

Android タブレットクラスタを用いた並列処理に関する一考察

濱中真太郎^{†1} 中村優太^{†2} 野村駿^{†2} 坂本寛和^{†2} 山口実靖^{†1}

1. はじめに

タブレット PC が普及し、安価でこれらの端末を入手できる環境が整いつつある。タブレット PC にはフラッシュメモリが搭載されており、フラッシュメモリはハードディスクと比較して、ランダムアクセス性能が高い。よって、フラッシュメモリが搭載されたタブレット PC を複数台用いてクラスタリングを行うことで、極めて高いスループットの I/O 処理環境を構築することが可能であると考えられる[1]。

本稿ではまず、単一 Android のタブレットの I/O 性能の評価を行う。次に、Android タブレットのクラスタの構築方法を示し、性能評価により Android タブレットクラスタが非常に高いランダム I/O スループットを実現できることを示す。最後に、Android タブレットクラスタにおける並列ランダムアクセス処理のボトルネックに関する考察を行う。

2. 単一端末の I/O 性能

本章で単一の Android タブレット端末のランダム I/O 性能の評価を行う。

2.1 測定方法

タブレット端末のストレージ上に用意した 10GB のファイルに対してランダムアクセス要求(システムコール read())を連続して発行するプログラムを実行してその処理時間を測定することにより Android タブレットのランダム I/O 性能を評価した。測定は Nexus7 2013(Android 4.4.2, CPU Qualcomm Snapdragon S4 Pro 1.5 GHz, RAM 2GB, Storage SK Hynix H26M51003EQR 16 GB eMMC NAND Flash)で行った。

2.2 測定結果

測定結果を図 1 に示す。図の横軸は並列実行スレッド数であり、縦軸は得られた I/O スループットで単位は [Transaction/Sec] である。各線は、システムコール read() のブロックサイズを表している。図より、スレッド数を一定数(4 スレッド)以上増加させても総 I/O 性能の増加は無いことが分かる。また、ファイルシステムのブロックサイズが 4KB であるためシステムコールブロックサイズを 4KB 未満にしてもトランザクションスループットの向上が無いことが分かる。本測定より、タブレットに搭載されているフラッシュメモリのランダムアクセス性能は 2100[transaction/sec]程であることがわかった。

3. Android タブレットクラスタ

本章にて、Android タブレットクラスタの実装手法、そして同クラスタのランダム I/O 性能評価について述べる。

3.1 実装手法

本クラスタは 1 台のマスター端末と、n 台のスレイブ端末で構成される。図 2 に実装の詳細を示す。本稿の構築例では、n=1~8[台]、各スレイブ端末が 10[GB]のストレージ領域を持ち、クラスタ全体で最大 80GB のストレージ空間が存在する。無線 LAN ルータは NEC PA-WR9500N-

