For this case study I would need a bit more reading/consulting, but unfortunately I’m running out of time. Below is starting point of my thought on this topic.

I would try to make balance between expensive/safe/fast solution (for example AWS, Google could…) and less expensive/safe/fast solution (like Hetzner dedicated servers).

I would try to separate incoming data into several groups:

1. Critical data

In short this is data which can not be recreated if lost (game logs), and data which is needed in real time

* Critical data in term of availability (data needed in real-time for game running). I would expect that these databases are relatively smaller, so if needed some in memory solution could be recommended. Due to importance of this data, I would rely on some AWS, Google cloud solution. If this data should be calculated from some other database, I would rely on cloud infrastructure
* Not so critical data in term of availability (data used for analysis of game activity)

game logs/player activity data –some not in memory cloud solution seems to be suitable due to it reliability in term of storing and backups.

1. Not critical data

Data which can be recreated (aggregations from game logs) and data which is not needed real-time

Data for inhouse analyses and data science could be kept and processed through dedicated servers