

データセンターにおける省電力な 仮想ネットワーク構築手法

大阪大学大学院情報科学研究科
情報ネットワーク学専攻
○樽谷優弥 大下裕一 村田正幸

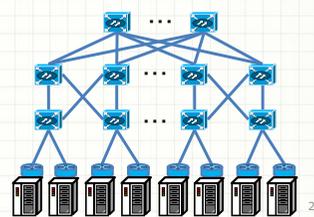
Advanced Network Architecture Research Group
http://www.anarg.jp/

データセンター

- 多数のサーバーとサーバーを結ぶネットワークで構成
- サーバー間が連携してデータを処理
 - 分散ファイルシステム
 - 分散コンピューティング

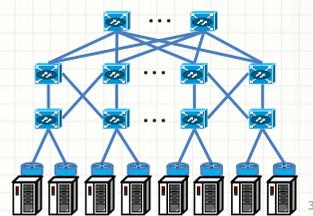
↓

- ネットワーク構造が処理性能に直結



データセンターネットワークにおける課題

- ネットワーク性能の確保と低消費電力の両立
 - データセンターの処理性能の確保
 - 帯域の不足や通信遅延がサーバーの連携へ影響
 - 従来型データセンターネットワークは冗長に機器を配置し、需要の変化に対応
 - データセンターネットワークの消費電力の削減
 - 低消費電力のためには使用機器数を削減することが必要



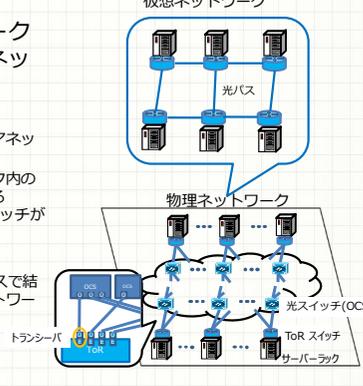
研究目的

- ネットワーク性能の確保と低消費電力の両立が可能なデータセンターネットワークの構築方法の確立
 - ネットワーク性能の確保
 - ネットワーク性能の要求に応じてネットワークを動的に構築
 - 制約条件
 - サーバーラック間に発生するトラフィックを輻輳なく収容可能
 - サーバーラック間の通信遅延を目標値以下に抑える
 - 低消費電力化
 - 電気スイッチの使用ポート数の最小化

光ネットワークを用いたデータセンターネットワーク

仮想ネットワーク

- 動的に仮想ネットワークの構造を変更可能なネットワーク
 - 物理ネットワーク
 - 光スイッチを用いたコアネットワークを構築
 - 各光スイッチにはラック内のサーバーと接続しているToR(Top of Rack)スイッチが接続
 - 仮想ネットワーク
 - ToRスイッチ間を光バスで結ぶことにより仮想ネットワークを構築

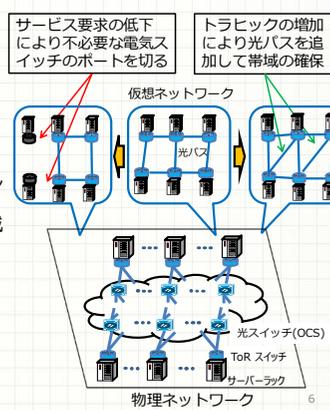


仮想ネットワーク制御による省電力化

- 光スイッチの消費電力は電気スイッチより著しく少ない
 - サービス要求の低下により不必要な電気スイッチのポートを切る
 - トラフィックの増加により光バスを追加して帯域の確保
- 電力削減の方針
 - 仮想ネットワークに使用しないToRスイッチのポートをOFFにして消費電力削減

↓

- 環境変動に応じて、必要な電気ポートの数が少ない仮想ネットワークを構成



仮想ネットワーク制御手法

- 従来型仮想ネットワーク制御の適用は困難
 - 大規模なデータセンターネットワークでは計算時間が膨大
 - データセンターの頻繁に発生するトラフィック変動に対応が困難
- データセンターネットワークに適した仮想ネットワーク制御
 - パラメーターの調整によって経路フロー数、最大ホップ数、リンク数が調整できるネットワーク構造
 - 短い計算時間で適切なネットワーク構造に移行可能
 - 既存のデータセンターネットワーク構造をベースにネットワーク性能と必要リンク数の調節をより細やかな粒度で行うことができる構造を提案
 - 性能要件を満たすための構築パラメーター決定
 - 性能要件を満たすネットワーク構造を少ないリンク数で構築する構築パラメーターを決定

