Maximize Database Performance with Flash

Mughees A. Minhas
VP, Product Management
Oracle Corp.

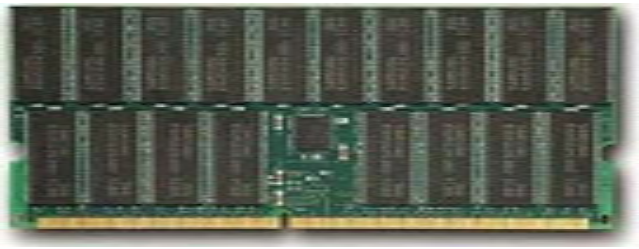
The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Agenda

- Flash and Database
- Oracle's Innovations in Flash
- Q & A

The Promise of Flash

- Replace expensive 15K RPM disks with fewer Solid State (Flash) Devices
  - Reduce failures & replacement costs
  - Reduce cost of storage subsystem
  - Reduce energy costs
- Lower transaction & replication latencies by eliminating seek and rotational delays
- Replace power hungry, hard-to-scale DRAM with denser, cheaper devices
  - Reduce cost of memory subsystem
  - Reduce energy costs



Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Applications of Flash in Databases

- As an additional storage layer (caching)
  - Stage active database objects in flash
  - Accelerate reads and writes to these objects
- For Data/Log Files
  - Improves user transaction response time
  - Increases overall throughput for IO intensive workloads

