Improving energy efficiency in networks

Christina Morfopoulou and Erol Gelenbe
Dept. of Electrical & Electronic Engineering
Imperial College, London SW7 2BT, UK
{c.morfopoulou,e.gelenbe}@imperial.ac.uk

1. INTRODUCTION

Energy efficiency in Information and Communication Technology (ICT) has become a major priority especially for the data center networks operators and internet service providers. ICT alone is responsible for 2% of the global carbon footprint production [1], a figure similar to the one of the airline industry. Especially in the case of more developed countries this figure rises up to 10% and is expected to grow significantly in the future. On the other hand, network devices are highly overprovisioned in order to endure peak time demands and failures and thus they are underutilized in large percentage of time, which reveals an opportunity for large energy savings. This research area has recently drawn much attention [2] but there is no systematic approach developed yet. Thus, our goal is to study and identify opportunities and solutions for energy efficiency in networks while taking into careful consideration the network performance and the provided Quality of Service (QoS).

2. POWER AWARE ROUTING CONTROL

We consider the case where there is a given store and forward packet network with given capacities and power profiles of the network devices. In this case, routing control can be used as a means to reduce the total energy consumption of the network. The constraints that arise in this problem is that one has to remain aware of QoS considerations and network performance. We have proposed a method [3] that uses a queueing theoretic analysis and optimisation technique to distribute traffic so as to reduce a cost function that comprises both energy and QoS. Consider a network with N queues denoted $\mathbf{N} = \{1, \ldots, N\}$ which is carrying a set of *user traffic classes* $\mathbf{U}$. A subset $\mathbf{R}$ of these nodes will be the usual store and forward nodes or routers, while the remaining set of nodes $\mathbf{L}$ will represent links which connect store and forward nodes, i.e. $\mathbf{N} = \mathbf{R} \cup \mathbf{L}$. This separation of the N nodes into the set of routers and the set of links has the advantage that allows us to model separately the impact of routers and links on QoS and on the power consumption. In this model, apart from the usual user traffic we also have control traffic which may be seen either as physical flows of control packets that are sent out from certain decision nodes, to routers where it may be necessary to re-route traffic, or as a virtual and mathematical representation of the rate at which control decisions are made at this router. Apart from the default routing protocol's decisions probabilities P, the routing control probabilities are also included in the model denoted as Q.

The model that we propose [4] is a special case of G-Networks with triggered customer movement, where the control classes embody the triggers of the mathematical model [5], thus we are able to apply the corresponding theory. So, the steady-state probability that a router r or link l contains at least one packet of user class k is given by:

$$q(r,k) = \frac{\Lambda_R(r,k)}{\mu_r + \Lambda_R^-(r,(r,k))},\ if\ r \in \mathbf{R} \tag{1}$$

$$q(l,k) = \frac{\Lambda_L(l,k)}{\mu_l},\ if\ l \in \mathbf{L} \tag{2}$$

where μ_r and μ_l are the service rates of the routers and the links respectively. $\Lambda_R^-(r,(r,k))$ is the total arrival rate of control packets of class (r,k) to the node r where they are entitled to act. The total arrival rates of user packets of class k to routers $r \in \mathbf{R}$ and links $l \in \mathbf{L}$ are given by

$$\Lambda_R(r,k) = \lambda(r,k) + \sum_{l \in \mathbf{L}} q(l,k)P(l,k,r)\mu_l \tag{3}$$

$$\Lambda_L(l,k) = \sum_{r \in \mathbf{R}} [P(r,k,l)q(r,k)\mu_r + \Lambda_R^-(r,(r,k))q(r,k)Q(r,k,l)] \tag{4}$$

where $\lambda(r,k)$ is the external arrival rate of user packets of class k at router r , probabilities P represent the default routing probabilities while Q represent the control decisions exercised by the control packet. Similar formulas are obtained for the control traffic.

Using the traffic formulas we are able to get the queue lengths at each node and the total average delay through the network $\overline{T_N}$. The power consumption of the network P_N is comprised of the power consumption of each node and link, which is composed of a fixed value (the static power consumption) and an increasing function of the load. The routing optimization can now be expressed as the minimization of a cost function f that includes both the network power consumption and the average delay:

$$f = P_N + c\overline{T_N} \tag{5}$$

where c is a constant that establishes the relative importance of delay with respect to power. The minimisation can be achieved by selecting appropriate route control parameters $Q(i, k, j)$. Since we are interested in gradual optimisation in the presence of ongoing flows, we build a gradient descent algorithm that reduces the cost function at a given operating point of the network and is determined by its n^{th} computational step:

$$Q_{n+1}(i,k,j) = Q_n(i,k,j) - \eta \frac{\partial f}{\partial Q(i,k,j)}|_{Q(i,k,j)=Q_n(i,k,j)} \quad (6)$$

where $\eta > 0$ is the "rate" of the gradient descent and the partial derivative is computed with the nth updated values of the weights. For a N node network we show that the optimization can run in $O(N^3)$ time complexity. Note that the presented model and algorithm can also be used to optimize other performance metrics of the network and includes the effect of the control traffic on the performance [6].

