UNIVERSITEIT VAN AMSTERDAM

System and Network Engineering

# Bits-nets-energy. Transporting bits or transporting energy: does it matter?

Dr. Paola Grosso
Email: p.grosso@uva.nl
URL: http://staff.science.uva.nl/~grosso

COMMIT/

Cluster Green Software

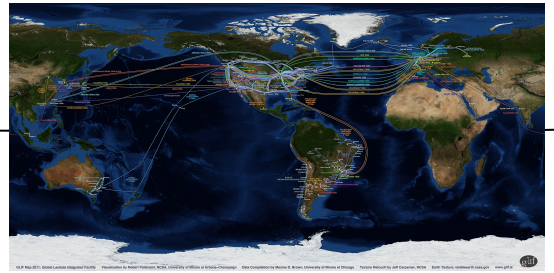
GigaPort

# Research goal

*Cloud computing gives users much freedom on where they host their computation and storage.*

*However the CO2 emission of a job depends on the location and the energy efficiency of the data centers where it is run.*

Local data center

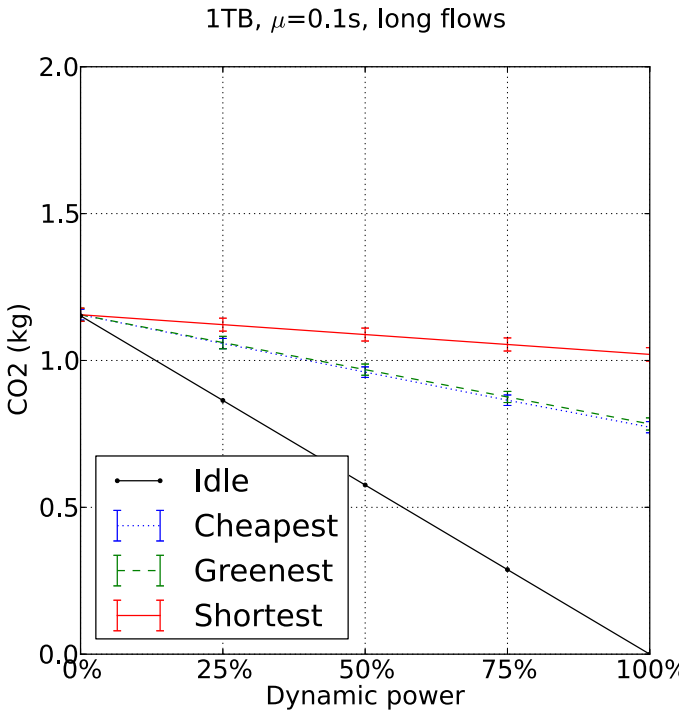


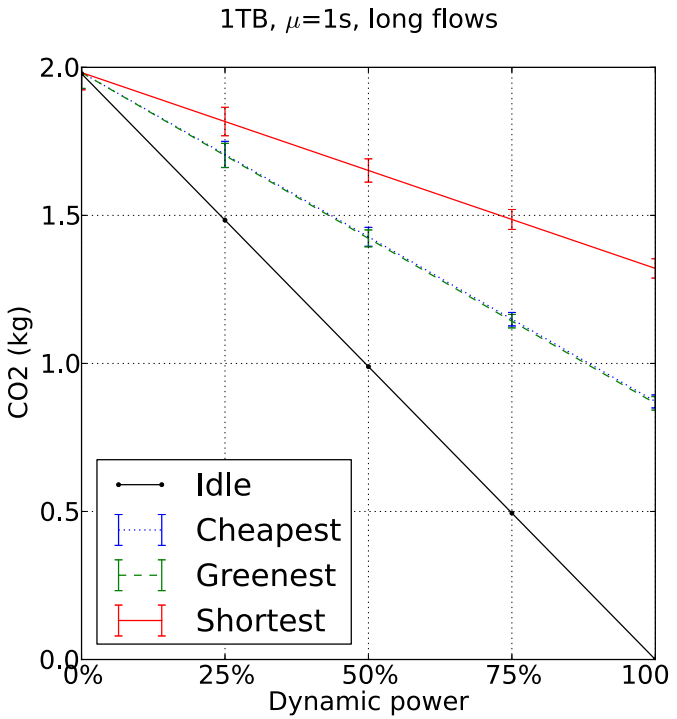
Remote data center

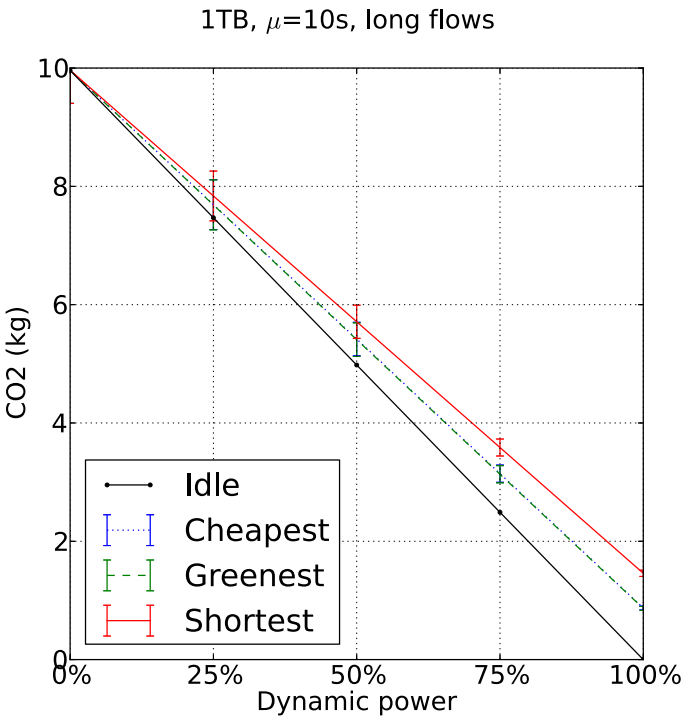
- Does it matter where you run?
- Does the network transporting the data have a significant contribution?
- Can software plan data movements more intelligently?

# Networks and CO2

- Take a network (Esnet, working on using SURFnet data)
- Define the traffic model running on it
