Calculating the Energy Consumption of a Website

June/July 2017

Anouk Boukema
Supervisor: Maarten de Waard

Motivation

Environmental Concerns

- Awareness
- Insight
- Motivation

greenhost

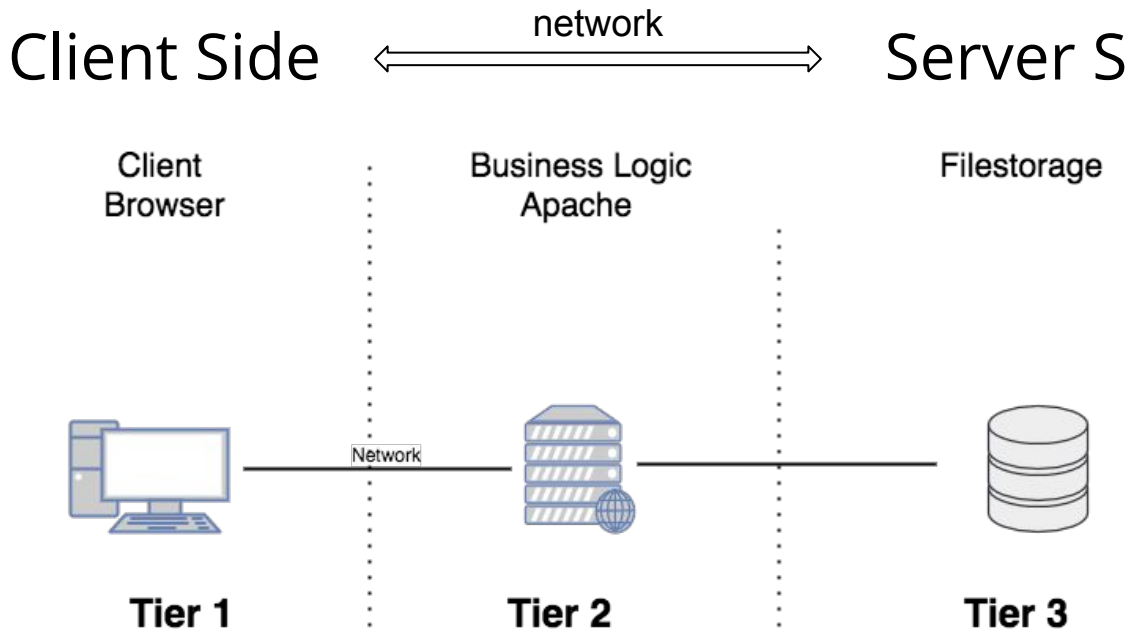
Research Question

How to calculate the energy consumption of a website?

Sub questions:

- What are the energy consuming components of a website?
- What data can be measured at these components?
- How does this data relate to the total energy usage of the underlying machine

