SAND93-0587
Unlimited Release
Distribution Category UC-905
Printed June 1993

Downsizing a Database Platform for Increased Performance and Decreased Costs

Marc M. Miller and Lawrence F. Tolendino

Sandia National Laboratories
Communications Department
Albuquerque, NM 87185

January, 1993

Abstract

Technological advances in the world of microcomputers have brought forth affordable systems and powerful software that can compete with the more traditional world of minicomputers. This paper describes an effort at Sandia National Laboratories to decrease operational and maintenance costs and increase performance by moving a database system from a minicomputer to a microcomputer.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Trademarks

3COM is a trademark of 3COM Corporation.

Apollo is a registered trademark of Hewlett-Packard Company.

DEC and **VAX** are trademarks of the Digital Equipment Corporation

Ethernet is a trademark of Xerox Corporation.

INGRES is a trademark of ASK Computer Systems, Inc.

INTEL is a registered trademark of Intel Corporation.

PostScript is a trademark of Adobe Systems, Inc.

Sco and **Open Desktop** are registered trademarks of The Santa Cruz Operation, Inc.

UNIX is a registered trademark of AT&T in the U.S.A. and other countries.

Table of Contents

I. Management Summary 6
II.Introduction 6
    Background 7
    Project Genesis 7
III. Implementation 8
    Migration and Development 8
    Configuration 8
    Primary and secondary nightly data exchanges 9
    Remote printing 10
IV.Operational Results 10
    Performance Testing 10
V. Future Developments 12
VI. Conclusion 13
Appendices 14
    Appendix A - VAX Hardware and Cost 14
    Appendix B - PC Hardware and Associated Costs 15
    Appendix C - Software and Associated Costs 16

I. Management Summary

Sandia National Laboratories (Sandia) has utilized two DEC VAX systems (development and production) extensively for the past seven years to document the organization's communications circuits, both voice and data. These VAX systems have grown to keep pace with application development and the increasing number of customers. However, the systems have reached saturation and their aging hardware has become unreliable. The remarkable technological development of Intel-based microcomputers and high capacity disk drives coupled with the emergence of reliable UNIX operating systems promised a cost effective way of replacing the aging VAX systems.

The inspiration for proposing this migration occurred when an aging workstation-based file/print server was replaced with a 486/33 PC running UNIX. By chance, included with this operating system was the same database software used on the VAX systems. As a trial run, the VAX-based database and applications were ported to the 486/33 with minimal effort and tested extensively. The results were very encouraging. Response was shown to be over seven times faster running typical database inquiries. Multiple users could be supported with serial and network access.

With these results in hand, a project to migrate the databases was initiated. First, a production system and a backup were purchased. Second, the telephone database was migrated to the new production system due to the smaller size of this database compared to the much larger data communications database. After a shakedown period, the data communications database will be migrated to the new production system. Finally, after all the bugs are worked out, the VAX systems will be retired.

The telephone database has been successfully migrated to the production system and performance has met all expectations.

II.Introduction

Sandia National Laboratories (Sandia) has utilized two DEC VAX systems extensively for the past seven years to document the organization's communications circuits, both voice and data. These VAX systems have grown to keep pace with application development and the increasing number of customers. Typically 20 customer sessions will be active on the systems simultaneously performing work order, trouble ticket, or circuit documentation functions. These VAX systems have evolved to become the cornerstones for providing Sandia communications services. However, the systems have reached saturation and their aging hardware has become unreliable. The remarkable technological development of Intel based microcomputers and high capacity disk drives coupled with the emergence of reliable UNIX operating systems promise a cost effective way of replacing the aging VAX systems.

Background

The current database systems are implemented on two VAX 8350 computer systems. One is primarily a development system and the other is for production use. The database applications are the result of a continuous evolutionary process that began with a flat file database running on a CDC6600 timeshare system. The applications have been developed over a period of seven years and have evolved to provide critical documentation services. Currently there are about 14,000 data circuits and 10,000 voice circuits documented by custom Ingres applications in 100 MByte and 36 MByte databases respectively. Indeed, Sandia has come to rely heavily on these services to such an extent that the providing of voice and data services has become an essentially paperless process.

However, the applications have outgrown the VAX systems as more functions were included and the number of customers increased. Typically 20 technicians may have concurrent sessions. In addition, Sandia is rapidly assuming responsibility for its own phone service, and as it does, the amount of data to be stored will increase by an order of magnitude. Unfortunately the VAX systems have not proved capable of supporting these additional functions and system responsiveness has degraded to an unacceptable level. While the VAX systems were becoming unacceptable, Intel based PCs were growing more capable as the raw CPU power of a 33 Mhz 486 approached 17 MIPS and disk capacity exceeded 1 Gigabyte per drive. With support costs for the VAX systems exceeding $30,000 per year, it appeared very attractive to migrate the applications to a different platform.

