Technical Report

UCAM-CL-TR-784
ISSN 1476-2986

Number 784

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Characterizing 10 Gbps network interface energy consumption

Ripduman Sohan, Andrew Rice,
Andrew W. Moore, Kieran Mansley

July 2010

15 JJ Thomson Avenue
Cambridge CB3 0FD
United Kingdom
phone +44 1223 763500

http://www.cl.cam.ac.uk/

Technical reports published by the University of Cambridge Computer Laboratory are freely available via the Internet:

http://www.cl.cam.ac.uk/techreports/

ISSN 1476-2986

Characterizing 10 Gbps Network Interface Energy Consumption

Ripduman Sohan, Andrew Rice, Andrew W. Moore
Computer Laboratory, University Of Cambridge
William Gates Building, 15 JJ Thomson Avenue
Cambridge, CB3 0FD, UK
Email: `firstname.lastname@cl.cam.ac.uk`

Kieran Mansley
Solarflare Communications
Westbrook Centre
Cambridge, CB4 1YG, UK
Email: `kmansley@solarflare.com`

***Abstract*—Understanding server energy consumption is fast becoming an area of interest given the increase in the per-machine energy footprint of modern servers and the increasing number of servers required to satisfy demand. In this paper we (*i*) quantify the energy overhead of the network subsystem in modern servers by measuring, reporting and analyzing power consumption in six 10 Gbps and four 1 Gbps interconnects at a fine-grained level; (*ii*) introduce two metrics for calculating the energy efficiency of a network interface from the perspective of network throughput and host CPU usage; (*iii*) compare the efficiency of multiport 1 Gbps interconnects as an alternative to 10 Gbps interconnects; and (*iv*) conclude by offering recommendations for improving network energy efficiency for system deployment and network interface designers.**

1. Introduction

Our dependency on information technology in our daily lives has led to computing infrastructure becoming a significant consumer of energy. A study commissioned in Japan in 2006, for example, showed that communications and computing infrastructure accounted for 4% of all total electricity production [1], while in the USA and UK it has been shown that communications and computing infrastructure account for 3% [2] and 10% [3] of countrywide electrical energy consumption respectively.

Of the energy consumed by computing infrastructure, a significant amount is consumed within servers; recent studies have estimated this figure as approximately 1.5% of all power consumption in the USA [4]. With the continual growth in both the size and number of servers required to service ever-increasing demand, it is important to optimize and minimize server energy usage.

Communication is a fundamental function of the modern server; the energy efficiency of any server is intrinsically linked to how *quickly* and *efficiently* data can be moved between it and other devices. Considering that the amount of data being transmitted is continually increasing both over the Internet and private networks [5] it follows that a power efficient network subsystem can result in significant runtime energy cost savings [6].

There is also great emphasis on increasing the power usage effectiveness of large scale datacenters by reducing the power consumed providing support functions such as power distribution and cooling [7]. As the energy optimisations associated with support functions reach physical limits, it is likely that, in the future, better power usage efficiency will depend on reducing the power consumption of servers.

An important first step in optimizing energy consumption is quantifying its use. In this work, we set out to examine the energy efficiency of 10 Gbps (10G) server interconnects. In particular, we make the following contributions: (*i*) we measure and characterize the idle and active power consumption for a number of production 10 Gbps Network Interface Cards (NICs) of varying makes, models, architectures and utilizing different physical media; (*ii*) we compare their energy efficiency from a throughput and host CPU utilisation perspective; (*iii*) we outline the absolute energy efficiency of all the measured NICs and (*iv*) we compare the cost and power efficiency of 10G NICs to single, dual and quad port 1 Gbps (1G) configurations.

The remainder of this paper is structured as follows: Section 2 outlines our measurement hardware, software, infrastructure and methodology. Section 3 details our measurement results and analysis in the areas of idle (Section 3.1) and active (Section 3.2) energy efficiency while Section 3.3 compares the energy efficiency of 1G and 10G configurations. Section 4 presents an analysis of the absolute energy efficiency of the measured NICs. Section 5 speculates on how system designers and NIC designers may be able to improve energy efficiency. Section 6 discusses related work and Section 7 concludes.

2. Measurement Platform

The results in this paper are derived from independent measurements conducted by the authors. There are three reasons we chose to conduct independent measurements: (*i*) our measurements serve to validate those of the manufacturer; (*ii*) carrying out our own measurements enables us to instrument at a finer level and according to our specific requirements and, most importantly, (*iii*) using standardized measurement infrastructure and methodology enables us to compare our results across different NICs. The remainder of this section describes our NIC test set, testbed, measurement platform and methodology in detail.

