Gestión de Centro de Cómputos

Prof. Cristian Cabanellas

Prof. Gerardo Riveros

Facultad Politécnica, Universidad Nacional de

AGENDA

- ***Flashaback (Topologías de redes)***
- ***Green IT***
- ***PUE***
- **Simulador Greencloud**

AGENDA

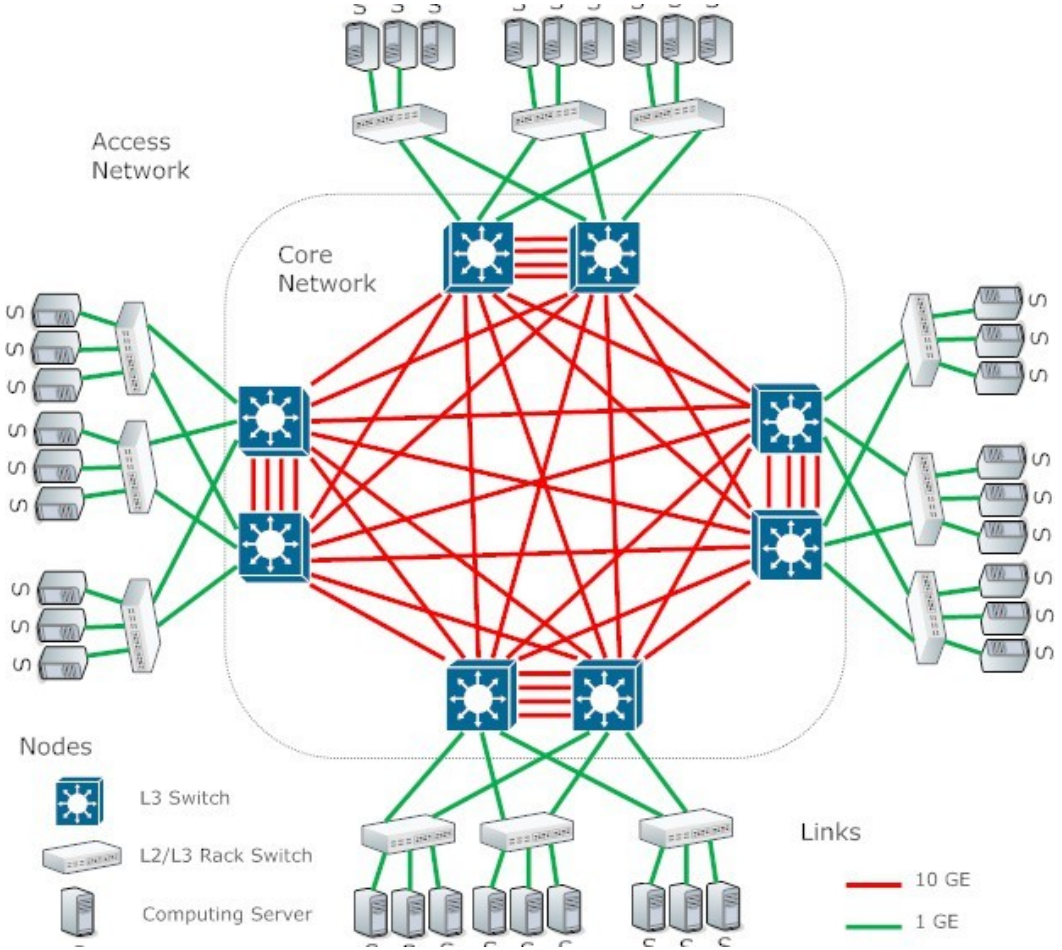
- ***Flashaback (Topologías de redes)***
- **Green IT**
- **PUE**
- **Simulador Greencloud**

TWO-TIER ARCHITECTURE

The computing servers are physically arranged into racks interconnected by layer-3 switches providing full mesh connectivity.

