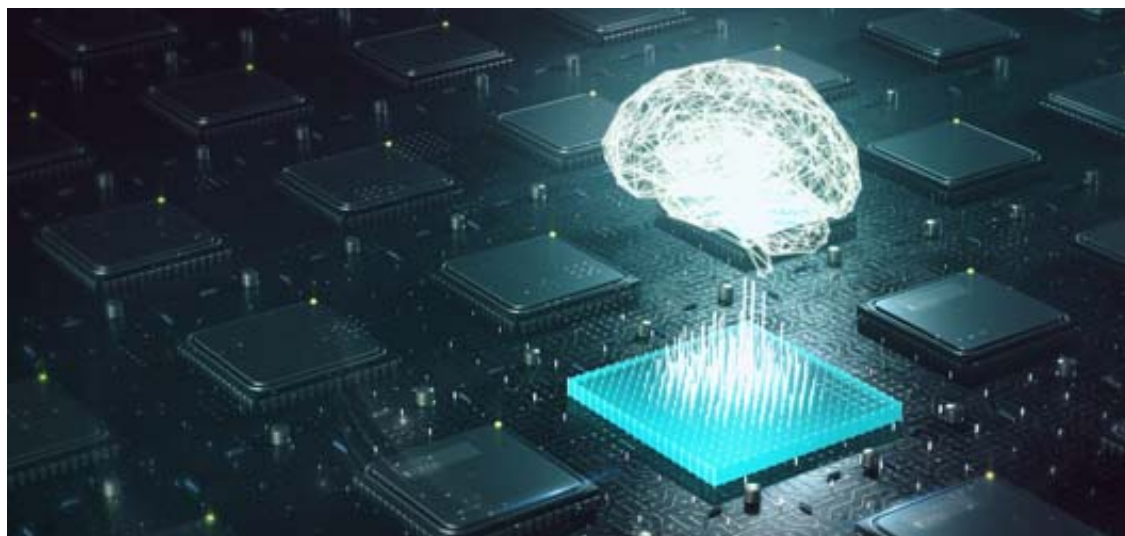# IBM Systems MEDIA

How to Get Started With PowerAI

A step-by-step guide to getting started with IBM's artificial intelligence framework.

By Jaqui Lynch

01/01/2019

Over the past year or two, Power Systems users have begun hearing more about deep learning and machine learning—and how PowerAI can help make these more accessible. In order to understand and deploy PowerAI it is helpful to understand the concept of deep learning and also what the PowerAI offerings are.

What is Deep Learning?

Deep learning is the use of algorithms to train software to perform tasks like speech and image recognition. It allows you to derive information about relationships within vast amounts of data that are just too large for regular programming. It also allows you to analyze visual and auditory data. PowerAI distributed deep learning (DDL) is a Message Passing Interface (MPI)-based communication library that is optimized for deep learning training. Applications are integrated with DDL so that they can run in parallel across a cluster of systems to get the best performance.

Another concept that is often discussed is large model support (LMS). LMS loads the neural model and dataset in system memory and caches activity to GPU memory. This allows the models and training batch size to scale as it takes advantage of both system and GPU memory to support it. LMS is provided with Caffe, TensorFlow and PyTorch.

Deep learning and LMS are finding homes in many areas such as manufacturing, financial services, healthcare and life sciences, retail, utilities and hospitality. In the medical area they are being used for patient triage, health management, real time alerts, diagnostics and disease identification. Financial services are using it for risk analysis and credit checks. These are just a few of the common uses.

PowerAI and PowerAI Enterprise

PowerAI is designed to provide an end-to-end, deep learning platform for data scientists. It is an architecture designed for deep learning and consists of a distribution of open source frameworks that supports up to 4 nodes for deep learning using software such as TensorFlow, Caffe, Keras, Pytorch and Chainer. PowerAI Enterprise includes those frameworks, libraries and tools that come with PowerAI and supplements them with additional components to support data management and ETL, training visualization and monitoring, multitenancy support and security, user reporting and charge back, dynamic resource allocation, external data connectors, plus many other components. It comes with Support Line services (http://www-01.ibm.com/common/ssi/ShowDoc.wss?docURL=/common/ssi/rep_ca/7/899/ENUSLP18-0087/index.html&lang=en&request_locale=en)and can scale well beyond four nodes to hundreds of nodes.

