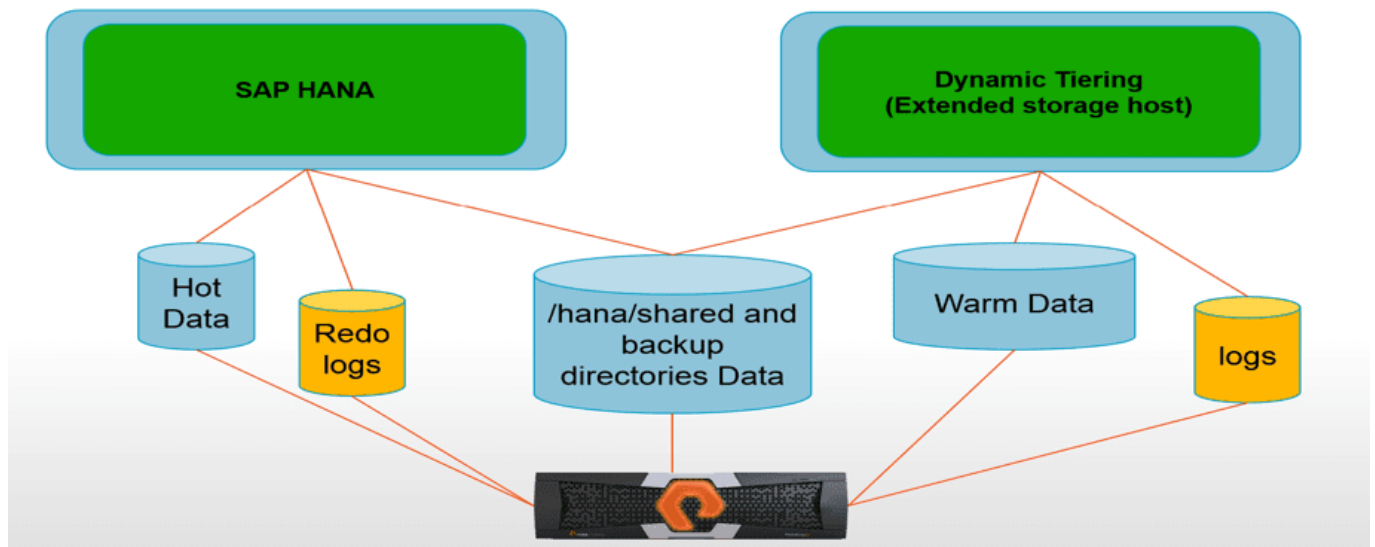


SAP HANA/Dynamic Tiering Calculation View's Pure Performance on Flash Storage



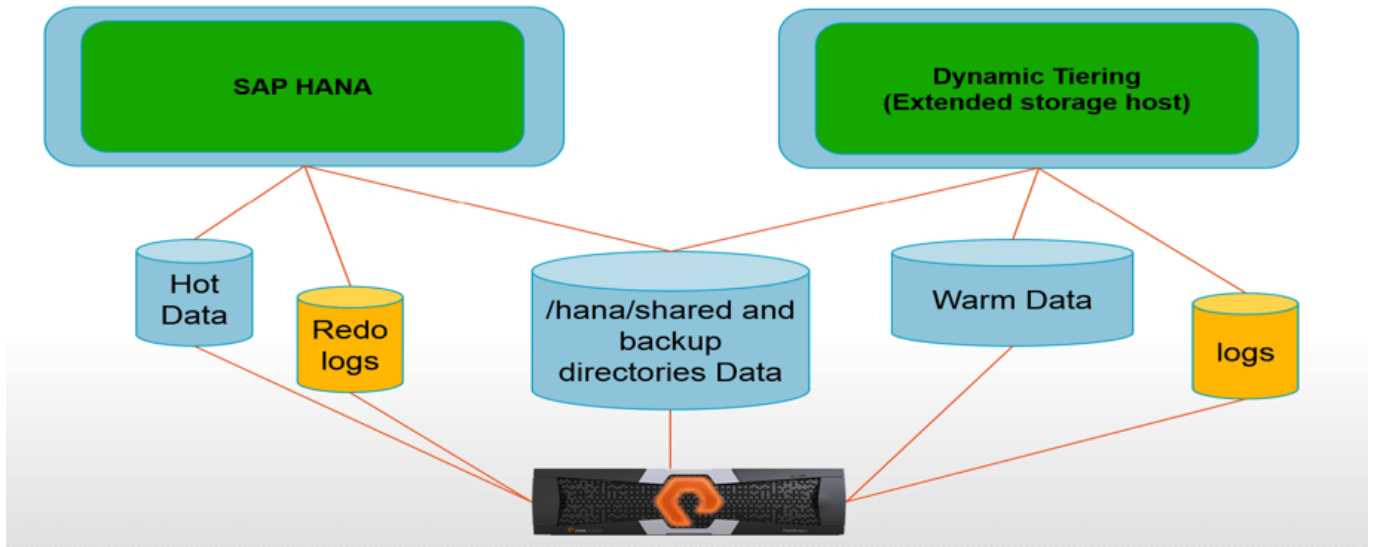
In this blog, I am going to explain how I tested SAP HANA dynamic tiering at SAP COIL lab Palo Alto. It has been an extremely interesting experience to see how fast data models were on dynamic tiering based on Flash storage, which, in this case was Pure Storage®.

SAP HANA is an in-memory data platform. The value of in-memory computing is proven with many live customers. But RAM is a very expensive resource, SAP HANA has tools to help customers lower cost by moving data to storages based on data temperature. The hot data will always reside in main memory, while warm and cold data can be moved to various storage devices. But if storage is very fast like Flash storage, the performance drop of queries runtimes is quite negligible.

This lowers the overall cost of adopting a SAP HANA solution when warm/cold data can be moved to Dynamic tiering based on Flash storage and there is also not much loss of performance in terms of business queries' runtimes.

What are the benefits of dynamic tiering?

Dynamic tiering allows data-aging rules to be defined by administrators, and to be executed automatically at scheduled intervals.



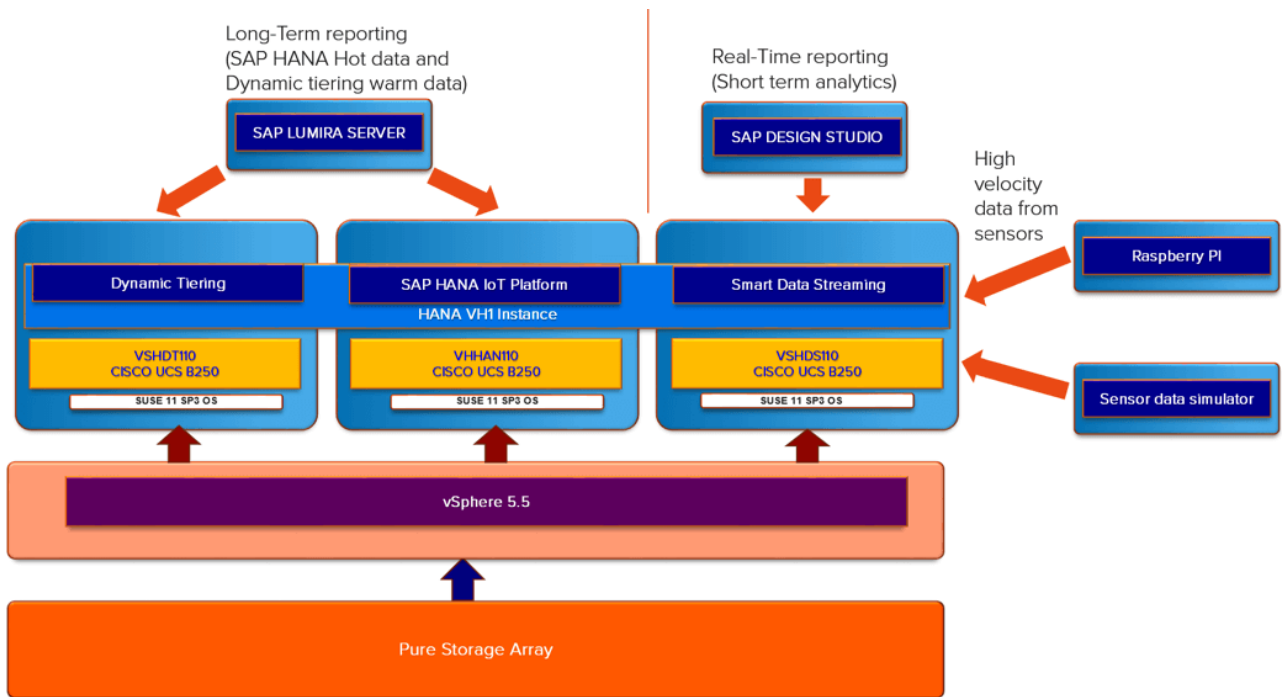
System Landscape

SAP HANA SPS11, and SAP HANA dynamic tiering; SAP BusinessObjects Design Studio; SAP HANA Cloud Platform, SAP Smart Data streaming; SAP BusinessObjects Lumira, server version for teams; and Pure Storage FlashArray//m.

