CHAPTER 1

INTRODUCTION

Low power is one of the main constraints for the design of battery-operated embedded systems. However, this design objective has come into attention for high performance and data centre systems. The main reasons are power constraint of the processor and physical constraint of the chip as the semiconductor industry has reached its physical scaling limits. Continuous increase in the number of transistors on a chip has led to the so-called “dark silicon” phenomena, where the power density does not allow all the transistors to turn on simultaneously. There is a large body of research on harnessing dark silicon or maximizing performance under power constraints.

Cost and environmental reasons are other motivations to govern energy-efficient and low power design. Hardware design companies have considered energy efficiency as one of the main design concerns and have provided mechanisms to ease developing green applications. Intel provides RAPL interface which enables the software developers to measure and control the power consumption at different domain, including core, package, DRAM and embedded graphic. ARM has introduced big. LITTLE technology, which allows migrating applications between simple and complex cores based on work load demands. IBM has employed low power little cores in Blue Gene/Q to increase power efficiency. It is evident by the latest developments; the paradigm shift has been occurring from the performance centric to energy efficient centric design methodologies in the industry.

The energy demand of data centres that support Map Reduce model is increasing rapidly, which is the main obstacle for their scalability. Energy consumption in data centres contributes to major financial burden and prolongs break-even point (when a data canter makes a profit), designing energy-efficient data centres is becoming very important. Current server designs, based on commodity high-performance processors are not an efficient way to deliver green computing in terms of performance/watt. The embedded processors that are designed and developed based on energy efficiency metrics are finding their way in server architectures. Microservers employ embedded low power processors as the main processing unit. The platforms are promising solution to enhance energy-efficiency and reduce cost in data centres.

Several companies and academics have developed cluster architectures based on ARM or Intel Atom cores. FAWN (Fast Array of Wimpy Nodes), which composed of a large number of embedded and efficient Intel Atom cores where each core is low power dissipating only a few watts of power. X-Gene platform developed by Applied Micro is another example of a server-class SoC which is designed for cloud and enterprise servers based on ARM v8 64-bit core architecture. HP low-power Moon-shot servers uses ARM and Atom embedded cores on a single rack. Due to the wide adoption of x86-based architectures in servers.

The world of big data is changing constantly and producing a large amount of data that creates challenges to process them using existing solutions. Big data applications heavily rely on deep machine learning and data mining algorithms, running complex database software stack with significant interaction with I/O and OS. The Apache Hadoop framework, a defacto standard for analytics, assists the processing of large datasets in a distributed computing environment. Numerous big data applications rely on using the Hadoop MapReduce framework to perform their analysis on large-scale datasets. Research works have reported the performance analysis of Hadoop MapReduce applications on high performance servers such as Xeon.

The little core is more energy-efficient than big core in almost all studied applications, and in particular for compute-intensive applications. Low power embedded architectures can provide significant energy efficiency for processing big data analytics applications compared to conventional big high performance core. There have been works on characterizing Hadoop MapReduce applications or optimizing them for performance or power.

The main focus is on performance optimization, ignoring energy-efficiency, or mainly deployed on high performance big Xeon core. The performance and power of Hadoop MapReduce applications is sensitive to various tuning parameters at application (application type, data size per node), system (HDFS block size, number of mappers running simultaneously per micro server node) and architecture levels operating voltage and frequency of core.

`

CHAPTER 2

SYSTEM ARCHITECTURE

Apache Hadoop is an open-source Java-based framework of MapReduce implementation. It assists the processing of large datasets in a distributed computing environment and stores data in highly fault-tolerant distributed file system, HDFS. When an application is submitted for scheduling, Hadoop splits its input data into a fixed data blocks where each block is assigned to a map task. A map task transforms the input data into intermediate key value pairs. These generated intermediate values are transferred from the mappers to the appropriate reducers in the merge stage. Shuffle and sort of key-values are done in this stage.

