

並列I/O SPFSの性能評価

3G-3

村上 岳生、藤崎 直哉、新開 慶武
(株)富士通研究所

{takeo,naoya,shinkai}@flab.fujitsu.co.jp

1 はじめに

SPFS(Scalable Parallel File System) [1, 2] はスケラブルで高性能な並列I/O機構である。本稿では当社の並列コンピュータ AP3000 上で行なった、並列IO SPFSの性能評価について述べる。

2 実験システムの構成

性能測定に使用した実験システム AP3000 の構成図を図 1 に示す。AP3000 は各ノードを高速な専用ネットワーク (AP-Net) で接続した並列計算機システムである。各ノードでは Solaris OS が動作する。各ノードにはローカルディスクが接続されている。

ノードはアプリケーションプログラムが計算を行なう計算ノードとアプリケーションからのファイル I/O 要求を処理する I/O ノードに分けて使用した。

全ノードで SPFS デーモンが動作する。SPFS ライブラリをリンクしたアプリケーションプログラムを計算ノード上で実行させ、その延長で I/O ノード上に SPFS サーバが起動される。SPFS サーバは Solaris のローカルファイルシステム (ufs) を介して、ローカルディスク上のサブファイルにアクセスする。

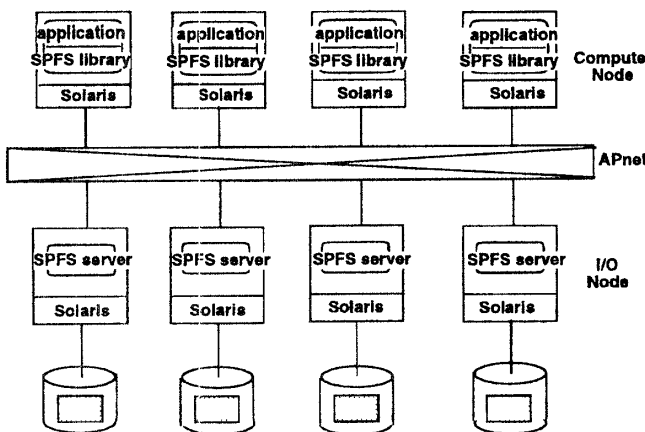


図 1: 実験システム構成

3 性能測定の狙い

今回は SPFS の持つ様々な機能のうち性能向上に関連する以下の3点に着目した。

I/O ノードストライピング 逐次および並列アプリケーションがデータを連続シークエンシャルアクセスする際に、複数 I/O ノードにアクセスを分散させ、処理速度の向上を図る。

ストライドアクセス 計算ノードから発行される非連続なリスト I/O を一括して処理する事により処理速度の向上を図る。

コレクティブ I/O 並列アプリケーションの各プロセスが配列データを分担してアクセスする際に、各プロセスからのアクセス要求を一括して処理する事により処理速度の向上を図る。

これらの効果を検証するために、逐次アプリケーションが連続シークエンシャルアクセスを行うケースと並列アプリケーションが配列データを離散シークエンシャルアクセスするケースについて性能測定を行なった。

4 連続シークエンシャルアクセスの性能

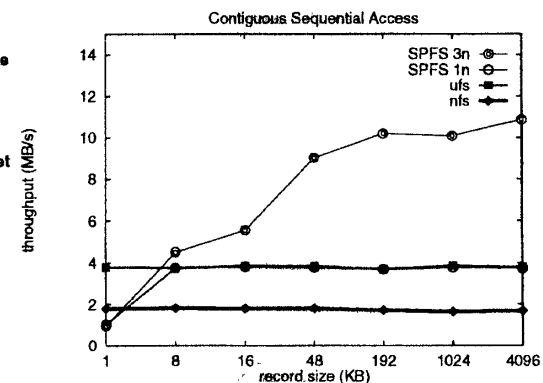


図 2: 連続シークエンシャルアクセスの性能

図 2 に大容量のファイルを連続シークエンシャルアクセスした際の性能を示す。比較のため NFS, ローカルファイルシステム (ufs) の性能も示した。SPFS 1n はファイルのストライピングを行わずにクライアント 1、サーバ 1 の構成で測定した性能であり、ローカルファイルシステム (ufs) に匹敵する性能を達成している。また同構成の NFS に比べほぼ倍の性能であ

