

NFS G-BASE の性能評価

3Q-9

安尾 典子, 平岡 昭夫, 飯沢 篤志
(株) リコー ソフトウェア研究所

1 はじめに

G-BASEは(株)リコーにおいて開発されたデータベース管理システムである。[1] DBMSの分散化の必要性から、G-BASEに対してもその作業が進められている。その最初のステップとして、G-BASEはNFS(Network File System)[2]により共有されたファイルシステム上でデータベースを構築し、それを共有する複数のマシンからアクセスできるよう改良された。[3,4]

本稿ではそのNFSに対応したG-BASEの性能評価を、オリジナルなG-BASEとの比較及びLM (Lock Manager), DD (Data Dictionary), データベースのサイト依存性という観点から行なう。NFS対応による変更はDM (Data Manager)とLMに限られている。

マシンはローカルサイト(そのマシン上にDMが存在する)としてVAX Station 2000、NFS serverとしてVAX 8650を使用した。

2 NFS G-BASEの主な改良点

NFS対応とすることによる最も大きな変更は、LMプロセスとDMプロセスの間の通信手段である。(LMは同時実行制御を司るプロセスで、DMはデータベースを直接操作するプロセスである。)従来の同一ホストに限定された手法の代わりにRPC (Remote Procedure Call)を用いるように変更された。[5] この変更によってNFSをサポートしている任意の機種種のUNIXマシン間でのデータベースの共有が可能になった。またDDもネットワーク環境に対応するように改良された。

3 評価の方法及びその結果

3.1 データ

データはWisconsin-Madison大学のBitton等がDBMSの一般的なベンチマーク用に提唱しているものの一部を用いた。[6] 1レコード182バイトで、そのうちわけは2バイトの整数13個と52バイトの文字列3組である。13個の整数は以下のような数列である。

I0	I1	I2	I3	I4	...	I11	I12
999	999	1	9	19	...	499	999
927	0	1	4	19	...	298	34
210	1	0	2	0	...	3	662
000	2	0	1	2	...	190	180
852	3	1	8	15	...	76	11
...	...	...	...	...	...	...	...
123	999	1	0	1	...	381	856

Table 1: 整数データ

Performance of NFS G-BASE
Noriko YASUO, Akio HIRAOKA, Atsushi IIZAWA
Software Research Center
Ricoh Co., Ltd.

表1の2行めの数字は数列の上限を示している。I1, I2はユニークな数列であり、I1はランダムにI2は順に並んでいる。

3組の文字列(s0, s1, s2)は52個の文字からなり、1, 27, 52文字目に任意の大文字のアルファベットを置き、それ以外をxで埋めつくした形をしている。s0, s1はユニークな文字列の並び、s2はそのうちの4種類のみがランダムに並んだものである。

データは1000件で測定した。

3.2 NFS G-BASEとオリジナルG-BASEの比較

NFS G-BASEとオリジナルG-BASEの性能の比較のために、同一のDD及びデータベースについて、同じ操作を行なった。

操作は:

- ・1フィールドへの射影
 - ・(1~3)個の条件を与えた検索
 - ・あるフィールドのレコード全体に及ぶ更新
- 以上はインデックス付きインデックスなしの両ケースに対して行なう。
- ・レコードの挿入、削除
 - ・スキーマの再構成
- (データベースの移動、インデックス、フィールドの付加)である。

通常の使い方では、要する時間の比はオリジナルG-BASEのものを1として、検索で2.2、挿入で2.1、削除で1.1、更新で2.1という結果であった。この場合はレコードのオカレンスごとにロックしながら実行される。一方明示的にデータベース全体に排他的ロックをかけ同様の測定を行なった結果は、検索で1.1、挿入で1.8、削除で1.1、更新で1.0でありロックにかかる時間が特に増大していることが分かる。

スキーマ再構成について同様の比は、データベースの物理的な移動で1.3、1フィールドに対するインデックスの付加で1.3、1フィールドの付加で1.3という結果であった。

3.3 LM, DD, データベースのサイト依存性について

LM, DD, データベースのサイト依存性は、それぞれがローカルサイトとリモートサイトにある場合の $2^3 = 8$ 通りについて前節と同じ操作行ない比較することで評価した。このうちのLM, DD, データベースが三者ともローカルサイトにある場合が前節のNFS G-BASEのデータに相当する。

