

期限付バックアップ中のディスク読み取り量の再計算による ユーザのファイル転送時間の短縮

高橋 風太¹ 串田 高幸¹

概要：課題は、ファイルサーバによるバックアップサーバへのファイルのバックアップ中にユーザがファイルサーバへファイルを転送すると、ファイルサーバがバックアップ中でない場合と比較してユーザのファイル転送時間が増加することである。提案では、バックアップの可用帯域幅に上限を設けバックアップ期限にバックアップが終了するよう定期的に再計算することで、ユーザが使用できる帯域幅を増加させ、ユーザのファイル転送時間を短縮する。可用帯域幅は、転送すべきバックアップのファイルサイズから、期限を超過しないバックアップの転送速度を予測することで定める。基礎実験では、ファイルサーバによる 100GB のファイルのバックアップ中にユーザ A, B がそれぞれ 100GB, 50GB のファイルを転送した場合のユーザ A のファイル転送時間は 26 分 21 秒、ユーザ A のみがファイルを転送した場合のファイル転送時間は 15 分 36 秒であった。したがって、課題の発生により、ユーザ A のファイル転送時間は約 40.8% 増加した。評価では、提案方式の適用の有無で、ユーザ A のファイル転送時間を比較する。

1. はじめに

背景

個人や企業は、ハードウェアやソフトウェアの障害、災害が発生した場合にデータの損失を避けるために、データをリモートでアーカイブすることが求められる [1]。自然災害によるデータの損失を防ぐため、バックアップは期間内に終了するようにする [2]。企業は通常、バックアップが業務に影響することを防ぐため、バックアップを真夜中に実行する。バックアップ期間は企業によって異なる。本研究では、バックアップの期限を企業の業務開始時刻とする。

バックアップの利用例として、東映アニメーションが挙げられる。ここでは映像やテレビ、キャラクター商品の開発事業が展開され、これらのコンテンツの制作のため、画像ファイルや動画ファイルを業務に用いるデータ(以下業務データ)として共有する。複数のコンテンツ制作を同時進行する場合があるため、オンプレミスのシステムは 24 時間稼働している^{*1}。

ユーザ間のデータの共有や管理は、ファイルサーバや Network Attached Storage を使用して行われる [3]。ファ

イルサーバのデータを保護するため、ファイルサーバの業務データはバックアップサーバへバックアップされる。

ファイルサーバがバックアップサーバにファイルをバックアップしているときにユーザがファイルサーバにファイルを転送すると、ファイルサーバがバックアップしていないときに比べてユーザのファイル転送時間が増加する。

既存研究では、バックアップの終了時間が業務開始時間を超過しない程度にバックアップ速度を制限した [4]。ユーザが利用できるディスク帯域幅が増加することで、バックアップ時のユーザのファイル転送時間を約 4.5% 短縮した。この研究では、3つの条件が前提となっている。1つ目は、バックアップ速度を減少させてもバックアップ時間が期限を超過しないことが条件である。2つ目は、バックアップとユーザのファイル転送が同時に発生することが条件である。3つ目は、ファイルサーバにファイルを転送するユーザが同時に 2 人以上にならないことが条件である。

マルチユーザ環境では、各ユーザに対するサービス品質は公平である必要がある [5-7]。

課題

ユーザが同時に 2 人以上ファイルサーバへファイルを転送すると、ファイル転送時間が増加する。ユーザのファイル転送時間が他ユーザのファイル転送やバックアップによって 38% 増加すると、ユーザはイライラする [8]。

図 1 にファイル転送時間増加の概要を示す。

¹ 東京工科大学大学院 バイオ・メディア研究科コンピュータサイエンス専攻
〒192-0982 東京都八王子市片倉町 1404-1

^{*1} Tech Target Japan ”社外とも簡単データ共有、「東映アニメーション」のアニメ制作を下支えするツールとは”<https://techtarget.itmedia.co.jp/tt/news/1704/12/news07.html>(参照 2023-7-4)

