

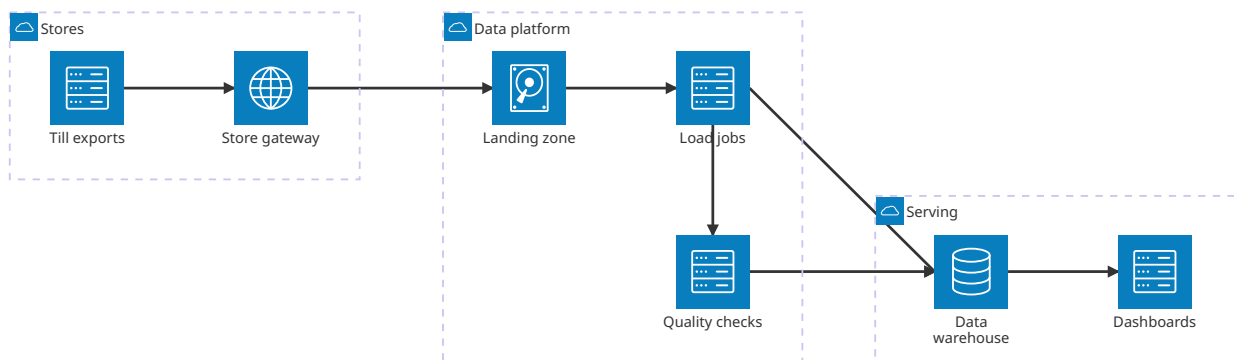
Runbook: the nightly data warehouse load

Version 4.1. Reviewed 26 May 2026. Owner: the data platform lead. On call: data platform, plus store systems for the till exports.

This runbook covers one thing: the nightly load that takes the day's sales from 340 stores in 9 countries into the data warehouse and refreshes the dashboards the buyers read at 08:00. Payroll, stock replenishment and the finance close have their own runbooks.

The load starts at 01:00 central European time and must finish by 06:30. Stores in the eastern-most time zone close at 22:00 local, which is 20:00 central, so every till export is in by 23:30 central.

Where the load runs



The store gateway is the only path from a store into the platform. A till export that fails the gateway's checksum is rejected and re-sent by the store system; nothing is repaired by hand.

What the load has to move

Source	Rows a night	Share
Till transactions	3,100,000	71.3 %
Stock movements	980,000	22.5 %
Online orders	215,000	4.9 %
Returns	56,000	1.3 %
Total	4,351,000	100.0 %

Capacity

Fact	Value
Load window	01:00 to 06:30 central European time, 5.5 hours
Load workers	8
Rows per worker per hour	160,000
Combined rate	1,280,000 rows an hour

Fact	Value
Time needed at the full rate	3 hours 24 minutes
Spare inside the window	2 hours 6 minutes
Rate with one worker lost	1,120,000 rows an hour, 3 hours 53 minutes
Rate with two workers lost	960,000 rows an hour, 4 hours 32 minutes

The load survives the loss of two workers inside the window. With three lost it misses 06:30, and that is the escalation trigger below.

Warning

Do not add workers during a load to catch up. Above eight workers the warehouse's write path becomes the limit and every worker slows down. Hold the late stores for the morning catch-up load instead.

Before the load starts

- ☒ Confirm the store gateway reports every store
- ☒ Confirm the landing zone has free space for twice a normal night
- ☐ Confirm all eight workers report healthy
- ☐ Confirm yesterday's quality report has no open failure

If the load falls behind

1. At 04:00, check the count of loaded rows. Below 3,000,000 means the load is behind.
2. Count the running workers. Seven is acceptable. Six or fewer is not.
3. If three or more workers are lost, page the data platform on call.
4. If the load cannot finish by 06:30, publish yesterday's dashboards with a banner and tell the buying teams before 07:30. A stale dashboard is acceptable. A wrong one is not.
5. Never restart the load from the beginning. Each source is loaded in order and a restart doubles the rows already written.

All organizations in the sample documents are invented. Names and figures are illustrative.