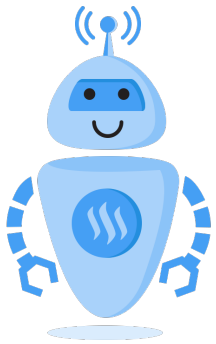
CEPHALOCON APAC 2018

THE FUTURE OF STORAGE

22-23 March 2018 | BEIJING

# Automatic Operation Bot for Ceph

You, Ji

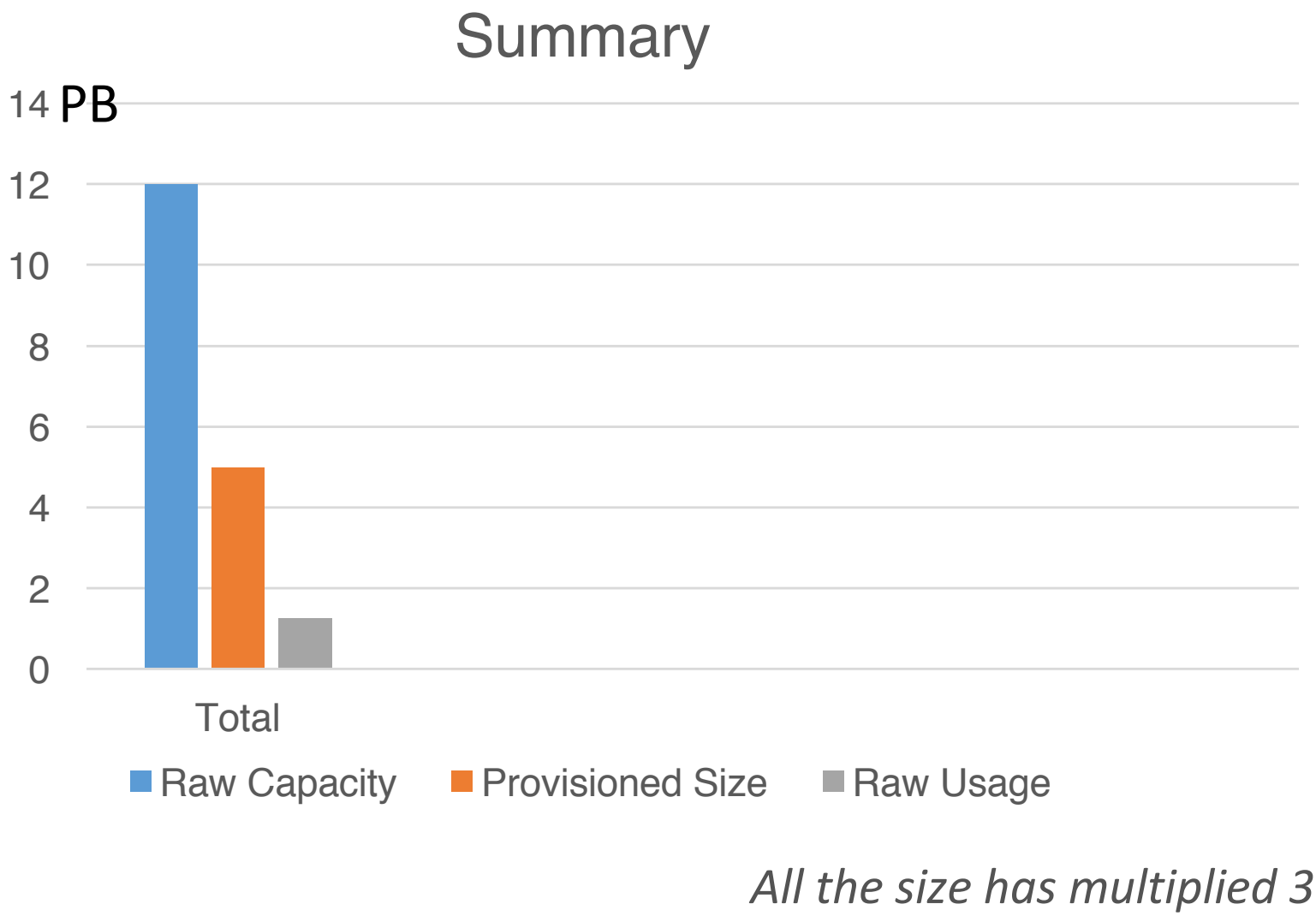


- **Ceph in eBay**
- Ops for Ceph Cluster
- Automatic opertion Bot

| Cluster | Capacity | Nodes | Disks | SSD pool | Raw Usage | Provision Usage | Used IOPS (Avg/Max) |
|---|---|---|---|---|---|---|---|
| LVS01 | 1.4PB | 18 | 400 | 0 | 4.6% | 13% | 1.10K/7.5K |
| LVS02 | 3.24PB | 54 | 950 | 3.3T | 12.7% | 60.7% | 11.5K/24K |
| PHX02 | 3.49PB | 55 | 1000 | 6.6T | 8.8% | 41.9% | 9K/16K |
| SLC07 | 3.77PB | 58 | 1100 | 6T | 12.5% | 36.4% | 14K/41.3K |
| CAL CephFS | 18T | 36 | 36 | 18T | 19.19% | 29.6% | 3K/5K |