Characteristics:

- **Up to 5500 nodes**
- **Access & core layers**
- **1/10 Gb/s links**
- **Full mesh**
- **ICMP load balancing**



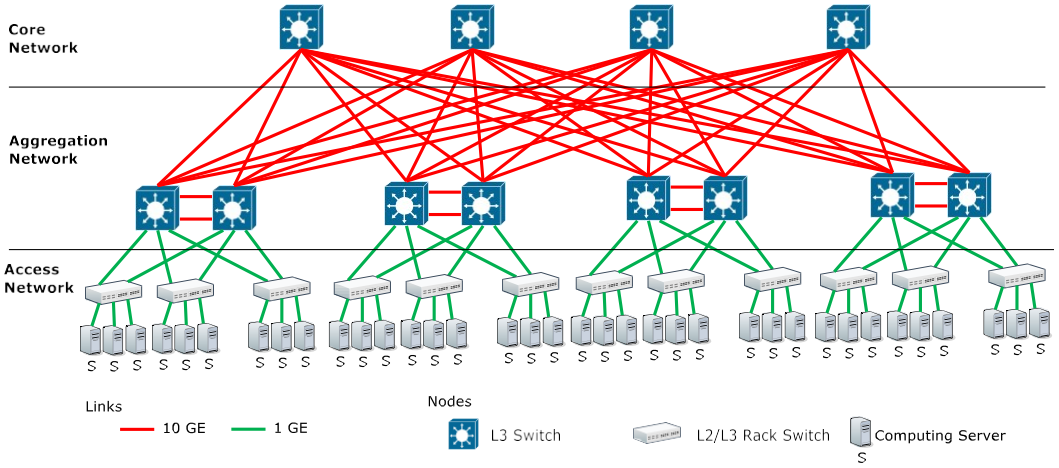
Facultad Politécnica, Universidad Nacional de

THREE-TIER ARCHITECTURE

Being the most common nowadays, three-tier architecture interconnects computing servers with access, aggregation, and core layers increasing the number of supported nodes while keeping inexpensive layer-2 switches in the access.

Characteristics:

- **Over 10,000 servers**
- **ECMP routing**
- **1/10 Gb/s links**

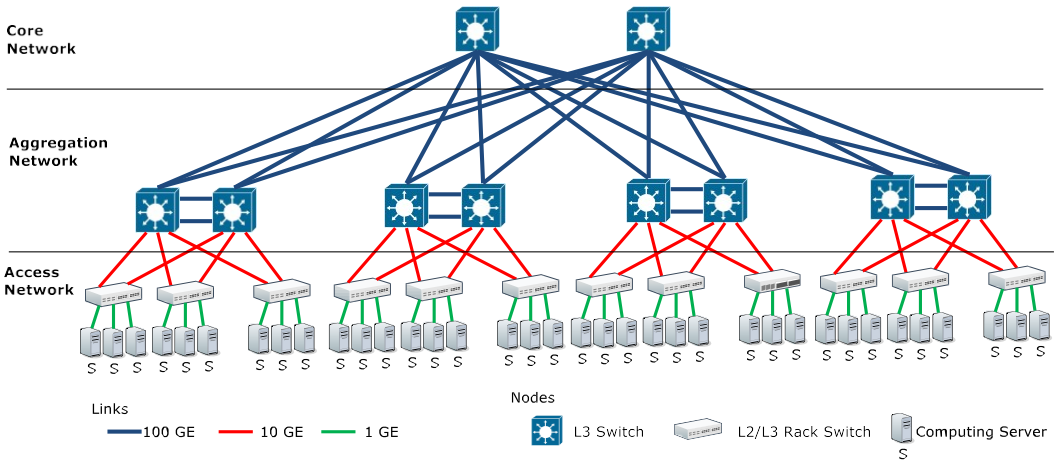


THREE-TIER HIGH-SPEED ARCHITECTURE

With the availability of 100 GE links (IEEE 802.3ba) reduces the number of the core switches, reduces cablings, and considerably increases the maximum size of the data center due to physical limitations.

Characteristics:

- **Over 100,000 hosts**
- **1/10/100 Gb/s links**



AGENDA

- *Flashaback (Topologías de redes)*
- **Green IT**
- *PUE*
- *Simulador Greencloud*

WHAT IS GREEN IT?

Green IT also referred as Green computing is a study and practice of using computing resources in an efficient manner such that its impact on the environment is as less hazardous as possible.

The primary goals of Green IT are to make sure that least amount of hazardous materials is used; the computing resources are used efficiently in terms of energy and to promote recyclability.

Source: http://en.wikipedia.org/wiki/Green_computing

WHY IS GREEN IT IMPORTANT?