IoT scenario to test the performance of Dynamic tiering on Flash

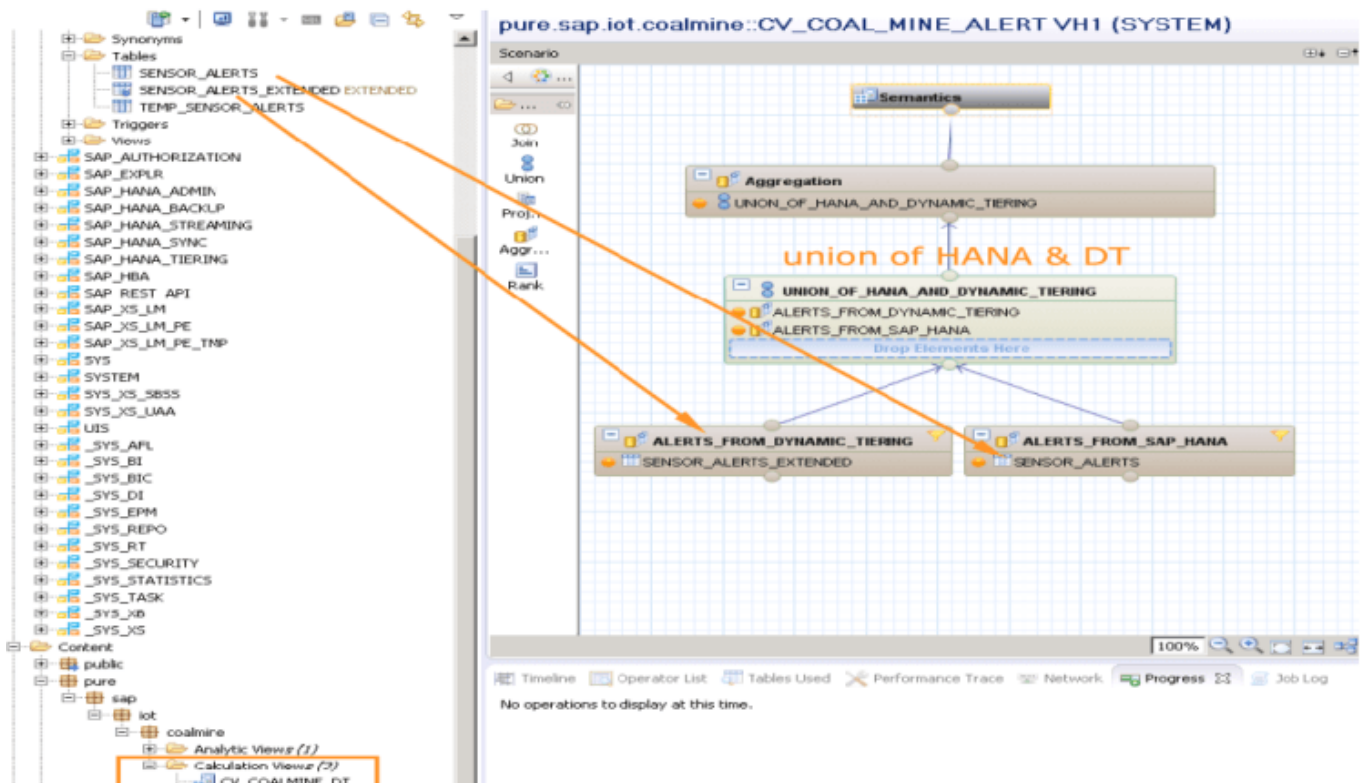
IoT solutions generate a large amount of data and it is the high-velocity data which needs to be collected and analyzed at real-time. I implemented an IoT application that demonstrates a coal mining company leveraging IoT applications for a ground-level alert system. Coal mines are complex work environments where the safety of workers is of utmost priority; a number of precautions have to be in place to avoid any mishaps. To aid this process, various networked sensors monitor the environment closely, reporting any deviation from an ideal situation.

As illustrated in the following diagram, the high-velocity data is collected by SAP Smart Data Streaming for doing real-time analysis and then sent to SAP HANA. SAP Smart data streaming issues and captures alert when the system senses smoke, fire, earthquake, equipment failure, and so on. Also, short-time or real-time analytics are performed on SAP Smart data streaming. The long-term analytics are done on SAP HANA and Dynamic tiering which has data collected over a period of time.



The Data model

In order to test the performance of dynamic tiering for analytical models, I created tables in SAP HANA called **SENSOR_ALERTS** (300 bytes) and an extended table -**SENSOR_ALERTS_EXTENDED**(same structure as **SENSOR_ALERTS**) on dynamic tiering. On these tables, I created a calculation view(OLAP model) which basically performs a union of the SAP HANA table and the dynamic tiering table.



On this data model (calculation view) I created OLAP queries which were for SAP Lumira dashboard.

