DVG
Data Vault Generator

generated
processes

generated
dependencies

Start control

Processes with status

Dependencies

Scheduler

new jobs can be started?
(dependencies fulfilled)

planned

scheduled

running

ok

Parallel execution

Executoren
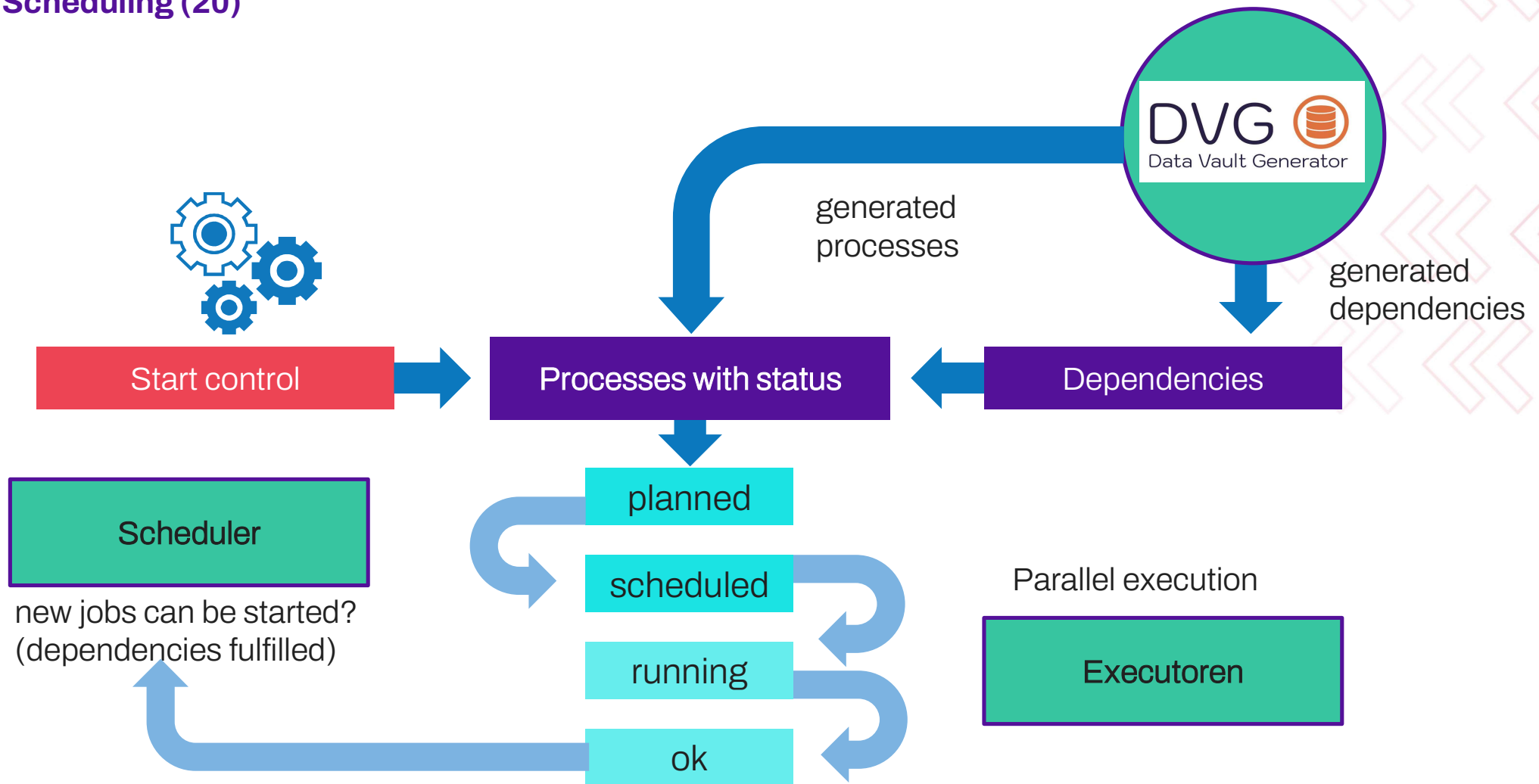
Scheduling (20)

» The DVG automatically generates a list of all necessary loading processes and derives the process dependencies from the data lineage.

» This information is transferred to metadata tables during deployment.

» When the DWH team plans a load, all necessary load processes for the respective load time are transferred to another metadata table with the status "planned" and the database-internal scheduler and executor tasks are activated. The number of executors and therefore the degree of parallel loading can be configured.

» The scheduler regularly checks whether there are any processes that no longer have any open predecessor processes and transfers these to the "scheduled" status.

» Free executors execute the processes with the "scheduled" status. The "running" status is set first. Successfully executed processes receive the status "ok". (If they fail, "error" is set)

» The scheduler can now set further processes to "scheduled". The cycle continues until all processes have the status "ok". Finally, the entire load time is set to the status "ok" and the database-internal scheduler and executor tasks are deactivated.

note:

• Several loading times can be loaded simultaneously in the DWH. Dependencies between the different loading times are taken into account.

• If the final completion of a loading time in the DWH is delayed (e.g. because a delivery system delivers extremely late), the loading of the next loading time can still be started.

Scheduling (20)

Example of monitoring the process status during loading

» Three loading times are loaded in parallel

» the individual loading processes have different status values

aktueller Ladezustand

● ok ● planned ● running ● scheduled

TGV	Anzahl
2023-06-01 00:00:00	243
2023-06-15 00:00:00	243
2023-06-30 00:00:00	243

Prozesse mit Fehlern bzw. running — 23 rows

TGV	PROCESS	STAT	MESSA
2023-06-01 00:00:00.000	LOAD_DM_V_OUT_AGG_BESTELLUNG_2_MART_D_DIMBESTELLUNG	running	
2023-06-01 00:00:00.000	CRYPT_WEB_KUNDE	running	
2023-06-15 00:00:00.000	LOAD_TEMP_WEB_S_SATBESTELLUNG	running	
2023-06-15 00:00:00.000	LOAD_DM_V_OUT_AGG_TERMINTREUE_2_MART_D_DIMTERMINTREUE	running	
2023-06-15 00:00:00.000	LOAD_RV_V_VAULT_ROAD_R_RS_BESTELLUNG_TO_WEB_F_POSITION_PRODUKT_EK_F	running	

Show 18 more

Übersicht Ladetage — 3 rows

TGV	STAT	LOAD_TYPE	PLAN_TIME	END_TIME	LAUFZEITN
2023-06-01 00:00:00.000	planned	standard	2023-08-08 14:13:52.376	null	
2023-06-15 00:00:00.000	planned	standard	2023-08-08 14:13:58.712	null	
2023-06-30 00:00:00.000	planned	standard	2023-08-08 14:14:04.657	null	