

Network and Telecom Equipment - Energy and Performance Assessment

Test Procedure and Measurement Methodology

Draft 1.0.2, Oct 14 2008

Revision Record:

Version	Date	Author	Reason
0.2	12/12/2007	dkh	First draft
0.3	03/22/2008	dkh	Input from METI "Small Routers" document
0.5	06/11/2008	aalimian	Extended product classes, algorithm graph
0.6	06/15/2008	dkh	Changes per first ECR trials across different product types
0.8	06/28/2008	dkh	Input from Verizon VZ-TPR-9205 spec
0.9	07/15/2008	dkh	Changes and comments per input from Bruce Nordman (LBNL)
1.0	10/12/2008	dkh	Input from ITU-T Focus Group on ICT's and Climate Change

Areg Alimian, IXIA Corp

Daniel Kharitonov, Juniper Networks, Inc.

Bruce Nordman, Lawrence Berkeley National Lab

aalimian@ixiacom.com

dkh@juniper.net

BNordman@lbl.gov

Table of Contents

List of Acronyms	4
Purpose.....	5
Background.....	5
Theoretical basis	5
Scope.....	5
Measurement procedure.....	6
Metric computation.....	7
* Sometimes, for convenience purposes energy efficiency is also reported in Gbps/Watt under the name of EER; $EER = 1 / (ECR)$ Reporting	7
Reporting.....	8
REFERENCES	9
APPENDIX A. Measurement Conditions.....	10
APPENDIX B. Proposed Product Classes & Test Applications	11
Class 1 - Routers	11
C1.1 Core routers	11
C1.2 Edge Routers	11
C1.3 Multipurpose routers	11
Class 2 - WAN/Broadband Aggregation Device.....	12
C2.1 BRAS devices	12
C2.2 BSR/Common Edge devices	12
Class 3 - Ethernet L2/L3 Switches.....	12
C3.1 Carrier Ethernet Platforms	12
C3.2 Generic Ethernet Platforms	13
Class 4 - Experimental	13
Class 5 - Security appliances (DPI, Firewalls, VPN Gateways etc).....	13
Class 6 - Application Gateways (Layer 5-7 accelerators, load balancers, etc).....	13
APPENDIX C. Notes and rationale.....	14

List of Acronyms

ECR	- Energy Consumption Rating (scalar metric, Watts/Gbps)
ECRW	- Energy Consumption Rating Weighted (synthetic metric)
ECR (class Y)	- ECR metric computed for class equipment Y
T_f	- Measured maximum effective throughput (full-duplex, Gbps)
E_f	- Energy consumption under full load, Watts
E_h	- Energy consumption under half load, Watts
E_i	- Energy consumption under idle conditions, Watts

Purpose

The purpose of this document is to define a standard framework for first-order approximation of energy efficiency for network and telecom equipment.

Background

As energy efficiency in datacenter and network and telecom space becomes an topic of increasing interest, it spurs significant activity across commercial companies and government agencies to define the sets of energy efficiency criteria for qualification and homologation purposes.

So far such work has been mostly oriented towards private testing and qualification – an approach best suitable for early adopters and organizations with extensive qualification labs and test budgets. This document complements these efforts with an open efficiency criteria that can be independently estimated and reported by equipment vendors and test labs for ease of selection and comparison.

The purpose of this work is tri-fold:

1. Define a test procedure for measurement and estimate of energy efficiency for network and telecom equipment
2. Establish a common energy efficiency metric for the network and telecom industry
3. Promote energy awareness and competition between OEM vendors

Theoretical basis

This document defines energy efficiency as energy consumption normalized to effective throughput. Such approach goes in line with high-level methodology suggested in [SAINT 2008] and [VZ.TPR.9205] documents.

In other words, this document assumes the more energy-efficient network system to be the one that can move more data (in bits) using the same energy budget (in Joules).

