

## Characterization and Prediction of Host Load in a Google Data Center

Characterization and prediction of system load is essential for optimizing its performance. We characterize and predict host load in a real-world production Google data center, using a detailed trace of over 25 million tasks across over 12,500 hosts. In our characterization, we present valuable statistics and distributions of machine's maximum load, queue state and relative usage levels. Compared to traditional load from Grids and other HPC systems, we find that Google host load exhibits higher variance due to the high frequency of small tasks. Based on this characterization, we develop and evaluate different methods of machine load prediction using techniques such as autoregression, moving averages, probabilistic methods, and noise filters. We find that a linear weighted moving average produces accurate predictions with a 80%-95% success rate, outperforming other methods by 5%-20%. Surprisingly, this method outperforms more sophisticated hybrid prediction methods, which are effective for traditional Grid loads but not data center loads due to its more frequent and severe fluctuations.