From Field Data to DC Quality: Data-Driven Quality Ecosystem

PRESENTER

Youngjune Kang youngjune.kang@sk.com

Co-Author. Beomseok Choi, Inho Cho

Corresponding Author. Myoungkwan Cho, VP of SK hynix

Agenda

- **Background**
- **Data-Driven Quality Ecosystem (DDQE) Introduction**
  - Quality parameters from data center :
    - Failure Rate
    - Drive Operation Log / Drive Health Log
    - Quality Problem List
  - Feedback to data center
    - Automatic outlier detection
    - Concurrent SSD data log analysis and feedback
    - Failure Prediction based on SSD log data
- **Things to note to move forward**

Background

- As the scale of data center increases, eSSD volume in data centers is significantly increasing. And usage type of eSSDs become more diversified.
- To guarantee stable quality and provide concurrent and transparent quality support, SK hynix introduces a **DDQE(Data-Driven Quality Ecosystem**).

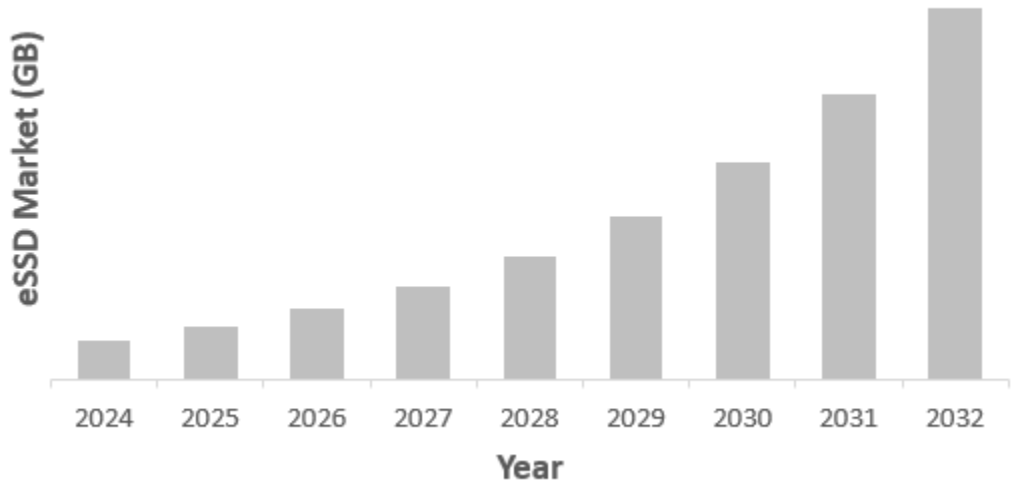


Fig 1. eSSD Market Growth Outlook

Applications
- Cloud Service
- AI
- Storage
- Computing

Location
- Latitude
- longitude
- Etc.

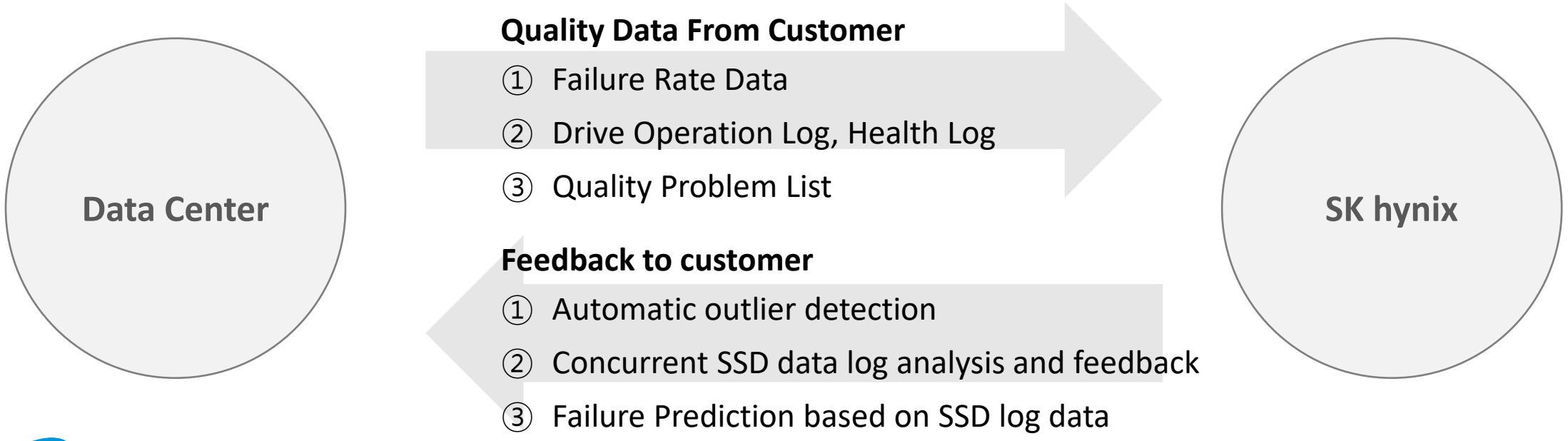
Server
- ODM sites
- Server Gen.
- Operating System