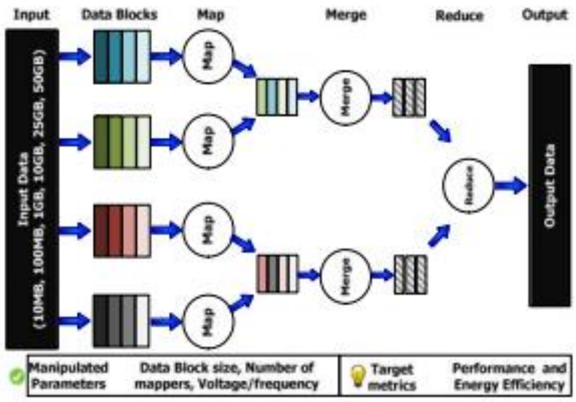


Figure 1: A simple conceptual view of the Hadoop data flow

As different subset of intermediate key-value pairs are assigned to each reducer, the reducers consolidate data into the final output. There are a number of parameters that directly impact the MapReduce application performance and energy-efficiency. The parameters including the number of mappers, operating voltage and frequency of the core, HDFS block size, and the size of data per node that can be tuned by the user, scheduler or the system and are impacting the energy-efficiency. Application Diversity Hadoop cluster hosts a variety of big data applications running concurrently.

Four bench marks have been used. These micro-benchmarks stress-test different aspects of a micro server cluster. It includes two real-world applications namely Naive Bayes -NB and Collaborative Recommendation Filtering-CF) in our study by incorporating mahout library. Hadoop micro-bench-marks and real-world applications for this study along with their particular domain and data type. Interdependent Tuning Parameters.

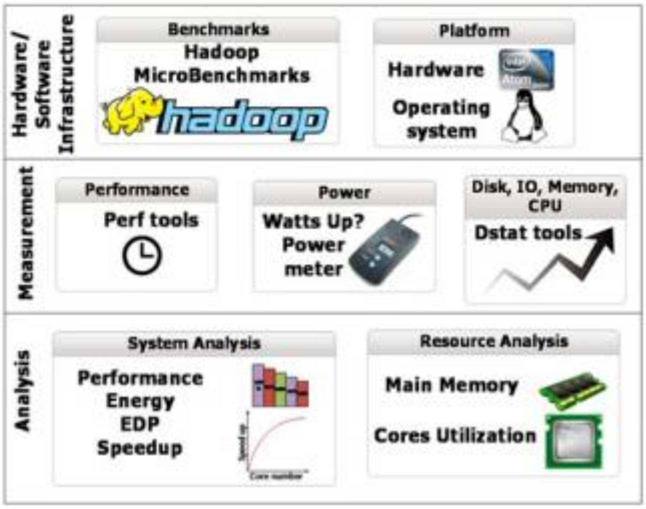


Figure 2: Methodology used.

The impact of the system, application, and architectural level performance and power tuning parameters including the HDFS block size (32, 128, 256, 512, 1024 MB), input data size of the application (10, 100 MB, 1, 10, 25 and 50 GB), number of mappers that run simultaneously on a single node (1, 2, 4 and 8), and frequency settings (1.2, 1.6, 2.0, 2.4 GHz) to evaluate how these parameters affect energy efficiency of big data applications on microserver. Moreover, we thoroughly analyse the impact of these parameters on memory system and processor utilization.

CHAPTER 3

TECHNOLOGY USED

3.1 ARM big. LITTLE

ARM big. LITTLE is a heterogeneous computing architecture developed by ARM Holdings, coupling relatively battery-saving and slower processor cores with relatively more powerful and power-hungry ones. Typically, only one "side" or the other will be active at once, but all the cores have access to the same memory regions, so workloads can be swapped between Big and Little cores on the fly. The intention is to create a multi-core processor that can adjust better to dynamic computing needs and use less power than clock scaling alone. ARM's marketing material promises up to a 75% savings in power usage for some activities. ARM big. LITTLE architectures are used to create a multi-processor system-on-chip (MPSoC).

