Lenovo AI 验证设计（适用于在 ThinkSystem 服务器上构建灵活、可扩展的 AI 基础结构）

上次更新时间：**23 November 2018**

2.0 版

介绍使用 Intel Optimized Caffe 和 TensorFlow 的 AI 基础结构参考架构

基于利用 Intel Xeon SP CPU 和 NVIDIA GPU 的强大的、多功能 Lenovo ThinkSystem SD530、SR650 和 SR670 服务器的解决方案

针对高性能、经济高效和可扩展解决方案的部署注意事项

包含适用于不同服务器和相关网络的详细硬件配置

Peter Seidel (Lenovo)

Ajay Dholakia (Lenovo)

Nathan Stewart (Lenovo)

目录

1 简介 5

2 业务问题及业务价值 6

2.1 业务问题 6

2.2 业务价值 6

3 要求 7

3.1 功能性要求 7

3.2 非功能性要求 7

4 架构概述 8

5 组件模型 9

5.1 AI 框架概述 9

5.1.1 Caffe 9

5.1.2 Intel Caffe 10

5.1.3 TensorFlow 10

5.2 软件注意事项 11

5.2.1 操作系统 11

5.2.2 库 11

5.3 用于训练数据的大数据存储 11

5.3.1 Cloudera Enterprise 11

5.4 使用已训模型的推理 13

5.4.1 TensorFlow Serving 13

5.5 联想智能超算平台（LiCO） 14

6 运行模式 16

6.1 硬件描述 16

6.1.1 Lenovo ThinkSystem SR650 服务器 16

6.1.2 Lenovo ThinkSystem SD530 服务器 17

6.1.3 Lenovo ThinkSystem SR670 18

6.1.4 Lenovo RackSwitch G8052 18

6.1.5 Intel Omni-Path (OPA) 100 系列 48 端口非管理型边缘交换机 19

6.1.6 Mellanox SB7800 系列 36 端口 InfiniBand 交换机 19

6.1.7 Lenovo ThinkSystem NE10032 机架交换机 20
6.1.8 Lenovo ThinkSystem NE2572 机架交换机 20
6.2 集群节点 20
6.2.1 训练节点 21
6.2.2 头节点 22
6.2.3 大数据节点 23
6.2.4 推理节点 23
6.3 Lenovo XClarity 24
6.4 网络 25
6.4.1 数据网络 25
6.4.2 硬件管理网络 26
6.5 预定义的训练集群配置 27
7 部署注意事项 28
7.1 提高集群性能 28
7.2 低成本设计 28
7.3 扩展注意事项 28
8 附录：物料清单 30
8.1 训练节点 31
8.2 头节点 35
8.3 管理网络交换机 36
8.4 数据网络交换机选项 36
8.5 联想智能超算平台（LiCO） 37
9 附录：示例训练工作负载 38
9.1 AI 框架 38
9.1.1 Intel Caffe 38
9.1.2 TensorFlow 38
9.2 训练数据集 38
9.3 训练模型 38
9.4 训练长度 38
9.5 训练数据 39
9.6 训练网络比较 40
9.7 推理框架 40

9.8 大数据框架 41

资源 42

文档历史记录 44

1 简介

本文档介绍在 Lenovo ThinkSystem 服务器上构建灵活、可扩展的人工智能（AI）基础结构的参考架构。它提供一个预定义且经过优化的硬件基础结构，适用于多种使用场景下的数据访问、模型训练和推理。此参考架构提供使用 Lenovo 产品构建 AI 基础结构的规划、设计注意事项以及最佳实践信息。

随着企业可获得的数据越来越多、种类越来越丰富、速度越来越快，如何从中获得最大价值已成为一大挑战。要应对这一挑战，必须使用在经过调整的硬件平台上运行的合适的数据处理和分析软件。随着下一代分析算法——机器学习（ML）和深度学习（DL）等 AI 技术的普及，企业正采用这些技术来现代化其应用程序和 IT 架构。尤其对于使用图像识别和模式检测技术的应用程序而言，通过开发和部署合适的 DL 模型可实现大幅度性能提升。

AI 应用的过程包含以下关键步骤：

- 数据访问
- 模型训练
- 推理

提供数据访问的任务必然涉及与多种数据存储库的连接。数据存储库的例子包括数据仓库、Hadoop 集群、云存储及其他数据存储系统等。

AI 应用的另一个关键步骤是了解及选择深度学习模型。典型模型一般基于深度神经网络（DNN），并需要大量计算资源用于训练。对于此类模型训练用例而言，使用横向扩展集群设计的硬件基础结构是实现 DL 应用的一个关键要求。

评估 DL 模型需要数据科学团队运行多个训练作业，然后比较结果以选择最佳模型及其关联的参数值。联想智能超算平台（LiCO）是一个拥有简单易用的 GUI、用于计划及安排此类模型训练作业的工具。

推理步骤着眼于在目标应用程序环境中部署并使用经过训练的模型。经过训练的模型用于对图像或模式等新数据进行决策。对于经过良好调试的推理阶段而言，其目标是让决策的准确度达到与在模型测试阶段中所观察到的相似。部署推理系统时，必须将吞吐量和延迟等指标考量在内。

本参考架构的目标受众是 IT 专业人员、技术架构师、销售工程师以及协助通过 Lenovo 硬件来规划、设计、实施高级分析解决方案的顾问。本文档假设您熟悉常见 AI 软件组件及功能。有关详细信息，请参阅本文档末尾的“资源”一节。

2 业务问题及业务价值

本部分介绍与 AI 基础结构相关的业务问题，以及 Lenovo 硬件和软件提供的价值。

2.1 业务问题

不论企业规模如何，数据爆炸为其创造了从数据中获得价值的巨大机会。如利用得当，机器学习和深度学习等新技术可带来更快、更好的深度见解。企业当今面临一大挑战，即如何建立一个可帮助存储、处理及使用其所获得的海量数据的 IT 基础结构。这一巨大的 AI 应用挑战包括以下部分：需从各种存储库中访问相关数据；需建立一个环境供数据科学家探索各种 ML 和 DL 模型、评估相关训练算法，以找到适合目标应用程序使用的模型；以及需在目标应用程序中部署经过训练的模型。

此外，他们还面临另一挑战，即如何保障在企业内部拥有足够算力和存储容量以应对海量的企业数据。尤其对于训练 DL 模型而言，它需要大量运算，因此必须提供合适的硬件基础结构供数据科学团队使用。

最后，要满足 DL 模型评估的要求，还迫切需要易用的工具集，以便数据科学家快速提高工作效率。如何在规模合理、配置得当的硬件基础结构上管理及安排各种 DL 模型训练作业始终至关重要。

2.2 业务价值

此参考架构中所述的 AI 基础结构解决方案为数据存储库、训练 DL 模型以及部署已训模型进行推理提供了预先配置的起点，可帮助解决上述主要的业务挑战。在确定合适的硬件基础结构和软件工具来创建适合的 AI 实施环境方面，用户可以省去大量的尝试工作。因此，能够快速使用合适的环境以及简单易用的工具可提供巨大的业务价值。

此参考架构所述的解决方案可为企业的数据科学团队提供灵活、可扩展的硬件基础结构。根据 DL 模型集合的不同，该解决方案受目标应用域、数据可用性和随着时间推移的增长驱动，提供采用横向扩展集群组织形式的计算和存储服务器配置，以帮助提供关于寻找规模合适、内部部署的硬件这一难题的解决方案。

此外，该解决方案还介绍了易于使用且可高效管理和安排 DL 模型评估作业的软件工具。

3 要求

本部分介绍此参考架构的功能性和非功能性要求。

3.1 功能性要求

AI 基础结构解决方案满足以下关键功能要求：

- 可处理各种数据存储库以满足训练时的数据访问
- 可处理 DL 领域内的各种模型
- 可部署行业标准 AI 软件框架
- 可处理各种类型的海量数据
- 可支持在各种应用程序环境中部署经过训练的模型
- 各种客户端接口

3.2 非功能性要求

客户要求 AI 模型训练解决方案具有简单、可靠、快速的特点。以下非功能性要求至关重要：

- 简单：
  - 易于开发
  - 易于规模化管理
  - 高级作业管理
  - 多用户
  - 支持不同用户类型轻松访问数据
- 可靠：
  - 具有快照和镜像数据保护功能
  - 自动进行自我修复
  - 可深入查看软件/硬件运行状况和问题
  - 高可用性（HA）和业务连续性
- 快速：
  - 提供顶级性能
  - 可扩展性
- 安全且受控：
  - 强大的认证和授权功能
  - 支持 Kerberos
  - 提供数据保密性和完整性

4 架构概述

Lenovo AI 基础结构解决方案基于一个灵活、可扩展的参考架构。数据存储库可以是各种存储系统，包括 Hadoop 集群在内。主要硬件基础部分是在 ThinkSystem SD530、SR670 或 SR650 服务器上实施的训练节点。SD530、SR670 或 SR650 服务器构成的集群相互连接，以满足为 DL 训练以及模型评估环境提供最佳性能所需的总计算能力的要求。对于推理部分，应用环境可基于仅包含 CPU 算力或支持 GPU 算力的 Lenovo 服务器。此外，推理端点也可以是工作站、电脑及其他设备。

图 1 显示了使用 Lenovo ThinkSystem 硬件基础结构的 AI 参考架构的概览。

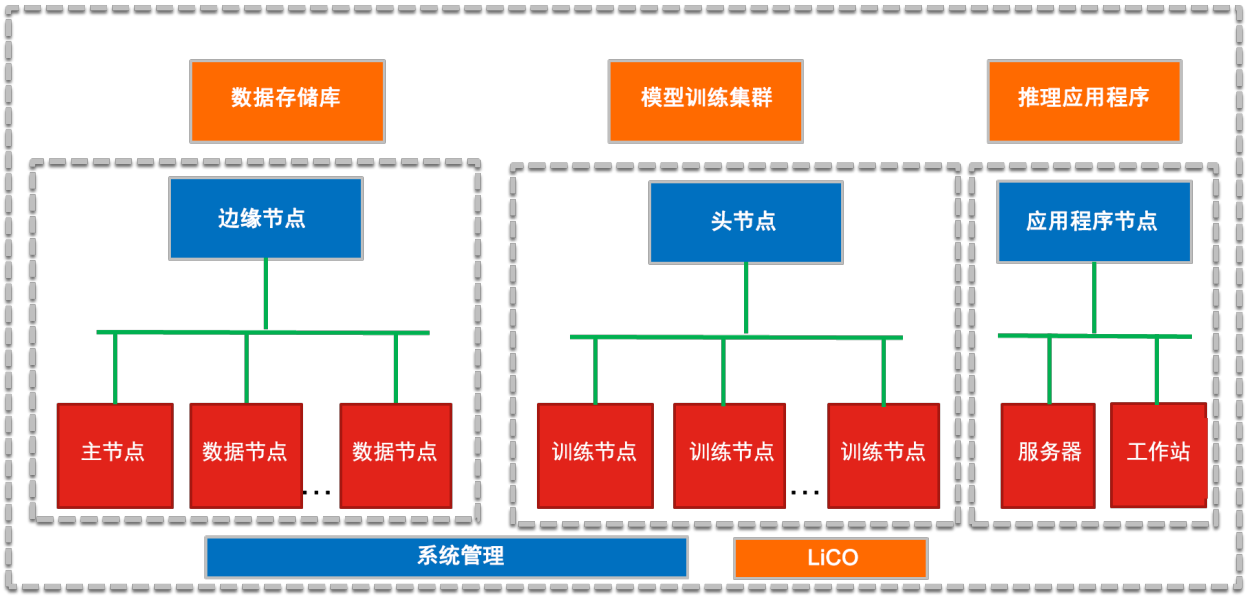

图 1. Lenovo AI 参考架构概览。

数据访问是成功 AI 应用项目的关键部分。企业的数据可能位于不同的存储库中。一个灵活的基础结构必须支持跨所有存储了相关数据的可用数据存储库进行数据访问。图 1 显示了以 Hadoop 集群为例的数据存储库示例。该集群包含主节点、数据节点和可选的边缘节点。