- Use the energy monitoring information and energy costs data
- Compare path selection strategies : shortest, cheapest and greenest




*"A motivation for carbon aware path provisioning for NRENs" (submitted to eEnergy2014)*

Universiteit van Amsterdam

System and Network Engineering

# Transporting Bits or Transporting Energy: Does it matter?

A comparison of the sustainability of local and remote computing

# Three scenarios

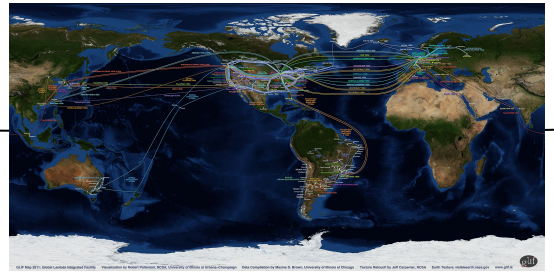
- **Processing (CPU intensive)**
  - *Input data and core hours*
- **Software interactive**
  - *Input data($D_{in}$) and core hours plus output data($D_{out}$)*
- **Storage**
  - *Data amount in ($D_{in}$) and data amount read($D_{out}$)*
  - *Type of storage (cold or hot)*
  - *Retention time(RT) and download rate*
  - *Local and remote RAID configuration*

# Bits-to-nets cost

$$Q = 1KWh \sim X \text{ gr } CO_2$$

Three components:

- Cost of local network at source data center
- Cost of local network at destination data center
- Cost of transport network