SSD
- TLC/QLC
- Interface
- Controller
- Density

Fig. 2 Various usage conditions of SSDs used in data centers

Data-Driven Quality Ecosystem for customer support

- SK hynix established the DDQE.
  - To detect outlier from failure rate along worldwide region, workload, application and platform.
  - To analyze the SSD health data during operation in data center.
  - To predict the potential failure in advance.
- All process/data is customized for each customer and managed as confidential.



Failure Rate Monitoring

- DDQE can detect the quality problems across region, application, workload and platform, and feedback to data centers concurrently.

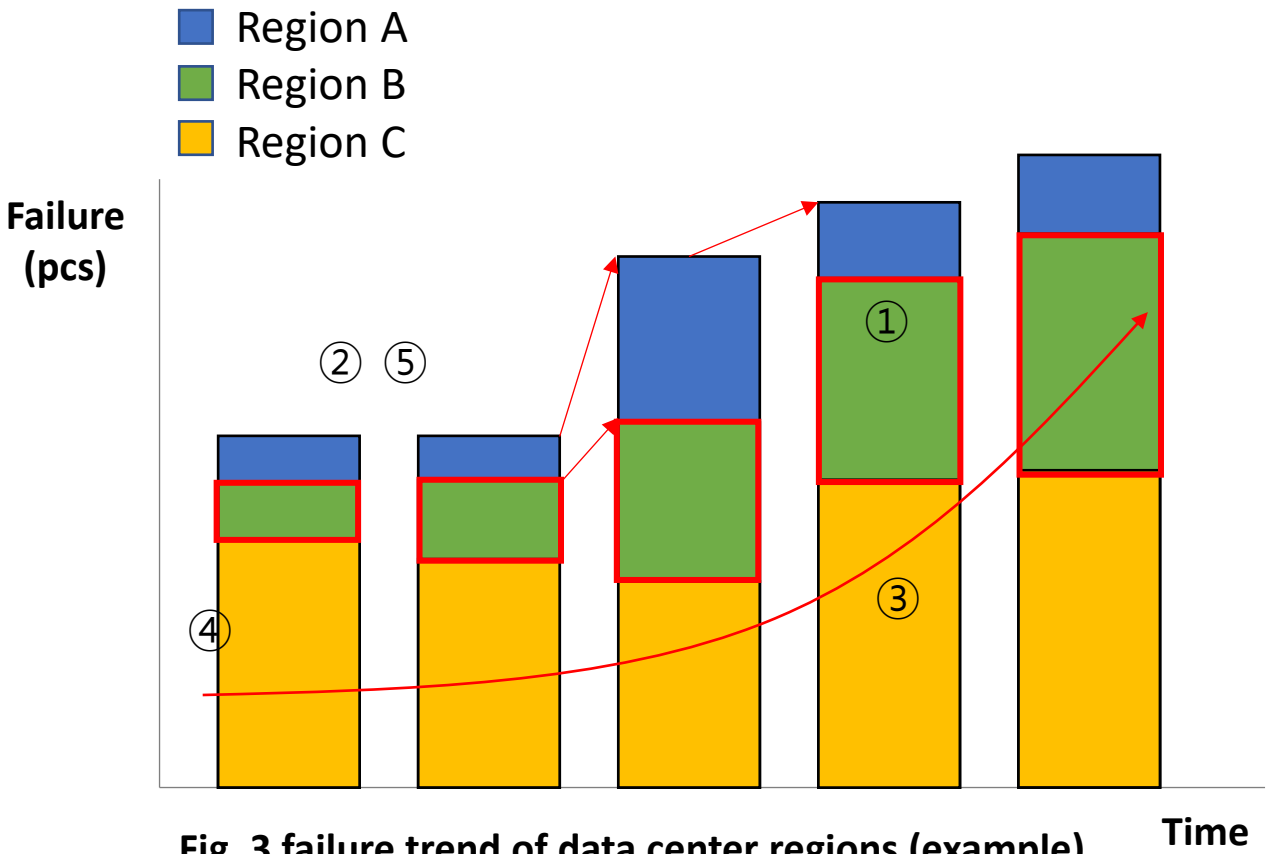


Fig. 3 failure trend of data center regions (example)

