Fingrid Datahub Oy / Marko Juslin 6.5.2024

Fingrid Datahub Measurement data continuous delivery WS I - 08/2024

Time 13.-14.8.2024 10:00–15:00

Place Fingrid Datahub Oy, Läkkisepäntie 23, Helsinki

Present	Marko Juslin	Fingrid Datahub Oy (cm)
	Laura Markkanen	Fingrid Datahub Oy
	Tuomas Aunola	Fingrid Datahub Oy
	Otto Kuuranne	Fingrid ICT
	Remco Nederpel	CGI NL
	Gerold Slagter	CGI NL
	Jan Loman	CGI Finland Oy (sec)
	Antti Kirjola	Caruna Oy
	Anssi Vanhatalo	Elenia Oyj
	Vesa Hulttinen	Hansen Technologies Finland Oy
	Pyry Suomala	Helen Oy
	Jami Kosunen	TietoEvry Oy
	Ville Kuusela	Solteq Oyj
	Mikko Vähäsaari	Savon Voima Verkko Oy
	Lari Saarinen	Aidon Oy
	Elina Kontro	Landis+Gyr Oy
Not present		
	Jari Rusanen	Kajave Oy

Day 1

1 Workshop Introduction

Marko went through the agenda and gave a short introduction to the workshop.

Introduction round where everyone introduced themselves.

2 Background and targets

Marko went through the background and targets of continuous metering delivery.

Questions & comments:

1. Question: Is it possible to use current and the new channel parallel?

– Answer: Need to be decided during the workshop. CGI assumes parallel usage will be allowed.

2. Question: Do only OK statuses need to be delivered in 6 hours or do all statuses?

   – Answer: Will be discussed during the workshop.

3. Comment: Enabling of the event channel is important. Enable a solution that opens doors, and not too restricted. Those parties who are active on the balance market will be interested in getting the data asap (i.e. streaming out).

4. Question: Why 6 hours?

   – Comment: Just an arbitrary number. Less than 24 hours but there was no exact knowledge of how fast new meters can deliver the data on the time when the “Älyverkkotyöryhmä” workgroup discussed about this.

5. Comment: Downtime requirements vs buffering requirements from DSO side needs to be taken into account. E.g. downtime of 30 minutes ends up to potential millions of events being buffered.

3 Datahub and performance characteristics

Remco presented datahub shortly and also performance characteristics based on PoC phase 1 analysis.

6. Question: Can slides be shared?

   – Answer: Yes. Slides will be shared in the Teams.

7. Question: How much overhead there is in the message?
   – Answer: Will return to this on day 2 of the workshop. Xml is quite heavy.

8. Question: What is the quality of the current 5 min readings DSO are collecting?
   – Answer: If a meter sends something it is ok. -> Need to consider re-sends. Data-hub does not want duplicate readings by default.

9. Question: Is there a limit for message size?
   – Answer: Recommended maximum size is currently 100 MB.

10. Comment: Datahub could handle the 6-hour requirement already after optimization and if DSOs bundle the data.

11. Question: Does the supplier want an increased number of messages in event channel?
    – Comment: One message into datahub means in many cases one or more messages out from datahub to many parties.
    – Comment: One potential requirement. Suppliers could decide if they want to receive streaming data. The supplier's opinion of frequency and number of messages out from datahub needs to be studied more.

4 Workshop Business requirements

Workgroup members were asked to write down the pre-tasks on post it papers. Then they were divided to the three groups to discuss about the pre-tasks and amend them. After that groups summarized and presented those for a common discussion with all workshop members.

Pre-tasks:
What kind of benefits do you see becoming with the continuous metering data delivery from different actor roles point of view?

- What are the short-term needs?
- What kind of longer-term goal/visions could be in the horizon?
- What kind of constraints do you see?

Workgroup members details for above questions can be found in appendix 1.

12. Question: How is defined the continuous delivery?
    – Answer: 15min/1h resolution of data but delivery frequency should be from 1 min to 6 hours

13. Question: Do we need both channels?
    – Answer: Current solution/channel does not support high frequency deliveries with small messages.

