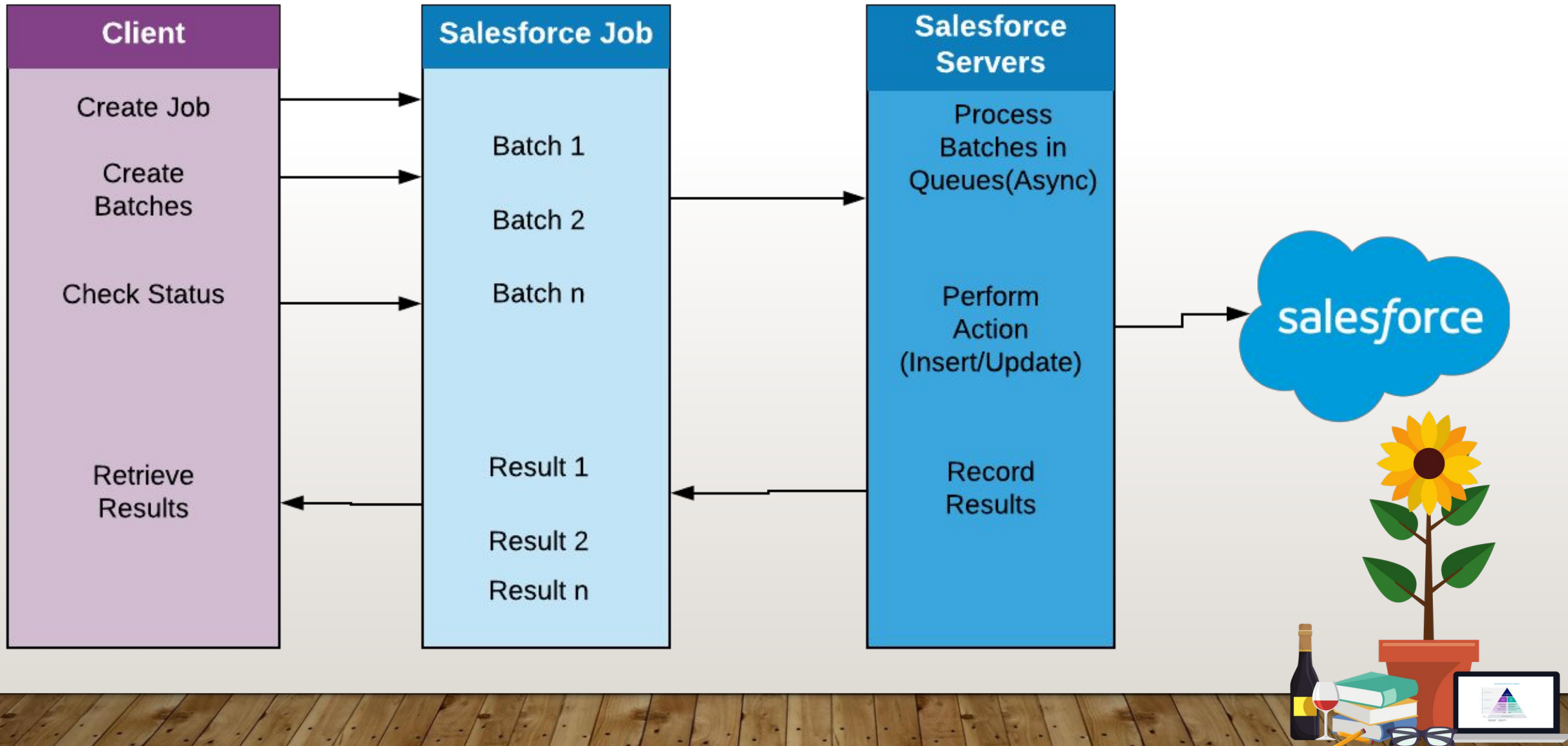


Introducing Bulk API

- Programmatic option from Salesforce to load data into an org at speed
- Optimised for loading and deleting Large data volumes
- Based on REST principles
- Lets you query, insert, update, upsert and delete records in batches
- Runs in Serial and Parallel mode
- Asynchronous processing



How does it work?