3.2 RAPL interface

Intel provides RAPL interface which enables the software developers to measure and control the power consumption at different domain, including core, package, DRAM and embedded graphic. RAPL provides a way to set power limits on processor packages and DRAM. This will allow a monitoring and control program to dynamically limit max average power, to match its expected power and cooling budget. In addition, power limits in a rack enable power budgeting across the rack distribution. By dynamically monitoring the feedback of power consumption, power limits can be reassigned based on use and workloads. Because multiple bursts of heavy workloads will eventually cause the ambient temperature to rise, reducing the rate of heat transfer, one uniform power limit can’t be enforced. RAPL provides a way to set short term and longer term averaging windows for power limits. These window sizes and power limits can be adjusted dynamically.

3.3 Low power little cores in Blue Gene/Q

Blue Gene is an IBM project aimed at designing supercomputers that can reach operating speeds in the PFLOPS (peta FLOPS) range, with low power consumption. Trading the speed of processors for lower power consumption. Blue Gene/L used low frequency and low power embedded PowerPC cores with floating point accelerators. While the performance of each chip was relatively low, the system could achieve better power efficiency for applications that could use large numbers of nodes. Dual processors per node with two working modes: co-processor mode where one processor handles computation and the other handles communication; and virtual-node mode, where both processors are available to run user code, but the processors share both the computation and the communication load.

CHAPTER 4

IMLEMENTATION

The methodology in which the experiments are conducted is presented The methodology is divided into three major steps.

- Hardware/Software Infrastructure

Study conducted on Intel Atom C2758 server that has 8 processing cores per node and two levels of cache hierarchy. The operating system is Ubuntu 13.10 with Linux kernel 3.11. The experiments are performed on eight-node Atom server with Hadoop 1.2.1. It is important to note that while network overhead in general is influencing the performance of studied applications and therefore the characterization results, for big data applications, a modern high speed network introduces only a small percent performance overhead. A high speed 1 Gbit/s network to avoid making it a performance bottleneck. There could be certainly more parameters for performance and power tuning, however, attempts to provide an in-depth understanding of how concurrent tuning of these highly accessible and easy tunable parameters at various levels can significantly impact the performance and energy efficiency.

Hardware Type	Parameter	Value
Motherboard	Model	Super micro A1SRM-2758F
CPU (*BW = Bandwidth)	Model	Intel Atom C2758
	# Core	8
	Hyper-Threading	No
	Base Frequency	1.9 GHz
	Turbo Frequency	No
	TDP	20 W
	L1 Cache	24 KB
	L2 Cache	4 * 1024 KB
	Memory Type	DDR3 1600 MHz
	Max. Memory BW*	25.6 GB/s
	Max. Memory Channels	Dual Channel
Disk (HDD)	Model	Seagate ST1000DM003-1CH1
	Capacity	1000 GB
	Speed	7200 RPM
Network Interface Card	Model	ST1000SPEXD4
	Speed	1000 Mbps

Figure 3: Experimental Micro server Platform

- Measurement

  To capture the performance characteristics of the studied applications. Perf is a Linux profiler tool that records hardware performance counters data. Perf exploits Performance Monitoring Unit (PMU) in the processor to measure performance as well as other hardware events at turn time. For measuring power consumption, Watts up PRO power meter measures and records power consumption at one second granularity. The power reading is for the entire system, including core, cache, main memory, hard disks and on-chip communication buses. The average power consumption of the studied applications and subtracted the system idle power to estimate the power dissipation of the core. The same methodology is used in for power and energy analysis. Idle power is measured using Watts up power meter when the server is not running any application and is in the idle state. Dstat is used for main memory, disk and CPU utilization analysis. Dstat is a system-monitoring tool, which collects various statistics of the system.

- Analysis

  The resource utilizations including CPU utilization and memory footprint are saved at run-time in CSV file and then processed by R, an environment for statistical analysis. MapReduce execution breakdown, including setup, map, reduce and clean up phases is obtained through parsing the log files of Hadoop framework. The main analysis of this work includes performance, EDP, MapReduce execution time breakdown, CPU utilization and main memory footprint.

