Volume 3, Issue 11, November 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Proposed Design for Tolerant Bulk Data Transfer on Internet

Varsha Sagare
M.E (CSE) BSCOER,
Pune University, Pune, India

S.M Chaware
BSCOER,Pune
India

Abstract—In communication network there has been renewed interest in the problem of transferring bulk data (terabyte) use commercial ISPs. The key insight underlying the recent work was that many of applications are delay-tolerant and hence the bulk data can be transferred at minimal cost, from diurnal traffic patterns, using store and forward through intermediate storage nodes.In this paper proposed new concept on store and forward method which used to be minimized load on data centre and also reduce time to transfer data on the Internet. The general problem of finding a cost-optimal transfer of the bulk data using minimum cost flow on network. An important characteristic of this solution approach is the ability to handle nodes with storage. To consider nodes with storage that varies over time in terms of capacity and cost. Here use 95 percentile pricing schema for bandwidth scheduling show that there exists a huge potential for the cost savings in real-world networks with time-varying costs for both link capacities and node storage & also new concept to be minimized load on data centre & ISP required minimum time & memory to send data .

Keywords— Bandwidth Scheduling, ISP, E2E Network, 95 percentile pricing schema, Store-and-Forward Networks.

I. INTRODUCTION

Internet has been a communication network, its development has been driven by the assumption that connections and data transfers are sensitive to delay. In recent years however, the network has progressively shifting from communication to content dissemination. Unlike communication, content dissemination can often tolerate much larger delays, e.g. in the order of hours. This higher tolerance to delay allows scheduling to go beyond congestion avoidance. Here, briefly illustrate how to use store-and-forward scheduling & E2E scheduling to perform bulk data transfer that may be impossible or, under current pricing schemes for bandwidth, prohibitively expensive.

The growth of social networking has also crucially changed how to interact, share and consume information. The observation that a vast amount of multimedia content downloaded is not consumed right away, and so, is delay-tolerant (DT) has opened the possibility of offering bulk downloads as a service that the ISPs can offer.[1],[2] This has meant that ISPs have had to rethink their networks beyond merely routing and forwarding packets. ISPs can enable a variety of services for a range of applications that take advantage of bulk-data transfers, both for consumers and for businesses.

There are several applications for such a DT service. The popularity of services like Netflix has meant that as a next generation service, movies may be available for download from the Netflix queue to an Xbox [17] or a similar device rather than via snail mail. The enormous growth of social networking sites has meant that companies need to synchronize their data warehouses both within the country and across continents. News organizations need to move multimedia content to their web-servers across the country so that data may be accessed from the closest server.

II. LITERATURE SURVEY

Following concept are required to design tolerant bulk data :-

A. Bulk Data

Residential and corporate bulk data have fuelled an unprecedented increase in overall Internet traffic over the last few years. On the end-user side, these include high-definition movies from commercial Web sites or peer-to-peer (P2P) networks, large-scale software updates and remote backups.[1],[2] Adding to this, data centres hosting cloud computing applications exchange large amounts of synchronization, accounting and data-mining traffic, while large corporate and government organizations contribute increasing numbers of economic, engineering and scientific datasets.

B. Bulk Bottlenecks under Flat-Rate Pricing

Several ISPs throttle P2P traffic from flat-rate residential customers during peak hours in order to free capacity for interactive traffic that is valued more highly by most users.[6],[8] Popular P2P applications like Bit Torrent tend to be targeted, but in the future it may be any other bulk application that becomes popular among flat-rate residential customers. [4],[6]Such throttling bottlenecks introduced by cascades of traffic-shaping devices can have a severe impact on flows across multiple ISPs with different throttling times. In Figure 1 give an example of a difficult case of combined throttling using hypothetical uplink and downlink rates of a sender and receiver at different access ISPs. In particular, to observe that by chaining such throttling bottlenecks, the combined throttling of receiver (valley (u)) and sender (valley (v)) results in a small transferred volume indicated by the shaded area behind the two rates. In the extreme case of sender and receiver pairs with long, non-overlapping valleys, the transfer could be throttled across the day. [9],[10].