*3 copies for all cluster*

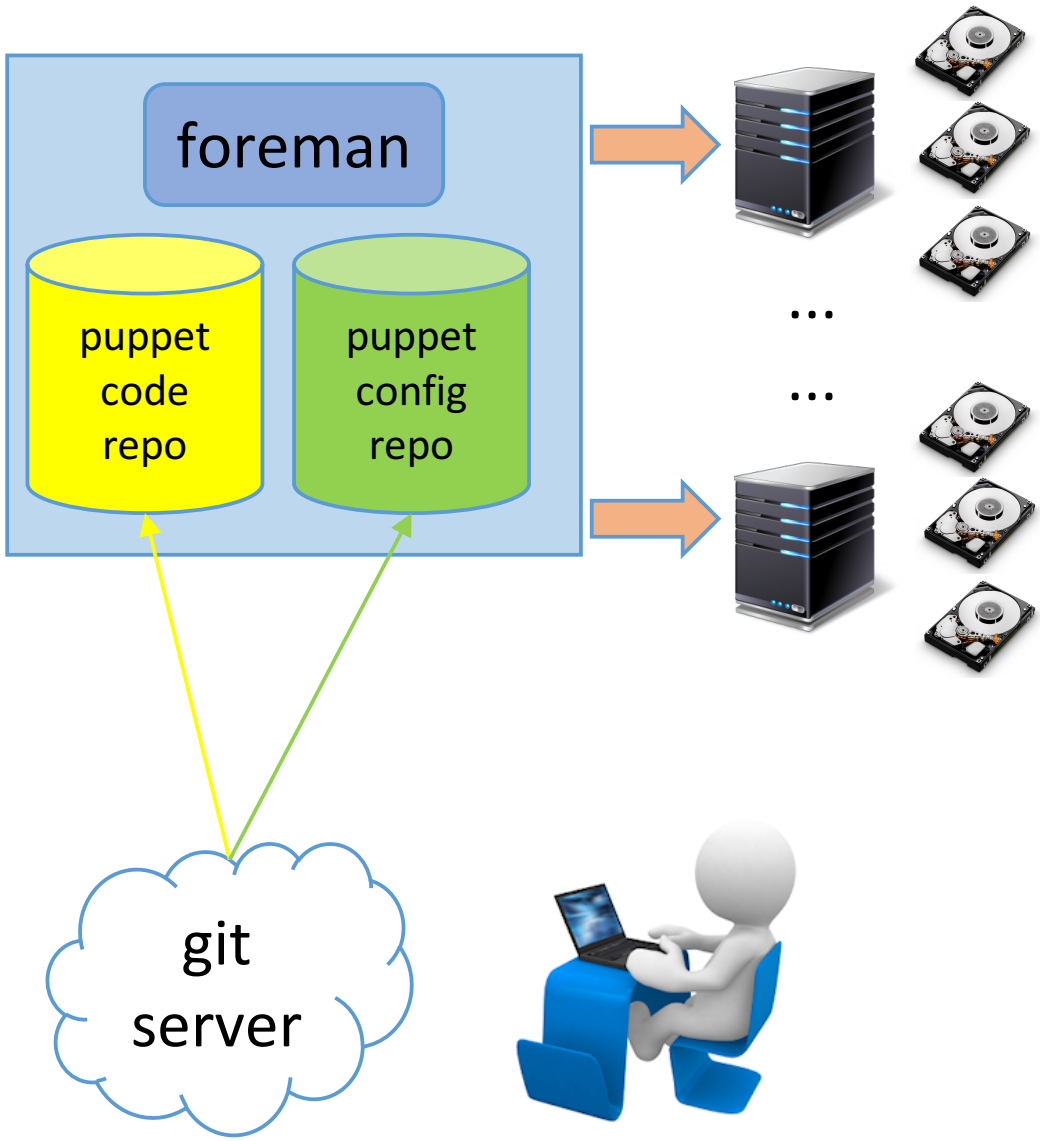
ebay™




*All the size has multiplied 3*

- Ceph in eBay
- **Ops for Ceph Cluster**
- Automatic opertion Bot

# Deploy




- *Use code to deploy*
- *Differences kept in Config Repo*
- *puppet auto add free disks*
- *NO auto weight for new OSD*