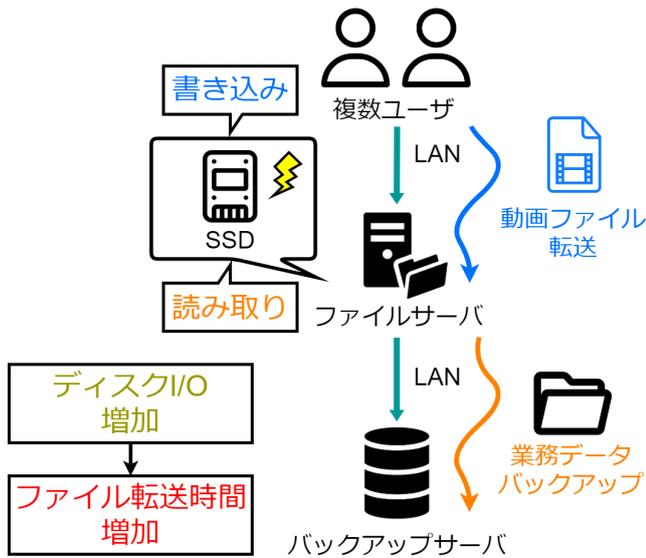


図 1 ファイル転送時間増加の概要

既存研究の提案方式を適用した場合のバックアップ終了時間の平均値は、1683[s]であった [4]。この時のバックアップ期限は1800[s]であり、平均値との差は約6.5%である。したがって、バックアップ終了時間をバックアップ期限に近づけることで、ユーザのファイル転送時間をより短縮する必要がある。

各章の概要

第2章では、関連研究について述べる。第3章では、本研究の提案方式とユースケースについて述べる。第4章では、提案方式の実装と評価のための実験方法について述べる。第5章では、提案方式の評価と分析について述べる。第6章では、本研究が残す議論について述べる。第7章では、本研究の成果とまとめについて述べる。

2. 関連研究

効率と公平性に基づいたマルチユーザー光グリッドネットワーク上の大容量ファイル転送のための動的スケジューリングアルゴリズムを提案している研究がある [6]。大規模ファイル転送の問題に対処するための4つの動的スケジューリングアルゴリズムが提案されているが、マルチユーザー光グリッド環境を前提としている。したがって、ユースケースを映像制作会社としている本研究には適用できない。

信頼性の高いマルチキャストベースのアプローチを提案している研究がある [9]。課題は、データセンターにおけるデータダウンロードでは、中央のファイルシステムがボトルネックになることである。また、この課題が複数のVMによるファイルサーバからのダウンロードでも発生することを指摘している。これを解決するため、MCDシステムを提案している。MCDシステムは、物理ネットワーク

トポロジに基づいて論理マルチルートツリーを構築し、2フェーズアプローチの論理ビューを使用する。論理ビューへの変換により、階層構造を置き換える。この階層構造により、ルーターレベルでのパケット損失を処理する。2フェーズアプローチの最初のフェーズでは、データはすべてのノードにマルチキャストされる。2番目のフェーズでは、信頼性の低いマルチキャストの欠点を補うため、論理ツリーを使用して効率的なエラー訂正を行う。評価では、提案方式により、既存のシステムより9.9倍から14.29倍高速になったことを示している。しかし、データダウンロードのみに焦点を当てており、アップロード時間について評価していない。また、バックアップ時間が期限を超過する可能性について触れていない。本研究のユースケースである映像制作会社では、ファイルサーバがバックアップによりアップロードをする想定である。したがって、この研究により本研究の課題を解決することができない。

チャンクベースのファイルバックアップのためのスケラブルな並列重複排除を提案している研究がある [10]。課題は、ランダムな順序で到着するファイルの重複排除の処理に要する時間が、ディスクのスループットのボトルネックにより低くなることである。ファイルの類似性を利用し、ファイルごとのチャンク探索に1つのディスクアクセスのみを行うことで、ボトルネックの問題を軽減している。評価実験の結果、フラットチャンクインデックスに必要なRAMと提案手法に必要なRAM使用量の比率は86.56:1であった。ディスクのボトルネックの原因はRAM使用量がオーバーフローするとディスクにページングする必要があることであり、筆者らはボトルネックの問題を軽減している。しかし、ファイルサーバでの重複排除中にユーザがファイルを転送した場合、ユーザのファイル転送時間は増加する。

