COMPARE 3 DWH ARCHITECTURES

LUKAS SCHULT, LUANA JUHL

COMPARISON TABLE

	DW Only	DM Only	DW & DM	Virtual DW	Explanation
Administration expenses	+	-	0	-	Run & Manage/Maintain cost
Implementation	-	+	--	++	Cost and time needed to implement the Architecture including all components and necessary hardware
Performance	0	+	+	+	Speed at which Data can be accessed
Size	>100 GB	<100 GB			

COMPARISON TABLE

	DW Only	DM Only	DW & DM	Virtual DW	Explanation
Flexibility	-	+	0	++	Adaptability in response to changes
History	+	+	+	--	Storage of historical data used to determine data trends
Security	-	+	0	+	Management of access to data stored in the data model
Data Quality & Consistency	+	-	+	-	

IMPLEMENTATION

DW Only	DM Only	DW & DM	Virtual DW
• Big amounts of data ➢ Large amounts of resources necessary ➢ Implementation process more complex • Months – Years • High cost	• Less data ➢ Implementation more simple • Days - Months • Cost efficient	• DW & DM need to be implemented + Connections • Long implementation time • Very high cost	• Simple structure ➢ Small complexity • Little implementation time • Low cost

FLEXIBILITY

DW Only	DM Only	DW & DM	Virtual DW
• Defined by various domains ➢ Adapting to changes more difficult	• Defined by single subject matter • Small data model ➢ Changes easy & quick	←	• Defined for varying formats and structures ➢ Changes easy & quick

DATA QUALITY & CONSISTENCY

DW Only	DM Only	DW & DM	Virtual DW
• High • Data conversion into common format ➢ No discrepancies • "Single source of truth"	• Redundant Data in various DMs	• Uniform data format	• Data quality logic manually created

SECURITY

DW Only	DM Only	DW & DM	Virtual DW
• Central repository ➢ Data access not limited	• Separate repositories ➢ Data access limited	• ←	• Security permissions defined in meta data ➢ Data access controlled

ADMINISTRATION EXPENSES

DW Only	DM Only	DW & DM	Virtual DW
It uses a centralized system design which simplifies Management and Backup capabilities.	Decentralized processes like backup, updates have to be done for each DM. In combination with a high error risk, this results in higher costs than DW.	Data is centralized, which allows safer management of DMs. When used together, the costs level out with savings from DW and more expenses from DMs.	Requires Views to the underlying Databases to be managed. Stacking Views on Views can also require extensive computing resources. These factors make the architecture very expensive.

SIZE

DW Only	DM Only	DW & DM	Virtual DW
DW size range is 100 GB to 1 TB+.	DM size is usually less than 100 GB.	Based on data from the Datawarehouse the size range is 100 GB to 1 TB+.	Comparable to tradidional DW.
Data Warehouse is a large repository of data collected from different sources.	DM only has a specific data to work with.	DM uses a subset of data from the DW.	The Virtual DWs size is easily modified, and Auto-scaling is available.

PERFORMANCE

DW Only	DM Only	DW & DM	Virtual DW
With a large amount of data stored in the DW the processing time increases and the performance is suffering.	Allow efficient access because the amount of data is smaller.	DMs improve the performance of a DW because they can take over processing tasks.	Virtual views on the data provide a fast query time, but the required computing resources are high.

HISTORY / LOGGING

DW Only	DM Only	DW & DM	Virtual DW
Has a dedicated location to store the History data.	←	←	Due to the acquisition of data during runtime and no central data storage the history data is not saved.
Retention times can vary between days, weeks, months, etc.	←	←	-

CONCLUSION

- DW and DM in combination address each other's weaknesses and work well in combination
- Virtual DWs provide visualization of Data stored in distributed physical environments through abstraction.
  - faster access and scaling but is expensive and has no historical data storage

QUELLEN

- https://www.geeksforgeeks.org/difference-between-data-warehouse-and-data-mart/
- https://intellipaat.com/blog/tutorial/data-warehouse-tutorial/merits-and-demerits-of-using-data-warehouse/
- http://mbenhaddou.com/2020/01/16/advantages-and-disadvantages-of-a-data-mart/
- https://blog.unbelievable-machine.com/en/virtual-data-warehousing-efficient-data-processing
- https://www.astera.com/de/type/blog/types-of-data-marts/
- https://www.intricity.com/whitepapers/physical-vs-virtual-tables
- https://www.guru99.com/data-warehouse-vs-data-mart.html
- https://wisdomschema.com/virtual-data-warehouse/
- https://www.talend.com/resources/cloud-data-warehouse-architecture/