Joint ITU / IEEE Workshop on Ethernet - Emerging Applications and Technologies

(Geneva, Switzerland, 22 September2012)

Energy Efficiency in IEEE Ethernet Networks –

Current Status and Prospects for the Future

Michael J. Bennett
Energy Sciences Network / Lawrence Berkeley National Laboratory
mjbennett@lbl.gov

Lowell D. Lamb
Broadcom Corporation
lowell.lamb@broadcom.com

Agenda

- *Current Status of IEEE Energy Efficiency Solutions*
  Lowell D. Lamb
- *Future Directions in Standards for Energy Efficiency*
  Michael J. Bennett
- *Discussion*

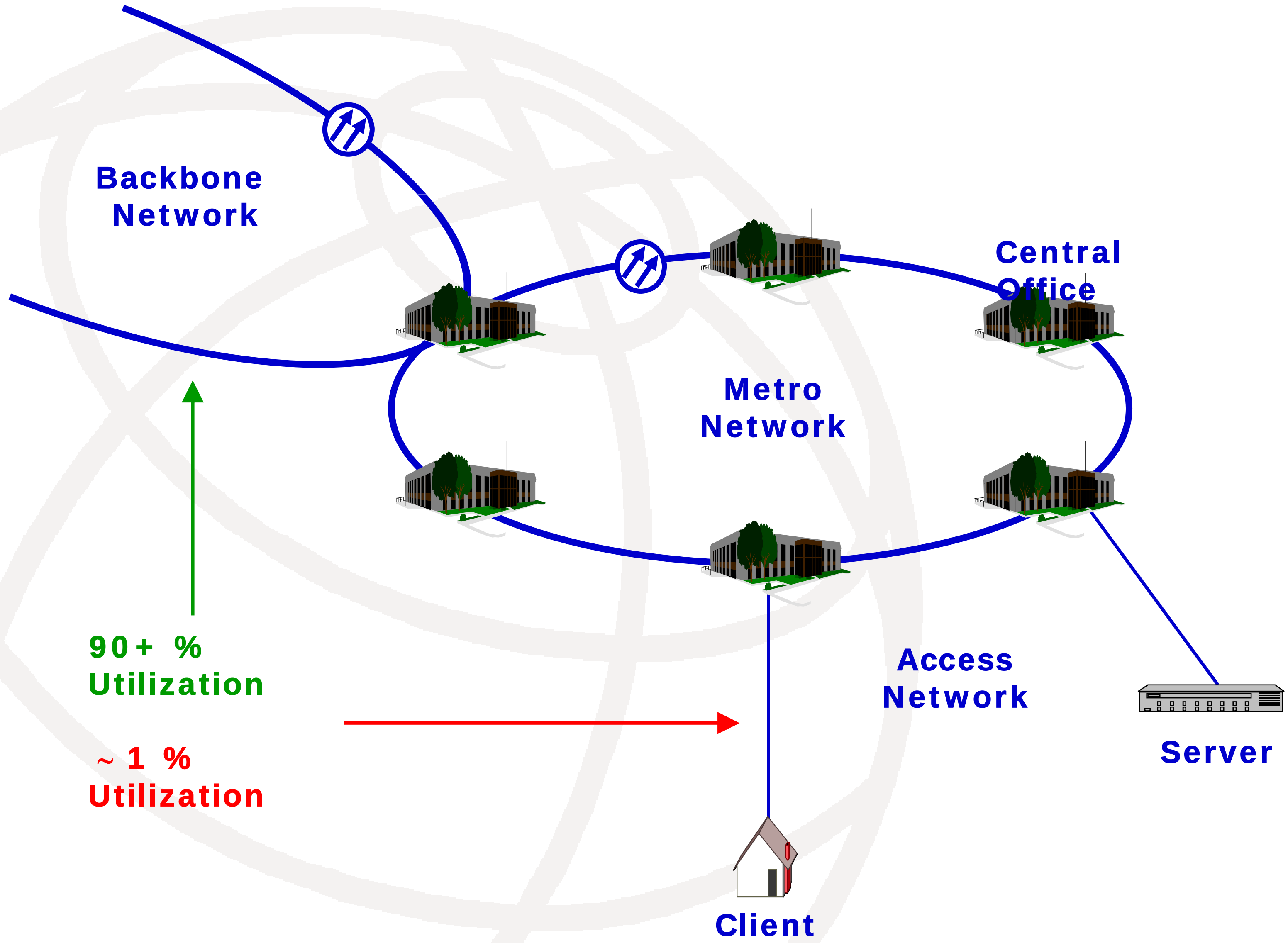
Current Status of IEEE Energy Efficiency Solutions

Lowell D. Lamb, Broadcom

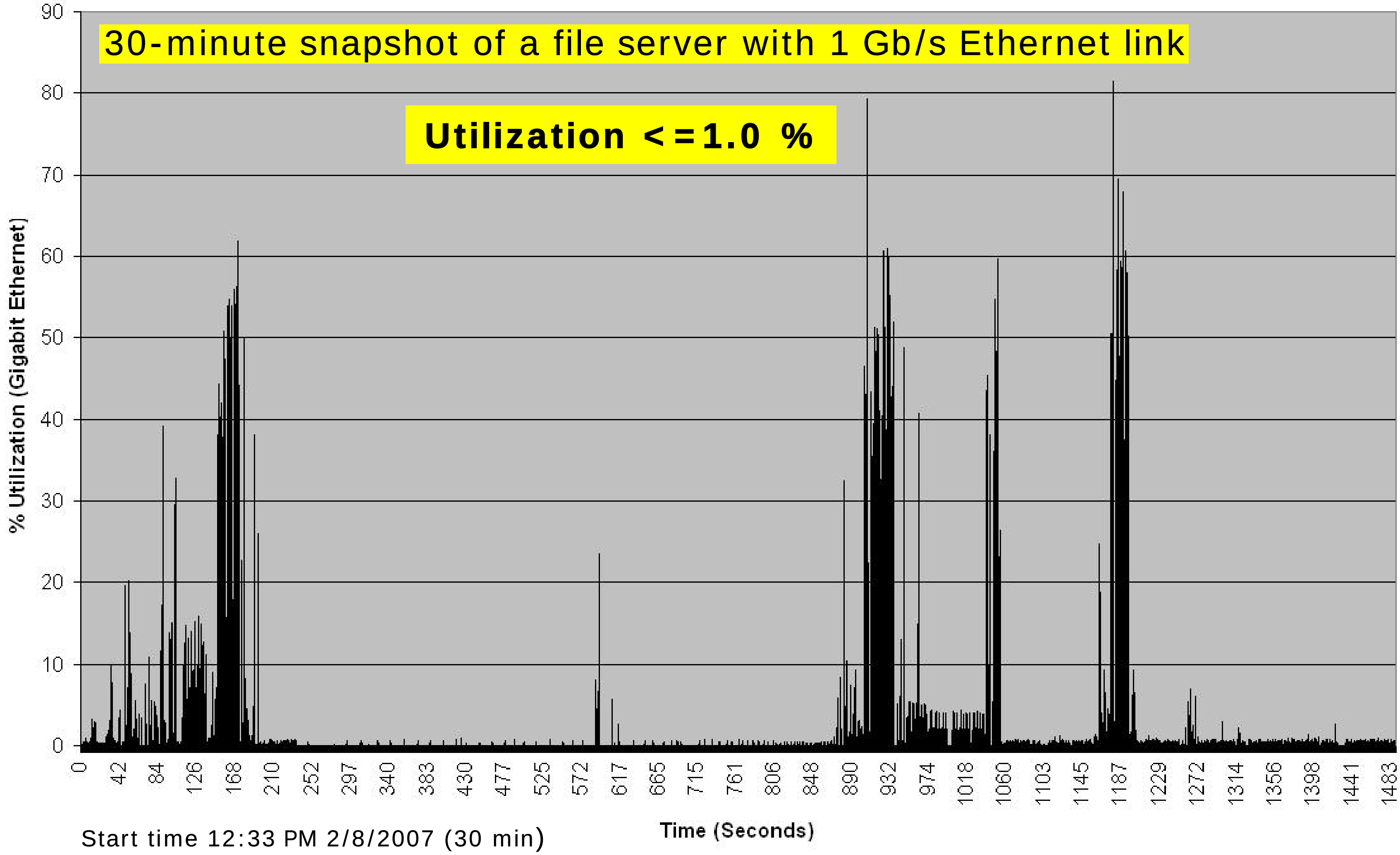
With input from
Wael William Diab, Broadcom