Monitor Jobs

Bulk Data Load Jobs - salesforce.com - Develop...

Find: Previous Next Options

Home Chatter Campaigns Leads Accounts Contacts Opportunities Forecasts Contracts Cases Solutions Products Reports Dashboards +

Search All Setup... Expand All | Collapse All

Force.com Home

Administer

- Manage Users
- Manage Apps
- Company Profile
- Security Controls
- Domain Management
- Communication Templates
- Translation Workbench
- Data Management
- Mobile Administration
- Desktop Administration
- Email Administration
- Google Apps

Build

- Customize

Monitor Bulk Data Load Jobs

Help for this Page

Monitor the status of recent bulk data load jobs. These jobs are created by Data Loader and other Bulk API client applications.

Quota

Your organization has processed 0 batches in the last 24 hours. Your organization can process 10,000 batches in a 24-hour period.

Resource used in the last 24 hours:
CPU: 0 milliseconds
IO: 0 bytes
Disk: 0 bytes

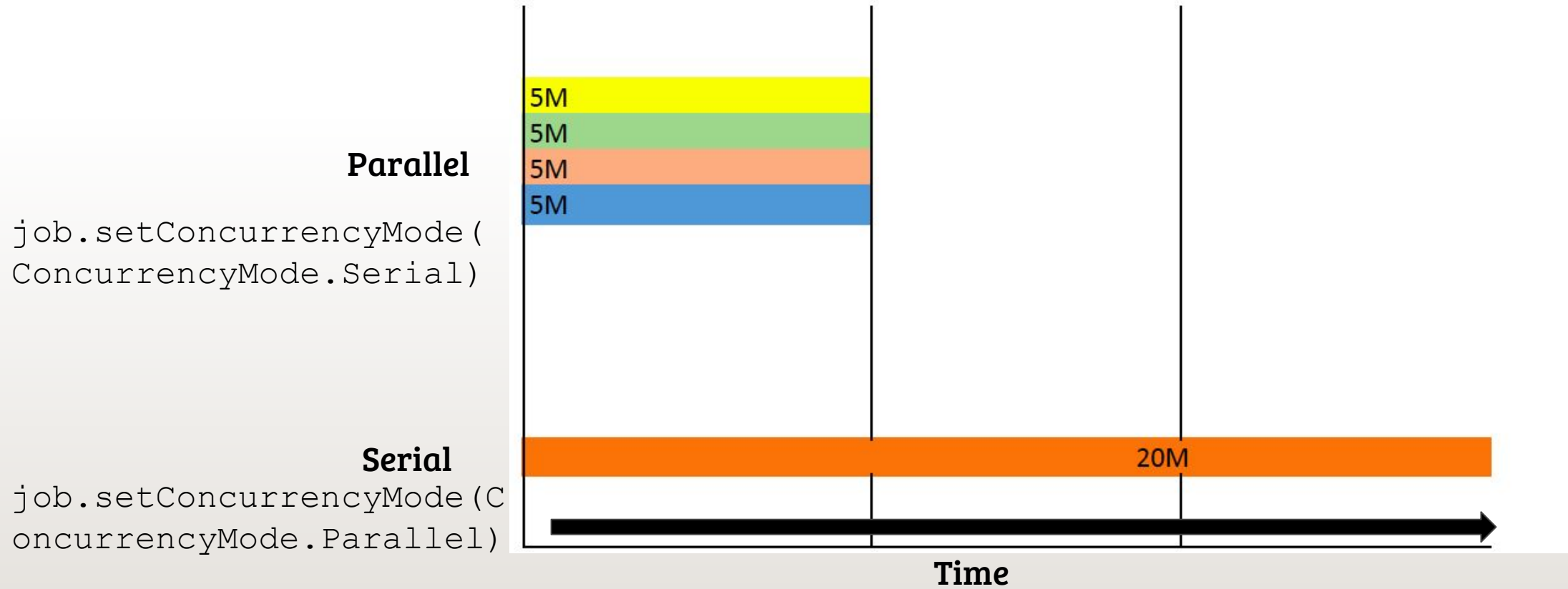
In Progress

Job ID	Submitted By	Start Time +	Status	Operation	Object	Records Processed	Records Failed	Progress
No records to display.								

