Gmail

Rudresh G P <rudreshgp18@gmail.com>

Note for DBT discussion

1 message

kim chen <mymycen@gmail.com> Tue, Dec 6, 2016 at 10:43 PM
To: Seema Bharadwaj <seemabharadwaj29@gmail.com>, Rudresh G P <Rudreshgp18@gmail.com>, kumar awanish <awanish00@gmail.com>, 李静 <jingli.818@hotmail.com>, Sonali Nayak <sonalirosy91@gmail.com>

	Logical	physical
	Easy for caching and indexing	Difficult to work with caching and indexing
	The size of string used to represent address is small	The size is big as it contains device id, disk id, cylinder id, sector and block id
	Flexibility to move data in the memory and fix the pointers	

Hash join

(Need to have equal condition; eg: R.r=2, then we could do hash join)

If does not fit with the memory, then we have another two options-Grace hash, Hybrid Hash- Need to have enough memory to store the data.

Grace hash-Don't require memory bcz nothing to do with disk; divide partition.

Hash join--Memory for small table.

Non-equal =>Merge Join

Does not fit with the memory=> Nest Loop join

NLJ: Outer source is small and inner is indexed

Big data=> Scan

Sequential Scan- When table size is big, table/sequencial scan is better as only 1 I/O for 1 disk reading

Index Scan- Indexing scan is more costly as it takes one I/O for each record and is not feasible for large set of records

B(R) + T (R) * B(S) (Clusted)

Bitmap –Limited attribute; based on the

One dimension have many same value,

--
Kim (陳小均)