Energy-Efficiency in Wired Communication Networks

Anne-Cécile Orgerie
ENS de Lyon – RESO – LIP, France
annececile.orgerie@ens-lyon.fr

www.cost804.org
Action IC0804

10th March 2011, COST Training School, Brasov - Romania

ENS DE LYON

Why do we need to be Green?

- The number of Internet users has been multiplied by 5 between 2000 and 2009. (http://www.internetworldstats.com)
- "Transmitting data through Internet takes more energy (in bits per Joule) than transmitting data through **wireless** networks."
  (Gupta & Singh – *Greening of the Internet* – SIGCOMM 2003)
- "By 2015, **routers** will consume 9% of Japan's electricity." (Michiharu Nakamura (Hitachi) - Nature Photonics Technology Conference 2007)

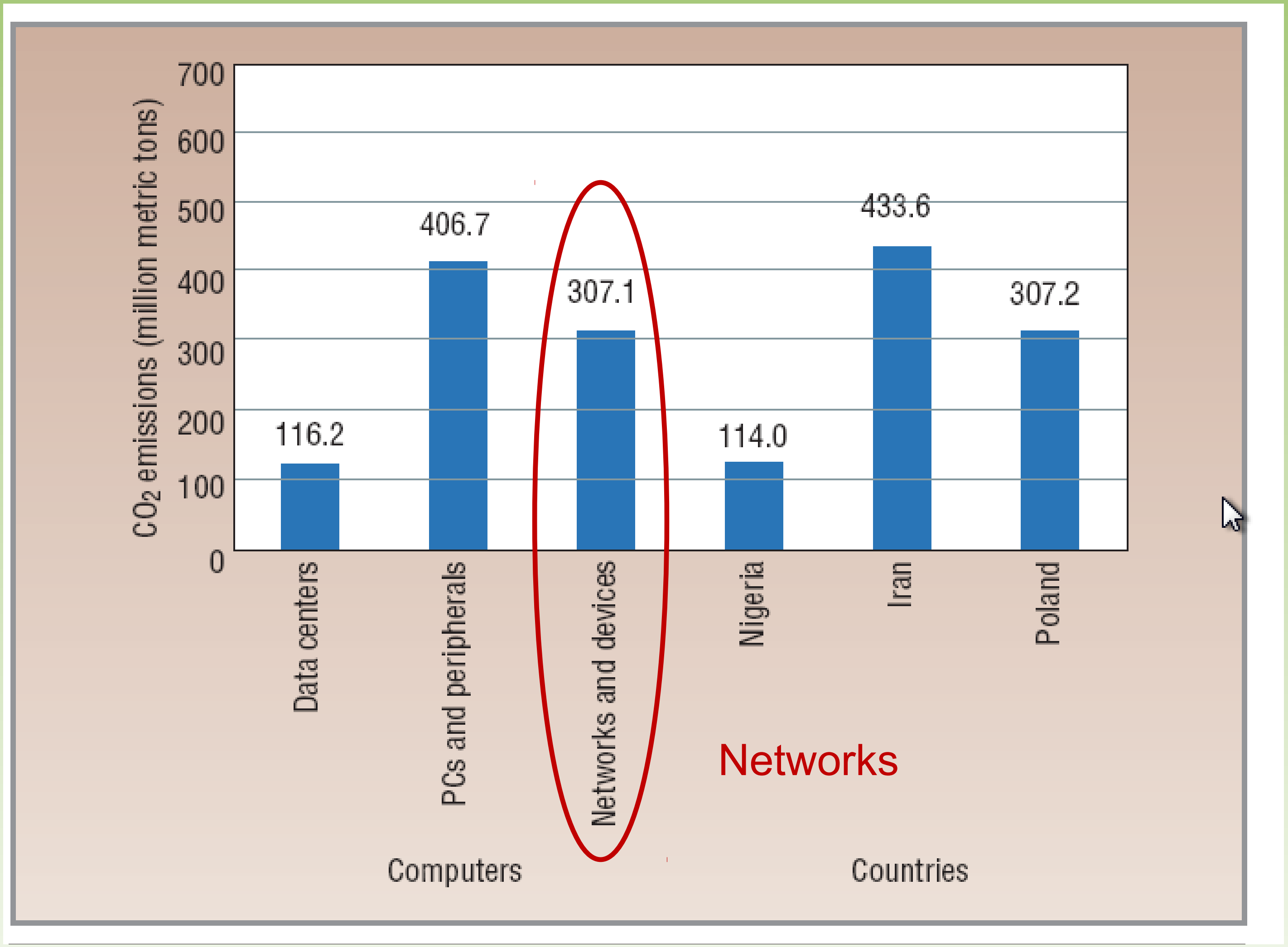
Why do we need to be Green?

“The global information and communications technology (ICT) industry accounts for approximately 2 percent of global carbon dioxide (CO_2) emissions, a figure equivalent to aviation.”

- Gartner Group, Inc. (2007)

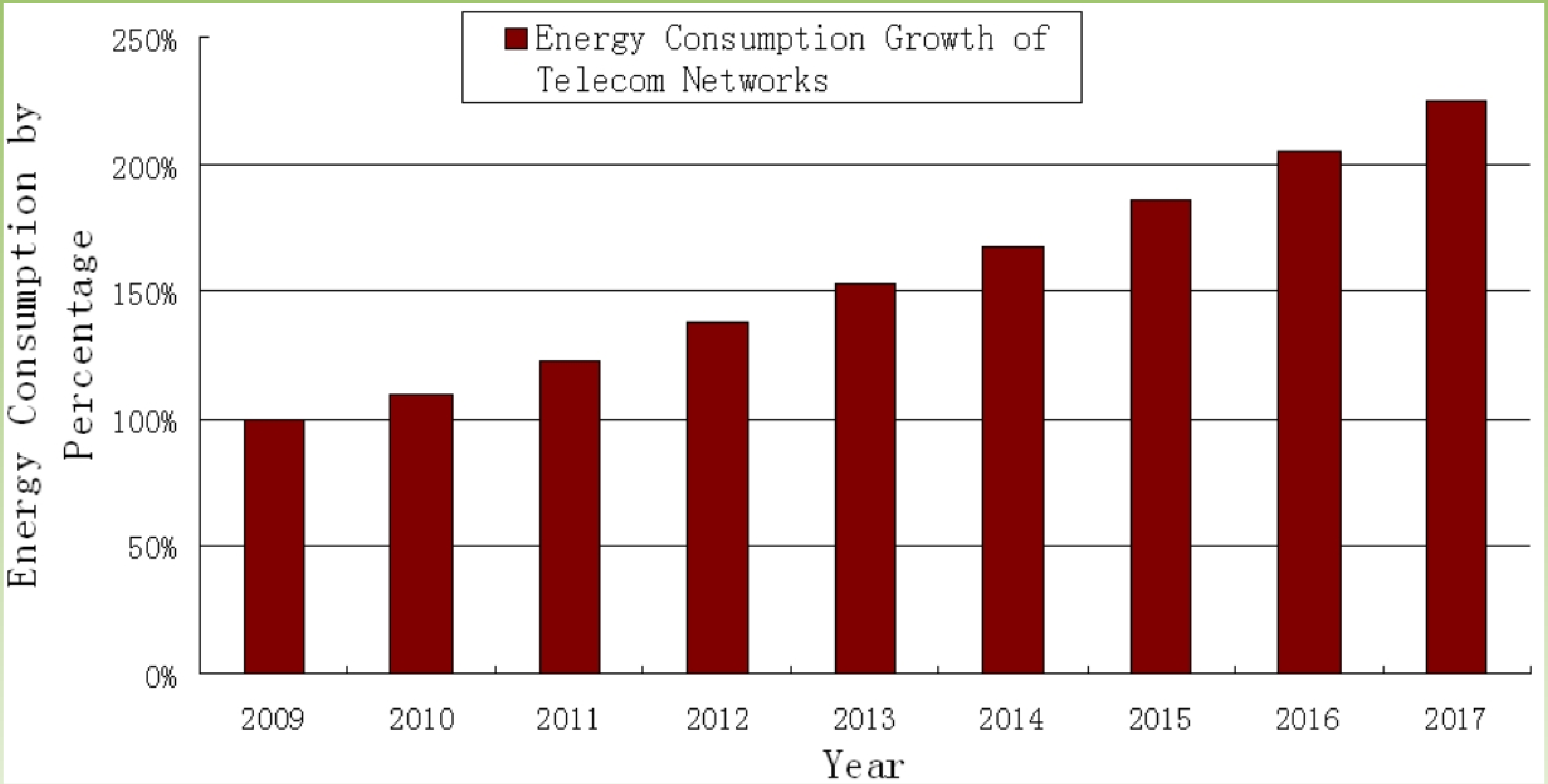
→ ICT use grows faster than airline traffic

The 2%



"*Some Computer Science Issues in Creating a Sustainable World*" J. Mankoff, IEEE Computer, 2008 .

Energy consumption forecast

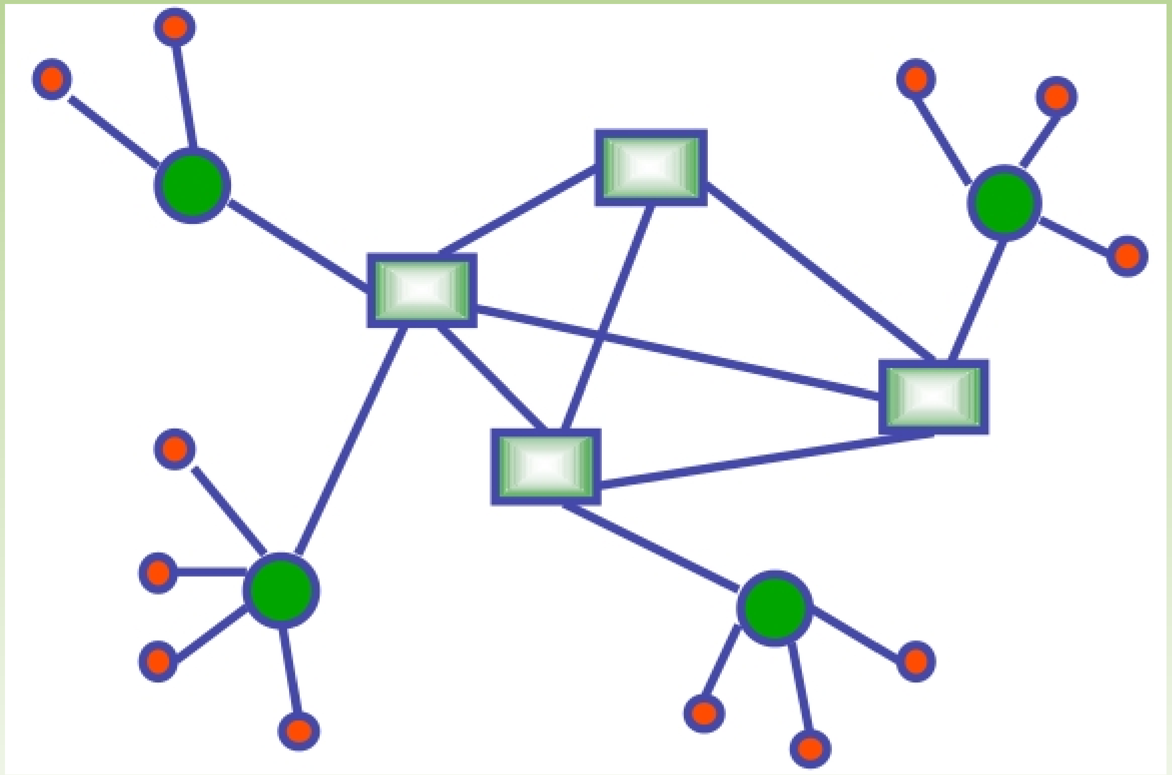


→ Increase in demand, increase in infrastructure

"*Energy Efficiency in Telecom Optical Networks*", Y. Zhang et al., IEEE Communications Surveys & tutorials, 2010.

Network Architecture

- Core (Backbone), Metro, Access Networks



Energy consumption

(A) 2015-2020 NETWORK FORECAST: DEVICE DENSITY AND ENERGY REQUIREMENTS IN THE BUSINESS-AS-USUAL CASE (BAU). EXAMPLE BASED ON THE ITALIAN NETWORK.

	power consumption [W]	*number of devices* [#]	*overall consumption* [GWh/year]
Home	10	17,500,000	1,533
Access	1,280	27,344	307
Metro/Transport	6,000	1,750	92
Core	10,000	175	15
		Overall network consumption	1,947

"*The potential impact of green technologies in next-generation wireline networks*", R. Bolla et al., IEEE Communications, 2011.