Performance Testing

Initially, all the data is present in SAP HANA table (SENSOR_ALERTS) as SAP Smart data streaming dumps the data into SAP HANA. I populated 500 million records in SAP HANA table (SENSOR_ALERTS) and no data was present in dynamic tiering (SENSOR_ALERTS_EXTENDED). All the OLAP queries were doing 500 million records scan. But initially all the data was present in SAP HANA and then I started moving chunks of data in batches of 25% (or 125 million records) to the Dynamic tiering table.

The runtimes of the queries to my surprise did not drop much as queries started fetching data from dynamic tiering. This is obviously due to the fact the dynamic tiering tables are present in this case on an extremely fast flash storage like Pure storage FlashArray//m.

I picked one of the analytical queries to show the query runtimes and comparing with the query runtimes when all the data is present in the memory (SAP HANA) to the different cases as more and more data was moved to Dynamic tiering. For more details regarding this analysis look at the white paper published by SAP.

Below is the graph which shows the performance of a query which does 500 million records scans against the baseline. The baseline is when all the data for this query is present in SAP HANA and no data is present in Dynamic tiering.

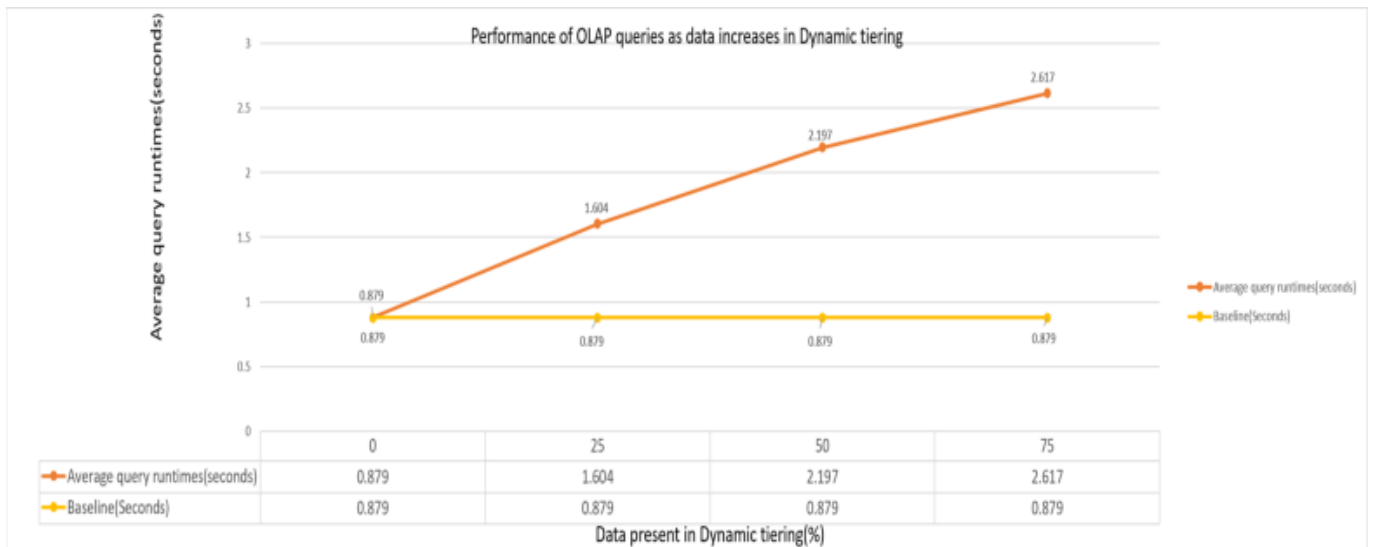
Below graph shows four data points

100% data (500 million records) in SAP HANA and 0% data in dynamic tiering — Baseline query performance

75% data (375 million records) in SAP HANA and 25% data(125 million records) in dynamic tiering

50% data (250 million records) in SAP HANA and 50% data (250 million records) in dynamic tiering

25% data (125 million records) in SAP HANA and 75% data (375 million records) in dynamic tiering



As seen above the drop in query performance is marginal, it only slows down by 0.5-0.7 seconds for every 125 million records moved to dynamic tiering from SAP HANA. This shows how good the query performance can be for the data models even when a lot of data is moved to Dynamic tiering. This helps in reducing the memory footprint of SAP HANA and helps in saving millions of dollars on SAP HANA licensing.

Here are the other benefits of deploying Pure Storage for SAP HANA and dynamic tiering applications:

- Data reduction: Reduces the storage footprint of SAP HANA and Dynamic tiering
- Accelerated warm data readiness: Fast movement of data between SAP HANA and dynamic tiering
- Cost and performance analysis: With a little or marginal loss in performance how much SAP HANA licensing money can be saved.

[You can refer to the white paper published by SAP for more details about all the benefits.](#)