Problem Statement:
Energy Consumption in Telecom Networks

Link Utilization Varies



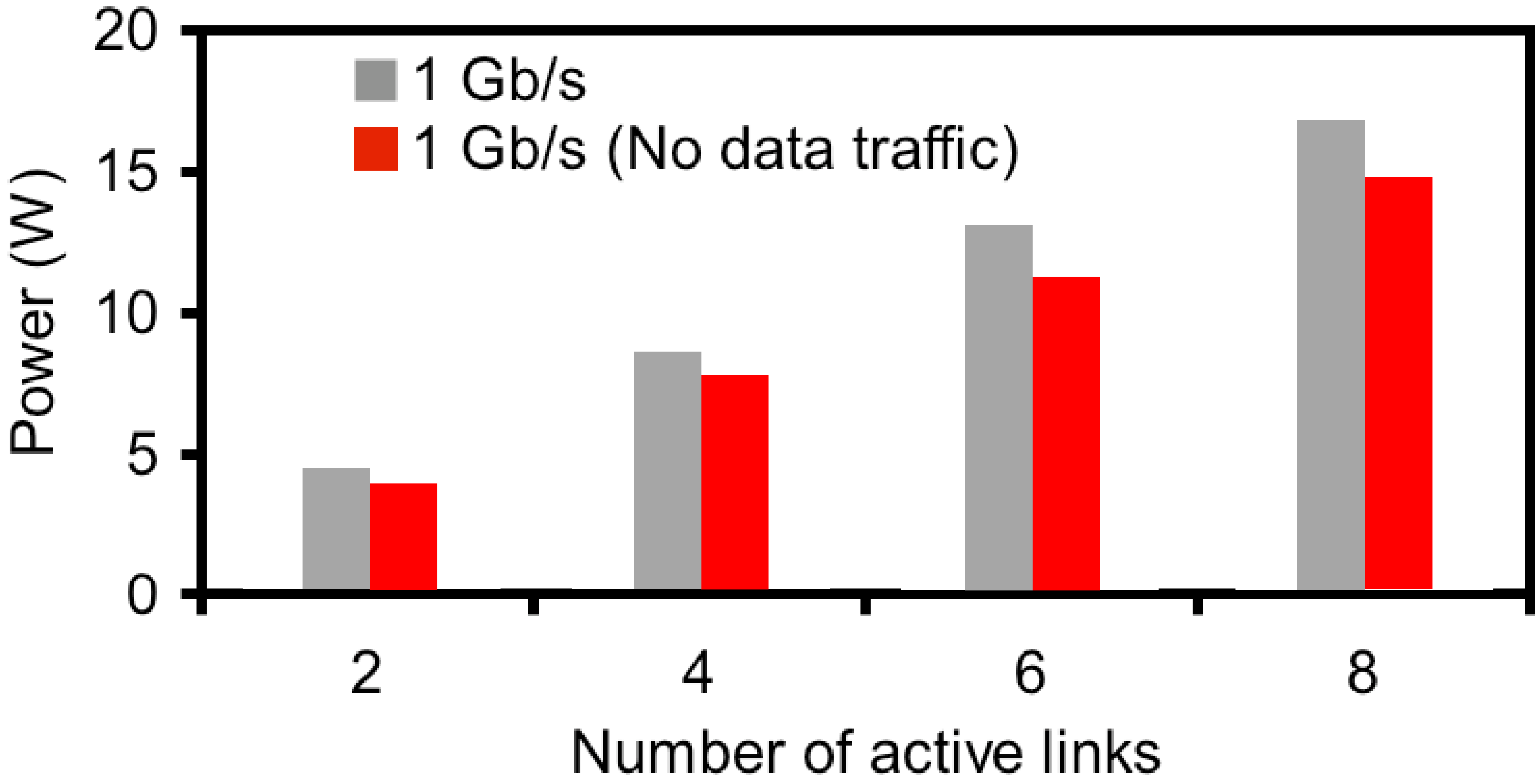
Example Traffic Profile



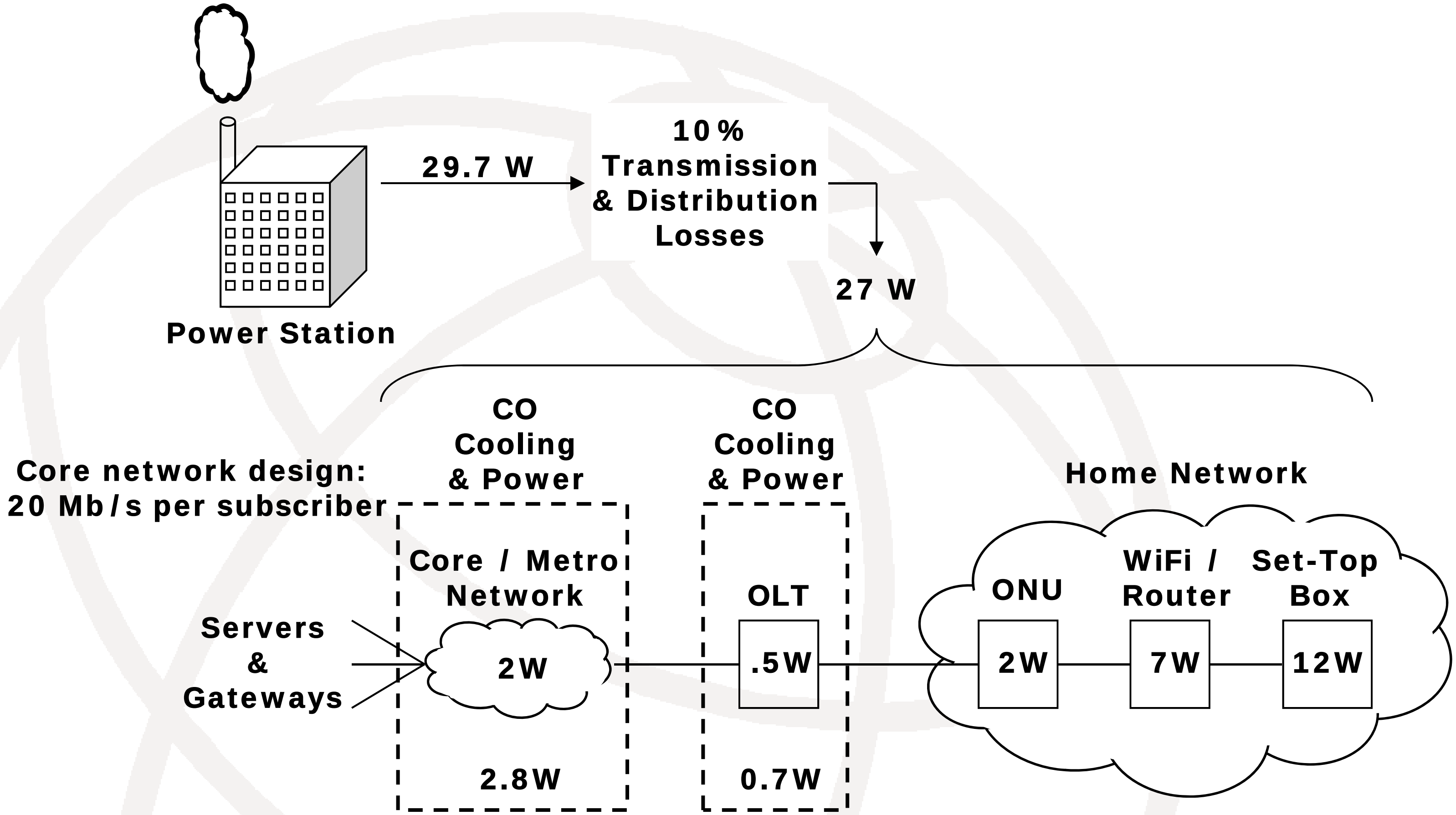
Take-home message: Most access links are idle most of the time.

Idle Connections Consume A Lot of Power

Typical Ethernet switch with 24 ports of 1 Gb/s



Residential Broadband Energy Requirements



29.7 W-Year = 429 kg Coal / Year = 1, 573 kg CO_2 / Year per subscriber

Almost all of this energy is generated and consumed to power idle access links.

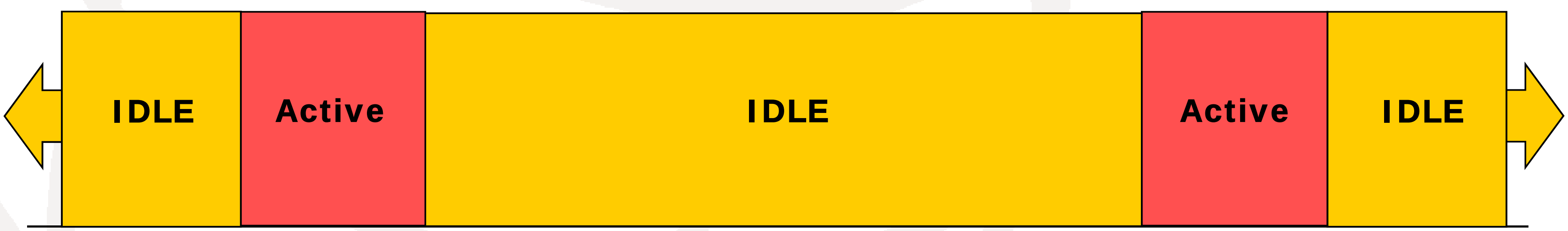
Lowell D. Lamb, *The Future of FTTH – Energy Consumption of Broadband Access Networks: Challenges and Opportunities (Invited)*, **NOC 2009 Proceedings: 14th European Conference on Networks and Optical Communications**, Valladolid, June 2009.

