RDBMS on Linux for Accelerator Control

A.Yamashita, T.Fukui, M.Kodera, T.Masuda and R.Tanaka
SPring-8, Mikazuki, Hyogo, 678-5198 Japan

Abstract

Recent developments on Unix on PC platform begin to impact not only academic computing society but also commercial software developers. Major relational database management system (RDBMS) vendors announced to port their products to Linux and some have already shipped with low price or free.

These products make possible to migrate database servers that are running on expensive Unix platforms to inexpensive PC platforms running reliable operating system.

We ported a relational database for SPring-8 storage ring control system to a PC server running Linux and investigated possibility of RDBMS on an open source operating system focusing on performance and reliability.

1 INTRODUCTION

Unix-like operating systems running on personal computer platforms have grown rapidly in recent years. Especially Linux is becoming a major PC operating system in server field with its cost, performance and reliability. Software industries cannot ignore this trend. Most of major RDBMS vendors released their main products for Linux.

In the meantime, PC hardware is available with much lower cost of commercial Unix servers. Performance of RDBMS on PC platform may have reached to the level of commercial Unix server with far better price and performance ratio.

In this article we compare a RDBMS on a PC platform to an existing server which manages SPring-8 storage ring (SR) control database.

2 RDBMS ON SPRING-8

A RDBMS running on Hewlett-Packard (HP) server manages the data of SPring-8 storage ring and beamline [1]. It stores parameters on 226 tables, 26 GB of log data and 9014 points of on-line data. And it is still growing with new equipment on SR, new beamlines and new accelerators. New SUBARU, which began its commissioning last fall (1998), depends on the database for its parameter data management and data logging. The data of the injector synchrotron will begin to be managed by the database system since January 1999. Although RDBMS vendor guarantees its scalability, it will require more resources of computer like CPU, memory and disks as growing number of data points and clients.

Development of Linux, dropping cost of personal computer hardware with high performance and release of commercial RDBMS may provide an alternative solution to expensive Unix platforms.

3 RDBMS TO TEST

We chose Sybase Adaptive Server Enterprise (ASE) release 11.03 for a candidate to compare an existing commercial Unix server. Sybase ASE on a HP K250 has been serving SPring-8 data successfully and it satisfied performance, reliability and easy of maintenance required for an accelerator control. If Sybase ASE on Linux is comparable to one running on HP server, it will reduce future cost of hardware according to expansion of number of data points and clients.

3.1 Installation

We obtained Sybase ASE 11.03 package for Linux from a web server[1]. Rpm utility[2] expand the package to /opt area. Installation of ASE itself is similar to one of HP-UX version except two points [2].

1.Kernel parameter for the maximum size of shared memory must be changed to match the server requirements. We expanded its value from 32MB (default) to 128MB.

2.Linux does not support *raw* (character) device for disk access. It means RDBMS must access disk controller through operating system file system. Thus, ASE uses standard Unix file or *cooked* (block) disk partition for its disk device.

The second can cause a trouble i.e. RDBMS *thinks* the data has been written on the disk, but actually the data exist in operating system's buffer instead of the disk. If the server stops with trouble like power failure, it may yield inconsistency between RDBMS and actual data on the disk. On the other hand, disk buffer makes an advantage for faster disk access.

---

[1] http://www.redhat.com/
[2] Provided by Red Hat Software. Inc.

4 TEST PLATFORMS

Among the database serves for accelerator control, entering and retrieving accelerator log gives heaviest load to the RDBMS. We compared RDBMS performance on Linux on PC and HP server with loading data read/write test. Platform for this test is summerlized as follows.

4.1 Linux platform

Pentium II 450MHz single CPU,
256MB memory,
128MB configured for ASE,
SCSI disk. block device partition used,
Linux Kernel 2.0.36[3],
Red Hat 5.1 distribution with glibc,
ASE for Linux 11.03,
100BaseT Ethernet.

4.2 HP server

K260 series HP PA8000 180MHz 4CPU,
3CPU configured for ASE,
2GB memory,
1.4GB configured for ASE,
SCSI disk character device used,
HP-UX 10.20,
ASE for HP-UX 11.5.1,
10BaseT Ethernet.

We also measured the performance of K260 configured with 128MB memory and one CPU for ASE. Even this lower configuration, the HP server has an advantage to the Linux server because the operation system can use remained resources.

Other parameters for database configuration were remained as defaults.

4.3 Network

These two servers are connected to the same dual speed 100Mbit/s capable Ethernet hub with loading machine shown as Fig.1. No other user processes except server processes were running during the test.

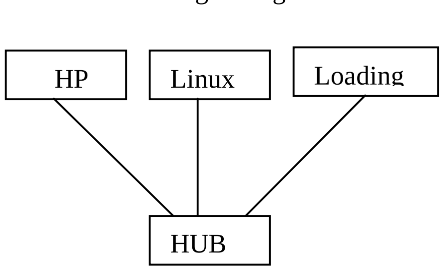

Fig.1 Network configuration.

[3] Kernel before 2.0.36 has network problem to use ASE.

4.3 Data loading machine

A dedicated Linux client machine writes and reads data to/from both severs. Test programs written in Python [3] with ct-library extension [4] issue queries via remote procedure calls (RPC) to servers. RDBMS on the server returns results for queries using stored procedures. Another Python script reads data from the servers. One writing process was running for one table. Thus 60(number of tables)*2(read/write) process were running in data loading machine. These processes gave light load to the loading machine, so that there were no interrupts to the test due to the overload of client machine.

