Cs621 current paper 2024.solved by Izhar from swat mng01 campus.

MapReduce has three main phases:

1. Map
2. Sort
3. Reduce

MapReduce Example (based on Three Phases)

The canonical MapReduce Example: Word Count • Example corpus: Jane likes toast with jam Joe likes toast Joe burnt the toas

Decomposition Techniques:

i. Recursive Decomposition
ii. Data Decomposition
iii. Exploratory Decomposition
iv. Speculative Decomposition

Static topology Connection Networks:

I. • RingHypercube (Loop) networks
II. • Mesh
III. • Torus
IV. • Tree networks
V. • Hypercubenetwork.

Introduction to GPU:

“Graphics Processing Unit ()GPU) is a chip or electrnic circuict capable of rendering graphics for

Display on an electronic device. GPUs work by using a method called parallel processing, where

Multiple processors handle seprate parts separate parts of the same task.”

The world's first GPU, the GeForce 256, was marketed by NVIDIA in 1999. These GPU chips
can process a minimum of 10 million polygons per second and are used in nearly every computer
on the market today.

Client-Centric Consistency Models Types:Names

i. Eventual Consistency
ii. Monotonic Writes
iii. Read Your Writes
iv. Writes Follow Reads.

Data-centric Consistency Models Types Names:

I. Continuous Consistency
II. Consistent Ordering of Operations
III. Causal Consistency
IV. Grouping Operations

Types of mutex

1. A normal mutex
2. A recursive mutex
3. Error check mutex.

Architecture of GPU

Parallel coprocessor to conventional CPUs
Implement a SIMD structure, multiple threads running the same code.

- Thread local rigister
- Block local memory and control
- Global memory

Gris of Block of Threat

Global memory The CPU is the conventional multi-core processor with limited parallelism to exploit. The GPU has a many-core architecture that has hundreds of simple processing cores organized as multiprocessors. Each core can have one or more threads. Essentially, the CPU's floating-point kernel computation role is largely offloaded to the manycore GPU. The CPU instructs the GPU to perform massive data processing.

- Scale code to hundreds of cores running thousands
- The task runs on the GPU independendly from the CPU.

MPI Basics:

Many parallel programs can be written using just these six functions, only two of which are nontrivial;

–MPI_INIT

MPI_FINALIZE

–MPI_COMM_SIZE

–MPI_COMM_RANK

–MPI_SEND

–MPI_RECV

Advantages/Benefits of the Message Passing Interface(MPI)

- **Portability** (There is no need to modify your source code when you port your application to a different platform that supports (and is compliant with) the MPI standard)
- **Functionality:(** Over 115 routines are defined in MPI-1 Al**)**
- **Scalability:** (MPI is designed to scale to large numbers of processors, which makes it well-suited for high performance computing)
- **Availability:(** A variety of implementations are available, both vendor and public domain.)

Parallel I/O Tools

Collections of system software and libraries have grownup to address I/O issues:

• At Parallel file systems • MPI-IO

• High level libraries

• Relationships between these are not always clear. Choosing between tools can be diffcult.

Replication and scaling Technique:

Replication and caching for performance are widely applied as scaling technique.Replicating the data and moving it closer to where it is needed helps to solve the scalability problem.

When systems scale: The first Problems to surface are those associated with performance as the systems get bigger, they get often slower. Another problem is how to efficiently synchronize all of the replicas created to solve the scalability issue.

Message Passing Programming Modes:

Different communication modes that can be used for message passing programming are:

- Synchronous/asynchronous
- Blocking/non-blocking
- Buffered/unbuffered

Types of process-to-resource binding

• Binding by identifier (e.g., URL, (IPaddr:Port))

• Binding by value (e.g., standard libraries)

• Binding by type (e.g., monitor, printer)

Performance of GPU:

Bill Dally of Stanford University considers power and massive parallelism as the major benefits of GPUs over CPUs for the future.

Two Aspects:

- ➢ Data Access Rate Capability
- ✓ Bandwidth
- ➢ Data Processing Capability
- ✓ How many ops per sec.

How big is the MPI library?

- ✓ Huge (125 Functions).
- ✓ Basic (6 Functions).

Where to get MPI library?

Standard message-passing library includes best of several previous libraries.

• MPICH (WINDOWS / UNICES)

• http://www-unix.mcs.anl.gov/mpi/mpich/

• Open MPI (UNICES) • http://www.open-m.

Buffered:

A routine that uses buffered message passing is a program that sends and receives messages using a buffer. In this mode, messages are queued in the buffer until they are ready to be sent or Recived.

Un-buffered:

A routine may be used to perform a specific message passing task, such as sending or receiving a message between two processes. Does not use a buffer. Messages are sent and received immediately without any buffering. This mode is useful when low latency is requird.

Synchronous Message Passing:

A synchronous communication is not complete until the message has been received.

• Completes once ack is received by sender.

Communication upon synchronization. Hoare's Communicating Sequential Processes (1978). BLOCKING send and receive operations.

• Unbuffered communication.

• Several steps in protocol- synchronization, data movement, completion.

• Delays participating processes.

Asynchronous Communication:

An asynchronous communication completes before the message is received. It has three modes: • Standard send: completes once the message has been sent, which may or may not imply that the message has arrived at its destination. • Buffered send: completes immediately, if receiver not ready, MPI buffers the message locally. • Ready send: completes immediately, if the receiver is ready for the message it will get it, otherwise the message is dropped silently.

Performance Metrics: Efficiency:

Efficiency is a measure of the fraction of time for which a processing element is usefully employed Mathematically, it is given by E= S/T Following the bounds on speedup, efficiency can be as low as 0 and as high as 1.

Performance Metrics: Super-linear Speedups:

- The speedup of adding numbers on proccersssor is given by

$$S = \frac{n}{logn}$$

- Efficiency is given by

$$E = 0\left(\frac{n}{logn}\right)$$

$$= 0\left(\frac{1n}{logn}\right)$$

Independent MPI-IO

For independent I/O each MPI task is handling the I/O independently using non collective calls like MPI_File_write() and MPI_File_read()

Fined-grained

Decomposition of computation into a large number of tasks results in fine-grained decomposition.

Coarse Grained

Decomposition of computation into a small number of tasks results in a coarse grained decomposition.

File I/O: Lustre File System

Metadata Server (MDS) makes metadata stored in the MDT(Metadata Target) available to Lustre clients.

Object Storage Server(OSS) provides file service, and network request handling for one or more local OSTs.

Object Storage Target (OST) stores file data (chunks of files).

Latency:

Total time to send a message.

Bandwidth:

Number of bits transmitted in a unit of time.

Bisection:

The number of connections that need to be cut to partition the network in 2.

I/O bottleneck:

There are three main reason for I/O bottleneck: • Increasing CPU Speed as compared to I/O • Increase in number of CPUs • New application domains that increasing I/O demand.

MapReduce:

Simple programming model that enables parallel execution of data processing programs •Executes the work on the data near the data •In a nutshell: HDFS places the data on the cluster and MapReduce does the processing work.