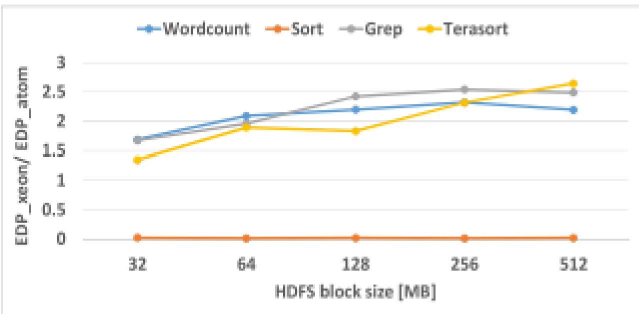


Figure 4: EDP ratio of Hadoop applications on Xeon to Atom at various HDFS block sizes.

Energy efficiency analysis on Xeon versus atom present energy efficiency analysis of the studied applications when changing the frequency on two very distinct micro architectures Intel Xeon- conventional approach to design a high-performance server and Intel Atom- microserver that advocates the use of a low-power core. the EDP results on Atom and Xeon are shown. For each workload, the EDP values are normalized to the EDP result on Atom at the lowest frequency of 1.2 GHz and with 512 MB HDFS block size.

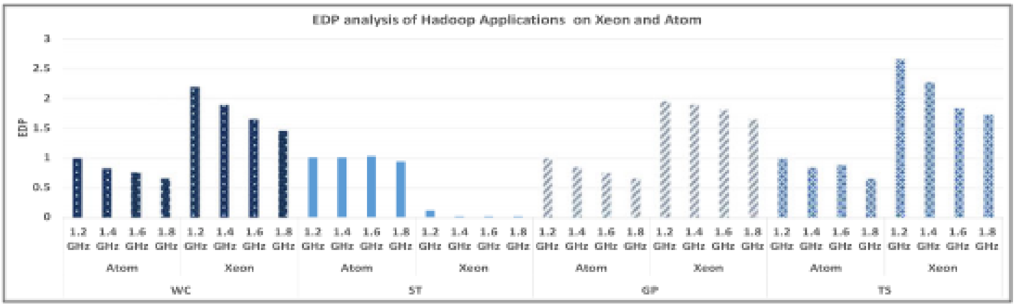
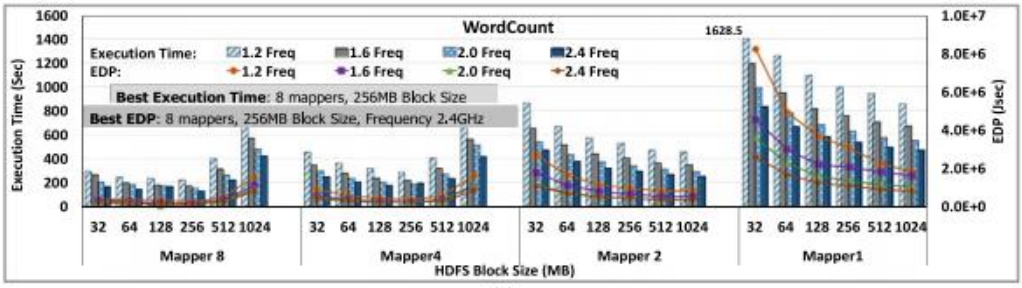
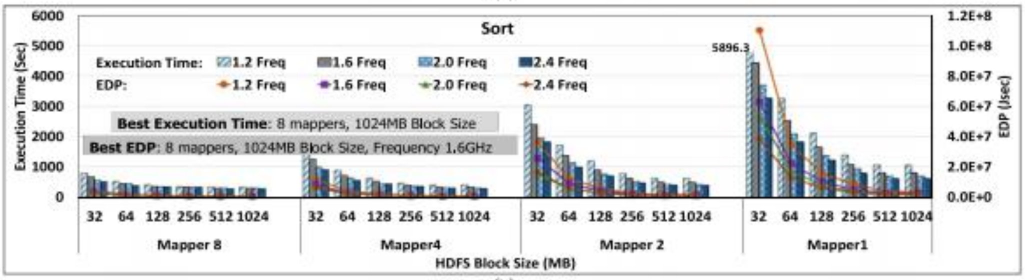


Figure 5: EDP analysis of Hadoop applications on Xeon and Atom with frequency scaling.

