PREDICTIVE MODEL for PUE CALCULATION

- Parikshit Bhaduri, Vice President – Product Development,

GreenField Software Private Limited

Introduction

Data center is a key infrastructure of a computing environment. Centralised data centres are replacing standalone IT infrastructure, to gain efficiency and cost effectiveness. The **ACE** (*Availability Capacity & Efficiency*) are metrics, which are important for sustained, future ready and efficient Data Centre operations. The challenge for any organization is, how to optimally balance these three metrics for providing effective and efficient IT services. The energy cost of running a midsized Data center, on an average, amounts to 20% of the Total Cost of Ownership (TCO). Hence, conservation of energy, both from a cost and reduction of GHG emissions perspective, is extremely important. PUE (Power Utilization Effectiveness) is one of the key parameters to effectively monitor and achieve power efficiency targets by the IT managers.

The fundamental first step is to prioritize energy-saving opportunities by gaining an understanding of data center energy consumption. It also requires an analysis of demand side systems compared with supply side systems. Demand side includes processors, server power supplies, storage, and networking devices. These components usually account for about half of total consumption. The other half encompasses supply-side systems, such as uninterruptible power supply, power distribution, cooling, lighting, and building switchgear. Supply-side equipment is not an independent consumer of power; its power consumption depends on power demand.

PUE definition

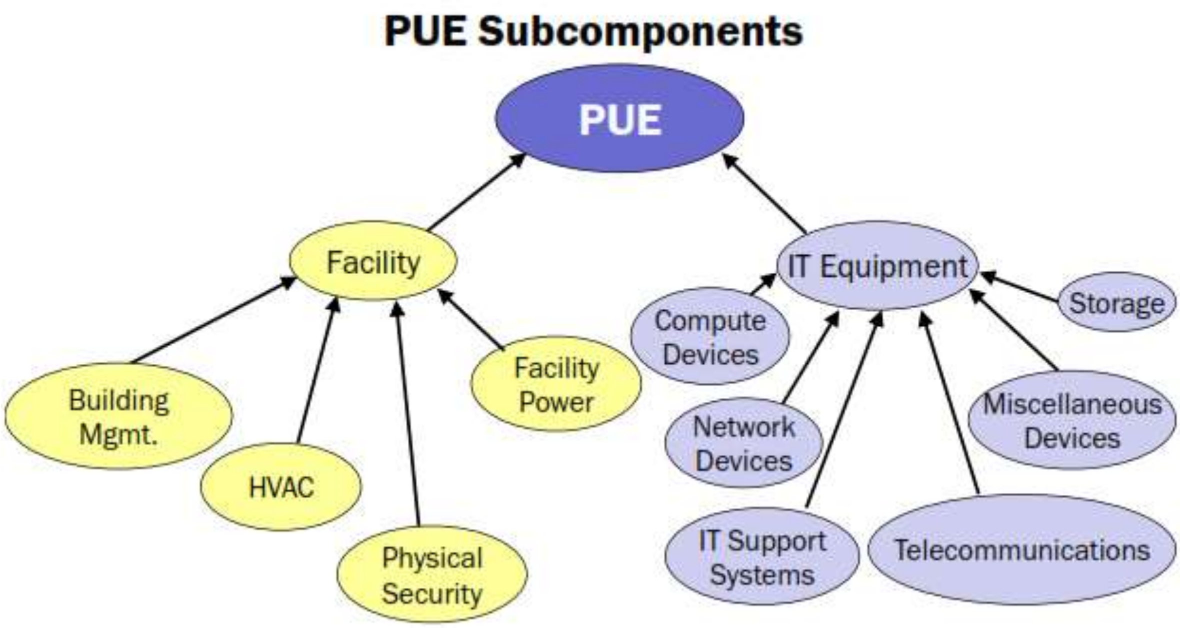
Power usage effectiveness (PUE™) has become the industry-preferred metric for measuring infrastructure energy efficiency for data centers. It was developed by The Green Grid Administration, a non-profit, open industry consortium of end users, policy makers, technology providers, facility architects, and utility companies working to improve the resource efficiency of information technology and data centers throughout the world. Since its original publication in 2007, PUE has been globally adopted by the industry.

