Exadata Unique Smart Data Caching

A Guide for Storage Caching Tier Implementation

October, 2025, Version 1.0

Public

ORACLE

Table of contents

Introduction

Oracle Exadata is an engineered database platform designed to deliver optimized performance for Oracle Database workloads, including AI, online transaction processing (OLTP), and analytics. Exadata combines specialized hardware and software, integrating high-performance compute servers, intelligent storage servers, and high-speed networking with Remote Direct Memory Access (RDMA) for low-latency data access. The resulting extreme performance enables improved resource efficiency, enabling leading scalability and consolidation of all workloads. The Exadata platform supports 100% compatible deployments across on-premises, Oracle Cloud Infrastructure (OCI), and multicloud environments (Azure, AWS, and Google Cloud), allowing users to develop once and deploy anywhere.

This technical brief focuses on Exadata's Smart Data Cache, a capability comprised of multiple caching components in the Exadata storage server. Exadata Smart Data Cache leverages advanced caching mechanisms, including Exadata RDMA Memory (XRMEM) and Exadata Smart Flash Cache, to optimize database performance. The Exadata Smart Flash Cache, integrated within storage servers, serves as a high-speed cache layer that intelligently stores frequently accessed "hot" data, such as tables or indexes, to reduce latency and accelerate I/O operations. XRMEM, a memory-based caching technology, complements this by providing ultra-low-latency access to the hottest data, further enhancing performance for I/O intensive workloads.

Exadata Smart Data Cache includes numerous optimizations to further boost performance by intelligently caching data. In-Memory Columnar Cache on Storage Servers (Columnar Cache) accelerates analytics on row-oriented operational data by storing in flash cache a columnar representation of the data. Smart Flash Log boosts redo log performance by storing redo directly in flash. These smart data caching capabilities uniquely leverage the cache in ways conventional architectures cannot, delivering extreme performance gains.

Exadata Storage Offloading

Exadata Smart Scan or storage offloading is a core capability of Exadata that leverages the intelligence of Exadata storage servers to offload database processing tasks from the database servers to the storage servers. This capability is enabled by the intelligent Exadata System Software, which runs on Exadata storage servers and works in tandem with the Oracle Database on the database servers. Unlike traditional storage systems containing hard disks or all-flash arrays with no built-in intelligence, Exadata storage servers actively process data, reducing the load on database servers, minimizing network traffic (only relevant rows and columns are returned), and boosting the performance for all types of workloads. For example, a traditional storage system, when an analytic query scans a table, the entire dataset is sent across the network to the database server, causing bottlenecks in bandwidth and CPU, whereas Exadata processes queries at storage tier, returning only filtered results. Exadata also offloads other critical tasks to storage servers like decryption, backup and recovery offloading, capabilities not available on traditional systems.

Exadata Smart Data Cache Concepts

Exadata leverages advanced caching mechanisms to optimize database performance, particularly for AI, analytics, and transactional processing workloads. The caching is achieved using a tiered structure that utilizes Exadata RDMA Memory (XRMEM), Exadata Smart Flash Cache, and the underlying storage (High Capacity disk or capacity-optimized flash). Exadata Smart Flash Cache in conjunction with XRMEM is one of the essential technologies that enables a single storage server on Oracle Exadata Database Machine to process up to 2,800,000 SQL Read I/O operations per second (IOPS), and scan data in flash at a rate of up to 100 GB/s. Advanced software capabilities, including Columnar Cache and Smart Flash Log leverage the cache to further accelerate specific database operations. All layers of Exadata Smart Data Cache are compatible with encrypted data that make use of Transparent Data Encryption (TDE), ensuring reliability and security. The encrypted blocks are cached transparently in both XRMEM and Smart Flash Cache.

Exadata RDMA Memory (XRMEM) Data Accelerator

Exadata RDMA Memory (XRMEM) Data Accelerator is an ultra-fast capability in Exadata Storage Servers that leverages Remote Direct Memory Access (RDMA) to provide ultra-low latency and high-throughput data access. By utilizing RDMA, XRMEM allows database clients to directly read from memory in the storage servers, bypassing traditional network and I/O stacks, which eliminates CPU interrupts and context switches. This results in significantly faster response times and lower read latencies compared to conventional storage solutions. XRMEM Data Accelerator enables a single storage server in Exadata to scan columnar data at a rate of up to 500 GB/s. Combined with RDMA, XRMEM Data Accelerator reduces OLTP read latency to as low as 14 microseconds. Real-world database workloads, utilizing the shared XRMEM Data Accelerator, can achieve up to 25.2 million OLTP Read IOPS (8K IOs) in a single rack. XRMEM is configured automatically, with no user interaction required, and is automatically managed.

Exadata Smart Flash Cache

Exadata Smart Flash Cache (Flash Cache) is a high-performance cache that uses flash storage to accelerate database read and write operations. It is designed to reduce I/O latency and improve throughput for frequently accessed data. Flash Cache resides in the storage servers acting as an intermediary between slower disk storage (or capacity-optimized flash), and database servers. While copies of the hottest data are kept in XRMEM, the Flash Cache automatically caches additional frequently accessed data on high-performance PCIe flash cards. Each storage server has up to 27.2 TB of performance-optimized flash to accelerate read and write operations.

