Power-aware Routing with Rate-adaptive Network Elements

Spyridon Antonakopoulos
Bell Laboratories, Murray Hill NJ
Email: spyros@research.bell-labs.com

Steven Fortune
Bell Laboratories, Murray Hill NJ
Email: sjf@research.bell-labs.com

Lisa Zhang
Bell Laboratories, Murray Hill NJ
Email: ylz@research.bell-labs.com

***Abstract*—Current Internet service-provider networks are typically over-provisioned, with the actual traffic through a network element often being much less than the capacity of the network element. However, current network element power consumption is largely independent of actual traffic. This presents an opportunity to reduce network power usage. Such an opportunity may be exploited locally, by redesigning individual network elements to make them rate-adaptive, or globally, by power-aware traffic routing.**

Instantiating either approach requires significant engineering effort. We attempt to quantify, as realistically as possible, the power-savings opportunity that can be obtained using these two approaches, in isolation or together. In particular, we investigate whether power-aware routing provides any additional benefit if network elements are rate-adaptive.

We adopt a fairly simple model of network power use, where power consumption is attributed to links. A link may be turned off, in which case its power usage is zero. Otherwise, a link that is turned on consumes no less than a certain amount of power, called base power, and power use increases further with traffic. A significant parameter of the model is the ratio of base power over full-capacity power. Since it is difficult to estimate feasible values for this ratio, we investigate multiple scenarios in which its value ranges from 0 to 100%.

We demonstrate that the combination of rate-adaptivity and power-aware routing saves a significant fraction of network power consumption, for a wide variety of network topologies, traffic loads, and base power ratios. More specifically, if the base power ratio is 50% or more, then power-aware routing appears to be of significant additional benefit over rate adaptivity alone; if the ratio is 25% or less, then power-aware routing offers a relatively small additional benefit.

I. INTRODUCTION

Current Internet backbone networks are provisioned with much more capacity than average traffic [1]. Two major reasons for this over-provisioning are the significant daily variation in traffic load [11] and the need for redundant capacity to handle network element or link failures. However, the electricity use of current network elements (routers and switches) appears to be largely constant, independent of actual traffic [4]. This presents an opportunity to reduce backbone electricity use by making it more sensitive to traffic load.

One method to achieve this goal is via local power management within each network element, by making the power consumption of the element *rate-adaptive*, i.e. dependent on the traffic through the element. For example, a router typically has a network processor that processes packets for IP address lookups, congestion control, and policy enforcement. Techniques that are well-known in the computer industry for power management (speed scaling, clock gating, sleep modes, etc.) can be applied to network processors as well, and significant power savings are predicted to follow [7].

Power management can also take place globally, at the network management level. Traffic within the network may be routed in a way that minimizes the total power used by the network. Such traffic management assumes that the relationship between traffic and power consumption is well understood for all network elements, and that it is feasible to compute a routing that achieves substantial power savings.

Instantiating either of these approaches presents formidable challenges: the local case presupposes a redesign of network element hardware and software; the global case requires a management system that is capable of network-wide traffic rerouting in response to traffic fluctuations.

The goal of this paper is to obtain as realistic as possible a prediction of the power savings that could be obtained by using one or both of these approaches. Both approaches address the same opportunity, so a specific question is whether power-aware routing is still useful even if network elements have been redesigned to be rate-adaptive. We estimate power consumption by developing models for the power requirements of network elements, and then simulating various network topologies under varying traffic loads.

The power use of a typical router can be attributed to packet-processing, to the switch fabric, and to the physical interface within every line card, with each of the above consuming 25-50% of the total. We do not yet know the full extent to which each function can be made rate-adaptive; packet-processing seems to be the easiest and the physical interface the most difficult to optimize.

We employ a relatively simple model in which total network power consumption is allocated on a per-link basis, and consists of a fixed component called *base power*, which is required merely to keep the link operational, as well as an additional component that is dependent on traffic. (Note that the ratio of base power to maximum power use is in effect the complement of the *energy proportionality index* (EPI) [8]—expressed in percent, base power ratio plus EPI sum to 100.) Future feasible values for base power ratio depend upon the detailed engineering of a network element and thus are difficult to estimate. Hence, we consider several possible values for said ratio, ranging from zero to 100%. We also assume that a link may be turned off completely, using no power at all.

This study is explicitly scale-invariant: we predict relative but not absolute power savings. Some estimates on router power consumption are available [4], [12], which together with absolute traffic estimates can predict absolute power savings.