Energy Efficient Ethernet (EEE)

Ethernet IDLE Symbols

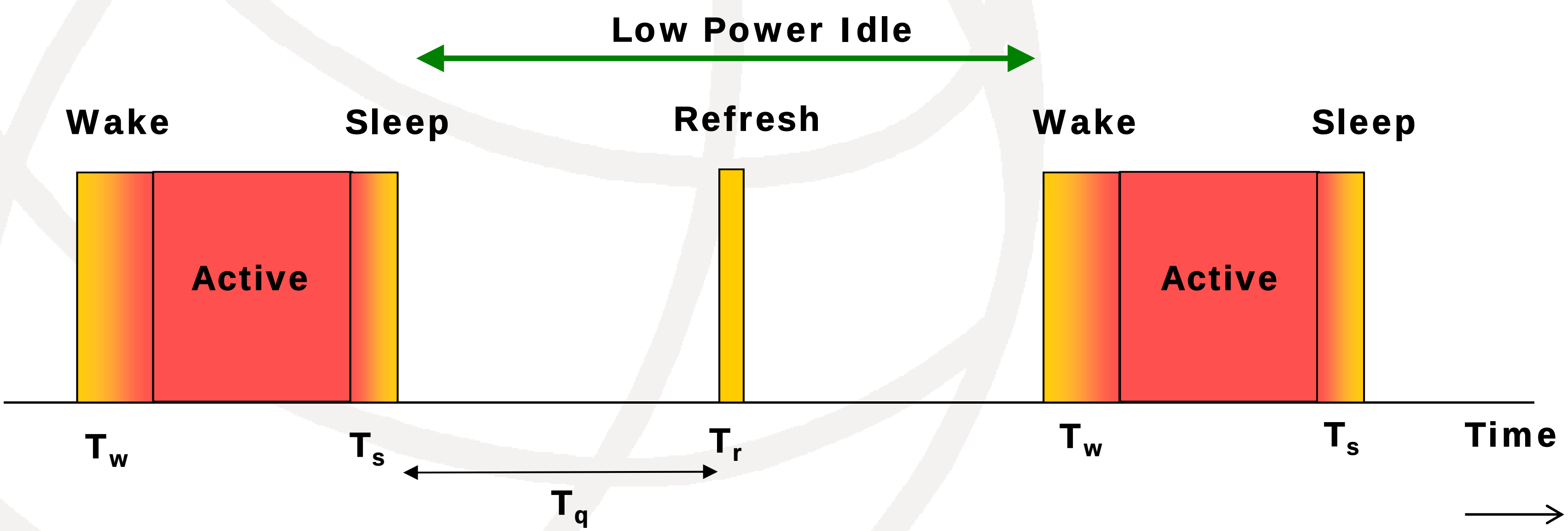
- **In the absence of traffic, IDLE symbols are sent continuously to**
  - **Provide a good clock to the link partner's receiver**
  - **Facilitate the detection of new traffic**

- ***On access links, IDLE symbols account for most of the power consumption***



Energy Efficient Ethernet (IEEE Std 802.3az-2010)

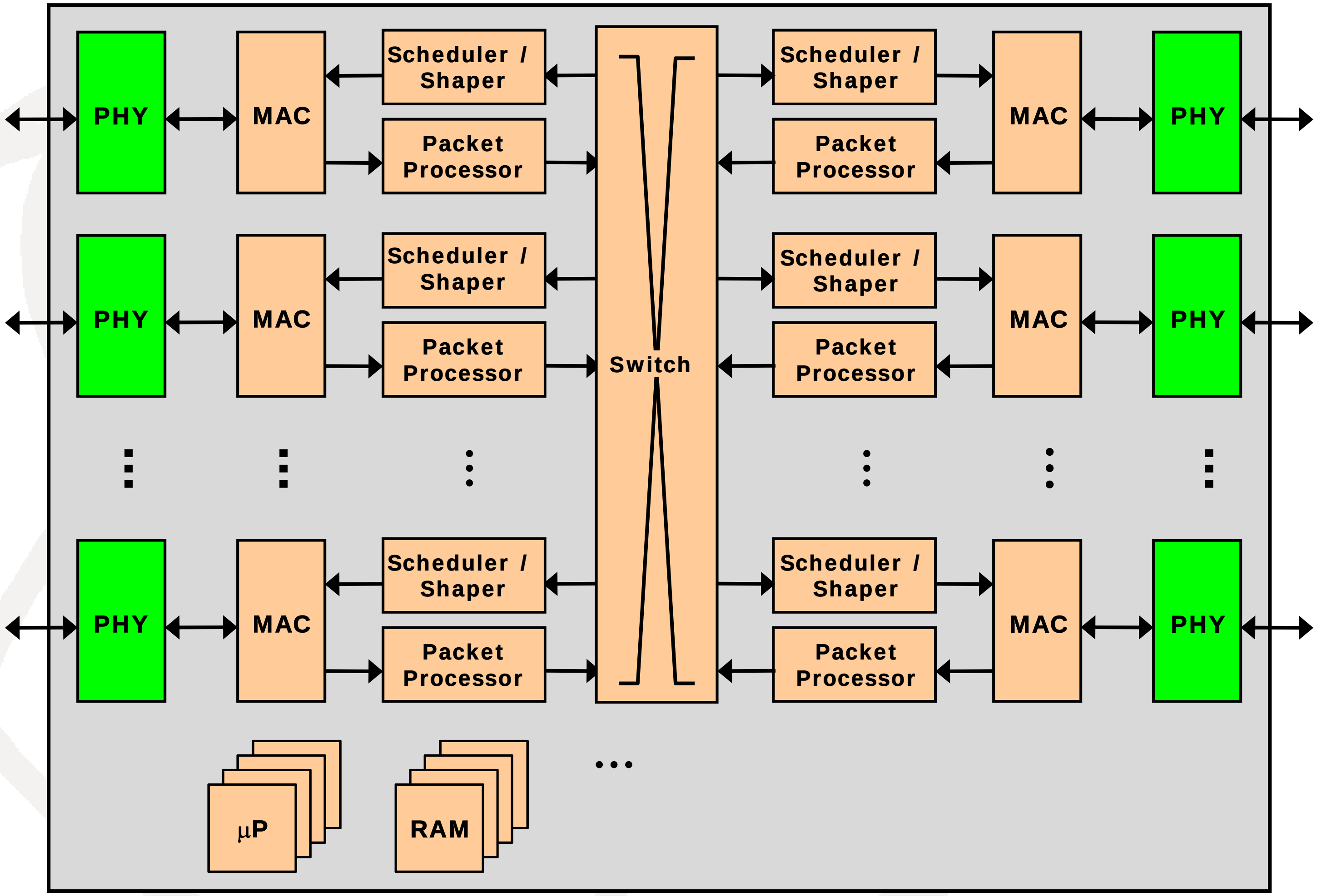
- Energy Efficient Ethernet (EEE) enables transitions to and from a low-power state in response to changes in network demand;
- The "BASE-T PHYs" and Backplane PHYs are included:
  - 100BASE-TX (Full Duplex), 1000BASE-T (Full Duplex), 10GBASE-T
  - 10GBASE-KR, 10GBASE-KX4, 1000BASE-KX
- *An LLDP-based layer-2 protocol allows for additional savings*



Doing More:

Energy Efficient Networks (EEN)

Inside a Communications ASIC



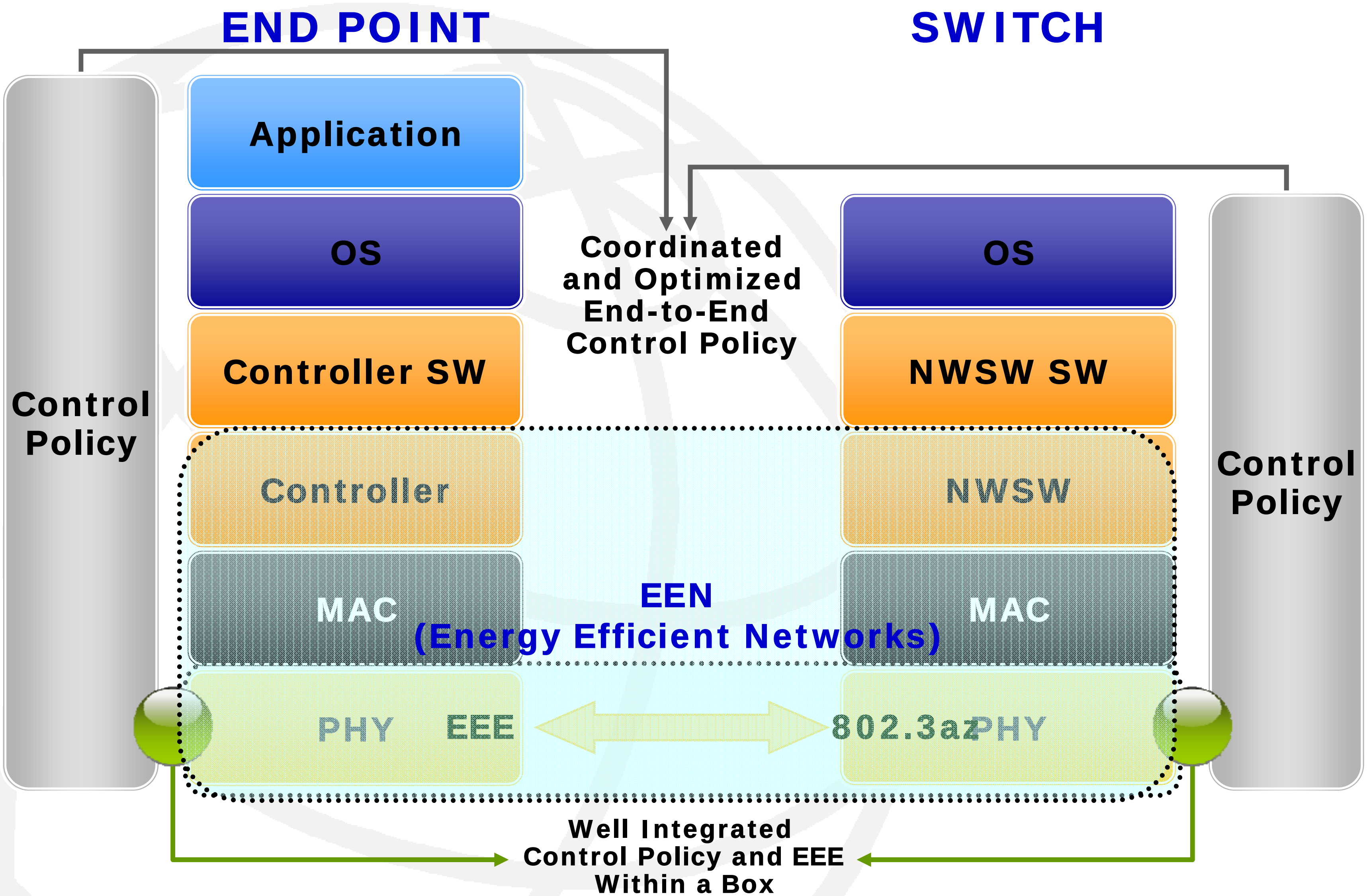
PHY ←→ **802.3az (EEE) does a good job here.**

What about other subsystems inside the ASIC?
What about other devices outside the ASIC?