Scope

This definition is best suited for medium- to large-scale network and telecom systems primarily serving data streams. It is less relevant to small office, SOHO and multi-purpose devices, where throughput is less relevant and efficiency criteria need to be more complex and involved, such as described in [METI 2008].

In general, this document is applicable to many types of network and telecom

equipment, including, but not limited to – routers, L2/L3 switches, optical shelves, security devices, load balancers.

Class-specific requirements are outlined in Appendix B.

Measurement procedure

The actual measurement cycle is designed to be simple, fast and inexpensive to run. It can be fully automated and is design to reflect the utilization profile and conditions frequently experienced in the field

There is no SUT configuration change allowed any time during the test. All energy savings adjustments (if done) by the SUT should be automatic

The procedure consists of four main steps.

System under Test (SUT) Preparation.

SUT is configured according to class requirements and offered load defined in the class requirements (Appendix B).

Router tester equipment is used to simulate the load and collect the results.

Step 1

First run determines the maximum system throughput T_f (full duplex, measured in Gbps) with methodology similar to RFC2544 within a selected application (at zero packet loss, full-mesh port configuration). There is no time limit for this run. The run is complete after T_f is determined.

Step 2 (full load)

Second run offers the load T_f (identified at step 1) to SUT for period of 1200 seconds. Energy consumption is being sampled for the entire period, and average consumption E_f calculated*.

Step 3 (half load)

Third run reduces the load T_f twice ($T_h = 0.5 \times T_f$) and runs for another 1200 seconds. Energy consumption is being measured for the entire period, and average consumption E_h calculated. Load reduction is achieved by reducing packet rate on all configured ports.

Packet loss during second or third run (if seen) invalidates the measurement and resets testing to first run to provide a better T_f estimate

*Please refer to Appendix A for measurement conditions and qualifications

Also, refer to Appendix C for graphical representation of the algorithm

Step 4 (*idle load*)

Idle run removes the load and runs for another 1200 seconds minutes. Energy consumption is being measured for the entire period, and average consumption E_i calculated. Load reduction is achieved by idling packet rate on all configured ports, or disabling ports on packet tester side, at vendor discretion.

Metric computation

One primary and one secondary metric are calculated with results from the aforementioned measurement algorithm.

The primary is a peak ECR metric, which is calculated according to the following formula:

$$\text{ECR} = E_f / T_f \quad (\text{expressed in Watts per 10Gbps})$$

Where: T_f = maximum throughput (Gbps) achieved in the measurement
 E_f = energy consumption (Watts) measured during running test T_f

ECR is normalized to Watts/10Gbps and has a physical meaning of energy consumption to move 10Gbits worth of user data per second. This reflects the best possible platform performance for a fully equipped system within a chosen application and relates to the commonly used interface speed*

Second metric is a weighted (synthetic) metric that takes idle mode into account. It is used in addition to the primary metric to estimate power management capabilities of the device.

$$\text{ECRW} = ((\alpha \times E_f) + (\beta \times E_h) + (\gamma \times E_i)) / T_f \quad (\text{dimensionless})$$

Where: T_f = maximum throughput (Gbps) achieved in the measurement
 E_f = energy consumption (Watts) measured during running test T_f
 E_h = energy consumption (Watts) measured during half-load test
 E_i = energy consumption (Watts) measured during idle test
 α, β, γ = weight coefficients to reflect the mixed mode of operation

ECRW reflects the dynamic power management capabilities of the device, which match energy consumption to the actual work accomplished. An ideal system following the Barroso principle of energy-proportional computing should yield $\text{ECRW} = | 0.55 * \text{ECR} |$ or better result. A system with no power management should yield $\text{ECRW} = | \text{ECR} |$

In this document, we specify $\alpha = 0.35, \beta = 0.4, \gamma = 0.25$

* Sometimes, for convenience purposes energy efficiency is also reported in Gbps/Watt under the name of EER; $\text{EER} = 1 / (\text{ECR})$