The low power characteristics of the Atom results in a lower EDP on Atom compared to Xeon for most applications with the exception of the Sort. This is due to the fact that the performance gap (in terms of execution time) for the I/O bound benchmarks is very large between Atom and Xeon. Since EDP is the function of the execution time and power, the total EDP on Xeon is lower for the Sort benchmark. In addition, the results show that increase in the frequency reduces the total EDP. While increasing the frequency increases the power consumption, it reduces the execution time of the application and consequently the total EDP. In addition, a sensitivity analysis of EDP ratio of the applications on Xeon to Atom.

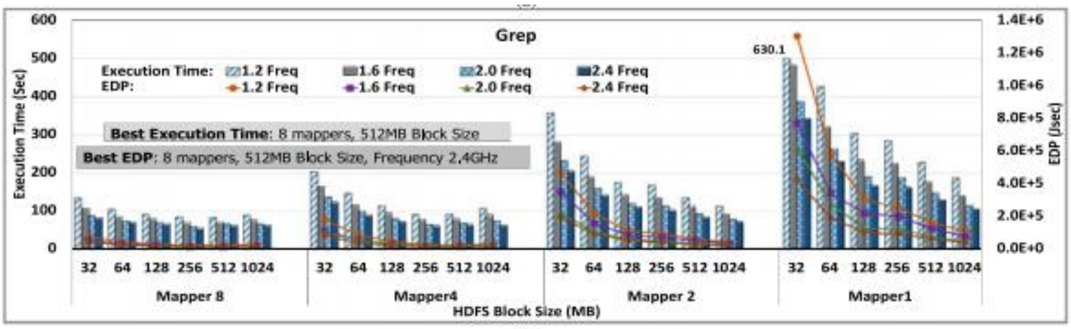


(a)

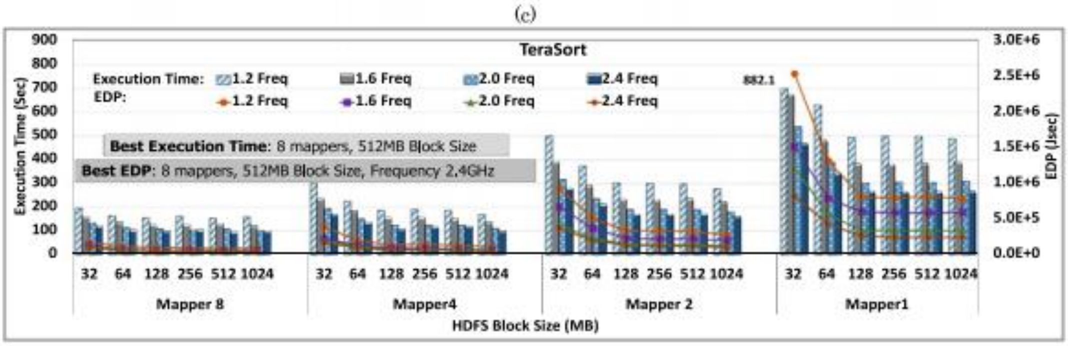


(b)

Figure.6 (a) Execution Time and EDP of Word Count with various mappers, HDFS block size and operating frequencies. (b) Execution Time and EDP of Sort with various mappers, HDFS block size and operating frequencies.



(c)



(d)

Figure7: (c) Execution Time and EDP of Grep with various mappers, HDFS block size and operating frequencies. (d) Execution Time and EDP of Terasort with various mappers, HDFS block size and operating frequencies.

The EDP change with respect to the HDFS block size for a frequency of 1.8 GHz. The results show that increasing HDFS block size increase the EDP gap between Atom and Xeon. Since in Atom, the performance bottleneck exists in the memory subsystem, improving memory subsystem performance by increasing HDFS block size enhances its performance more significantly compared to Xeon, and reduces the performance gap between the two architectures. Overall, Atom has shown to be significantly more sensitive to tuning parameters. Therefore, the performance gap between the two architectures can be reduced significantly through fine-tuning of the system and architectural parameters on Atom, allowing maximum energy efficiency.

