Energy Savings of Sleep Modes enabled by 5G Software-Defined Heterogeneous Networks

M. Klapež, Carlo A. Grazia, **Maurizio Casoni**
maurizio.casoni@unimore.it

UNIMORE Networking Lab
netlab.unimore.it

Department of Engineering *Enzo Ferrari*
University of Modena and Reggio Emilia

Palermo - Italy - EU, September 10-13, 2018

RTSI 2018

Main sources of energy consumption

The vast majority of mobile networks' energy consumption is given by operating the facilities, i.e. **base stations** and technology centers.

The major environmental impact and energy usage of base stations comes from electricity consumption in the use phase.

For a typical urban base station site in Europe, 84% of its overall energy consumption is generated by its use, compared to 14% generated by production and 2% from logistics.

$\Downarrow$

The energy amount needed to power a mobile network is mainly determined by *the amount of geographical coverage that it provides (i.e.* **the number of base stations***), instead of by the generated traffic.*

Imminent threats to energy savings

"Real" 4G (LTE Advanced) / 5G

$\Downarrow$

Higher frequencies

$\Downarrow$

Smaller cells

$\Downarrow$

More cells!!!

Traffic load in mobile networks

The load in mobile networks is anything but linear in both the spatial and temporal domains.

- Temporal domain: **the networks' utilization remains low most of the time**, commonly for more than 12 consecutive hours in outdoor sites, while peak loads last regularly less than 30 minutes and are generated by a limited subset of users.
- Spatial domain: **different sites are almost never under peak loads at the same time, but are frequently in minimum usage conditions simultaneously**. Furthermore, different sites serve on average different traffic amounts.

Just use sleep modes!

It follows that the average mobile networks' efficiency (the rate of actual resource utilization of the network capacity) remains low.

$\Downarrow$

Just use **sleep** modes then! Computers and smartphones do that all the time. Even cars! (Full Hybrid cars)

The challenge

The sleep concept makes sense. Moreover, straightforward.

Execution IS NOT.

Why?

The challenge

The sleep concept makes sense. Moreover, straightforward.

Execution IS NOT.

Why?

Because you need to provide network services 24h.

Today BSs need to permanently signal their presence and to continuously listen to the radio environment in order to provide seamless coverage to end users.

$\Downarrow$

You make a BS sleep, you drop **coverage**, capacity and throughput.

A likely scenario

BS ‘Green’ is operating at its maximum output power, and it does not detect terminals requesting network services for a certain amount of time.

It then enters sleep mode, where it behaves like it’s off but it’s actually ready to go back to a normal state very quickly, consuming only a little bit of energy.

$\Downarrow$ 10 minutes later $\Downarrow$

A terminal ‘Angry’ requesting network services enters the area formerly covered by BS ‘Green’.

$\Downarrow$ What happens now $\Downarrow$

A likely scenario

BS ‘Green’ is operating at its maximum output power, and it does not detect terminals requesting network services for a certain amount of time.

It then enters sleep mode, where it behaves like it’s off but it’s actually ready to go back to a normal state very quickly, consuming only a little bit of energy.

⇓ 10 minutes later ⇓

A terminal ‘Angry’ requesting network services enters the area formerly covered by BS ‘Green’.

⇓ What happens now ⇓

nothing

What?

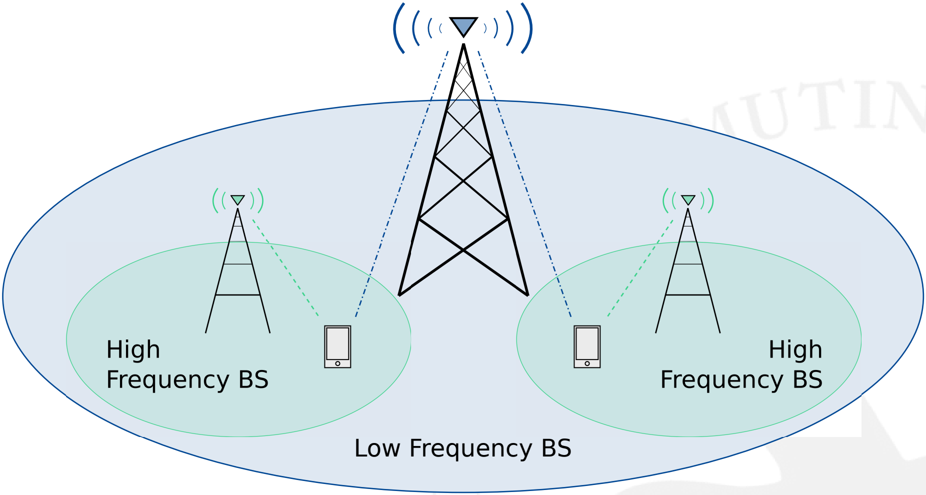
BS ‘Green’ can’t detect terminal ‘Angry’ even if the latter is in its usual range.

