

The long-term costs of artificial intelligence are not high considering the benefits for companies that make strategic decisions about AI compute and storage.

# Future Proofing Hardware for Artificial Intelligence

August 2022

**Written by:** Penny Madsen, Senior Research Director, IDC BuyerView

## I. Introduction

Successful companies are choosing to carry out artificial intelligence (AI) projects in-house, developing solutions for performance and compliance requirements. These companies are laying a foundation for disruptive change across the business, with AI promising benefits ranging from increased customer satisfaction to better efficiency and decision making. AI, however, can require special infrastructure requirements.

Companies deploying AI need a solution that can deal with large data sets, as well as a highly skilled workforce. The other element is hardware. Luckily, there are many ways to acquire and manage infrastructure in the market today, and storage and compute infrastructure costs are much more easily managed. This is especially true for those companies deploying AI at scale, introducing incremental benefits across the business.

## II. The Argument for AI Investment

For companies that make the up-front investment, the benefits of AI don't take long to surface. Improved customer satisfaction, for example, quickly impacts the bottom line. The simulation of business scenarios and prediction of business risks can also rapidly change the direction in which an organization moves.

Companies with AI already in production are much more likely to have formed enterprise-wide strategies around its use. This places an even greater emphasis on the choice of technologies used to underpin its delivery. It is not only core compute environments that need to be considered. According to IDC's AI InfrastructureView 2021, about three quarters of companies deploying or evaluating AI will look to store data for AI workloads at the edge. Many are also using the edge to increase compute performance.

## AT A GLANCE

### WHAT'S IMPORTANT

- » Investments in new compute and storage for AI are happening at scale.
- » IDC AI InfrastructureView 2021 research shows that high up-front costs are an inhibitor, however, for 32% of companies deploying AI.
- » 81% of companies evaluating AI today plan to upgrade their storage, compared with 77% of those in production.
- » Many in production have already deployed new hardware for AI. They are now more likely to see benefits such as improved customer satisfaction and better decision making as well as lower overall infrastructure costs. These companies are not as impacted by high up-front costs.
- » Attaining these benefits is only possible, however, when organizations factor in future AI use cases and consider options that will meet future AI needs.

### III. It's All About Data

Mature adopters of AI have learned to define data differently than those just starting out. Their broad definition covers data in use by an AI platform, AI applications, and AI-enabled applications as well as data prepared for AI training and inferencing. They are more likely to have central repositories of integrated data collected from more disparate sources.

According to IDC's AI InfrastructureView 2021, 40% of companies with AI in production have data sets of a size that is beyond the capabilities of traditional relational databases, and these are constantly growing. Mature AI users are also deploying content analytics, discovery, and search among other tools that combine business intelligence with business analytics across digital content. While the use of multi-layered machine learning is not as high, it is increasing year over year. IDC's AI InfrastructureView 2021 recognizes that more than 70% of companies using AI data sets said these are about 50TB and 1PB in size. This places an enormous emphasis on the type of storage and compute deployed across the organization, for those with AI in production and those looking to make first investments.

### IV. Maturing Workloads

The workloads used for AI processing determine the types of hardware deployed. The most common workloads in production for mature AI users today are in critical areas such as security, customer relationship management, and networking. Then you have data-focused workloads such as structured data management, application development and testing, and a quickly growing number of specific AI applications that have their own dense workload requirements.

As a first mover in the AI space, a leading Canada-based Nature Fresh Farms places its AI requirements at the heart of all of its infrastructure decisions. The company owns 200 acres of greenhouses that produce vegetables in an AI-optimized way — AI is used to generate better yields and to enhance delivery across its supply chain. It has built out an AI systems using a hub-and-spoke approach, with edge locations (grower locations) feeding back into a single datacenter core at its large farm in Ontario. It is constantly reassessing its infrastructure requirements. For this reason, it has architected its systems to scale using containerized hyperconverged infrastructure with graphics processing unit (GPU) options.

Nature Fresh Farm's hardware choices are about simplifying and accelerating the AI experience. Many of Nature Fresh Farm's projects are delivered in an agile fashion, and each new project brings additional data requirements. The average rate of data Nature Fresh Farms processes daily is up to 1TB. On top of this, Nature Fresh Farm must also guarantee that its AI systems can connect with more traditional applications and environments.

### V. Benefits

Companies that focus on future proofing their hardware investments for future AI use will be able to ensure they have environments that are optimized for the delivery of AI projects, from the core to new edge capabilities. The right approach to planning for AI rollouts ensures that companies can easily scale their AI environments to support incoming data while ensuring consistent operations without the risk of downtime at times of capacity challenges. Thinking ahead will allow AI users to establish the required vendor relationships and ensure a supply chain for delivery of hardware (important once AI becomes a more agile and in-demand practice across the business).

Many companies carrying out AI projects speak of increasing demand, and capacity requirements that can increase in waves. The other benefit of future proofing is that it allows companies to form organization-wide strategies around AI's use and deployments, offering a more coordinated vision across the company and helping promote future use of AI.

## VI. Considerations

There are a number of considerations companies must evaluate before they make infrastructure investments across the business. These include:

- » **Future workloads, applications, and required capacity.** This starts at the pilot phase. Individual pilots can often be run on traditional hardware, but multiple pilots can place significant strain on existing systems. This is where most companies are typically forced to start considering additional hardware requirements, from GPU acceleration to scalable storage systems. In many cases, part of this thought process should include a longer-term vision for what hardware will be used and how AI will be used across the business.
- » **Training and staff.** Are you properly equipped to run the new infrastructure purchased? Understand the capabilities of your IT department to get proper business buy-in.
- » **Payment models.** Are up-front investments or flexible payment/rollout models better for your business and use cases? If flexibility is the chosen route, carefully consider order and delivery cycles, and keep this in mind as you schedule new projects.
- » **Learning from your peers.** Consider who else in the market has carried out similar AI projects. In many cases, companies are happy to share findings from pioneering AI projects. There are also many industry associations, research laboratories, and vendors that are willing to offer support when it comes to planning your AI projects and infrastructure purchases.
- » **Asking your vendors about the next iterations of storage and compute designs.** Hardware designed specifically for AI has come a long way in a short amount of time, and new developments are happening all the time in processor, storage I/O, GPU, and other technologies. Ask your vendor what is best for you right now, and what could be better for you in the future. Also consider if you require hardware and software solutions that can help you scale faster.
- » **Budgeting accordingly/considering return on investment.** Understand that AI is an up-front commitment that will continue as it scales, but that its benefits often equate to efficiencies, savings, or both across other parts of the business.
- » **How your environment will scale.** As well as where it will be housed, you need to know what network requirements will be required and what remote management and other tools like security will come into play. You also need to know how your AI environment will interact with traditional IT and who at your business will manage all of this.

## VII. Conclusion

The road to efficient, scalable, and successful AI deployments can be full of pitfalls that can lead to inefficiencies and increase cost if proper due diligence is not carried out. Business road map plans, solution considerations, and vendor discussions are highly important when embarking on an AI journey. Once a company sets out on a path, the benefits of proper planning will multiply as capabilities become requested across the organization. Laying the right foundation early allows companies to better respond to these requests, not only improving buy-in from across the organization but also driving better efficiencies and project outcomes in the long term.

Luckily, options in the market today mean that storage and compute infrastructure costs can be contained over the long term.

## About the Analyst



### ***Penny Madsen, Senior Research Director, IDC BuyerView***

Penny Madsen is a senior research director for IDC's BuyerView, a portfolio of primary research products that provide insights into end user use and adoption of AI, cloud, edge and related infrastructure platform technologies and services. Her research covers trends, investments and purchase decisions, helping leading vendors and infrastructure providers develop strategies for future customer deployment scenarios.

## MESSAGE FROM THE SPONSOR

AI users clearly see many great benefits thanks to AI technology. However, it is not too late for those who have not yet 'jumped on the bandwagon.'

1. AI Users are leveraging their expertise and infrastructure widely. With the Dell Validated Design for AI, you can quickly build AI architectures for your outcomes. Focus on deploying a comprehensive strategy, simplifying IT to provide faster, deeper insights.
2. AI leaders focus on performance, scale and cost. With the Validated Solution for AI, you can run many AI workloads in your existing VMWare® environment and power your workloads with a best-in-class on-premises to cloud infrastructure.
3. AI Leaders accelerate results with on-premises GPU accelerators. Dell and NVIDIA solutions help organizations harness the power of AI from anywhere, from core to edge easily.

No matter where you are on your journey, Dell Technologies helps organizations become AI leaders. Visit [Dell.com/AI](https://Dell.com/AI) to learn more.



The content in this paper was adapted from existing IDC research published on [www.idc.com](http://www.idc.com).

**IDC Research, Inc.**  
140 Kendrick Street  
Building B  
Needham, MA 02494, USA  
T 508.872.8200  
F 508.935.4015  
Twitter @IDC  
[idc-insights-community.com](http://idc-insights-community.com)  
[www.idc.com](http://www.idc.com)

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2022 IDC. Reproduction without written permission is completely forbidden.