# Monitoring & Alerts

AlertManager

*Grafana*

Prometheus *local disk*

Prometheus *remote disk*

ceph-exporter

pool usage

cluster availability

cluster reliability

PG status

*ceph admin node*

node-exporter

admin-sockets

megacli

CPU

Mem

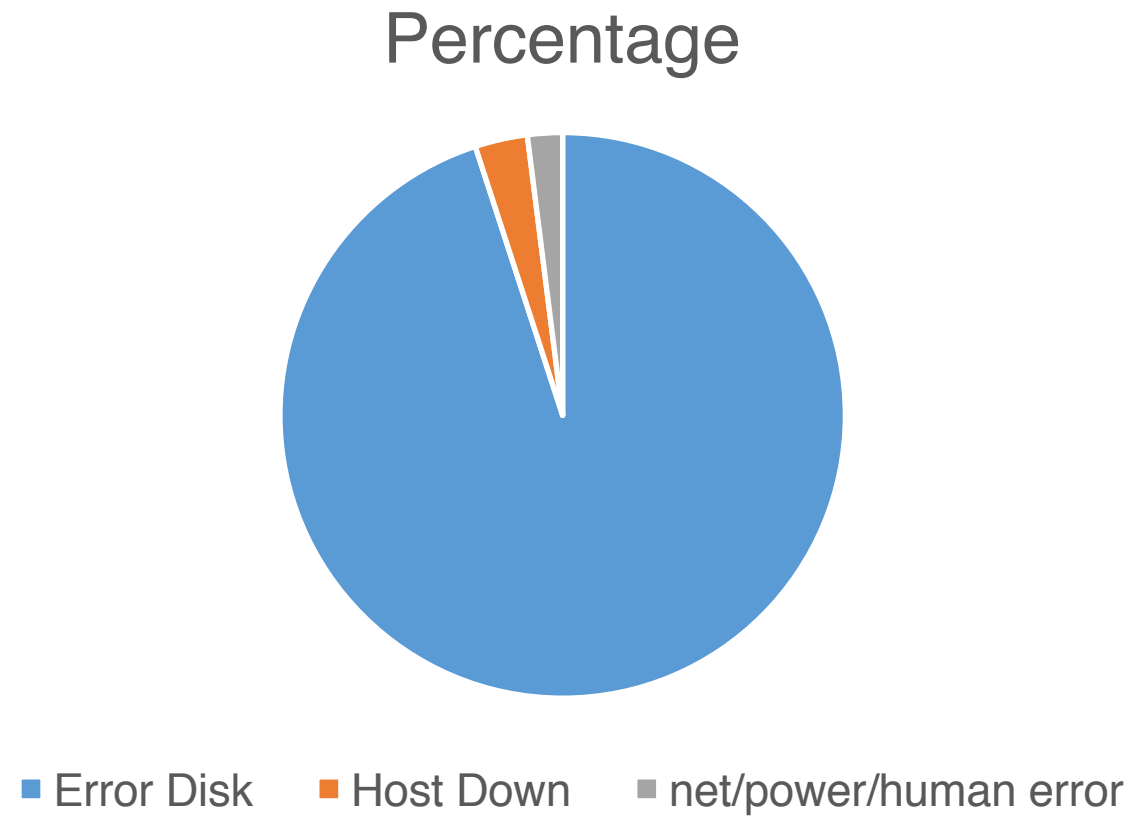
Disk

Net

iostat

*ceph node*

# Remediation

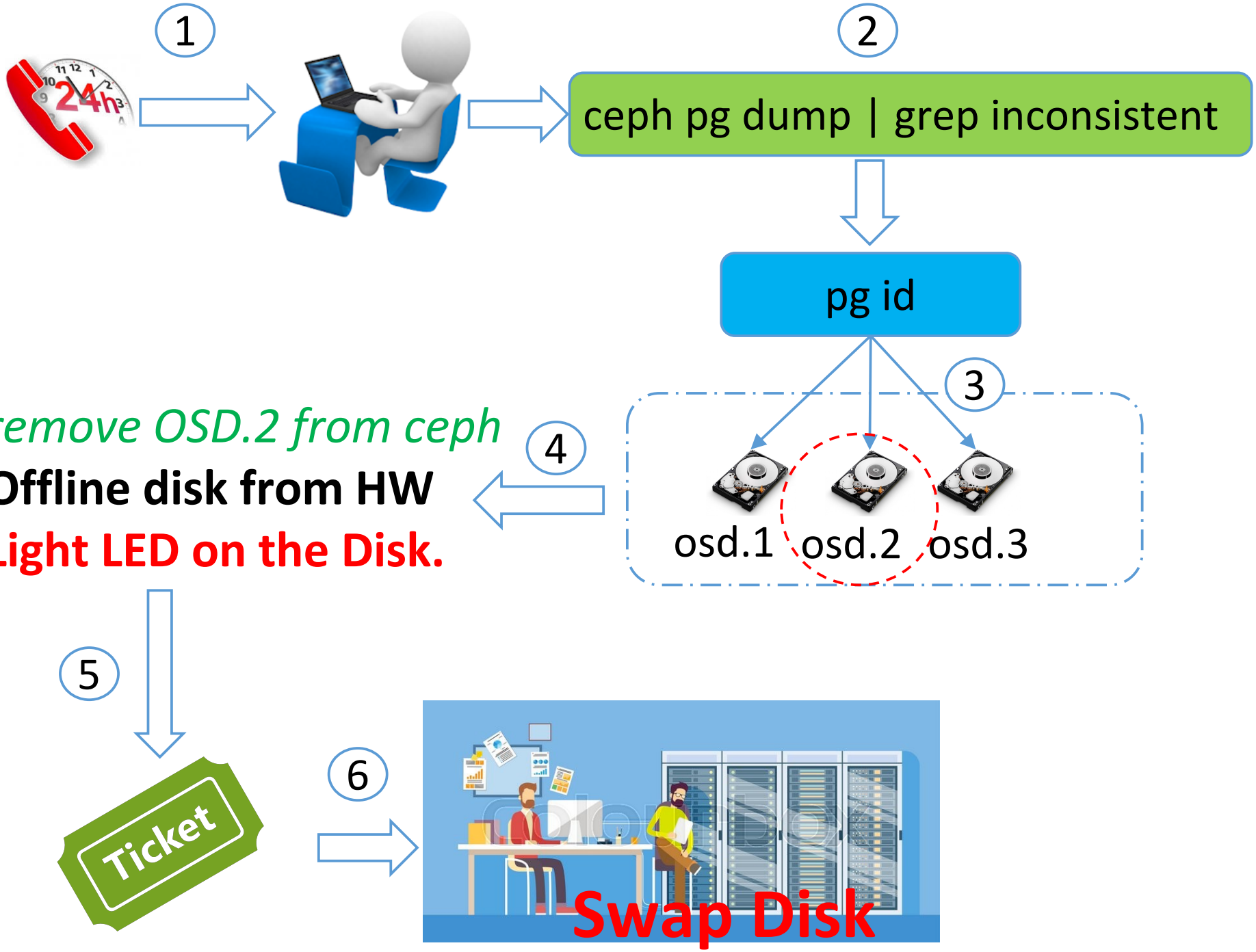



- *~ 95% warning/error caused by bad disks.*
- *~3% caused by host down*
- *~2% other errors, such as network, electric power, human error*

