INTERNATIONAL TELECOMMUNICATION UNION

FOCUS GROUP ICT&CC

TELECOMMUNICATION STANDARDIZATION SECTOR

STUDY PERIOD 2005-2008

English only

Original: English

22 September 2008

INPUT DOCUMENT

Source: Juniper Networks, Inc, Daniel Kharitonov

Lawrence Berkeley National Lab, Bruce Nordman

Ixia Corporation, Areg Alimian

Title: **Network and Telecom Equipment - Energy and Performance Assessment**

Network and Telecom Equipment - Energy and Performance Assessment

Test Procedure and Measurement Methodology

Draft 0.97, Sep 22 2008

Revision Record

Version	Date	Author	Reason
0.2	2/12/2007	dkh	First draft
0.3	3/22/2008	dkh	Input from METI “Small Routers” document
0.5	6/11/2008	aalimian	Extended product classes, algorithm graph
0.8	6/28/2008	dkh	Input from Verizon VZ-TPR-9205 spec
0.9	07/15/2008	dkh	Changes and comments per input from LBNL

Prepared by

Areg Alimian, IXIA corp aalimian@ixiacom.com
Daniel Kharitonov, Juniper Networks, Inc. dkh@juniper.net
Bruce Nordman, Lawrence Berkeley National Lab BNordman@LBL.gov

Table of Contents

List of Acronyms 5
Background 6
Key Definitions 6
Scope 6
Efficiency Measurement Methodology 7
Metric computation 8
Reporting 9
REFERENCES 10
APPENDIX A. Measurement Conditions 11
APPENDIX B. Proposed Product Classes & Test Applications 12
- Class 1 - Routers 12
  - C1.1 Core routers 12
  - C1.2 Edge Routers 12
  - C1.3 Multipurpose routers 12
- Class 2 - WAN/Broadband Aggregation Device 13
  - C2.1 BRAS devices 13
  - C2.2 BSR/Common Edge devices 13
- Class 3 - Ethernet L2/L3 Switches 13
  - C3.1 Carrier Ethernet Platforms 13
  - C3.2 Generic Ethernet Platforms 14
- Class 4 - Experimental 14
- Class 5 - Security appliances (DPI, Firewalls, VPN Gateways etc) 14
- Class 6 - Application Gateways (Layer 5-7 accelerators, load balancers, etc) 14

APPENDIX C. Generalized measurement algorithm 15
APPENDIX D. Using ECR/EER for agency label/compliance 16
APPENDIX E. Notes and rationale 16

List of Acronyms

ECR	- Energy Consumption Rating (scalar metric, Watts/10Gbps)
EER	- Energy Efficiency Rating (weighted metric, Gbps/kWatt)
ECR (class Y)	- ECR metric computed for class equipment Y
$\mathbf{T_f}$	- Measured maximum effective throughput (full-duplex, Gbps)
$\mathbf{E_f}$	- Energy consumption under full load, Watts
$\mathbf{E_h}$	- Energy consumption under half load, Watts
$\mathbf{E_i}$	- Energy consumption under idle conditions, Watts
SUT	- System Under Test

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

Network and telecom systems form a special class of ICT equipment that is designed exclusively for data transfer (as opposed to data processing, storage or hybrid functions).

The purpose of this work is tri-fold:

1. Establish a common energy efficiency metric for the network and telecom industry
2. Define a test methodology for measurement and estimate of energy efficiency for network and telecom systems
3. Promote energy awareness and progress at individual, business and national levels.

Key Definitions

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 per second) using the same energy budget (in Watts).

Note, that this definition can be recursively applied on all levels of the actual multi-level data exchange infrastructure – an Internet in the whole, core ISP network, edge ISP network, access network and individual telecom devices constituent to any network type.

Since the energy performance of the network within a chosen topology and link technology will patently depend on energy performance of individual building blocks (devices), further discussion is limited to energy efficiency of such devices.

Scope

This document is oriented towards for medium- to large-scale network and telecom (Enterprise, Residential and Business) systems. It is less relevant to consumer-level, SOHO and small CPE cases, where throughput is less relevant and efficiency criteria need to be based on functionality, and not on the revenue-generating throughput (see [METI 2008] and [EC 2008] for examples of consumer-level network energy metrics).