Project Genesis

The DEC VAX 8350, which functions as the development system, was purchased January, 1988. A companion 8350 was purchased shortly after as the production system. The predicted lifetime of these systems was four years and by 1992 they were obsolete and overloaded. The VAX system hardware costs are shown in Appendix A.

A separate project, within the Sandia communications organization, had identified the utility of TCP/IP as a communications standard for the interconnection of heterogeneous computer systems (ref. document SAND91-1168). Concurrently, the communications organization upgraded the LAN file and print server from an Apollo workstation running a proprietary version of UNIX to an Intel 486/33 based system running SCO Open Desktop (ODT, a UNIX operating system based on AT&T UNIX System V). This new print and file server performed very well, providing TCP/IP connections to a heterogeneous network of workstations and PCs. In addition, the costs associated with the network server were very low. Main memory and mass storage costs were particularly attractive and the reliability of the winchester disk drives (over 10 years MTBF) was outstanding.

Coincidentally, a copy of Ingres was bundled with SCO ODT. It was only a matter of a short time before staff members were experimenting with this version of Ingres and considering whether this Intel 486/33 based PC could replace a VAX 8350. Preliminary tests were very encouraging indicating a 3 fold performance increase using simple SQL queries. Other tests indicated that over 20 simultaneous customers could access the system via Ethernet. With such promising performance results in hand a cost analysis was performed.

The initial hardware cost analysis indicated that a 486 system, capable of replacing the VAX 8350, would cost about $10,000. This system would include; 2 each 1 GIGABYTE drives for the database, a 500Mbyte system drive, a cartridge tape backup system and 64 Mbytes of main memory. This price was so attractive that it was decided to purchase a 486/50 system, migrate the telephone communication database application to this system, and perform a detailed evaluation. The breakdown of hardware and software costs for the 486/50 evaluation system is detailed in Appendixes B and C.

In order to evaluate the database software, it was also necessary to upgrade the bundled Ingres software on the 486/33 server machine to a full development system. The 486/33 was then used to port the database and applications to UNIX from the VAX 8350 running the VMS operating system.

III. Implementation

Migration and Development

The first step in evaluating the effectiveness of operating the database on a 486 based machine was to port the software. Since the Ingres database software is essentially the same whether running on VMS or UNIX, no substantial recoding of our applications would be necessary. We found that it was a fairly painless procedure to bring the database and source code from the VAX to the 486/33 via Ethernet and FTP. Some editing of the source code was required to change directory specifications from VMS to UNIX. This was accomplished by writing a script that opens each source file and replaces the existing VMS directory specifications with the UNIX equivalents. Also, Ingres script files used to reload the database into UNIX had to be modified slightly to operate as shell scripts versus VMS DCL command files and to correct directory specifications. Other than a few modifications required to some source modules to correct for a newer version of Ingres on the 486/33, the applications were compiled successfully with no changes. Performance on the development 486/33 was very encouraging. The applications ran smoothly, quickly, and flawlessly.

Configuration

Once the applications were verified to operate successfully on the development machine, the 486/50 system was readied. The system had been purchased in pieces. Not all requirements could be fulfilled by a single vendor. For example, it was necessary to

provide approximately sixteen simultaneous asynchronous connections. This necessitated the integration of a mutli-port serial card into the system, which was not available from the system vendor. In addition, a cartridge tape drive, ethernet interface, and external hard drives were added.

The system was designed to overcome hardware failures quickly, with minimal loss of data. The database backup consists of a daily checkpoint (a snapshot of the database) and journals (continuous recording of transactions designed to allow recovery of all transactions up to the last checkpoint). This configuration is similar to our production VAX, where the database is kept on three disks in three different formats: system software installed on the internal hard disk, database files installed on the first external hard disk, and database backup installed on the second external hard disk. This configuration allows for recovery from many hardware, software, or human errors. Data can be recovered from the checkpoint (with journaling bringing the restoration up to the last successful transaction). If this fails due to a corrupted checkpoint, the database can be recovered from the unload files. This is a "last ditch" effort as the process is very time consuming. In addition, a tape backup is made weekly and stored offsite, offering yet another restoration option.

