ON BANDWIDTH ON DEMAND PROBLEM

E.P. Stepanov[1,a], R.L. Smeliansky[1,b]

[1] *Lomonosov Moscow State University, Moscow, GSP-1, 1-52, Leninskiye Gory, 119991, Russia*

E-mail: [a] estepanov@lvk.cs.msu.ru, [b] smel@cs.msu.ru

The paper discusses approaches to organizing the service "Bandwidth on Demand" for data transfer between Data Centers (DC). In this paper we present an analysis of the routes' aggregation protocols, consider routing that uses aggregation and construct a mathematical model of the data flow distribution among generated routes. This model makes possible to determine the existence of a flow load distribution for the case of a single flow and some particular cases of the solution absence in the multiflow case.

Keywords: Bandwidth on Demand, Multipath Protocols, Quality of Service

1. Introduction

Recently, there has been a trend in the growth of backbone traffic between Data Centers (DCs). According to TeleGeography [1], by the end of 2017, the share of such traffic on the most popular route across the Atlantic Ocean had reached 75%, and in 2023 it should exceed 93%. This can be explained by the development of the global cloud services market, which is currently concentrated in North America and Europe. Therefore, the growth of traffic between DCs is provided mainly by the data centers of cloud providers and corporate data centers that use hybrid clouds.

However, cloud data centers impose special requirements on channel bandwidth allocation and charging policies. The most promising approach to meet these requirements is to provide channel bandwidth according to the "Pay as you go" model - only when there is a need for it, i.e. bandwidth on demand (BoD).

The paper discusses approaches to the implementation of BoD services in the network of the Internet provider. We assume that from a certain data center it is necessary to transfer a certain amount of data to another specific data center in a time not exceeding a certain value. We will also assume that the involved data centers are connected to each other through a transport network.

The implementation of BoD services can be divided into two components: route aggregation and fair distribution of client traffic flows among these routes. Aggregation must be carried out, since the free resources of each individual route may not be enough to meet the needs of users. Route aggregation can be carried out using protocols of different levels, which will be discussed in section 2. In section 3 we present a mathematical model that allows you to determine the existence of a solution to the problem of one flow fair distribution, and whether it is unique or not. Then, the mathematical formulation for the case of multiple flows is considered, and options for solving it are proposed. Section 4 presents the conclusion and directions of future work on the problem considered in this paper.

2. Route aggregation

By route aggregation we mean a service that allows you to transfer data between the same pair of DCs, using several different routes at the same time. By route we mean a sequence of physical links in a transport network that does not contain cycles and connects the data center to each other. The task of route aggregation does not address the issue of ensuring the necessary quality of service, how BoD does. To implement route aggregation, several problems must be solved.

First you need to determine the number of routes whose resources can be combined. To solve this problem, it is necessary to determine what properties routes between entry and exit points should have. We assume that the main requirement for routes is the absence of intersections on physical links for the routes of one client. This constraint is related to the operation of congestion control algorithms. If the flow routes have an intersection in a bottleneck, then at the moment of congestion a synchronization effect may occur, that will lead to a simultaneous substantial decrease in the flow rate. The number of disjoint routes between two points can be determined using the Menger theorem [2], which states that the largest number of edge-disjoint routes from vertex $\boldsymbol{u}$ to vertex $\boldsymbol{v}$ is equal to the smallest number of edges in the $< \boldsymbol{u}, \boldsymbol{v} >$ cut.

It is worth to note that the requirement that there is no intersection on routes physical links may be redundant when the intersection of routes has sufficient available bandwidth, i.e. there will be no network bottleneck. In addition, there may be a situation, when in the network topology there are no alternative disjoint routes between entry and exit points. In this case, we can reduce the problem to the previous one if we transform the graph of the network topology in such way that the edge corresponding to the physical link with high throughput is replaced by several edges between the same vertices with lower throughput. An alternative solution could also be to search for routes with the least number of intersections, as MCMF [3] does.

After determining $\boldsymbol{k}$ - the number of disjoint routes, we can apply algorithms for generating $\boldsymbol{k}$ routes between two vertices in the network topology graph. However, not every algorithm for $\boldsymbol{k}$ disjoint routes generation can provide a solution. For example, the greedy algorithm [4], which at each step generates a new shortest route along the unused edges of the graph, cannot always construct all $\boldsymbol{k}$ routes. This is because disjoint routes are not required to be the shortest in the graph. Therefore, it is better to generate all disjoint routes at the same time, for example, by reducing the problem to finding the maximum flow in the network, as MCMF [3] does.

