Introduction to MPP Fundamentals

Zhenghua Lyu

May 31th, 2018

Outline

Outline

What's in a GPDB Running Instance

- Many Postgres Instances(Only one Master, many Segments)
- QD & QE
- Share Nothing(local storage, computing is coupled with storage)
- Dispatch & InterConnect

Outline

SPMD

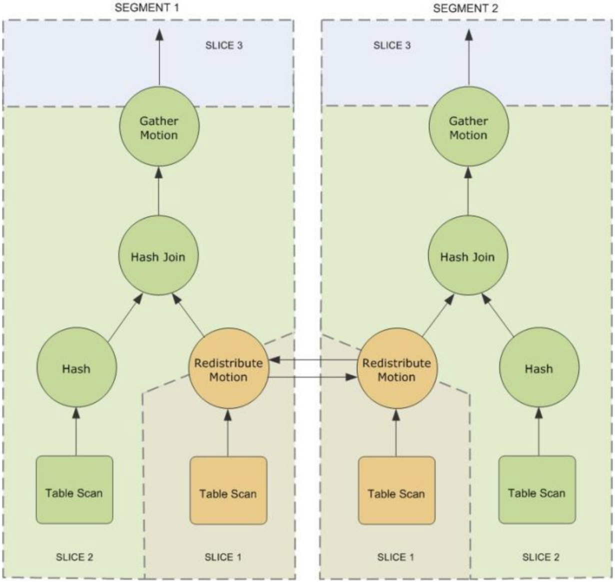
- SPMD: single program, multiple data
- A example of MPI programming
- Take plan as a program, and greenplum as a VM
- MPI communication APIs VS MPP Motion

Outline

Slice

- Motion Node breaks plan into different parts
- a single part is called slice
- a slice handles part of data processing of the whole query
- Image slice as function abstraction for a distributed language
- A motion node is contained in both two slices it breaks

Slice Example



Gang

- A group of one or more qExec processes working on the same task, and given identical instructions (or nearly so), is called a gang.
- A slice is executed by a gang
- Take a look at different gang types and examples

Motion

- A special kind of Node in Plan of Greenplum DB
- Motion Type: MOTIONTYPE_HASH, MOTIONTYPE_FIXED, MOTIONTYPE_EXPLICIT
- Redistribute Motion, Broadcast Motion, Gather Motion, Explicit Redistribute Motion

Outline

DISTRIBUTION_POLICY

- A concept of relation(static)
- GpPolicyType
- The catalog gp_distribution_policy
- Examples

LOCUS

- A concept of distribution of tuples across processes
- Examples & Understand each type's meaning
- Important code to read: cdbpath_motion_for_join

Flow

- A concept of plan
- Destribe how the tupe flows

Dispatch

- Dispatch plan & query
- Dispatch command for Transaction Management