Exadata In-Memory Columnar Cache on Storage Servers

Exadata In-Memory Columnar Cache on Storage Servers (Columnar Cache) utilizes Flash Cache and XRMEM to significantly accelerate analytical query performance while supporting transactional workloads. This capability automatically converts frequently accessed data, including Hybrid Columnar Compression (HCC) formats, into a pure columnar format stored in the Flash Cache and XRMEM, enabling faster scans, joins, and aggregations by accessing only required columns. Optimized for Exadata, the Columnar Cache uses in-memory optimizations like SIMD vector processing reducing I/O and CPU usage for analytics while maintaining row-based caching for online-transaction processing.

Exadata enhances analytical query performance by automatically promoting frequently accessed columnar data from the Columnar Cache in Flash Cache to XRMEM. This promotion leverages XRMEM's ultra-low-latency access via RDMA, , significantly accelerating columnar data reads for analytical workloads. This capability is fully automated, driven by Exadata's intelligent software that identifies hot columnar data based on access patterns, requiring no manual configuration. This tiered caching optimizes mixed OLTP and analytical workloads, with XRMEM providing high-speed access to critical data and Flash Cache handling larger datasets, all while maintaining transactional consistency.

The metadata for the Columnar Cache in Flash Cache is preserved across reboots, ensuring that the cache remains intact and functional after maintenance or storage server restart, improving reliability and minimizing the need to repopulate the cache.

Exadata Smart Flash Log

In an Oracle Database, all redo log entries must be written to the underlying disk drives in high-capacity (HC) storage or capacity-optimized flash in Extreme Flash (EF) storage. In an OLTP workload, fast response time for database log writes is crucial. Log writes to hard disks can become a performance bottleneck due to latency and I/O bandwidth limitations. Exadata Smart Flash Log enhances database performance and reliability by intelligently leveraging flash storage to optimize redo log writes in Exadata. This capability writes redo to high-performance flash devices in the Exadata storage servers, reducing latency for critical write operations compared to traditional disk-based logging. However, writes to flash may occasionally incur high latency, outliers that disrupt performance. The Smart Flash Log

Writeback capability eliminates these outliers by writing redo log data concurrently to Flash Cache on another flash device, and asynchronously flushing the writes to disk. It acknowledges the data as committed as soon as either the Smart Flash Log or the flash cache persistently store the data. Since the writes are to different flash devices and it is extremely unlikely to experience write latency outliers on two devices at the same time, this minimizes commit latency for mission-critical OLTP transactions. By minimizing I/O bottlenecks and accelerating transaction processing, this technology significantly boosts performance for applications requiring high transactional throughput, making it a key component of Exadata's high-performance architecture. Smart Flash Log Writeback works with Oracle Data Guard Primary and Standby database to provide fast, high throughput reads and writes for both online redo log files and standby redo log files.

Exadata Smart Data Cache Hardware and Performance

The hottest data is stored in the XRMEM tier of the Smart Data Cache. XRMEM is comprised of banks of DRAM, that can be directly read using RDMA technology. Each Exadata X11M High Capacity and Extreme Flash storage server has 1.25 TB of XRMEM and in conjunction with the RDMA over Converged Ethernet (RoCE) private network provide ultra-low latency OLTP reads of as low as14 microseconds and a scan rate of 500 GB/s.

Exadata Smart Flash Cache uses the latest PCIe performance-optimized flash technology rather than flash disks. PCIe flash greatly accelerates performance by placing flash directly on the high-speed PCIe bus rather than behind slow disk controllers and directors. Each Exadata Storage Server includes four 6.8 TB PCIe performance-optimized flash cards with a total capacity of 27.2 TB of flash capacity. A single Exadata X11M rack can hold up to 17 storage cells, including 68 PCIe flash cards providing 462.4 TB of flash capacity. The solid-state storage delivers dramatic performance advantages with Exadata storage. It implements automatic caching of database reads and writes and can deliver over 1,000,000 8K Flash Write IOPS in a single storage server. The performance scales as more database and storage servers are added to the infrastructure.

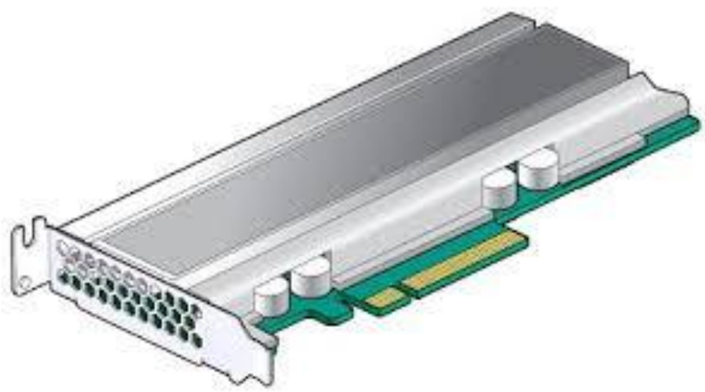


Figure 1 Oracle Flash Accelerator F680 PCIe Gen 5 Card

Exadata provides the choice of High Capacity (HC), Extreme Flash (EF), High Capacity-Z (HC-Z), and eXTended (XT) storage servers. HC, HC-Z and XT storage servers use high-capacity disk drives to store data, whereas EF servers store data on capacity-optimized flash devices. A X11M HC storage server consists of 12 x 22 TB disk drives, 1.25 TB of XRMEM and 4 x 6.8 TB of performance-optimized flash. A X11M EF storage server consists of 4 x 30.72 TB capacity-optimized flash drives, 1.25 TB of XRMEM and 4 X 6.8 TB of performance-optimized flash. A X11M HC-Z storage server consists of 6 x 22 TB disk drives, 576 GB of XRMRM, 2 x 6.8 TB flash. A X11M-XT storage server consists of 12 x 22 TB of disk drives, 2 x 6.8 TB of flash. Both types of storage servers leverage the Exadata Smart Data Cache to boost performance. While the capacity-optimized flash in EF servers is significantly faster than the disk storage in HC servers, the performance-optimized flash used by the Flash Cache in both servers is faster than capacity-optimized flash. The data, when resident in performance-optimized flash, provides a similar performance profile when running workloads on either HC or EF storage infrastructure.