EEN Adds a Control Policy to EEE

- **What is it?**
  - Control of when the *system* enters and exits a low-power state
- **How does it work?**
  - Via the LLDP-based layer-2 protocol defined by EEE (IEEE Std 802.3az-2010)
- **Where does it sit and what is its scope?**
  - It sits above EEE (IEEE Std 802.3az-2010)
  - At a minimum, the control policy will
    - Determine when to enter or exit a low-power state
    - Determine which subsystems enter a low-power state
    - Dynamically manage the traffic profiles (buffering depth, etc.)
- **Key features**
  - Gives network operators and IT managers fine-grained control of
    - Network power consumption
    - The efficiency-performance tradeoff

End-to-End Savings



Recommended Reading

EEN White Paper

- http://www.broadcom.com/collateral/wp/EEE-WP101-R.pdf (English)
- http://www.broadcom.com/collateral/wp/EEE-WP100-R-ja.pdf (Japanese)
- http://www.broadcom.com/collateral/wp/EEE-WP100-R-ko.pdf (Korean)
- http://www.broadcom.com/collateral/wp/EEE-WP100-R-zh_cn.pdf (Simplified Chinese)
- http://www.broadcom.com/collateral/wp/EEE-WP100-R-zh_tw.pdf (Traditional Chinese)

Wikibon

- http://wikibon.org/wiki/v/Networks_Go_GrEEN

Industry Awards

IEEE and Broadcom (shared), *Electronic Design* 2010 Best Electronic Design Award

- http://electronicdesign.com/article/news/Electronic-Design-Announces-2010-Best-Electronic-Design-Award-Winners.aspx

Energy Efficiency in IEEE EPON

P1904.1 SIEPON Specification (Table 4.1)

Power-Saving Modes are defined for all 3 Packages

Item	Feature	Package A	Package B	Package C
EDP	EPON Data Path	N/A	N/A	shall implement EDP per Annex 7A
RF	REPORT MPCP format	shall implement REPORT MPCPDU format per 8.4.1.3	shall implement REPORT MPCPDU format per 8.4.3.3	shall implement REPORT MPCPDU format per 8.4.2.3
RLC	report queue length calculation	shall implement queue length calculation per 8.4.1.2	shall implement queue length calculation per 8.4.3.2	shall implement queue length calculation per 8.4.2.2
QSD	queue service discipline	shall implement queue service discipline per 8.4.1.1	shall implement queue service discipline per 8.4.3.1	shall implement queue service discipline per 8.4.2.1
DCQ	discovery and configuration of queue parameters	N/A	should implement discovery and configuration of queue parameters per 8.4.3.4	N/A
USM	ONU transceiver status monitoring	shall implement transceiver status monitoring per 9.1.3	should implement transceiver status monitoring per 9.1.5	shall implement ONU transceiver status monitoring per 9.1.4, associated alarms and warnings per 9.1.6
TSM	OLT transceiver status monitoring			shall implement OLT transceiver status monitoring per 9.1.4
PLD	UNI port loop detection	N/A	N/A	shall implement UNI port loop detection per 9.1.8
RPC	remote ONU transmitter power supply control	N/A	N/A	shall implement remote ONU transmitter power supply control function per 9.4
E	events	shall implement events per 9.2.6	shall implement events per 9.2.6, 9.2.7 and 9.2.8	shall implement events per 9.2.3, 9.2.4, 9.2.5, and 9.2.6
LPTK	optical link protection, trunk type	N/A	should implement trunk optical link protection per 9.3.5	should implement trunk optical link protection per 9.3.3
LPTE	optical link protection, tree type	N/A	N/A	should implement tree optical link protection per 9.3.4
DE	data encryption	shall implement data encryption and integrity protection mechanism per 11.2.2	shall implement data encryption and integrity protection mechanism per 11.2.3	N/A
AU	ONU authentication	shall implement ONU authentication and secure provisioning per 11.3.3	shall implement ONU authentication and secure provisioning per 11.3.4	shall implement ONU authentication and secure provisioning per 11.3.2
MG	management	shall implement eOAM based management per 13.4	shall implement eOAM based management per 13.3	shall implement eOAM based management per 13.2
DCD	device and capability discovery	shall implement device discovery and capability discovery per 12.2.3	shall implement device discovery and capability discovery per 12.2.2	shall implement device discovery and capability discovery per 12.2.1
SU	software update	shall implement software update mechanism per 12.3.3	shall implement software update mechanism per 12.3.2	shall implement software update mechanism per 12.3.1
ME	management entities	shall implement management entities per 14.4	shall implement management entities per 14.3	shall implement management entities per 14.2
PS	power saving	shall support power saving per 10.5.3	should support power saving per 10.5.4	shall support power saving per 10.5.5
PM	performance monitoring	N/A	N/A	shall implement performance monitoring per 8.5
TVM	OLT VLAN modes	OLT shall support VLAN modes defined in 7.2.2.3	OLT shall support VLAN modes defined in 7.2.2.1.1, 7.2.2.1.3, and 7.2.2.1.5	OLT shall support VLAN modes defined in 0 – 7.2.2.2.7
UVM	ONU VLAN modes	ONU shall support VLAN modes defined in 7.2.2.3	ONU shall support VLAN modes defined in 7.2.2.1.2, 7.2.2.1.4, and 7.2.2.1.6	ONU shall support VLAN modes defined in 0 – 7.2.2.2.7
TTM	OLT tunneling modes	OLT shall support Tunneling modes defined in 7.3.2	N/A	N/A
UTM	ONU tunneling modes	ONU shall support Tunneling modes defined in 7.3.2	N/A	N/A
MC	multicast connectivity	shall support multicast operation as defined in 7.4.5	shall support multicast operation as defined in 7.4.2	shall support multicast operation as defined in 7.4.3 and 7.4.4
MCC	multicast connectivity, coexistence	shall support multicast connectivity coexistence per 7.4.1.1.2		
MA	MAC aging	N/A	N/A	shall implement MAC aging function as defined in 7.2.2.2.8
PSL	Port Selective Loopback	shall support Port Selective Loopback per 9.1.9	N/A	N/A