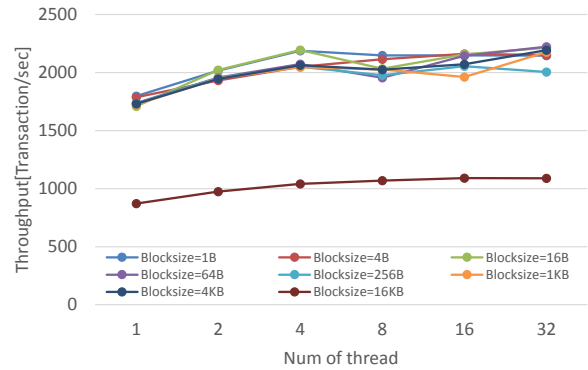


図 1 I/O 性能評価結果

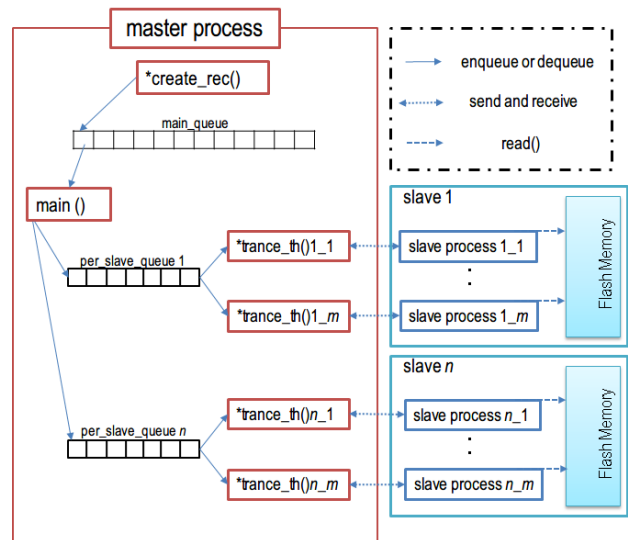


図 2 タブレットクラスタのシステム構成

HP(IEEE802.11n/a/g/b)を利用した。マスター端末は要求の発行、要求の振り分け、要求の転送、応答の受信を行う。スレイブ端末はマスター端末から受け取った要求アドレスに対して指定サイズの I/O 読み込みを行いマスター端末へ応答を送り返す。

■ マスタプロセス

マスタプロセスはマスター端末上で動作し、1 本の main_queue と n 本の per_slave_queue、1 個の main スレッドと n*m 個の転送スレッドを有している。すなわち、per_slavequeue1 本あたり m 個の転送用スレッドが存在する。main_queue にはクライアントプログラムが発行した仮想ストレージへの I/O 要求が格納される。通常、I/O 要求発行は別プロセスにて行われるが、実験用の本実装ではマスタプロセス内で発行を行った。per_slave_queue には、スレイブ端末ごとの I/O 要求が格納され、main_queue から per_slave_queue への移動は main スレッドにより行われる。

転送用スレッドは per_slave_queue から I/O 要求を

Android Tablet Cluster for High Throughput I/O Processing
^{†1} 工学院大学 情報通信工学科
 Information and Communication Technology, Kogakuin University
^{†2} 工学院大学大学院 電気電子工学専攻
 Electrical Engineering and Electronics, Kogakuin University Graduate School

取り出し、それをスレイブ端末のスレイブプロセスに転送する。転送用スレッドは後述する各スレイブ端末のスレイブプロセスと TCP コネクションを確立する。転送は TCP コネクションを用いて行われる。main_queue に対して発行される I/O 要求には仮想ストレージ全体におけるアドレスが格納されている。例えば、要求アドレスが 20GB 以上 30GB 未満のアドレスである場合は、当該 I/O 要求は 3 本目の per_slave_queue に移動され、スレイブ端末 3 に転送される。

■ スレイブプロセス

スレイブプロセスはマスタ端末のマスタプロセスから I/O 要求を受け取り、送られてきた要求アドレスに対して I/O システムコールを行う。そして、応答をマスタプロセスに送信する。スレイブプロセスは、マスタ端末の転送用スレッドと同数だけ起動され、転送用スレッド 1 個とスレイブプロセス 1 個で TCP コネクションが 1 個確立される。

3.2 測定結果

read() システムコールのブロックサイズを 1 バイト、スレイブ端末 1 台ごとの転送用スレッドの数を 1~32 個として並列実行したときの I/O 性能測定結果を図 3 に示す。図の横軸はスレイブ端末の台数で、縦軸は得られた I/O スループットである。各線は、スレイブ端末 1 台あたりの転送用スレッドの数である。スレイブ端末の台数に着目すると、スレイブ端末数を増やすに伴い I/O スループットが向上していることが分かる。スレイブ端末 1 台 32 スレッド実行時のスループットに着目すると、前章で得たフラッシュストレージの限界性能 (2100[transaction/sec]) に近い性能が達成できていることが分かる。しかし、スレイブ端末が 2 台以上のスループットに着目すると、スレイブ端末 1 台あたりの I/O 性能は上記性能より低いことが分かる。このことより、マスタプロセスからスレイブプロセスに送られる I/O 要求の量が十分でなく、フラッシュメモリの性能を飽和させることができていると予想できる。また、スレイブ端末ごとのスレッド数を増加させても十分な性能が得られないことから、タブレット間の Wi-Fi 通信におけるパケット数が処理のボトルネックとなっていることも原因の 1 つとして考えることができる。そこで、1 回 (1 パケット) あたりの転送 I/O 要求集約数を 10 個に変更して I/O スループットを測定した。測定結果を図 4 に示す。スレイブ端末数と合計スループットに着目すると、スレイブ端末数に比例に近い合計 I/O スループットが実現されていることが分かる。得られた I/O スループットは最大で約 16,000[Transaction/sec] であり、非常に高いスループットが実現できていることがわかる。並列実行スレッド数を一定数 (8 個) 以上増やしてもスループットの向上は見られないことから並列実行スレッド数 8 個以上でタブレットに搭載されているフラッシュメモリ性能を飽和するに十分な要求数が転送されているといえる。このことから I/O 要求および応答を集約して転送することにより Wi-Fi 通信を用いるタブレットクラスタにおいて非常に高いランダム I/O スループットを実現できることが分かる。Wi-Fi は CSMA 方式の一つである CSMA/CA 方式を用いて動作しており、多端末環境で平均フレームサイズが小さい通信を行うと高い性能が得られない [2]。よって集約により性能向上が実現されたと考えられる。

