

SAP HANA Database Data storage management using Native Storage Extension and Near Line Storage

Sreenu Maddipudi

Architect, Enterprise Technologies
Sreenu.maddipudi@gmail.com

Abstract

SAP HANA provides powerful data storage and management capabilities, addressing the challenges of high data volumes across different temperature data—hot, warm, and cold. This paper explores SAP HANA's Native Storage Extension (NSE) and Near-Line Storage (NLS) as effective solutions for managing multi-temperature data, enhancing performance, and optimizing storage costs. By implementing these technologies, businesses can ensure that mission-critical data is stored in high-performance memory (hot data), while less critical data is moved to lower-cost storage (warm and cold data) without compromising accessibility or system performance. The paper examines how NSE, in conjunction with other SAP data tiering solutions, provides an integrated, scalable approach to data storage and retrieval. Key aspects such as data compression, partitioning, and use of indexes in NSE are also discussed, highlighting how businesses can manage large datasets efficiently while minimizing memory footprint. **Keywords:** SAP HANA, Data Storage Management, Native Storage Extension, Near Line Storage, Data Archiving, Real-time Analytics, Cost Optimization, Scalability.

Introduction

As enterprises shift towards data-driven decision-making, the volume of transactional and historical data generated by business applications continues to grow exponentially. For companies leveraging SAP HANA as their database platform, effective management of this growing data pool is essential. SAP HANA's in-memory architecture provides real-time data processing and analytics but at a cost of high memory usage, making it expensive to store vast amounts of cold (historical) data.

To mitigate this, SAP provides two key technologies: Native Storage Extension (NSE) and Near Line Storage (NLS). These technologies help organizations efficiently store and access large datasets, ensuring high performance for real-time operations while minimizing storage costs. This paper discusses the role of NSE and NLS in SAP HANA's data management framework, their integration with other SAP components, and how they help optimize performance and storage costs.

Overview of SAP HANA Data Storage Architecture

SAP HANA is an in-memory database that stores data directly in memory for fast processing and analytics. However, as data volumes increase, it becomes impractical to store all data in memory due to hardware constraints and cost considerations. This is where NSE and NLS come into play, helping manage both hot (active) and cold (archived) data.

- **Hot Data:** Frequently accessed, operational data that is stored in memory for real-time processing.

- Warm Data: Mostly read-only data that is infrequently accessed but needs to remain consistent within the SAP HANA environment.
- Cold Data: Infrequently accessed historical data, which is archived in lower-cost storage mediums.