SIEPON Low-Power-Mode State Machines

Package A | Package B | Package C

OLT

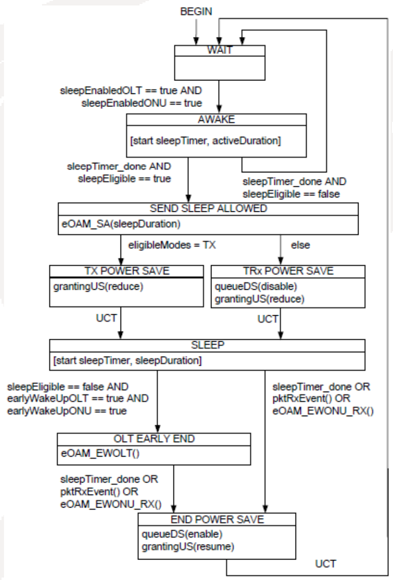


Figure 10-6—Power saving mechanism state diagram in the OLT

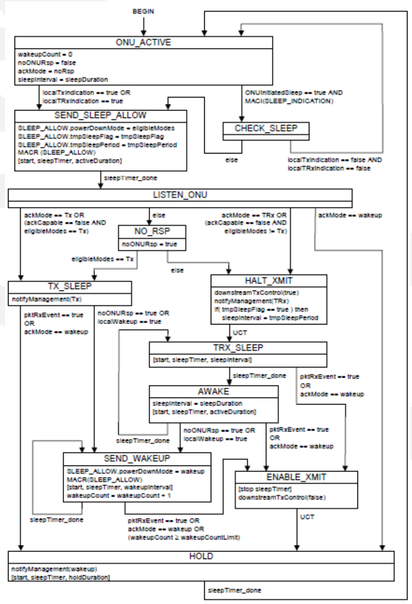


Figure 10-14—OLT power saving mechanism state diagram

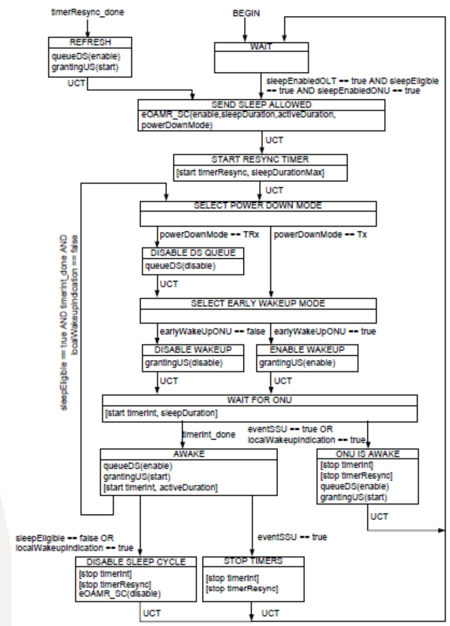


Figure 10-17—OLT power saving mechanism state diagram

ONU

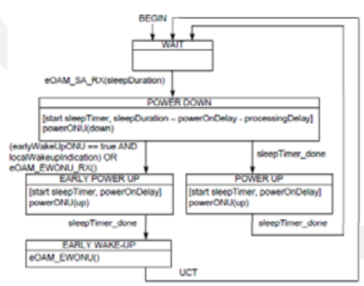


Figure 10-7—Power saving mechanism state diagram in the ONU

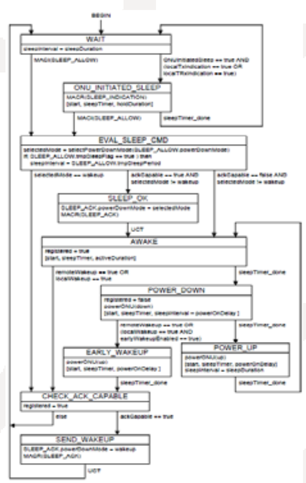


Figure 10-15—ONU power saving mechanism state diagram

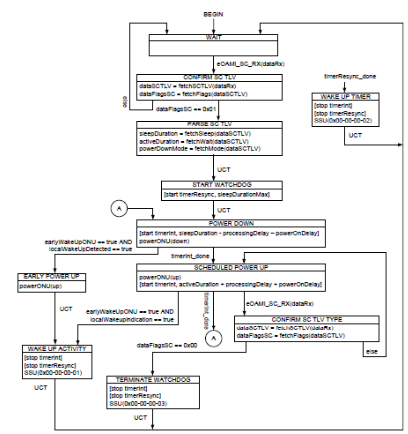
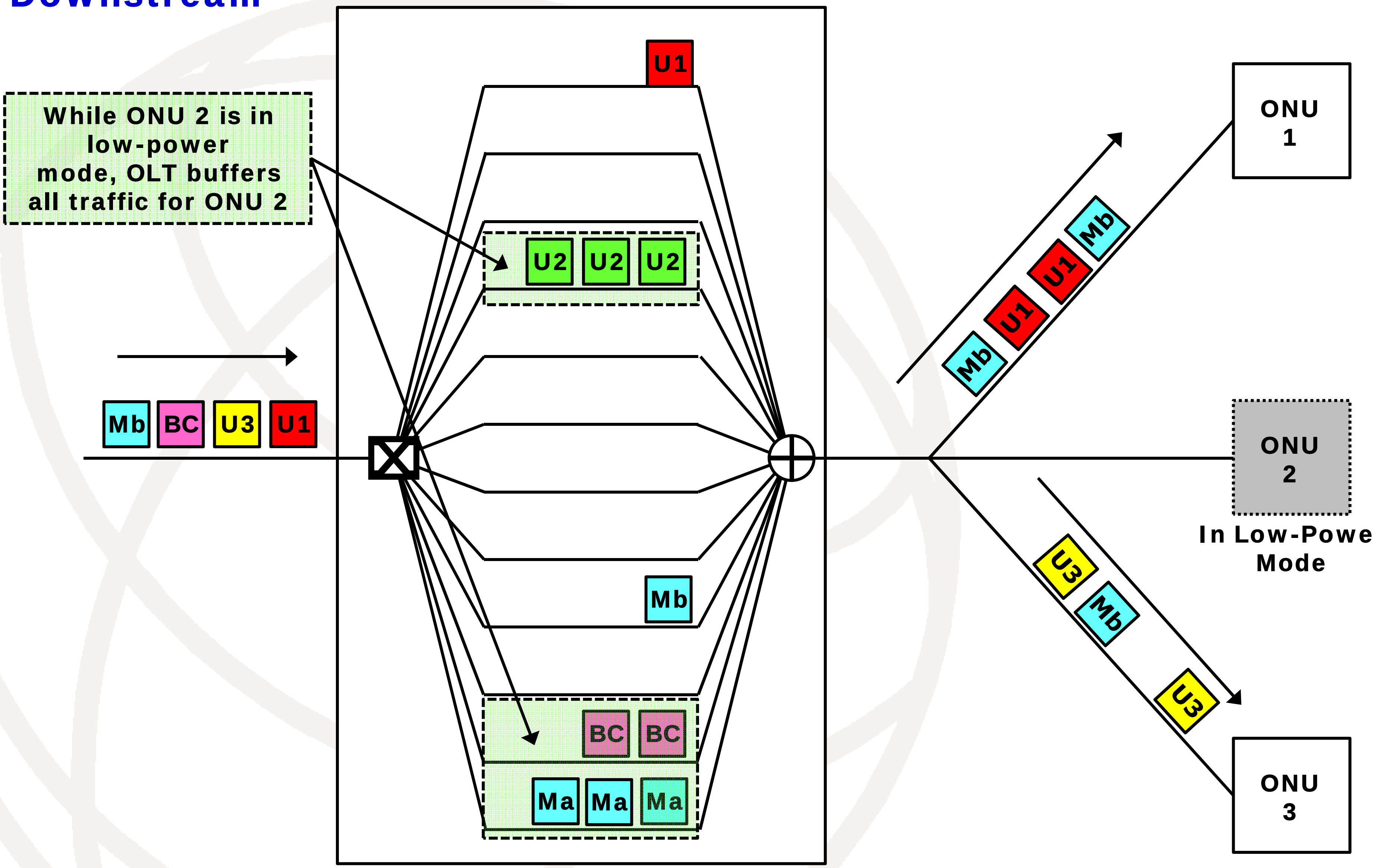


Figure 10-18—ONU power saving mechanism state diagram

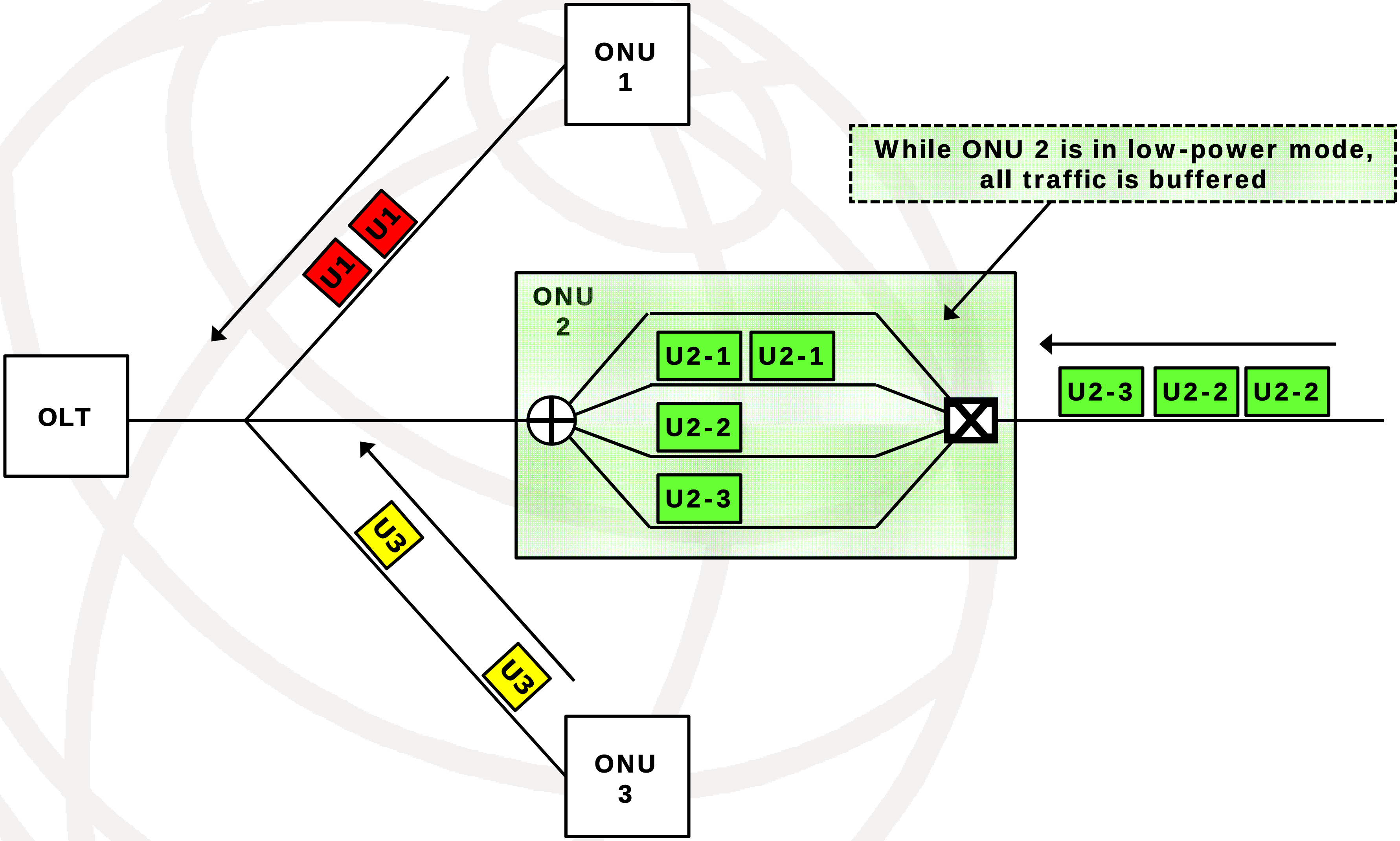
System-Level Requirement #1: No Packet Loss

