Progress Exchange

2013 – Boston, MA

Progress Very Large Databases (VLDB)

Presented by: Dan Foreman, Progress Expert

BRAVEPOINT

- **Progress User since 1984**
- **Since V2.1 (There was no commercial V1)**
- **Author of some Progress Books:**
  - *Progress Performance Tuning Guide*
  - *Progress Database Admin Guide*
  - *Progress System Tables Guide*
  - *promon – debghb* **Soon to be released!**
- **And Database Admin Tools:**
  - **ProMonitor/ProCheck/LockMon**
  - **Pro Dump&Load**

BRAVEPOINT

Audience Survey – Production DB Version

- **V8 or earlier**
- **V9**
- **V10.0**
- **V10.1**
- **V10.2**
- **V11.0**
- **V11.1**
- **V11.2**
- **V11.3**

BRAVEPOINT

Audience Survey – Database Size

- **Single Largest Database Size**
  - **> 10 TB**
  - **> 5 TB**
  - **> 1 TB**
  - **> 500 GB**
- **Progress VLDB is a very exclusive "club"**
  - ☺

- **Minimum of 500gb**
  - **250gb minimum in 2011**
  - **100gb minimum in 2007**
- **Single Database (not a set or group)**
- **Progress only (no Oracle allowed)**

VLDB Criteria

- **Some sites had qualifying DBs but split decided to split their VLDB into multiple, smaller DBs for performance reasons**

- **Example:**

  - **Reads per Second beyond a certain point would not improve for a single DB regardless of –spin or –B values**

  - **This problem was prior to the extensive latch changes made in V10.1C and –lru*skips parameters in V10.2B**

BRAVEPOINT

Site	Description
Broder	T-Shirts, etc.
Alt	Mortgage Servicing
DTW	Mortgage Lending
FT	Financial Services
XXX	Payroll Services
CM	Claims Management
BMRB	Big Modern Retail Bank in Russia
FI	Business Support System DB somewhere in Europe

BRAVEPOINT

Database Sizes

Company	Progress	Size
Broder 2007	**10.1A0205**	**290GB**
Broder 2009	**10.1C02**	**554GB**
Broder 2011	**10.1C02**	**760GB**
Broder 2013	**10.2B07**	**893GB**
XXX 2010	**10.1C03**	**4.7TB**
XXX 2011	**10.1C03**	**5.5TB**
Alt 2012	**10.2B06**	**688GB**
Alt 2013	**10.2B06**	**940GB**

BRAVEPOINT

Database Sizes

Company	Progress	Size
BMRB 2009	10.1C0442	1.5TB
BMRB 2011	10.2B0401	2.8TB
BMRB 2013	???	Mgmt Unwilling
DTW	10.2B06	714GB (1.3TB)
FI	10.2B06	2.3TB

- **I am not personally aware of any VLDBs on Progress V11**

- **I used to have some V9 VLDB customers but no longer**

BRAVEPOINT

Largest Table(s)

Site	Records	Size	Same Table?
BMRB 2009	7.04 Billion	691GB	Yes
BMRB 2011	13.7 Billion	1.3 TB	Yes
XXX 2011	719 Million	1.2 TB	Yes
FI	8.6 Billion	1.9 TB	Yes

BRAVEPOINT

Largest Table in XXX

| Table | Records | Size | -Record Size (B)- | | | ---Fragments--- | | Scatter |
			Min	Max	Mean	Count	Factor	Factor
PUB.iegrecord	718932941	1.2T	205	3283	1760	890066584	1.0	2.2

Another Large Table (BMRB) 2009 & 2011

PUB.DataLine	7040294464	690.6G	38	16494	105	-1546825661	1.0	1.0
PUB.DataLine	13687002345	1.3T	38	21498	102	806235942	1.0	1.0

Another Large Table (FI)

| Table | Records | Size | -Record Size (B)- | | | ---Fragments--- | Scatter | |
			Min	Max	Mean	Count	Factor	Factor
PUB.D	8590845958	1.9T	176	319	243	8590846399	9426.3	1.0

BRAVEPOINT

- **Sun (BMRB, Alt, FI) – For some reason I see more Sun Servers outside of the USA**

- **IBM (XXX, Broder, NFCU, CM, DTW)**

- **HP/UX > Windows (BP)**
  - **BP <> BravePoint or British Petroleum but it is a UK Company**

- **Linux – I personally have not seen a VLDB yet but I have heard of a few**

- **DTW has some of their DBs entirely on SSD (Solid State Disks); one of their other databases is on Hierarchical Storage**

- **SANs can be "dangerous" because they are frequently used as shared storage for multiple servers; at one VLDB customer we found that a very high I/O Data Warehouse was on the same physical disks as the production databases; not a good configuration**

- **Avoid the ZFS File System**

- **No VLDBs can be found on NetApp**

BRAVEPOINT

DB Buffer Cache (-B)

Site	-B	DB Blk	Memory
DTW	3,800,000	8K	30gb
Broder	500,000	8K	4gb
Alt	5,373,952 917,504 (-B2)	8K	43gb 7.3gb