The performance that XRMEM and Smart Flash Cache provide at the database level for successive generations of Exadata with 8 database servers and 14 storage servers (formerly known as a full rack) is shown in the following table.

Metrics		Exadata Database Machine X7-2 Full Rack[1]	Exadata Database Machine X8-2 Full Rack[1]	Exadata Database Machine X8M-2 Full Rack[1]	Exadata Database Machine X9M-2 Full Rack[1]	Exadata Database Machine X10M[2]	Exadata Database Machine X11M[2]
Raw Flash Data Bandwidth	HC	350 GB/s	350 GB/s	350 GB/s	630 GB/s	630 GB/s	1,400 GB/s
	EF[3]	350 GB/s	560 GB/s	560 GB/s	1,050 GB/s	840 GB/s	1,400 GB/s
Maximum SQL Exadata RDMA Memory Bandwidth	HC					5,320 GB/s	7,000 GB/s
	EF					5,320 GB/s	7,000 GB/s
Database Flash Read IOPS[4]	HC	4,776,000	4,776,000	12,000,000	22,400,000	22,400,000	22,400,000
	EF	4,776,000	4,776,000	12,000,000	22,400,000	22,400,000	22,400,000
Database Flash Write IOPS[4]	HC	4,352,000	4,352,000	6,580,000	8,596,000	12,824,000	14,000,000
	EF	4,352,000	4,352,000	6,580,000	8,596,000	12,824,000	14,000,000

[1] Each rack contains 8x database servers and 14x storage servers in a single rack
[2] Each rack contains 8x database servers and 14x storage servers in a multi-rack configuration
[3] Exadata X10M introduced a new architecture for Extreme Flash Storage Servers using 4x performance-optimized flash cards and 4x capacity-optimized flash drives. X9M and earlier EF servers utilized 8x performance-optimized flash cards in a combined data cache and persistence model
[4] Database flash read and write IOPS require database servers to drive OLTP workload on the storage servers

Table 1 XRMEM and Flash I/O Performance X7-2 to X11M

The Exadata Smart Data Cache Advantage

Exadata's Smart Flash Cache capabilities are unique. Unlike traditional storage systems where flash may be added as a disk replacement for better performance, Exadata integrates intelligent software that determines how and when flash storage should be used as part of a coordinated data caching strategy. With scalable Exadata storage, flash performance, and XRMEM combined with RDMA benefits can be delivered directly to the application layer. In contrast, traditional storage arrays suffer from internal and network bottlenecks that limit the benefits of flash, even when flash is added to these systems.

Advanced Architecture & Intelligent Data Tiering

Exadata's Smart Flash Cache is designed with unique software intelligence that dynamically determines the optimal use of flash and RDMA memory. Unlike traditional storage arrays with no built-in intelligence resulting in frequent system bottlenecks—Exadata integrates flash storage directly into a coordinated data caching strategy. Hot data is cached automatically into flash and XRMEM for high performance, while less frequently accessed (cold) data remains on disk. This advanced tiering means Exadata commonly achieves cache hit rates of over 90%, resulting in an effective flash capacity up to 10 times the physical hardware.

Superior Performance & Scalability

Exadata delivers unmatched throughput and IOPS scalability at the database level. A single storage server can provide up to 100 GB/s of flash throughput and 500 GB/s from XRMEM for the hottest data—far exceeding the capabilities of traditional storage arrays, which are often limited by controller bottlenecks to just a few GB/s, even with flash added. Exadata's architecture supports up to 2,800,000 read IOPS per X11M database and storage server, and adding more servers increases overall system capacity and performance. This design leverages scalable RoCE networking, direct RDMA memory access, and can scan from memory, flash, and disk in parallel.

Enterprise-Grade Reliability & Write Optimization

Exadata only uses enterprise-grade flash for resilience and endurance, avoiding the risks and degradation issues seen with consumer-grade flash. The Write Back Flash Cache capability adds another dimension by caching database block writes, eliminating disk bottlenecks for large-scale OLTP and batch workloads. With a single X11M storage server providing 1,000,000 8K write IOPS, and fully redundant persistent write caching, Exadata's I/O performance rivals that of entire arrays with thousands of disk drives.

Shared Resource Efficiency & Robust Data Protection

While typical external flash cards are bound to individual servers and can't be shared, Exadata enables flash and memory to be shared across the entire system, supporting advanced configurations like Oracle RAC. The architecture supports seamless consolidation, scalability, and high availability. Exadata software automatically detects and bypasses underperforming flash, rerouting I/O if failures or degradations occur, preventing database outages and reducing administrative complexity—capabilities that local flash deployments lack.