US year 2000

"*Greening of the Internet*", M. Gupta and S. Singh, SIGCOMM 2003

Device	Est. Numbers	Cons. TWh	Costs	Equiv.
Hubs	93.5 Mio	1.6		
LAN-Sw.	95.000	3.2		
WAN-Sw.	50.000	0.2		
Router	3.257	1.1		
Total (US)		6.1	10^9 US $	1 Nucl. Reactor
Total (World)		144		

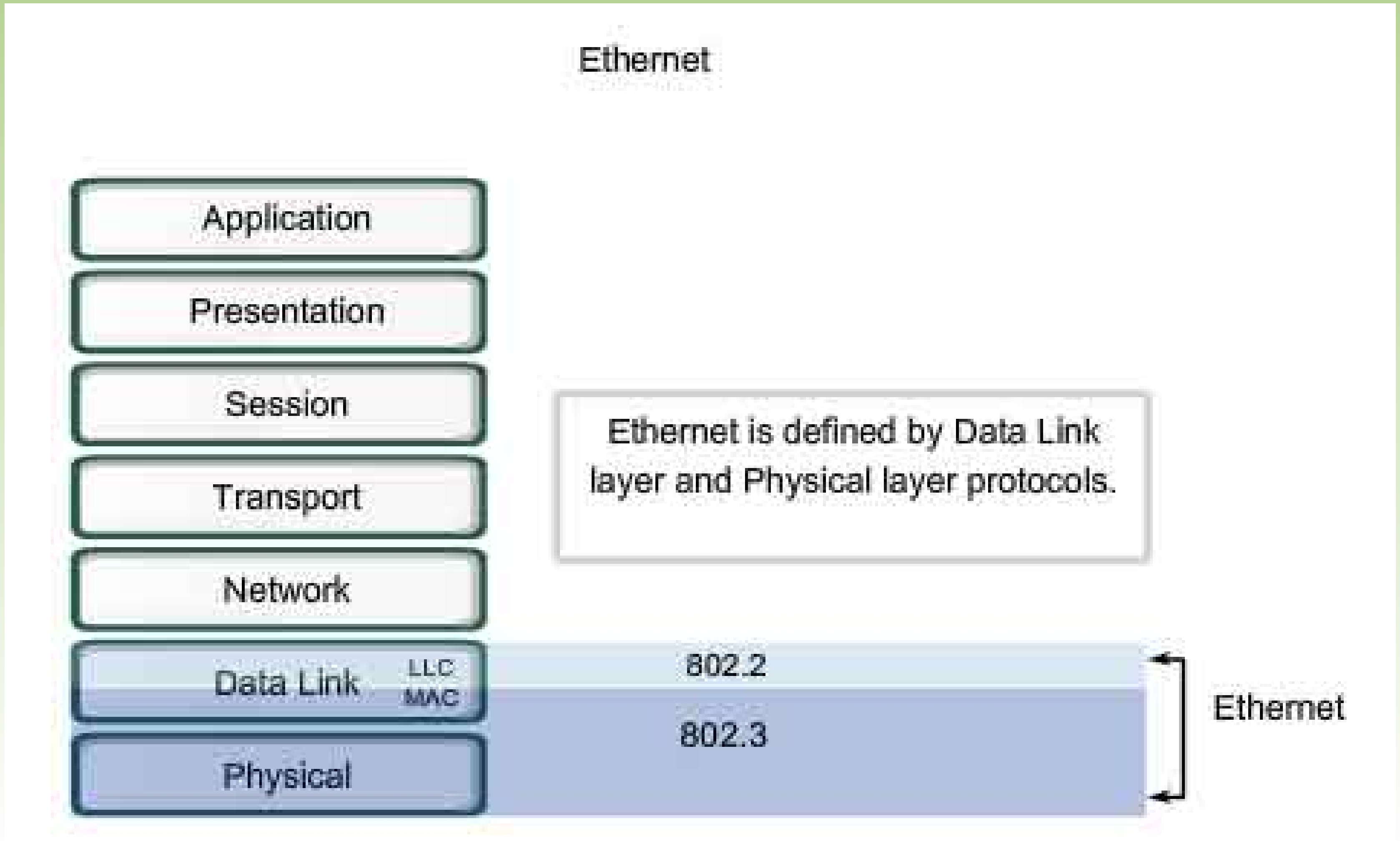
Energy consumption

(B) TRAFFIC AND TOPOLOGICAL DATA (AVERAGE FIGURES) FOR THE 2015-2020 PERSPECTIVE NETWORK. SOURCE: TELECOM ITALIA.

Home access	*number of customers per DSLAM*	640
	usage of a network access (user up-time)	30%
	link utilization when a user is connected	10%
Core, transport and metro	*redundancy degree for metro/transport devices* (η_t)	*13%*
	redundancy degree for core devices (η_c)	100%
	redundancy degree of metro/transport links (ψ_t)	100%
	redundancy degree of core device links (ψ_c)	50%
	link utilization in metro networks	40%
	link utilization in core networks	40%

Ethernet

- IEEE 802.3
- MAC: Media Access Control
- LLC: Logical Link Control



Plan

1. Issues
2. Energy-efficient techniques
3. Background
4. Bandwidth reservation scheduling
5. Conclusion and Perspectives

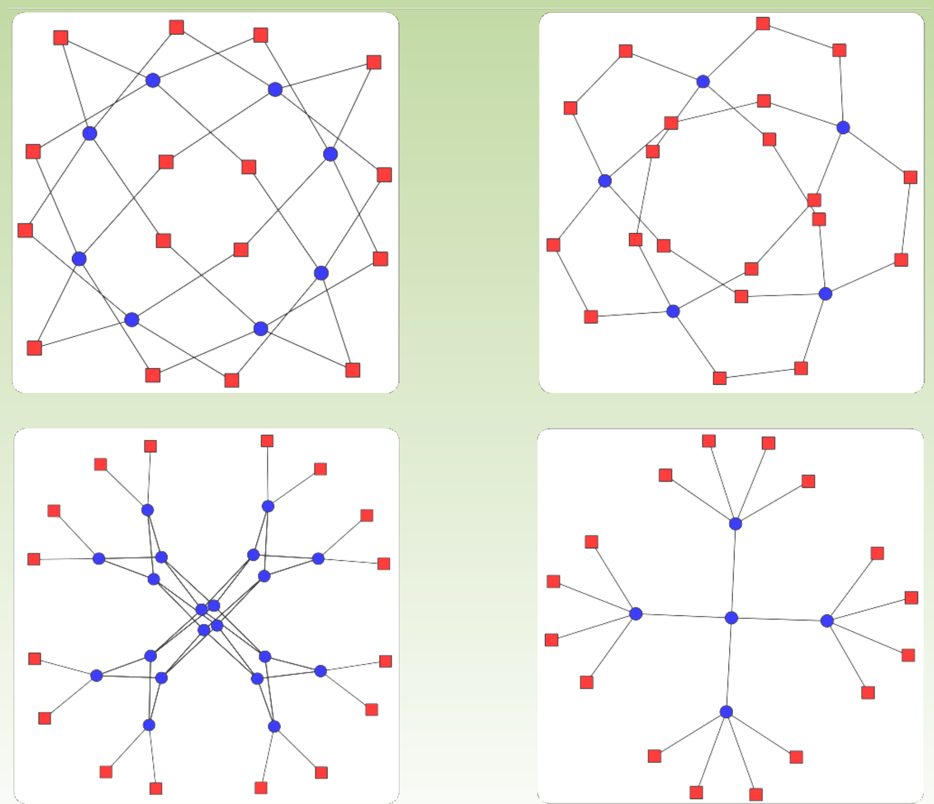
Issues

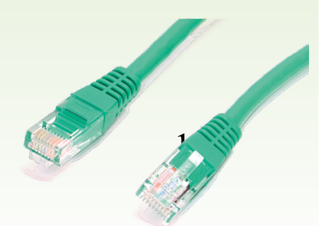
Design Issues

- Energy consumption is not taken into account in the **design** of wired network components.

  → high redundancy for fault tolerance

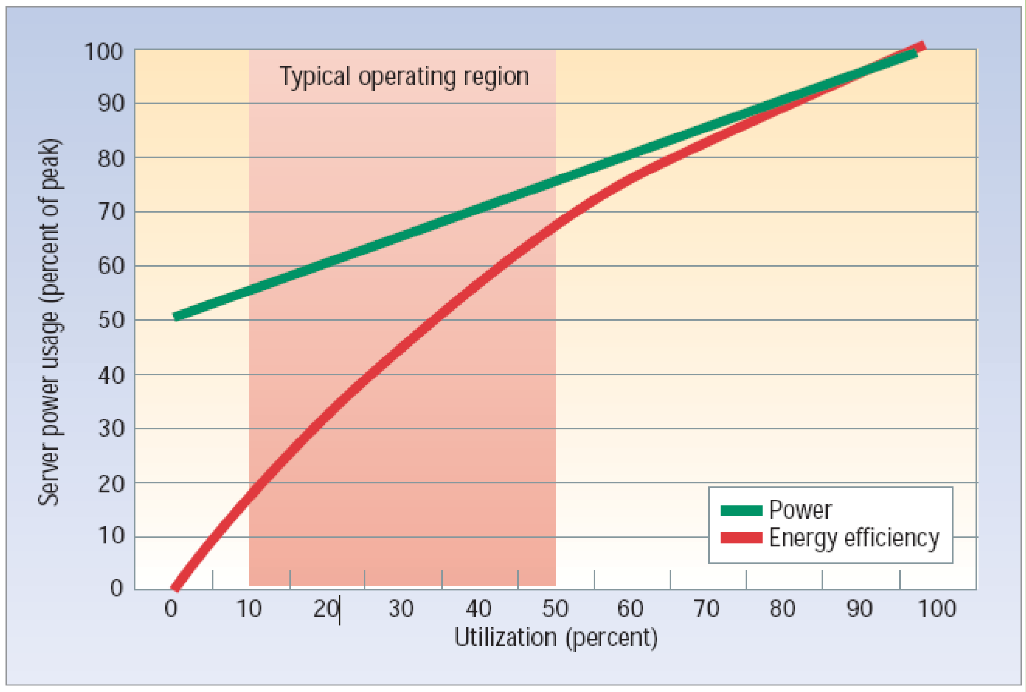
“*How can architecture help to reduce energy consumption in data center networking?*” L. Gyarmati and T. Anh Trinh, e-Energy 2010.

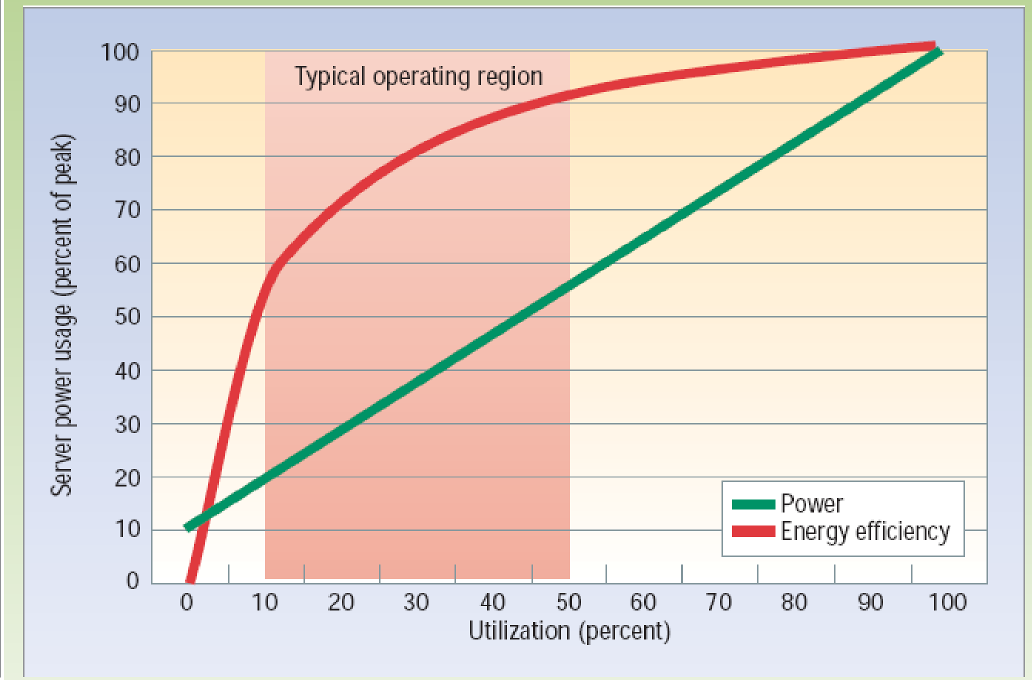




Proportionality Issue

- Energy consumption of network equipments is not **proportional** to usage (high base power).



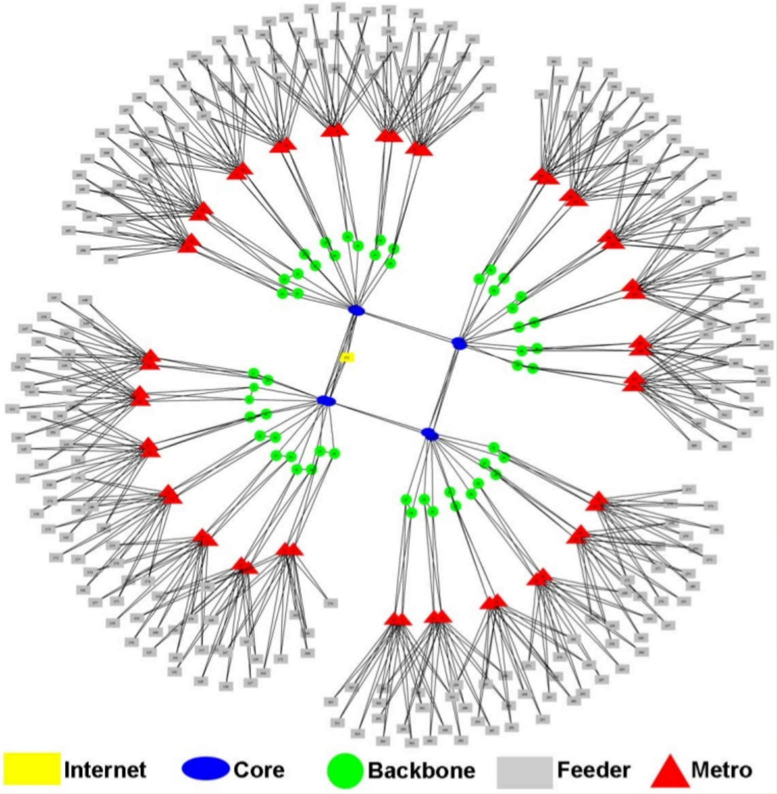


"*The case for energy-proportional computing*" L. Barroso and U. Hölzle, Computer, 2007.

Consumption Issue

- Network equipments stay always **fully on** even if unused.