Reporting

Results can be variably reported based upon a class definition, or a combination of application and packet size, such as: ECR (class A) = Y, or ECR_x (B) = Z, Where A = equipment class, B= payload type, x = packet size; Y, Z = calculated efficiency

For instance, ECR (Class 1) = 12 Watts/Gbps; ECR₂₅₆ (IPv6) = 5.2 Watts/Gbps

For comparison purposes, the data can be collected in tables to reflect head-to-head competitive situation typical to RFP qualification, for example:

	Product A	Product B	Product C	Product D
Product class	Core	Core	Core	Core
Nominal Capacity	640G	1.28T	1.6T	3.2T
ECR (Class 1)	10 W/Gbps	12 W/Gbps	9 W/Gbps	12 W/Gbps
ECRW	9.1 W/Gbps	11 W/Gbps	7 W/Gbps	10 W/Gbps

REFERENCES

[VZ.TPR.9205] Verizon NEBS™: Energy Efficiency Requirements for Telecommunications Equipment. Todd Talbot, Ludwig C. Graff
<http://www.verizonnebs.com/TPRs/VZ-TPR-9205.pdf>

[METI 2008] METI Final Report (Small Routers, L2 Switches) by Router, etc. Evaluation Standard Subcommittee, Energy Efficiency Standards Subcommittee of the Advisory Committee for Natural Resources and Energy
http://www.eccj.or.jp/top_runner/pdf/tr_small_routers-apr_2008.pdf

[SAINT 2008] Proceedings of IEEE Symposium on Applications and the Internet. Luc Ceuppens, Daniel Kharitonov, Alan Sardella “POWER SAVING STRATEGIES AND TECHNOLOGIES IN NETWORK EQUIPMENT OPPORTUNITIES AND CHALLENGES, RISK AND REWARDS”

[EC 2008], Code of Conduct on Energy Consumption of Broadband Equipment, Draft Version 3 JOINT RESEARCH CENTRE Institute for the Environment and Sustainability Renewable Energies Unit European Commission

APPENDIX A. Measurement Conditions

A.1 Temperature

The equipment shall be evaluated at an ambient temperature of $25^{\circ}\text{C} \pm 3^{\circ}\text{C}$. The SUT itself should stay offline or operate at this air temperature for no less than three hours prior to the test. No ambient temperature changes are allowed until the test is complete.

A.2 Humidity

The equipment shall be evaluated at a relative humidity of 30% to 75%

A.3 Air Pressure

The equipment shall be evaluated at site pressure between 860 to 1060 mbar

A.4 DC Voltage

The input to the SUT (all active feeds) shall be at a nominal DC voltage of $\pm 0.5\%$

A.5 AC Voltage

The input to the SUT (all active feeds) shall be the specified voltage $\pm 1\%$ and the specified frequency $\pm 1\%$

A.6 Metrology requirements

Every active power feed should have the power (amp) meter installed in-line, with desired accuracy no less than $\pm 1\%$ of the actual power consumption

A.7 Sampling frequency

Ef, Eh and Ei calculations are based on averaging multiple readings over the course of measurements. Power meter(s) should be able to produce no less than 100 evenly-spaced readings in every 1,200 sec test cycle.

APPENDIX B. Proposed Product Classes & Test Applications

Disclaimer: For the purposes of public testing, all platforms should be tested with publicly available (shipping) software images, publicly available (shipping) board hardware revisions and fully documented and supported configurations

Class 1 - Routers

C1.1 Core routers

Description. Core routing platforms are systems with Terabit (half-duplex) or higher capacity. They are designed to provide line-rate performance in network cores with minimum functions (packet lookup and forwarding/switching). Core routing platforms come in various form factors, in standalone and multichassis enclosures.

