

大規模データ収集システムの高効率化に対する DCB の適用性評価

Applicable Evaluation of DCB for High Efficiency of a Large-Scale Data Acquisition System

西村 俊彦[†]
Toshihiko Nishimura今池 怜生[†]
Reo Imaike岩井 慎之介[†]
Shin-nosuke Iwai長坂 康史[†]
Yasushi Nagasaka

1. はじめに

近年の情報通信技術の発展により、ネットワーク上には多種多様なデータが大量に発生し、それらを活用するために、高速かつ効率的に収集、解析することが求められている。その利用形態の一つとして、現在様々な分野でネットワーク分散型データ収集システムが活用されている。

これらの状況は基礎科学実験分野においても同様である。特に ATLAS 実験^[1]などに代表される大量のデータを取り扱う高エネルギー物理学実験において、実験データの収集のために用いられるデータ収集システムでは、処理効率を高めるための並列処理が行われ、一つの事象を多くの測定器で測定するため、そのトラフィックは多対一通信となる。

しかし、このトラフィックの特殊性から、現状では通信経路中での輻輳によるフレームロスが発生し、通信速度の公平性の低下が確認されている^[2]。この問題に対してロスレスイーサネット技術である DCB (Data Center Bridging) を用いる手法が提案・検証されているが、実験規模の拡大時におけるデータ収集システムの大規模化に対する適用性は確認されていない^[3]。

そこで本研究では、データ収集システムの通信効率向上のため、システムの大規模化に対する DCB の適用性を評価することを目的とする。

2. データ収集トラフィックの特殊性

一般的なネットワークトラフィックは、ネットワーク上の複数の送信端末から任意の受信端末に対して発生し、状況に応じて宛先の受信端末が変化する。すなわち、多対多の非同期トラフィックである性質を持つ。

これに対して、本研究が対象とする ATLAS 実験などで用いられる高エネルギー物理学実験におけるデータ収集システムでは、同時刻に発生したデータをまとめて処理するために、千台規模の複数のコンピュータ上に散在する測定データを、一斉に一台の解析用コンピュータに集約する必要がある。そのため、データ収集システムにおいて発生するトラフィック (データ収集トラフィック) は、多対一の同期トラフィックとなる。そのトラフィックの総量は、例えば ATLAS 実験の場合では、毎秒約 100 GByte 以上となる。

3. DCB (Data Center Bridging)

DCB (Data Center Bridging) は、フレームロスの発生しない確実なパケット転送を保証する FibreChannel のフレームを、イーサネット上で通信する FibreChannel over Ethernet (FCoE) を実現するためのロスレスイーサネット技術である。DCB の主な特徴として、従来のネットワークでは困難であったフレームロスの完全な抑制、すなわち無損失なパケット転送が可能な点や、従来のイーサネットおよび TCP/IP

との互換性を持つ点が挙げられる。

DCB におけるロスレス機構の実現には、従来のイーサネットの機能を拡張した一連の規格が用いられる。現在主要なものとして、優先度に基づいてトラフィックを区別する Priority-based Flow Control (PFC) や、輻輳の検知・制御を行う Congestion Notification (CN)、優先度に基づいた通信帯域の保障を行う Enhanced Transmission Selection (ETS) が存在する。

4. 性能評価

4.1 概要

本研究では、千台規模の機器が介在する大規模なネットワークを持つデータ収集システムに対する DCB の適用性を検証するため、計算機シミュレーションを行った。シミュレーションは、分散イベント駆動型ネットワークシミュレータである Network Simulator 3 (NS3) に、Brent ら^[4]によって開発された拡張パッケージを導入することで行った。

4.2 シミュレーション条件

大規模データ収集システムに対する DCB の有効性を検証するため、データ収集システムの構成を再現した多対一のスイッチングネットワークをシミュレータ上で構築した。シミュレーションに用いた計算機の仕様を表 1 に示す。

シミュレーションは、N 台のクライアントから 1 台のサーバに向けて同時に TCP トラフィックを送信することで行う。このとき、クライアント数 (N) は 1 から 2,048 までの累乗ずつ増やし、N 台のクライアントの平均スループットと合計フレームロス数を算出する。N が 64 の場合のネットワーク構成を図 1 に示す。

表 1 計算機仕様

CPU	Intel(R) Xeon(R) CPU E5-2620 v3
Memory	16,384 Mbyte
OS	Ubuntu 14.04 LTS x64_86
HDD	1 TByte

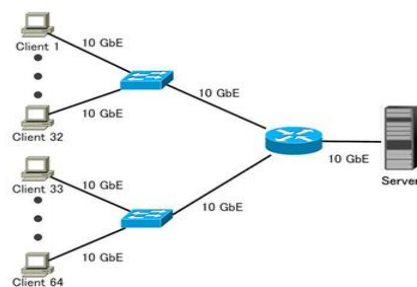


図 1 シミュレーションネットワーク構成 (N = 64)

[†] 広島工業大学, Hiroshima Institute of Technology

本シミュレーションでは、48 port のスイッチの利用を前提とし、N の増加に応じてスイッチ数を増やす。例えば、N が 64 である図 1 のネットワークの場合、スイッチ数は計 3 台であり、N が 2,048 の場合、スイッチ数は計 45 台となる。

また、シミュレーション時間は 60 秒で、通信帯域はすべて 10 Gbps、MTU を 1,500 Byte、送信データサイズは実際のデータ収集システムにおける典型的な値である 1,024 Byte とする。そして、DCB を無効にした従来のイーサネットを用いた TCP/IP 通信との比較を行う。