On the other hand, terminal ‘Angry’ has no means to signal its presence..

$\Downarrow$

Deadlock!

Heterogeneous Networks



Low vs High frequencies

Low frequencies

- Long wavelengths
- Few information
- Signal carried over long distance

High frequencies

- Short wavelengths
- Lot of information
- Signal carried over short distance

In order to offer the same coverage, high-frequency macro cells would need a dramatic amount of transmission power to overcome signal degradation.

For instance, **an LTE network may have to consume around 60 times the energy expended by a 2G network to offer the same coverage.**

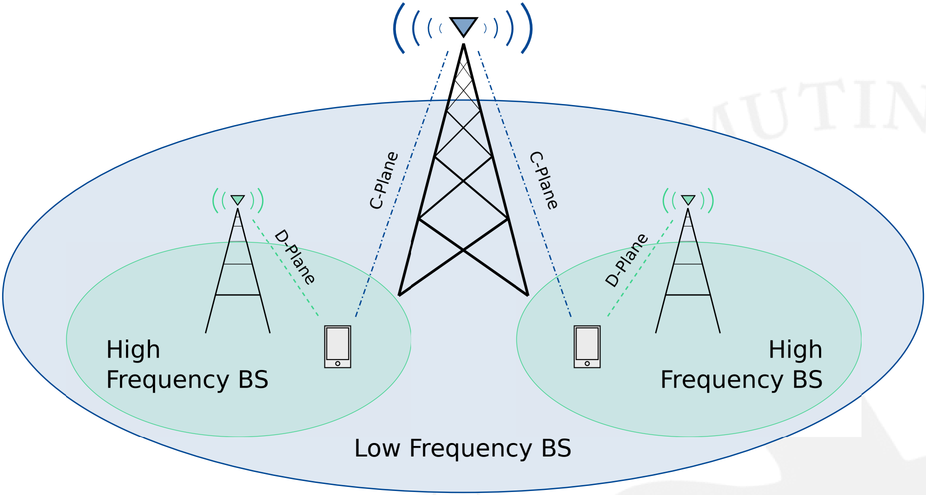
Abstractions

"Abstraction is at the center of much work in Computer Science. It encompasses finding the right interface for a system as well as **finding an effective design for a system implementation**." - B. Liskov

SDN (Software-Defined Networking) is entirely based on **one key abstraction: the separation between Control Plane and Data Plane**.

Why not introducing the same in HetNets radio access?

‘Abstracted’ HetNets (SDHNs)



A new scenario

The area covered by the HF BS ‘Green’ is also covered by LF BS ‘Blue’. BS ‘Green’ is operating at its maximum output power, and it does not detect terminals requesting network services for a certain amount of time.

It then enters sleep mode, consuming only a little bit of energy.

$\Downarrow$ 10 minutes later $\Downarrow$

A terminal ‘Happy’ requesting network services enters the area formerly covered by BS ‘Green’.

$\Downarrow$ What happens now $\Downarrow$

A new scenario

The area covered by the HF BS ‘Green’ is also covered by LF BS ‘Blue’. BS ‘Green’ is operating at its maximum output power, and it does not detect terminals requesting network services for a certain amount of time.

It then enters sleep mode, consuming only a little bit of energy.

⇓ 10 minutes later ⇓

A terminal ‘Happy’ requesting network services enters the area formerly covered by BS ‘Green’.

⇓ What happens now ⇓

BS ‘Blue’ **wakes up** BS ‘Green’. 👍

Applicable from LTE onwards

The signal generated by cells operating legacy technologies, e.g. 3G, is usually capable to penetrate indoor even if those cells are deployed outdoor.

3G also lacks the interference management capabilities of LTE.

Particularly if high frequency bands are used, instead, the signals of LTE, LTE-Advanced and especially mm-wave (i.e. 5G) outdoor cells are mostly confined outdoor.

