联想高性能计算
解决方案白皮书

手机专线 400 819 6776
座机专线 800 830 6776
dcg.lenovo.com.cn

咨询电话 400 819 6776

英特尔®至强®可扩展平台 加速探索与获取洞察

CONTENTS
目录

1 联想高性能计算

> 高性能计算多年来一直是科技综合实力竞争的至高点，也在一定程度上反映了各大服务器厂商系统研发方面的实力。作为行业的技术领先者，联想公司在这一领域积累了长达 20 多年的丰厚经验，并在关键技术领域不断创新，取得大量里程碑式的成果。

联想从 1999 年进军高性能服务器领域，是最早针对高等院校和科研院所进行产品开发与市场拓展的厂商，并在市场中一直处于技术领先地位。截止 2018 年，先后为数万个用户成功实施了高性能集群。曾经两次承担了中国科学院网络计算中心主节点的建设任务，并且成功地与威廉姆斯车队进行合作，成为国产品牌中最早将高性能业务拓展到海外的企业。以此为契机，联想顺应国际主流技术发展趋势，以市场需求为驱动，吸收国内外最新技术成果，进行了大量创新性研发，突破包括系统设计与优化、系统基础架构、系统软件等在内的一大批高性能服务器的关键核心技术，开发出一系列可扩展、易管理、好使用、稳定可靠的高性能服务器产品，并配备可满足用户个性化需求的行业解决方案，提供从系统层到应用软件层的全面解决方案和技术服务。

2002 年 7 月，联想研制成功“深腾 1800”万亿次集群系统，安装在中科院数学与系统科学研究院。这是世界上第一个实际速度超过 1 万亿次的大规模集群系统。曾入选新华社 2002 年中国十大新闻及两院院士评选的 2002 年中国十大科技进展，并荣获 2004 年国家科技进步二等奖。2002 年末，另一套深腾 1800 大规模集群系统安装在中科院大气物理所国家重点实验室。

2002 年 12 月 30 日，联想深腾 1800 中标大庆油田，使该油田第一次在国内实现三维叠前深度偏移地震资料处理。


国家科学技术进步奖

证 书

为表彰国家科学技术进步奖获得者，特颁发此证书。

项目名称：联想深腾 1800 大规模计算机系统

奖励等级：二等

获 奖 者：联想（北京）有限公司

证书号：2004-J-254-2-03-D01

2003 年，联想成功研制“国家网格主节点—联想深腾 6800 超级计算机”，安装在中科院计算机网络信息中心。这是当时世界上 Linpack 效率（78.5%）最高的高端通用计算机，其组合查询性能名列当时所有大型服务器的第四位， 其典型应用 MM5 的测试结果在 2004 年 3 月列世界所有超级计算机的第一位。该机荣获 2005 年国家科学技术进步二等奖、2005 年国家重点新产品奖、2004 年信息产业重大技术发明奖。联想深腾 6800 自 2004 年初在网络中心对外服务以来，一直 7 X 24 小时稳定运行，在双星计划、气候模式计算、油藏模拟、材料科学计算、流体力学计算等领域取得了 150 多项重要计算成果。


国家科学技术进步奖

证 书

为表彰国家科学技术进步奖获得者，特颁发此证书。

项目名称：国家网格主结点—联想深腾 6800 超级计算机

奖励等级：二等

获 奖 者：联想（北京）有限公司

证书号：2005-J-226-2-07-D01

联想深腾系列高性能计算机成为最早进入世界 TOP500 的一批国产计算机， 分列当时世界 TOP500 的第 14、43、98 和 299 名。 这是一个历史性的突破，联想深腾系列高性能计算机已成为国际知名国内主流的品牌。联想在推动高性能技术产业化方面取得了突破性进展，联想的高性能计算机广泛应用于许多关键领域，在国民经济和社会发展中发挥重要作用。

目前，集群已成为世界高性能计算机体系结构的主流，联想深腾 1800、深腾 6800 和深腾 7000 为这一趋势的形成做出了重要贡献。2002 年 8 月初，世界上主流并行编程环境 MPI-ch 的发明人、美国阿贡实验室 William Cropps 参观联想深腾 1800 后写道：“We see the future of clustering computing”。

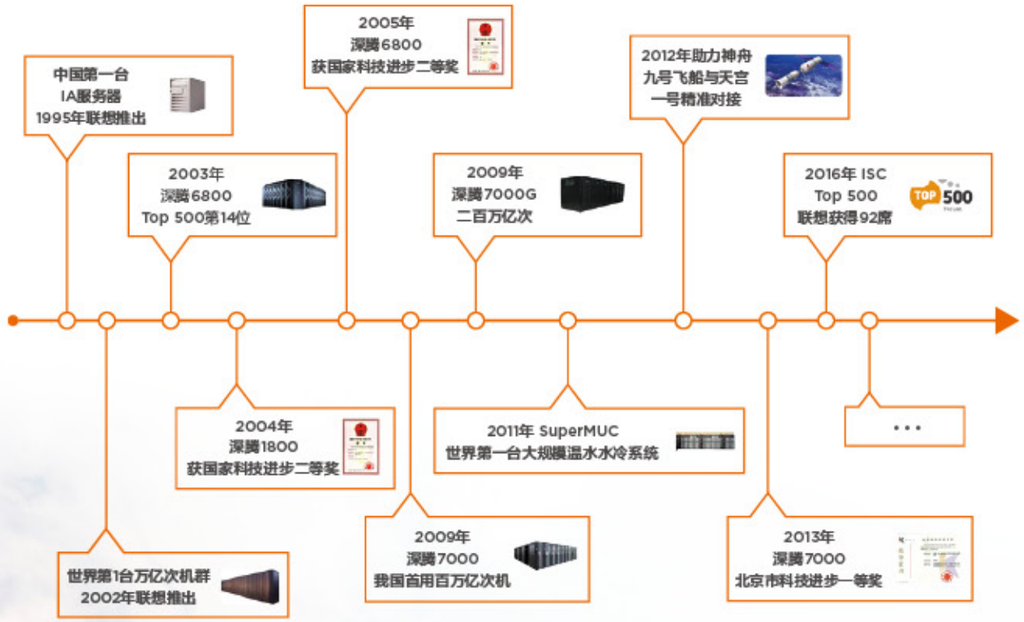
联想在高性能服务器基础技术方面有着长期的积累，有齐全的产品线和严格的质量控制体系，为高性能计算机的研制和生产奠定了坚实的基础。

在产品设计上，联想坚持用户导向的原则，同时结合对新技术的深入理解和消化吸收，始终遵循模块化设计思想，在充分综合考虑各模块精密配合和整机系统合理整合的基础上，先设计出最佳性价比、最稳定的产品方案，然后对方案进行工程计算仿真，同时不断地结合验证性实验，最终才形成可行的开发方案，从而保证为用户在最短的时间里开发出最贴近的具有竞争力的产品。

在研究开发上，联想建立了与国际接轨的两级研发体系，即公司级研发平台和各事业部研发中心。公司级研发平台由联想研究院、软件中心、板卡中心和工业设计中心组成。事业部研发中心隶属于各事业部，直接承担具体的专项技术开发工作。联想在高性能服务器技术上已突破并拥有了自己的核心技术，拥有自主知识产权的系统设计与优化技术、系统监控技术、系统管理技术、高可用和负载均衡技术以及基础架构技术等关键技术，在高性能计算机系统技术方面已申请国家发明专利 85 项，其中，46 项已获授权。

在工程技术上，联想拥有针对服务器的部件及整机进行专业性测试的全套技术。部件测试包含外观、结构、功能、兼容性、可靠性、安全性、性能和环境 8 个方面的测试，以保证所有部件符合联想服务器技术特性和质量标准的要求，对服务器的核心部件如电源、内存，还建立了专业化的实验室，实现了部件的自动测试。如全球技术领先的自动电源测试实验室和自动内存测试实验室，国内功能最全面、技术最先进的系统测试实验室，以及高温实验室、电磁兼容检测实验室、噪音实验室、湿热实验室等等，所有产品需要在这些实验室中通过一系列的严格检测，只有通过了这一系列的严格检测的服务器产品，才可以顺利出厂，提供给客户。联想始终严格执行国际标准的质量控制体系，是国内最早通过 ISO9000 – 2000 版质量认证体系的服务器厂商。

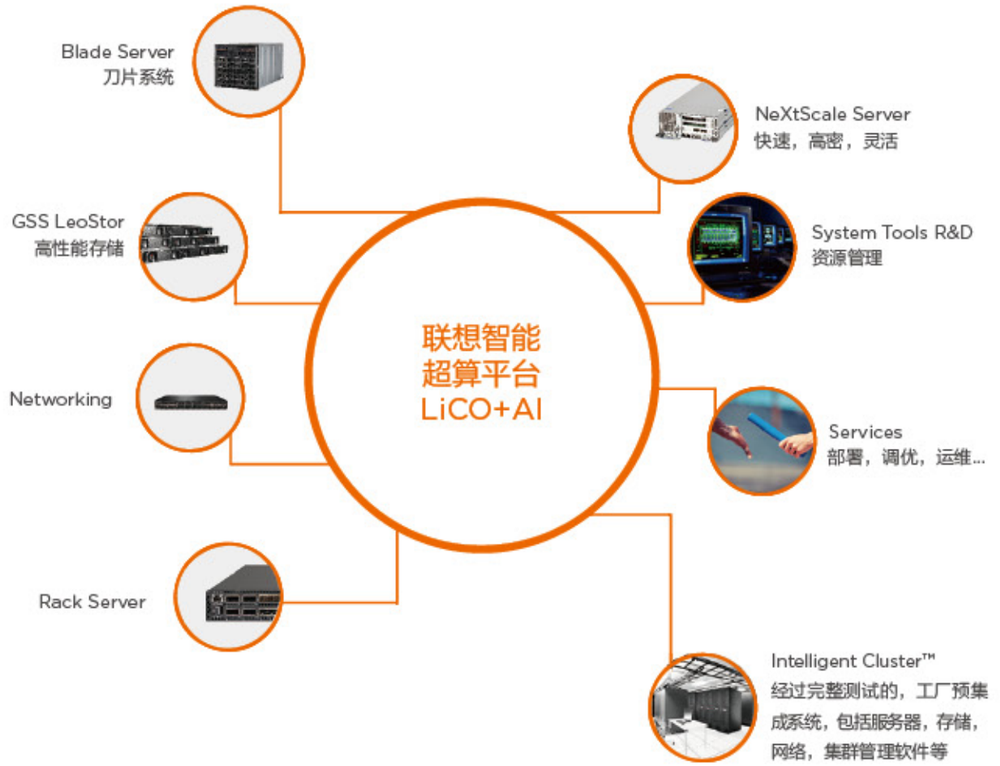
在技术服务与方案上，联想服务器应用方案中心拥有雄厚的技术力量，在硬件平台、操作系统、数据库、软件、网络、存储、集群技术等方面有着多年的技术和经验积累，可以分别从不同的技术层面为用户提供有效的产品应用和方案支持服务。中心拥有先进的实验环境，包括方案集成实验室、性能评测实验室、数据中心、客户实验室四个部分，为用户提供方案开发、测试、方案移植、优化以及培训、咨询等服务，及时、快速、可靠地解决用户系统在使用过程中所遇到的技术问题，使客户的系统可以更加安全稳定地运行，以保障和促进客户业务的顺利开展并取得更大的成功。



2014 年 9 月 29 日， 联想宣布完成对 IBM x86 业务的收购，从此，具有丰富的高性能计算方面经验的原 IBM x86 大批 HPC 专家加入了联想。

原 IBM x86 部门的熟悉应用的行业专家非常了解行业用户的需求，他们会针对行业的具体情况，与行业应用软件开发商密切配合，提供切实可行的解决方案，便于最终用户使用。

2016 年 7 月 1 日，从 ISC 2016 凯旋归来的联想集团再度吹响集结号，在北京隆重召开了以“开启 E 级计算新篇章”为主题的首届全球超算峰会。本次大会联想正式发布了其自主研发的，面向 E 级计算的高性能计算机系统深腾 x8800。



2017 年 6 月 30 日，联想正式升级高性能计算机系统为深腾 x8810，这个平台以联想只能超算软件 LiCO 为核心，是联想面向智能超算的统一平台，该平台涵盖传统高性能计算和人工智能技术。这是我们向人工智能方向迈出的一大步。

2018 年，联想以 140 套的绝对优势，荣登 TOP500 排行榜第一名。

2018 年，联想的 LiCO AI 产品赢取了国际大奖， 这也是中国的高性能计算产品首次获得的国际大奖。

2018 年，联想凭借在温水水冷方面的多年的技术投入和技术沉淀，获得了 2018 年数据中心科技成果杰出奖。

2019 年联想 HPC/AI 全年实现盈利性增长，利润率提升 5 倍，继续保持中国区 Top-100 的第一名，赢取了南方科技大学（Top-500 排名 127 位）、中芯国际、辽河油田、航空动力等主要高性能项目同时，联想超算高性能计算机平台进行了升级，目前的平台不但涵盖传统高性能计算和人工智能技术同时也将大数据平台进行了有机融合，形成了联想统一计算平台。

2 联想智能超算 LiCO 平台

概述

随着互联网的普及和 IT 业的高速发展，高性能计算（HPC）已经不再是少数大公司或大型科研机构的专属要求，而是被越来越多的包括政府，教育科研，石油石化，制造，军工和生命科学类的客户所需要和接受， 针对 HPC 方案的复杂性，用户需要一个易用的 HPC 集群管理和作业管理平台。

同时，最近几年人工智能（AI）快速发展，大多数人工智能模型的训练需要使用 GPU，而如 NVIDIA Tesla V100 等 GPU 十分昂贵，不同的部门和人员有不同的模型训练需求，为每个部门和人员购买独占的 GPU 是对资源的浪费，所以一个集中共享式的 AI 模型训练平台也成为了越来越多客户的需求。

联想智能超算平台（Lenovo intelligent Computing Orchestration 以下简称 LiCO）是联想数据中心集团（DCG）开发的，针对高性能计算（HPC）和人工智能（AI）的一站式解决方案，在一套集群中通过统一的资源调度，可以同时支持 HPC 作业和 AI 作业的运行。

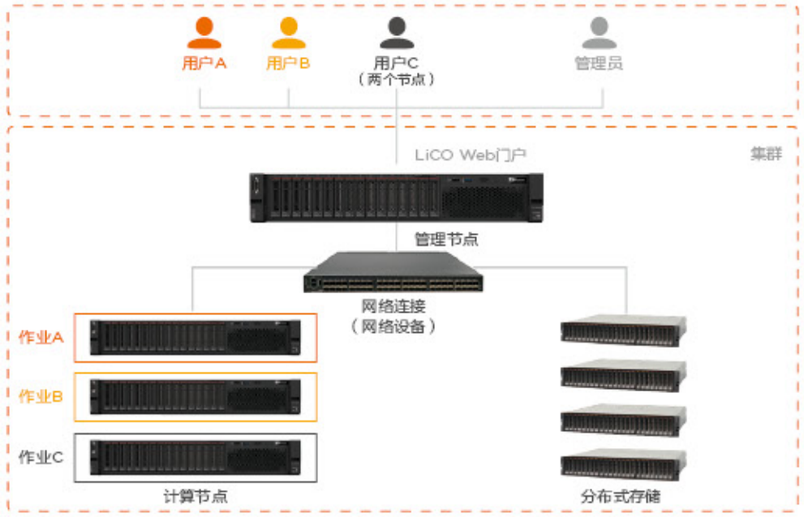
LiCO 集成了集群需要的集群调度软件、监控软件、计算库、分布式文系统等，使用 LiCO 可以快速的部署好一个 HPC 和 AI 集群。

LiCO 提供了统一的 web 访问接口，集群管理员可以使用 LiCO 方便的管理集群，HPC 用户可以使用 LiCO 方便的提交和管理 HPC 作业，AI 用户可以使用 LiCO 进行 AI 模型的训练。

方案介绍

在如下的一个物理集群上：

- LiCO 可以作为一个 HPC 的平台；
- LiCO 可以作为一个 AI 模型训练的平台；
- LiCO 可以作为一个 HPC+AI 模型训练的平台。

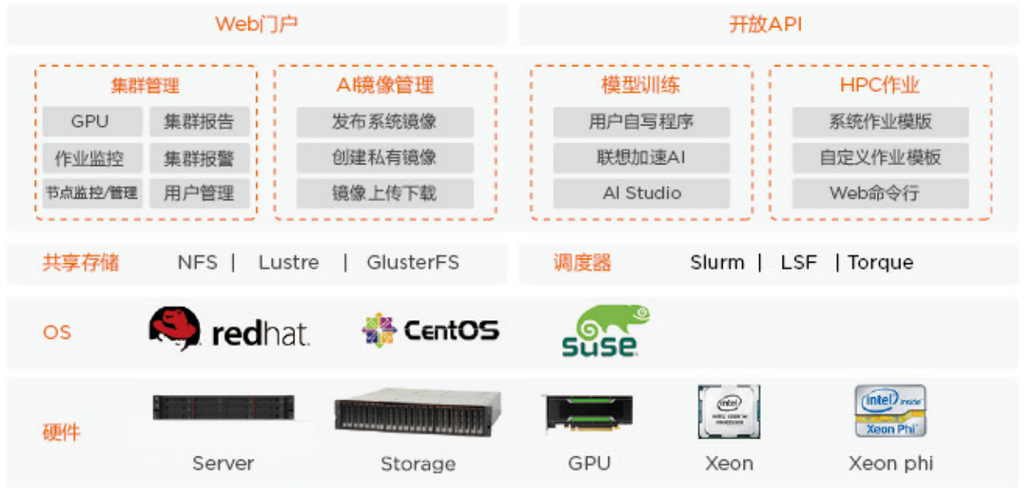


LiCO 平台提供了统一的 Web 门户：

- 集群管理员可以使用 LiCO 管理集群；
- HPC 用户可以使用 LiCO 提交和管理 HPC 作业；
- AI 用户可以使用 LiCO 进行 AI 模型的训练。

LiCO 平台提供了开放 API：

- 方便用户通过 API 提交作业和与其它现有系统进行集成。



LiCO 集群管理功能介绍

集群管理

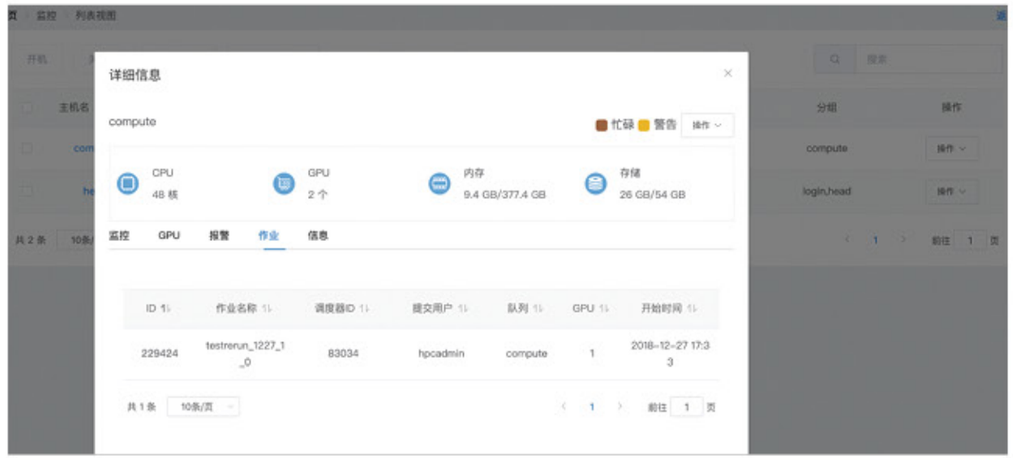
一目了然的集群总体状态图

系统仪表板显示了集群整体状况，包括 CPU、GPU、内存、网络、存储、作业、各类型节点使用情况、报警、调度系统工作状况等，方便管理员直观的了解集群的整体运行状况。

完善的节点信息展示

系统提供了查看单个节点详情信息的功能，节点详细信息包括：节点名、CPU、GPU、内存、硬盘等信息。不仅提供了节点各种监控指标（如温度、能耗、CPU 使用率，Load、内存使用率、网络吞吐等）的历史趋势图，也提供了节点各个 GPU 的使用率、显存占用率、温度的历史趋势信息。