Traffic engineering in service-provider networks has many competing objectives and constraints, e.g. delay minimization, quality-of-service guarantees, reliability, business agreements, etc. It is beyond the scope of this paper to study these constraints in combination with power minimization. We instead attempt to estimate the power-saving opportunity. Of course, exploiting the latter could pose new, significant technical hurdles; for example, an equipment failure may necessitate rapid reactivation of nonlocal links that have been turned off to save power.

The rest of this paper is organized as follows. Section II describes how we generated test instances, Section III the experiments, Section IV the corresponding results, and finally Section V presents our conclusion.

Related work: Power-aware routing has received considerable attention, including [4], [8], [9], [10]. On the other hand, we believe that we are the first to study systematically the combined effects of power-aware routing and rate-adaptivity with base power.

II. Test-instance Methodology

A test instance consists of a network topology with capacitated links, a traffic matrix T, and for each link e, a power-traffic function $f_e(x)$ that determines the power needed to carry x units of traffic on e. We describe the methodology used to generate synthetic test instances. The goal was to obtain instances that model correctly-provisioned service provider networks; to do this we mimic the process by which a service provider might design its network. The instance-generation methodology assumes that only a network topology is given, and infers from it the traffic matrix, link capacities, and power-traffic curves, as explained below.

The network topologies in our test instances are derived from the datasets in the Rocketfuel study [2]. This study used packet-tracing technology to determine the router-level topologies of autonomous systems in Europe, Australia and the United States. From each Rocketfuel dataset we construct two network topologies. The first one, the *router graph*, is the router-level topology from the dataset, with minor cleanups (e.g. discarding nodes and links not in the largest connected component). However, the dataset also specifies the city containing each router, and typically there are multiple routers at each city. The second topology, the *city graph*, is obtained by collapsing all routers at a city together, and then merging together any resulting parallel links. The city and router graphs for Rocketfuel AS 3967 are shown in Figure 1. Unsurprisingly, city and router graphs constructed from the same dataset tend to be structurally similar; nevertheless, the router graph generally has more links than the city graph, and hence there are more opportunities for a power-aware routing algorithm to turn links on and off.

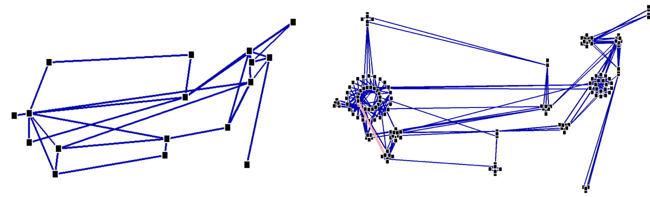

Fig. 1. City graph (left) and router graph (right) from the Rocketfuel AS 3967.

The traffic matrix is based on a variant of the population-distance model, in which the traffic between two nodes is proportional to the product of their respective populations, divided by a distance factor [6]. We estimate the “population” of each city by the number of routers located at that city, under the assumption that the service provider has approximately the right number of routers to handle all traffic generated by the city. In particular, the traffic entry $T[a,b]$ is

$$T[a,b] = \frac{|\text{routers}(a)| \cdot |\text{routers}(b)|}{e^{\text{distance}(a,b)/1500}} .$$

Link capacities are chosen so that traffic can still be routed even if any single node fails. Such link capacities are easy to compute: we remove each individual node in turn from both the network and the traffic matrix, and then route the remaining traffic in the network using shortest-hop routing. The capacity of each link is set to the maximum load on the link, over all choices of removed node.

The power-traffic function of each link e depends upon an instance-specific parameter β, representing the base power ratio, and the capacity c_e of the link. The function linearly interpolates between βc_e for zero traffic and c_e for full traffic, but exhibits a discontinuity at 0, where its value also becomes zero (meaning that link e is turned off). Formally, the power-traffic function $f_e(x)$ for a link e with capacity c_e is

$$f_e(x) = \begin{cases} 0 & \text{if } x = 0 \ ; \\ \beta c_e + (1-\beta)x & \text{if } 0 < x \leq c_e \ . \end{cases} \quad (1)$$

Notice that the marginal power per unit traffic is independent of the capacity of the link, but does depend upon β (in fact, it is $1-\beta$). In Figure 2, plots 2–6 show five power curves for $\beta = 100\%, 75\%, 50\%, 25\%, 0\%$. The left-most plot is the “always-on” power-traffic function

$$f_e(x) = c_e \quad \text{if } 0 \leq x \leq c_e \ , \quad (2)$$

which models the situation where power consumption is completely oblivious to fluctuations in traffic.

