Optimizing Energy Efficiency for Bulk Transfer Networks

[Extended Abstract]

Matthias Herlich
University of Paderborn
Pohlweg 47-49
33098 PADERBORN
GERMANY
herlich@uni-paderborn.de

Holger Karl
University of Paderborn
Pohlweg 47-49
33098 PADERBORN
GERMANY
holger.karl@uni-paderborn.de

ABSTRACT

Usually Network traffic is modeled using flows of data, but flows are an abstraction for aggregated data transfers. By analysing the transfers themselves, a more detailed understanding can be created.

For example, content delivery networks have to transfer large amounts of data with relatively few transfers. The available data rate will affect the time the transfer takes. The energy needed to transfer the data may be more important than the time needed, but there will usually be a trade-off between time and energy.

The formalization of the problem is a graph with energy consumption and capacity given for each edge as well as a list of the requested transfers.

We solve the problem of a network transferring predefined amounts of data with a minimal amount of energy using mixed integer programming.

The model and solution technique allow to calculate the most energy-efficient routing and can further be applied to study how minimal time routing and minimal energy routing relate to each other.

Categories and Subject Descriptors

C.4 [**Computer Systems Organization**]: Performance of Systems

General Terms

Performance

1. INTRODUCTION

When modeling a computer network the data transfers are often aggregated into flows. Flows are a good abstraction for data streams and aggregations of many small transfers, but for large transfers they fail to model the dependence of supplied data rate to future requests. Hence, flows change over time and these changes have to be modeled, too. The model of bulk transfers explains how the requested flows change.

A bulk transfer is a request to transmit a predefined amount of data from a source to a destination. When the amount of data has been transmitted, the bulk transfer is over. Hence, doubling the data rate of a bulk transfer will cut the time needed in half. This implies that the transfers a network has to handle depend on how it handled transfers in the past.

Examples for networks that handle bulk transfers are content delivery networks [4], which have to update their remote servers, or content-centric networks [2, 3], which can handle requests for blocks of data.

We explain how bulk transfers can be modeled and show how the most efficient way of executing the bulk transfers can be calculated.

2. MODEL

We define a *bulk transfer* to be the transmission of a predefined amount of data from a single source to a single destination. In contrast to bulk transfers, a *flow* specifies a required data rate, which will be used for an unknown amount of time. Today a bulk transfer is finished ASAP, but more energy-efficient methods exist.

We assume the underlying hardware can be disabled, to conserve energy. As disabling and enabling the hardware at arbitrary points in time makes the solution for minimal energy consumption infeasible, we define *slices* of time in which the assigned data rates are constant. A single bulk transfer can use multiple slices and can have a different data rate in each slice, but will eventually have transmitted all its data after a finite number of slices.

Given an undirected graph $G = (V, E)$ with a set of vertexes V and set of edges E, we define two additional parameters for each edge: energy consumption $\varepsilon : E \to \mathbb{Q}$ and capacity $c : E \to \mathbb{Q}$. The *capacity* $c(e)$ of an edge e is the maximum data rate it can transfer. The *energy consumption* $\varepsilon(e)$ is the amount of energy used when the edge e is active. The energy consumption does not depend on the amount of data transferred in a slice, but only whether data was transferred or not. This definition reflects that a line card consumes nearly the same amount of energy when it is idle as when it is actively transmitting, as current network hardware lacks sleep modes. The only way it can conserve energy is by being disabled completely [1].

As load model we assume a set B of bulk transfers. Each of these bulk transfers $b \in B$ has a source s_b, a destination d_b, and an amount of data a_b to transfer.

For a given set of slices T needed for all bulk transfers, $f_b^t : E \to \mathbb{Q}$ is a function that specifies how much data of bulk transfer $b \in B$ is sent over the given edge in slice $t \in T$.

