

EMC Information Sharing: Direct Access to MVS Data from UNIX and NT

Walt Kohler
EMC Corporation
171 South Street
Hopkinton, MA 01748
508 435-1000

kohler_walt@emc.com

ABSTRACT

In this extended abstract we briefly describe EMC's information sharing technology that enables UNIX and NT systems to directly access MVS mainframe datasets and how this technology can be used to directly access an MVS DB2 database.

Keywords

Information sharing, database extractor, direct access, MVS, DB2, UNIX, NT.

1. INTRODUCTION

Timely access to information is critical to most businesses, but many are making strategic and tactical decisions based on stale data. Valuable transactional information is frequently stored in MVS mainframe DB2, VSAM, and IMS databases, while data warehouses and data marts for business analysis and decision making are being built on open system UNIX and NT platforms using Oracle, Sybase, Informix, and SQL Server. Unloading data from MVS databases and moving it to a UNIX or NT system can take days and consume business critical CPU cycles and network bandwidth. Access to MVS datasets through an Oracle, Sybase, Informix or SQL Server gateway can avoid the data unload step, but it still consumes CPU cycles on the mainframe and network bandwidth to move the data to the UNIX or NT system.

What is needed is an efficient way to run critical online transaction processing applications on the MVS mainframe while providing UNIX and NT data warehouse applications with access to the mainframe data. EMC's information sharing technology does this by enabling direct read-only access to MVS datasets stored on EMC's Symmetrix Enterprise Storage from one or more UNIX and NT systems. Using this technology, an open system database such as Oracle, Sybase, Informix, or SQL Server could directly access the MVS data through a gateway or the MVS data could be extracted by a separate open system application and loaded into the open system database without consuming mainframe CPU cycles or network bandwidth.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGMOD '99 Philadelphia PA

Copyright ACM 1999 1-58113-084-8/99/05...\$5.00

2. Direct Access to MVS DB2

An example of how EMC's information sharing technology can be used to directly access an MVS DB2 database is shown in Figure 1. The middle box depicts a Symmetrix Enterprise Storage system containing a DB2 database on one or more MVS standard disk volumes. The Symmetrix also contains a separate set of MVS DB2 Business Continuance Volumes (BCV's). The standard volumes containing the DB2 data are mirrored on BCV's using EMC's TimeFinder software. In order for a UNIX or NT open system to access the DB2 datasets while the mainframe application is online, the user temporarily splits off the BCV volumes from the standard volumes before the open system application reads the data from the BCV volumes. In order to get a consistent, up-to-date copy of the DB2 database on the BCV volumes, the user must also flush data in host buffers before performing the split.

Once the BCV volumes are split, the open system host can directly access the BCV volumes using EMC's InstaShare software. InstaShare enables shared, read-only access to MVS sequential datasets on CKD formatted mainframe volumes from UNIX and NT systems. An InstaShare agent residing on the MVS platform provides MVS user authentication and dataset access verification via industry-standard SAF packages such as RACF. TCP/IP network traffic is minimal, since the network is only used to seek approval to access the datasets. The MVS data itself is read directly from the Symmetrix by the UNIX or NT open system. InstaShare provides an API based on the IBM C/370 library functions (fopen(), fread(), fseek(), fclose(), etc.) for use by layered applications.

One of the layered applications under development by EMC is a DB2 database extractor with an ODBC SQL/CLI interface. The extractor implementation is based on the seven layer model shown in Figure 2. InstaShare provides the basic I/O level. The physical level consists of a file manager and buffer manager. The page level implements DB2-specific functionality, including index page set navigation and data page set navigation, as well as some internal DB2 data conversions. The row level is responsible for navigating the records and index entries within the data and index pages respectively. The logical level interprets the data in the rows for the user tables and for the catalog information. At the SQL level the SQL parser and optimizer parses the query and decides on an execution plan. The API level is responsible for managing the interface to the client application. ODBC was chosen since it is a widely

implemented call-level interface that defines a standard SQL grammar.

EMC, Symmetrix, TimeFinder and InstaShare are registered trademarks of EMC Corporation. All other trademarks are registered trademarks of their respective owners.

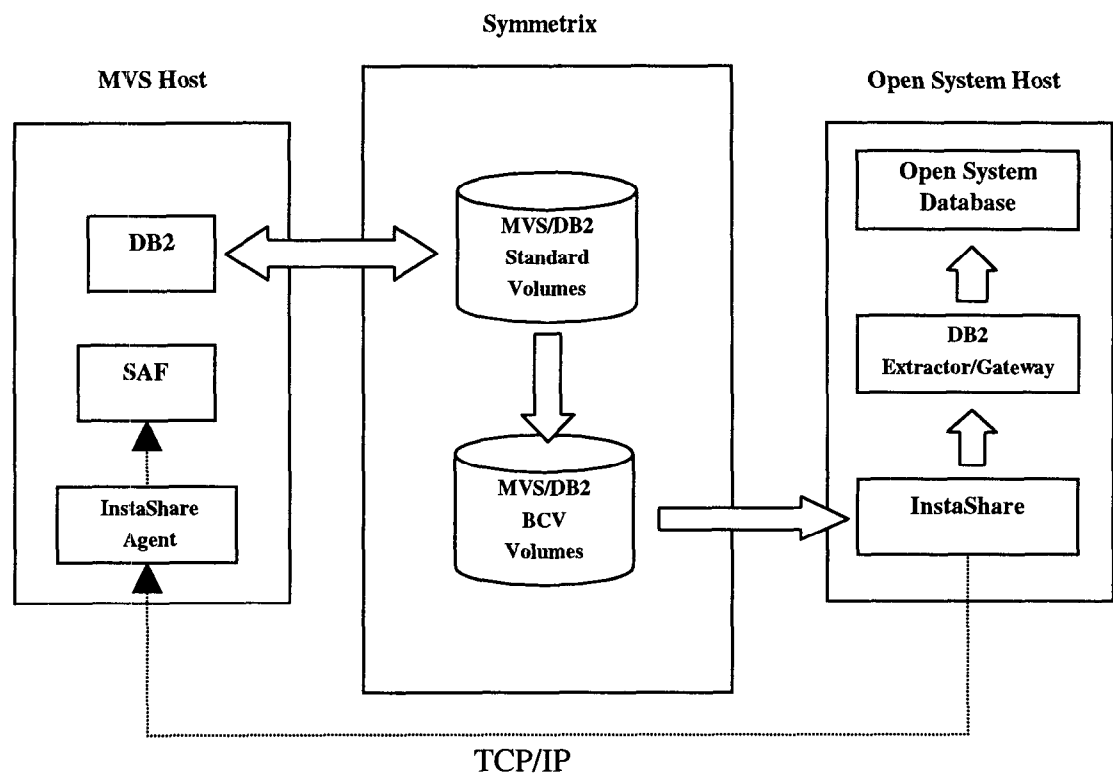


Figure 1. MVS DB2 Information Sharing

ODBC - SQL/CLI		API
Simple Parser & Optimizer		SQL
Catalog	User Tables	Logical
Data	Indexes	Row Level
Data	Indexes	Page
Files	Buffers	Physical
InstaShare		I/O

Figure 2. Seven Layer Extractor Implementation