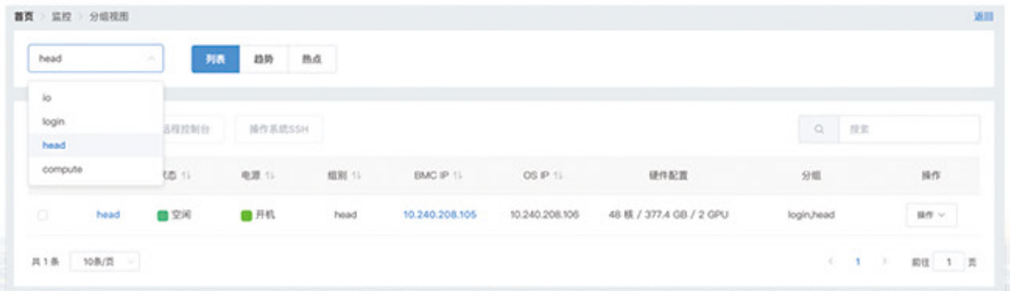
节点上当前作业的列表。



节点上当前报警信息的列表。

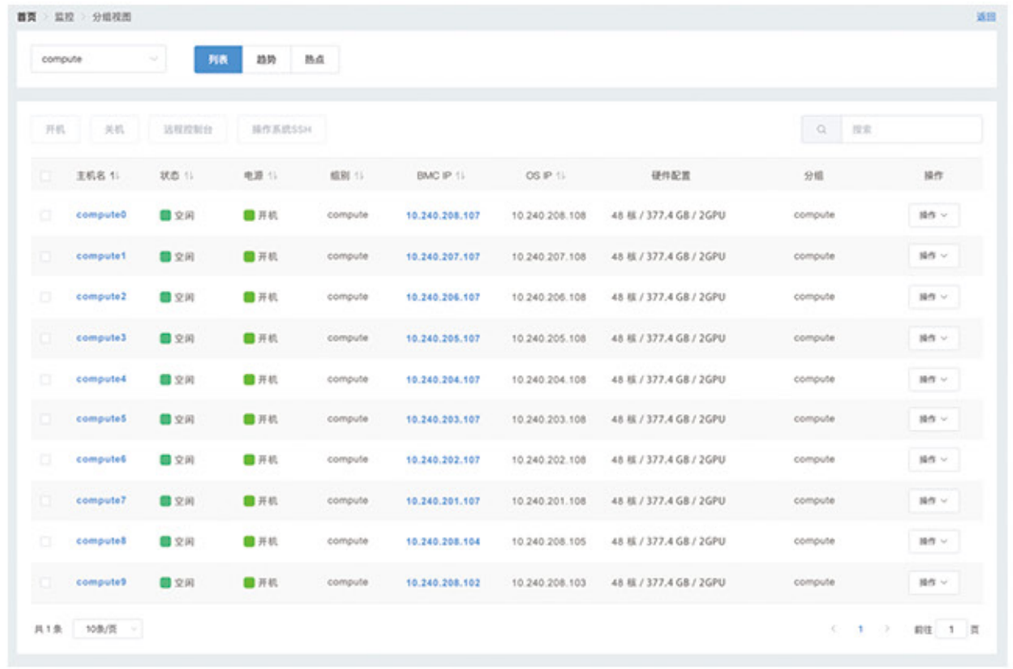
灵活的节点分组策略

系统提供了节点分组功能，管理员可以根据需要将集群节点进行逻辑分组，以便后面对不同的分组进行批量监控和管理。



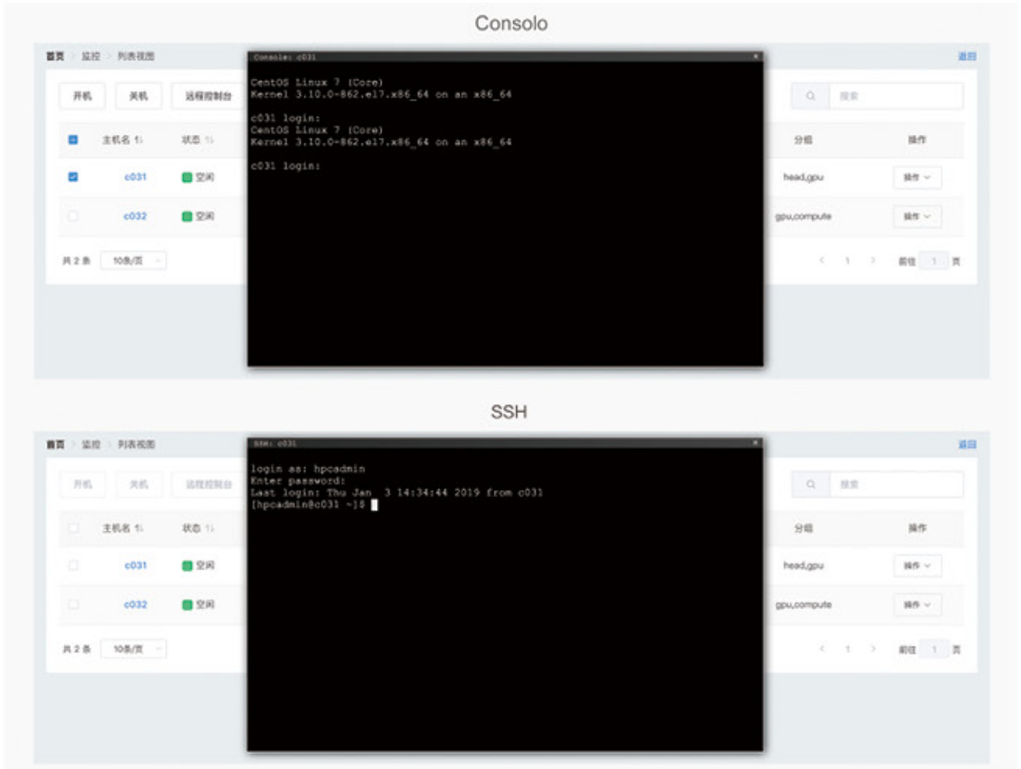
批量电源操作

管理员在 web 页面上可以选中多个节点进行批量电源操作。



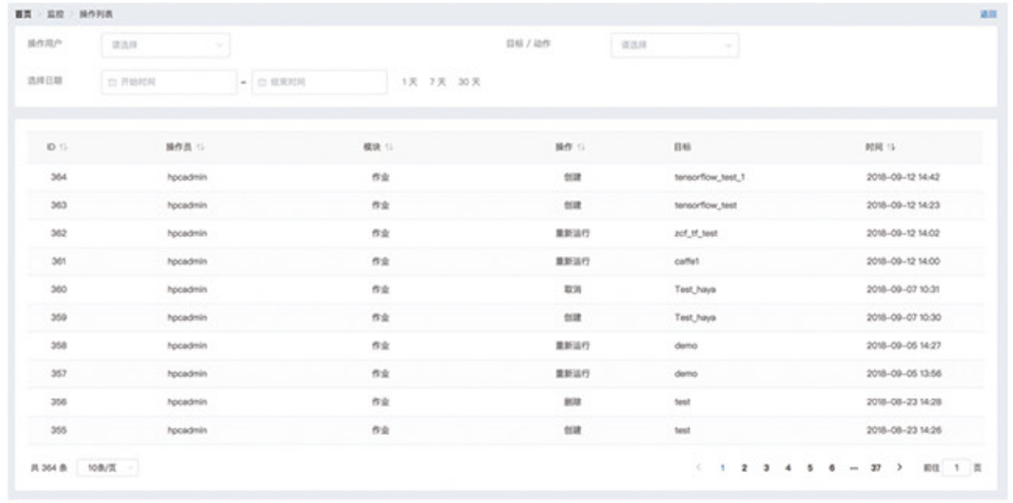
批易用的节点管理 Web Console 和 WebSSH

无需下载任何插件，直接在 Web 界面打开节点的 Console、SSH 来访问和管理节点。



完成的集群操作日志和操作日志报告

通过操作日志功能，管理员可以查看用户在何时对集群进行了何种操作。

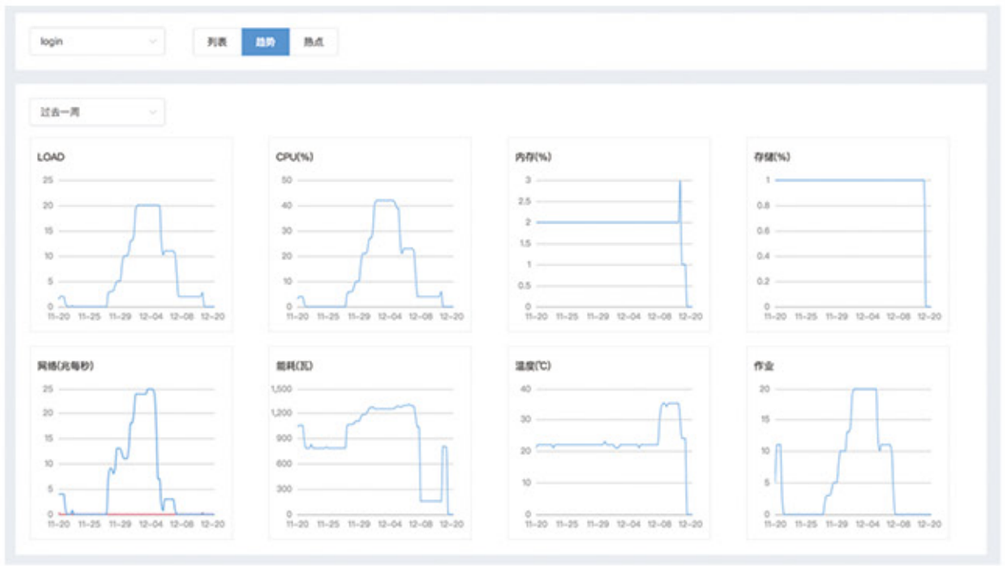


管理员可以通过操作日志报告来导出操作记录。

集群监控

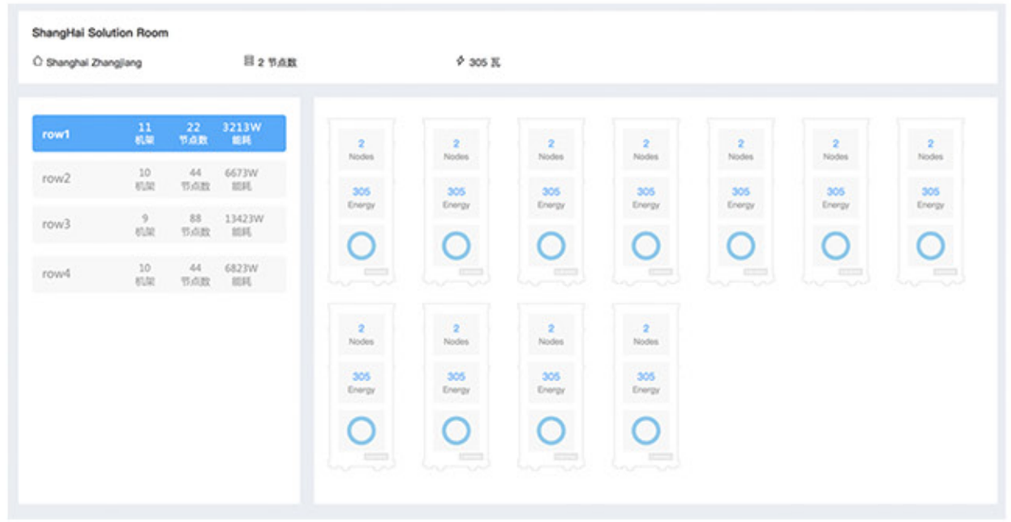
丰富的监控指标

系统支持对多种指标进行监控，如 CPU 使用率、Load、内存使用率、硬盘使用率、网络吞吐、温度、能耗、作业负载等。



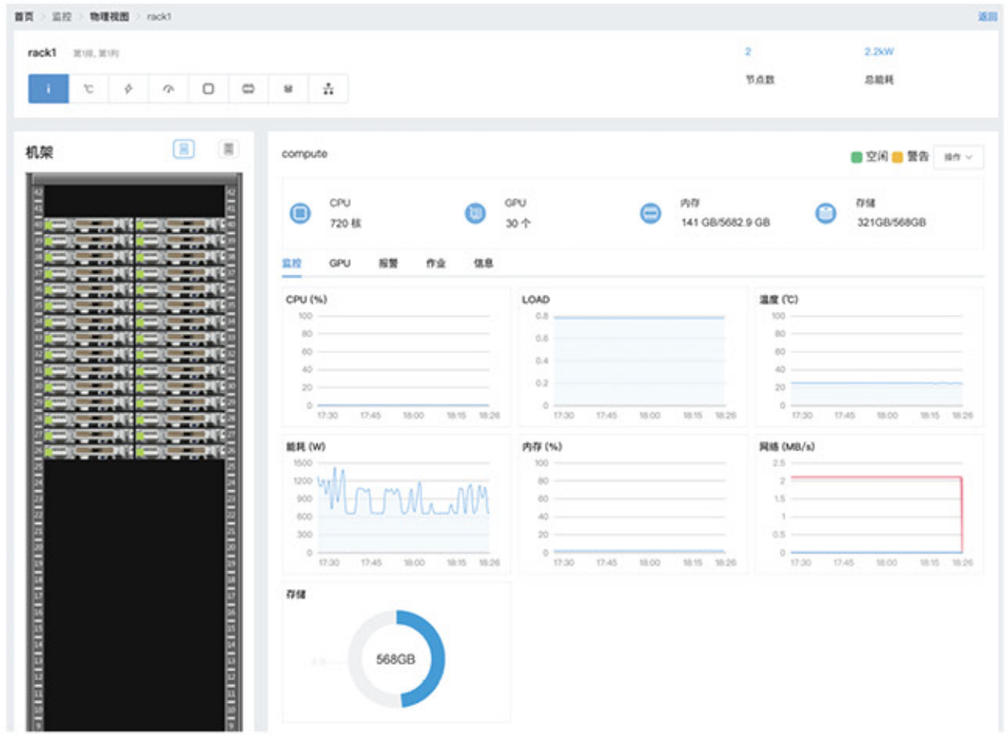
机房物理视图

机房物理视图直观的展示了机房中各个机架的具体位置、名称、能耗、节点的使用及报警统计。

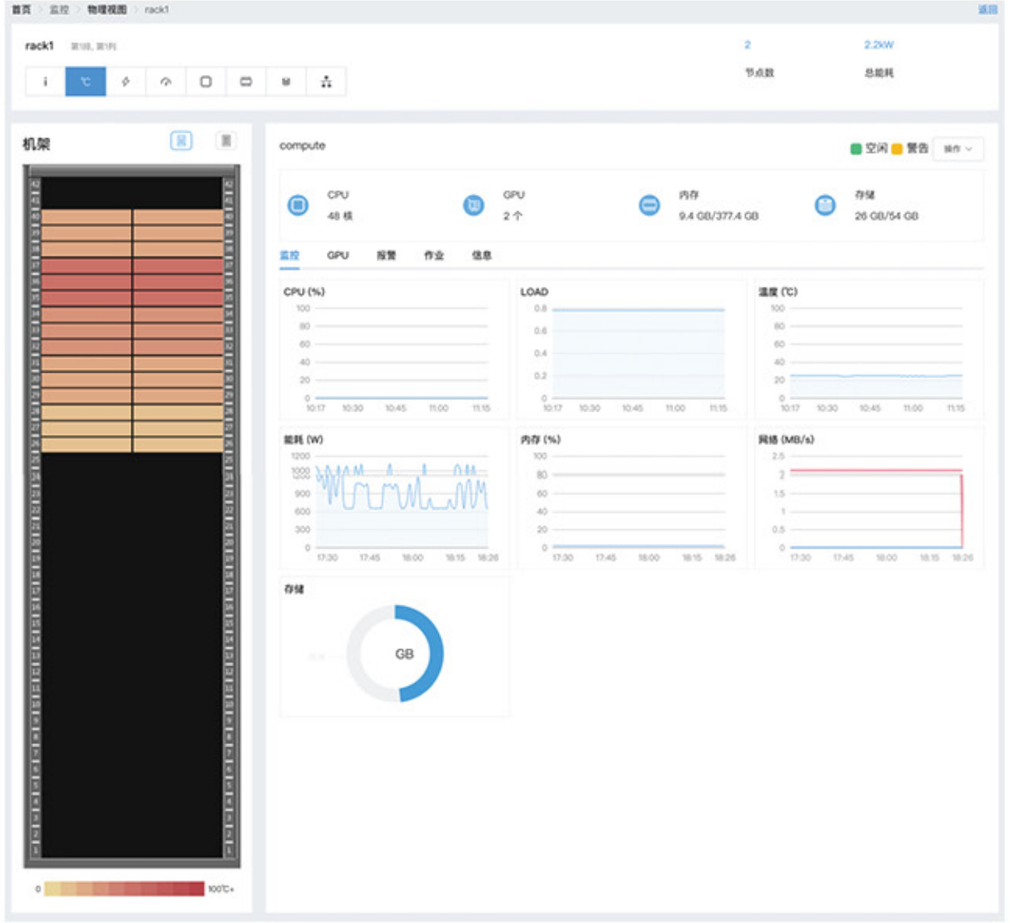


机架物理视图

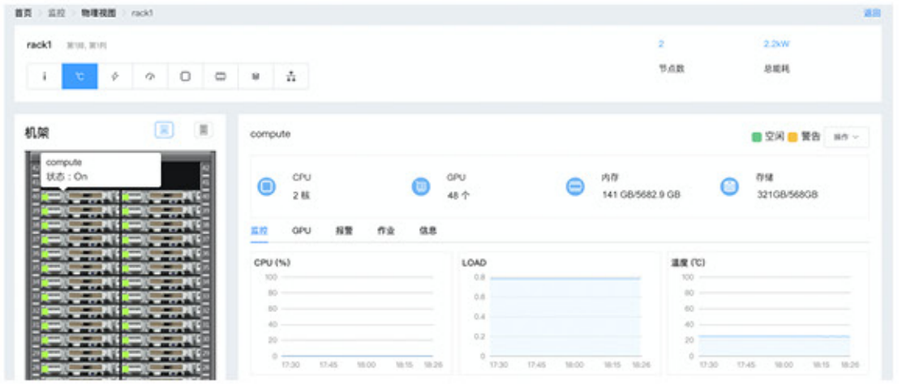
机架物理视图按节点在机架上的实际安装的情况来进行展示。



机架视图可以按照多种监控指标的数值通过颜色来标注节点。



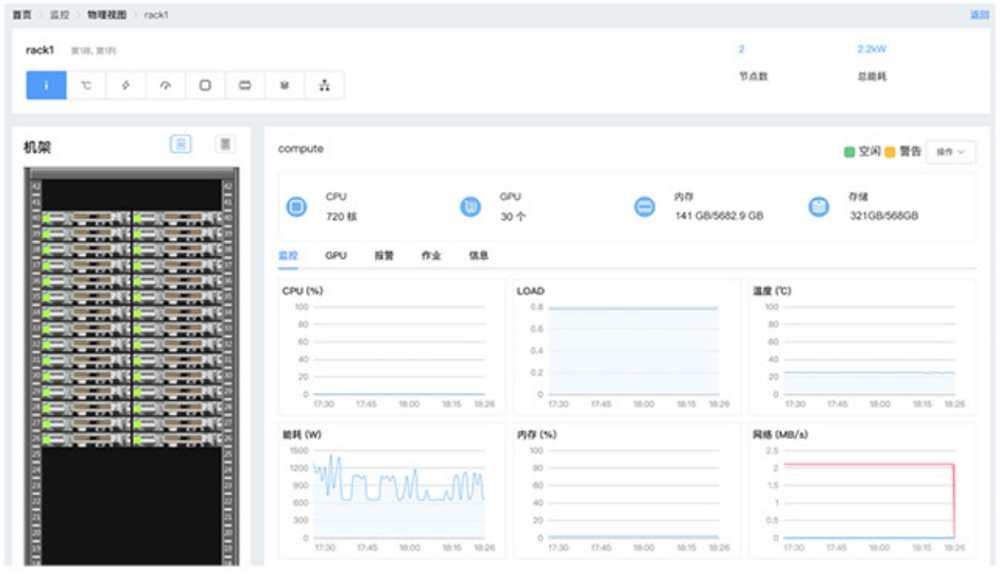
管理员可以点击机架图上具体节点来查看详细监控信息。



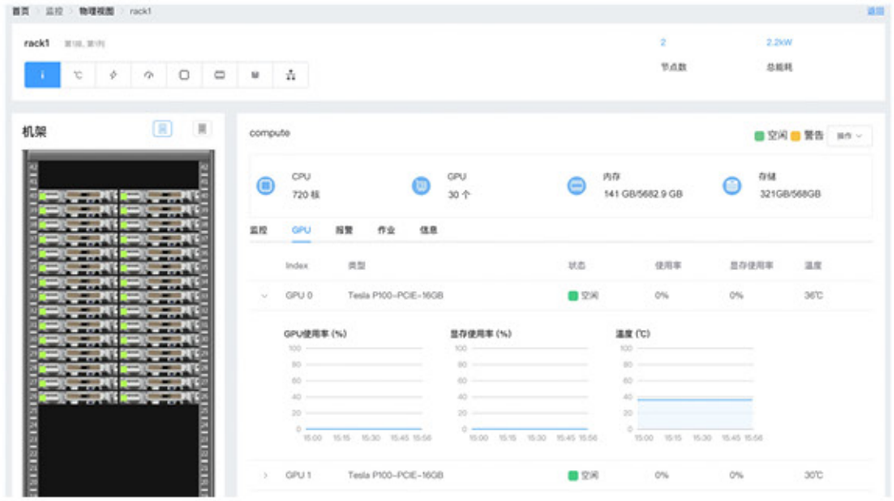
单节点性能监控

系统提供针对单节点详细监控面板，通过该面板可以查看节点的静态配置信息，如 CPU、GPU、内存、硬盘、机器型号、网络 IP 等；并可以对节点进行电源操作或访问节点的 console、ssh。

通过监控面板可以查看各个指标的历史趋势，如 CPU 使用率、Load、内存使用率、硬盘使用率、网络吞吐、温度、能耗等。

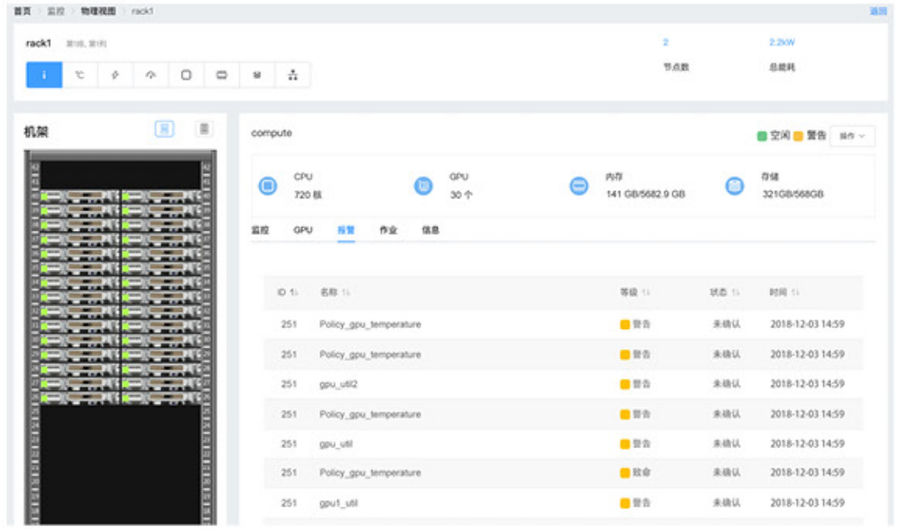


通过 GPU 面板查看节点 GPU 使用情况。



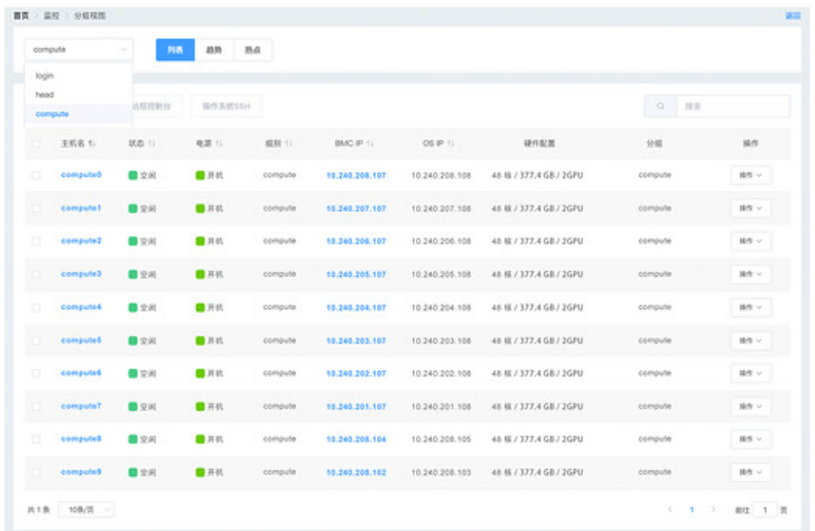
通过作业面板查看节点上正在运行的作业。

查看节点上未处理的报警记录。

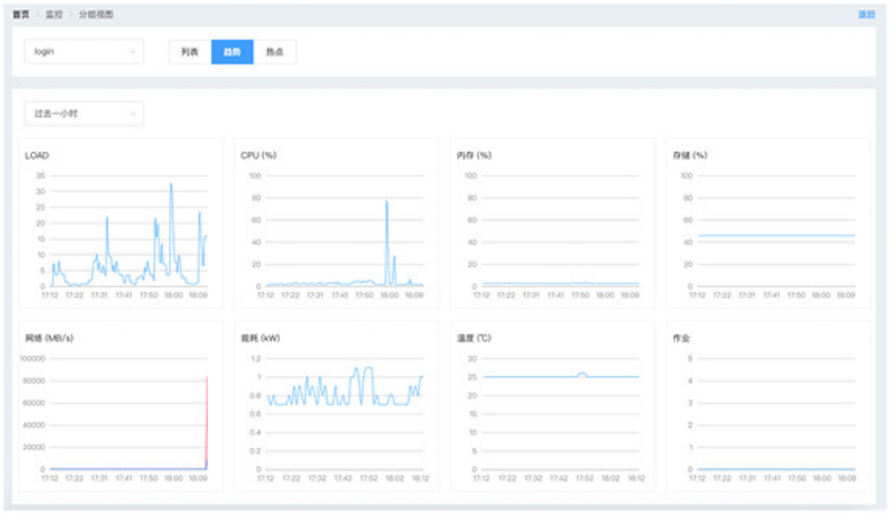


灵活的节点分组

系统支持将集群中的节点分成不同的逻辑分组（单个节点可以从属于多个逻辑分组），管理员可以通过逻辑分组来进行批量的监控和管理操作，为管理大型或超大型集群提供了灵活性和便利性。



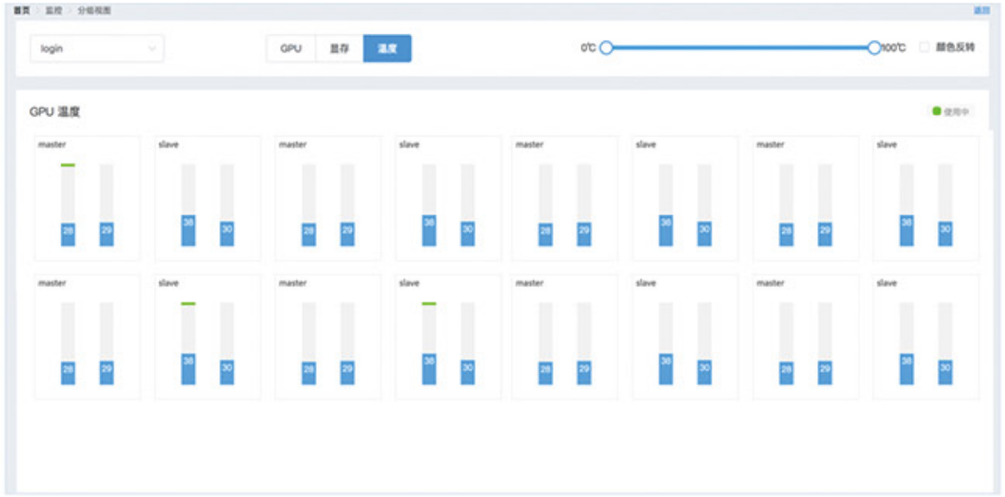
系统提供对逻辑分组整体的 CPU 使用率、Load、内存使用率、温度、能耗、作业负载等指标进行监控。



逻辑分组的监控热力图功能通过颜色深浅来表示组内节点的指标实时值，直观的展示了组内各个节点的运行状况并方便管理员快速定位问题节点。

GPU 视图

系统支持按照集群逻辑分组来查看每个 GPU 的使用率、显存占用率、温度等实时值，帮助管理员直观的了解各个 GPU 的运行状况，并提供自定义的阈值颜色过滤功能，能够快速定位问题 GPU。