Home

- A typical desktop computer can consume 200-300W of power.
- This results in emission of about 220Kg of CO2/annum

Data center

- Many processing units (Server), data storage units and network communication units.
- As of 2006 the data centers in US used 61 billion kilowatt- hours (kWh), or 1.5 percent of total U.S. electricity consumption.

SOME INTERESTING NUMBERS

- **The total estimated energy bill for data centers in 2010 is $11.5 billion**
- **Data Centers produce 170 million metric tons of CO2 worldwide currently per year.**
- **670 million metric tons of CO2 are expected to be emitted by data centers worldwide annually by 2020.**
- **32 percent of all servers are running at or below 3 percent peak and average utilizations, wasting energy spinning and cooling, and doing virtually no work.**

Source: http://lushtech.com/why_go_green/

AGENDA

- *Flashaback (Topologías de redes)*
- *Green IT*
- **PUE**
- *Simulador Greencloud*

ENERGY CONSUMPTIONS

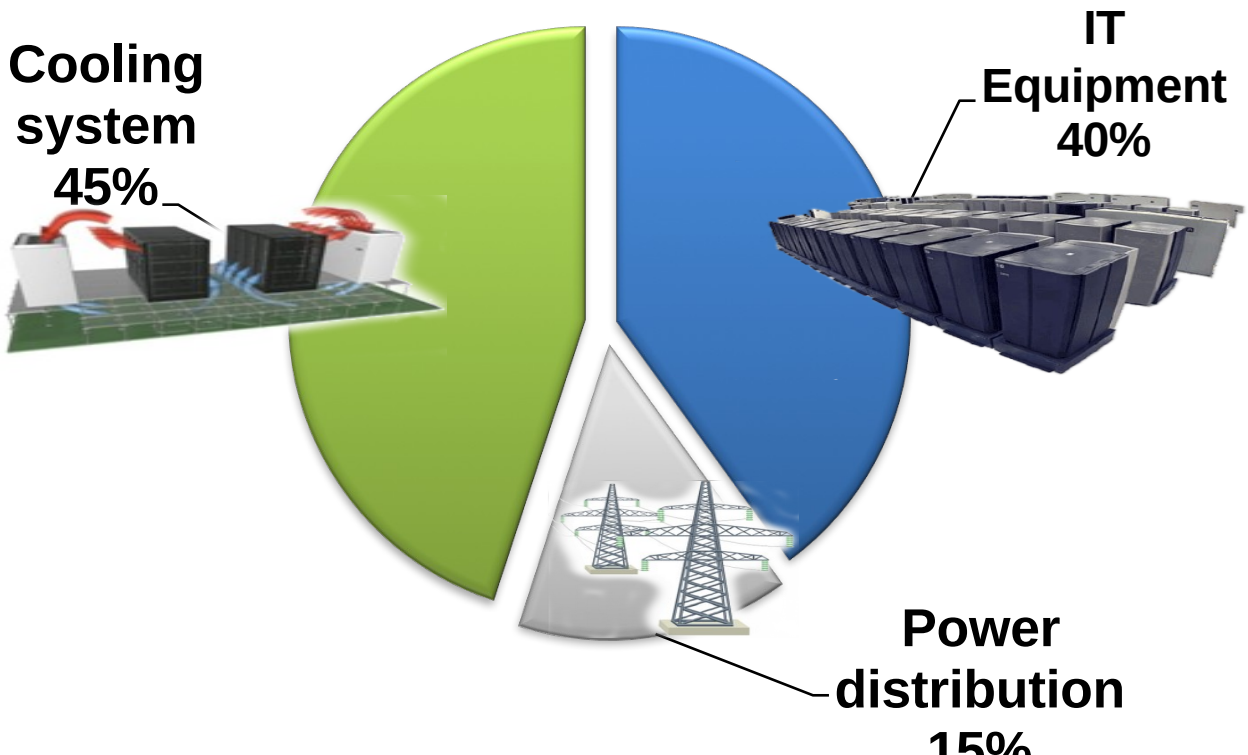
The major IT companies, such as Microsoft, Google, Amazon, and IBM, pioneered the field of cloud computing and keep increasing their offerings in data distribution and computational hosting.