III. Experiments

For each test instance, we wished to determine the relative benefits of rate adaptivity and power-aware routing. To do this, we computed a scaled traffic matrix T_α by multiplying every entry in T by α, for values of α in the range 10% to 100%. We then attempted to compute a power-aware routing

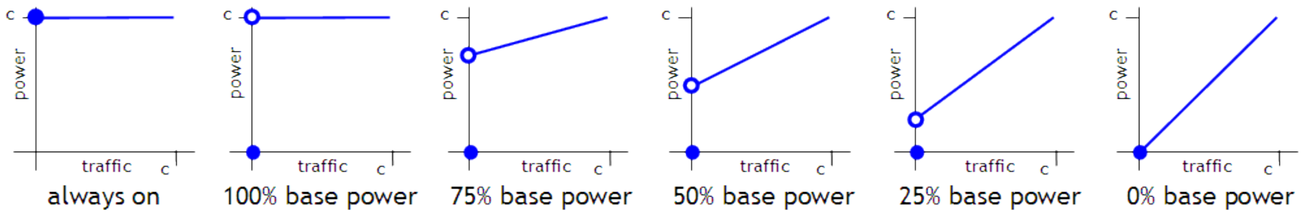

Fig. 2. The "always-on" power-traffic function, and linear power-traffic functions with various base power ratios.

for each traffic matrix T_α. A generic form of the result is given in Figure 3. The top curve, always constant, denotes power consumption when links are always on, as implied by power-traffic function (2). The second curve is the improvement due to rate adaptivity. The third curve shows the further improvement due to power-aware routing. Finally, the bottom curve is a straight line indicating perfect rate adaptivity (i.e. 0% base power); this constitutes a bound on the improvement attainable by power-aware routing.

Computing a routing that is globally optimal with respect to power consumption is an NP-hard problem, and we did not attempt to find such a routing. Instead we used the following heuristic, which might be called *min-power iterative greedy routing*. Initially, the network does not carry any traffic and no links are turned on. Each demand is considered one by one, and routed on the path of least total *marginal cost*, which is simply the extra power required to carry the demand. Thus, the marginal cost on a link is proportional to the size of the demand, plus the base power of the link if the latter has not been turned on yet. During the first pass, a route for each demand is computed, although some demands may not be routable because of capacity constraints. In every subsequent pass, the route of each demand is recomputed assuming existing routes of other demands, plus cost penalties that prevent filling links to capacity while there are still unroutable demands. This process halts when it reaches a point where any additional passes would not be able to reduce network consumption further, nor route any more demands.

The theoretical underpinnings of the above routing heuristic stem from work by Charikar and Karagiozova [5]. Even though we cannot provide a worst-case guarantee on the quality of the solution produced, nevertheless this heuristic is quite simple, always succeeded in routing all demands of our test instances, and did yield substantial power savings when there was opportunity to do so.

IV. Simulation Results

We constructed test instances from a number of Rocketfuel datasets; to save space we only report on AS 3967 and AS 1755. The router versions of these two instances have about 150 nodes and 350 links; the city version of AS 3967 has 17 nodes and 27 links (Figure 1), while the city version of AS 1755 has 23 nodes and 38 links.

The power usage curves as functions of traffic load for the city instance of AS 1755 are given in Figure 4. The plot in the upper left corner (100% base power ratio) assumes that there

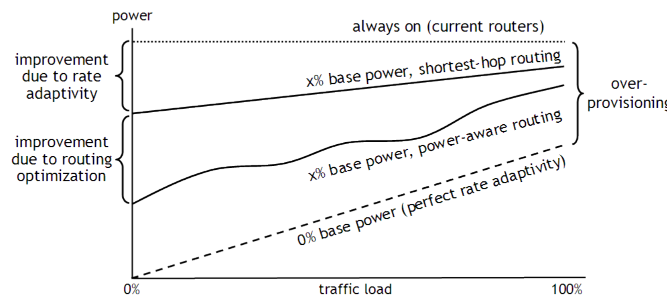

Fig. 3. Qualitative plot of power consumption using non-power-aware (shortest-hop) and power-aware routing, including the extreme cases of no rate adaptivity (always on) and perfect rate adaptivity.

is no rate adaptivity in a link, thus the benefit is solely from turning links off; in conditions of light load (30%), the benefit is substantial—roughly half the power. As the base power is decreased, the benefit of turning links off also decreases, so that in the lower right hand corner (25% base power ratio) there is little additional benefit in turning links off.

