1. What I've Been Asked to Do

I was asked to look into what might go wrong when we move from using the current heartbeat table to a newer intervalstable for downstream reporting and analytics.

This is especially important for:

- People tracking **how much time users spend** in Worlds
- People watching **embodiment**, **presence**, **RUX/editor usage**, and other behaviors

The goal is to **stop using heartbeat**, but before that, we must make sure intervals is solid and safe for others to depend on.

---

2. What I'm Trying to Find Out

I'm looking into these key things:

- **Is the intervals table missing any data?**
- **Are the sessions it shows very different from heartbeat?**
- **Do time spent, sessions, or counts go down or up suddenly?**
- **Do the columns in intervals match what teams are using from heartbeat?**
- **Will reports and dashboards break if we switch?**
- **How do we backfill old data safely before switching?**
- **What data gaps or bugs might go unnoticed until it's too late?**

---

3. What Problems We Might Face

Here are some problems I'm expecting:

a. Missing data

Some older dates might not have clean intervals data. This could lead to under-reporting.

b. Sessions counted differently

Heartbeat logs every few seconds. Intervals summarizes into start and end times. So total sessions or time might look different.

c. Columns are different

Heartbeat has things like timestamp, while intervals has start_time and end_time. Dashboards doing joins may break.

d. Latency

Heartbeat data arrives fast. Intervals might be delayed. Alerts and live dashboards may stop being “live”.

e. Some metrics may no longer be accurate

Metrics like “active seconds”, or PDP engagement, could look off if we switch to intervals without tweaking definitions.

f. Risk of double-counting

If someone uses both heartbeat and intervals at the same time by mistake, numbers could be inflated.

---

4. What I Need From the Team

To do this job properly, I need help from a few teams:

From Data Platform / Infra:

- How far back is intervals data guaranteed to be accurate?
- What is the delay (latency) between real-time and intervals table updates?
- Are there edge cases where intervals does not capture something heartbeat does?

From Metrics / Dashboards Owners:

- Which dashboards or metrics currently use heartbeat?
- Are they sensitive to latency or small metric changes?
- Are joins using heartbeat timestamps or session IDs?

From Analytics / Product:

- What are the **top 10 use cases** of heartbeat today?
- Is anyone doing manual aggregations that need to be updated?
- Are teams already testing intervals somewhere?

---

5. Where I Plan to Start

Here’s my plan of action:

1. **Understand the schema** of both tables – heartbeat and intervals

2. **Pick a few key metrics** like session count, time spent, DAU, and see how they look when pulled from both tables
3. **Do a backfill test**: Try pulling intervals data from June 2024 to now, compare with heartbeat
4. **Create a comparison sheet or dashboard**
    - Total sessions
    - Avg session length
    - Time spent per user
5. **List all mismatches** and reach out to teams to ask: “Will this affect you?”

---

6. Questions I Will Be Asking

Here’s a running list of questions I’ll be asking across teams:

1. Is intervals data trustworthy before a certain date (e.g. June 2024)?
2. Are there any edge cases where intervals drops short sessions that heartbeat would capture?
3. Are intervals always available per world? Or are some worlds missing it?
4. Does embodiment or RUX rely on timestamp-level precision heartbeat gives?
5. Is the intervals table updated frequently enough for near real-time use?
6. Is there any audit or data quality dashboard already running for intervals?
7. Who are the top consumers of heartbeat table today?
8. Can we agree on a cutoff date for when everyone should switch?