

自動マウントを利用した外部ストレージの電源制御

根本 利弘[†] 喜連川 優[‡]

[†] 東京大学地球観測データ統融合連携研究機構 〒153-8505 東京都目黒区駒場 4-6-1

[‡] 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

E-mail: {nemoto,kitsure}@tkl.iis.u-tokyo.ac.jp

あらまし 自動マウントを用いた外部ストレージシステムの電源制御について述べる。ネットワークを通じて電源の制御が可能なコンセントを利用し、このコンセントにディスクアレイ装置を接続してホストからの電源のオン・オフ制御を可能とするとともに、自動マウントを利用し、ホストからのアクセスリクエストに応じて自動的にディスクアレイ装置の電源制御を行う機能の実装を行った。自動マウントにより電源が制御されているディスクアレイ装置の消費電力を実際に測定するとともに、シミュレーションにより自動マウントを用いた際のディスクアレイ装置の消費電力削減効果の評価を行った。

キーワード 自動マウント, 電源制御, グリーン・ストレージ

Power Management of External Storage Systems using Automount

Toshihiro NEMOTO[†] and Masaru KITSUREGAWA[‡]

[†] EDITORIA, The University of Tokyo 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan

[‡] Institute of Industrial Science, The University of Tokyo 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan

E-mail: {nemoto,kitsure}@tkl.iis.u-tokyo.ac.jp

Abstract Power management of external storage systems using automount is described. Using remote controllable power strip, we implemented the mechanisms that power of disk array system is turned on or off automatically according to access request from host server through automount. We measured the power consumption of power managed disk array system and evaluated the effectiveness of power management mechanisms with automount by simulations.

Keyword Automount, Power Management, Green Storage

1. はじめに

近年、地球温暖化が極めて重要な問題となり、温室効果ガスの削減が重要課題となっている。IT機器の消費電力量が急増しており、さらに、その中においてもストレージの占める割合は増加傾向にあり[1]、ストレージ機器の省電力化は急務である。このような背景の下、アクセスがないときにディスクドライブの回転を下げたり、停止させたりすることにより消費電力を低下させる MAID (Massive Arrays of Inactive Disks) 機能を有するディスクアレイ装置が開発され、発売されている。しかしながら、MAID 機能を有するディスクアレイ装置においても、ホストからの要求に対応するために、ディスクアレイ装置のコントローラや個々のディスクドライブの回路は通電されており、多数のディスクドライブを有するディスクアレイ装置では、消費電力は、まったくアクセスがない状態においても、ディスクドライブが通常回転時にあるときの 40~80%程度となっている[2][3]。そこで、より大幅な消費電力の削減を実現するため、リモートからのオン・オフ制御

が可能な電源コンセントを利用し、アクセス要求に応じてホスト側からディスクアレイ装置の電源を制御する機構の実装を行った。本稿では、本機構の実装法について述べ、実験システムにおいて消費電力、応答性能を測定した結果を報告するとともに、シミュレーションにより本機構による消費電力削減効果について評価を行う。

2. 実装方式

リモートからのオン・オフ制御を可能とする電源コンセントなどを用い、ホストからディスクアレイ装置の電源制御を可能とする。このような電源コンセントには、シリアルポートを通じて制御を行うもの、ネットワークを通じて行うものなどが製品化されており、ネットワークを利用するものには、TELNET, HTTP, メールなどを使用するものがあるが、いずれかの方式により、接続されたディスクアレイ装置の電源のオン・オフをホストからコマンドより制御可能とする。すなわち、シリアルポートを利用する場合には、シリ