This behavior was qualitatively similar over all the instances that we tested. Figure 5 shows two of the plots for the city instance of AS 3967. In this case, the relative benefit of turning links off was somewhat less, but still significant, particularly during light load.

For each AS, the router graph has similar structure to the city graph, but with more links. We expected that the extra links would create more opportunities for saving power by turning links off, and therefore provide a greater advantage to the power-aware routing heuristic. Figure 6 shows the behavior for the router graph of AS 3967. Indeed, power-aware routing performed slightly better than in the city-graph case (cf. Figure 5), however the difference is not large.

Recall that our test-instance methodology creates links with arbitrary capacities. Real network links only have a discrete choice of capacities (1Gbps, 10Gbps, 40Gbps, etc.). To see whether having discrete link capacities would change the results, we retested some instances after rounding their link capacities up to the next highest integral power of 2. The results for AS 3967 are plotted in Figure 7. The upwards rounding markedly increases the initial over-provisioning of the network, however the relative behavior of rate-adaptation and power-aware routing was essentially unchanged.

Finally, we experimented with other models for the power-traffic function. Modern processors use dynamic voltage and

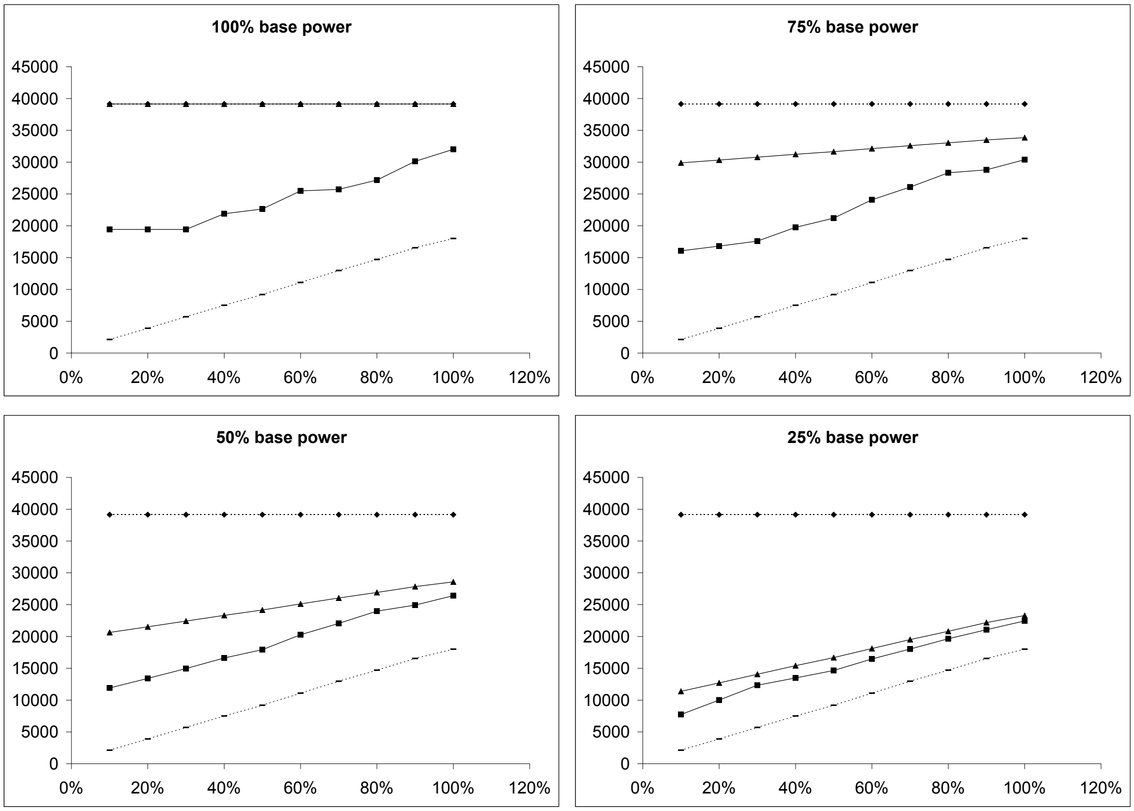

Fig. 4. Simulation results for the city graph of Rocketfuel AS 1755. In all plots, the x-axis indicates network traffic load as a percentage, and the y-axis represents network power consumption. Triangular (▲) and square (■) points denote results for non-power-aware (shortest-hop) and power-aware routing, respectively. The non-rate-adaptive (⧫) and perfectly rate-adaptive (–) cases are also plotted for reference.