NIC	Link Rate (Gbps)	Physical Medium	Part Number
Solarflare(Fibre)	10	Fibre	SFE4002
Solarflare(Base-T)	10	Base-T	SFE4001
Solarflare(CX4)	10	CX4	SFE4003
Broadcom(Fibre)	10	Fibre	PE10G2T-SR
Intel(CX4)	10	CX4	PE10G2I-CX4
Intel(Base-T)	10	Base-T	PE10G1-T
Intel 1G	1	Base-T	EXPI9400PT
Broadcom Multiport(2x1G)	1	Base-T	NC380T
Intel Multiport(2x1G)	1	Base-T	EXPI9402PT
Intel Multiport(4x1G)	1	Base-T	PEG4I-RoHS

Table 1: NIC Test Set

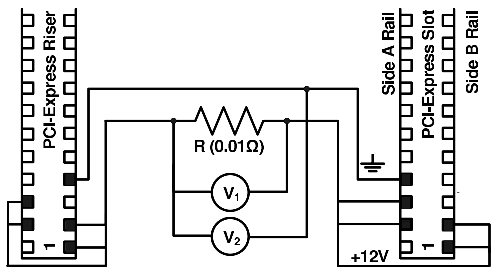

Figure 1: 12v Power Measurement Apparatus

2.1. NICs

Table 1 lists all the NICs measured in this work. We measured six production 10G NICs from four manufacturers and an additional four 1G NICs for the 1G-10G comparison discussed in Section 3.3. For verification purposes, we provide the part numbers of all measured NICs. All the Solarflare NICs are second generation devices based on an identical reference design with differences due only to adapting the NIC for different physical media.

The Broadcom(Fibre) NIC (version 1.5) is manufactured by Silicom. The Intel(CX4) and Intel(Base-T) NICS are both based on the Intel 82598EB and are manufactured by Silicom. However, the Intel(Base-T) NIC (version 1.4) has a physical layer manufactured by Teranetics.

For the 1G NIC set, the Intel 1G is based on the Intel 82572GI Gigabit Controller. The Broadcom Multiport(2x1G) NIC is manufactured by HP and based on a pair of BCM5706 CKFBG controllers. The Intel Multiport(2x1G) is based on the Intel 82571GB chipset while the Intel Multiport(4x1G) is based on the Intel 82571EB chipset but is manufactured by Silicom.

The 10G measurements span the most common physical media types: CX4 (IEEE standard 802.3ak), short range fibre (IEEE standard 802.3ae) and Base-T (IEEE standard 802.3an). This is of interest because there is a clear tradeoff between the cost of the NIC and the physical media: CX4 is a simple, low power copper wire standard designed to connect over short distances of up to 15 meters. The simplicity of the standard means the physical layer of the NIC is cheap to implement, however, the interconnect cables are complex and expensive to manufacture. Base-T is able to utilize existing cheap twisted pair cabling, however the signal processing overheads at 10G result in complicated NIC physical layer designs. Finally, fibre is a relatively cheap interconnect but mandates the use of expensive transceivers for data transmission.

2.2. Hardware and Software

All measurements were taken on a pair of SuperMicro machines consisting of an 6025W-NTR+B server board based on the Intel 5400 chipset, equipped with two Xeon 5482 dual die 3.20 GHz quad core CPUs for a total of 8 logical processors. Every core has 32KB of level one data cache and every die has 6MB of shared level two cache. The system was equipped with 4GB of RAM on a quad-pumped 1600 MHz memory bus. NICs interface with the host through a PCI-Express (version 2) bus and are connected via an 8 lane connector theoretically capable of sustaining a 8GB/s transfer rate.

For the duration of the measurements the operating system used was Windows Server 2008/Enterprise running in 32bit mode. Every NIC was measured using the latest drivers available (as provided on the product support website) at the time of measurement. Ethernet frame size was 1500 bytes. We used the IXIA Chariot [8] tool to generate realistic traffic streams when taking measurements that required the NIC to be active (i.e. transferring data).

2.3. Measurement Apparatus & Methodology

We measure energy consumption by measuring the power used by the NIC. PCI-Express connectors provide voltage at two levels, 3.3v and 12v. By intercepting both voltage supply lines we are able to determine the current (and, by extension, power) used by the device.

Figure 1 illustrates one half of the measurement apparatus in detail. PCI-Express connectors supply a (single sourced) 12v voltage on pins 2,3 of the Side A rail and 1,2 on the Side B rail of the connector. We intercept these pins and common them, feeding the resulting line through a 0.01Ω series resistor, R, before re-splitting the line to feed identical pins on a riser card into which the NIC is fitted.

We use this apparatus to calculate the power consumed in the 12v circuit as follows: Using Ohm's law we are able to determine the current in the circuit, I, by measuring the potential difference across R using voltmeter V_1. As the current in the circuit is constant, it follows that the power being consumed in the circuit may be calculated as the product of I and the potential difference across the entire circuit[1] as measured by voltmeter V_2. [2]

A similar setup forms the other half of the measurement apparatus by binding pins 9,10 of the Side A rail and 8,10 of the Side B rail thereby enabling the calculation of power drawn on the 3.3v circuit. Some of our analysis required measurement of whole server power consumption. For this measurement we used two standard off-the-shelf digital power meters with a resolution of 0.1W.