训练集群是中心组件，包括头节点和支持设计扩展的多个训练节点。上图中还显示了其他组件，例如数据存储库、系统管理节点、LiCO 客户端节点和推理应用程序节点。这些基础部分共同组成了一个完整的 AI 应用环境，其中包括数据收集、准备、模型训练以及在目标应用程序中部署模型以进行推理。本文档重点介绍上图所示的模型训练集群。

图 1 中的推理集群显示了不同类型的推理节点。这些节点可以是服务器、包括电脑在内的工作站以及其他设备。可以在包含不同类型推理端点的一个或多个节点上实施应用程序环境。

5 组件模型

本部分介绍 AI 训练解决方案的高级软件模型。如今 AI 软件栈的演化日新月异，几乎每个月都会出现升级的新框架。在无数开源方案中进行选择十分耗时且困难。如图 2 所示，此参考架构主要采用经过 Lenovo 测试的框架，并在 Lenovo ThinkSystem 硬件基础结构上进行配置。

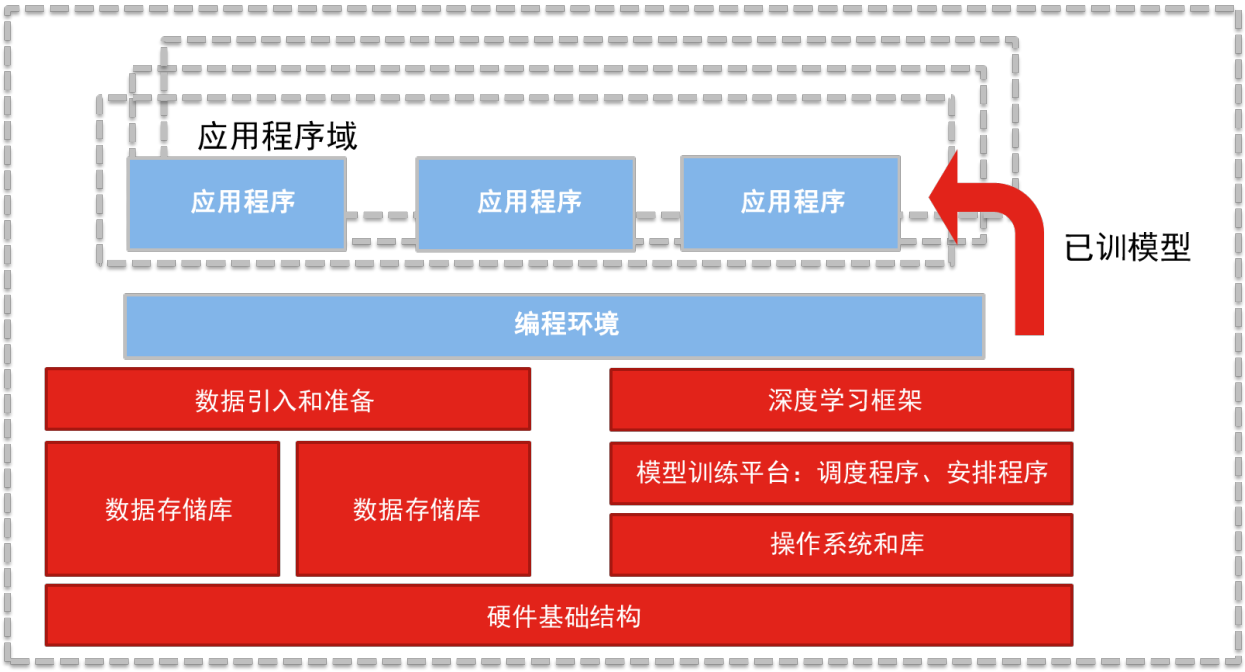

图 2. Lenovo 模型训练解决方案组件模型。

5.1 AI 框架概述

5.1.1 Caffe

Caffe 是一个由 Berkeley AI Research（BAIR）和社区参与者开发的深度学习框架。Caffe 支持若干种针对图像分类和图像分割的不同深度学习架构。

Caffe 的优势包括：

- 表达：使用纯文本形式而非代码来定义模型和优化。
- 速度：对于高级模型以及当今研究和行业中获得的海量数据来说，训练速度至关重要。
- 模块化：灵活性高，易于应用新任务和新设置。
- 开放性：具备推动科研进步所需的公共代码、参考模型和可重复性。
- 社区：学术研究、初创原型以及工业应用均可从共同讨论和 BSD-2 项目的开发成果中受益。

5.1.2 Intel Caffe

软件优化对于高计算利用率和性能提升来说至关重要，而 Intel Caffe 针对深度学习训练进行了优化。数年前，由于软件优化有限且计算利用率较低，Intel CPU 上的深度学习性能未能达到最优。过去两年中，Intel 已完成深度学习功能的优化，实现了高利用率，让深度学习科研人员可使用通用 Intel CPU 进行 DL 训练。

Intel Caffe 适用于单节点和多节点操作。并行化有两种常规方法：数据并行和模型并行。Intel 的 Caffe 发行版使用的是数据并行。

数据并行

数据并行方法在各节点上运行不同的数据批。数据在节点间拆分，但所有节点使用同一训练模型。这意味着，一次迭代总的数据批大小等于所有节点上各个数据批大小之和。

包含 Machine Learning Scaling Library（MLSL）的 Intel Caffe 发行版提供两种使用数据并行进行多节点训练的方法：

- 缺省 – Caffe 对梯度执行 Allreduce 运算，然后各节点在本地执行随机梯度下降（SGD），然后对权重增量执行 Allgather。
- 分布式权重更新 – Caffe 对梯度执行 Reduce-Scatter 运算，然后各节点在本地执行 SGD 下降，然后对权重增量执行 Allgather。

数据分发

一种方法是将训练数据集划分为大小大致相等的非连续子集。然后将各子集分发到各训练节点上。使用进行了相应准备的数据层运行多节点训练，即需要准备独立的原型配置或将各子集为各节点放在完全相同的位置。

还有一种更简单的方法（此参考架构中使用的方法），即将完整数据集分发给所有节点，然后配置数据层将不同子集抽取到各节点。prototxt 中对训练阶段设置了参数 shuffle:true。由于每个节点都具有其唯一的随机化种子，因此会抽取唯一的图像子集。

节点通信

Intel Caffe 采用 Intel Machine Learning Scaling Library（MLSL），该库提供数据并行和模型并行的通信原语、SGD 的通信模式以及分布式的权重更新。它针对 Intel Xeon 处理器进行了优化，并支持 Intel Omni-Path 架构、InfiniBand 和以太网。

5.1.3 TensorFlow

TensorFlow 是另一个深度学习框架。它是开源软件库，可用于高性能数字计算。其灵活的架构可帮助跨多种平台（Cpu、Gpu、TPUs）以及从台式机到服务器集群，再到移动和边缘设备轻松进行计算部署。该框架最初由 Google AI 组织中 Google Brain 团队的研究人员和工程师开发，可以为机器学习和深度学习提供强有力的支持，灵活的数字计算核心被用于多种其他科学领域。

TensorFlow 的部分主要功能包括：

- 对深度神经网络等机器学习方法的编程支持。
- 对涉及多维数组（称为“张量”）的数学表达式进行高效计算。
- 独立于硬件的编程，且同一代码可用于 CPU 或 GPU 计算。
- 跨机器和海量数据集的高计算可扩展性。

5.2 软件注意事项

5.2.1 操作系统

此参考架构基于 Red Hat Enterprise Linux（RHEL）7.4 部署。开始安装前，用户需要确保部署的所有节点上已安装最新的官方补丁，并已执行所有必要的升级。

Intel 和 Red Hat 的合作带来了具有以下特点的平台：

- 加密/解密硬件卸载、操作系统级别的安全性以及 RHEL 和 Security Enhanced Linux（SELinux）中的身份管理带来的多层安全性
- 适用于集群环境和 Intel 组件的驱动程序、库和扩展带来的性能优化

5.2.2 库

- Intel Math Kernel Library（Intel MKL）是一个优化数学函数库，其中仅有部分函数适用于深度学习。大部分深度学习模型建立在有限的基础模块上，这些模块称作原语，它们对张量进行操作。这些基础模块或底层深度学习函数在 Intel MKL 库中针对 Intel Xeon 产品系列进行了优化。
- Intel MKL-DNN 库包含一组专用于深度学习的 MKL 函数。它以 Apache 2 许可证的形式开源发布，包含创建复杂模型所需的所有关键基础模块。Intel MKL-DNN 允许业界和院校深度学习开发者分发该库并向该库贡献新函数或改进的函数。
- NVIDIA CUDA 深度神经网络库（cuDNN）是一个适用于深度神经网络的基元 GPU 加速库。cuDNN 可为标准例程（例如前后卷积、池化、标准化和激活层）提供高度精调的实施。cuDNN 是 NVIDIA 深度学习 SDK 的一部分。

5.3 用于训练数据的大数据存储

训练数据的存储可能位于不同的存储环境。此类数据存储库之一就是使用 Cloudera Enterprise 软件在 Lenovo 硬件上实施的 Hadoop 集群。此外，Lenovo 硬件还提供其他 Hadoop 集群解决方案（请参阅本文档末尾的“资源”部分）。

5.3.1 Cloudera Enterprise

Cloudera Enterprise 提供多种功能和能力，可支持不同行业和企业中的任务关键型和实时大数据分析，例如金融服务、零售、媒体、医疗保健、制造、电信等行业以及政府组织、顶尖的财富 100 和 Web 2.0 企业。适用于大数据的 Cloudera 平台可用于多种用例，既包括使用 MapReduce 或 Spark 处理数据源（例如点击流）的批处理应用程序，也包括使用传感器数据的实时应用程序。

用户可在端口 22 上使用 安全 Shell（SSH）从防火墙外部登录到 Cloudera 客户端，以从企业网络访问 Cloudera 解决方案。Cloudera 提供若干接口供管理员和用户根据其角色和访问级别来执行管理和数据功能。Hadoop 应用程序编程接口（API）可用于访问数据。Cloudera API 可用于集群管理和监控。Cloudera 数据服务、管理服务和其他服务在集群中的节点上运行。存储是集群中各数据节点的一个组件。根据客户的需要，可通过 Hadoop API 或网络文件系统（NFS）将数据合并到 Cloudera Enterprise 存储。