Energy consuming components of a website

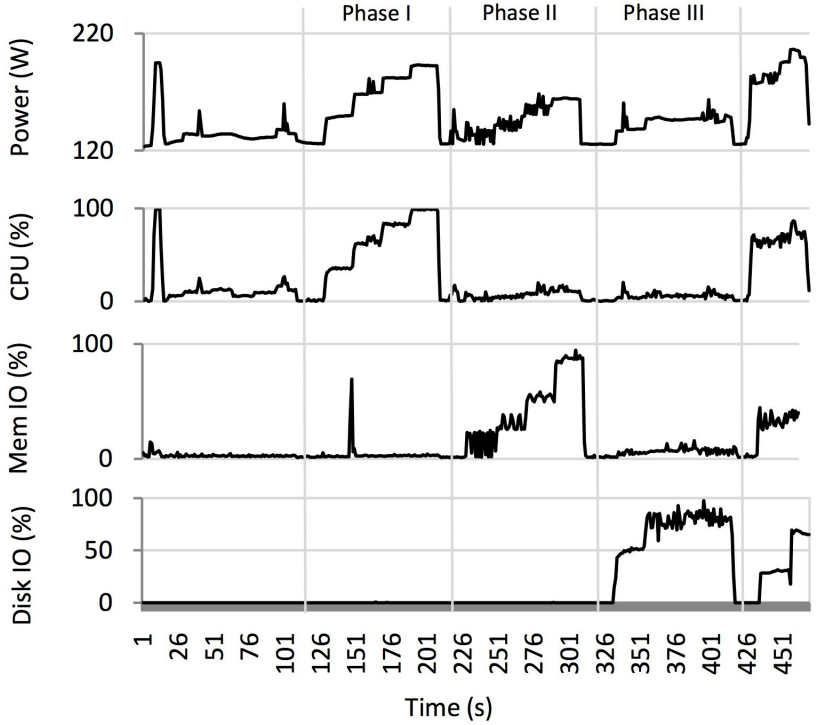


[1] Maarten van Steen Andrew S. Tanenbaum. Distributed Systems Principles and Pradigms. Vrije Universiteit, 2016.

Related Work

Relative power impact of different resources on dynamic power consumption.

CPU = 58%
MEM = 28 %
Disk = 14% [1]



[1] Aman Kansal, Feng Zhao, Jie Liu, Nupur Kothari, and Arka A Bhattacharya. Virtual machine power metering and provisioning. In Proceedings of the 1st ACM symposium on Cloud computing, pages 39–50. ACM, 2010.

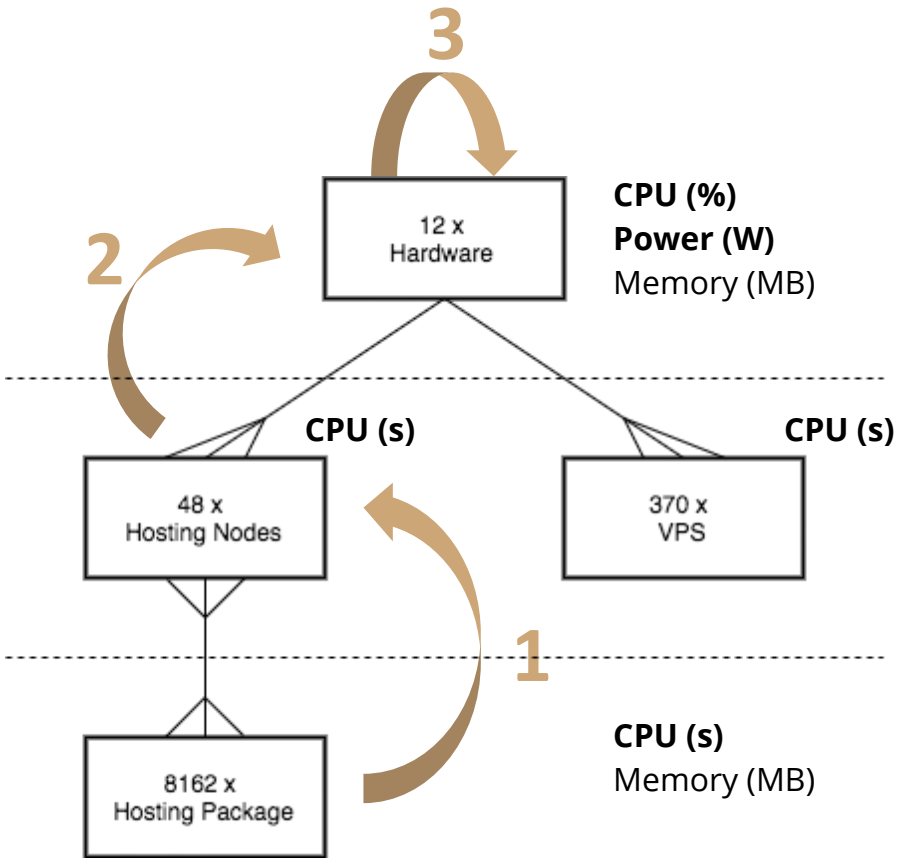
Architecture

CPU second (s)
- Active processing of one core

CPU usage (%)
- percentage of the total CPU's capacity

Power (W)
- Total amount of Wattage going into baremetal machine

Memory (bytes)
- In buffer & cache



Approach

Assumption:
The data of each layer is correlated with the others over time.

