Energy Efficiency of MapReduce

Laura Keys, Yanpei Chen, Randy H. Katz

Overview

Motivation

• Energy costs for datacenters increasing rapidly.

• Ongoing power/cooling costs ≈ initial purchase cost over 3-5 year lifetime of equipment.

• MapReduce is a key datacenter workload.

• Need to understand relative energy consumption of various system components.

• Any predictive models would be invaluable for other workloads.

Approach

• **Measure** energy consumption for a variety of MapReduce workloads and configurations.

• **Model** energy consumption of various system components.

• **Apply** findings to predict and reduce energy consumption.

Areas of Investigation

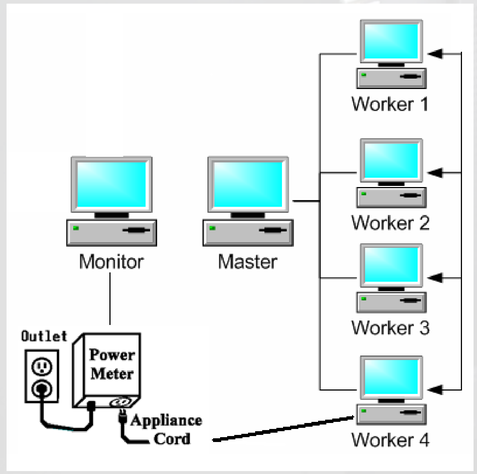
• Effects on energy by varying the following parameters:
- Number of nodes
- Workload type
- Dataset size
- Different hardware configurations

• Tradeoffs between power, energy, time-to-finish

Energy Measurement

Setup

• Power meter on a single machine, out of the plug
• 1W accuracy, measurements every second
• Multiple runs for each configuration
• Collect power for both master and slave nodes



Performance Metrics

• Total energy
• Aggregate power
• Power per machine
• Job duration time

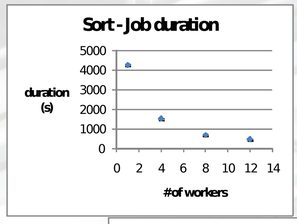
Workloads

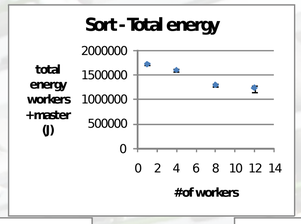
Real-World Workloads
• Sort
• RandomWrite
• Web crawl

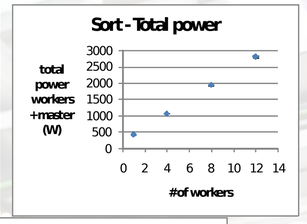
Synthetic Workloads
• HDFS Read
• HDFS Write

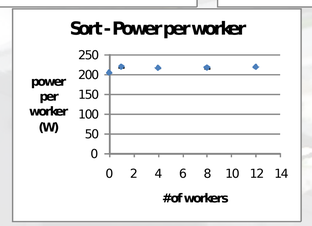
Results

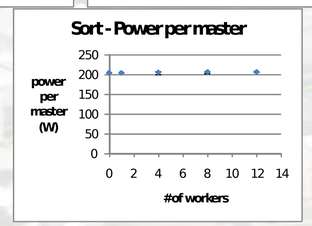
Sort



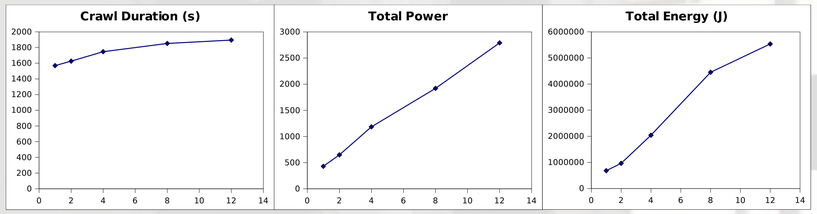






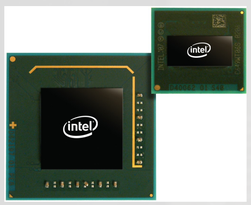


Nutch Web Crawl



Switching Gears...

Porting workloads to 4-node Atom cluster



Power Measurements
- Measured with a Kill-a-Watt and verified with ACme meter
- Each processor runs at ~26 W when idle
- ~27 W with fully utilized CPU

Job Duration Times
- Sort jobs take on average 3-5x longer than on RadLab Opteron Cluster

Overall Effect on Energy
- Sort job run on Atom cluster used around 2/3 the energy of the R Cluster!

Conclusions

- Shorter job duration → less energy
- From **Sort**, more nodes → faster job completion
- From **Crawl**, more nodes → longer job completion when
- From **Atoms**, running at lower power for longer times

results in lower energy usage

When a set of nodes is powered on regardless of its workload, it is best to complete the job as fast as possible.

The best way to ensure that a job will complete quickly is to select a dataset size per worker node that is sufficient to overshadow any setup and communication overhead time.