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

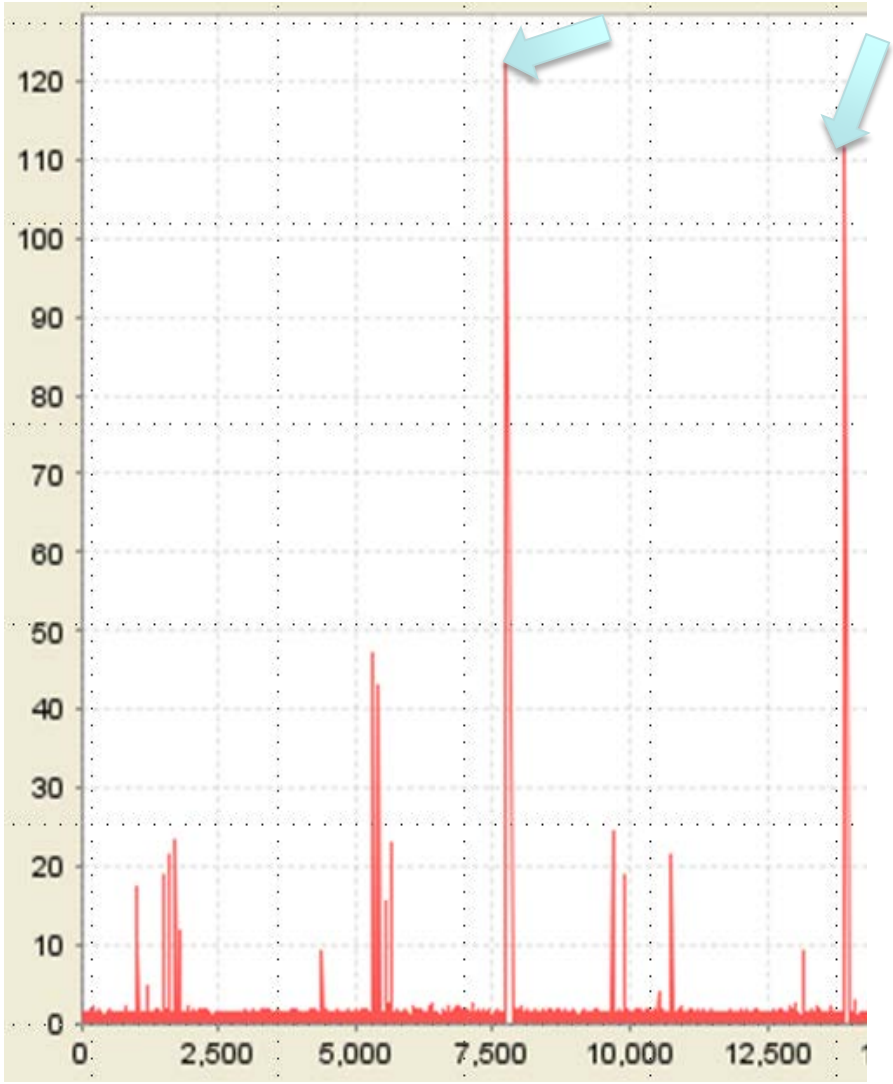
To Cache or Not to Cache

- Random reads against tables and indexes
  - Cached: more likely to have subsequent reads
- Sequential read tables, or Scans
  - Not Cached: sequentially accessed data is unlikely to be followed by reads of the same data
- Backups, mirrored copies of the block
  - Not Cached: Why?

But most general purpose flash solutions are database agnostic and cache all the above workloads

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

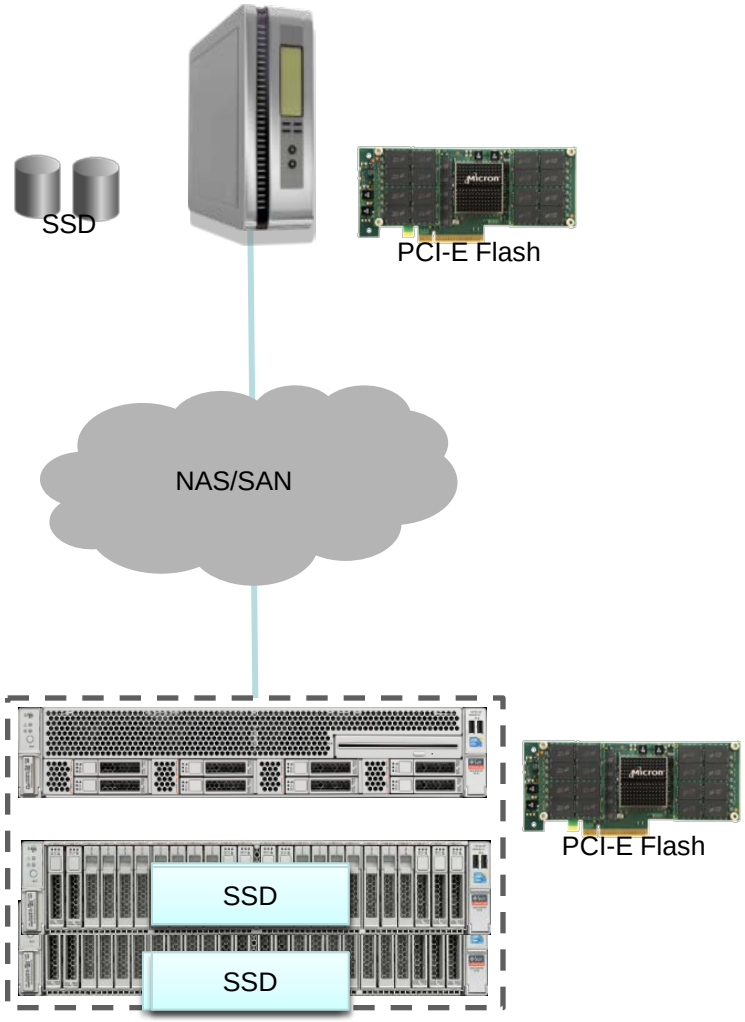
Flash And Database Logs



- Flash has very good *average* write latency
- Greatly improves user transaction response time
- Flash occasional outliers, one or two orders of magnitude slower
  - Garbage collection, etc., contribute to that delay
- OLTP workloads dislike such large variations

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Where To Introduce A Flash Device



- Direct Attached
  - Mount flash inside your server
  - PCI-E or SSD
- Networked Storage
  - Share device on the network (FC or 10GE)
  - Popular implementations:

Tiered Storage: Multiple tiers of disk drives (SSD, FC, SAS); various performance characteristics and data moves between these tiers

Hybrid Storage: Combination of Direct Attached flash on the storage controller and HDD in expansion shelves

All Flash Arrays: All storage is some form of flash device – either SSD or custom Flash modules