Test:

1. Plot
2. Fit - Linear Regression on Training set
3. Test - accuracy (mean squared error) on Test set

Answer research question

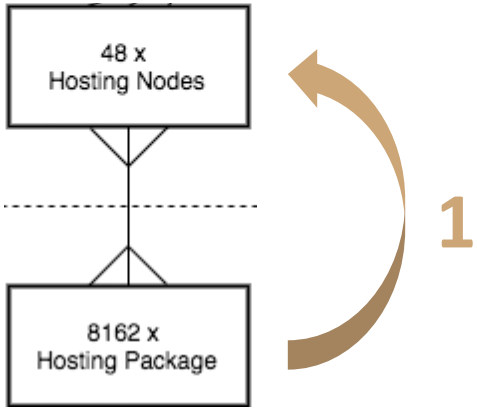
- Creating formula translating the CPU(s) of a Hosting Packages → Power used by the hardware.

Part 1 - Pre-processing

+ Hosting Nodes only contain packages
+ Known which packages run on which hosting node

$CPUhn_i \approx \sum CPUpack_i$

Points in interval 776
Hosting nodes 48
---------- x
Data points 37,248

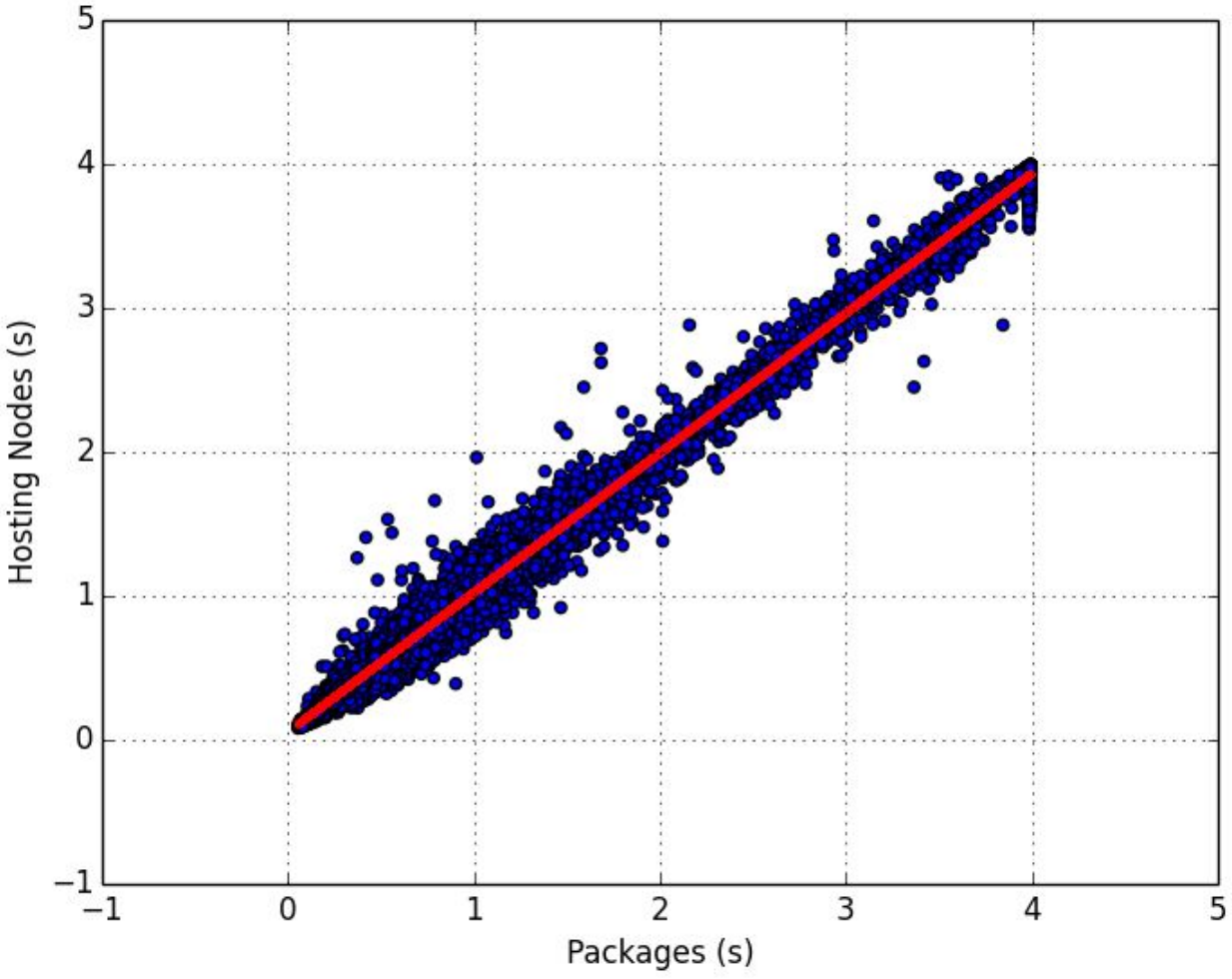


Part 1 - Results

Datapoint (x,y) =
$(CPUhn_i , \sum CPUpack_i)$

$CPUhn_i =$
$0.97 \times \sum CPUpack_i + 0.054$

Mean Squared Error = 0.0054

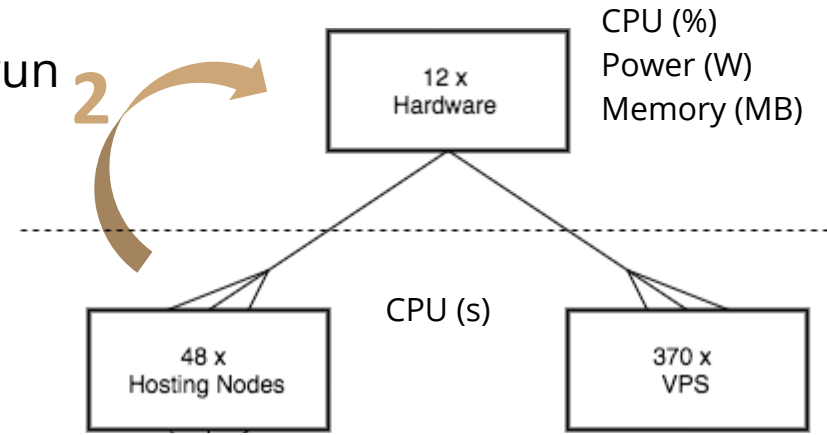


Part 2 - Pre-processing

+ Hardware nodes only run HN + VPS
- No knowledge on which HN and/or VPS's run on which Hardware node.
- CPU of hardware nodes is measured in percentages instead of seconds.