A faster version of this algorithm [7],[8] limited to a single step of the gradient descent in order to provide fast computation has been evaluated using our 23 node topology. The power profiles were measured from the PC-based routers of the testbed and random delay weights were added to the links. The results in figure 1 compare the power consumption of the network using shortest paths and the power optimised routing ($c = 0$ in (5)), in the presence of 5 flows with distinct source-destination pairs. Varying the input traffic of the flows, the optimisation yields an average saving of 36 Watts, at the cost of an increase in average end-to-end delay of 4.4ms, but with maximum delay increase of 13ms. By

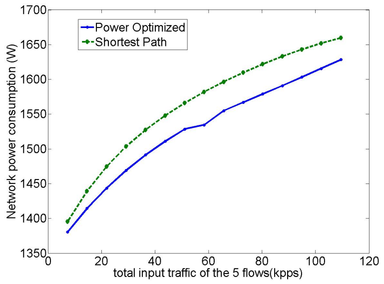

Figure 1: Power consumption against varying traffic load for power-optimised versus shortest path

adjusting c in (5) we can increase the relative importance of the parameters in the cost function and limit the increase in network delay as shown in figure 2. Note that the power

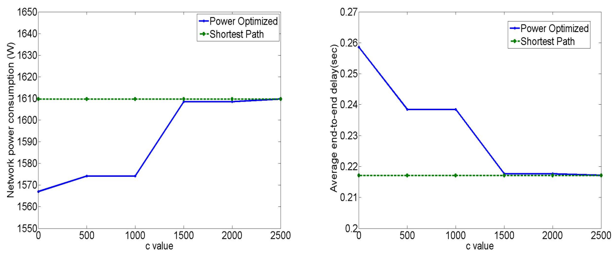

Figure 2: Power-delay trade-off varying c in (5)

profiles of the devices measured and used in the results have a large static power consumption, but still the savings come without turning off/on the machines. The savings of energy aware routing will be larger in a larger testbed and in case of more energy-proportional devices.

3. POWER AWARE ADMISSION CONTROL

We also examine the usage of an admission control for improving network energy efficiency. We have proposed an admission control mechanism [9] which aims to keep the power consumption at the lowest possible level by delaying users that would increase significantly the power consumption of the network, and admit them at a later time when their effect in power consumption will be acceptable. Thus, since some of the operation areas are more power efficient than others, by using admission control we could reduce the total network energy consumption. Some first experimental results show the effectiveness of the method and reveal room for potential energy savings, but these energy savings come at the expense of increased waiting delay for users before being admitted into the network. As further work, the effect of the waiting limitations of the flows as well as the possibility of turning off and on the machines are being examined.

4. CONCLUSIONS

There is increasing interest in reducing energy consumption for all areas of human activity, and ICT systems are major consumers. In this work we address the opportunities and propose solutions for energy savings in networks, through analytical modelling, optimization algorithms and experimental evaluation. We seek to integrate methods and implementation techniques so as to provide guidelines and practical means for energy optimization, while remaining aware of network performance and QoS considerations.

5. REFERENCES

[1] M. Webb. Smart 2020: Enabling the low carbon economy in the information age. Technical report, The Climate Group, London, 2008.

[2] A. Bianzino, C. Chaudet, D. Rossi, and J. Rougier. A survey of green networking research. *Communications Surveys Tutorials, IEEE*, PP(99):1 –18, 2010.

[3] E. Gelenbe and C. Morfopoulou. A framework for energy-aware routing in packet networks. *The Computer Journal*, 54(6):850–859, 2011.

[4] E. Gelenbe and C. Morfopoulou. Routing and G-networks to optimise energy and quality of service in packet networks. In *Energy-Efficient Computing and Networking*, volume 54 , pages 163 – 173. Springer 2011.

[5] E. Gelenbe. G-networks with instantaneous customer movement. *Journal of Applied Probability*, 30(3):742–748, 1993.

[6] C. Morfopoulou. Network routing control with G-networks. *Performance Evaluation*, 68(4):320 – 329, 2011.

[7] E. Gelenbe and C. Morfopoulou. Gradient optimisation for network power consumption. In *GreenNets 2011*.

[8] E. Gelenbe and C. Morfopoulou. Power savings in packet networks via optimised routing. *Mobile Networks and Applications*, 17:152–159, 2012.

[9] G. Sakellari, C. Morfopoulou, T. Mahmoodi, and E. Gelenbe. Using energy criteria to admit flows in a wired network. In *27th International Symposium on Computer and Information Sciences, (ISCIS)*, 2012.