Qualification. 1Tbps or better forwarding capacity

Test Application: IPv4, IPv6 or MPLS forwarding at discretion of the vendor;
packet size: 64B; forwarding over any types of forwarding entries (static, connected, IGP, EGP) – no less than active 16K routes.

Interface types: 10GE or 100GE as designated by the vendor, SR optics

Other notes. For the purposes of testing, all redundant components (fabric, routing engines, power supplies etc) should be present in the system

Metric awarded: ECR64(MPLS, IPv4, IPv6)

C1.2 Edge Routers

Description. Edge routing platforms

Qualification. 200Gbps or better forwarding capacity

Test Application: IP VPN, PWE, or VPLS forwarding at discretion of the vendor;
packet size: 256B; forwarding over any types of forwarding entries across all VPN instances, no less than 2K VPN instances active (PWE circuits, VPLS instances, IP VPN VRFs)

Interface types: at vendor discretion

Other notes. For the purposes of testing, all redundant components (fabric, routing engines, power supplies etc) should be present in the system

Metric awarded: ECR256 (PWE, VPLS, IP VPN)

C1.3 Multipurpose routers

Description. Routing platforms of variable purposes (enterprise, edge, etc)

Qualification. L3 forwarding

Test Applications: IPv4 or IPv6 forwarding at vendor discretion. packet size: 576B; forwarding over any types of forwarding entries, no less than 16K active routes.

Interface types: electrical or optical at vendor discretion

Other notes. For the purposes of testing, redundant component may be removed

Metric awarded: ECR576 (IPv4, IPv6)

Class 2 - WAN/Broadband Aggregation Device

C2.1 BRAS devices

Description. Legacy broadband aggregation devices

Qualification. Any capacity, PPPoE, PPPoA, PPP, per-subscriber QoS

Test Applications: PPPoE, PPPoA, PPP forwarding at discretion of the vendor; packet size: 256B; forwarding over any types of per-subscriber entries, no less than 64K subscribers with no less than four (4) queues assigned to each.

Interface types: SR optical at vendor discretion

Other notes. For the purposes of testing, all redundant components (fabric, routing engines, power supplies etc) should be present in the system

Metric awarded: ECR256 (PPP, PPPoE, PPPoA)

C2.2 BSR/Common Edge devices

Description. Broadband aggregation devices, Ethernet-oriented

Qualification. Any capacity, PPPoE, PPP, IP DHCP, per-subscriber QoS

Test Applications: IP/DHCP, PPPoE, PPP forwarding at discretion of the vendor; packet size: 256B; forwarding over any types of per-subscriber entries, no less than 64K subscribers with no less than four (4) queues assigned to each.

Interface types: SR optical at vendor discretion

Other notes. For the purposes of testing, all redundant components (fabric, routing engines, power supplies etc) should be present in the system

Metric awarded: ECR256 (PPP, IP DHCP, PPPoE)

Class 3 - Ethernet L2/L3 Switches

C3.1 Carrier Ethernet Platforms

Description. Carrier-grade ethernet switching platforms

Qualification. L2 (Ethernet) forwarding, MPLS forwarding, IPv4, or IPv6 forwarding

Test Application: Ethernet or MPLS forwarding at vendor discretion. Payload packet size: 256B frames; forwarding over any types of forwarding entries and encap type.

Interface types: SR optical (10/100/1000/10GE) at vendor discretion

Other notes. For the purposes of testing, redundant components must be present
Metric awarded: ECR256 (Ethernet, MPLS, IPv4, IPv6)

C3.2 Generic Ethernet Platforms

Description. Ethernet switching platforms

Qualification. L2 (Ethernet) forwarding, MPLS forwarding, IPv4, or IPv6 forwarding

Test Application: Ethernet or MPLS forwarding at vendor discretion. Payload packet size: 1500B frames; forwarding over any types of forwarding entries and encaps type.