14. Comment: One solution could be that DSO streaming service sends datahub and datahub separate reports for not ok and missing data to DSO and DSO's MDM system sends only the changed data and new data, not all data.

15. Question: Does the law apply to 100% of readings or 95% i.e. do we need to get missing values in 6 hours?
    – Comment: DSO needs to collect the data from the meter in 6 hours
    – Fingrid comment: and data shall be in datahub in 6 hours.
    – No answer at the moment, to be discussed later.

16. Question: Should we prevent community calculations immediately and do those once a day or some other frequency if DSO cannot send all community values in one message?
    – Comment: Should be possible to send only ok or all statuses

17. Comment: No acknowledgements (event handled in datahub). Reminders will be sent to DSO and DSO's MDM system sends the value to Datahub.

5 Ending of day 1

First day of the workshop was about getting acquainted with the participants, getting familiar with the topics and getting the first thoughts out there regarding vision, requirements and definitions.

When just focusing on being compliant with the new legislation for publishing measurement data within 6 hours it is agreed upon that the existing batch channel would probably suffice when a consensus between datahub and DSOs is reached on new guidelines on how to bundle the measurements and possibly after some optimization on datahub.

Having the measurement data near-real-time in datahub has not yet view on strong business needs, the participants agree on the potential it could have in the future developments.

Conclusion on day 1 is that a Timeseries Event Channel is desired in time and steps to introduce one should be taken. However, it is expected that most DSOs won't be able to implement the interface prior to 01-01-2026 and will use the existing batch interface to comply with the new legislation.

Note:

Event is defined as a single measurement for a specific resolution (1h | 15m) for a single accounting point.

Batch is defined as multiple measurements (e.g. 1 day) with a specific resolution (1h | 15m) for multiple accounting points.

Meeting ended at 15:45.

Day 2

6 Datahub Event Channel Introduction

Remco presented the approach and result of:

- Study of capabilities - what does the current batch channel and environment can handle regarding the increases exchange of measurement data).

- Study of technologies – If Datahub is to receive individual measurement values per accountingpoint and interval (roughly 480 mio events/messages per day) what would be the best way to introduce that capability?

18. Question: Can DSO send the data right away?
    – Answer: DSOs can send the data immediately after they get it.

19. Question: Does it matter where data comes from in PoC?
    – Answer: No, but production data cannot be used. Datahub uses a “NodeJS” driver to send in high amounts of measurements per second for the Proof of Concept. The code, including instructions, will be shared with the workgroup participants for reference purposes.

20. Question: Should the data be anonymized?
    – Answer: No production data.

21. Comment: Focus has been on ingestion. Forwarding side stream channel has been out of scope in this minimum value product scope. Suppliers will not get the events as is. Event messages sent by DSO will be sent to suppliers as batch files. During Day 1 workshop was discussed that maybe forward side could be configurable.

22. Question: Can events be bundled on DSO side?
    – Answer: Yes. DSO can decide to send bursts or frequent events. Single message/event includes one accounting point and one value only.

23. Question: How big of a change it would be if DSO can send multiple values in one event?
    – Answer: Datahub would need to split the event -> More load to datahub

7 Tell & Share

Currently Phase 2 is in progress. Meaning Fingrid and CGI are executing two PoCs to validate the result of the technology study. Remco presented the progress on the PoCs and that current results show that Datahub, with the proposed enhancements, can process up to 50k events per second which should be enough to receive all measurements timely. Besides the progress on the PoCs the Test driver package was explained in more detail which will be shared with the workgroup participants for reference purposes.

24. Comment: Meter does not have all information shown in the example format compared to currently used format. Need to be decided where data can be enriched.

25. Question: If one event is one metering point and one value. Why multiple values cannot be put to one event.
    – Answer: It becomes quickly a batch interface. Kafka handles one event as one record. Datahub would need to split the events.

26. Question: Is there a roadmap for supplier/3rd party side streaming in datahub sending?
    – Answer: No on existing roadmap and no on MVP of event channel. That could be the logical next step.

8 Recap Day 1

Main conclusions of day one was presented by Remco for verification by the group. This would then decide the focus for the afternoon.

Main conclusions from day 1 were.

