insideHPC

insideHPC Guide to

How Expert Design Engineering and a Building Block Approach Can Give You a

Perfectly Tailored AI, ML or HPC Environment

by Keith Shaw

Brought to you by

Silicon Mechanics

Illustration courtesy of SiliconMechanics

White Paper

Contents

Introduction

When considering a large complex system, such as a high-performance computing (HPC), supercomputer or compute cluster, you may think you only have two options — build from scratch from the ground up, or buy a pre-configured, supercomputer-in-a-box from a major technology vendor that everyone else is buying. But there is a third option that takes a best-of-both-worlds approach. This gives you "building blocks" expertly designed around network, storage and compute configurations that are balanced, but also flexible enough to provide scalability for your specific project needs.

Across the artificial intelligence (AI), machine learning (ML) and HPC landscape, organizations are moving from proof-of-concept to production projects that require software and hardware beyond off-the-shelf components or cookie-cutter server infrastructures. Most AI and ML projects demand that computing power, storage capacity and network infrastructure work seamlessly together to help avoid bottlenecks. For example, the fastest processors available won't matter if your storage network is too slow.

Several companies, including NVIDIA, offer complete, supercomputer-in-a-box systems that harness the power of the NVIDIA HGX™ A100 GPU and its related components. The NVIDIA DGX™ A100 POD, for example, offers a complete system that provides great performance and several options for those looking for the latest features.

In many cases, customers could end up with features or hardware that they don't need, or fall short in areas where they could use some extra power. That's where working with the expert design engineers at Silicon Mechanics can help. Alternatives exist to the DGX A100 POD that can provide the same amount of performance, but with the additional bonus of having specific customizations that directly connect with a company's AI, ML, or deep learning project.

In this guide, we will present things to consider when building a customized supercomputer-in-a-box system with the help of experts from Silicon Mechanics.

An Example of a Flexible Configuration Using Building Blocks

Silicon Mechanics has developed a specific configuration, the Silicon Mechanics Atlas AI Cluster™, which can be modified to meet several high-end HPC and AI workload needs. While this configuration is designed to be a turnkey system that can be plugged in to get up and running fast, it can also be quickly scaled up to a supercomputer level, and/or customized to optimize specific workloads or data types. Whatever project you have in mind, this Linux, building-block style configuration provides enough power for state-of-the-art AI, ML or deep learning (DL) projects.

This complete, rack-scale system starts with the NVIDIA A100 GPU, which offers amazing acceleration to power scalable applications for AI, data analytics, ML and HPC environments. With up to 20 times higher performance over the previous generation, the A100 and its Ampere architecture can run the largest models and datasets for even the most demanding organizations.

- The Silicon Mechanics Atlas AI Cluster integrated product simplifies DL and AI deployments by using Silicon Mechanics' Rackform A354NV or Rackform A380A servers (depending on the cluster size), NVIDIA Mellanox networking, and NVIDIA NGC software with the goal of minimizing time to production.
- The Silicon Mechanics' configuration is based on AMD EPYC™ CPUs and 8X NVIDIA HGX A100 GPUs using both NVIDIA NVLINK® and NVIDIA NVSwitch™. It also supports optional GPU Direct RDMA with up to 8 Mellanox ConnectX®-6 Virtual Protocol Interconnect® (VPI) HDR InfiniBand adapters. In addition, this system supports optional storage nodes based on NVMe storage servers and NVIDIA Mellanox® Spectrum® 2000 Gigabit Ethernet HDR switches.

With up to 20 times higher performance over the previous generation, the A100 and its Ampere architecture can run the largest models and datasets for even the most demanding organizations.

- Software for the Silicon Mechanics Atlas AI Cluster configuration includes Silicon Mechanics' AI Stack, Silicon Mechanics' Scientific Computing Stack, and support for popular HPC frameworks. This includes software such as:
  - **Weka**
  - **Lustre**
  - **S3-compliant object storage**
  - **Ubuntu 20.04 operating system**
  - **TensorFlow**
  - **PyTorch**
  - **Keras**
  - **R**
  - **NVIDIA software tools, including CUDA, cuDNN, and NGC GPU-accelerated containers**

Key Consideration #1: Scalability

Design flexibility

Whether you're building a small, proof-of-concept project or aiming for something bigger from the start, you want to protect your investment and be sure that the system will adapt and grow as your project grows. Design flexibility is key here, with the ability to add nodes or racks to the hardware as needed.

Investment

If you use a customized configuration, your initial investment goes further – you don't need to spend extra money on overhead built into similar, more well-known all-in-one systems that may not be as easy to expand upon incrementally.

Intelligent scalability

It's important to scale intelligently. It does you no good to have a ton of computing boxes with no ability to feed them the data required for training the model. This approach allows you to pay for what you need and not set yourself up for a very expensive solution that gets bottlenecked on either the compute, storage, or networking. This requires intelligent scalability.

As your problem set grows, or if compute and storage requirements change, update, or evolve, the system is designed to scale together seamlessly.