Downstream



System-Level Requirement #1: No Packet Loss

Upstream



Other System-Level Requirements

- **ONUs must maintain registration while in low-power mode**
- **Independent power-savings modes for upstream and downstream must be supported**
- **SLAs must be enforced**
  - **Downstream scheduling and ONU sleep cycles are synchronized**
  - **Upstream DBA and ONU low-power modes are synchronized**
- **Impact on system design**
  - **Buffers, schedulers, & DBA must be implemented to accommodate low-power modes**
- **Impact on network operation**
  - **SLAs must be "energy-efficiency" aware**

Measured Values

- Power-down ONU optics and SERDES whenever traffic is light / absent
- Processor, PHY, and upstream packet buffer remain active
- OLT – All downstream packets are buffered during low-power mode
  - Unicast, multicast, & broadcast traffic
- ONU – All upstream packets are buffered during low-power mode
- No traffic is lost while ONU is sleeping
- ONU maintains registration during low-power periods
- Power-savings mode can be configured so that SLAs are maintained for all services

Measured values for prototype* ONU with single 100M UNI – Duty cycle: 30 msec ON / 200 msec OFF

Active State		Low-Power State	
PON Xcvr	= .8W	PON Xcvr	= .2W (75% savings)
Ethernet PHY	= .3W	Ethernet PHY	= .3W
PON SOC	= .7W	PON SOC	= .45W (36% savings)
Power Supply	= .6W	Power Supply	= .6W (more efficient supply could save .2-.4W)
Total	**= 2.4W**	**Total**	**= 1.55W (35% savings)**

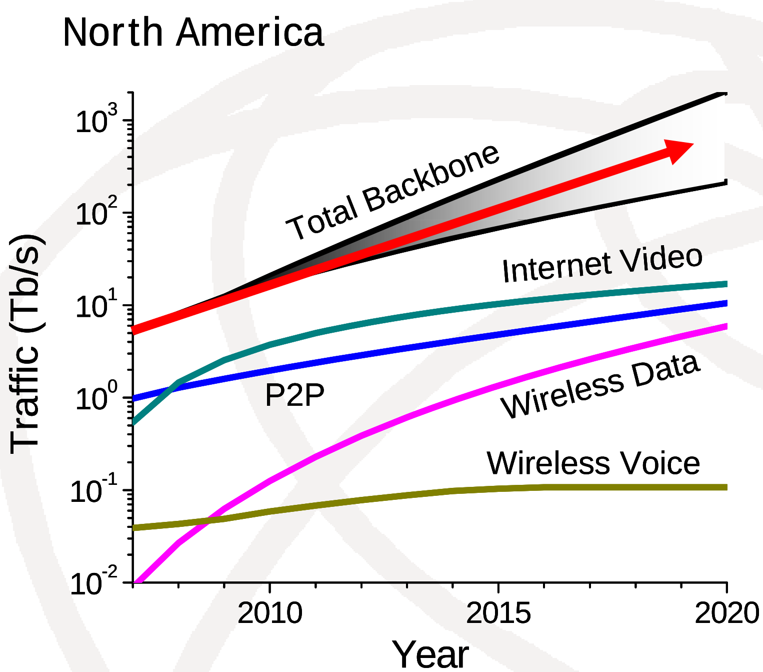
*Better performance should be possible with equipment optimized for energy-efficient operation

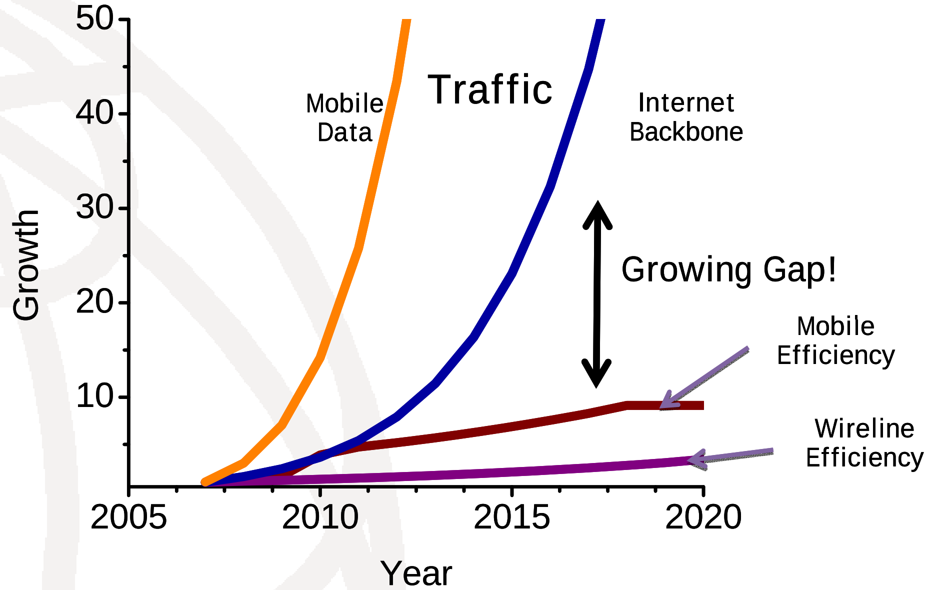
Future Directions in Standards for Energy Efficiency

Michael Bennett

With input from
Inder Monga (ESnet)
Thierry Klein (Bell Labs)
Wael William Diab (Broadcom)

Growing Network Energy Gap





Data from: RHK, McKinsey-JPMorgan, AT&T, MINTS, Arbor, ALU, and Bell Labs Analysis: Linear regression on log(traffic growth rate) versus log(time) with Bayesian learning to compute uncertainty

INTERNET

=5th

HIGHEST COUNTRY

If the internet was a country: energy consumption is higher than Russia and a little less than Japan

ENERGY

+27%

INCREASE

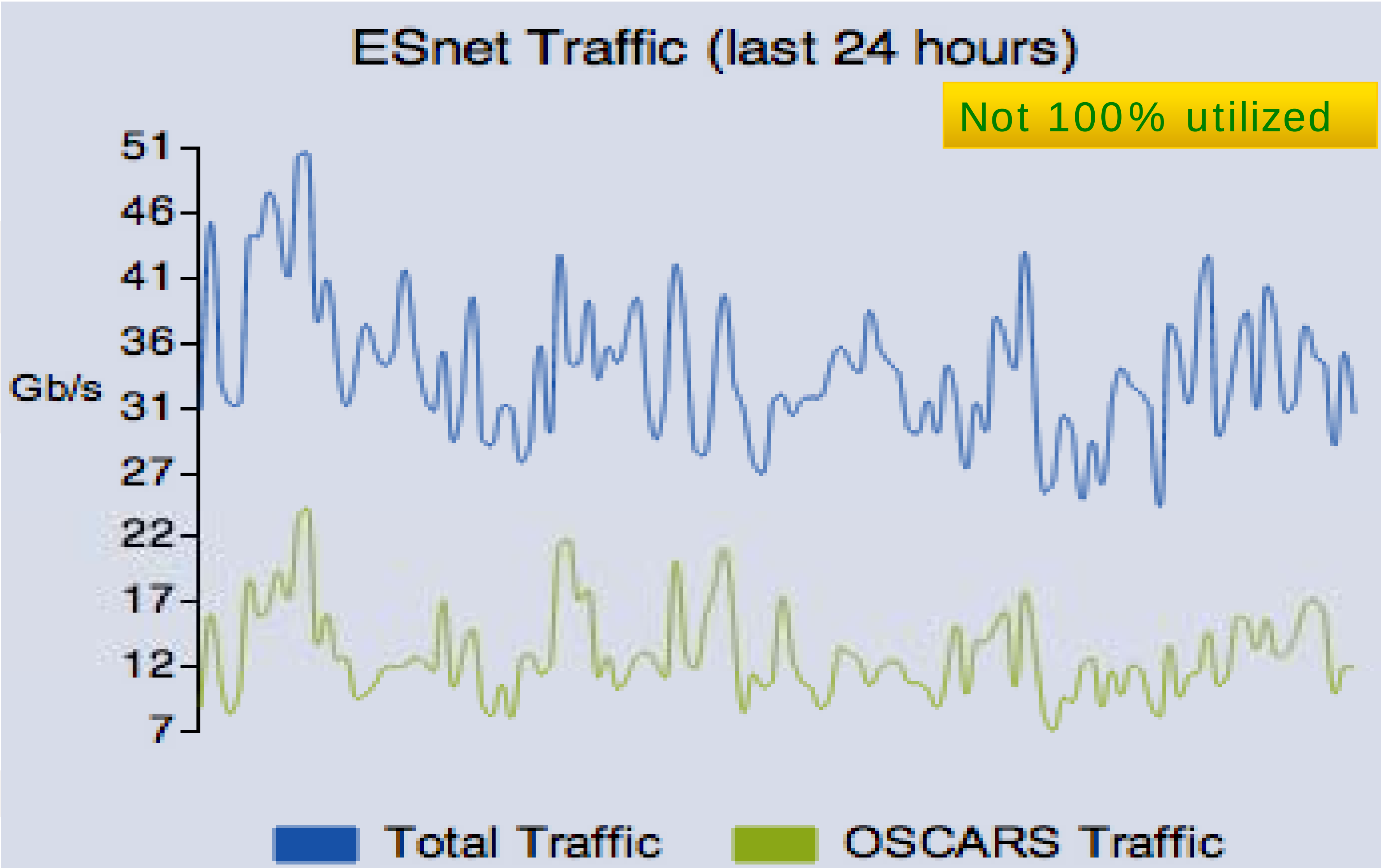
Energy consumption in communications service provider (CSP) networks is forecast to increase by 27% from 2012 to 2016

Courtesy Thierry Klein

Why continue to develop standards for EEE?

