Lenovo Validated Design for AI Infrastructure on ThinkSystem Servers

Last update: **09 November 2018**
Version 2.0

Reference architecture for AI with Intel Caffe and TensorFlow

Solution based on Lenovo ThinkSystem servers with Intel Xeon processors and NVIDIA GPUs

Deployment considerations for high-performance, cost-effective and scalable solutions

Contains the detailed hardware configurations for servers and associated networking

Peter Seidel (Lenovo)
Ajay Dholakia (Lenovo)
Nathan Stewart (Lenovo)

Table of Contents

1 Introduction

This document describes the reference architecture for a flexible and scalable Artificial Intelligence (AI) infrastructure on Lenovo ThinkSystem servers. It provides a predefined and optimized hardware infrastructure for data access, model training and inference under various usage scenarios. The reference architecture provides planning, design considerations, and best practices for implementing the AI infrastructure with Lenovo products.

With the ever-increasing volume, variety and velocity of data becoming available to an enterprise comes the challenge of deriving the most value from it, this task requires the use of suitable data processing and analytics software running on a tuned hardware platform. With AI techniques such as Machine Learning (ML) and Deep Learning (DL) gaining popularity as the next-generation analytics algorithms, enterprises are modernizing their application and IT architectures by employing these components. In particular, applications using image recognition and pattern detection techniques can realize significant performance improvements by developing and deploying suitable DL models.

The AI adoption journey involves the following key steps:

- Data access
- Model training
- Inference

The task of providing data access entails connection with various data repositories. Examples of data repositories are data warehouses, Hadoop clusters, cloud storage and other data storage systems.

Another key step in the AI adoption journey is exploration and selection of models for deep learning. Typical models are based on deep neural networks (DNNs) and require a significant amount of computational resources for training. Using hardware infrastructure designed as a scale-out cluster for such model training use cases is a key requirement for enabling DL adoption.

The task of evaluating DL models requires a data science team to run several training jobs and compare the results to select the desired model and its associated parameter values. Lenovo Intelligent Computing Orchestration (LiCO) is a tool that provides an easy to use GUI for scheduling and orchestrating such model training jobs.

The inference step is aimed at deploying and using the trained model in the target application environment. The trained models are used for making decisions on new data such as images or patterns. The goal of a well-tuned inference stage is to enable decisions with accuracies similar to those observed during the model testing step. It is important to keep in mind the metrics of throughput and latency when deploying an inference system.

The intended audience for this reference architecture is IT professionals, technical architects, sales engineers, and consultants to assist in planning, designing, and implementing advanced analytics solution with Lenovo hardware. It is assumed that you are familiar with popular AI software components and capabilities. For more information, see “Resources” at the end of this document.

2 Business problem and business value

This section describes the business problem that is associated with AI infrastructure and the value that is offered by Lenovo hardware and software.

2.1 Business problem

Data explosion is creating a significant opportunity to extract value from it for enterprises of all sizes. New technologies like Machine Learning and Deep Learning are promising faster and better insights if they can be properly exploited. Enterprises today face the challenge of creating an IT infrastructure that can help store, process and use the vast amount of data available to them. As part of this AI adoption challenge, there is the need to access relevant data from various repositories, set up an environment for data scientists to explore various ML and DL models and evaluate associated training algorithms in their quest for selecting the appropriate models for use in target applications, and deploy the trained models in the target applications.

Another challenge is having the necessary compute power and storage capacity available on-premises where a significant amount of enterprise data resides. In particular, training DL models is a compute-intensive task, requiring appropriate hardware infrastructure to be made available to the data science team.

Finally, there is also a critical need to match DL model evaluation requirements with a user friendly toolset that allows data scientists to become productive very quickly. Managing and orchestrating the various DL model training jobs on a suitably sized and configured hardware infrastructure remains a key problem.

2.2 Business value

The solution for AI infrstructure described in this reference architecture helps address the key business challenges outlined above by providing preconfigured starting points for data repositories, for training the DL models and then deploying them for inference. Users can save a lot of experimentation typically required for identifying appropriate hardware infrastructure as well as software tools to create a suitable AI implementation environment. Having quick access to such an environment with easy to use tools provides significant business value.

The solution described in this reference architecture provides a flexible and scalable hardware infrastructure for use by the data science team in an enterprise. Depending on the collection of DL models, driven by the target application domain, data availability and growth over time, the solution provides the configuration of compute and storage servers organized as scale-out clusters to help address the challenge of finding right-sized on-premises hardware.

Furthermore, the solution also describes the use of software tools for managing and orchestrating DL model evaluation jobs in a user-friendly and efficient manner.

3 Requirements

The functional and non-functional requirements for this reference architecture are desribed in this section.

3.1 Functional requirements

An AI infrastructure solution supports the following key functional requirements:

- Ability to handle various data repositories for access to data for training
- Ability to handle various models from within DL domains
- Ability to deploy industry-standard AI software frameworks
- Ability to handle large volumes of data of various data types
- Ability to support deployment of trained models in various application environments
- Various client interfaces

3.2 Non-functional requirements

Customers require their AI model training solution to be easy, dependable, and fast. The following non-functional requirements are key:

- Easy:
  - Ease of development
  - Easy management at scale
  - Advanced job management
  - Multi-tenancy
  - Easy to access data by various user types
- Dependable:
  - Data protection with snapshot and mirroring
  - Automated self-healing
  - Insight into software/hardware health and issues
  - High availability (HA) and business continuity
- Fast:
  - Superior performance
  - Scalability
- Secure and governed:
  - Strong authentication and authorization
  - Kerberos support
  - Data confidentiality and integrity

4 Architectural overview

The Lenovo AI infrastructure solution is based on a flexible and scalable reference architecture. The data repositories can be a variety of storage systems including, for example, a Hadoop cluster. The primary hardware building block is the training node implemented on ThinkSystem SD530, SR670 or SR650 servers. A cluster of SD530, SR670 or SR650 servers are connected together to meet the desired total compute capacity required to deliver the best performance for DL training and model evaluation environment. For inference, the application environment can be based on Lenovo servers with CPU-only or GPU-enabled compute capabilities. Additionally, inference end-points can also be workstations, PCs and other devices.

Figure 1 shows the architecture overview of the AI reference architecture that uses Lenovo ThinkSystem hardware infrastructure.

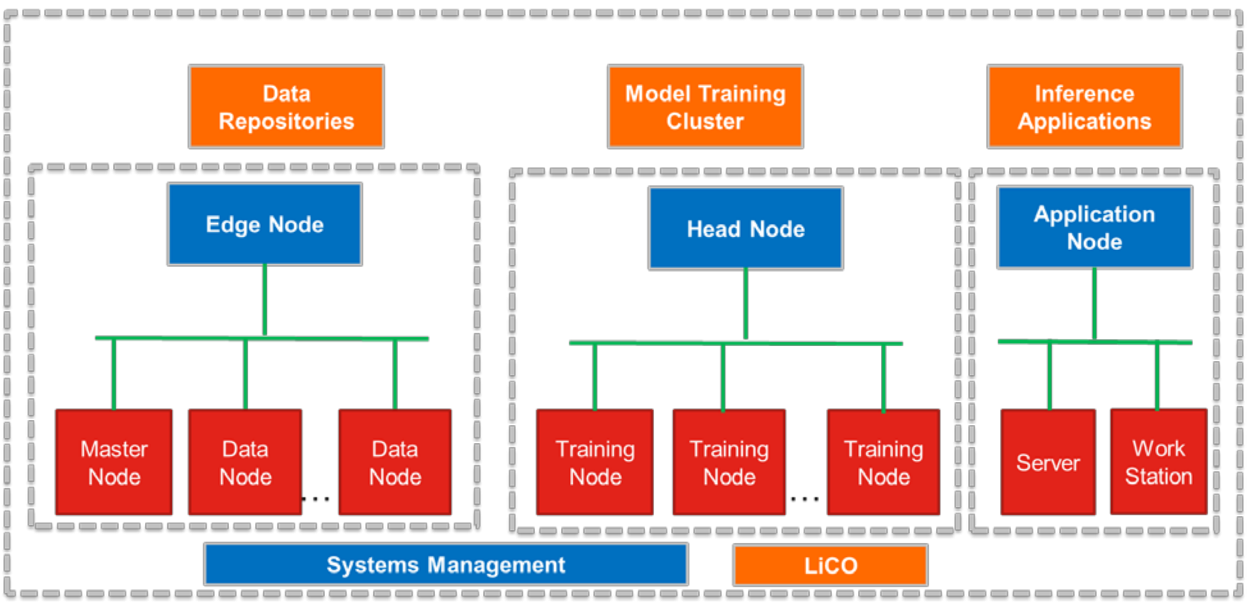

Figure 1. Lenovo AI reference architecture overview.

Access to data is crucial for a successful AI adoption project. The data in an enterprise can reside in a variety of repositories. A flexible infrastructure must allow access to data across all available data repositories where data of interest is stored. Figure 1 shows a Hadoop cluster as an example of a data repository. The cluster includes master nodes, data nodes and option edge nodes.

The training cluster is the central box containing the head node and multiple training nodes that allow the design to be scalable. Additional components such as the data repositories, systems management node, LiCO client node and the inference application nodes are also shown in the figure above. Collectively, these building blocks comprise a complete environment for AI adoption from data collection, preparation, model training and model deployment for inference in the target application. The focus of this document is the model training cluster shown above.

The inference cluster in Figure 1 illustrates different types of inference nodes. These can be servers, work stations including PCs and also devices. The application environment can be implemented on one or more nodes with a collection of inference end-points of various types.

5 Component model

This section describes the high-level software model of the AI training solution. The AI software stack is quickly evolving today with new and updated frameworks popping up almost monthly. Filtering through the numerous open source options can be time consuming and confusing. As shown in Figure 2, this reference architecture highlights frameworks Lenovo has tested and configured on Lenovo ThinkSystem hardware infrastructure.

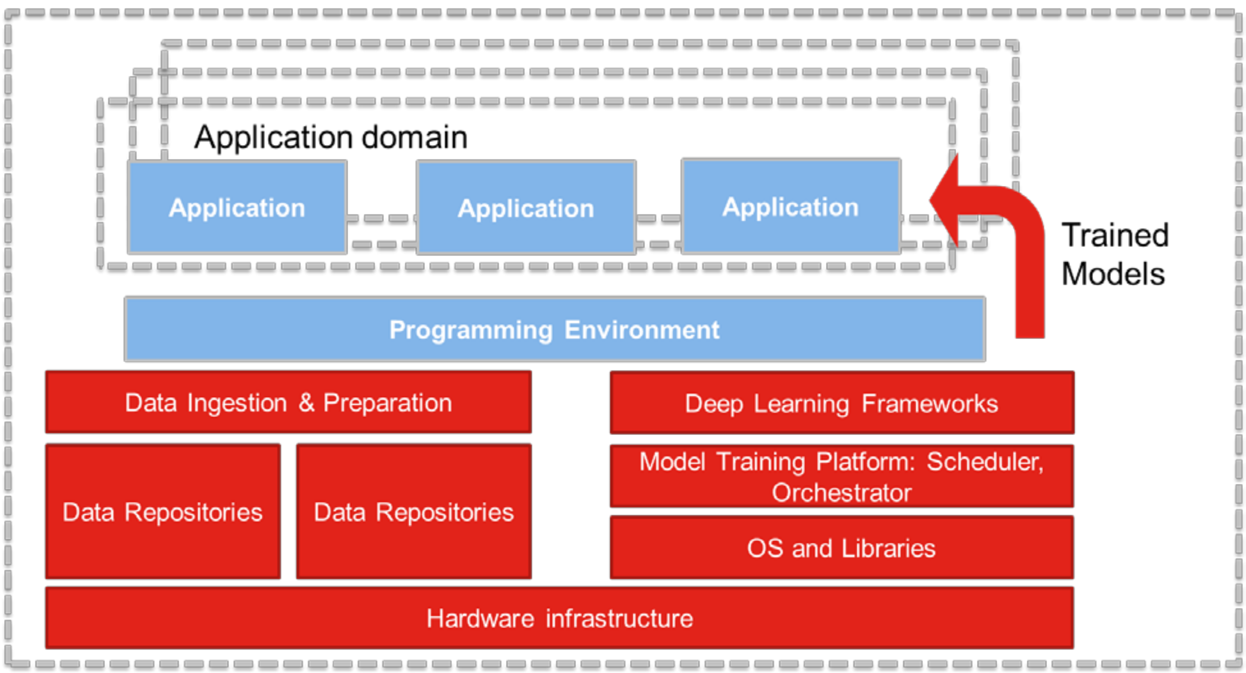

Figure 2. Lenovo model training solution component model.

