# On The Break-even Point Between Cloud-Assisted and Legacy Routing

Prasun Kanti Dey
Dept of Electrical and Computer Engineering
University of Central Florida
Email: prasun@knights.ucf.edu

Murat Yuksel
Dept of Electrical and Computer Engineering
University of Central Florida
Email: yuksem@gmail.com

***Abstract*—As more than 40K service providers are advertising 600K or more IP prefixes, scalability of routing has emerged to be a matter of great concern. In this paper, to explore a spectrum of designs, we consider a Cloud-Assisted Routing (CAR) framework which follows a hybrid and opportunistic approach by keeping the high priority tasks at the router and use an adaptive router-cloud integration when beneficial. In particular, it maintains *most* of the control plane functions at the cloud and *least* of it at local router and vice versa for the data plane. Comparing the performance and monetary cost benefits between CAR, we discuss: *i*) What is the break-even point? *ii*) What are the key components of CAR to be monetarily beneficial? *iii*) What are the constraints that will make the traditional routing favorable than CAR?**

## I. Introduction

With the rise of ubiquitous provisioning of Internet connectivity, many Internet Service Providers (ISPs) are delivering services, applications and even virtualized infrastructures for customers. The advent of software-defined networking (SDN) [7] attests the success of *offloading control complexity from the network to software platforms* for centralized optimizations rather than leaving them to the core components like routers. Recent consensus is to do that with open interfaces, as in OpenFlow [12], allowing outsourcing of control plane tasks to remote platforms that are potentially managed by enterprises other than network operator, such as the cloud [14].

As complexity of network layer is increasing while cost for cloud computing services is declining consistently, offloading more of the networking functions to the cloud seems very natural. For performance intensive packet level functions like routing, local hardware platform cannot entirely be excluded from the design in order to maintain the lookup times in the order of nanoseconds. Yet, the routing complexity is growing further and need of software platforms seems inevitable.

While the price of router hardware (e.g., DRAM) is not consistently declining, the cloud service prices are following a trend of 3-10% reduction each year, which our study in Figure 1 also supports. But, the main question still exists, *"By leveraging the cloud's memory and computation resources, can we actually remedy the routing scalability issue?"*

Most of this work has been done when authors were in University of Nevada, Reno.

This work was supported in part by NSF award CNS-1321069.

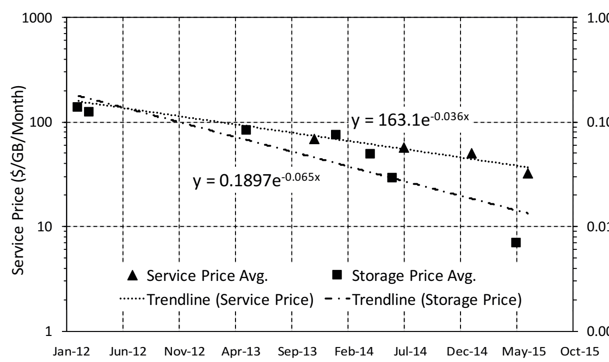



Fig. 1: Cloud storage vs. service price trends

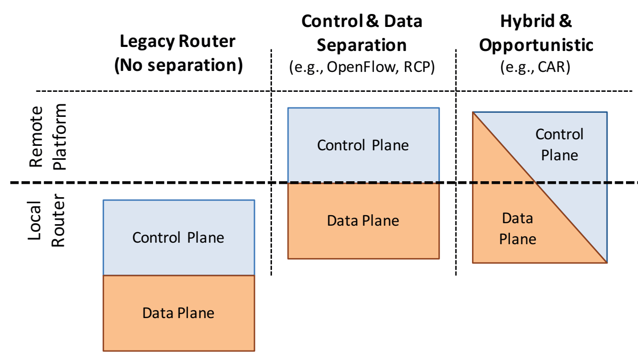



Fig. 2: Hybrid vs. full-separation of routing functions

For an answer to this question, we consider Cloud-Assisted Routing (CAR) [10] which advocates for a hybrid (Figure 2) and opportunistic placement of data and control plane functions between the local router hardware and remote cloud. We think, hybrid approaches that maintain high priority tasks at the router with *partial* offloading to remote platforms and employ an adaptive cloud-router integration framework have more likelihood of addressing future routing scalability and flexibility trade-offs. This enables us to explore the possible break-even points between legacy and cloud-supported routing as well.