Base PowerAI is targeted towards development and small environments, and includes everything to allow you to quickly get the environment operational and performing. PowerAI Enterprise adds the functionality to rapidly scale the environment and includes Spectrum Conductor which takes care of the job scheduling across all the nodes. This reduces training time and increases efficiency.

The third piece of software is PowerAI Vision. This is used to analyze images and video and allows users to focus on rapidly identifying and labeling data sets so they can then train and validate the models in a GUI interface. It is an add-on to PowerAI Enterprise.

One key advantage of the PowerAI environment is that everything comes pre-compiled and optimized for the IBM POWER platform. This allows you to develop and deploy models much faster without having to go through the compile yourself.

Prerequisites

PowerAI Enterprise is optimized for the Power System AC922 and the S822LC for High Performance Computing. A minimum or 2 cores, 128GB memory, NVIDIA P100 (S822LC) or V100 (AC922) GPUs are required. Without the GPUs, it's not possible to run PowerAI. It runs on RHEL 7.5 LE (little endian) or higher with additional required NVIDIA software such as CUDA and the latest NVIDIA GPU drivers. The current release of PowerAI is 1.5.4 and PowerAI Enterprise and PowerAI Vision are 1.1.2 (as of 11/16/2018). Prior to 1.5 PowerAI ran on Ubuntu 16.04 on the S822LC. 1.5 requires RHEL on either the S822LC or the AC922.

Specific requirements for PowerAI v5.1 include:
Red Hat Enterprise Linux (RHEL) 7.5 (architecture: ppc64le).
V1.5.4 also supports Ubuntu 18.04 on the host running Docker containers
NVIDIA GPU driver Version 410.72
NVIDIA CUDA Version 10.0.130
NVIDIA cuDNN (CUDA deep neural network) 7.3.1
NVIDIA NCCL 2.3.5
Anaconda 5.2
AC922 at the latest firmware from IBM

For shared filesystems, the options are Spectrum Scale (was GPFS) and NFS. As of the end of October 2018 RHEL 7.6 was released but is not yet supported on PowerAI – this means care must be taken when using yum for

updates. Otherwise, you could inadvertently update RHEL to an unsupported level. You can use the subscription manager on the server to hard set it to 7.5 to avoid this issue until RHEL 7.6 is supported:

```
sudo subscription-manager release --set=7.5
```

POWER9 AC922

The AC922 is a system that is designed to support high performance computing (HPC) and deep learning. The processor module consists of 16, 18 , 20 or 22 POWER9 cores with up to four threads per core. 512KB or L2 cache and up to 10MB of L3 cache is shared by each pair of cores and streaming memory bandwidth is between 108 GBps and 170GBps. The system uses DDR4 memory with a maximum of 2TB per server.

The AC922 can have up to six NVIDIA Tesla V100 GPUs that are connected using NVLink v2.0. This means the bandwidth between GPUs is 200GBps bidirectional and between the GPUs and the POWER9 chip is 300GBps bidirectional—this compares to 32GBps GPU to chip for traditional PCIe3 systems.

The AC922 comes with four PCIe Gen4 slots plus an IPMI port on the BMC card for control of the server. Power is redundant and requires 200-240V. There are also two disk bays that can be used for SATA disks or SATA SSDs.

There are two models of the AC922: the 8335-GTH and the 8335-GTX,

The 8335-GTH has two POWER9 processors (each with 16 or 20 cores) and zero, two or four GPUs. It is air cooled. The minimum required for this server is two processor modules, 128GB memory, two power supplies that must both be connected to power and rack mount hardware for a 19" rack.