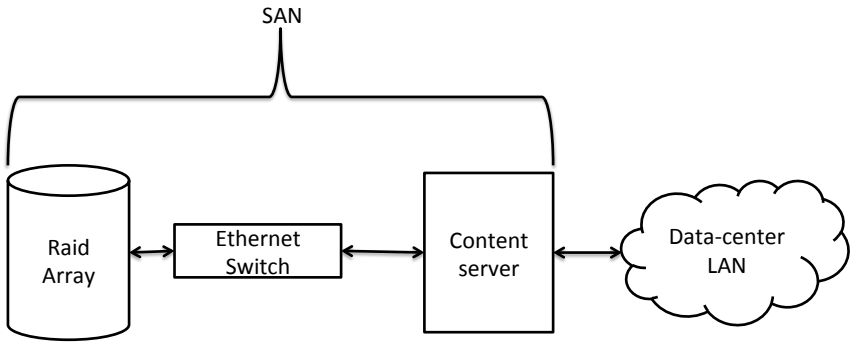
Local data center

Remote data center

$$Cost_{bits-to-nets} = Cost_{LAN-source-data-center} + Cost_{transport-network} + Cost_{LAN-destination-data-center}$$

$$Cost_{bits-to-nets} = Cost_{LAN-source-data-center} + Cost_{transport-network} + Cost_{LAN-destination-data-center}$$

# Data center cost - storage scenario




We account for the different power requirements of idle and active disks.

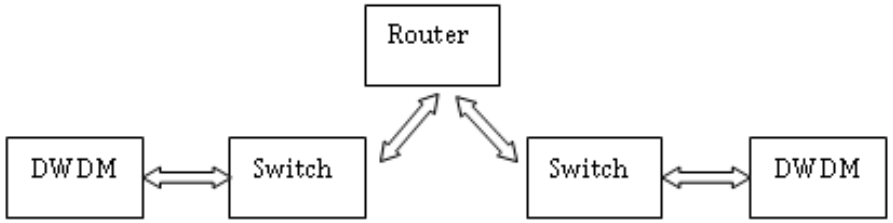
We also include the information on the redundancy level of the different RAID systems.

$$E_{total\_storage}(D_{in}, RT, D_{out}) = E_{write}(D_{in}) + E_{store}(D_{in}, RT) + E_{read}(D_{out})$$

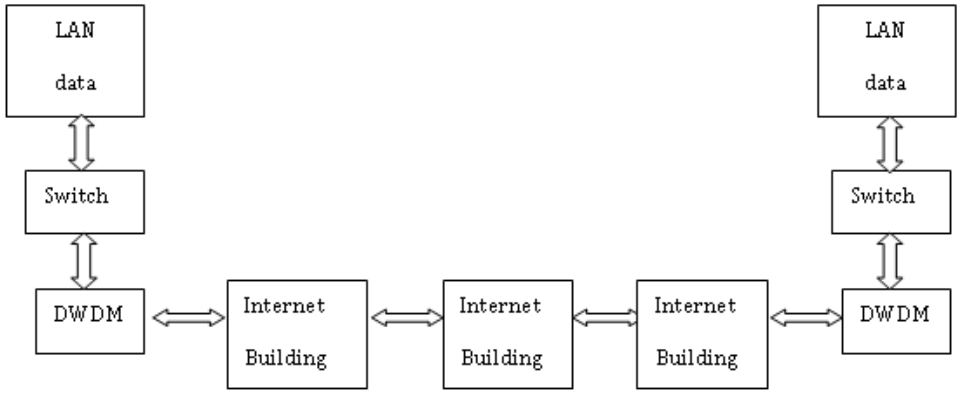
$$Cost_{bits-to-nets} = Cost_{LAN-source-data-center} + Cost_{transport-network} + Cost_{LAN-destination-data-center}$$