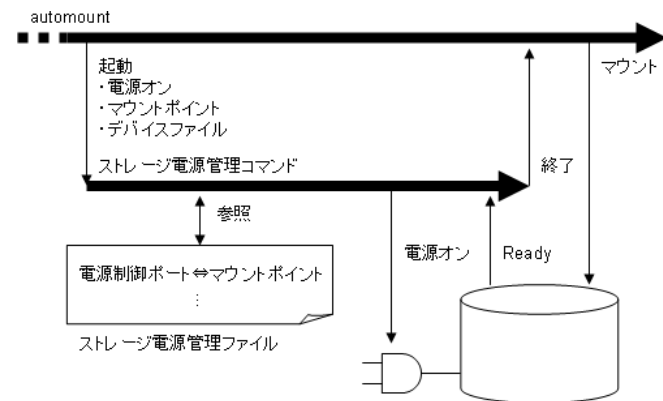


図 1 アクセス要求時の動作

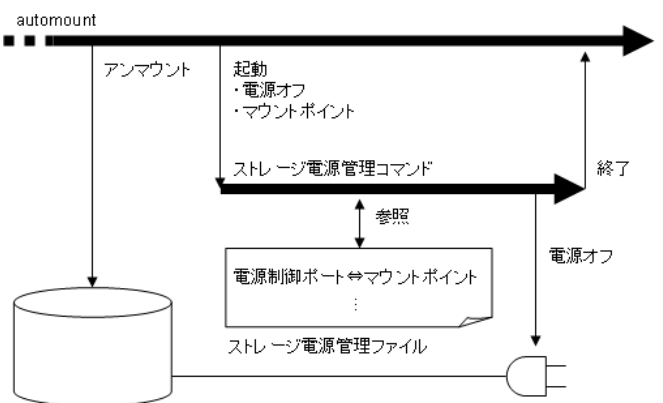


図 2 アクセス終了時の動作

アルポートより電源制御コマンドを送信する，HTTP を利用する場合には，電源制御を行う HTTP リクエストを発行するコマンドを作成する．

アクセス要求に対する自動的なディスクアレイ装置の制御には，UNIX，Linux などにおいて利用されている automount (自動マウント) デーモンを使用する． automount に対し，アクセス要求によりディスクのマウントが行われる直前，および，アクセス終了後，ディスクがアンマウントされる直後にストレージ電源管理コマンドを起動するよう変更をする．ストレージ電源管理コマンドは，マウント時には電源オンの指定，マウントポイント，デバイス名を引数として起動され，アンマウント時には，電源オフの指定，マウントポイントが引数として起動される．ストレージ電源管理コマンドは，マウントポイントと電源制御ポートの対応が記述されているストレージ電源管理ファイルを参照し，要求されたマウントポイントに対応する電源制御ポートに対して，電源オンコマンドを発行し，ディスクアレイ装置を起動する．1つのディスクアレイに対して複数の電源が接続されている場合には，ストレージ電源管理ファイルに複数の対応を記述し，ストレージ電源管理コマンドは，全てのポートに対して電源オンコマンドを発行する．その後，ストレージ電源管理コマンドは，引数により指定されたデバイスが使用可能となるまで待ち，その後，終了する (図 1)．アンマ

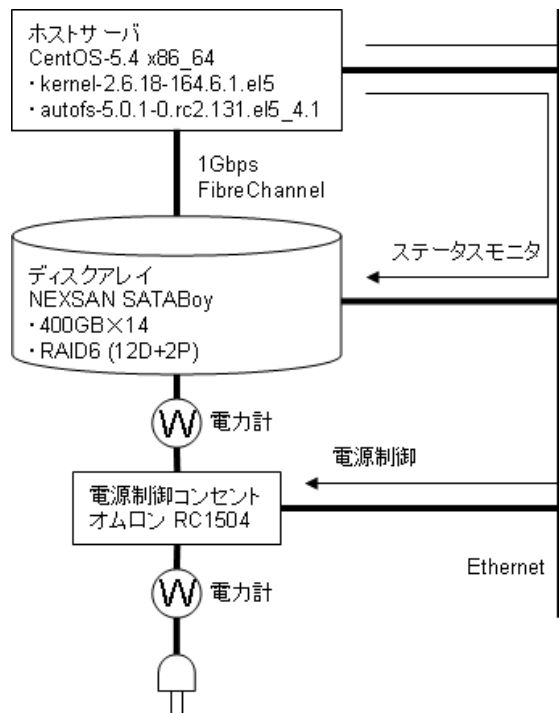


図 3 実験システムの構成

ウント時には，ストレージ電源管理コマンドは，1つのコンセントに対して複数のディスクアレイ装置が接続されている場合を考慮し，マウントポイントに対応する電源制御ポートに関して，要求されたマウントポイント以外の全てのマウントポイントにディスクがマウントされていない場合に，電源制御ポートに対して電源オフコマンドを発行し，終了する．このように，アクセス要求に対してディスクアレイ装置の電源制御を実現する (図 2)．