5.1 AI Framework Overviews

5.1.1 Caffe

Caffe is one deep learning framework, developed by Berkeley AI Research (BAIR) and by community contributors. Caffe suports several different types of deep learning architectures targeted towards image classification and image segmentation.

Some of the advantages of Caffe are:

- Expression: plaintext schemas instead of code are used to define models and optimizations.
- Speed: training speed is crucial for state-of-the-art models and massive data found in research and industry today.
- Modularity: flexibility and adoption of new tasks and settings.
- Openness: common code, reference models, and reproducibility are required for scientific progress.

- Community: academic research, startup prototypes, and industrial applications all benefit from joint discussion and development in a BSD-2 project.

5.1.2 Intel Caffe

Software optimization is essential to high compute utilization and improved performance, and Intel Caffe is optimized for deep learning training. A few years ago, deep learning performance was sub-optimal on Intel CPUs as software optimizations were limited and compute utilization was low. Over the past two years, Intel has optimized deep learning functions achieving high utiliazation and enabling deep learning scientitist to use general-purpose Intel CPUs for DL training.

Intel Caffe is designed for both single-node and multi-node operation. There are two general approaches to parallelization: data parallelism and model parallelism. The Intel Distribution of Caffe uses data parallelism.

Data Parallelism

The data parallelization method runs training on different data batches on each of the nodes. The data is split among the nodes, but the same training model is used on all nodes. This means that the total batch size in a single iteration is equal to the sum of the individual batch sizes of all nodes.

The Intel Distribution of Caffe with the Machine Learning Scaling Library (MLSL) offers two approaches for multi-node training using data parallelism:

- Default - Caffe does Allreduce operation for gradients, then each node is doing stochastic gradient descent (SGD) locally, followed by Allgather for weights increments.
- Distributed weights update - Caffe does Reduce-Scatter operation for gradients, then each node is doing SGD locally, followed by Allgather for weights increments.

Data Distribution

One approach is to divide your training data set into disjoint subsets of roughly equal size. Distribute each subset onto each node used for training. Run the multinode training with the data layer prepared accordingly, which means either preparing separate proto configurations or placing each subset in exactly the same path for each node.

An easier approach, which was used in this RA, is to simply distribute the full data set to all nodes and configure the data layer to a draw a different subset on each node. The parameter shuffle:true was set for the training phase in prototxt. Since each node has its own unique randomizing seed, it will draw a unique image subset.

Node Communication

Intel Caffe utilizes the Intel Machine Learning Scaling Library (MLSL), which provides communication primitives for data parallelism and model parallelism, communication patterns for SGD, and distributed weight updates. It is optimized for Intel Xeon processors and supports the Intel Omni-Path Architecture, InfiniBand, and Ethernet.

5.1.3 TensorFlow

TensorFlow is another deep learning framework. It is an open source software library for high performance numerical computation. Its flexible architecture allows easy deployment of computation across a variety of platforms (CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices. Originally developed by researchers and engineers from the Google Brain team within Google's AI organization, it comes with strong support for machine learning and deep learning and the flexible numerical computation core is used across many other scientific domains.

Some of the main features of TensorFlow are:

- Programming support for machine learning methods including deep neural networks.
- Efficient calculation of mathematical expressions involving multi-dimensional arrays called tensors.
- Programming is hardware independent and the same code can be used CPU or GPU computing.
- High computation scalability across machines and huge data sets.

5.2 Software considerations

5.2.1 OS

The reference architecture is based on Red Hat Enterprise Linux (RHEL) 7.4 deployments. Users need to ensure that the latest official patches are installed on all nodes of the deployment and any necessary upgrades have been executed before starting the installation.
Collaboration between Intel and Red Hat yields a platform built for:

- Layered security due to hardware offload of encryption/decryption, OS-level security, and identity management in RHEL and Security Enhanced Linux (SELinux)
- Optimized performance with a combination of drivers, libraries, and extensions for cluster environments and Intel components

5.2.2 Libraries

- The Intel Math Kernel Library (Intel MKL) is a library of optimized mathematical functions of which only some of these functions are used for deep learning. Most deep learning models are built from a limited set of building blocks known as primitives that operate on tensors. These building blocks or low-level deep learning functions have been optimized for the Intel Xeon product family inside the Intel MKL library.
- The Intel MKL-DNN Library contains a subset of MKL functions targeted for deep learning. It was released open-source under an Apache 2 license with all the key building blocks necessary to build complex models. Intel MKL-DNN allows industry and academic deep learning developers to distribute the library and contribute new or improved functions.
- The NVIDIA CUDA Deep Neural Network library (cuDNN) is a GPU-accelerated library of primitives for deep neural networks. cuDNN provides highly tuned implementations for standard routines such as forward and backward convolution, pooling, normalization, and activation layers. cuDNN is part of the NVIDIA Deep Learning SDK.

5.3 Big Data Storage for Training Data

Data storage for training data can reside in a variety of storage environments. One such data repository is a Hadoop cluster implemented using Cloudera Enterprise software on Lenovo hardware. Other Hadoop cluster solutions are available on Lenovo hardware (see Resources section at the end of this document).

5.3.1 Cloudera Enterprise

Cloudera Enterprise provides features and capabilities to support mission-critical and real-time big data analytics across different industries, such as financial services, retail, media, healthcare, manufacturing, telecommunications, government organizations, and leading Fortune 100 and Web 2.0 companies. The Cloudera platform for big data can be used for various use cases from batch applications that use MapReduce or Spark with data sources, such as click streams, to real-time applications that use sensor data.

Users can log into the Cloudera client from outside the firewall by using Secure Shell (SSH) on port 22 to access the Cloudera solution from the corporate network. Cloudera provides several interfaces that allow administrators and users to perform administration and data functions, depending on their roles and access level. Hadoop application programming interfaces (APIs) can be used to access data. Cloudera APIs can be used for cluster management and monitoring. Cloudera data services, management services, and other services run on the nodes in cluster. Storage is a component of each data node in the cluster. Data can be incorporated into Cloudera Enterprise storage through the Hadoop APIs or network file system (NFS), depending on the needs of the customer.

Figure 3 shows the Cloudera Enterprise key capabilities that meet the functional requirements of customers.

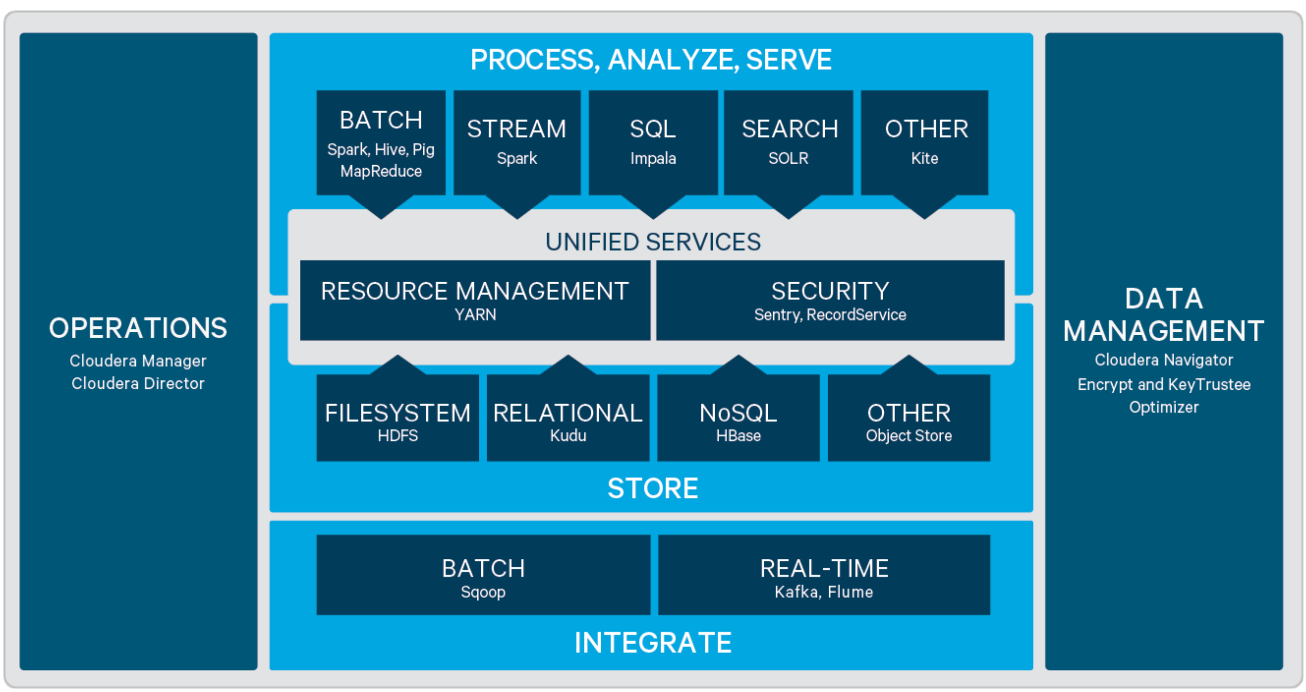

Figure 3. Cloudera Enterprise key capabilities

Cloudera Enterprise solution contains the following components relevant for use as a repository of data for model training:

- Core Components: HDFS, MapReduce and YARN are core components of Apache Hadoop that are included and supported in the Cloudera Enterprise edition. HDFS is a distributed file system designed to turn a cluster of servers into a massively scalable pool of storage. HDFS accepts data in any format regardless of schema and can scale linearly as more cluster nodes with storage capacity are added. MapReduce is a data processing software framework designed to match the massive scale of HDFS and Hadoop. YARN is designed to provide open source resource management for Hadoop to enable a diverse set of workloads like interactive SQL, advanced modeling and real-time streaming.
- Search Engine: Cloudera Search

  Cloudera Search is Apache Solr that is integrated with Cloudera Enterprise, including Apache Lucene, Apache SolrCloud, Apache Flume, Apache Tika, and Hadoop. Cloudera Search also includes valuable integrations that make searching more scalable, easy to use, and optimized for near-real-time and batch-oriented indexing. These integrations include Cloudera Morphlines, which is a customizable transformation chain that simplifies loading any type of data into Cloudera Search.
- NoSQL - HBase

  A scalable, distributed column-oriented datastore. HBase provides real-time read/write random access to very large datasets hosted on HDFS.
- Stream Processing: Apache Spark

  Apache Spark is an open source, parallel data processing framework that complements Hadoop to make it easy to develop fast, unified big data applications that combine batch, streaming, and interactive analytics on all your data. Cloudera offers commercial support for Spark with Cloudera Enterprise. Spark is 10 – 100 times faster than MapReduce which delivers faster time to insight, allows inclusion of more data, and results in better business decisions and user outcomes.

For more details on the Lenovo Big Data Validated Design for Cloudera Enterprise, see:

https://lenovopress.com/lp0776-lenovo-big-data-validated-design-for-cloudera-enterprise-and-vmware-thinksystem

5.4 Inference using Trained Models

Inference is the process of applying a trained model to new data. If the model is well trained on a large data set, the model output on new data should have a similar accuracy to the test data during training. There are several inference frameworks, but for this RA TensorFlow Serving was used.

5.4.1 TensorFlow Serving

TensorFlow Serving is a flexible, high-performance serving system for machine learning models, designed for production environments. TensorFlow Serving makes it easy to deploy new algorithms and experiments, while keeping the same server architecture and APIs. TensorFlow Serving provides out of the box integration with

TensorFlow models, but can be easily extended to serve other types of models.

TensorFlow Serving has several key concepts:

- **Servables** are the central abstraction in TensorFlow Serving. Servables are the underlying objects that clients use to perform computation (for example, a lookup or inference). TensorFlow Serving represents a **model** as one or more servables. A machine-learned model may include one or more algorithms (including learned weights) and lookup or embedding tables.
- **Loaders** manage a servable's life cycle. The Loader API enables common infrastructure independent from specific learning algorithms, data or product use-cases involved. Specifically, Loaders standardize the APIs for loading and unloading a servable.
- **Sources** are plugin modules that find and provide servables. Each Source provides zero or more servable streams. For each servable stream, a Source supplies one Loader instance for each version it makes available to be loaded.
- **Managers** handle the full lifecycle of Servables, including:
  - loading Servables
  - serving Servables
  - unloading Servables
- **TensorFlow Serving Core** manages (via standard TensorFlow Serving APIs) the following aspects of servables:
  - Lifecycle
  - Metrics

  TensorFlow Serving Core treats servables and loaders as opaque objects.

Figure 4 shows the high level framework for TensorFlow Serving

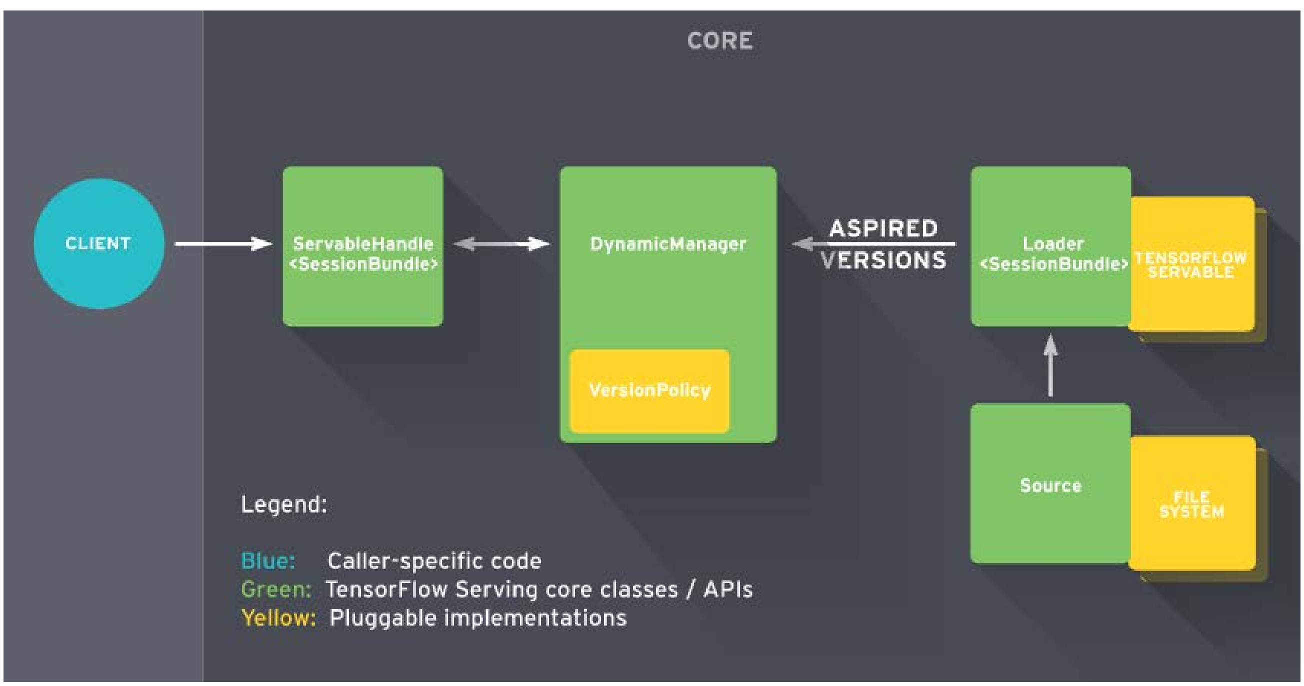

Figure 4. TensorFlow Serving Framework

Broadly speaking:

1. Sources create Loaders for Servable Versions.
2. Loaders are sent as Aspired Versions to the Manager, which loads and serves them to client requests.

For more information, see the TensorFlow Serving Overview:
https://www.tensorflow.org/serving/overview

5.5 Lenovo Intelligent Computing Orchestration (LiCO)

Lenovo intelligent Computing Orchestration (LiCO) is a software solution that simplifies the deployment, management, and use of distributed scale-out clusters for high performance computing (HPC) and Artificial Intelligence (AI) development.

By providing an intuitive interface (Figure 5), LiCO puts all your cluster resources at the fingertips of your users. From academic research to enterprise data centers, LiCO ensures a quicker time-to-value for your researchers, engineers, and data scientists, and reduces work for your cluster administrators and operational support.

LiCO accelerates AI and HPC development by reducing complexity for users, without restricting any of the robust orchestration tools to which your cluster users may be accustomed.

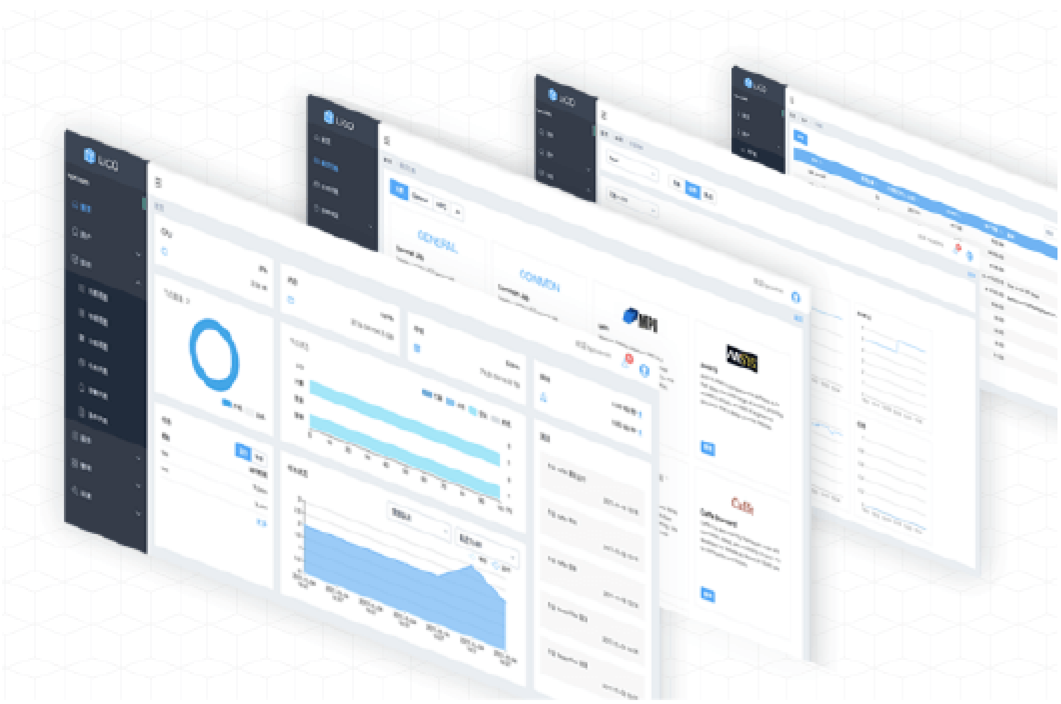

Figure 5. Lenovo Intelligent Computing Orchestration (LiCO)

LiCO offer a range of capabilities, including:

- Role-based access
- Easy management of data sets and DNN models

- Easy scheduling of training jobs
- Results review of training job results

For more information, see the Lenovo LiCO Datasheet:

https://lenovopress.com/datasheet/ds0029-lenovo-intelligent-computing-orchestrator

6 Operational model

This section describes the hardware infrastructure aspects of the AI reference architecture.

To show the operational model for different sized customer environments, this reference architecture describes several node configurations for supporting different DL training requirements.

6.1 Hardware description

6.1.1 Lenovo ThinkSystem SR650 Server

The Lenovo ThinkSystem SR650 is an ideal 2-socket 2U rack server for small businesses up to large enterprises that need industry-leading reliability, management, and security, as well as maximizing performance and flexibility for future growth. The SR650 server is particularly suited for ML & DL applications due to its rich internal data storage, large internal memory and selection of high performance Intel processors. It is also designed to handle general workloads, such as databases, virtualization and cloud computing, virtual desktop infrastructure (VDI), enterprise applications, collaboration/email, and business analytics.

The SR650 server supports:

- Up to two Intel Xeon Scalable Processors
- Up to 3.0 TB 2666 MHz TruDDR4 memory (certain CPU part numbers required),
- Up to 24x 2.5-inch or 14x 3.5-inch drive bays with an extensive choice of NVMe PCIe SSDs, SAS/SATA SSDs, and SAS/SATA HDDs
- Flexible I/O Network expansion options with the LOM slot, the dedicated storage controller slot, and up to 6x PCIe slots

Figure 6. Lenovo ThinkSystem SR650

Combined with the Intel Xeon Scalable Processors (Bronze, Silver, Gold, and Platinum), the Lenovo SR650 server offers an even higher density of workloads and performance that lowers the total cost of ownership (TCO). Its pay-as-you-grow flexible design and great expansion capabilities solidify dependability for any kind of workload with minimal downtime.

The SR650 server provides high internal storage density in a 2U form factor with its impressive array of workload-optimized storage configurations. It also offers easy management and saves floor space and power consumption for the most demanding use cases by consolidating storage and server into one system.

This reference architecture recommends the ThinkSystem SR650 for the following reasons:

- **Performance**: This hardware supports the latest Intel Xeon Scalable processors and TruDDR4 Memory.
- **Flexibility**: Server hardware uses embedded storage, which results in simple scalability (by adding nodes).
- **PCIe slots**: Up to 7 PCIe slots are available if rear disks are not used, and up to 3 PCIe slots if the Rear HDD kit is used. They can be used for network adapter redundancy and increased network throughput.
- **Higher power efficiency**: Titanium and Platinum redundant power supplies that can deliver 96% (Titanium) or 94% (Platinum) efficiency at 50% load.
- **Reliability:** Outstanding reliability, availability, and serviceability (RAS) improve the business environment and helps save operational costs

For more information, see the Lenovo ThinkSystem SR650 Product Guide:

https://lenovopress.com/lp0644-lenovo-thinksystem-sr650-server

6.1.2 Lenovo ThinkSystem SD530 Server

The Lenovo ThinkSystem SD530 is an ultradense and economical two-socket server in a 0.5U rack form factor. With four SD530 servers installed in either the ThinkSystem D2 Enclosure or ThinkSystem Modular Enclosure, you have an ideal high-density 2U four-node (2U4N) platform for enterprise and cloud workloads.

2U4N systems have gained popularity in a variety of data centers, from large enterprises to service providers, because their small footprint and inherent density make them ideal for building solution-based appliances at a low cost. The combination of the Lenovo ThinkSystem SD530 and D2 Enclosure is engineered to deliver these types of solutions.

Figure 7: Four ThinkSystem SD530 servers installed in a D2 Enclosure

For more information, see the Lenovo ThinkSystem SD530 Product Guide:
https://lenovopress.com/lp0635-thinksystem-sd530-server

6.1.3 Lenovo ThinkSystem SR670

The ThinkSystem SR670 is a purpose-built 2 socket 2U node for GPU enabled workloads. Supporting the latest NVIDIA GPUs and Intel Xeon Scalable processors, the SR670 has been designed for optimal performance for the high-end computation required by both Artificial Intelligence and High Performance Computing workloads. The SR670 delivers up to 4 GPUs per node and provides the optimal hardware for the LICO platform.

Figure 8. Lenovo ThinkSystem SR670

For more information, see the Lenovo ThinkSystem SR670 Product Guide:
https://lenovopress.com/lp0923-thinksystem-sr670-server

6.1.4 Lenovo RackSwitch G8052

The Lenovo System Networking RackSwitch G8052 (as shown in Figure 9) is an Ethernet switch that is designed for the data center and provides a simple network solution. The Lenovo RackSwitch G8052 offers up to 48x 1 GbE ports and up to 4x 10 GbE ports in a 1U footprint. The G8052 switch is always available for business-critical traffic by using redundant power supplies, fans, and numerous high-availability features.

Figure 9. Lenovo RackSwitch G8052

Lenovo RackSwitch G8052 has the following characteristics:

- Forty-eight 1 GbE RJ45 ports
- Four standard 10 GbE SFP+ ports
- Low 130W power rating and variable speed fans to reduce power consumption

For more information, see the Lenovo RackSwitch G8052 Product Guide:
https://lenovopress.com/tips1270-lenovo-rackswitch-g8052

6.1.5 Intel Omni-Path (OPA) 100 Series 48-port Unmanaged Edge Switch

The Intel Omni-Path Edge Switch consists of two models supporting 100 Gb/s for all ports, an entry-level 24-port switch for small clusters and a 48-port switch. The larger switch, in addition to enabling a 48-port fabric in 1U, can be combined with other edge switches and directors to build much larger multitier fabrics. These Intel Omni-Path Edge Switches are members of the Intel Omni-Path Fabric 100 series of switches, host adapters, and software delivering an exceptional set of high-speed networking features and functions.

For further information on the OPA 100 series switch, visit this link:

https://www.intel.com/content/www/us/en/products/network-io/high-performance-fabrics/omni-path-edge-switch-100-series.html

Figure 10. Intel Omni-Path (OPA) 100 Series 48-port Unmanaged Edge Switch

6.1.6 Mellanox SB7800 Series 36-port InfiniBand Switch

Built with Mellanox's Switch-IB™ 2 InfiniBand switch device, the SB7800 series provides up to 36 100Gb/s full bi-directional bandwidth per port. SB7800 smart network switch is designed to enable in-network computing through the Co-Design Scalable Hierarchical Aggregation Protocol (SHARP) technology. The Co-Design architecture enables the usage of all active data center devices to accelerate the communications frameworks using embedded hardware, resulting in order of magnitude applications performance improvements.

For further information on the InfiniBand SB7800 series switch, visit this link:

http://www.mellanox.com/page/products_dyn?product_family=225&mtag=sb7800

Figure 11. Mellanox SB7890 InfiniBand Switch

6.1.7 Lenovo ThinkSystem NE10032 Rack Switch

The Lenovo ThinkSystem NE10032 RackSwitch that uses 100 Gb QSFP28 and 40 Gb QSFP+ Ethernet technology is specifically designed for the data center. It is ideal for today's big data, cloud, and enterprise workload solutions. It is an enterprise class Layer 2 and Layer 3 full featured switch that delivers line-rate, high-bandwidth switching, filtering, and traffic queuing without delaying data. Large data center-grade buffers help keep traffic moving, while the hot-swap redundant power supplies and fans (along with numerous high-availability features) help provide high availability for business sensitive traffic.

Figure 12. Lenovo RackSwitch NE10032

For more information, see the Lenovo RackSwitch NE10032 Product Guide:
https://lenovopress.com/lp0609-lenovo-thinksystem-ne10032-rackswitch

6.1.8 Lenovo ThinkSystem NE2572 Rack Switch

The Lenovo ThinkSystem NE2572 RackSwitch that uses 25 Gb SFP28 and 10 Gb SFP+ Ethernet technology is specifically designed for the data center. The NE2572 is an extremely cost-effective switch for high-end 25GbE hybrid cloud data center applications.

Figure 13. Lenovo RackSwitch NE2572

For more information, see the Lenovo RackSwitch NE2572 Product Guide:
https://lenovopress.com/datasheet/ds0010-lenovo-thinksystem-ne2572-rackswitch

6.2 Cluster nodes

The AI reference architecture is implemented on a set of nodes that make up a cluster.

Training nodes run Caffe or TensorFlow services for model training. Head nodes are primarily used for storage including the training dataset and pre / post run training models. Head nodes are connected to the same high speed data network as training nodes for maximum performance. Head Nodes can use the Lenovo ThinkSystem SR650 servers with locally attached storage. Training nodes can use the Lenovo ThinkSystem SR650 servers, the SD530 servers for increased density, or the AI optimized ThinkSystem SR670 servers.

Inference nodes run an inference framework such as TensorFlow Serving. Inference nodes can be one or more Lenovo ThinkSystem SR650 or SD530 servers. For specific applications, a Lenovo ThinkStation P Series Tiny or ThinkPad P series laptop can be used for inference.

6.2.1 Training Nodes

Table 1, Table 2, &

Table ***3*** list the recommended system components for training nodes demonstrated in this reference architecture.

Table 1. SR650 Training node configuration

Component	Worker node configuration
Server	ThinkSystem SR650
Processor	2x Intel Xeon -processors: 6148 Gold, 20-core 2.4Ghz
Memory – base	192 GB: 12x 16GB 2666MHz RDIMM 384 GB: 12x 32GB 2666MHz RDIMM
Disk (OS)	Dual M.2 128GB SSD with RAID1
Disk (data)	2x 960GB 2.5" SATA SSD
Storage controller	OS: M.2 RAID1 mirror enablement kit Data: ThinkSystem 430-16i 12Gb controller
Hardware storage protection	OS: RAID1 Model Storage: None (JBOD).
Hardware management network adapter	Integrated 1G BaseT XCC management controller
Data network adapter	Intel OPA Adapter 100Gb or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter**
Accelerator (Optional)	2x NVIDIA Tesla V100 32GB PCIe GPUs or 2x NVIDIA Tesla P100 16GB PCIe GPUs

**NOTE: If SR650 is populated with two GPUs & a Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter, the network bandwidth is limited to 56Gb/s.

Table 2. SD530 Training node configuration

Component	Worker node configuration
Server	ThinkSystem SD530 with D2 Enclosure
Processor	2x Intel Xeon -processors: 6148 Gold, 20-core 2.4Ghz
Memory – base	192 GB: 12x 16GB 2666MHz RDIMM 384 GB: 12x 32GB 2666MHz RDIMM
Disk (OS)	Dual M.2 128GB SSD with RAID1
Disk (data)	2x 960GB 2.5" SATA SSD
Storage controller	OS: M.2 RAID1 mirror enablement kit Data: ThinkSystem on-board 6Gb SATA controller

Hardware storage protection	OS: RAID1 Model Storage: None (JBOD).
Hardware management network adapter	Integrated 1G BaseT XCC management controller
Data network adapter	Intel OPA Adapter 100Gb or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter
Accelerator (Optional)	2x NVIDIA Tesla V100 32GB PCIe GPUs or 2x NVIDIA Tesla P100 16GB PCIe GPUs

Table 3. SR670 Training node configuration

Component	Worker node configuration
Server	ThinkSystem SR670
Processor	2x Intel Xeon -processors: 6148 Gold, 20-core 2.4Ghz
Memory – base	192 GB: 12x 16GB 2666MHz RDIMM 384 GB: 12x 32GB 2666MHz RDIMM
Disk (OS)	Dual M.2 128GB SSD with RAID1
Disk (data)	2x 960GB 2.5” SATA SSD
Storage controller	OS: M.2 RAID1 mirror enablement kit Data: ThinkSystem on-board 6Gb SATA controller
Hardware storage protection	OS: RAID1 Model Storage: None (JBOD).
Hardware management network adapter	Integrated 1G BaseT XCC management controller
Data network adapter	Intel OPA Adapter 100Gb or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter
Accelerator	4x NVIDIA Tesla V100 32GB PCIe GPUs

The Intel Xeon Scalable Processor recommended in the above tables will provide a balance of performance vs. cost for DL training nodes. Higher core count and frequency processors are available for compute intensive workloads. The OS is loaded on a dual M.2 SSD memory module with RAID1 mirroring capability. Data disks are JBOD configured for fast training data throughput.

6.2.2 Head Nodes

The Head node is the storage center of the DL training cluster and supports several other key functions needed on a Deep Learning cluster. Table 3 lists the recommended components for a Head node, which can be customized according to client needs.

Table 4. Head node configuration

Component	Master node configuration
Server	ThinkSystem SR650
Processor	2x Intel Xeon -processors: 6130 Gold, 20-core 2.1Ghz
Memory – base	192 GB: 12x 16GB 2666MHz RDIMM 384 GB – 12x 32 GB 2666MHz RDIMM

Disk (OS / local storage)	OS: Dual M.2 128GB SSD with RAID1 Data: 24x 960GB 2.5" SATA SSD
Storage controller	OS: M.2 RAID1 mirror enablement kit Data: 3x ThinkSystem RAID 930-8i 4GB Flash 12Gb controller
Hardware storage protection	OS: RAID1 Data: RAID 5
Hardware management controller	Integrated 1G BaseT XCC management controller
Data network adapter	Intel OPA Adapter 100Gb or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter

The Intel Xeon Scalable Processors and minimum memory specified in Table 4 is recommended to provide sufficient performance as a Head node. The OS is loaded on a dual M.2 SSD memory module with RAID1 mirroring capability. Data disks are RAID configured for reliable data storage.

6.2.3 Big Data Nodes

The Cloudera reference architecture is implemented on a set of nodes that make up a cluster which includes two main node types: Worker nodes and Master nodes. Worker nodes use either the ThinkSystem SR650 servers with locally attached storage, or SD530 dense compute nodes with external SAS storage. Master nodes use ThinkSystem SR630 servers.

Worker nodes run data (worker) services for storing and processing data.

