

No Power Struggles: Coordinated Multi-level Power Management for the Data Center

Nishad Gothoskar
ngothosk@andrew.cmu.edu

1: Summary

This paper deals with the problem that datacenters have so many variables that they want to optimize. Power, efficiency, bandwidth, latency, etc. There is loads of research and work in optimizing the factors. We have talked about many of them throughout the class. But they are all of major importance to these companies. However, when we deploy solutions individually for each of these they don't work well together. They may interfere with each other not allowing them to function properly and could even make these problems worse.

With a multivariate optimization problem, they are looking to do a gradient descent style method that every so often adjusts parameters of the model to optimize the variables. The design is basically a set of managers that work together. Those managers are efficiency controllers, power cappers, group power cappers, and virtual machine controllers.

Also, like the other papers, they use trace-driven evaluation. In this case they used 180 different traces. And evaluated performance on each. They found capabilities of savings to be +60%. They also saw that the effectiveness of each of the components was different at different utilizations and loads. The VCM was useful at lower loads but at higher loads the EC become more useful while the VCM was less.

2: Strengths

- Takes into account more than just cost like the previous paper
- Because there are many parameters of importance to companies running data centers

3: Weaknesses

- Didn't discuss enough about the overhead of managing all this. Is it worth it?

4: Future Directions

- Allow system for companies to choose what they want to optimize and what they are willing to give up in terms of performance/latency and all those metrics.