Chapter 4: Patterns (Part 2)

Bucket Pattern

Back to the Question

Correct Option:

- Our system ingests thousands of log lines each day for each host it monitors.

  Trying to keep one document per host or one document per log line may not work very well. Keeping a group of log lines per host sounds like a good design, however, knowing the details about the use cases and queries will help to confirm that this is a good model.

Incorrect Options:

- Our system handles 1 million IOT devices.

  This is not sufficient to say that we should be using this pattern. If the devices are not sending data on a regular basis that need to be grouped, then we would not need the pattern. For example, we may just have to keep track of the device, its configuration, and its state.

- Our system ingests 10 million pieces of data per day from 1 million devices, with 20% coming from 10 devices.

  The issue seems more related to the fact that a minority of devices have different behavior than the majority of devices. Consider looking at the Outlier pattern before opting to use the Bucket Pattern in this situation.

- Our system must embed a one-to-many relationship in one of our models, however, some of the result documents would be too big.

  There is not enough information to tell us how we could group the many side of the relationship into buckets. The most likely optimal action to take is to link the relationship instead of embedding it.

- Our system performs sums and averages over all elements of certain arrays.

  It sounds like we should look into the *Computed Pattern* instead.

Proceed to next section