Generally, 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 and load balancers.

Class-specific notes and requirements are independent from metric computation methodology and are outlined in Appendix B.

Efficiency Measurement Methodology

Per definition adopted above, SUT efficiency is a function of energy consumption and actual (revenue-generating) throughput, both parameters being a subject of measurement (verification).

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 $\mathbf{T_f}$ (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 $\mathbf{T_f}$ is determined.

Step 2 (full load)
Second run offers the load $\mathbf{T_f}$ (identified at step 1) to SUT for period of 1200 seconds. Energy consumption is being sampled for the entire period, and average consumption $\mathbf{E_f}$ calculated*.

Step 3 (half load)
Third run reduces the load $\mathbf{T_f}$ twice ($\mathbf{T_h}$ = 0.5 x $\mathbf{T_f}$) and runs for another 1200 seconds. Energy consumption is being measured for the entire period, and average consumption $\mathbf{E_h}$ 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 $\mathbf{E_i}$ 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

Two metrics are proposed to be calculated based on the aforementioned measurement algorithm.

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

ECR = $\mathbf{E_f/T_f}$ (expressed in (k)Watts per (Nx)Gbps)

Where: $\mathbf{T_f}$ = maximum throughput (Gbps) achieved in the measurement
$\mathbf{E_f}$ = energy consumption (Watts) measured during running test Tf

ECR is usually scaled to Watts/10Gbps and thus 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 reflects the fact, that most telecom systems have significant peak-on and peak-off times and that network traffic exhibits inherently fractlal behavior [Fractal 1996]. Therefore, it is is designed to promote the real-time power management capabilities.

$\mathbf{EER = T_f / ((\alpha \times E_f) + (\beta \times E_h) + (\gamma \times E_i))}$ (expressed in Gbps per kWatt)

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

EER is expressed in Gigabits per kWatt and represents the amount of data the SUT can move within a fixed power budget.

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

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 EERx (B) = Z,
Where A = equipment class, x = packet size; Y, Z = calculated efficiency

For instance, ECR (Class 1) = 120 Watts/10Gbps; EER256 (IPv6) = 50 Gigabits/kWatt

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 I	**Product J**	Product W
Product class	Core	Core	Core	Core
Nominal Capacity	640G	1.28T	1.6T	3.2T
ECR (Class 1)	100W/10Gbps	120W/10Gbps	**90W/10Gbps**	120W/10Gbps
EER64 (IPv4)	120Gbps/kW	110Gbps/kW	**150Gbps/kW**	100Gbps/kW

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 Comission

[Fractal 1996], Proc. 25th Allerton Conference on Communication, Control and Computing**. Bo Ruy** "Fractal network traffic modeling: past, present, and future**"**

APPENDIX A. Measurement Conditions

A.1 Temperature
The equipment shall be evaluated at an ambient temperature of 25°C ±3ºC. 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 ±0.5%

A.5 AC Voltage
The input to the SUT (all active feeds) shall be the specified voltage ±1% and the specified frequency ±1%

A.6 Metrology requirements
Every active power feed should have the power (amp) meter installed in-line, with desired accuracy no less than ±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, EER64 (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, EER256 (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, EER576 (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, EER256 (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, EER256 (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, EER256 (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 encap 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, EER1500 (Ethernet, MPLS, IPv4, IPv6)

Class 4 - Experimental

Description.
Qualification.
Test Application:

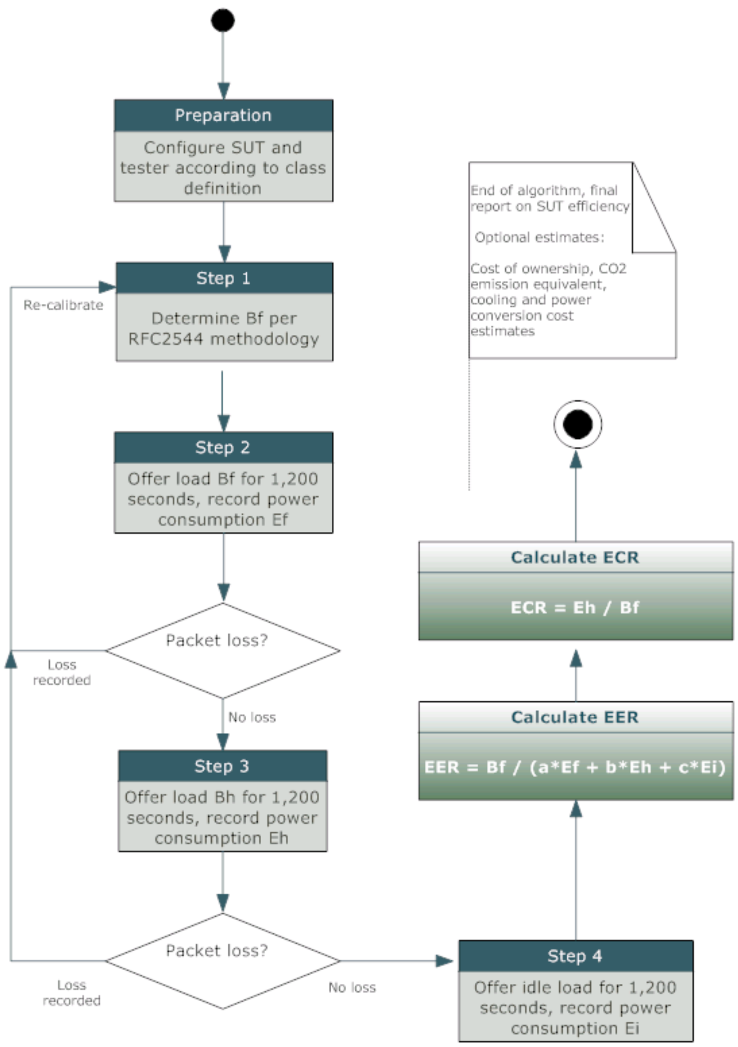
Class 5 - Security appliances (DPI, Firewalls, VPN Gateways etc)

Need precise class definitions here
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, EER512 (IPSec DES, IPSec 3DES, HTTPS, DPI, IDS, etc)

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

Need precise class definitions here
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, EER512 (SLB, TCP acceleration, compression, etc)

APPENDIX C. Generalized measurement algorithm



APPENDIX D.   Using ECR/EER for agency label/compliance

Due to their sheer simplicity and direct physical meaning, ECR/EER metrics can also be employed to set goals in compliance and approval programs. In such an application, metric syntax may change to reflect the authority name, compliance revision and product class, for example:

ECR-2008, Class C1.1 (green label) – awarded to qualifying platforms within 25% of the reference metric of 150 Watts/10Gbps (or better).

ECR-2008, Class C1.1 (yellow label) – awarded to qualifying platforms within 35% of the reference metric of 150 Watts/10Gbps (or better).

ECR-2008, Class C1.1 (red label) – non-compliant are platforms within 45% of the reference metric of 150 Watts/10Gbps (or worse).

Future-looking labels can be constructed in accordance to the goals of government certification, environment goals (i.e. CO2 reduction), local business or regulatory needs, i.e.:

ECR-2010, Class C2.2 (green label) – awarded to qualifying platforms within 25% of the reference metric of 80 Watts/10Gbps (or better). Metric constitutes the minimum requirement for equipment to be put in service in 2010 and later.

APPENDIX E.   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 spec different from [METI 2008] and [EC 2008] documents?

A. The main difference is in the focus and equipment class. METI and EC work was primatily 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 light bulb – 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 Tf?

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), exception 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 EER metric design promotes automatic (intelligent) power management schemes.

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

A. Absolutely. Hybrid devices are the systems that 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 neither required, nor should be encouraged. Instead, the vendors of hybrid devices are expectant to obtain the ECR/EER metrics in all relevant classes the same SUT 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. X percent traffic in Class A and Y 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))

A similar transform can be employed for EER.