[Failure Symptoms : Automatically Detectable]

① Continuously Increasing

Failure rate continuously rise over a “n” period.

② Significantly Increasing

Failure rate increases above criteria.

③ Uptrend in long term period

Failure rate of this week is larger then past weeks.

④ Failure concentrated in certain group

If *Cramer V coefficient between groups is greater than criteria, it indicates a abnormal symptom in certain group.

⑤ Failure portion change

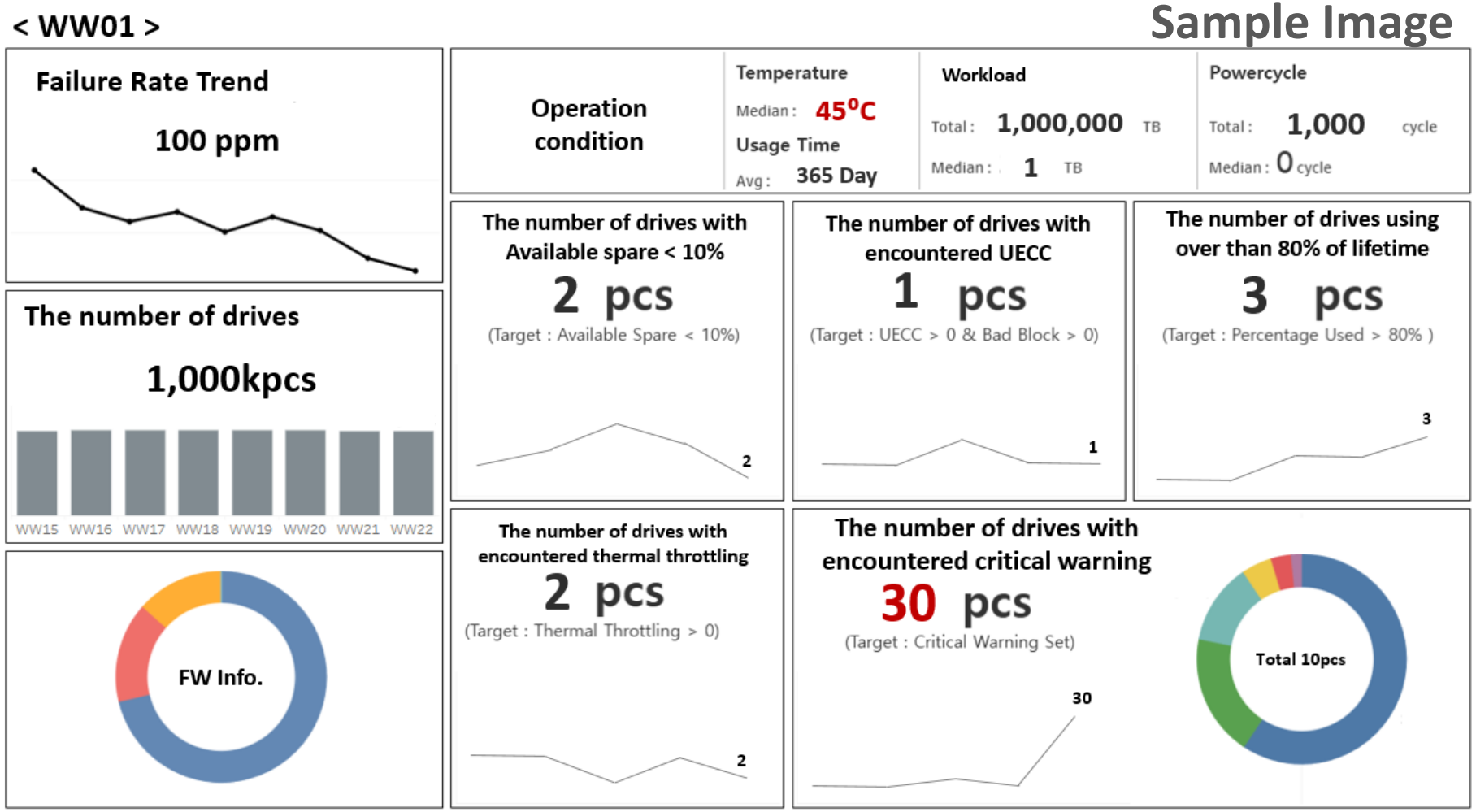
If Cramer V coefficient between the group is significantly changed from last week, it indicates abnormal change occurred.