ペタスケールコンピューティングシステムにおける、進行中のI/O要求を調整するための新しいI/O対応バッチスケジューリングフレームワークを提案している研究がある [11]。課題は、複数のアプリケーションから同時にストレージへアクセスすることにより引き起こされるI/Oの輻輳が原因で、パフォーマンスが4倍抑制されていることである。実験結果として、IBM Blue Gene/Qシステムと呼ばれるスーパーコンピュータのジョブパフォーマンスを約30%向上させている。しかし、本研究は映像制作会社がユースケースであるため、スーパーコンピュータを導入し、環境を変更する必要がある。

クラスターの集中ファイルサーバにおいて、新しいヒントベースのI/Oメカニズムを提案している研究がある [12]。課題は、共有ストレージに複数の並列アプリケーションがアクセスすると、異なるクライアントからのI/O要求が干渉し、ファイルサーバのI/Oパフォーマンスが低下することである。評価では、NFS(Network File System)と比

較している。結果として、読み取り速度が約2倍増加している。しかし、本研究のユースケースでは、ファイル転送を行うユーザが1人の場合であっても、バックアップ中のユーザのファイル転送時間が増加する。したがって、異なるクライアントからの要求を前提としているこの研究では本研究の課題を解決できない。

3. 提案

提案方式

既存研究では、バックアップの帯域幅を絞っている [4]。その分ユーザが帯域幅を使用できるので、ファイル転送時間を短縮できる。しかし、実際にバックアップされるファイルサイズが予測通りにならず、予測している転送済みのファイルサイズと、実際に転送済みのファイルサイズの差が時間の経過とともに大きくなる。したがって、制御する帯域幅の計算をバックアップ中に再度行う。以下の図2に、提案の概要を示す。

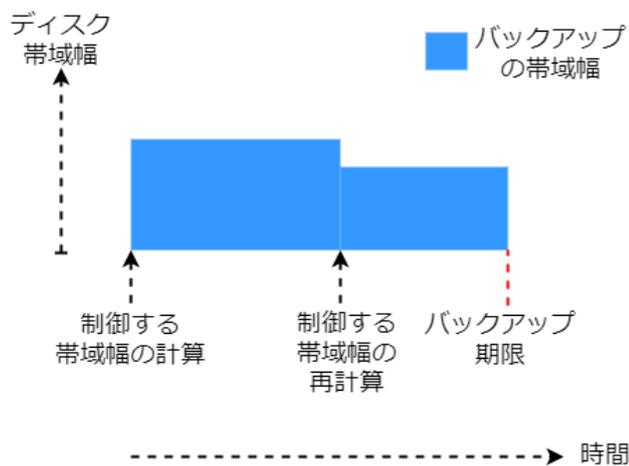


図2 提案の概要

帯域幅の計算を再度行うことで、時間の経過とともに大きくなる予測している転送済みのファイルサイズと、実際に転送済みのファイルサイズの差を0にすることができる。これにより、帯域幅の制御がより正確になる。

再計算は定期的に行われる。定期的にバックアップの終了時間を予測するが、これがバックアップの期限となる時刻を超えていた場合、その時点で再計算が行われる。これにより、バックアップの終了時間が期限を超過することを防止する。

既存研究と同様、提案方式はバックアップ期間中にユーザがファイルを転送した場合に適用される [4]。バックアップの帯域幅の制御のため、ディスクの読み取り量に上限を設ける。1秒あたりのディスク読み取り量の上限を S [MB/s] とする。バックアップ予定の業務データの容量を C [MB] とする。バックアップ期限までの時間を L [s] とする。バックアップ速度の最大値と、ユーザからのファイル転送中の