图 3 显示了满足客户功能需求的 Cloudera Enterprise 关键能力。

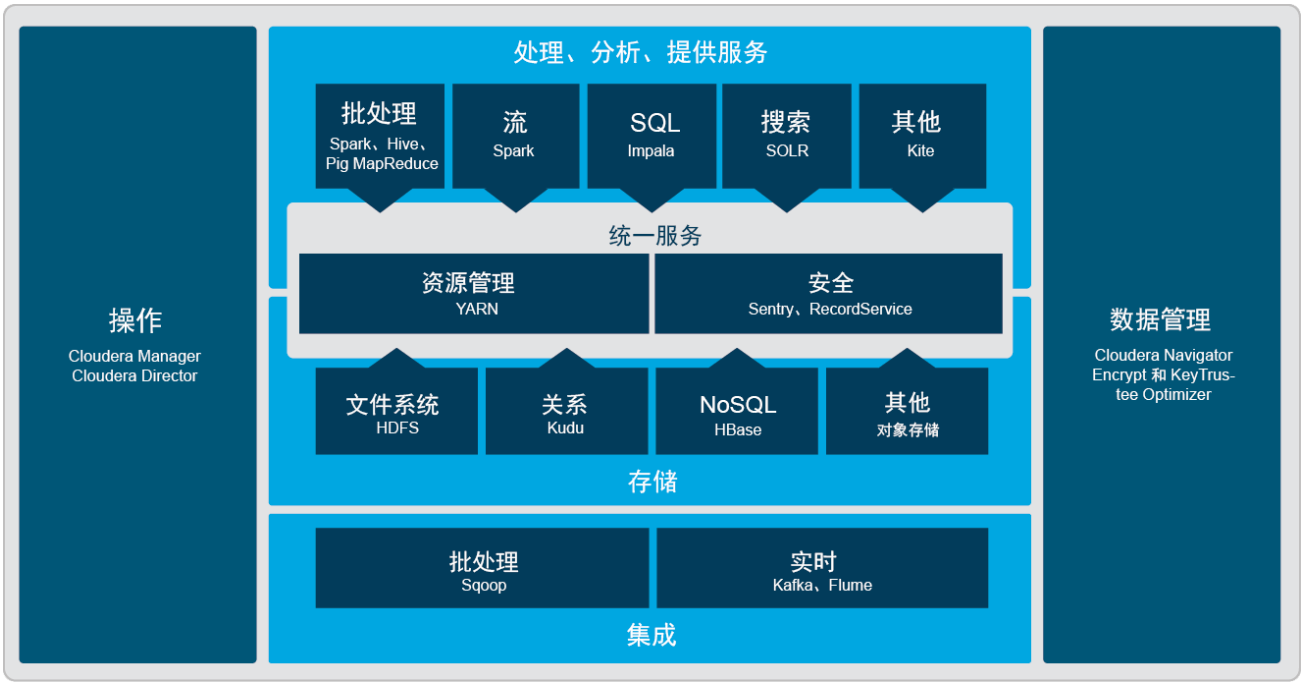

图 3. Cloudera Enterprise 的关键能力

Cloudera Enterprise 解决方案包含以下与作为模型训练数据存储库使用相关的组件：

- 核心组件：HDFS、MapReduce 和 YARN 是 Apache Hadoop 的核心组件，Cloudera Enterprise 版包含并支持这些组件。HDFS 是一个分布式文件系统，用于将服务器集群变成一个海量可扩展存储池。HDFS 可接受任何格式的数据（不论其使用何种架构），并可通过加入具有存储容量的集群节点实现线性扩展。MapReduce 是一个数据处理软件框架，用于满足 HDFS 和 Hadoop 的巨大规模。YARN 旨在为 Hadoop 提供开源资源管理，以支持 Interactive SQL、高级建模以及实时流式处理等多样化的工作负载。
- 搜索引擎：Cloudera Search

  Cloudera Search 是与 Cloudera Enterprise 集成的 Apache Solr，包含了 Apache Lucene、Apache SolrCloud、Apache Flume、Apache Tika 和 Hadoop。Cloudera Search 还包含了一些十分有用的其他集成，让搜索更易于扩展和使用，并经过了优化以实现几乎实时且面向批处理的索引。这些集成中包括 Cloudera Morphlines，其是一个可定制的转换链，可简化将任何类型的数据加载到 Cloudera Search 的过程。
- NoSQL - HBase

  一个可扩展、分布式的列式数据存储。HBase 提供对在 HDFS 上托管的特大数据集进行实时读/写随机存取的功能。

- 流式处理：Apache Spark

  Apache Spark 是一个开源并行数据处理框架，它是对 Hadoop 的补充，让开发对全部数据进行批处理、流式传输及交互式分析的快速、统一的大数据应用程序更为容易。Cloudera 通过 Cloudera Enterprise 对 Spark 提供商业支持。Spark 的速度是 MapReduce 的 10 - 100 倍，让您可更快获得分析结果、纳入更多数据，从而做出更明智的业务决策并带来更好的用户效果。

有关适用于 Cloudera Enterprise 的 Lenovo 大数据验证设计的详细信息，请访问：
https://lenovopress.com/lp0776-lenovo-big-data-validated-design-for-cloudera-enterprise-and-vmware-thinksystem

5.4 使用已训模型的推理

推理是将经过训练的模型应用于新数据的过程。如果已使用较大数据集对模型进行了良好的训练，则模型对新数据得到的输出准确度应与对训练测试数据得到的准确度相似。目前有多种推理框架，但此处使用的是 RA TensorFlow Serving 框架。

5.4.1 TensorFlow Serving

TensorFlow Serving 是一个用于机器学习模型的灵活而高性能的服务系统，专为生产环境而设计。TensorFlow Serving 可轻松部署新算法和实验，同时保持相同的服务器架构和 API。TensorFlow Serving 提供与 TensorFlow 模型的即时集成，但也可轻松扩展来服务其他类型的模型。

TensorFlow Serving 具有几个重要概念：

- **Servable** 是 TensorFlow Serving 中的核心抽象。Servable 是客户端用于执行计算（例如一次查找或推理）的底层对象。TensorFlow Serving 以一个或多个 Servable 来代表一个**模型**。机器学习模型可能包含一个或多个算法（包括习得权重）以及查找表或嵌入表。
- **加载器（Loader）** 管理 Servable 的生命周期。加载器 API 让通用基础结构独立于涉及的特定学习算法、数据或产品用例。具体而言，加载器对加载和卸载 Servable 的 API 进行标准化。
- **源（Source）**是查找并提供 Servable 的插件模块。每个源提供零个或多个 Servable 流。对于每个 Servable 流，源为各版本提供一个加载器实例让其可供加载。
- **管理器（Manager）**处理 Servable 的整个生命周期，包括：
  - 加载 Servable
  - 服务 Servable
  - 卸载 Servable
- **TensorFlow Serving 核心**（通过标准 TensorFlow Serving API）管理 Servable 的以下方面：
  - 生命周期
  - 指标

  TensorFlow Serving 核心将 Servable 和加载器视作不透明对象。

图 4 显示了 TensorFlow Serving 的高层框架

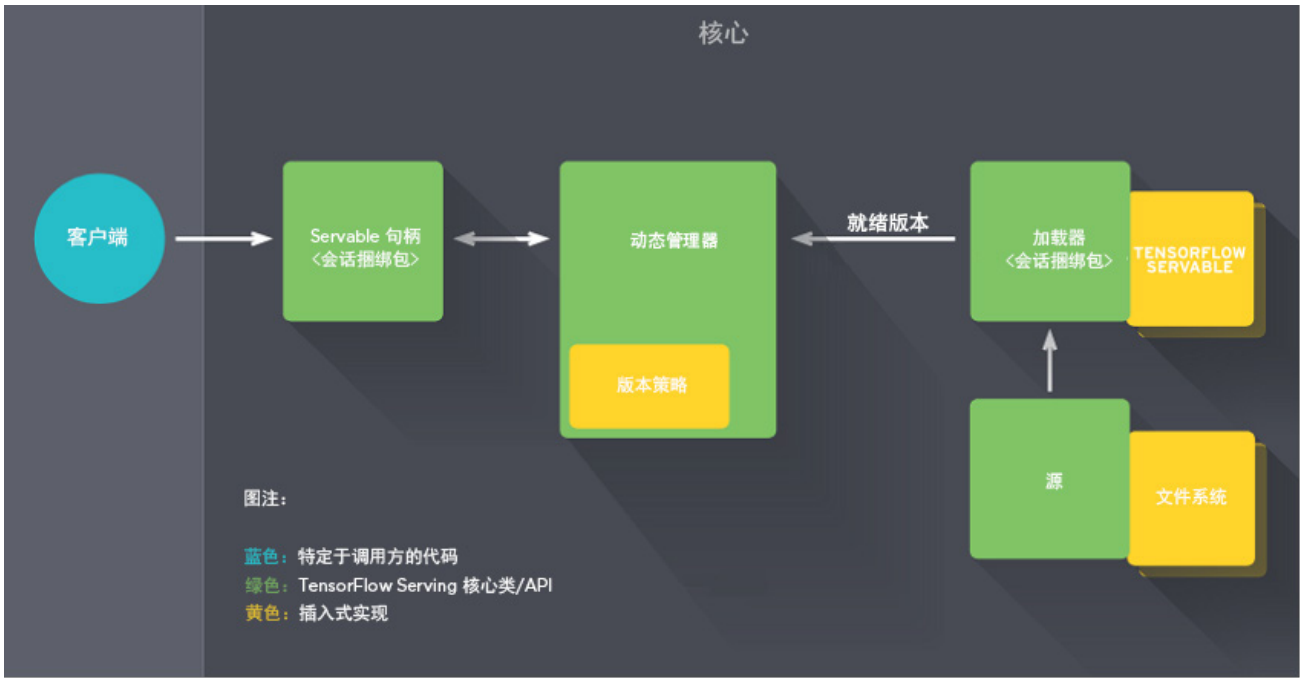

图 4. TensorFlow Serving 框架

概括地说：

1. 源创建 Servable 不同版本的加载器。
2. 加载器作为就绪版本（Aspired Version）发送给管理器，后者会将其加载并提供给客户端请求。

有关详细信息，请参阅《TensorFlow Serving 概述》：
https://www.tensorflow.org/serving/overview

5.5 联想智能超算平台（LiCO）

联想智能超算平台（LiCO）是一个软件解决方案，可简化高性能计算（HPC）以及人工智能（AI）开发的部署、管理和分布式横向扩展集群的使用。

通过提供直观的界面（图 5），LiCO 让您的用户只需动动手指就能管理所有集群资源。不论是学术研究还是企业数据中心，LiCO 都可以让您的研究人员、工程师、数据科学家在更短的时间内获得更大价值，并减少集群管理员和运行支持人员的工作量。

通过降低使用复杂度，同时不限制您的集群用户可能惯用的任何可靠业务流程工具，LiCO 可加快 AI 和 HPC 的开发速度。

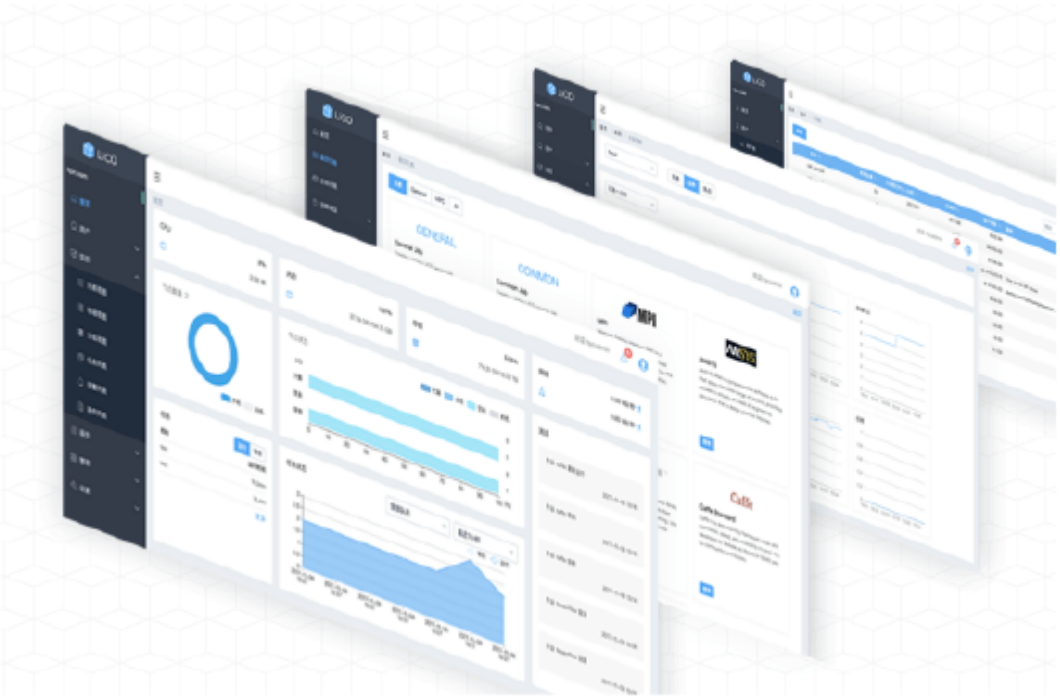

图 5. 联想智能超算平台（LiCO）

LiCO 提供多种功能，包括：

- 基于角色的访问
- 轻松管理数据集和 DNN 模型
- 轻松安排训练作业
- 检查训练作业的结果

有关详细信息，请参阅 Lenovo LiCO 数据表：

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

6 运行模式

本部分介绍 AI 参考架构的硬件基础结构方面。

为介绍适用于不同规模客户环境的运行模式，此参考架构描述了几个支持不同 DL 训练要求的节点配置。

6.1 硬件描述

6.1.1 Lenovo ThinkSystem SR650 服务器

Lenovo ThinkSystem SR650 一款 2 插槽 2U 机架服务器，对于需要行业领先的可靠性、管理性、安全性，以及最大程度满足未来增长所需性能和灵活性的大中小型企业来说，这是一款理想之选。SR650 服务器具有丰富内部数据存储、大容量内存以及高性能 Intel 处理器，非常适合 ML 和 DL 应用程序。此外，它也适用于处理常规工作负载，如数据库、虚拟化和云计算、虚拟桌面基础结构（VDI）、企业应用程序、协作/电子邮件以及业务分析。