5 TEST

5.1 Online database

SPring-8 on-line database consists of 60 tables. Each table contains 2000 to 4000 rows in it. One row is updated in one data taking cycle. For normal SPring-8 operation, one data taking process loads up to 250 data to a table in every one to 30 seconds. Tables have 9152 columns in total. And data acquisition processes enter data into tables in rate of 6.4 Kbytes/sec. The on-line database occupies total of 160MB disk space including index and log segment. The size is larger than data cache of ASE on the Linux server.

One on-line database consists of columns of sequential number, time and data as shown in Fig. 2.

Seq_no	Data 1	,,,,	Data n	time
1	…	,,,,	…	…
:	:	:	:	:
2000 to 4000 rows				
2000	…	…	…	…

Fig. 2 Structure of an on-line database table

Each table has two indexes of time columns and sequential number columns. Writing process updates time and data columns using sequential number index.

5.2 Measurement

Two kinds of measurement were performed to measure Linux and HP servers.

One is checking CPU load on the Linux server writing/reading the same amount of dummy data with same rate of real SR run. Processes running on the data-loading machine issue RPC to read or write data in the time interval according to real SR data taking. Total of 120 users connected to the server machine.

We measured the maximum performance for writing as the second test. Total of 60 writing processes running on the data-loading machine sends RPC to read or write data as soon as the server returns the answer for the previous RPC. The average speed and maximum time was measured for 120 minutes. The Linux and HP-UX machine performed this test separately.

6 RESULTS

6.1 CPU load at normal operation

We observed 1 minutes CPU load average with /proc/loadavg for the Linux server. Fig. 3 shows 120 minutes of load of the CPU. One unit of load means one process on the queue to use the CPU or I/O. The server process consumes less than 30% of CPU power running on Linux machine. The test load continues for 14 days for Linux server with no deterioration of performance and any trouble.

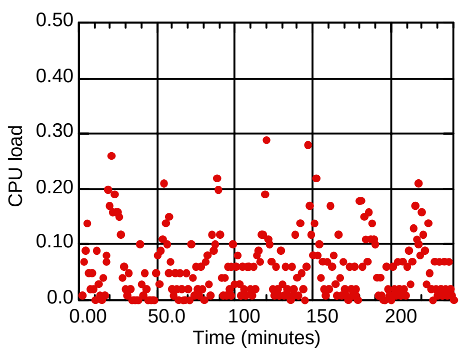

Fig.3 CPU load of Linux server at normal operation

A system stored procedure, 'sp_sysmon', on server monitored 3 minutes statistics for I/O rate of disk/network during the run. The statistics shows access to database log segment consumed two third of disk I/O and access to table segment used one third of disk I/O.

6.2 Maximum I/O rate

Linux ASE server wrote 186Kbytes of table per seconds. Server took 0.20 seconds to write a row of 200 columns in average and 5.6 seconds in longest. Results are summarised in Table 1.

	Linux	HP	HP (*)
Average insertion time (s) to 200 column table	0.2	0.2	1.21
Max delay time (s)	5.6	0.47	101
Total rate (Kb/s)	186	179	44.3

(*) 1CPU and 128MB configuration

Table 1. Summery of performance.

Fig. 4 shows the relation between the number of columns per table versus the number of inserted rows per second for Linux and HP servers.

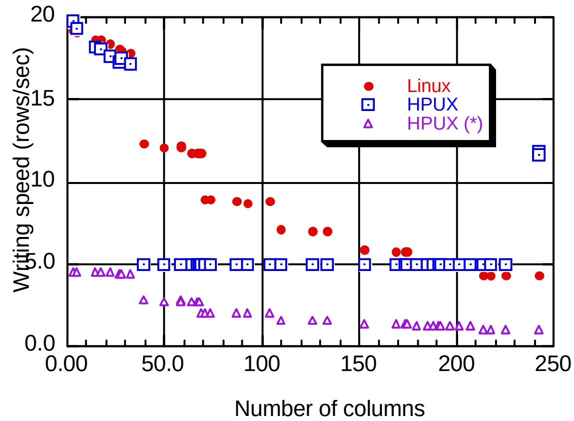

Fig.4 Writing speed vs. the number of columns

The one CPU and 128MB Linux server has comparable performance to 4CPU and 2GB HP machine and much faster than the HP with same number of CPU and memory.

As a result, Linux and HP have nearly equal speed in writing on-line database.

7 CONCLUSION

ASE on Linux shows good performance compared to HP machine in accelerator log reading/writing test. It handles over 120 users logged simultaneously with no significant delay. ASE itself has good reliability as far as continuous run for two weeks.

Still, we have not tested following items on ASE running on Linux server.

1.Its scalability at SMP (symmetric multi processors) configuration.

2.Large memory capability. Can we install memory of 2Gbytes?

3.Large database handling in the order of 100G bytes.

Though large database operation requires the test above, RDBMS on Linux PC proved a good performance on writing/reading on-line data. RDBMS on PC platform is promising for a database server to use at the size of SPring-8 SR.

REFERENCES

[1] A.Yamashita et al., Proc. of ICALEPCS ’97, Beijing, China, 1997, P.427.
[2] http://www.mbay.net/~mpeppler/linux.html
[3] http://www.python.org/
[4] http://starship.skyport.net/~pgodman/