To calculate the total energy consumption for all bulk transfers, we introduce a helper function $x^t : E \to \{0, 1\}$ that expresses whether an edge e was used in a time slice t or not:

$$x^t(e) = \begin{cases} 1, & \text{if } \forall b \in B : f_b^t(e) = 0 \\ 0, & \text{else} \end{cases}$$

A linear constraint for the helper function x^t can be created by defining each $x^t(e)$ as a binary variable and adding the linear constraints

$$\forall t \in T, b \in B, e \in E : f_b^t(e) \leq x^t(e)c(e),$$

where $c(e)$ is a constant.

It is necessary that flows in each slice obey the usual constraints for a flow network: capacity constraint, skew symmetry and flow conservation. The capacity constraint

$$\forall t \in T, e \in E : \sum_{b \in B} \max(0, f_b^t(e)) \leq c(e)$$

guarantees that the maximum data rate of no edge is exceeded (assuming a full duplex transmission). The skew symmetry

$$\forall b \in B, t \in T, (u, v) \in E : f_b^t((u, v)) = -f_b^t((v, u))$$

guarantees a consistent model. And the flow conservation

$$\forall b \in B, t \in T, u \in V \setminus \{s_b, d_b\} : \sum_{v \in N(u)} f_b^t((u, v)) = 0$$

makes sure that each bulk transfer $b \in B$ only starts or ends in its source s_b or destination d_b. $N(v)$ denotes the set of vertexes incident to vertex v.

In addition, we need a constraint which guarantees that each bulk transfer has finished, i.e. all its data was transferred. We only need to guarantee that the data leaves the source s_b, as the flow conservation implies that it must end up at the destination d_b. Hence the transfer constraint can be written as "all data a_b leaves the source s_b"

$$\forall b \in B : \sum_{t \in T} \sum_{v \in N(s_b)} f_b^t((s_b, v)) = a_b$$

or "all data a_b reaches the destination d_b"

$$\forall b \in B : \sum_{t \in T} \sum_{v \in N(d_b)} f_b^t((v, d_b)) = a_b$$

The total energy consumption for all bulk transfers is given by

$$\sum_{t \in T} \sum_{e \in E} x^t(e)\varepsilon(e).$$

With the given constraints and the function for energy consumption, mixed integer programming can be used to calculate the most energy-efficient routing.

The routing that uses the least number of time slices of all solutions that use the minimal amount of energy can be calculated by adding a term that describes the number of slices in which data was transferred divided by a sufficiently large number. The number of slices in which data was transferred can be calculated as a sum of helper function analogous to x^t.

3. CONCLUSION

We introduced a new model for network analysis that focuses on bulk transfers instead of flows, but is able to combine both of them. In addition we explained the idea of slices to make the calculation of a solution feasible using mixed integer programming. Given the energy consumption and the capacity of each edge, we can calculate energy optimal routing schemes for requested bulk transfers.

4. REFERENCES

[1] M. Gupta and S. Singh. Greening of the Internet. In *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, page 26. ACM, 2003.

[2] V. Jacobson, M. Mosko, D. Smetters, and J. Garcia-Luna-Aceves. Content-centric networking. *Whitepaper, Palo Alto Research Center*, 2007.

[3] V. Jacobson, D. Smetters, J. Thornton, M. Plass, N. Briggs, and R. Braynard. Networking named content. In *Proceedings of the 5th international conference on Emerging networking experiments and technologies*, pages 1–12. ACM, 2009.

[4] G. Pallis and A. Vakali. Insight and perspectives for content delivery networks. *Communications of the ACM*, 49(1):106, 2006.

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Computer Networks Group

Optimizing Energy Efficiency in Bulk Transfer Networks

Matthias Herlich, Holger Karl

Problem Statement

Given a network and a set of bulk transfers (instead of flows),

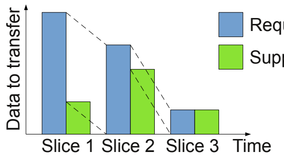
what is the most energy-efficient route for the bulk transfers?

