

Acxiom Squeezes Strategic Advantage out of Sorting

PipeSort Helps Data Products Company Slash MVS Elapsed Time by 40 Percent

By John Kador

When you make a credit card purchase, chances are the resulting records are massaged by the giant computers of Acxiom Corporation.

Founded in 1969, Conway, Arkansas-based, Acxiom is one of the largest data warehouse and database marketing firms serving credit card issuers. Acxiom currently serves 17 of the nation's top 25 credit card issuers. In fact, *Fortune Magazine* recently rated Acxiom as number 19 of the 100 best places to work in the United States

The volume of records processed by Acxiom's IBM ES9000 systems is staggering. On just two machines alone, the company sorts an average of 5 Terabytes of data in 5,000 sort jobs every day, according to Systems Analyst Brad Smith. Moreover, the files can range in size from two records to over 3 billion records. In response to this high volume of sorting, Acxiom has had to evolve a sophisticated sorting strategy.

Sorting is an essential requirement of almost every element of data center operations, yet most data centers do not consider sorting a strategic activity. Big mistake, Smith says, because fine tuning the sorting operation can deliver major benefits to customers and can pay for itself many times over in terms of reduced overhead, faster processing and more predictable operations. "Central to our strategy has been our decision to use SyncSort as our sorting engine to support our marketing, decision support and data warehouse services for our customers," he says. Acxiom has relied on SyncSort Inc., (Woodcliff Lake, N.J.) for over five years.

Extract Value

Acxiom's customers use its data management services to extract value from very large volumes of data. The company is keenly interested in being able to measure how its investment in information technology supports its mission. To its credit, Acxiom practices for itself what it preaches to its customers.

Acxiom has an unusually sophisticated approach to data cen-



ter metrics, as well as ambitious programs in place to measure various aspects of data center operations performance, monitor service level objectives and calculate costs. These measurements support the continued investment the company makes in SyncSort as Acxiom expands its operations and range of services.

For example, SyncSort is used exclusively to support the company's "bread and

butter" dataset backup operations. Acxiom operates nightly dataset backups directly from tape, mostly in the form of a 350 Gigabyte RAID (Redundant Array of Independent Disks) 5 farm from EMC. Almost all the backup processing is I/O-bound, according to Smith. "RAID 5 devices are expensive. We have a strong financial incentive to reduce the time these machines wait for sorting operations to finish," he adds.

In measurement after measurement, Smith demonstrated that SyncSort provides dramatic wall-clock savings. "In our experience, sort jobs processed by SyncSort generally finish in less than half the elapsed time as the same jobs run by other utility programs and cost an average of 80 percent less," he says. Executed Channel Program (EXCP) metrics are generally accepted as a reliable indicator of a product's overall processing efficiency. EXCP savings with SyncSort generally exceed 96 percent, with respect to other utilities and even custom Assembler routines. "When we use SyncSort to backup data, the numbers allow us to predict cost savings of as much as 95 percent," Smith says.

To understand how Acxiom uses SyncSort to achieve a strategic processing advantage on behalf of its clients, let's follow how the product is used to meet the needs of a major department store. The department store captures an enormous volume of information from its point of sale and credit card applications. This data combines customer purchase activity with information from the credit card database to generate valuable information that includes items purchased, geographical data and information about the purchaser's age, income

and interests. The client looks to Acxiom to take this volume of data and extract profiles – subsets of data that reveal patterns and trends that department store managers can act on to create new value.

One of Acxiom's responsibilities is to sort all of this data into a manageable form that can serve as the input to a decision support system, executive information system or other analytical application. These applications make it easy for department store knowledge workers and executives to do "what-if" analyses, perform drill down exercises and populate spreadsheets for further analysis. Another common application that Acxiom is asked to do is create unique mailing lists from a number of disparate sources. To do so, Acxiom must perform a number of sort operations to identify duplicate names and eliminate redundancies. In the mail processing business, this operation is called "merge and purge."

Another common requirement is to then re-sort the mailing list information into postal worker street order sequence. Such a list is presorted to follow precisely, by street and house number within streets, a mail delivery person's delivery route. Not only does this get the mail delivered faster and with more accuracy, but the post office rewards the presorting operation, an operation it would otherwise have to do itself, with lower postal rates for the department store.