The results reported in this paper are the average of at least three independent measurements. All related measurements are

NIC	Offload	Media	Idle Power (W) 3.3v	12v	Total
Intel(Base-T)	No	Base-T	6.0	15.2	21.2
Solarflare(Base-T)	No	Base-T	1.0	17.0	18.0
Broadcom(Fibre)	Yes	Fibre	5.9	7.2	13.1
Solarflare(Fibre)	No	Fibre	2.6	3.1	5.7
Intel(CX4)	No	CX4	5.6	0.0	5.6
Solarflare(CX4)	No	CX4	1.6	3.0	4.6

Table 2: 10 Gbps NICs - Idle Power Consumption

NIC	Link Speed	Media	Number Of Active Links	Idle (W) Power
Broadcom(Fibre)	10 Gbps	Fibre	0	11.1
			1	12.1
			2	13.1
Intel Multiport(4x1G)	1 Gbps	Base-T	0	7.9
			1	9.0
			2	10.1
			3	11.1
			4	12.3

Table 3: Multiport NICs - Idle Power Consumption

verified to be within 3% of one other. All results are rounded up to one decimal place.

3. Characterizing Energy Consumption

In characterizing the energy consumption of the NICs we focus on three areas: we begin by analyzing the idle energy consumption of the NICs (Section 3.1), followed by an analysis of active or in-use energy efficiency (Section 3.2). Finally we conclude by comparing the runtime energy costs and power efficiency of 1G and 10G NIC deployments (Section 3.3) in servers.

3.1. Idle Energy Consumption

Idle energy is defined as the energy consumed by the card when powered, with all links connected (and operating system driver loaded) but not transferring any data. In practice it is the least amount of energy required to keep the card functional. Table 2 lists the idle power profiles of the 10G NICs in our test set. Our measurements lead us to make the following observations:

3.1.1. NICs may contribute significantly to server energy consumption: Typical modern servers have a baseline power draw of between 150–250W depending on hardware configuration. The measured NICs, on the other hand, show a power consumption of between 5–20W. Thus, the addition of a 10G NIC adds between 2.0–13.3% on baseline power consumption.

While NIC power consumption may seem insignificant on first glance, it is high enough that we consider it worth factoring in when designing large server farms. For example, the difference in idle power consumption between the most (Intel(Base-T)) and least (Solarflare(CX4)) expensive 10G NICs listed in Table 2 is 16.6W. This equates to an increased running cost of $14 per-annum[3] for the Intel(Base-T) device compared to the Solarflare(CX4). For a datacenter of 1000 machines, this results in an additional cost of $14,000 per year – a figure large enough to warrant careful consideration of which 10G interconnect should be used in the servers. This issue may be compounded even further by high throughput applications (e.g. video processing) which require multiple 10G interconnects.

3.1.2. Physical media influences power consumption: As Table 2 shows, there is an order of magnitude difference in the idle power consumed by all the NICs in the test set. Various reasons may account for this difference, most significantly the internal design of the NIC and the CMOS processing technology may significantly influence power draw.

To determine the power consumption attributable to adaptation for the physical layer we focus on the Solarflare series of NICs. As explained in Section 2.1 all the Solarflare NICs measured in this paper are based on an identical internal design and manufactured using the same CMOS processing technology. Discussion with the manufacturer revealed that while there are differences in the circuitry and internal firmware in the three variations measured, the changes are mostly minor bug fixes which have no impact on power consumption. The only major differences in the design of the measured NICs are due to adaptation for the physical layer.

The results highlight that the CX4 variation has the lowest power consumption due to the simple and straightforward wire-like design of the CX4 physical protocol. This is followed closely by the Fibre variation which consumes an additional watt due to the transceiver (as explained in Section 3.1.4). Finally, the Base-T variation consumes the most energy due to the power consumed in the signal processing component of the card which is responsible for generating the pulse-amplitude-modulated waveform in the physical media.

While our physical media analysis are based on the Solarflare NICs, results in Table 2 verify our claims. In general, for all measured cards CX4 devices consume the least energy followed by fibre and Base-T variations respectively.

3.1.3. Offload is more power expensive: A common design optimization involves offloading network processing onto the NIC for the purposes of increased performance or reduced host CPU usage. It is commonly expected that the increased functionality and complexity of offload NICs will result in devices that have a significantly larger power footprint than more conventional designs.

While our NIC test set only includes a single offload device (Broadcom(Fibre)), our measurements confirm expectations. This device has an order of magnitude larger power draw than any other NIC adapted for CX4 or Fibre. The increased power consumption is due primarily to relatively high power usage in the 12v circuit. This is attributable to the CPU and RAM on the NIC which continue to draw power even when the NIC is idle.

3.1.4. Link connection status has little effect on power consumption: Multiport NICs composed of multiple physical links on the same device are becoming increasingly popular due to the need for increased server network capacity. Multiport NICs are preferred over single-port NICs due to the economic (they are cheaper per-port), space (they only require a single PCI-Express slot) and management (they only require a single driver) savings they offer.

We set out to measure the power consumption of multiport NICs with respect to link connection status. Specifically, we quantified device power consumption in relation to the number of active links. We tested by physically removing the transceiver in the case of Fibre and disconnecting the link in the case of Base-T. We measured the power consumption for all (1G and 10G) NICs in our test set and observed that connection state has very little impact on the overall idle power consumption of the device.