In this paper, we explore this break-even point by analyzing the trend of last 30 years DRAM price and the declining trend of cloud storage and service prices. Developing a comparative

978-1-5090-5093-2/16 $31.00 © 2016 IEEE
DOI 10.1109/CloudNet.2016.32

model for CAR and traditional routing costs, we identify break-even points between CAR and legacy routing, and determine whether CAR presents a viable and promising architecture for solving the growing routing complexity. Since CAR requires less physical memory at local router, this essentially reduces the FIB (Forwarding Information Base) cache size as well. We also show how much elasticity we can achieve regarding FIB cache size in near future if we implement CAR.

The rest of the paper is organized as follows: Section II discusses our modeling work on CAR cost versus legacy routing cost and details our analyses in order to arrive at the model. Then, we identify break-even points between CAR and legacy routing in Section III. Finally, Section IV summarizes our work and discusses possible future work.

### A. Motivation

BGP FIB size has already reached to 600K entries [3], exceeding TCAM (Ternary Content-Addressable Memory) capacities and most conservative forecasts point to a quadratic growth of the routing table size [4]. This growing table size combined with longer prefixes (due to shrinking IPv4 space and growth of IPv6) and line rates (45Mbps in 1998 to 100Gbps in 2013) become challenging to address [9]. Time budget for forwarding lookups is getting more stringent due to increase in backbone speeds, and further, more complex lookups are in the horizon due to more flexibility and friendliness expected from network routing.

While FIB compression and aggregation [17] efforts are shown to gain 60% reduction in SRAM (Static Random-Access Memory) requirements [13], it is questionable whether or not these efforts alone can continue fitting the growing routing tables into legacy SRAMs while matching the shrinking budget for forwarding lookups. Although SDN [7] and OpenFlow [12] offer great flexibility and programmability, software platforms may not be sufficient in addressing the scalability concerns of routers. CAR deviates from typical SDN design by placing most of the control at remote platform and the least of it at the local router.

Armbrust et al. 2010 [6] discusses about the elasticity of the cloud system by showing an example of pay-per-use pricing for cloud. From business perspective, it is beneficial to have the ability to add or remove resources at any time. Methods of cloud pricing is always an issue between the user and cloud system [15]. But, earlier works primary focused on the computational expenses caused by exploring distributed systems on cloud rather than evaluating the storage price.

Packets delegated from the router will be transmitted over the actual physical link towards the cloud, which will cause transit cost. Depending on link type (i.e., dedicated or shared), the cost will vary. Odlyzko [8] has proposed a model on quality of service. We use that for developing our cost model.

### B. Architecture

An architectural view of a hybrid "CAR router" is illustrated in Figure 3. As discussed earlier, CAR aims to find a middle ground where it can exploit both the local hardware to scale router performance and a completely cloud-based approach for highly flexible routing service. Similar to how virtual memory systems use secondary storage to maintain full content, CAR uses cloud to implement the full functionality of *Router X (RX)*, and keeps *RX* as 'active' while *Proxy Router X (PRX)* as 'passive'. CAR designers should follow these two principles:

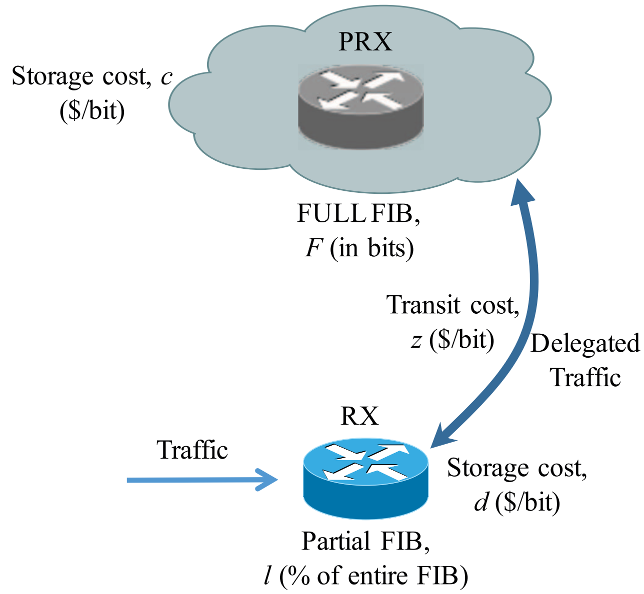



Fig. 3: CAR architectural view

1) CAR should offload computationally heavy but not-so-urgent tasks (e.g., BGP table exchange, shortest-path calculation) to cloud to the extent possible. *PRX* should hold the full FIB table and act as the default point of service for data and control plane functions that can not be handled at *RX*.
2) Keep data plane closer to the router while some of the control plane operations such as on-demand route computations due to failures or collection of flow-level simple statistics will still be done at *RX*. However, CAR will orchestrate heavy routing optimizations at *PRX*.