SR650 服务器支持：

- 最多两个 Intel Xeon 可扩展处理器
- 最高 3.0 TB 2666 MHz TruDDR4 内存（需要特定 CPU 部件号）
- 最多 24 个 2.5 英寸或 14 个 3.5 英寸驱动器插槽，可使用 NVMe PCIe 固态硬盘、SAS/SATA 固态硬盘和 SAS/SATA 硬盘等丰富选择
- 灵活的 I/O 网络扩展选择，包括 LOM 插槽、专用存储控制器插槽和最多 6 个 PCIe 插槽

图 6. Lenovo ThinkSystem SR650

通过与 Intel Xeon 可扩展处理器（铜牌、银牌、金牌和铂金）配合使用，Lenovo SR650 服务器可提供更高密度的工作负载和性能，从而降低总拥有成本（TCO）。其按需付费的灵活设计以及强大的扩展能力可增强任何类型工作负载的可靠性并最大程度地降低停机时间。

SR650 服务器通过卓越的工作负载优化存储配置选择，在 2U 外形规格下提供高密度的内部存储。此外，通过将存储和服务器整合到一个系统，在最严苛的用例中也能提供轻松管理功能，节省占地空间与功耗。

基于以下原因，此参考架构推荐使用 ThinkSystem SR650：

- **性能**：此硬件支持最新的 Intel Xeon 可扩展处理器和 TruDDR4 内存。
- **灵活性**：服务器硬件使用嵌入式存储，因此可轻松扩展（通过增加节点）。
- **PCIe 插槽**：如未使用背面磁盘，最多可使用 7 个 PCIe 插槽；如使用了背面硬盘套件，最多可使用 3 个 PCIe 插槽。它们可用于网络适配器冗余，增加网络吞吐量。

- **更高能效**：钛金级和白金级冗余电源模块在 50% 负载时可提供 96%（钛金级）或 94%（白金级）的效率。
- **可靠性**：卓越的可靠性、可用性和可维护性（RAS）可改善业务环境并有助于节约运营成本

有关详细信息，请参阅《Lenovo ThinkSystem SR650 产品指南》：
https://lenovopress.com/lp0644-lenovo-thinksystem-sr650-server

6.1.2 Lenovo ThinkSystem SD530 服务器

Lenovo ThinkSystem SD530 是一款 0.5U 机架外形规格的超高密度经济型双插槽服务器。四个 SD530 服务器安装在 ThinkSystem D2 机柜或 ThinkSystem 模块化机柜中，是处理企业和云工作负载的理想高密度 2U 四节点（2U4N）平台。

不论是大型企业还是服务提供商，2U4N 系统在各种数据中心中广受欢迎，其占地面积小和固有高密度的特点使其非常适合用于以低成本构建基于解决方案的设备。Lenovo ThinkSystem SD530 和 D2 机柜的组合专为提供此类解决方案而设计。

图 7. 四个 ThinkSystem SD530 服务器安装在一个 D2 机柜中

有关详细信息，请参阅《Lenovo ThinkSystem SD530 产品指南》：
https://lenovopress.com/lp0635-thinksystem-sd530-server

6.1.3 Lenovo ThinkSystem SR670

ThinkSystem SR670 是专门用于支持 GPU 的工作负载的 2 插槽 2U 节点。SR670 支持最新版本的 NVIDIA GPU 和 Intel Xeon 可扩展处理器，可为人工智能和高性能计算工作负载所需的高端计算提供最优性能。SR670 可最多为每个节点提供 4 个 GPU，还可为 LICO 平台提供最佳硬件。

图 8. Lenovo ThinkSystem SR670

有关详细信息，请参阅《Lenovo ThinkSystem SR670 产品指南》：
https://lenovopress.com/lp0923-thinksystem-sr670-server

6.1.4 Lenovo RackSwitch G8052

Lenovo System Networking RackSwitch G8052（如图 9 所示）是专为数据中心设计的以太网交换机，可提供简单网络解决方案。Lenovo RackSwitch G8052 在 1U 空间中最多可提供 48 个 1 GbE 端口和 4 个 10 GbE 端口。通过冗余电源模块、风扇和各种高可用性功能，G8052 交换机可始终满足业务关键型流量的要求。

图 9. Lenovo RackSwitch G8052

Lenovo RackSwitch G8052 具有以下特点：

- 四十八个 1 GbE RJ45 端口
- 四个标准 10 GbE SFP+ 端口
- 130 瓦的低额定功率和变速风扇，可降低功耗

有关详细信息，请参阅《Lenovo RackSwitch G8052 产品指南》：
https://lenovopress.com/tips1270-lenovo-rackswitch-g8052

6.1.5 Intel Omni-Path (OPA) 100 系列 48 端口非管理型边缘交换机

Intel Omni-Path 边缘交换机包含两种对所有端口支持 100 Gb/s 的型号，一种是用于小集群的入门级 24 端口交换机，另一种是 48 端口交换机。较大的交换机除在 1U 空间中支持 48 端口构造外，还可与其他边缘交换机及控制器结合使用以构建更大的多层构造。这些 Intel Omni-Path 边缘交换机是 Intel Omni-Path Fabric 100 系列交换机、主机适配器和软件的成员，它们可提供一套卓越的高速网络特性和功能。

有关 OPA 100 系列交换机的详细信息，请访问以下链接：

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

图 10. Intel Omni-Path（OPA）100 系列 48 端口非管理型边缘交换机

6.1.6 Mellanox SB7800 系列 36 端口 InfiniBand 交换机

SB7800 系列配备 Mellanox Switch-IB™ 2 InfiniBand 交换机设备，可为每个端口提供最高 36 100Gb/s 的完整双向带宽。SB7800 智能网络交换机旨在通过共同设计的可扩展分层聚合协议（SHARP）技术启用网络内部计算。共同设计架构允许使用所有有效的数据中心设备，通过嵌入式硬件加快通信框架，进而改善重大应用程序性能。

有关 InfiniBand SB7800 系列交换机的详细信息，请访问以下链接：

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

图 11. Mellanox SB7890 InfiniBand 交换机

6.1.7 Lenovo ThinkSystem NE10032 机架交换机

Lenovo ThinkSystem NE10032 RackSwitch 采用 100 Gb QSFP28 和 40 Gb QSFP + 以太网技术，专为数据中心而设计。它是如今的大数据、云和企业工作负载解决方案的理想之选。该产品是一款功能完整的企业级第 2 层和第 3 层交换机，可提供无延迟数据的线速、高带宽切换、过滤和流量排队。大型数据中心级别缓冲区可帮助维持流量移动，同时热插拔冗余电源模块和风扇（以及众多高可用性功能）可为业务敏感流量提供高可用性。

图 12. Lenovo RackSwitch NE10032

有关详细信息，请参阅《Lenovo RackSwitch NE10032 产品指南》：
https://lenovopress.com/lp0609-lenovo-thinksystem-ne10032-rackswitch

6.1.8 Lenovo ThinkSystem NE2572 机架交换机

Lenovo ThinkSystem NE2572 RackSwitch 采用 25 Gb SFP28 和 10 Gb SFP+ 以太网技术，专为数据中心而设计。NE2572 是一款极具性价比的交换机，适用于高端 25GbE 混合云数据中心应用程序。

图 13. Lenovo RackSwitch NE2572

有关详细信息，请参阅《Lenovo RackSwitch NE2572 产品指南》：
https://lenovopress.com/datasheet/ds0010-lenovo-thinksystem-ne2572-rackswitch

6.2 集群节点

此 AI 参考架构在组成一个集群的一组节点上实施。

训练节点运行 Caffe 或 TensorFlow 服务来进行模型训练。头节点主要用于存储，包括存储训练数据集和运行前/后的训练模型。头节点连接到与训练节点相同的高速数据网络以获得最佳性能。头节点可使用带本地连接存储的 Lenovo ThinkSystem SR650 服务器。训练节点可使用 Lenovo ThinkSystem SR650 服务器、SD530 服务器（用于增加密度）或经过 AI 优化的 ThinkSystem SR670 服务器。

推理节点运行 TensorFlow Serving 等推理框架。推理节点可为一个或多个 Lenovo ThinkSystem SR650 或 SD530 服务器。对于特定应用程序，推理节点可使用 Lenovo ThinkStation P 系列 Tiny 或 ThinkPad P 系列笔记本电脑。

6.2.1 训练节点

表 1、表 2 和 表 3 列出了在此参考架构中演示的推荐用于训练节点的系统组件。

表 1. SR650 训练节点配置

组件	工作节点配置
服务器	ThinkSystem SR650
处理器	2 个 Intel Xeon 处理器：6148 金牌，20 核 2.4Ghz
内存 – 基础	192 GB：12 个 16GB 2666MHz RDIMM 384 GB：12 个 32GB 2666MHz RDIMM
磁盘（操作系统）	具有 RAID1 功能的双 M.2 128GB 固态硬盘
磁盘（数据）	2 个 960GB 2.5 英寸 SATA 固态硬盘
存储控制器	操作系统：M.2 RAID1 镜像支持套件 数据：ThinkSystem 430-16i 12Gb 控制器
硬件存储保护	操作系统：RAID1 模型存储：无（JBOD）。
硬件管理网络适配器	集成式 1G BaseT XCC 管理控制器
数据网络适配器	Intel OPA 适配器 100Gb 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器**
加速器（可选）	2 个 NVIDIA Tesla V100 32GB PCIe Gpu 或 2 个 NVIDIA Tesla P100 16GB PCIe Gpu

**注：如果 SR650 装有两个 GPU 和 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器，网络带宽将限于 56Gb/s。

表 2. SD530 训练节点配置

组件	工作节点配置
服务器	配备 D2 机柜的 ThinkSystem SD530
处理器	2 个 Intel Xeon 处理器：6148 金牌，20 核 2.4Ghz
内存 – 基础	192 GB：12 个 16GB 2666MHz RDIMM 384 GB：12 个 32GB 2666MHz RDIMM
磁盘（操作系统）	具有 RAID1 功能的双 M.2 128GB 固态硬盘
磁盘（数据）	2 个 960GB 2.5 英寸 SATA 固态硬盘
存储控制器	操作系统：M.2 RAID1 镜像支持套件 数据：ThinkSystem 板载 6Gb SATA 控制器
硬件存储保护	操作系统：RAID1 模型存储：无（JBOD）。
硬件管理网络适配器	集成式 1G BaseT XCC 管理控制器
数据网络适配器	Intel OPA 适配器 100Gb 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器
加速器（可选）	2 个 NVIDIA Tesla V100 32GB PCIe Gpu 或 2 个 NVIDIA Tesla P100 16GB PCIe Gpu

表 3. SR670 训练节点配置

组件	工作节点配置
服务器	ThinkSystem SR670
处理器	2 个 Intel Xeon 处理器：6148 金牌，20 核 2.4Ghz
内存 – 基础	192 GB：12 个 16GB 2666MHz RDIMM 384 GB：12 个 32GB 2666MHz RDIMM
磁盘（操作系统）	具有 RAID1 功能的双 M.2 128GB 固态硬盘
磁盘（数据）	2 个 960GB 2.5 英寸 SATA 固态硬盘
存储控制器	操作系统：M.2 RAID1 镜像支持套件 数据：ThinkSystem 板载 6Gb SATA 控制器
硬件存储保护	操作系统：RAID1 模型存储：无（JBOD）。
硬件管理网络适配器	集成式 1G BaseT XCC 管理控制器
数据网络适配器	Intel OPA 适配器 100Gb 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器
加速器	4 个 NVIDIA Tesla V100 32GB PCIe GPU

以上表格中推荐使用的 Intel Xeon 可扩展处理器可为 DL 训练节点实现性能与成本之间的平衡。对于计算密集型工作负载还可选择更高核数和频率的处理器。操作系统在具有 RAID1 镜像功能的双 M.2 固态硬盘内存条上加载。数据磁盘是为高速训练数据吞吐量配置的 JBOD。