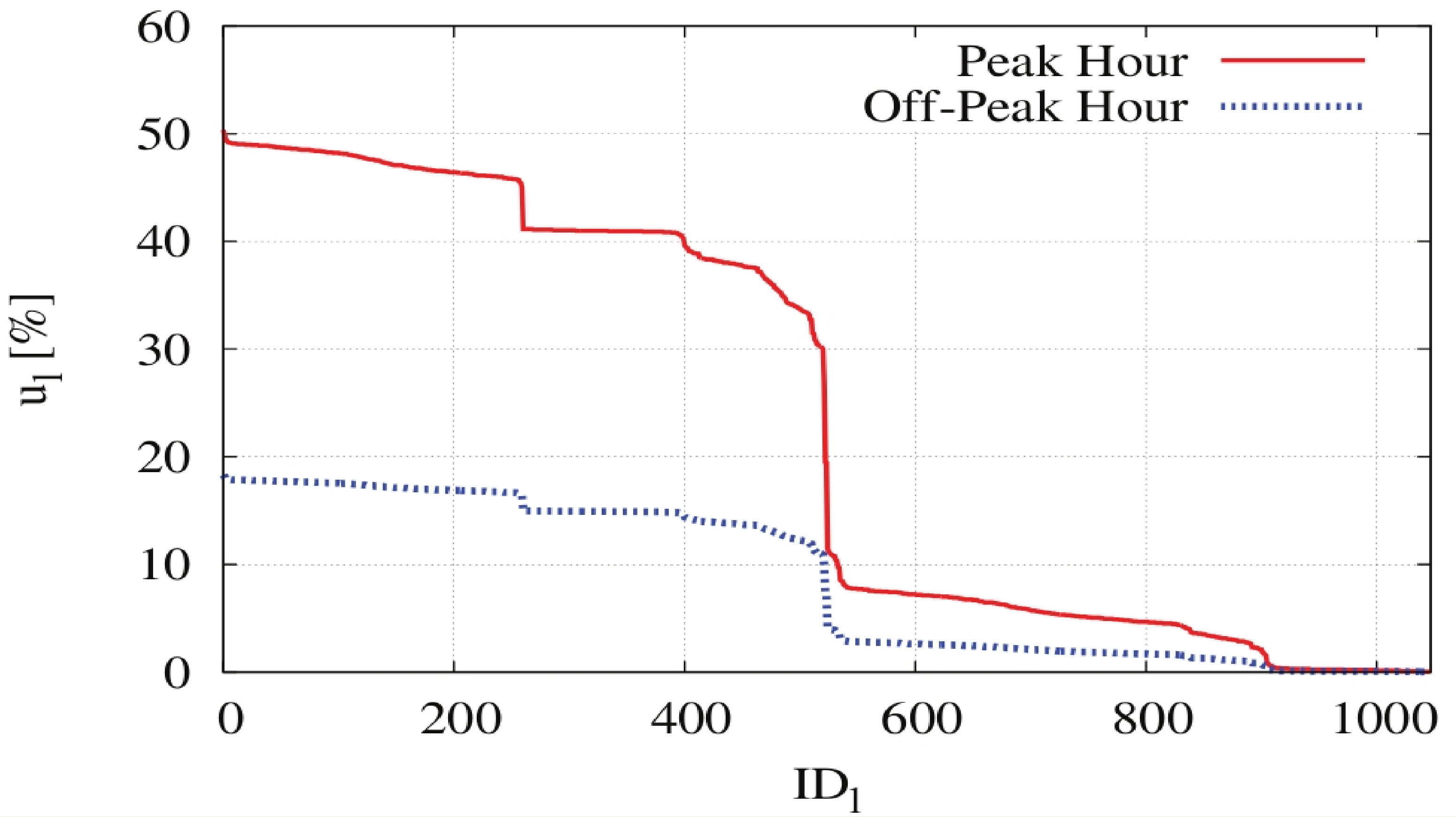
Node Type	Power [kW]	Fraction of Total Node Power
Core	10	9.46%
Backbone	3	19.03%
Metro	1	6.32%
Feeder	2	65.19%



Traditional	Energy-Aware	Saving
12.47 GWh/year	9.54 GWh/year	23.5%

"*Energy-Aware Backbone Networks: a Case Study*"
L. Chiaraviglio et al., GreenComm Workshop, 2009.

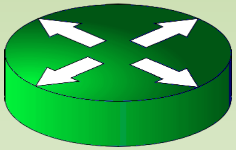
Networks are Lightly Used



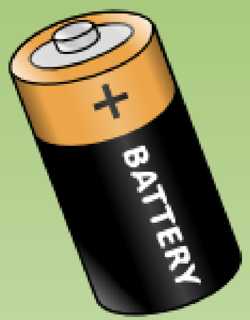
"Data Networks are Lightly Utilized, and Will Stay That Way" O. Andrew, 1999.

"*Energy Saving and Network Performance: a Trade-off Approach*" C. Panarello et al., e-Energy 2010.

Energy-Efficient Techniques



From battery-based networks



Battery based solutions :
- Ad-hoc networks
- Sensors networks
- Mobile devices

→ adapted softwares, environments, protocols and services

To take it or not ?

ex: Energy-aware routing protocols : averaging energy usage in sensors infrastructures to avoid unbalanced energy usage on specific nodes

State of the Art: 4 main approaches

1. Optimizing hardware components

2. Shutdown approach: put components to sleep

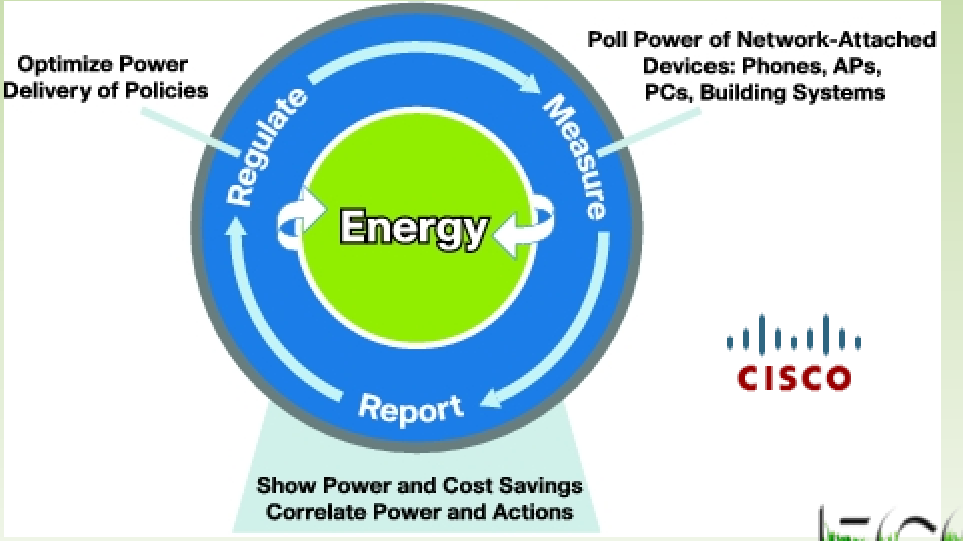
3. Slowdown approach: adaptation to the needs

4. Coordination approach: network-wide management and global solutions

Initiatives and consortiums

- Network equipment manufacturers start to launch green initiatives and green products.

ECR INITIATIVE™ *The Energy Consumption Rating (ECR) Initiative*
a framework for measuring the energy efficiency of network and telecom devices.



IEEE 802.3az Task Force

- Cisco, LBNL, Broadcom, Teranetics, Intel, Fujitsu, HP, IBM, Nexans, LSI, ...
- 3 billions of Ethernet cards in the world
- Adaptive Link Rate
- Low Power Idle
- Standard adopted en 2010

Energy Efficient Ethernet

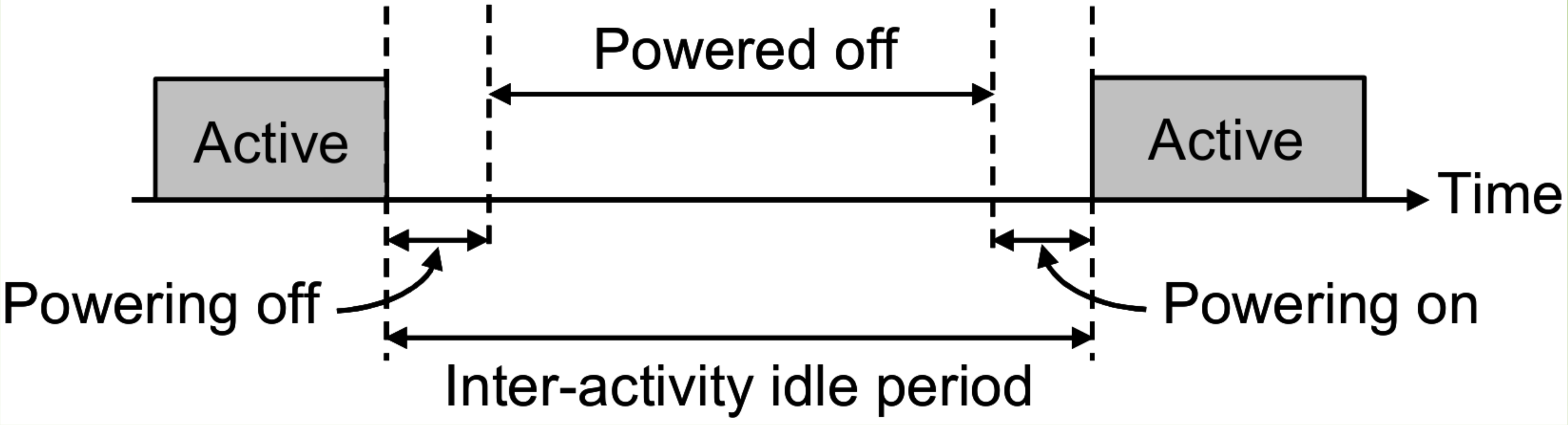
"*Achieve Energy Efficiency in Ethernet Receivers* " P. Reviriego et al., 2009.

Optimizing hardware components

- Power saving by cable length
- Smart cooling FAN
- Power on/off button
- Wake On LAN
- Low power modes
- Port auto power down

The shutdown approach

- Network links are lightly utilized
  → switch off unused equipments (sleeping mode)



Sleeping

- **Uncoordinated**: link layer approach where an interface sleeps based on local decisions alone.
- **Coordinated**: network-wide approach in which the routing protocol aggregates traffic into few routes during low load to explicitly enable some interfaces to sleep.

  → consider topologies that allow for adaptation

  → can save energy of end-hosts too

"*Greening of the Internet*", M. Gupta and S. Singh, SIGCOMM 2003

Issues with the shutdown approach

- The connectivity is lost → connection-oriented applications, rerouting strategies, proxying
- It takes time and energy to switch on and off → loss of reactivity, energy benefits?
- When to switch off? → time threshold?
- How to switch on again? → periodic wake up? Magic packet?

Impact on routers

- Modify routing protocols to not treat link sleeping as link failure
- Restrict Hello messages to awake interfaces only
- Use an algorithm that identifies minimal number of links needed while satisfying QoS needs of the supported flows (non-trivial, costs time)
- Routers need to be able to predict low load and know when to go to sleep (and incorporate this into routing protocols)

Proxying Ethernet Adapter

Such a proxy might be hosted within a NIC and would:

1) Cover for desktop computer for simple routine protocol events

2) Wake-up desktop computer when its resources are needed

3) Possibly exchange state information with computer OS

→ NIC can be powered-on when computer sleeps

"*The next frontier for communication networks: power management* ",
K. Christensen et al., Computer Communications, 2004.

Proxying Ethernet Adapter

Messages arriving to an idle PC

Packet type	# in trace	% in trace
SNAP	835	12.55
ARP	2143	32.20
BRIDGE	882	13.25
DHCP	433	6.51
NetBIOS Datagram	568	8.53
NetBIOS Name service	483	7.26
OSPF Hello	177	2.66
SAP	302	4.54
Remote cache service	69	1.04
IGMP	150	2.25
Groove	88	1.32

"*The next frontier for communication networks: power management* ", K. Christensen et al., Computer Communications, 2004.

Modify routing algorithm to allow on/off

- Modification of OSPF

  1) Election of “exporter” routers, they calculate their shortest path trees (Dijkstra);

  2) “Importer” routers use the exporter trees to compute their path trees and fix the links that have to be switched off;

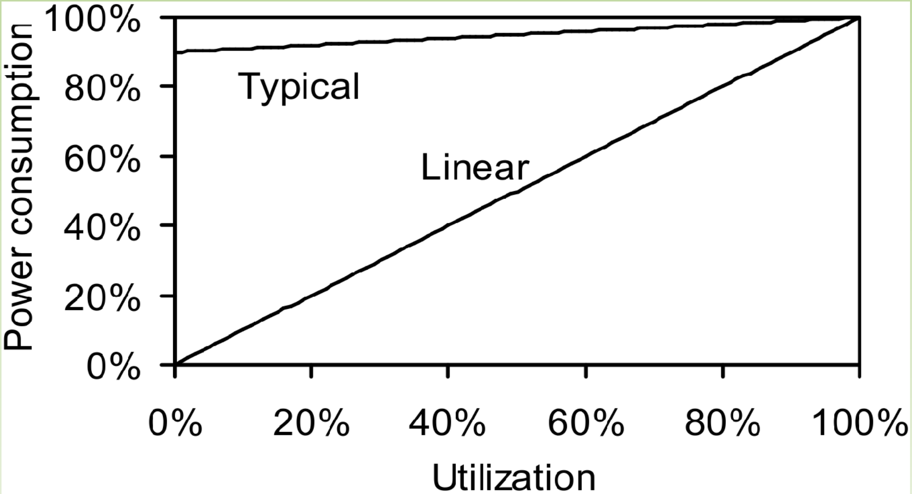
  3) Each router computes its paths using Dijkstra algorithm on the residual network topology.

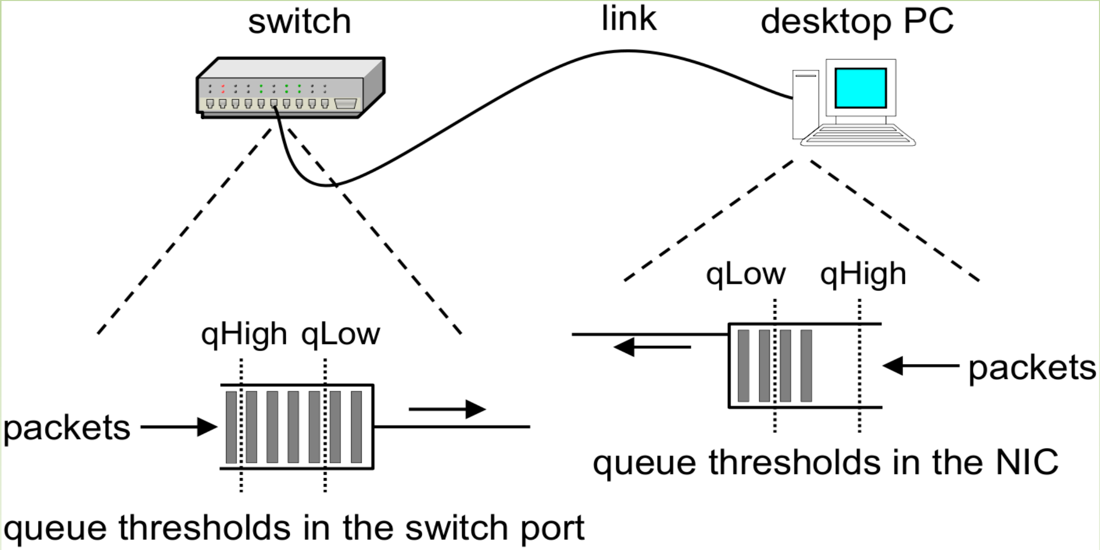
“*An energy saving routing algorithm for a green OSPF protocol* ”,
A. Cianfrani et al., INFOCOM Workshop, 2010.

The slowdown approach

- Energy consumption is not proportional to the usage.