- Batch channel will most likely be used by many DSO and therefore requirements, guidelines and potential impact need to be determined to make sure this channel is ready for 1.1.2026.
- Timeseries Event Channel is seen as an important step to enable future developments and therefore is determined required from datahub perspective.
- Terminology like: Batch, Event and Continuous delivery is agreed upon.
    - Event is defined as a single measurement for a specific resolution (1h | 15m) for a single accountingpoint.
    - Batch is defined as multiple measurements (e.g. 1 day) with a specific resolution (1h | 15m) for multiple accounting points.
    - Continuous delivery is currently described as. Variable frequency (1 min - 6 hours). Standardised interval. (15m | 1h)

27. Comments of day 1 business requirement workshop conclusion:
    – One channel per DSO
    – Survey to DSOs who could be able to use the event channel and when.

- Focus to 6 hours law requirement after that we could focus on event (more frequent or continuous) channel. Event channel very good to have in the future. Parallel development.
- Define what is the business problem we are trying to solve as DSOs do not have many actual business needs and suppliers/3rd parties are not in the present in the workshop.
- Some DSOs would need to make batches to events for estimated values.
- Most of the DSOs would probably not be ready to use event channel 1.1.2026.
- Capability of the MDM of market parties for the more frequent delivery is a question mark.
- Focus should be on event channel development.
- During the workshop, definitions and descriptions were given which made it clearer what is meant with continuous delivery and possible consequences for a DSO to connect with Timeseries Event Channel. At the time of the previous survey this was not clear yet and therefore it is expected some DSO's might have a different opinion now regarding the exchange of measurement data at a much higher frequence. It is also becoming clearer this will affect retailers as the measurements are being distributed to them. A new survey for more detailed questions would help to get more insight in who will actually use the new channel if it becomes available and by taking into account the different market roles some questions tailored to the retailers could help identifying their needs.

9 Workshop Business Requirements

Event channel planning and priority was discussed first in three groups before presenting that to the common discussion with whole group:

- Group 1:
    - 2031 mandatory deadline for generation 2 meters. 2032 mandatory to send information via event channel to give transition period after 2031. Communication network may not be fully ready to support the event channel earlier. Bandwidth limitations.
    - How to define rules for meters that are a part of this law change but cannot deliver data as required (communication issue etc.)?
    - Event channel available 1.1.2026 for those who want to use it.
    - Reason for mandatory transfer to event channel?
        - Possible extra maintenance from two channels (e.g. if changes to current format)
        - No benefit if only few uses the event channel
- Group 2:

- Focus on batch channel to meet the law requirements. Time for survey.
- Continue event channel but at a slower pace.

- Group 3:
  - Event and batch channel available during 2026 from Datahub side.
  - 2028 - 2032 mandatory to send information via event channel.

28. Discussion:
    - Datahub should provide delivery frequency and message size instructions and DSOs can choose the delivery frequency.
    - max frequency 15 min
    - min frequency 3-5 hours
    - Report/view to DSOs on how many metering points have been delivered in 6 hours.
    - At least two vendors have indicated that there are interests to use the event channel from 1.1.2026.

29. Next steps:
    1. Workshop and testing day 24-25.9 and if needed separate online meetings for testing.
    2. Week 36 progress call for parties that will join the testing day.
    3. Two weeks upfront the next meeting in September the survey draft will be shared to workgroup.
    4. Two weeks upfront the next meeting in September the first draft of the technical specification will be shared with the workgroup.

10 Next meetings

24th and 25th September

11 Ending of day 2

The second day was about getting some focus on the consequences of implementing the 2026 legislation. During the workshop part of the day the groups focused on the planning of impacting the current batch interface and the introduction of the new Timeseries Event Channel. There were some differences between the groups regarding the prioritization of the event channel.