6.2.2 头节点

头节点是 DL 训练集群的存储中心，支持深度学习集群所需的若干其他重要功能。表 3 列出了推荐的头节点组件，可根据客户需求进行定制。

表 4. 头节点配置

组件	主节点配置
服务器	ThinkSystem SR650
处理器	2 个 Intel Xeon 处理器：6130 金牌，20 核 2.1Ghz
内存 – 基础	192 GB：12 个 16GB 2666MHz RDIMM 384 GB – 12 个 32 GB 2666MHz RDIMM
磁盘（操作系统/本地存储）	操作系统：具有 RAID1 功能的双 M.2 128GB 固态硬盘 数据：24 个 960GB 2.5 英寸 SATA 固态硬盘
存储控制器	操作系统：M.2 RAID1 镜像支持套件 数据：3 个 ThinkSystem RAID 930-8i 4GB 闪存 12Gb 控制器
硬件存储保护	操作系统：RAID1 数据：RAID 5
硬件管理控制器	集成式 1G BaseT XCC 管理控制器
数据网络适配器	Intel OPA 适配器 100Gb 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器

推荐使用表 4 中指定的 Intel Xeon 可扩展处理器和最低内存以确保头节点可提供足够的性能。操作系统在具有 RAID1 镜像功能的双 M.2 固态硬盘内存条上加载。数据磁盘采用 RAID 配置以提供可靠的数据存储。

6.2.3 大数据节点

Cloudera 推理架构在组成一个集群的一组节点上实施，该集群主要包含两种节点：工作节点和主节点。工作节点使用具有本地连接存储的 ThinkSystem SR650 服务器，或具有外部 SAS 存储的 SD530 密集型计算节点。主节点使用 ThinkSystem SR630 服务器。

工作节点运行存储及处理数据的数据（工作）服务。

主节点运行以下几种服务：

- 协调和管理集群的管理控制服务
- 文件和 Web 服务等杂项和可选服务

被称作边缘节点、系统管理节点或网关节点的节点安装在集群数据网络上，但不直接运行 Cloudera Enterprise 软件。其作用是将 Cloudera 集群连接到外部网络以进行远程管理访问、从外部源引入数据或运行访问 Cloudera Enterprise 集群的最终用户应用程序软件。

单个系统管理/网关节点在此参考架构中配置为最小单位节点，用于 Linux 操作系统的远程管理以及硬件维护。根据高速数据引入集群和边缘节点应用程序的具体要求，可增加此服务器的 CPU、内存、存储及网络能力。

有关 Cloudera 大数据集群的详细说明，请参阅《适用于 Cloudera Enterprise 的 Lenovo Press 参考架构》：

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

6.2.4 推理节点

推理节点在 Lenovo ThinkSystem SR650 和 SD530 上实施。表 5 和表 6 列出了在此参考架构中演示的推荐用于训练节点的系统组件。

表 5. SR650 推理节点配置

组件	工作节点配置
服务器	ThinkSystem SR650
处理器	2 个 Intel Xeon 处理器：6148 金牌，20 核 2.4Ghz
内存 – 基础	192 GB：12 个 16GB 2666MHz RDIMM
磁盘（操作系统）	具有 RAID1 功能的双 M.2 128GB 固态硬盘
存储控制器	操作系统：M.2 RAID1 镜像支持套件
硬件存储保护	操作系统：RAID1
硬件管理网络适配器	集成式 1G BaseT XCC 管理控制器
数据网络适配器	Intel OPA 适配器 100Gb 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器**
加速器（可选）	2 个 NVIDIA Tesla V100 32GB PCIe Gpu 或 2 个 NVIDIA Tesla P100 16GB PCIe Gpu

**注：如果 SR650 装有两个 GPU 和 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器，网络带宽将限于 56Gb/s。

表 6. SD530 推理节点配置

组件	工作节点配置
服务器	配备 D2 机柜的 ThinkSystem SD530
处理器	2 个 Intel Xeon 处理器：6148 金牌，20 核 2.4Ghz
内存 – 基础	192 GB：12 个 16GB 2666MHz RDIMM
磁盘（操作系统）	具有 RAID1 功能的双 M.2 128GB 固态硬盘
存储控制器	操作系统：M.2 RAID1 镜像支持套件
硬件存储保护	操作系统：RAID1
硬件管理网络适配器	集成式 1G BaseT XCC 管理控制器
数据网络适配器	Intel OPA 适配器 100Gb 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器
加速器（可选）	2 个 NVIDIA Tesla V100 32GB PCIe Gpu 或 2 个 NVIDIA Tesla P100 16GB PCIe Gpu

6.3 Lenovo XClarity

集群的*系统管理* 包括硬件管理、操作系统管理和 Caffe 应用程序管理。

硬件管理 使用的是 Lenovo XClarity Administrator，它是一种集中式资源管理解决方案，可减少复杂度、加快响应速度并提高 Lenovo 服务器系统和解决方案的可用性。XClarity 用于将操作系统安装到新的工作节点、更新各集群节点的固件、记录硬件警报以及在需要进行修复操作时发出报告。

图 14 显示了 Lenovo XClarity Administrator 的界面，服务器、存储、交换机及其他机架组件在该界面中进行管理，且状态在仪表板上显示。Lenovo XClarity Administrator 是一个可快速导入到服务器虚拟化环境的虚拟设备。

图 14. XClarity Administrator 界面

此外，极端集群管理工具包（xCAT）可提供一个可扩展的分布式计算管理和配置工具，该工具提供一个用于硬件控制、发现和操作系统部署的统一界面。它可用于简化或自动化集群节点的管理。有关 xCAT 的详细信息，请参阅第 20 页上的“资源”一节。xCAT 可与 Confluent 配合使用以支持其他管理功能以及连接到集群的图形化界面。有关 xCAT 和 Confluent 的详细信息，请参阅本文档末尾的“资源”一节。

6.4 网络

此参考架构包含两个网络：高速数据网络和管理网络。

6.4.1 数据网络

数据网络创建多节点间的专用集群，用于训练节点间的高速数据传输以及将数据导入集群。推荐使用的 OPA 交换机是可提供 32 个 100Gb OPA 端口的 Intel Omni-Path（OPA）100 系列 48 端口非管理型边缘交换机。建议的 IB 交换机是 Mellanox SB7890 IB 2 EDR 交换机，可提供 36 个 100Gb IB 端口。建议的 100GbE 交换机是 Lenovo NE10032，可提供 32 个 100GbE 端口 。如果使用 25GbE 数据网络，则建议使用可提供 48 个 25GbE 端口的 Lenovo NE2572 交换机。

图 15 显示了用于大数据、训练和推理的网络设计。建议大数据集群的节点连接使用 25 GbE，以将节点连接到配置为 25 GbE 的 NE10032 交换机或 25 GbE 原生 NE2572 交换机。训练集群的节点连接使用 100Gb 时性能最佳（如图 17 所示）。推理节点可使用 25GbE 或 100Gb（如用于高吞吐量应用程序）进行连接。

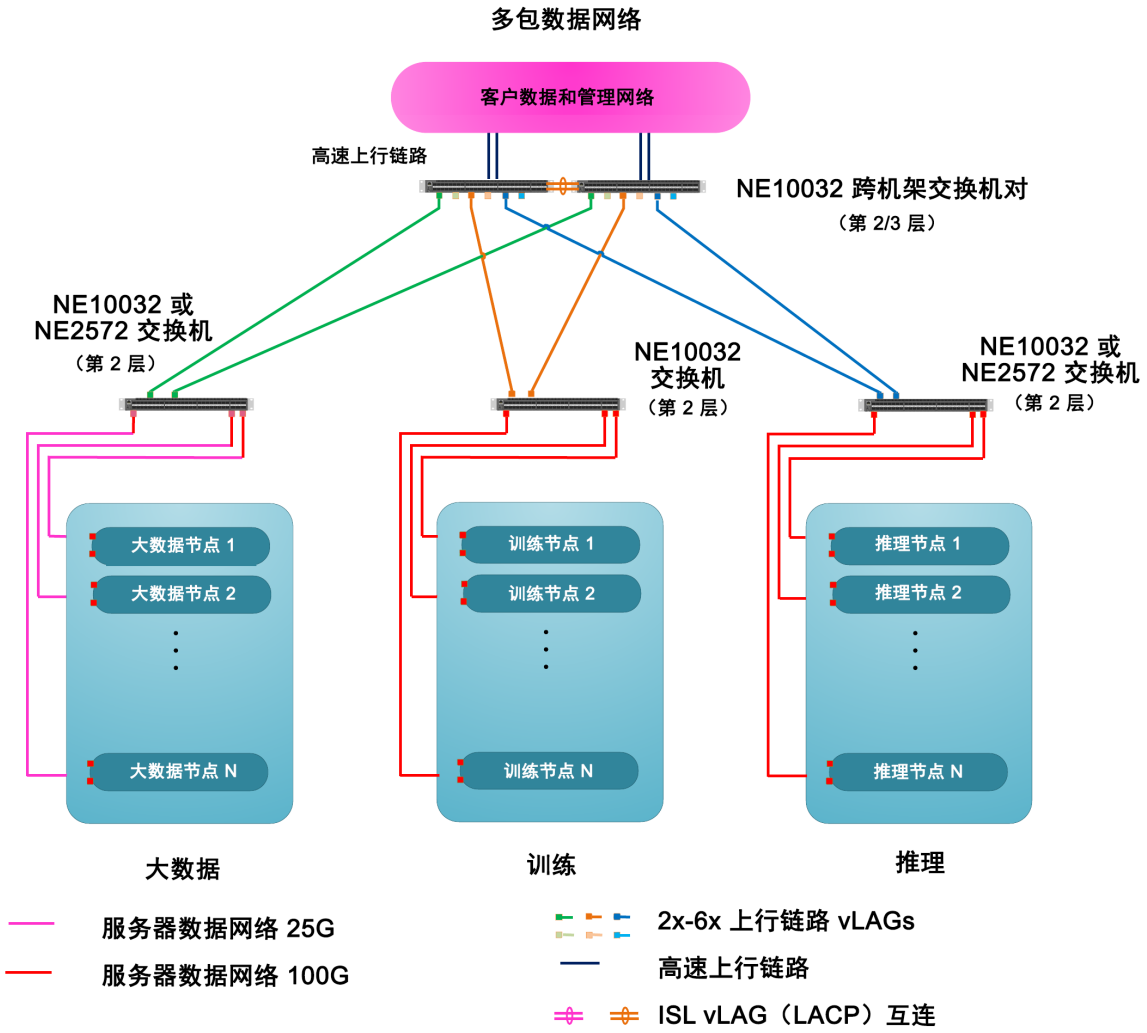

图 15. 数据网络图

6.4.2 硬件管理网络

硬件管理网络是用于带外硬件管理的 1GbE 网络。推荐使用的 1GbE 交换机是具有 10Gb SFP+ 上行链路端口的 Lenovo RackSwitch G8052。通过 ThinkSystem SR650 和 SD530 服务器中的 XClarity Controller 管理模块（XCC），带外网络支持对集群节点进行硬件级别的管理，例如部署节点、配置 UEFI 固件、查看硬件故障状态以及远程控制节点电源。

Intel Caffe 的功能不依赖于 XCC 管理控制器。集群数据和操作系统管理网络可通过 XCC 硬件管理网络进行共享，也可通过各自交换机上的 VLAN 进行分隔。然后，集群和硬件管理网络通常会直接连接到客户的现有管理网络以便对集群进行远程维护。

6.5 预定义的训练集群配置

此参考架构所提供的预定义配置旨在简化客户的初始布局过程，并提供适用于不同大小训练工作负载的示例起点。这些配置包括训练节点、头节点、网络交换机、连线和机架硬件。对于少于四个节点的 DL 训练集群，推荐使用 SR650。为降低小集群设计的成本，SR650 可作多用途使用，既可用作训练节点，也可用作头节点。对于较大的群集，SD530 可提供密集型解决方案，相比 SR650 可节省 75% 的机架空间。下面的表 7 显示了各预定义集群大小所需的硬件。