*Cramer V Coefficient : Association measurement between multiple nominal values.

Drive Log Monitoring : Operation summarize

- DDQE summarizes the behavior of SSDs by analyze drive log weekly.

Sample Image



Failure Prediction from drive operation log and health log

- DDQE can predicts potential failure through the ML/DL model by training the good/bad drives log data.
- Data center can replace the potential failure drives before real failure encounter. : improve service interrupt.

Normal Group

Day 1 : Operation / Health log

Day 2 : Operation / Health log

⋮

Day 60 : Operation / Health log

Bad Group

Day 1 : Operation / Health log

Day 2 : Operation / Health log

⋮

Day 60 : Fail !

Time Series
Deep Learning Model

Predict
Potential Failure Drive

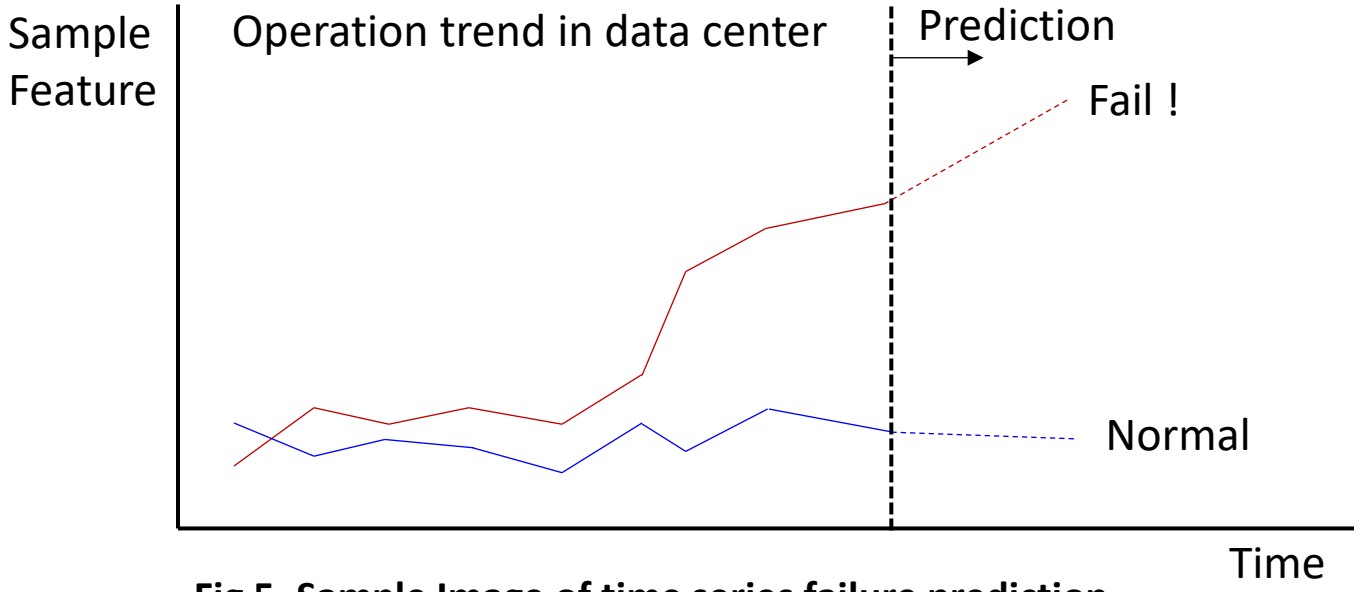
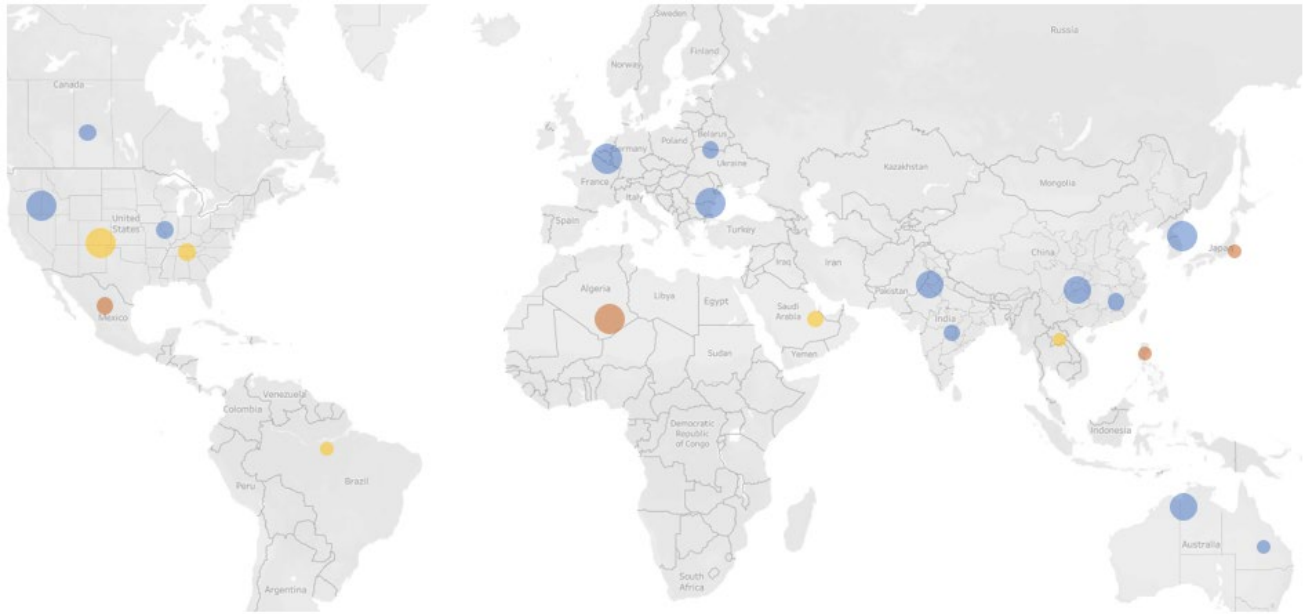