$$\sum CPUhw \approx a \times (\sum CPUhn + \sum CPUvps) + b$$

Data points = 776

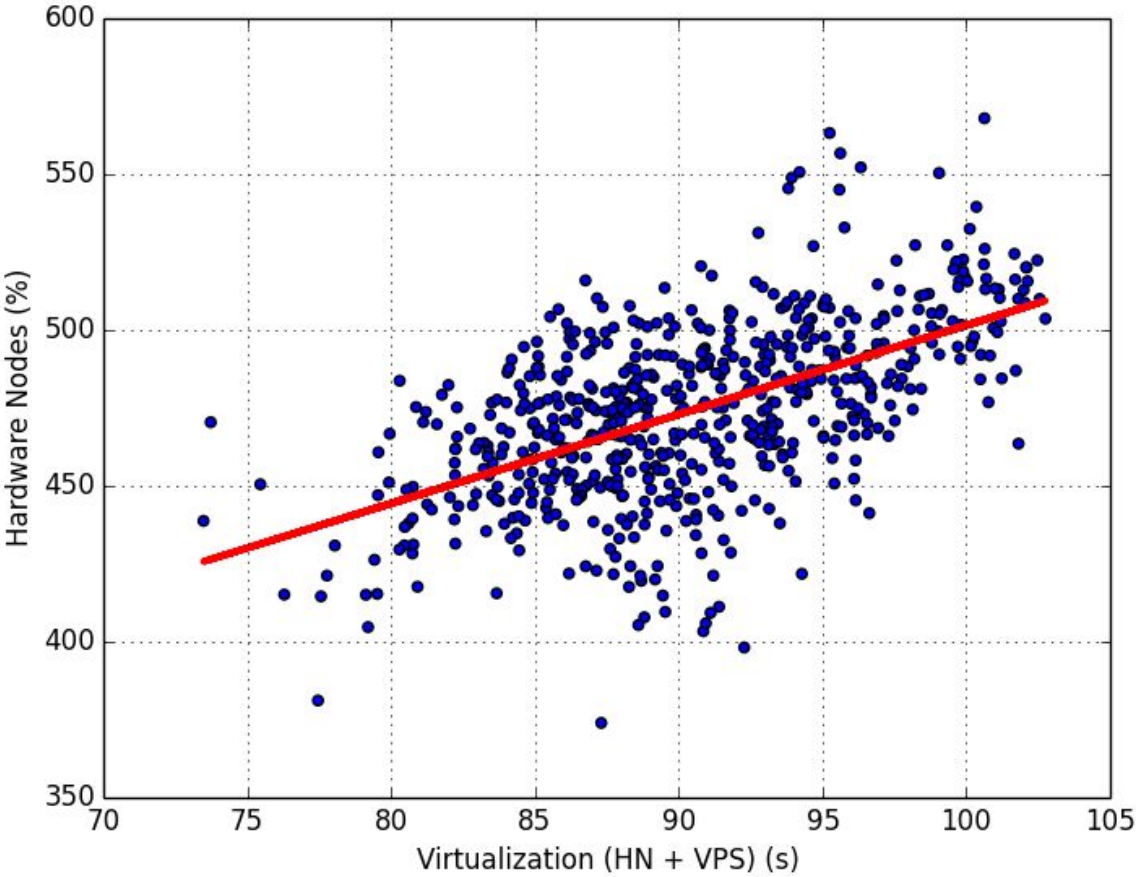


Part 2 - Results

datapoint (x,y) =
(∑CPUhw , ∑CPUhn + ∑ CPUvps)

