

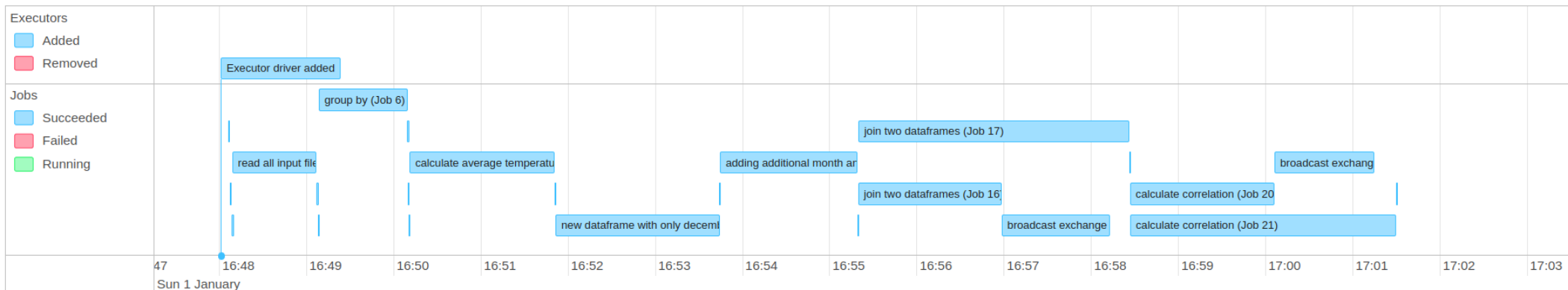
Spark – Bauernregeln

- 2 Executors
- Executor 0
 - 1 GB RAM, 12 cores
 - Complete Tasks: 253
 - Total Task time: 23 min 28 sek
 - Total Job time:
 - Input Data: 15,4 GB
- Executor 1
 - 1 GB RAM, 8 cores
 - Complete Tasks: 76
 - Total Task time: 33 min 29 sek
 - Total Job time:
 - Input Data: 5,6 GB
- Total Task Time: Tasks sind parallel ausgeführt: 1 task auf 1 Core – also sprechen wir nicht von “wall clock time” – also die Zeiten sind nicht vergleichbar da ein Rechner mehr Cores hat
- Frage: Wie viel ist jeder Kern ausgelastet? Erkennen ob es Situationen gibt, in denen einige Kerne schlafen, während andere arbeiten.

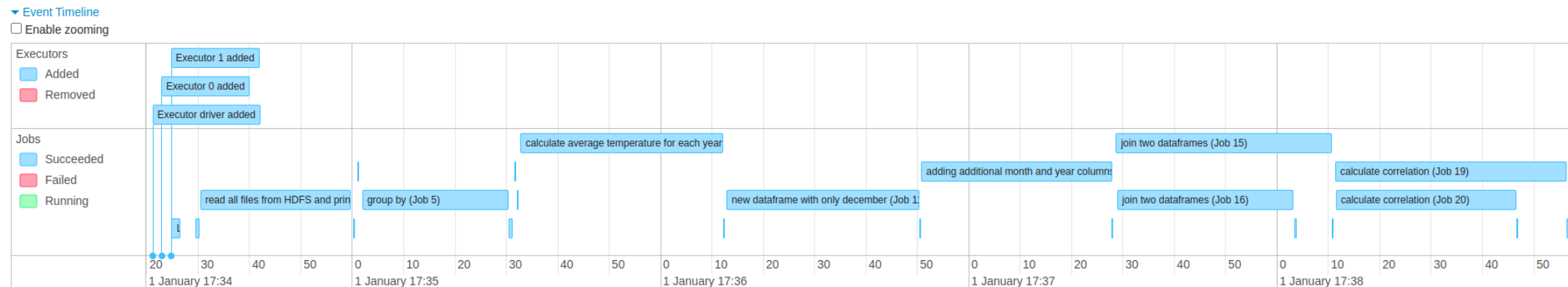
	Stage 2		Stage 5		Stage 10		Stage 13		Stage 16	
	Input Data	Total time	Input Data	Total time	Input Data	Total time	Input Data	Total time	Input Data	Total time
Executor 0	1.5 GB	1.6 min	1.5 GB	1.5 min	1.5 GB	2.1 min	1.5 GB	2.2 min	1.5 GB	2.2 min
Executor 1	808 MB	3.8 min	826 MB	3.6 min	808 MB	5.0 min	826 MB	4.7 min	808 MB	4.8 min

	Stage 19		Stage 20		Stage 25		Stage 26		TOTAL	
	Input Data	Total time	Input Data	Total time	Input Data	Total time	Input Data	Total time	Input Data	Total Time
Executor 0	1.5 GB	2.6 min	2.3 GB	3.7 min	1.5 GB	2.6 min	2.3 GB	3.8 min	15.4 GB	22.3
Executor 1	808 MB	5.3 min	-	-	808 MB	5.8 min	-	-	5.6 GB	33

Single Node:
start: 16:48:08
end: 17:01:29
total time: 13 minutes 21 seconds



two nodes:
start: 17:34:30
end: 17:38:56
total time: 4 minutes 26 seconds



Tasks in Stage 2 (Job 2)

12 Cores in Executor 0 schlafen für ca. 20 sekunden?