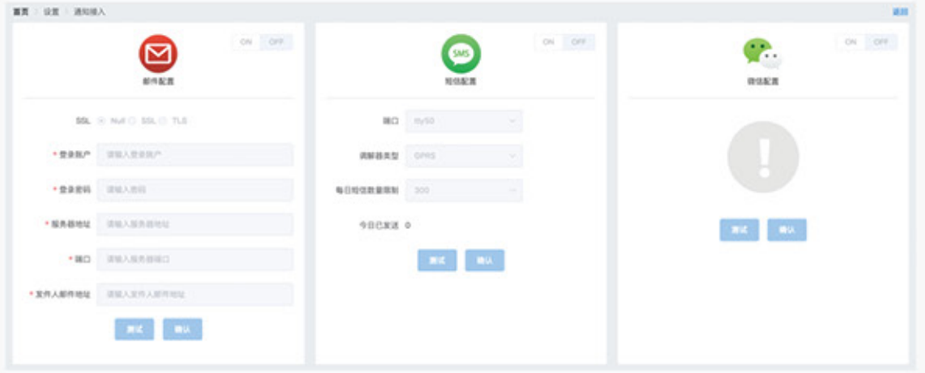
集群报警

报警策略管理

系统提供报警策略管理功能，管理员可以增加、修改、删除策略，并可以灵活的启用、停用。

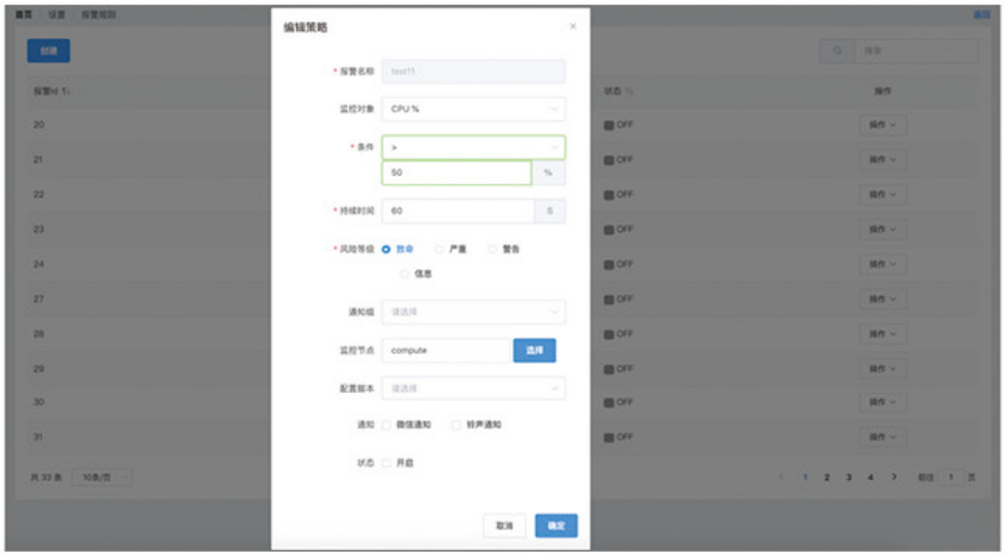
修改策略

支持通过邮件、短信、微信等方式发送报警



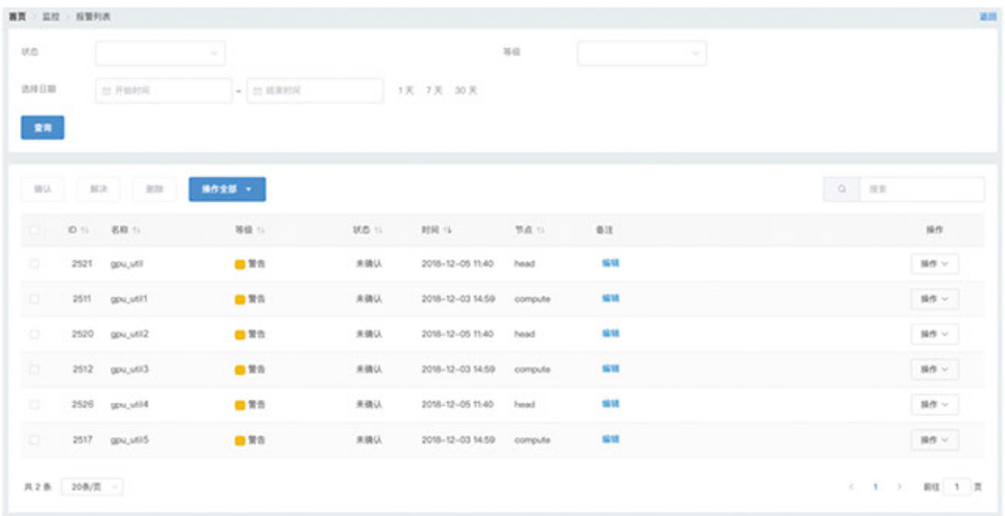
灵活的报警触发机制

系统提供报警策略管理功能，添加报警策略时能自定义报警指标的报警阈值，报警等级，以及触发的邮件或短信通知。



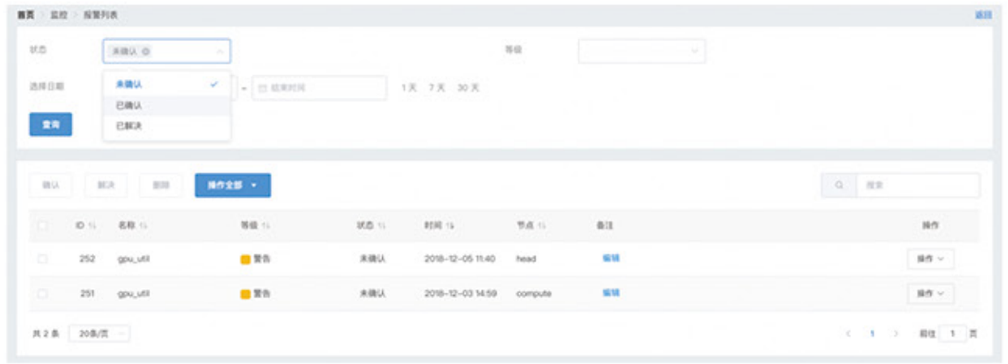
报警查询

系统提供对报警记录的查询功能，并能够按照报警等级、报警时间、报警状态等信息进行查询及排序。



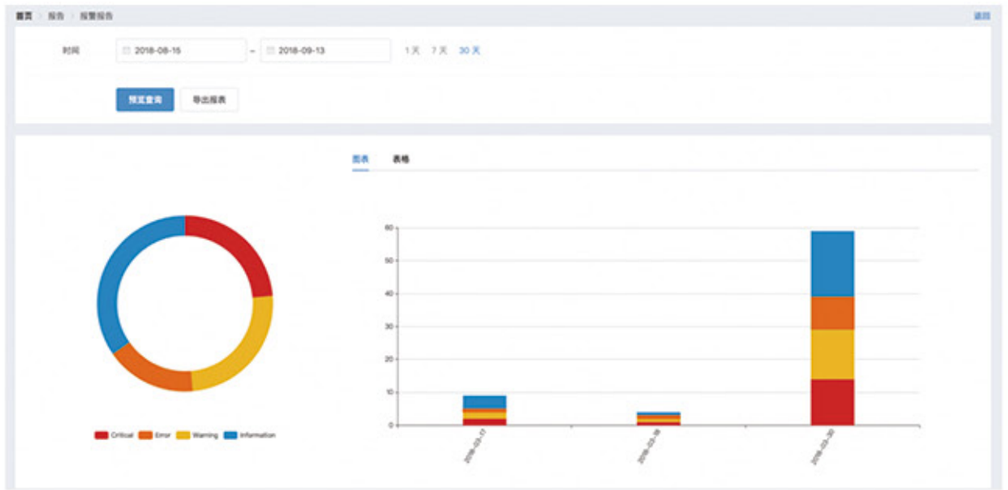
多样的报警处理

系统提供对报警进行确定、解决、删除等状态操作，并可以对报警查询结果进行批量处理。



实时报警查询

系统可以在线查看报警统计表。

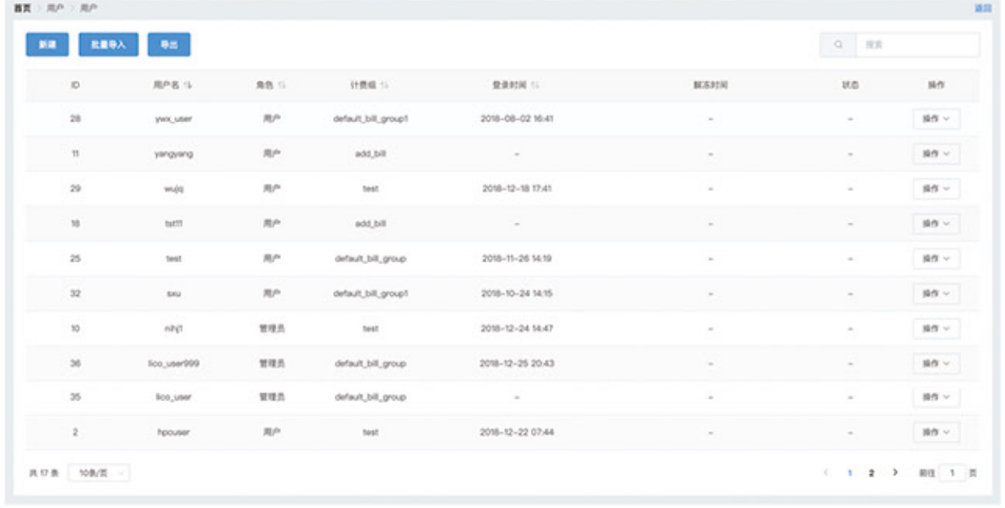


也可以导出统计表和详情表两种报告模板，并根据多种条件进行过滤。

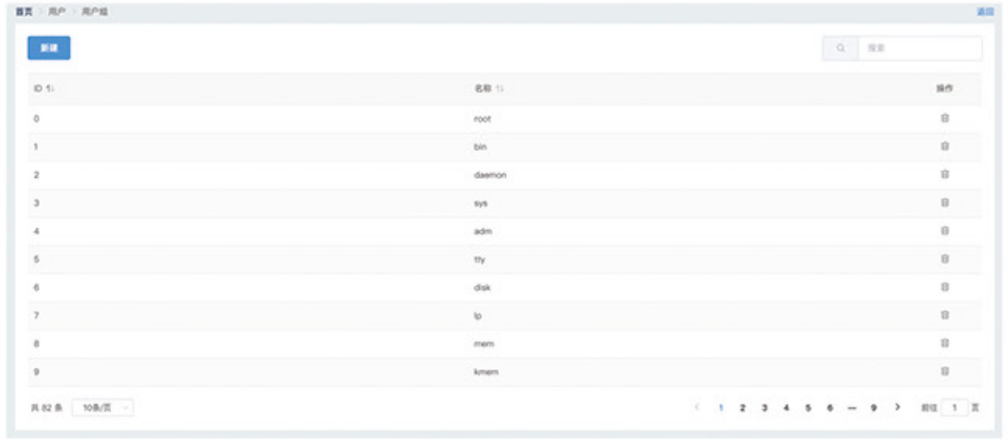
用户管理

用户和用户组管理

系统提供了用户管理功能，可以直接管理本地用户或通过 LDAP 来实现对集群用户的管理。

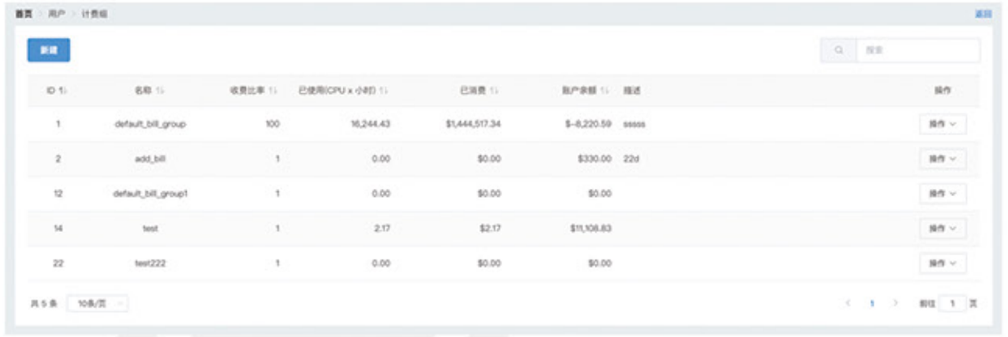


系统提供了用户组管理功能，可以直接管理本地用户组或通过 LDAP 来实现对集群用户组的管理。

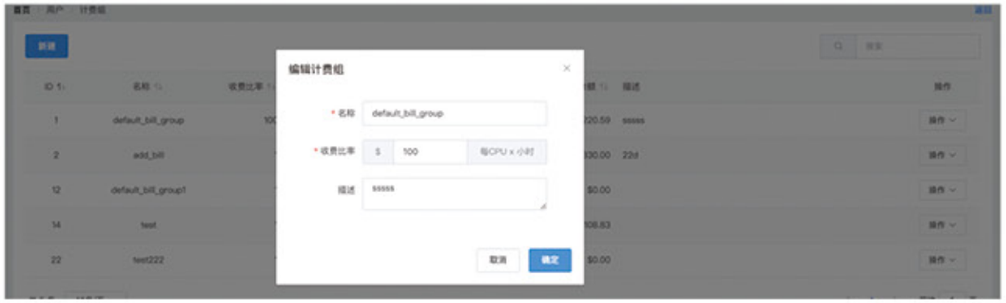


计费组管理

系统提供计费组管理功能，支持查看、新增、删除、修改计费组。



针对不同的计费组可以设置不同的费率。



支持预付款和欠费，系统提供计费组的费用管理功能，管理员可以对计费组的费用账户进行预付费充值、扣款操作。账户支持欠费，在欠费情况下用户也可以提交作业并计费。

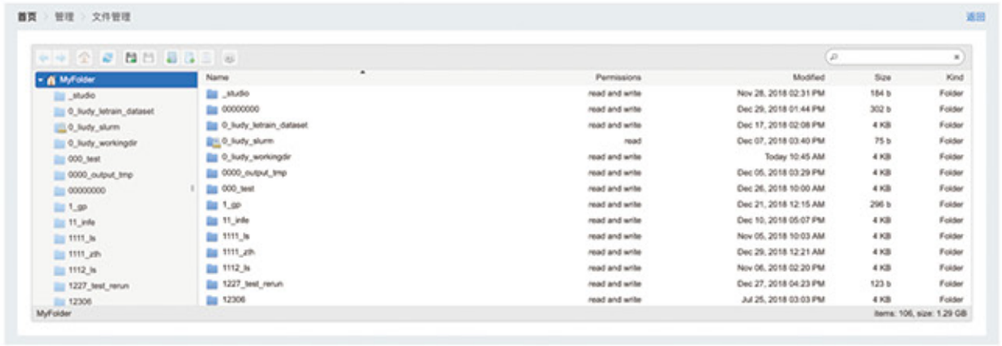
LiCO HPC 功能介绍

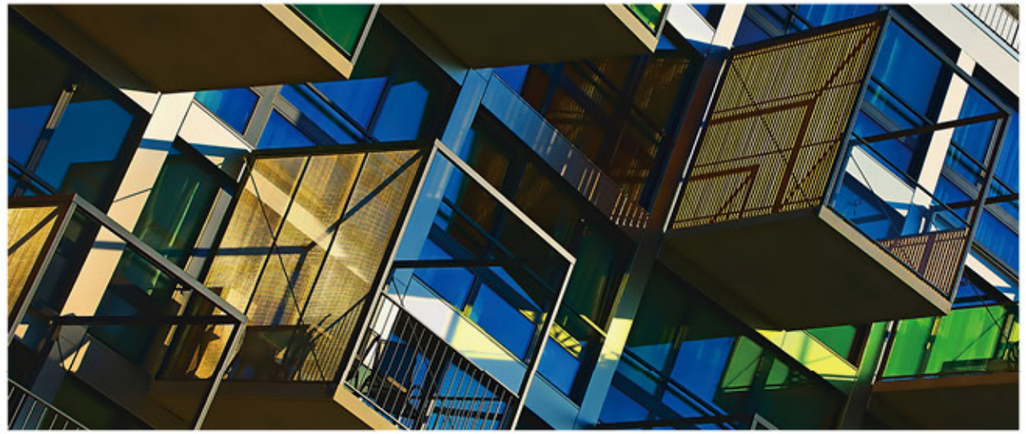
支持多种主流调度器

系统支持多种主流调度器：SLURM、LSF、Torque 等。

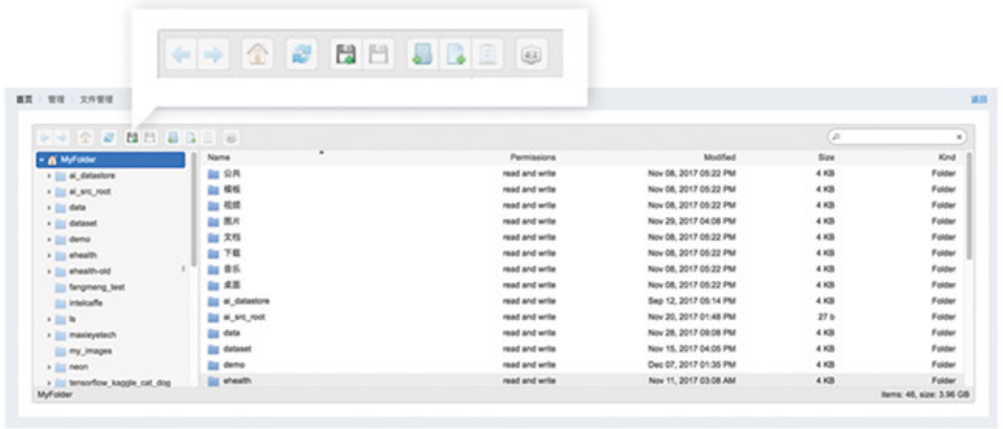
易用的 Web 文件系统

支持多种分布式文件系统，如 Lustre、GPFS 等；支持用户空间隔离，用户通过 Web 可以操作自己空间中的文件。





系统提供基于 Web 的文件管理系统，用户在自有文件系统空间内进行创建文件和文件夹、编辑、删除、上传、下载、重命名等操作。



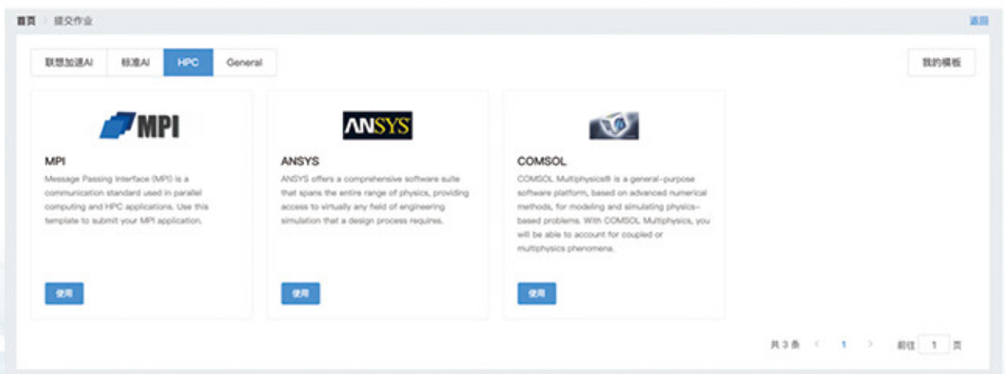
多种作业提交方式

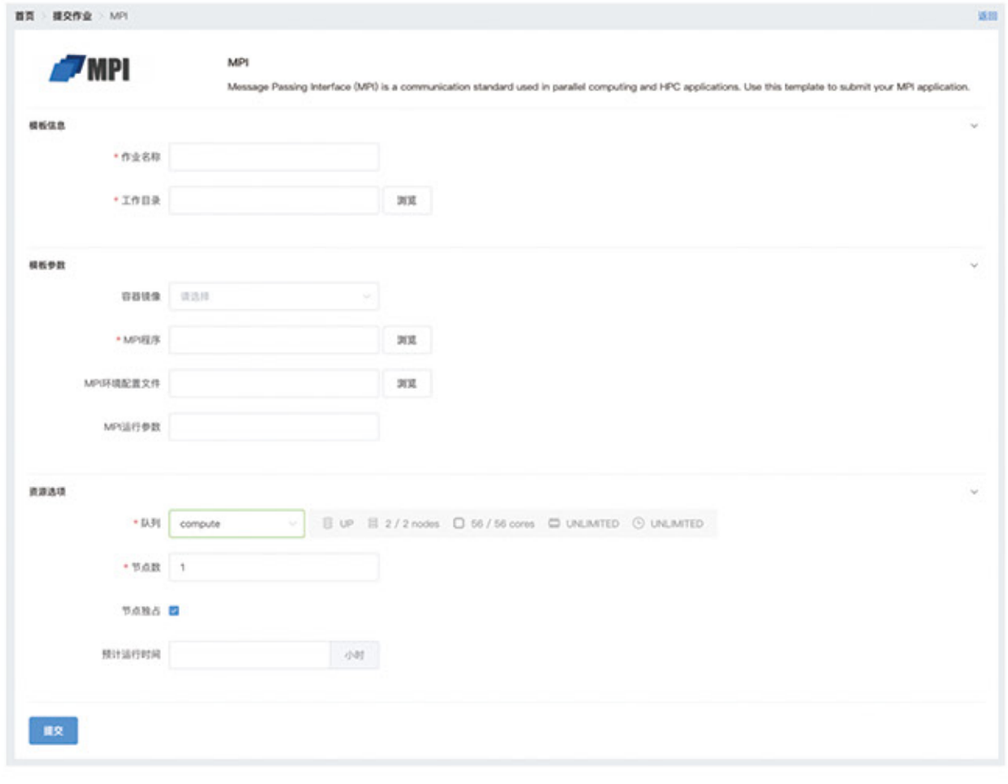
通过命令行提交作业

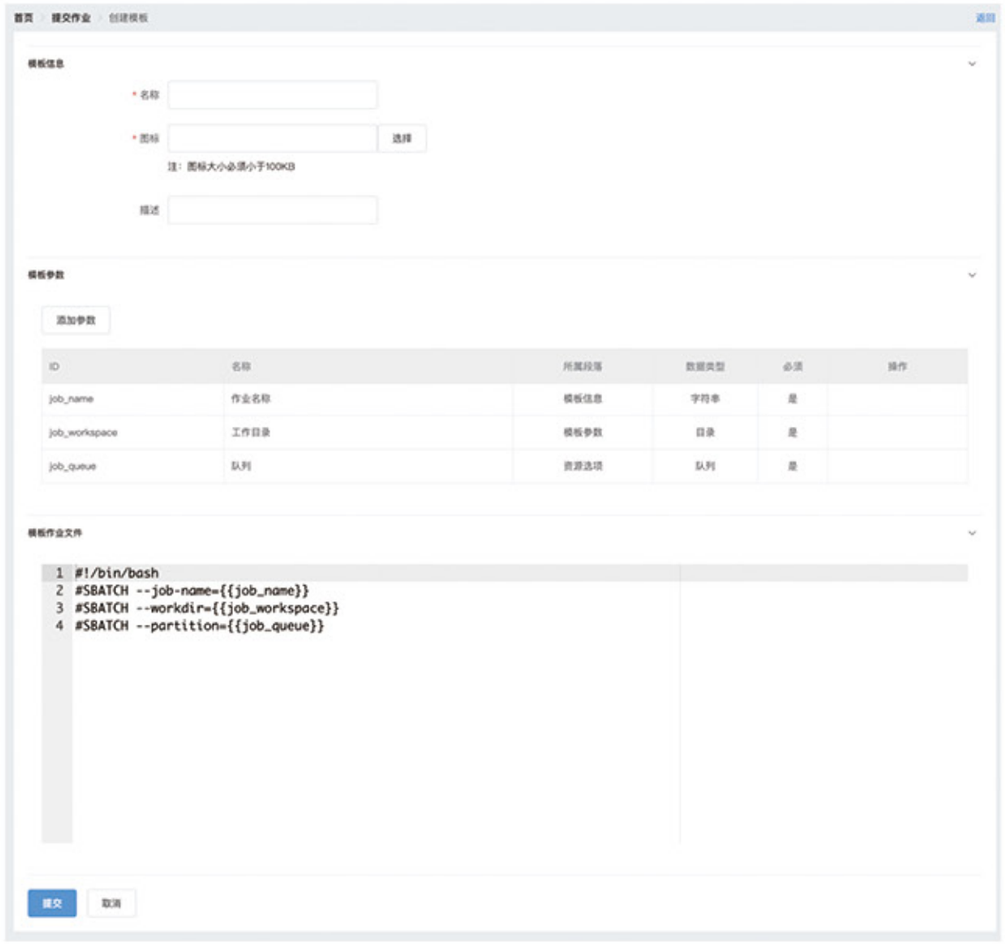
```
[demo@login01 lustre]$ qsub job.pbs
71.mgt
[demo@login01 lustre]$ qstat
Job ID                    Name             User            Time Use S Queue
------------------------- ---------------- --------------- -------- - -----
67.mgt                     job              demo             00:00:15 C batch
68.mgt                     job              demo             00:00:15 C batch
69.mgt                     job              demo             00:00:15 C batch
70.mgt                     job              demo             00:00:15 C batch
71.mgt                     job              demo                    0 R batch
[demo@login01 lustre]$
```

