APPLICATION NOTE

AN-118

RDS Intranet Database Growth & Performance

Overview

This note describes predicted growth and performance of the RDS Intranet database. It confirms that servers specified by Reedholm are capable of handling RDS Intranet application needs well into the future.

Each server is delivered with six RAID slots configured with three active, and one hot-spare, 36GB drives. This default configuration provides roughly 72GB of striped and redundant storage that can be doubled by plugging in two more 36GB disks.

Database storage calculations are based on test cases because size is highly dependent on data types and storage schemes. The RDS Enterprise Server was designed for a fab outputting 20,000 wafers per month while gathering parametric data at the following rate:

- Test every wafer, with 25 wafers per lot
- Probe 5 die per wafer, with 10 intradie per die
- Execute 20 tests per intradie, or 200 per die

With that model, the server can easily store ten years worth of data with dozens of concurrent users.

Selection of Microsoft SQL Server

Microsoft was chosen because it was reasonably priced for small applications and because it was integrated with development tools used in creating the RDS Intranet product.

According to pundits, Microsoft has made great progress in improving database performance and is delivering a product that can handle large volumes of data without significant performance degradation. Recent evaluations and comparisons have shown SQL Server 2000 to outperform both Oracle and IBM SQL database offerings.

Performance of the RDS Intranet database depends heavily on the hardware platform. To maximize performance, the standard RDS Enterprise server has a RAID Level 5 disk subsystem, 1GB of memory, and dual processors.

Performance Criteria

Maximizing database performance requires a lot of RAM, a fast storage system, and powerful processors.

1) RAM size needs to store enough of the database in memory (stored procedures, indexes, data) so that disk requests are minimized during queries.
2) Exchanging data between RAM and disk is dependant on speed of disk access, disk controller design, backplane bandwidth, etc. Thus, efficiency of the entire storage system is important, not just disk speed and disk access time.
3) Dual processors take advantage of multi-tasking and multi-threading capabilities of the OS, Web Server, and SQL Server software. A simplistic assessment is that one processor can run the RDS Intranet application while the other handles database requests. Of course, what is really going on with the instruction handling is much more complicated.

Performance Hits

Performance hits are events when the system (application, database, and server) has a steep reduction in performance, usually caused by there not being enough RAM to handle increased index sizes.

Performance thresholds are not easily predicted, and gathering information to perform predictions is frustrating since database vendors are not inclined to reveal information that might divulge proprietary protocols. Reedholm took the approach of over specifying hardware with a plan to ugrade when performance plateaus were reached. So far, that has not happened.

While not a hardware issue, large transaction files cause performance degradation. That is because every query has to go through pertinent transaction data. A nightly backup (these shrink or eliminate transaction files) is usually enough to keep database size in check. However, higher volume operations need to have backup run more often.

Database Size Background

Data storage in an SQL database is not easily predicted. While many parameters are involved, most of the disk space is used for table storage, so that is the focus of these calculations (based on "Estimating the Size of a Table" from the SQL Server Help files).

Record Size

Record size is determined from a few calculations.

1) Sum of each column data type with a fixed size (bit types are grouped in eights).
2) [# varchar columns x 2] + [Sum of varchar lengths] + 2. Set to 0 if no varchar's exist.
3) 4 bytes for a row header
4) Null bitmap data: Int{[(#columns + 7)/8] + 2}

Record sizes greater than 8060 bytes are not possible except for special data types (text, graphics, etc.), none of which the RDS Intranet database uses. Table 1 illustrates calculating record size for one database table with static, bit, and variable field lengths.

Column Name	Data Type	Bytes
AcquireScreen_ID	Int	4
ModifyTime	Smalldatetime	4
Modify_User_ID	Int	4
Locked	Char	1
Locked_User_ID	Int	4
Sum of static columns		**17**
Advanced	Bit	1/8th
Sum of bit data in groups of eight		**1**
AcquireScreen_Name	nvarchar(30)	60
ScreenDesc	nvarchar(50)	100
# Variable columns x 2 + Length + 2		**166**
Four bytes per row header		**4**
Null bitmap data (8 + 7) / 8 + 2		**3**
TOTAL RECORD SIZE		**191**

Table 1 – Example Record Size Calculation

Records Per Page & Total Pages

SQL Server stores data in 8192 byte pages with a 96 byte header. The number of records stored per page is the integer value of 8096/[record size + 2]. For the example in table 1, 41 records can be stored per page.

Page content can only contain data from one object or table. The number of pages used per object/table is found by executing the SQL command "DBCC CHECKALLOC".

RDS Intranet database tables do not use clustered indexes, so these were not calculated.

Determining Database Space Usage

Because of the relationship of pages and record sizes, disk space usage is computed with large record quantities, e.g. 1000s of die. Any database size calculations made with small data sets will not correlate with the disk space used. Correlation can be seen once the database is >100MB.

Device Space

Using values from table 2, 1,000 devices (50 pages x 8192) would require 407kB of disk storage.

Database Table Name	Bytes per Record	Records per Page	Pages per 1K
Device_Library	313	25	40
Device_Names	75	105	10

Table 2 – Device Database Tables

Test Plan Space

Worst case is that all tests are equation tests or have attached equations, and that all variable length columns are maxed out. This results in 6,317 bytes per test as shown in table 3.