3. 実験システムによる測定結果

3.1. 実験システム

図 3 に，実験システムの構成を示す．ホストサーバの OS は CentOS-5.4 である． automount は，2009 年 12 月現在，CentOS-5.4 用に提供されている最新バージョンである autofs-5.0.1-0.rc2.131.el5_4.1 に含まれるものに2節で述べた実装方式による変更を加えた．ディスクアレイ装置は NEXSAN 社製の SATABoy である [4]．サーバとは 1Gbps のファイバチャネルにより接続されている．本実験システムで使用した SATABoy は日立グローバルストレージテクノロジーズ社製の 400GB SATA ディスク Deskstar 7K400 [5] を 14 本搭載しており，ストライプサイズを 128KB とし，12D+2P の RAID6 で構成し，全空間を一つのボリュームとして ext3 ファイルシステムを作成した．SATABoy はディスク自体に AutoMAID と呼ばれる MAID 機能を有している．アクセスが終了し，一定時間が経過すると，各ディスクドライブを通常状態 (active) から，ヘッドを退避させる

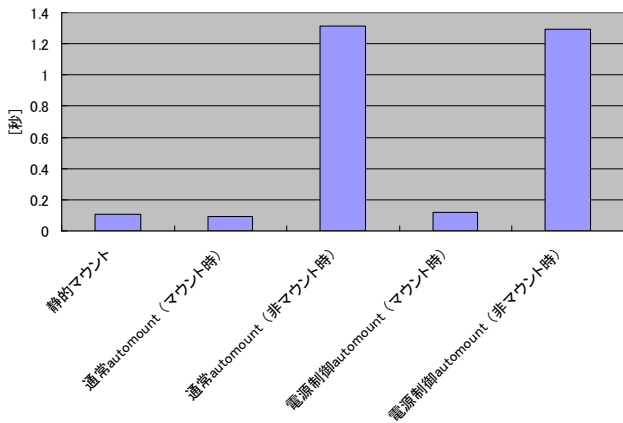


図 4 active 状態におけるアクセス時間

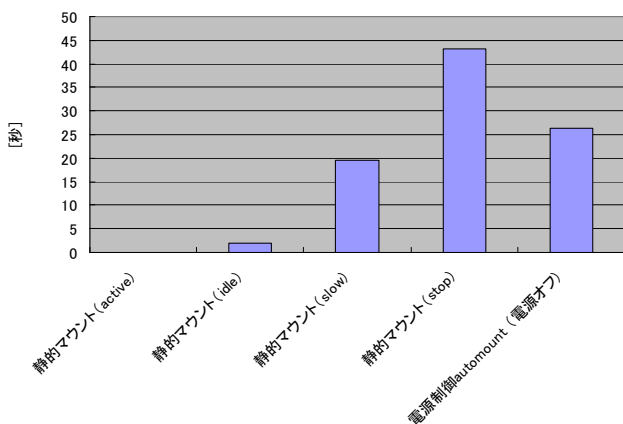


図 5 ディスクアレイの各状態からのアクセス時間

(idle), 回転数を落とす (slow), 回転を停止する (stop) の各状態へ移行し, 消費電力の低減を図る. 各ディスクの状態は, SATABoy が提供する HTTP による GUI を通じて確認可能である. 電源制御には, オムロン社製の RC1504 を用いた [6]. RC1504 はシリアルポートによる制御, TELNET による制御, HTTP による制御が可能であるが, 本実験システムでは, ホストサーバから RC1504 に対して HTTP リクエストを発行することで RC1504 の各ポートのオン・オフ制御を行う. また, RC1504 の入力部および出力部に電力計を挿入し, ディスクアレイ装置単体の消費電力, および電源制御装置を含めた消費電力を測定している.

3.2. 応答時間

ディスクアレイ上の全てのディスクドライブにデータがまたがるように 1.5MB のファイルを作成し, このデータを読み出すのに要する時間を測定した. AutoMAID 機能を有効にし, 全ディスクドライブの状態が一致していることを確認した後, 読み出しを行っている. 測定はそれぞれ 10 回行い, 平均値を求めている.

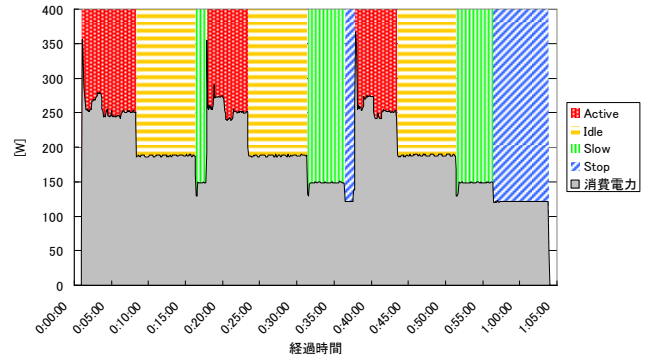


図 6 ディスクアレイ装置の消費電力の推移

図 4 は, ディスクアレイ装置のディスクドライブが active 状態であるときに, 1.5MB のデータを読み出すのに要した時間である. それぞれ, automount を使用せずに静的にマウントした場合, 通常の automount を使用し, 既にマウントされている状態でアクセスをした場合および非マウント状態でアクセスをした場合, 電源制御機能を付加した automount を使用し, 既にマウントされている状態の場合および非マウント状態の場合の結果である. 静的マウント, 通常 automount 使用時, 電源制御機能を付加した automount 使用時のそれぞれの比較より, automount の使用, および電源制御機能の付加によるオーバーヘッドはほぼないことが確認できる. また, マウント状態でのアクセスと非マウント状態からのアクセスの差より, マウント動作に約 1.2 秒要することがわかる.

図 5 は, ディスクアレイ装置のディスクドライブの各状態において, 1.5MB のデータを読み出すのに要した時間である. ディスクアレイの電源制御を行わない場合はディスクアレイを静的にマウントし, 電源制御を行う場合には automount を利用している. ディスクアレイ装置の電源がオフの状態でのアクセスには約 26 秒を要するのに対し, ディスクアレイ装置の電源オンで, 各ドライブが回転停止状態 (stop) の場合には平均で 40 秒以上を要している. これは, 電源投入時には全てのディスクドライブが同時に通常状態 (active) になるのに対し, 回転停止状態でリクエストを受けたときには, 必要なドライブのみの回転を試みるため, 各ドライブの起動が逐次的に実行されていることがあるためと考えられる.

3.3. 消費電力

図 6 はディスクアレイ装置の消費電力の推移である. ここでは, 全ディスクドライブが, ディスクアレイ装置電源オフ状態→アクセスリクエスト発行→active→idle→slow→アクセスリクエスト発行→active→idle→slow→stop→アクセスリクエスト発行→active→idle→slow→stop→ディスクアレイ装置電源オフと