Interface types: Copper or SR optical (10/100/1000/10GE) at vendor discretion

Other notes. For the purposes of testing, redundant components may be removed

Metric awarded: ECR1500 (Ethernet, MPLS, IPv4, IPv6)

Class 4 - Experimental

Description.

Qualification.

Test Application:

Class 5 - Security appliances (DPI, Firewalls, VPN Gateways etc)

Description. Security platforms of variable purposes (IP Sec VPN, HTTPS, DPI, IDS etc)

Qualification. L3 forwarding, security features

Test Application: IP SEC or HTTPS, minimum number of firewall or DPI forwarding rules at vendor discretion; 512B payload packets

Interface types: at vendor discretion

Other notes. For the purposes of testing, redundant component may be removed

Metric awarded: ECR512 (IPSec DES, IPSec 3DES, HTTPS, DPI, IDS, etc)

Class 6 - Application Gateways (Layer 5-7 accelerators, load balancers, etc)

Description. Application platforms of variable purposes (SLB, accelerators, compressors)

Qualification. Application-specific features

Test Application: User traffic at vendor discretion (need more qualification for setup); 512B payload packets

Interface types: at vendor discretion

Other notes. For the purposes of testing, redundant components may be removed

Metric awarded: ECR512 (SLB, TCP acceleration, compression, etc)

APPENDIX C. Notes and rationale

Q. General - What is the difference between first-order and second-order efficiency approximation?

Wikipedia: *First-order approximation* is the term scientists use for a further educated guess at an answer

A. Second-order approximation would require application-specific configurations and traffic profiles. Since we cannot match SUT to the exact configuration, load and conditions it will experience in any single deployment case, we provide a first-order approximation of what the energy efficiency can be.

Q. General – Why is ECR methodology different from [METI 2008] and [EC 2008] documents?

A. The main difference is in the focus and equipment class. METI and EC work was primarily targeted at consumer-level equipment where performance is not a differentiator. Instead, METI and EC documents define a fixed set of energy allowances for every product class and functionality option supported. The fact that this option may not operate at line rate typically does not matter in home and SoHo environments. As a result, consumer-level network and telecom equipment can be massively oversubscribed from the bandwidth perspective without noticeable impact on usability. For instance, it does not matter if the home DSL router cannot operate all wireless or wired LAN ports at line rate, as sustained performance is not required for domestic LAN operation. In fact, consumer-grade network device can be easily compared to a lightbulb – it fills a basic need at a fixed energy cost.

Carrier-class network and telecom equipment, on the other hand, presents a different case, where functions are delivered across many ports at high speed and revenue generation depends on performance. In the carrier world, an oversubscribed platform is not equal to line-rate device application-wise, and thus, it cannot be fairly compared from the energy consumption perspective.

This is where network equipment loses the analogy to light bulbs – the amount of commercial payload (system capacity) must be factored into efficiency estimates. Internet traffic can, in fact, be compared to commercial freight payload – it naturally takes more energy to transport an increasing volume of cargo. This is why ECR is defined as an efficiency metric, which positions every SUT relative to its competition on the energy/performance grid.

Q. General - What is homologation and what agencies are involved?

A. Homologation is conforming equipment to national or international standards. We expect ECR methodology to influence homologation practices in EU (IEC), USA (EPA) and Japan (METI)

Q. General - In test procedure, the SUT is equipped up to the maximum. However, in many applications, it won't be the case. Would the measured ECR metric still be relevant?

A. Modular telecom platforms are rarely deployed in full configurations from the start; instead, they typically reach their service ceiling midlife, when the network goes through expansion and upgrade rounds.

To estimate the effect of the partial configuration, we can represent the power draw of a modular router or switch to be a sum of a fixed part F (chassis, host system, fabric, clocking) and a variable part V (which represents removable linecards, interface ports and physical line drivers), $E = F + V$

It is trivial to demonstrate that a system with more efficient fixed and variable parts (as normalized by throughput) in a full configuration will also remain more efficient across all partial configurations. If this condition is not true, a crossover point can be found, where a previously less effective system may become more efficient with proportional reduction of removable components (typically, a fairly degraded configuration). For most practical cases, partial configurations will never change the relative standing of comparable platforms; moreover, a higher utilized system will yield a better energy efficiency in the first place.