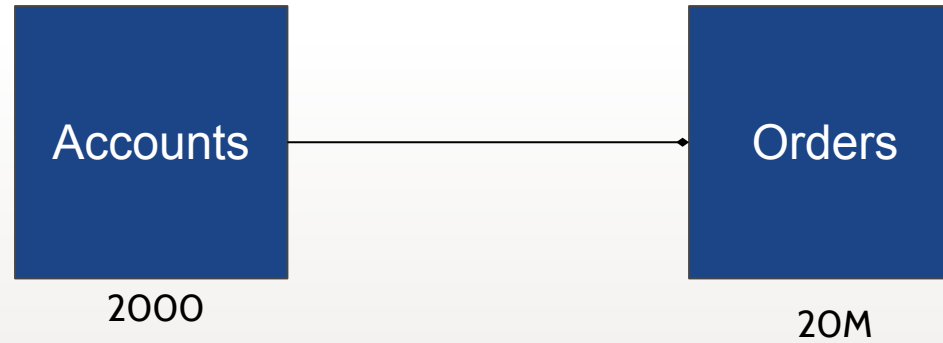
Completed last 7 days

Job ID	Submitted By	Start Time	End Time +	Status	Operation	Object	Records Processed	Records Failed	Time to Complete ([hh:mm:ss])
750i0000000n4dU	Regan, Sean	2/20/2014 7:58 AM	2/20/2014 9:06 AM	Closed	Insert	Order	1000000	0	01:07:54

Serial and Parallel Modes

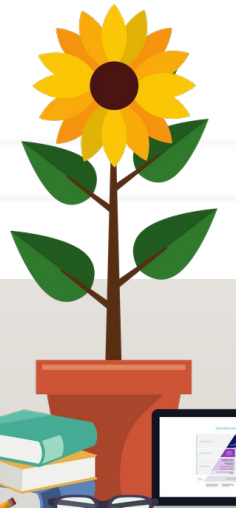


Use Case



Lookup Options

Related To	<u>Account</u>	Child Relationship Name	Orders
Related List Label	Orders		
Required	<input type="checkbox"/>		
What to do if the lookup record is deleted?	Don't allow deletion of the lookup record that's part of a lookup relationship.		



Concepts

Degree of parallelism

An integration job that inserts 100,000 records in 5 minutes. Your batch size is 10,000 records, so you have 10 total batches. For simplicity's sake, say that it takes Salesforce 1 minute to process each batch. 10 minutes of work was completed in 5 minutes
The overall degree of parallelism for the Bulk API job is 2

**degree of
parallelism =**

amount of work completed (as a duration)

actual amount of time it took to complete that work

For Serial loads limited to 1

For parallel loads as much as server allows - the idea is to maximise.

