

# SCI StorInt™ Dispatch - IBM080102

## Announcing IBM's purchase of XIV Storage

This Silverton Consulting (SCI) Storage Intelligence (StorInt™) Dispatch provides a summary of IBM's recent purchase of XIV Storage, an independent storage subsystem provider based in Israel.

### XIV Purchase Summary

IBM has attempted for years to construct enterprise class storage subsystems out of common hardware and software. For example, the ESS was based on AIX server hardware and modified AIX software. XIV takes this approach much farther with Intel commodity servers, common networking hardware, and Linux derived software.

There are many NAS, iSCSI, and archive storage products out today based on Linux software and commodity Intel server hardware but XIV is one of the few products doing this for high performance FC attached storage. XIV also supports the iSCSI interface today and IBM will offer this as well as FC. Today, XIV has a number of large customers in the US and the Middle East.

### XIV architecture

The XIV system is based on multiple commodity Intel Xeon 64-bit servers, a Nextra O/S (Linux derivative), GigE Ethernet switches, and SATA drives with interface module and data module servers. A typical system is a rack full of servers consisting of 3 interface module servers, 8 data module servers, 3 UPSs, and commodity Ethernet switches. The data modules hold software for data management, from 4 to 64GB of cache, and up to 15 SATA drives. The interface modules have target FC and iSCSI HBAs and software to support request routing to the data modules.

The system architecture is based on mirroring data across data modules, i.e., all data is copied to two data modules. Data modules can be configured with different capacity drives and act independently to spread the data assigned to them across all drives in the module. Any single LUN's data is spread across all the data modules and drives in a system in a "pseudo-random" fashion.

The interface modules provide coarse grain addressing for the data and directs data requests to the two data modules that hold the data. Each data module handles the fine grain addressing for the data it controls.

As the data is spread throughout a data module and across two data modules for redundancy, rebuild activity for drive failures can take advantage of potentially all other drives residing in the other data modules. IBM/XIV claims that a 750 GB drive can be rebuilt in 20 minutes. IBM/XIV also claimed that a failed data module could be completely rebuilt in less than 3 hours from redundant data in remaining data modules.

Data sparing is done on a raw capacity basis. As such, one needs to set aside enough spare capacity to support rebuilding the largest data module and with 1TB drives this is 15TB. Then one must divide the remaining capacity in half for mirrored redundancy. All of this would leave ~50TB of data space in a typical XIV system with 15-1TB drive data modules and 8 data modules/rack.

Moreover, depending on performance requirements, one may wish to have more or less data modules per rack. In this case, XIV supports multi-rack subsystems with 10GigE link(s) between them. It's entirely possible to have one rack with only interface modules and multiple subsequent racks with only data module servers.

## Announcement significance

It may be that IBM has found the next storage subsystem architecture with potential to also handle non-open block storage. Accordingly, XIV can readily be altered to support NFS/CIFS and potentially mainframe protocols. XIV architecture is a significant change from IBM's current storage architectures and it could conceivably be used in other parts of their storage portfolio as well. The key questions for the storage industry are - where are XIV's weak points and how far can IBM take it.

---

---

*Silverton Consulting, Inc. is a Storage, Strategy & Systems consulting services company, based in the USA offering products and services to the data storage community*