→ Adjust the link rate to the actual traffic: ALR

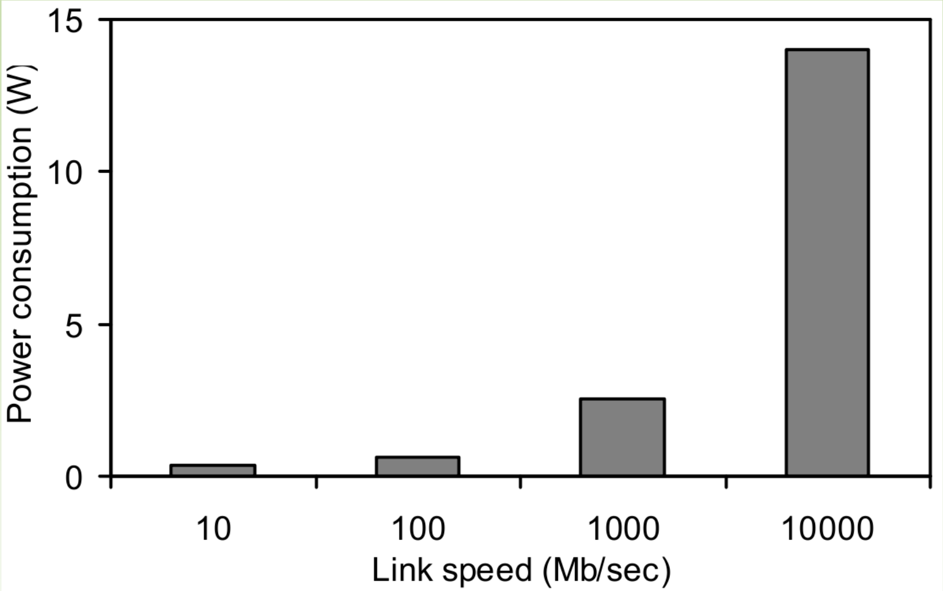




"*Managing Energy Consumption Costs in Desktop PCs and LAN Switches With Proxying, Split TCP Connections, and Scaling of Link Speed* ", C. Gunaratne et al., International Journal of Network Management, 2005.

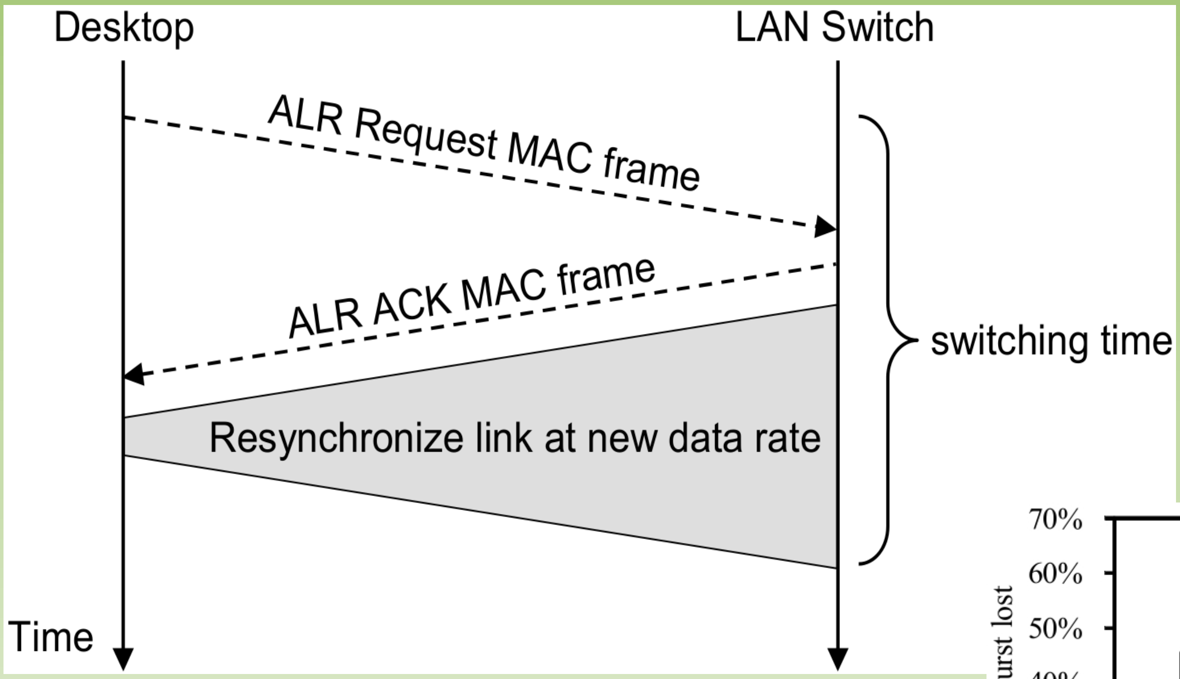
Issues with the slowdown approach

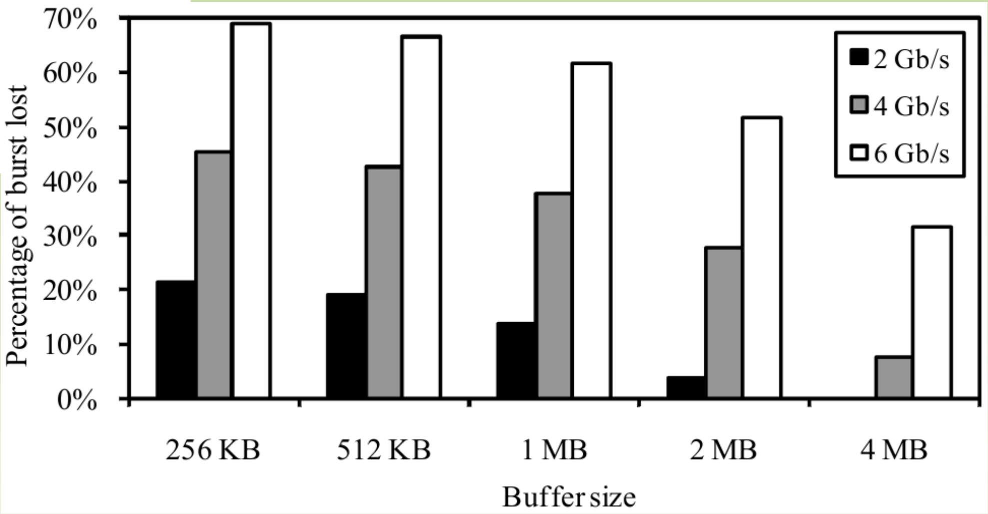
- Rate switching takes time (synchronization). → packet loss, lower reactivity
- How to define the thresholds? → Markov models, live adaptation?
- Oscillations



Ethernet NIC

Switch Data Rate





"*An Initial Performance Evaluation of Rapid PHY Selection (RPS) for Energy Efficient Ethernet*" F. Blanquicet, LCN 2007.

ADSL2

- ADSL: Asymmetric Digital Subscriber Line
- Several power modes

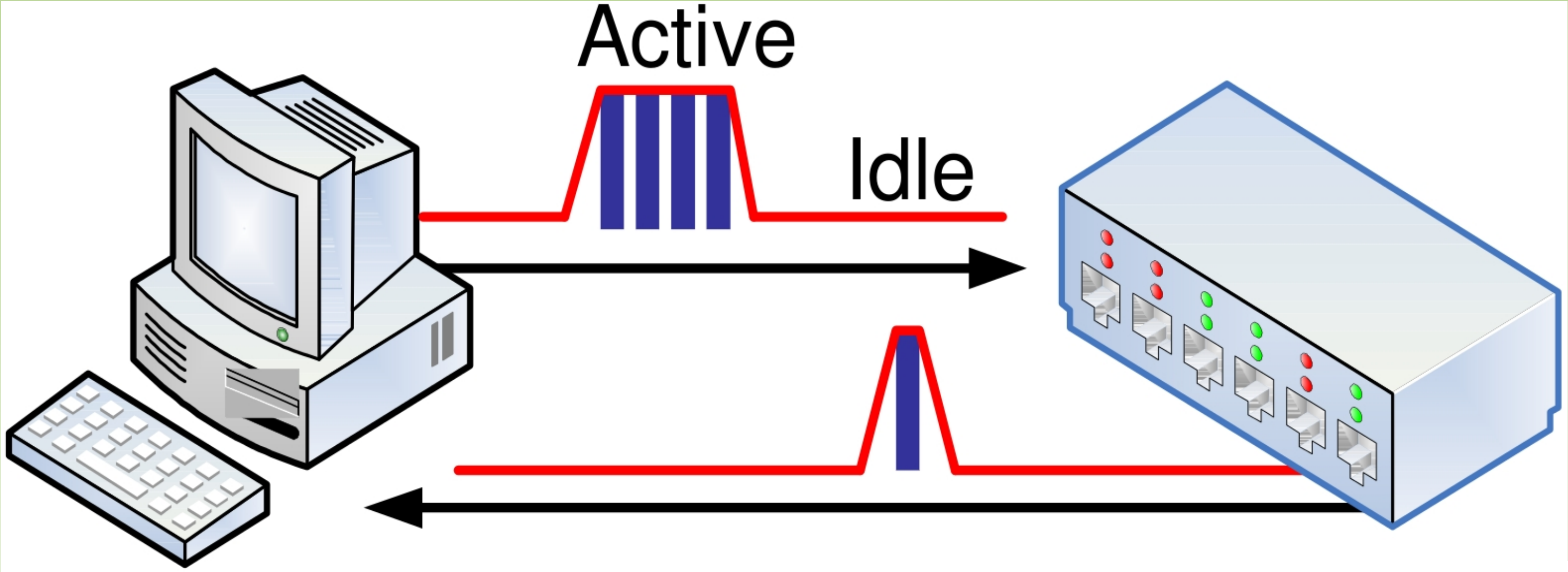
State	Name	Description	
L0	Full On	The ADSL link is fully functional.	3 W
L2	Low Power	The ADSL link is active, but a low-power signal conveying background data is sent from the ATU-C to the ATU-R. A normal data carrying signal is transmitted from the ATU-R to the ATU-C.	0.5 W
L3	Idle	There is no signal transmitted on the line. The ATU may be powered or unpowered in L3.	0.3 W

- No standardized policy to determine when to switch mode → capabilities not used

"*Low Power Modes for ADSL2 and ADSL2+*" G. Ginis, Broadband Communications Group White Paper (Texas Instrument), 2005.

Low Power Idle

- Transmit data as fast as possible, return to low power idle



“*Introduction to Energy Efficient Ethernet*” A. Healey, 802.3az meeting, 2010.

Low Power Idle

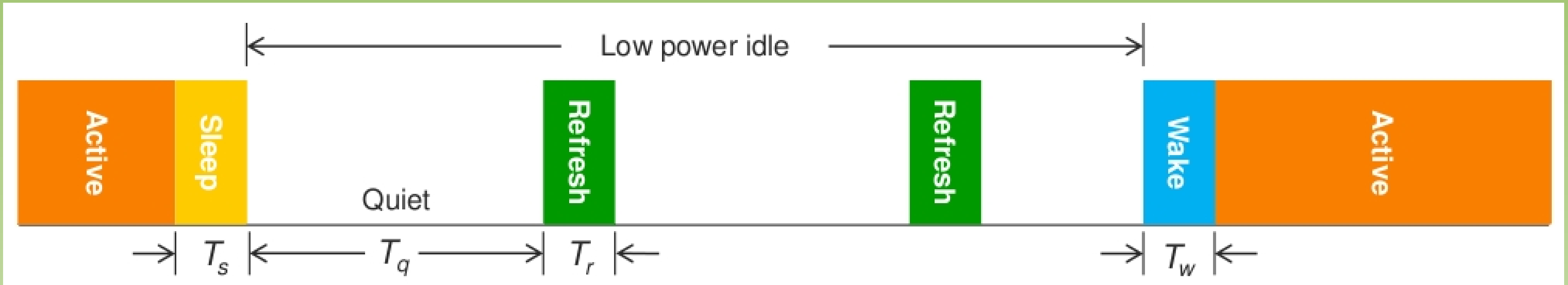
- Transmit data as fast as possible, return to low power idle

  - highest rate provides the most energy-efficient transmission (Joules/bit)

  - low power idle consumes minimal power (Watts)
- Power is reduced by turning off unused circuits (portions of PHY, MAC, interconnects, memory, CPU)
- Transmitter initiates low power idle transitions
- No loss

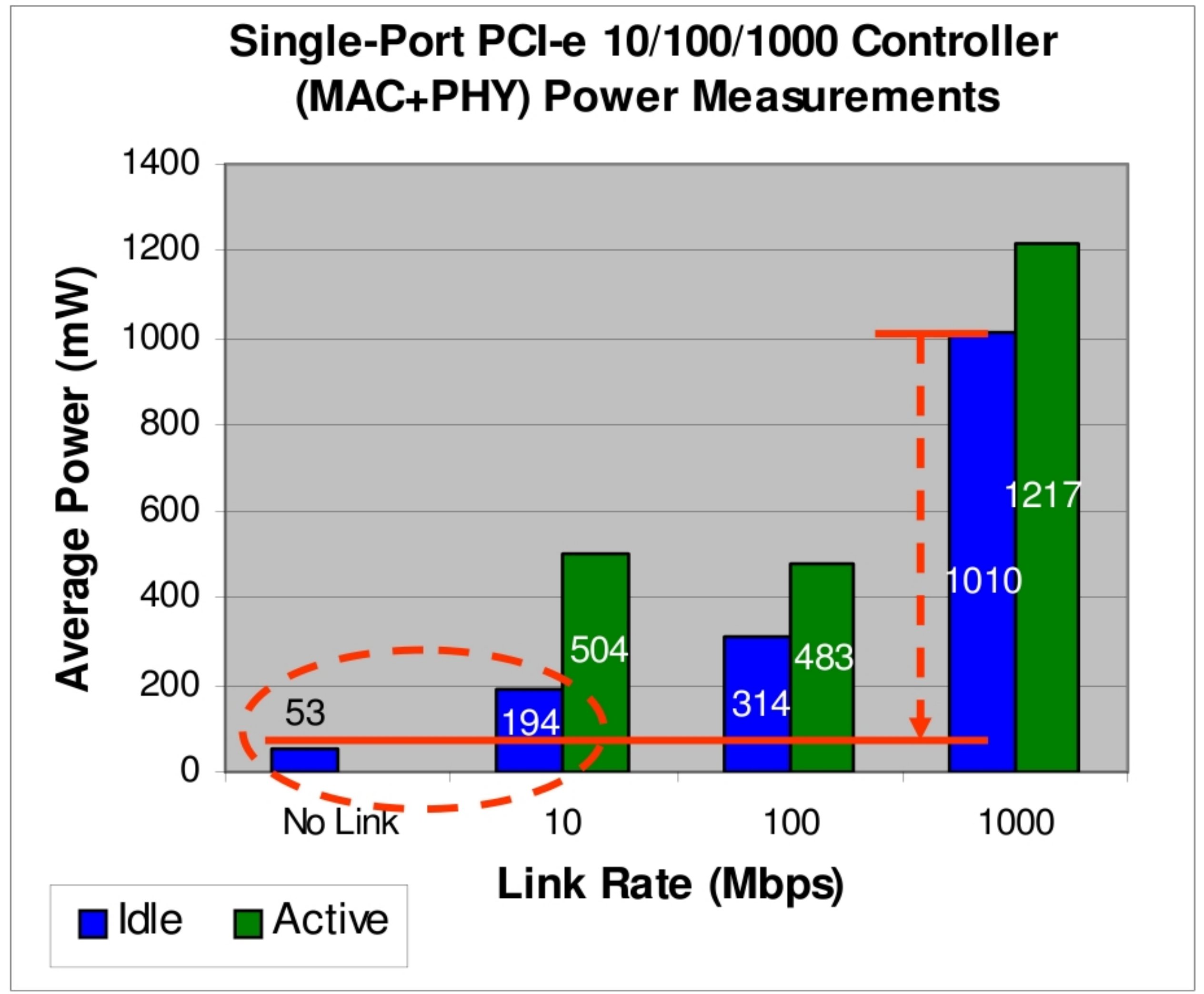
Energy Efficient Ethernet

Low Power Idle



- Periodical refreshment of the receiver (heartbeat to detect failures, facilitate fast transition
- Enable power saving in the PHY
- Transmissions are deferred to per-defined time
- T_w = 11 μs, T_s = 5 μs, T_q = 1.7 ms
- System may voluntarily increase the deferral time for deep sleep modes

Low Power Idle



Source: Intel labs. Intel® 82573L Gigabit Ethernet Controller, 0.13µm, “Idle” = no traffic, “Active” = line-rate, bi-directional

Energy Efficient Ethernet

Low Power Idle Expectations

“*IEEE 802.3az: The Road to Energy Efficient Ethernet*”
K. Christensen et al., IEEE Communications Magazine, 2010.

	1 Gb/s	10 Gb/s	Total
Assumptions			
Savings per link — no data (W)	1	5	—
Link utilization (%)	1%	3%	—
Active links (millions)	250	65	315
Electricity cost ($/kWh)	0.10	0.10	—
Results — EEE Savings			
Per link (%)	81%	82%	—
Per link (W)	0.81	4.10	—
Total (MW)	200	270	470
Total (TWh/year)	1.8	2.3	4.1
Total (million $/year)	180	230	410
Results — Ideal Savings			
Total (TWh/year)	2.2	2.8	5
Total (million $/year)	220	280	500

Global solutions

To be energy efficient, solutions should be:

- adapted to the **network topology** (redundancy, multi-paths)
- adapted to the **traffic** (bursts or always low, ...)
- adapted to the **scenario usage** (P2P, Web servers)
- realistic in terms of **technology** (compatibility, interoperability)
- scalable, reliable, fast, fault-tolerant, performant, secure, ...

Cooperation example

- *GreenCoop: Cooperative Green Routing with Energy-efficient Servers*, L. Chiaraviglio et al., e-Energy 2010.
- Cooperation between one ISP (Internet Service Provider) and one CP (Content Provider)
- ISP is the owner of the network infrastructure, while CP manages a set of servers connected to the network (replicated resources).
- Users ask for CP resources under QoS constraints.

Cooperation example

- Users ask for CP resources under QoS constraints.
- Pre-computation of two disjoint paths for each source-destination pair.
- Maximization of resources utilization (links and servers)
- 20% of energy savings

→ issues with security, privacy

→ scalability, adaptation to new servers, links

→ common objective function

→ crucial server placement

Centralization example

- *Optimisation of Power Consumption in Wired Packet Networks*, E. Gelenbe and S. Silvestri, Quality of Service in Heterogeneous Networks (Springer Book), 2009.
- Inspired by wireless approaches
- Centralized approach for routing and resource management
- Promising results related to QoS (delay, packet loss, jitter)
- Validation on a 44-router network → scalability ?

Centralization example

- Energy Management System:

1) Observe traffic flows and monitor node status and energy use

2) Select the network configuration that offered an acceptable level of QoS to ongoing and predicted flows with lower energy use

3) Manage and sequence dynamic changes in links and nodes and reroute traffic to reduce power consumption

Protocol Cost

- "*Computational Energy Cost of TCP*" B. Wang and S. Singh, INFOCOM, 2004.
- Analysis of the energy consumption of TCP on laptop with FreeBSD and Linux
- 60-70 % of the energy is due to the kernel-NIC copy operation
- 15% for the copy from user space to kernel space and 15% for TCP processing cost.
- Among these 15%, 20-30% accounts for the cost of computing checksums.

Clean-slate approaches

- IP was designed in the 1970s
- 200 connected hosts in 1980, about 570 million in 1998

“The basic idea of the clean-slate approach consists in temporarily ignoring the strong design constraints imposed by the existing Internet in order to better understand requirements and explore potential benefits. The objective is to define the ideal architecture that best supports present and anticipated future applications while correcting known problems with IP in relation to issues like security, QoS, and mobility.”

“*The clean-slate approach to future Internet design: a survey of research initiatives*” J. Roberts, Annals of Telecommunications, 2009.

Surveys

- "*Energy Efficiency in Telecom Optical Networks*" Y. Zhang et al., IEEE Communications Surveys & Tutorials, 2010.

- "*Energy Efficiency in the Future Internet: A Survey of Existing Approaches and Trends in Energy-Aware Fixed Network Infrastructures*" R. Bolla, IEEE Communications Surveys & Tutorials, 2011.

Background

"*Energy-Efficient Bandwidth Reservation for Bulk Data Transfers in Dedicated Networks* " A.-C. Orgerie et al., Journal of Supercomputing, to appear in 2011.

Bulk Data Transfers with Advance Reservations in Dedicated Networks

- **BDT (Bulk Data Transfers)** → large volumes of data to transfer, moldable/malleable, deadline
- **ABR (Advance Bandwidth Reservations)** → bandwidth provisioned for the transfer (no resource competition, no congestion)
- **Dedicated Networks** → private, managed by one entity, secure

Why dedicated networks are relevant

In **2007**, to distribute the entire collection of **Hubble telescope data** (about 120 Terabytes) to various research institutions, scientists chose to copy these data on hard disks and to send these hard disks via **mail**.
It was faster than using the network.

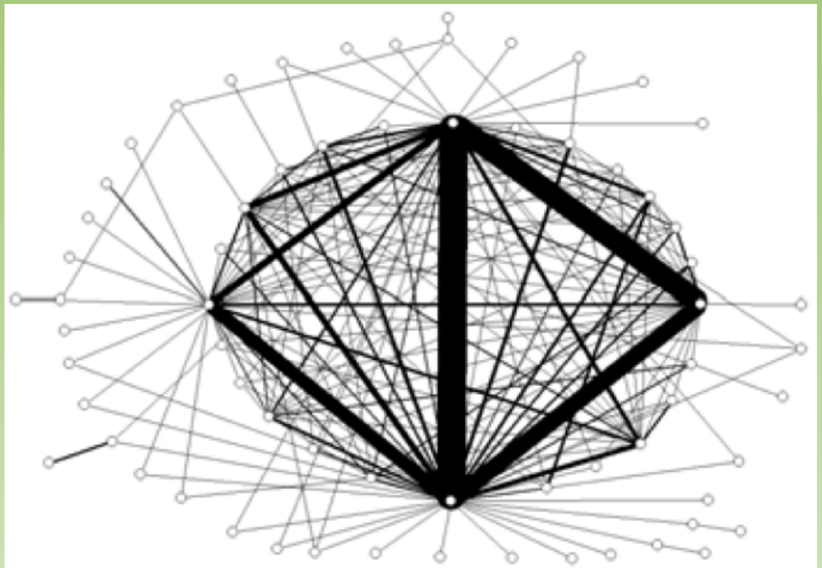
Cyrus Farivar. Google's Next-Gen of Sneakernet. [online] http://www.wired.com/science/discoveries/news/2007/03/73007, 2007.

The **Large Hadron Collider** (LHC) produces 15 million Gigabytes of data every year.

http://lcg.web.cern.ch/lcg/public/default.htm

Examples

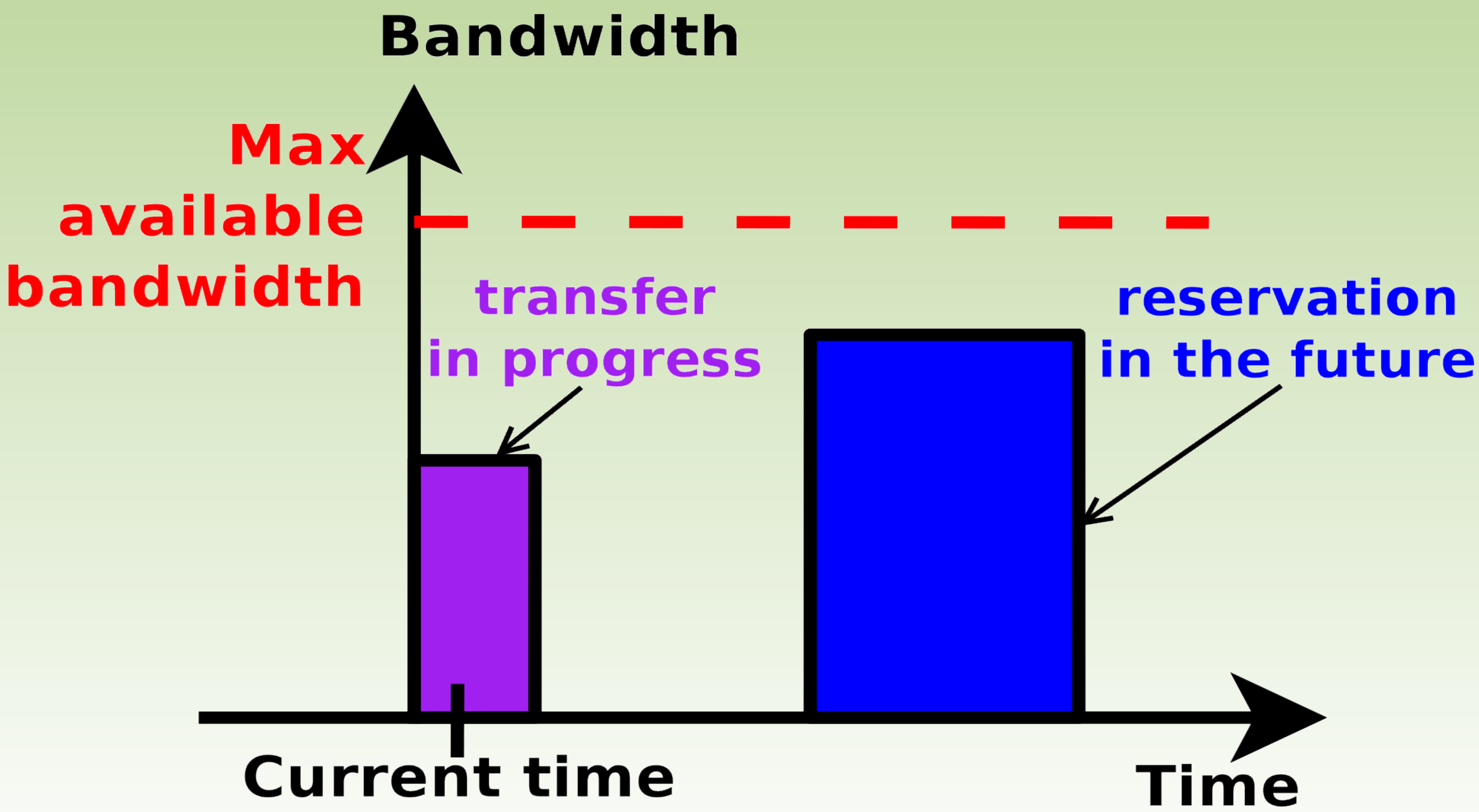
- Backup networks
- Enterprise networks
- E-Science networks
- Interbank networks
- Data center networks



Fedwire Interbank Payment Network

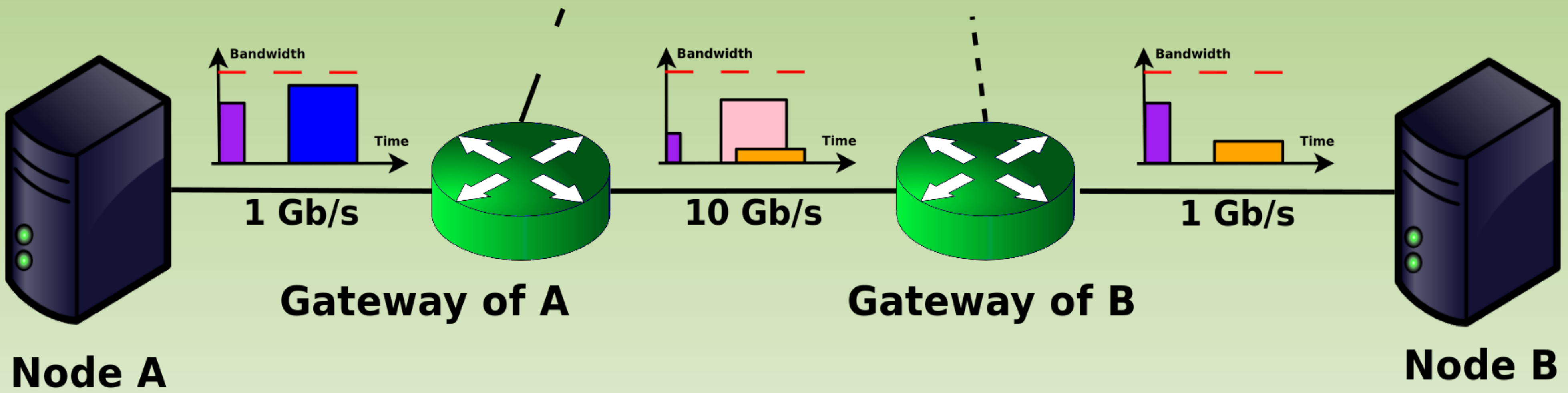
Advance Bandwidth Reservations

- One agenda per port and one per router
- End-to-end reservation (the whole path, at the same time, with identical bandwidth for all the links)



End-to-end reservation

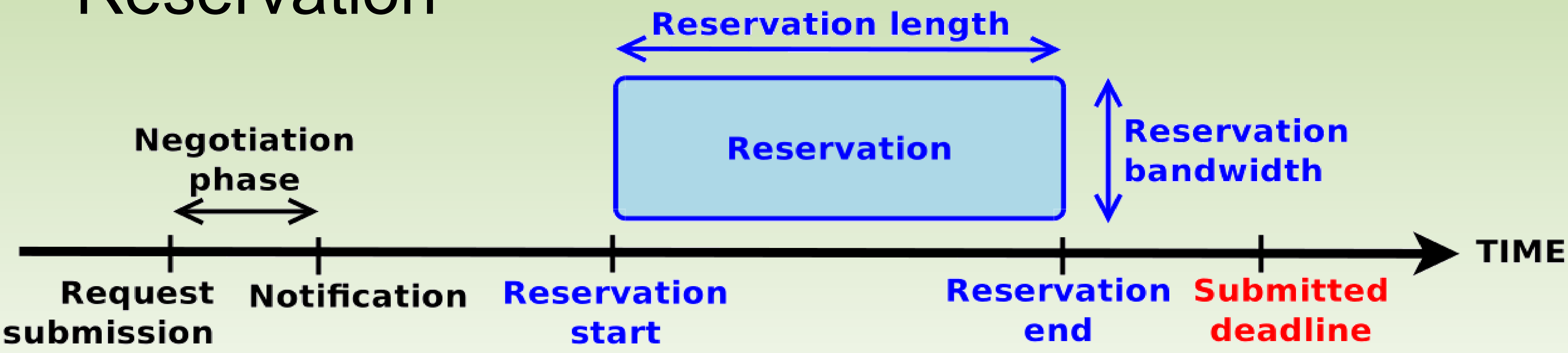
- Scheduling on all the agenda of the path



- Not store-and-forward approach

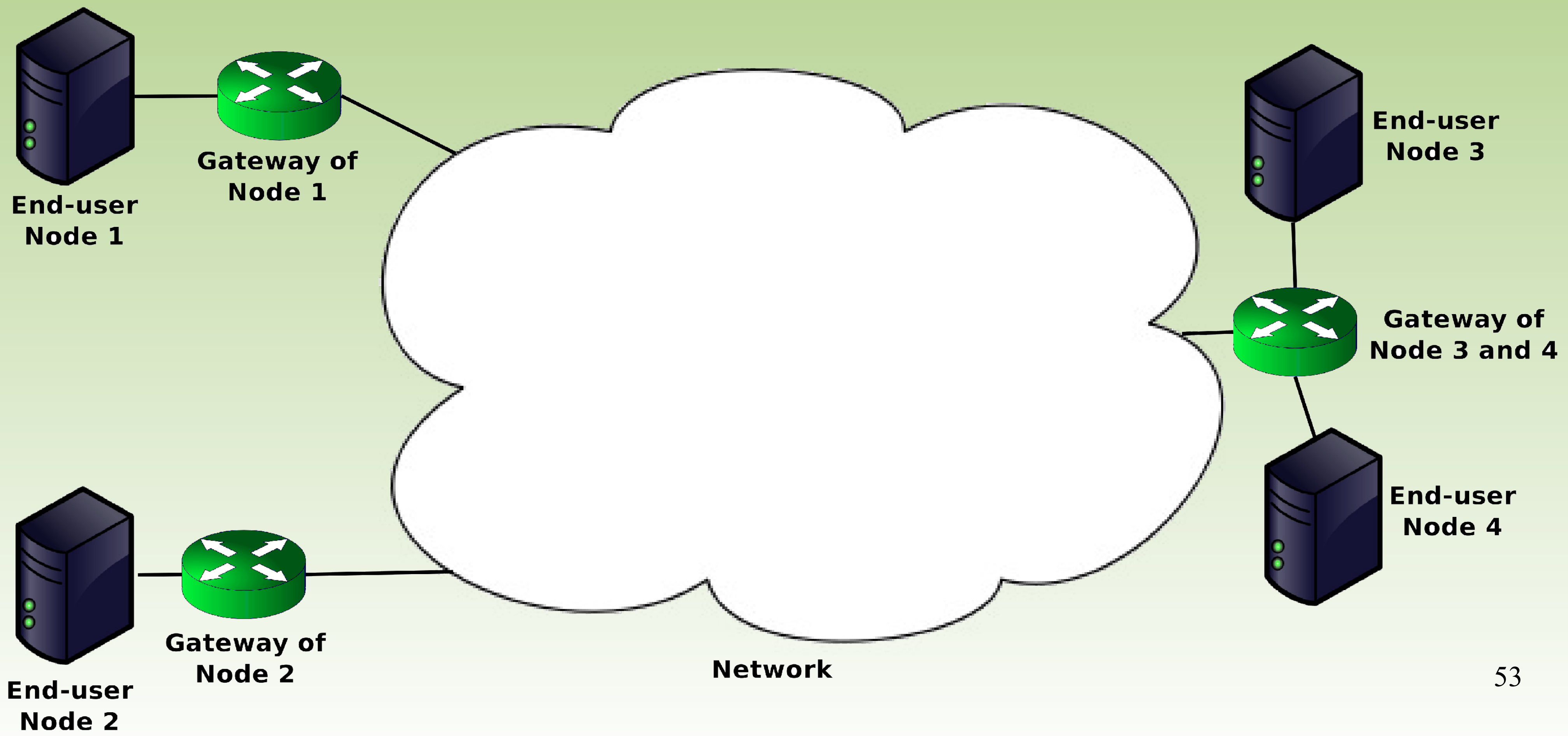
Reservation Process

- Request: at least Volume and Deadline
- Negotiation (admission control, availability check)
- Notification : OK or earliest possibility
- Reservation



Global architecture & scenario

- End users want to send BDT to other end users.
- End users are connected to gateways.



Underlying assumptions

- Routers are ALR-enabled and can be switched off and on.
- Static and symmetric routing, symmetric network
- End-to-end energy consumption can be computed using ECOFEN.

→ <u>Goal:</u> to find a good trade-off between performance (# of granted reservations) and energy consumption of the network.

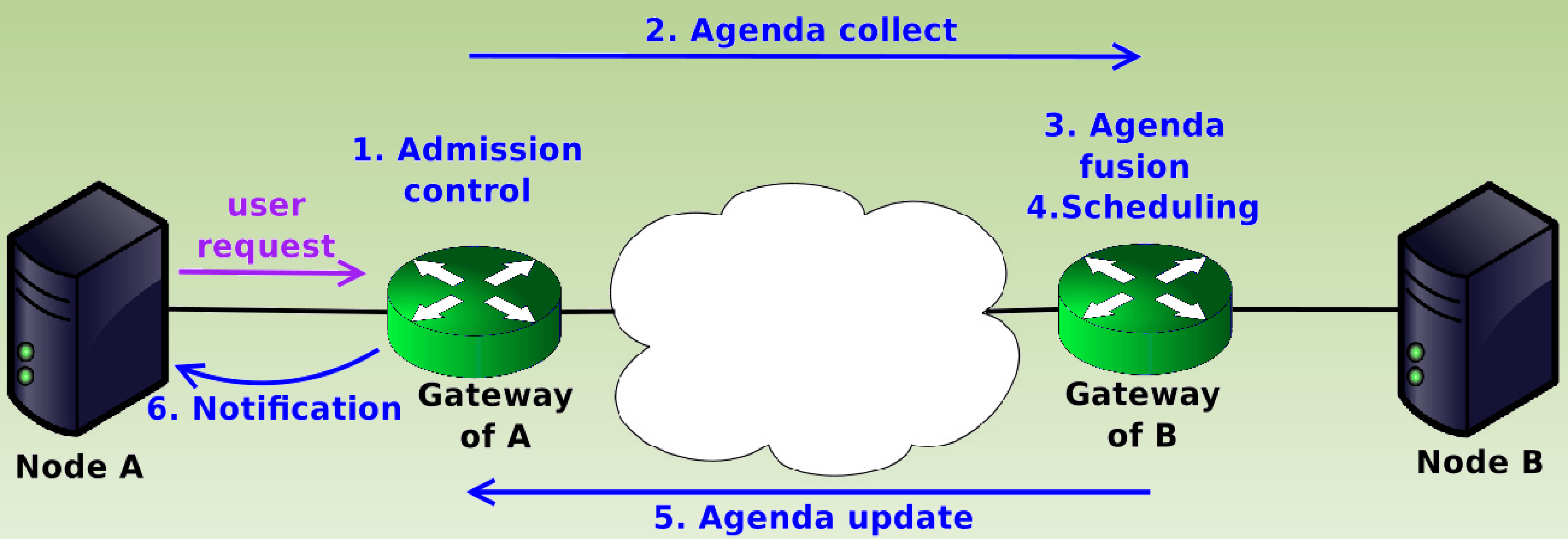
HERMES: High-level Energy-awaRe Model for bandwidth reservation in End-to-end NetworkS

Main characteristics

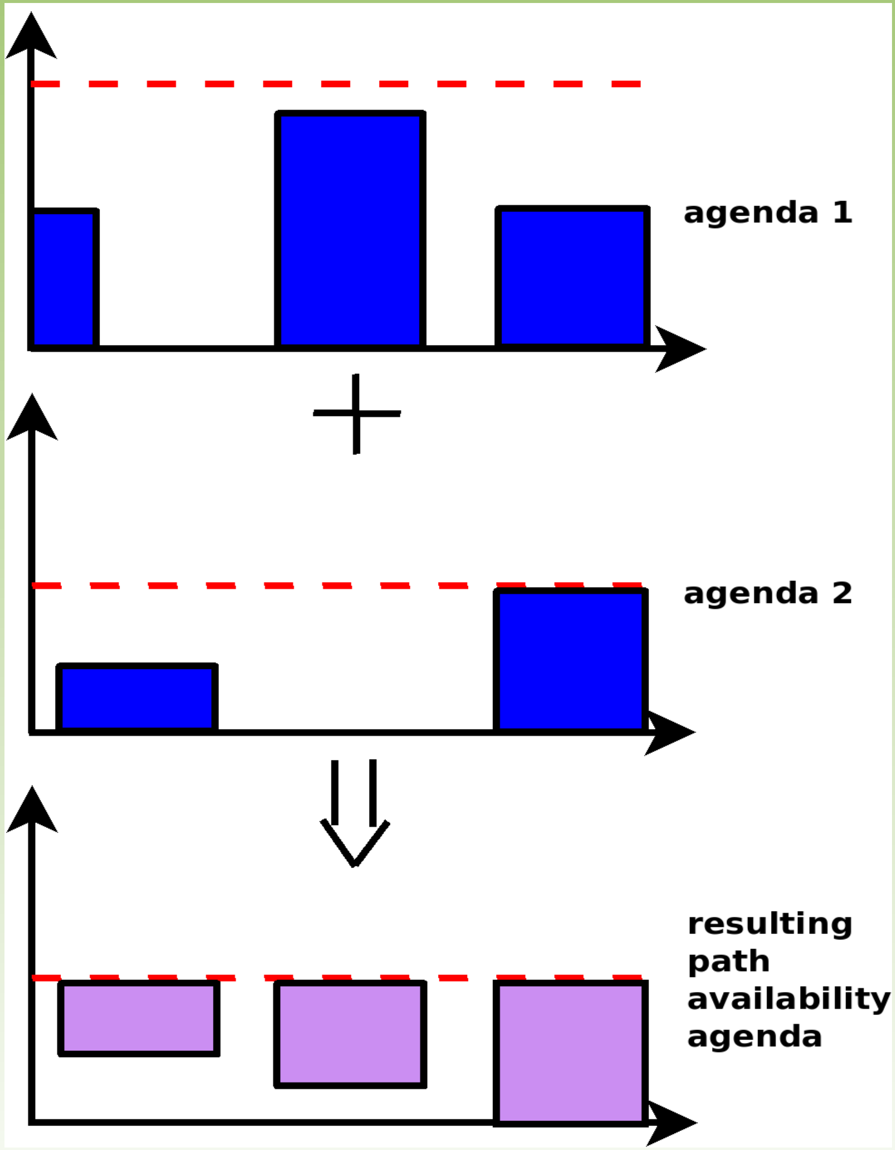
- Switching off unused nodes
- Distributed network management
- Energy-efficient scheduling with reservation aggregation
- Usage prediction to avoid on/off cycles
- Minimization of the management messages
- Usage of DTN (Disruptive-Tolerant Network) for network management purpose

Agenda collect and fusion

- One round-trip aggregated message

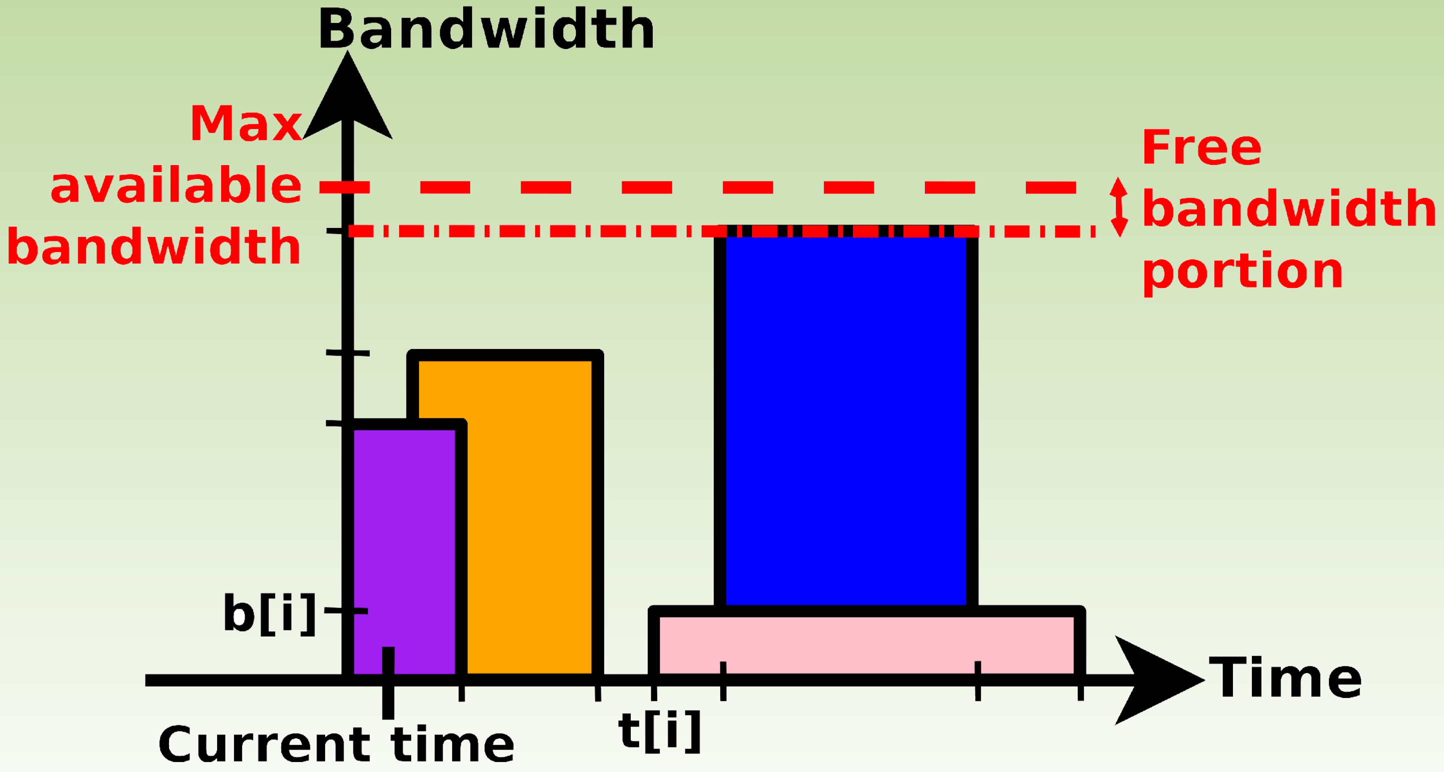


Agenda fusion



ABR scheduling

- Try to put the reservation after and before each event, and estimate the energy consumption for each one
- Chose the less energy consuming option



Prediction and switching off

- At the end of a reservation, for each resource:
  - if there is a reservation soon in the agenda
    → stay powered on
  - else
    → predict the next reservation and stay on if it soon, otherwise switch off.