## II. A Comparative Cost Model

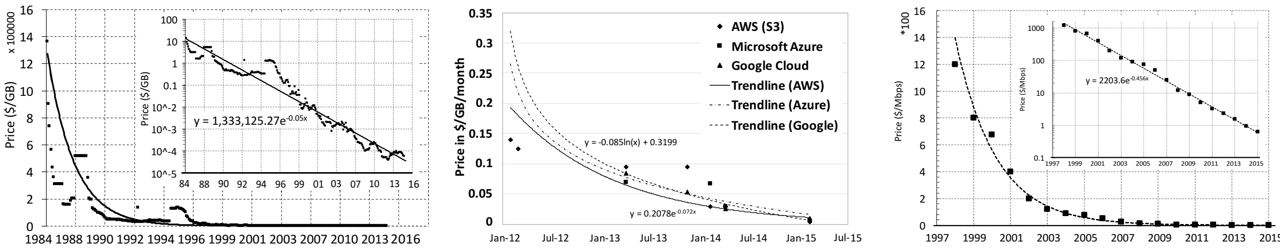
A natural question that arises is whether or not CAR will be monetarily beneficial in comparison to offloading the complex routing functions to local hardware solutions, e.g., putting the full forwarding table to a enlarged DRAM within the router rather than to a cloud. We explore answers for such economic viability questions in this section.

We base our analysis to two simplistic cost models:

$$C = dF \tag{1}$$

$$\hat{C} = dFl + cF + z\rho(l) \tag{2}$$

$C$ is the cost for traditional routing and $\hat{C}$ is for CAR. $d$ and $c$ are the costs of \$/bit storage at the local DRAM and the cloud over one unit of time, respectively. $z$ is \$/bit cost




(a) Historical DRAM price

(b) Cloud storage price

(c) Transit price

Fig. 4: Price trends

for transferring packets delegated to the cloud. Furthermore, $F$ is the FIB size (in bits), $l$ is the percentage of FIB size that needs to be stored in the local router (basically, FIB cache) to sustain an acceptable (i.e., close to traditional routers) average delay for packets forwarding time. where, $\rho(l)$ is the amount of traffic delegated to the cloud and is a decaying function of $l$ due to the significant locality in traffic (Zipf's law for Internet traffic). We now delve into modeling the three cost components $d$, $c$, and $z$.

*1) Memory Prices: $d$ and $c$:* Forecasting memory pricing is uncertain due to its heavy dependency on market. For modeling, DRAM price $d$ in our model, we collected DRAM prices starting from 1984 [2], shown in Figure 4a. It is not always true that the DRAM price is going to be lower in future. Yet, *favoring the traditional routers*, we model the DRAM price with an exponential decay respect to time, $t$:

$$d(t) = 1,333,125.27\ e^{-0.05t} \tag{3}$$

To model the cloud storage price, $c$, we analyzed the prices of the most used cloud storage providers (i.e., Amazon Simple Storage Service (AWS S3), Google Cloud, Microsoft Azure) [5], [11], [16] shown in Figure 1 and Figure 4b. It is quite clear that price of storage at cloud is strictly reducing and gives us the following exponential decay model:

$$c(t) = 0.1897\ e^{-0.065t} \tag{4}$$

*2) Transit Price, $z$:* Transit service prices by ISPs are known to decline at an exponential pace according to the work by Fishburn and Odlyzko [8]. Using their work and based on our collection of actual transit service price values [1], we plot transit costs in Figure 4c. The trend is certainly decaying, which means that the cost of delegating data traffic to the cloud will eventually reduce and favor the CAR approach. Fitting the data to an exponential decay gives us the following model for the transit cost:

$$z(t) = 2,522.4\ e^{-0.496t} \tag{5}$$

## III. Economic Viability

### A. Break-even Point

Based on the cost models (1) and (2), we find the break-even points between CAR and No CAR cases in terms of the costs. Though, Figure 4c shows the transmit cost as $\$/Mbps$, for our calculation, we normalized all the values to gigabyte. As a parameter into this comparative analysis, we consider the percentage of FIB that is to be stored in $RX$, i.e., the FIB cache size.