Throughput

Records processed per second - Measure of performance

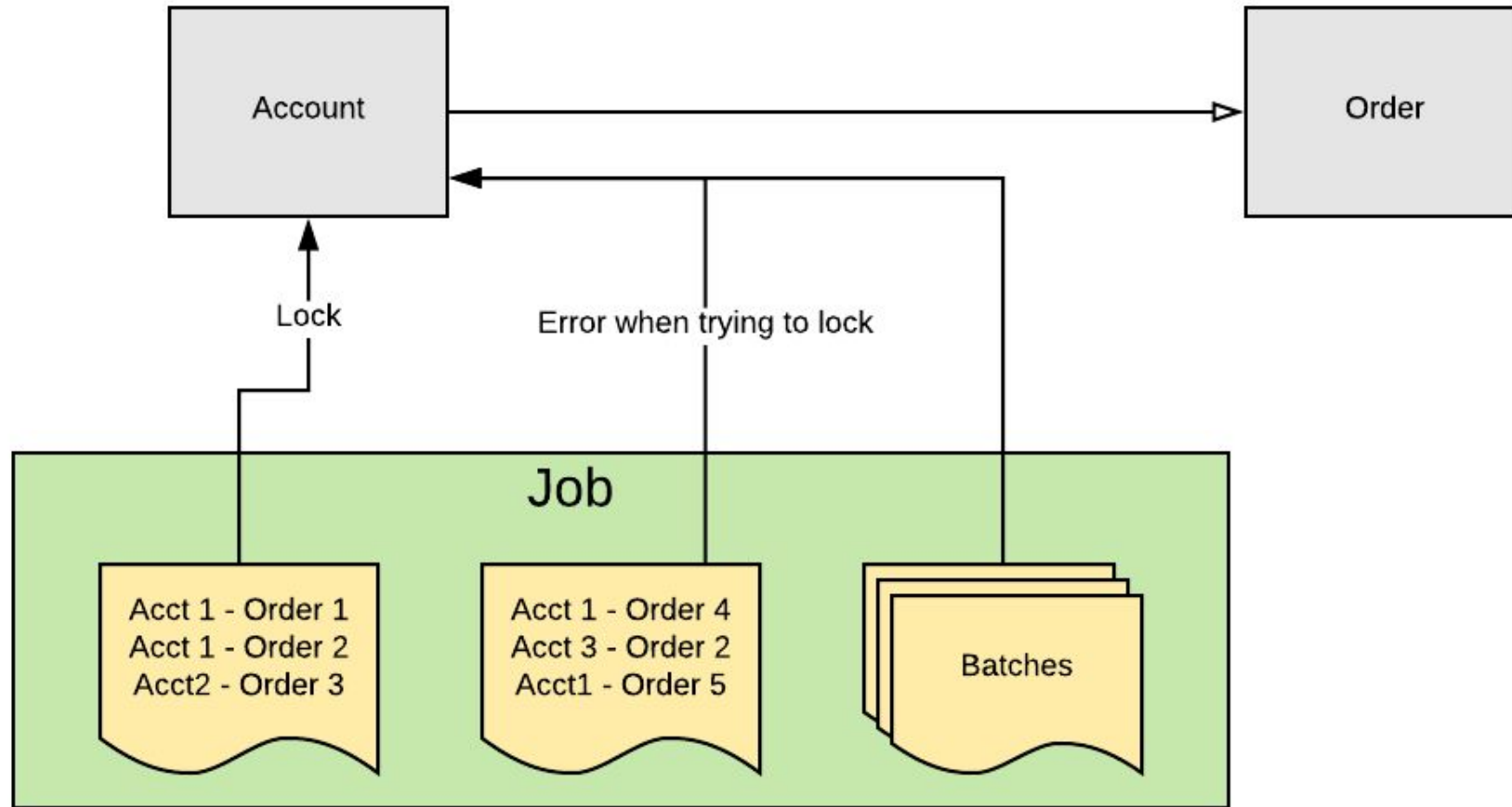


Parallel Processing

- Can increase throughput if done right
- Can produce unexpected degree of parallelism if not architected well
- Locks can significantly impact parallel processing by
 - Wasted processing time
 - Causing failures
 - Reduced throughput



What is lock contention?



What causes a lock contention?

- Master detail relationship
- Certain types of lookup relationship
- Roll -up Summary fields
- Triggers
- Workflows
- Group Memberships
- Overlapping runs