バックアップ速度の最小値の差を R [MB/s] とする。この時、ディスク読み取りの制限決定の条件を式1に示す。

$$S = \frac{C}{L} + R \quad (1)$$

R は事前に基礎実験をすることにより、バックアップ速度の最大値及びユーザからのファイル転送中のバックアップ速度の最小値を算出する。これは、実際に転送されるファイルサイズが実際に転送されるファイルサイズより小さくするためである。これにより、バックアップの終了時間がバックアップの期限を超過することを防止している。既存研究では、実際にバックアップの終了時間がバックアップの期限を超過することはなかった [4]。

ユースケース・シナリオ

本研究の提案方式は、映像制作会社での適用を想定している。下記の図3に、ユースケースの概要を示す。

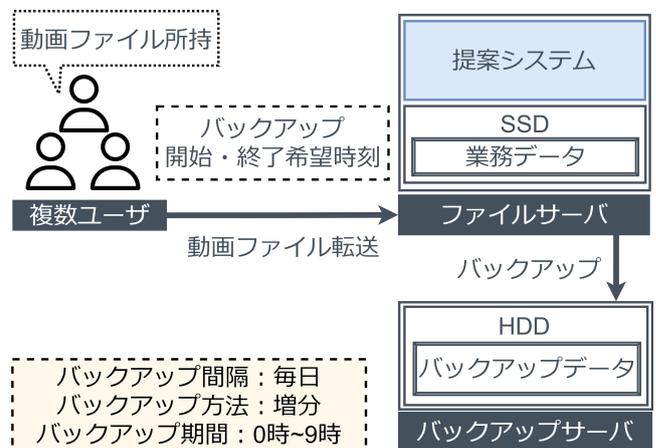


図3 ユースケースの概要

納期が近づき、映像制作会社の社員であるユーザは深夜残業を行う*2。ユーザは動画ファイルを共有するため、バックアップの実行中であるファイルサーバへ動画ファイルを転送することを想定する。バックアップ期間は0時9時であり、9時までにバックアップを終了する必要がある。これは、業務に支障を来さないためである。ユーザはバックアップの開始・終了希望時刻を提案システムに入力することで、課題の原因であるファイル転送時間の増加及びバックアップ期限の超過を防ぐことができる。ファイルサーバのディスクはSSDである。これは、処理速度を優先するためである。バックアップサーバのディスクはHDDである。これは、バックアップ容量を優先するためである。ファイルサーバはバックアップサーバに1日の差分を転送する。データの流出を避けるため、ローカルでのバックアップを想定する。

*2 キャリアガーデン "映像制作会社社員の勤務時間・休日・残業は多い?" <https://careergarden.jp/eizouseisakukaisha/kinmujikan/>(参照 2023-7-4)

4. 実装

提案方式を実装したソフトウェアをファイルサーバに配置する。ファイルサーバに配置されたソフトウェアは、ユーザからのファイル転送をトリガーに、バックアップ速度を期限が過ぎない程度に確保する。これにより、ユーザのファイル転送時間がバックアップ速度を変更しない場合と比較して短縮される。以下の図4に、提案ソフトウェア配置の概要を示す。