Common outdoor building materials, in fact, present high penetration resistance to mm-waves, so as human bodies do.

$$\Downarrow$$

HetNets are expected to be surely deployed in outdoor-indoor settings.

Why yet another Metric?

Several already discussed (ECI, ECR, PIs, APC, γ, ECG, AGE, $etc.$).

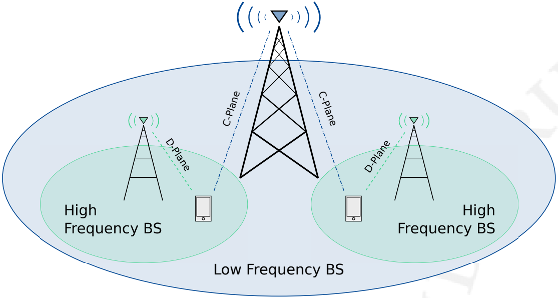
Often, they evaluate correlations under a single point of view. e.g.

- Service perspective: energy consumption VS
  - capacity
  - **data rate**
  - blocking probability
- Deployment perspective: energy consumption VS
  - BS number
  - **coverage area**
  - max. number of users

Coverage area and **data rate** often belong to different metrics.
Even when together, **it is not trivial to vary and study them separately**.

Basics

An SDHN network is composed of a number n of **clusters** where $n > 0$.



A **cluster**

A **cluster** is a set of cells that provide signaling and data to users in a defined geographical area, through a low-frequency BS for the C-Plane and a number m of high-frequency cells for the D-Plane.

Instantaneous Power of a SDHN: *Clusters*

The overall power required by the network:

$$P_{SDHN} = \sum_{x=1}^{n} P_{cluster_x} \tag{1}$$

where $cluster_x$ represents a cluster and $P_{cluster_x}$ the power it needs.

The latter, in turn, is defined as:

$$P_{cluster_x} = P_C + P_D \tag{2}$$

where P_C is the power needed by the cluster C-Plane and P_D that needed by the cluster D-Plane.

Instantaneous Power of a SDHN: *C-Plane*

P_C **is function of coverage area**, P_D is function of traffic load.

$$P_C = P_{max_{LF}} \cdot PC_{LF}(a) \quad (3)$$

- $P_{max_{LF}}$ is the maximum power needed by the Low-Frequency cell.
- a is the coverage area that the cell should serve.
- The Power Coefficient function $PC_{LF} : \mathbb{R} \rightarrow \mathbb{R}$ is defined by $PC_{LF} : a \mapsto [min_{PC}, 1]$ where min_{PC} is the minimum scale factor that the cell can reach by modulating its output power (e.g. in the case of LTE, $min_{PC} \approx 0.5$, as around 50% of an eNodeB power consumption may scale linearly with the output power).

(The notation $cluster_x$ is not included for clarity of presentation)

Instantaneous Power of a SDHN: *D-Plane*

P_C is function of coverage area, P_D **is function of traffic load**.

$$P_D = \sum_{y=1}^{m} P_{D_y} \tag{4}$$

with m as the number of High-Frequency cells in the cluster and with P_{D_y} as the power needed for a High-Frequency cell y to provide D-Plane capabilities.

The power needed for the D-Plane must also account for cells in sleep mode.

(The notation $cluster_x$ is not included for clarity of presentation)

Instantaneous Power of a SDHN: *D-Plane* (2)

P_C is function of coverage area, P_D **is function of traffic load**.

$$P_{D_y} = S_{0_y} \cdot \left[P_{max_{HF_y}} \cdot PC_{HF_y}(l_y) \right] + S_{1_y} \cdot \left[PS_y \right] \quad (5)$$

- $P_{max_{HF_y}}$ denotes the maximum power needed by a cell y.
- l_y is the y traffic load.
- The Power Coefficient function $PC_{HF_y} : \mathbb{R} \to \mathbb{R}$ is defined by $PC_{HF_y} : l_y \mapsto [min_{PC_y}, 1]$ where min_{PC_y} is the minimum scale factor that the cell y can reach by modulating its output power.
- PS_y is the power needed by the cell y while in sleep mode.

S_0 **simply blanks the term that expresses the power consumed by an active cell** y **if** y **itself is sleeping, while** S_1 **blanks** PS_y **if the cell is active.**

(The notation $cluster_x$ is not included for clarity of presentation)
(Square brackets for clarity only, they denote what depend on cell sleeping or not)

Instantaneous Power of a SDHN: *D-Plane* (3)

P_C is function of coverage area, P_D **is function of traffic load**.