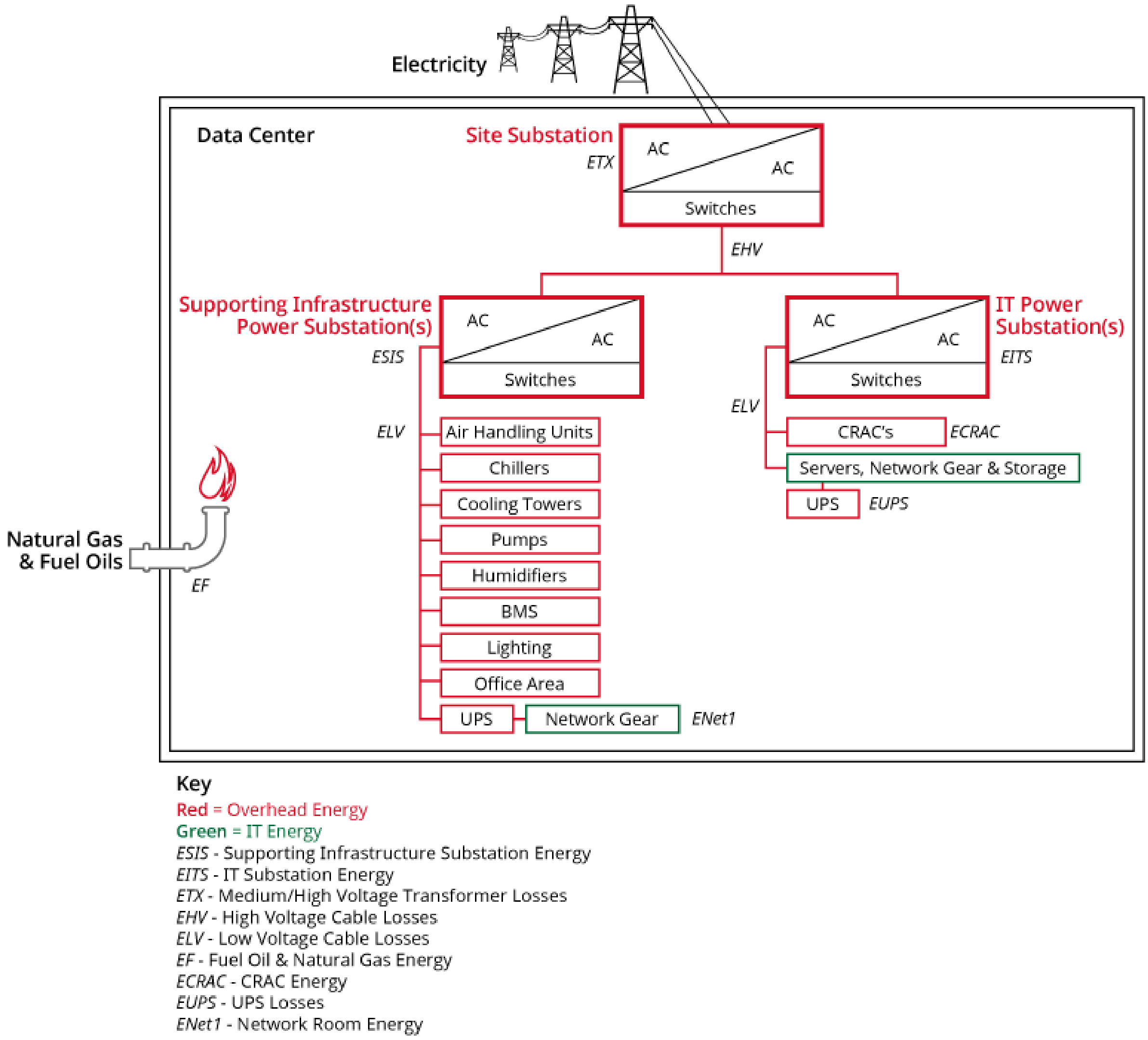
PUE is an excellent metric for understanding how well a data center is delivering energy to its information technology equipment. The metric is best applied for looking at trends in an individual facility over time and measuring the effects of different design and operational decisions within a specific facility. Therefore, comparing two data centers based on public reports of their PUE results was not initially recommended because many attributes of data center design, engineering, implementation, and operations affect PUE. PUE measures the relationship between the total facility energy consumed and the IT equipment energy consumed. When viewed in the proper context, PUE provides strong guidance for and useful insight into the design of efficient power and cooling architectures, the deployment of equipment within those architectures, and the day-to-day operation of that equipment.