Fig 5. Sample Image of time series failure prediction

On time dashboard for quality problem

- DDQE can also provide the dashboard with data centers quality problems.
- Data center can manage the quality problems with SSD supplier synchronously.



- Location : Europe, Asia, ..
- Application : Cloud, AI, DB, …
- Problem example : FW update Fail, no available spare, …

Fig 6. Various issues from various group

On time dashboard for quality problem

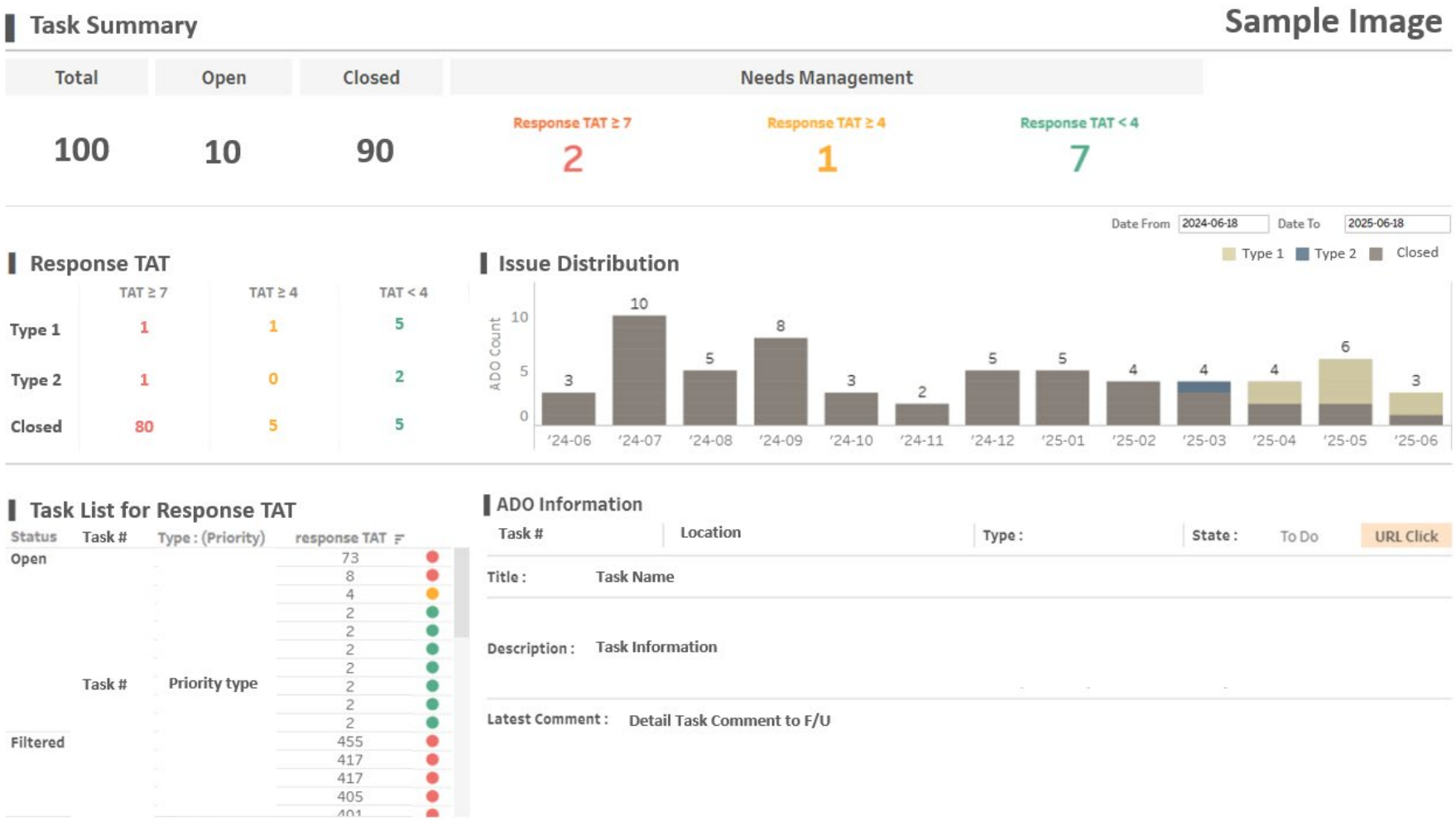


Fig 7. Task support dashboard for data center

Summary

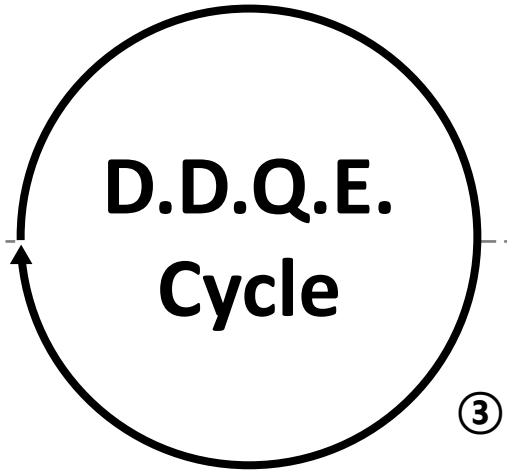
- The function of DDQE system in data center life cycle.

① Normal Operation Phase

- Monitor SSD operation condition and failure rate across region, application, workload and platform

② Failure Symptom Occurred Phase

- Early detect abnormal symptom.
- Analysis operation log and health Log.



④ After Problem Phase

- Train the failure symptom to failure prediction model.
- Predict potential failure before real failure encounter

③ Quality Problem Handling Phase

- On time Dashboard can support data center to manage quality problems with SSD supplier synchronously.

Things to note to move forward

- As data center's application being diversify with future technology, Drive supplier needs more elaborated data from data center.
- Strong collaboration between data center and supplier is inevitable
- SK hynix is ready to cooperate with customers.

Author

- First Author : Youngjune Kang, youngjune.kang@sk.com
- Co-author : Beomseok Choi, beomseok.choi@sk.com, Inho Cho, inho25.cho@sk.com
- Corresponding Author : Myoungkwan Cho, VP of SSD/Solution QRA, myoungkwan.cho@sk.com

Booth #207

Meet the future of memory.
Just steps from the entrance.

Innovation starts here,
Literally.

SK hynix