S_0 simply blanks the term that expresses the power consumed by an active cell y if y itself is sleeping, while S_1 blanks PS_y if the cell is active.

S_{0_y} and S_{1_y} are therefore piecewise functions defined as follows:

$$S_{0_y} = \begin{cases} 0 & \text{if } y \text{ is sleeping} \\ 1 & \text{if } y \text{ is active} \end{cases} \qquad S_{1_y} = \begin{cases} 0 & \text{if } y \text{ is active} \\ 1 & \text{if } y \text{ is sleeping} \end{cases}$$

(The notation $cluster_x$ is not included for clarity of presentation)

Energy Consumption of a SDHN in T: *Clusters*

$$P_{SDHN} = \sum_{x=1}^{n} P_{cluster_x}$$

$$\Downarrow$$

$$E_{SDHN} = \sum_{x=1}^{n} E_{cluster_x} \tag{6}$$

$$P_{cluster_x} = P_C + P_D$$

$$\Downarrow$$

$$E_{cluster_x} = E_C + E_D \tag{7}$$

Energy Consumption of a SDHN in T: *C-Plane*

$$P_C = P_{max_{LF}} \cdot PC_{LF}(a)$$

$$\Downarrow$$

$$E_C = \int_{t_0}^{t_1} \Big(P_{max_{LF}} \cdot PC_{LF}(a(t))\Big) d(t) \quad : \quad t \in [t_0, t_1] \tag{8}$$

where $a(t)$ is the area coverage to offer at time t and $[t_0, t_1]$ is the time interval for which $T = t_1 - t_0$.

Energy Consumption of a SDHN in T: *D-Plane*

$$P_D = \sum_{y=1}^{m} P_{D_y}$$

$$\Downarrow$$

$$E_D = \sum_{y=1}^{m} E_{D_y} \tag{9}$$

with E_{D_y} as the energy consumed by a cell y in the time interval T.

Energy Consumption of a SDHN in *T*: *D-Plane* (2)

$$P_{D_y} = S_{0_y} \cdot \left[P_{max_{HF_y}} \cdot PC_{HF_y}(l_y) \right] + S_{1_y} \cdot \left[PS_y \right] \quad (5)$$

$$\Downarrow$$

$$E_{D_y} = \int_{t_0}^{t_1} \left(S_{0_y} \cdot \left[PA_y \right] + S_{1_y} \cdot \left[PS_y \right] \right) dl_y$$

$$PA_y = P_{max_{HF_y}} \cdot PC_{HF_y}(l_y(t)) \quad (10)$$

$$: \quad t \in t_0 t_1$$

Note that contrarily than in Equation (5), in Equation (10) the parameter $l_y(t)$ also accounts for its variability in time.

Scenario

It represents office buildings in an urban area.

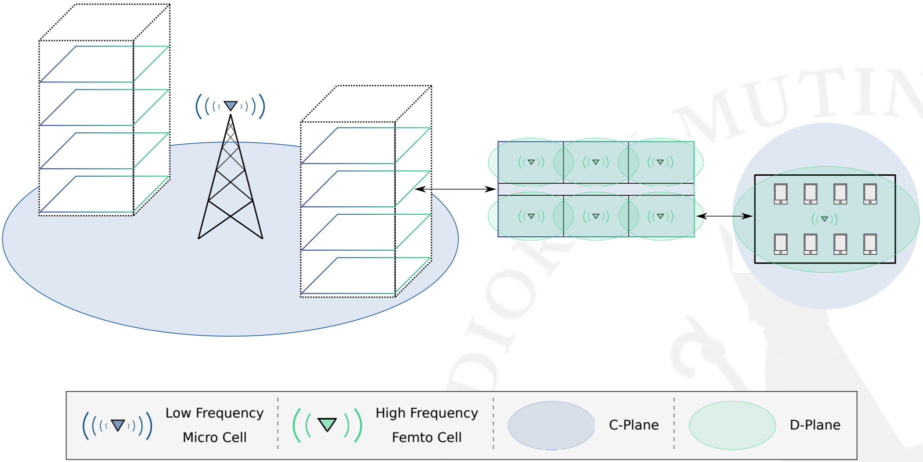
Rationale for the choice:

- Data-communication-intensive.
- Typically empty for at least half a day.
- Few shadowing effects caused by people movements.

