

# 分散大容量メモリシステム DLM の初期性能評価

緑川 博子<sup>†</sup> 小山 浩生<sup>†</sup> 黒川 原佳<sup>‡1</sup> 姫野 龍太郎<sup>‡2</sup>

<sup>†</sup>成蹊大学 工学研究科, <sup>‡</sup>理化学研究所, <sup>‡1</sup>情報基盤センター, <sup>‡2</sup>次世代計算科学研究開発プログラム

64bitOS で広がった大きなアドレス空間を生かし、ローカル物理メモリサイズに制限されずに、大規模データを扱うプログラムの実行を可能にする分散大容量メモリシステム DLM とそのコンパイラを構築した。DLM はクラスタの各ノードの遠隔メモリを集めて仮想的に大容量メモリとする。以下の初期評価では、計算ノードの物理メモリよりも大きいデータ領域を用いるテストプログラムで、swap ファイルを使用する通常の C プログラムと、DLM メモリサーバを使用するプログラムとで性能比を調べた。2つのテストプログラムでは、1GbEther クラスタで、遠隔/搭載物理メモリのサイズ比 2 程度で、通常プログラムの 5 倍、10GbEther クラスタでは、遠隔/搭載物理メモリサイズ比 0.15 倍程度で、10 倍以上の性能が得られた。また従来、swap ファイルサイズの制限を受け、実行不可能であったプログラムを実行することができた。DLM は、特定のハードウェアや、システムライブラリや OS などに手を入れることを前提とせず、一般ユーザがユーザレベルソフトウェア（移植性のある DLM ライブラリとコンパイラ）のみで、容易に分散大容量メモリを手に行けることを示した。

## DLM 設定ファイル例 hostfile

```
calhost 2048 //2GB
memhost1 8192 //8GB
memhost2 4096 //4GB
memhost3 1024 //1GB
:
```

## プログラム実行コマンド例

```
prog -- -n 4 -f hostfile
```

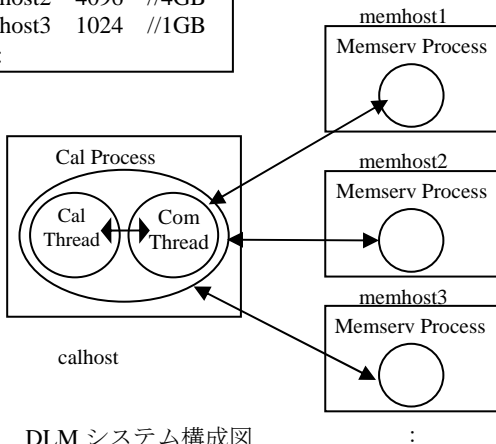


図 1 DLM システム構成図

```
// 行列ベクトル積 matv.c DLM プログラム
#include <stdio.h>
#define N 16384 // total memory 2048MB + 32KB
dml double a[N][N], x[N], y[N]; // DLM 使用宣言
int main(int argc, char *argv[])
{ int i,j;
  double temp;
  for(i = 0; i < N; i++) // 行列 a を初期化
    for(j = 0; j < N; j++) a[i][j] = i;
  for(i = 0; i < N; i++) x[i] = i; // x を初期化
  for(i = 0; i < N; i++){ // a[N][N]*x[N]=y[N] 計算
    temp = 0;
    for(j = 0; j < N; j++) temp += a[i][j]*x[j];
    y[i] = temp;
  }
  return 0;
}
```

通常プログラムに **dml** を加えるだけ

**dml** コンパイラでコンパイルすると利用可能

図 2 DLM 使用プログラム例 (matv.c) dlmc-0.0.3

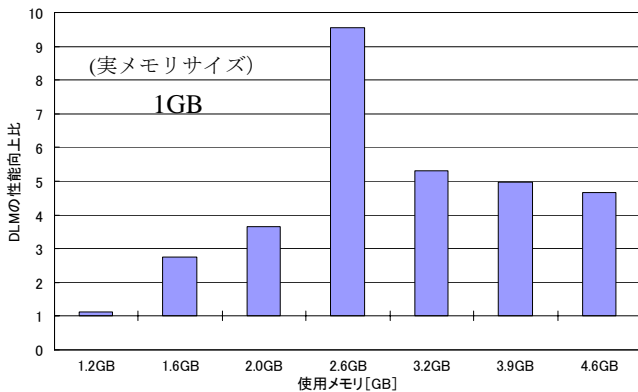


図 3 matv の速度向上比 (1GbE クラスタ)

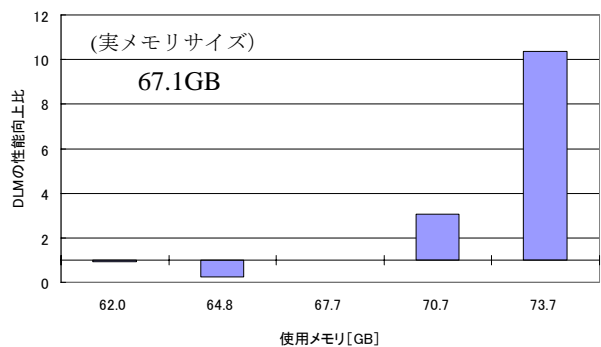


図 4 matv の速度向上比 (10GbE クラスタ)

表 1 性能測定環境

	10GbEther クラスタ (CSLM)*	1GbEther クラスタ (hp-cluster)
Cluster	HP DL585 G2 x 5 Nodes	HP ML150G2 x 8 Nodes
Node CPU	Opteron 2.8GHz x 4 (8Cores)	Xeon 2.8GHz x 2CPU HyperThread
NodeMemory	64GByte (67.1GByte)	1GByte (L2cache 1MB/CPU)
Network ,NIC	10GbE protocol, Myri-10G	1GbE, Broadcom 5721 PCI-Express Giga NIC
Switch	Fujitsu XG1200(10GbE Switch)	CentreCom GS924GT(1GbE Switch)
Hard Disk	SAS 147GB 10krpm RAID1	ST3808110AS 80GB 7200rpm

## 参考文献:

緑川, 小山, 黒川, 姫野, “分散大容量メモリシステム DLM の設計と DLM コンパイラの構築,” 信学会研究報告, CPSY, Dec.2007.

\* CSLM: 理化学研究所 次世代計算科学研究開発プログラム所有のクラスタ