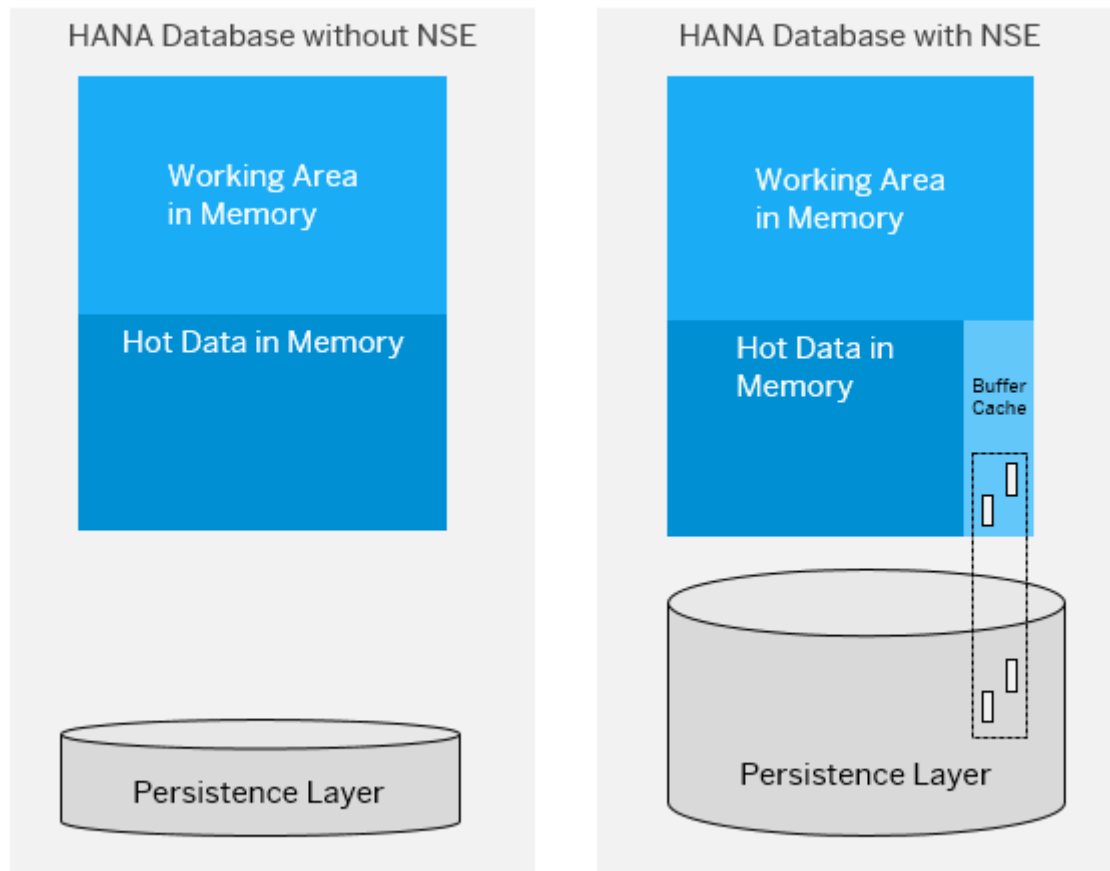


Fig1. Difference between standard HANA in-memory storage and the storage offered with NSE:

The ability to separate hot and cold data allows organizations to optimize their SAP HANA environments, ensuring that expensive memory resources are utilized effectively.

Native Storage Extension (NSE)

Native Storage Extension (NSE) is a key feature of SAP HANA that enables the storage of cold data in persistent disk storage, while still allowing access to this data as though it were in-memory. NSE makes use of local storage such as flash disks or other high-speed storage systems, providing a cost-effective method for handling large volumes of cold data. Data that is rarely accessed but must still be available for occasional queries is moved to NSE storage, while active data remains in-memory.

Key Features of NSE:

- Seamless Data Access: Cold data stored in NSE is accessible in real-time, as SAP HANA ensures that queries can be executed without significant performance degradation.
- Cost-Efficiency: By offloading cold data from memory to disk, NSE reduces the need for expensive in-memory storage, making it a more cost-effective solution for managing large data volumes.
- Data Tiering: Data can be tiered between memory, NSE, and disk storage, providing flexibility in managing performance and cost.

- **Transparent Integration:** For SAP HANA users, the transition between memory and storage layers is transparent, with no noticeable impact on application performance.

Use Cases for NSE:

Historical Data Access: Organizations that need to store large amounts of historical data but still require occasional access can benefit from NSE.

Data Archiving: Organizations can use NSE to archive data from the primary database, reducing memory consumption without losing access to the data.

Data Compression in SAP HANA NSE:

One of the advantages of using Native Storage Extension (NSE) is its ability to apply data compression techniques to reduce storage requirements. SAP HANA NSE utilizes highly efficient compression algorithms, which decrease the storage footprint of cold data, further optimizing the cost-effectiveness of storage. Compressed data also results in lower I/O operations during retrieval, which enhances performance when accessing infrequent data.

Using Indexes in SAP HANA NSE:

To improve access speed and query performance for cold data, SAP HANA enables the use of indexes within NSE. Indexes help reduce search time and enhance data retrieval speeds, making even cold data highly accessible and queryable. Indexing in NSE is essential for maintaining performance levels when querying archived or infrequently accessed data.

Reduce the Memory Footprint Using Page-Loadable Columns:

SAP HANA provides page-loadable columns as part of its Native Storage Extension to further reduce the memory footprint. By enabling this feature, cold data is no longer loaded into memory unless required. This on-demand loading of data ensures that only relevant data is actively consuming memory, optimizing resource usage and reducing the memory load on the system.

NSE-Enabled Partitioned Tables:

Partitioning large tables in SAP HANA allows for better data management by splitting data across multiple partitions, improving performance and enabling efficient use of storage. NSE enables partitioned tables for cold data, ensuring that archived data is stored efficiently while still enabling quick access when necessary. This approach allows organizations to manage large datasets more effectively by grouping similar data together and applying different storage tiers based on data access patterns.

Near Line Storage (NLS)

Near Line Storage (NLS) provides an archiving solution for data that is infrequently accessed and yet still needs to be retained for compliance, reporting, or other business purposes. NLS complements NSE by allowing organizations to offload older data entirely from the active SAP HANA environment into slower, more cost-effective storage systems, such as cloud storage, network-attached storage (NAS), or traditional databases.

