

ZIFF, a provider of advanced AI and deep learning techniques for B2B and B2C businesses, uses FlashBlade™ from Pure Storage® as the core data platform to support its most demanding AI workloads, which can now be processed with unprecedented speed.



BUSINESS TRANSFORMATION

ZIFF is able to deliver unprecedented levels of AI and deep learning services to its clients, saving months of time and hundreds of thousands of dollars while producing answers to questions that even global enterprises have not been able to solve.

GEO

North America

INDUSTRY

Technology

“At ZIFF we want to help CEOs fall in love again with their business using transformational AI.”

David Gonzalez, CEO

ZIFF BREAKS THROUGH THE AI ANALYTICS BARRIER WITH FLASHBLADE FROM PURE STORAGE

There’s an adage among innovative CEOs that “tame business leaders rarely make history.” Those willing to apply disruptive thinking to gain entirely new insights outside the scope of traditional business practices often make or shape industries. That’s why today’s savvy CEOs are looking to artificial intelligence (AI) and deep learning as exciting new pathways to success. That said, they often encounter a barrier: most organizations, even global enterprises, lack the human and capital resources capable of applying AI and deep learning techniques to their unique business problems. Helping companies overcome this barrier is the mission of ZIFF Inc.

“ZIFF is built on the proposition that most product companies do not have the analytic horsepower to solve their user-experience problems with AI or machine learning,” said David Gonzalez, co-founder and CEO of ZIFF. Added Ben Taylor, the other co-founder and Chief AI Officer, “There are many reasons for this. The process is very complex; there are many frameworks to choose from; and a few very large companies have scooped up virtually all the talent. Companies want to do this work in-house, but they simply lack the people and experience. We know of several Fortune 500 companies who have been working on projects for two years or more and still haven’t succeeded.”

ZIFF’s value proposition to CEOs is to allow organizations with insufficient in-house analytic capabilities to incorporate AI into their core products and services. ZIFF applies its expertise to help companies in all three phases of an AI project: Gathering a data set that is large enough to provide accuracy; analyzing that data; and then deploying the results of that analysis to the real-world problem the client is trying to solve.

For example, the CEO of a large U.S. auto insurer wanted to reduce the time and cost of assessing vehicle damage by trying to automate the process. This involved more than four million photos of damaged cars compiled over several years that had to be labeled by company employees. The company’s data scientists examined this data set for two months and concluded that the human labels were not accurate enough to form the basis of an AI project.

ZIFF applied its data-curation software and generated labels for the four million photos that were 30% more accurate than the human labels, all in less than 24 hours. The new algorithm has now been applied to the insurer’s claims-assessment process and is producing three times the number of estimates per day compared to previous methods.

COMPANY:

ZIFF
ziff.ai

USE CASE:

- FlashBlade – Artificial intelligence, deep learning
- Database – Postgres®

CHALLENGES:

- Data-storage I/O was a persistent bottleneck.
- Petabyte-scale data files overtaxed storage infrastructure.
- Clients want ever-more-complex questions answered.

IT TRANSFORMATION:

- I/O has disappeared as an issue of concern.
- High-value data scientists are more productive by maintaining higher utilization levels.
- Clients receive answers and actionable insights in less time and at lower cost.

“We have tripled the workload we can handle since installing FlashBlade.”

Ben Taylor, *Chief AI Officer*

DATA BOTTLENECKS POSED A HUGE CHALLENGE

Currently, ZIFF applies its curation software to data sets built from images, video and audio, and each data set can run into the terabytes or even petabytes. Handling that data efficiently has been a persistent challenge for ZIFF’s co-founders, long before they started their new company.

As a former chief data scientist, Taylor built a custom system using multiple graphics processing units (GPUs) and a combination of hard-disk and solid-state storage devices. “We had a lot of great researchers working on exciting deep learning projects, but we were always running out of storage. You would have one researcher who wanted to kick another researcher’s project off the system. We had to negotiate all those conflicts, and as a result we were spending hours, even days, just shuffling data on and off the system, which added no value and frustrated everyone.”

At another point, Taylor added, “we had an NFS server, and I remember having all sorts of problems with it, like NFS mounts going down all the time.”

Recalled Gonzalez: “There never was a time when data I/O was not a top-of-mind problem. And if you can’t get data in and out quickly, you lower the utilization of your most valuable resources — your GPUs and your data scientists.”

FLASHBLADE: MODERN DATA PLATFORM FOR AI

When it came time to build the IT infrastructure for ZIFF, Taylor looked at a number of options for storage and found one uniquely suited for his company’s needs: FlashBlade from Pure Storage. FlashBlade is the first all-flash data platform purpose-built to handle massively parallel problems such as data analytics, artificial intelligence and deep learning.