Performance and Cost Advantages Over Flash-Only Arrays

Some vendors have responded to bottlenecks by creating costly flash-only arrays, but these sacrifice efficient data tiering and limit performance benefits to the size of expensive flash. Exadata's architecture provides both greater effective capacity and better performance. It achieves this through optimized data movement, SQL offload processing, and direct database integration—features that flash-only solutions and traditional storage arrays can't match.

Management of the Exadata Smart Flash Cache

Exadata Smart Flash Cache's capabilities are offered right out of the box and do not require manual configuration. For example, database objects are loaded automatically in the Flash Cache depending on how frequently they are accessed. The frequent access of the data plays a key role in its operation, particularly in managing cache content when it is full. Exadata System Software ensures that frequently accessed data such as index blocks, control files, and file headers is prioritized, while non-repeatable I/O (e.g., backups, Data Pump, etc.) is not cached to avoid cache pollution. There are cases, especially with mission-critical application tables and indexes where manual tweaking ensures frequently accessed objects are loaded in the flash cache even before they are scanned.

Automatic tiering of Cache in Exadata Storage Server

Automatic tiering of cache on Exadata storage servers optimizes performance by dynamically managing data across multiple cache tiers, including XRMEM, and flash. Exadata's intelligent storage architecture automatically caches the most frequently accessed data in the fastest tiers, such as XRMEM, and flash cache, which provides rapid access to large volumes of hot data. For the most frequently accessed data, the system employs RDMA over the Exadata private network to efficiently move data between storage servers and database nodes, enabling database servers to directly read from XRMEM, bypassing traditional I/O bottlenecks and eliminating context switches. Exadata continuously monitors workload patterns, prioritizing data placement across flash, XRMEM, and persistent storage (hard drives or capacity-optimized flash as the case may be) to maximize performance and minimize latency, ensuring optimal resource utilization for demanding database applications.

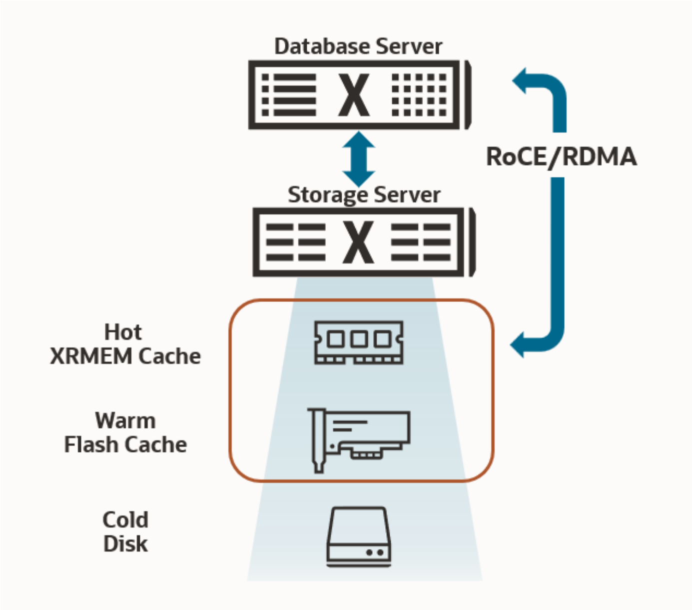


Figure 2 Auto tiering of cache in a storage cell

Automated Management of the Exadata Smart Flash Cache

The Exadata Smart Flash Cache holds frequently accessed data in very fast flash storage, while most of the data is kept in very cost-effective disk storage. This happens automatically and transparently, without the administrator having to take any action. The Flash Cache is smart because it knows when to avoid trying to cache data that will never be reused or will not fit in the cache.

When the database sends a read or write request to the Exadata storage servers, it includes additional information in the request about whether the data is likely to be read again and whether it should be cached. Based on the information the database sends, the Exadata System Software intelligently decides which data is likely to be re-read, and is worth caching, versus data that would just waste cache. Random reads and writes against tables and indexes are likely to have subsequent reads and normally will be cached and have their data delivered from the Flash Cache. In addition, Flash Cache operates in writeback mode and is persistent across server restarts and will not require any warmup period.

Knowing what not to cache is of great importance to realize the performance potential of the cache. For example, when writing backups or a mirrored copy of a block since these blocks will not be re-read in the near term, there is no reason to devote valuable cache space to these objects or blocks. Only the Oracle Database and Exadata System Software have this visibility and understand the nature of all the I/O operations taking place on Exadata. Having the visibility through the complete I/O stack allows optimized use of the Flash Cache hardware to store only the most relevant data.

Exadata System Software automatically caches objects read, in large table scans, in the Flash Cache based on how frequently the objects are read. The algorithm considers the size of the object, the frequency of access of the object, the displacement of the object in the cache based on the access pattern, and the type of scan being performed by the database. Depending on the Flash Cache size, and the other concurrent workloads, all or only part of the object is cached. For objects that are larger than the size of the flash cache, the software automatically caches only a portion of the object and eliminates thrashing the flash cache.

Starting with Exadata System Software 25ai, release 25.2, a storage server will automatically discard TEMP data from the Flash Cache immediately after the database completes the operation using it. Previously, TEMP data was written to the disk when hotter data required the flash space. This capability, which is unique to Exadata, optimizes Flash Cache space utilization and improves overall system performance by avoiding unnecessary disk I/Os.

Exadata System Software 25.2 also introduces enhancements to caching algorithms, including how large writes are handled within the Exadata Smart Flash Cache. Under this arrangement, priority is granted to large writes associated with specific operations (for example, TEMP and flashback logging) that consistently deliver the greatest performance benefits. Prioritization of large writes in Flash Cache occurs only when flash cache utilization exceeds 50%, either in terms of overall flash cache space usage or in terms of the flash cache space allocated to the associated Exadata I/O Resource Management (IORM) group. When utilization is below the 50% threshold, all large writes have the same opportunity to use the cache as before. While utilization exceeds the 50% threshold, only prioritized large writes are allowed to use the cache.

Automatic Loading of Objects in the Flash Cache

Preferential treatment over which database objects are cached is also provided with the Exadata System Software and Oracle Database. For example, objects can be kept in the cache and always be cached, or an object can be identified as one which should never be cached. This control is provided by the storage clause attribute, `CELL_FLASH_CACHE`, which can be assigned to a database table, index, partition, and LOB column.

There are three values to which the `CELL_FLASH_CACHE` attribute can be set:

1. `DEFAULT` specifies that the cache used for an object is automatically managed as described in the previous section.
2. `NONE` specifies that the object will never be cached.
3. `KEEP` specifies that the object should be kept in cache, once it is there.

Most workloads are best served by letting Exadata automatically manage the objects in cache. Overriding the default settings prevents Exadata from optimizing use the cache and is not recommended.

The following command could be used to direct the `CUSTOMERS` table to remain in Exadata Smart Flash Cache:

```
ALTER TABLE customers STORAGE (CELL_FLASH_CACHE KEEP)
```

This storage attribute can also be specified when the table is created.

In the above example, the Exadata Storage Server will cache data for the `CUSTOMERS` table and will keep it in flash while other tables that have not had the `KEEP` attribute specified will be aged out of cache. In the normal case where the `CUSTOMERS` table is spread across many Exadata Storage Servers, each Exadata cell will cache its part of the table in its own flash. Generally, there should be more flash cache available than the objects `KEEP` is specified for. This leads to the table being completely cached over time.

If `KEEP` has been specified for an object, and it is accessed via an offloaded Smart Scan, the object is kept in and scanned from cache. Another advantage of the flash cache is that when an object that is kept in the cache is scanned, the Exadata software will simultaneously read the data from both flash and disk to get a higher aggregate scan rate than is possible from either source independently.

Exadata System Software 24ai version 24.1.0 and above in conjunction with Oracle Database 23ai, database objects configured with the `KEEP` option are automatically loaded into Exadata Smart Flash Cache. This capability benefits online transaction processing (OLTP) workloads by avoiding cache misses on database objects configured with the `KEEP` option. Previously, the `KEEP` option was only effective after the data was read into the cache. The capability is helpful in certain use cases where an object must reside in the flash cache, however, the default setting is recommended.

Exadata I/O Resource Management

Consolidating multiple databases onto a single Exadata is a cost-saving solution for customers. Exadata I/O Resource Manager (IORM) can be used to prioritize the use of flash or XRMEM for the different databases running on Exadata. IORM protects the latency of critical OLTP I/O requests in flash cache. When table scans are running on flash concurrently with OLTP I/O requests, the OLTP latency may be impacted. IORM queues and throttles the table scan, and other lower priority requests. IORM can provide predictable performance by guaranteeing space in Flash Cache for different workloads. In a consolidated Exadata environment sharing the storage, Flash Cache space becomes a critical resource. IORM can reserve flash cache for critical databases or PDBs while preventing less important (test, development etc.) entities from consuming the entire flash cache. IORM plans can be extended across multiple Oracle Grid Infrastructure clusters sharing the same storage server resources including flash. This empowers administrators to reserve flash for the most performance critical databases and clusters.

When Flash Cache is allocated according to the IORM plan, the caching of redo log files is included in the space accounting for each multitenant (Container Database) or a non-CDB database. Exadata RDMA Memory (XRMEM) can be allocated in a similar manner as flash cache where higher priority databases can benefit from low-latency RDMA memory.

Exadata Storage Cache Observability

Oracle Database's Automatic Workload Repository (AWR) report has been enhanced from the time the observability was introduced in Exadata System Software 20.1. A new section – Flash Cache Redo Caching is added to showcase this new capability.

Flash Cache Redo Caching

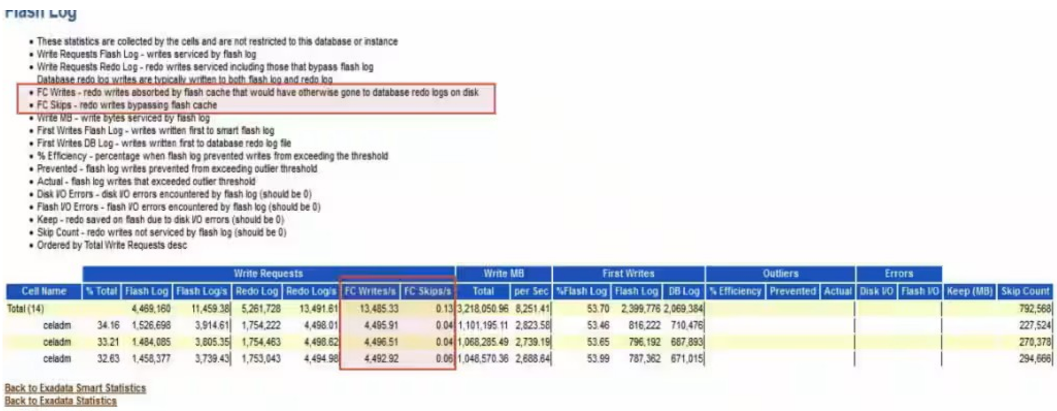
- These statistics are collected by the cells and are not restricted to this database or instance
- Flash cache may cache redo writes that would otherwise go to disk
- This works with, and does not replace, the use of flash log
- Total Writes - redo writes submitted, does not included population writes
- FC Writes - redo writes absorbed into flash cache
- Pop Writes - population writes into flash cache for redo reads

	Read Reqs					Write Reqs					Read MB					Write MB				
Cell Name	Total	per Sec	Hits/s	Misses/s	Skips/s	Total	per Sec	FC Writes/s	Skips/s	Pop Writes/s	Total	per Sec	Hits/s	Misses/s	Skips/s	Total	per Sec	FC Writes/s	Skips/s	Pop Writes/s
Total (14)						5,259,332	13,485.47	13,485.33	0.13							3,821,414.71	9,798.50	9,798.50	0.00	
celadm						1,753,651	4,496.54	4,496.51	0.04							1,273,866.13	3,266.32	3,266.32	0.00	
celadm						1,753,418	4,495.94	4,495.91	0.04							1,273,818.12	3,266.20	3,266.20	0.00	
celadm						1,752,263	4,492.98	4,492.92	0.06							1,273,730.47	3,265.98	3,265.98	0.00	

Back to Exadata Flash Cache
Back to Exadata Smart Statistics

Figure 3 Flash Cache Redo Caching section in AWR

Figure 3 outlines the section where the Flash Cache Redo Caching metric writes the total MBs in addition to individual storage statistics of Write Requests of redo logs in the flash cache including the skips.



Flash Log

- These statistics are collected by the cells and are not restricted to this database or instance
- Write Requests Flash Log - writes serviced by flash log
- Write Requests Redo Log - redo writes serviced including those that bypass flash log
  Database redo log writes are typically written to both flash log and redo log
- FC Writes - redo writes absorbed by flash cache that would have otherwise gone to database redo logs on disk
- FC Skips - redo writes bypassing flash cache
- Write MB - write bytes serviced by flash log
- First Writes Flash Log - writes written first to smart flash log
- First Writes DB Log - writes written first to database redo log file
- % Efficiency - percentage when flash log prevented writes from exceeding the threshold
- Prevented - flash log writes prevented from exceeding outlier threshold
- Actual - flash log writes that exceeded outlier threshold
- Disk I/O Errors - disk I/O errors encountered by flash log (should be 0)
- Flash I/O Errors - flash I/O errors encountered by flash log (should be 0)
- Keep - redo saved on flash due to disk I/O errors (should be 0)
- Skip Count - redo writes not serviced by flash log (should be 0)
- Ordered by Total Write Requests desc

	Write Requests							Write MB		First Writes			Outliers			Errors			
Cell Name	% Total	Flash Log	Flash Log/s	Redo Log	Redo Log/s	FC Writes/s	FC Skips/s	Total	per Sec	%Flash Log	Flash Log	DB Log	% Efficiency	Prevented	Actual	Disk I/O	Flash I/O	Keep (MB)	Skip Count
Total (14)		4,469,160	11,459.38	5,261,728	13,491.61	13,485.33	0.13	3,218,050.96	8,251.41	53.70	2,399,776	2,069,384							792,568
celadm	34.16	1,526,698	3,914.61	1,754,222	4,498.01	4,495.91	0.04	1,101,195.11	2,823.58	53.46	816,222	710,476							227,524
celadm	33.21	1,484,085	3,805.35	1,754,463	4,498.62	4,496.51	0.04	1,068,285.49	2,739.19	53.65	796,192	687,893							270,378
celadm	32.63	1,458,377	3,739.43	1,753,043	4,494.98	4,492.92	0.06	1,048,570.36	2,688.64	53.99	787,362	671,015							294,666

Back to Exadata Smart Statistics
Back to Exadata Statistics

Figure 4 Flash Log section in AWR with Writeback enabled

Figure 4 highlights the Flash Cache (FC) Writes and Skips columns within the Flash Log section of the AWR report showing the number of redo log Writes to Flash Cache.

Exadata XRMEM Cache Observability

The AWR report also contains information relating to XRMEM cache.

RDMA being software agnostic where database server memory reads directly from the memory of the storage servers, the accounting of RDMA I/Os is not there, therefore the Oracle Database statistics are critical in understanding the use of XRMEM cache. The XRMEM cache statistics in the Exadata-specific sections of the AWR report only include I/Os that are serviced by cellsrv, which are not RDMA I/Os.

The XRMEM Cache Configuration section contains summary information including the caching mode -- Write-Through or Write-Back (X8M and X9M) -- and overall size. The XRMEM Cache Space Usage section provides summary statistics on space usage in XRMEM cache.

Exadata Smart Flash Cache Observability

Exadata Smart Flash Cache can intelligently determine the data that is most useful to cache based on data usage, access patterns, and hints from the database that indicate the type of data being accessed. It is important to monitor the flash cache usage amongst all the databases running on Exadata Database Machine.

