Optimizing Hash Partitioning for Solid State Drives

in heol Shin
Dept of omp ter S ien e
onsei niversit
onsei ro Seodaem n g
Seo l orea
smanioso onsei a fallsmal s

Hong han oh
Dept of omp ter S ien e
onsei niversit
onsei ro Seodaem n g
Seo l orea
onsei a

onmoo ng
le troni s
ang ae daero gil
Seogho g Seo l ep i li
of orea
onmoo ng lge om

Sangh n Par *
Dept of omp ter S ien e
onsei niversit
onsei ro Seodaem n g
Seo l orea
sangh n s onsei a r

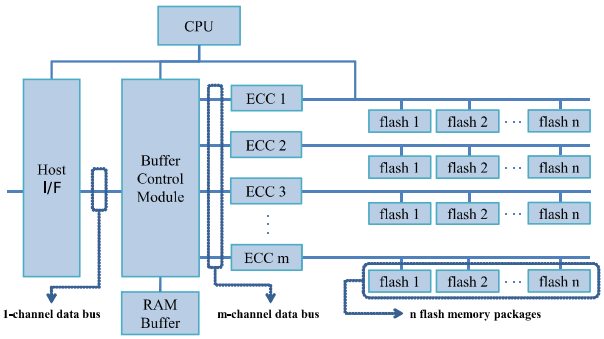
S

The use of flashSSDs has increased rapidly in a wide range of areas due to their superior energy efficiency, shorter access time and higher and idth hen ompared to HDDs he internal parallelism created by multiple flash memory packages embedded in a flashSSDs, is one of the unique features of flashSSDs. Many new DBMS technologies have been developed for flashSSDs, but query processing for flashSSDs have dra n less attention than other D S te hnologies Hash partitioning is pop larl sed in er pro essing al gorithms to materialize their intermediate results in an effi ient manner n this paper e propose a novel hash par titioning algorithm that e ploits the internal parallelism of flashSSDs. The devised hash partitioning method outperforms the traditional hash partitioning te hni e regardless of the amo nt of availa le main memor independentl from the buffer management strategies (blocked I/O vs page sized O e implemented o r method ased on the so r e ode of the PostgreS storage manager PostgreS relation files created by the TPC-H workload were employed in the e periments O r method as fo nd to e p to times faster than the traditional method ith lo ed O and times faster than the traditional method ith page sized O

S on epts

• nformation s stems → er operators

* orresponding a thor el mail address sangh n s onsei a r



ig re lashSSD ar hite t re

e ords

er e tion Hash oin Hash Partitioning lash Stor age Devices; Internal Parallelism of flashSSDs

OD O

In the last decade, flashSSDs have been rapidly adopted in a ide range of areas e a se of their s perior feat res hen ompared to HDDs Personal sers have emplo ed flashSSDs in their laptop and desktop computers in order to oost operating s stem and speed p a ess to fre entl used applications. Data centers have also adopted flashSSDs to improve energy efficiency and performance. Many DBMS technologies have b en developed for flashSSDs. While majority of previous studies have focused on the buffer and index management of DBMSs on flashSSDs, query processing algorithms for flashSSDs have drawn less attention from resear hers

he nternal Parallelism of lashSSDs

A typical flashSSD architecture is shown in Figure 1. A flashSSD includes a CPU, a host interface (host I/F), a RAM buffer, Error-Correcting Code(ECC) modules, multiple data transfer channels and multiple flash memory pack-