Background

Comparing Bulk Transfers to Flows

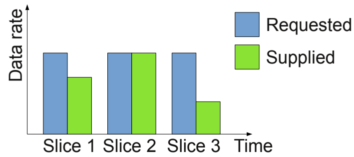
Bulk transfers

- Model for large amounts of data (e.g. CDN)
- Specify amount of data to transmit
- Done after predefined amount of data has been transmitted



Flows

- Model for streams or aggregation of many small transfers
- Specify required data rate
- Continue for an unknown time



Slices of time

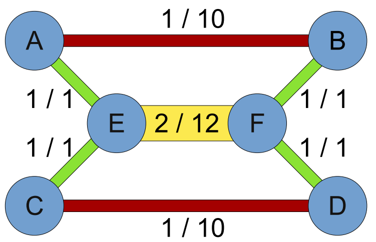
- Needed to optimize energy efficiency using mixed integer programming
- Constant data rate in a single slice
- A bulk transfer can take several slices to finish

Assumption

Network hardware can be disabled to conserve energy

Model – Definition of the Network and Bulk Transfers

Network Graph



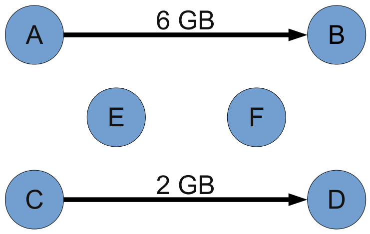
Bandwidth (width) [GB/slice] / Energy Consumption (color) [J/slice]

Undirected graph
$G=(V,E)$

Capacity of edges
$c: E \rightarrow \mathbb{Q}$

Energy consumption of edges
$\varepsilon: E \rightarrow \mathbb{Q}$

Requested Bulk Transfers

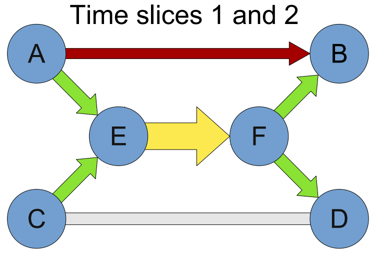


Set of bulk transfers: B

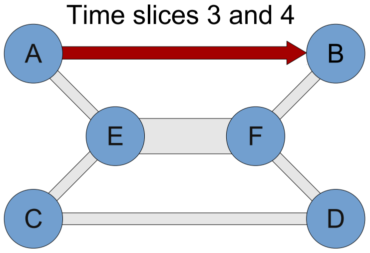
$\forall b \in B$: Source $s_b \in V$; Destination $d_b \in V$; Transfer size $a_b \in \mathbb{Q}$

Solution – Finding Routes with Optimal Energy Efficiency using MIP

Time slices 1 and 2



Time slices 3 and 4



Optimal solution needs multi-path routing!

Minimize $\sum_{t \in T} \sum_{e \in E} x^t(e)\varepsilon(e)$ (total energy consumption)

Where $x^t(e) = \begin{cases} 0, & \text{if } \forall b \in B: f_b^t(e)=0 \\ 1 & \text{else} \end{cases}$ (edge activity)

T set of needed slices

$f_b^t: E \rightarrow \mathbb{Q}$ (flow transmitted for $b \in B, t \in T$)

With flow constraints:

$\forall t \in T, e \in E: \sum_{b \in B} max(0, f_b^t(e)) \leq c(e)$ (capacity constraint, full duplex)

$\forall b \in B, t \in T, (u,v) \in E: f_b^t((u,v)) = -f_b^t((v,u))$ (skew symmetry)

$\forall b \in B, t \in T, u \in V \setminus \{s_b, d_b\}: \sum_{v \in N(u)} f_b^t((u,v)) = 0$ (flow conservation)

$\forall b \in B: \sum_{t \in T} \sum_{v \in N(s_b)} f_b^t((s_b,v)) = a_b$ (complete transfer)

(with $N(v)$ being the set of vertices incident to v)

Contact person

Matthias Herlich
+49 5251 60–5382
herlich@upb.de

University of Paderborn
Pohlweg 47–49
33098 PADERBORN
GERMANY
http://wwwcs.upb.de/cs/cn/

Head of group

Prof. Dr. Holger Karl
+49 5251 60–5375
holger.karl@upb.de