The 8335-GTX also has two POWER9 processors (each with 18 or 22 cores) and 4 or 6 GPUs. It is water cooled. The minimum required for this server is 2 processor modules, 128GB memory, four NVIDIA Tesla V100 GPUs, 2 power supplies that must both be connected to power, and rack mount hardware for a 19" rack.

Summary

The combination of PowerAI and the AC922 server is a great improvement on previous deep learning platforms. The move to the NVIDIA V100 GPUS combined with NVLink v2 provides much better memory bandwidth between GPUS and the POWER9 chip. This means data gets in and out of memory much faster, which improved performance for the

models. Familiarity is needed with RHEL Linux, especially installing and updating utilities and packages. If you are interested in deep learning then the combination of PowerAI and the AC922 is a very powerful way to implement it. You can start small and then upgrade to PowerAI Enterprise and integrate additional nodes as needed, bringing in PowerAI Vision when you are at Enterprise and are ready to integrate analysis of video and images into the system.

References

For more information on PowerAI, check out the following reading materials:

IBM PowerAI Enterprise 1.1 Announcement
http://www-01.ibm.com/common/ssi/ShowDoc.wss?docURL=/common/ssi/rep_ca/9/897/ENUS218-139/index.html&request_locale=en (http://www-01.ibm.com/common/ssi/ShowDoc.wss?docURL=/common/ssi/rep_ca/9/897/ENUS218-139/index.html&request_locale=en)

PowerAI Redbook
http://www.redbooks.ibm.com/redbooks/pdfs/sg248409.pdf (http://www.redbooks.ibm.com/redbooks/pdfs/sg248409.pdf)

AC922 Technical Overview
http://www.redbooks.ibm.com/redpapers/pdfs/redp5494.pdf (http://www.redbooks.ibm.com/redpapers/pdfs/redp5494.pdf)

IBM Power UK User Group PowerAI Getting Started
Handout: http://public.dhe.ibm.com/systems/power/community/aix/PowerVM_webinars/70_PowerAI_Getting_Started.pdf (http://public.dhe.ibm.com/systems/power/community/aix/PowerVM_webinars/70_PowerAI_Getting_Started.pdf)

Replay:
http://public.dhe.ibm.com/systems/power/community/aix/PowerVM_webinars/70_PowerAI_Getting_Started.wmv (http://public.dhe.ibm.com/systems/power/community/aix/PowerVM_webinars/70_PowerAI_Getting_Started.wmv)

PowerAI Home page at IBM
https://www.ibm.com/us-en/marketplace/deep-learning-platform (https://www.ibm.com/us-en/marketplace/deep-learning-platform)

PowerAI FAQ
https://developer.ibm.com/linuxonpower/deep-learning-powerai/faq/ (https://developer.ibm.com/linuxonpower/deep-learning-powerai/faq/)

PowerAI Developerworks User Forum

https://developer.ibm.com/answers/topics/powerai/ (https://developer.ibm.com/answers/topics/powerai/)

PowerAI Resources
https://developer.ibm.com/linuxonpower/deep-learning-powerai/ (https://developer.ibm.com/linuxonpower/deep-learning-powerai/)

PowerAI Prerequisites
https://www.ibm.com/support/knowledgecenter/SSFHA8_1.1.2/powerai_hwandsw_reqs.html (https://www.ibm.com/support/knowledgecenter/SSFHA8_1.1.2/powerai_hwandsw_reqs.html)

About the author
Jaqui Lynch has over 38 years of experience working with a projects and OSes across vendor platforms, including IBM Z, UNIX systems and more.

Related Content

Systems management (/systems-management) Centralize AIX LPAR and Server Management With NIM → (https://ibmsystemsmag.com/Power-Systems/06/2009/centralize-aix-lpar-server-management-nim)
Application development (/application-development) Building REST Web Services With RPG → (https://ibmsystemsmag.com/Power-Systems/02/2012/rest-web-services-with-rpg)
IT infrastructure (/it-infrastructure) POWER9 Hardware and More to Look Forward to in 2018 → (https://ibmsystemsmag.com/Trends/01/2018/look-forward-to-2018)

IBM Systems MEDIA