Oracle Exadata System Software 24.1 introduced the Exadata Storage Cache Statistics utility (ecstat), which provides real-time statistics about Exadata Smart Flash Cache. The ecstat utility improves cache observability on each Exadata storage server and reports the following:

- I/O operations balance across the main storage and Exadata Smart Flash Cache
- I/O operations (small reads, large reads, small writes, and large writes) on each storage server
- I/O throughput by main storage and Exadata Smart Flash Cache
- Sizing of Exadata Smart Flash Cache. Is it undersized?

The ecstat utility in Exadata System Software 25.1 has the following additional capabilities:

- Top reasons why I/O was not satisfied using Exadata Smart Flash Cache
- ecstat now reports performance statistics for top I/O consumers. Consumers may be databases that are explicitly specified in the IORM plan
- Top IOPS consumers sorted by I/O throughput

The ecstat utility is useful for determining the root cause of cache misses or performance issues. It provides the exact reason why a particular I/O bypassed the Flash Cache (e.g., nocache, large write etc.).

The AWR contains a wealth of information about Exadata Smart Flash Cache. The AWR report includes a separate section on Flash Cache Configuration and Space Usage, which provides summary statistics on the configuration (Write-Back or Write-Through) and space usage in Exadata Smart Flash Cache.

Flash Cache Configuration

- These statistics are collected by the cells and are not restricted to this database or instance
- Size (GB) - configured size for Flash Cache

Mode	Compression	Status	Size (GB)	Cells
WriteBack		normal	23845.81	All

Back to Exadata Flash Cache
Back to Exadata Smart Statistics

Flash Cache Space Usage

- These statistics are collected by the cells and are not restricted to this database or instance
- Space is at the time of the end snapshot
- Ordered by Space (GB) desc

		Default								Keep			
		OLTP			Large Writes					OLTP			
Cell Name	Space (GB)	%Clean	%Synced	%Unflushed	%Temp Spill	%Data/Temp	%Write Only	%Scan	%Columnar	%Clean	%Unflushed	%Scan	%Columnar
Total (3)	71,351.14	26.36	23.15	21.91	0.15	0.76	19.08	6.24	2.35	0.00	0.00		
dbm0celadm01	23,783.71	26.32	23.15	21.86	0.15	0.78	19.07	6.31	2.37	0.00	0.00		
dbm0celadm03	23,783.71	26.50	23.05	22.00	0.16	0.83	19.01	6.14	2.31	0.00	0.00		
dbm0celadm02	23,783.71	26.25	23.25	21.88	0.14	0.68	19.17	6.26	2.36	0.00	0.00		

Figure 5 AWR Report: Flash Cache Configuration and Space Usage

Flash Cache User Reads

The Flash Cache User Reads sections show information about read requests, read throughput, and read efficiency from database clients. The statistics show the I/Os against different areas of Exadata Smart Flash Cache:

- Small I/O - relates to block requests
- Scan - relates to scan requests
- Columnar - relates to columnar cache requests
- Keep - relates to read requests against the KEEP pool

Flash Cache User Reads

- These statistics are collected by the cells and are not restricted to this database or instance
- Total - total number of reads from Flash Cache
- OLTP/Scan/Columnar reads include reads on keep objects
- Misses/Partial Hits - number of read request misses and partial hits
- Ordered by Total Hit Read Requests desc

	Read Requests									Read Megabytes							
Cell Name	Total Hits	Hits per Sec	OLTP	Scan	Columnar	Keep	Misses Partial Hits	Active Secondary Hits	Active Secondary Misses	Total Hits	Hit MB/s	OLTP	Scan	Columnar	Keep	Active Secondary Hits	Active Secondary Misses
Total (3)	184,127	93.85	184,020	107			26,231			11,506.77	5.86	11,486.26	20.51				
dbm0celadm11	79,162	40.35	79,124	38			6,423			4,947.43	2.52	4,940.37	7.06				
dbm0celadm10	52,976	27.00	52,950	26			7,320			3,309.45	1.69	3,304.27	5.18				
dbm0celadm12	51,989	26.50	51,946	43			12,488			3,249.89	1.66	3,241.62	8.27				

Flash Cache User Reads Per Second

- These statistics are collected by the cells and are not restricted to this database or instance
- Total - total number of reads per second from Flash Cache
- OLTP/Scan/Columnar reads include reads on keep objects
- Ordered by Total Hit Read Requests per Second desc

	Read Requests per Second						Read MB per Second				
Cell Name	Total Hits	OLTP	Scan	Columnar	Keep	Misses	Total Hits	OLTP	Scan	Columnar	Keep
Total (3)	93.85	93.79	0.05			13.37	5.86	5.85	0.01		
dbm0celadm11	40.35	40.33	0.02			3.27	2.52	2.52	0.00		
dbm0celadm10	27.00	26.99	0.01			3.73	1.69	1.68	0.00		
dbm0celadm12	26.50	26.48	0.02			6.36	1.66	1.65	0.00		

Figure 6 AWR Report: Flash Cache User Reads

Flash Cache User Writes

The Flash Cache User Writes section shows information about write requests and write throughput for Exadata Smart Flash Cache in Write-Back mode.