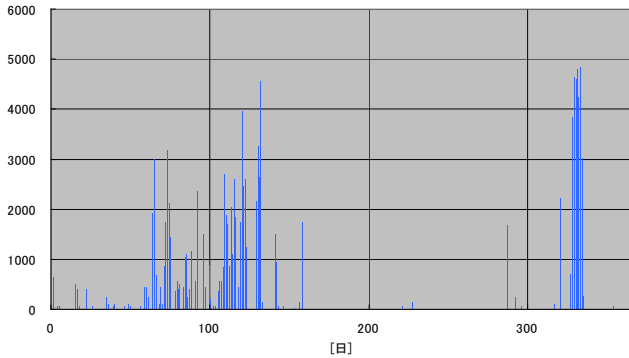


図 7 1日あたりのリクエスト数

変化した場合である。ディスクドライブの状態変化に従い、消費電力が削減されていることがわかる。しかし、全ディスクドライブの回転を停止させている場合においても約 120W の電力を消費しており、通常状態の消費電力の 45～50%程度にとどまっている。ディスクアレイ装置の電源制御を行う場合には、ディスクアレイ装置自体は電力を消費せず、電源制御装置の入力部、出力部の電力計の差より、電源制御装置の消費電力は 4W 未満であり、ディスクアレイ装置の電源制御を行うことにより、アクセスがない場合には大幅に消費電力を削減することが可能であることがわかる。

4. シミュレーションによる評価

4.1. シミュレーション条件

電源制御によるディスクアレイ装置の電力削減効果を確認するため、シミュレーションにより評価を行った。シミュレーションに用いたアクセスリクエストは、2008 年 12 月から 2009 年 11 月までの 1 年間の東京大学生産技術研究所においてアーカイブしている MTSAT 画像データに対する ftp 経由の 114071 件のアクセスリクエストである。図 7 に 1 日あたりのリクエスト数を示す。

ディスクアレイ装置は 100MB/s の転送帯域を持ち、SATABoy を想定して AutoMAID 機能を有するものとした。ディスクドライブが active 時にはリクエスト発行直後に転送が開始されるとし、ディスクドライブが idle, slow, stop 状態の場合には active 状態になるまでにそれぞれ 1.8 秒、19.4 秒、43.1 秒を要し、電源オフの状態からは 26.1 秒を要するものとした。また、active, idle, slow, stop 状態における消費電力は 260W, 190W, 150W, 120W とし、電源制御機能を有効にしている場合には、電源制御装置分として常に 3W を消費するものとしている。シミュレーションは、AutoMAID 機能、電源制御機能を用いず、ディスクドライブが常に active 状態である場合、電源制御機能のみを用い、アクセス終了後、電源がになるまでの時間を 120 秒およ

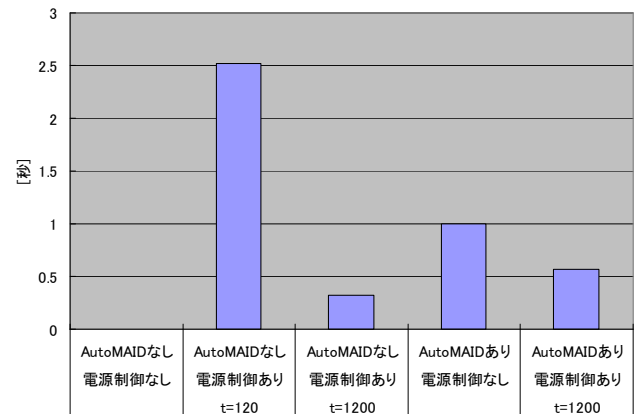


図 8 応答遅延時間

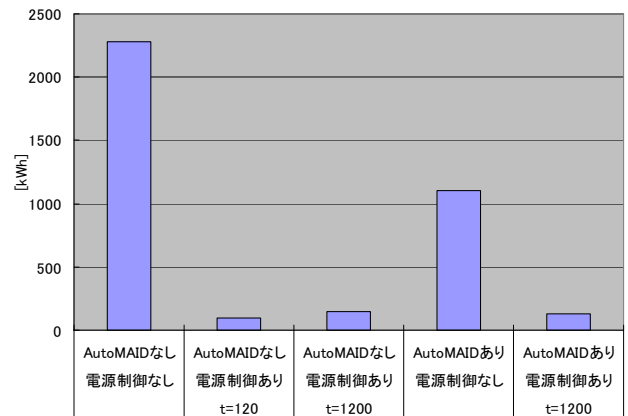


図 9 消費電力量

び 1200 秒とした場合、AutoMAID 機能のみを用いた場合、および AutoMAID 機能を用い、かつアクセスが終了してから 1200 秒後に電源をオフとする場合の 5 通りについて行った。AutoMAID 機能は、SATABoy において設定可能な最短時間とし、アクセスが終了してから 120 秒後、600 秒後、900 秒後にそれぞれディスクドライブが idle, slow, stop 状態になるとしている。