a = 2.82
b = 219.81

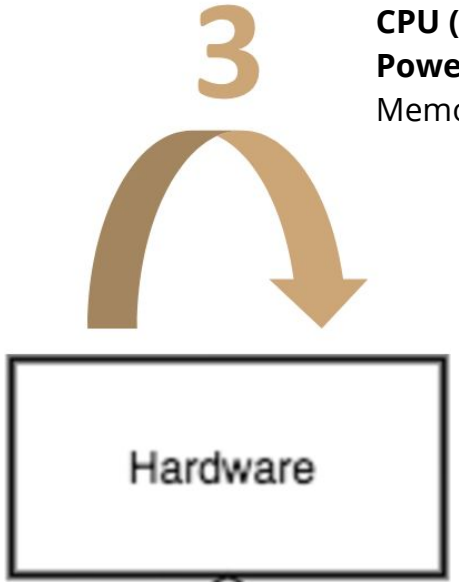
mean squared error = 530.83
C.a. 23 %



Part 3 - Pre-processing

$Phw_i \approx a \times CPUhw_i + b$

Points in interval =	776	
Hardware nodes =	12	x
Data points	= 9.312	

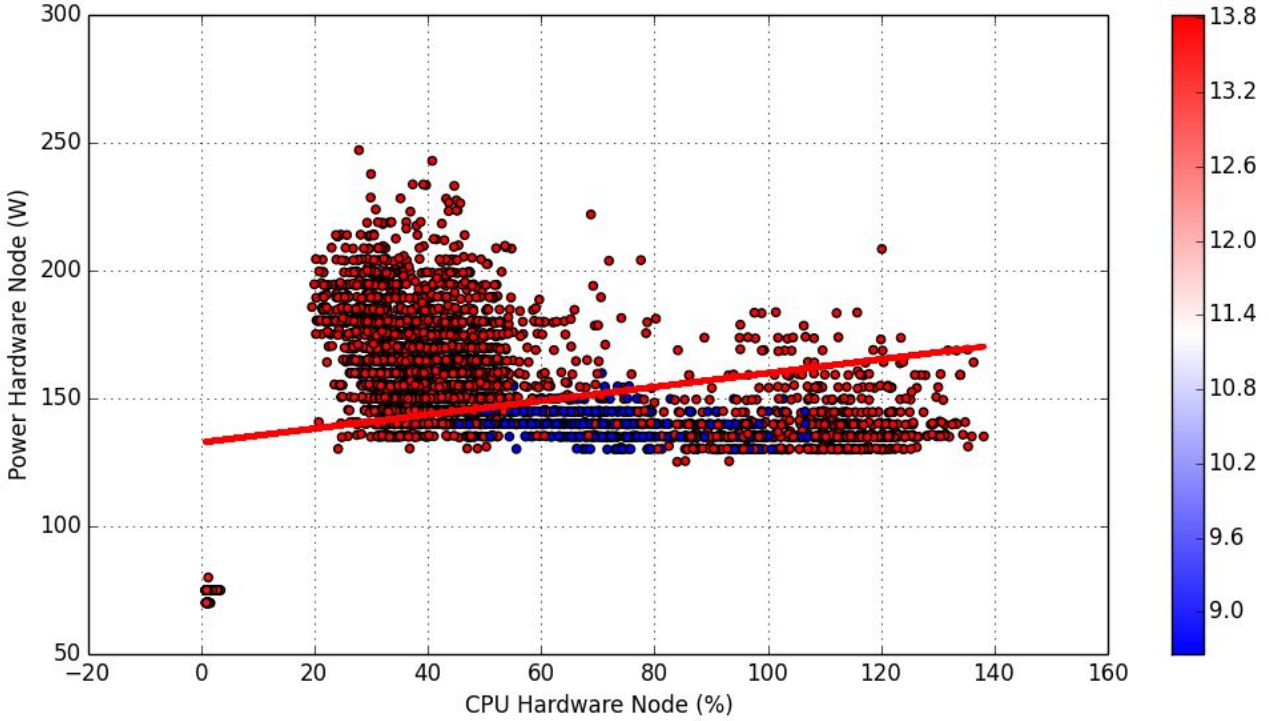


CPU (%)
Power (W)
Memory (MB)

Part 3 - Results

Datapoint (x,y,z(color)) =

$(Phw_i , CPUhw_i , MEMhw_i)$



mean squared error = 934.62 (c.a. 30 W)