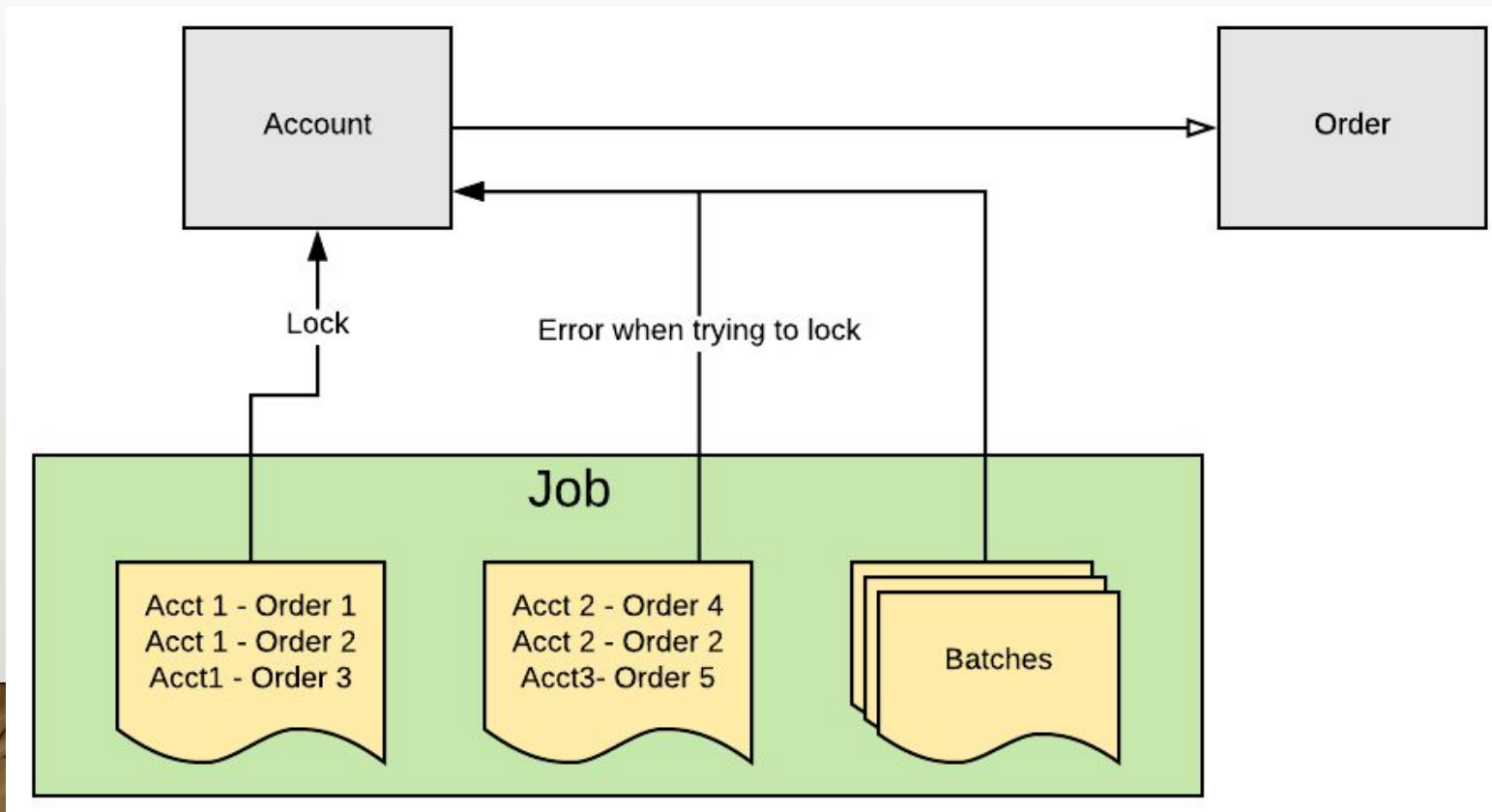


Strategies for resolving lock contention

- Do not lock lookup records - Change lookup type config

“Don't allow deletion of the lookup record that's part of a lookup relationship”

- Order records to be loaded by parent record ID



Strategies for resolving lock contention

- Controlled feeds, chunking data into more jobs less batches?
 - Avoid low controlled feed loads if possible
 - Reduce number of Jobs and increase number of batches
 - Keep at least 20 batches in the queue to optimise usage of worker threads
- Disable Validation Rules, Triggers, Workflows while loading
- Lookout for rollup summary field, use summary reporting if possible
- Defer sharing calculations for mass group membership changes