Even though not many DSOs are expected to actually use it to comply to 1.1.2026 legislation it is concluded that Timeseries Event Channel is preferred to be ready prior to 1.1.2026 or at the latest early in 2026. Workgroup did agree a deadline should be introduced for all DSOs to use the new event channel. This is conceptually set at the deadline

Katuosoite Läkkisepäntie 21 00620 Helsinki
Postiosoite PL 530 00101 Helsinki
Puhelin 030 395 5000
Faksi 030 395 5196
Y-tunnus 2745543-5, ALV rek.
etunimi.sukunimi@fingrid.fi
www.fingrid.fi

for the introduction of the generation 2 meters which is in 2031 and might have transition period (e.g. one year) after that.

This would allow the DSOs to start using the Event Channel between expected introduction in late 2025 and 2031.

Requirements for the Event Channel aren't clear yet and due to the focus on if the channel is needed and when it is needed there wasn't enough time to really determine the business, system and integration requirements for the interface. To keep the progress the following points are agreed on:

- Prevent datahub impact on the existing batch channel if possible. By changing the guidelines on how to use the channel system impact can be prevented. Line of thought for the guidelines and the impact on the batch channel:
    - Introduce average message size/ average measurements per message it is possible to keep the total number of messages down even when a day of measurement data is spread out in smaller portions to allow compliancy with the 6 hours.
    - Batch is not meant for single time intervals (e.g. every 15 minutes). So, periods of 3-5 hours are expected to present in the message. There will be no constraint on when these messages can be sent in (1min - 6hours) as long as they are sent in fast enough to comply to 6 hours.
    - Datahub can provide information on the % of measurement data being processed within 6 hours of the reading per DSO. DSO could use this information to decide if they are compliant by comparing the information with their generation 2 meter population.
- Workshop in September will be changed to focus on getting the system requirements signed of for batch channel and the timeseries event channel.
    - Draft Technical specification for event channel will be provided with a focus on minimal viable product. This might result in sub optimal working for distribution channel and so on.
    - Survey will be held to get more feedback on planning, distribution and scope.
    - Draft Revised Batch channel guidelines will be provided to propose the new guidelines in order to keep the impact for both DSOs and Datahub to a minimum.

Meeting ended at 14:24.

Appendices Appendix 1 in this document.
Appendix 2 Workshop Frequent Delivery metering data - August 2024 v2.0

APPENDIX 1

Group 1: (raised topics based on the pre-tasks)

- Metering system infrastructure is in DSO companies prepared for law requirements (6 hours)
- If there is delivery of data to end customers, then it shall be delivered to datahub also.
- No matter if meter is generation 1 or 2, there is always read the changes after delivery.
- 2031 mandatory deadline for gen 2 meters. 2032 mandatory to send information via event channel to give transition period after 2031. Communication network may not be fully ready to support the event channel earlier. Bandwidth limitations.
- To be checked if there is no data for all metering points, has Datahub mandate to define if rules to allow ones where is no need to have data?
- Event channel requires pretty much development from vendor.
- Is estimated data coming in same channel.
- There are always some batch processing needs.
- If there is no connection to meter, there is need for batch processing anyway.
- Should correction be a part of the event interface or not? Distinguish between these also regarding distribution?
- Will more recent data result in more use of the customer portal -> impact on usage.
- Enabler to provide flexibility services (via aggregator role)?
- Ingestion vs. distribution: Are recipients ready to receive more frequent or even streaming data? Not all parties benefit from real time data, but some do.
- Constraints:
    - Only ok values be included.
    - Industry has lots of legacy systems, procedures and data formats.
- Short term business requirements:
    - Fulfill law requirements.
    - More modern API than before (not SOAP, e.g. JSON)
    - Modern data format
- Long term business requirements:
    - from 6 hours to streaming
    - from REST to straight pipe
- Systems are designed to support the current delivery model and moving to continuous delivery require significant changes to various systems by 2026.
- End customer the opportunity to monitor their electricity usage more efficiently without the need to buy any additional devices. Another way to control devices (not just HAM port or datahub).
- Do the MDR/DSOs rely on an explicit acknowledge after sending a (set of) values?
- What measures does the DSO need to take in case of unavailability of the streaming channel?
- Need manage batch data delivery from gen 1 meters as batch parallel to more frequent data delivery from gen 2 meters.
    - where the info is needed if more frequent data delivery is possible (or not)
    - does consumer need the information of that?
    - customer services at least need it.
    - system delivering the data needs it.
- More frequent data delivery
    - what is the impact of more frequent data on estimation? Or on statuses used?

