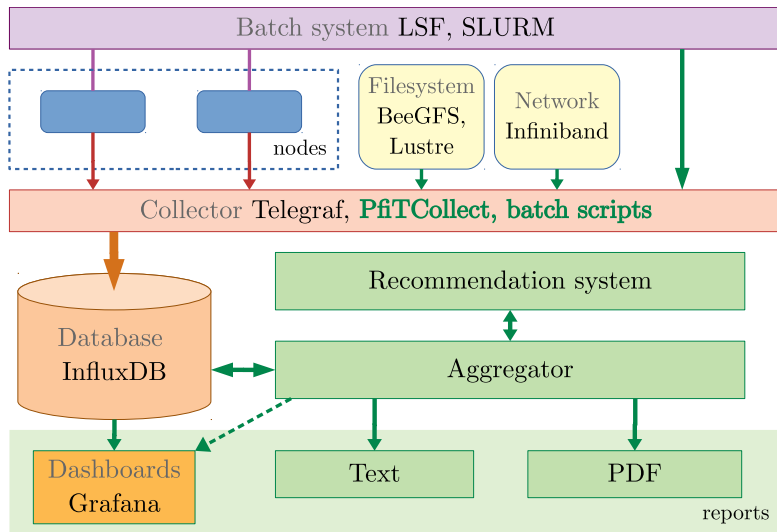


# Architecture



## Main focuses



- shared env** we gather process based metrics which allow to use the toolkit with shared nodes
- analysis** the recommendation system analyzes gathered metrics to deliver comprehensive performance information
- lightweight** components of the toolkit like *collectors* and *report generators* are designed to be lightweight
- reports** multiple report types (*text*, *pdf*, *dashboards*) ensure that users get convenient depth of information based on their requirements

## Nodes

- CPU
- Memory
- System

## Process

- CPU times
- Memory
- IO

## Filesystem

- Lustre
- BeeGFS
- Local

## Others

- Job
- GPU
- Network (Infiniband)

## Works with high load

- 1 month retention policy
- 10 seconds interval
- 400 GB of data
- 500 nodes
- 25 million series

## Another high load scenario

- 1 week retention policy
- 1 minute interval
- 70 GB of data
- 2000 nodes
- 7 million series

## Definition

Recommendations are suggestions for users to help with improving efficiency of their jobs

Recommendation system is designed to be completely separated from other parts of the toolkit. Only input data should be provided.