Specialized Sorting Applications

One of the ways Acxiom has responded to the reality of specialized sorting applications is by installing PipeSort. PipeSort is a specialized sort engine that enables SyncSort to run multiple sort operations simultaneously on common input data. PipeSort is especially effective in processing tip-to-tail applications, where the output of one sort operation is treated as the input to another, which feeds the input of still another.

"Acxiom's elaborate data manipulation projects occasionally require the sort process to read a common set of data once, and then distribute the input records to multiple, simultaneous SyncSort executions," Smith said.

PipeSort can direct a common set of input records to up to eight concurrent sort jobs.

Acxiom has successfully applied PipeSort to help create a decision support system for one of its retailer customers. Acxiom uses PipeSort when the customer needs to view several profiles

of its sales or marketing data. "We create these profiles by passing the records once to PipeSort, which then runs three or four sorts. The sorts perform summations by count or by field and create profiles by region, state and other criteria. Running PipeSort to create the information for this decision support system saves us 40 percent in MVS elapsed time," Smith says.

SyncSort has become central to many of the applications supporting the profiling systems that customers rely on to make sense of mountains of data. Acxiom systems analysts use SyncSort to set up the sort operations by making sure that records are, for example, physically close, so as to reduce processing time.

"We also use it, on occasion, to perform summations based on particular types of keys, not necessarily sort keys, but keys we have generated within the records themselves," Smith says.

Sorting out Benefits of Host and Client/Server

Like most organizations that have evolved a mainframe-centered information-processing environment, Acxiom is investigating the benefits of migrating to a client/server environment. Because such a change in processing infrastructure would obviously have a major impact on the data center's ability to process data, including sorting, Smith and his colleagues undertook an analysis of the migration plans of the 10 business groups most demanding of the data center.

The survey they sent out asked about plans to migrate from the host to client/server, and to what extent the groups expected sorting and other processing requirements to change in the next two years.

The results of the surveys indicated that nine of the 10 business groups intended to stay with the host-centric model of computing. These groups concluded that the client/server systems could not adequately scale to meet the high processing volumes. Furthermore, management processes such as backup, security and data integrity were deemed to be more mature and reliable in the host environment, relative to the client/server world.

Competing for Resources

Acxiom supports over 100 business work groups, all of whom are competing for data center resources. This pressure forces Smith and his colleagues to squeeze every ounce of performance out of their systems. "The more efficiently we can operate, the faster we can get work through the data center, the better we can support our customers," Smith notes. "It's not just that our customers can get their work processed faster, but we can give them a more defined, more reliable schedule." Acxiom knows that most customers value being able to rely on commitments more than beating deadlines. Of course, Acxiom is always raising the bar on its service level objectives.

To that end, Acxiom continues to measure its data center operations. The company's metrics emphasize CPU cycles and EXCPs, not memory cycles, and wall-clock elapsed time. According to Smith, EXCP counts have been reduced considerably, by as much as 50 to 80 percent, he says. A large part of this reduction he attributes to SyncSort's efficiency. "Such reductions would be welcome even if you're running only a few sort operations per day, but Acxiom is in the neighborhood of 5,000 sorts per day on just two machines," Smith says. "The savings more than justify the investment in the SyncSort solution."

Elapsed time continues to be the biggest constraint, more critical than the actual cost of the job, according to Smith. There are only 24 hours in a day and the batch window is constantly besieged by the encroachment of real-time operations. It sometimes becomes problematic to get all the batch work done in the time available. That's

why SyncSort's contribution to reducing the processing required to backup, process and otherwise sort large batch applications is so important. "SyncSort allows us to reduce the time required to process large batch applications by 50 percent, and reduces costs by 80 percent," Smith notes. ●

Like most organizations that have evolved a mainframe-centered information-processing environment, Acxiom is investigating the benefits of migrating to a client/server environment.

About the Author:

John Kador is a freelance writer in Geneva, Ill. He can be reached at jkador@compuserve.com.



50 Tice Blvd., Woodcliff Lake, NJ 07677
201-930-8200 www.syncsort.com