- Netting calculation to be done also more frequently?
- More frequent data delivery evens the load on the system.
- Data is faster available for the consumers/market parties.
- Streaming data -> real time aggregations? would BRPs benefit as well? These are no party at the moment (eSett)
- Has one invoice morel impact on the metering data delivery (no?)?
- How to tell to datahub that this accounting point is moving to continuous delivery (visible in datahub somehow)?
- Validation of data only in datahub?
- what validation included in datahub?
- datahub also calculating values?
- Raw data delivery without validation in MDM system?
- what status is used for raw data?
- Can all parties deliver continuously in future, mandatory for all?
- Continuous delivery for those parties that have capability for that.
- Measurement data available immediately for supplier/DSO or in datahub (also netted data)
- Is there need for two different ways to delivery metering data?
- netted data shall be calculated immediately and published and delivered to market.
- How to inform end customers which way metering data is delivered for that accounting point?

- **Summary** (and topics to be discussed)**:**
- What means continuous delivery?
- Two parallel ways to deliver data needed. Support both batch and new event channel to enable parties to step in when ready?
- Is there needed to store/manage data of metering point level delivery frequency (batch or continuous)?
- DSO do not have interest for continuous delivery, but suppliers and 3rd parties might have business cases.
- End customers have other way to utilize the continuous data than reaching it from datahub.
- Should only ok values accepted, and missing ones not even send in the event channel?
-

Group 2:

- DSO: Consumption forecast for next 2 weeks.
- Demand response management and services require real time measurements.
- DSO: Demand response forecast for tomorrow. Find the best way to optimize consumption and production.
- Preparation for more frequent resolution in future: 5 min, 1 min
- Real time consumption-based end customer communication: e.g. warning about larger than normal usage on high price time.
- Real time data to the exchange spot price-based customers about consumption
- MDM system should not have breaks. If MDM can collect and deliver data real time, it should be allowed. -> possible buffering in datahub end.
- Supplier should be able to choose, do I want measurement in real time or not.
- Restrictions:

    - electricity usage is changing, is everybody need to make better estimation ex. spot price usages.
    - system downtime -> lot of data when everything working after problem or maintenance.
- Overall: No bottlenecks due to big batches anymore
- Limitation: Cost vs. benefit, especially for DSOs
- BRB: Infra day trading /forecast
- Future dev idea: estimation and validation to be done in datahub.
- consumer: Understanding of what the invoice is based on.
- consumer: consumption “shaving”, peak shaving
- Short term:
    - more shrunken format more power full in processing
    - DSO doesn’t need to info if account is.
    - datahub working or not.
    - data delivery should be datahub based on DSO system.
    - no benefits from steam
    - short term packets can be simpler.
    - stream is more to events data which timeseries is not.
    - provide a better service to customers and market parties.
    - enable lighter system requirements.
- Supplier/3rd party impact when exchanging measurement data more frequently?
- What can be said about performance per minute in combination with downtime?
- Additional requirements needed? Does this mean a shift in responsibilities?
- What kind of insights does a DSO want to get regarding the processing of the timeseries event?
- are functional acks still relevant to XXX, order of processing.
- Is there a requirement that values for an accounting point interval could be send via both event and batch channel?
- What problem are we solving when metering timeseries are more frequently available?
- what requirements there are for updates of values?
- Long term:
    - provide a “real” time E2E data flow.
- Benefits:
    - DSO can get real time data to own purposes from datahub without own database: -> Long vision: data must be more than markets needs:
        - phase by phase
        - settable interval: 15 min -> 5 min -> 1 min
    - el. -> usage is changing -> from data can be made more real time calculation and analysis to DSO’s systems.
    - customers need more from data:
        - elect cars charger
        - production
        - usages changes -> what can tell to customer.
- What pattern/protocol is preferred? HTTP (REST)?
- What interval would we like to publish? 5 min – 15 min?
- How long should it take for a value to be available?
- What number of timeseries are we expect to exchange? Initial value + x updates?
- Consumer: Peak shaving when understanding of

- DSO: Demand response forecast for tomorrow.