As a result of the relatively low cost of PC hardware and software, it was decided to utilize a primary and a backup system (it was not possible to have a backup system using the VAX platform because of the high cost). Thus almost any failure scenario can be corrected quickly. This includes catastrophes such as CPU, motherboard, or power supply failures which completely cripple the system. For example, if the primary system power supply fails, we can disconnect the primary external hard drives and remount them on the secondary machine. After some brief file system checking and rebuilding by the operating system due to the crashing of the primary machine, the secondary would be up and running with no discernable difference between it and the primary other than the node name.

Primary and secondary nightly data exchanges

In order to keep data current on the secondary machine, data is copied nightly. A script is executed via the UNIX *at* command for copying Ingres unload files to the secondary machine. If all other preferable recovery techniques fail, the secondary machine can be used with the unload files from the previous night. Current account information is maintained on the secondary system so that in the event the secondary machine must be pressed into service, users will be able to access it with current usernames and passwords. Account information is written to a special text file (including encrypted passwords). This file is then transferred to the backup machine and used to restore account information. On the primary machine, the UNIX *cron* command is used to generate the account profile file and copy it to the secondary machine. The following entry in root's crontab file accomplishes this task:

```
0 23 * * * /tcb/bin/ap -d > /profile.acct;rcp /profile.acct sco81:/profile.acct
```

At the 0th minute of the 23rd hour (11:00pm), on any day of the week and any month of the year, the program *ap* is executed by root. This program generates the account profile file, calling it profile.acct. The file is then copied to the secondary machine using the TCP/IP copy program *rcp*. If it becomes necessary to restore account information on the secondary machine, the program *ap* is again used:

/tcb/bin/ap -r -f /profile.acct

Remote printing

The production 486 system was intended to handle database activity only, thus other functions such as file sharing, network services, and printing were to be handled by a machine dedicated to these functions. A means of directing print jobs to another machine was needed. SCO ODT provides for printing via lpsched. However, this facility does not include an inherent means of printing to remote machines. Remote printing is accomplished by creating a local "printer" and having it execute a remote shell script to the desired remote machine, passing the print file. This technique proved to be successful until attempts were made to print fairly large files which are typical for PostScript files. It was found that this remote shell process would hang and no print would take place. Also, permission had to be granted for each user on the remote print server.

A more robust and ubiquitous print service is the lpr/lpd software. This is designed for network printing. Included with the SCO software is an lpr/lpd implementation which is almost fully functional. SCO refers to this package as rlp. We found that rlp works fine when printing to a remote machine, however, we were unable to make the software print locally on our print server! Therefore, the production 486/50, while it could send out a print job, had no remote node that could accept and print the job.

The solution was to try an lpr/lpd package found on the Internet that was a fully functional implementation ported to SCO. This software has an interface to lpsched such that any jobs that are to be printed locally can be passed to lpsched automatically. The production 486/50 can now print to our local 486/33 development machine, or print to remote printers via a dedicated print server also running lpr/lpd software.

IV. Operational Results

Performance Testing

Several performance tests were run to compare the 486-based machines with our VAX 8350s. These tests were informal in nature, designed to get a "feel" for the performance differences. We wanted to see roughly how the cpu speeds, disk access, and overall performance differed, especially if the differences were close.

To roughly compare disk access, a simple C program was created which opens a file and writes 80,000 characters (80 characters per line, 1000 lines). A companion program opens this file and reads it. Performance was then compared by simply timing the execution of these programs on various platforms.

The results of executing these programs on the VAX systems and Intel system are as follows:

Computer	Write (sec)	Read (sec)
VAX8350 (development)	**10**	**4**
VAX8350 (production)	**17**	**8**
486/33	**0.1**	**0.2**

NOTE: the times for the VAX systems were determined by simply watching a clock. Due to the length of time, accuracy greater than 1 second was not deemed necessary. The time for the 486/33 was generated by utilizing the UNIX *time* command.

Comparing the program execution times shows that the 486/33 writes about 100 times faster than the VAX, and reads about 20 times faster.

Performance was also tested with an Ingres SQL database script. The script contains typical database manipulations that are regularly performed by the database application software. The databases on the VAX and the 486 machines were identical with the same number of rows in each table and the same number of tables in the database. In addition, care was exercised to ensure that all database parameters, including indices, were the same. The SQL script is:

```
select date('now');
select * from customers where name like 'NAME';
select * from phones where tn='8451234';
select * from circuits where id ='C8020';
select * from conductors where id='C8020';
update phones set tn='9999999' where id='C8020';
update phones set tn='8451234' where id='C8020';
select date('now');
```

The times for executing the script on the two systems are the result of averaging 10 runs. The results for all runs were essentially identical.