Q. Test procedure – how is the system probed for effective throughput T_f ?

A. We do not define the exact probing and search algorithm for zero-loss operation. We suggest using RFC2544 methodology and applications for doing so – i.e. binary search for correct load profiles.

Q. Test procedure – why is zero-loss operation required? RFC2544 allows for configurable percentage of packet loss.

A. Indeed, there are cases, where application class prevents lossless operation at exact line rate (theoretical physical line limit). Examples would be – interfaces with byte stuffing (i.e. SONET), non-payload traffic leaving the router/switch (IGP/EGP updates), etc. In this case, the RFC2544 procedure needs to be instructed to top at a safe load level – for instance, 98 or 99 percent theoretical line load. This should not affect relative platform standings as all equipment belonging to the same class would have to be tested in a similar way. On the other hand, random (even minimal) packet loss is very undesirable to modern routers and switches and should be avoided at all costs.

Q. Test procedure – why is the test run defined at 1,200 seconds?

A. We need a compromise between the accuracy and speed. Currently suggested value of 1,200 seconds in most cases will allow full ECR test suite to complete in approximately 1.5 hours. Longer test runs would increase the run time accordingly and risk tying up expensive resources for extended period of time. Shorter test runs may bring the danger of overestimating the SUT. For example, some excess traffic (or traffic bursts) can be wrongfully accounted as delivered, while it could be actually buffered inside the device under test. Also, SUT's energy ratings may be reasonably affected by the state of its active cooling system, which might require a certain temperature threshold to activate (i.e. spin fans at full speed).

Q. Test procedure – why can't SUT be reconfigured between the test runs?

A. This requirement is there to reflect the dynamic nature of internet traffic and associated load profiles. While, indeed, it is often possible to statically alter configuration to match the relaxed load (i.e. remove unused ports, fabric cards, lookup engines, etc), this is not a viable case in the field situation, where the effective load can change at any second. Our ECRW metric design promotes automatic (intelligent) power management schemes.

Q. Metric calculation – can ECR be used to compare hybrid devices?

A. Absolutely. Hybrid devices are the systems that patently can operate in different equipment classes – i.e. a router and a firewall. Because of existence of such systems, some vendors make claims that they should form a separate class. This is not required. Instead, the vendors of hybrid devices are encouraged to obtain the ECR metrics in all relevant classes the device can be certified, for example ECR(Class A) and ECR(Class B). From those metrics, the end-user can easily deduct the projected ECR metric in a mixed mode of operation (i.e. 20 percent traffic in Class A and 80 percent Class B).

Further, when comparing a hybrid device (A) against a tandem of single-purpose devices (B, C), the ECR logic remains intact. The comparison becomes:

Max (ECR (A, Class A), ECR(A, Class B) vs. (ECR (B, Class A) + ECR (C, Class B))

Q. Can ECR be manipulated to achieve better readings?

A. In the design of ECR, we put significant effort to make sure the system-under-test performance cannot be skewed during the test phase. Since ECR is open and peer-reviewed, this also ensures that potentially wrongful ratings can be independently verified and weeded out.

Q. By focusing on components, isn't ECR ignoring the larger picture?

A. Having the best-in-class network components with respect to energy efficiency is not enough to build sustainable networks, but it's a necessary first step. It should be followed with responsible network design (which includes non-telecom infrastructure) and network operation. ECR can be considered a metric for building the bricks. It's possible to build a bad house with good bricks, but if the original goal is to build a good house, it's much easier to do with proper elements.

Interestingly enough, ECR methodology can be also extended over to the network level, although this is not covered in this document.