- **Summary** (and topics to be discussed)**:**
- Concern:
  - How to get the DSOs together on using datahub to facilitate in interval data as input for their local processes.
  - Cost vs benefit
  - What problem we are solving?
- Long term:
  - Enable suppliers with more accurate and near real time.
  - Enable DSOs with a central real time measurement.
- Requirement:
  - No buffering on DSO
  - allow multiple.
- Ambition:
- Why do it local when it can be done central?
  - estimation
  - validation
  - etc.

Group 3:
- Limitation:
  - Netting and energy community calculations working in continuous data transfer: What is acceptable delay?
- Constraints:
  - Could there be situation that the data has already been send once, but datahub asks to send the data again for example last three days? What happens then?
  - How are errors patched in continuous delivery?
  - What should we do if datahub cannot take the readings in? Or datahub is down -> How should the data be stored?
  - Possible cost vs. benefits
  - if complexity rises too much cost to DSO/datahub can be too much
  - only new meters and depending on DSO contract not all meters' stream data the same way.
  - Hes/MDM which systems would do? Which role?
  - All MDM are very optimized for once-a-day intervals.
  - Current way of doing MDM should stay and streaming is extra.
  - Lobbying against the legislation because small DSOs find the system changes too expensive compared to the benefits.
- Limitations:
  - data collection: Telco networks require some randomized delays (for e.g. 0-15 min)
  - estimation in MDM system
  - calculation in MDM system (from cum values to usage)
  - are there negative effects on trust if values are delivered more frequently (more missing data visible)?

    - Benefit:
        - Continuous delivery could be lighter way to more data around.
        - although keeping track that what has been delivered already can be pretty hard. -> How could we keep track that that readings have been sent?
- Benefits per actor:
    - 3rd parties' access to data
    - Metering service provider: Trouble shooting?
    - DSO?
- Longer term visions:
    - data stream delivery method should be scalable for higher resolution based on DSO's goals.
    - if we really want to end user to have value from this the authorization should be easier for example through strong authentication and APIs
    - 5 min measurement interval
    - continuous aggregation
    - almost real time acting of electricity markets.
    - to supply datahub and end customers more data from their usage" Hey 10 min ago your L1 connection spied heavily."
    - smaller scale production unit usage data during the day
    - Raw data to datahub
    - Virtual HAN port from datahub
    - 1-5 min series (per phase)?
    - Flexibility + connected with relay commands?
    - data available -> get I when it suits you.
    - near real time monitoring of electricity consumption
    - new smart grid products
    - Load control products.
- Short term needs:
    - Fulfill the legislative requirements.
    - specification for: Data model, requirements for storing data/backups (datahub down)
    - possibility to deliver data when it is most suitable to each system without need to buffer heavily.
    - DSO: To comply with the law requirement without making it too complicated and costly.
    - supplier: possible to get intraday data for electricity market purposes.
    - TSO: comply with law and possibly get intraday data from whole customer base
    - good PR from the industry Justification for costs.
    - Customers can use the data to notice faults or problems "oh my stove was left on."
    - Enable measurement data delivery with current interface with more frequent delivery cycle.
    - Possibility to deliver measurement data from two different systems (by DSO)
    - Optimize current system and DSOs to the bundling to meet the requirements from the law. Streaming could be introduced when solution is ready and market parties can start to use streaming channel at their own place.
    - Parallel usage of the channels

Summary (and topics to be discussed)**:**

- Short term:
  - Law requirements need to be fulfilled.
  - Supplier needs to be found out.
  - DSO based options.
- Long term:
  - DSO options
  - end user data authorization.
  - virtual HAN port
  - resolution based on party's possibilities.
- Benefits:
  - 3rd party data
  - sending lighter
  - intraday forecast more accurate
  - aggregators
  - smart grid products
  - flexible hub
- Constraints:
  - streaming data is difficult to validate.
  - technology solution for two parallel technologies
  - management of parallel deliveries
  - error management and fixing of those.
  - HES/MDM roles and responsibilities
  - cost/benefit
  - netting and energy community calculations with stream data
  - Customer service load /trust