



Crossload
Consulting GmbH



Jens Gleichmann
Geschäftsführer
jens.gleichmann@x-l-c.de

HANA NSE details

Warum NSE die HoP TCO verbessert

07.10.21



www.linkedin.com/in/jens-gleichmann
www.xing.com/profile/Jens_Gleichmann

Agenda

1. NSE – technical details
 - a) Overview
 - b) Criteria
 - c) Implementation

2. Side effects

3. Project Review
 - a) PoC customer industry: insurance
 - b) Implementation customer industry: retailer

4. Take-aways

NSE

Native Storage Extension (warm)

Data Tiering - NSE

Overview

SAP HANA Native Storage Extension



SAP HANA Native Storage Extension (NSE) is a general-purpose, built-in warm data store in SAP HANA that lets you manage less-frequently accessed data without fully loading it into memory. It integrates disk-based or flash-drive based database technology with the SAP HANA in-memory database for an improved price-performance ratio.

Needed license: SAP HANA, **Standard Edition**

Used memory by NSE has to be licensed like hot store memory

Store data according to temperature	<ul style="list-style-type: none">• Provide options for native data tiering for the storage of warm data which is less frequently accessed• Provide support for non-volatile memory hardware to store hot data in persistent memory
-------------------------------------	--

Data Tiering - NSE

Overview



Hot store

Persistent memory (PMEM/vPMEM) extends the in-memory storage capacity for hot data in SAP HANA.

Warm store

Native storage extension is an intelligent, built-in disk extension for the SAP HANA in-memory database. It is the primary warm store option for SAP HANA on premise and SAP Cloud Platform, SAP HANA service.

Cold store

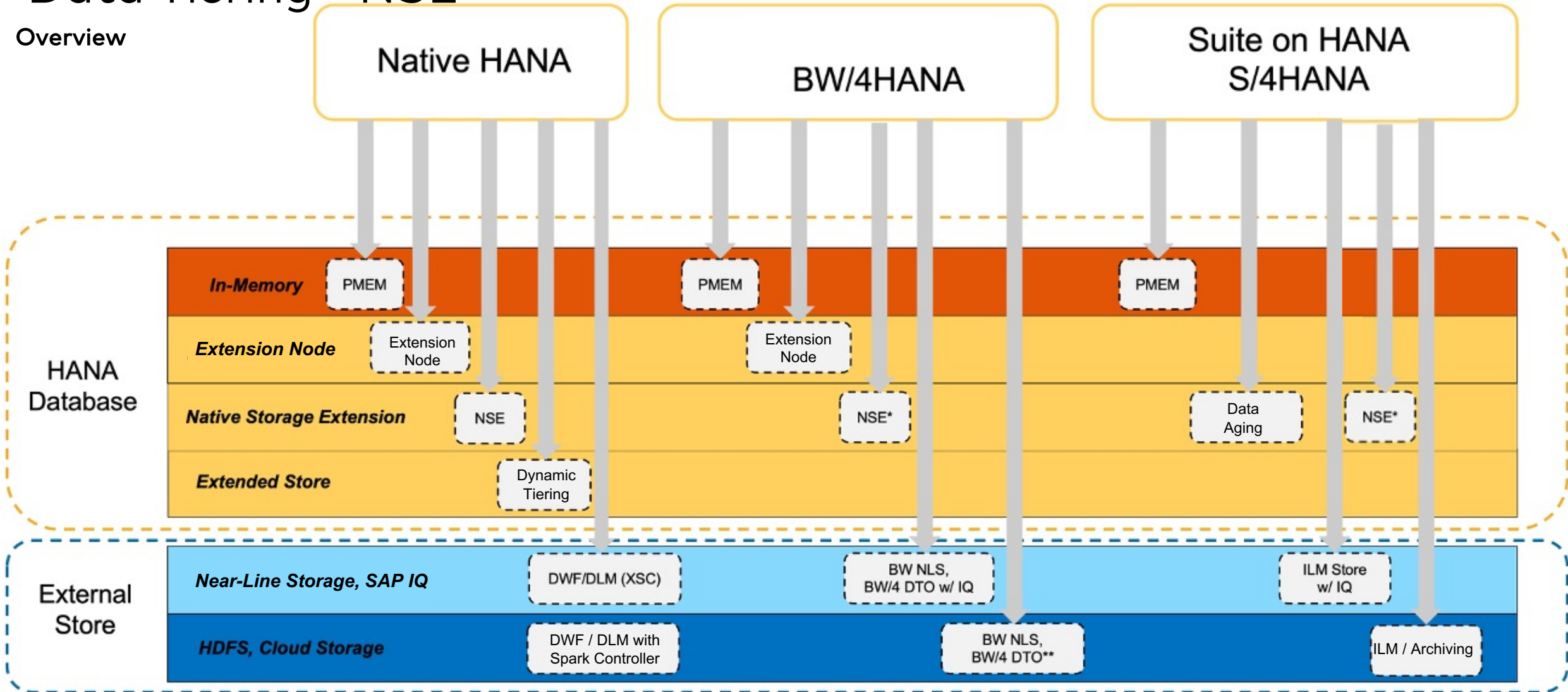
SAP HANA cold data tiering provides persistence capabilities for SAP HANA cold data in external data stores, such as HDFS, Azure Data Lake, and Big Data services from SAP.

- Data remains inside HANA
- Data are inside HSR, backup (memory / storage costs)
- System growth means longer startup time, backup time, recovery time and initial sync of HSR as well the take over duration

- Data stored externally (low costs storage) away from HANA

Data Tiering - NSE

Overview

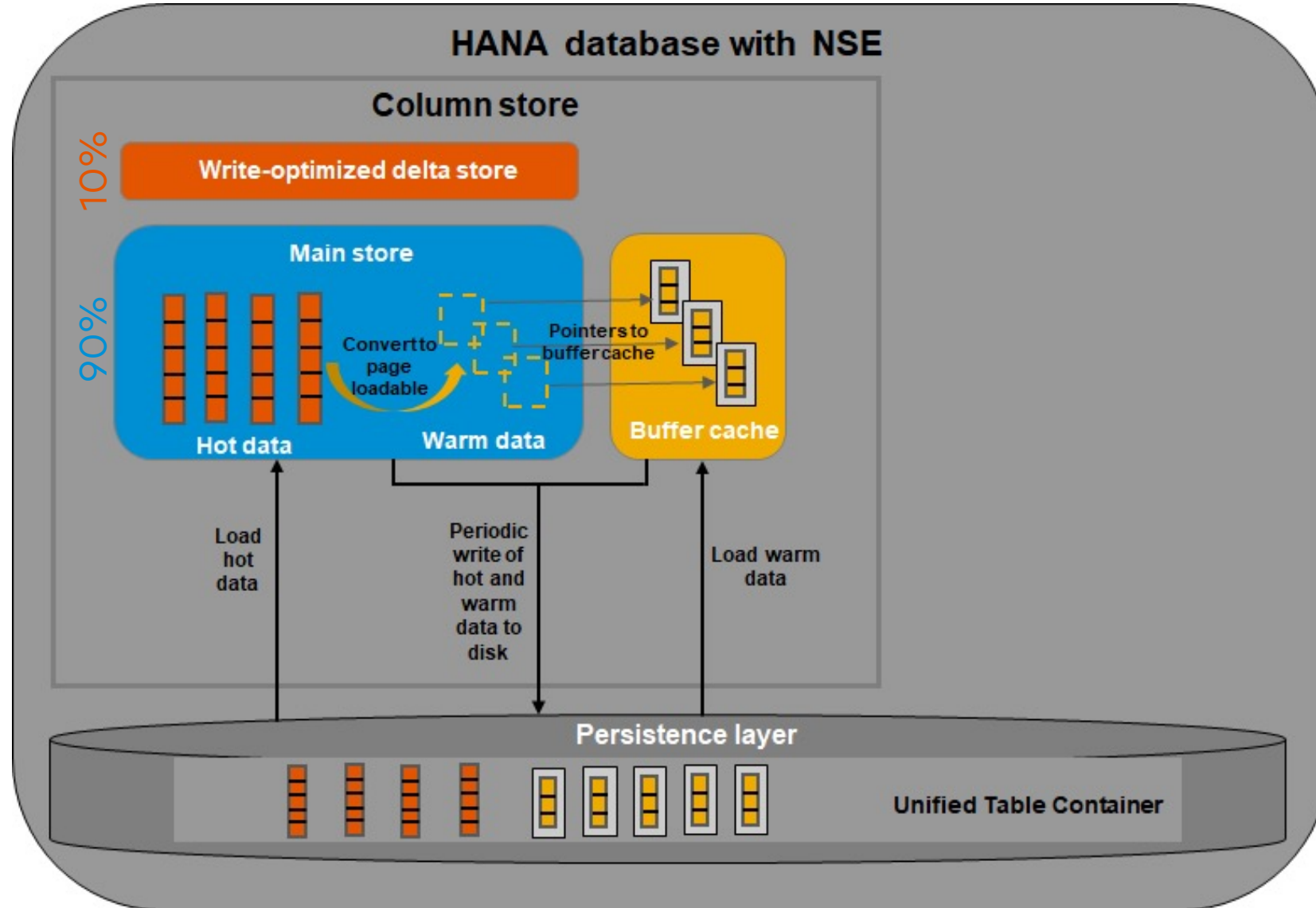


* Exclusively available for SAP BW/4HANA, S/4HANA and BSoH (Note:2816823)