表 7. 预定义的 DL 集群配置

	单个节点	单个机箱	两个机箱	四个机箱	16 个机箱
训练节点模型	SR650	SD530	SD530	SD530	SD530
训练节点	1	4	8	16	64
D2 机柜	0	1	2	4	16
头节点	0	1	1	1	2
头节点存储	24 个 960GB	24 个 960GB	24 个 960GB	24 个 960GB	48 个 960GB

将 GPU 添加到 SR650、 SD530 或 SR670 可显著缩短特定工作负载类型所需的训练时间。对于少于两个 GPU 的 DL 训练集群，推荐使用 SR650。对于较大的集群，SD530 提供密集型解决方案，比 SR650 节省 50% 的机架空间。SR670 专为配备 GPU 的训练而设计，最多可为每个 2U 系统装载四个 GPU。下面的表 8 显示了使用 GPU 加速器时各预定义集群大小所需的硬件。

表 8. 预定义的 DL 集群配置与 GPU 加速器

	单个节点	单个机箱		两个机箱		四个机箱		16 个机箱	
训练节点模型	SR650	SR670	SD530	SR670	SD530	SR670	SD530	SR670	SD530
训练节点	1	1	2	2	4	4	8	16	32
D2 机柜	0	0	1	0	2	0	4	0	16
头节点	0	1		1		1		2	
头节点存储	24 个 960GB	24 个 960GB		24 个 960GB		24 个 960GB		48 个 960GB	
加速器（GPU）	2	4		8		16		64	

7 部署注意事项

本部分介绍部署 AI 解决方案的其他注意事项。

预定义配置是一组基准配置，可根据其内容实施，也可根据具体客户要求（如降低成本、改善性能以及提高可靠性等）进行修改。

修改预定义配置时，必须了解关于未来将如何使用集群的几个关键方面。对于训练而言，必须了解用于训练的数据当前及未来的类型和大小，以及训练速度是否是关键要素。此外，还需考虑集群其他用例的计算需求，以及是否需要多个集群并行运行。

设计深度学习集群基础结构时，建议使用代表性数据和工作负载执行必要测试和概念验证，以确保提出的设计可达到必要的成功条件。以下部分提供关于定制预定义配置的信息。考虑对预定义配置进行定制时，请咨询具有设计 DL 集群基础结构经验的系统架构师。

7.1 提高集群性能

有几种方法可提高集群的训练性能：增加节点数，以及选择更高性能的硬件。增加训练节点（横向扩展设计）是马上或在未来提高集群性能的最简单方法，在数据训练要求发生改变的情况下尤其如此。选择更高性能的 CPU 或添加 GPU 需要在集群设计阶段对训练要求有更好地理解。一旦为一个集群选择了 CPU 或 GPU 型号后，建议在所有训练节点中都使用该型号。此外，提升性能通常意味着成本增加，因此进行高性能设计时必须权衡成本与效益。

7.2 低成本设计

要降低 DL 参考架构解决方案的成本，可在若干关键之处进行修改。考虑低成本方案时，必须确保客户了解较低成本的设计可能会导致较低性能，这点非常重要。此 DL 参考架构的低成本版本可通过以下方式实现：使用较低成本的节点处理器、减少每个数据节点的内存容量，以及使用较低成本的存储驱动器。节点处理器可替换为 Intel Xeon SP 系列中的其他处理器。选择其他处理器可能要求较低频率的内存，这也可降低解决方案中每个节点的成本。

为每个数据节点使用较小的内存容量可降低设计成本。但是，由使用较低内存容量的数据节点组成的集群，其性能可能较低。因此，在概念验证评估阶段应使用真实用户数据进行测试以了解其带来的性能影响。

可将数据节点上的存储驱动器替换为较低容量的驱动器或标准硬盘。但应如前文所述，在概念验证测试阶段评估使用这些低成本的存储选件对数据节点性能造成影响。

7.3 扩展注意事项

达到现有基础结构容量极限时，可通过增加更多训练节点以及头节点（如必要）来扩展集群。达到现有机架容量极限时，可向集群增加新机架。部分训练工作负载可能不会线性扩展。

设计新深度学习解决方案参考架构实施时，未来的横向扩展是在初始设计阶段便需考虑的重要注意事项。必须考虑网络和管理这两个互相联系的方面。这两方面对于集群的运行均至关重要，且随着集群基础结构的增长而变得更为复杂。

上面“网络”部分所述的网络模型专为提供集群内可靠的机架网络互连而设计。随着机架数量的增加，预定义的网络拓扑可保持平衡与对称。如果计划将集群扩展到一个以上的机架，那么集群的最初设计便应针对多个机架，即使最初的节点数量可能只需一个机架即可容纳。以多个机架为设计起点可确保实施适当的网络拓扑，防止未来需要进行重新配置或更改硬件。

此外，随着集群中节点数量的增加，许多管理集群的任务也会增加，例如更新节点固件或操作系统。将建立集群管理框架作为初始设计的一部分，以及提前考虑管理大型集群可能遇到的挑战最终将带来显著的回报。

主动规划未来的横向扩展以及将集群管理框架的开发纳入初始集群设计可为未来增长奠定良好的基础，这将最大限度地减少随着集群增长而需要重新配置硬件的可能性以及集群管理遇到的问题。

8 附录：物料清单

本部分介绍 AI 解决方案不同硬件配置的物料清单（BOM）。本部分分为训练节点、头节点和网络三个小节。

BOM 包括部件号、组件描述和数量。表 7 和表 8 列出了每个预定义配置中定义的各核心组件的数量。

本附录包含的 BOM 列表并非穷举，必须使用订购配置工具进行检验。价格、支持和维护选项不在本文档的讨论范围内。

本 BOM 信息适用于美国，其中的部件号和描述在其他国家/地区可能有所不同。可从 Lenovo 销售团队获取其他配置示例。组件可能发生改变，恕不另行通知。

8.1 训练节点

表 9 列出了 SR650 培训节点的 BOM，表 10 列出了配备 GPU 的 SR650 培训节点的 BOM。

表 9. 训练节点：SR650

代码	描述	数量
7X06CTO1WW	SR650 训练：ThinkSystem SR650 - 3 年保修	1
AUVX	ThinkSystem SR650 2.5 英寸机箱，含 8 个或 16 个插槽	1
AWDX	Intel Xeon 金牌 6148 20C 150 瓦 2.4GHz 处理器	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz（2Rx4 1.2V）RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5 英寸 SATA/SAS 8 插槽背板	2
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2（配有镜像支持套件）	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps 非热插拔固态硬盘	2
B0ZR	ThinkSystem 2.5 英寸 Intel S4600 960GB 主流 SATA 6Gb 热插拔固态硬盘	2
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH 转接卡 1 套件	1
AURC	ThinkSystem SR550/SR590/SR650（x16/x8）/（x16/x16）PCIe FH 转接卡 2 套件	1
AUKK	ThinkSystem 10Gb 4 端口 SFP+ LOM	1
AU0B 或 ASWQ	Intel OPA 100 系列单端口 PCIe 3.0 x16 HFA 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器	1
AXCA	ThinkSystem 免工具滑轨	1
AURD	ThinkSystem 2U 左侧 EIA 标准滑锁	1
AVWG	ThinkSystem 1600W（230V）白金级热插拔电源模块	2
AURQ	Lenovo ThinkSystem 2U 3FH 转接卡支架	1
AURP	Lenovo ThinkSystem 2U 2FH 转接卡支架	1
AUQB	Lenovo ThinkSystem 主流 MB - 2U	1
B173	XClarity™ Controller Standard 到 Enterprise 工厂升级	1
AUSF	Lenovo ThinkSystem 2U MS CPU 性能散热器	2

表 10. 培训节点：配备 GPU 的 SR650

代码	描述	数量
7X06CTO1WW	SR650 训练：ThinkSystem SR650 - 3 年保修	1
AUVX	ThinkSystem SR650 2.5 英寸机箱，含 8 个或 16 个插槽	1
AWDX	Intel Xeon 金牌 6148 20C 150 瓦 2.4GHz 处理器	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz（2Rx4 1.2V）RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5 英寸 SATA/SAS 8 插槽背板	2
AUNM	ThinkSystem 430-16i SAS/SATA 12Gb HBA	1
AUMV	ThinkSystem M.2（配有镜像支持套件）	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps 非热插拔固态硬盘	2
B0ZR	ThinkSystem 2.5 英寸 Intel S4600 960GB 主流 SATA 6Gb 热插拔固态硬盘	2
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH 转接卡 1 套件	1
AURC	ThinkSystem SR550/SR590/SR650（x16/x8）/（x16/x16）PCIe FH 转接卡 2 套件	1
AUKK	ThinkSystem 10Gb 4 端口 SFP+ LOM	1
AU0A 或 ASWQ	Intel OPA 100 系列单端口 PCIe 3.0 x8 HFA 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器	1
AXCA	ThinkSystem 免工具滑轨	1
AURD	ThinkSystem 2U 左侧 EIA 标准滑锁	1
AVWG	ThinkSystem 1600W（230V）白金级热插拔电源模块	2
AURQ	Lenovo ThinkSystem 2U 3FH 转接卡支架	1
AURP	Lenovo ThinkSystem 2U 2FH 转接卡支架	1
AUQB	Lenovo ThinkSystem 主流 MB - 2U	1
B173	XClarity™ Controller Standard 到 Enterprise 工厂升级	1
AULQ	ThinkSystem 1U CPU 性能散热器	2
AX4F	ThinkSystem SR650 GPU 导风管	1
B34S 或 B0M1	ThinkSystem NVIDIA Tesla V100 32GB PCIe 被动式 GPU 或 ThinkSystem NVIDIA Tesla P100 16GB PCIe 被动式 GPU	1/2

表 11 列出了 SD530 训练节点的 BOM，表 12 列出了配备 GPU 的 SD530 训练节点的 BOM。表 13 列出了 D2 机柜的 BOM。

表 11. 训练节点：SD530

代码	描述	数量
7X21CTO1WW	Stark：ThinkSystem SD530 - 3 年保修	1
AUXN	ThinkSystem SD530 计算节点	1
AWEW	Intel Xeon 金牌 6148 20C 150 瓦 2.4GHz 处理器	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz（2Rx4 1.2V）RDIMM	12
AUYJ	ThinkSystem SD530 2x2 SAS/SATA 背板	1
AUYL	ThinkSystem SD530 SW RAID 套件	1
B0ZR	ThinkSystem 2.5 英寸 Intel S4600 960GB 主流 SATA 6Gb 热插拔固态硬盘	2
AUMV	ThinkSystem M.2（配有镜像支持套件）	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps 非热插拔固态硬盘	2
AUYM	ThinkSystem SD530 正面 VGA/USB KVM 分支模块	1
AVUT	ThinkSystem XClarity Controller 标准到高级升级	1
AU0B 或 ASWQ	Intel OPA 100 系列单端口 PCIe 3.0 x16 HFA 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器	1

表 12. 训练节点：配备 GPU 的 SD530

代码	描述	数量
7X21CTO1WW	节点：ThinkSystem SD530 - 3 年保修	1
B0M3	ThinkSystem SD530 GPU 托盘计算节点	1
AWEW	Intel Xeon 金牌 6148 20C 150 瓦 2.4GHz 处理器	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz（2Rx4 1.2V）RDIMM	12
AUYJ	ThinkSystem SD530 2x2 SAS/SATA 背板	1
AUYL	ThinkSystem SD530 SW RAID 套件	1
B0ZR	ThinkSystem 2.5 英寸 Intel S4600 960GB 主流 SATA 6Gb 热插拔固态硬盘	2
AUMV	ThinkSystem M.2（配有镜像支持套件）	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps 非热插拔固态硬盘	2
AUYM	ThinkSystem SD530 正面 VGA/USB KVM 分支模块	1
AVUT	ThinkSystem XClarity Controller 标准到高级升级	1
AU0B 或 ASWQ	Intel OPA 100 系列单端口 PCIe 3.0 x16 HFA 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器	1
B0MU	ThinkSystem SD530 GPU 托盘	1
B13L	ThinkSystem SD530 GPU 转接卡模块	1
B34S 或 B0M1	ThinkSystem NVIDIA Tesla V100 32GB PCIe 被动式 GPU 或 ThinkSystem NVIDIA Tesla P100 16GB PCIe 被动式 GPU	1/2