Future growth

These are some of the reasons why the flexible Silicon Mechanics Atlas AI Cluster configuration is designed to support future growth. With each storage node and compute node that you add, the performance of the cluster scales linearly, and can be added seamlessly down the road. As your problem set grows, or if compute and storage requirements change, update, or evolve, the system is designed to scale together seamlessly.

Key Consideration #2: Storage

Scalability without limitations

Large data sets are required to deliver accurate AI results. Having this data drives incredibly large storage demands, and managing these data sets requires a system that can quickly scale without limitations.

"AI is akin to building a rocket ship. You need a huge engine and a lot of fuel. The rocket engine is the learning algorithms but the fuel is the huge amounts of data we can feed to these algorithms."
– Andrew Ng, "The Inevitable: Understanding the 12 Technological Forces That Will Shape Our Future"

This often means lots of compute, but it also means being able to feed that compute. Traditional Network Attached Storage (NAS) is bandwidth limited, so AI projects need to leverage an AI-first storage solution to effectively pull in data. Because the compute is so incredibly powerful, you need a storage solution that is purpose built for AI training scenarios.

HPC-focused systems

High-performance computing has similar issues, but can use traditional parallel file systems that are capable of large streaming data sets. While the two storage systems might end up looking similar physically, an HPC-focused system is more likely to use a Lustre solution, versus an AI system that might use an AI-specific storage solution such as that provided by Weka and an S3-compliant object storage tier.

Storage tiering

Storage tiering is another area that companies need to consider with their system, since it helps ensure minimized cost and maximized availability, performance and recovery. However, not all storage tiering is equal. The key to tiering is to keep things as cost-effective as possible — you don't want to suffer a performance penalty. But keep in mind that not every system needs Ferrari-like storage. An optimized system will help make sure your project has enough space for hot data, balancing the rest with less expensive data storage to meet regulatory or persistence requirements as needed.

Key Consideration #3: Networking

Consider leading-edge technology that can help you get the best possible I/O for all that data. Two examples are below:

1 NVIDIA GPUDirect®

When moving data through an AI or ML algorithm, or training a neural network, you need the highest data throughput possible. GPUs are able to consume data much faster than CPUs, and as GPU computing power increases, so does the demand for IO bandwidth. NVIDIA GPUDirect® can enhance data movement and access for NVIDIA GPUs. With GPUDirect, network adapters and storage drives can directly read and write to/from GPU memory. This eliminates unnecessary memory copies, decreases the CPU overhead, and reduces latency, all resulting in significant performance improvements. Through a comprehensive set of APIs, customers can access GPUDirect Storage, GPUDirect Remote Direct Memory Access (RDMA), GPUDirect Peer to Peer (P2P) and GPUDirect Video.

2 Connect-IB InfiniBand

Connect-IB InfiniBand adapter cards from Mellanox provide the highest performing and most scalable interconnect solution for server and storage systems. Maximum bandwidth is delivered across PCI Express 4.0 leveraging HDR 100 or 200 Gbps InfiniBand, together with consistent low latency across all CPU cores. They also offload CPU protocol processing and data movement from the CPU to the interconnect, maximizing the CPU efficiency and accelerating parallel and data-intensive application performance. It supports data operations such as noncontinuous memory transfers, which eliminate unnecessary data copy operations and CPU overhead. Storage nodes also see improved performance with the higher bandwidth, and standard block-and-file access protocols can leverage InfiniBand RDMA for even more performance.

Taking a Holistic Approach – The Silicon Mechanics Perspective

The Silicon Mechanics Atlas AI Cluster configuration is more than just a combination of the nodes, hardware components and software that sit within the system. Because of the role of GPUs, NVIDIA has been heavily involved in developing this ecosystem to specifically address the requirements for organizations involved in AI, ML, DL, and HPC projects. But, as this system shows, there are many other elements that go into the design, holistically, that allow the functionality required for these workloads.

Everything within the system — the hardware and software — works together to achieve the goals above and enable AI, ML, and HPC. That's why the Silicon Mechanics team takes a holistic approach to each customer's project. From the beginning, they focus on building a balanced solution that is optimized for the particular type of data being produced via the AI project. For example, an AI project that only relies on images will have different storage requirements than a data set with text and images. This might require some customization of the system to fit the project.

Silicon Mechanics engineers have the skills and experience to build this balanced system from the ground up, rather than just relying on a set of pre-configured lists of components that are built to support a vague, generic concept. As the Silicon Mechanics Atlas AI Cluster configuration shows, knowing what elements work together allows the adoption of a building-block model that can quickly — but still effectively — adjust to specific workloads. In addition, the building blocks are already optimized for purpose on their own and, in many cases, tested to work with related technologies, giving you the best of both worlds. The key is having engineers who know all the component elements, all the vendors, new technology, and how they work together.

The result is that you can be confident that your system will give you everything you need for a successful AI, ML and HPC project — not just for now, but for the years ahead.

For more information on this or other use case-specific clusters made using a building block model, visit SiliconMechanics.com/clusters/atlas-ai-cluster.