Master nodes run the following types of services:

- Management control services for coordinating and managing the cluster
- Miscellaneous and optional services for file and web serving

Known as Edge, System Management or Gateway nodes, these are installed on the cluster data network but do not run Cloudera Enterprise software directly. Their purpose is to connect the Cloudera cluster to an outside network for remote administration access, for ingesting data from an outside source, or for running end user application software which accesses the Cloudera Enterprise cluster.

A single system management/gateway node is configured in this reference architecture as a minimal node configured for remote administration of the Linux OS and for hardware maintenance. Based on the particular requirements of the cluster for high speed ingesting of data and edge node applications, the CPU, memory, storage, and network capability of this server can be increased.

For a detailed description of the Cloudera big data cluster, see the Lenovo Press reference architecture for Cloudera Enterprise:

https://lenovopress.com/lp0776-lenovo-big-data-validated-design-for-cloudera-enterprise-and-vmware-thinksystem

6.2.4 Inference Nodes

The inference nodes are implemented on Lenovo ThinkSystem SR650 and SD530. Table 5 & Table 6 list the recommended system components for training nodes demonstrated in this reference architecture.

Table 5. SR650 Inference node configuration

Component	**Worker node configuration**
Server	ThinkSystem SR650
Processor	2x Intel Xeon -processors: 6148 Gold, 20-core 2.4Ghz
Memory – base	192 GB: 12x 16GB 2666MHz RDIMM
Disk (OS)	Dual M.2 128GB SSD with RAID1
Storage controller	OS: M.2 RAID1 mirror enablement kit
Hardware storage protection	OS: RAID1
Hardware management network adapter	Integrated 1G BaseT XCC management controller
Data network adapter	Intel OPA Adapter 100Gb or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter**
Accelerator (Optional)	2x NVIDIA Tesla V100 32GB PCIe GPUs or 2x NVIDIA Tesla P100 16GB PCIe GPUs

**NOTE: If SR650 is populated with two GPUs & a Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter, the network bandwidth is limited to 56Gb/s.

Table 6. SD530 Inference node configuration

Component	**Worker node configuration**
Server	ThinkSystem SD530 with D2 Enclosure
Processor	2x Intel Xeon -processors: 6148 Gold, 20-core 2.4Ghz
Memory – base	192 GB: 12x 16GB 2666MHz RDIMM
Disk (OS)	Dual M.2 128GB SSD with RAID1
Storage controller	OS: M.2 RAID1 mirror enablement kit
Hardware storage protection	OS: RAID1
Hardware management network adapter	Integrated 1G BaseT XCC management controller
Data network adapter	Intel OPA Adapter 100Gb or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter
Accelerator (Optional)	2x NVIDIA Tesla V100 32GB PCIe GPUs or 2x NVIDIA Tesla P100 16GB PCIe GPUs

6.3 Lenovo XClarity

Systems management of a cluster includes hardware management, Operating System, and Caffe applications management.

Hardware management uses the Lenovo XClarity Administrator, which is a centralized resource management solution that reduces complexity, speeds up response and enhances the availability of Lenovo server systems and solutions. XClarity is used to install the OS onto new worker nodes; update firmware across the cluster nodes, record hardware alerts and report when repair actions are needed.

Figure 14 shows the Lenovo XClarity Administrator interface in which servers, storage, switches and other rack components are managed and status is shown on the dashboard. Lenovo XClarity Administrator is a virtual appliance that is quickly imported into a server virtualized environment.

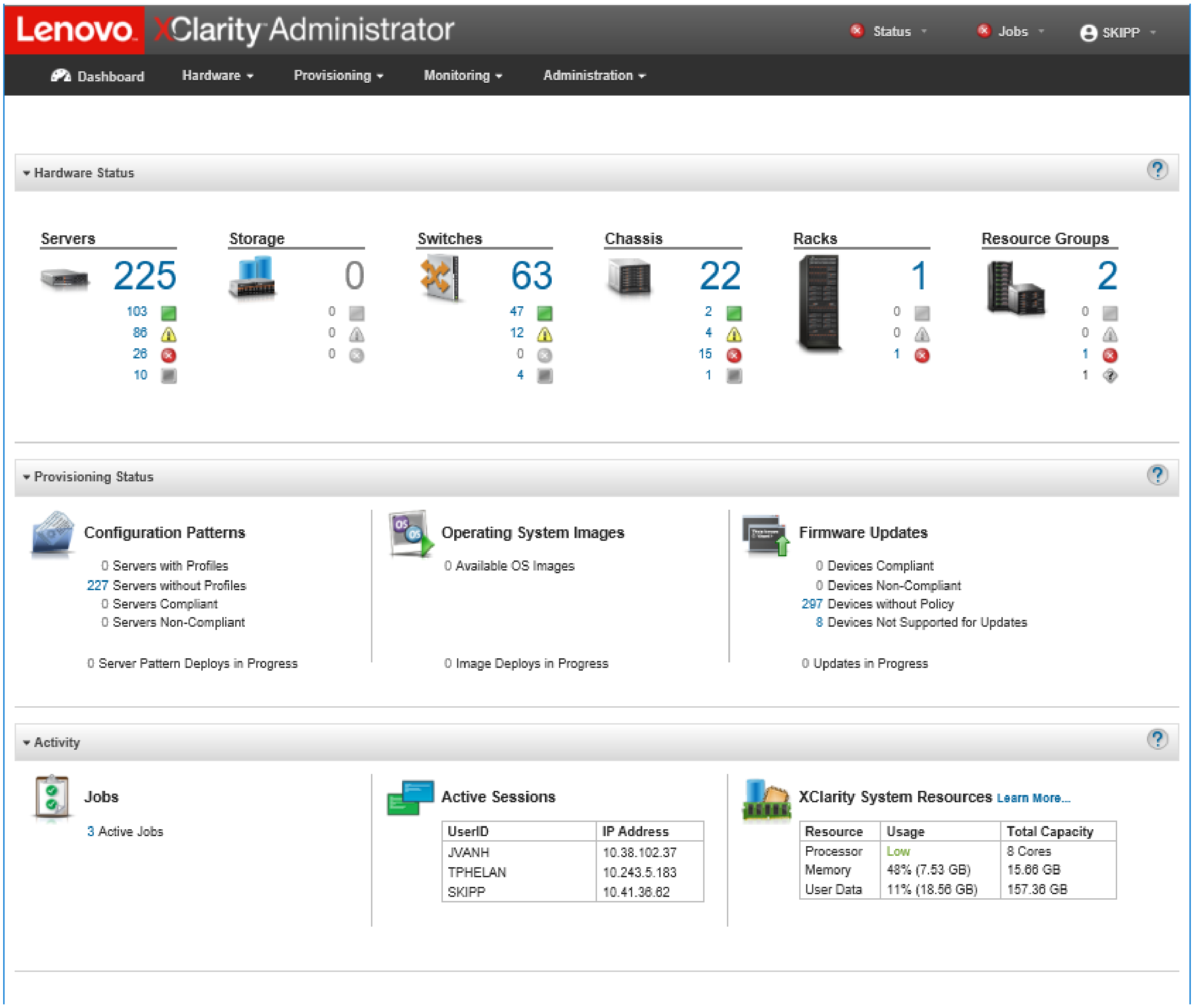

Figure 14. XClarity Administrator interface

In addition, the Extreme Cluster Administration Toolkit (xCAT) provides a scalable distributed computing management and provisioning tool that provides a unified interface for hardware control, discovery and operating system deployment. It can be used to facilitate or automate the management of cluster nodes. For more information about xCAT, see “Resources” on page 20. xCAT can be combined with Confluent to enable additional management features and a graphical interface to the cluster. For more information about xCAT & Confluent, see “Resources” on at the end of this document.

6.4 Networking

The reference architecture specifies two networks: a high-speed data network and a management network..

6.4.1 Data network

The data network creates a private cluster among multiple nodes and is used for high-speed data transfer across training nodes, and also for importing data into the cluster. The recommended OPA switch is the Intel Omni-Path (OPA) 100 Series 48-port Unmanaged Edge Switch that provides 32 100Gb OPA ports. The recommended IB switch is the Mellanox SB7890 IB 2 EDR Switch that provides 36 100Gb IB ports. The recommended 100GbE switch is the Lenovo NE10032 that provides 32 100GbE ports. If 25GbE is used for the data network, the recommended switch is the Lenovo NE2572 switch that provides 48 25GbE ports.

Figure 15 shows a network design for big data, training, and inference. It is recommended the big data cluster use 25 GbE for node connections to a NE10032 switch configured for 25GbE or the 25GbE native NE2572 switch. The training cluster performs best using 100Gb for node connections, as shown in Figure 17. Inference nodes can be connected with 25GbE or 100Gb, if used in a high throughput application.

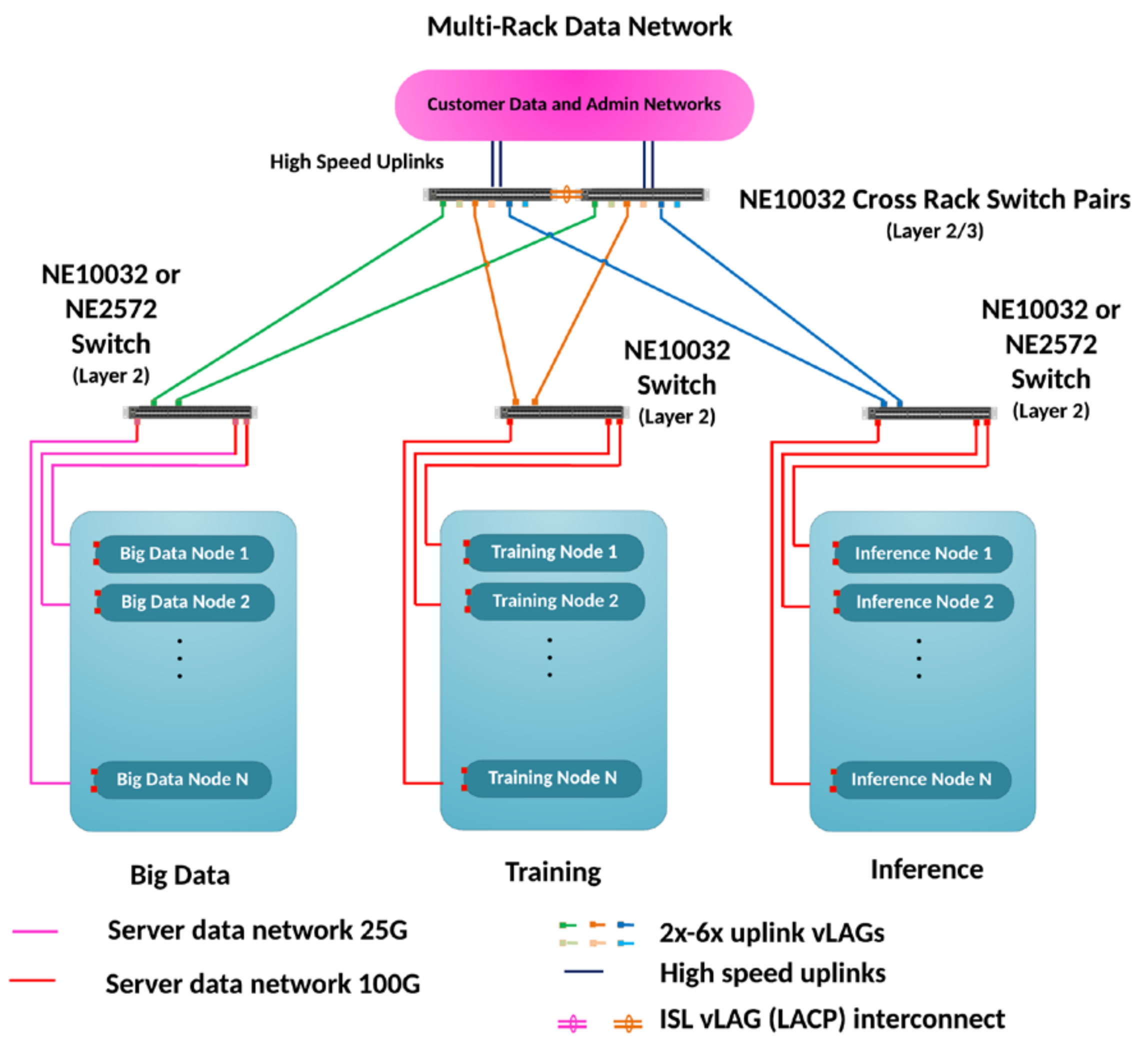

Figure 15. Data Network Diagram

6.4.2 Hardware management network

The hardware management network is a 1GbE network for out-of-band hardware management. The recommended 1GbE switch is the Lenovo RackSwitch G8052 with 10Gb SFP+ uplink ports. Through the XClarity Controller management module (XCC) within the ThinkSystem SR650 and SD530 servers, the out-of-band network enables hardware-level management of cluster nodes, such as node deployment, UEFI firmware configuration, hardware failure status and remote power control of the nodes.

