IntegriChain® Unleashes Idle CPUs to Bolster Its Customers' Business Intelligence

Pharmaceutical business intelligence provider knocks down I/O bottleneck and multi-threads for a 16X more powerful system.

Solution Focus

- MySQL
- Big Data
- Software as a Service (SaaS)
- Healthcare

Summary of Benefits

- End-to-end solutions support from Dell and SanDisk
- 16X faster overall throughput
- 3X faster jobs than a CPU-constrained in-memory database
- 4-5X faster ingest jobs
- Deliver products weekly instead of monthly
- Guarantee faster visualization jobs for all customers, at any time of day

The Challenge

IntegriChain's flagship product offering, *Demand Network Analytics™ (DNA)* helps big pharma companies aggregate and analyze their inventory and sales to better understand their business. Founded five years ago, IntegriChain has eight of the top 12 US pharmaceutical companies using its *DNA* tools. IntegriChain processes over 45 percent of US channel data (by dollar volume) including cleansing, correlation, and delivery to their customers, who then slice and dice it.

As *DNA*'s customer base grew, Senior Director of Data Management, David Dunn found that the company's MySQL-based systems were reaching their limits. David explained, "Most of our products generate workloads that consist of small transactions in high volume and high frequency. For example, one of our customers generates 142 million transactions every month on queries that analyze historical data. The 450 IOPS capability of hard disks just couldn't keep up."

IntegriChain needed a solution that could help *DNA* with the following:

1. Improve application performance to keep up with a growing customer base with expanding data needs, without requiring major code or architecture changes
2. Free IT to stop tuning system performance so it could support new products and product functionality, increasing the value *DNA* offers customers
3. Cost-effectively scale to support future innovation

The SanDisk® Solution

End-to-End Solution Support from the Experts

David explained how Dell's staff helped him choose the best servers, processors, and Fusion ioMemory™ devices to meet the company's ambitious objectives: "Dell was fantastic to work with, giving us servers with the compute power to meet our processing needs. They connected us with knowledgeable SanDisk® support that helped us implement the Fusion ioMemory™ ioDrive® devices to optimally use Dell server resources. The end result was a very happy customer."

Painlessly Blowing Away the I/O Bottleneck

In its own case study on IntegriChain, Dell described how combining Fusion ioMemory™ devices with upgraded Dell servers and Intel processors resulted in a drop-in solution that achieved three times the performance with 30% lower latency.
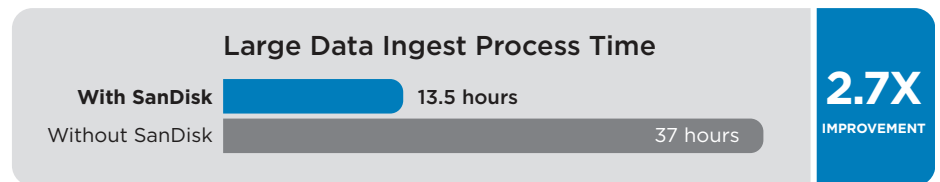
"We tried an in-memory database on a Dell 910, but the application had a formatting

and encryption pipe that spiked the CPU at 2.3 GHz," David said. "The processor on the Dell 710 was much faster and Fusion ioMemory enabled the application to use these CPU cycles more effectively. The test job we ran completed three times faster, reducing the job time from 9 minutes to 3 minutes."

Putting CPUs Back to Work

David then described IntegriChain's experience after deploying the SanDisk powered Dell systems into production. "After seeing the proof-of-concept results, we realized implementing ioDrive was an easy hardware fix to some of our other I/O bottlenecks," David said. "We deployed ioDrive Duo into our ETL (Extract, Transform, Load) engine. Immediately, this cut a process from 37.5 hours to 13.5. Further adjustments to the ETL software brought the run time to less than 7 hours."

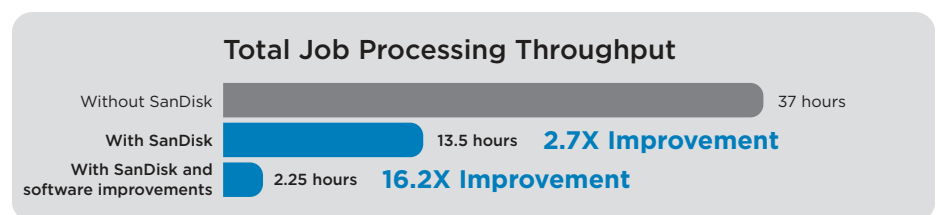This chart shows results of one of their largest ingest jobs.

Large Data Ingest Process Time

With SanDisk	13.5 hours
Without SanDisk	37 hours

2.7X IMPROVEMENT

But IntegriChain was just getting started. David continued, "Our previous system had met its architectural limits. The ioDrive gave us a drop-in solution that improved performance by three times. This bought us the time to properly re-architect our system to be multi-threaded."