Each Building has four floors, and each floor includes six office rooms. Each office is considered of being sized approximately $60m^2$ and, during working day hours, to contain an average of 8 active terminals served by an LTE femto cell deployed in the room.

384 users total.

Scenario


Low Frequency
Micro Cell
High Frequency
Femto Cell
C-Plane
D-Plane

Model

For simplicity, C-Plane is exclusively served through the micro cell and the **D-Plane is exclusively served by the femto cells**.

Without this constraint, **the micro site could also serve low traffic requests while femto cells are sleeping, saving even more energy**.

	Micro Cell	Femto Cell
Vendor	NEC	Fujitsu
Model	MB4420	LS100
Environment	Outdoor	Indoor
Max. Power Consumption	150 W	13.5 W
Max. Data Rate	N/A (Only used as C-Plane)	75 Mbps UL, 112 Mbps DL
Transmit Output Power	Adjustable, 1-5 W	Adjustable, ~50 mW
Bandwidths	5, 10, 15, 20 MHz	5, 10, 15 (DL only) Mhz
MIMO	2×2	2×2 (DL only)
Max. Number of Users	N/A (Only used as C-Plane)	16

TABLE I: Cells Technical Specification

Model

Application	9-13, 14-18	13-14	8-9, 18-19
Web browsing	25%	15%	10%
VoIP & IM	15%	10%	5%
Downloads & updates	20%	0%	0%
Social networking	10%	10%	10%
Video streaming	10%	25%	10%
Total	80%	60%	35%

TABLE II: Applications Traffic Load

From mobile traffic estimates of the Radiocommunication Sector of the International Telecommunication Union

Traffic Load	Femto Cells	Micro Cell
100% (or PM OFF)	13.5 W (649 Wh)	132-150 W
80%	11.9 W (571.2 Wh)	132-150 W
60%	10.3 W (494.4 Wh)	132-150 W
35%	8.2 W (393.6 Wh)	132-150 W
0% (Sleep, if ON)	0.9 W (43.2 Wh)	132-150 W

TABLE III: Cells Energy Consumption

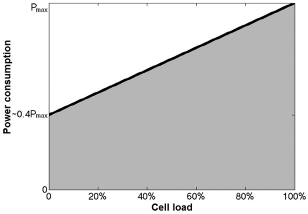
Models a typical 8-hour workday with some flexibility.
The resulting estimated traffic pattern is consistent with real average traffic patterns reported by network operators.

Model

Simplified "on-off" power profiles where an active BS is assumed to consume a fixed amount of energy while a sleeping BS consumes almost none are often used in the literature.

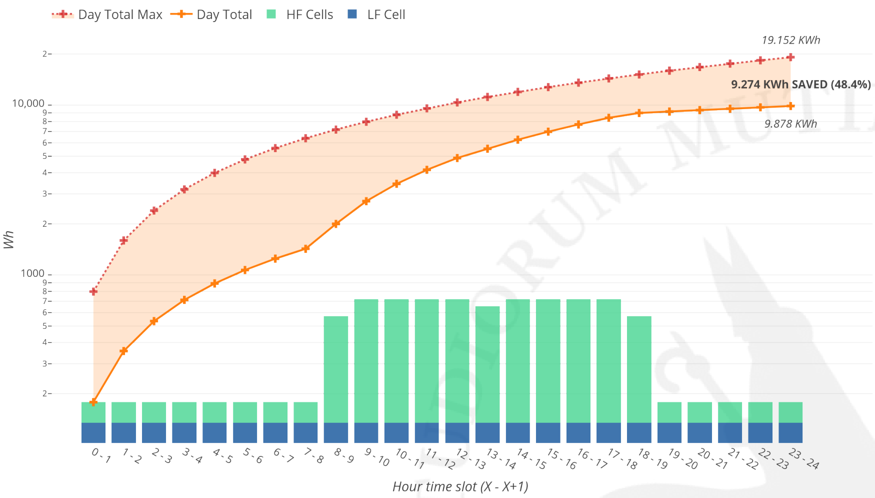
Very strong simplification.

Plus, as network densification results in more and more deployments of micro, pico and femto cells using **digital power amplifiers with large peak-to-average power ratios, the traffic-dependent portion of energy consumption becomes significant**.



For an LTE cell, estimates indicate that approximately 50-60% of the BS power usage scales linearly with the traffic load.