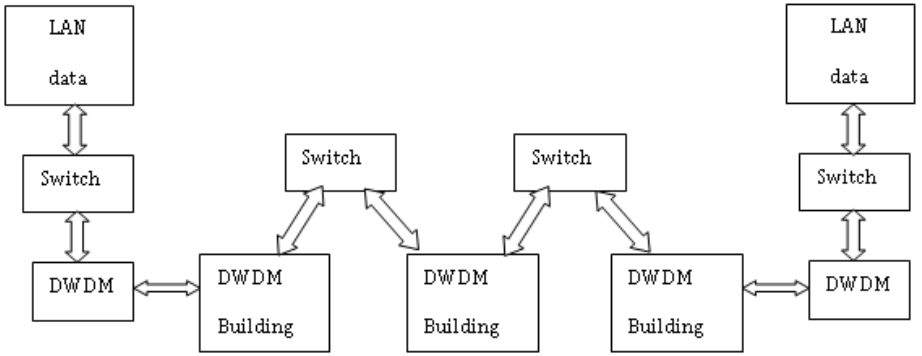
# Network cost: a building block approach

- Internet path.
- Lightpaths



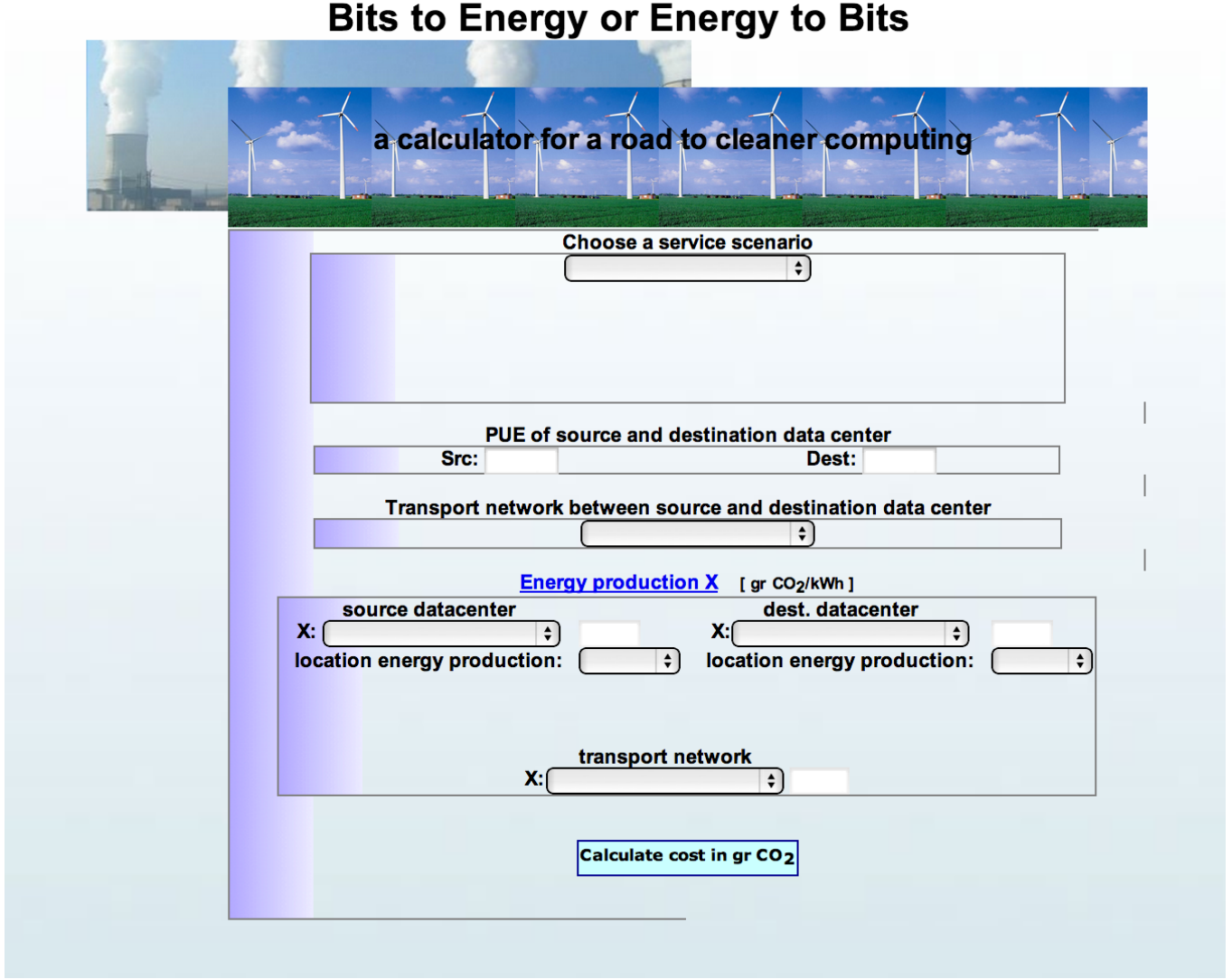




# Bits-Nets-Energy

http://sne.science.uva.nl/bits2energy/



# The output

**Local data center**

PUE=1.2
Natural gas
X: 340 gr. $CO_2$ per kWh

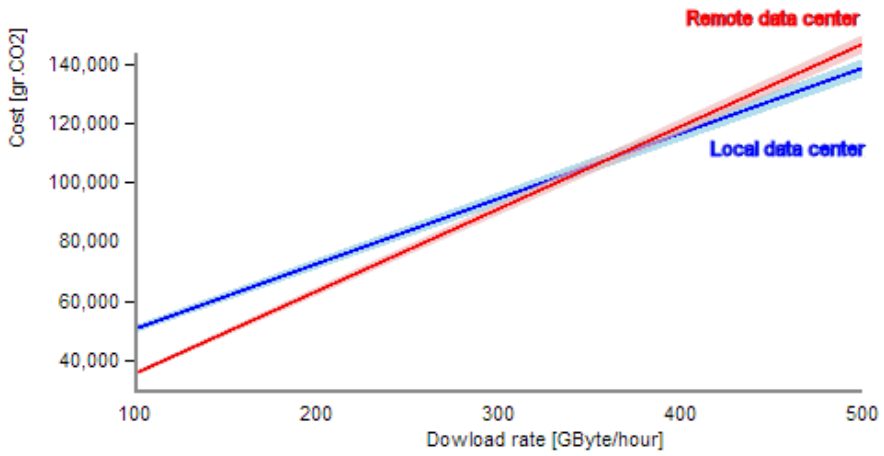
**Network**

Internet long distance

X: 520 gr. $CO_2$ per kWh

**Remote data center**

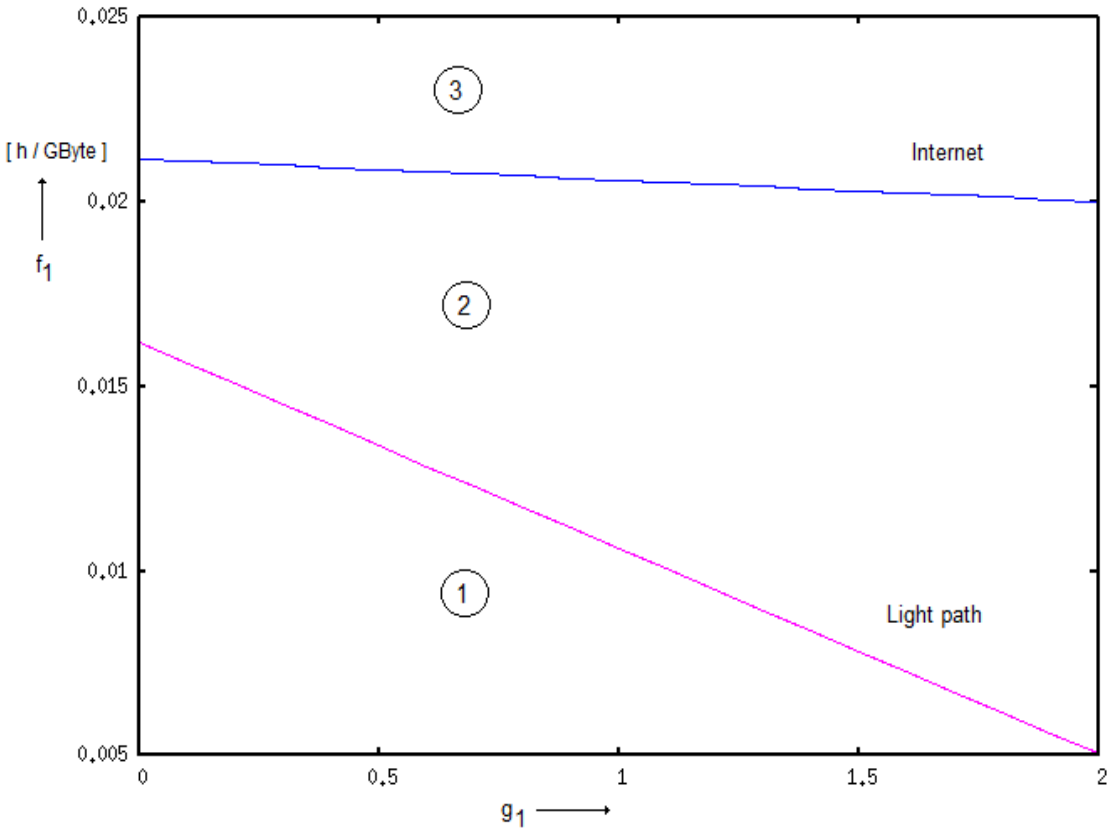
PUE=1.5
Hydroelectricity
X: 15 gr. $CO_2$ per kWh




| Cost local data center | |
|---|---|
| | g $CO_2$ (kWh) |
| LAN (input data) | 1425.28 (4.1920) |
| Storage | 72145.28 (212.1920) |
| LAN (output data) | 59861.76 (176.0640) |
| Energy prod. loss | 5390.49 (15.8544) |
| Total | 138822.81 (408.3024) |

| Cost to remote data center | |
|---|---|
| | g $CO_2$ (kWh) |