The upgraded application, combined with Fusion ioMemory's ability to give multiple cores simultaneous access to the flash, delivered astounding results. David said, "Now, not only do jobs run faster, but we can use the six processor cores to run six jobs simultaneously, fully using the multicore CPUs. This essentially improved our overall throughput by over 16 times."
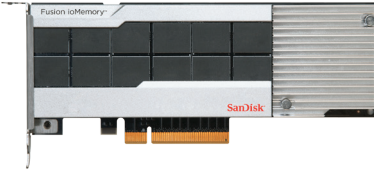
Delivering More Products and Services on the Same Hardware

Total Job Processing Throughput

Without SanDisk	37 hours	
With SanDisk	13.5 hours	**2.7X Improvement**
With SanDisk and software improvements	2.25 hours	**16.2X Improvement**

Upgrading their application to put multicore CPUs to work had an unexpected but significant benefit to the ROI on IntegriChain's purchase. "We initially bought capacity based on single-threaded computing," David said, "but after multi-threading, we found we could do much more with the same hardware. Since deploying ioDrive, we have expanded from a single product delivered in 37 hours to delivering three different product spaces a week for each of our twenty customers— all on the same hardware."
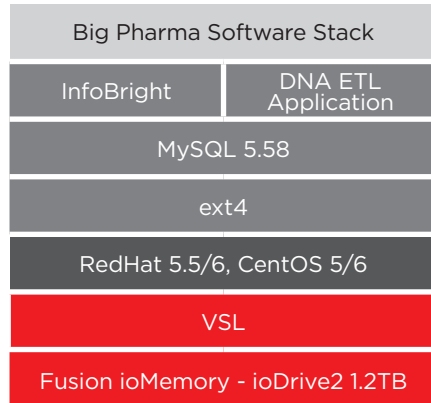
Performance with a Purpose: Better Customer Intelligence

IntegriChain's goal for *DNA* was to make it a single source that customers could use to quickly and accurately view and generate reports for all of their product distribution channels. David summarized how IntegriChain is using *DNA*'s lightning-fast performance to help its customers, "Jobs we used to deliver monthly, we now deliver weekly—and are looking at delivering daily. We can now return views to every customer we have within a three-hour window, under any load. This enables our customers to be better informed so they can respond more quickly and intelligently to the needs of the market, and gives them a competitive advantage."

"The ioDrive devices gave us a drop-in solution that improved performance by three times. This bought us the time to properly re-architect our system to be multi-threaded. Now, we can run six jobs simultaneously [improving] our overall throughput by over 16 times."

**David Dunn,
Senior Director of Data Management,
IntegriChain**

Fusion ioMemory™ - ioDrive®2

Performance Density

		18.4X
With SanDisk		IMPROVEMENT
Without SanDisk		

Workload support. 7/3 = 2.3X more products for 2X customers, 4X faster (monthly > weekly). 2.3*2*4 = 18.4X.

SanDisk Powered Software Stack

Big Pharma Software Stack	
InfoBright	DNA ETL Application
MySQL 5.58	
ext4	
RedHat 5.5/6, CentOS 5/6	
VSL	
Fusion ioMemory - ioDrive2 1.2TB	

Contact information

sales-dell@sandisk.com

Western Digital Technologies, Inc.
951 SanDisk Drive
Milpitas, CA 95035-7933, USA
T: 1-800-578-6007

Western Digital Technologies, Inc. is the seller of record and licensee in the Americas of SanDisk® products.

SanDisk Europe, Middle East, Africa
Unit 100, Airside Business Park
Swords, County Dublin, Ireland
T: 1-800-578-6007

SanDisk Asia Pacific
Suite C, D, E, 23/F, No. 918 Middle
Huahai Road, Jiu Shi Renaissance Building
Shanghai, 20031, P.R. China
T: 1-800-578-6007

For more information, please visit:
www.sandisk.com/dell

SanDisk®
a Western Digital brand

At SanDisk, we're expanding the possibilities of data storage. For more than 25 years, SanDisk's ideas have helped transform the industry, delivering next generation storage solutions for consumers and businesses around the globe.
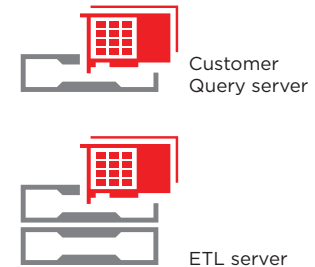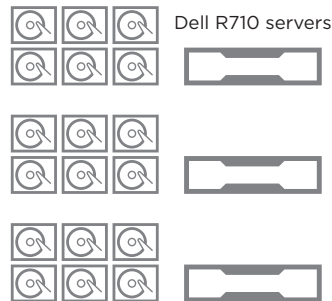
System Overview

System Before	System After
2 to 3 products	7 products
10 customers	20 customers
Delivered weekly/monthly	Daily and weekly

Dell R710 servers

- 3 x Dell R710 servers,
  - Quad-core 3.46 GHz, Intel processors
  - 96GB RAM
- OS: RedHat 5.3
- Application
  - MySQL
  - Custom ETL application
- Hard disks: 6 x 600GB 15K drives

Customer Query server

ETL server

- 1 x InfoBright Customer Query server, Dell T710, hex-core 3.46 GHz Intel processor, 192GB RAM
- 2 x ETL servers, Dell T710, hex-core 3.46 GHz Intel processor, 192GB RAM

Summary

Implementing Fusion ioMemory gave IntegriChain the following benefits:

- End-to-end solutions support from Dell and SanDisk
- 16X faster overall throughput
- 3X faster jobs than a CPU-constrained in-memory database
- 4-5X faster ingest jobs
- Deliver products weekly instead of monthly
- Guarantee faster visualization jobs for all customers, at any time of day

"The platform's capability continues to expand our ability to deliver," David said. "The ioDrive devices have enabled us to generate full ETLs on customer data for four different products—and we have more products coming to the market."

About IntegriChain and Demand Network Analytics™ (*DNA*)

IntegriChain, Inc. is the developer of Demand Network Analytics™ (*DNA*). Launched in 2007, *DNA* is the world's largest data management and application cloud, processing hundreds of billions of dollars in annual channel commerce. The leading healthcare and consumer brands count on *DNA* to put real-time channel to work in their commercial businesses. Located in Princeton, New Jersey, IntegriChain has been recognized on both the Inc 500 and the Red Herring 100 lists of America's Fastest Growing Companies, and is nineteenth on the Forbes 2011 list of America's Most Promising Companies.