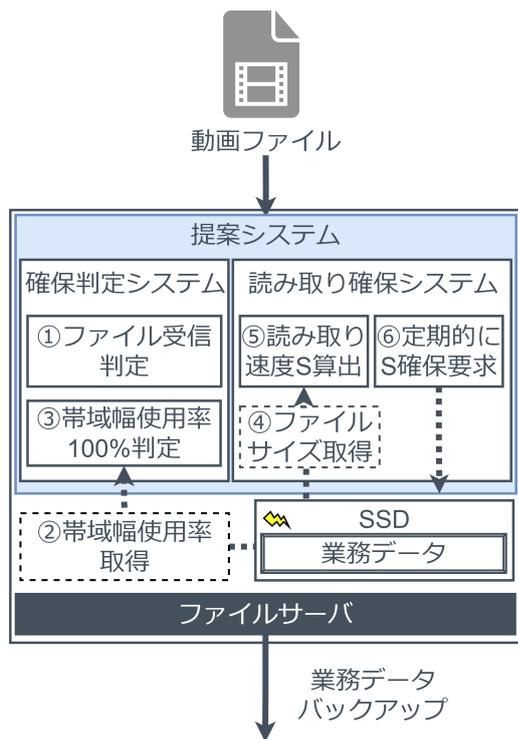


図4 提案ソフトウェア配置の概要

確保判定システムの①ファイル受信判定では、動画ファイルがバックアップ期間中に受信されるかを判定する。②帯域幅使用率取得により、確保判定システム内の③帯域幅使用率が100%かを判定する。帯域幅の使用率の取得には、`iowait` の`%util`を使用する。①及び③をどちらも満たした場合、提案システムは読み取りの確保を開始する。読み取り確保システムの⑤でバックアップの読み取り速度を算出するため、④でファイルサイズを取得する。読み取り速度を算出した後、⑥で定期的にバックアップの読み取り速度の確保を要求する。

5. 実験

実験では、ユーザのファイル転送時間が他ユーザのファイル転送やバックアップによって38%以上増加するか、すなわち実験環境において課題が発生するかを確認する。

課題の基礎実験として、2つの実験を行い課題を確認する。この2つの実験からユーザのファイル転送時間を比

較することで、ファイル転送時間増加の原因が他ユーザのファイル転送やバックアップによるものかを確認する。実験①は、単一ユーザのファイル転送時間を計測する実験である。実験②は、課題となるユーザのファイル転送時間を計測する実験である。

`rsync` コマンドを用いて動画ファイルを転送する。転送する動画ファイルのサイズは、映像制作会社において100GBの動画ファイルを転送する例より、ユーザ、ファイルサーバともに約100GBであることを想定する*3。実験での動画ファイル及び業務データは、実際の動画ファイルとのファイル転送時間に有意な差はないという前提のもと、ランダムなテキストがデータとして書き込まれているダミーファイルを使用する。以下の図5に、課題の基礎実験の概要を示す。

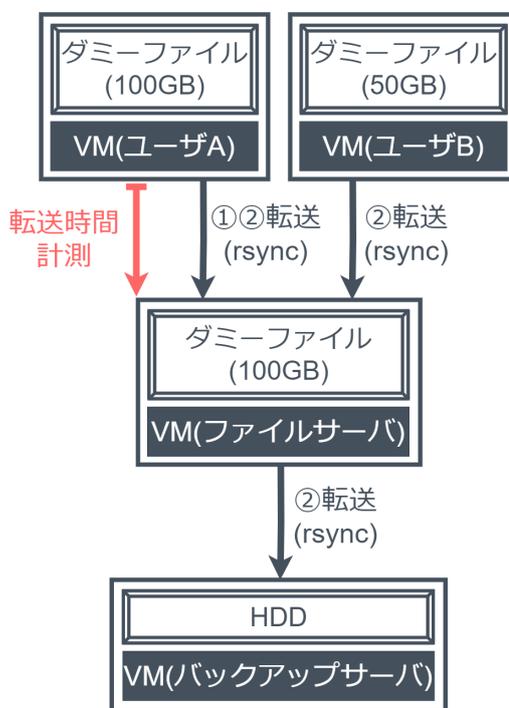


図5 課題の基礎実験の概要

図5の実験①では、ユーザAが100GBの動画ファイルを想定したダミーファイルを所持している。ユーザAはファイルサーバにファイルを転送する。その際、ユーザAがファイルを転送を開始してから転送が終了するまでの時間を計測する。

図5の実験②では、ユーザAが100GBの動画ファイルを想定したダミーファイルを所持している。ユーザAはファイルサーバにファイルを転送する。その際、ユーザAがファイルを転送を開始してから転送が終了するまでの時間を計測する。実験①と異なり、同時にユーザBが50GBの