While calculating PUE all the non-ICT loads and power losses from all the loads associated with running a Data center are considered as over heads. The effort is to minimize these overheads and achieve PUE closer to one.

$$PUE = \frac{Total\ Facility\ Energy}{IT\ Equipment\ Energy}$$

Total facility energy is defined as the energy dedicated solely to the data center (e.g., the energy measured at the utility meter of a dedicated data center facility or at the meter for a data center or data room in a mixed use facility). The IT equipment energy is defined as the energy consumed by equipment that Is used to manage, process, store, or route data within the compute space. PUE can be calculated using Green Grid Standards metric approach.





Predictive Modelling

While it is important to constantly monitor all the parameters impacting PUE, it is also necessary to come up with a proper predictive model to forecast the probable operating efficiency. This require a rigorous statistical model, which can talk about the significance of several predictors in a prediction scenario. In this line, the analysis of under-utilization of IT load has to be carried out to understand the scope for improving efficiency intelligently.

We introduce a linear model to forecast the probable PUE given an IT load. The model can be scheduled per day to update with current data stream. The model parameter (slope) indicates the possibility of underutilization of IT load or unregulated cooling. The data cleaning requires clipping of noises from predictor variables, for example here we have clipped IT load data those which lies outside of $\pm 10\%$ of the mean. The summary measure has been reported to understand the average IT load required for DC, along with maximum and minimum. The liner model has been generated keeping IT load as independent whereas PUE as the dependent variable. The model can be extended to Multivariate linear regression to understand the significance of other predictor variable in determining PUE for example temperature, humidity, time of the day etc.

$$PUE = f(\% \, of \, IT \, Load) + \varepsilon$$

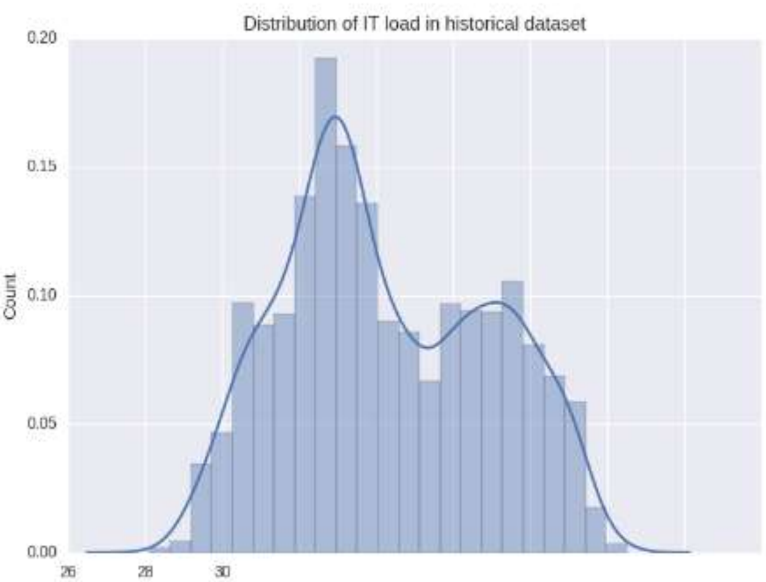
There is also possibility that for different range of IT Load the nature of linearity will change. There can be provision to run piecewise linear regression to understand the nature of the data and to detect the inflection points.

