2023-01-30 Architecture WG Meeting Notes

Date

30 Jan 2023

ZOOM Meeting Information:

Monday, January 30, 2023 at 11:30am PT/2:30pm ET.

Join Zoom Meeting

https://zoom.us/j/7904999331

Meeting ID: 790 499 9331

Antitrust Policy Notice

Linux Foundation meetings involve participation by industry competitors, and it is the intention of the Linux Foundation to conduct all of its activities in accordance with applicable antitrust and competition laws. It is therefore extremely important that attendees adhere to meeting agendas, and be aware of, and not participate in, any activities that are prohibited under applicable US state, federal or foreign antitrust and competition laws.

Examples of types of actions that are prohibited at Linux Foundation meetings and in connection with Linux Foundation activities are described in the Linux Foundation Antitrust Policy available at http://www.linuxfoundation.org/antitrust-policy. If you have questions about these matters, please contact your company counsel, or if you are a member of the Linux Foundation, feel free to contact Andrew Updegrove of the firm of Gesmer Updegrove LLP, which provides legal counsel to the Linux Foundation.

openIDL

Attendees:

- Sean Bohan (openIDL)
- Mason Wagoner (AAIS)
- Dale Harris (Travelers)
- Peter Antley (AAIS)
- Ken Sayers (AAIS)
- Faheem Zakaria (Hanover)
- Jeff Braswell (openIDL)
- David Reale (Travelers)
- Tsvetan Georgiev (Senofi)
- Brian Hoffman (Travelers)
- Milind Zodge (Hartford)
- Ash Naik (AAIS)
- Yanko Zhelyazkov (Senofi)

Agenda:

- Update on Large Carrier POC (Ken Sayers)
- Update on ND POC (Ken Sayers)
- Update on openIDL Testnet (Jeff Braswell)
- Update on internal Stat Reporting with openIDL (Peter Antley)
- Update on Infrastructure Working Group (Sean Bohan)

- Future Topics

Notes:

- Large Carrier POC (see notes from TSC call last week)
  - Senofi the Infrastructure Partner
- ND Uninsured Motorist POC
  - Opening up nodes for carriers and ND on Weds Feb 1
  - Until Feb 10 to load data
  - Will work with ND DOI to finish data call and create report
  - Hardening and increasing throughput
  - Diff model than most data calls, most small amount of data, all aggregated - different in that EP includes unique VINs, handle that kind of load, required server configs
  - How to config servers if diff use cases have different load needs, what to take into account to manage diff types of data calls - stretched the Architecture a little
- openIDL Testnet
  - Going well, up and running, more updates to come
- Internal Stat Reporting demo
  - Working through issues, schemas, stat plan converters, decoders, (mobile home owners
  - This project will coincide with large carrier POC, dont need all the stat data to do that, will capture in the same format (stat plan format)
  - Big Picture - test phase for all lines by end of year
  - Demo video continues (graphics, script, production of footage by end of week)'
- IWG
- Future Topics / AOB
  - PA
    - PA working on ETL (last week's discussion), key point: doing data loads and if there are errors in the load to not commit /rollback - looking for resilient loading solution
    - When we talk about a job to load something, way it works today, stat records come in positional format, characters at the end, first two characters represent product line
    - easy as a carrier - giant load of stat objects, process it - DO WE WANT JOBS TO BE MULTI-LINE? - trying to figure out where to put reconciliations and checksums so loads happen correctly
  - DH - one load per line not a huge inconvenience, as it is they sep some auto records due to size of them, when talking about load, the Carriers resp is to get it into the HDS, not somethign openIDL needs to worry about - thinking about load in terms of HDS
  - PA - talk about doing a load of data, have error tolerances, come down to jurisdiction and line of insurance (correct per DH) - if you have lines that have errors, only pass the ones w/o errors and reject the ones with errors - break it up by lines
  - DH - never get a situatin where a line of business is "error free"
  - PA - 2 DIFF parts of the load, PA is talking about assuming gets 10 records, ensure they end up in proper tables, with checksums, etc. - NOT addrressing edit package - input/output control mechansim - edit package before this - this process assumes post edit
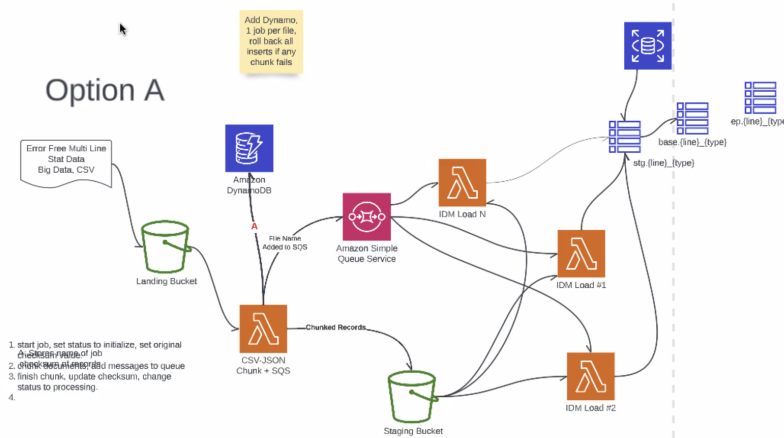
  - PA
  - PA - work to move data from inbound bucket into DB, once sure correct in staging table, do transact to do "select from insert", move files, flash the staging info - Per ND, going with serverless, cognizant of memory, take large files, chunk up, into queue, series of lambda functions, spawn concurrent lambdas, staging bucket, open file, load into staging table
  - PA - option A gets all files in there, gets tricky to know when/where file lost - orchestrate load
  - .