*3 DIREXT!EXTREME ”導入企業インタビュー映像制作プロジェクト” <https://www.directextreme.com/case/interview2.html>(参照 2023-7-9)

ファイルを転送する。また、ファイルサーバは 100GB の業務データを想定したダミーファイルをバックアップサーバの HDD へバックアップする。

実験環境

実験に用いるサーバは、ハイパーバイザの ESXi から作成される仮想マシン (以下 VM) である。以下に、課題の基礎実験に用いる VM の構成情報を示す。vCPU は、VM に割り当てる CPU である。4つのサーバを作成し、役割をそれぞれユーザ A、ユーザ B、ファイルサーバ、バックアップサーバとする。

- VM 構成情報 (ユーザ A)
OS Ubuntu-20.04.2
vCPU 2 コア
RAM 8GB
HDD 120GB
- VM 構成情報 (ユーザ B)
OS Ubuntu-20.04.2
vCPU 2 コア
RAM 8GB
HDD 120GB
- VM 構成情報 (ファイルサーバ)
OS Ubuntu-20.04.2
vCPU 2 コア
RAM 8GB
HDD 300GB
- VM 構成情報 (バックアップサーバ)
OS Ubuntu-20.04.2
vCPU 2 コア
RAM 8GB
HDD 120GB

実験結果と分析

図 6 は、ユーザ A が 100GB、ユーザ B が 50GB のファイルをファイルサーバへ転送し、ファイルサーバが 100GB のファイルをバックアップサーバへ転送した際の、ユーザ A のファイル転送の増加時間である。図 6 では、ユーザ A のみのファイル転送では、ユーザ A のファイル転送時間は 15 分 36 秒である。ユーザ A のファイル転送時にユーザ B 及びファイルサーバのファイル転送が重なった場合、ユーザ A のファイル転送時間は 26 分 21 秒である。ユーザ A、B 及びバックアップの場合のユーザ A のファイル転送時間は、ユーザ A のみのファイル転送の場合より 40.8%増加している。

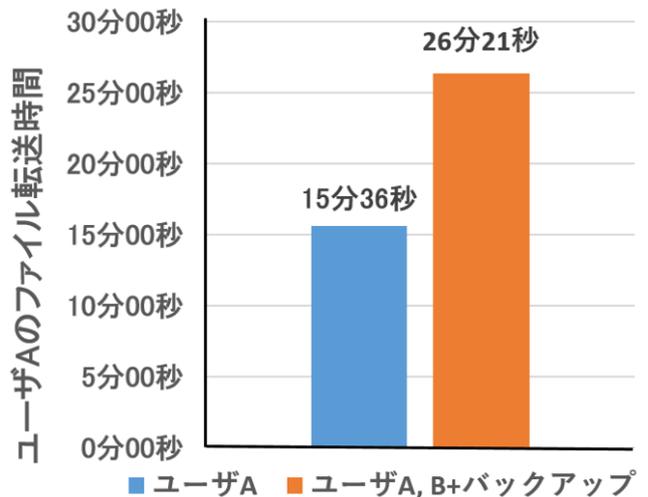


図 6 ユーザ A のファイル転送の増加時間

6. 議論

提案方式では、バックアップの帯域幅を再計算することで、最もファイル転送時間が長いユーザのファイル転送時間を短縮することができる。しかし、これは各ユーザのファイル転送時間を平等にすることで成り立つ提案方式であるため、ファイル転送時間が遅くなるユーザの不満が発生する可能性がある。ユースケースは映像制作会社であるため、転送ファイルがどのプロジェクトによるものかを把握し、最も進めるべきファイルを優先して転送することで不満を解消することができる。

7. おわりに

ファイルサーバがバックアップサーバにファイルをバックアップしているときにユーザがファイルサーバにファイルを転送すると、ファイルサーバがバックアップしていないときに比べてユーザのファイル転送時間が増加する。バックアップの帯域幅を絞り定期的に再計算することで、ユーザが使用できる帯域幅が増加し、ファイル転送時間を短縮できる。ファイル転送時間の短縮により、ユーザのイライラを抑制することに貢献できる。