** SAP HANA Data Lake (IQ) for DTO is planned

Data Tiering - NSE

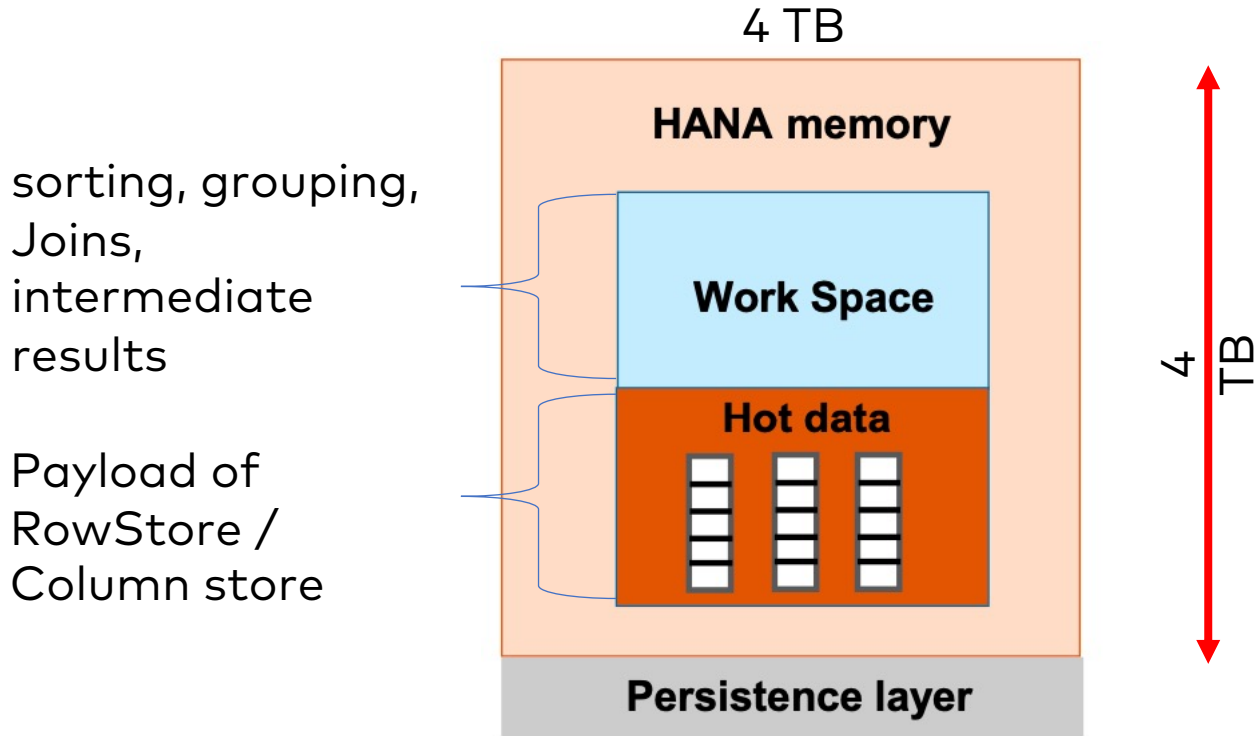
Overview



Data Tiering - NSE

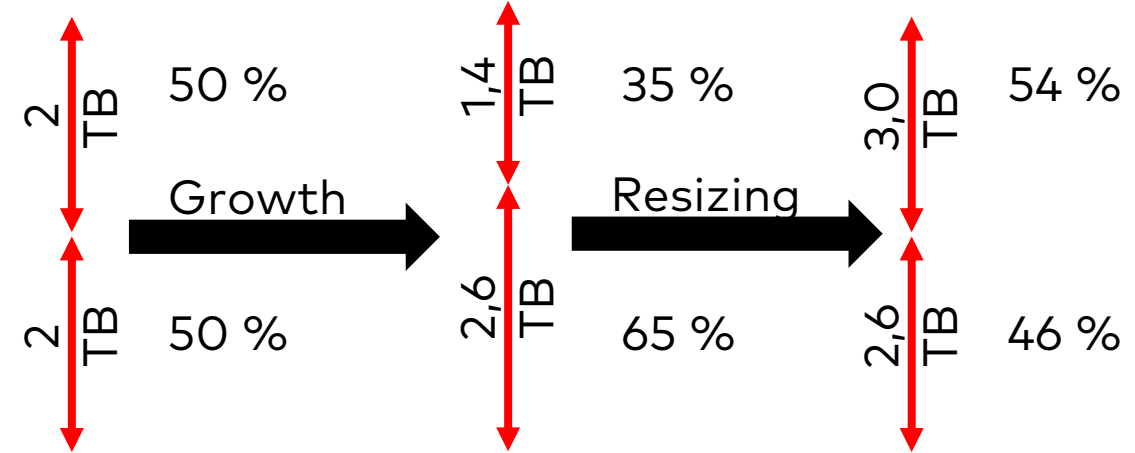
Overview

Normal growth of a HANA system



initial Sizing

Max. usage

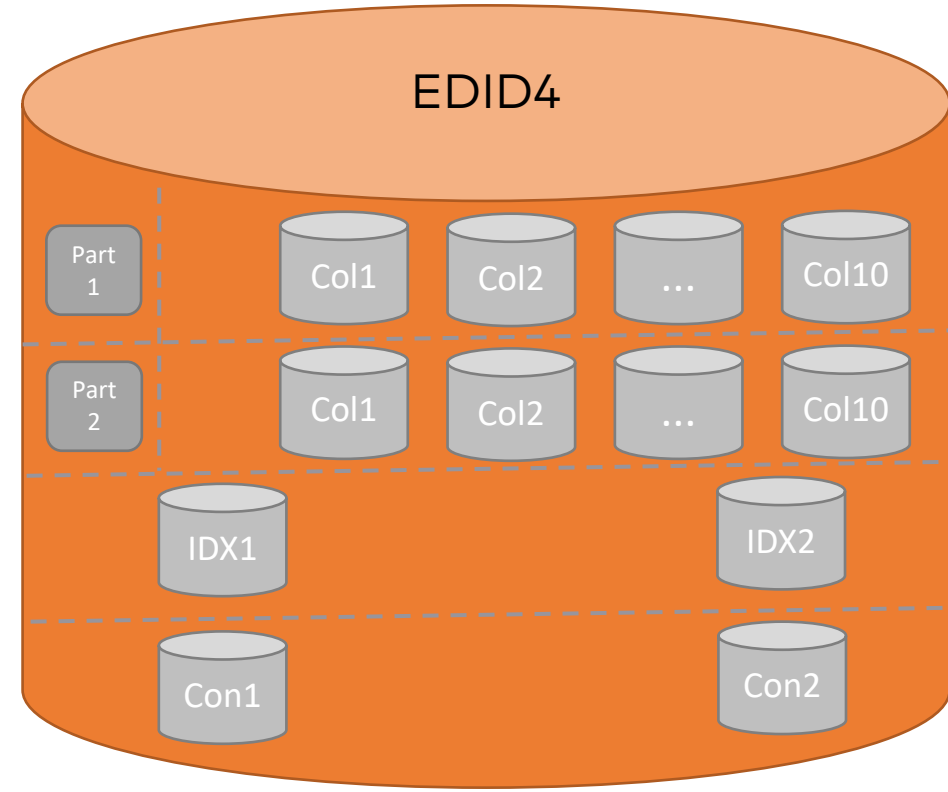
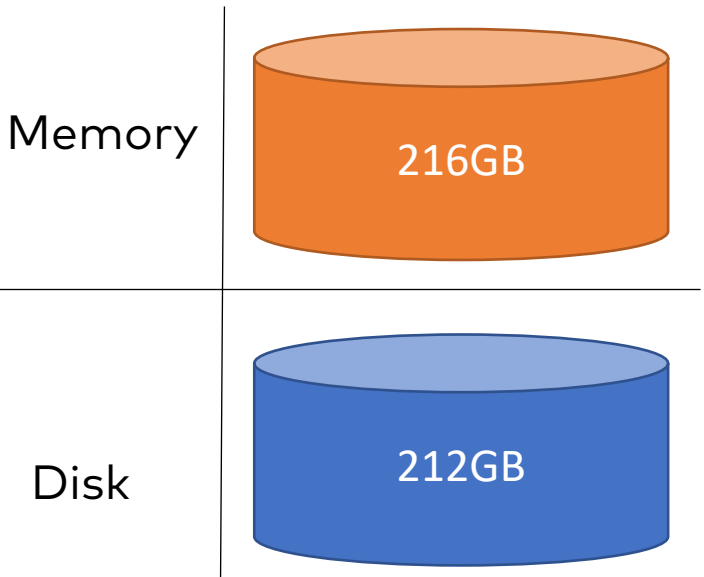
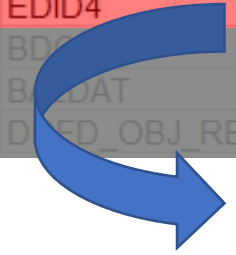


SAP Peaksizing of HANA Payload: **100%** of all CS+RS data, even if not loaded!

Data Tiering - NSE

Overview

Name of Table	Tabl Store Type	Table Rank	No. Columns	No. Records	Total Disk Size	Tot Cur Mem Siz	Cumulated %	Tab Size in Mem
EDID4	C	1	10	4,487,116,462	212.65	216.24	21.53	165.05
...	C	5	14	868,953,040	44.82	46.18	58.85	31.68
...	C	6	8	81,243,304	32.72	33.18	62.16	31.33
...	C	25	15	182,181,324	6.77	7.32	85.18	3.27

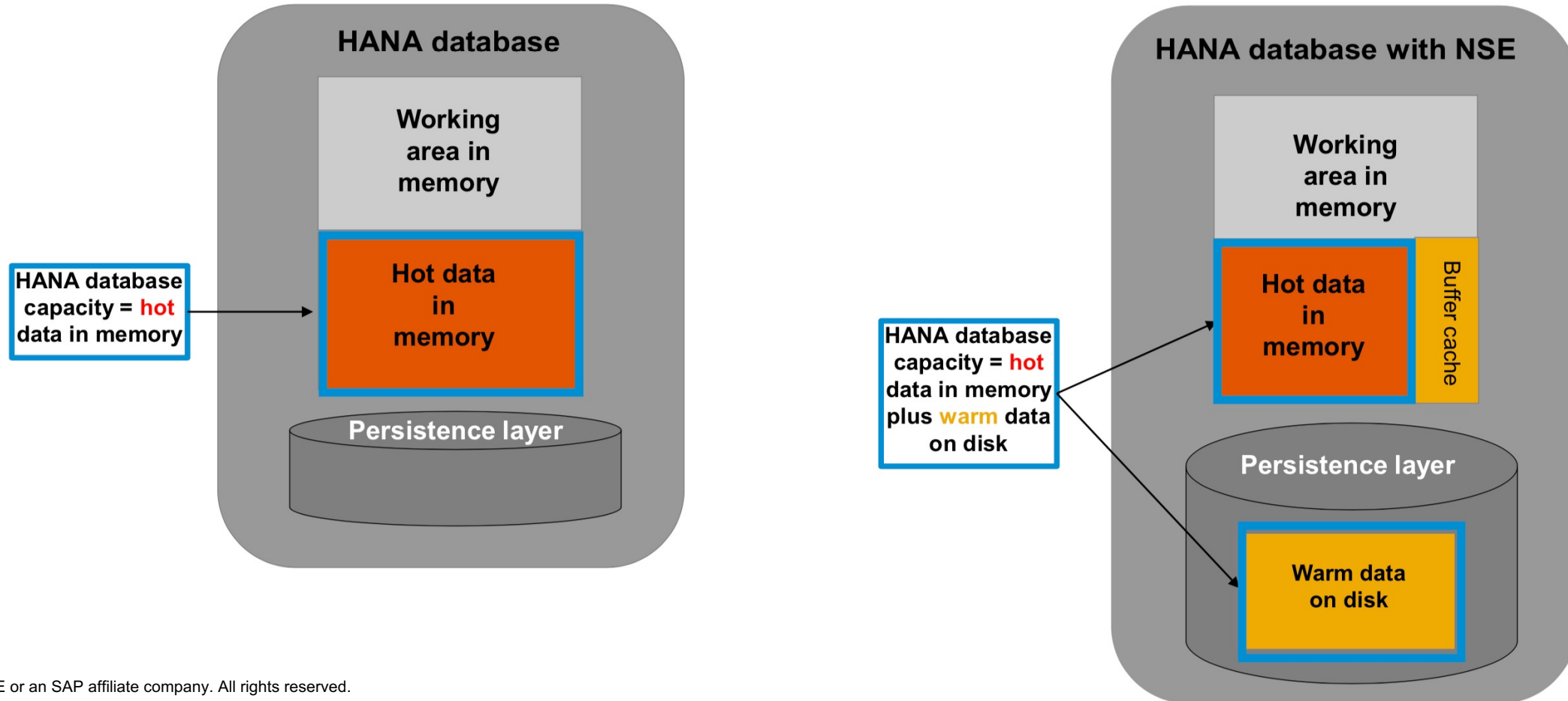


Legend:
 Col = Column
 IDX = Index
 Con = Concat Attribute

Data Tiering - NSE

Overview

NSE manages “page loadable” warm data in the HANA database with expanded disk capacity, and an intelligent buffer cache to transfer pages of data between memory and disk.



Data Tiering - NSE

Overview

Name of Table	Tabl Store Type	Table Rank	No. Columns	No. Records	Total Disk Size	Tot Cur Mem Siz	Cumulated %	Tab Size in Mem
EDID4	C	1	10	4,487,116,462	212.65	216.24	21.53	165.05
BDCP2	C	5	14	868,953,040	44.82	46.18	58.85	31.68
BALDAT	C	6	8	81,243,304	32.72	33.18	62.16	31.33
DRFD_OBJ_REP_STA	C	25	15	182,181,324	6.77	7.32	85.18	3.27

Field	Data Type	length (Decimals)	Description
MANDT	CLNT	3	Client
DOCNUM	NUMC	16	IDoc number
COUNTER	NUMC	3	Counter in cluster table
SEGNUM	NUMC	6	Number of SAP segment
SEGNAM	CHAR	30	Name of SAP segment
PSGNUM	NUMC	6	Number of the hierarchically higher SAP segment
HLEVEL	NUMC	2	Hierarchy level
DTINT2	INT2	5	Length field for VARC field
SDATA	LCHR	1000	Application data



Partitioning Attribute

4,5 bln Rows
= 14 partitions => 320 mln per partition

216 GB Memory
= 14 partitions => 15 GB per partition

Data Tiering - NSE

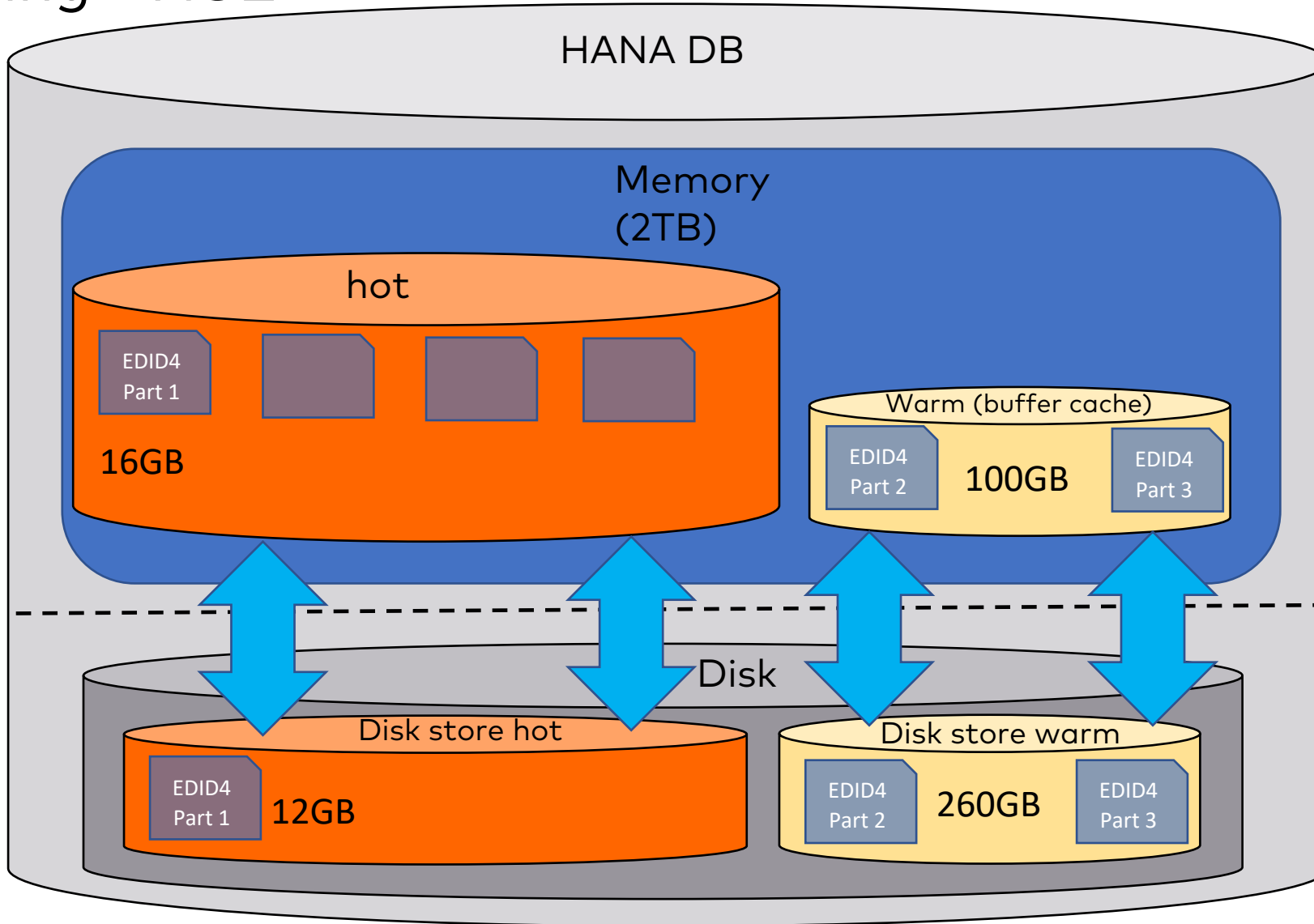
Overview

EDID4:
216GB memory
212GB disk

NSE:
200GB of EDID4

Hot:
16GB memory
12 GB disk

Tables / columns can be loaded partially or full

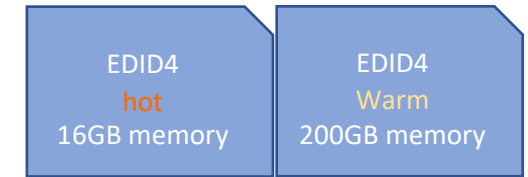


Working memory= Rest of memory which is not allocated by buffer cache or hot memory