Database Table Name	Bytes per Record	Records per Page
Test_Names	337	23
Test_Facts	598	13
Test_Switches	36	213
Test_Rules	35	218
InstrValues_Facts	2213	3
CalcValues_Facts	274	29
User_Equations	2621	3
Process_Facts	203	39
Total Bytes per Test	**6317**	**-**

Table 3 – Test Tables

The next step is to calculate the number of pages needed for each record type for a specific number of tests. That is done in table 4 for 100 and 10k tests.

Database Table Name	Pages per 100 Tests	Pages per 10k Tests
Test_Names	5	435
Test_Facts	8	770
Test_Switches	1	47
Test_Rules	1	46
InstrValues_Facts	34	3334
CalcValues_Facts	4	345
User_Equations	34	3334
Process_Facts	3	257
Total	**90**	**8568**

Table 4 – Pages Required for 100 and 10k Tests

Multiplying the number of pages in table 4 by 8,192 bytes/page results in these totals:

- 100 tests = 738kB
- 10k tests = 70.2MB

Test List Space

Because test records are only referenced by test list tables, test list size is a separate calculation. The first two database tables in table 5 have one record per test list, while the last one has one record for every test in the test list. That leads to these calculations:

- One test list with 100 tests = 24,576 bytes
- 10k test lists with 100 tests each = 58.4MB
- 10k test lists with 1k tests each = 553MB

Database Table Name	Bytes per Record	Records per Page
IntraDieList_Names	79	99
IntraDieList_Details	245	32
IntraDieList_Facts	52	149

Table 5 – Test List Tables

Die Pattern Space

The first two die pattern tables in table 6 have one record per pattern while the last has one record per die:

- One die pattern with 1k die = 115kB
- One die pattern with 10k die = 934kB
- 1k die patterns with 10k die = 915MB
- 1k die patterns with 100 die = 13.9MB

Database Table Name	Bytes per Record	Records per Page
DiePattern_Names	75	105
DiePattern_Details	397	20
DiePattern_Facts	87	90

Table 6 – Die Pattern Tables

Lot Data Space

Lot data has two parts, header data and test results. Table 7 totals the number of pages required for 9600 lots, or 240,000 wafers. Each page requires 8,192 bytes, so one year of lot headers requires only 168MB.

Database Table Name	Pages for 9600 Lots
Tested_Lots	370
Lot_Details	3200
LotPatterns	17
LotPatternFacts	237
LotTests	15239
LotDie	179
LotWafers	1305
Total	**20547**

Table 7 – Pages Used for Lot Headers

The first three tables in table 8 contain one record per lot. The other tables contain one record:

1) Per unique intradie.
2) Per test (and in order they are executed).
3) For each tested die position.
4) Per wafer tested.
5) For each die tested per wafer.
6) For each non-sweep data point.
7) For each wafer, die, intradie, and test order sequence tested. This data is lot independent and plateaus once every combination is tested.

Database Table Name	Bytes per Record	Records per Page
Tested_Lots	302	26
Lot_Details	2510	3
LotPatterns	12	578
LotPatternFacts (1)	18	404
LotTests (2)	62	126
LotDie (3)	28	269
LotWafers (4)	42	184
LotBinData (5)	22	337
Test_Data_Facts (6)	28	269
DataPointIDs (7)	20	368

Table 8 – Lot Test Tables

The majority of the space is used by the table Test_Data_Facts, which resolves to 240 million points or 892,194 pages per year for 9,600 lots. At 8,192 bytes per page, this equates to 7.3GB per year. Obviously, header data per lot is dwarfed by the test results.

Note that the quantity of data points and space consumed are the same for 10,000 die with one test per die as for one die with 10,000 tests.

Conclusion

The standard RDS Enterprise server has 72GB of space, enough to store ~10 years of data. That volume can be expanded to 144GB by simply adding two more 36GB disks. Those can be added at any time, thereby providing roughly 20 years of space.

Calculations show that a large fab putting out 20,000 wafers per month and gathering parametric test data from every wafer would not run out of disk space for almost twenty years. Starting with larger disk capacity is always an option, but that just brings costs forward and adds capacity that may not be needed for years down the road, if ever.

Higher test volume and/or more users than described might require:

- More RAM and different software versions
- Different backup system
- Expanded disk capacity

Maximum RAM

The maximum memory SQL Server Standard can use is 2GB, which requires 4GB installed in the server since the OS will take the first 2GB. To utilize more than 2GB of RAM requires upgrading to SQL Server Enterprise edition. Windows must also be upgraded to either Enterprise or Datacenter editions.

Tape Backup

The current tape backup is limited to 36GB per tape and would be rather slow for data in excess of that, plus require swapping tapes. In fact, Reedholm recommends backing up the database to either external hard drives or over the network so that IS handles it.

The tape unit is provided for smaller facilities where backing up the database will be the responsibility of the process engineers.

Larger Disk System

As an example of needing more disk space, an increase in sampling from 5 to 25 die per wafer results in an increase to 1.2 billion data points per year, or enough to use 72GB's of capacity in two years.

Determine Requirement's of Each Server

Prior to each RDS Intranet order, an estimation of total operational needs is used to determine whether or not a larger server is required. That analysis is used to predict when an upgrade might be needed 2, 3, or 5 years later, if ever.