Figure 5 plots the graph for cloud storage and DRAM costs.

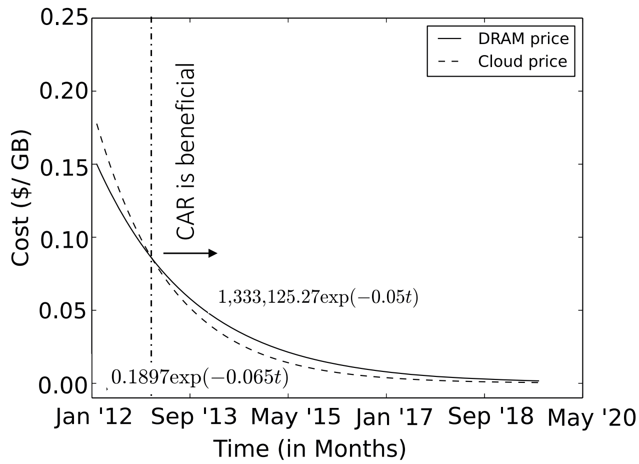



Fig. 5: Cloud storage price vs. DRAM price

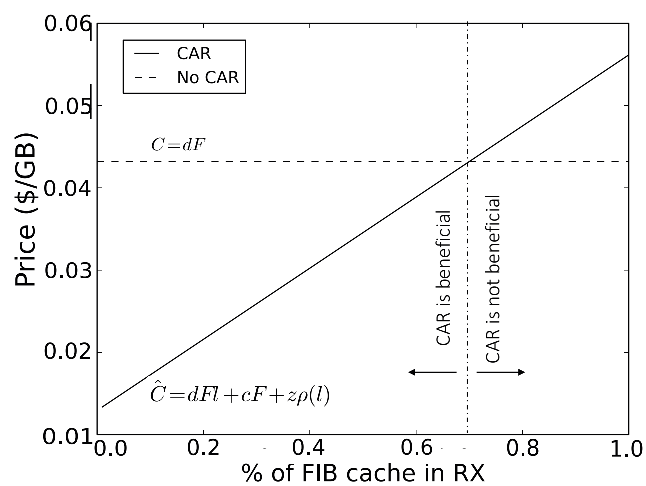



Fig. 6: CAR vs. No CAR after 96 months

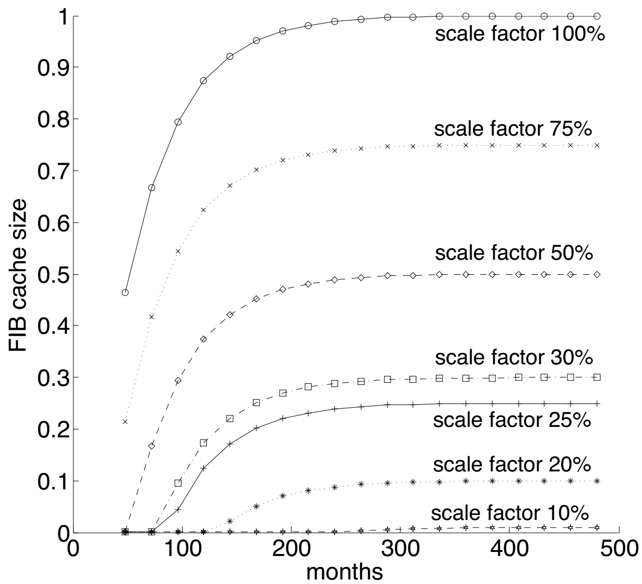



Fig. 7: CAR savings compared to legacy routing with no CAR

After 2013, cloud storage price is cheaper than DRAM and it will be more economic in near future. Based on Section II-1, since DRAM does not clearly follow the decaying trend, this graph will certainly be profitable for cloud storage.

Figure 6 illustrates the comparison between $C$ and $\hat{C}$. $\hat{C}$ increases as FIB cache size increases, while $C$ maintains the constant value as it is FIB cache size independent. Break-even happens for this two model when FIB cache size is 70%, which means, $\hat{C}$ is more economical as long as we keep the FIB cache size less than or equal to 70% of the total FIB size.

### B. CAR Savings

So far, we have explained how much savings CAR can offer at any specific time in future. We have extended our work in Figure 7 to find out how much buffer CAR can give in FIB cache size reduction over a period of time. Here, scale factor is the percentage of CAR savings over traditional routing i.e., scale factor 75% means, $\hat{C}$ is 25% cheaper than $C$. All the curves are concave, meaning they will reach to a certain maximum point. FIB cache size can not be reduced beyond that threshold for a given scale factor.