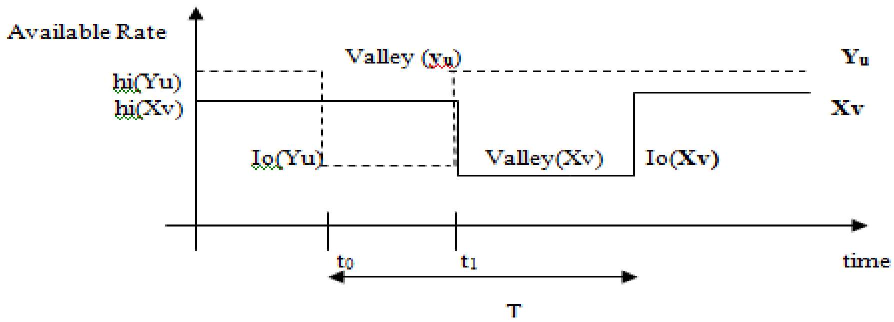

Figure1:E2E and SnF transfer data between sender ISP V & receiver ISP U

C. Bulk Bottlenecks under Percentile Pricing

Similar problems can arise under the 95-percentile pricing often applied to corporate customers or hosting services that pay based on (nearly) peak usage. Such pricing is justified by the fact that the cost of networking equipment depends on the maximum load it has to carry with a certain quality of service (QoS). Given that customers pay according to peak traffic, and granted that loads typically exhibit strong diurnal patterns, this leaves much already-paid-for off peak capacity that can be used to send additional bulk data at no extra cost. The end result could be additional transit costs as the bulk flow cannot avoid increasing the peak load.

D. Store 'n' Forward to the Rescue

As such bottlenecks (due either to pricing or throttling) become more prevalent, it argue that to restore the performance of bulk transfers and minimize transmission costs will require a new 'Store' n ‘Forward' (SnF) service based on 'temporal redirection' techniques. Existing 'spatial redirection' techniques like native and overlay routing perform path selections over short periods to avoid bottlenecks at the Internet’s core, but have no way of escaping complex accumulated constraints that can occur in the future via the combination of various bottlenecks in different time zones.

While experimented taken from [3],[11] with several nodes, present results from a setup of Figure 3 for ease of explanation [3],[6],[11]. In this figure, node V is the source in Western Europe and node U is the destination in South America. Intermediate nodes T1, T2 and T3 are three large well-provisioned transit ISPs, also in Europe. In associate storage nodes with all the transit ISPs.Obtained all the traffic loads information from one of the largest transit providers in Europe that peers with ISPs from all continents. Use data from real traffic links.

Consider the cost of storage at the rate of $0.40/GB/Month [3] (about thrice the cost of Amazon’s S3). Using this as a basis, to associate storage cost into our network model. In addition, Consider recent transit costs of $30K /Gbps/month [8][10] as a basis for calculating the data transfer cost. To make our study richer, assumed that data transfer rate between the transit links T2 and T3 was about half the transit cost above. Work with one day’s worth of data. The original data in 5-minute intervals was converted to hourly data. The data in all links was found to have non-coinciding peaks and valleys. Consider the following three scenarios:-

- First, consider how much it would cost to send data without storage nodes at current data transit rates.
- As a second step, add storage nodes at the intermediate ISPs and
- Finally, consider a variable transit cost model where the cost of data transfer rises during peak hours.

For simplicity, to consider the transit cost during peak hours to be twice the normal transit cost.

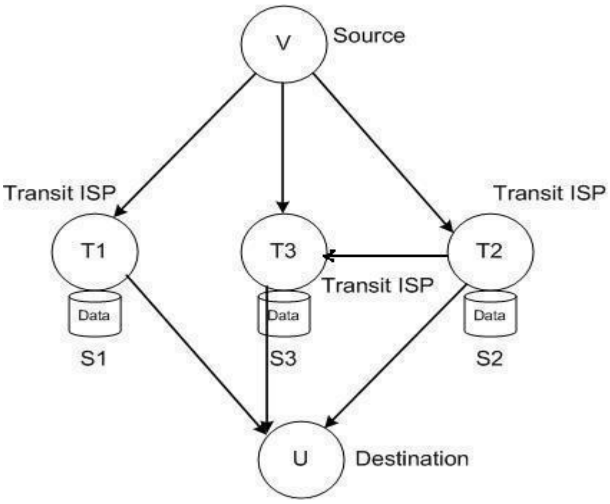

Figure 2: Topology used in the sending terabytes from a sender V in Europe to a destination U in South America.