- Prediction using the history.

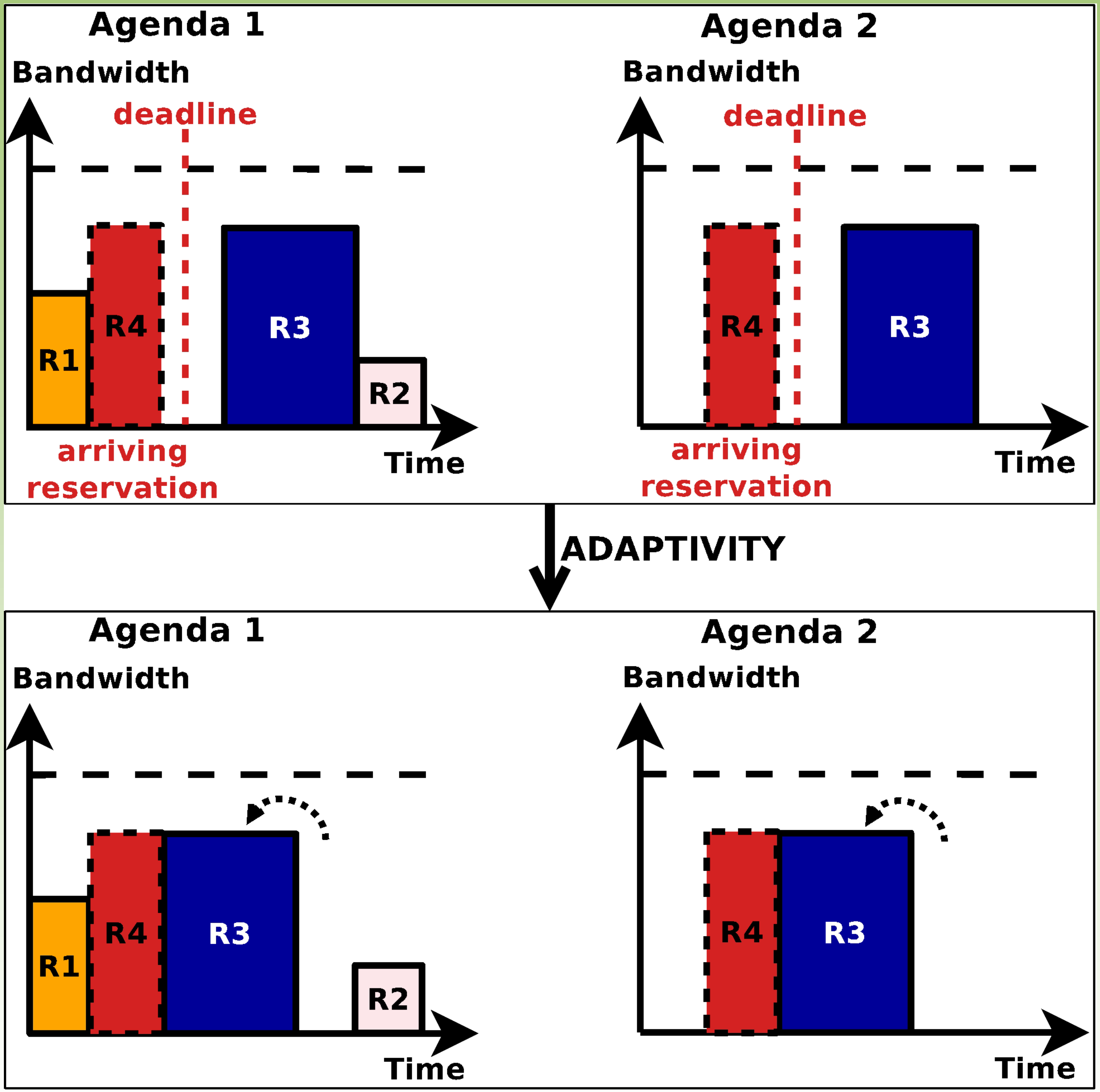
Network switched off by pieces: DTN usage

- Each reservation request has a TTL

  - if TTL = 0 → request to compute now, answer to give as soon as possible

  - otherwise, users can wait for the answer. The request moves forward into the network hop-by-hop waiting for the nodes to wake up. If the TTL is expired, the whole path is awaken.

Adaptativity



Discussion on the scheduling algorithm

Example: node A wants to send 200 Mb to node B

- At 1 Gb/s, it takes 0.2 seconds and consumes:

$$E_{transfer} = P_{EthernetCard}(NodeA, 1\,Gb/s) \times 0.2 + P_{EthernetCard}(NodeB, 1\,Gb/s) \times 0.2$$

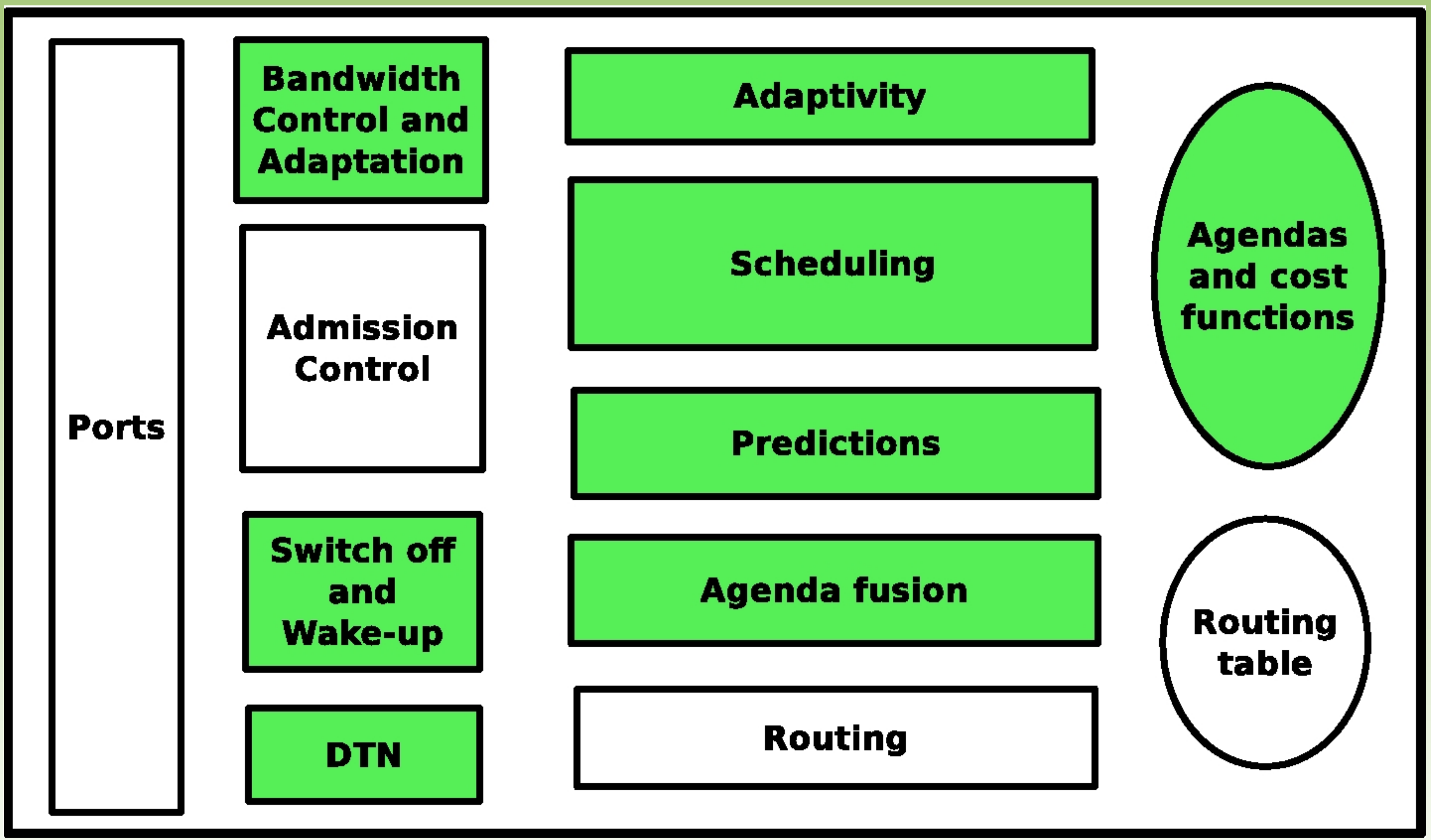
- At 100 Mb/s, it takes 2 seconds and consumes:

$$E_{transfer} = P_{EthernetCard}(NodeA, 100\,Mb/s) \times 2 + P_{EthernetCard}(NodeB, 100\,Mb/s) \times 2$$

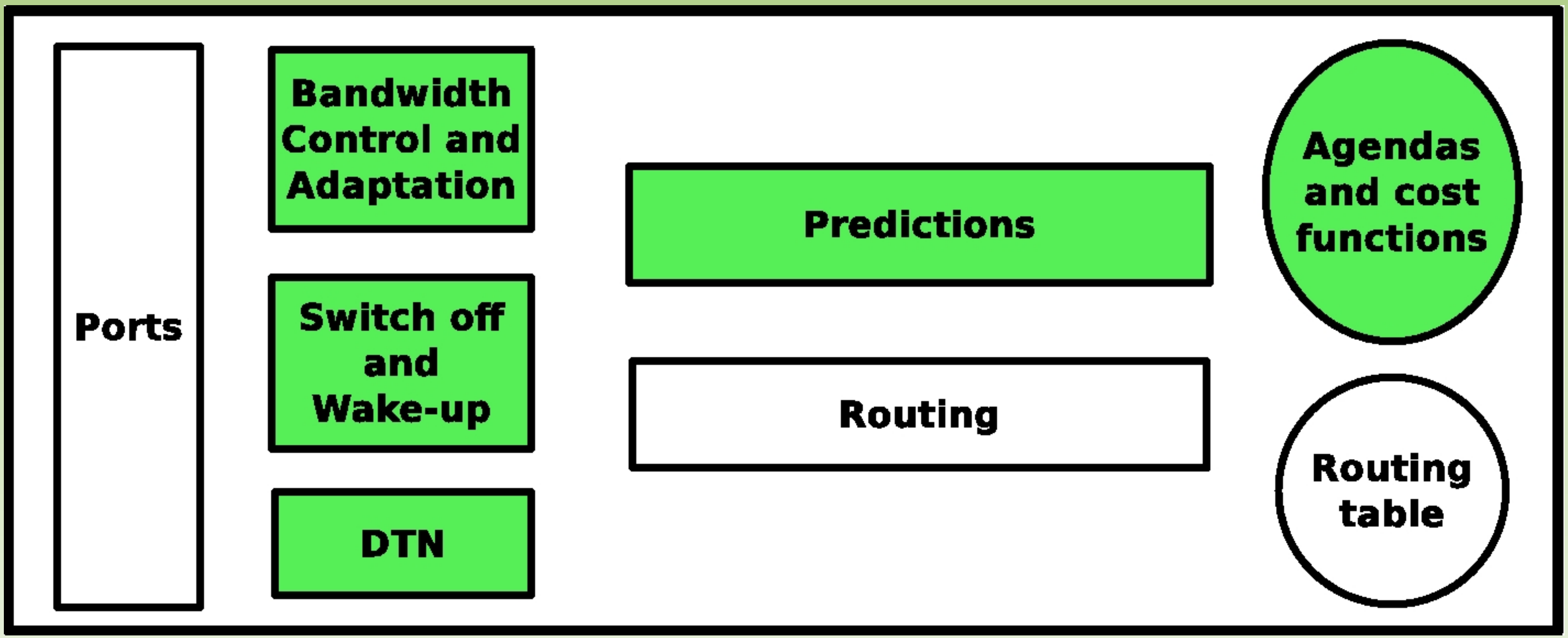
→ So, to be energy efficient, we should have:

$$P_{EthernetCard}(1\,Gb/s) < 10 \times P_{EthernetCard}(100\,Mb/s)$$

Gateway Architecture



Router Architecture



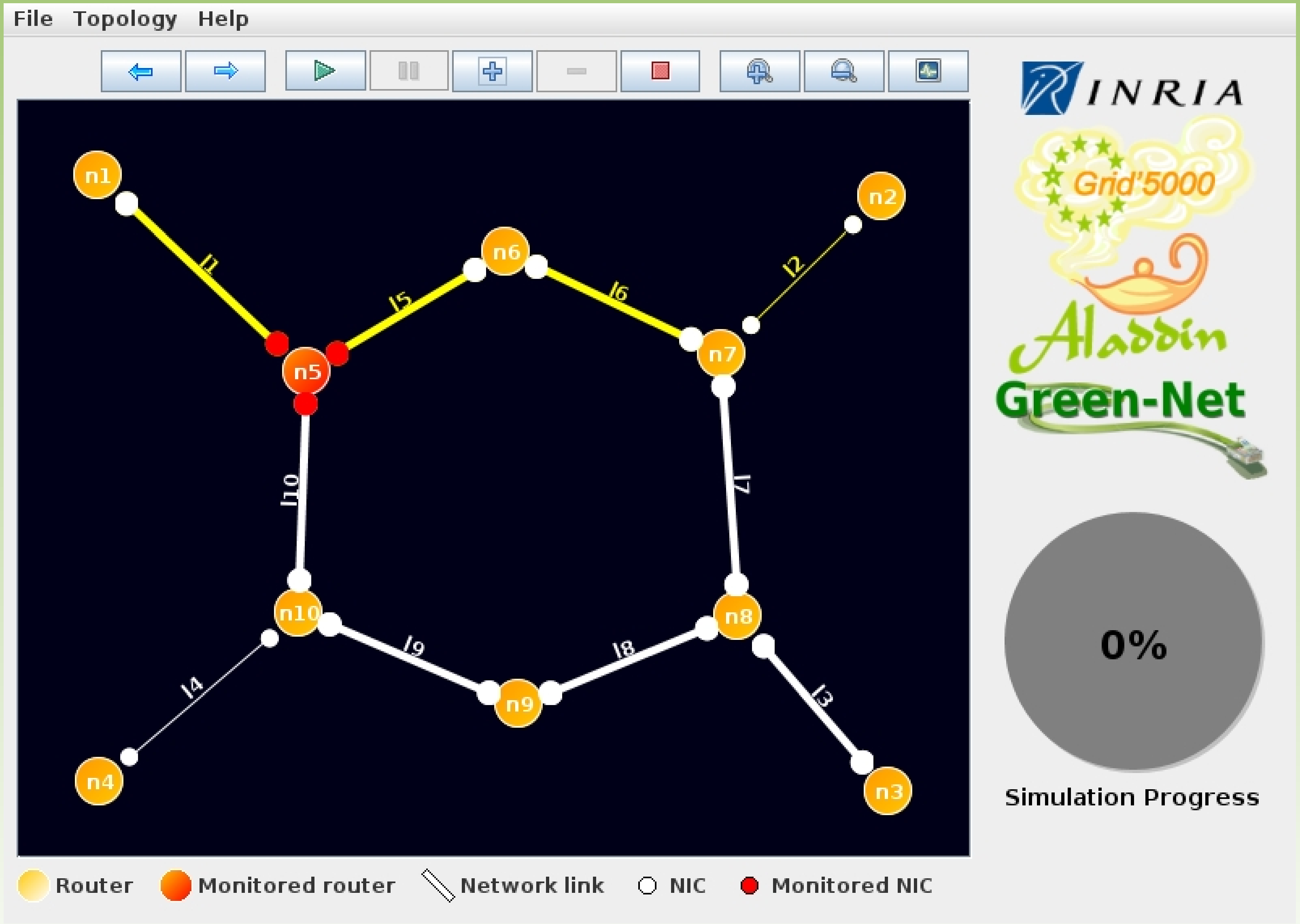
HERMES Evaluation

Simulation Results

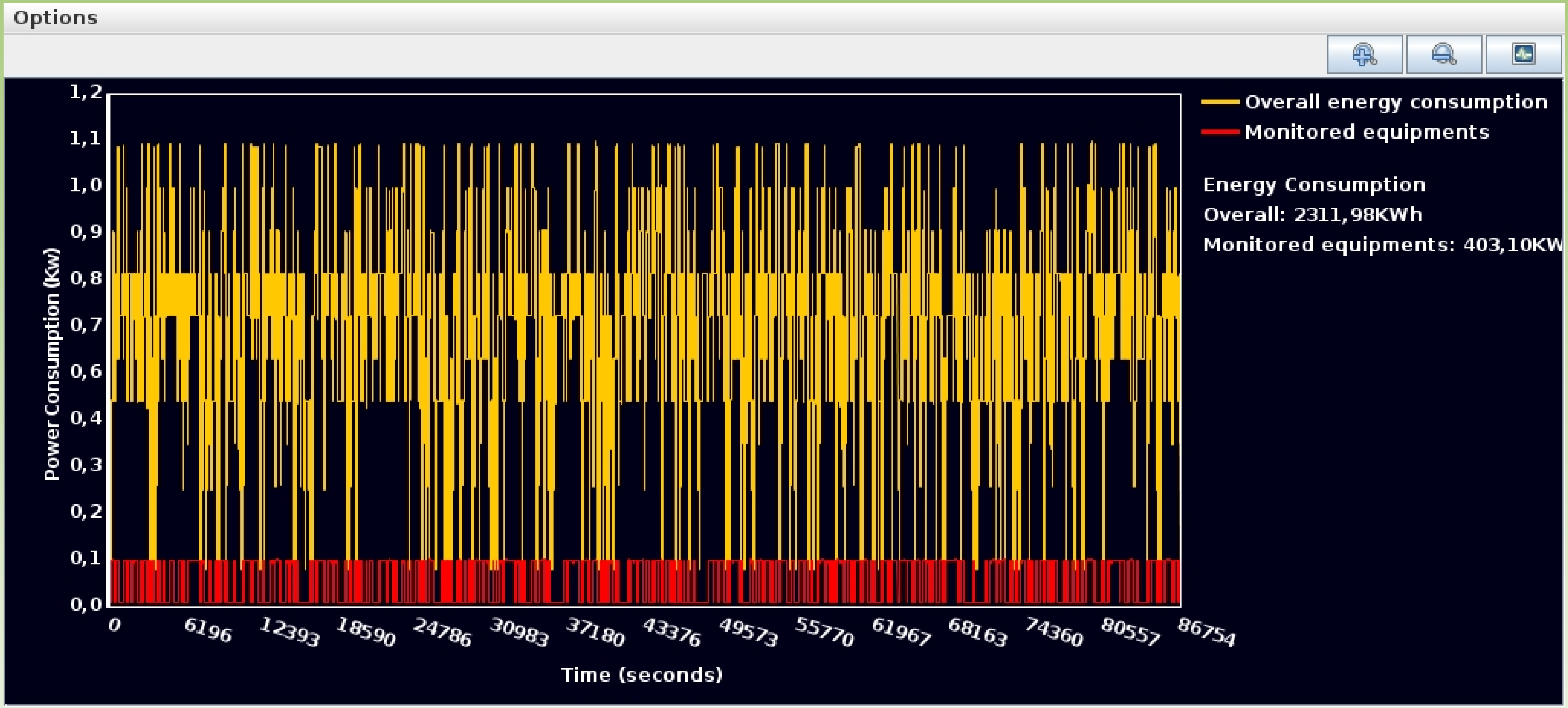
- BoNeS (Bookable Network Simulator)
- Written in Python (5,000 lines)
- Generates random network with the Molloy & Reed method or uses configuration file
- Generates traffic according to statistical laws:

  - submission times (log-normal distribution)
  - data volumes (negative exponential)
  - sources and destinations (equiprobability)
  - deadlines (Poisson distribution)

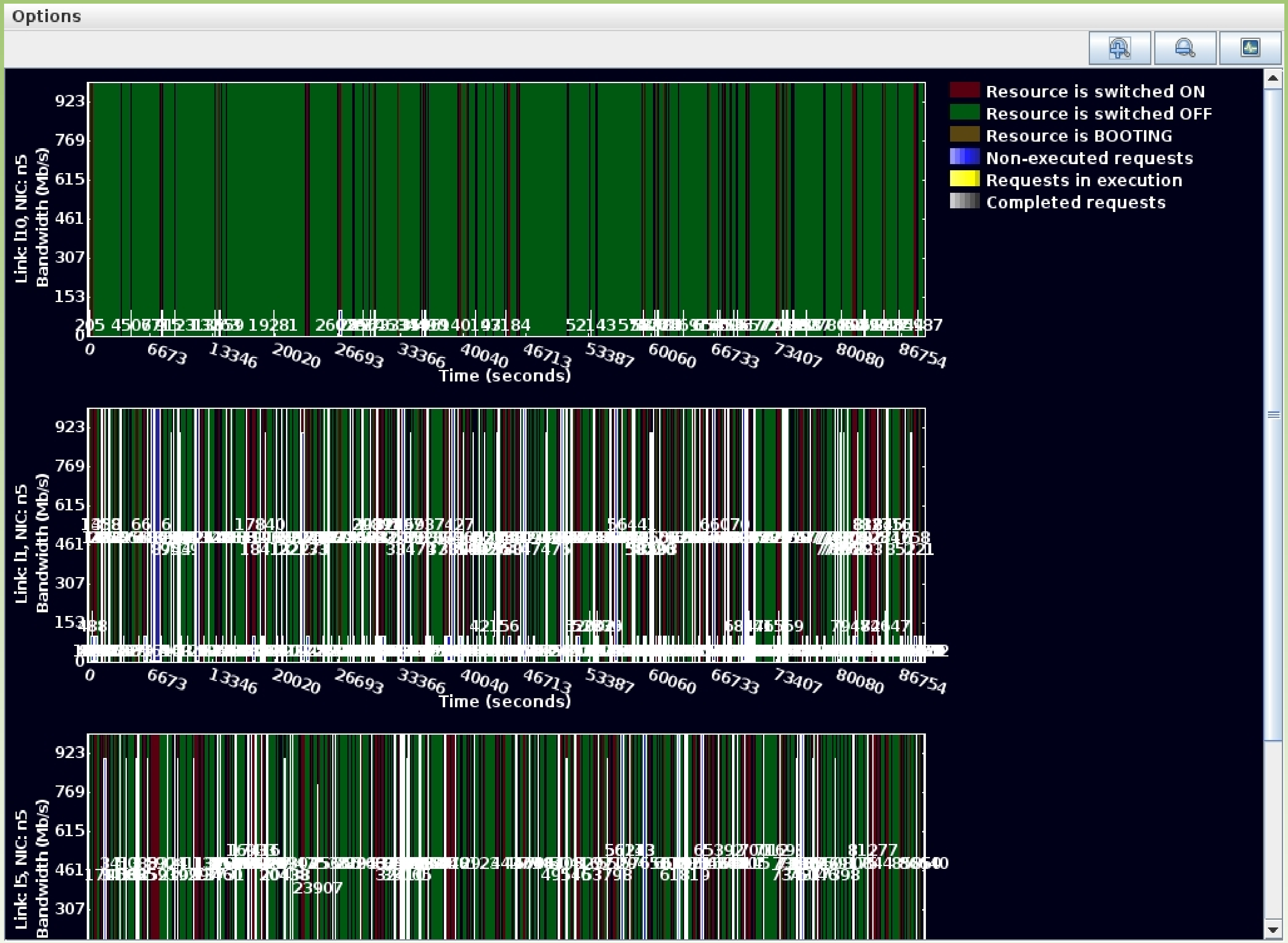
Visualization tool



Visualization tool



Visualization tool



Comparison with other schedulings

- **First**: the reservation is scheduled at the earliest possible place;
- **First green**: the reservation is aggregated with the first possible reservation already accepted;
- **Last**: the reservation is scheduled at the latest possible place;
- **Last green**: the reservation is aggregated with the latest possible reservation already accepted;
- **Green**: HERMES scheduling;
- **No-off**: first scheduling without any energy management.

→ always before deadline

Simulations

- Network simulated: 500 nodes, 2 462 links.
- Random Network (Molloy & Reed method)
- All the nodes can be sources and destinations.
- Time to boot: 30 s.; time to shutdown: 1 s.
- 1 Gbps per port routers

Component	State	Power
Chassis	ON	150 W
	OFF	10 W
Port	1 Gbps	5 W
	100 Mbps	3 W
	idle, 10 Mbps	1 W

Results with a 31% workload

- 80 experiments for each value
- Four hour period of simulated time for each experiment
- Energy consumption in Wh

Scheduling	No off	First	First green	Last	Last green	Green
Average	412 306	205 270	203 844	204 949	196 260	203 342
Standard deviation	2 685	2 477	1 938	2 375	2 695	2 145
Accepted volume (Tb)	2 148	2 148	2 128	2 014	1 853	2 149
Cost in Wh per Tb	191.92	95.55	95.78	101.74	105.92	94.60

Results with different workloads

- 31%, 46% and 61%
- Average occupancy per link
- Compared to current case (no-off), HERMES could save 51%, 46% and 43% of the energy consumed depending on the workload

Workload	No off	First	First green	Last	Last green	Green
31%	191.92	95.55	95.78	101.74	105.92	94.60
46%	149.84	81.61	81.95	87.74	92.40	80.63
61%	130.45	74.73	74.91	80.09	84.63	73.79

Contributions and Perspectives

- State of the art for green wired networks (I hope it will be useful for your research)
- Complete and energy-efficient ABR framework for BDT including scheduling, prediction and on/off algorithms called HERMES
- Validation of HERMES with simulations
- Perspective: to encourage network equipment manufacturers to design new equipments able to switch on and off and to boot rapidly.

Long-term challenge

- Networks are numerous, heterogeneous and interconnected; they use various technologies, functionalities and protocols
- IPv6 designed in 1990 (IETF), starts to be widely applied due to urgent constraints (we are running out of addresses)

  → need of standards

  → time to deploy, apply new protocols around 15 years

  → interoperability and backwards compatibility can compromise energy savings

  → protocols designed now to solve current problems will be applied in 15 years with future traffic and applications

Beyond the networks

- Networks can help increase the energy savings in other domains
- Ex: Teleconference, video-conference
- All the side effects should be taken into account.

  - build teleconference systems (infrastructure, equipments)

  - build new networks with guaranteed QoS

“Virtual meetings and Climate Innovation in the 21st Century” WWF Study

Are we going on the good way?

- New functionalities
- Create new practices and needs
- Multiplication of the devices
- Capability overlap
- Health issues

Useful link

http://www.greenit-conferences.org

Green ICT conferences/workshops/journals list

COST Training School Webpage

- It's here:

  http://www.ens-lyon.fr/LIP/RESO/COST-TrainingSchool-2011/

- Slides and deliverable

Thank you for your attention!

Questions?

Anne-Cécile Orgerie
annececile.orgerie@ens-lyon.fr