通过系统自带的作业模板来提交作业

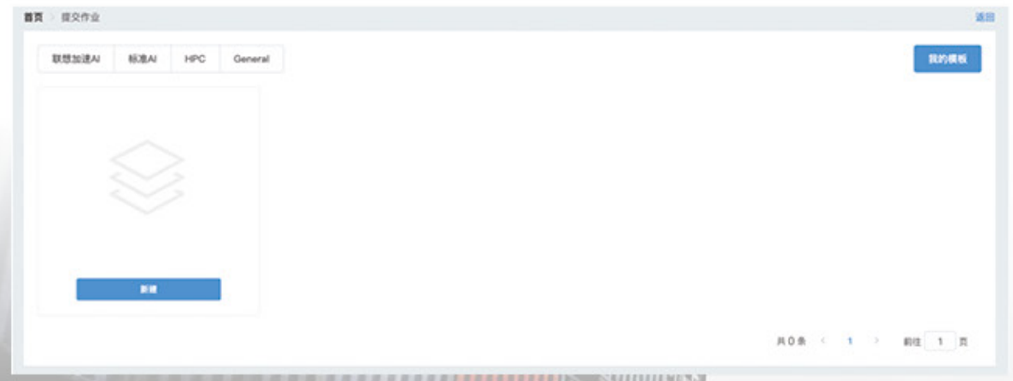
系统的模板商店功能提供了多种面向不同应用的作业模板，方便用户提交作业。

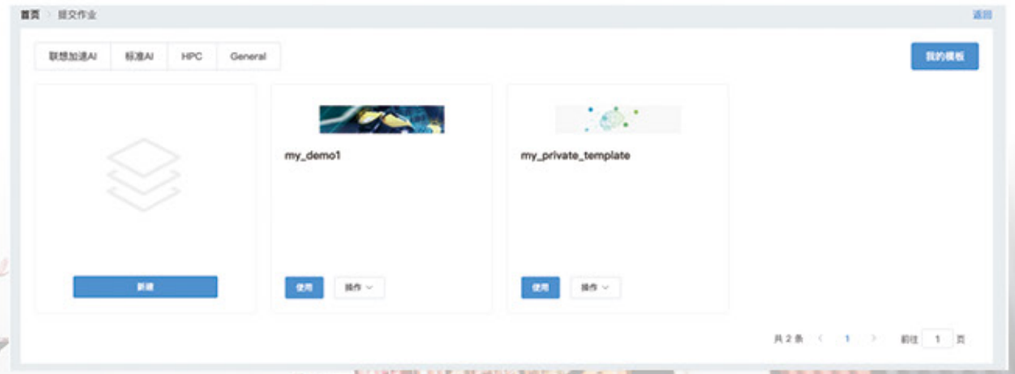


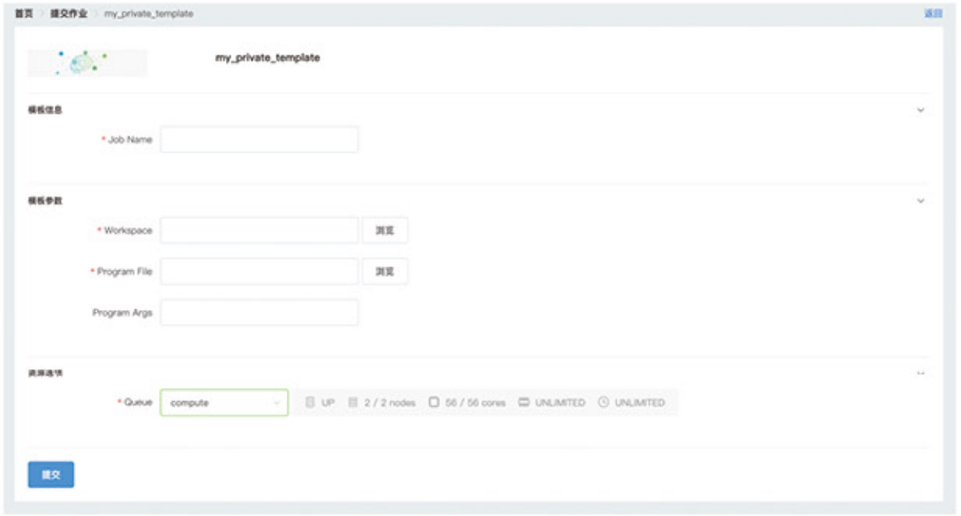




创建定制化模板并通过定制化作业模板提交作业



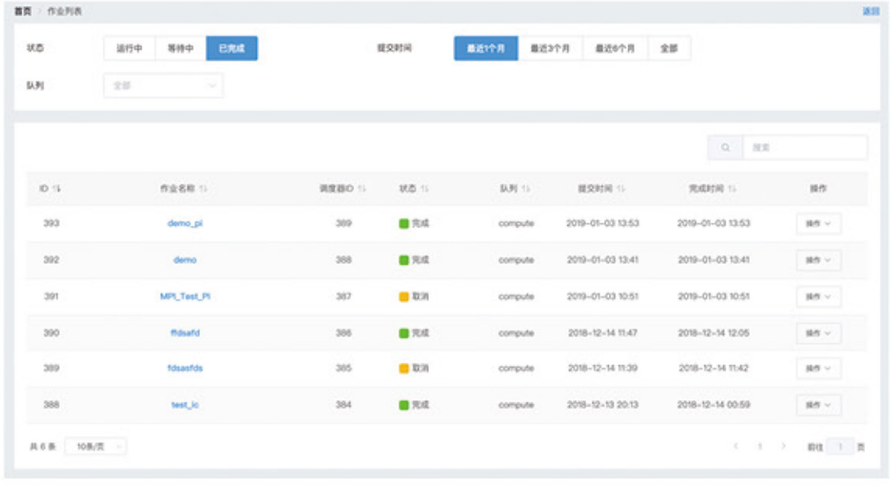




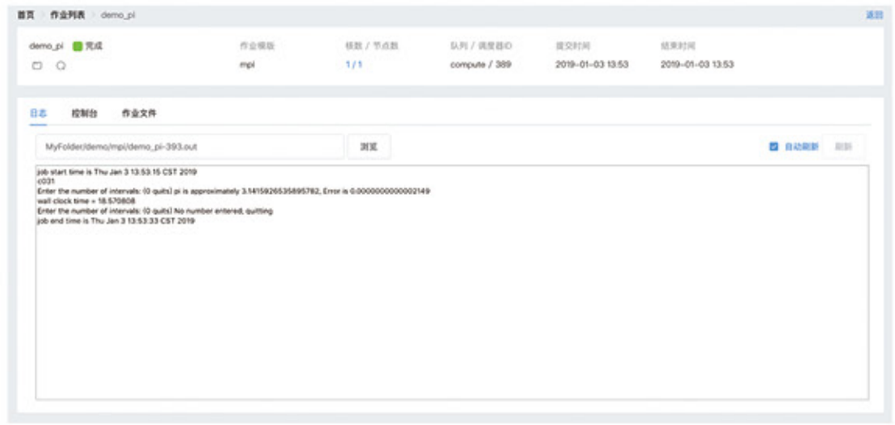
作业管理和监控

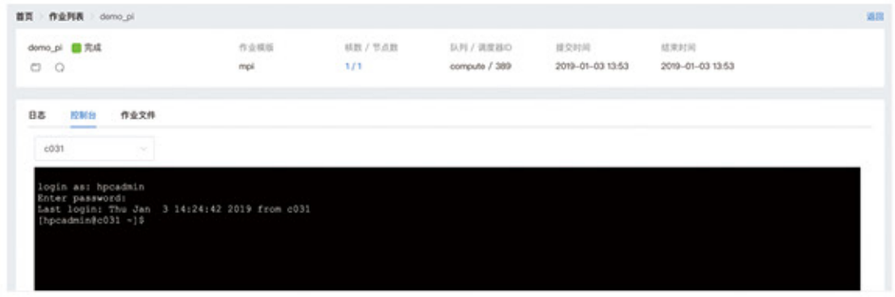
通过多种方式提交的作业，可以在 LiCO 界面上进行统一的管理

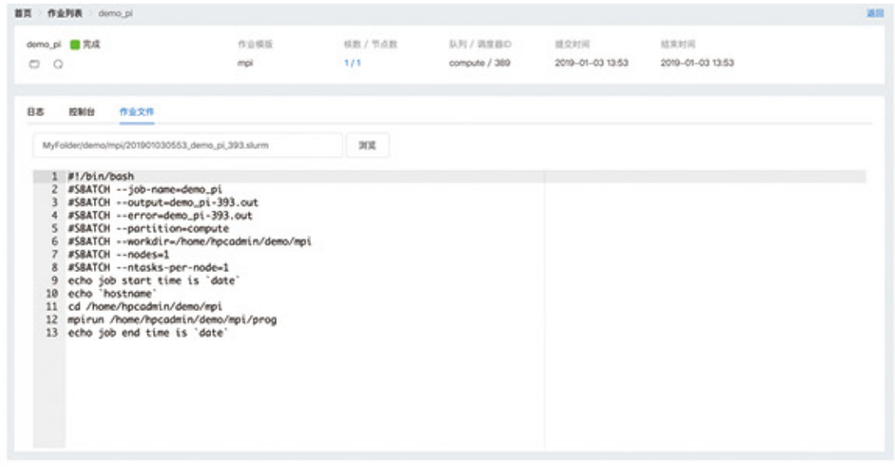
用户可以通过作业管理功能查看当前正在排队、运行、完成的作业。



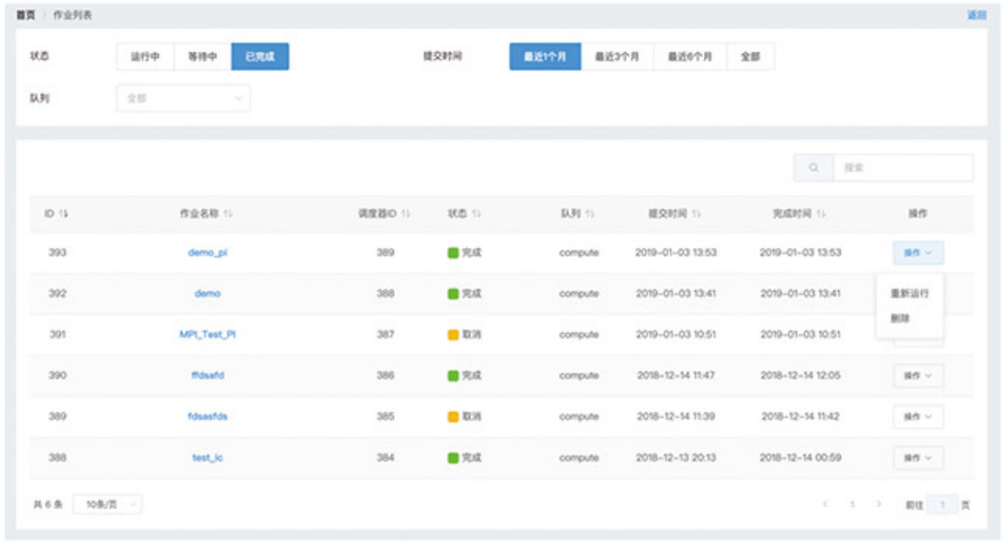
作业详情功能提供了作业运行日志的监控、作业运行状态的查看和作业文件查看等。







用户可以取消、删除、重新运行作业。

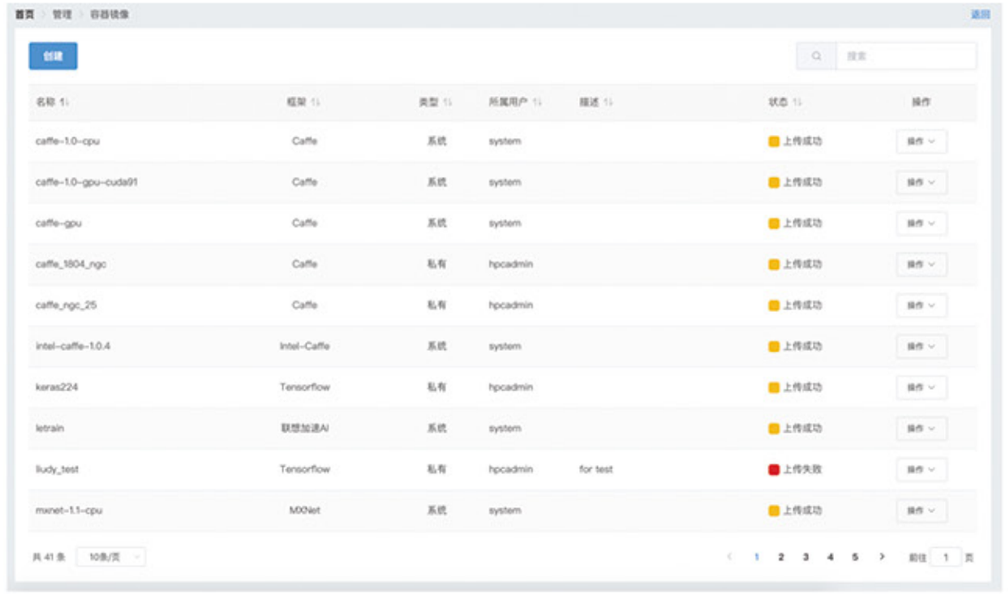


LiCO AI 功能介绍

- 平台提供统一资源调度，在一个集群中，支持 HPC 作业和 AI 作业同时运行。
- 平台支持 CPU, Xeon Phi, GPU 等多种处理器，支持 GPU 使用率、GPU 显存和 GPU 温度的监控。
- 平台支持多种 AI 框架和分布式训练，包括但不限 TensorFlow, Caffe, Intel-Caffe, Neon，MXNet 等。
- 平台提供镜像管理，系统提供预设镜像，同时支持用户创建和使用自定义镜像运行作业。
- 使用自写程序进行模型训练：用户提供程序和数据，使用系统自带作业模板（单机，多机分布式）提交 AI 作业并监控作业运行情况。
- 使用联想加速 AI 进行模型训练：使用联想加速 AI 的自带模型，用户不需要编写程序，可以多机分布式进行 AI 模型训练，并使用训练好的模型进行推理。
- 使用 AI Studio 进行模型训练：面向计算机视觉，提供图片分类、物体识别、实例分割 场景的端到端模型训练。

镜像管理

AI 作业均运行在 Singularity 容器下，系统提供了对容器镜像的管理功能，用户除了可以使用系统提供的基础镜像之外，也可以上传自定义的私有镜像。

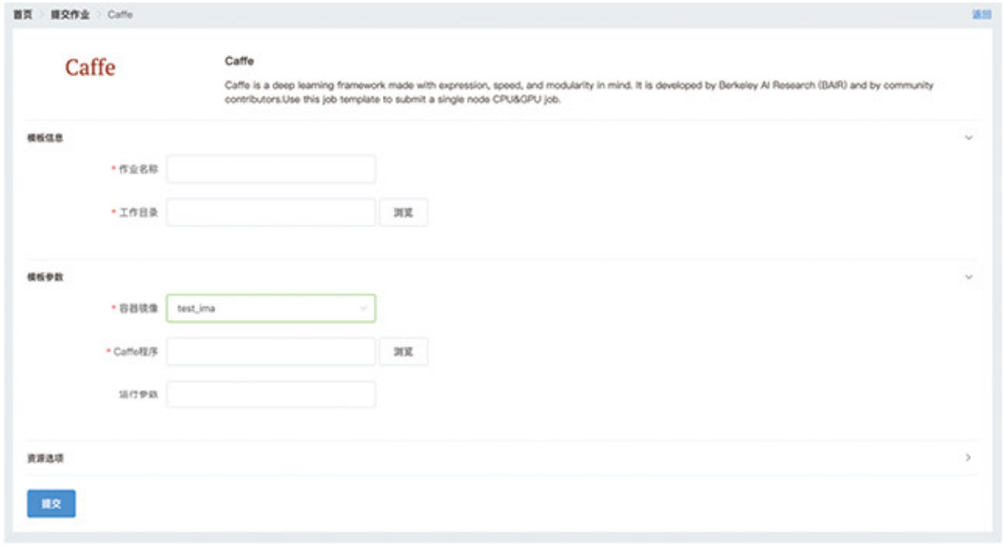


使用用户自写程序训练模型

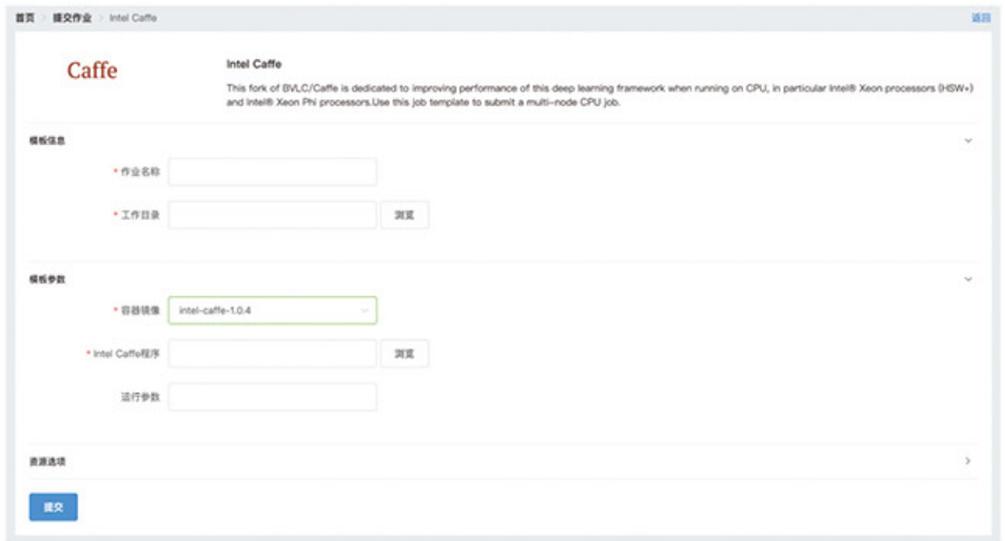
标准 AI 作业模板提供了常用 AI 框架的作业模板，用户提供基于不同 AI 框架编写的模型训练程序，然后通过作业模板方便的提交作业来运行其程序。

Caffe 及 Intel Caffe 作业

用户可以通过 Caffe 作业模板提交在单机运行的 Caffe 程序，支持使用 CPU 或 GPU。

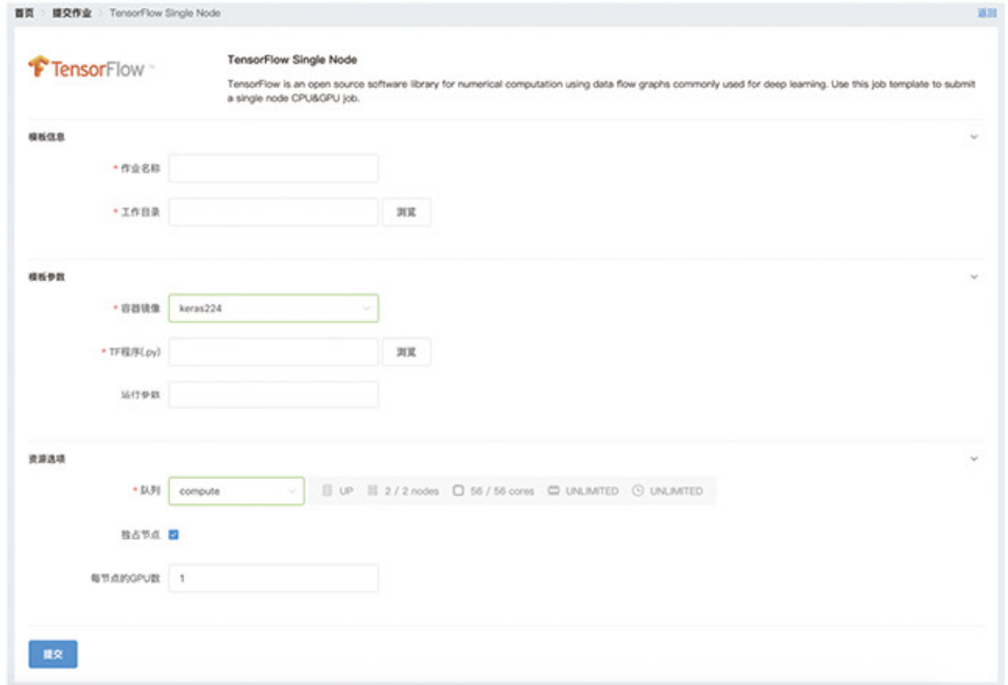


用户可以通过 Intel Caffe 作业模板提交单机或分布式运行的 Intel Caffe 程序，Intel Caffe 是基于 MKL/MKL-DNN 优化的，通过分布式运行可以在 Xeon、Xeon Phi 上获得理想的训练速度。

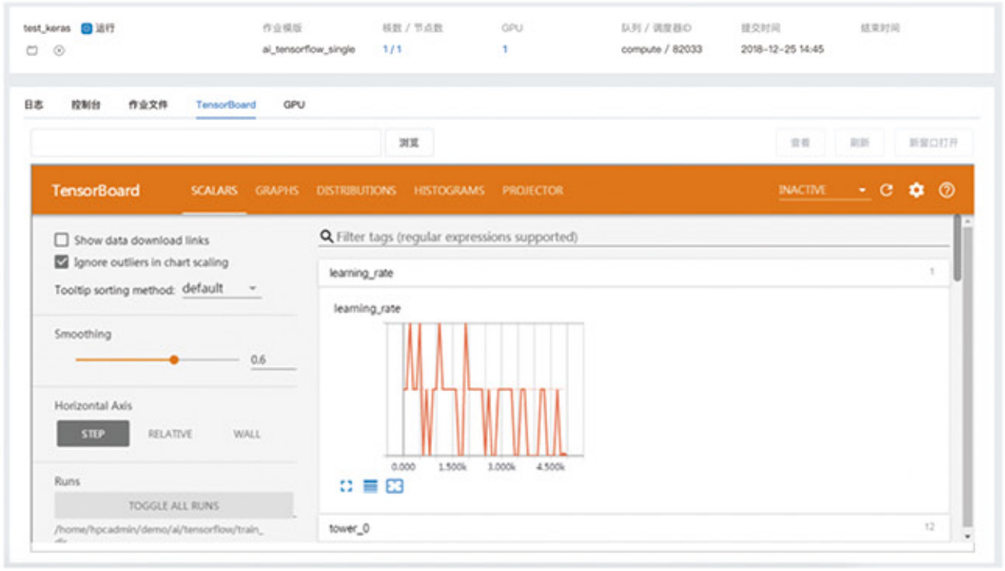


TensorFlow 作业

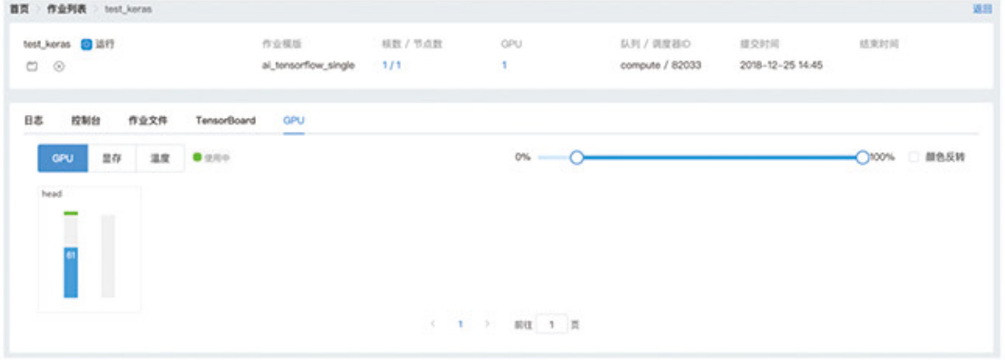
系统提供 TensorFlow 单机及分布式两种作业模板，两种模板均支持使用 CPU 或 GPU。