4.2. シミュレーション結果

4.2.1. 応答遅延時間

図 8 はリクエストが発行されてディスクドライブが active 状態となるまでに要する、各リクエストの応答遅延時間の平均値である。1200 秒後に電源をオフにする場合には、平均応答遅延時間は 1 秒未満である。AutoMAID によりディスクドライブが stop 状態である場合よりも電源オフの状態からの方が応答遅延時間が小さいため、AutoMAID 機能を用いない場合の方が平均の応答時間も短くなっている。電源オフまでの時間を 120 秒と短くすると、平均応答遅延時間は約 2.5 秒と劣化する。なお、電源オフとなった回数は、AutoMAID 機能の有無に関わらず電源オフまでの時間

が 1200 秒の場合には 243 回、120 秒の場合には 2169 回であった。

4.2.2. 消費電力量

図 9 はシミュレーション期間中のディスクアレイ装置の消費電力量である。電源制御を行う場合には、電源制御装置の消費電力量も含んでいる。AutoMAID 機能のみを用いた場合には約 50% の電力量が削減可能であるのに対し、電源制御機能を用いることで、90% 以上の電力量の削減が可能である。一般に MTSAT 画像をはじめとする地球観測データへのアクセスは高いバースト性を有しており、電源制御機能による消費電力削減効果が極めて有効であると考えられる。

5. おわりに

本稿では、リモートからのオン・オフ制御が可能な電源コンセントを使用し、automount を用いて、アクセス要求に応じてホスト側からディスクアレイ装置の電源を制御するの実装法について述べ、実験システムにおいて消費電力、応答性能を測定した結果を報告するとともに、シミュレーションにより、本機構による消費電力削減効果について評価を行った。電源制御を行うことにより、ディスクアレイ装置の消費電力を大幅に削減できることを示した。

参 考 文 献

- [1] U.S. Environmental Protection Agency ENERGY STAR Program, “Report to Congress on Server and Data Center Energy Efficiency, Public Law 109-431”, Aug. 2007.
- [2] 富士通株式会社, “富士通 ETERNUS ディスクアレイの MAID 技術による省エネルギー”, http://storage-system.fujitsu.com/jp/products/diskarray/download/pdf/MAID_whitepaper.pdf, Nov. 2007.
- [3] Greg Schulz, “MAID 2.0: Energy Saving without Performance Compromises, Energy Saving for Secondary and Near-line Storage Systems”, <http://www.nexsan.com/solutions/energysavings/whitepaper.php>, Jan. 2008.
- [4] ネクサン・テクノロジーズ・インク, “ハイパフォーマンス SATA ストレージ”, <http://www.nexsan.jp/sataboy.html>.
- [5] Hitachi Global Storage Technologies, “Feature Set Description, Hitachi Deskstar 7K400”, [http://www.hitachigst.com/tech/techlib.nsf/techdocs/427842CE2002C16D86256ECF006D773C/\\$file/Feature_Set.pdf](http://www.hitachigst.com/tech/techlib.nsf/techdocs/427842CE2002C16D86256ECF006D773C/$file/Feature_Set.pdf).
- [6] オムロン株式会社, “マルチコントロールコンセント RC1504 取扱説明書 詳細版”, http://www.omron.co.jp/ped-j/dengen/download/rc1504/rc1504_manual_j04.pdf.
- [7] 竹内渉, 根本利弘, 金子隆之, 安岡善文, “WWW を利用した MTSAT データ処理・可視化・配信システムの構築”, 写真測量とリモートセンシング, vol.46, no.6, pp.42-48, Nov.2007.