

CASE STUDY

Cogo Labs

Cogo Labs deploys Pure Storage FlashBlade™ products as a data hub, accelerating the move to containers on Kubernetes and allowing the company to improve its rigorous and creative application of data science in the creation of new business opportunities.

Business Transformation

Data scientists and software developers are substantially more productive, supporting the company's core mission of launching successful new companies. Storage management costs are sharply reduced, accelerating the move from cloud to on-premise solutions.

Challenges

- Legacy storage had reached capacity and the vendor's suggested replacement couldn't match data-management requirements.
- A new data center required a modern, infinitely scalable IT infrastructure.

Solution Benefits

- "Blindingly fast storage" enables data scientists and developers to be considerably more productive.
- Tight integration between Kubernetes and FlashBlade accelerates the transition to a new containerization management platform.
- Workloads and data can be pulled back from the cloud, lowering operating costs.



Geo

North America

Industry

Business Services

Company

Cogo Labs

www.cogolabs.com

Use Case

- Pure Storage FlashBlade™
- Pure Storage Orchestrator™ container storage
- VMware® vSphere® virtual server infrastructure
- Kubernetes container management



Cogo Labs Accelerates Formation of New Businesses with Kubernetes on Pure Storage FlashBlade

Cogo Labs is a unique incubator. It incubates new business teams internally, then grows them to profitability through the efforts of dedicated personnel and Cogo's experienced engineering, analytics, and design teams. Once one of these internal business units is mature enough, it is spun out as its own company with Cogo Labs continuing to give them the support they will need to grow quickly and succeed.

"In contrast to the model of a venture-capital firm or a traditional incubator, all the companies in our portfolio are companies we have started ourselves with ideas generated internally," according to Ed Hammond, Principal Solutions Architect and a 12-year veteran of the Cambridge, Massachusetts, company. "We have 150 of the brightest quants [quantitative analysts] and engineers in the business."

"To identify new areas of business, Cogo will draw on its massive proprietary data surrounding current internet trends to identify high-potential business ideas where we see promise and think we can build a better mousetrap to better capitalize within that space," Hammond said.

Portfolio companies include:

- EverQuote, an online insurance marketplace
- Autotegrity, an auto shopping service that ADP has since acquired
- CareDash, an online portal for rating healthcare providers
- HopJump, a provider of personalized travel recommendations and content.

The common denominator? Rigorous and creative application of data science. As a result, data is one of Cogo's most valuable assets.

Opening New Data Hub Opportunities

With its mission-critical databases growing rapidly, storage management is a high priority for the IT staff. "While some of our systems use local storage, the data for all our virtual machines had been supported on two VNX SANs from DellEMC," Hammond noted. "Many of our high-performance users didn't have fast enough storage because their applications were pushing the I/O limits on the EMC. In addition, the VNX systems were reaching their capacity limits."

Hammond's first call was to DellEMC. "I wasn't impressed with what they offered, with all of the technology being at least three years old. Their offerings made no mention of NVMe and object storage was something on a roadmap nobody could quite give me," he said. "I had heard good things about Pure Storage, attended a local event, liked what I heard, and the rest is history. Cogo and Pure have had a great partnership these past three years."

“Using Pure Storage FlashBlade solution. FlashBlade is a scale-out, massively parallel data hub that enables consolidation of multiple workloads on a single platform. With unprecedented performance, scalability, and flexibility, FlashBlade empowers organizations to manage massive volumes of data and tackle data-intensive workloads ranging from data warehouses to artificial intelligence.

Hammond decided to go with a Pure Storage FlashBlade solution. FlashBlade is a scale-out, massively parallel data hub that enables consolidation of multiple workloads on a single platform. With unprecedented performance, scalability, and flexibility, FlashBlade empowers organizations to manage massive volumes of data and tackle data-intensive workloads ranging from data warehouses to artificial intelligence.

Hammond admits that the FlashBlade hub was “probably overkill” for his immediate need to support 400 virtual machines. But his team was impressed by the advanced architecture, scalability, and performance potential. “The concept of a data hub fascinated us, given our escalating data-management needs and focus on cost efficiency.”

Finding New Applications for “Blindingly Fast Storage”

After installing a FlashBlade configured with five 17TB blades, many of Cogo’s production applications were quickly moved onto it. Workloads include: Cassandra, Elasticsearch, Spark, RabbitMQ, Airflow, MySQL, PostgreSQL, Enterprise GitHub, etcd and control planes for Kubernetes, Developer VMs, Travis, CodeCov, Sentry, Ansible, Munin, and on-prem S3. All coexist on FlashBlade with no contention issues.