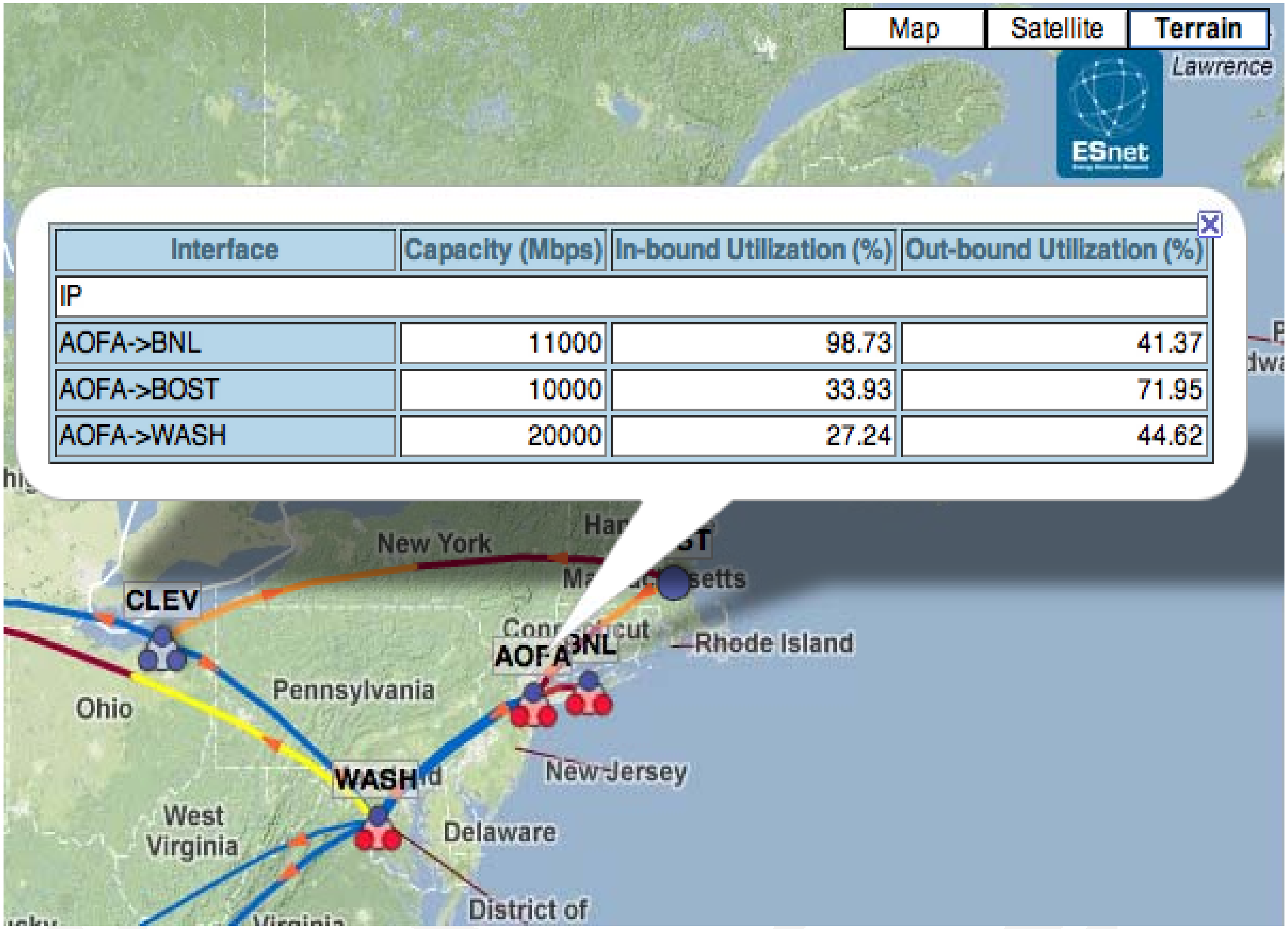
- ICT accounts for at least 1% of the world's energy-use
  - Roughly equivalent energy consumption of 15 million US homes
  - Equivalent to CO2 emissions of 29 million cars.
  - About one-third of this is due to network equipment
- IEEE Std. 802.3az-2010™
  - Provides tools to reduce energy consumed by network equipment
    - PHY power reduced roughly 70% or more
    - Much more savings possible in the system
  - Specified for copper ("BASE-T" and backplane) interfaces up to and including 10G
- We can do more
  - What about higher speeds or optical interfaces?
  - What's next?

Why continue to develop standards for EEE?



www.es.net

Why continue to develop standards for EEE?



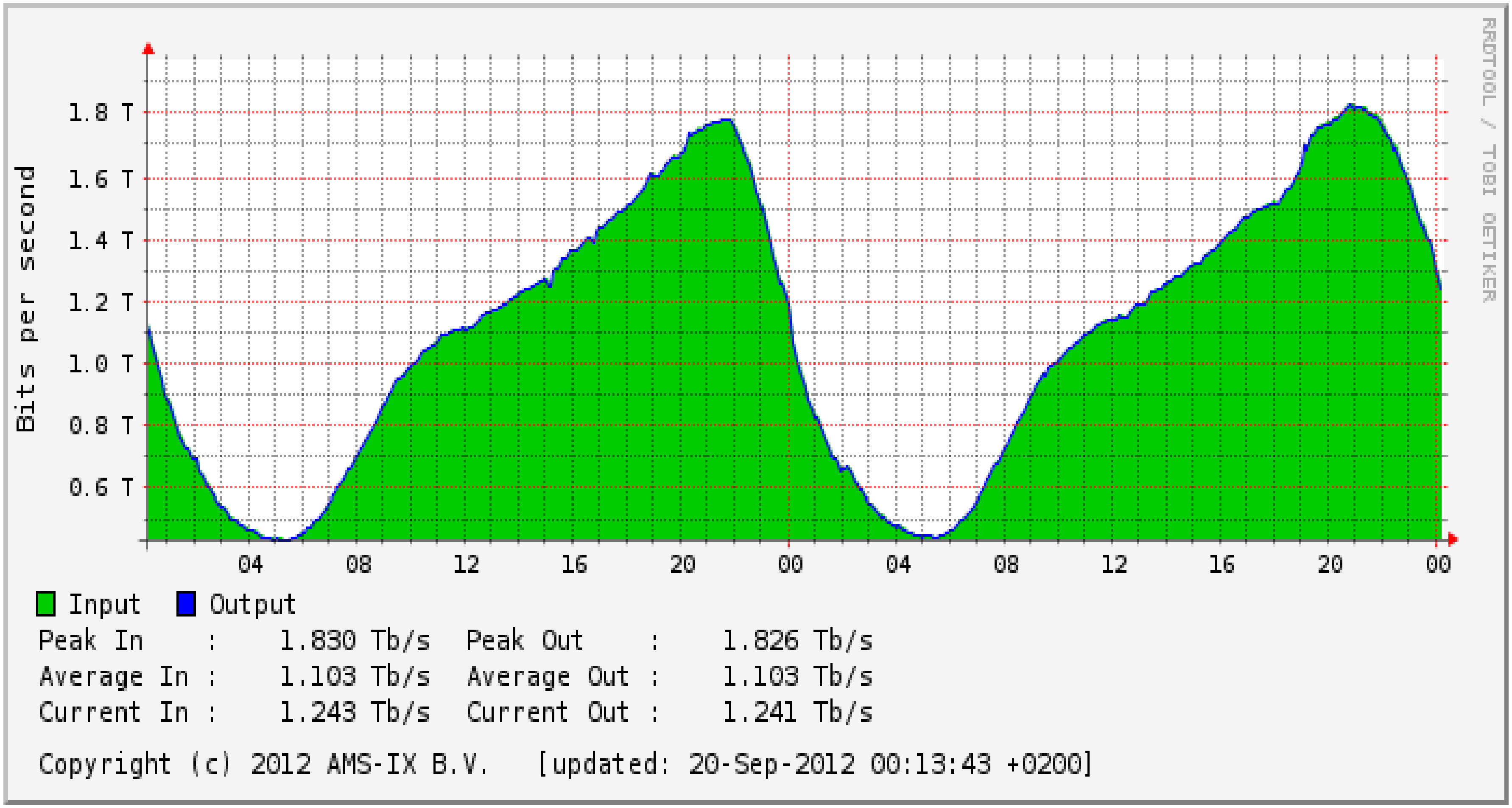
Interface	Capacity (Mbps)	In-bound Utilization (%)	Out-bound Utilization (%)
IP			
AOFA->BNL	11000	98.73	41.37
AOFA->BOST	10000	33.93	71.95
AOFA->WASH	20000	27.24	44.62

weathermap.es.net

- **High-speed networks have periods of low utilization**
  - **It is a matter of scale**
- **Consider using Low Power Idle (LPI)**
  - **Trade off is increased latency variation**

Why continue to develop standards for EEE?