Putting It All Together

- Flash devices in application tier (server attached flash) lacks enterprise class scalability and high availability
- Flash devices in traditional storage arrays are not efficient as storage controllers don't respond quickly enough to workload changes and are IO bound
- All Flash Arrays lack the features and stability of traditional arrays

Agenda

- Flash and Database
- Oracle’s Innovations in Flash
- Q & A

Oracle's Innovations in Flash

- Exadata Smart Flash Cache
- Exadata Smart Flash Log
- Exadata Smart Flash Cache Compression
- Exadata Smart Flash Cache Scan Awareness

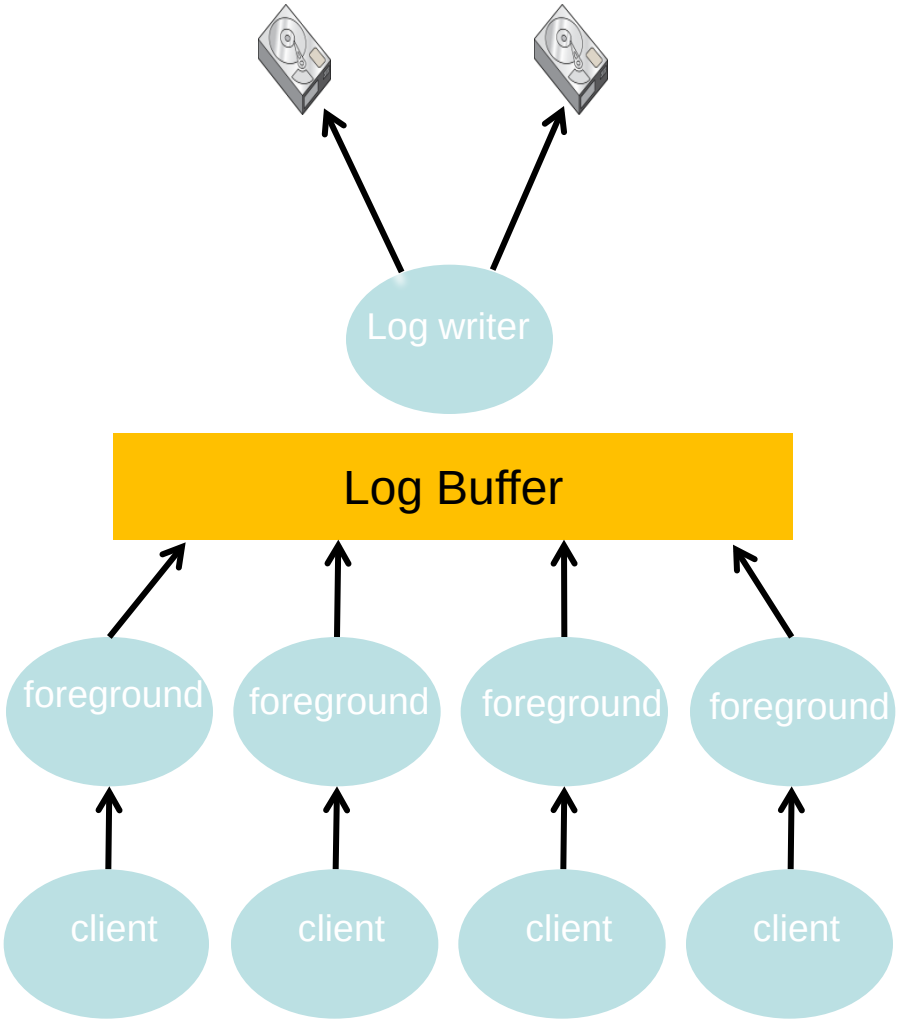
Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Exadata Smart Flash Cache

- Understands different types of I/Os from database
  - Skips caching I/Os to backups, data pump I/O, archive logs, tablespace formatting
  - Caches Control File Reads and Writes, file headers, data and index blocks
- Write-back flash cache
  - Caches writes from the database not just reads
- Cluster-aware (RAC)