- Ceph in eBay
- Ops for Ceph Cluster
- **Automatic opertion Bot**

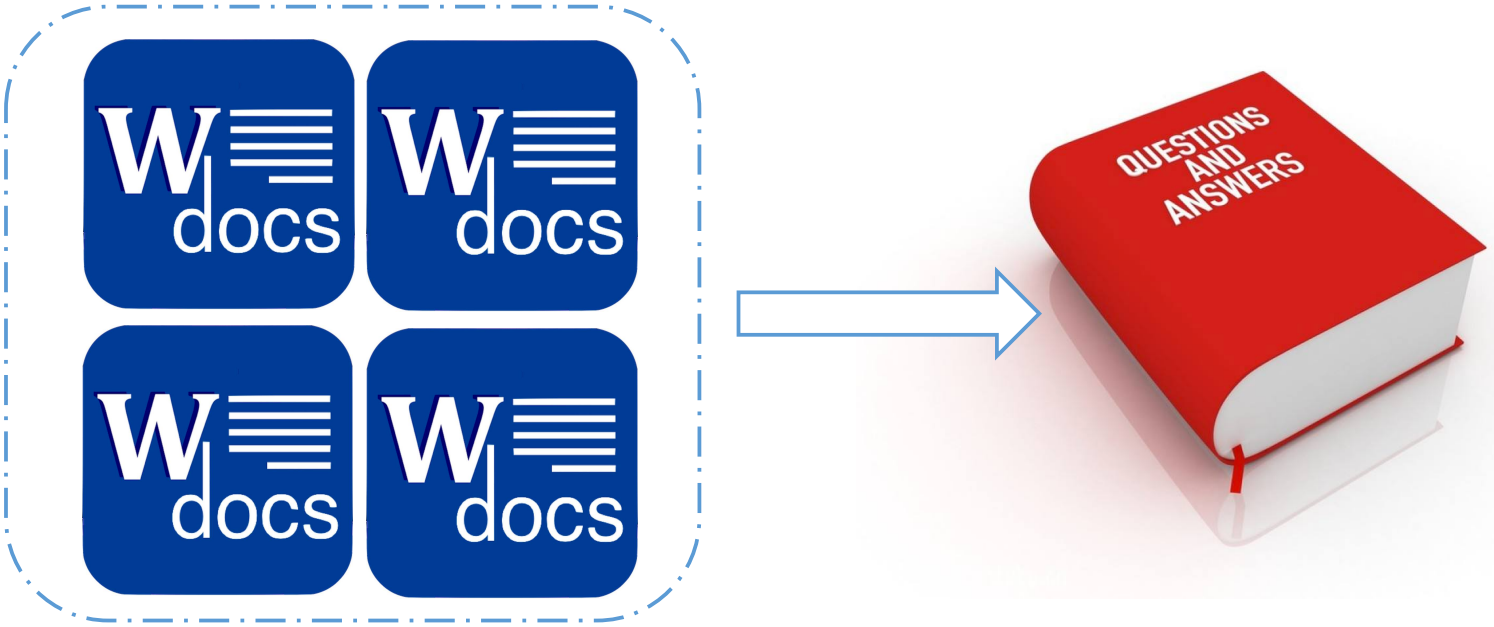
# How to build a automatic bot?

- **Handbook for operation**
- Interactive chat bot
- Automatic operation bot


1
2
ceph pg dump | grep inconsistent
pg id
3
osd.1
osd.2
osd.3
4
$ remove OSD.2 from ceph
$ Offline disk from HW
$ Light LED on the Disk.
5
Ticket
6
Swap Disk