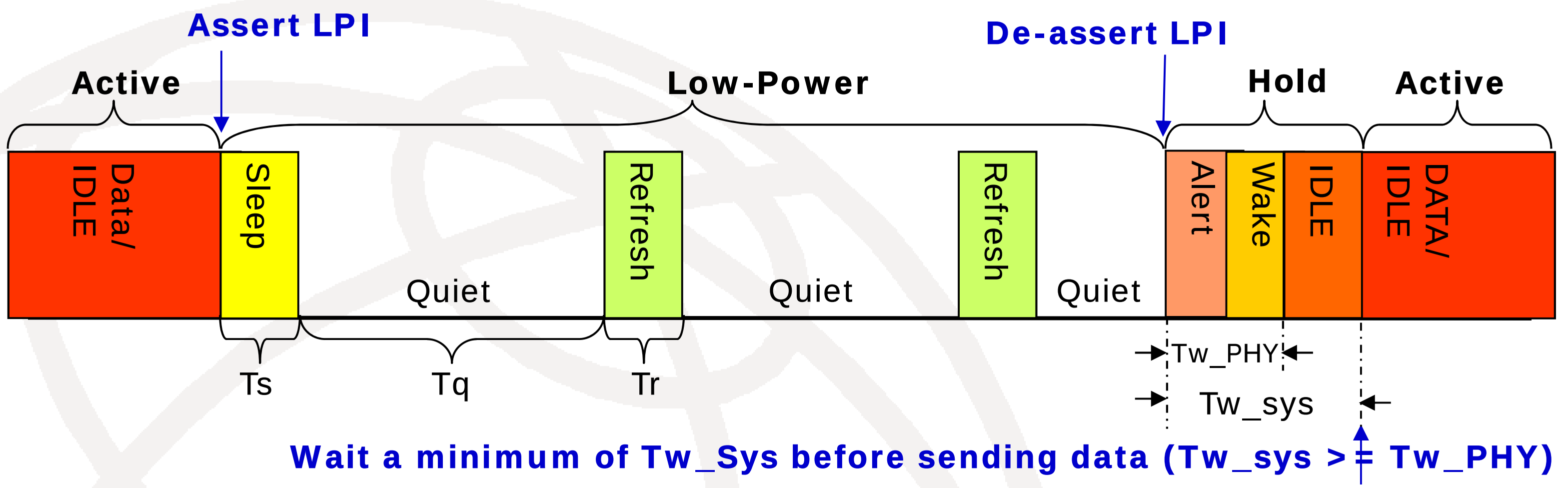
Daily graph



Aggregate Traffic at AMS-IX (24 hours)

Source: https://www.ams-ix.net/statistics/

Low-Power Idle (LPI)



- Low-Power Idle (LPI): PHY powers down during idle periods (does NOT send IDLE symbols)
- During power-down, maintain coefficients & synchronization to allow rapid return to Active state
- Asymmetric and Synchronous Modes
  - Asymmetric: One direction can enter a quiet state independent of the other
  - Synchronous: Both directions enter and leave the low power state together (1000 Mb/s is synchronous)
  - This is a PHY layer distinction. Higher layers may be asymmetric
- Wake times for the respective twisted-pair PHYs:
  - 100BASE-TX: Tw_PHY <= 30 usec
  - 1000BASE-T: Tw_PHY <= 16.5 usec
  - 10GBASE-T: Tw_PHY <~ 8 usec (2 modes)
- **PHY power in LPI mode ~20-40% of normal (depends on type and implementation)**

EEE progress in IEEE

- P802.3bj 100G Backplane and Copper Cable
  - Changed scope to include first generations PHYs
  - In Task Force review (D1.1)
  - Developing Fast wake (sub microsecond wake times)

- P802.3bm 40 Gb/s and 100 Gb/s Operation Over Fiber Optic Cables
  - Preparing to discuss changes similar to P802.3bj
  - Fast wake only (no PMD shutdown)
  - No auto-negotiation

EEE progress in IEEE

- Reduced Twisted-Pair Gigabit Ethernet
  - Added objective to do EEE
  - Study Group is preparing for Task Force phase

- P802.3bn EPON Protocol over Coax
  - EEE (in some form) being discussed

- Next Generation BASE-T
  - First Study Group Meeting
    - Just starting the conversation

Supporting Efforts

- Examining EEE
  - **Suggestion for improved efficiency though a buffer and burst policy and a look at how EEE might work for 100G Optical Ethernet**
  - **Reference**: P. Reviriego, K. Christensen, J. Rabanillo, J.A. Maestro, "**An Initial Evaluation of Energy Efficient Ethernet**", IEEE Communications Letters (ISSN: 1089-7798), Vol. 15, No 5, May 2011, pp. 578-580.
  - P. Reviriego, B. Huiszoon, V. López, R.B. Coenen, J.A. Hernández, J.A. Maestro, "**Improving Energy Efficiency in IEEE 802.3ba High-Rate Ethernet Optical Links**", IEEE Journal of Selected Topics in Quantum Electronics (ISSN 1077-260X), Vol. 17, No 2, March/April 2011, pp. 419-427.

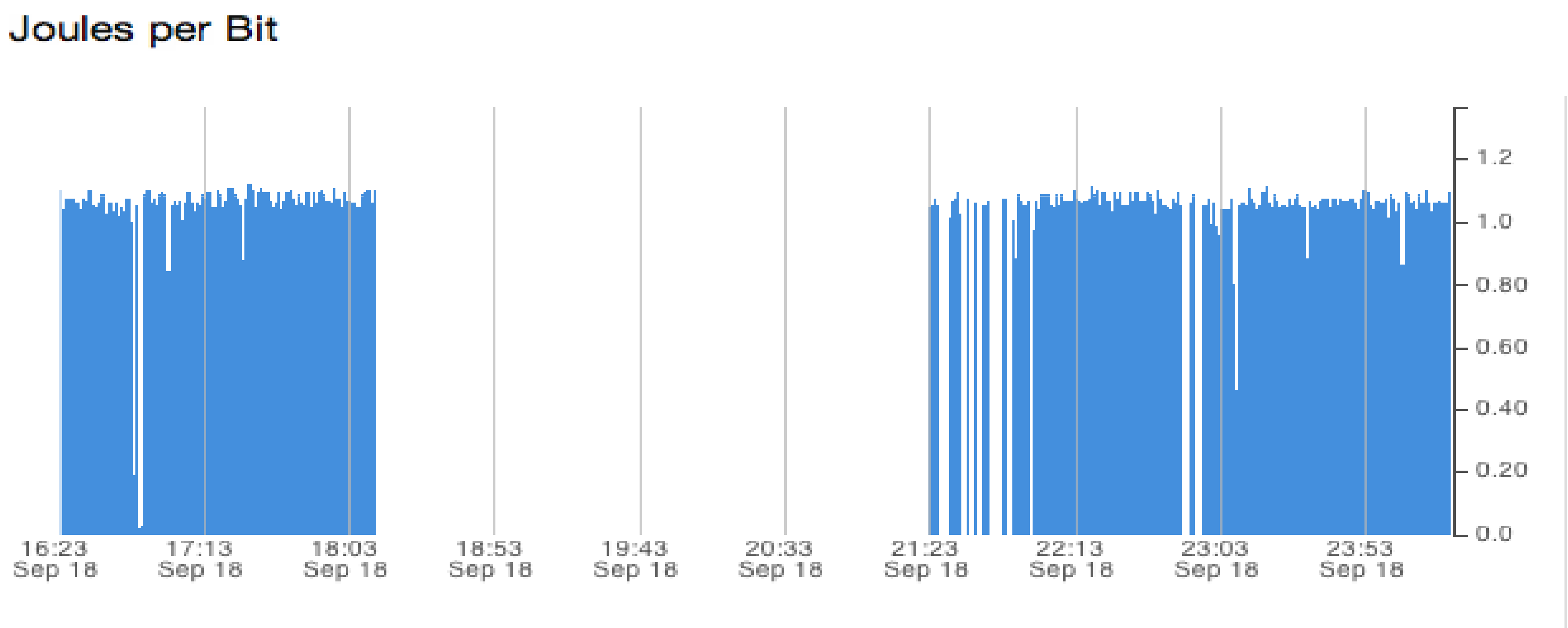
Supporting Efforts

- **Models being developed**
  - **May help evaluate proposals**
  - **Reference:** Marsan, M.A.; Anta, A.F.; Mancuso, V.; Rengarajan, B.; Vasallo, P.R.; Rizzo, G.; , **"A Simple Analytical Model for Energy Efficient Ethernet**," Communications Letters, IEEE , vol.15, no.7, pp.773-775, July 2011

- **Baseline network energy use**
  - **Energy Sciences Network (ESnet) is developing measurement infrastructure to collect network energy use baseline data**
    - **Reference:** http://www.es.net/RandD/green-networking/ https://my.es.net/ani/power

Supporting Efforts

- **Baseline network energy use**
  - **Measuring the power being used by the entire network to do "useful work" moving bits of information from ingress to egress**
  - **We're interested in measured network power / energy-use**
    - **Anyone else collecting this kind of data?**



Summary

- **EEE is a necessary tool in the ICT energy-reduction toolbox**
  - **Feature is becoming more available as projects progress**
  - **It does not solve all of the ICT energy-use problems**
  - **Lowell provided a good example of feature capability with the EEN and control policy**

- **More work to do**
  - **Many in the ICT industry are doing their part**
  - **Can use more help, for example**
    - **More control policy development**
    - **More baseline network energy measurements**
    - **? (your great idea goes here ...)**

Thank you!