Intel Caffe functions have no dependency on the XCC management controller. The cluster Data and OS management networks can be shared with the XCC hardware management network, or can be separated via VLANs on the respective switches. The cluster and hardware management networks are then typically connected directly to the customer's existing admin network to facilitate remote maintenance of the cluster.

6.5 Predefined training cluster configurations

The intent of the predefined configurations provided in this reference architecture is to ease initial sizing for customers and to show example starting points for different sized training workloads. These consist of training nodes, head nodes, network switches, cabling, and rack hardware. The SR650 is recommended for DL training clusters requiring fewer than four nodes. To reduce cost for these small cluster designs, the SR650 can be multi purposed as a training node and head node. For larger clusters, the SD530 provides a dense solution requiring 75% less rack space vs the SR650. Table 7 below, shows the hardware required for each predefined cluster size.

Table 7. Predefined DL cluster configurations

	Single Node	Single Chassis	Two Chassis	Four Chassis	16 Chassis
Training Node Model	SR650	SD530	SD530	SD530	SD530
Training Nodes	1	4	8	16	64
D2 Enclosures	0	1	2	4	16
Head Nodes	0	1	1	1	2
Head Node Storage	24x 960GB	24x 960GB	24x 960GB	24x 960GB	48x 960GB

Adding GPUs to SR650s, SD530s, or SR670s can greatly decrease the training time required for certain types of workloads. The SR650 is recommended for DL training clusters requiring fewer than two GPUs. For larger clusters, the SD530 provides a dense solution requiring 50% less rack space vs the SR650. The SR670 is specifically designed for training with GPUs, and can hold up to four GPUs per 2U system. Table 8 below, shows the hardware required for each predefined cluster size when using GPU accelerators.

Table 8. Predefined DL cluster configurations with GPU Accelerators

	Single Node	Single Chassis		Two Chassis		Four Chassis		16 Chassis	
Training Node Model	SR650	SR670	SD530	SR670	SD530	SR670	SD530	SR670	SD530
Training Nodes	1	1	2	2	4	4	8	16	32
D2 Enclosures	0	0	1	0	2	0	4	0	16
Head Nodes	0	1		1		1		2	
Head Node Storage	24x 960GB	24x 960GB		24x 960GB		24x 960GB		48x 960GB	
Accelerators (GPUs)	2	4		8		16		64	

7 Deployment considerations

This section describes various other considerations for deploying the AI solution.

The predefined configurations represent a set of baseline configurations that can be implemented as specified or modified based on specific client requirements, such as lower cost, improved performance, and increased reliability.

When you consider modifying the predefined configuration, you must understand key aspects of how the cluster will be used. In terms of training, you must understand the current and future type and size of data to be used for training, and whether training speed is a significant factor. Consider also the computing needs of other uses of the cluster and whether many of them will need to run in parallel.

When designing the Deep Learning cluster infrastructure, we recommend conducting the necessary testing and proof of concepts against representative data and workloads to ensure that the proposed design will achieve the necessary success criteria. The following sections provide information about customizing the predefined configuration. When considering customizations to the predefined configuration, work with a systems architect who is experienced in designing DL cluster infrastructures.

7.1 Increasing cluster performance

There are a few approaches that can be used to increase cluster training performance: increasing the node count, and selecting higher performing hardware. Adding additional training nodes (scale-out design) is the easiest way to increase cluster performance now or in the future, especially if data training requirements change. Selecting higher performance CPUs or adding GPUs, requires a better understanding of the training requirements during the cluster design phase. Once a CPU or GPU model is selected for a cluster, it is recommended that model be used in all training nodes. Often, improving performance comes at increased cost, and you must consider the cost-to-benefit trade-offs of designing for higher performance.

7.2 Designing for lower cost

There are several key modifications that can be made to lower the cost of a DL reference architecture solution. When lower-cost options are considered, it is important to ensure that customers understand the potential lower performance implications of a lower-cost design. A lower-cost version of the DL reference architecture can be achieved by using lower-cost node processors, reducing the amount of memory capacity per data node and using lower-cost storage drives. The node processors can be substituted with other processors in the Intel Xeon SP family. Selecting a different processor may lead to a lower frequency memory, which can also lower the per-node cost of the solution.

The use of a smaller memory capacity per data node can provide a lower-cost design. However, the performance of a cluster with data nodes using the lower memory capacity can be lower. Testing during proof-of-concept evaluation should be done with real user data to understand the performance implications.

The storage drives on a data node can be changed to a lower-capacity drive or a standard HDD. The impact on the performance of data nodes using these lower-cost storage options should be evaluated during proof-of-concept testing as mentioned before.

7.3 Scaling considerations

When the capacity of the existing infrastructure is reached, the cluster can be scaled out by adding more training nodes and, if necessary, head nodes. As the capacity of existing racks is reached, new racks can be added to the cluster. Some training workloads might not scale linearly.

When you design a new Deep Learning solution reference architecture implementation, future scale out is a key consideration in the initial design. You must consider the two related aspects of networking and management. Both of these aspects are critical to cluster operation and become more complex as the cluster infrastructure grows.

The networking model described in the section “Networking” in the previous section is designed to provide robust network interconnection of racks within the cluster. As more racks are added, the predefined networking topology remains balanced and symmetrical. If there are plans to scale the cluster beyond one rack, initially design the cluster with multiple racks, even if the initial number of nodes might fit within one rack. Starting with multiple racks will enforce proper network topology and prevent future reconfiguration and hardware changes.

Also, as the number of nodes within the cluster increases, many of the tasks of managing the cluster also increase, such as updating node firmware or operating systems. Building a cluster management framework as part of the initial design and proactively considering the challenges of managing a large cluster will pay off significantly in the end.

Proactive planning for future scale out and the development of cluster management framework as a part of the initial cluster design provides a foundation for future growth that will minimize hardware reconfigurations and cluster management issues as the cluster grows.

8 Appendix: Bill of Material

This section includes the Bill of Materials (BOMs) for different configurations of hardware for the AI Solution. There are sections for training nodes, head nodes, and networking.

The BOM includes the part numbers, component descriptions and quantities. Table 7 & Table 8 list the quantity of each core component defined in each of the predefined configurations.

The BOM lists in this appendix are not meant to be exhaustive and must be verified with the ordering configuration tools. Any discussion of pricing, support and maintenance options is outside the scope of this document.

This BOM information is for the United States; part numbers and descriptions can vary in other countries. Other sample configurations are available from your Lenovo sales team. Components are subject to change without notice.

8.1 Training Node

Table 9 lists the BOM for the SR650 training node, and Table 10 lists the BOM for the SR650 training node with GPUs.

Table 9. Training node: SR650

Code	Description	Qty
7X06CTO1WW	SR650 Training : ThinkSystem SR650 - 3yr Warranty	1
AUVX	ThinkSystem SR650 2.5" Chassis with 8 or 16 bays	1
AWDX	Intel Xeon Gold 6148 20C 150W 2.4GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	2
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
B0ZR	ThinkSystem 2.5" Intel S4600 960GB Mainstream SATA 6Gb Hot Swap SSD	2
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AU0B or ASWQ	Intel OPA 100 Series Single-port PCIe 3.0 x16 HFA or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter	1
AXCA	ThinkSystem Toolless Slide Rail	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
AVWG	ThinkSystem 1600W (230V) Platinum Hot-Swap Power Supply	2
AURQ	Lenovo ThinkSystem 2U 3FH Riser Bracket	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
B173	XClarity™ Controller Standard to Enterprise Upgrade in factory	1
AUSF	Lenovo ThinkSystem 2U MS CPU Performance Heatsink	2

Table 10. Training node: SR650 with GPUs

Code	Description	Qty
7X06CTO1WW	SR650 Training : ThinkSystem SR650 - 3yr Warranty	1
AUVX	ThinkSystem SR650 2.5" Chassis with 8 or 16 bays	1
AWDX	Intel Xeon Gold 6148 20C 150W 2.4GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	2
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
B0ZR	ThinkSystem 2.5" Intel S4600 960GB Mainstream SATA 6Gb Hot Swap SSD	2
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AU0A or ASWQ	Intel OPA 100 Series Single-port PCIe 3.0 x8 HFA or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter	1
AXCA	ThinkSystem Toolless Slide Rail	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
AVWG	ThinkSystem 1600W (230V) Platinum Hot-Swap Power Supply	2
AURQ	Lenovo ThinkSystem 2U 3FH Riser Bracket	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
B173	XClarity™ Controller Standard to Enterprise Upgrade in factory	1
AULQ	ThinkSystem 1U CPU Performance Heatsink	2
AX4F	ThinkSystem SR650 Air Duct for GPU	1
B34S or B0M1	ThinkSystem NVIDIA Tesla V100 32GB PCIe Passive GPU or ThinkSystem NVIDIA Tesla P100 16GB PCIe Passive GPU	1/2

Table 11 lists the BOM for the SD530 training node, and Table 12 list the BOM for the SD530 training node with GPUs. Table 13 lists the BOM for the D2 enclosure.