| LAN local (input data) | 2850.56 (8.3840) |
| Network (input data) | 3172.69 (6.1013) |
| LAN remote (input data) | 78.60 (5.2400) |
| Storage dest. | 3978.60 (265.2400) |
| LAN remote (output data) | 3301.20 (220.0800) |
| Network (output data) | 133253.12 (256.2560) |
| Energy prod. loss | 297.27 (19.8180) |
| Total | 146932.04 (781.1193) |

# Results




In region 1 the task should be performed locally, independently of the type of transport network.

In region 2 the task can be performed remotely provided that the connection is a light path.

In region 3 the task should be done remotely for both types of transport networks.

Given different network paths we can identify decision boundaries as function of the task complexity.

# Current work

## The network matters!

( if the transport network is powered by dirtier energy than the data, the contribution of the network to the total cost in gr. CO2 for moving data can be significant ).

Cluster Green Software

How can cloud software use the information on the network to make the ‘right’ decision?

How can cloud software use the information on the network to make the ‘right’ decision?

GigaPort

# More information

Calculator and report are online:

- http://sne.science.uva.nl/bits2energy/

- http://www.surf.nl/en/knowledge-and-innovation/knowledge-base/2013/research-report-transporting-bits-or-transporting-energy-does-it-matter.html

***Publications:***

- "Storage to Energy: modeling the carbon emission of storage task offloading between data centers" presented at the CCNC conference( Las Vegas Jan. 2014)

- " A decision framework for placement of applications in clouds that minimizes their carbon footprint" in Journal of Cloud Computing (Springer Open Journal)