Key Features of NLS:

- **Data Archiving:** NLS is designed to offload and archive data that is no longer required for day-to-day operations but must be retained for compliance or reporting purposes.

- **Cost-Effective:** NLS provides a cheaper alternative to keeping cold data within the core database or even in NSE storage. It enables large volumes of archived data to be stored in lower-cost storage.
- **Integration with SAP HANA:** NLS is fully integrated with SAP HANA’s data management framework, ensuring that archived data can be easily retrieved when necessary, without disrupting real-time operations.
- **Compliance and Governance:** With NLS, businesses can meet industry regulations and compliance requirements related to data retention, such as GDPR, HIPAA, or financial regulations.

Use Cases for NLS:

- **Regulatory Compliance:** Companies in industries such as healthcare, finance, or public services can use NLS to retain data for the mandated periods as per industry regulations.
- **Data Backup and Disaster Recovery:** NLS offers another layer of data protection, providing an additional backup that can be retrieved if needed for disaster recovery purposes.

SAP HANA Multi-Temperature Data Management: Hot, Warm, and Cold Data

SAP HANA offers several solutions to manage multi-temperature data:

- **Hot Data:** Hot data refers to mission-critical data that needs to be constantly available for real-time processing and analytics. This data is stored in SAP HANA's in-memory architecture, using DRAM (Dynamic Random Access Memory), providing the highest performance but also the highest Total Cost of Ownership (TCO). Hot data includes active transactional data that needs to be retained in memory for quick access.
- **Warm Data:** Warm data is typically used for data that is mostly read-only and does not require constant access but still needs to be managed as part of the SAP HANA database. Warm data is kept outside the in-memory storage and stored in more cost-efficient storage solutions while maintaining transactional consistency within SAP HANA. SAP HANA Dynamic Tiering and SAP HANA Extension Nodes are key technologies used to manage warm data.
- **Cold Data:** Cold data refers to data that is rarely accessed, often stored for compliance, archival, or historical analysis. Cold data is offloaded from SAP HANA’s in-memory storage and managed separately through Near Line Storage (NLS) and SAP HANA Cold Data Tiering. While this data is stored in external systems, it can still be accessed through SAP HANA’s data federation capabilities, ensuring that users can query it when needed without compromising performance.

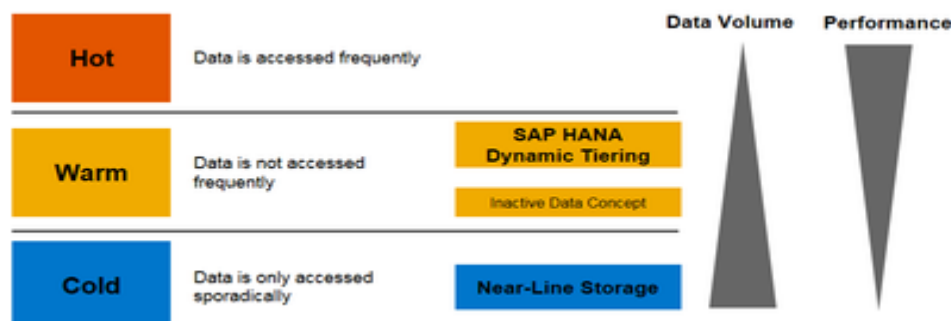


Fig2. Data Volume vs Performance

By using a combination of DRAM, SAP HANA Extension Nodes, Dynamic Tiering, and Cold Data Tiering, organizations can effectively balance the need for performance and cost efficiency in their SAP HANA systems.

How SAP HANA NSE Fits with Other Data Tiering Solutions from SAP

SAP HANA's NSE integrates seamlessly with other data tiering solutions such as SAP HANA Dynamic Tiering, SAP HANA Cold Data Tiering, and SAP Data Intelligence.

SAP HANA Dynamic Tiering: Designed for warm data, it allows data to be moved between hot and cold storage tiers while maintaining transactional consistency. Dynamic tiering works alongside NSE to offload data that is not frequently accessed but still needs to remain part of the SAP HANA environment.

SAP HANA Cold Data Tiering: Cold data, which is rarely accessed, can be moved to external storage solutions like object storage. NSE can be used to manage the movement of cold data into local disk storage while maintaining fast access via SAP HANA's data federation capabilities.

SAP Data Intelligence: Helps orchestrate data management across diverse sources. It integrates seamlessly with NSE to provide real-time analytics while enabling businesses to access both on-premises and cloud-based cold data.