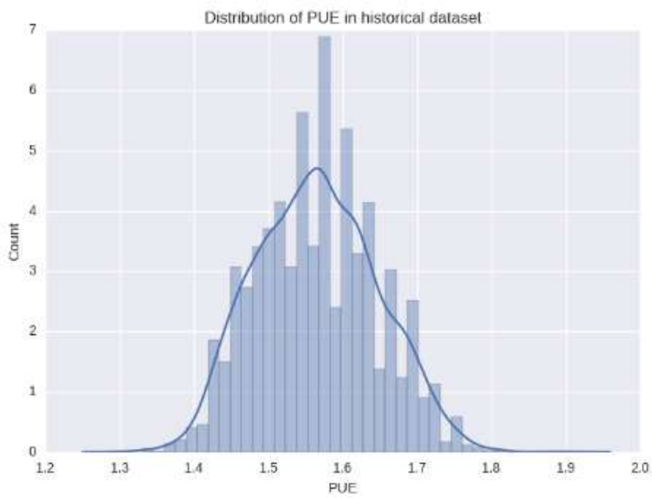
Experiment:

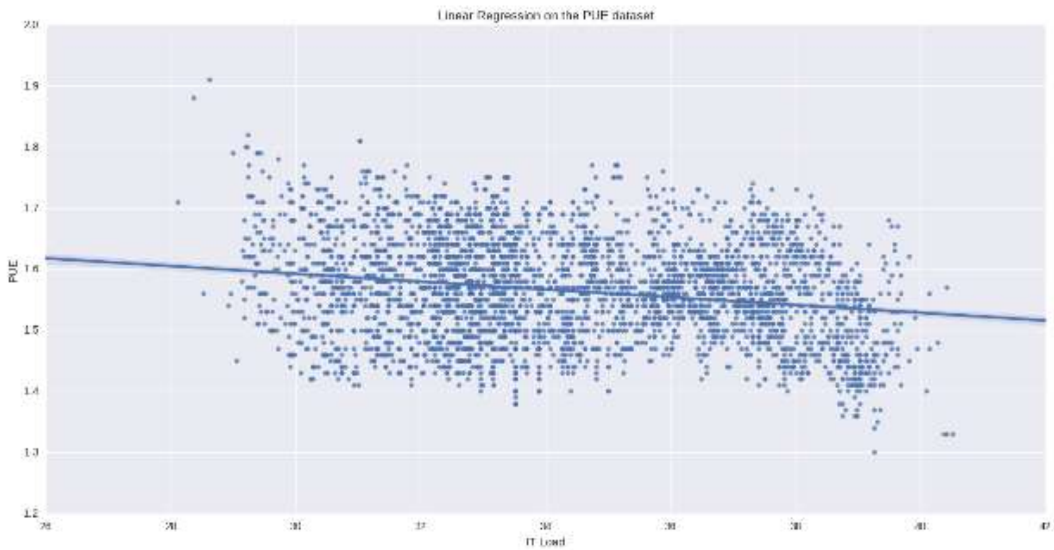
We have run the regression on actual two month's data from a data center and generated following model:

$$PUE = 1.7829465 - (0.0063633 * \%IT\_Load)$$

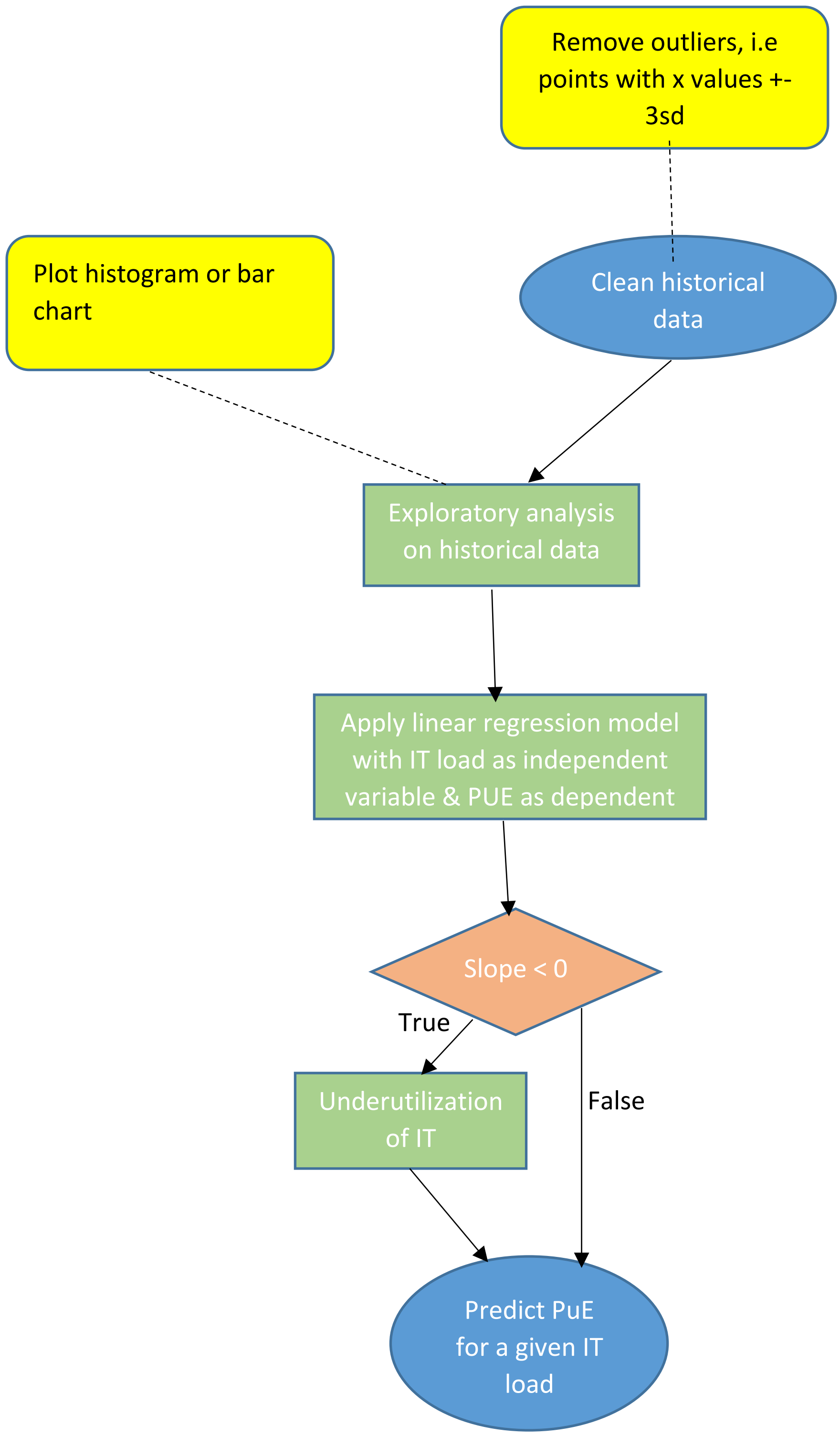
with the predictor variable has p-value $< 2{\times}10^{-16}$. Few observations are reported here.







Flow chart of PUE predictive analysis



Conclusion

PUE is an important metric in assessing the overall energy efficiency of a data center. To maintain the desired level of PUE in a dynamic environment of a Data Centre and to adjust to various ICT and non-ICT loads associated with Data centre operations, there is a requirement of a predictive modelling where various operational Data from a DCIM software is captured, based on which parameters can be forecasted accurately, various measures can be initiated proactively to attain the desired energy efficiency.