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Exadata Smart Flash Log

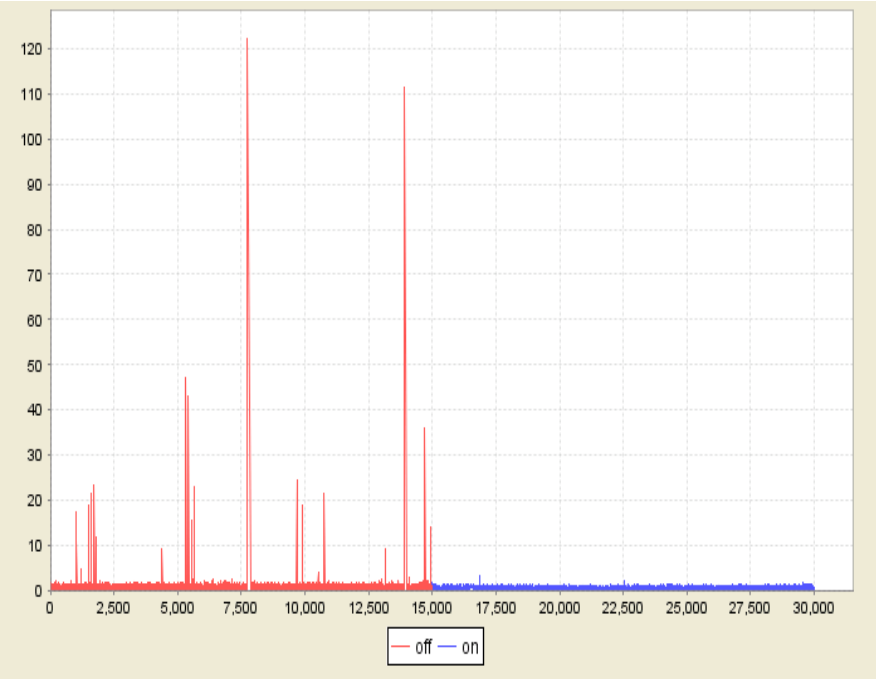


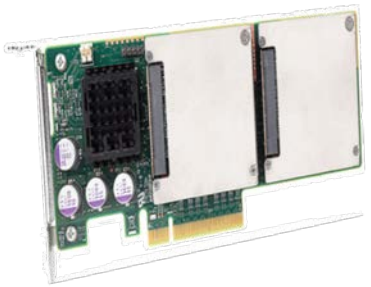
- Outliers in log IO slow down lots of clients
- Outliers from any one copy of mirror affect response time
- Performance critical algorithms like space management and index splits are sensitive to log write latency
- Legacy storage IO cannot differentiate redo log IO from others

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Exadata Smart Flash Log

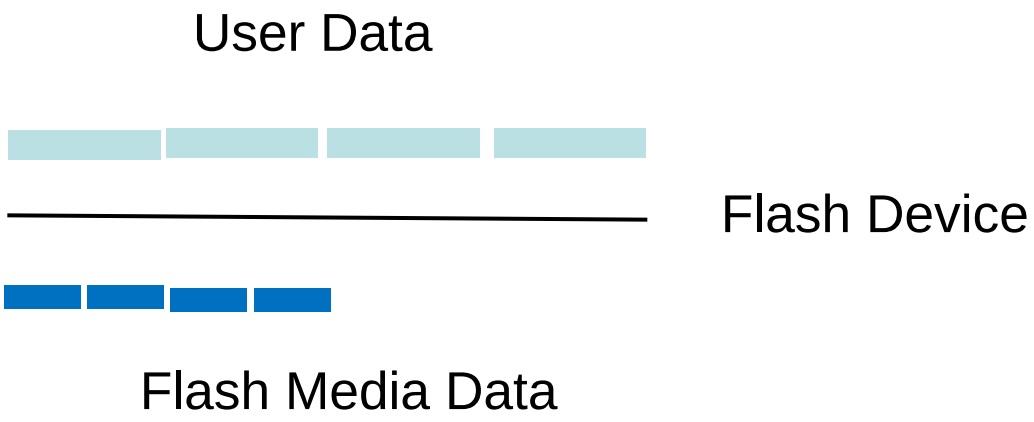
- Smart Flash Log uses flash as a parallel write cache to disk controller cache
- Whichever write completes first wins (disk or flash)
- Reduces response time and outliers
  - "log file parallel write" histogram improves
  - Greatly improves "log file sync"
- Uses almost no flash capacity (< 0.1%)
- Completely automatic and transparent





Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

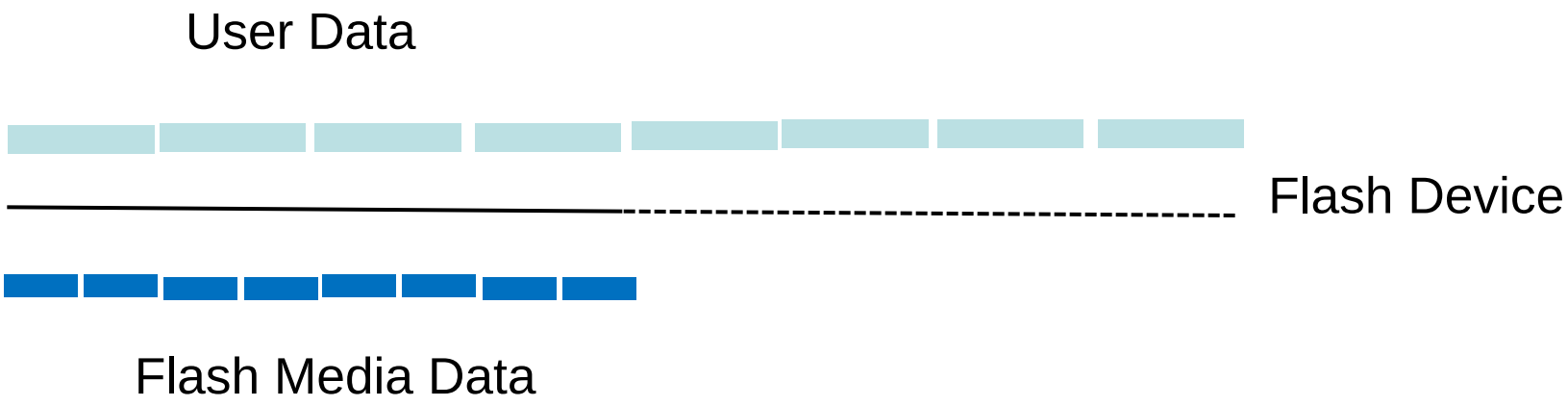
Exadata Smart Flash Cache Compression



- As user adds more data, data is compressed and written
- Flash device has no logical space at the end for user data but has lot of physical space

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Exadata Smart Flash Cache Compression

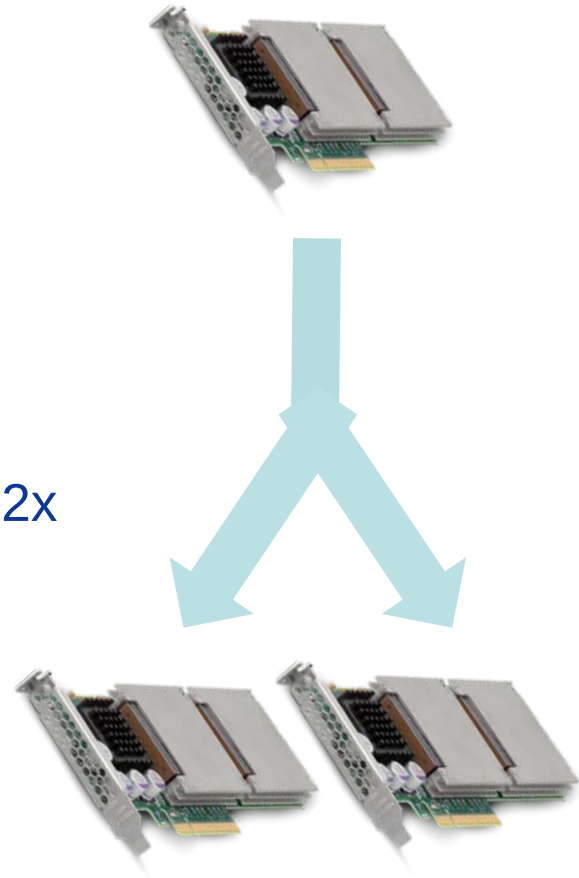


- Extend the logical address space to store more data

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Exadata Smart Flash Cache Compression