Spin (-spin)

Site	-spin	Remarks
DTW	30000	
Alt	100000	
Broder	5000	30000 > 10000 > 5000

- **FI: probkup online, Backed up 372036190   db blocks in 58:27:32 (2.97GB DB)**

- **Broder: probkup online made on DB replicated with AI files (DR server). 3 hours.**

- **DTW uses an online probkup <u>and</u> a separate SAN Flash Copy**

Database Replication

- **None (but they are using After Imaging)**

- **AI Log Based Replication**

- **OE Replication**

- **None of my VLDB customers are currently using SAN Replication**

BRAVEPOINT

- **Longest time without a DB Shutdown: Broder: "*We did not have any downtime for about 2 years*"**

- **CM**
  - **Every 3 months; 12 hours maximum**
- **Broder**
  - **15 minutes every night**
- **BMRB**
  - **Twice a month during weekend night; Length of the window is not more than 6 hours**
- **Alt**
  - **Every Sunday**
- **DTW**
  - **Every scheduled shutdown in involuntary**

Concurrent DB Connections (using the .lic file)

- **DTW:** **6,007**
- **Alt:** **3,894**
- **Broder:** **1,389**

- **OE Management**

- **ProMonitor**

- **ProTop**

- **promon**

- **Homegrown**

BRAVEPOINT

Dump & Load

- **Broder: Pro Dump & Load**
- **Wachovia: 2007 - "Are you kidding?"**
- **Wachovia: 2008 - Pro Dump & Load**
- **BMRB: "Never happened"**
- **CM: 10 years ago; now do selective table level D&L & idxcompact**
- **Alt Pro Dump & Load**
- **QL: proutil tablemove and/or BUFFER-COPY**

Database Administrators

- **Broder:**
  - **I am 0.25 Progress DBA and 0.75 everything else**
  - **0.25 Oracle DBA + MS SQL Server DBA**
  - **0.50 Unix admin and problem resolution mediator**
- **Alt – 4.5 full time DBAs**
- **QL – 4 full time DBAs**
- **CM – 1 full time DBA**

- **"Online schema changes or better to say the absence of it. That is by far the biggest one."**

- **"Progress does not seem to make a lot of improvements for VLDBs. All VLDB should be 24 by 7 and thus everything should be done online. For example, Progress made recent improvements to idxbuild in 10.2B. But that is an off-line only utility. I would rather see improvements for idxactivate."**

- **Table Partitioning**

- **Better Non-Uniform Memory Architecture (NUMA) Support**

- **Routine DB Maintenance shouldn't require refreshing the OE Replication Target**

BRAVEPOINT

- **Broder:**
  - **I would say 9GB in a month growth.**
  - **The total size of all after-image files generated in a day is 20GB.**
  - **We did not have unplanned downtime for 5-6 years. We did not have any downtime for about 2 years. Mostly because scripts that alert of .bi HWM and Lock Table growth ahead of critical levels. And I guess our hardware is good too. We did switch to DR site two years ago due to SAN hardware problem.**

Until this customer I had never seen ONE TRILLION DB Requests before

```
   DB Requests: 1,290,367,893,766    DB Reads: 7,556,608,277    Hit Ratio: 171:1
   Checkpoints: 6,480                 Flushed: 56,494            Cluster: 32768K
     DB Writes: 75,013,159         APW Writes: 74,128,217           APW%: 99%
     BI Writes: 22,686,946         BIW Writes: 8,314,061            BIW%: 37%
     AI Writes: 27,017,750         AIW Writes: 26,541,304           AIW%: 98%
  BI Busy Buff: 795,259         BI Empty Buff: 0              BI Partial: 9,270,86
  AI Busy Buff: 29,390          AI Empty Buff: 0              AI Partial: 400,093
•
       Commits: 894,922,810                     Undos: 104,245
   Rec Updates: 218,136,939         Record Creates: 81,876,673
   Record Reads: 401,729,738,665    Record Deletes: 51,467,284
   Record Locks: 2,328,657,181        Record Waits: 158,488
   Locks in Use: 5            Lock High Water Mark: 15,141
   Block Extends: 24,640                      BI Size: 3,637.25mb
   Empty Buffers: 0
      Sem Waits: 46,163,131            Latch Timeouts: 1,840,963,530
```

BRAVEPOINT

- **Progress can handle VLDBs especially with V10.1C and later**

- **A high quality Server & Storage Device are essential components**

- **I have never personally seen a Terabyte sized Progress DB on Windows; that doesn't mean they don't exist…maybe just hiding**

- **If you are anticipating a VLDB, it might be good to avoid niche Operating Systems**

BRAVEPOINT

Questions?

- **Vragen (Dutch)**

- **Shitsumon (Japanese)**

- **КЕТЦАЛЬ (Russian)**

  - $\alpha =+ A\square=3;!\_$

Contact Info:

- **Dan Foreman (dforeman@bravepoint.com)**