4.1 Results

- Execution Time Analysis shows the execution time of the studied Hadoop applications with respect to the number of mapper slots (cores), HDFS block size and operating frequency with the fixed input data size of 10 GB per node for Hadoop micro-benchmarks and real-world applications, respectively for instance, 10 GB input data size per node presents 80 GB input data size processed by application in an 8-node cluster. Hadoop exploits cluster-level infrastructure with many nodes for processing big data applications, ho the experimental data should be collected at the node level to understand how various optimizations and scheduling decisions affects the performance, architectural parameters and energy-efficiency at the node level.
- Energy-Efficiency Analysis EDP is a fair metric to compare various architectures, or even the impact of changing optimization knobs in an architecture. EDP (or PxDxD) represents a trade-off between power and performance. Without EDP and just using energy metric for comparison, simply reduce the voltage and frequency in an architecture, and reduce its energy, at a cost of lowering the performance (increased execution time). Performance along with energy is important to find out the impact of optimization parameters. In order to characterize the energy efficiency, evaluate Energy Delay Product (EDP) metric to investigate trade-off between power and performance when tuning Hadoop and processor parameters

- MapReduce Phase Breakdown Analysis is essential to profiling and characterizing the application behavior. The performance of various phases of MapReduce application to analyze the frequency impact on various phases of MapReduce application, while tuning parameters at the application, system, and architecture levels. Our results show that reduce phase of Grep and Map phase of the sort application are less sensitive to the frequency as these phases are I/O intensive in nature.
- System Resources Profiling and Utilization the real time system resources profiling (CPU utilization and memory footprint) to understand the runtime behavior and resource utilizations of Hadoop micro-benchmarks. Real-world applications have not been included in the system resource utilization.
- Input Data Size Sensitivity Analysis Hadoop exploits cluster-level infrastructure with many nodes for processing big data applications, to understand the impact of various parameters and how their interplay impacts EDP, single node characteristics analysis is required. The number of mappers is fixed at 8 with the default HDFS block size (64 MB) and governor is set as on-demand. The results show that the execution time is proportional to the input data size. Power consumption also increases slightly as the size of input data increases

CHAPTER 5

APPLICATIONS

The speedup obtained when increasing the number of available cores on a microserver node outweighs the power overhead associated with increasing the number of cores, making a configuration that uses the maximum number of available cores per node the most energy-efficient across all studied applications. Unlike microservers, for traditional high performance server the power consumption increase, as the number of mappers' increases, outweighs the performance gains. Microservers introduces a new trade-off to process the Big data applications for maximum energy-efficiency.