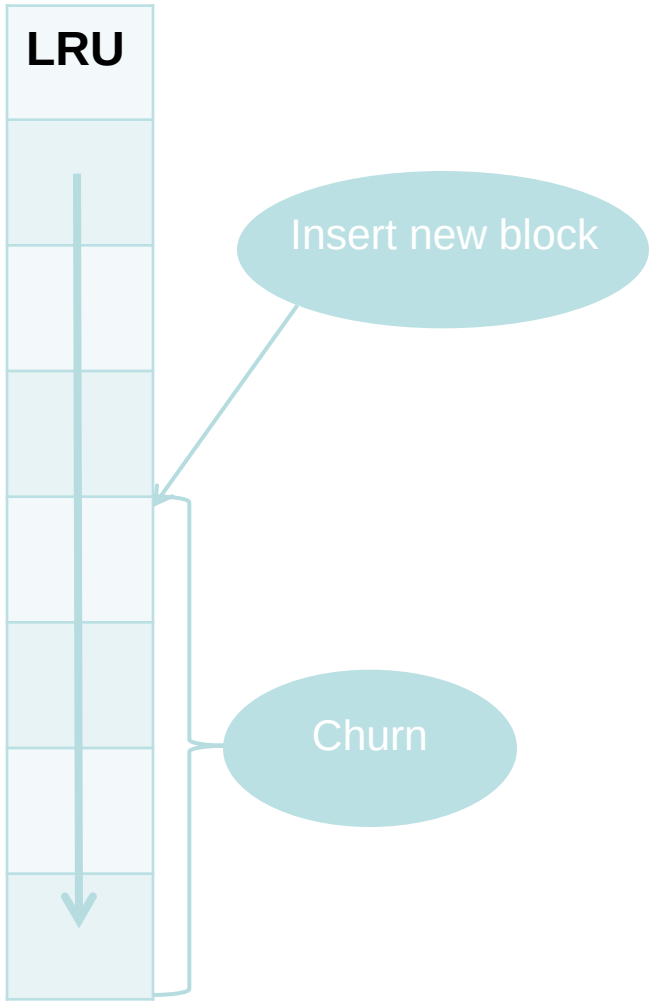
- Exadata automatically Compresses all data in Smart Flash Cache
  - Compression engine built into flash card
  - Zero performance overhead on reads and writes
  - Logical size of flash cache increases upto 2x
  - User gets large amount of data in flash for same media size
  - Elasticity of flash cache is completely automatic and transparent



Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Exadata Smart Flash Cache Scan Awareness

- On a traditional cache, if you scan dataset larger than cache size
  - Blocks 0,1,2,3 brought into cache, cache is full
  - Block 20,21,22,23 say replaces 0,1,2,3
- Repeat the same scan
  - Block 0,1, 2, 3 will replace blocks 20,21,22,23
  - Block 20,21,22,23 will again replace block 0,1,2,3
- Traditional caches churn with no actual benefit
- Some implementations call the insertion of new block in the middle scan resistant



Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Exadata Smart Flash Cache Scan Awareness

- Exadata Smart Flash Cache is scan resistant
  - Ability to bring subset of the data into cache and not churn
  - OLTP and DW scan blocks can co-exist
- Nested scans bring in repeated accesses
  - Repeat, For each item in large table, scan small table
  - Smart enough to pull the small table into flash since it is accessed repeatedly even though the size of large table alone is larger than flash cache
- No need to set “KEEP” attribute in data warehouses
- Happens automatically, no tuning or configuration needed

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Exadata Smart Flash Benefits

- Smart Flash Cache is database aware
- Smart Flash Logging avoids redo log outliers
- Smart Flash Compression doubles flash media capacity
- Smart Flash Cache Scan provides subset scannning and is table scan resistant
- Happens automatically, no tuning needed

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

Agenda

- Flash and Database
- Oracle's Innovations in Flash
- Q & A


ORACLE®