系统内置 TensorBoard 可以对 TensorFlow 程序运行过程进行图形化监控。

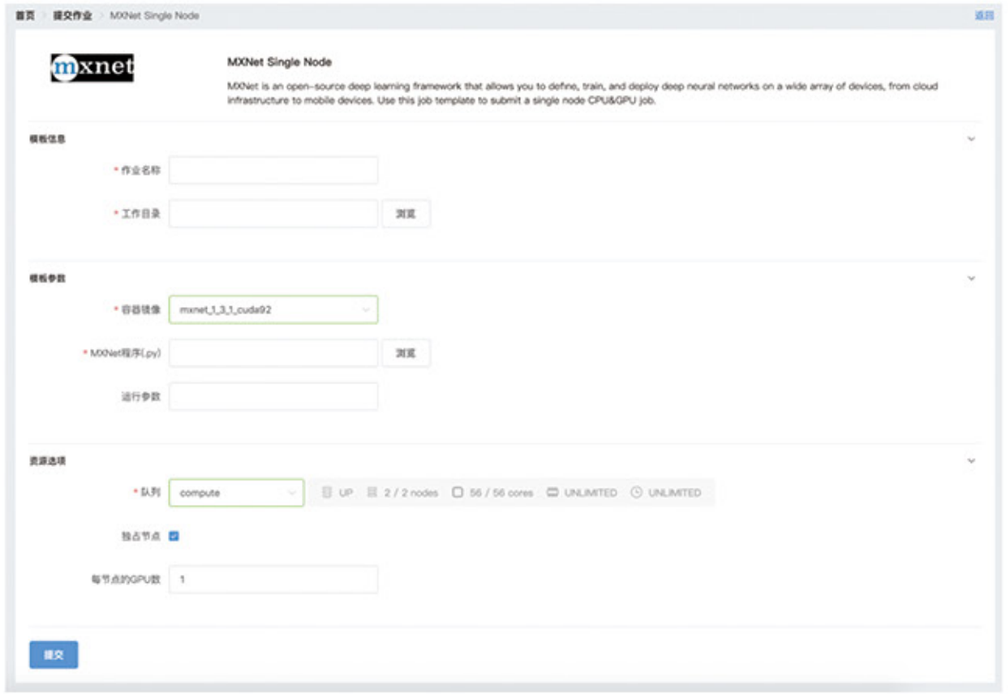


可以在作业列表中查看作业进展，GPU 使用情况。



MXNet 作业

系统提供 MXNet 单机及分布式两种作业模板，两种模板均支持使用 CPU 或 GPU。

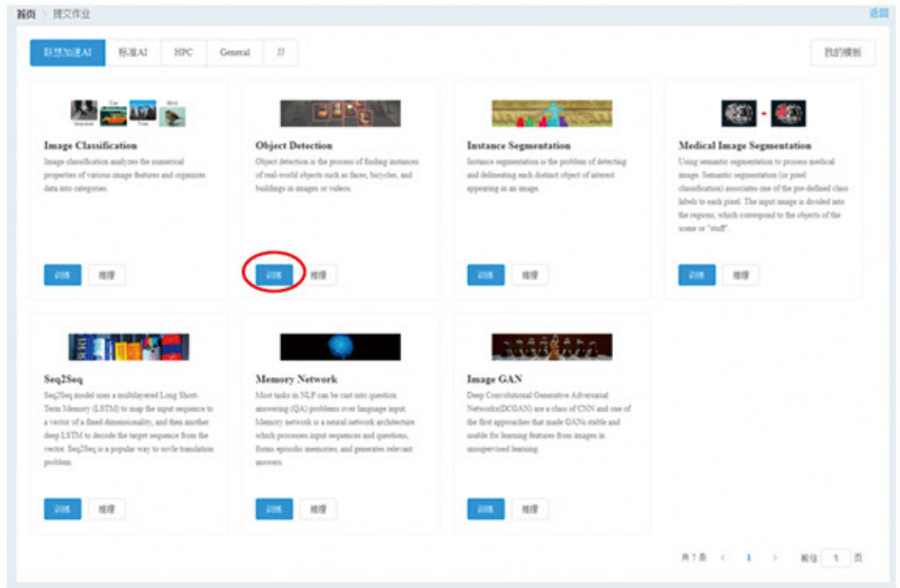


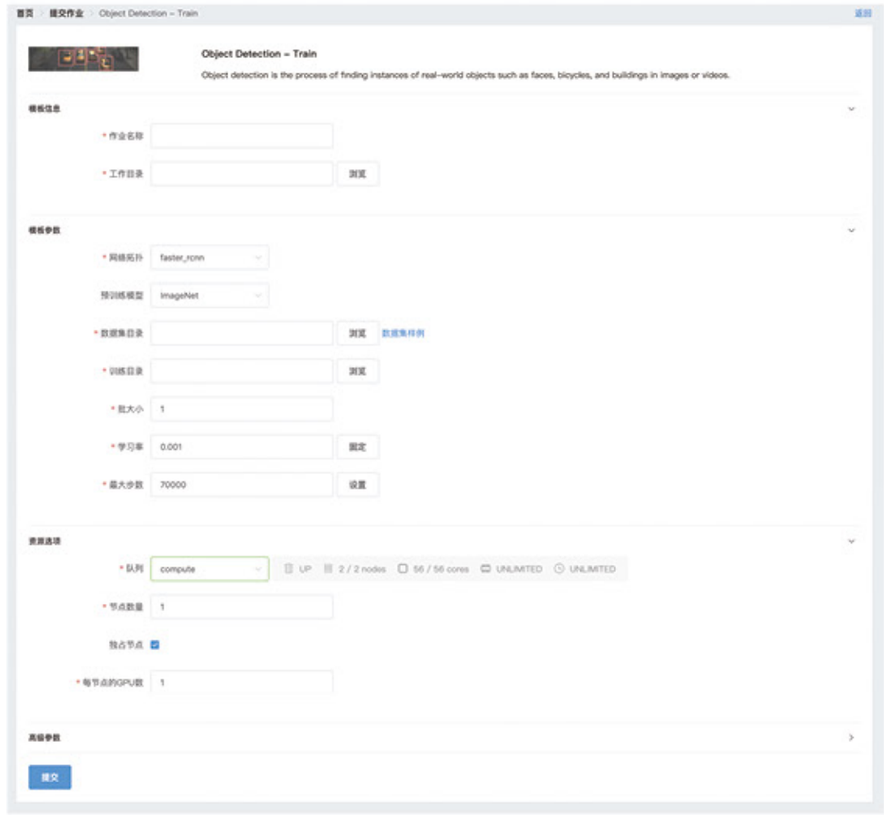
联想加速 AI

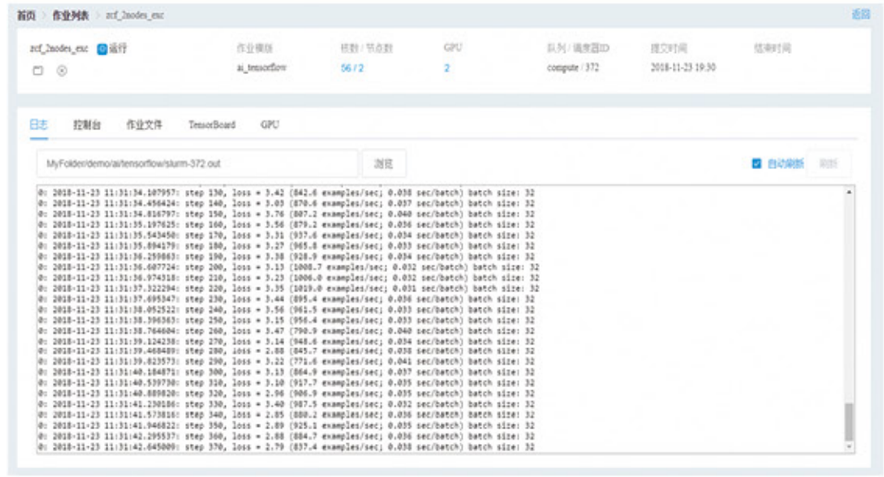
联想加速 AI 提供了图片分类，物体识别，实例分割 ，医疗图像分割，seq2seq，自然语言问答，对抗网络等 7 大类 AI 应用。每大类应用集成了常用的模型，比如 Lenet, AlexNet, Googlenet，VGG, ResNet, Inception, Unet, Faster R-CNN, Mask R-CNN，Memory Network, Seq2Seq, GAN 等。用户不需要编写程序，只需要提供数据，就可以通过 LiCO 提供的模板直接使用这些模型进行训练。联想加速 AI 训练好的模型，用户可以直接使用 LiCO 提供的模板进行预测，不需要编写程序。

下面是一个使用联想加速 AI 进行物体识别模型的训练，并使用训练好的模型进行预测的例子：

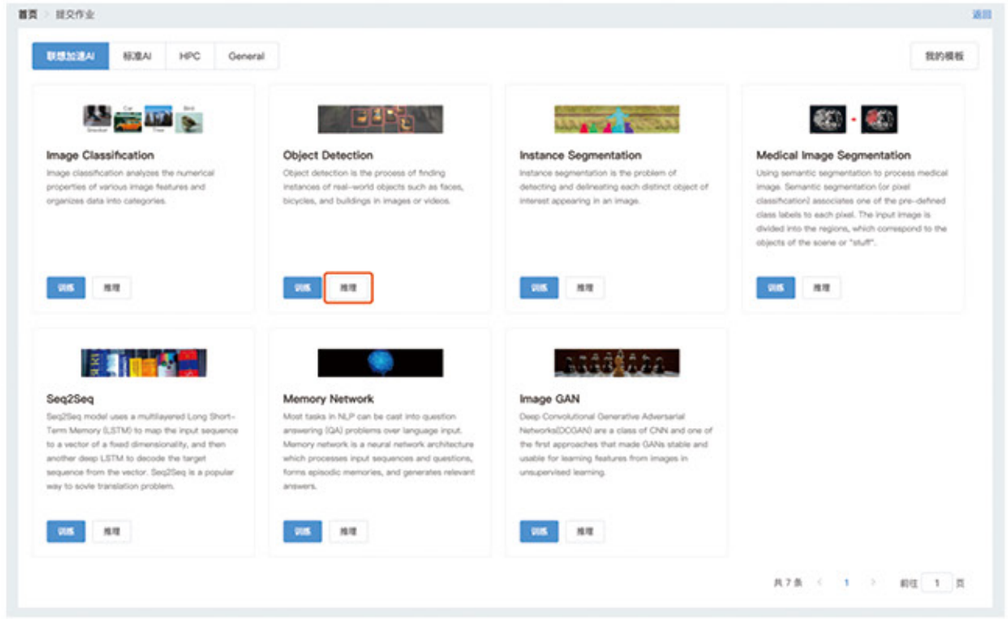
- 通过物体识别的模板提交作业，用户只需要提供数据，然后选择一个系统自带的模型就可以开始进行模型的训练。

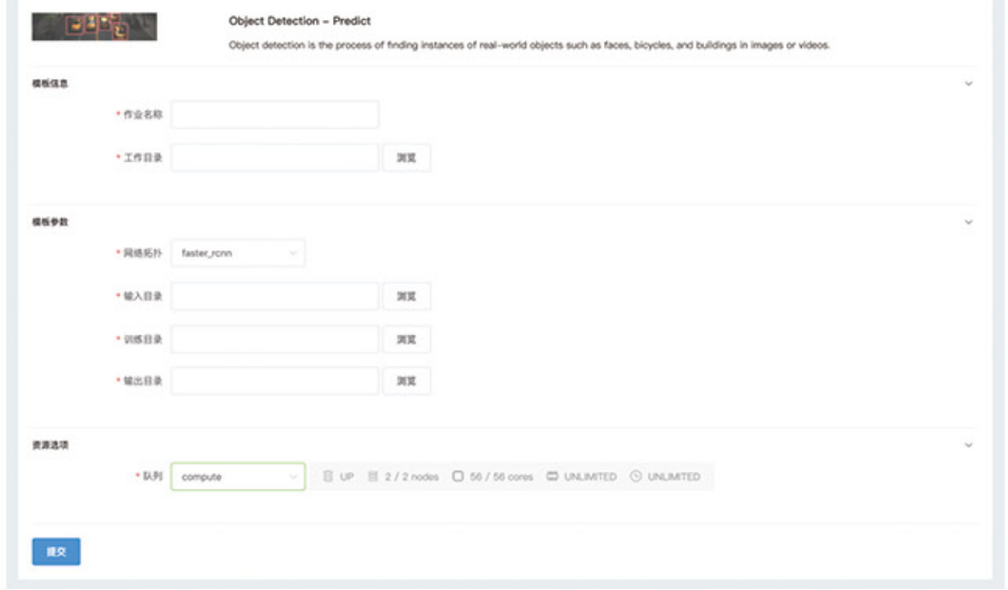


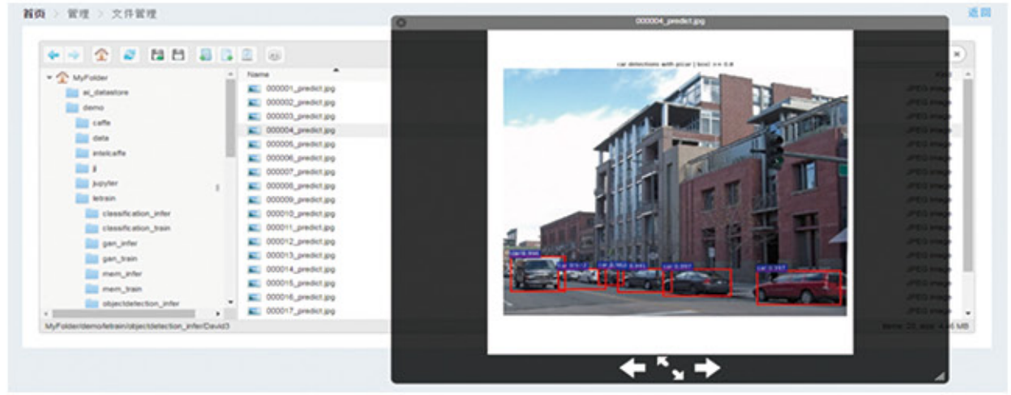




- 模型训练完成后使用系统自带的推理模板进行预测，用户上传需要预测的图片，选择已训练好的模型进行预测。

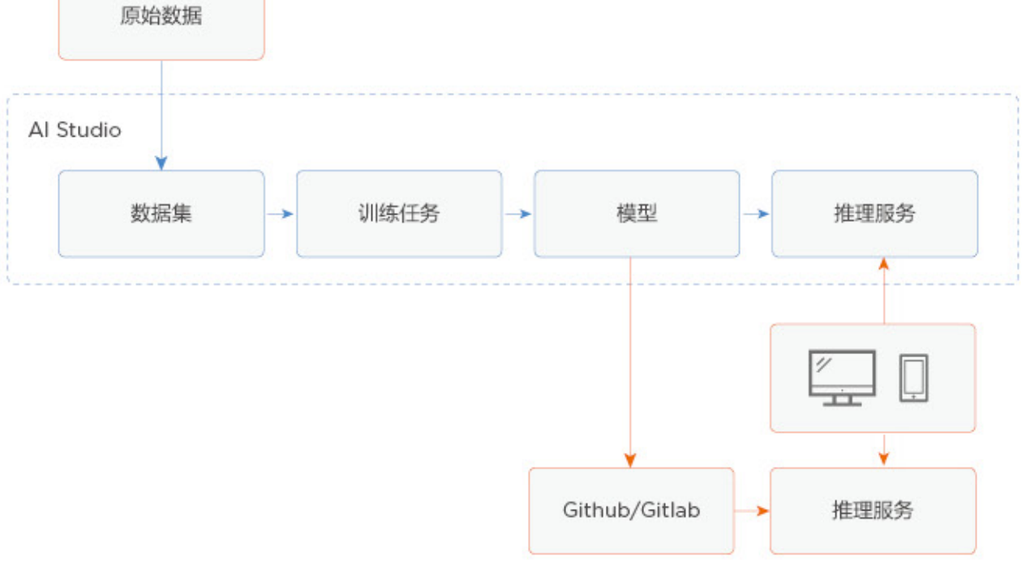






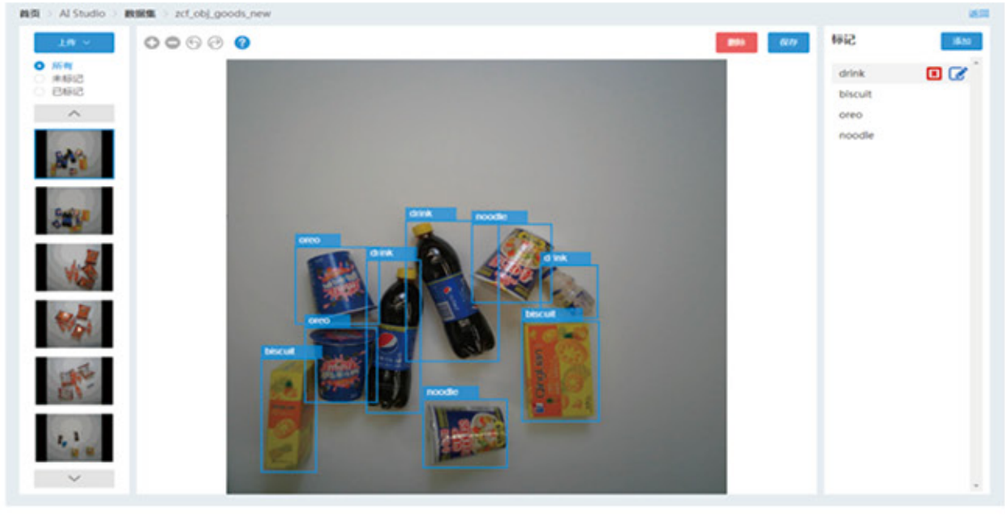
AI Studio 端到端的模型训练

面向计算机视觉三种场景：图像分类、物体识别、实例分割 ， AI Studio 提供了端到端的模型训练。



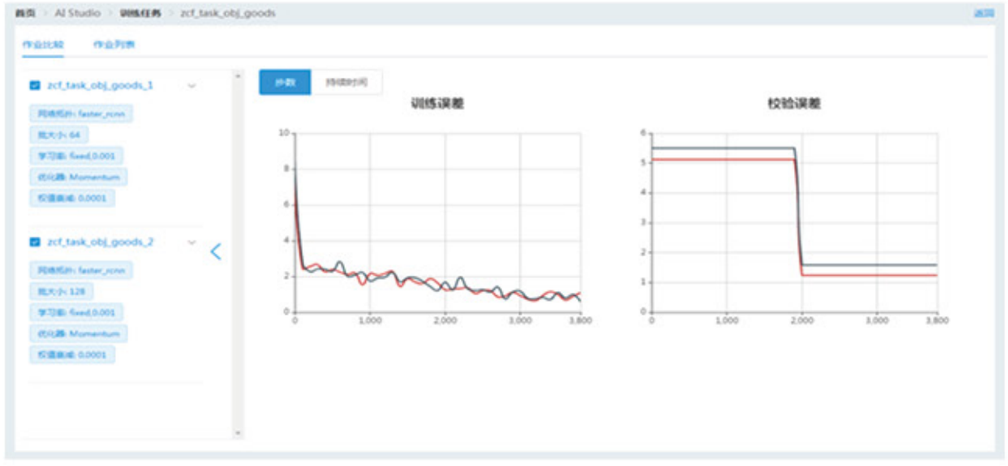
数据集管理

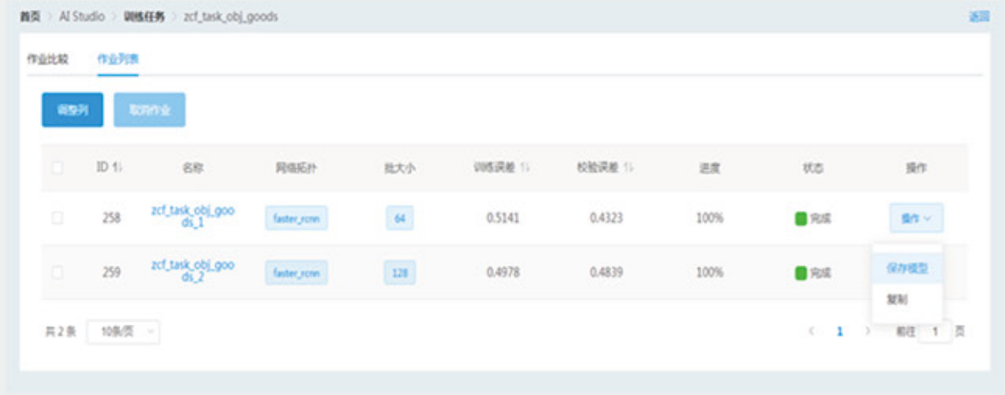
数据集管理支持图片分类、物体识别、实例分割 三种场景的数据集。支持数据集在线标注和生成数据集， 也支持导入格式匹配的数据集。



训练任务

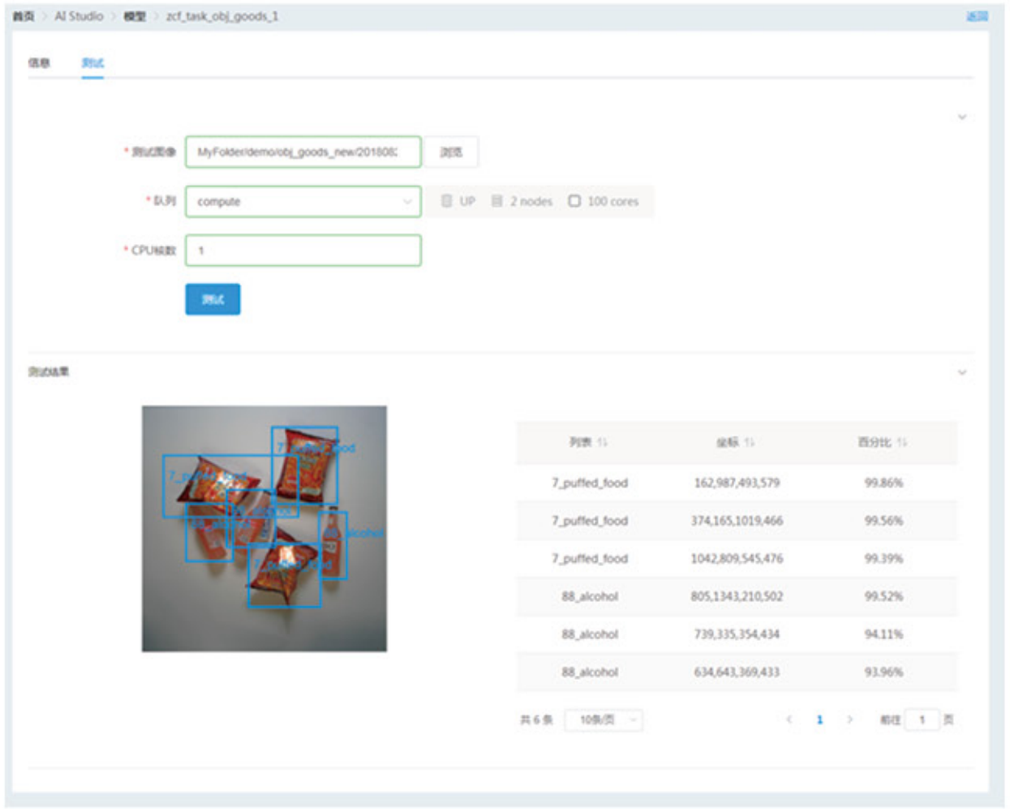
模型训练是一个不断调节超参尝试，训练最优模型的过程。一个训练任务可以批量发起多个作业，一个作业对应于一组超参的组合，大大加快模型训练的过程。训练好的模型可以进行保存。

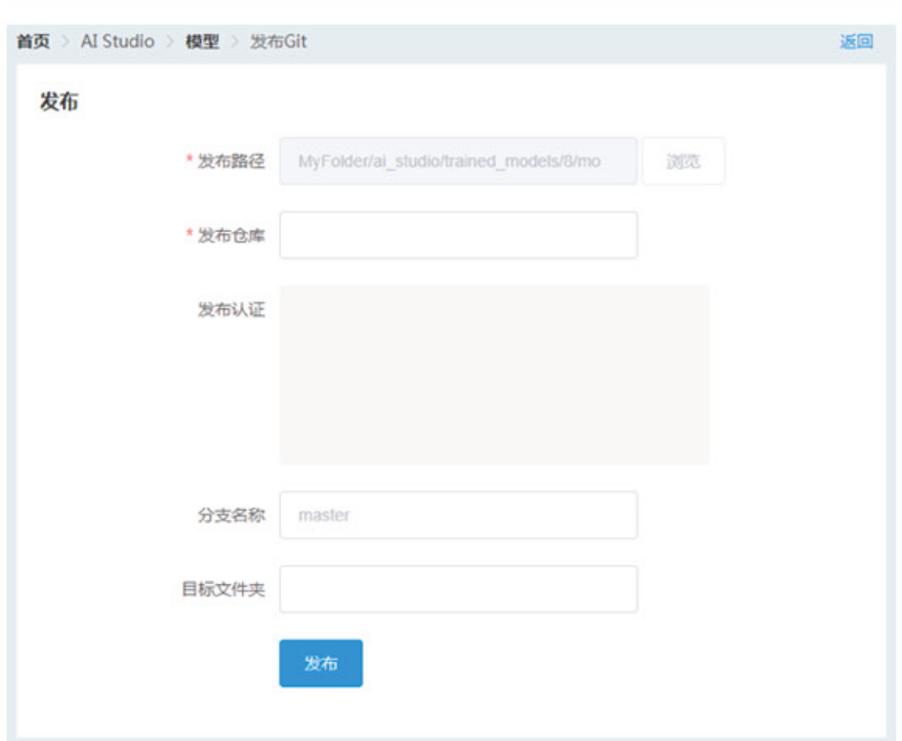




模型

列出保存的模型，可以输入图片测试模型的预测结果，支持将模型部署为推理服务或者将模型发布到 Github/Gitlab。





推理服务

- 推理服务列出了上一步部署的服务，这些服务提供 Web API，外部程序通过调用 Web API 实现图片的预测。

```python
import requests
import base64
test_image = "D:\\20180824-13-43-21-401.jpg"
with open(test_image, 'rb') as f:
    content = base64.b64encode(f.read())

data = {
    "image": content,
    "image_type": "BASE64"
}
api_key = "30tsMG8by-4x28WRcifPLBY55SPSfdV28gms3niHzDU="
header = {
    'Accept': 'application/json',
    'authorization': 'token ' + api_key
}
res = requests.post(
    "https://10.240.208.102/openapi/v1/service/0211bc8e5a1111e9a8ec7cd30ade3d48/",
    headers=header,
    json=data,
    verify=False
)
print res.json()
```

