elastic

Daily Elastic Observability B(y|i)te
Index Lifecycle Management & data streams

David Pilato (@dadoonet)

Before data streams

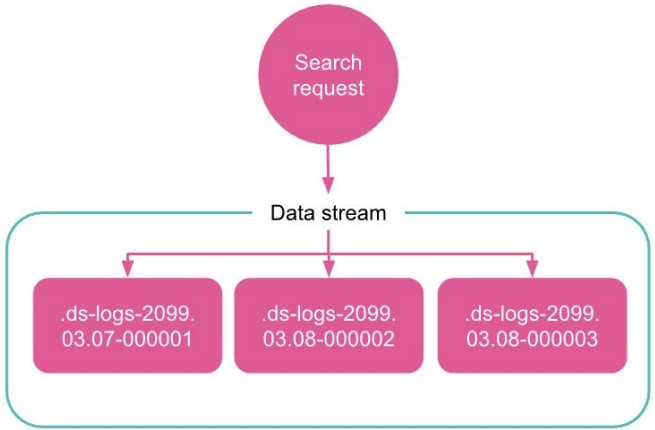
- One single index for a use case (for example **logs**)
  - mixing **nginx** logs with **smtp** logs (different fields)
  - with thousand of fields (10k limit)
  - most fields are empty (bad compression because of sparsity)
  - very big index mapping

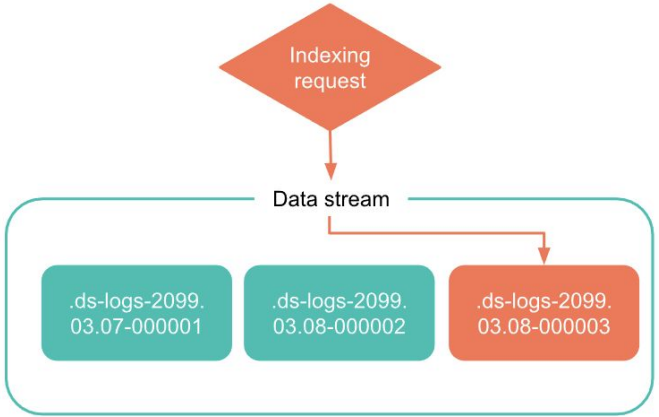
Data streams

- One datastream per dataset
- Naming convention **<type>**-**<dataset>**-<namespace>
  - **logs**-**nginx**-default, **logs**-**smtp**-default
  - **metrics**-**linux.iostat**-default, **metrics**-**system.diskio**-default
- Data streams **must** have a matching **index template**
- Documents **must** contain the **@timestamp** field (configurable)
- for **append-only** time series data

Data streams

- backing index **.ds-<data-stream>-<yyyy.MM.dd>-<generation>**
- rollover and alias
  - many search indices
  - one single write index





elastic

Before ILM

- Time based indices like **filebeat-2021.10.28**
  - some are very small
  - what happens for big events (like christmas)?
  - what happens when you have unpredictable traffic?

ILM

- an ILM policy:
  - Might use rollover
  - Might be moved to warm, cold, frozen phases
  - Might be removed
- ILM can be applied to data streams

In action