Gartner group estimates energy consumptions to account for up to 10% of the current data center operational expenses (OPEX), and this estimate may rise to 50% in the next few years.

Along with the computing-based energy high power consumption generates heat and requires an accompanying cooling system that costs in a range of $2 to $5 million per year. There are a growing number of cases when a data center facility cannot be further extended due to the limited available power capacity offered to the facility.

Facultad Politécnica, Universidad Nacional de

DISTRIBUTION OF ENERGY CONSUMPTION



Facultad Politécnica, Universidad Nacional de

THE DISTRIBUTION OF DATA CENTER POWER CONSUMPTION

Parameter	Power consumption (kW h)		
	Two-tier (2T)	Three-Tier (3T)	Three-tier high-speed (3Ths)
Data center	477.8	503.4	508.6
Servers	351	351	351
Switches	126.8	152.4	157.6
Core (C_1)	51.2	25.6	56.8
Aggregation (C_2)	–	51.2	25.2
Access (C_3)	75.6	75.6	75.6

On average, the data center consumption is around 432 kW h during an hour of the runtime. On the yearly basis, it corresponds to 4409 MW h or $441 thousand with an average price of 10 c per kW h.

Facultad Politécnica, Universidad Nacional de

POWER USAGE EFFECTIVENESS (PUE)

The Green Grid organization for green computing provides the following standard metric for calculating energy efficiency of the data centers:

Power Usage Effectiveness (PUE)

PUE = Total Facility Power/IT Equipment Power

Facultad Politécnica, Universidad Nacional de

WHAT DOES PUE MEAN?

- **PUE shows the relation between the energy used by IT equipment and energy used by other facilities such as cooling needed for operating the IT equipment.**
- **For example, a PUE of 2.0 indicates that for every watt of IT power, an additional watt is consumed to cool and distribute power to the IT equipment.**
- **At present the PUE of a typical enterprise data center is around between 1 to 3.**

Facultad Politécnica, Universidad Nacional de

SO WHAT CAN WE DO?

There are 2 ways in which the energy used by the data centers can be reduced

- One way is to reduce the energy used by other factors such as cooling to maintain the IT equipment.
- Other way is to reduce the energy used by the IT equipment .

AGENDA

- *Flashaback (Topologías de redes)*
- *Green IT*
- *PUE*
- ***Simulador Greencloud***