LiCO 开放 API

LiCO 开放 API 提供如下接口：

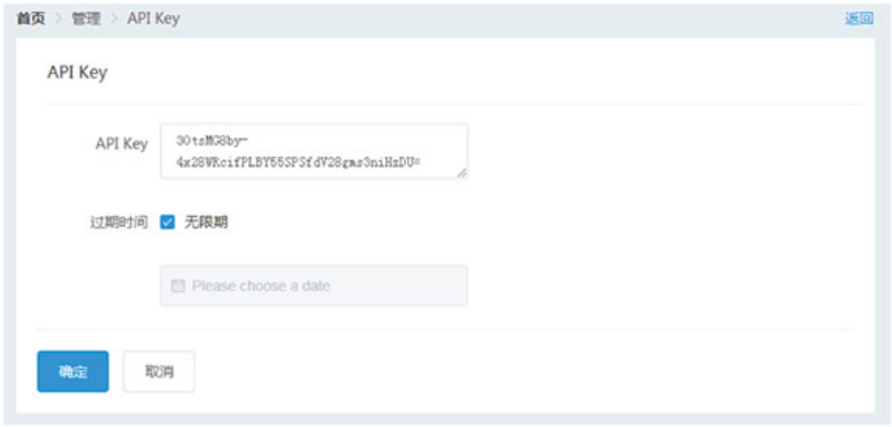
提交作业（HPC 作业和 AI 作业）的 API

模型发布到 github/git-lab 的 API

模型推理服务的 API

使用 LiCO 开放 API 可以方便实现 LiCO 与现有其它系统的集成，LiCO 开放 API 的使用方法：

使用 LiCO 生成 API Key。



使用 API Key 调用 API， API 接口提供了 hook，可以控制在作业开始或者完成时回调 hooks 提供的 url。

```python
import json
import requests

url = "https://10.240.208.102/openapi/v1/job/submit/"
api_key = "30tsMG8by-4x28WRcifPLBY55SPSfdV28gms3niHzDU="
headers = {
    "Authorization": "token "+api_key,
    'Content-Type': 'application/json'
}
job = {
    "template_id": "letrain_instance_segmentation",
    "parameters": {
        "job_name": "shape_segmentation",
        "job_workspace": "MyFolder/demo_0228",
        "model_name": "maskrcnn",
        "data_dir": "MyFolder/demo_0228/dataset",
        "train_dir": "MyFolder/demo_0228/train_dir",
        "batch_size": 1,
        .
        .
        .
    },
    "hooks": [
        {
            "type": "complete",
            "url": "http://10.104.207.106:8080/generic-webhook-trigger/invoke?token=train_finish",
        }
    ]
}
res = requests.post(
    url=url, headers=headers, data=json.dumps(job), verify=False
)
print(res.status_code)
print(res.content)
```

3 联想高性能计算产品及特点介绍

联想高性能计算硬件和环境配套

高性能计算服务器产品

（1）Lenovo ThinkSystem SD530 高密度刀片服务器

Lenovo ThinkSystem SD530 是一款采用 0.5U 机架、密度极高、经济实惠的新型刀片双路服务器。在 ThinkSystem D2 刀片机箱或 ThinkSystem 模块化刀片机箱中安装四台 SD530 刀片服务器时，您将获得一个用于高性能计算，人工智能和处理企业和云工作负载的理想高密度 2U 四节点（2U4N）刀片平台。

2U4N 系统已在众多数据中心（包括大型企业和服务提供商）得到普及，因为它们占地面积较小，密度较高，因此非常适用于以较低成本构建基于解决方案的设备。Lenovo ThinkSystem SD530 与 D2 刀片机箱的组合旨在提供这些类型的解决方案。是替代传统的刀片中心的最佳解决方案。

建议用途：云、MSP、CSP、HPC、AI 和超融合解决方案、分支机构或远程办公室

下图显示了安装在 D2 刀片机箱中的四台 ThinkSystem SD530 服务器。

图 安装在 D2 刀片机箱中的四台 ThinkSystem SD530 服务器

相关信息

SD530 组合了刀片的效率和高密度与机架式服务器的性能和简便性。它旨在运行搭载最多内核的至强铂金级别处理器，可处理要求最严苛的 HPC/ 技术计算 /AI 工作负载。

此外，通过添加 GPU 托盘，SD530 还可支持两个高性能 GPU。此服务器支持一系列不同的 NVIDIA GPU。

主要特性

ThinkSystem SD530 高密度刀片产品在 ThinkSystem D2 刀片机箱或 ThinkSystem 模块化刀片机箱中安装了四台热插拔 SD530 服务器。每个这样的刀片机箱仅占用 2U 空间（每台服务器 0.5U），并且为配置大量内部存储留出了空间。由于采用了整体化设计，这款解决方案的价格极为低廉，总体拥有成本（TCO）较低。

可扩展性和性能

SD530 服务器和刀片机箱提供了各种特性，可提高性能、增强可扩展性并降低成本：

一个 2U 刀片机箱最多支持四个节点，每个节点搭载两个英特尔® 至强® 可扩展家族处理器，最多配置 16 个 DIMM，6 个驱动器托架和两个 PCIe 插槽。这是一款高度密、可扩展、成本低廉的服务器。

- 支持一系列英特尔® 至强® 可扩展家族处理器 - 旨在运行搭载最多内核的、成本高效的英特尔® 至强® 铂金级别处理器。
- 处理器最多支持 28 个内核，内核速度高达 3.6 GHz，TDP 额定功率最高为 165W。
- 每台服务器最多支持两个处理器、共 56 个内核和 112 个线程，可最大限度地提高多线程应用的并行执行性能。刀片机箱中安装四个节点时，仅 2U 的机架空间总共可支持 224 个内核。
- 采用英特尔睿频加速技术 2.0（Intel Turbo Boost Technology 2.0）提供智能、适应性的系统性能，可允许 CPU 内核在峰值负载期间以最高速度运行，暂时超出处理器热设计功率（TDP）。
- 英特尔超线程技术（Intel Hyper-Threading Technology）可在每个处理器内核中支持同步多线程（每个内核最多两个线程），从而提高多线程应用的性能。
- 英特尔虚拟化技术（Intel Virtualization Technology）集成了硬件级虚拟化嵌套，便于操作系统供应商更合理地将硬件用于处理虚拟化工作负载。

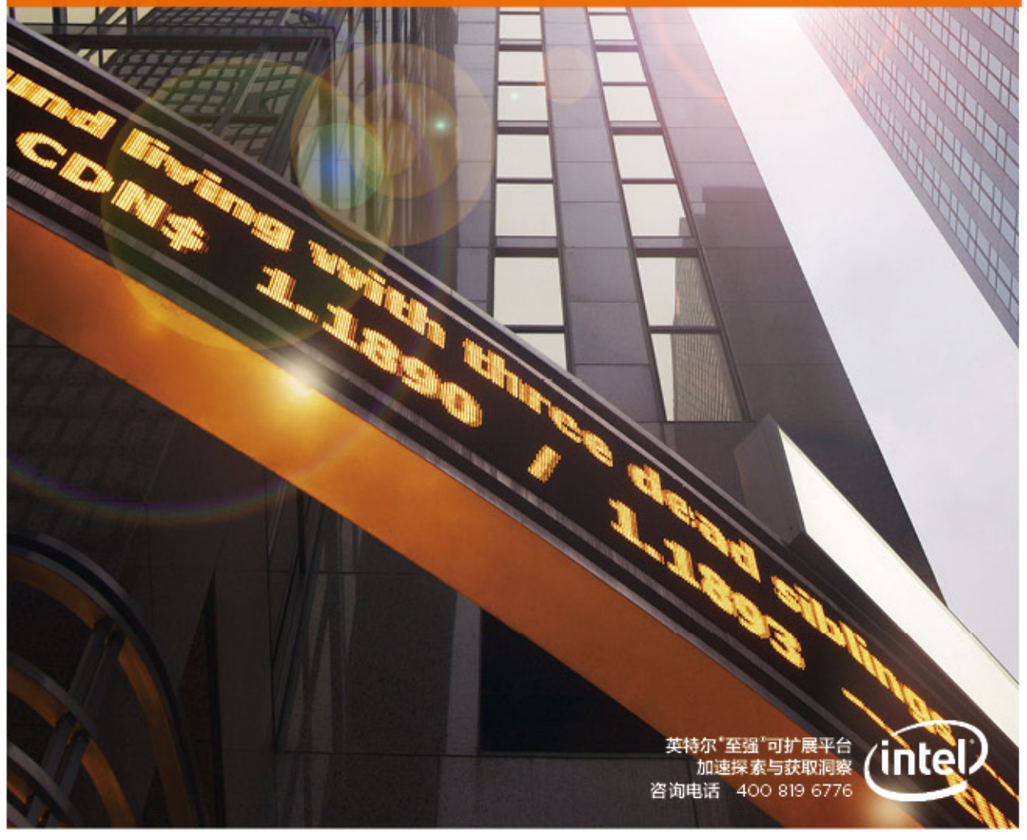
- Intel Advanced Vector Extensions 512（AVX-512）可用于加速企业级工作负载，包括数据库和企业资源规划。
- 每个处理器具有六条内存通道，内存速度高达 2666 MHz，可最大限度地提高系统性能。
- 最多支持 16 个 DIMM，可最大限度地提高内存容量，进而支持 1 TB（使用 16 个 64 GB LRDIMM）或 1.5 TB（使用 12 个 128 GB 3DS RDIMM）内存。
- 添加 1U GPU 托盘后，最多可支持两个 GPU，帮助提高处理能力。
- 与 6Gb SAS 解决方案相比，12 Gbps SAS 内部存储连接可将数据传输速率提高一倍，最大限度地提高了存储密集型应用的性能。
- 每台 SD530 服务器最多支持六个 2.5 英寸热插拔驱动器。两个驱动器托架可配置为支持 NVMe 驱动器，可提高吞吐量、带宽并缩短延时，最大限度地改进 I/O 性能。
- 使用 7.68 TB 2.5 英寸 SAS 热插拔 SSD 时，每台 SD530 最高支持 46TB 的内部存储。
- 支持按照联想专利设计的新型 M.2 适配器，便于轻松启动操作系统。在 RAID 1 配置中，可用的 M.2 适配器支持一个或两个 M.2 适配器，以提高启动盘的性能和可靠性。
- 使用固态硬盘（SSD）或与传统硬盘驱动器（HDD）结合使用，可提高 I/O 性能。与普通 HDD 相比，SSD 支持的每秒 I/O 读取操作数（IOPS）可达 100 倍以上。
- 本服务器具有两个可选的 10 Gb 以太网端口（10GBASE-T 或 SFP+），可将嵌入式 X722 控制器连接到刀片机箱后端的可选 8 端口 EIOM 模块。
- 配置一个 PCIe 3.0 x16 或两个 PCIe 3.0 x8 插槽，这提高了 I/O 灵活性。
- 与上一代 PCI Express 2.0 相比，PCI Express 3.0 I/O 扩展功能可将最大理论带宽提高 60%。

可管理性和安全性

强大的系统管理特性可简化 SD530 的本地和远程管理：

- 本服务器通过 XClarity Controller（XCC）来监控服务器可用性。可以选择升级到 XCC 高级版本以支持远程控制（键盘、视频、鼠标）功能。可以选择升级到 XCC 企业版本，以额外支持远程媒体文件（ISO 和 IMG 映像文件）装载、启动捕获和能耗封顶。
- Lenovo XClarity Administrator 提供全面的硬件管理工具，可延长无故障运行时间，降低成本，并通过高级服务器管理功能提高生产效率。
- 基于 UEFI 的新型 Lenovo XClarity Provisioning Manager 可在启动时通过 F1 键访问，可提供系统库存信息、图形化 UEFI 设置、平台更新功能、RAID 设置向导、操作系统安装功能以及诊断功能。
- 支持 Lenovo XClarity Energy Manager，该工具会实时获取服务器的功耗和温度数据，并提供自动控制来降低能源成本。

- 集成式可信平台模块（TPM）2.0 支持高级加密功能，如数字签名和远程认证。
- 支持安全启动（Secure Boot）模式，确保仅使用具有数字签名的操作系统。支持 HDD 和 SSD，以及 M.2 驱动器（安装在 M.2 适配器中）。
- 支持行业标准高级加密标准（AES）NI，可提供更快速、更强大的加密。
- 与辅助操作系统结合使用时，Intel Execute Disable Bit 功能可防止某些类型的旨在导致缓冲区溢出错误的恶意攻击。
- Intel Trusted Execution Technology 可基于硬件阻止恶意软件攻击，这样应用即可在它自己的隔离空间中运行，而不受系统上运行的所有其它软件的干扰，从而增强安全性。
- 利用刀片机箱中安装的 SMM 管理模块，只需一个以太网连接即可为所有四台 SD530 服务器和整个刀片机箱提供远程系统管理功能。
- 此外，本刀片机箱还支持双以太网端口 SMM 管理模块，使用该模块时，可以通过一个以太网连接以菊链方式连接 7 个刀片机箱和 28 台服务器，从而显著减少管理整个 SD530 服务器机架和刀片机箱所需的以太网交换机端口数量。



高能效

- SD530 服务器和刀片机箱提供以下高能效特性，可帮助节省能源，降低运营成本，提高能源可用性并实现绿色环保。
- 某些配置符合 ASHRAE A4 标准，可在 45° C 的数据中心环境下运行。
- 高能效平板组件有助于降低运营成本。
- 高效电源通过了 80+ 铂金（Platinum）认证。符合 Energy Star 2.1 要求。
- Intel Intelligent Power Capability 可根据需要接通或断开各处理器单元的电源，以降低功耗。
- 与 1.35 V DDR3 DIMM 相比，低压 1.2 V DDR4 内存 DIMM 的功耗要低 20% 以上。
- 与 2.5 英寸 HDD 相比，SSD 的功耗要低 80% 以上。
- Lenovo XClarity Energy Manager（可选）提供了高级数据中心功耗通知、分析和基于策略的管理，可帮助减少散热，降低冷却需求。
- 本服务器采用六角形通风孔，与圆形通风孔相比，六角形通风孔可以排列得更加紧密，便于气流更高效地通过系统。

可用性和可维护性

SD530 服务器和刀片机箱提供了许多简化可维护性并延长系统无故障运行时间的特性：

- 本服务器支持单设备数据纠正（SDDC，也称为 Chipkill）、自适应双设备数据纠正（ADDDC，也称为冗余位迁移或 RBS）、内存镜像和内存列备用功能，以在出现无法恢复的内存错误时提供冗余。
- 本服务器提供了支持 RAID 冗余的热插拔驱动器，可实现数据保护并延长系统无故障运行时间。
- 双 M.2 启动适配器支持 RAID-1，这有助于将两个已安装的 M.2 驱动器配置为冗余对。
- D2 刀片机箱和模块化刀片机箱均支持两个热插拔电源，它们将构成一个冗余对，可确保关键业务应用的可用性。
- 无需工具即可进行升级以及更换可维修部件，如风扇、适配器、CPU 和内存。
- 主动平台预警（包括 PFA 和 SMART 预警）：处理器、稳压器、内存、内部存储（SAS/SATA HDD 和 SSD）、风扇、电源、RAID 控制器以及服务器环境和子组件温度。可以通过 XClarity Controller（XCC）向管理器（如 Lenovo XClarity Administrator、VMware vCenter 和 Microsoft System Center）发放预警。这些主动预警有助于您在出现可能的故障之前采取相应措施，进而延长服务器无故障运行时间并提高应用程序可用性。
- 与传统的机械式 HDD 相比，SSD 可显著提高可靠性，从而延长无故障运行时间。
- 内置的 XClarity Controller 会持续监控系统参数，触发预警，并在出现故障时执行恢复操作，以最大限度地缩短停机时间。
- 使用 Lenovo XClarity Provisioning Manager 在 UEFI 中提供内置诊断功能，以快速完成故障排除任务，进而缩短维护时间。
- Lenovo XClarity Provisioning Manager 支持诊断功能，且会将服务数据保存到 U 盘或远程 CIFS 共享文件夹中，以便排除故障并缩短维护时间。
- 如果交流电源暂时断电，服务器将会自动重启（基于 XClarity Controller 服务处理器中的电源策略设置）。
- 支持在受支持的智能手机上运行并通过服务化 USB 端口连接到服务器的 XClarity Administrator Mobile 应用，以提供其他本地系统管理功能（需要可选的 KVM 分支模块）。
- 提供三年客户可更换单元和现场有限保修，下一工作日 9x5 服务。服务可升级（可选）。

（2）Lenovo Flex

Lenovo Flex System 高性能计算平台是一款集安全、效率、可靠性等特性为一体的服务器节点。

（3）Lenovo 四路服务器

Lenovo 四路服务器，4U 机架式服务器，适用于 HPC、数据库应用、商业分析、企业应用、云计算和虚拟化等多个领域。该节点的无盖 4U 机架优化高密度设计更适合海量的网络计算应用，且不仅简化了设备维护，同时又具备容错和高可用等特性。四路服务器最多可支持 4 颗 Intel Xeon 高性能处理器，内存支持可达 6TB，非常适合作为大内存节点。

（4）Lenovo 两路服务器

Lenovo 两路服务器，2U 机架式服务器；支持 Intel 系列处理器；支持 24 个 DIMM，最大内存可达 384GB；支持 24 个 2.5 英寸 SATA/SAS 硬盘或 14 个 3.5 英寸 +2 个 2.5 英寸 SATA/SAS 混出模式。

（5）水冷服务器

NeXtScale 液冷刀片：联想 NeXtScale 直接水冷刀片系统是业界最新的高密度计算解决方案。NeXtScale 以其灵活、开发和简化数据中心的特点为高性能计算、网格计算、仿真和分析以及大规模云计算等应用提供了优秀的基础架构。

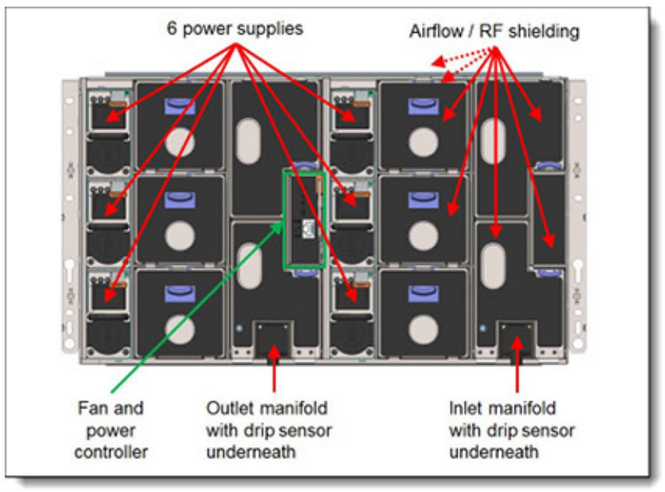
NeXtScale 直接水冷刀片系统面对目前数据中心的局限性进行密度和性能优化而设计，可以在一个传统 42U 机柜内轻松部署 72 个节点，以及相应的网络设备。

NeXtScale 直接水冷刀片系统设计运行于“温水”环境，节点冷却进水温度在 18-45 摄氏度。这意味着用户可以大幅降低 IT 系统的制冷能源消耗，最大化提高能效，降低 TCO 总拥有成本。

NeXtScale 液冷型机箱：由于采用液冷设计，使得 NeXtScale 刀片机箱可以抛弃传统刀片机箱中转为刀片散热而提供的大量散热风扇，这一方面降低了风扇消耗的能耗，另一方面大幅降低了系统运行的噪音，测试表明在 0.5 米距离上，NeXtScale 刀片机箱的运行噪音仅为 56 分贝。

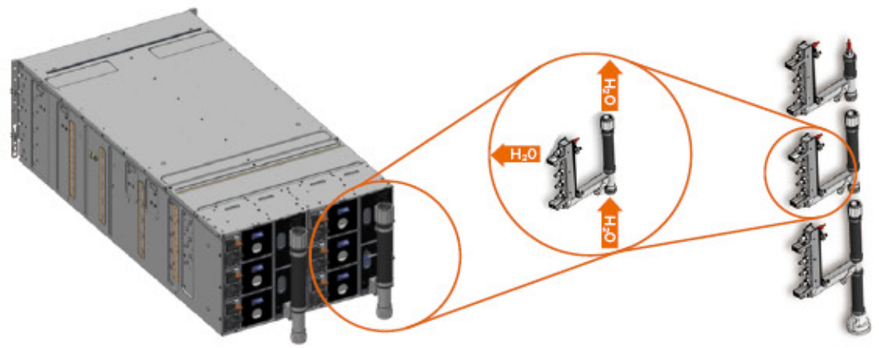
NeXtScale 机箱正视图

1 台液冷机箱 6 个标准宽度节点插槽，一个标准宽度节点包括两台 SD650 计算节点，即一个机箱可安装 12 台计算节点。



NeXtScale 机箱后视图

机箱后部已取消散热风扇，只有六个电源模块，管理模块和供水、回水接口。



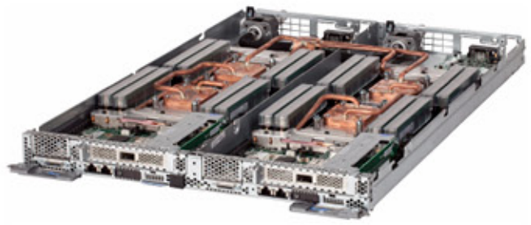
NeXtScale 机箱液冷组件连接

节点插入机箱后，直接通过快速接头与液冷组件连接。

机柜视图

一个机柜中六个机箱的液冷组件直接连接，构成整个机柜的供水与回水通道，正两个管路再分别连接到 CDU（冷水分配单元）的供水与回水管路，构成完整的二次侧水循环。

SD650 液冷节点：

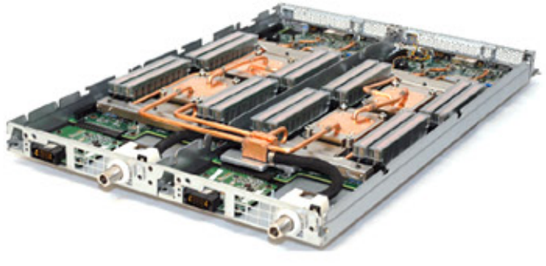


SD650 液冷节点内部视图

SD650 液冷节点采用通用的 Intel CPU，具有 16 条内存插槽，支持 TruDDR4 内存；可以安装 2 个 SSD 硬盘，集成双千兆网卡。1 台 NeXtScale 液冷机箱内部可以安装 12 个 SD650 液冷节点，每两个节点为一组。

SD650 液冷是“冷板式”设计，即发热部件不直接接触液体，而是先把热量传输给装有液体的铜制冷板，再通过液体循环带出设备。

SD650 节点的主要发热部件：CPU、内存和 I/O 板卡都有导热件通过，节点中大部分的热量都通过冷却液体带走。



SD650 液冷节点后视图

每对内存中间插入一个金属导热件，以高效吸收热量，并传递到水冷管件。每对内存条上有一个塑料盖，用以确保没有空气流入来吸收热量。

液冷节点总体效果：通过在机箱和节点上的多重液冷设计手段，NeXtScale 液冷系统可以将 85%-90% 的发热量通过液冷系统带走，而剩余的 10%-15% 的热量只需要机房环境空调即可轻松解决。

整机柜交付：联想 NeXtScale 采用整机柜交付方式，一个机柜提供 6 个水冷机箱，可安装 72 个水冷节点，管理网和监控网的千兆连接在机柜内实现，整个机柜对外连接最多可达 16 个 10Gb 端口。