Power(w) = 0.32 × CPUhw + 3.3 × MEMhw + 87.34

Final Formula

$\sum Phw = a*\sum CPUpack + b* \sum CPUvps + c * \sum MEMhw + d$

a = 0.867663

b = 0.895096

c = 3.30113

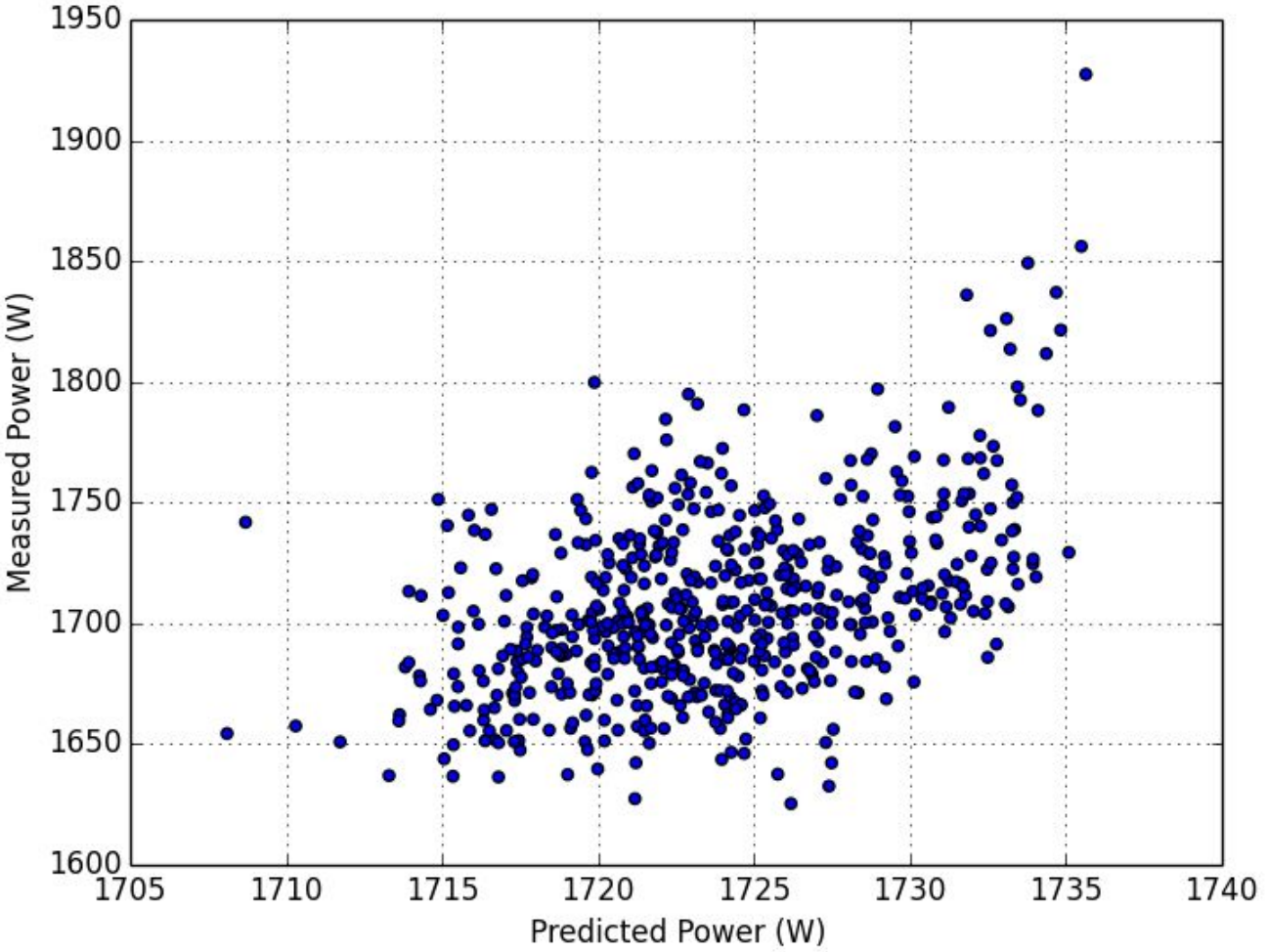
d = 1118.6

Verify this formula by plotting measured power at a certain time against the predicted power at the same time, and calculate the mean squared error

Power Prediction

Datapoint (x,y) =
$(\sum Phw_i , \sum Phwpredict_i)$

Mean Squared Error:
1536.22
C.a. 40 Watt



Energy Consumption of a Website

$\sum Phw_{predict} = a*\sum CPUpack + b* \sum CPUvps + c * \sum MEMhw + 1118.6$

Assumption 1: MEMhw = MEMpack
$virtualization_{tot}$ = 418
hn_{tot} = 48 → 48/418 = 11 %

$\sum Phw = a*\sum CPUpack + c *\sum MEMpack + 0.11*1118.6$

$packages_{tot}$ = 8162

$Phw = a*CPUpack + c * MEMpack + (0.11*1118.6)/8162$

power min	0.768 W
power average	4.23 W
power max	12.25 W

Conclusion

With an accuracy of ± 40 W it is possible to estimate the energy consumption of a website given the CPU in seconds, and Memory in bytes of that website.

Discussion & Future work

Calculated Energy Consumption might differ from reality:

- Other resources/processes might influence the power consumption
- Linear regression might not be sophisticated enough to calculate power consumption from the data
- Relationship MEMhw and MEMpack should be researched

- Look at other tiers for complete power consumption
- Generalize for other hosting companies

Questions?