結果は表2のとおりで、値は秒単位である。

4 考察

NFS G-BASE とオリジナル G-BASE の相違について考えられることは3点ある。1点めは NFS G-BASE が RPC 及び XDR を使用しているということである。測定に使用したマシンが両方とも VAX であることから byte order の変換が多少はオーバーヘッドになっているかもしれない。2点めは下位プロトコル自体の速度である。(NFS G-BASE の使用している) inet domain の UDP/IP と、(オリジナル G-BASE の使用している) UNIX domain のスピードを比較すると、後者のほうがかなり速いことが知られている。3点めは DM と LM の通信に使用しているハンドシェイクプロトコルの所要時間である。データグラムによる通信であることから、信頼性を増すためにはかなり厳密なプロトコルが必要である。しかしながらスピードアップの可能性のある部分であるから、この所要時間を必要最小限に抑えるためには十分に検討しなくてはならない。

今回の測定に用いたマシンのローカル対リモートの性能比は 1:6 程度である。DD, LM, データベースのサイト依存性について顕著なのは、まずその性能による影響である。例えば、検索、挿入、削除で最も速いのは、DD も LM もリモートにある場合であった。データベースがローカルにあるかリモートにあるかはこれらの場合余り重要ではないようである。すなわちこれらの操作についてはサイト依存性はほとんどないといえる。ただし更新に関しては、DD, LM がリモートサイトにある場合が他に比べて速いのは同様であるが、データベースがリモートサイトにあるとローカルサイトにある場合の 3.8 倍程度の時間がかかっている。

以上の結果から、DM, LM を他のマシンに置いて NFS を用いてそれ等を使用することに問題はなく、むしろ小さなワークステーションでは DM, LM, データベースが 1 スピンドルのディスクにアクセスするほうが問題であるといえる。それに対して挿入や更新が多いデータベースはローカルサイトに置くべきで、参照のみであればローカルでもリモートでも同程度の応答速度が得られることが分かる。すなわち、各ユーザは参照するだけの共有データは能力の高いサーバに、個人データはローカルサイトに置くという方針で分散環境を構築すれば実用的なシステムを形成することができると期待される。

5 今後の課題

今後の課題として、

- 1.) データベースをローカルサイト(q)とリモートサイト(r)に置いて、 $q \times q, r \times r, q \times r$ のジョインのデータを取る。
 - 2.) G-BASE には単純検索の他にターゲットという概念があり一度ターゲットファイルにデータを集めてそれに対して更に操作を行なうという機能があるが、そのターゲットに対する分散化の影響を見積もる。
 - 3.) 同じデータベースにアクセスする DM の数を増やしてゆき競合の影響を調べる。
 - 4.) ファイルアクセス、ロックに関する通信をモデル化し、シミュレーションを行なう。
- 等が計画されている。

尚、この調査は分散化する際のシステムの最適化に用いられる予定である。

References

- [1] (株) リコー編, G-BASE システムマニュアル
- [2] R.Samberg, D.Goldberg, S.Kleiman, D.Walsh, B.Lyon, "Design and Implementation of the Sun Network Filesystem", USENIX Summer, 1985, pp563-568
- [3] 飯沢篤志, 平岡昭夫, "データベース管理システム G-BASE の NFS 対応(1)", 情報処理学会第 36 回 (昭和 63 年前期) 全国大会
- [4] 平岡昭夫, 飯沢篤志, "データベース管理システム G-BASE の NFS 対応(2)", 情報処理学会第 36 回 (昭和 63 年前期) 全国大会
- [5] Sun Microsystems, Inc., "Remote Procedure Call Protocol Specification",
- [6] D.Bitton, D.J.DeWitt, C.Turbyfill "Benchmarking Database Systems A Systematic Approach", Proceedings of the VLDB Conference, 1983, pp8-19

DD	LM	DB	射影	検索	挿入	削除	更新
L	L	L	41.58	17.15	13.23	3.31	699.38
L	L	R	39.53	17.23	17.00	3.49	2253.54
L	R	L	29.60	12.07	9.46	2.61	491.15
L	R	R	27.16	11.68	10.77	2.44	1859.11
R	L	L	42.31	17.46	13.33	2.93	698.75
R	L	R	39.31	17.10	13.90	2.82	2292.22
R	R	L	27.25	11.58	9.02	2.36	488.23
R	R	R	27.11	11.90	10.63	2.43	1850.46

Table 2: NFS G-BASE における DD, LM, DB の位置依存性 (L: local, R: remote)