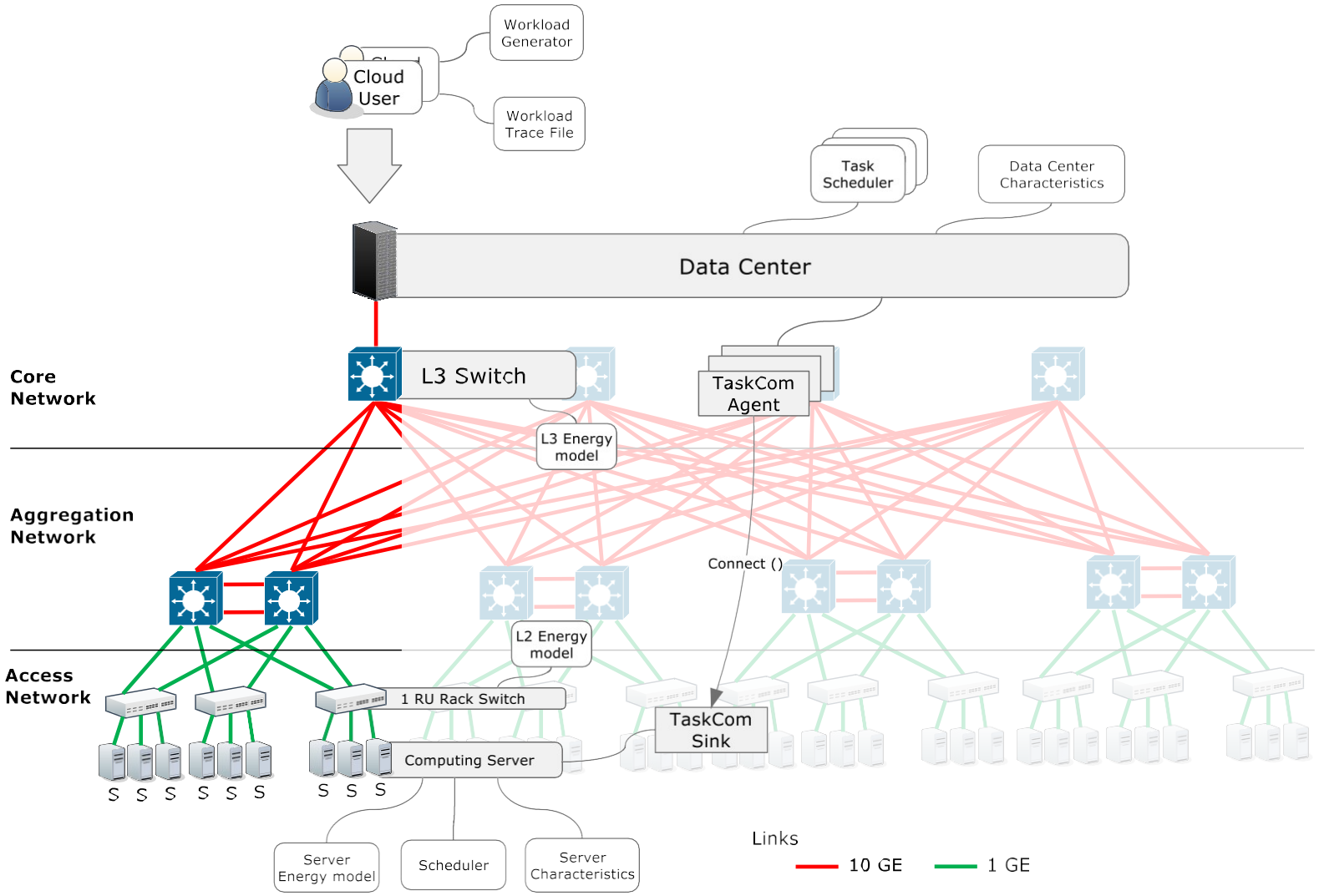
GREENCLOUD SIMULATOR

GreenCloud is a simulation environment for energy-aware studies of cloud computing data centers, developed as an extension of a packet-level network simulator Ns2. It offers a detailed fine-grained modeling of the energy consumed by the elements of the data center, such as servers, switches, and links.

From the energy efficiency perspective, a cloud computing data center can be defined as a pool of computing and communication resources organized in the way to transform the received power into computing or data transfer work to satisfy user demands.

Facultad Politécnica, Universidad Nacional de

SIMULATOR ARCHITECTURE



Facultad Politécnica, Universidad Nacional de

SIMULATOR COMPONENTS

- **Servers**
- **Switches and Links**
- **Workloads**
  - *Computationally Intensive Workloads (CIWs)*
  - *Data-Intensive Workloads (DIWs)*
  - *Balanced Workloads (BWs)*

POWER SAVING TECHNIQUES

- **Dynamic Network Shutdown (DNS)**
- **Dynamic Voltage and Frequency Scaling (DVFS)**