FlashBlade met all of ZIFF’s criteria for a modern data platform: unmatched I/O; performance; virtually infinite scalability; a small footprint that reduced operating costs; and effortless management. ZIFF installed a FlashBlade chassis with seven 8TB blades, for usable capacity of more than 70TB.

The impact of FlashBlade on ZIFF’s business operations has been significant.

“FlashBlade eliminated the I/O issue,” Gonzalez reported. “We have all the performance we need, and more. We were able to index four million images for an auto insurer in 24 hours. There’s no way we could have done that in so short a time without FlashBlade.”

By slashing processing times, ZIFF is not only able to take on more customer projects, it can handle more complex ones as well. “We tripled the workload since installing FlashBlade.” Taylor noted. “We are no longer shuffling data sets through the cloud or on a SAN. With FlashBlade, millions of images can be loaded in an hour, whereas the alternative was either much, much longer, or possibly not at all.” It saved time and eliminated interruptions to productivity.

One of the most significant impacts FlashBlade has had is its ability to support ZIFF’s “deep indexing,” which uses deep learning techniques to curate huge volumes of images in a small amount of time with high accuracy. “FlashBlade can handle huge data sets without any problem. It allows us to do very fast indexing — around 3,000-4,000 images per second — giving us the ability to handle projects for clients who have tens or even hundreds of millions of images,” said Taylor.

The power of ZIFF’s software and FlashBlade was evident in a project ZIFF took on for Chatbooks, the popular app for automatically turning collections of photos into printable photobooks. Because Chatbooks stores images that may never be used in a photobook,

it wanted a way to quickly but accurately identify “garbage” that could be filtered out. A data scientist at Chatbooks spent six weeks manually labeling images, and eventually compiled a data set of 20,000 images, with the goal of using that subset as a way to train the rest of the images in Chatbooks’ database of millions of images.

Using the ZIFF to provide AI Assisted Curation on FlashBlade a 14 million image training set was cleaned and new candidate models were trained all in less than 24 hours. “With FlashBlade, the combination of higher throughput and the elimination of data shuffling has reduced the time-to-value for our clients,” Taylor said.

DELIVERING RELIABILITY AND CLOUD-LIKE SIMPLICITY

Other significant advantages of FlashBlade, according to ZIFF’s co-founders, are its reliability and minimal management requirements. “I no longer have to go to the data center and fiddle with drives or deal with an NFS mount crashing,” Taylor said. “So, the reliability of FlashBlade is a huge benefit for us since uptime translates directly into scaling our workloads.

“I call it ‘forgetability’. FlashBlade is so easy to use and so reliable, that we can just forget it is there.”

The biggest benefit of FlashBlade, Gonzalez added, “is that it keeps our data scientists productive. They’re working on real business problems, not spending time debugging hardware or waiting for data to be shuttled off and on the cloud. Their productivity is the most important cost factor in our business and keeping them busy on interesting work is the best way to retain them as members of our team.”

An important impact of FlashBlade’s performance is one that may go unnoticed, Gonzalez observed. “With an AI project, a win always follows many failures, and you are always iterating to try out new approaches or improve accuracy. So, the victory is not being able to train a model on 20 million images. The real victory is being able to do that six or seven times on different data sets. FlashBlade allows us to do that, by reducing the time it takes to train a model to as little as 24-48 hours, down from as much as two or three weeks.”

In addition to taking on more projects for clients, ZIFF is able to use FlashBlade for internal workloads as well. “We have been able to load open-source code and internal data sets onto FlashBlade that we would not have been able to do with another storage approach. This has allowed our data scientists to iterate more and learn faster,” Taylor said.

Furthermore, Gonzalez noted, the ability to put multiple, varied data sets and workloads on FlashBlade “allows us to keep our GPUs busy. In addition to handling all of our client work, we’re able to feed internal research projects and other models, and all of them run on FlashBlade concurrently.”

Looking ahead, ZIFF sees real-time analysis as the next big step in AI. “You’ve always had what we call the ‘ceremony’ in AI projects — getting the data, training it and then consuming it,” Taylor said. “With the rate at which the FlashBlade platform moves data, we see the ability to stream data through an AI engine at the same time someone is consuming the results, even if they are a mobile user. That could have huge implications for healthcare, call centers, or national security.”

“I call it ‘forgetability’.
FlashBlade is so easy to use
and so reliable, that we can
just forget it is there.”

Ben Taylor, *Chief AI Officer*



info@purestorage.com
www.purestorage.com/customers