Computer System	Script Execution Time (sec)
VAX 8350 (production)	**92**
486/33	**14**
486/50	**12**

Comparing the script execution times shows that the 486 systems have more than seven times the overall performance of the VAX 8350.

An additional test was conducted to compare the times required to enter and exit an Ingres application. These times are typically what a user experiences day-to-day. For example, the following table illustrates times required to enter and exit the Ingres interactive SQL monitor:

Computer System	Enter ISQL (sec)	Exit ISQL (sec)
VAX8350	**11**	**3**
486/33	**5**	**<1**
485/50	**2**	**<1**

Obviously, there is a significant improvement in the time required to begin and end an application when running on the 486 systems.

V. Future Developments

The first attempt at utilizing a 486 platform involved our telephone circuit documentation. Having proved successful, the next step will be to migrate our data communications database and associated applications. Presently our data communications database is much larger than the telephone database. This will place a greater load on the 486 system, both in the size of the database and in the number of users. However, we believe the system will still be able to successfully support our requirements.

The PC hardware platform is evolving quickly. Within months or even weeks, newer, faster models are becoming available. The 486/33 and 486/50 systems described here have quickly been surpassed by the 66 Mhz version. The old AT bus, still used in the latest PCs, is now being challenged by the new VESA local bus, promising higher throughput within the PC. The next generation Intel chip will soon be on the market, bringing the PC to still higher performance levels. As PC performance improves we can obtain improved application performance by taking advantage of these affordable upgrades. In the minicomputer world, investments have been so high and upgrades so costly that keeping up with technology has been difficult. For example, our VAX 8350's have long been surpassed by many newer VAX models, however, due to budget constraints, we have had to remain with the 8350's as long as possible. The low cost of microcomputer hardware allows us to upgrade pieces or entire systems in reasonable time periods. We can more easily afford to replace or supplement a disk drive that is being squeezed for space, or try a new backup medium such as large-density tape drives or optical storage.

VI. Conclusion

We have successfully replaced an aging VAX minicomputer with a microcomputer that has both higher performance and lower costs. We have shown that an Intel-based PC running the UNIX operating system is able to support multiple users running database applications originally developed for a VAX minicomputer. The lower cost of hardware and software allows us to more easily keep up with the fast-paced changes in computer technology.

Appendices

Appendix A - VAX Hardware and Cost

DEC VAX 8350 Production System
2-CPU 8350
32 Mbytes system RAM
400 Mbyte system hard drive
400 Mbyte user hard drive
400 Mbyte database hard drive
200 Mbyte backup hard drive
9-track magnetic tape drive

Total cost: $140,000

Appendix B - PC Hardware and Associated Costs

Gateway 2000 50 Mhz 486 DX/2 PC	$4915
tower case	
1.2 Mbyte 5.25" floppy	
1.44 Mbyte 3.5" floppy	
16 Mbyte system RAM	
500 Mbyte SCSI hard drive	
Adaptec 1740 SCSI Host Adapter	
ATI Ultra Video card	
1 parallel/2 serial ports	
Gateway 101 key ANYKEY keyboard	
Gateway 14" Crystal Scan Super VGA monitor	
System Memory Upgrade	
48 Mbytes addition system RAM	$1920
Ethernet Card	
3-COM EtherLink II, model 3C-503	$200
SCSI Internal Tape Drive	
Archive Viper 2150s, 1/4" cartridge drive (QIC-150)	$465
Serial Expansion Card	
Stargate Plus-8 multi-port serial card, model P8A005	$420
External SCSI Hard Drives	
Micropolis MC1598 1.2Gbyte	$1792 ea.
External Case for Hard Drives	
JMR four-bay, model MIDT-002	$300
Total cost:	$11804

Appendix C - Software and Associated Costs

Software Maintenance Costs:

486/50 Production System

Software maintenance: $1861.50/year

This maintenance includes telephone support as well as dial-up support and information.

System and Application Software Costs:

486/33

SCO Open Desktop	$1100 (base operating system)
SCO Server Upgrade	$1121 (multi-user)
Ingres	$0 (included in the base SCO distribution)
Ingres upgrade to 6.4	$250
total	$2471

Copy to:

1900 D. L. Crawford
1951 P. W. Dean
1951 Y-S Yu
1954 M. O. Vahle
1955 D. M. Darsey
1955-1 J. P. Sena
1955 M. M. Miller (5)
1955 S. D. Nelson
1955 L. F. Tolendino
7141 Technical Library (5)
7151 Technical Publications
7613-2 Document Processing for DOE/OSTI (10)
8523-2 Central Technical Files

END

DATE FILMED

9/23/93