“The ability of FlashBlade to host such a wide variety of protocols and workloads is one of its most impressive features,” Hammond said.

At the time, Hammond was in the process of pulling back some workloads and data from the cloud, making the S3 integration on FlashBlade very valuable. “Having on-prem S3 compliant storage was key for us to start our migration of some workloads to our co-lo from the cloud. Had we kept things as they were, it would be prohibitively expensive to run all our jobs in the cloud,” he explained. “The storage is expensive and you are buying the required compute hardware every three months.”

From day one, Cogo Labs experienced what Hammond calls “blindingly fast storage” that eliminated previous I/O concerns. “Everyone benefited. We were able to move our entire platform onto the FlashBlade product,” he said. “What we expected to be a refresh of the EMC wound up being a single FlashBlade, which has performed beyond our wildest expectations.”

Having solved Cogo’s immediate need for faster storage, Hammond noted that “Just the load of our VMs was not pushing our FlashBlade at all, so we started looking at other ways we could better leverage it.”

ED HAMMOND,
PRINCIPAL SOLUTIONS ARCHITECT

“The ability of FlashBlade to host such a wide variety of diverse workloads is one of its most impressive features.”

ED HAMMOND,
PRINCIPAL SOLUTIONS ARCHITECT

At the time, Cogo was in the early stages of a massive move to Kubernetes as the management platform for its containerization strategy. “It was a game-changer to learn about Pure Storage Orchestrator (PSO) and how much that could simplify things for us,” Hammond recalled. “We followed the instructions, fired it up, had an issue, submitted a support ticket, realized the fault was ours and had it up and running in less than the 12 mins it took support to reply with the right solution. With PSO, you can simply have Kubernetes fire up a container, tell it how much transient/persistent storage is required and you’re off to the races. We were able to triple our database workloads right away.” The Kubernetes deployment incorporates Elasticsearch, Cassandra, InfluxDB, Spark, Prometheus, and Grafana.

Joe Presbrey, Cogo’s Chief Engineer, observed that FlashBlade is ideal for high-volume transactional workloads. “We have one InfluxDB instance backed by Pure Storage that’s doing about 300 million to 350 million events a day. And we’re persisting that for months and doing instantaneous rollups,” he explained. “There are trillions of data points per instance. Having the performance of the Pure FlashBlade makes this a no-brainer for us. We have moved from I/O wait on storage IOPS to compute-bound limitations and shipping delays on Intel CPUs.”

At any one time, about fifteen Cogo-incubated companies are drawing on insights from the InfluxDB database using Grafana for end-user visualization and monitoring.

“The combination of FlashBlade and Kubernetes through Pure Storage Orchestrator was such a success that we’re basing our entire infrastructure at our new data center on it,” Hammond reported. The new data center will initially house 200 Kubernetes compute nodes with four FlashBlades using Pure Storage Orchestrator (PSO) dynamic provisioning with almost no local storage.

PSO delivers container storage-as-a-service to build and deploy scale-out, microservice-based applications on demand for dynamic containerized environments. PSO functions as the control plane, integrating seamlessly with container-optimization frameworks like Kubernetes.

CASE STUDY

“The performance, resiliency, and deep turn-key integration of FlashBlade allow our developers to iterate effortlessly, focus on delivering business value, and sleep at night, knowing nothing is going to go wrong.”

JOE PRESBREY, CHIEF ENGINEER

“The number-one benefit we saw was the reduction of complexity and the cloud-like experience PSO gives us, even though we’re actually on bare metal and saving a ton of money being on-premises versus in the cloud,” Presbrey said. “The performance, resiliency, and deep turn-key integration of FlashBlade allow our developers to iterate effortlessly, focus on delivering business value, and sleep at night, knowing nothing is going to go wrong.”

Cogo Labs has found FlashBlade to be a highly cost-effective solution for moving large volumes of data with exceptional ease-of-use and long-term scalability.

“Everything we have thrown at FlashBlade hasn’t even begun to push its limits,” Hammond said. “VM storage has proven to be a great solution for almost all our storage needs, with infinite possibilities to come.”

purestorage.com

800.379.PURE



©2019 Pure Storage, Inc. All rights reserved. Pure Storage, the P logo mark, FlashBlade, and Pure Storage Orchestrator are trademarks or registered trademarks of Pure Storage, Inc. All other names may be trademarks of their respective owners.

ps_cs_cogolabs_01

cogolabs.com | 781.228.5678 | info@cogolabs.com