Thus, we have generated a set of routes containing free resources that can be used to provide the BoD service. The next problem is to find the approach to simultaneously use the resources of these routes for user traffic. There are a large number of protocols and technologies that can help in this problem. The choice may be based on the following parameters, depending on the capabilities and priorities of the BoD service provider:

- *scalability* - the provided service should be able to adapt to a different number of routes between entry and exit points. For example, in optical networks, 80 channels with different wavelengths can be used simultaneously for transmission;
- *granularity* – transport flows as well as packets of one transport flow can be distributed among routes. Each of the options has its own advantages and disadvantages. For example, in the case of balancing different transport flows, the rate of one transport flow will be limited by resources of the route along which it will be transmitted. In the case of balancing packets of one transport flow, it is necessary to take into account the work of congestion control algorithms, since different delay on routes can lead to erroneous congestion recognition;
- *adaptability* - dynamically change the number of used routes. The reasons to change the number routes may be different. For example, changes in the topology or in the physical links load, so it is necessary to use other routes;
- *resource allocation delay* - how much time it takes to pre-configure network devices until the client starts to use the network;
- *fault tolerance* - in case of some route failure, it should be possible to switch to a backup route. To provide fault tolerance, failure detection mechanisms, automatic switching and maintaining some reserve are important;
- *guarantee* - there is a reserve of resources for each client or routes are shared by several clients, which, with an increase in load / number of clients can lead to insufficient bandwidth (in this case an additional route should be involved).

All protocols for route aggregation can be divided based on the TCP/IP model layers. We consider the protocols of the TCP/IP model from top to bottom and start with the transport layer, since the data transport is not considered at the application level. The most suitable protocols at the transport level are multipath protocols, which allow you to distribute a single ordered stream of application bytes among several transport subflows, thus providing packet balancing.

There are two approaches to multipath routing: static and dynamic. The MPTCP static approach [5] involves a priori allocation of a certain number of transport subflows among which data stream segments are distributed. The dynamic approach of FDMP [6] involves the dynamic allocation of a subflow at the request of a transport agent, depending on the correspondence of the total allocated subflows throughput to the application demand. It is worth noting that in a traditional TCP/IP network for multipath protocols, the term transport subflow is used instead of the term route. This is due to the fact that the multipath protocol itself does not guarantee that transport subflows will pass through the network using different routes, this problem is assigned to the network. However, through the use of own congestion control algorithms and a packet distribution module, it becomes possible to balance packets between transport subflows, which cannot be optimally done within the network.

At the network level, transport flow balancing techniques such as ECMP [7], MPLS-TE [8] together with the RSVP resource reservation protocol [9] can be applied. In this case, it is worth paying special attention to the capabilities of distributing flows amount routes, since large "elephant flows" and small "mouse flows" can be transmitted simultaneously. Simple hashing of the packet header part in this case will not work, since it does not reflect the flow load. It is also necessary to pay attention to the routes used in balancing. For example, ECMP routes should be with equal cost value (for example, the shortest). These constraints may not allow you to use the routes defined in the first step. Therefore, to balance flows, it is more profitable to look towards UCMP, where routes can be with varying cost.

In the case of the link layer, the protocols may vary depending on the physical environment of the used network. However, the main constraint for all link layer aggregation protocols is to use only those routes that pass through the same network devices. Most network equipment for working with Ethernet networks supports both static configuration of link aggregation and dynamic control using the LACP, PAgP protocols [10]. Just as in balancing at the network level, the question about flows' distribution arises.

Similar to LACP channel aggregation techniques can be found for other types of networks, although they can have a completely different physical basis. So, LTE-Unlicensed wireless networks can simultaneously transmit using several WiFi channels [11]. Channel Bonding technology is also described in the 802.11 standard [11]. In the case of OTN networks, the VCAT inverse demultiplexing technique can be used in conjunction with the method of dynamically changing the link capacity LCAS [12], which will allow you to distribute the required throughput across multiple routes.

Thus, the choice of protocols for route aggregation depends on the physical environment of data transmission, the capabilities of network equipment, and the priorities of the BoD service provider. Sharing protocols at different levels is a good solution, since none of them individually will be able to use all available routes for simultaneous data transfer.

3. Flow distribution

After the route aggregation we should consider the problem of client flow load distribution among individual routes. The first question that interests us is whether there are enough free resources to meet the needs of a customer of BoD services. If there are enough resources, then how should they be distributed?

To solve this problem, we present a mathematical model of the simultaneous data transmission across several routes for the case of a single flow, where under the term flow we mean all the traffic of one client from the entry point to the exit point of transport network. The input data of the mathematical model are the following:

- $G = (V, E) -$ directed graph without cycles
- $\mathrm{C} = \{\mathrm{c}_{ij}\} -$ set of available bandwidth for each arc $\mathrm{e}_{ij} \in E$. The available throughput for the arc can be estimated based on the load statistics of the corresponding physical link, and in addition prediction methods can be used to estimate the available throughput in the near future.
- $R -$ required throughput for the flow
- $P = \{P_k\} -$ set of disjoint routes that do not contain cycles

Based on input data we can define the following quantities:

- $\Delta_j = \min_{e_{km} \in P_j} c_{ij}$ – available bandwidth on the route P_j
- $\Delta = \sum_j \Delta_j$ – available bandwidth on all routes

The knowledge of Δ allows you to answer the question - is there enough free network resources to provide the required bandwidth R. Conditions under which it is possible to distribute the load the a flow among free network resources:

$$\Delta \geq R$$

It is worth noting that in case of $\Delta = R$ there is only one way to distribute the load, in which all the free bandwidth of the routes will be occupied. If $\Delta > R$, then for real values of the subflow load (by the subflow we mean the part of the initial flow that follows its own route) there will be infinitely many options for the flow load distribution among the given routes. One of the distributions can be generated with the help of progressive filling algorithm, which is used to generate a max-min fair distribution [13]. Then the flow load will be evenly distributed among the routes specified in the first step.

In the case when we have more than one client, the question arises whether it is possible to simultaneously satisfy the requests of all customers or not. Even if part of the network free resources is already distributed, and we know that the flows of all clients can be accommodated, then it is possible to reallocate the resources in such a way that the requests of all clients are satisfied. For this case, the input data of the model should take into account all clients, therefore we include in input data a set of flows $\{f_i\}$ (entry and exit points), their bandwidth requirements $\{R_i\}$ and routes $\{P_{ik}\}$. Based on the given input data, we can determine some constraints under which there is definitely no solution to the load distribution problem:

$$\exists f_i\colon \Delta_{f_i} < R_i$$

A complete answer to the question of a solution existence can be obtained, for example, by reducing the described problem to a linear programming problem, however, the properties and limitations of this approach are the subject of future research.

4. Conclusion

In this paper we considered the problem of providing the BoD service, which was divided into the route aggregation problem and the flow load distribution problem. Various implementation options for the route aggregation problem were analysed according to the network parameters, the desires of the service provider and the capabilities of the network equipment. In the case of flow load distribution problem we presented mathematical model, that allows to determine the existence of a solution to this problem.

A further area of research is the study of methods for solving the problem of multiflow load distribution. There is also the question of maintaining a free resources reserve to meet the needs of new customers and maintaining the necessary quality of service for current customers in the event of a change in flows load.

5. Acknowledgement

This work is supported by Russian Ministry of Science and Higher Education, grant #05.613.21.0088, unique ID RFMEFI61318X0088.

References

[1] Telecommunications market research that's data-driven, TeleGeography. Retrieved from https://www.telegeography.com/

[2] Böhme, Thomas, Frank Göring, Jochen Harant. "Menger's theorem." *Journal of Graph Theory* 37.1 (2001): 35-36.

[3] Stepanov, E., R. Smeliansky. "On Analysis of Traffic Flow Demultiplexing Effectiveness" *2018 International Scientific and Technical Conference Modern Computer Network Technologies (MoNeTeC)*. IEEE, 2018.

[4] Navin Kukreja, Guido Maier, Rodolfo Alvizu, Achille Pattavina. SDN based automated testbed for evaluating multipath TCP. In IEEE International Conference on Communication, ICC 2015, London, United Kingdom, June 8-12, 2015, Workshop Proceedings, pages 718–723, 2016.

[5] Costin Raiciu, Christoph Paasch, Sebastien Barre, Alan Ford, Michio Honda, Fabien Duchene, Olivier Bonaventure, Mark Handley. How hard can it be? designing and implementing a deployable multipath tcp. In Presented as part of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI 12), pages 399–412, San Jose, CA, 2012. USENIX.

[6] Evgeny Chemeritskiy, Evgeny Stepanov, Ruslan Smeliansky. Managing network resources with flow (de) multiplexing protocol. MATHEMATICAL and COMPUTATIONAL METHODS in ELECTRICAL ENGINEERING, 53:35–43, 2015.

[7] Chiesa, Marco, Guy Kindler, Michael Schapira. "Traffic engineering with equal-cost-multipath: An algorithmic perspective." *IEEE/ACM Transactions on Networking (TON)* 25.2 (2017): 779-792.

[8] D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell, J. McManus, Requirements for Traffic Engineering Over MPLS, RFC 2702, Sep. 1999.

[9] Awduche, Daniel, et al. "RSVP-TE: extensions to RSVP for LSP tunnels." (2001).

[10] Irawati, Indrarini Diah, Sugondo Hadiyoso, Yuli Sun Hariyani. "Link Aggregation Control Protocol on Software Defined Network." *International Journal of Electrical and Computer Engineering* 7.5 (2017): 2706.

[11] Dilmore, Michael, Angela Doufexi, George Oikonomou. "Analysing Interface Bonding in 5G WLANs." *2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*. IEEE, 2018.

[12] Bernstein, Greg M. "IP Bandwidth on Demand and Traffic Engineering via Multi-Layer Transport Networks." *2006 IEEE First International Workshop on Bandwidth on Demand*. IEEE, 2006.

[13] Bertsekas, Dimitri P., Robert G. Gallager, Pierre Humblet. *Data networks*. Vol. 2. New Jersey: Prentice-Hall International, 1992.