![Monash Information Services]

Design Choices in MPP Data Warehousing

Lessons from DATAllegro V3

A Monash Information Services Bulletin

by

Curt A. Monash, Ph.D.

May, 2007

Sponsored by:

![DATAllegro — DATA AT THE SPEED OF BUSINESS]

Introduction

Index-light MPP is the best approach to high-end data warehousing.

In a recent white paper*, we laid out the case for *index-light MPP (Massively Parallel Processing) data warehouse appliances.* Embraced to varying degrees by DATAllegro, DATAllegro's hardware partners, Teradata, IBM, Netezza, HP, and Greenplum/Sun, this story hinges on three technical factors:

> *1. Shared-nothing MPP.* Loosely-coupled systems are significantly cheaper than tightly-coupled ones, for the same level of raw component performance.
> *2. Reduced use of indices.* By minimizing redundant references to information, index-light systems can store up to 7X less data than index-heavy ones. This produces enormous savings both in hardware and in administrative costs.
> *3. Avoidance of random disk reads.* Disk rotation speeds have only improved 12.5-fold in the past 50 years, making random disk lookup the greatest constraint on conventional RDBMS performance. Index-light systems largely evade this bottleneck.

And this approach works. Index-light systems beat their *index-heavy SMP* counterparts on several major criteria:

- Performance
- Price/performance
- Consistency of performance
- Administration costs

* *Index-Light MPP Data Warehousing, March 2007*

There are three major types of appliance implementations.

We've divided computing appliances (data warehouse or otherwise) into three major categories:

- Type 0 – based on custom chips or FPGAs (Field-Programmable Gate Arrays). Only Netezza among data warehouse appliance providers follows this strategy.
- Type 1 – custom hardware assembled from standard parts, without chip-level programming. Teradata follows this strategy. DATAllegro did too, in earlier versions of its product.
- Type 2 – utterly standard hardware, perhaps carefully configured. DATAllegro is now in this camp, partnering with Dell, Cisco and EMC. So are most other data warehouse software suppliers, at least to the extent they have appliance offerings at all.

In this paper, we'll use DATAllegro's new product generation to illustrate issues in data warehouse appliance design.

Data warehouse appliance design choices

MPP data warehousing stresses hardware and software designs.

Often managing tens or even hundreds of terabytes of data, MPP data warehouse appliances are among the most powerful IT systems around. Accordingly, hardware designs and the related software optimizations are stressed at multiple points.

- There needs to be lots of processor capacity. Ditto RAM. Also important are the speed and capacity of Level 2 (i.e., on-processor) cache.
- Hugely important in an MPP design is interprocessor communication speed.
- With hundreds of hard-working disks in a typical large system, disk reliability is a big issue too.

Accordingly, we find it helpful to consider data warehouse appliance hardware design choices from at least five different vantage points.

You need lots of raw power – i.e., lots of components.

First, let's acknowledge the obvious – a big and busy data warehouse requires lots of computing power. I.e., it requires lots of computer parts – processors, other chips, disk drives, and so on. But systems with lots of components scale better in MPP/shared-nothing configurations than when everything is more tightly linked together. As discussed in *Index-Light MPP Data Warehousing,* this is the reason almost every serious data warehouse software supplier, Oracle and Microsoft excepted, has chosen an MPP architecture.

They should be carefully balanced.

Beyond raw compute power, there's the question of computing power *balance.* We've researched appliances in multiple areas of software application. In most cases, vendors talk about the important tuning exercise of properly balancing the mix of parts – processors, RAM, disks, etc. In DATAllegro's products, much of the processing involves data streaming off of disks in table or partition scans. The various chips need to be sufficient to comfortably accept all this data. And even once the data is in memory, there must be enough RAM to efficiently handle various kinds of intermediate query processing.

Vendors optimize for multiple chip elements.

An increasingly important theme – again, in multiple software areas, not just data warehousing – is optimization for specific chip architectural elements. Level 2 cache, memory busses, multiple processor cores – using these to their full potential is non-trivial. For example, Intel works closely with DATAllegro to optimize DATAllegro's software, as it does with many other software vendors.

Accelerated networking is crucial.

Hugely important is the acceleration of networking. Teradata, the oldest data warehouse appliance maker, has a whole proprietary switching/networking subsystem. DATAllegro, in an unusual design choice, relies on Infiniband, currently via standalone Cisco SFS 7008 boxes. Appliance vendors in many networking- and security-related areas make similar choices, whether it's just a few extra Ethernet ports in an otherwise standard hardware design, all the way up to using special network processing and encryption chips.

Disk reliability is important and improving.

Fifth and last, there's the important matter of disk reliability. In any busy computer system, *some* disk or other will be spinning most of the time. But in a busy data warehouse system, especially one with an index-light architecture, *most* of the disks may be spinning most of the time. RAID mirroring is a standard technique for dealing with this, but even so distressingly frequent disk replacements may be needed. In V3, DATAllegro goes a step further via the EMC CX310. It uses only 12 of the 13 (the extra is a hot spare) disks in a 6+6 RAID 10 mirrored arrangement. If a disk fails, there's a quick automatic cutover to the hot spare disk. Only if there's a second failure in the same CX310 will there be performance degradation or an urgent need for disk replacement.

DATAllegro now achieves all this on standard hardware.

Traditionally, networking accelerations and computing power balance have been major reasons for adopting Type 1 appliance strategies. In its first two product generations, DATAllegro was no exception. But in its V3 strategy, DATAllegro cleverly retains most of these benefits, yet also offers the advantages of a Type 2 "virtual appliance" approach. (The biggest sacrifice we know of was some encryption acceleration.) The company reports huge Moore's Law performance enhancements, dwarfing any negative effect from moving to standard hardware. And now it has the reliability engineering, standard serviceability, and general customer comfort that come with relying on high-volume name-brand products.

About the Author

For more than a quarter-century, Curt Monash has been a leading analyst of and strategic advisor to the software industry. Praised by Lawrence J. Ellison for his "unmatched insight into technology and marketplace trends," Curt was the software/services industry's #1 ranked stock analyst while at PaineWebber, Inc., where he served as a First Vice President until 1987. Since 1990 he has owned and operated Monash Information Services, a highly acclaimed technology analysis firm focused on enterprise software. He has been extensively published and quoted in the technology and general business press, and has been a regular columnist for Application Development Trends, Software Magazine, and Computerworld. To get Curt's latest research, please see www.monash.com/feed.php .

Prior to his business career, Curt earned a Ph.D. in Mathematics (Game Theory) from Harvard University at the age of 19. He has held faculty positions in mathematics, economics and public policy at Harvard, Yale, and Suffolk Universities. For more information please see www.monash.com .

About the Sponsor

DATAllegro entered the market in 2003 with the goal of making data warehousing more affordable and more valuable to companies than any other offering. After researching the technology available at that time, DATAllegro invented a new way of distributing data across a number of servers and then running queries in parallel. Integrated with hardware, storage and a database, the end result was a data warehouse appliance that represented a true breakthrough in data warehouse price/performance. Instead of paying millions for a traditional system, companies could achieve a 10-100x improvement in query performance, at a fraction of the cost of other providers.

The company can be reached via www.datallegro.com .

Further Reading

For more research on the subjects of this white paper, please see www.dbms2.com , specifically www.dbms2.com/category/relational-database-management-systems/rolap/. Future research may be found via the free RSS and e-mail subscriptions at http://www.monash.com/feed.php .