Energy consumption difference ($min_{PC_y} = 0.5$)



Thank you

Questions?

UNIMORE Networking Lab - **netlab.unimore.it**

FAQ

Can you quantify the switching time necessary to sleep/awake a femtocell?

- This is currently under investigation.
- Can depend on manufacturer.
- Should be negligible, because no collaborative gradual transitions and associated handovers are necessary (as in cell wilting-blossoming).

What happens to the users that are being served during the switch off?

- No switch off happens if users are being served.

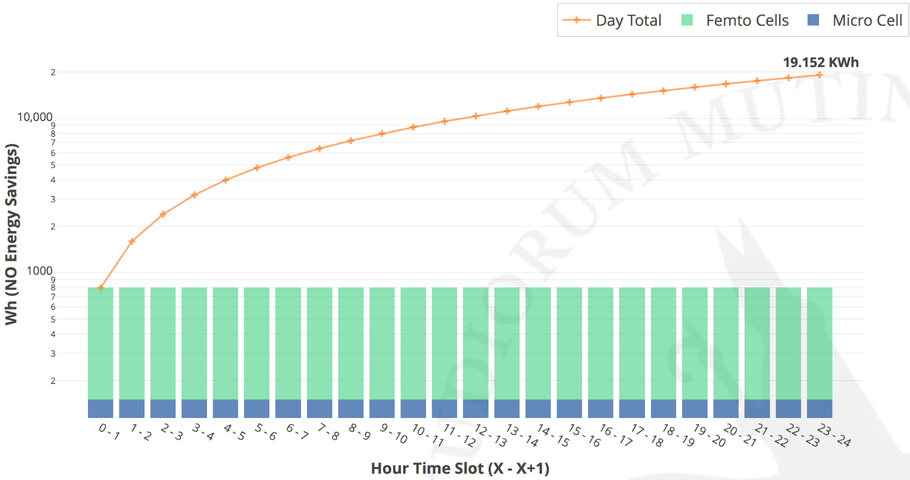
Isn't using microcells to serve C-Plane only an unrealistic assumption?

- No.
- Besides, if microcells start serving low-rate D-Plane too, you save even more energy.

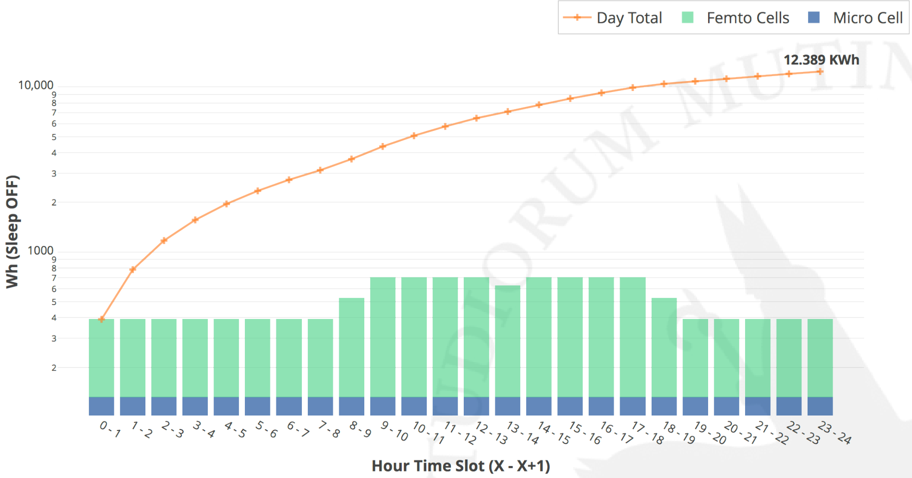
Isn't a comparison with other approaches in literature missing?

- Yes, impossible to even think about starting it in 6 pages.