In this AWR section, 'First Writes' indicate new data being written to Exadata Smart Flash Cache, while 'Overwrites' indicate data being overwritten in Exadata Smart Flash Cache. First Writes also require writing additional flash cache metadata.

Flash Cache User Writes

- These statistics are collected by the cells and are not restricted to this database or instance
- Total - total number of write requests or write megabytes to Flash Cache
- First Writes/Overwrites also include Keep Writes and Large Writes
- Ordered by Total Write Requests desc

	Write Requests											
	Total						per Sec					
Cell Name	Total	First Writes	Overwrites	Partial Writes	Keep	Large Writes	Total	First Writes	Overwrites	Partial Writes	Keep	Large Writes
Total (3)	581,100,526	4,933,642	558,665,341	17,501,543		340	296,177.64	2,514.60	284,742.78	8,920.26		0.17
dbm0celadm10	194,980,366	1,632,570	187,515,502	5,832,294		67	99,378.37	832.09	95,573.65	2,972.63		0.03
dbm0celadm12	193,278,721	1,710,956	185,730,887	5,836,878		78	98,511.07	872.05	94,664.06	2,974.96		0.04
dbm0celadm11	192,841,439	1,590,116	185,418,952	5,832,371		195	98,288.20	810.46	94,505.07	2,972.67		0.09

	Write Megabytes											
	Total						per Sec					
Cell Name	Total	First Writes	Overwrites	Partial Writes	Keep	Large Writes	Total	First Writes	Overwrites	Partial Writes	Keep	Large Writes
Total (3)	4,091,157.91	180,603.81	3,383,079.48	527,474.62		21.26	2,085.20	92.05	1,724.30	268.85		0.01
dbm0celadm10	1,368,968.39	59,910.65	1,133,010.90	176,046.84		4.19	697.74	30.54	577.48	89.73		
dbm0celadm12	1,362,516.62	60,948.97	1,126,965.98	174,601.67		4.88	694.45	31.06	574.40	88.99		
dbm0celadm11	1,359,672.90	59,744.19	1,123,102.59	176,826.11		12.19	693.00	30.45	572.43	90.13		

Figure 7 AWR Report: Flash Cache User Writes

Mission Critical Availability of the Exadata Smart Flash Cache

The hardware used for the Exadata Smart Flash Cache is very reliable, but all hardware is subject to failure. Spreading the flash cache across 4 PCIe cards mitigates some of this risk. If there is a failure of one of the flash cards, the Exadata System Software automatically detects the loss of the card and takes the failed portion of the flash cache offline. During this period the Exadata cell continues to operate and serve data from the remaining cache. This allows replacement of the failed flash card to be deferred until a convenient time when the Exadata cell can be taken offline, and the flash card replaced. After the card is replaced, the Exadata System Software automatically detects the presence of the new card and automatically starts using the additional flash. If there were "dirty" blocks in the failed flash card, which were not yet written to disk, then the Exadata System Software in conjunction with ASM or Exascale will automatically retrieve the mirrored copies from the other storage cells to recover the latest copy of the data.

If a storage server goes offline or a flash device fails, then the databases redirect the I/Os to use secondary mirrors for read operations.

Oracle Exadata System Software, in the background, prefetches the secondary mirrors of the OLTP data that is most frequently accessed into the Flash Cache on the replaced flash device, or storage server once it has been restarted. Oracle Exadata System Software automatically manages the secondary mirrors in the flash cache in an optimal way so that newer or more active secondary mirrors replace the cold data in the cache. When a flash device is replaced, the new flash device is warmed up before an I/O operation is redirected to it. The capability guarantees high availability for OLTP application workloads.

Exadata System Software 25ai, release 25.2, adds enhancements to the caching algorithms, enabling a storage server to perform a partner read to a remote cache rather than more expensive disk read on the local server. This capability effectively transforms a local flash cache miss into a remote cache hit, improving read I/O performance.

Conclusion

The Exadata Smart Data Cache is a key capability of Exadata systems, designed to greatly enhance performance for OLTP applications by caching frequently accessed data in high-speed flash storage and XRMEM (Exadata RDMA Memory). This intelligent caching mechanism lowers latency, speeds up I/O operations, and improves query performance by offloading disk I/O to faster storage layers. As a result, it reduces bottlenecks, increases throughput, and optimizes resource use.

For analytical workloads, Oracle Database on Exadata leverages the Columnar Cache, which automatically converts frequently accessed data—including data stored in Hybrid Columnar Compression (HCC) formats—into a purely columnar format within the flash cache. This enables faster scans, joins, and aggregations by allowing the database to access only the necessary columns. The most frequently accessed columnar data is automatically promoted to Exadata RDMA Memory for the lowest possible latency.

The Smart Data Cache ensures that Exadata is scalable, reliable, and efficient—making it essential for organizations running mission-critical database applications. It delivers outstanding IOPS for demanding workloads, can more than double scan rates for data warehousing and reporting, and provides special support for critical database logging functions.

With the ability to intelligently select and manage cached data, Oracle Database on Exadata—powered by the Exadata Smart Data Cache—is the first and only database optimized for both flash and RDMA access to XRMEM, setting a new standard for performance and efficiency.

ORACLE

Connect with us

Call +**1.800.ORACLE1** or visit **oracle.com**. Outside North America, find your local office at: **oracle.com/contact**.

blogs.oracle.com   facebook.com/oracle   twitter.com/oracle