GFB (Generalized Flattened Butterfly)

- 構築パラメーターの調整によって、様々なネットワーク構造が構築可能

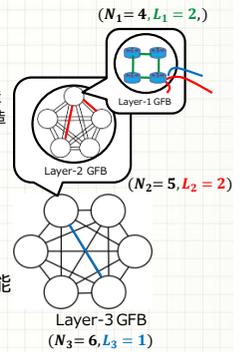
- 構築パラメーターの調整によってさまざまなネットワークを構築可能
 - Flattened Butterfly のような接続が密な構造
 - Switch-based DCell のような接続が疎な構造

- 構築パラメーター

- k : 階層数
- N_k : 階層 k で接続する $(k-1)$ 層の GFB の数
- L_k : 階層 k で利用する各 ToR スイッチあたりのリンク数

- 構築パラメーターから以下が計算可能

- ToR スイッチ間の最大ホップ数
- 各リンクのフロー数

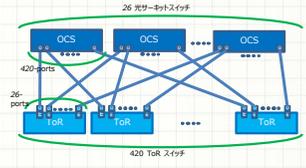


性能要件を満たすための構築パラメーター決定手法

- 最大ホップ数の目標値からそれを満たすことが可能な階層数の候補を決定
 - 階層数によって達成可能な ToR スイッチ間の最大ホップ数の最小値は以下の式で計算可能
$$H_k = 2^k - 1$$
- 各階層数の候補に対して以下を満たす構築パラメーターを求める
 - ToR スイッチ間の最大ホップ数が目標値以下
 - ToR スイッチからのトラフィックを輻輳なく収容
- 各階層数の候補の中から使用リンク数が最小の構築パラメーターを選択

使用リンク数の評価

- 評価環境
 - ToR スイッチの台数: 420
 - 各リンクの帯域: 10Gbps
 - 目標最大リンク使用率: 0.6
 - ToR スイッチから送信されるトラフィック量は最大 6Gbps
 - ToR スイッチ間のトラフィック量は VLB (Valiant Load Balancing) [1] を用いて負荷分散
- 比較対象
 - 固定のネットワーク構造
 - トラフィック量の最大値を流すことができる GFB を固定のトポロジーとして与える
 - 固定のネットワーク構造からトラフィックにあわせて、リンクを削減する手法 [2] を用いた場合

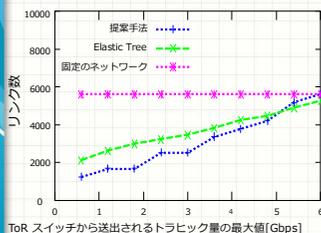


[1] M. Kodialam, T. Lakshman, and S. Sengupta, "Efficient and robust routing of highly variable traffic," in Proceedings of NetNets, Nov. 2004.

[2] B. Heller, S. Satharman, P. Mahadevan, Y. Yakoumis, P. Sharma, S. Banerjee, and N. McKeown, "ElasticTree: saving energy in data center networks," in Proceedings of USENIX, pp. 17-32, Apr. 2010.

リンク数の評価結果

- 提案手法はトラフィック量が少なくなるにつれリンク数が削減可能
 - トラフィック量にあわせたネットワークを構築しているため
- 固定のネットワークからリンクを削減する手法よりも低消費電力
 - 固定のネットワークにおいて機器の電源を落とすだけでは不十分



まとめと今後の課題

- データセンターにおける省電力な仮想ネットワーク構築手法を提案
- トラフィックを収容するために必要となるリンク数を評価
 - 固定のネットワーク構造より多くのリンクを削減
 - 固定のネットワークからリンクを削減するよりもリンクを削減
- 今後の課題
 - 自律分散的な仮想ネットワーク制御