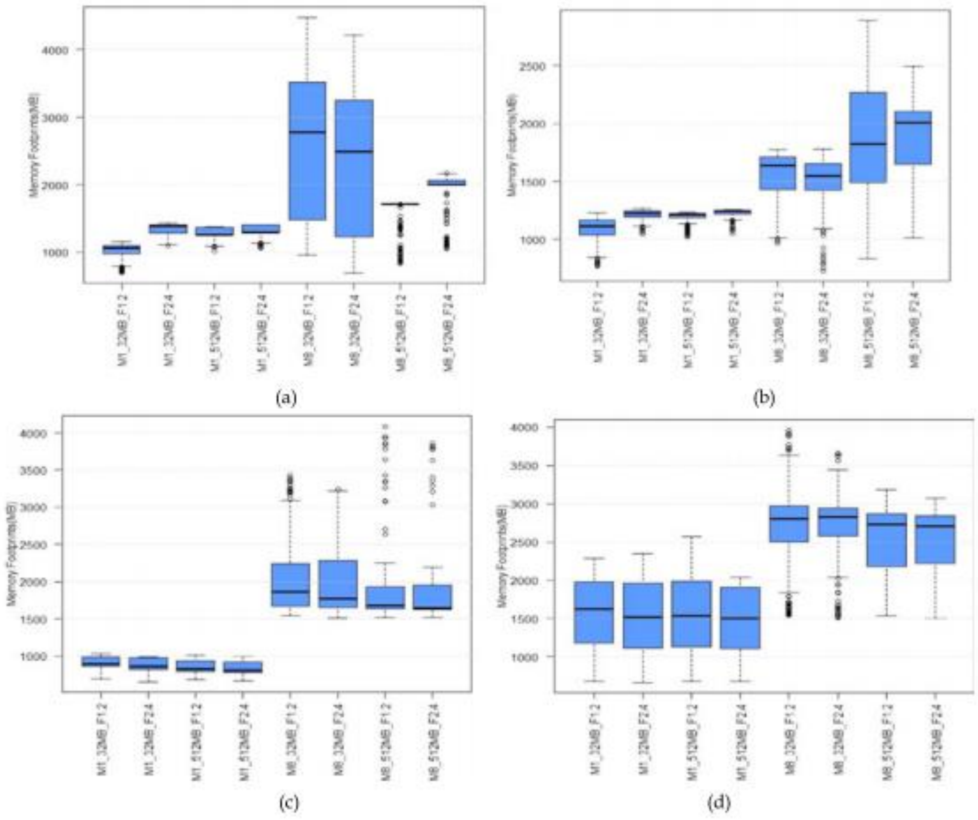


Figure 8: (a) Memory Footprints (MB) of Word Count. (b) Memory Footprints (MB) of Sort. (c) Memory Footprints (MB) of Grep. (d) Memory Footprints (MB) of TeraSort.

- Increasing the number of mappers/cores, improves performance and reduces the CPU utilization. In all studied cases using maximum number of cores produces best results in terms of both performance and energy-efficiency. It was also observed that if the number of mappers exceeds available cores, mapper tasks are buffering which potentially reduces the performance and impact the energy-efficiency.
- Although utilizing all available cores on each microserver node provides maximum energy-efficiency across all studied applications, concurrent fine-tuning of frequency and HDFS block size reduces the reliance on the maximum number of cores, and instead make a configuration with fewer number of cores to be energy-efficient competitive with the maximum number of cores. This helps freeing up cores on each node to accommodate scheduling incoming applications in a cloud-computing environment.
- Tuning the block size significantly affects the performance and energy-efficiency of the system. I/O bound Hadoop applications provide the optimal execution time and EDP with the largest HDFS block size. Default HDFS block size of 64 MB is not optimal, neither for power nor for the performance.
- The speed up improvement is more when the HDFS block size is larger. I/O bound applications can run at a lower frequency to save power. Performance loss can be compensated to a significant extend by increasing the number of mappers.
- Increasing the number of mappers and the number of active cores result in drastic reduction in average core utilization. In other words, with more number of mappers most of the times the cores are becoming idle and dissipate leakage power. This motivates employing Dynamic Power Management (DPM) techniques for big data applications when running large number of mappers.
- Default Hadoop configuration parameters are not optimal for maximizing the performance and energy efficiency. With fine tuning the Hadoop parameters along with the system configurations, a significant gain in performance and energy-efficiency can be achieved.

CHAPTER 6

FUTURE SCOPE

Recently, there have been a number of efforts to understand the behaviour of big data and cloud scale applications by benchmarking and characterizing them, to find out whether state-of-the-art high performance server platforms are suited to process them efficiently. The most prominent big data benchmarks, includes Cloud Suite, HI Bench, BigDataBench, Link Bench and Cloud Rank-D which mainly focus on the applications characterization on high performance servers Cloud Suite benchmark was developed for Scale out cloud workloads. HI Bench is a benchmark suite for Hadoop MapReduce. The BigDataBench was released recently and includes online service and offline analytics for web service applications.

Link Bench is a real-world database benchmark for social network application. Cloudscape use a systematic approach to benchmark various components of the cloud to compare cloud providers. These works analyse the application characterization of big data applications on the Hadoop platform, but they do not discuss the Hadoop configuration parameters for energy efficiency. Recent works have investigated the energy efficiency in the Hadoop system The focus of these works is on the reduction of operating cost of data centres for energy efficiency.

The study is different as it focuses on tuning Hadoop parameters to improve the performance and energy efficiency. The impact of Hadoop configuration parameters has not studied the impact of frequency scaling and its interplay on Hadoop specific parameters such as HDFS block size and the number of mappers for optimizing the energy efficiency. Hadoop has focused on the resource utilization for performance and energy efficiency on Amdahl blades running Hadoop.