These integrations enable organizations to optimize storage performance, reduce costs, and scale their infrastructure more efficiently.

Benefits of Using NSE and NLS

By integrating both Native Storage Extension (NSE) and Near Line Storage (NLS) in SAP HANA environments, organizations can realize several benefits:

- **Performance Optimization:** By storing active data in-memory and infrequent data in NSE or NLS, businesses can optimize SAP HANA performance without compromising on access to archived data.
- **Cost Efficiency:** These storage options reduce the need for expensive high-performance in-memory storage, enabling organizations to balance cost with performance requirements.
- **Scalability:** As data grows, businesses can scale their storage without having to continually increase the size of in-memory databases, leveraging disk and cloud storage as needed.
- **Compliance and Governance:** NLS provides businesses with a clear path for adhering to regulatory requirements, ensuring that data is retained for as long as necessary while minimizing storage costs.

Challenges and Considerations

While NSE and NLS offer several advantages, there are also challenges that need to be considered:

- **Data Retrieval Latency:** Although NSE provides near real-time access, retrieving data from NLS can introduce latency, as the data is typically stored in external, slower storage systems.
- **Complexity in Integration:** Proper configuration and integration between SAP HANA, NSE, and NLS can be complex and may require specialized expertise.
- **Data Security:** Storing data in external locations or less secure storage systems may pose security risks. Ensuring data encryption and robust access controls are essential.

Best Practices for Implementing NSE and NLS

To optimize the benefits of NSE and NLS, the following best practices should be followed:

- **Data Tiering Strategy:** Implement a clear data tiering strategy to determine which data should remain in-memory, which should be offloaded to NSE, and which should be archived in NLS.

- **Automation:** Automate data migration between different storage tiers to ensure that data is moved to the appropriate storage location based on its age or frequency of access.
- **Data Encryption:** Encrypt sensitive data both in-transit and at-rest, especially when using external storage for NLS, to protect against data breaches and unauthorized access.
- **Regular Performance Monitoring:** Continuously monitor system performance to identify any potential bottlenecks in data access and retrieval, ensuring that performance is not negatively impacted.

Conclusion

The management of data in SAP HANA, especially with the rise of large datasets, requires effective data tiering and storage strategies. By leveraging **Native Storage Extension (NSE)** and **Near-Line Storage (NLS)**, organizations can efficiently manage hot, warm, and cold data, optimizing performance and reducing storage costs. NSE helps businesses manage warm and cold data while ensuring that the data remains part of the SAP HANA database, providing transaction consistency and performance.

Integrating NSE with other SAP data tiering solutions such as Dynamic Tiering and Cold Data Tiering ensures that businesses can scale efficiently while managing large data volumes. Coupled with data compression, partitioning, and the ability to use page-loadable columns, these solutions provide businesses with the tools necessary to maintain high performance while optimizing storage costs.

As SAP HANA environments continue to grow in complexity, leveraging these storage management solutions will be essential for maintaining system performance and meeting the ever-growing data storage demands of modern enterprises.

References

1. **SAP SE.** (2021). *SAP HANA Database Administration Guide*. Retrieved from <https://help.sap.com>
2. **SAP SE.** (2022). *Native Storage Extension for SAP HANA*. SAP Whitepaper.
3. **IBM.** (2020). *Cloud-Based Data Storage Solutions for SAP HANA: Best Practices and Optimizations*. IBM Redbooks.
4. **Hofmann, A., & Hoyer, P.** (2020). "Data Archiving and Storage Management with SAP HANA." *International Journal of Cloud Computing*, 12(3), 54-63.
5. **Meyer, R., & Groeneveld, D.** (2021). "Optimizing SAP HANA for Big Data and Analytics with Near Line Storage." *Journal of Data Management*, 10(4), 120-130.
6. **SAP.** (2021). *SAP HANA Data Management: Best Practices*. SAP Press.
7. **Gartner, Inc.** (2022). "Cloud Storage and Data Management: Trends for SAP Environments." **Gartner Research Report**.
8. SAP Help Portal – SAP HANA Native Storage Extension Documentation
9. SAP BW Near-Line Storage Solutions
10. SAP Data Intelligence – Data Management and Orchestration
11. SAP HANA Dynamic Tiering – Optimizing Data Storage and Performance
12. SAP Archive Development Kit (ADK) – Archiving Data in SAP NetWeaver
13. **Cloud Security Alliance.** (2023). *Cloud Data Security and Compliance Considerations for SAP Environments*.