Table 3 illustrates the results for the Broadcom(Fibre) and Intel Multiport(4x1G) devices. As the table shows, for both the 10G Fibre and 1G Base-T devices, link connection only marginally increases power consumption (approximately 1W). For the sake of brevity we omit reporting the results of the other multiport NICs in the test set. However, we verify that we observed similar results in all cases.

Our measured results indicate that between 40–85% (Intel Multiport(2x1G) and Intel Multiport(4x1G) respectively) of the overall power consumed by multiport NICs is attributable to the system electronics and remains constant regardless of the number of in-use links.

3.2. Active Energy Efficiency

This section studies the energy efficiency of the 10G NICs in the measurement test set. We present results listing the active energy consumption of the NICs and analyze the results to determine the most energy efficient NICs with respect to throughput and host CPU used.

3.2.1. Active Energy Consumption: Active power consumption is obtained by measuring the NIC power usage while transferring data over 5 bidirectional TCP streams. Table 4 lists the results for the active power consumption of the 10G NICs in the test set coupled with the host CPU required to sustain the maximum achievable throughput. The total amount of host CPU available in the system is 800%, defined as 8 logical processors each of which can be fully dedicated to the experiment.

There is very little difference in the power usage of an active NIC compared to an idle one. For all measured NICs the difference in power usage is less than 1W with the largest delta being only 0.9W (Broadcom(Fibre)). This leads us to conclude that very little energy is required to actually transmit data and that the majority of the energy is expended in powering the NIC system electronics.

Finally, the results also show that throughput performance varies tremendously across the measurement set (ranging from 11–18.7 Gbps) . However, there is no correlation between power usage and performance – some low performing NICs have a high power draw while other higher performing NICs have a low power draw.

3.2.2. NIC Performance Per Watt: For any set of NICs able to sustain a required level of performance, the most power efficient can be defined as the one that is able to provide the most performance for the least amount of energy consumed. Using this requirement, we define the performance per watt of a NIC as the throughput in Gbps per watt of energy consumed.

We analyzed all the 10G NICs in our test for the purposes of determining NIC performance per watt. Figure 2 provides the results. As the figure indicates, the best performance is provided by the Solarflare(CX4) due to its high throughput and low power footprint. This is followed by the Solarflare(Fibre) which has near identical performance to the CX4 variation of the NIC but consumes 1W more of power in the physical layer due to the fibre transceiver (Section 3.1.4).

While the Broadcom(Fibre) has the best throughput performance of all measured NICs, it fares poorly from a performance per watt perspective due to the high energy consumption of the offload engine on the NIC. Unsurprisingly, the Base-T NICs have the lowest performance in the measured set due to their high power overhead at the physical layer.

3.2.3. Server Performance Per Watt: Conventionally, all data transferred through the NIC is subject to processing in the host operating system's network layer. As has been measured previously, this processing requires substantial amounts of host CPU, especially for high speed links [9]. High host CPU usage may be problematic as it means less CPU is available to service applications.

As stated previously, high host CPU usage has inspired the development of offload NIC designs [10] which move some or all network processing onto the network card for the purpose of reducing host CPU utilization. However, Section 3.1.3 has also highlighted that offload designs have higher energy consumption.

There is clearly a tradeoff between the throughput performance of the NIC, the amount of power it consumes and the amount of host CPU used to service the network interconnect. An ideal NIC will provide high throughput, use little power and consume a minimum amount of host CPU.

Given a set of NICs that can be serviced within a maximum threshold of host CPU dedicated to network processing, an administrator will likely select the one able to provide the best performance for the least power *and* host CPU consumption. However, correlating NIC power consumption, throughput and host CPU consumption is non-trivial; all three parameters are independent variables as listed in Table 4.

We introduce *server* performance per watt as a simple metric that enables reasoning about NIC host CPU consumption. Server performance per watt is defined as the throughput obtained per watt of *server* energy consumed. It is based on our observations that: (*i*) an idle powered server has a constant power draw and (*ii*) server power consumption increases in proportion to CPU load[4]. In effect, this metric incorporates the utilisation of host CPU for servicing the network. If the NIC requires a large amount of host CPU server power consumption increases and server performance per watt reduces.

We analyzed all the 10G NICs in our test set to determine

NIC	Active Power (W)	Throughput (Gbps)	CPU Usage(%)
Intel(Base-T)	21.4	11.0	369.6
Solarflare(Base-T)	18.2	15.8	508.3
Broadcom(Fibre)	14.0	18.7	264.7
Solarflare(Fibre)	5.9	15.9	508.3
Intel(CX4)	5.6	10.3	302.3
Solarflare(CX4)	4.9	16.5	484.4