表 13. ***训练节点：D2 机箱***

代码	描述	数量
7X20CTO1WW	D2 机箱：ThinkSystem D2 机柜 - 3 年保修	1
AUXM	ThinkSystem D2 机柜	1
AUY9	ThinkSystem D2 10Gb 8 端口 EIOM SFP+	1
AUY8	ThinkSystem D2 4 插槽 x16 平移台	1
AUYC	ThinkSystem D2 滑轨	1
AUZ2	ThinkSystem D2 2000W 白金级 PSU	2
A4VP	1.0 米，C13 转 C14 跳线，机架电源线	2

表 14 列出了配备 GPU 的 SR670 训练节点的 BOM。

表 14. ***训练节点：配备 GPU 的 SR670***

代码	描述	数量
7Y37CTO1WW	SR670 训练：ThinkSystem SR670 - 3 年保修	1
B3XX	L1 底座条幅 CTO1WW	1
AWDX?	Intel Xeon 金牌 6148 20C 150 瓦 2.4GHz 处理器	2
AWDW	Intel Xeon Gold 6142 16C 150 瓦 2.6GHz 处理器	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz（2Rx4 1.2V）RDIMM	12
B42K	1U 2.5 英寸 SATA/SAS 8 插槽背板	1
B3Y8	Blaze 转接卡	1
AUMV	ThinkSystem M.2（配有镜像支持套件）	1
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps 非热插拔固态硬盘	2
B0ZR	ThinkSystem 2.5 英寸 Intel S4600 960GB 主流 SATA 6Gb 热插拔固态硬盘	2
AUZY	ThinkSystem I350-T2 PCIe 1Gb 2 端口 RJ45 以太网适配器	1
AU0B 或 ASWQ	Intel OPA 100 系列单端口 PCIe 3.0 x16 HFA 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器	1
B47V	条幅滑轨	1
B3YC	2000W 白金级 PSU	2
B3Y7	TLA Sonic 2 平板 MB 条幅	1
AVUT	ThinkSystem XClarity Controller 标准到高级升级	1
B34S	ThinkSystem NVIDIA Tesla V100 32GB PCIe 被动式 GP	1-4

8.2 头节点

表 15 列出了头节点的 BOM。

表 15. 头节点：SR650

代码	描述	数量
7X06CTO1WW	SR650 头节点：ThinkSystem SR650 - 3 年保修	1
AUVV	ThinkSystem SR650 2.5 英寸机箱，含 8 个、16 个 或 24 个插槽	1
AWEN	Intel Xeon 金牌 6130 16C 125 瓦 2.1GHz 处理器	2
AUND	ThinkSystem 32GB TruDDR4 2666 MHz（2Rx4 1.2V）RDIMM	12
AURA	ThinkSystem 2U/Twr 2.5 英寸 SATA/SAS 8 插槽背板	3
AUNJ	ThinkSystem RAID 930-8i 2GB 闪存 PCIe 12Gb 适配器	3
AUMV	ThinkSystem M.2（配有镜像支持套件）	1
AUR3	ThinkSystem SR550/SR590/SR650 x16/x8 PCIe FH 转接卡 1 套件	1
AURC	ThinkSystem SR550/SR590/SR650（x16/x8）/（x16/x16）PCIe FH 转接卡 2 套件	1
AUKK	ThinkSystem 10Gb 4 端口 SFP+ LOM	1
AU0B 或 ASWQ	Intel OPA 100 系列单端口 PCIe 3.0 x16 HFA 或 Mellanox ConnectX-4 1x100GbE/EDR IB QSFP28 VPI 适配器	1
AXCA	ThinkSystem 免工具滑轨	1
AURD	ThinkSystem 2U 左侧 EIA 标准滑锁	1
AVWG	ThinkSystem 1600W（230V）白金级热插拔电源模块	2
B0ZR	ThinkSystem 2.5 英寸 Intel S4600 960GB 主流 SATA 6Gb 热插拔固态硬盘	24
AUUV	ThinkSystem M.2 CV3 128GB SATA 6Gbps 非热插拔固态硬盘	2
AURQ	Lenovo ThinkSystem 2U 3FH 转接卡支架	1
AURP	Lenovo ThinkSystem 2U 2FH 转接卡支架	1
AUNP	FBU345 超级电容器	3
AUQB	Lenovo ThinkSystem 主流 MB - 2U	1
5977	选择存储设备 – 无需已配置的 RAID	1
AUSA	适用于 EIA 的 Lenovo ThinkSystem M3.5" 螺钉	4
AURR	适用于转接卡的 ThinkSystem M3.5 螺钉 2x2 件和 SR530/550/558/570/590 Planar 5 件	4
B0ML	功能：在主板上启用 TPM	1
6311	2.8 米，10A/100-250V，C13 转 IEC 320-C14 机架电源线	2
B173	XClarity™ Controller Standard 到 Enterprise 工厂升级	1

8.3 管理网络交换机

表 16 列出了管理网络交换机的 BOM。

表 16. 管理网络交换机

代码	描述	数量
7159HC1	Lenovo RackSwitch G8052（从后到前）	1
ASY2	Lenovo RackSwitch G8052（从后到前）	1
A3KR	适用于 442 毫米 RackSwitch 的通风口导风管	1
A3KP	可调节的 19 英寸 4 柱导轨套件	1

8.4 数据网络交换机选项

表 17 列出了 OPA 数据网络交换机的 BOM。

表 17. OPA 数据网络交换机

代码	描述	数量
0449HCR	交换机：Intel OPA 100 系列 48 端口非管理型边缘交换机（PSE）	1
AU08	Intel OPA 100 系列 48 端口非管理型边缘交换机（PSE）	1
AU0C	Intel OPA 100 系列边缘交换机管理卡	1
AU0L	3 米 Intel OPA 100 系列无源铜缆 QSFP28（每个节点一根）	1

表 18 列出了 IB 数据网络交换机的 BOM。

表 18. IB 数据网络交换机

代码	描述	数量
0724Y13	交换机：Mellanox SB7890 EDR IB 非管理型交换机（PSE）	1
00MP560	3 米 Mellanox EDR IB 无源铜缆 QSFP28	1

表 19 列出了 100GbE 数据网络交换机的 BOM。

表 19. 100GbE 以太网数据网络交换机

代码	描述	数量
7159HE1	交换机：Lenovo ThinkSystem NE10032 RackSwitch（从后到前）	1
AV17	Lenovo ThinkSystem NE10032 RackSwitch（从后到前）	1
A3KP	可调节的 19 英寸 4 柱导轨套件	1
AV20	Lenovo 3 米无源 100G QSFP28 DAC 线缆	1

表 20 列出了 25GbE 数据网络交换机的 BOM。

表 20. 25GbE 以太网数据网络交换机

代码	描述	数量
7159HE3	交换机：Lenovo ThinkSystem NE2572 RackSwitch（从后到前）	1
AV19	Lenovo ThinkSystem NE2572 RackSwitch（从后到前）	1
A3KP	可调节的 19 英寸 4 柱导轨套件	1
AV1X	Lenovo 3 米无源 25G SFP28 DAC 线缆	1

8.5 联想智能超算平台（LiCO）

表 21 列出了 LiCO 的订购选项。

表 21. LiCO

代码	描述	数量
B1YC	Lenovo HPC AI LiCO 软件 90 天评估许可证（无支持或更新）	1
B1Y9	Lenovo HPC AI LiCO 软件，包含 1 年“服务和支持”	1
B1YA	Lenovo HPC AI LiCO 软件，包含 3 年“服务和支持”	1
B1YB	Lenovo HPC AI LiCO 软件，包含 5 年“服务和支持”	1

9 附录：示例训练工作负载

本部分介绍此 AI 训练解决方案验证中使用的示例环境。

9.1 AI 框架

9.1.1 Intel Caffe

使用了 Intel Caffe 1.1.0 版、Intel Math Kernel Library 库（MKLML）2018.0.1.20171227 和 Intel MKL-DNN v0.12。Intel Caffe 使用 Intel Skylake 优化编译器标志通过 gcc 进行编译。

9.1.2 TensorFlow

TensorFlow 版本 1.8 和 CUDA 版本 9.0.176 用于 GPU 培训。TensorFlow 下载时已经过预编译。

9.2 训练数据集

使用了 ImageNet 数据库。ImageNet 是按照 WordNet 层次结构（目前仅包含名词图像）组织的图像数据库，该层次结构的每个节点使用成千上万张的图像进行描述。用于训练的 1.28M ImageNet 数据库是全球研究者和数据科学家均在使用的大型图像数据库。

9.3 训练模型

训练了一个 50 层残差网络（ResNet-50）神经网络。ResNet-50 由微软研究院于 2015 年发布，用于 ImageNet 和 MS-COCO 竞赛。

9.4 训练长度

ResNet-50 模型在 1.28M ImageNet 数据集上训练了 90 期（epoch）。一期包含对训练集的一个完整训练循环。训练运行了 90 期的意思是每张图像被训练模型遍历了 90 次。在 1.28M ImageNet 数据集上运行 90 期即相当于处理了 1.15 亿张图像。

9.5 训练数据

图 16 显示了几种不同系统配置的图像训练数据。此图表可用作调整指南，以了解达到所需图像处理速度需要何种集群配置。蓝色条形代表没有配备 GPU 的仅 CPU 训练节点，适用于所需速度低于 1000 张图像/秒的工作负载。绿色条形表示配置了 P100 GPU 的节点，其性价比对于低于 2000 张图像/秒的工作负载是个不错的选择。橙色和红色条形表示配置了 V100 GPU 的系统，其使用 8 个 GPU 可处理约 5000 张图像/秒。右侧坐标轴使用的数据相同，但将数据单位转换为百万张图像/天。

让我们以 2000 张图像/秒的训练要求为例。在图 16 中找到 2000 张图像/秒，可发现有多种方法可满足此训练要求：在四个 SD530 节点中使用 8 个 P100，在两个 SD530 节点中使用 4 个 V100，或在一个 SR670 系统中使用 4 个 V100。对于本示例，Lenovo 的建议是在一个 SR670 系统中使用 4 个 V100。此配置需要最少的 CPU 数和内存即可达到 2000 张图像/秒的训练要求，因此成本最低。