## IV. Summary and Discussion

To address the alarming increase in routing complexity, particularly in inter-domain level, we have proposed a new architectural approach, Cloud Assisted Routing. Leveraging high computation and memory power of cloud services to reduce the complex routing functions such as forwarding and flow level policy management from physical router, CAR will certainly improve the routing performance. While covering the possible benefits of the system, we also outlined the architectural principle and key components of CAR.

Several future works and security concern resolutions are needed to be implemented. Beside using just the cloud storage for keeping the entire FIB, we can explore the in-memory service (like REDIS) in CAR scenario to get faster response and better performance. Apart from bridging the gap between hardware router and software based routing services, CAR can allow more centralized cloud-based intra-domain traffic engineering optimization(e.g., OSPF link weight setting, MPLS path computation and establishments). For inter-domain routing, cloud providers can host Proxy Routers for any physical Router as a trusted third party for traffic balancing and different SLA management between multiple autonomous systems.

By involving cloud-aware backbone in forwarding, CAR will certainly create a wide range of research opportunity in failure-triggered traffic delegation method as well. All these aspects including the price decaying trend in cloud and uncertain DRAM cost in future are surely critical and worthy of further exploration.

## References


[1] Internet transit prices - historical and projected. Available online: http://drpeering.net/white-papers/Internet-Transit-Pricing-Historical-And-Projected.php.

[2] Memory prices (1957-2015). Available online: http://www.jcmit.com/memoryprice.htm.

[3] The CIDR Report. Available online: http://www.cidr-report.org.

[4] BGP routing growth in 2011. APNIC, October 2011. Available online: http://labs.apnic.net/blabs/?p=25.

[5] Amazon s3, ms azure and google cloud storage pricing comparison. CloudBerry Lab Blog, April 2016. Available online: http://www.cloudberrylab.com/blog/amazon-s3-azure-and-google-cloud-prices-compare/.

[6] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, et al. A view of cloud computing. *Communications of the ACM*, 53(4):50–58, 2010.

[7] N. Feamster, J. Rexford, and E. Zegura. The road to SDN: An intellectual history of programmable networks. *ACM Queue*, 2013.

[8] P. C. Fishburn and A. M. Odlyzko. Dynamic behavior of differential pricing and quality of service options for the internet. *Decision Support Systems*, 28(1):123–136, 2000.

[9] J. Herzen, C. Westphal, and P. Thiran. Scalable routing easy as PIE: A practical isometric embedding protocol. In *Proceedings of IEEE ICNP*, pages 49–58, October 2011.

[10] H. T. Karaoglu and M. Yuksel. Offloading routing complexity to the cloud(s). In *Proceedings of IEEE ICC Workshop on Cloud Convergence (WCC)*, pages 1367–1371, Budapest, Hungary, June 2013.

[11] S. Lelii. Amazon, google, microsoft cut cloud storage prices again, December 2012. Available online: http://searchcloudstorage.techtarget.com/news/2240174898/Amazon-Google-Microsoft-cut-cloud-storage-prices-again.

[12] T. A. Limoncelli. OpenFlow: A radical new idea in networking. *ACM Queue*, June 2012.

[13] G. Rétvári, J. Tapolcai, A. Kőrösi, A. Majdán, and Z. Heszberger. Compressing IP forwarding tables: Towards entropy bounds and beyond. In *Proceedings of the ACM SIGCOMM*, pages 111–122, 2013.

[14] J. Sherry, S. Hasan, C. Scott, A. Krishnamurthy, S. Ratnasamy, and V. Sekar. Making middleboxes someone else's problem: Network processing as a cloud service. In *Proceedings of ACM SIGCOMM*, Helsinki, Finland, August 2012.

[15] H. Wang, Q. Jing, R. Chen, B. He, Z. Qian, and L. Zhou. Distributed systems meet economics: pricing in the cloud.

[16] K. Ward. Amazon cuts cloud storage prices to nearly nothing, March 2015. Available online: https://virtualizationreview.com/articles/2015/03/31/amazon-cuts-cloud-storage-prices-to-nearly-nothing.aspx.

[17] M. Zec, L. Rizzo, and M. Mikuc. DXR: Towards a billion routing lookups per second in software. *ACM SIGCOMM Computer Communication Review*, 42(5):29–36, 2012.