Table 4: Measured 10 Gbps NICs - Active Power, Throughput And CPU Usage

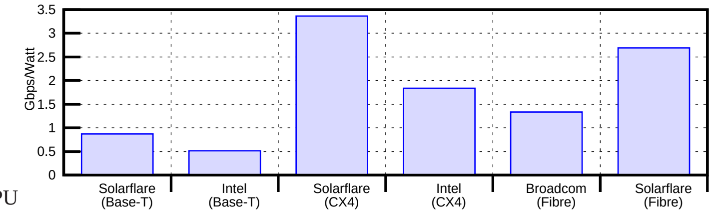

Figure 2: 10 Gbps NICs – NIC Performance Per Watt

server performance per watt. Figure 3 presents the results. The Broadcom(Fibre) NIC is the most efficient from a whole-server perspective due to its low CPU load in relation to the extra power consumed by its offload engine. This is followed by the Solarflare NICs which have a better server performance per watt result than the Intel NICs in spite of consuming more CPU due to their higher throughput characteristics. Finally the Intel NICs have the lowest performance overall due to their low throughput. The server performance per watt results are also interesting as they show that, from a system perspective, overall server energy efficiency is still dominated by host CPU utilisation.

Note however that CPU manufacturers are employing more and more aggressive optimizations to reduce per-core power consumption with the aim of building more concurrent (a larger number of cores) CPUs. Hence, it is possible that within 3–5 years server power consumption will no longer be dominated by the CPU but by other components.

3.3. Multiport 1G vs 10G

System designers seeking a desired level of throughput performance have the choice of using multiple or multiport 1G NICs or a single 10G NIC. Provided a set of NICs able to sustain a required level or performance, it is in the interest of the designer to choose the one that provides the best performance for the least cost.

In this section we compare a number of single and multiport (dual and quad) 1G configurations with the 10G NICs in our test set in order to determine those that provide the best performance-to-power ratios. We focus on NICs adapted for the Base-T physical layer as this is the most prevalent wiring infrastructure in modern datacenters. Table 5 presents the characteristics of our measured 1G NICs. The results lead to the following observations:

3.3.1. Throughput efficiency decreases as the number of ports increase: Our measurements show that achievable throughput does not scale in relation to the number of ports. While it is unlikely that any NIC will achieve its theoretical throughput (due to host and protocol overheads), we found that the single port NIC is able to achieve 85% of theoretical bandwidth, the dual ports are able to achieve 82.5% of theoretical bandwidth but the quad port device is only able to achieve 70% of theoretical bandwidth. In comparison our 10G NICs are able to achieve up to 93.5% (Broadcom(Fibre)) of theoretical bandwidth.

While a detailed explanation of this drop in throughput efficiency would require finer instrumentation and measurement, it is likely to be due to an increase in the overheads (e.g. interrupts, context switching, bus contention) associated with transferring data over multiple physical links.

3.3.2. Power consumption increases in correlation to the number of ports: As Table 5 illustrates, the power footprint of the multiport NICs increases in relation to the number of ports on the device. Focusing on the Intel single and dual port NICs (chosen as devices from the same manufacturer are likely to contain common design elements and be implemented using similar technology), we notice that the the average active power[5] consumed per port remains approximately the same (1.8–1.9W) for the single and dual port variations. Furthermore, power consumption actually increases to 3.125W for the quad port NIC. However, this increase is likely to be due to the fact that the quad port NIC is manufactured by Silicom and thus uses a different physical layer implementation to the single and dual port NICs

While confirmation would require detailed instrumentation and measurement, power consumption measurements suggest that there is little or no electronic integration or scaling on the device. Visual inspection of the NICs and controller datasheets confirm a single controller but physical power draw seems to suggest a duplication of functionality (and associated electronics) in a single packaging. In some cases (e.g. Broadcom Multiport(2x1G)) the multiport NIC is actually composed of multiple 1G NICs coupled on the same printed circuit board. From a technical perspective, the only advantage of using 1G multiport devices in comparison to single port NICs is the PCI-Express slot savings efficiency.

3.3.3. 1G NICs Possess Similar NIC Performance Per Watt Characteristics as 10G NICs: Next, we evaluated the efficiency of the multiport 1G NICs by calculating their efficiency in terms of NIC performance per watt. Figure 4 provides the results of this analysis. As illustrated, the relatively low power consumption and high throughput achieved by the Intel Multiport(2x1G) NIC ensures it has the highest performance per watt of the measured set. This is followed closely by the Intel 1G and then the Solarflare(Base-T). The Broadcom Multiport(2x1G) and Intel Multiport(2x1G) both have much lower performance per watt due to their low throughput and high power draw.

While the NIC performance per watt metric provides a simple, efficient and abstract mechanism for comparing the power

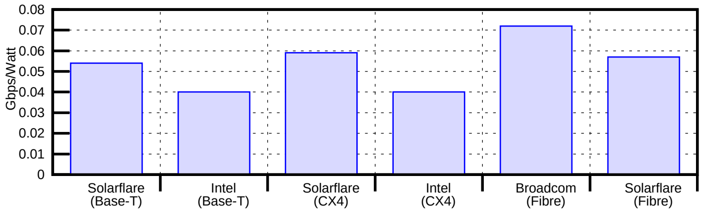

Figure 3: 10 Gbps NICS – Server Performance Per Watt

NIC	Media	Throughput (Gbps) Theoretical	Throughput (Gbps) Actual	Active Power (W)
Intel 1G	Base-T	2	1.7	1.9
Broadcom Multiport(2x1G)	Base-T	4	3.3	7.0
Intel Multiport(2x1G)	Base-T	4	3.3	3.6
Intel Multiport(4x1G)	Base-T	8	5.7	12.5

Table 5: 1G NICs - Performance And Power Characteristics

efficiency of different NICs it is important to note that practical factors may influence the range of available choices. For example, even though the Intel Multiport(2x1G) device has the best performance per watt, delivering throughput approaching 10 Gbps with this NIC configuration requires a motherboard with five PCI-Express slots (most only contain one or two).