参考文献

- [1] Rahumed, A., Chen, H. C., Tang, Y., Lee, P. P. and Lui, J. C.: A secure cloud backup system with assured deletion and version control, *2011 40th International Conference on Parallel Processing Workshops*, IEEE, pp. 160-167 (2011).
- [2] Yao, J., Lu, P. and Zhu, Z.: Minimizing disaster backup window for geo-distributed multi-datacenter cloud systems, *2014 IEEE International Conference on Communications (ICC)*, pp. 3631-3635 (online), DOI: 10.1109/ICC.2014.6883885 (2014).
- [3] Mesnier, M., Ganger, G. and Riedel, E.: Object-based storage, *IEEE Communications Magazine*, Vol. 41, No. 8, pp. 84-90 (online), DOI:

- 10.1109/MCOM.2003.1222722 (2003).
- [4] 高橋風太, 串田高幸: ファイルサーバにおける期限付バックアップ中のユーザの使用帯域幅の動的制御, 技術報告45, 東京工科大学コンピュータサイエンス学部, 東京工科大学コンピュータサイエンス学部 (2023).
- [5] Su, G.-M., Han, Z., Wu, M. and Liu, K. R.: A Scalable Multiuser Framework for Video Over OFDM Networks: Fairness and Efficiency, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 16, No. 10, pp. 1217–1231 (online), DOI: 10.1109/TCSVT.2006.883513 (2006).
- [6] Hu, M., Guo, W. and Hu, W.: Dynamic Scheduling Algorithms for Large File Transfer on Multi-user Optical Grid Network Based on Efficiency and Fairness, *2009 Fifth International Conference on Networking and Services*, pp. 493–498 (online), DOI: 10.1109/ICNS.2009.37 (2009).
- [7] Qin, C., Guo, W., Sun, W., Jin, Y. and Hu, W.: Scheduling strategies for multiple optical grid applications based on scheduling span and fairness, *Network Architectures, Management, and Applications VI* (Hu, W., Liu, S.-K., Ichi Sato, K. and Wosinska, L., eds.), Vol. 7137, International Society for Optics and Photonics, SPIE, p. 713715 (online), DOI: 10.1117/12.803530 (2008).
- [8] Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J. and Shneiderman, B.: Determining Causes and Severity of End-User Frustration, *International Journal of Human-Computer Interaction*, Vol. 17, No. 3, pp. 333–356 (online), DOI: 10.1207/s15327590ijhc1703.3 (2004).
- [9] Kaiser, J., Meister, D., Gottfried, V. and Brinkmann, A.: MCD: Overcoming the Data Download Bottleneck in Data Centers, *2013 IEEE Eighth International Conference on Networking, Architecture and Storage*, pp. 88–97 (online), DOI: 10.1109/NAS.2013.18 (2013).
- [10] Bhagwat, D., Eshghi, K., Long, D. D. E. and Lillibridge, M.: Extreme Binning: Scalable, parallel deduplication for chunk-based file backup, *2009 IEEE International Symposium on Modeling, Analysis & Simulation of Computer and Telecommunication Systems*, pp. 1–9 (online), DOI: 10.1109/MASCOT.2009.5366623 (2009).
- [11] Zhou, Z., Yang, X., Zhao, D., Rich, P., Tang, W., Wang, J. and Lan, Z.: I/O-Aware Batch Scheduling for Petascale Computing Systems, *2015 IEEE International Conference on Cluster Computing*, pp. 254–263 (online), DOI: 10.1109/CLUSTER.2015.45 (2015).
- [12] Chen, H., Xiong, J. and Sun, N.: A novel hint-based I/O mechanism for centralized file server of cluster, *2008 IEEE International Conference on Cluster Computing*, pp. 194–201 (online), DOI: 10.1109/CLUSTR.2008.4663771 (2008).