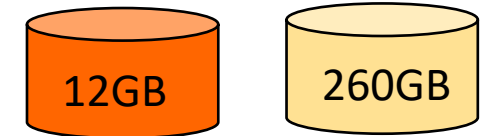
Data Tiering - NSE

Overview

Disk	default	hot	warm
Disk	216 GB	12GB	260GB
Sum	216 GB	272GB	
Savings		+ 56 GB	



Memory	default	hot	warm
memory	212 GB	16GB	100GB (*limited by buffer cache)
Sum	212 GB	116 GB	
Savings		- 96GB	



* Worst case scenario, if EDID4 is the only table in NSE. Efficiency calculation scales with the buffer cache sizing and the number of unused data

Data Tiering - NSE

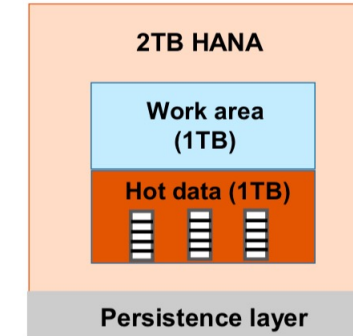
Overview

On-premise sizing

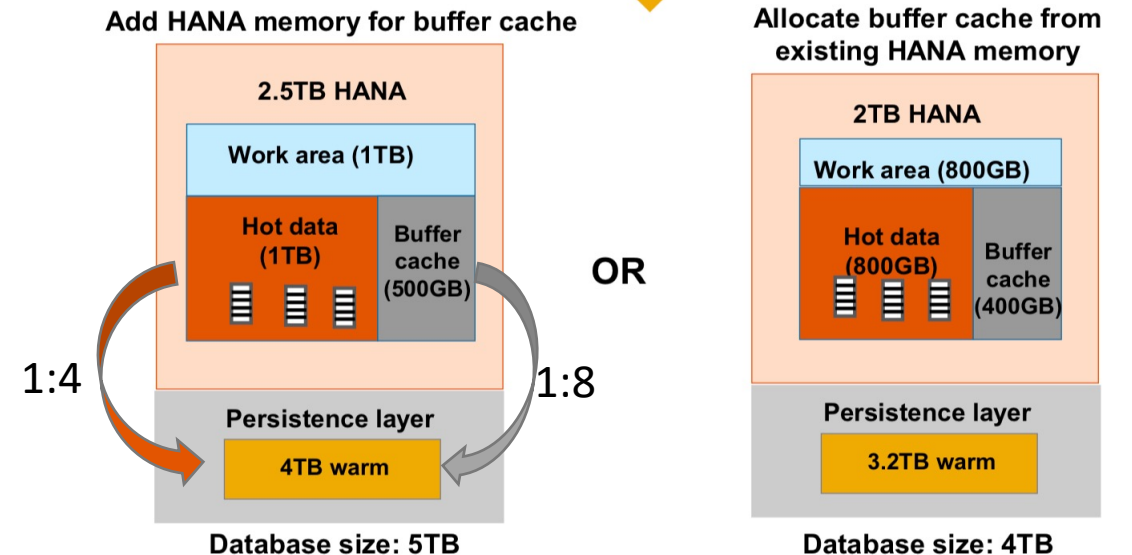
- Determine volume of warm data to add to the HANA database
- May add as much warm storage as desired - up to 1:4 ratio of HANA hot data in memory to warm data on disk
- Divide volume of warm data by 8 – this is size of memory buffer cache required to manage warm data on disk
- Either add more HANA memory for buffer cache, or use some of existing HANA memory for buffer cache (will reduce hot data volume)
- Work area should be same size as hot data in memory (equivalent to HANA with no NSE)

© 2019 SAP SE or an SAP affiliate company. All rights reserved. | PUBLIC

Example: begin with 2TB HANA system
Database size: 1TB



Expand HANA database capacity with warm data that is 4x the size of hot data



Data Tiering - NSE

Overview

Table level	Column level	Partition level	Index level
complete table	single/multiple columns can be selective relocated to NSE	single/multiple partitions can be selective relocated to NSE	single/multiple indexes can be selective relocated to NSE
complete table in memory (column loadable) = hot	complete column in memory (column loadable) = hot	complete partition in memory (column loadable) = hot	complete index in memory (column loadable) = hot
complete table is NSE managed (page-loadable) = warm	complete column is NSE managed (page-loadable) = warm	complete partition is NSE managed (page-loadable) = warm	complete index is NSE managed (page-loadable) = warm

Data Tiering - NSE

Criteria

Two possibilities for an analysis

NSE Advisor

Advisor is running for a defined time (min. 7 days) in the prod. system

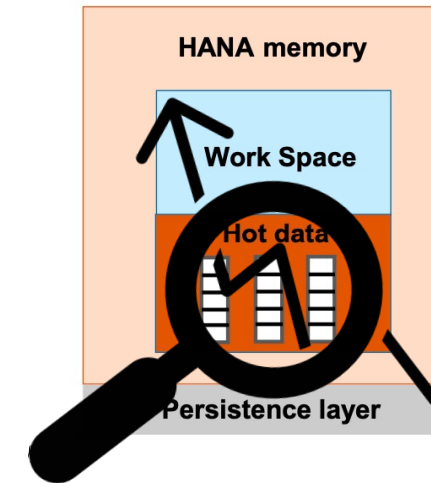
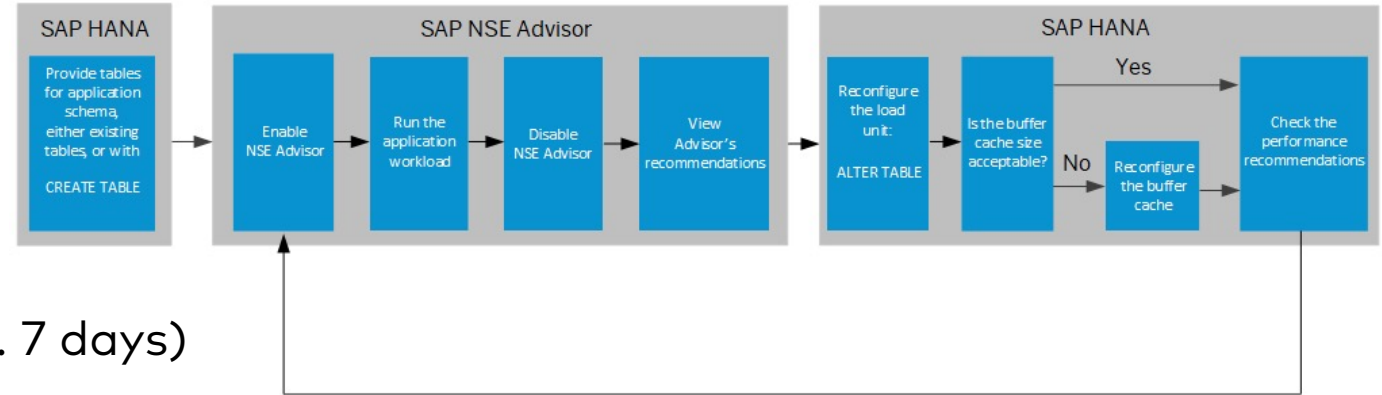
20-30% Performance Overhead

=> No exact scenario, general analysis

Manually

Top20 tables/indexes incl. SQL- und process analyses

=> Exact scenario, application and selects are known

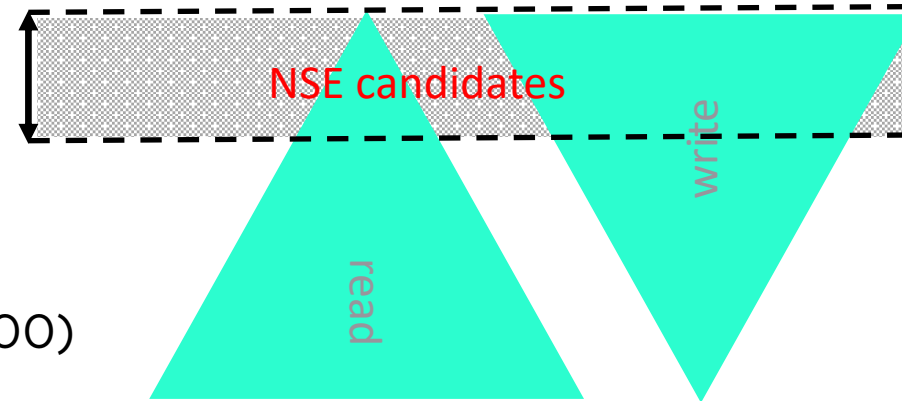


Data Tiering - NSE

Criteria

Selection criteria

high growing rate => many **write** I/O
low read rate => less **read** I/O
no core components (e.g. ACDOCA)
no usage in complex scenarios
no LOBs (known issues – fixed >= Rev. 55.00)

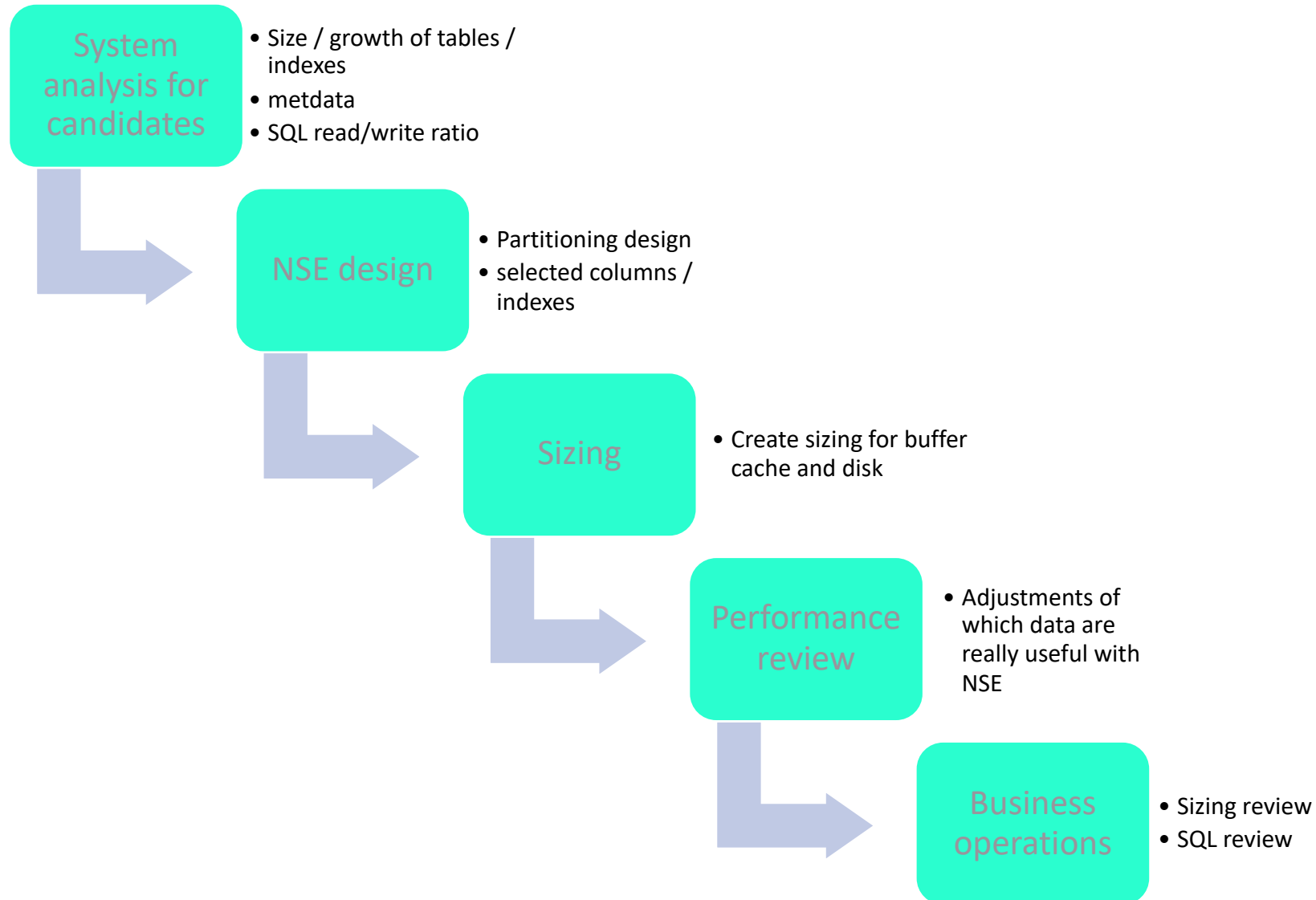


Example:

Change documents (CDPOS, CDHDR), Application Log (BALDAT), IDocs (EDID4), Archive Info Sources (ZARIX*).