Similarly, this analysis does not account for host CPU consumed servicing the network. However the per-packet processing overheads associated with multiple slower (compared to 10G) links may lead to inordinately high host CPU requirements. For example, we measured the host CPU required to service the Intel 1G and Intel Multiport(2x1G) links and extrapolated that it would require 1085% and 759% of host CPU respectively to service throughput equivalent to that provided by the Solarflare(Base-T). In comparison the Solarflare NIC only requires 508.3% of host CPU.

In summary while the performance per watt offered by 1G and 10G NICs is similar, practical issues concerning physical capacity and the amount of host CPU required to service the network render the 10G NICs the most sensible choice for the majority of configurations.

4. Contextualizing Energy Efficiency

In this section we compare the energy efficiency of the NICs in the test set using an absolute energy efficiency metric. This is useful because it is allows us to objectively compare the overall energy consumption of the devices, their relative efficiency (with respect to each other) and it provides indications as to the energy efficiency of the devices with respect to theoretical lower and upper bounds.

Our analysis utilizes the absolute energy efficiency metric defined by Parker et al. for network energy efficiency [11]. This work defines the logarithmic unit dBε, which allows comparison across different network technologies and architectures. Moreover, as it is based on a physical constant the measure is transparent, transportable and scalable.

While full details on the background, assumptions and design of the metric are available in the papers that introduces it [12], [11], we note that the absolute energy efficiency in dBε is calculated as:

$$dB\epsilon = 10log_{10}\left(\frac{Power/BitRate}{kTln2}\right) \quad (1)$$

Here, Power is the power consumed in Watts, Bit Rate is the data rate in bits per second, k is the Boltzmann constant (1.381×10^{-23} Joules/Kelvin) and T is the absolute temperature in Kelvin.

Table 6 presents the absolute energy efficiency results of all the NICs in our test set (calculated using a value of 300K for T). While the relative absolute efficiency values for the NICs in the test set loosely mirror performance per watt, the results show that there is almost an order of magnitude difference between the absolute energy efficiency of the most (Solarflare(CX4)) and least (Intel Multiport(4x1G)) efficient NICs. The results also show that, generally, the 10G NICs and 1G NICs (as a group) have similar absolute energy efficiencies. However, it is interesting to note that while the 10G NICs consume more energy, they are approximately 5 times more energy efficient than the 1G NICs.

Comparing the calculated values in Table 6 with similar results calculated in the work that defines the metric [11], we find that per-bit transported, the most efficient 10G NIC in the test set has an absolute energy efficiency figure that is 8 times more efficient than the most-efficient CPU they measured (119.9dBε). Furthermore, we find that, generally, the 10G NICs in our test set are more efficient (111-118dBε) than the varying networking and computing equipment tested by the authors of the metric (115-130dBε).

5. Towards Increased Energy Efficiency

In this section we discuss, based on observed results, how system designers can ensure maximum efficiency of deployed systems and speculate on optimisations that would be useful for increased NIC efficiency.

5.1. Ensuring Deployed Systems Are Efficient

System designers deploying large scale systems should consider the cost of deploying physical media in relation to the running cost of the NIC over the time span of the deployment. While common, cheap media such as Base-T has a lower deployment cost, Section 3.1.1 showed the running cost of the NIC is higher due to its larger power draw. Designers should also account for rising or falling trends in energy costs with time. Finally, designers should also account for the ever-decreasing cost and power footprint of NICs as the technology matures as this will influence the running cost of the system when machines begin to get replaced at the end of their deployment cycle.

As NIC power consumption is approximately constant regardless of load, system designers should design and appropriate the system to maximize link utilization considering other

NIC	Link Rate (Gbps)	Absolute Energy Efficiency (dBε)
Solarflare(CX4)	10	110.1
Solarflare(Fibre)	10	111.1
Intel(CX4)	10	112.8
Broadcom(Fibre)	10	114.1
Intel Multiport(2x1G)	1	115.8
Intel 1G	1	115.9
Solarflare(Base-T)	10	116.0
Intel(Base-T)	10	118.3
Broadcom Multiport(2x1G)	1	118.7
Intel Multiport(4x1G)	1	118.8

Table 6: Absolute Energy Efficiency 1 and 10G NICs

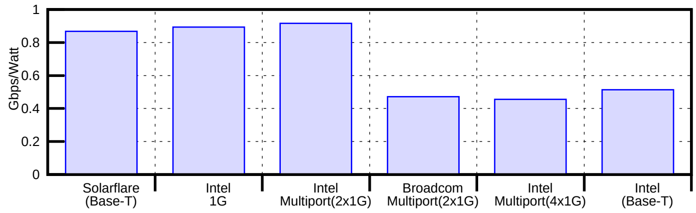

Figure 4: 1 vs 10 Gbps - Performance-Per-Watt

resources (e.g. CPU or RAM) are not an issue. This leads to increased NIC power efficiency, which, in turn increases the overall efficiency of the system. Similarly, utilising all ports on a multiport NIC results in increased efficiency as the power cost is ammortized over a larger number of active ports as explained in Section 3.1.4.