4.3 シミュレーション結果

クライアント数の増加に対する通信速度とフレームロスを図 2 に示す。図 2 より、N が 1 ~ 2,048 までの全ての場合において、通信速度に差異が見られなかった。このとき、DCB 無効時の合計フレームロス数は、N の増加に比例して増加傾向にあり、最小値が N が 1 の場合の 6,225 Frame で、最大値が N が 1,024 の場合の 61,9069 Frame であった。これに対して、DCB 有効時は、N がどのような場合においても発生しなかった。

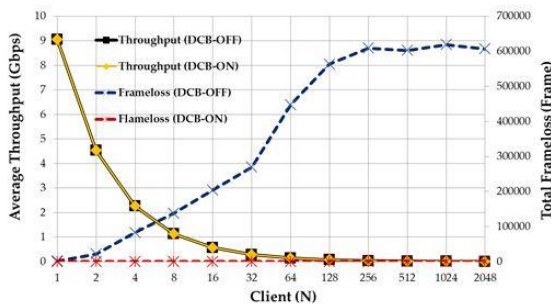


図 2 クライアント数に対する通信速度とフレームロス

4.4 Fairness Index による公平性評価

データ収集システムの通信効率の向上には、各フローにおける通信速度のばらつきの少ない状態、すなわち公平性の高い通信が必要とされる。そのため、シミュレーション結果における各フローの通信速度の公平性を評価する。本研究では通信速度の公平性を測る指標として、Fairness Index を用いる^[5]。

Fairness Index の値 f は、ネットワーク全体の合計フロー数を k 、各 TCP コネクション i のスループットを x_i として定義したとき、式(1)によって算出される。

$$f = \frac{(\sum_{i=1}^k x_i)^2}{k \sum_{i=1}^k x_i^2} \quad (1)$$

f は 0.0 から 1.0 までの値となり、1.0 に近いほど x_i が公平、すなわち各フロー間の公平性が高いことを示す。

Fairness Index の算出結果を図 3 に示す。図 3 より、DCB 無効時のフロー数が 1 から 128 までの場合は f が 0.9 以上であり、高い公平性を維持している。しかし、フロー数 256 以降は急激に低下し、2,048 のとき、約 0.075 と極めて公平性が低い結果となった。一方、DCB 有効時はフロー数の増加に関わらず、約 1.0 と高い公平性を維持していることがわかる。

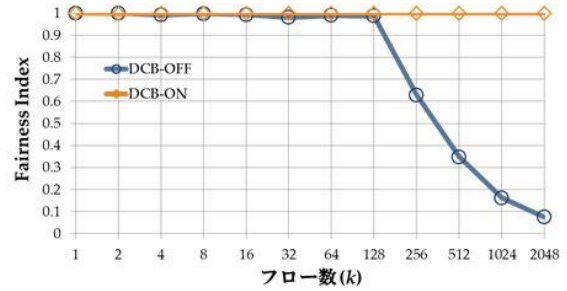


図 3 フロー数に対する Fairness Index

5. 考察

計算機シミュレーションの結果より、送信クライアント数の増加に対する DCB の有効、無効による通信速度の有意な差が見られなかった。このことから、DCB 有効時は無効時と同様の通信速度を維持できるといえる。さらに、DCB 有効時はどの環境においてもフレームロスが発生しないことから、特殊性の高いデータ収集システムネットワークにおいても、DCB が有効に機能すると考えられる。

また、通信速度の公平性評価の結果より、DCB 有効時は無効時と比較して、フロー数の増加に対して約 1.0 と極めて高い公平性を維持していることがわかる。特に、フロー数 2,048 の場合においては、無効時と比べて Fairness Index が約 0.925 の高いことから、DCB の有効化によって各フローの通信速度のばらつきを大幅に抑え、公平性の高い通信が可能となると考える。

これらのことから、大規模データ収集システムに対して DCB を適用することで、既存の通信方式と比べて大幅な信頼性、公平性の向上による通信の高効率化が期待できる。

6. まとめ

本研究では、データ収集システムが抱えるフレームロスによる通信効率の低下という課題に対して、ロスレスイーサネット技術である DCB を用いて解決することを目的としている。そのため、従来では困難であったデータ収集システムの大規模化に対する DCB の適用性を計算機シミュレーションによって検証した。その結果、最大 2,048 対 1 の環境において、既存のシステムが抱える問題を解決し、通信効率の向上が可能になることが明らかとなった。

今後の課題として、データ収集システム環境での DCB の確実な有効性を追求するため、実環境を用いた DCB の性能評価が必要である。また、DCB 環境での通信効率向上のため、最適な輻輳制御アルゴリズムの開発・評価を行う。

参考文献

- [1] Jinlong Zhang, *et al.*, "ATLAS Data Acquisition", Real Time Conference, pp.240-243 (2009).
- [2] 沖恭志, 長坂康史, "大規模データ収集システムにおける通信の効率化に関する研究", 広島工業大学紀要 研究編, Vol.47, pp.127-131 (2013).
- [3] 山本戸啓亮, 西村俊彦, 長坂康史, "大規模データ収集システムにおける DCB を用いた通信効率の評価", 2016 年電子情報通信学会総合大会講演論文集, pp.9 (2016).
- [4] Brent Stephens & Alan L. Cox, Ankit Singla, John Carter & Colin Dixon & Wesley Felter, "Practical DCB for improved data center networks", Proceedings of IEEE INFOCOM, pp.1824-1832 (2014).
- [5] D.-M. Chiu & R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks", Computer Networks and ISDN Systems, vol. 17, pp. 1-14 (1989).