SIMULATION SETUP

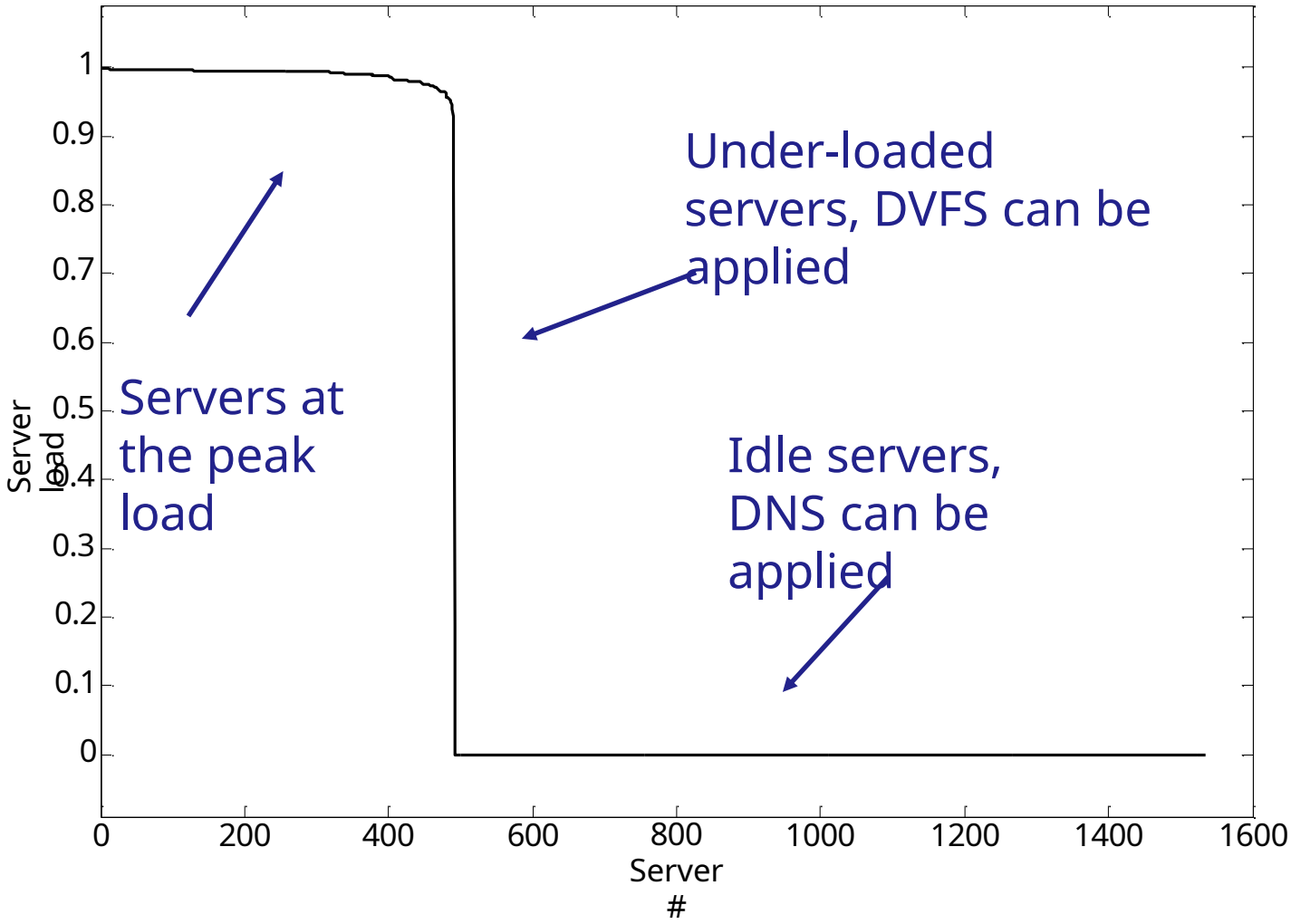
The data center composed of 1536 computing nodes employed energy-aware “green” scheduling policy for the incoming workloads arrived in exponentially distributed time intervals. The “green” policy aims at grouping the workloads on a minimum possible set of computing servers allowing idle servers to be put into sleep.

SETUP PARAMETERS

Parameter	Data center architectures		
	Two-tier	**Three-Tier**	**Three-tier high-speed**
Topologies			
Core nodes (C_1)	16	8	2
Aggregation nodes (C_2)	-	16	4
Access switches (C_3)	512	512	512
Servers (S)	1536	1536	1536
Link (C_1-C_2)	10 GE	10 GE	100 GE
Link (C_2-C_3)	1 GE	1 GE	10 GE
Link (C_3-S)	1 GE	1 GE	1 GE
Link propagation delay	10 ns		
Data center			
Data center average load	30%		
Task generation time	Exponentially distributed		
Task size	Exponentially distributed		
Average task size	4500 bytes (3 Ethernet packets)		
Simulation time	60.minutes		