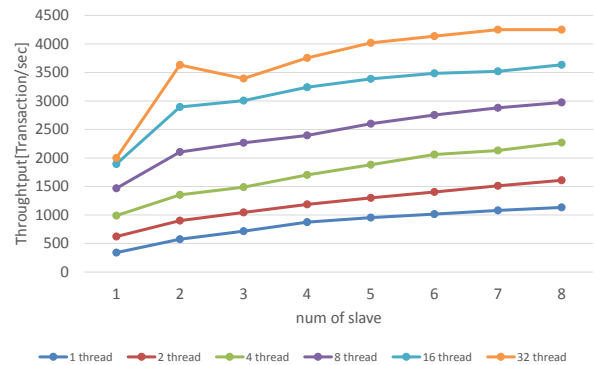


図 3 I/O 性能評価結果 (タブレットクラスタ)

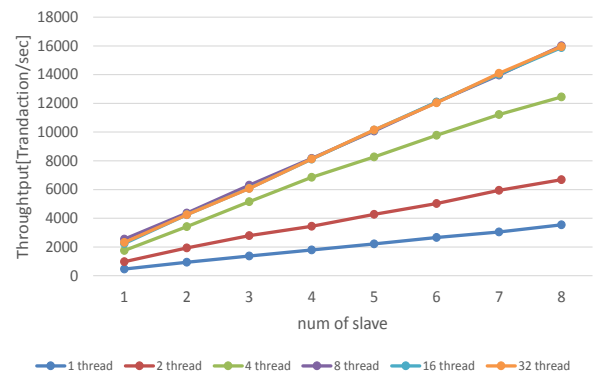


図 4 I/O 性能評価結果 (タブレットクラスタの要求集約数 10)

4. おわりに

本稿では、高いランダムアクセス性能を持つフラッシュメモリを搭載しており安価である Android タブレットに着目し、Android タブレットをクラスタリングして高スループットランダムアクセス処理環境を構築することについての考察を行った。単一 Android タブレットと Wi-Fi 接続の Android クラスタのランダムアクセス処理性能の評価を行ったところ、I/O 要求の集約などにより Android タブレットクラスタにより高いスループットのランダムアクセス処理システムの構築が可能であることがわかった。

今後は、write 処理に関する考察、CPU を用いる処理の性能に関する考察を行っていく予定である。

謝辞

本研究は JSPS 科研費 (24300034, 25280022, 26730040) の助成を受けたものである

参考文献

- 1) 早水悠登, 合田和生, 喜連川優, "フラッシュストレージ環境におけるアウトオブオーダー型データベースエンジン OoODE の実験的クエリ処理性能評価", 電子情報通信学会第 6 回データ工学と情報マネジメントに関するフォーラム (DEIM2014), A2-3, 2014
- 2) Bianchi, G., "Performance analysis of the IEEE 802.11 distributed coordination function," Selected Areas in Communications, IEEE Journal on, vol.18, no.3, pp.535,547, March 2000