Data Tiering - NSE

Implementation



Data Tiering - NSE

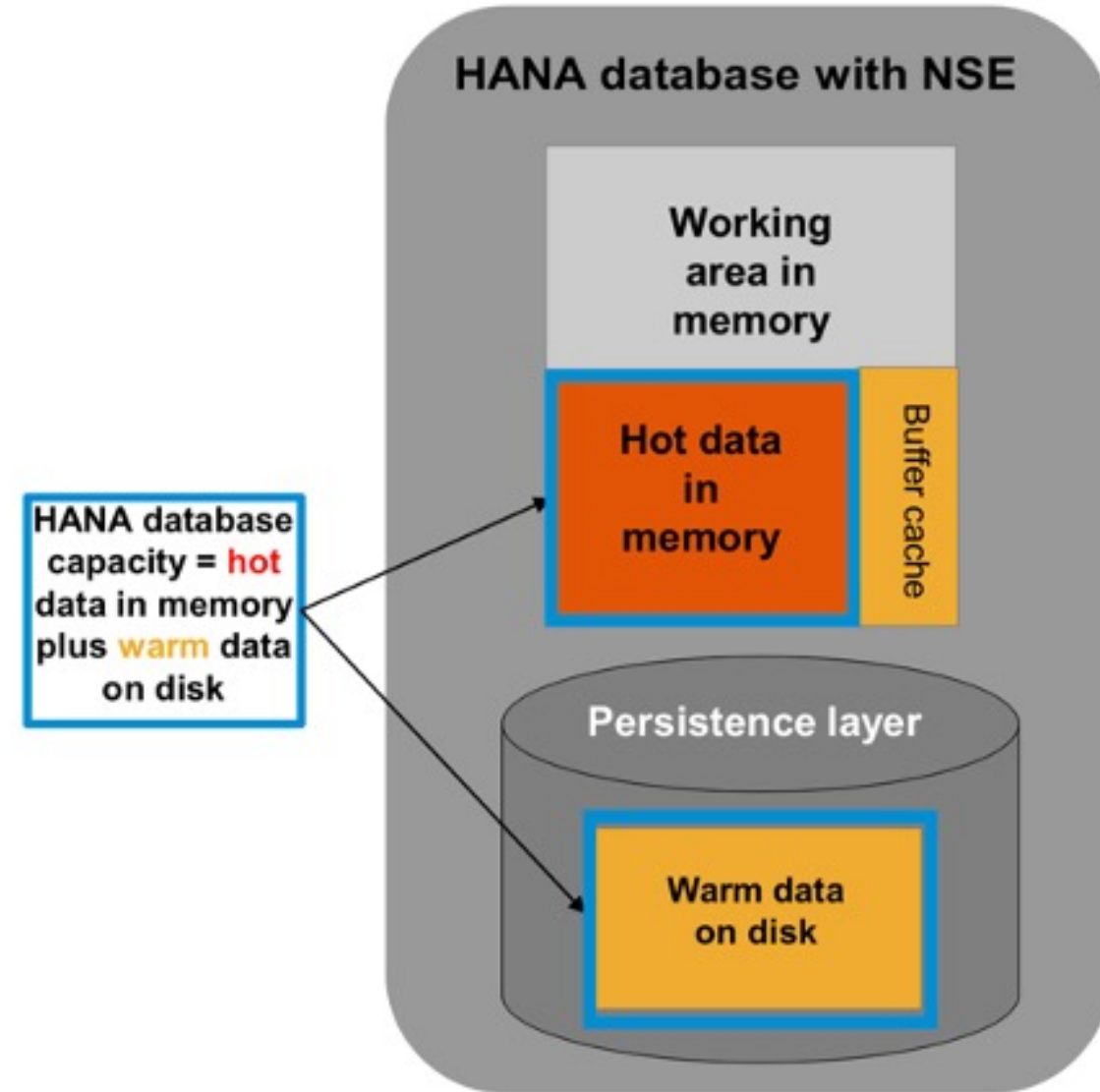
Implementation

Prerequisite

Min. HANA 2.0 SPS04

New persistence format :

```
SELECT * FROM M_CS_COLUMNS_PERSISTENCE
WHERE PERSISTENCE_TYPE = 'VIRTUAL_FILE' or
PERSISTENCE_TYPE = 'VIRTUAL_PAGED'
```



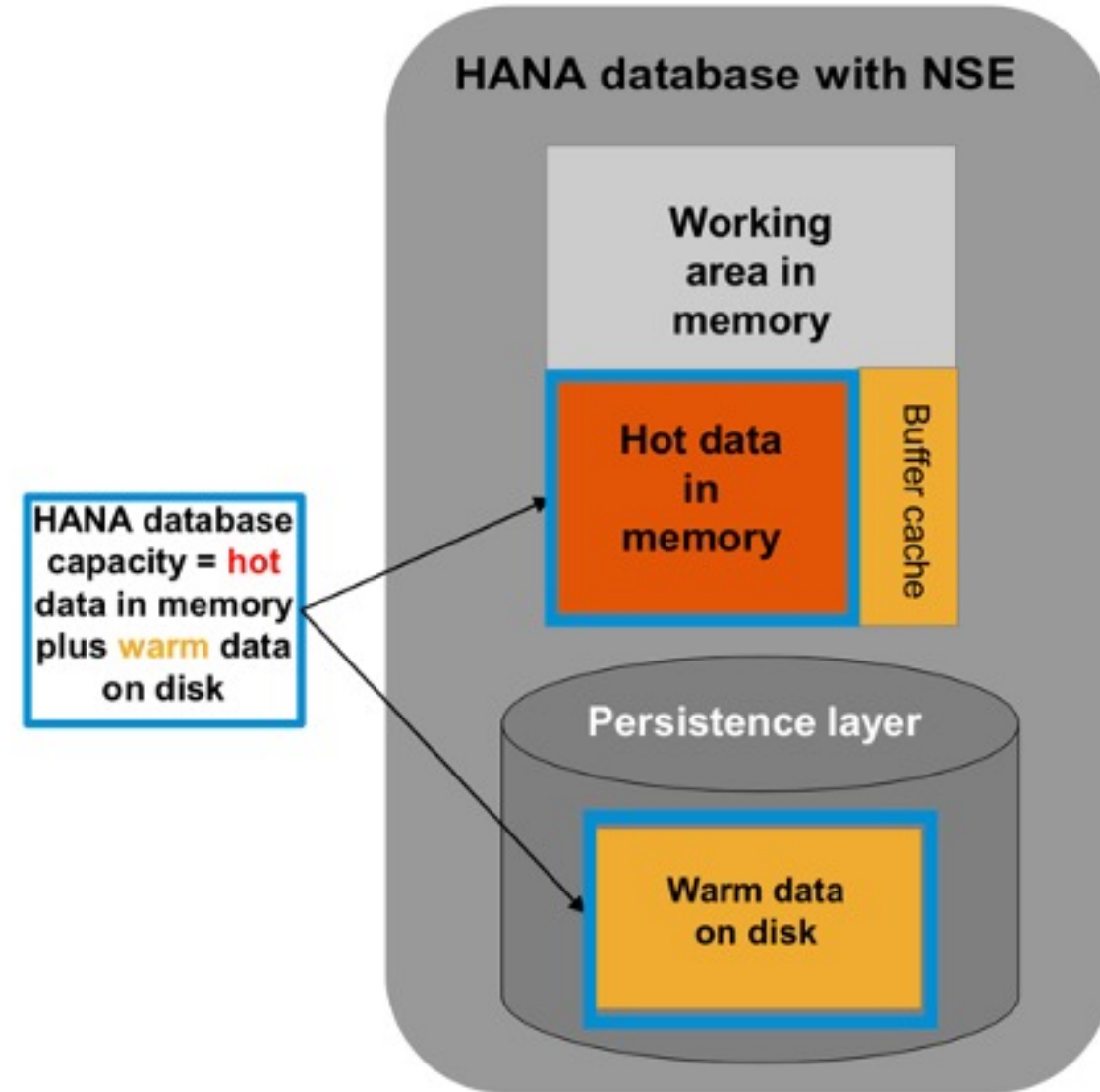
Data Tiering - NSE

Implementation

NSE support

By default buffer cache size: 10% of PAL*
 Activation via SQL Statement per Table, Partition, Column or Index
 Full ABAP DDIC support: ABAP Platform 2020 (SAP_BASIS 7.55) Partly ABAP DDIC support:

Software Component Version	Support Package
SAP_BASIS 740	SAPKB74025
SAP_BASIS 755	SAPK-75501INSAPBASIS
SAP_BASIS 750	SAPK-75020INSAPBASIS
SAP_BASIS 751	SAPK-75112INSAPBASIS
SAP_BASIS 752	SAPK-75208INSAPBASIS
SAP_BASIS 753	SAPK-75306INSAPBASIS
SAP_BASIS 754	SAPK-75404INSAPBASIS



*PAL: Process Allocation Limit => in this context of the hdbindexserver

Data Tiering - NSE

Implementation

Supported partitioning schemes

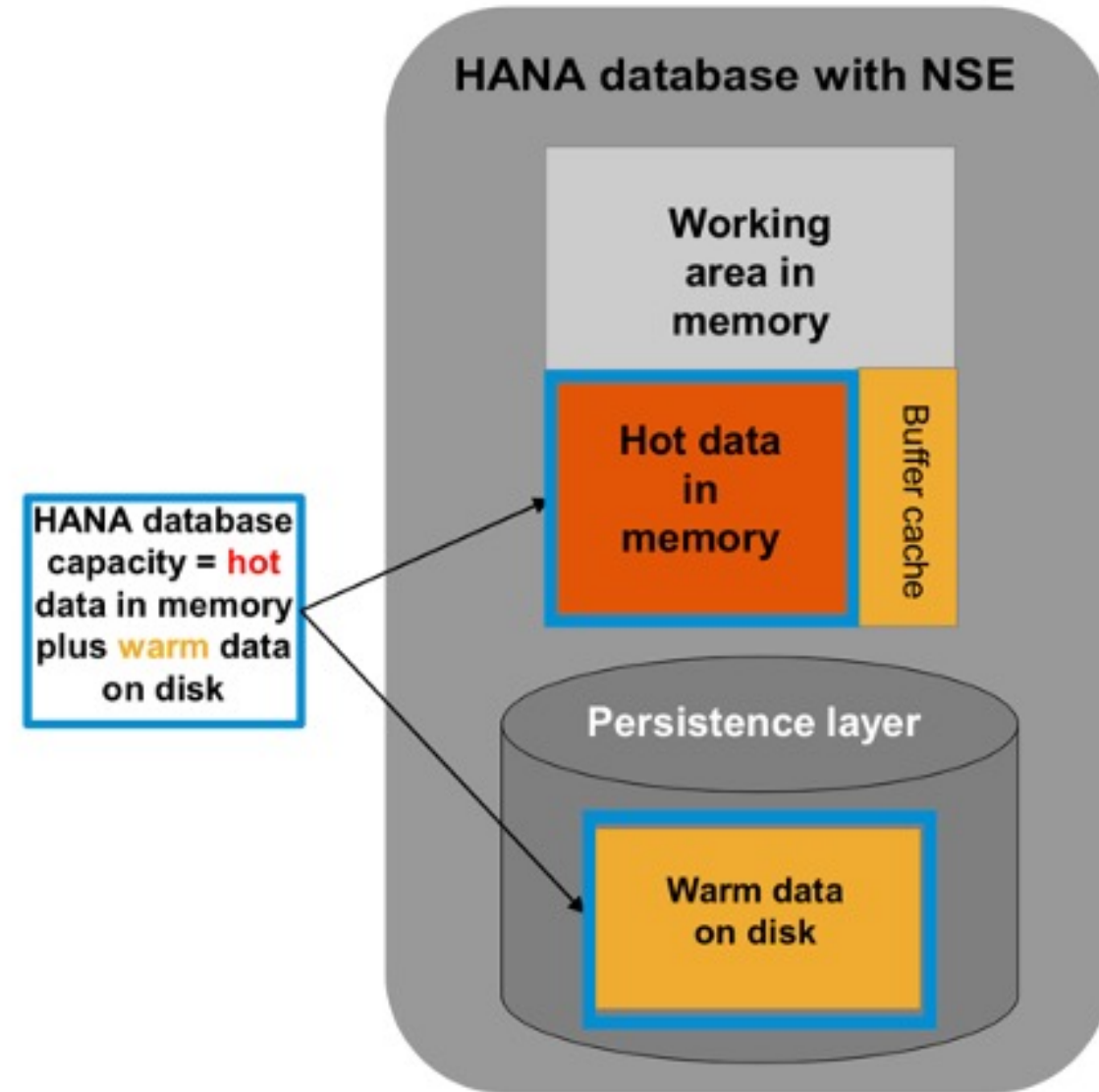
NSE supports the following partitioning scheme on **non-heterogeneous** partitions:

- Hash-range

NSE supports the following partitioning schemes on **heterogeneous** partitions:

- Range
- Range-range
- Range-hash

For all **other partitioning** schemes used in NSE tables, load units can be specified only on column, table, and index.



Data Tiering - NSE

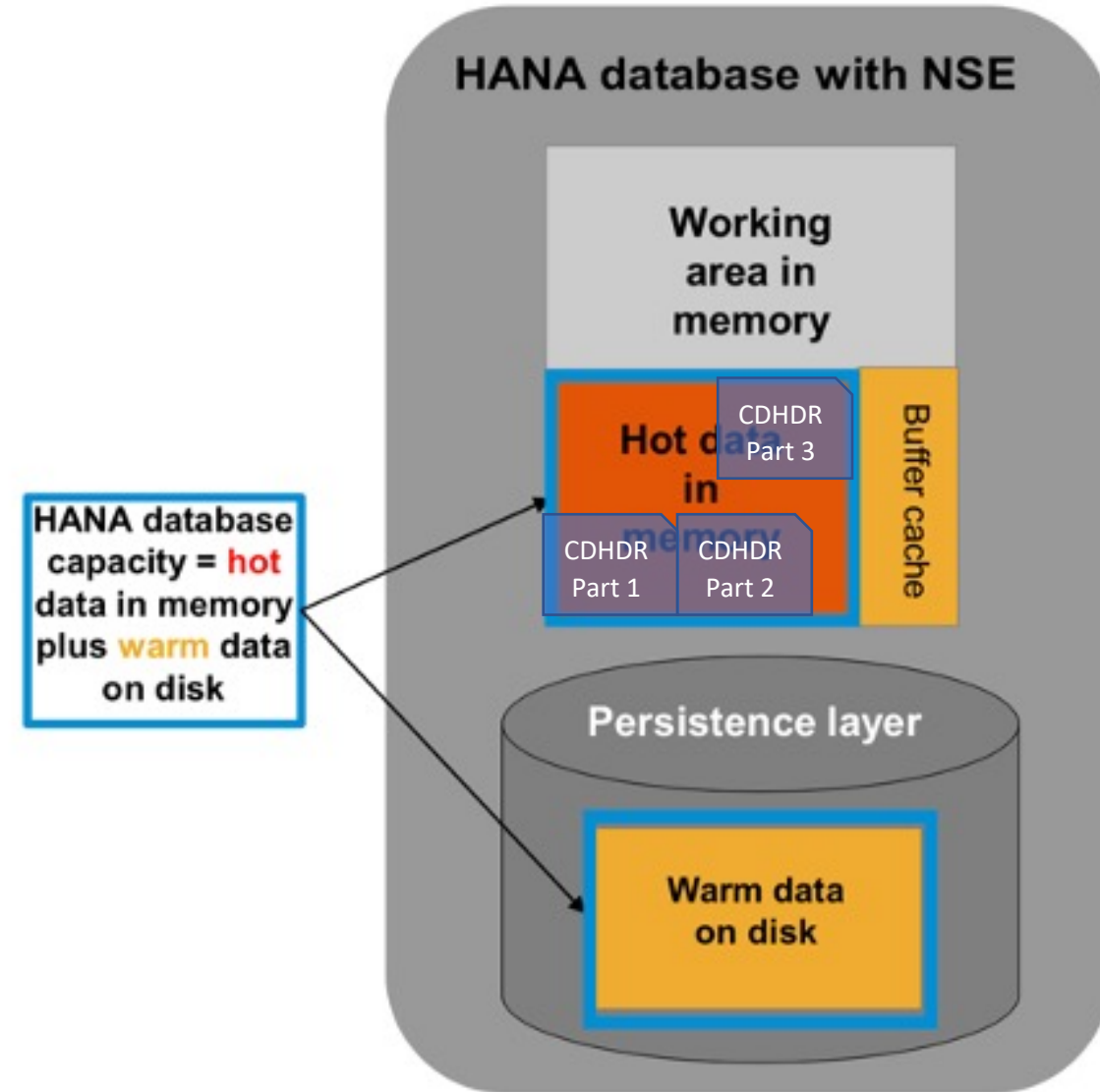
Implementation

Partitioning

Heterogeneous Partitioning:

```
ALTER TABLE CDHDR PARTITION BY
RANGE(UPDATE)
```

```
( (PARTITION '20170101' <= VALUES < '20180101',
PARTITION '20180101' <= VALUES < '20190101',
PARTITION '20190101' <= VALUES < '20200101',
PARTITION '20200101' <= VALUES < '20210101',
PARTITION '20210101' <= VALUES < '20220101',
PARTITION OTHERS
));
```



Data Tiering - NSE

Implementation

Note: Internal tests with changing the load_unit property of a table with 200 GB in size runs up to 15 minutes.

Activate NSE

Activation on **Table level** (recommended **offline**):

```
ALTER TABLE "<table_name>" PAGE LOADABLE [CASCADE];
```

Activation on **Column level** (recommended online):

```
ALTER TABLE <table_name> ALTER ("<column>" <data_type>
PAGE LOADABLE); (blocking concurrent changes)
ALTER "<table_name>" ALTER ("<column>" ALTER PAGE
LOADABLE); (non-blocking conversion)
```

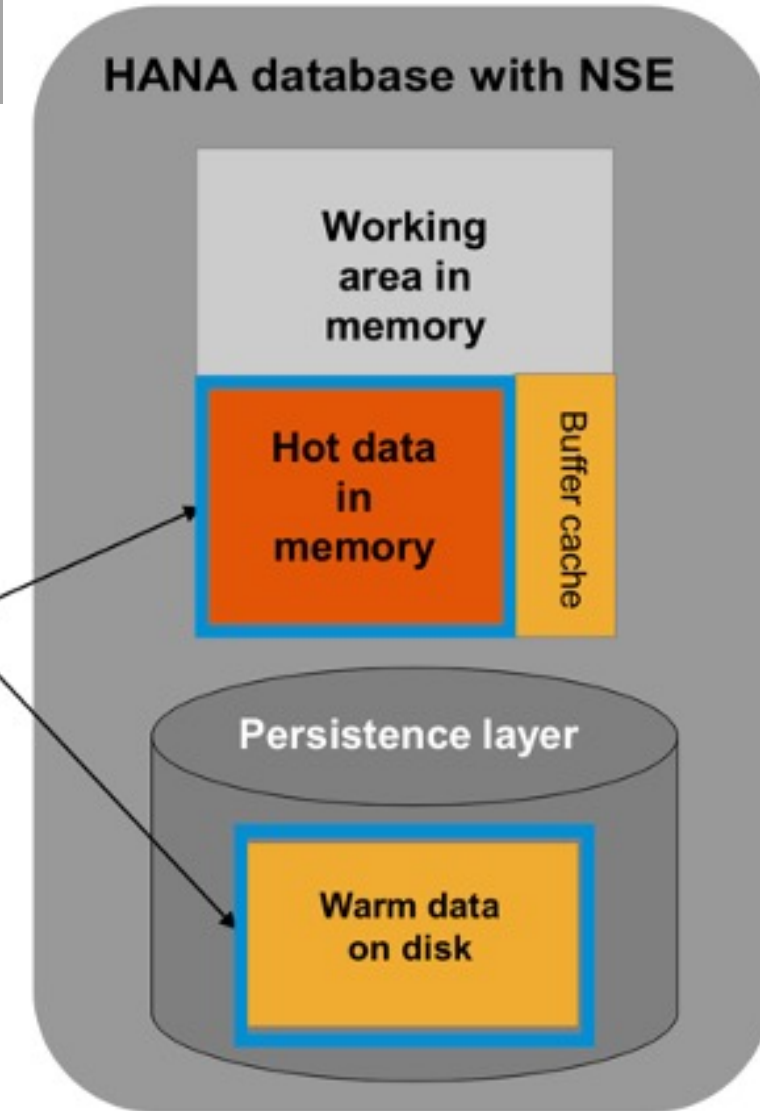
Activation on **Partition level** (recommended online):

```
ALTER TABLE <table_name> ALTER PARTITION <part_id>
PAGE LOADABLE;
```

Activation on **Index level** (recommended online):

```
ALTER "<schema_name>". "<index_name>" PAGE LOADABLE;
```

HANA database capacity = hot data in memory plus warm data on disk



Data Tiering - NSE

Implementation

Control NSE

```
SELECT * FROM M_CS_TABLES WHERE LOAD_UNIT = 'PAGE';
```

```
SELECT * FROM M_CS_ALL_COLUMNS WHERE LOAD_UNIT = 'PAGE';
```

```
SELECT * FROM M_CS_COLUMNS_PERSISTENCE;
```

```
SELECT * from M_BUFFER_CACHE_STATISTICS;
```

PORT	VOLUME_ID	CACHE_NAME	STATE	REPLACEMENT_POLICY	MAX_SIZE	ALLOCATED_SIZE	USED_SIZE	BUFFER_REUSE_COUNT	HIT_RATIO
30040	2	CS	ENABLED	IMPROVED LRU	16106127360	16106126208	15381664896	2052655	99.9974136352539

```
SELECT * from M_BUFFER_CACHE_POOL_STATISTICS;
```

CACHE_NAME	BUFFER_SIZE	REPLACEMENT_POLICY	GROWTH_PERCENT	TOTAL_BUFFER_COUNT	FREE_BUFFER_COUNT	LRU_LIST_BUFFER_COUNT	HOT_BUFFER_COUNT	BUFFER_REUSE_COUNT
CS	4096	IMPROVED LRU	1	244	0	206	38	60737
CS	16384	IMPROVED LRU	1	2791	2684	61	46	23806
CS	65536	IMPROVED LRU	1	2186	2079	36	71	30946
CS	262144	IMPROVED LRU	8	55667	1369	1788	52509	1529184
CS	1048576	IMPROVED LRU	8	1209	115	1038	56	407982
CS	4194304	IMPROVED LRU	1	0	0	0	0	0
CS	16777216	IMPROVED LRU	1	0	0	0	0	0

Side effects

Data Tiering - NSE

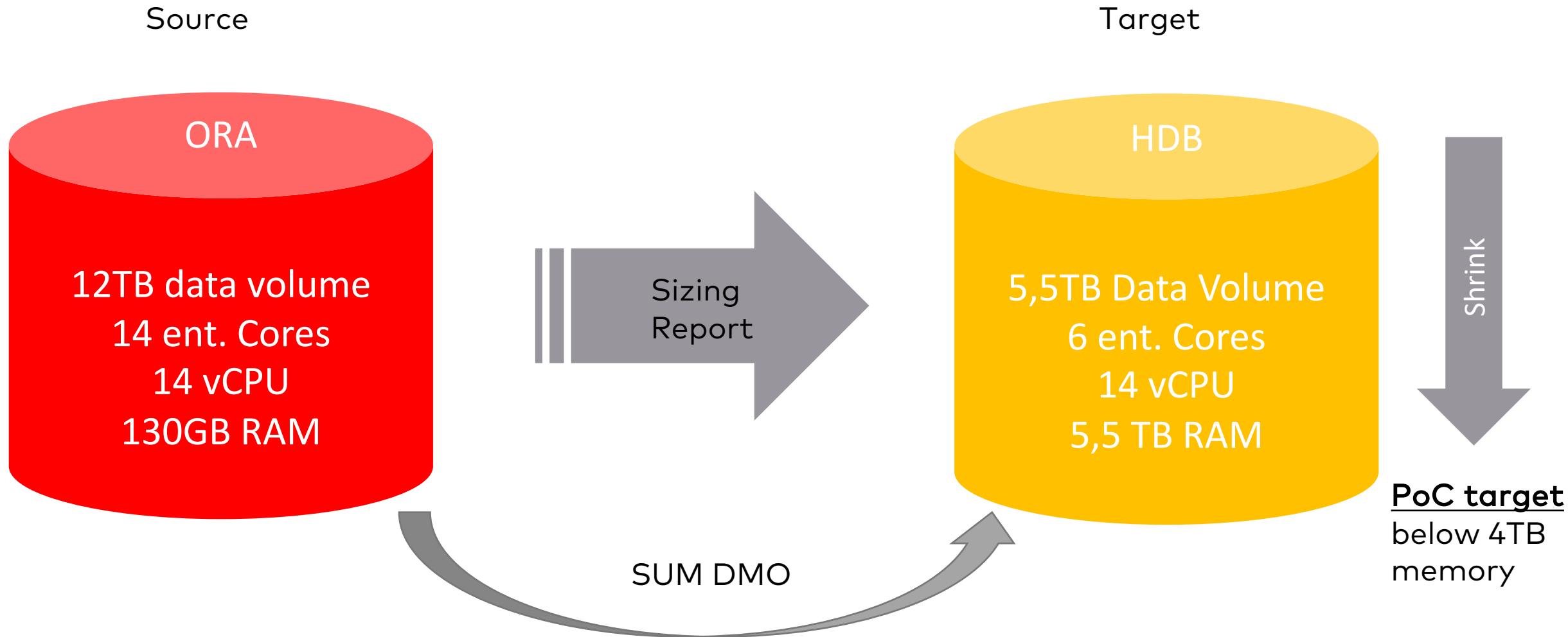
Benefits / side effects

Benefits	Side effects
Enormous potential for main memory savings (TCO)	more disk space needed
No changed behavior for end users (functionality)	Sizing including disk + cache must be constantly reviewed
no change for HA / DR or system copies	new coding has to be reviewed for high read rates of NSE tables
Coding changes are not required	
no additional infrastructure required (besides some more disk space)	

Project Review PoC customer industry: insurance

Data Tiering - NSE

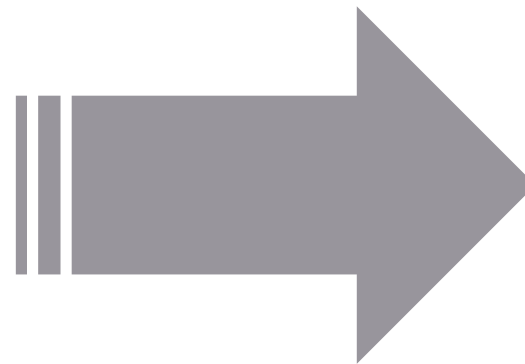
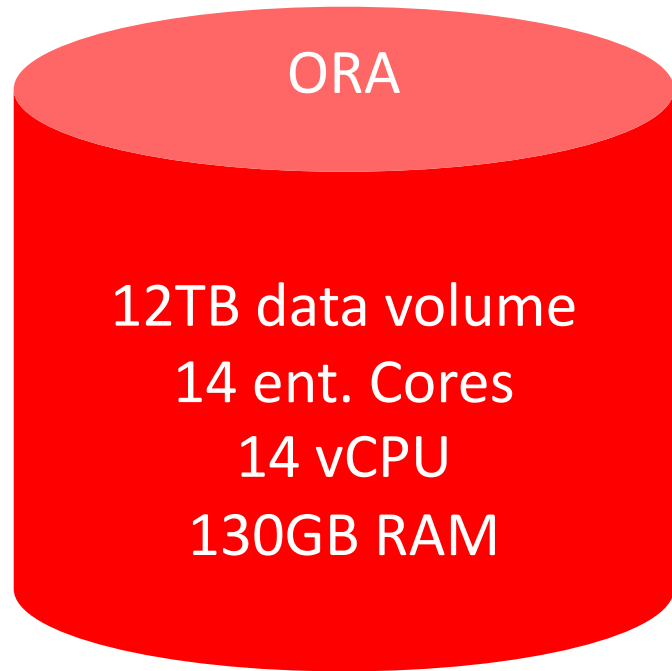
Project Review



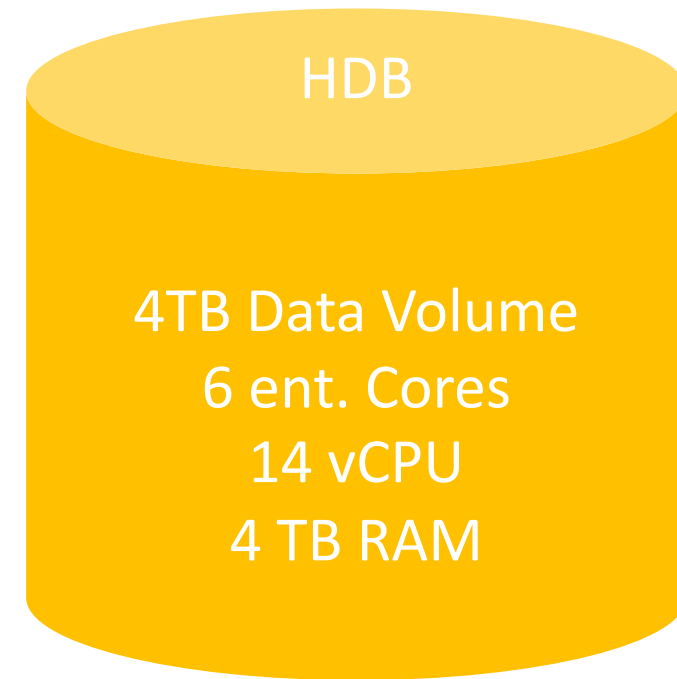
Data Tiering - NSE

Project Review

Source

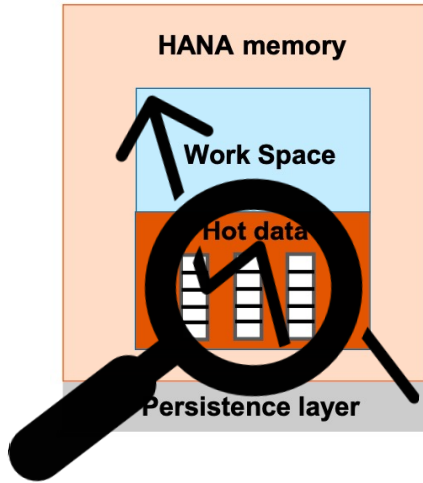


Target



Data Tiering - NSE

Project Review



- 12 tables are identified as possible candidates
- in total about 1TB memory (all 12 tables)
- Migration via SUM DMO without software packages
- Partitioning done before (phase preprocessing) migration (no additional downtime needed)
- Repeatable tests via Fallback Snapshot
- Review performance aspects
- Review duration

Data Tiering - NSE

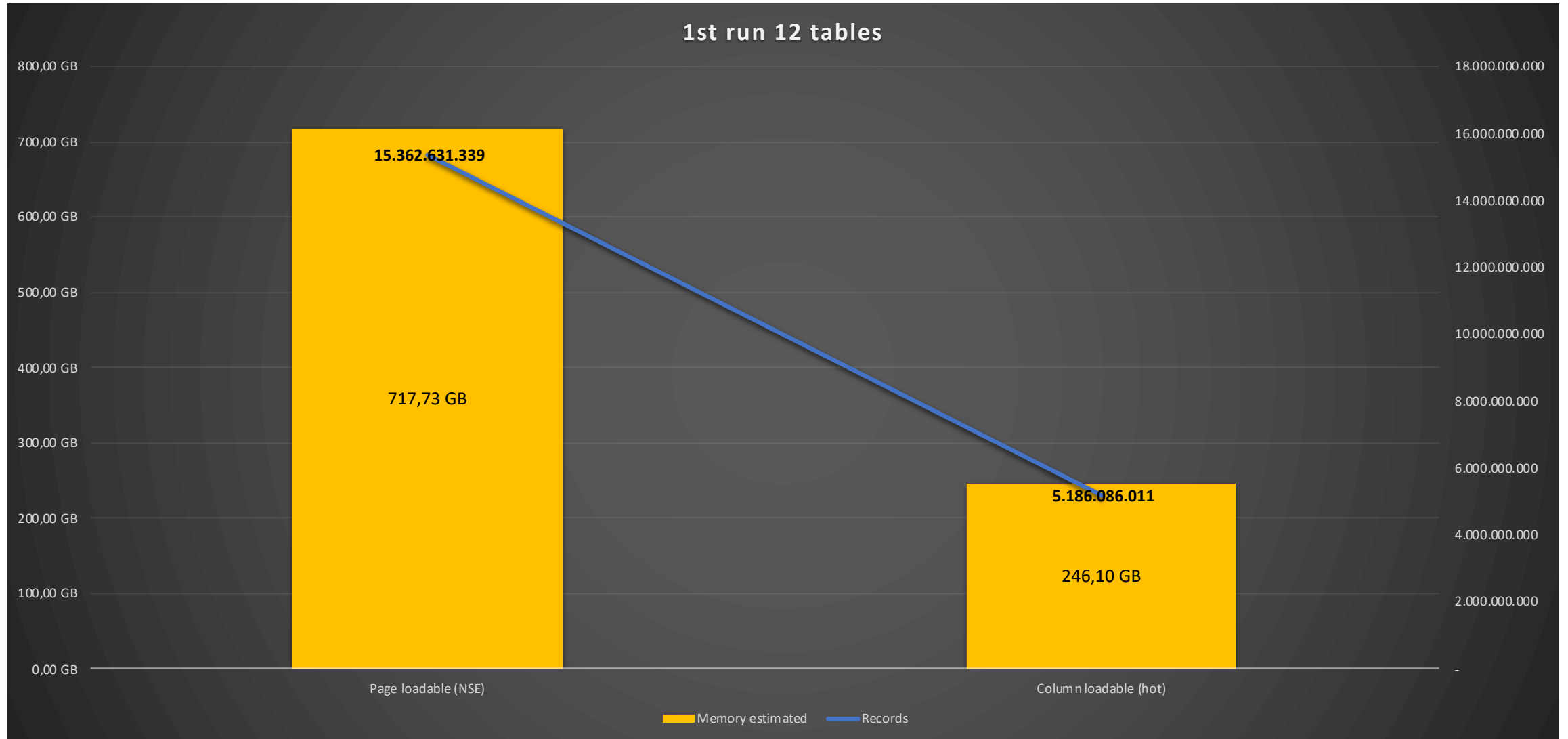
Project Review

Read / write ratio

TABLE_NAME	INSERT_COUNT	DELETE_COUNT	UPDATE_COUNT	REPLACE_COUNT	MERGE_COUNT	SELECT_COUNT	READ_WRITE_RATIO	LAST_MODIFY_TIME	LAST_SELECT_TIME
KONV	88.061.170	2.412.651	0	0	0	3.916.113	23,102964	13.04.2021 17:24:52.228199	19.04.2021 14:55:26.941172
WBRF	16.895.232	0	0	0	0	19.186	880,602105	13.04.2021 13:34:57.158052	13.04.2021 13:33:10.152985
BSIS	1.956.704	489.209	732	0	0	1.262	1.938,704437	07.04.2021 13:40:10.358711	26.03.2021 18:05:16.962331
BSAD	854.593	0	172	0	0	274.758	3,110974	13.04.2021 10:27:35.838888	13.04.2021 10:27:36.467528
WBRR	794.266	28.836	802.324	0	0	991.387	1,639547	13.04.2021 13:34:57.154694	16.04.2021 15:33:33.679695
CDHDR	730.050	1	0	0	0	417.498	1,748633	20.04.2021 12:21:22.861674	20.04.2021 14:22:20.185403
WBRK	648.028	0	16.039.797	0	0	42.042.952	0,396923	19.04.2021 20:28:48.105788	20.04.2021 08:23:14.250068
BSAS	369.111	0	0	0	0	0	?	07.04.2021 13:39:49.088579	?
/ABT_DUPLI	345	0	0	0	0	200	1,725	12.04.2021 15:59:57.663977	12.04.2021 15:59:55.639775
/AB_PASSAT	0	0	0	0	0	7.391	0	?	26.03.2021 11:45:13.5978
/AB_MAUTVL	0	0	0	0	0	274	0	?	26.03.2021 11:38:56.845241
/AB_ABS_DE	0	0	0	0	0	9.472	0	?	13.04.2021 11:57:22.647562
/AB_MAUTU	0	0	0	0	0	15.110	0	?	13.04.2021 11:57:22.649333