TABLE I:
Transit cost for daily data transfer from source (V) to destination (U)

Data	6 TB	18 TB	30 TB
No storage	0.56 Million	1.8 Million	4 Million
Storage	30K	360 K	1.2 Million
Storage + Varying cost	30K	400 K	1.4 Million

To summarize the results from the analysis in Table I [6] will generate from Fig.2. It gives the result of monthly transit cost (in us dollars) for daily data transfer volume of 6tb, 18tb and 30tb from source v in Europe to destination, u in south America. The following observation is :

- With no storage at the ISP, transfer several TB of data everyday with an almost linearly increasing transit cost.
- With storage, the cost of data transfer does not increase linearly as the amount of data transferred.
- This shows that the data transfer is taking advantage of non-coinciding peaks and valleys between two links to send majority of the data during non-peak hours at the receiver without a major increase in transit cost till about 18TB. For the 30TB transfer, the transit cost contributes significantly to the data transfer. The US $30K is the normal monthly transit cost for daily 6TB transfer.[16]

III. LIMITATIONS OF TOLERANT BULK DATA

- Several ISPs throttle P2P traffic from flat-rate residential customers during peak hours in order to free capacity for interactive traffic that is valued more highly by most users.
- Similar problems can arise under the 95-percentile pricing often applied to corporate customers or hosting services that pay based on (nearly) peak usage.
- Pricing is justified by the fact that the cost of networking equipment depends on the maximum load it has to carry with a certain quality of service (QoS).

IV. PROPOSED DESIGN

Here proposednew concept is that when data centre send data to receiver through SnF. SnF use storage node which use on technique if supposed in following Fig. 2 supposed three receiver wants same data at particular time zone then there is SnF sending data to first receiver as 30% similarly for second receiver 37% & then for third there is 39% it will go continue up to 100% but in SnF side there is to remove that average data i.e 30% data from storage node & free memory for next request. If suppose for next time zone it require same data then it go to request to data centre. This shows in fig 2

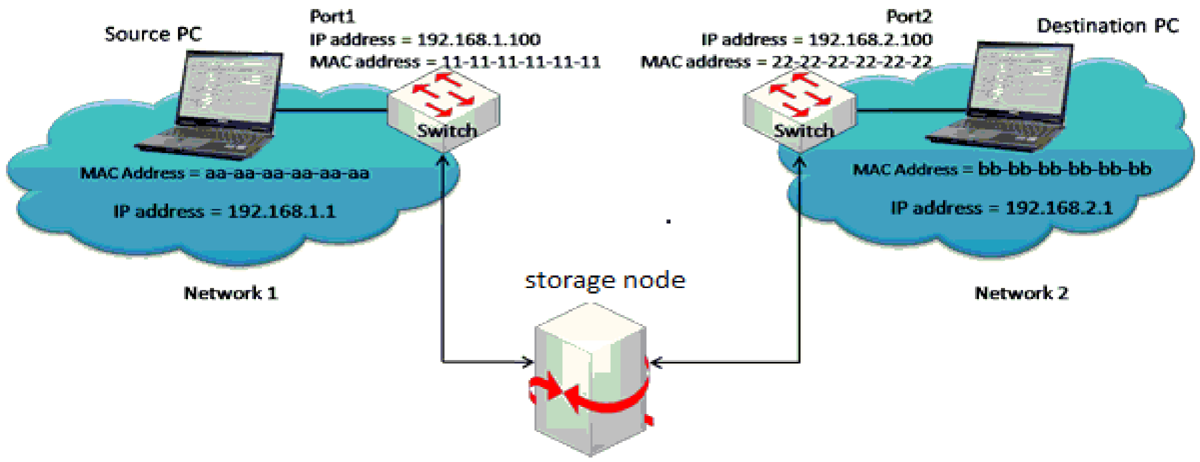

Figure. 3 proposed Design for SnF transfer between sender and receiver through storage node

V. PROPOSED SOLUTION

- In this paper present new concept is that when SnF send data to receiver then it automatically calculate mean of percentage data was transferred & that data automatically remove from memory & then it will free memory for next request.
- Given that the cost for the AP is determined by the peak-hour traffic, the individual user contribution towards the aggregate cost is not a linear function of its byte usage.