Facultad Politécnica, Universidad Nacional de

WORKLOAD DISTRIBUTION



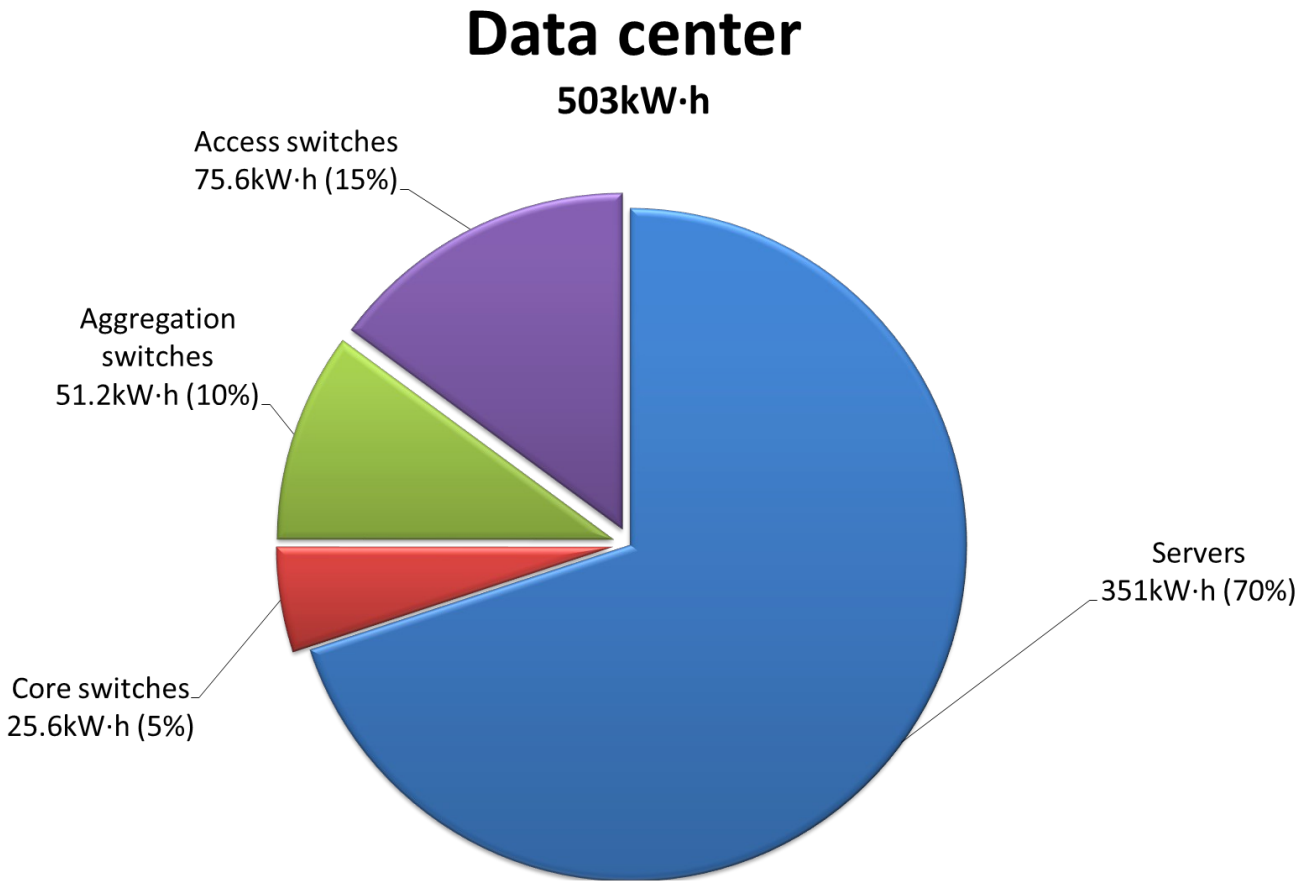
SIMULATION RESULTS

The dynamic shutdown shows itself equally effective for both servers and switches, while DVFS scheme addresses only 43% of the servers' and 3% of switches' consumptions.

Parameter	Power consumption (kW·h)			
	No energy-saving	DVFS	DNS	DVFS+DNS
Data center Servers Switches	503.4 351 152.4	486.1 (96%) 340.5 (97%) 145.6 (95%)	186.7 (37%) 138.4 (39%) 48.3 (32%)	179.4 (35%) 132.4 (37%) 47 (31%)
Energy cost/year	$441k	$435k	$163.5k	$157k