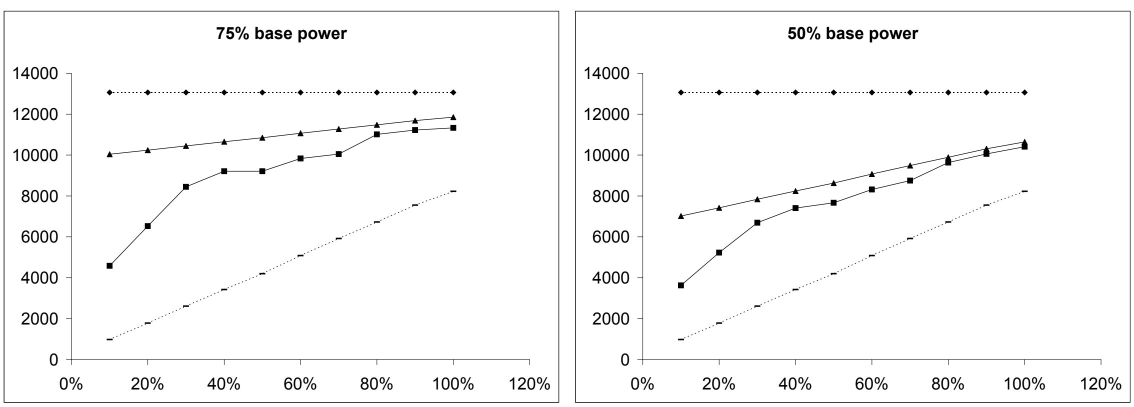

Fig. 5. Simulation results for the city graphs of Rocketfuel AS 3967.

frequency scaling to match their clock rate to processing requirements [3]. In such circumstances, power use increases superlinearly with clock rate. We modeled this behavior with a convex power-traffic function including base power, specifically

$$f_e(x) = c_e \left(\beta + (1 - \beta) \left(x/c_e \right)^2 \right) ,$$

adjusted to be zero at $x = 0$. Figure 8 shows the results, plotted simultaneously with those for the corresponding linear power-traffic curve, both on AS 3967. Network power consumption is predicted to be less in the convex case than in the linear case (since this is true on a per-link basis); nevertheless, the relative improvement due to power-aware routing is comparable in the two cases.

V. Conclusion

We believe that there is significant opportunity to save electricity in service-provider IP networks using rate adaptivity and power-aware routing. Current network elements have poor rate adaptivity, with base power ratio close to 100%. In this case, our results suggest significant benefit of power-aware routing, especially for low traffic. As rate adaptivity

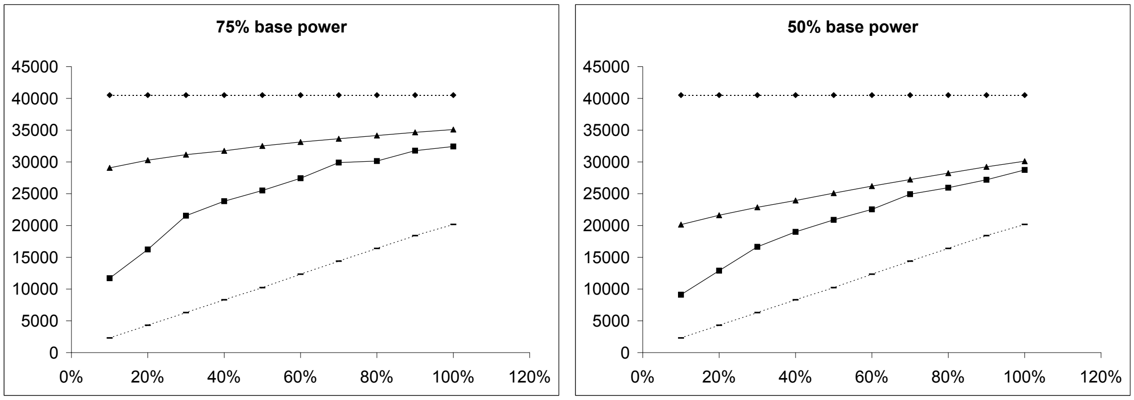

Fig. 6. Results for the router graph of AS 3967.