Table 11. Training node: SD530

Code	Description	Qty
7X21CTO1WW	Stark : ThinkSystem SD530 - 3yr Warranty	1
AUXN	ThinkSystem SD530 Computing Node	1
AWEW	Intel Xeon Gold 6148 20C 150W 2.4GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AUYJ	ThinkSystem SD530 2x2 SAS/SATA BP	1
AUYL	ThinkSystem SD530 SW RAID Kit	1
B0ZR	ThinkSystem 2.5" Intel S4600 960GB Mainstream SATA 6Gb Hot Swap SSD	2
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
AUYM	ThinkSystem SD530 Front VGA/USB KVM Breakout Module	1
AVUT	ThinkSystem XClarity Controller Standard to Advanced Upgrade	1
AU0B or ASWQ	Intel OPA 100 Series Single-port PCIe 3.0 x16 HFA or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter	1

Table 12. Training node: SD530 with GPUs

Code	Description	Qty
7X21CTO1WW	Node : ThinkSystem SD530 - 3yr Warranty	1
B0M3	ThinkSystem SD530 Computing Node for GPU Tray	1
AWEW	Intel Xeon Gold 6148 20C 150W 2.4GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AUYJ	ThinkSystem SD530 2x2 SAS/SATA BP	1
AUYL	ThinkSystem SD530 SW RAID Kit	1
B0ZR	ThinkSystem 2.5" Intel S4600 960GB Mainstream SATA 6Gb Hot Swap SSD	2
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
AUYM	ThinkSystem SD530 Front VGA/USB KVM Breakout Module	1
AVUT	ThinkSystem XClarity Controller Standard to Advanced Upgrade	1
AU0B or ASWQ	Intel OPA 100 Series Single-port PCIe 3.0 x16 HFA or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter	1
B0MU	ThinkSystem SD530 GPU Tray	1
B13L	ThinkSystem SD530 GPU Riser Module	1
B34S or B0M1	ThinkSystem NVIDIA Tesla V100 32GB PCIe Passive GPU or ThinkSystem NVIDIA Tesla P100 16GB PCIe Passive GPU	1/2

Table 13*. **Training node: D2 Chassis***

Code	Description	Qty
7X20CTO1WW	D2 Chassis : ThinkSystem D2 Enclosure -3yr Warranty	1
AUXM	ThinkSystem D2 Enclosure	1
AUY9	ThinkSystem D2 10Gb 8 port EIOM SFP+	1
AUY8	ThinkSystem D2 4-slot x16 Shuttle	1
AUYC	ThinkSystem D2 Slide Rail	1
AUZ2	ThinkSystem D2 2000W Platinum PSU	2
A4VP	1.0m, C13 to C14 Jumper Cord, Rack Power Cable	2

Table 14 lists the BOM for the SR670 training node with GPUs.

Table 14*. **Training node: SR670 with GPUs***

Code	Description	Qty
7Y37CTO1WW	SR670 Training : ThinkSystem SR670 - 3yr Warranty	1
B3XX	L1 Base Banner CTO1WW	1
AWDX?	Intel Xeon Gold 6148 20C 150W 2.4GHz Processor	2
AWDW	Intel Xeon Gold 6142 16C 150W 2.6GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
B42K	1U 2.5" SATA/SAS 8-Bay Backplane	1
B3Y8	Blaze Riser	1
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
B0ZR	ThinkSystem 2.5" Intel S4600 960GB Mainstream SATA 6Gb Hot Swap SSD	2
AUZY	ThinkSystem I350-T2 PCIe 1Gb 2-Port RJ45 Ethernet Adapter	1
AU0B or ASWQ	Intel OPA 100 Series Single-port PCIe 3.0 x16 HFA or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter	1
B47V	Banner Slide Rail	1
B3YC	2000W Platinum PSU	2
B3Y7	TLA Sonic 2 Planar Board MB Banner	1
AVUT	ThinkSystem XClarity Controller Standard to Advanced Upgrade	1
B34S	ThinkSystem NVIDIA Tesla V100 32GB PCIe Passive GP	1-4

8.2 Head Node

Table 15 lists the BOM for the head node.

Table 15: Head node: SR650

Code	Description	Qty
7X06CTO1WW	SR650 Head Node : ThinkSystem SR650 - 3yr Warranty	1
AUVV	ThinkSystem SR650 2.5" Chassis with 8, 16 or 24 bays	1
AWEN	Intel Xeon Gold 6130 16C 125W 2.1GHz Processor	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5" SATA/SAS 8-Bay Backplane	3
AUNJ	ThinkSystem RAID 930-8i 2GB Flash PCIe 12Gb Adapter	3
AUMV	ThinkSystem M.2 with Mirroring Enablement Kit	1
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH Riser 1 Kit	1
AURC	ThinkSystem SR550/SR590/SR650 (x16/x8)/(x16/x16) PCIe FH Riser 2 Kit	1
AUKK	ThinkSystem 10Gb 4-port SFP+ LOM	1
AU0B or ASWQ	Intel OPA 100 Series Single-port PCIe 3.0 x16 HFA or Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI Adapter	1
AXCA	ThinkSystem Toolless Slide Rail	1
AURD	ThinkSystem 2U left EIA Latch Standard	1
AVWG	ThinkSystem 1600W (230V) Platinum Hot-Swap Power Supply	2
B0ZR	ThinkSystem 2.5" Intel S4600 960GB Mainstream SATA 6Gb Hot Swap SSD	24
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps Non-Hot Swap SSD	2
AURQ	Lenovo ThinkSystem 2U 3FH Riser Bracket	1
AURP	Lenovo ThinkSystem 2U 2FH Riser Bracket	1
AUNP	FBU345 SuperCap	3
AUQB	Lenovo ThinkSystem Mainstream MB - 2U	1
5977	Select Storage devices - no configured RAID required	1
AUSA	Lenovo ThinkSystem M3.5" Screw for EIA	4
AURR	ThinkSystem M3.5 Screw for Riser 2x2pcs and SR530/550/558/570/590 Planar 5pcs	4
B0ML	Feature Enable TPM on MB	1
6311	2.8m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
B173	XClarity™ Controller Standard to Enterprise Upgrade in factory	1

8.3 Management network switch

Table 16 lists the BOM for the Management/Administration network switch.

Table 16. Management/Administration network switch

Code	Description	Qty
7159HC1	Lenovo RackSwitch G8052 (Rear to Front)	1
ASY2	Lenovo RackSwitch G8052 (Rear to Front)	1
A3KR	Air Inlet Duct for 442 mm RackSwitch	1
A3KP	Adjustable 19" 4 Post Rail Kit	1

8.4 Data network switch options

Table 17 lists the BOM for the OPA data network switch.

Table 17. OPA Data network switch

Code	Description	Qty
0449HCR	Switch : Intel OPA 100 Series 48-port Unmanaged Edge Switch (PSE)	1
AU08	Intel OPA 100 Series 48-port Unmanaged Edge Switch (PSE)	1
AU0C	Intel OPA 100 Series Edge Switch Management Card	1
AU0L	3m Intel OPA 100 Series Passive Copper QSFP28 Cable (one per node)	1

Table 18 lists the BOM for the IB data network switch.

Table 18. IB Data network switch

Code	Description	Qty
0724Y13	Switch : Mellanox SB7890 EDR IB Unmanaged Switch (PSE)	1
00MP560	3m Mellanox EDR IB Passive Copper QSFP28 Cable	1

Table 19 lists the BOM for the 100GbE data network switch.

Table 19. 100GbE Ethernet Data network switch

Code	Description	Qty
7159HE1	Switch : Lenovo ThinkSystem NE10032 RackSwitch (Rear to Front)	1
AV17	Lenovo ThinkSystem NE10032 RackSwitch (Rear to Front)	1
A3KP	Adjustable 19" 4 Post Rail Kit	1
AV20	Lenovo 3m Passive 100G QSFP28 DAC Cable	1

Table 20 lists the BOM for the 25GbE data network switch.

Table 20. 25GbE Ethernet Data network switch

Code	Description	Qty
7159HE3	Switch : Lenovo ThinkSystem NE2572 RackSwitch (Rear to Front)	1

AV19	Lenovo ThinkSystem NE2572 RackSwitch (Rear to Front)	1
A3KP	Adjustable 19" 4 Post Rail Kit	1
AV1X	Lenovo 3m Passive 25G SFP28 DAC Cable	1

8.5 Lenovo Intelligent Computing Orchestration (LiCO)

Table 21 lists the ordering options for LiCO.

Table 21. LiCO

Code	Description	Qty
B1YC	Lenovo HPC AI LiCO Software 90 Day Evaluation License (no support or updates)	1
B1Y9	Lenovo HPC AI LiCO Software w/1Yr S&S	1
B1YA	Lenovo HPC AI LiCO Software w/3Yr S&S	1
B1YB	Lenovo HPC AI LiCO Software w/5Yr S&S	1

9 Appendix: Example Training Workload

This section describes an example environment used in the validation of this AI training solution.

9.1 AI Frameworks

9.1.1 Intel Caffe

Intel Caffe version 1.1.0, Intel Math Kernel Library (MKLML) 2018.0.1.20171227, and Intel MKL-DNN v0.12 were used. Intel Caffe was compiled using gcc, with Intel Skylake optimized compiler flags.

9.1.2 TensorFlow

TensorFlow version 1.8 and CUDA version 9.0.176 were used for GPU training. TensorFlow was downloaded pre-compiled.

9.2 Training Dataset

The ImageNet database was used. ImageNet is an image database organized according to the WordNet hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. The 1.28M ImageNet database used for training is a large-scale image database used by researchers and data scientists around the world.

9.3 Training Model

A 50 layer Residual Network (ResNet-50) neural network was trained. ResNet-50 was released in 2015 by Microsoft Research for use in the ImageNet and MS-COCO competitions.

9.4 Training Length

The ResNet-50 model was trained on the 1.28M ImageNet Dataset for 90 epochs. One epoch consists of one full training cycle on the training set, so for a 90 epoch training run, each image has passed through the training model 90 times. This means a 90 epoch run on the 1.28M ImageNet Dataset processed over 115 million images.

9.5 Training Data

Figure 16 shows the image training data for several different system configurations. This chart can be used as a sizing guide for what type of cluster configuration is required to reach a desired image-processing rate. The blue bars are CPU only training nodes with no GPUs, which are good for workloads requiring less than 1000 images/sec. The green bars are nodes configured with P100 GPUs, and are a good cost to performance option for workloads of less than 2000 images/sec. The orange and red bars are for systems configured with V100 GPUs, which can process about 5000 images/sec using 8 GPUs. The right axis scale uses the same data, but converted to millions of images per day.

As an example, let's use a 2000 images/sec requirement for training. Looking at Figure 16 for 2000 images/sec, there are multiple ways to meet this training requirement: 8x P100s in four SD530 nodes, 4x V100s in two SD530 nodes, or 4x V100s in a single SR670 system. For this particular example, Lenovo's recommendation would be to use a single SR670 system with 4x V100s. This configuration has the least number of CPUs and lowest memory to meet the 2000 images/sec training requirement, and therefore should have the lowest.