The two applications with default Hadoop configuration parameters, study illustrates that default Hadoop configuration parameters (like HDFS block size of 64 MB) are not optimal for maximizing performance and energy efficiency. It is analysed that the performance and throughput with the scale-up and scale-out cluster environment to figure out which cluster configuration is suitable for Hadoop Map reduce jobs. Additionally, it has been presented that the optimization applied to Hadoop like concurrency, network, memory and reduce-phase optimization on the high performance server Xeon.

The energy usage and total cost of ownership for MapReduce applications has been analysed on the Xeon and ARM big. LITTLE architecture in. The impact of the Hadoop configuration parameters for performance and energy efficiency. HDFS block size is one of the key design parameters and vital to the performance and power optimization. Discuss the interplay of system, architectural and applications parameters nor study the resource profiling that is essential to understand the runtime behaviour and resource utilization of the Hadoop applications. The closest to our work as they conduct a study of microserver performance for Hadoop applications.

The main focus is on the assessment of five different hardware configuration clusters for performance, energy dissipation and cost Hadoop configuration parameters such as number of mappers, HDFS block size and data input size as well as a system parameter (frequency scaling) for the performance and energy efficiency on microserver. Hadoop configuration parameters that directly affect the MapReduce job performance, power and energy efficiency on emerging x86 based low power cores microservers and help to understand the interplay of the Hadoop system, architecture and application parameters to achieve the maximum performance and energy efficiency improvement.

CONCLUSION

A comprehensive analysis of the impact of Hadoop system configuration parameters, as well as application and architecture level parameters, and the interplay among them on performance and energy efficiency of various real-world big data applications running on Atom microserver, a recent trend in server design which advocates the use of low-power small cores to address the power and energy-efficiency challenges. The performance and energy efficiency of big data applications are highly sensitive to various Hadoop configuration parameters, as well as system and architecture level parameters, demonstrating that the baseline Hadoop as well as system configurations are not necessarily optimized for a given benchmark and data input size. Through performance and power measurements and analysis on Atom microserver shows that increasing the number of mappers that run simultaneously along with increasing the number of active cores help to maximize energy efficiency, shows that the overall energy efficiency is highly decided by the HDFS block size and is different for each benchmark, demonstrating that the default configuration parameters are not optimal. , explored the impact of scaling the operating frequency of the compute node for the performance and energy efficiency, show that big data applications become less sensitive to frequency with large number of mappers.

BIBLIOGRAPHY

[1] A. Venkat, et al., “Harnessing ISA diversity: Design of a heterogeneous-ISA chip multiprocessor,” in Proc. 41st Annu. Int. Symp. Comput. Archit., 2018, pp. 121–132.

[2] B. Raghunathan, et al., “Cherry-picking: Exploiting process variations in dark-silicon homogeneous chip multi-processors,” in Proc. Des. Autom. Test Europe Conf. Exhib., 2019, pp. 39–44.

[3] M. Malik, et al., “Big data on low power cores are low power embedded processors a good fit for the big data workloads?” in Proc. 33rd IEEE Int. Conf. Comput. Des., 2017, pp. 379–382.

[4] M. B. Taylor, “Is dark silicon useful? Harnessing the four horsemen of the coming dark silicon apocalypse,” in Proc. 49th ACM/ EDAC/IEEE Des. Autom. Conf., 2015, pp. 1131–1136.

[5] T. S. Muthukaruppan, et al., “Hierarchical power management for asymmetric multi-core in dark silicon era,” in Proc. 50th Annu. Des. Autom. Conf., 2016, Art. no. 174.

[6] Intel, Intel 64 and IA-32 Architecture Software Development Manual, Aug. 2019.

[7] P. Boyle, “The bluegene/q supercomputer,” PoS LATTICE2012 20, 2015.

[8] R. T. Kaushik and B. Milind, “Greenhdfs: Towards an energy-conserving, storage-efficient, hybrid Hadoop compute cluster,” in Proc. USENIX Annu. Techn. Conf., vol. 109, p. 34, 2017.