

FRED HUTCHINSON Cancer Research Center



Fred Hutchinson
Cancer Research Center Campus
Seattle, WA

“Silicon Mechanics is unique in that it really lives and breathes by its consultation skills and ability to not only provide its customers with solutions targeted at their needs, but also expert advice on that system going forward.”

Dirk Petersen
Director of Scientific Computing
Fred Hutchinson Cancer Research Center

Challenge

As the Director of Scientific Computing for the Fred Hutchinson Cancer Research Center in Seattle, Dirk Petersen’s internal IT organization stores and catalogues large amounts of unstructured and genomic data, all of which is critical to the organization and its many constituents. A data loss caused by computer system failure would be problematic for the organization and its researchers at best, and catastrophic at worst.

IT is given a limited budget each year to purchase and maintain researchers’ storage systems, making it difficult to afford the inherent cost and overhead of classic storage manufacturers for all data sets. And not just in terms of hard cost. These storage platform purchases can require significant up-front investment, hinder the ability to mix and match solutions along the way and force an organization’s IT team to conduct dreaded forklift upgrades that drain resources.

So when it came time to implement a new storage system, in part to meet the demands created by the recent explosion in genomic sequencing, but also to meet the growing IT needs of Fred Hutch’s 2,000-plus researchers, Petersen was in a seemingly difficult spot. The chosen system would need to satisfy the finance department, IT department and researchers’ desire for a cost-effective solution that was also reliable, scalable and customizable.

“Our goal was to create a scalable storage hardware infrastructure while keeping upfront costs low,” says Petersen. “This was extremely important because like most finance departments, ours does not enjoy surprises. And it was equally important that we have the flexibility for incremental growth – with the use of genomic data sets continuing to surge, a majority of our future infrastructure costs will be storage related. Maintaining control over the system infrastructure, affording us the ability to add or subtract pieces as necessary, is critical to remaining nimble.”

Solution

Petersen approached the possible solutions from a scientists’ perspective: researching the problem, talking to his peers, working with professional product testers and doing significant in-house testing. He considered various options including open source, distributed file systems, fully integrated and commercial hierarchal storage management solutions and commercial object storage vendors. None of the above, however, could meet Petersen’s cost, maintainability and access requirements.

Ultimately, he decided that a Swift implementation made the most sense. What is Swift? The OpenStack Object Store project, better known in the technology industry as Swift, offers cloud storage software that allows users to store and retrieve large volumes of data with a simple API. It is also compatible with the S3 API which means users can have an Amazon S3 in their own data center at 40-percent of the cost and are not be constrained by internet latency. Swift was built to help large organizations; it is optimized for durability, availability and concurrency across an entire data set. And in meeting Petersen’s needs even further, Swift offers an ideal platform for storing unstructured data in addition to genomic sequencing data. Unstructured data, which includes mostly text, but also rich media, currently has a growth curve substantially greater than that of structured data.

Fred Hutchinson Cancer Research Center

To implement the new Swift platform, Petersen engaged with SwiftStack, an enterprise level data storage solutions provider, and his current systems integrator, Silicon Mechanics, which provided system consultation and ultimately built Petersen's SwiftStack hardware configuration. In doing so, Petersen completely bypassed the original equipment manufacturers that have dominated the enterprise storage market for the past two decades, and put together a system that not only performs its duty in storing valuable data, but also provides his organization tremendous savings in terms of money, resources and downtime.

Says Petersen, "I'm bucking old school buying trends. In the IT world it used to be common knowledge that you wanted to avoid the road less travelled, but that is quickly changing. I think this trend is evident to others as well, because the white box companies -- system integrators like Silicon Mechanics -- are now increasing market share while the OEMs like IBM, HP and Dell are losing theirs. The value of Silicon Mechanics and SwiftStack to an IT organization like mine is that they offer variety -- I like to compare the OEM's to a static Playmobil set, where Silicon Mechanics offers more of an interchangeable collection of LEGO blocks."

Results

Petersen's Swift implementation has thus far kept his organization's data secure and the finance department happy. In fact, it has been so successful that he has been invited on multiple occasions to present the benefits of the architecture to industry conferences such as Bio-It World, most recently at its April 2015 event. But the most revealing implementation result shows on the faces of the IT staffers in the operations department.

"Smiles, all around," he says. "And now they look very confident too."

From a financial standpoint, Petersen's efforts certainly saved IT on its upfront infrastructure investment by engaging with SwiftStack and having Silicon Mechanics design and build the system. Since implementing Swift to store more than 300 TB of his organization's 1PB-plus of data, IT is now spending a mere \$4,000 per month in an era when storage costs are skyrocketing. Just on storage costs alone, he is saving his organization hundreds of thousands of dollars per year versus the next cheapest alternative.

Compared with other storage solutions, Petersen's evaluation proved that the SwiftStack/Silicon Mechanics option was:

- 76% cheaper per terabyte per month than NAS
- 60% cheaper per terabyte per month than Amazon
- 56% cheaper per terabyte per month than Google

Moreover, Petersen's selection has also saved the organization over the long haul by having the systems integrator provide ongoing consultative maintenance on that system.

"Silicon Mechanics is unique in that it really lives and breathes by its consultation skills and ability to not only provide its customers with solutions targeted at their needs, but also expert advice on that system going forward. In fact, they typically identify solutions to my needs that I hadn't thought of. And if, for some reason I do have a problem I simply call my rep and it gets fixed right away. The sales reps at Silicon Mechanics are technically are very savvy because they work very closely with the technicians."

Petersen is blazing a trail in the research industry by implementing a cloud-based, object storage system. In addition he is providing his peers in the research IT field a view of a storage implementation that performs on par with the industry standard's, but cost far less in terms of money, resources and headaches.

"In the first 4 months the new system grew by 300TB," says Petersen. "This alone saved us \$700,000 compared to using an Enterprise NAS system."

"The value of Silicon Mechanics and SwiftStack to an IT organization like mine is that they offer variety -- I like to compare the OEM's to a static Playmobil set, where Silicon Mechanics offers more of an interchangeable collection of LEGO blocks."

Dirk Petersen
Director of Scientific Computing
Fred Hutchinson Cancer Research Center