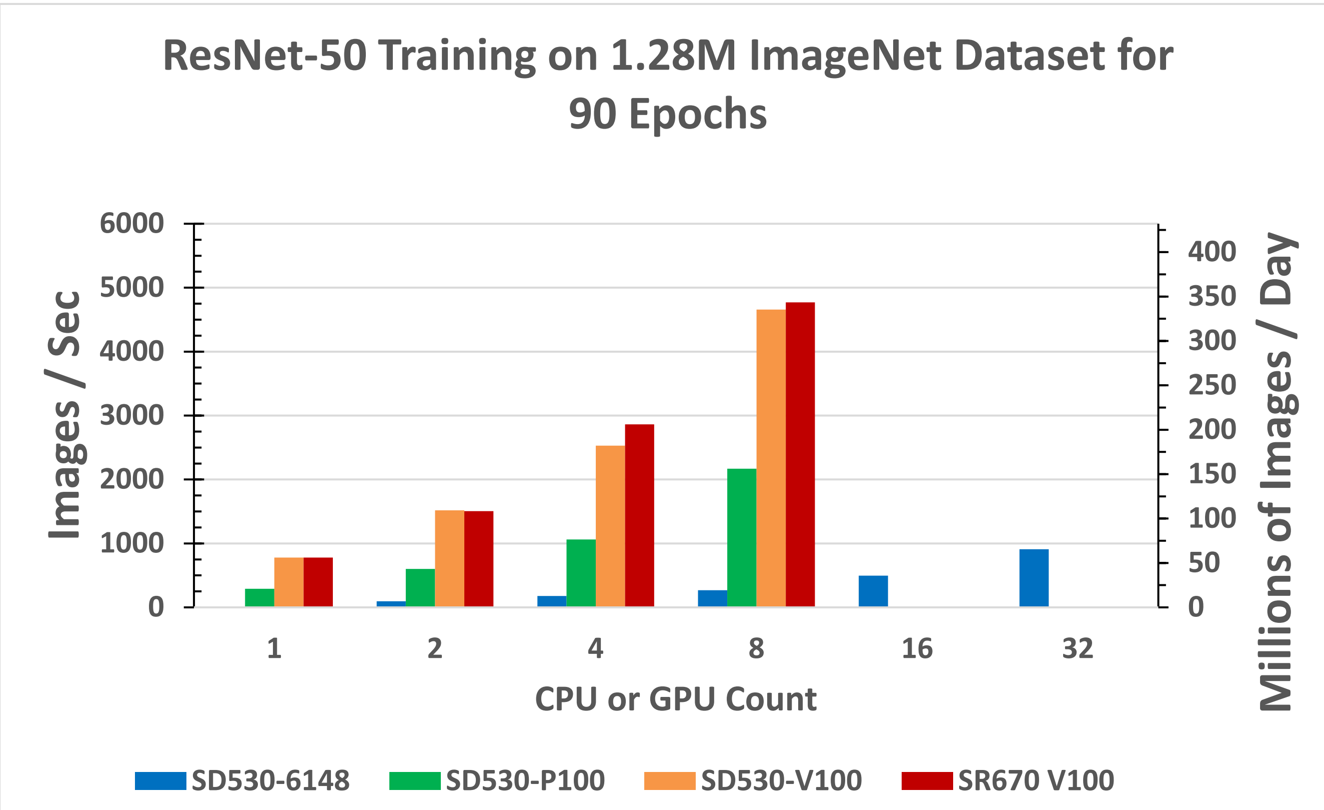

Figure 16. Training Data

Note: This data was collected using default Intel Caffe and TensorFlow settings.

9.6 Training Network Comparison

Figure 17 shows the image training data for several different network configurations. This chart shows training performance (images / sec) for several different network speeds including 10, 25, & 100 Gb/s Ethernet and

100 Gb/s OPA. Data was collected using 1, 2, and 4 SD530 servers each containing two V100 GPUs. All networks have the same performance for two GPUs, because this only requires one server and therefor does not require a data network. With two nodes using four GPUs the network starts to have an impact on training performance. The impact increases for four nodes using eight GPUs, which shows 10 Gb/s Ethernet has a 20% performance decrease compared to 100 Gb/s OPA. The decrease in performance for slower networks will continue to increase as the GPU and node count increase, so Lenovo recommends using a 100 Gb/s network when using 8 or more GPUs.

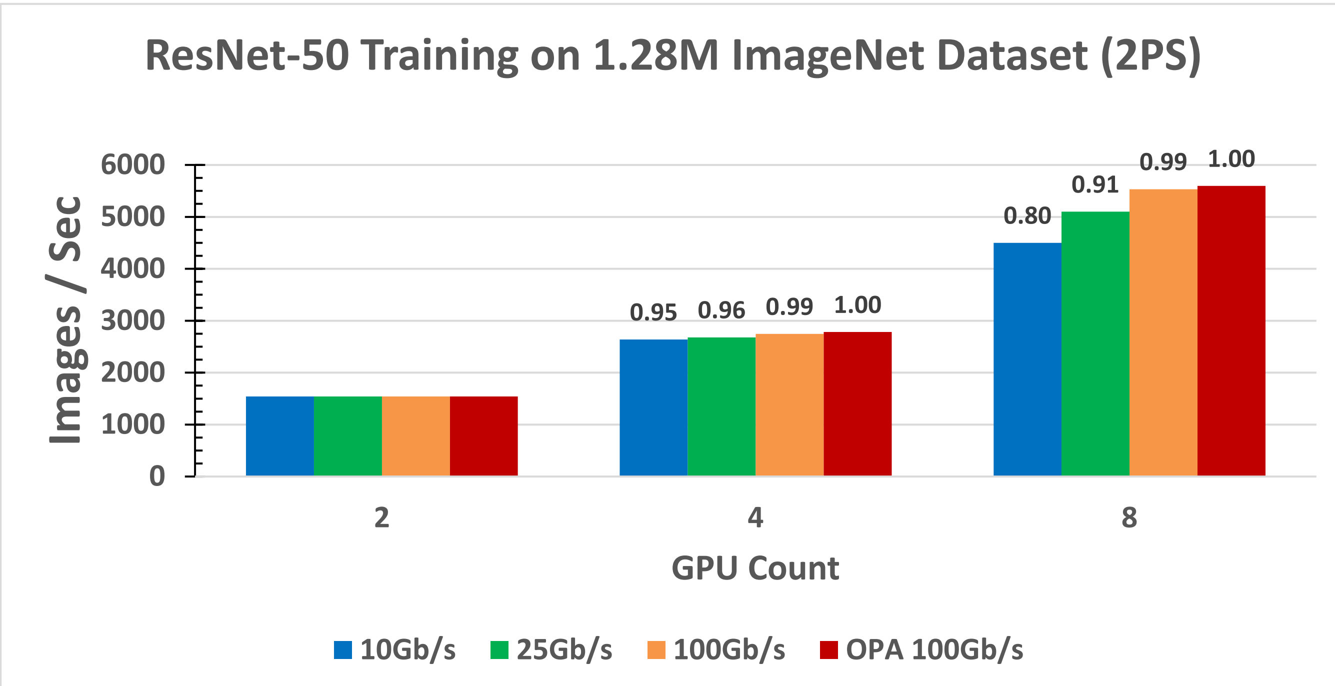

Figure 17. Training Network Comparison

9.7 Inference Framework

TensorFlow Serving release 1.11.1 was used for inference testing. TensorFlow Serving was downloaded as a Docker container, and was run in Docker-CE version 18.06.1. Models were trained in TensorFlow and exported using the SavedModel function. SavedModels are TensorFlow version independent and portable, unlike checkpoint files. SavedModels are compatible with TensorFlow Serving and can be loaded by the Docker container at start-up or embedded in the container.

9.8 Big Data Framework

Cloudera 5.9.1 was used for HDFS storage testing. The HDFS NFS Gateway service was installed on one of the management nodes, which provided a NFS mount path from the training nodes into the big data cluster. This service can be installed on multiple nodes for increased bandwidth into the HDFS storage cluster.

For more details on the Lenovo Big Data Validated Design for Cloudera Enterprise, see:

https://lenovopress.com/lp0776-lenovo-big-data-validated-design-for-cloudera-enterprise-and-vmware-thinksystem

Resources

For more information, see the following resources:

Lenovo ThinkSystem SR650 (Training/Head node):

- Lenovo Press product guide: https://lenovopress.com/lp0644-lenovo-thinksystem-sr650-server
- 3D Tour: https://lenovopress.com/lp0673-3d-tour-thinksystem-sr650

Lenovo ThinkSystem SD530 (Training node):

- Lenovo Press product guide: https://lenovopress.com/lp0635-thinksystem-sd530-server
- 3D Tour: https://lenovopress.com/lp0667-3d-tour-thinksystem-sd530

Lenovo ThinkSystem SR670 (Training node):

- Lenovo Press product guide: https://lenovopress.com/lp0923-thinksystem-sr670-server

Lenovo RackSwitch G8052 (1GbE Switch):

- Lenovo Press product guide: https://lenovopress.com/tips1270-lenovo-rackswitch-g8052

Lenovo RackSwitch NE10032 (100GbE Switch):

- Lenovo Press product guide: https://lenovopress.com/lp0609-lenovo-thinksystem-ne10032-rackswitch

Lenovo RackSwitch NE2572 (25GbE Switch):

- https://lenovopress.com/datasheet/ds0010-lenovo-thinksystem-ne2572-rackswitch

Lenovo XClarity Administrator:

- Lenovo Press product guide: https://lenovopress.com/tips1200-lenovo-xclarity-administrator

Lenovo Big Data Validated Designs:

- Lenovo Press reference architecture for Cloudera Enterprise: https://lenovopress.com/lp0776-lenovo-big-data-validated-design-for-cloudera-enterprise-and-vmware-thinksystem
- Lenovo Press reference architecture for Hortonworks Data Platforom: https://lenovopress.com/lp0828-lenovo-big-data-hortonworks-thinksystem-ra
- Lenovo Press reference architecture for MapR Converged Data Platform: https://lenovopress.com/lp0845-lenovo-big-data-validated-design-for-mapr-converged-data-platform

Intel OPA 100 (100Gb Switch)

- https://www.intel.com/content/www/us/en/products/network-io/high-performance-fabrics/omni-path-edge-switch-100-series.html

Mellanox SB7800 Series 36-port InfiniBand Switch (100Gb Switch)

- http://www.mellanox.com/page/products_dyn?product_family=225&mtag=sb7800

Intel Xeon Scalable Family Balanced Memory

- https://lenovopress.com/lp0742-intel-xeon-scalable-family-balanced-memory-configurations

NVIDIA GPUs

- V100 https://www.nvidia.com/en-us/data-center/tesla-v100/
- P100 https://www.nvidia.com/en-us/data-center/tesla-p100/

Caffe

- Caffe Tutorial: http://caffe.berkeleyvision.org/tutorial/

Intel Caffe

- Intel Distribution of Caffe: https://github.com/BVLC/caffe/tree/intel
- Multinode Guide: https://github.com/intel/caffe/wiki/Multinode-guide

TensorFlow

- TensorFlow: https://www.tensorflow.org
- TensorFlow Serving: https://www.tensorflow.org/serving/

ImageNet: http://image-net.org/index

ResNet-50: https://www.kaggle.com/keras/resnet50

Open source software:

- xCAT: xcat.org
- Lenovo xCAT & confluent http://hpc.lenovo.com/users/documentation/

Document history

Version 1.0	May 15, 2018	First version
Version 1.5	June 22, 2018	Second version • Added V100 GPUs in both SD530 and SR650 Training Node Configurations • Added TensorFlow AI software framework • Updated BOM tables to include configurations with V100 GPUs • Updated BOM tables to include configurations with InfiniBand switches and adapters
Version 1.6	July 27, 2018	Third version • Added P100 GPUs in both SD530 and SR650 Training Node Configurations • Updated BOM tables to include configurations with P100 GPUs
Version 1.7	September 7, 2018	Fourth Version • Added SR670 Training Node Configurations • Updated BOM tables to include configurations with SR670 • Updated BOM tables to include configurations with 100Gb Ethernet switches
Version 2.0	November 9, 2018	Fifth Version • Added inference • Added big data storage • Updated BOM tables to include configurations with 25Gb Ethernet switches

Trademarks and special notices

Information is provided "AS IS" without warranty of any kind.

All customer examples described are presented as illustrations of how those customers have used Lenovo products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information concerning non-Lenovo products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by Lenovo. Sources for non-Lenovo list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. Lenovo has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-Lenovo products. Questions on the capability of non-Lenovo products should be addressed to the supplier of those products.

All statements regarding Lenovo future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local Lenovo office or Lenovo authorized reseller for the full text of the specific Statement of Direction.

Some information addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in Lenovo product announcements. The information is presented here to communicate Lenovo’s current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard Lenovo benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Any references in this information to non-Lenovo websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this Lenovo product and use of those websites is at your own risk.