

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

Virtual Network Topology Control to Achieve Low Energy Consumption in an Optical Data Center Network

データセンターネットワークにおける低消費電力のための仮想ネットワーク制御手法

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 村田研究室
樽谷 優弥

Osaka University

データセンターネットワークにおける課題と解決策

- ネットワーク性能の確保と低消費電力の両立
 - ネットワーク性能がデータセンターの処理性能に大きな影響
 - 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
 - データセンター全体の消費電力のうちネットワークの消費電力が占める割合は大
 - 低消費電力のためには機器数を削減することが必要
- 解決策
 - 需要に応じて低消費電力なネットワークトポロジーを動的に構築

Osaka University

光ネットワークを用いたデータセンターネットワーク

- 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構築
 - 光バスの張り替えにより、要求に必要な性能を満たすトポロジーを構築
 - 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

Osaka University

低消費電力化のための仮想ネットワーク制御

- 目標：ToRスイッチの使用ポート数の最小化
- 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

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

本研究の目的
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手法の確立

Osaka University

研究方針

- 少数の構築パラメーターの計算による仮想ネットワークの制御
 - 少数の構築パラメーターの調整により構築可能なトポロジー
 - 少数の構築パラメーターの調整により性能要件に対して適切なトポロジーを計算可能
 - 大規模データセンターネットワークにおいても短い計算時間で適切なトポロジーに移行可能
 - 性能要件を満たすための構築パラメーター決定手法
 - 輻輳を防止する構築パラメーター決定手法
 - 頻繁に発生するトラフィック変動は負荷分散により対応
 - 負荷分散を考慮に入れた構築パラメーターを設定することにより、トラフィックを収容
 - サーバー間の通信遅延を抑えるための構築パラメーター決定手法
 - ホップ数・リンク負荷を一定以下に抑えることにより遅延の増大を防ぐ

Osaka University

GFB (Generalized Flattened Butterfly)

- Flattened Butterfly [1] を一般化した階層型トポロジー
 - 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
 - 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
- 以下の値を構築パラメーターから計算可能
 - 最大ホップ数 $H_k = (h_k + 1)H_{k-1} + h_k$
 - 各リンクを経由するToRスイッチ間フロー数 $X_k = \frac{\sum_{i=1}^{h_k} i s_i(i)}{L_k \prod_{i=1}^k N_i}$

[1] I. Kim, W. Dally, and D. Abts, "Flattened butterfly: a cost-efficient topology for high-radix networks," in Proceedings of ACM SIGARCH Computer Architecture News, vol. 35, pp. 126-137, June 2007.

Osaka University 7

輻輳を防止する構築パラメーター決定手法

- 短時間のトラフィック変動に対応するための負荷分散手法
 - VLB (Valiant Load Balancing) [2]
 - 宛先によらずランダムに経路する中継先を選択することにより負荷分散
 - 各ToRスイッチ間を流れるトラフィック量は以下で与えられる

$T_{VLB} = (V_{out} + V_{in}) / E$

E : ToRスイッチ数
 V_{out} : ToRスイッチが送信するトラフィック量
 V_{in} : ToRスイッチが受信するトラフィック量
 T_{VLB} : VLBを用いた場合における各ToRスイッチ間のトラフィック量
 X : 各リンクを経由するフロー数
 B : リンク帯域

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

- GFBの構築パラメーターの決定
 - ToRスイッチ間のフロー数が以下を満たすGFBの構築パラメーターを計算
$$X \leq B / T_{VLB}$$

Osaka University 8

サーバー間の通信遅延を抑えるための構築パラメーター決定手法

- 手法の方針
 - ホップ数、リンク負荷の両方を抑えることにより、遅延の増大を防止
- GFBの構築パラメーター決定方法
 - リンク負荷の抑制
 - 輻輳を防止する構築パラメーター決定手法により達成
 - ホップ数の抑制
 - 最大ホップ数が目標値以下となる構築パラメーターを計算

Osaka University 9

評価結果

- 既存のトポロジー
 - 既存のトポロジーの一部はリンク数を増やしても多くのトラフィックを収容不可
 - Flattened Butterflyでは収容するために多くのリンクが必要
- 提案手法で構築した仮想ネットワークトポロジー
 - 既存のトポロジーの半分のリンク数でトラフィックを収容可能
 - トポロジー全体ではなく階層毎に構築パラメーターを設定可能
 - 必要な箇所へのみリンクを追加するため少なくとも抑えることが可能

評価環境

- ToRスイッチ数: 420
- リンクの帯域: 10Gbps
- 各トポロジー構造の構築パラメーターはリンク数が最小となるように設定

Osaka University 10

評価結果

- 既存のトポロジー
 - 既存のトポロジーの一部はリンク数を増やしても最大ホップ数を小さくすることが不可
 - Flattened Butterflyは許容可能な最大ホップ数が大きくても必要なリンク数を下げることが不可
- 提案手法で構築した仮想ネットワークトポロジー
 - 許容可能なホップ数の大きさに関わらずリンク数を抑制可能
 - 既存のトポロジーにおける許容可能なホップ数に対して必要なリンク数が少ない箇所に追随

評価環境

- ToRスイッチ数: 420
- リンクの帯域: 10Gbps
- 各トポロジー構造の構築パラメーターはリンク数が最小となるように設定

Osaka University 11

まとめと今後の課題

- 低消費電力と目標性能の達成を両立する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手法を提案
- 既存のトポロジー構造と比較、評価
 - 既存のトポロジー構造よりも少ないポートで多くのトラフィックを収容可能
 - ノードの追加、削除を行った場合も目標性能を維持可能
- 今後の課題
 - 少ないリンクの張り替えによる環境変動への対応
 - ネットワークの構成による影響を考慮した制御タイミングの設定