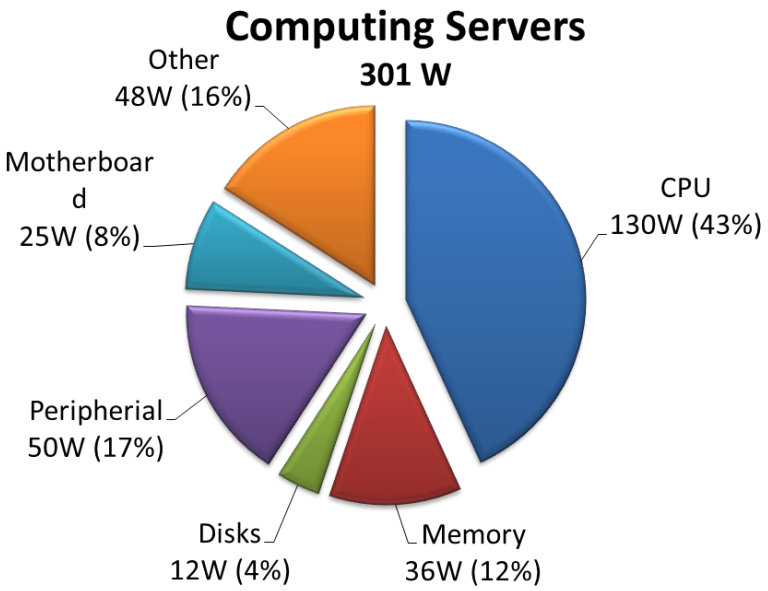
Facultad Politécnica, Universidad Nacional de

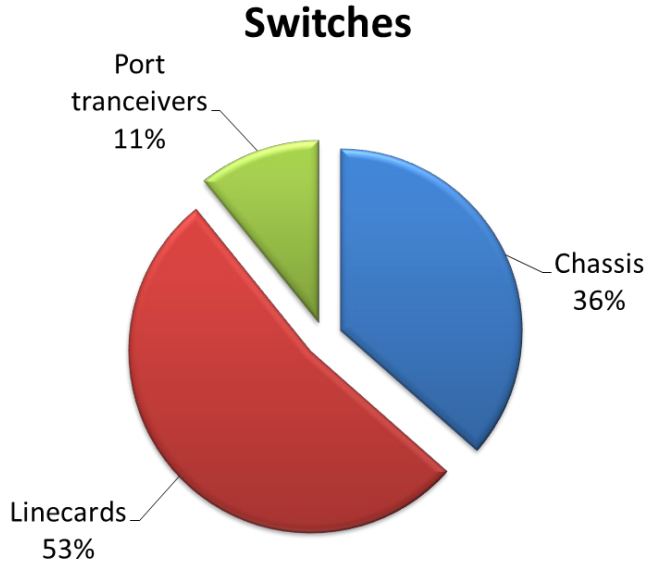
ENERGY CONSUMPTION IN DATA CENTER



Facultad Politécnica, Universidad Nacional de

ENERGY CONSUMPTION IN DATA CENTER





GREENCLOUD LAB

The simulation is setup using files located in tcl/scripts/ directory:

1. ***main.tcl:*** **determines the data center topology and simulation time.**
2. ***setup_params.tcl:*** **contains general configuration of servers, switches, tasks, monitoring and migration.**
3. ***topology.tcl:*** **creates the data center network topology.**
4. ***dc.tcl:*** **creates the data center servers and VMs.**
5. ***user.tcl:*** **defines behaviour of cloud users.**
6. ***record.tcl:*** **sets up runtime results reporting procedures.**
7. ***finish.tcl:*** **calculates and reports simulation statistics.**

Facultad Politécnica, Universidad Nacional de

GREENCLOUD LAB

1. **The simulation is run by executing the ./run/ script.**
2. **Output results can be accessed as raw values in trace files (./traces/) or the form of a dashboard generated from the raw values.**
3. **Multiple simulations can be run from a single run script. In such case, the results of different runs are accessible by tabs in the top of the dashboard.**

GREENCLOUD LAB

1. **Utilice la guía de laboratorio publicado en el Educa.**

FUENTES Y LECTURAS ADICIONALES

- **GreenCloud: a packet-level simulator of energy-aware cloud computing data centers. (Dzmitry Kliazovich · Pascal Bouvry · Samee Ullah Khan)**
- **GreenCloud: A Packet-level Simulator of Energy-aware Cloud Computing Data Centers. (University of Luxembourg)**
- **http://greencloud.gforge.uni.lu/**

THE END