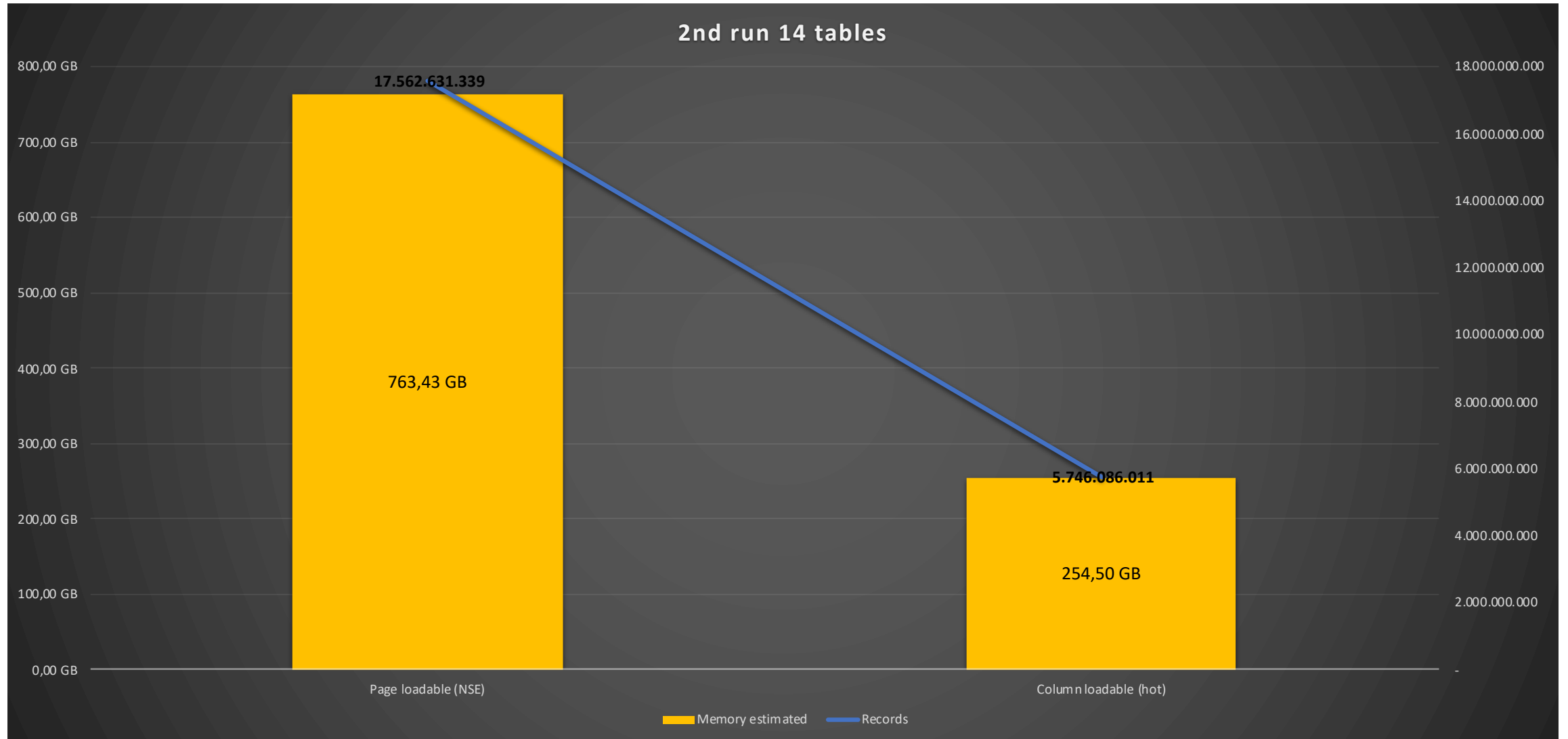
Data Tiering - NSE

Project Review



Data Tiering - NSE

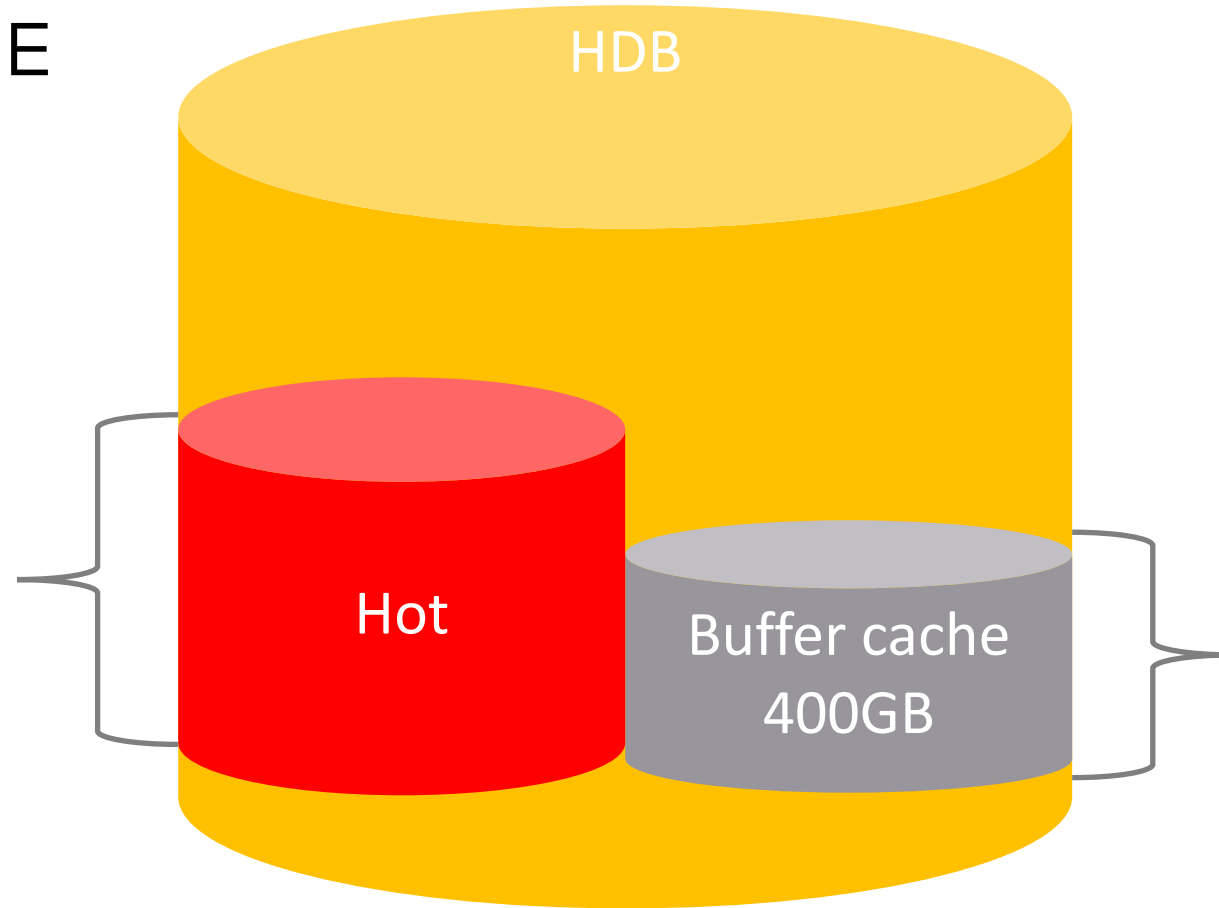
Project Review



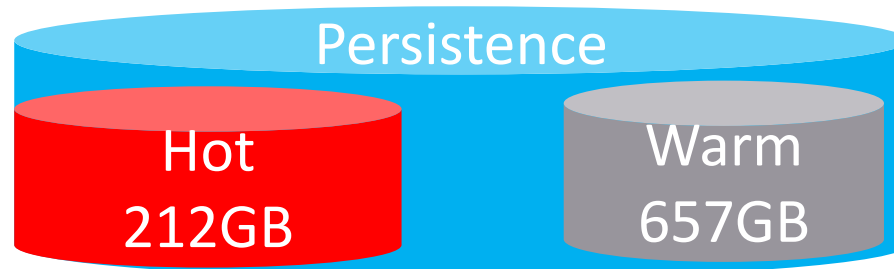
Data Tiering - NSE

Project Review

Tables type
Column:
246 GB RAM



Tables type
PAGE:
717GB RAM



Only the involved tables are shown

Data Tiering - NSE

Project Review

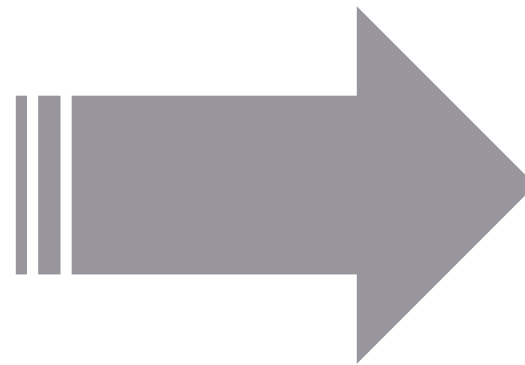
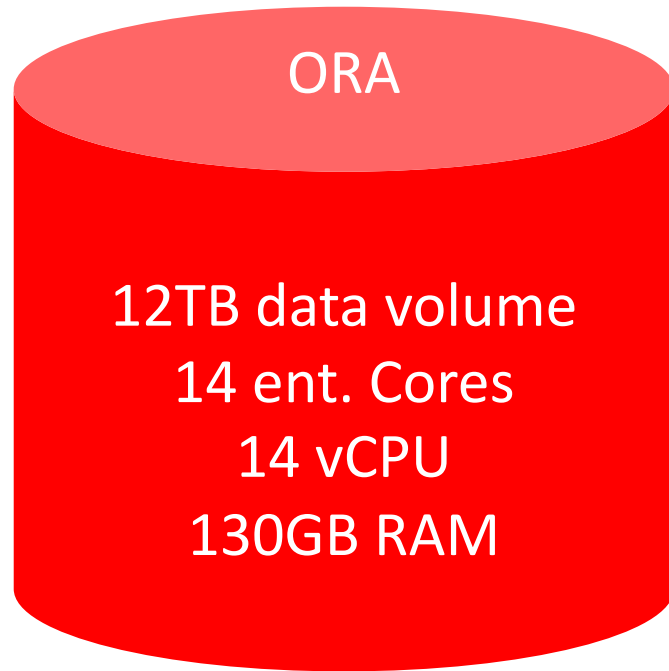
Disk	default	hot	warm
Disk	787 GB	212GB	657GB
Sum	787 GB	869GB	
Savings		+ 82 GB	

Memory	default	hot	warm
memory	951 GB	56,7GB	400GB (*limited by buffer cache)
Sum	951 GB	456,7 GB	
Savings		- 494,3GB	

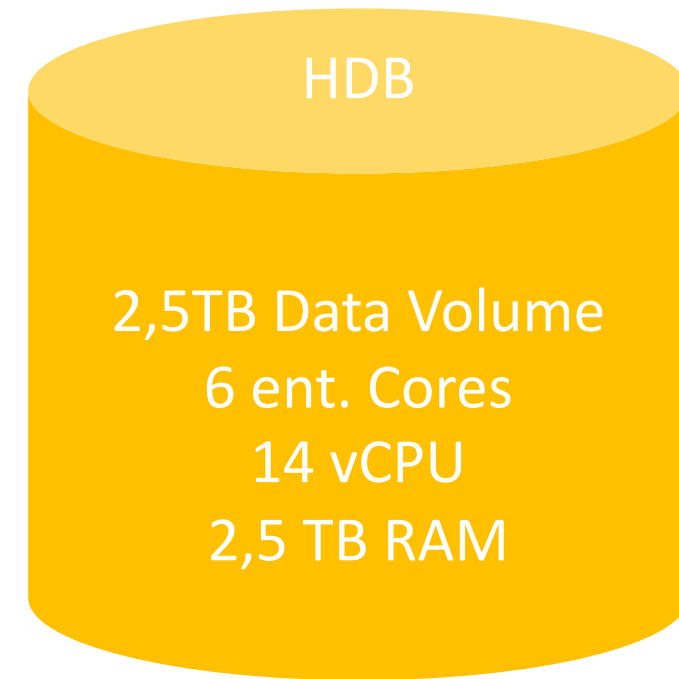
Data Tiering - NSE

Project Review

Source



Target



Data Tiering - NSE

Project Review

Summary:

- 0% Performance impact in test scenario ,billing run'
- 100% hit ratio of buffer cache => can be reduced
- More iterations needed to get the full benefit
- Savings about 500GB RAM => over 10% of HANA size
- Operation concept have to be adjusted => monitoring, partitioning, etc.

Project Review

Project implementation

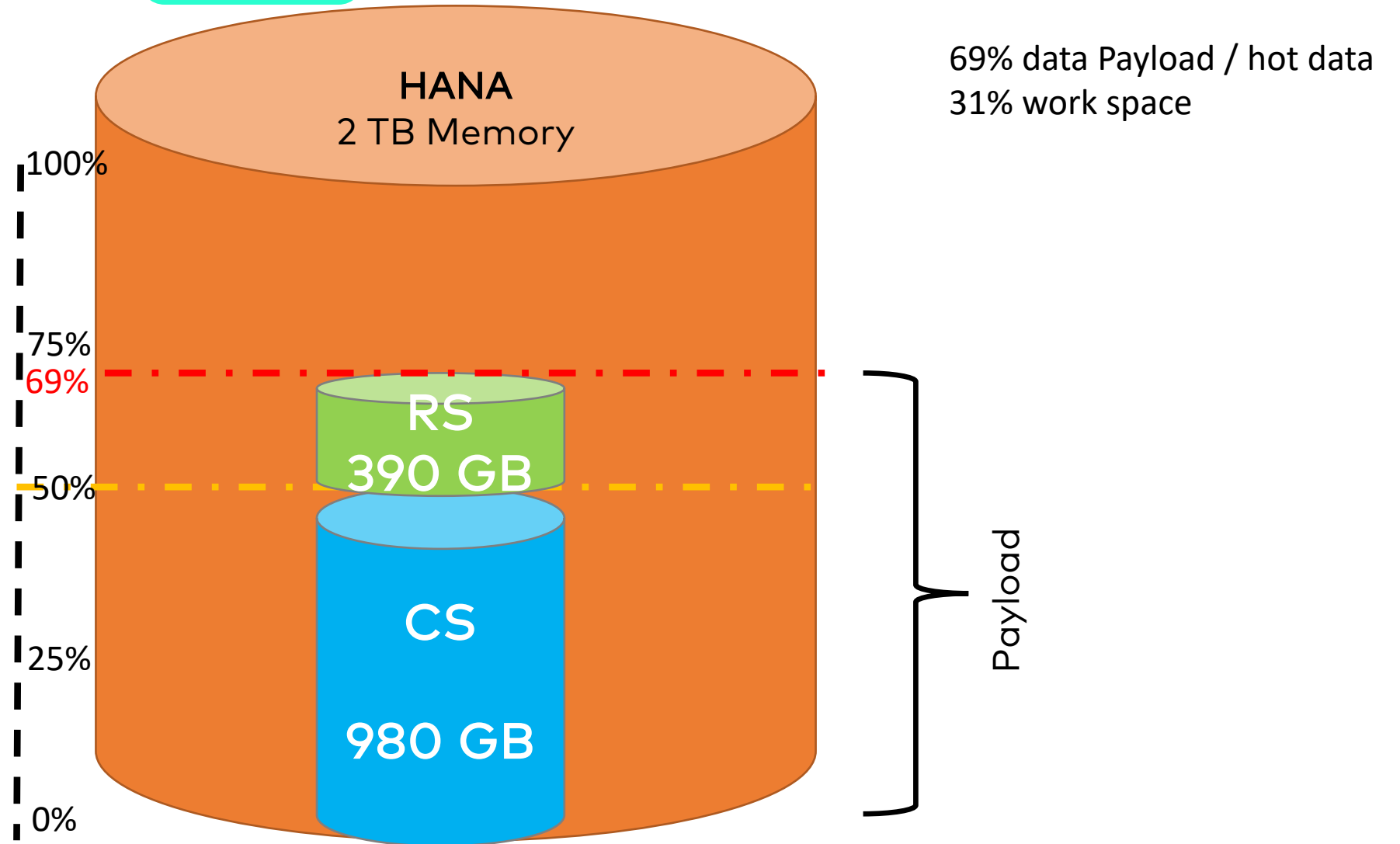
customer industry: retail

Data Tiering - NSE

Project Review

NSE design

- Partitioning design
- selected columns / indexes



Data Tiering - NSE

Project Review

NSE design

- Partitioning design
- selected columns / indexes

Attribute	FAGLFLEXA	CFIN_ACCCR	FAGL_SPLINFO_VAL	BSIS	BSEG
Records	4.931.709.995	2.122.091.327	1.567.028.839	746.143.847	870.499.889
loaded Records	822.744.984	708.167.227	522.804.806	746.143.847	870.499.889
loaded in %	16,68	33,37	33,36	100,00	100,00
Indexes	157.000,0MB	52.000,0MB	33.500,0MB	20.000,0MB	30.000,0MB
CS	139.000,0MB	36.000,0MB	18.000,0MB	27.000,0MB	26.300,0MB
Savings Cols	74.000,0MB	36.000,0MB	18.000,0MB	20.000,0MB	24.000,0MB
Savings Indexes	00,0MB	00,0MB	26.800,0MB	00,0MB	00,0MB
Records saved with NSE older than 2021	2.700.000.000	-	1.300.000.000	640.000.000	745.000.000
Saved by NSE Partitioning	118.937MB	35.991,0MB	15.363MB	23.073MB	22.524MB
Sum	296.000,0MB	88.000,0MB	51.500,0MB	47.000,0MB	56.300,0MB

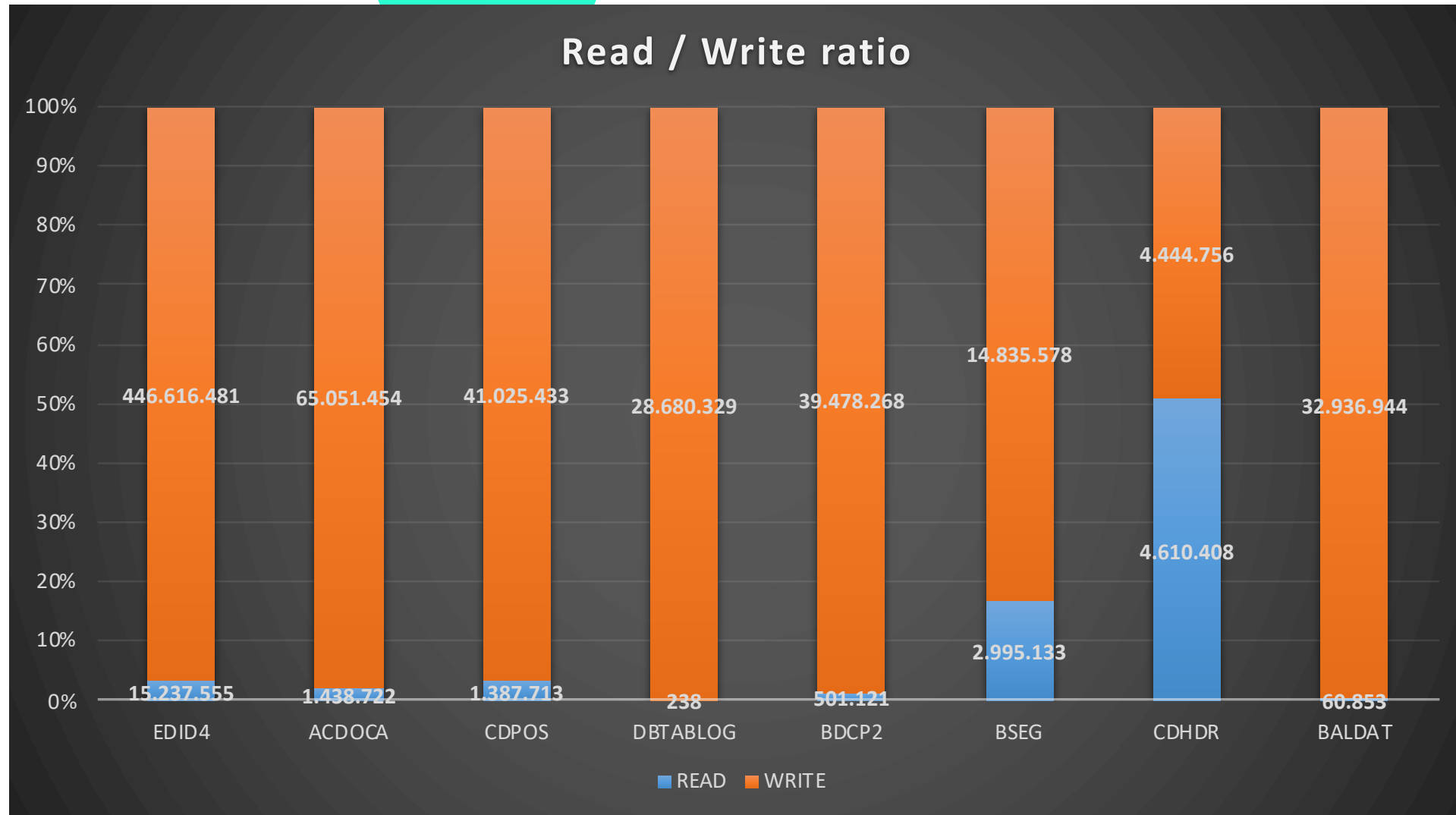
must be tested carefully
less frequently accessed