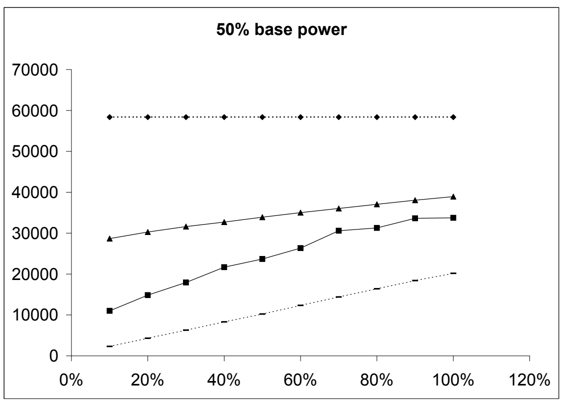

Fig. 7. The router graph of AS 3967 with quantized link capacities.

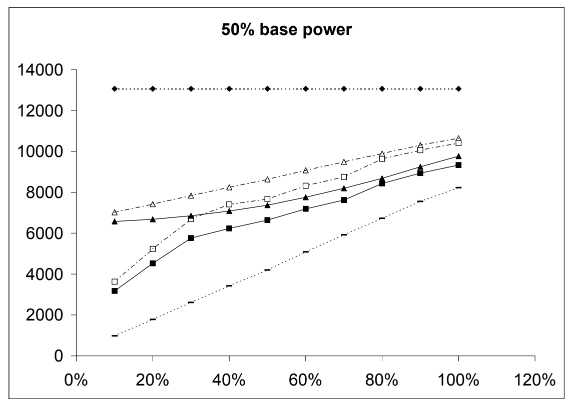

Fig. 8. The city graph of AS 3967 with the convex power-traffic function (see text). Hollow plot points show the corresponding power costs for a linear power-traffic function with the same base power ratio.

improves, the benefit of routing optimization diminishes. In the extreme case, if routers could be made perfectly rate-adaptive, then there would be no need for power-aware routing; unfortunately, this hypothesis seems rather unlikely.

Some graph-theoretic questions remain unanswered. For example, the power curves suggested some dependence on graph structure (cf. Figure 4 versus Figure 5), a fact for which we do not yet have a good explanation. Such an explanation could lead to criteria for designing network topologies that are particularly power efficient.

Acknowledgment. This R&D work was completed with the support of the U.S. Department of Energy (DOE), award no. DE-EE0002887. However, any opinions, findings, conclusions and recommendations expressed herein are those of the authors and do not necessarily reflect the views of the DOE.

REFERENCES

[1] International Internet Statistics. TeleGeography Research, 2005. Available at http://www.itu.int/dms_pub/itu-d/md/02/isap2b.1.1/c/D02-ISAP2B.1.1-C-0025!!PDF-E.pdf.

[2] Rocketfuel: an ISP topology mapping engine. Available at http://www.cs.washington.edu/research/networking/rocketfuel/.

[3] Enhanced Intel Speedstep technology for the Intel Pentium M processor. Intel White Paper 301170-001, 2004.

[4] J. Chabarek, J. Sommers, P. Barford, C. Estan, D. Tsiang, S. Wright. Power awareness in network design and routing. *Proceedings of IEEE INFOCOM*, 2008.

[5] M. Charikar, A. Karagiozova. On non-uniform multicommodity buy-at-bulk network design. *Proceedings of ACM STOC*, p. 176–182, 2005.

[6] S. Erlander, N. F. Stewart. *The gravity model in transportation analysis - Theory and extensions.* VSP, 1990.

[7] A. Francini, D. Stiliadis. Performance bounds of rate-adaptation schemes for energy-efficient routers. *Proceedings of IEEE HPSR*, 2010.

[8] P. Mahadevan, P. Sharma, S. Banerjee, P. Ranganathan. A power benchmarking framework for network devices. *Proceedings of IFIP Networking Conference*, 2009.

[9] S. Nedevschi, L. Popa, G. Iannaccone, S. Ratnasamy, D. Wetherall. Reducing network energy consumption via sleeping and rate-adaptation. *Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, pp. 323–336, April 2008.

[10] J. Restrepo, C. Gruber, C. Mas Machuca. Energy profile aware routing. *First International Workshop on Green Communications, IEEE International Conference on Communications (ICC)*, June 2009.

[11] M. Roughan, A. Greenber, C. Kalmanek, M. Rumescwicz, J. Yates, Y. Zhang. Experience in measuring Internet backbone traffic variability. *Proceedings of ACM SIGCOMM IMW*, p. 91–92, 2002.

[12] O. Tamm, C. Hermsmeyer, A.M. Rush. Eco-sustainable system and network architectures for future transport networks. *Bell Labs Technical Journal*, 14:311–327, 2010.