る。SPFS 3nはファイルを3つのノードにストライプ幅16KBでストライピングしてクライアント1、サーバ3の構成で測定した性能であり、レコードサイズが十分大きいケースではノード数に比例して性能が向上している。レコードサイズが1Kの場合にufs,NFSに比べSPFSの性能が落ちるのは、SPFSがクライアント上でキャッシングを行っていないので1Kのデータアクセス毎にサーバとの通信が入るためである。

5 離散シークエンシャルアクセスの性能

c8K

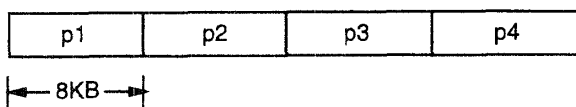


図 3: アクセスパターン

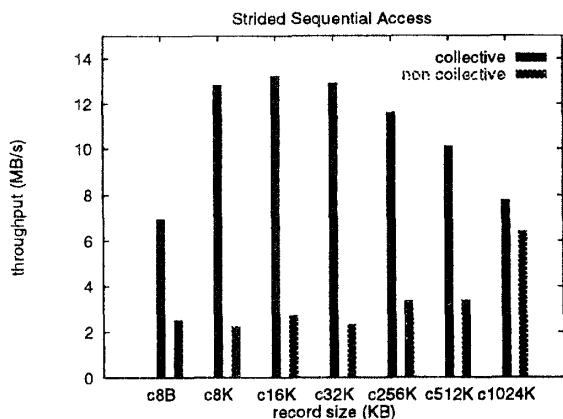


図 4: 離散シークエンシャルアクセスの性能

4並列のアプリケーションが配列データをアクセスするケースを想定して8種類の離散シークエンシャルアクセスパターンについて性能測定を行った。4プロセスそれぞれに1つの配列データの担当区画を分配し、各プロセスが自分の担当区画のデータをファイルに書き込む。それぞれのアクセスパターンは4台の計算ノード上の各プロセスが配列データのファイルを8byteから1024KBの単位でcyclicに分割してアクセスするものである。図3はcyclic 8K(c8K)のアクセスパターンにおけるプロセスp1~p4のファイルへのマッピングである。このパターンをファイルの先頭から32KB単位に繰り返す。1回のアクセス単位は1MBである。

データファイルは4台のI/Oノードに16KBのストライプ幅でストライピングされている。ファイルサ

イズは640MBである。

図4にSPFSのストライドアクセスインタフェースを使った性能測定結果を示す。collective(コレクティブI/O)は各プロセスの要求を一括する様に指定したケース、non collective(非コレクティブI/O)は指定していないケースである。

非コレクティブI/Oに比べコレクティブI/Oの性能が優れていることが分かる。コレクティブI/Oでは各計算ノードからのアクセス要求をI/Oノード上のI/Oバッファに集めて一括してディスクに書き込む事が出来る、これに対し、非コレクティブI/Oでは各計算ノードからのアクセス要求がI/Oノード上のマスタサーバでシリアルライズされ、I/Oバッファを使つてのディスクへのI/O量が4倍になってしまう。この差が測定結果に表れていると考えられる。

c8B(cyclic 8byte)のコレクティブI/Oの性能は他に比べて約半分である。これはアクセスパターンのパターン長が短く、I/Oノード上のSPFSサーバにおけるアクセスパターンの処理コストが大きくなる事が原因である。c256K以降のコレクティブI/Oの性能が低下して行くのは各クライアントプロセスの担当するデータがI/Oバッファサイズ(126K)以上に連続して続いたため一時点で1つの計算ノードにデータ転送処理が集中してしまう事が原因である。

6 おわりに

並列 IO SPFSの性能評価について述べた。SPFSの機能であるI/Oノードストライピング,ストライドアクセス,コレクティブI/Oの効果を確認する事が出来た。

今後、連続シークエンシャルアクセスのレコードサイズが小さいケースでの性能を改善するために部分的にクライアントキャッシングを導入する事を検討している。実アプリケーションをSPFSに適用しての評価も今後の課題である。

参考文献

- [1] 新開 慶武, 村上 岳生, 藤崎 直哉, " 並列 I/O SPFS の概要", 情報処理学会第54回全国大会論文集, 1997.
- [2] 藤崎 直哉, 新開 慶武, 村上 岳生, "並列 I/O SPFS の実装", 情報処理学会第54回全国大会論文集, 1997.