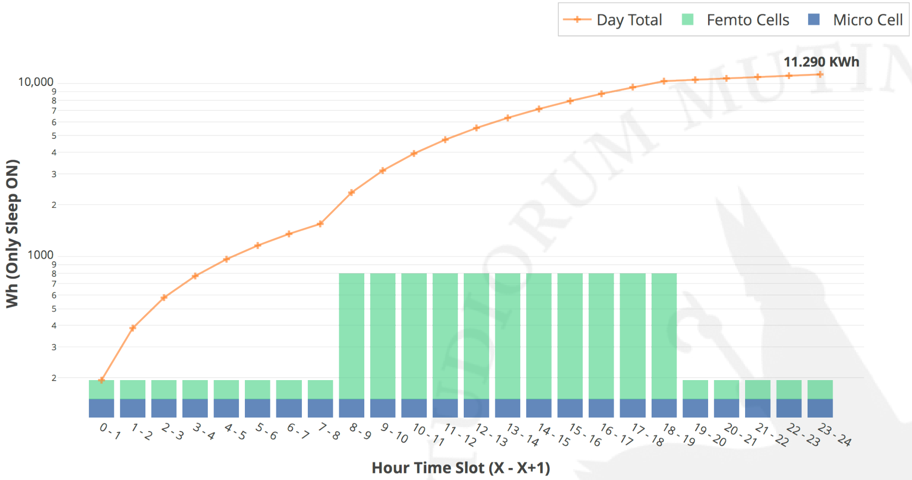
[EXTRA] Base case ($min_{PC_y} = 0.4$)



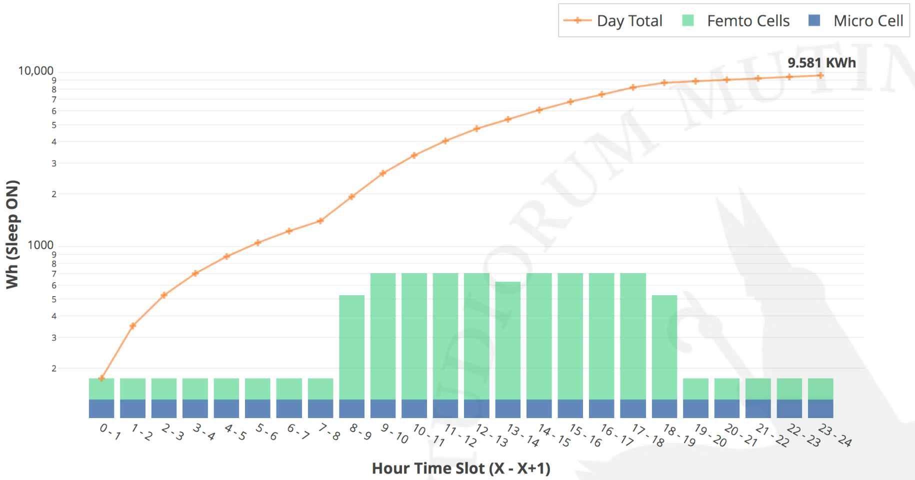
[EXTRA] Output Power Modulation ON and Sleep OFF ($min_{PC_y} = 0.4$)



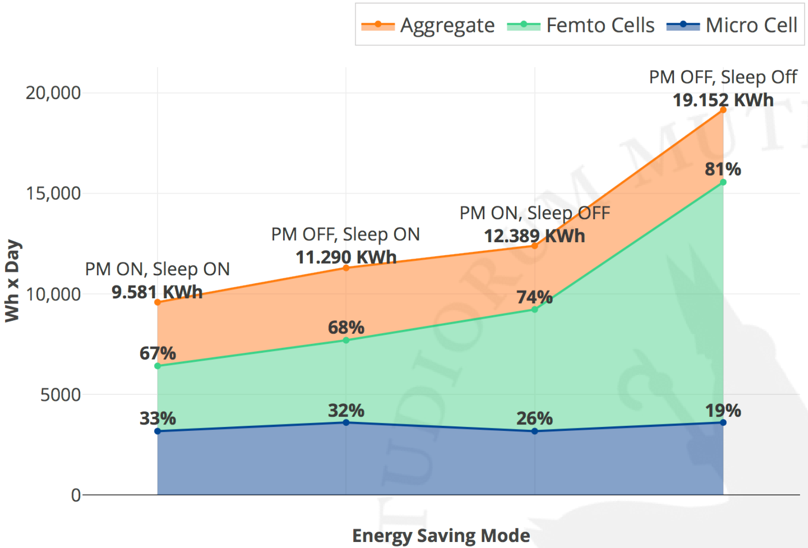
[EXTRA] Output Power Modulation OFF and Sleep ON ($min_{PC_y} = 0.4$)



[EXTRA] Output Power Modulation ON and Sleep ON ($min_{PC_y} = 0.4$)



[EXTRA] Energy consumption distribution ($min_{PC_y} = 0.4$)



[EXTRA] Energy consumption difference ($min_{PC_y} = 0.4$)