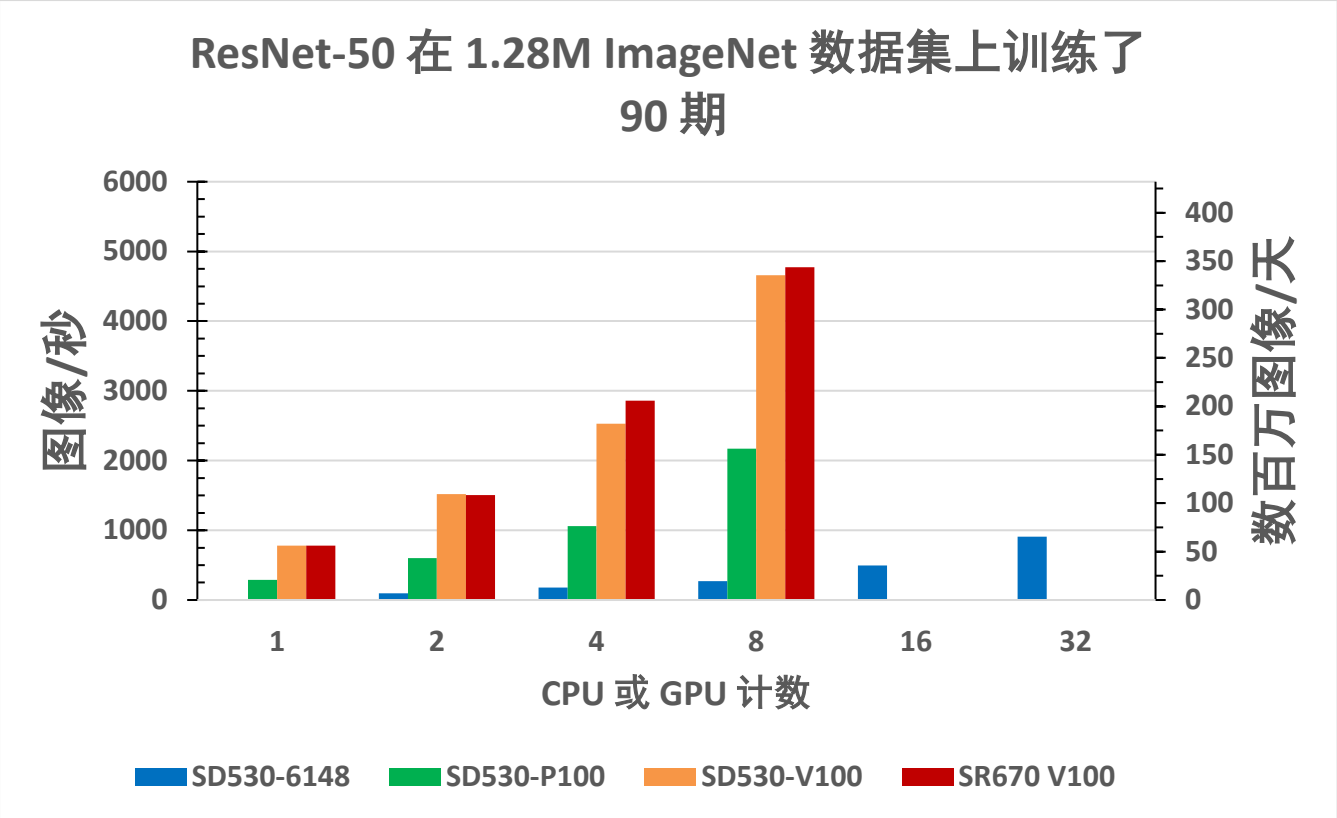

图 16. 训练数据

注：这些信息基于缺省的 Intel Caffe 和 TensorFlow 设置。

9.6 训练网络比较

图 17 显示了几种不同网络配置的图像训练数据。此图表显示不同网络速度（10、25、100 Gb/s 以太网和 100 Gb/s OPA）下的训练性能（图像/秒）。数据分别使用 1 个、2 个 和 4 个 SD530 服务器进行收集，每个服务器包含两个 V100 GPU。在只有两个 GPU 的情况下，所有网络的性能都相同，因为此时只需要一个服务器，因此无需数据网络。使用两个节点（即四个 GPU）时，网络开始对训练性能产生影响。使用四个节点（即八个 GPU）时影响进一步增大，可看到 10 Gb/s 以太网比 100 Gb/s OPA 的性能低 20%。随着 GPU 和节点数的增加，较慢网络的性能差距将继续扩大，因此 Lenovo 建议 GPU 的数量在 8 个及以上时，使用 100 Gb/s 网络。

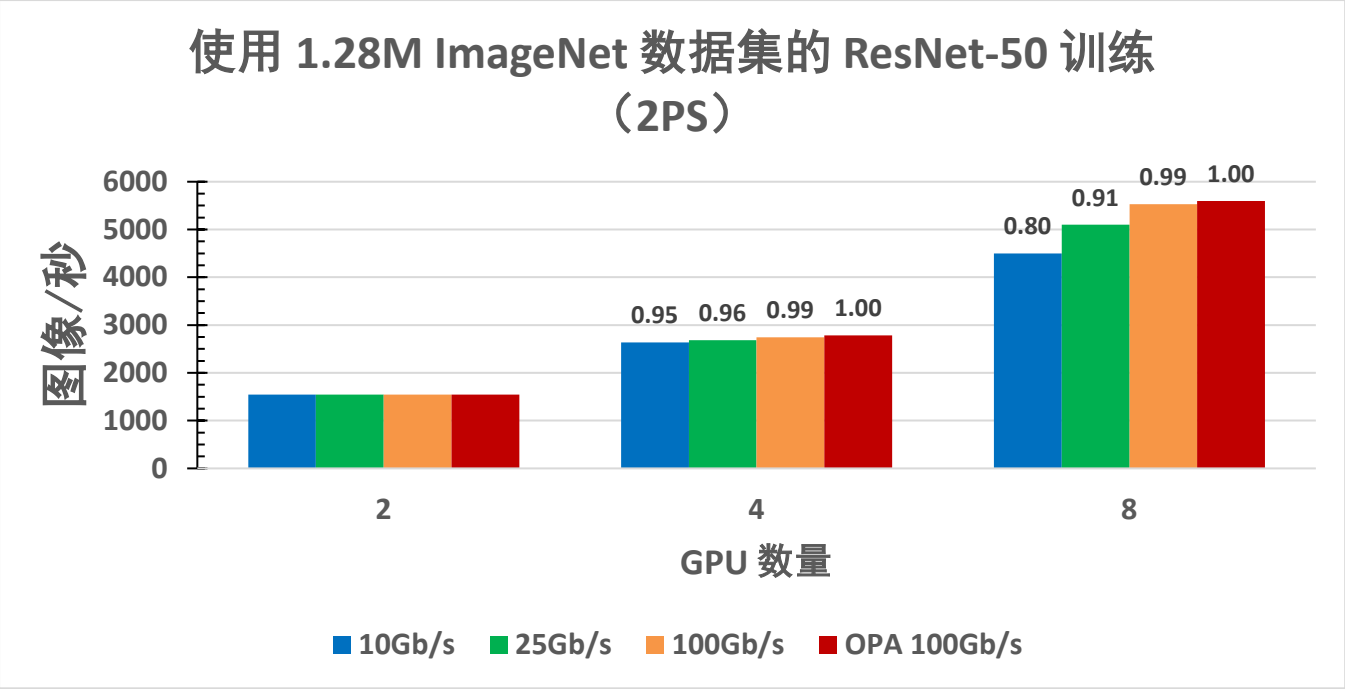

图 17. 训练网络比较

9.7 推理框架

推理测试使用的是 TensorFlow Serving 1.11.1 发行版。TensorFlow Serving 作为 Docker 容器下载，并在 Docker-CE 18.06.1 版中运行。模型在 TensorFlow 中训练并使用 SavedModel 功能导出。与检查点文件不同，SavedModel 独立于 TensorFlow 的版本并可移植。SavedModel 与 TensorFlow Serving 兼容，可在启动时由 Docker 容器加载或嵌入在容器中。

9.8 大数据框架

HDFS 存储测试使用的是 Cloudera 5.9.1。HDFS NFS 网关服务安装在其中一个管理节点上，提供从训练节点到大数据集群的 NFS 装载路径。可在多个节点上安装此服务以增加到 HDFS 存储集群的带宽。

有关适用于 Cloudera Enterprise 的 Lenovo 大数据验证设计的详细信息，请访问：

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

资源

有关更多信息，请参阅以下资源：

Lenovo ThinkSystem SR650（训练/头节点）：

- Lenovo Press 产品指南：https://lenovopress.com/lp0644-lenovo-thinksystem-sr650-server
- 3D 导览：https://lenovopress.com/lp0673-3d-tour-thinksystem-sr650

Lenovo ThinkSystem SD530（训练节点）：

- Lenovo Press 产品指南：https://lenovopress.com/lp0635-thinksystem-sd530-server
- 3D 导览：https://lenovopress.com/lp0667-3d-tour-thinksystem-sd530

Lenovo ThinkSystem SR670（训练节点）：

- Lenovo Press 产品指南：https://lenovopress.com/lp0923-thinksystem-sr670-server

Lenovo RackSwitch G8052（1GbE 交换机）：

- Lenovo Press 产品指南：https://lenovopress.com/tips1270-lenovo-rackswitch-g8052

Lenovo RackSwitch NE10032（100GbE 交换机）：

- Lenovo Press 产品指南：https://lenovopress.com/lp0609-lenovo-thinksystem-ne10032-rackswitch

Lenovo RackSwitch NE2572（25GbE 交换机）：

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

Lenovo XClarity Administrator：

- Lenovo Press 产品指南：https://lenovopress.com/tips1200-lenovo-xclarity-administrator

Lenovo 大数据验证设计：

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

Intel OPA 100（100Gb 交换机）

- 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 端口 InfiniBand 交换机（100Gb 交换机）

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

Intel Xeon 可扩展系列均衡内存

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

NVIDIA GPU

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

Caffe

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

Intel Caffe

- Caffe 的 Intel 分发版：https://github.com/BVLC/caffe/tree/intel
- 多节点指南：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

开源软件：

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

文档历史记录

1.0 版本	2018 年 5 月 15 日	第一版
1.5 版本	2018 年 6 月 22 日	第二版 • 在 SD530 和 SR650 训练节点配置中添加了 V100 GPU • 添加了 TensorFlow AI 软件框架 • 更新了 BOM 表，添加了配备 V100 GPU 的配置 • 更新了 BOM 表，添加了配备 InfiniBand 交换机和适配器的配置
1.6 版本	2018 年 7 月 27 日	第三版 • 在 SD530 和 SR650 训练节点配置中添加了 P100 GPU • 更新了 BOM 表，添加了配备 P100 GPU 的配置
1.7 版本	2018 年 9 月 7 日	第四版 • 添加了 SR670 训练节点配置 • 更新了 BOM 表，添加了 SR670 的配置 • 更新了 BOM 表，添加了 100Gb 以太网交换机的配置
2.0 版本	2018 年 11 月 9 日	第五版 • 添加了推理部分 • 添加了大数据存储部分 • 更新了 BOM 表，添加了 25Gb 以太网交换机的配置

商标和特别声明

© Copyright Lenovo 2018.

本文档中对 Lenovo 产品或服务的引用并非默认 Lenovo 有意在所有国家/地区提供这些产品和服务。

Lenovo、Lenovo 徽标、ThinkCenter、ThinkVision、ThinkVantage、ThinkPlus 以及 Rescue and Recovery 均为 Lenovo 的商标。

Microsoft、Windows、Windows NT 以及 Windows 徽标 是 Microsoft Corporation 在美国和/或其他国家或地区的商标。

Intel、Intel Inside（徽标）、MMX 和 Pentium 均为 Intel Corporation 在美国和/或其他国家或地区的商标。

Other company, product, or service names may be trademarks or service marks of others.

信息“按原样”提供，且不附带任何形式的保证。

所述的所有客户示例均为说明客户曾经使用 Lenovo 产品及其可能取得的结果之目的。实际的环境费用及性能特征可能因不同客户而异。

有关非 Lenovo 产品的信息来自此等产品的供应商、发布的公告材料或其他公开可获得的来源，均不构成 Lenovo 对这些产品的任何认可。非 Lenovo 价目表和性能数据来自公开可获得的信息来源，包括供应商公告和供应商全球主页。Lenovo 未对此等产品进行测试，且无法确认有关非 Lenovo 产品的性能、功能或任何其他表述的准确性。有关非 Lenovo 产品功能的问题应向此等产品的供应商提出。

有关 Lenovo 未来方向和意向的所有声明仅代表 Lenovo 之目标，且均有可能随时变更或撤销，恕不另行通知。若需获取具体方向声明的完整文本，请联系您当地的 Lenovo 办事处或 Lenovo 授权经销商。

部分信息述及了未来期望功能。此类信息不应被视为对任何未来产品的具体性能水平、功能或交付计划所作的明确承诺。此等承诺仅在 Lenovo 产品公告中作出。在此处提供此类信息旨在介绍 Lenovo 当前的投资和开发活动，以诚实信用为原则帮助我们的客户制定未来的规划。

性能基于使用标准 Lenovo 基准在受控环境下所得的测量值和预测值。任何用户将获得的实际吞吐率或性能将因多种因素而异，例如用户作业流中的多程序数量、I/O 配置、存储配置、处理的工作量等。因此，不保证用户能够实现相当于此处所述比例的吞吐率或性能改进。

所示图片仅为工程原型。生产型号可能会有所变更。

本资料中对非 Lenovo Web 站点的任何引用都只是为了方便起见才提供的，不以任何方式充当对那些 Web 站点的保证。那些 Web 站点中的资料不是 Lenovo 产品资料的一部分，使用那些 Web 站点带来的风险将由您自行承担。