System designers should consider whether throughput requirements are likely to increase with time. If throughput requirements are constant over the duration of the deployment, it may be more cost efficient to deploy multiport NICs due to the lower initial purchase cost (an efficient multiport NIC provides the same performance per watt as a good 10G NIC). However, if throughput requirements are likely to increase with time, the practical advantages of a 10G interface (Section 3.3.3) may result in a cheaper long-term solution.

5.2. Potential NIC Optimisations

Our measured results indicate that there are a number of potential NIC design optimisations that are likely to increase power efficiency. Section 3.1 showed that the NIC power consumption is approximately the same in idle mode as in active mode. One area of optimsation lies in reducing idle power consumption to reduce the energy footprint of an unused NIC.

Similarly, optimising NIC design so power draw increases in proportion to link utilisation (i.e. a lower transfer bit rate results in a lower energy consumption) would ensure that the NIC operates at optimum energy efficiency, regardless of load (some preliminary work on this issue has already been carried out by Popa et al. [13]).

Multiport designs benefit from power efficiencies achievable by electronic integration or scaling, e.g. using a single controller to manage all physical ports. Similarly, system logic to turn off the channel circuit when a port is disconnected would result in power savings for multiport NICs that are not in full use.

6. Related Work

To the best of our knowledge this is the first academic work that provides detailed measurements and comparative analysis detailing the power consumption and tradeoffs in modern 1G and 10G network interface cards over a range of design, manufacturer and physical media types. Previous work, however, has studied the power consumption of server, desktop and portable devices in various contexts as outlined below:

Stemm and Katz carried out one of the earliest studies on characterizing energy consumption in wireless NICs in hand-helds [14]. Like this work they measured the voltage and current drop across a small resistor for the purposes of determining power consumption. They show that the wireless interface accounts for a large proportion of the total power used in the hand-helds, and, similar to this work, the idle state dominates power consumption.

Ebert et al. measure the power consumption of wireless LAN interfaces for various operational and parameter settings for non-impaired radio frequency channels using a similar measurement setup to ours [15]. They derive the energy needed to transmit one bit of payload and build a simulation model of the interface, using it to investigate the effect of changing different operational parameters. Results show control of the power level and data rate adaptation make the biggest impact on energy efficiency.

Chandra uses the complementary technique of correlating network activity and published power information to build a detailed state model that is able to estimate the energy consumed for any sequence of traffic events [16]. While less accurate than direct measurement, this approach provides a simple and straightforward mechanism for estimating interface power consumption. Similarly, Hiaro et al. create whole system state models of the entire machine using published power consumption information of the major components [17].

Other work has examined the energy cost of the TCP networking protocol. Wang and Singh examine node level energy cost of TCP by measuring the power consumed as data moves through the networking stack [18]. Their results show that 60 – 70% of the energy cost is accountable in transferring data between the kernel and the NIC while 15% is due to user-kernel copy. Only 15% is due to TCP processing cost.

We note that fine-grained power measurement for the purposes of profiling energy consumption is being used to characterize other server components as well. Hylick et al. provide detailed measurements on disk drive power consumption [19]. Similarly, previous work has also profiled whole system power consumption [20], [21], [22].

Finally, the IEEE Energy Efficient Ethernet (IEEE standard 802.3az) task force [23] has been working on reducing the power consumption of 100 MB, 1G and 10G NICs by specifying mechanisms to put devices into sleep mode, stepping down link speeds in periods of low link utilisation and reducing transmit voltage for Base-T interfaces. This is an important development that should lead to increased energy efficiency and would make an excellent basis for a followup study. However, it will be some time before devices implementing the standard are commercially available.

7. Conclusions

This paper measured and analysed the power consumption of six 10 Gbps and four multiport 1 Gbps NICs spanning a range of design, manufacturer and physical media types. Our results found that, generally, 10 Gbps NICs consume between 4.5–20W of power depending on design and physical transmission media while 1 Gbps NICs consume between 2–13W. Furthermore, there is very little difference in the power consumption of an idle or loaded NIC. Higher link speeds have high host CPU requirements (between 250–500% CPU). Finally, the work determined that the current generation of 10 Gbps NICs are able to match mature 1 Gbps NICs in performance per watt energy efficiency.

Acknowledgements

The authors would like to extend their thanks to David Clarke and Steve Pope of Solarflare for technical assistance and Andy Hopper for financial and logistical support.

References

[1] M. Etoh, T. Ohya, and Y. Nakayama, "Energy consumption issues on mobile network systems," pp. 365–368, IEEE Computer Society, Aug 2008.