schema: openidl_etl
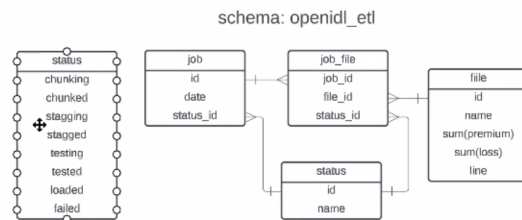
-
  - Atomicity of a file load
  - PA - once chunking - chunking and making the chunks by line, own tables, end up with weird math
  - KS - checksum before split em?
  - DH - how often extract data for mult lines?
  - PA - extract for mult lines is by EP, catastrophe call: lets look at residential property (3 lines - homeowners, dwelling properties and mobile homes) - way EPs work can look at any tables
  - KS - db tables are line sensitive, by loads
  - DH - confused about load - loading into HDS?
  - PA - loading into HDS, not required, if you want to set up node on testnet - one way to move data
  - KS - would give out as something to use or not - will have info about tables etc. - use it or buildi your own, meant as a astarting point - TRV will own load of the data, just trying to make it easier
  - PA - setting up for others to make it easier
  - DR - cleanroom products for cloud providers - this - multi tenant puts data into datastore, permissions and struc queries based on accounts, ex: DR puts data in S3 data lake, can be queried by Glue or Athena, and can allow AAIS to do agg queries, Hanover does lookup, can enable with few IM permissions
  - KS - consistent across implementations, ?
  - DR - asymmetric data, similar queries for Hartford and Trav b/c same schemas - snowflake live, Azure announced theirs, solving other problems they want to discuss - somethign to have an answer to - couple tutorials and videos, req access to the beta, see how it works, still in preview, gonna be very appealing to a lot of people already use these platforms - looks intriguing
  - Clean Rooms Link from Faheem
  - Blockchain based clean rooms link
  - https://learn.microsoft.com/en-us/azure/architecture/example-scenario/blockchain/multi-cloud-blockchain
  - https://resources.snowflake.com/webinars-thought-leadership/from-data-clean-rooms-to-data-marketplaces-how-to-securely-collaborate-across-your-business-ecosystem-2?utm_cta=website-workload-collaboration-use-case-webinar-business-ecosystem&_ga=2.175228194.1409119339.1674590263-711284629.1674590263&_gac=1.241968566.1674590264.EAIaIQobChMIqpn-6f7g_AIVlo7ICh1OjAGUEAAYAiAAEgLdxvD_BwE
  - https://www.theinformation.com/articles/microsoft-readies-service-to-let-cloud-customers-pool-data-safely?utm_source=ti_app
- Future Topics
  - Where is the edit package? Who makes/own/maintains it?
  - Edit Package discussion - how plugged in, how used by individual carriers (will they have their own tools? shared tools?)
  - Clean Rooms and openIDL
  - Security
  - Roles and Access
  - Governing Board

GMT20230130-1...1920x1080.mp4

Time	Item	Who	Notes

Documentation:

Notes: (Notes taken live in Requirements document)

Recording: