2022-05-23 Architecture WG Meeting Notes

Date

23 May 2022

Antitrust Policy

> ### 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)
- Joan Zerkovich (AAIS)
- Ken Sayers (AAIS)
- Eric Lowe (VA)
- Nathan Southern (openIDL)
- Jeff Braswell (openIDL)_
- Satish Kasala (The Hartford)
- David Reale (Travelers)
- Brian Hoffman (Travelers)
- Greg Williams (AAIS)
- Peter Antley (AAIS)
- Rajesh Sanjeevi (Chainyard)

Agenda:

1. Architecture WG Harmonized Data Store Task Force

- Host: Peter Antley, AAIS (for now)
- Tuesdays 9am (for now)
- Invites to go out (contact sbohan@linuxfoundation.org)

1. openIDL - Architecture - Member Requirements Files
2. Regulatory Reporting Requirements from Dale H (Travelers)
3. Requirements Table

Time	Item	Who	Notes

Action items

Notes:

- JeffB
- DaleH's document

- great input on aspects of arch requirements
- High level way of looking at the work that might be going forward
- Trad process: network stack, 1 party talking to another at different levels
- division between parties
- division between horizons (business and data)
- below the line - infrastructure implementing this
- definitions and stipulations (matrix)
- looking at requirements on the contributing side
- carrier node
- influence, impact on lower levels
- take a look, great contribution
- when Dale discussed last week, concept of granting consent to a data call
- orchestrated workflow transactions
- might be over the course of days
- data call req announced, on carrier side can it be provided, consent, other data? in aggregate not just one carriers business dominating the data
- business orchestration transacting at app level
- impact on how we implement in fabric
- good example, kind of req that will impact arch
- making sure on this call, work contributed
- KenS
- worth going thru it
- give folks a sense of what was contributed, see the value of it, restate what is not clear
- can walk thru it
- Dale - organized it - what are the diff things that need to happen for stat reporting on openIDL
- Reading column D
- D4
- DavidR - transactions will be determined by source system
- any changes to this source will be reflected in HDS
- Ken - can we clarify it?
- Satish - are we saying HDS will maintain time series on the policy AND the claim?
- Ken - yes, changes as they happen based on what source system says
- diff - if there is an error moving from source system to HDS, will not correct errors will replace (no onset-offset)
- Satish - only incremental or delta change?
- Ken - cant do snapshots, don't know what time period request is asked for
- need history as unfolded
- Satish - snapshot - point in time of a policy, if you take a point in time copy of the policy at 3pm on 5/23, take a copy of that same policy tomorrow there could be changes in between, question - are we capturing just changes or complete copy
- Jeff - data submitted in format of the data standard reqs, would reflect things that have and haven't changed - record format
- Ken -snapshot on pull, source data must have every possible snapshot, doesn't keep at snapshot
- David - HDS will be accurate rep of source systems in correct format at any time
- not incremental
- will look at entirety of period of relevance
- David - req captures that "these will be the records"
- Ken - policy and loss transactions
- D5 - data freshness
- D6 - 5 prior years + current in HDS
- Peter - strawman - need a decision on how long we retain data
- Ken
- D7 - data in the HDS will remain within 5% error tolerance per line and state based on openIDL edit package
- what? Data requested, data loaded? do we need to reevaluate every time we load data?
- Peter - would like someone to correct - NAIC handbook talking annual reporting
- Ken - constraining ourselves to annual reporting
- Jeff - mark for clarification from dale
- Brian - 5% consistent with AAIS stat plan
- Ken - we do it on load,
- Peter - individual load
- Ken - unclear of math / thresholds
- Jeff - biz requirement for arch
- D8 - access is permissioned
- Jeff - diff categories
- David - what Dale wants, any new use is consented to on a use by use basis
- if there for stat reporting, data not accessible without specific approval
- Ken
- permission on extraction
- David - you could be permissioned party for one use (annual report) not permissioned to pull data for another purposed - granular permission based on use
- Ken - use = extraction
- ask for another extraction
- Jeff - for the staged data, in extract data - single defined use - once you have the data cant use for another purpose (Column F ROW 2)
- David - nuance - cant extract data, you can only use for cert purpose - yes - use must be approved
- Satish - internal access? d0 we need to worry about internal access, from a carrier standpoint
- if you put data into HDS - can we see who put data into HDS from carrier standpoint
- Jeff - carrier would approve requests
- Ken - out of openIDL's hands - you own HDS< you control however you want
- David - HDS is carrier-owned asset they do with as they see fit
- Ken
- E2 - DATA shall be aggregated during extraction, E3 anonymized
- these two cells are related
- should we change from during?

- aggregated ON analytics node
- Jeff - not every detailed record supplied, some form as part of extraction query
- David - instead of on extraction - raw data doesn't leave carrier node in its raw form
- maybe cleaner to write - raw data doesn't leave carrier node
- Satish - original intent of request data, add requirements: consent, when, time period valid
- data shall be aggregated should be moved to extract
- David - columns are confusing stuff
- Ken - ignore columns for now
- David - raw data (no extraction pattern that pulls all data out)
- Carrier node to analytics node
- Ken
- E3 - anonymization happens on analytics node
- Jeff - does this also mean some of it does get transmitted for specific policies without unique identifiers
- David - where this goes - defining how specific extract patterns might behave
- don't want a lump travelers dataset sitting on analytics node indefinitely
- if pulled for purpose - would like to occur immediately, not stored as Travelers data in perpetuity
- Ken - carrier is the identifier
- Satish - not about PII but just carrier
- data source shall be anonymized
- David - PII should not be stored, etc.
- Eric - one of the things, having DL #s in HDS, VA wont get it but if not in the store - other things for useful research doesn't happen
- PII in and of itself
- Ken - DURING extraction
- Eric - not give me permission to get it
- confusing data request from smart contract that gives me permission
- Satish - clarification
- anonymized
- do we ever need to know from RR standpoint that this data is from Hartford, Travelers?
- Eric - either a data call or building HDS and go into other areas, if I call a market conduct of Hartford for this year send me all your info on your policies, you may pull from HDS
- smart contract negotiated - here is the data you have permission to pull it in this form
- option there going forward
- not limiting to what we see today but could happen in the future
- Jeff - stat report does not need to ID, in aggregate
- Eric - now stat rep, could be used for other reporting requirements, devil in this detail is the extraction
- David - lot of reqs driven by extraction pattern
- is EP says "combine x carriers by y carriers" will be stored that way
- specific use cases will define that
- concerned - storage of data will be determined by extraction pattern terms
- Eric - right now only for member companies, and right now AAIS gets it not totally anonymized
- you don't give me who carrier is, but AAIS has defined business need
- Jeff - combined at service but not available or reported
- Eric - AAIS not get or not reported (have it not reported)
- David - making sure data used for what it was requested for
- Eric -c ant envision every use case but if we have raw data and extraction patterns, no free reign for me (or other regulations)
- David - if intended data call says "all data aggregated with other carriers, noting from AAIS
- would have to approve at the time
- Eric - group needs to figure out min # of carriers to prevent regulators from de-identify who submitted it
- Ken - E4 - only aggregated and anonymous shall move past analytics node
- E5 no PII data transmitted beyond private analytics node
- Satish - do we ever need PII data published in stat reporting?
- other reporting?
- BrianH - not as part of stat reporting
- Satish - other data calls?
- Jeff - don't need actual PII
- David - one of the cores - if there is every a use case requiring PII, would want connection to be made prior to transfer out
- major benefits of this topology
- doesn't leave carrier node
- GDPR, CCPA makes it messy if it ever leaves
- Ken - didn't quite match
- could leave the carrier and be used for reporting but it never leaves analytics node
- David - major Arch req, in former design had to, this winter thought maybe we move it into the carrier node
- imp requirement - where is that line
- Jeff - req still stands
- not analytics node and not carrier
- Satish - example - age range of drive, sometimes requested, individual elements like date range - is that PII
- David - fact aggregated would obscure identifying things
- game is messy to play
- Ken - how do we from openIDL perspective meet this req
- Carriers not framework
- David - messy, say there is a req to put some sort of info considered PII in HDS, obfuscated and de-identified by extraction pattern, would set rule EP do not hold PII - nonfunctional req, needs to exist
- based on some use cases, will be necessity
- thinking
- Satish - in some cases, one of more PII elements, even if not contribute to constructing full ID of a person
- Jeff - what constitutes PII?
- David - safeguard req - PII doesn't leave carrier node (for now)
- Jeff - simplest approach
- most systems have the audit trail for this kind of data, don't want to create compliance problems

- David - ideally, don't wand PII in HDS to begin with, some abstraction preferable
- Travelers haven't considered all components of PII in HDS
- E6 - timeframes for aggregation need to be defined (when recorded, when extracted,)
- Jeff - full month?
- Ken - could be a day - need to know when
- Satish - start and end data mandatory as part of extraction pattern?
- Ken - are they all?
- David - getting into new realm of reqs here about extraction pattern requirements
- maybe higher level - extraction patterns shall have relevant parameters clearly defined - point to extract pattern document
- own thing, decouple EP reqs from overall requirements
- Jeff - int in active policies, but may reach back to prior years
- Ken - data calls tell you what you are asking for
- David - defined as its own document
- this should point to that
- E7 attributes used in aggregation shall be identified
- E8 logic for extracting data shall be defined
- E9 Calculations to be used for aggregation, analysis, reporting shall be defined
- E10 Specific use of info shall be defined
- Ken - when you create data call - what and why you want data
- David - define requirements of a data call (inc reason for it)
- reason sep - VERY granular - as to what those requirements could be
- data, retention policies, could get long
- concerned about data call and the EP is the implementation of data call reqs
- E11 permitted accessors to the info and users of the data shall be defined
- Ken - is there anything technical we expect to do to constrain
- Satish - do we need governance "I will used this data for this purpose"
- David - might be tech or governance - hold as req and progress which one it is
- Satish - part of the color coding?
- Ken part of the data call
- E12 proportion of carrier info to be used
- Ken - thing that kicked off what Jeff mentioned
- "I consented to it" but if you don't get 3 other carriers I don't consent
- if then - 2 phase
- David - starting to get into actual contracts and agreements
- language that needs to be agreed upon
- "use of the data shall conform to agreements"
- don't know the form of these things
- Brian - guess on Dales part
- bit of a swag estimate (25%)
- Ken put a place where you can say "this call i cant do x $"
- if you cant meet it, cancel the call
- Jeff - how would the arch come up with that answer
- David - messy part, smart contracts can do on the fly, most will be parsing of every call in language
- looking for req - the use and retention and access of all this data shall conform to the terms and conditions of this data call
- go back once arch more formalized "do it the way you promised"
- Ken - written down - spidey sense
- Jeff - not a smart contract, synchronous transact,
- may take longer to decide if some things can be consented to
- David - smart contract is short hand in my mind for extraction pattern
- E12 - only auth approvers may commit to a carrier data request
- ken - for all carriers?
- David - there needs to be granular permissions for different functions on the platform
- dont know all those tasks or user types
- need ability to have granular permissions
- Satish - is this openIDL or carrier can implement they way they feel
- David - openIDL side on the adapter at the least
- Ken - interacting with UI and credential, cred associated with roles, when you define creds in openIDL have to roles and perms
- Satish - should extend into HDS?
- Ken - federated
- David
- Ken takeaway
- connect and document the flow of a data call maps to things moving around to clarify
- David - nice if some docs uploaded, NAIC docs - nice if those could be parsed out as requirements

Recording:


openIDL_ArchWG...2322_video.mp4