整机柜交付水冷单柜

高性能计算 GPU 产品

协处理器是提升应用性能的重要手段，联想高性能计算系统的 GPU 节点通常选用 NVIDIA Telsa GPU 和 Intel Xeon Phi。协处理器的性能提升效果是严重依赖于应用本身代码的改造和数学模型 / 算法的特点的，因此不能保证协处理器会在任意应用和算法上都带来与标称性能相当的应用性能提升幅度。

高性能计算存储产品

（1）Lenovo DF 存储系统

DF 存储系统是一款针对海量数据存储应用而设计的大规模通用集群存储系统，采用通用硬件设备作为基本的构建单元，为应用提供全局统一的文件系统映像和完全与本地磁盘兼容的访问接口（POSIX 兼容）。

DF 存储系统能够支持超过 PB 级的存储容量，并根据用户应用发展的趋势，适时按需进行在线动态扩展；世界领先的元数据服务器集群技术消除了现有存储系统中所存在的单目录下文件数量、小文件处理速度等种种限制，提供了近乎无限的文件存储数量和极高的文件检索速度，是业界唯一一款能够高效支持千万级大目录的存储系统（单目录下可轻松创建千万数量级的文件，并能对文件进行高速随机检索）。同时 LeoStor 存储系统采用了自主研发的全系统规模数据高可用技术，彻底消除存储系统中的单点故障，结合特有的自动故障探测和快速故障恢复技术，确保用户的应用持续稳定地运行。

（2）Lenovo DSS

联想的 DSS（Distributed Storage Solution）。DSS 基于内置先进 RAID 技术的 GPFS 软件，和以可靠性和性能著称的联想 System X 服务器，不需要专用磁盘阵列控制器，简化了存储系统结构，提高了可靠性。同时 DSS 不像普通的分布式存储系统需要使用多份数据拷贝方式来保证数据安全性，保证了磁盘利用率与基于磁盘阵列的解决方案相同。

DSS 的另一个优势是秒级的 rebuild 时间，这使得磁盘故障和修复过程导致的系统性能下降，预报作业时间延长等现象不再存在。

借助于 GPFS 方便灵活的分级存储功能，选择不同配置的 DSS 组合在一起，就可以获得一个具有优秀分级存储功能的存储集群。

磁盘盘组中的磁盘和 2 个存储节点上的多个 SAS 接口构成多个环路，磁盘故障不会中断 SAS 环路，不会导致其余磁盘不可访问。磁盘盘组本身无控制器电路，只提供磁盘安装和环路连接，以及冗余的磁盘供电和风扇，也不构成单点故障点。

由于我们提供了热备空间和分布 RAID，我们的 rebuild 方式更加先进，从被动变成主动。每台 DSS 的一半盘中同时损坏三块磁盘后开始进行 critical rebuild，现有配置下需时 20-50 秒，通常为 30 秒。此过程不须更换磁盘，整个存储系统基本不会因为磁盘故障有性能损失，整个过程全自动化。

内置世界领先的 Decluster 技术，硬盘故障时，所有磁盘参与重建，将重构过程中的读写操作分散到各磁盘中，实现并行 I/O，从而提高重建速度及减小对存储性能的影响。块数据支持 8+2P，8+3P 的 Reed-Solomon 算法，元数据支持 3 副本，4 副本复制。

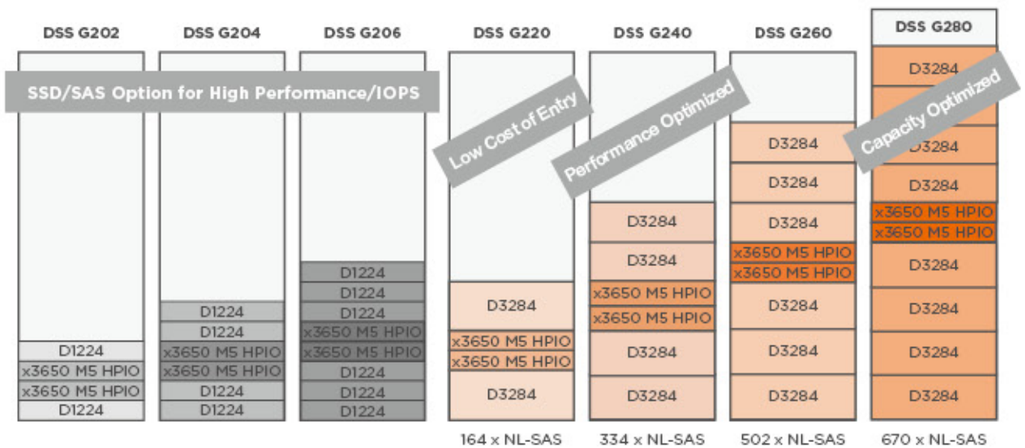
在数据安全性方面：

软件 RAID 模式支持同时坏掉 3 块硬盘而不中断业务，允许一个扩展柜损坏而数据不丢失、应用不中断；

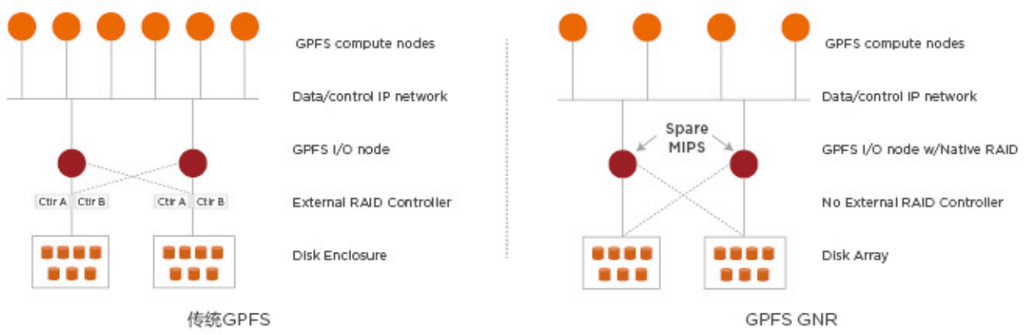
自动检测硬盘响应时间，当出现 timeout，I/O error 等情况能自动判断是硬盘本身还是路径原因导致，与此同时，系统重建由于错误而丢失的数据信息；

自动检测和维修硬盘的扇区错误，端对端的检错功能，完全避免磁盘“无声数据失效”。

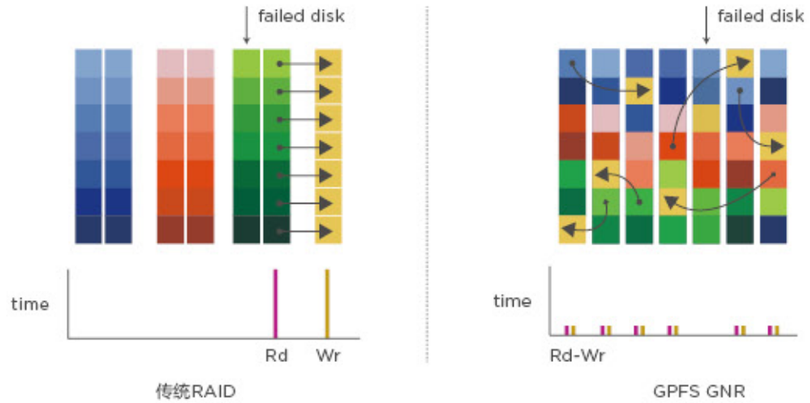
DSS存储系统产品系列



基于 GPFS GNR (General Native RAID) 技术的存储系统：



GPFS GNR 较传统的 GPFS 而言，省去了磁盘阵列控制器。而是采用 JBOD 的方式直接连接磁盘。这种连接方式不仅节约了成本，还在 I/O 性能和可靠性方面有更好的表现。尤其是在磁盘出故障需要重构时，GPFS GNR 的重构时间较传统磁盘阵列有明显优势。



较传统 RAID 而言，GPFS 采取 d-cluster 方式组织磁盘。这种方式能大大缩短磁盘损坏后系统重构的时间。上图分表描述了传统 RAID1 和 GPFS GNR RAID1 配置。黄色磁盘为预留的 spare 盘。在传统 RAID1 系统重构过程中，所有数据将从同一磁盘拷贝到 spare 磁盘。所需时间较长。而 GPFS GNR 中，每一对数据拷贝都分散分布在不同磁盘对中。当某个磁盘失效时，可以从各磁盘中读取相应的数据块，复制到预留的数据块中即可完成重构。其优势在于能将重构过程中的读写操作分散到各磁盘中，实现并行 I/O，从而缩短重构时间。

GPFS GNR 运行在 x86 平台下，采取的是 8+3p 模式。即 8 个数据盘加上 3 块校验盘。每份数据保存有 4 分拷贝，采用 d-cluster 的方式分布在 47 块磁盘上。当一组磁盘失效时，GPFS GNR 能很快重构，几乎不影响用户数据访问。即使有 3 组磁盘同时失效（非常少见），GPFS GNR 也能很快将其恢复在 2 组磁盘失效的模式，然后进行重构。GPFS GNR 在 x86 平台下，较传统的 RAID 而言在系统重构方面有很大优势。

研究发现，现行存储系统由于数据传输路径较长。从主机 CPU，到主机内存，光纤卡，存储控制器，最终到磁盘。在数据的传输和存储过程中，有一些细微的数据错误（如比特位翻转）很难通过常用的工具检测到。GPFS GNR 能提供企业级的点到点数据校验检测。存放在磁盘上的所有数据块都有相应的校验位。GPFS GNR 能对数据在传输和读写过程进行全面的检测，保证数据有效性和可靠性。

GPFS GNR 提供统一管理接口。磁盘重构，负载均衡，点对点数据校验检测等都有后台自动进行。很少需要用户手动干预。GPFS GNR 还能动态配置系统重构对整个磁盘带宽占用的比例，能保证关键应用的磁盘带宽。

GPFS GNR 提供全面的磁盘管理功能。能检测并管理磁盘连接，介质错误，磁盘失效等问题。针对介质错误或者点对点校验失效等错误，GPFS GNR 能在后台自动修复。根据磁盘工作状态，记录其“健康”状况，能在磁盘失效前，进行预警。

GNR 功能概览：

Declustered RAID

- 数据、奇偶校验信息以及热备盘空间统一条带化并分布到整个磁盘阵列
- 每个磁盘阵列磁盘数量无限制

2 路和 3 路容错支持

- Reed-Solomon 奇偶校验算法
- 2 或 3 路容错支持
- 3 或 4 镜像支持

端对端校验，错误写入检验

- 磁盘操作结果直接反映到 GPFS 用户 / 客户端
- 检测并纠正无声数据损坏问题

后台错误诊断机制（异步）

- 如果磁盘错误：尽可能验证并恢复
- 如果磁盘或磁盘柜路径错误：切换到另一条路径

支持磁盘热拔插

- 部分磁盘异常的情况下，保证系统服务不中断

Lenovo Intelligent Cluster 的介绍：

A. 要点

a. 复杂解决方案变得简易　　b. 行业最佳的技术，优化的解决方案设计

c. 享受 Lenovo 端到端支持的端到端解决方案

部署适用于技术计算、高性能计算（HPC）、存储和云环境的解决方案可能会给 IT 带来极其沉重的负担。Intelligent Cluster 利用 Lenovo 数十年的丰富经验，凭借预先集成、交付、得到全面支持的解决方案来降低部署复杂性，这些解决方案将行业最佳的组件与优化的解决方案设计紧密结合。借助 Lenovo Intelligent Cluster，客户可以集中精力最大限度提高业务价值，而不是耗费宝贵的资源来设计、优化、安装和支持满足业务要求所需的基础架构。

B. 行业最佳的技术，优化的解决方案设计

Intelligent Cluster 解决方案采用行业领先的 Lenovo® System X® 服务器、存储设备、软件和第三方组件，可以在集成交付的解决方案中支持广泛的技术选项。 Lenovo 针对可靠性、互操作性以及最高性能彻底测试和优化了每个解决方案，因此客户可以迅速部署系统并使投入工作，从而实现他们的业务目标。

C. 享受 Lenovo 端到端支持的端到端解决方案

Intelligent Cluster 解决方案由 Lenovo 构建、测试、交付和安装，并作为单一解决方案进行支持，而不是作为数以百计单独组件进行处理。 Lenovo 提供了包括 Lenovo 和第三方组件二者、单一联系点、解决方案层面的支持，可以在系统的整个生命周期内实现最高的系统可用性，因此客户可以花较少的时间维持系统，花更多的时间实现成果。

高性能计算网络产品

高性能计算网络部分包含多个不同协议产品，他们分别是：

（1）高速网络 EDR(100Gb)/HDR (200Gb) InfiniBand 网络：包括 36 端口、216 端口和 648 端口 EDR 交换机；40 端口、800 端口的 HDR 交换机。

（2）高速网络 Intel 100Gb OPA：包括 48 端口和 768 端口的交换机；

（3）16/32Gb FC 网络；

（4）12Gb SAS 网络；

（5）万兆网络；

（6）千兆网络。

高性能计算机房环境配套产品（配套的水冷和模块化数据中心）

（1）联想模块化数据中心

1）产品概述

目前数据中心基础设施形态千差万别，但一体化、标准化及模块化已经成为数据中心建设的主流趋势。在此背景下，联想推出了模块化数据中心（MDC），它将数据中心所必需的电气、制冷、机柜、监控、消防、布线、IT 设备及操作平台等软硬件集成在一个封闭的模块化空间内，在高度集成了计算能力的同时，还大大降低了对空间和能耗的需求，在具备高可靠性的同时提供极其灵活的可扩展能力。

联想针对不同客户类型，提供了完整的模块化数据中心解决方案，包括了模块化数据中心 Lenovo Smart Aisle、小型数据中心 Lenovo Smart Row，以及微型数据中心 Lenovo Smart Cabinet 解决方案等。

智慧数据中心		
Lenovo Smart Cabinet	Lenovo Smart Row	Lenovo Smart Aisle
		
机柜内的整合的智能型基础架构，适用于小型网络间，远程站点。 容量：3.5KW，单机柜	整排机柜的整合的智能型基础架构，适用于小型数据中心，灾后恢复。 容量：12.5KW，1-6个机柜	针对中大型数据中心基础设施的标准化、模块化解决方案，可快速部署、在线扩容、在线改造。 容量：25-200KW，7-35个机柜

2）Lenovo Smart Cabinet 介绍

Lenovo Smart Cabinet 是联想数据中心解决方案产品，应用于微小型数据中心、办公区域等室内环境。为客户提供完整的 UPS 不间断电源、机架配电、热量管理、机柜支撑和监控管理；600mm 宽 1200mm 深 2000mm 高，42U 高机柜（IT 可用空间 29U），重量约 300Kg，满足 3.5kW 设备散热及供电需求；采用机架式 UPS，满足后备 30 分钟；全封闭机柜，带来机房的进一步节能。

具有如下特点：

- 防尘降噪，高效节能
- 智能监控
- 高度集成、节省空间
- 良好的人机界面
- 快速交付，品牌无忧

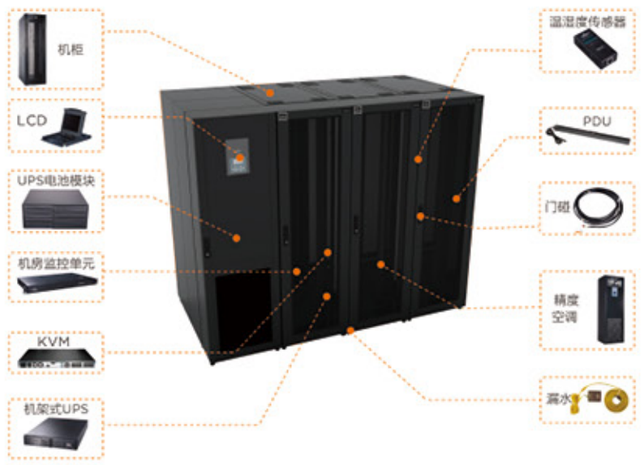
3）Lenovo Smart Row 介绍

Lenovo Smart Row 系列解决方案，是在已获得巨大市场成功的 Lenovo Smart Row 系列的基础上进行重大升级的全新解决方案，其应用仍然定位于 10~100 平米小型数据中心。符合用户对移动管理的需求、对机房环境的依赖度更低、适应更宽广的 IT 热密度范围。

为客户提供完整的 UPS 不间断电源、机架配电、热量管理、机柜支撑和监控管理，满足客户机房的基础架构需求。采用变频 12.5kW 的制冷空调，机架式 UPS，满足后备 30 分钟；采用全封闭机柜，带来机房的进一步节能；整体化方案具有高可靠、高智能和快速部署等优势；可为用户提供多种搭配方案。

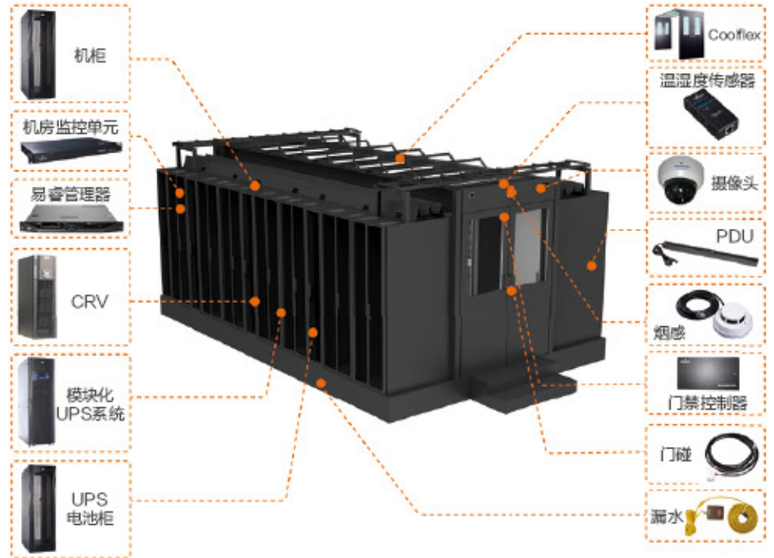
具有如下特点：

- 灵活设计与扩容
- 低建设成本
- 低运营成本
- 交付周期短
- 良好的人机界面
- 绿色节能



4）Lenovo Smart Aisle 介绍

Lenovo Smart Aisle 按照不同机柜数量、机柜密度，集成行间制冷，冷通道封闭，监控，机柜及配电并保证消防的一体化设计；供电上可灵活的采用集中 UPS 供电、分散模块化 UPS 供电等多种方式，实现 UPS 单路、双总线供电模式，并根据后期业务需要可逐渐扩容。



Lenovo Smart Aisle 模块化整体方案设计包括四大系统：供配电系统、空气调节系统、服务器机柜及附件系统和动力环境监控系统。

具有如下特点：

- 降低系统复杂度，提高系统可靠性 — 与传统数据中心相比，模块化设计可预先验证，可复制性强，可靠性不随系统数量增加而降低。
- 部署速度快 — 模块化数据中心可以非常迅速地部署，通常只需要几周时间；相比之下，建造传统的数据中心通常得需要 1 到 2 年，甚至几年时间。
- 降低运维成本（OPEX）— 相同的模块可以采用相同维护规程，降低维护人员要求，共享备品备件；采用行间空调及封闭冷通道技术，可以大大降低 PUE 值，带来数据中心整体运营成本降低。
- 易于部署 — 模块化数据中心可采用集装箱或预制单位构建方式部署在指定的任何位置，甚至可以移动；易于扩展 — 如果需要，可以添加更多的模块。

5）模块化数据中心整体方案特点

经济性

降低机房TCO。采用分期建设的思路，减少初期建设成本CAPEX占用；可提升单机柜容量，带来占地40%，方案美观紧凑；运营成本OPEX降低30%，更高效节能。

智能性

管理简单方便。通过统一监控平台实现机房智能监控、智能管理、智能供电和智能制冷，保障了对IT环境的智能控制。

高可用性

可用性超过99.999%。通过工程产品化，实现整体设计及交付，避免各系统之间耦合问题，大大提高数据中心整体可用性。

协同性

高度一体化。各系统风格、协议统一，一站式服务，实现快速部署和在线扩容，保障客户关键业务永远在线。

售后服务

联想提供整体机柜三年售后服务，且全国均有售后网点。

（2）直接水冷解决方案

1）设计规范、标准或依据：

《NeXtScale M5 WCT solution implementation general guide V3》

《Thermal Guidelines for Liquid Cooled Data Processing Environments》

2）主要设计参数和设计工况

设计指标及工况如下表：

NO.	项目	参数	备注
1	单机柜额定热负荷	33.5KW	节点按照最高功耗 CPU 算（165W） 规定其中 85% 的热负荷需要由水冷设备承担制冷
2	机柜数量	2	
3	单机柜最小热负荷	10KW	以 2*105W CPU、8*16GB MEM、1*IB、2*2TB 节点的 25% 负载估算
4	机柜进水温度	45℃	
5	单机柜额定水流量	≥36lpm	36x compute trays (dual node), 6x switches: Minimum flow rate = 36 liters per minute
6	单机柜水阻力	2.4psi	rack manifold and 6x DWC chassis
7	机柜工作水压范围	≤50psi	Pressure should be below 50psi
8	室外温度（额定点）	40℃	
9	电源形式	380V 50Hz 3N	最大为 4 路输入
10	室外温度范围	-14℃ ~ 40℃	在该范围内满足全负荷制冷要求，冬季防冻
11	冷水机组配置	不标配	实现完全自然冷
12	水质要求	参看文件	
13	机柜进出水接口形式	参看文件	
14	水过滤器要求	50μm	300 目

表 设计参数

3）通用设计要求

- 制冷系统具备高可靠性，满足液冷服务器长时间不间断运行。
- 制冷系统能够有效预防水温过低，防止在液冷服务器内出现凝露。
- 制冷系统具有必要的智能监控功能，能够对系统出现的异常情况做出告警和处理。

4）制冷系统设计方案

液冷服务器配套制冷系统如下图所示，本文的设计范围包括左边框体内的所有设备或功能，具体即中间换热单元、室外干冷器等设备或功能。

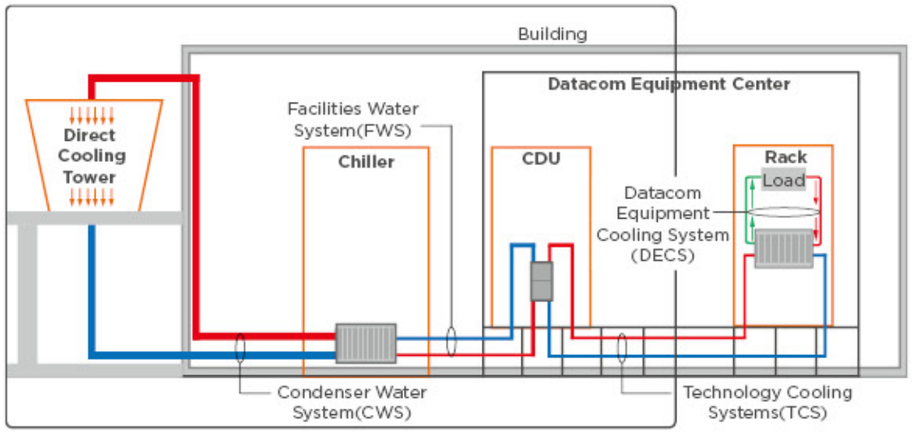


图 液冷服务器配套制冷系统

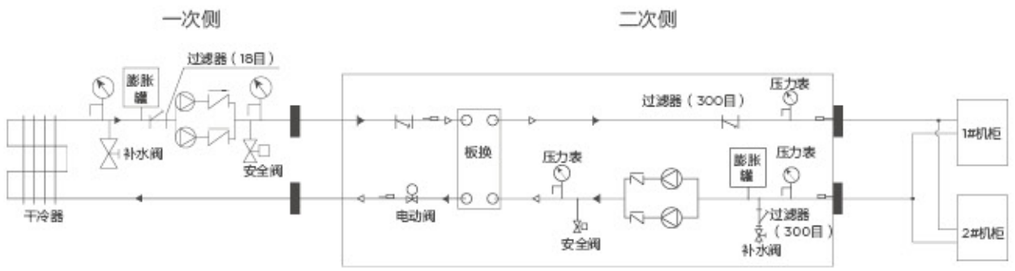


图 液冷服务器配套制冷系统设计方案

图所示为本建议书的设计方案，中间换热单元提供一定流量、一定温度的冷却水进入液冷服务器的换热芯体，带走CPU 发出的热量，被加热的冷却水流回中间换热单元的板式换热器，将热量释放到室外侧循环水，该部分热量再通过室外干冷器或者冷水机组释放到室外环境中，完成对液冷服务器的热量管理。

5）中间换热单元方案

中间换热单元在二次侧通过调节送入液冷服务器换热芯体的水温和水流量从而向末端机柜提供冷量，起到冷量分配的作用，此外为了便于实施最佳的室外侧方案及保持液冷服务器换热芯体内的水质，中间换热单元也起到系统隔离的作用。

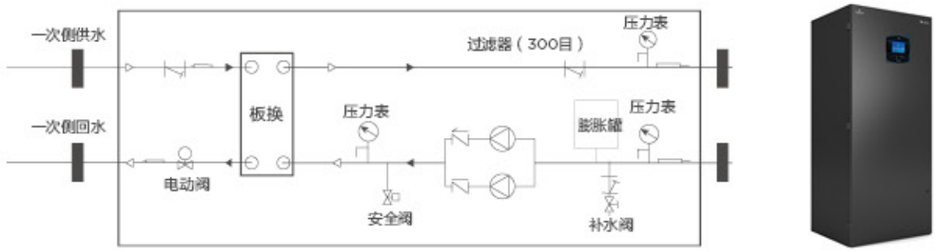


图 中间换热单元方案

6）中间换热单元系统设计

中间换热单元主要包括二次侧板式换热器、循环水泵、过滤器、膨胀罐、阀件、管路及电气和控制系统，如所示。

7）二次侧板式换热器

板式换热器选用钎焊式，采用国际知名品牌，不低于 SWEP、Danfoss。

为保证纯水水质，板式换热器采用 316L 不锈钢材质、换热器与外管路采用螺纹连接。

板式换热器表面敷有保温棉，消除壁板结露现象。

为了最大限度的使用自然冷源，中间换热单元按照 45℃出水、55℃进水进行选型计算，2 台机柜的散热量为 57kW（假设单机柜 33.5kW 的发热量 85% 全部被冷却水带走）。

Item	二次侧	一次侧
进水温度 ℃	55	43
出水温度 ℃	45	48
水流量 L/min	83	175
总压降 kPa	2.5	10.7
换热量 kW	68	

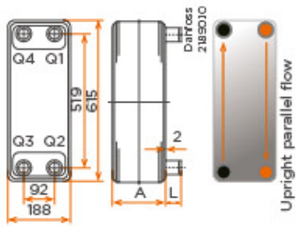


表 中间换热单元板换选型（一次侧工质为 30% 乙二醇水溶液）

8）二次侧循环水泵

采用一主一备双泵设计。

水泵选用采用高效、节能、紧凑的离心式清水泵。

水泵选用国际知名品牌，水泵电机可调速。

设计流量为 5.1m³/h，克服软管（假设 30m）、过滤器、机柜等器件的阻力，设计扬程为 33m。

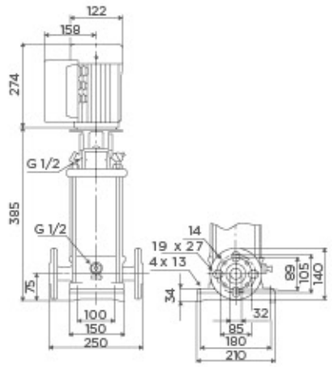


图 二次侧循环水泵

9）电动调节水阀

电动调节阀采用国际知名品牌：Siemens、Belimo、Johnson Controls 等品牌。

电动调节阀采用等百分比球阀形式。

执行机构应采用 24V AC/DC 供电，0.5 ~ 10VDC 或 4 ~ 20mA 信号控制，具有位置信号（0 ~ 10VDC 或 4 ~ 20mA）反馈。调节精度高，稳定可靠。

10）二次侧膨胀罐

膨胀罐选用隔膜式膨胀罐，作为定压膨胀装置，提供系统定压、缓冲水泵启停产生的水锤、容纳水温变化导致的管路内容积变化。

膨胀罐选用国际知名品牌，不低于 Aquasystem 和 Varem。

11）过滤器

过滤器选用 Y 型过滤器，不低于埃美珂、OR、天津格莱式等品牌。

12）橡胶软管

橡胶软管材质应为三元乙丙橡胶（EPDM）。

橡胶软管满足系统可能出现的最高压力。

橡胶软管采用国际知名品牌：Paker、Hansa、Manuli 等。

13）中间换热单元控制系统

中间换热单元采用艾默生机房空调专用控制器。

中间换热单元控制系统具有信号采集、逻辑计算、自检告警等功能，以保证系统的正常运行。

中间换热单元具有人机操作界面、数据通讯传输等交互途径，提供快捷的控制方式。

14）信号采集

一次侧进出水温度采集；　　漏水检测采集；

二次侧进出水温度采集；　　远程温湿度采集；

二次侧进出水压力采集；　　电动调节水阀开度采集。

15）二次侧出水温度控制

通过对一次侧电动调节水阀的控制，调节一次侧的水流量，使二次侧的出水温度达到设定的要求。

16）二次侧水流量控制

通过二次侧供回水压差、温差及进水温度对二次侧循环水泵的转速进行控制，向末端服务器提供充足的水流量。

17）防凝露控制

通过采集服务器周围环境的温湿度，实时监测出水温度使其高于露点温度。

18）告警

可设置高低水温告警、漏水告警、泵故障告警等多种属性告警。

19）水泵主备切换

当检测到运行水泵故障告警时，自动切换至备用水泵运行。

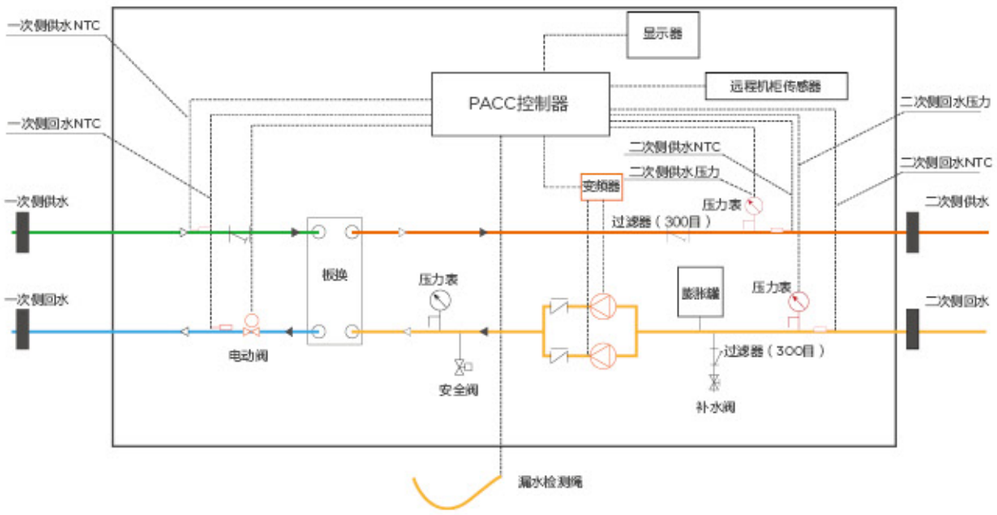


图 中间换热单元控制系统

20）室外散热方案

室外散热系统将中间换热单元传递出来的热量释放到室外环境中。

21）室外散热系统设计

室外散热系统主要包括干冷器、一次侧膨胀罐、水泵、过滤器、管路、电动调节水阀及电气和控制系统。如图所示。

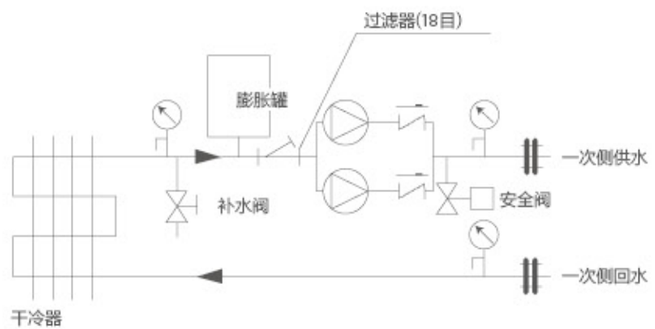


图 室外散热系统

22）干冷器设计

按照室外温度 35℃条件下能够提供满负荷散热量进行设计计算。

Item	空气侧	水侧
进水（风）温度 ℃	40	48
出水（风）温度 ℃	43.6	43
流量	54，000m^3/h	203L/min
换热量 kW	65.3	

表 40℃室外温度干冷器设计（30% 乙二醇水溶液）

图 干冷器

23）一次侧循环水泵

采用一主一备双泵设计。设计流量为 12.2m³/h，克服管路、过滤器、干冷器等器件的阻力，设计扬程为 22m。其他要求同二次侧水泵。

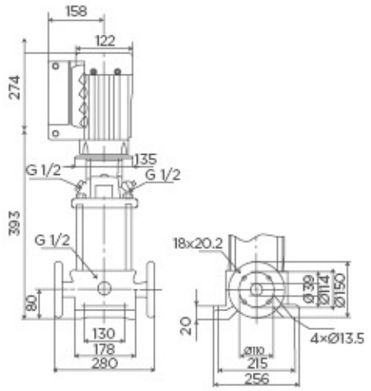


图 一次侧循环水泵

24）室外散热控制系统

室外散热系统的运行逻辑要尽可能的使用自然冷，以二次侧的出水温度为控制目标，一次侧循环水泵、干冷器风机、水阀按照由高到低的优先级进行调节，尽可能的降低能耗。

25）水质处理

二次侧循环水需要线监控水质、PH、硬度、腐蚀率等参数，一旦超出水质要求，就定期更换或者采用一部分水旁滤并进行补水的方法使总体水质满足要求。

对二次侧系统进行补水时，先经过一定的水处理，经过软化、去离子、反渗透（RO）等技术处理，由专业的水处理设备实现。

26）干冷器和闭式冷却塔的对比

干冷器和闭式冷却塔都是典型的室外散热设备，最本质的不同在于干冷器直接与室外空气进行热交换，而闭式冷却塔是在干冷器的基础上增加了一套喷淋设备，利用喷淋水的蒸发起到增强换热的作用。鉴于该特点：

- 闭式冷却塔需要专用的喷淋水供水系统，包括水泵、管路等，以使其能利用源源不断的喷淋水来制冷，同时该系统还需要水处理设备，以保证喷淋水不会在换热器上产生结垢、腐蚀等问题。因此闭式冷却塔的优点在于强化了换热，在室外温度较高时也能带来相当的制冷量，尤其是在炎热干燥的地区效果更明显，但相应的也带来了明显的缺点，冷却塔需要持续供应喷淋水，需要有水源，并且也产生了用水的费用，如果水质达不到要求，则需要定期的除垢换水，日常的维护工作较多，控制较复杂，另外在冬季低温时，还要将喷淋系统放水防冻。
- 干冷器没有喷淋系统，因此不存在由此带来的工程施工、额外设备投入等问题，不需要提供水源以及不会产生持续的用水费用，干冷器日常的维护也较简单。但由于干冷器没有喷淋水强化换热，冷源仅来自于与空气的强制对流换热，在室外温度较高时，需要更大的换热器才能达到更大的制冷量。

对于液冷服务器的散热需求，从应用上来看，当机柜数量较少即散热需求较小时，适合使用干冷器，投入小，施工维护简单；当机柜数量较多即总体散热需求较大时，需根据应用场地的环境温度、水源有无、占地面积等综合考虑选择冷却塔或是干冷器。

模块	设备名称	参数
中间换热单元系统部分	板式换热器	68kW（一次侧 55/45℃，二次侧 43/48℃）
	变频水泵	5m³/h，330kPa　2~9m³/h 调节
	膨胀罐及安全阀	2L，泄压阀
	补水阀	1 1/2
	止回阀	DN40
	水过滤器	300 目（50μm），DN40
	水过滤器	18 目，DN65
	球阀及执行器	DN50 螺纹，Kvs=40m³/h
	橡胶软管	EPDM，1”，15m
	快速接头	1”
中间换热单元控制部分	控制器 & 显示器	PACC 控制器
	NTC	一次侧供回水，二次侧供回水
	远程温度传感器	1# 机柜，2# 机柜
	压力传感器	二次侧供回水
	漏水检测绳	
	配电器件	主空开、变频泵空开 / 接触器、汇流排、成套线缆
室外干冷器	风机	27000m³/h，50Pa
	翅片管换热器	
室外水力模块	水泵	12.2m³/h，220kPa
	单向阀	DN65
	补水阀	1 1/2
	膨胀罐及安全阀	8L
	水过滤器	18 目，DN65
	压力表	
室外控制系统	控制板	PACC 控制器
	温度传感器	室外温度
	配电器件	主空开、泵空开 / 接触器、风机空开 / 接触器、冷水机组空开 / 接触器、汇流排、成套线缆

表 关键设备清单

联想高性能计算产品的特点

系统之性能和技术水平

高性能计算节点，登录管理节点，存储节点和图形处理服务器节点的内存，采用的是独具技术优势和特点的 TruDDRX 内存，相比其他友商，具有非凡的技术优势：

高于工业标准的 TruDDRX 内存：

- 高于行业标准的带宽工艺独有指标（POR+1）
- 仅选用 Tier 1 的原材料供应商
- 识别内存条是否经过认证，保护客户利益
- 内存防干扰设计
- 内存效率最高提高 15%
- 系统总体性能提升超过 10%

高于工业标准的TruDDRX（3，4....）

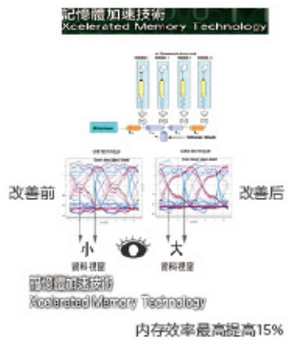
服务器

TruDDRX技术

- 高于行业标准的带宽工艺独有指标（POR+1）
- 仅选用Tier 1的原材料供应商
- 识别内存条是否经过认证，保护客户利益
- 内存防干扰设计

*仅16GB和LRDIMM

Type	Ranks Per DIMM and Date Width	DIMM Capacity (GB)		Speed (MT/s); Voltage (V); Slot Per Channel (SPC) and DIMM Per Channel (DPC)					
				1 Slot Per Channel	2 Slots Per Channel		3 Slots Per Channel		
				1DPC	1DPC	2DPC	1DPC	2DPC	3DPC
		4Gb	8Gb	1.2V	1.2V	1.2V	1.2V	1.2V	1.2V
RDIMM	SRx4	8GB	16GB	2133	2133	2133	2133	2133	1866
RDIMM	SRx8	4GB	8GB	2133	2133	2133	2133	2133	1866
RDIMM	DRx8	8GB	16GB	2133	2133	2133	2133	2133	1866
RDIMM	DRx4	16GB	32GB	2133	2133	2133	2133	2133	1866
LRDIMM	QRx4	32GB	64GB	2133	2133	2133	2133	2133	1866



内存效率最高提高15%

独特的风道分区技术：

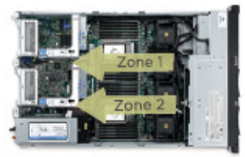
全时5-40摄氏度工作环境

服务器

联想服务器全时支持5-40摄氏度工作环境

- 提高空调温度，节约能耗
- 增加环境温度的裕度，提高可靠性
- 联想服务器对5-40摄氏度的工作环境是全年全天候满配的

刀片系统 风区设计



机架式 风区设计

高密度刀片系统 风区设计

高于工业标准的设计和工艺水平：

卓越工业设计和工艺

服务器

镀金含量远高出业界一倍

无烟设计

源于工业标准，高于工业标准

原料筛选

主机板原料选用与设计

主机板设计防止鉴井效应

讯号设计原理

英特尔®至强®可扩展平台
加速探索与获取洞察
咨询电话 400 819 6776

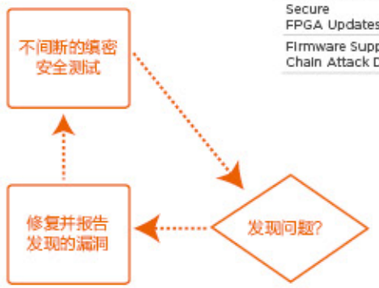
独有的安全信任机制：

领先的安全信任机制

服务器

	System x Host	System x Sys Mgt	Dell Host	Dell Sys Mgt	HP Host	HP Sys Mgt	Cisco Host	Cisco Sys Mgt	Super-micro Host	Super-micro Sys Mgt
Firmware Signing	√	√	√	√	√	√	√	√	?	N?
Dedicated TPM	√	√	O	N	O	N	O	N	?	N
Measured Boot	√	√	O	N	O	N	O	N	?	N
Verified Boot	√	√	√	N	√	N	√	N	?	N
Firmware Secure Rollback	√	N	√ (Partial)	N	N	N	N	N	?	N
Secure FPGA Updates	√	NA	N	NA	N	NA	N	NA	?	NA
Firmware Supply Chain Attack Detection	√	√	N	NA	Y	N	N	N	?	N?

Sys Mgt-IMM/CMM
N-Not supported
O-Optional feature (Customer need to order TPM option)
NA-Not Applicable



- 高于行业的安全机制
- 不仅仅在开机，在整个运行期间都测试，发现问题，修复，保证系统整个生命周期的安全

系统之可靠性、稳定性

高可靠性与稳定性是关键业务用户最关心的指标之一，也决定了系统在生命周期内能真正提供多少服务能力。

ITIC 调查表明 System X 平台可靠性同 IBM System p 不相上下，在 x86 平台中冠绝群雄：

- 在 4 小时以下意外宕机时间中，System X 可靠性是主要竞争对手的 2-5 倍
- 在 4 小时以上意外宕机时间中，System X 可靠性是主要竞争对手的 2-13 倍

权威评鉴的可靠性

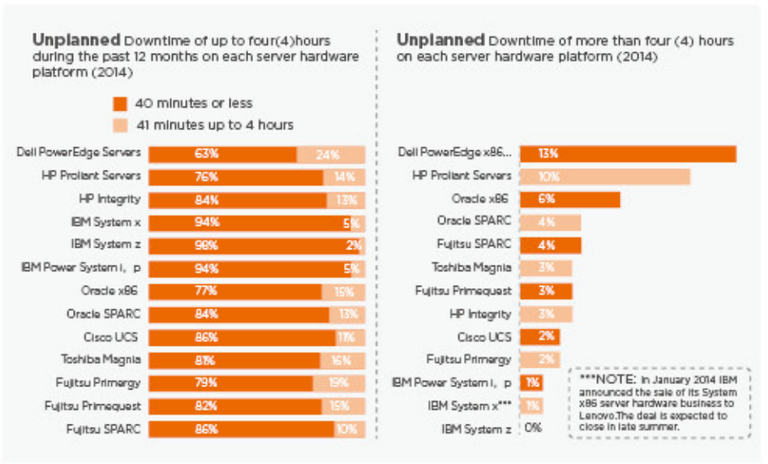
服务器

#1 x86可靠性 评价 [1]

ITIC Global Server Hardware, Reliablity Report-2014-2015

ITIC调查表明System X平台可靠性同 IBM System p 不相上下，在x86平台中冠绝群雄

- 在4小时以下意外宕机时间中，System X 可靠性是主要竞争对手的 2-5 倍
- 在4小时以上意外宕机时间中，System X 可靠性是主要竞争对手的 2-13 倍

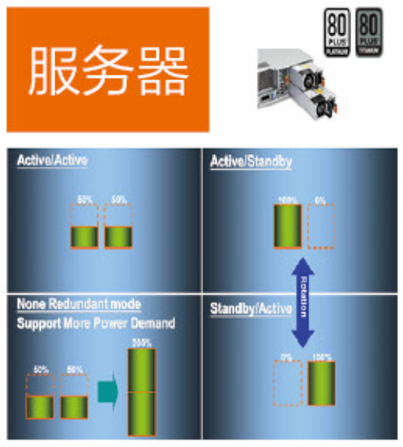


系统运行效率和节能设计

先进的电源节能设计

独有Active/Standby电源

服务器



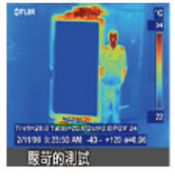
- 联想全系列机型选用多种符合80+认证标准的电源，从铜牌到钛金级别任君选择
- 多种功率，交直流电源支持，已经正式发布高压直流电源
- 联想服务器采用独有支持Active/Standby技术
  - 仅有一个电源在运行，减少损耗
  - 两个电源实现了真正意义上的冗余

双马达冗余热插拔对转风扇

服务器

借鉴航空涡扇设计原理

- 每一组风扇模块有两组独立的马达，本身就是冗余的
- 每一组风扇模块有两组叶片相反方向旋转，产生同向加强风力和对向抵消的力矩
- 热插拔风扇



联想服务器拥有业界最高的客户满意度

联想服务器在客户满意度方面——独领风骚

#1 总体客户满意度，另有22项#1涵盖了服务，产品和销售等多个评比指数

TBR满意度指数	排名
总体客户满意度	#1
服务满意度	#1
产品满意度	#1
销售责任心	#1

TBR #1 连续三年客户总体满意度第一（2014, 2015,2016）

联想高性能计算集群实施服务

此高性能计算集群实施和培训服务会将已经完成上架的各种硬件设备进行系统软件的安装和配置，从而真正成为一套可用于实际工作的集群系统。联想企业级业务集团专业服务团队的高性能计算实施工程师通过快速配置集群以满足客户的需求，并且针对集群日常管理所需的技能对客户进行培训。集群可由各种联想自有产品或联想 OEM 的产品组成，包括 System X 和 ThinkServer 机架式服务器、Flex System 刀片式服务器、NeXtScale 高密度服务器以及存储产品和网络产品等。联想企业级业务集团专业服务团队拥有专业的高性能计算领域的知识和技能，快速帮助客户将各种硬件组合成集群系统。

在项目的前期阶段就可以引入专业服务团队的HPC实施工程师，和客户进行充分的沟通和交流，了解客户的需求和想法，同时可以基于我们的实践经验，为客户提供最佳实践的意见和建议。而后针对客户的需求和想法，我们会撰写工作说明书，以明确实施服务的主要工作内容和范围。在和客户确认工作说明书后，我们就可以评估整个项目实施服务所需的工作量，从而提供报价给到联想的销售人员。

在高性能计算集群的硬件和服务下单后，我们会得到通知。实施工程师即可安排和客户沟通，制定具体实施计划以确保实施过程顺利。在实施阶段，根据集群规模大小，一个或者多个实施工程师会到达客户数据中心，完成所需的实施服务。在项目尾声，实施工程师会将正常工作的集群系统交付给客户，并且提供相应的安装实施文档以及交付培训。

实施速度

我们的实施工程师会凭借他们丰富的经验以最快的速度完成集群的实施工作。他们会使用经过多次验证的最佳实施经验以及软件组件版本，避免反复的试验和错误，快速让集群系统可以进入工作状态。

实施质量

我们的实施工程师不断的积累最佳实践方法，改进安装配置技术。他们还与联想的研发团队以及各地客户紧密配合，以帮助完善产品和实施服务的质量。

减少风险

我们的实施工程师会帮助您避免典型错误导致的性能问题、兼容性问题或保修问题。

选择联想的专业实施服务将帮助客户充分利用新系统的能力，而不会忽略一些重要的特性。客户不需要花费大量的时间去学习集群的知识，而是快速的从实施工程师的丰富经验中得到所需的知识和信息。

高性能计算集群实施服务通常包含以下内容：

每一个高性能计算集群实施服务都有特殊的地方，但是多数通常都包含以下内容：

- 一个准备和计划会议
- 管理节点的安装和配置。包含 RAID 配置、操作系统安装、BIOS/UEFI 设置、微码更新、群集管理软件安装（xCAT）
- 以太网配置和验证
- 节点（计算节点、存储节点、登录节点、其它用途节点等）RAID 配置
- 向节点（计算节点、存储节点、登录节点、其它用途节点等）分发操作系统，并配置 BIOS/UEFI、微码更新、驱动和软件包更新

根据硬件配置不同的可选服务包含：

- 高速网络（InfiniBand 或 10Gb）的配置和验证，以及相匹配的软件包安装，例如 OFED
- 存储系统配置和验证
  - 在管理节点上安装存储管理软件
  - 根据存储需求配置多路径软件
  - 存储控制器或 SAN 交换机 Zone 配置
  - 创建 RAID、LUNs 和文件系统
- 特殊用途节点配置，如 GPU 节点

根据软件配置不同的可选服务包含：

- 资源管理 / 调度软件的实施，如 Torque/Maui, Torque/Moab, PBS-Pro, LSF, Platform HPC, SLURM 等
- 集群监控或报警软件实施，如 Ganglia, Nagios
- GPFS 并行文件系统实施
- 特殊用途工具、软件、库的实施，如 Intel Compiler, Intel MKL, 特殊的 MPI bindings.

额外可选的附加服务包含：

- 协助解决方案的架构设计，审阅建议的解决方案
- 硬件测试：加电开机测试，包括 STREAM, PALLAS, IOzone, HPL 在内的基准测试
- 协助性能测试和验收测试
- 对于集群中各种组件的深入技术培训
- 协助客户进行应用软件和许可证管理器的安装
- 协助项目管理：实施进度安排、硬件和软件的订购和交付、预算管理

要订购此服务，请联系我们的商机管理员根据客户的实施需求撰写工作说明书（SOW，Statement of Work）。在中国地区，此服务可以通过以下编号进行订购。具体订购数量将由商机管理员根据工作说明书来确定。