[2] K. Kawamoto, J. Koomey, B. Nordman, R. Brown, M. Piette, M. Ting, and A. Meier, "Electricity used by office equipment and network equipment in the U.S.: Detailed report and appendices," Tech. Rep. LBNL-45917, Energy Analysis Department, Lawrence Berkeley National Laboratory, Feb 2001.

[3] "Green evangelist to call for big changes in computer use to aid environment." http://www.itwales.com/997539.htm. Retrieved 1st April 2010.

[4] J. G. Koomey, "Estimating total power consumption by servers in the U.S. and the world," tech. rep., Analytics Press, Feb. 2007.

[5] A. Odlyzko, "Internet traffic growth: sources and implications," in *Proceedings of SPIE (Optical Transmission Systems and Equipment for WDM Networking II)*, vol. 5247, pp. 1–15, Oct. 2003.

[6] K. Christensen, B. Nordman, and R. Brown, "Power management in networked devices," *IEEE Computer*, vol. 37, pp. 91–93, Aug 2004.

[7] "Going Green at Google — Clean Energy Initiatives." http://www.google.com/corporate/green/datacenters/measuring.html/. Retrieved 1st April 2010.

[8] "IxChariot - Test Device and System Performance under Real-World Load Conditions." http://www.ixchariot.com/. Retrieved 1st April 2010.

[9] A. Gallatin, J. Chase, and K. Yocum, "Trapeze/IP: TCP/IP at near-gigabit speeds," in *Proceedings of the 1999 USENIX Annual Technical Conference, FREENIX Track*, pp. 109–120, USENIX Association, Jun 1999.

[10] J. C. Mogul, "TCP offload is a dumb idea whose time has come," in *Proceedings of the 9th Workshop on Hot Topics in Operating Systems*, pp. 25–30, USENIX, May 2003.

[11] M. Parker, S. Nagraj, and S. Walker, "Absolute energy efficiency metric for carbon footprint resource management and network optimisation," in *The 2009 International Conference On Clean Electrical Power*, pp. 111–116, Jun 2009.

[12] M. Parker and S. Walker, "Differential temperature Carnot heat analysis shows that computing machines are thermodynamically irreversible," *Optics Communications*, vol. 281, pp. 3440–3446, Jul 2008.

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

[14] M. Stemm and R. Katz, "Measuring and reducing energy consumption of network interfaces in hand-held devices," *IEICE Transactions on Communications*, vol. 80, pp. 1125–1131, Aug 1997.

[15] J. Ebert, S. Aier, G. Kofahl, A. Becker, B. Burns, and A. Wolisz, "Measurement and simulation of the energy consumption of an WLAN interface," Tech. Rep. TKN-02-010, Technical University of Berlin, Telecommunications Networks Group, Jun 2002.

[16] S. Chandra, "Wireless network interface energy consumption," *Multimedia Systems*, vol. 9, pp. 185–201, Aug 2003.

[17] E. Hirao, S. Miyamoto, M. Hasegawa, and H. Harada, "Power Consumption Monitoring System for Personal Computers by Analyzing Their Operating States," in *Proceedings of the 4th International Symposium On Environmentally Conscious Design and Inverse Manufacturing*, pp. 268–272, Dec 2005.

[18] B. Wang and S. Singh, "Computational energy cost of TCP," in *The 23rd International IEEE Conference on Computer Communications*, vol. 2, pp. 785–795, Mar 2004.

[19] A. Hylick, R. Sohan, A. Rice, and B. Jones, "An analysis of hard drive energy consumption," in *IEEE International Symposium on Modelling, Analysis and Simulation of Computers and Telecommunication Systems, 2008*, pp. 1–10, Sep 2008.

[20] J. Lorch, "A complete picture of the energy consumption of a portable computer," Master's thesis, University of California at Berkeley, Faculty of Computer Science, 1995.

[21] A. Mahesri and V. Vardhan, "Power consumption breakdown on a modern laptop," vol. 3471 of *Lecture Notes in Computer Science*, pp. 165–180, Springer, Dec 2004.

[22] M. A. Viredaz and D. A. Wallach, "Power evaluation of a handheld computer," *IEEE Micro*, vol. 23, Jan 2003.

[23] "IEEE P802.3az Energy Efficient Ethernet Task Force." http://grouper.ieee.org/groups/802/3/az/public/index.html. Retrieved 1st August 2010.

[24] U.S. Energy Information Administration, "Average Retail Price Of Electricity to Ultimate Customers by End-Use Sector, by State." http://www.eia.doe.gov/cneaf/electricity/epm/table5_6_a.html. Retrieved 13th April 2010.

Notes

[1] Any ground pin may be used.

[2] While an absolutely accurate calculation for power will disregard the voltage consumed by the resistor (i.e. $(V_2 - V_1) \times R$), in practice the low resistance of R means the power drawn by the resistor is negligible and may be ignored for the purpose of simplicity

[3] Calculated using the average commercial US energy price index for electricity [24] as of January 2010

[4] We assume little or no disk activity as the workload is CPU bound

[5] Similar to the 10G NICs, active power is only marginally larger than idle power for the 1G NICs