- As well as experimental evaluations for the problem of delay on transferrin bulk data over a network with time varying capacities and costs
  - To focus on the effect of storage, where the nodes possess storage capabilities that also vary in capacity and cost with time. A summary of our major contributions is-To present a general SnF method for transforming any data with capacities and costs to network.
  - To show speed on network to transfer data & also shows cause of delay on network.
  - Bandwidth schedule for transfer of bulk-data can be reduced to minimum cost flow on the layered network.
  - To demonstrate capturing packets from node to calculate load on traffic.

VI. CONCLUSIONS AND FUTURE WORK

To build the system to increase the efficiency of network, sharing the bulk data without increasing load on servers. In this system manage the pricing issue according to pick on/off of network. Also handling delay data transfer between server and multiple clients using SnF (store and forward) method. This system also manages the price based data transfer according to package provided by network provider (ISP).

In recent times there has been much interest in theoretic formulations where there are multiple strategic entities each pursuing their own self-interest. It would be interesting to generalize our model to such situations. One possible approach is to consider nodes to be partitioned into groups (ISPs) and have a utility function for each group that captures the group's costs and benefits. Also need some charges which will affect to our system. The following work gives more powerful system to transfer data.

- To presented a general scheme for transfer data need to the internet in encoding form.
- When to sending data is loosed then for that requires error recovery.
- In sending data multiplexing concurrent DTB jobs required.
- Utilizing multiple up/down links for transfer.
- Survey of how changing market policies will affect the applicability of the model.

REFERENCES

[1] Ahuja, R., Magnanti, T., and Orlin, J., "*Network Flows: Theory, Algorithms and Applications*", Prentice-Hall, 1993.

[2] Amazon Import/Export. At http://aws.amazon.com/importexport/

[3] V. Aggarwal, A. Feldmann, and C. Scheideler, "*Can ISPs and P2P users cooperate for improved performance?*,"Comput. Commun. Rev.,vol. 37, no. 3, pp. 29–40, 2007.

[4] M. Armbrust, A. Fox, R. Griffith,A.D.Joseph,R.Katz,A.Konwinski, G. Lee, D. A. Patterson, A. Rabkin,I. Stoica, and M. Zaharia, "Above the clouds: A Berkeley view of cloud computing," UC Berkeley TR EECS-2009-28, 2009.

[5] *E.Brosh,S.A.Baset,D.Rubenstein, and H. Schulzrinne, "The delay-friendliness of TCP*," inProc. ACM SIGMETRICS, 2008, pp. 49–60.

[6] J. Chabarek, J. Sommers, P. Barford, C. Estan, D. Tsiang, and S. Wright, "*Power awareness in network design and routing,*" inProc. IEEE INFOCOM, 2008, pp. 457–465.

[7] A. Chaintreau, A. Mtibaa, L. Massoulie, and C. Diot, "*The diameter of opportunistic mobile networks*," inProc. ACM CoNEXT, 2007, Articleno. 12.

[8] C. Courcoubetis and R. Weber, Pricing Communication Networks: Economics, Technology and Modelling. Hoboken, NJ, USA: Wiley,2003.

[9] A. Dhamdhere and C. Dovrolis, "*ISP and egress path selection for mul-tihomed networks*," inProc. IEEE INFOCOM, 2006.

[10] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, "*Delegation for-warding,*" inProc. ACM MobiHoc, 2008, pp. 251–260.

[11] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, "*Diversity of forwarding paths in pocket switched networks*," in Proc. ACM IMC,2007, pp. 161–174

[12] Church. K, Grenberg, A, Hamilton, J. "Delivering Embarissingly Distributed Cloud Services. Proceedings of ACM HotNets – VIII.

[13] Feynman, R. "*The Feynman Lectures in Physics*", Addison-Wesley, 1970.

[14] Garey, M., and Johnson, D., "*Computers and Intractability: A Guide to the Theory of Incompleteness*", Freeman, 1970.

[15] Goldberg, A. Network optimization library. Available at http://www.avglab.com/andrew/soft.html

[16] Grotschel, M., Lovasz, L., and Schrijver, A., "*GeomeMultiplexing concurrent DTB jobs Utilizing multiple up/down links for transfer Survey of how changing market policies will affect the applicability of the model tric Algorithms and Combinatorial Optimization*" Springer-Verlag, 1988.

[17] Xbox live and Netflix. At http://www.xbox.com/en-US/live/netflix/default.html