Data Tiering - NSE

Project Review

NSE design

- Partitioning design
- selected columns / indexes



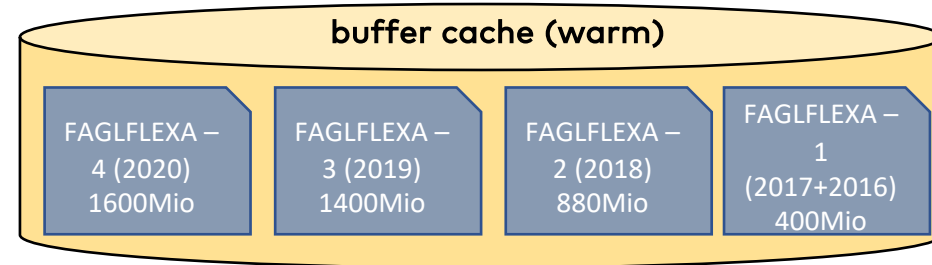
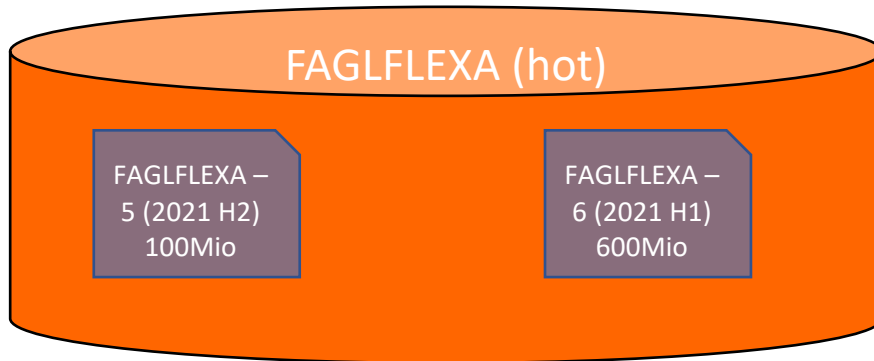
Data Tiering - NSE

Project Review



- Partitioning design
- selected columns / indexes

TABLE NAME	2016	2017	2018	2019	2020	2021	TOTAL
FAGLFLEXA	3.502.320	402.745.224	880.181.174	1.386.626.400	1.547.568.744	711.086.133	4.931.709.995
	Part 1		Part 2	Part 3	Part 4	Part 5/6	

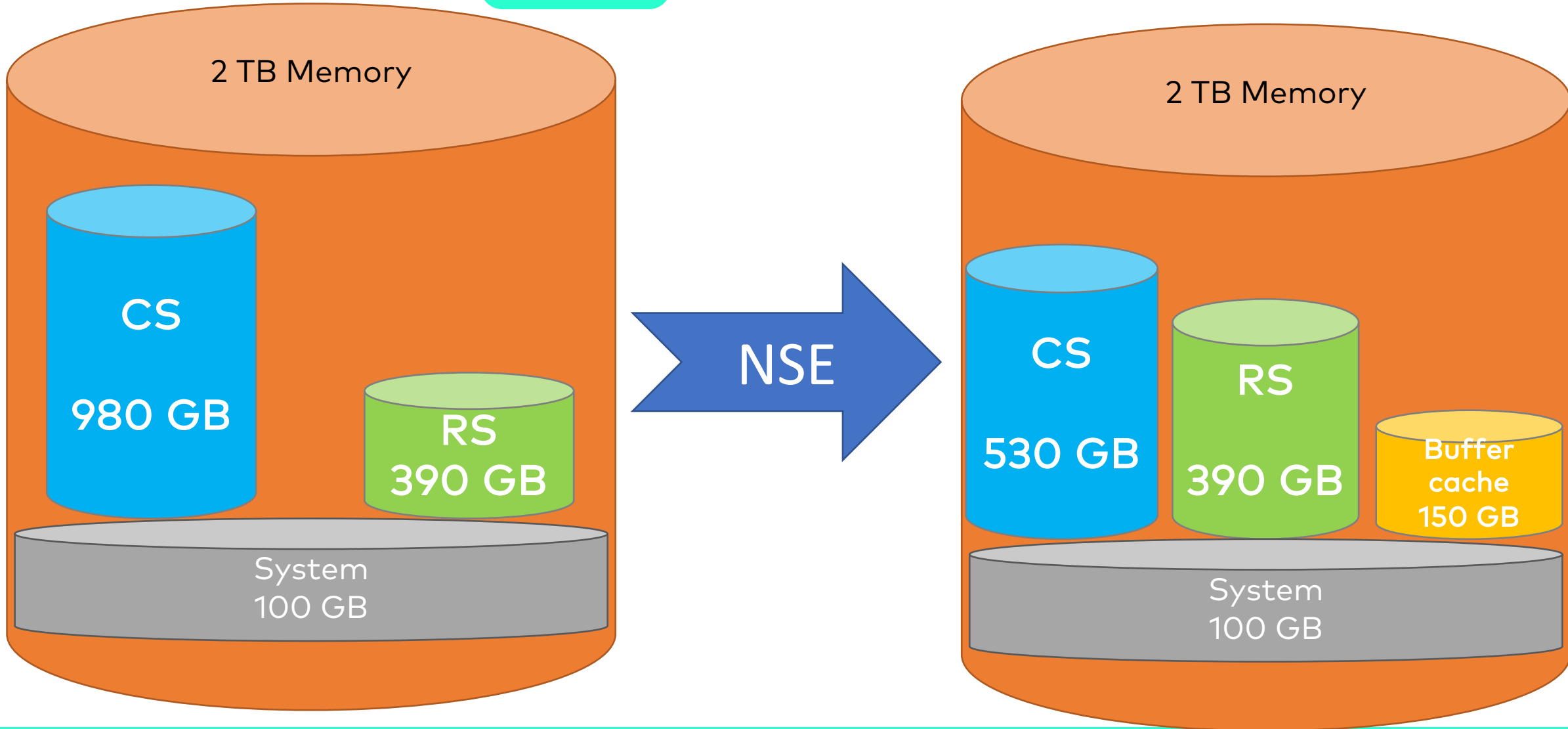


Data Tiering - NSE

Project Review

NSE design

- Partitioning design
- selected columns / indexes



Data Tiering - NSE

Project Review

Summary:

- **-10%** performance degradation due heterogeneous partitioning for selects on BKPF and BSEG
- Some partitions have to be moved back from warm to hot store (year 2020) due business scenarios => **no** performance degradation any more
- **10-20%** improvement of inserts, updates, deletes due to partitioning design
- **30-40%** improvements for delta merges in hot store due partitioning design

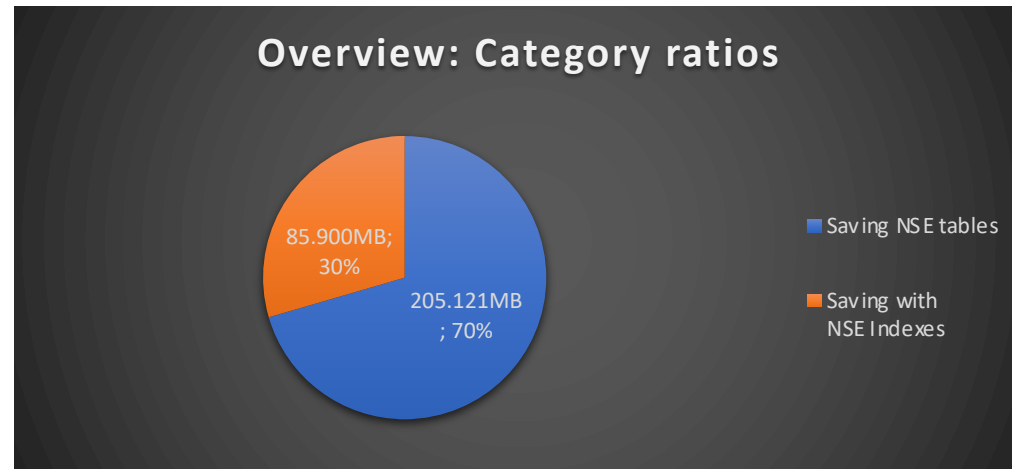
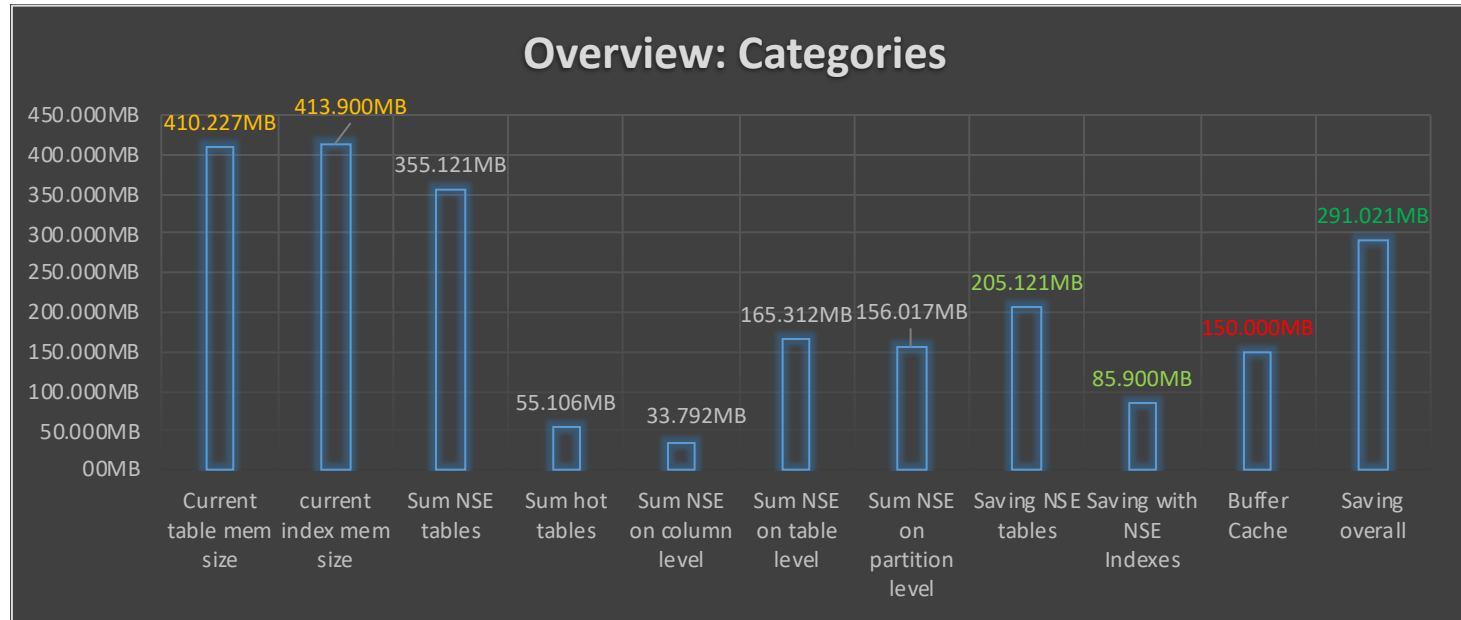
Data Tiering - NSE

Project Review

Summary:

- Overall memory savings:
300GB Payload
 = 600GB Peaksizing
 (incl. Work space)

exact numbers **450GB** savings
 = 900GB Peaksizing (incl. Work space)
 - 150GB CS Buffer Cache
 750GB savings



Take-aways

Data Tiering - NSE

Take-aways

1. No further licenses needed
2. Use NSE if savings are over 10% (=20% sizing value) of system size
3. NSE potential 20-30% for every SAP ERP system
4. Not NSE **or** archiving => NSE **AND** archiving
5. Proper system analysis is fundamental
6. Proper partitioning concept and monitoring is elementary
7. There is no reason not to use it

Data Tiering - NSE

Pitfalls

1. Wrong **partitioning design** => performance issues
2. Wrong **candidates** with too many reads => performance issues
3. Wrong **sizing for buffer cache** => too many unloads => performance issues
4. Wrong **data sizing on disk** => system availability
5. No frequent **monitoring of cache hit ratio** => too many unloads
6. No frequent **monitoring of partitions** => too much data in main memory



APPENDIX

Quellenangaben

2799997 - FAQ: SAP HANA Native Storage Extension (NSE)

2915190 - NSE: feature not supported: modifying partition property for non-HETEROGENEOUS partitioned table

2973243 - Guidance for use of HANA Native Storage Extension in SAP S/4HANA and SAP Business Suite powered by SAP HANA

[NSE Whitepaper](#)