# Write operations on the doc.

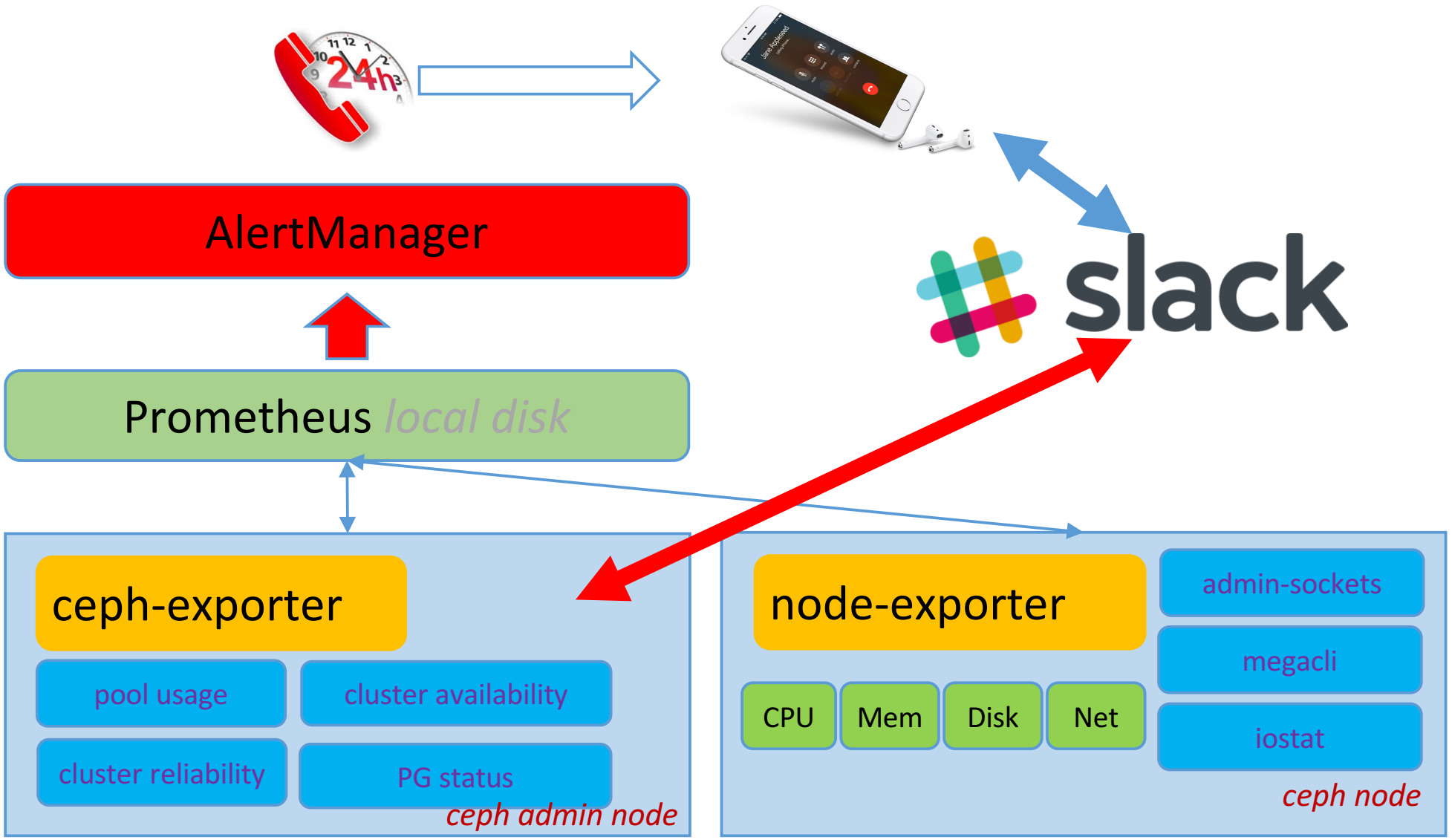
- *Record every failure of disks/node and related docs.*



# How to build a automatic bot?

- Handbook for operation
- **Interactive chat bot**
- Automatic operation bot

# Monitoring & Alerts



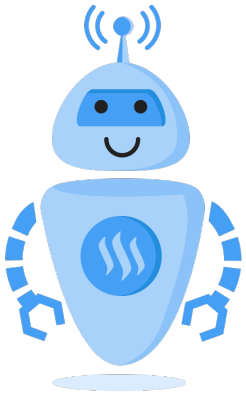
# Demo

# How to build a automatic bot?

- Handbook for operation
- Interactive chat bot
- **Automatic operation bot**

AlertManager

slack



Prometheus *local disk*

ceph-exporter

pool usage

cluster availability

cluster reliability

PG status

*ceph admin node*

node-exporter

CPU

Mem

Disk

Net

admin-sockets

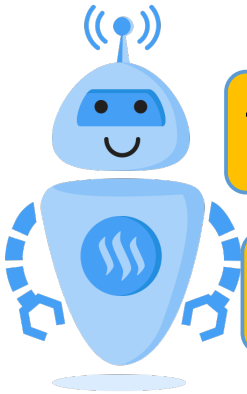
megacli

iostat

*ceph node*

AlertManager

slack



Task Queue

Executor

Prometheus *local disk*

ceph-exporter

pool usage

cluster availability

cluster reliability

PG status

*ceph admin node*

node-exporter

CPU

Mem

Disk

Net

admin-sockets

megacli

iostat

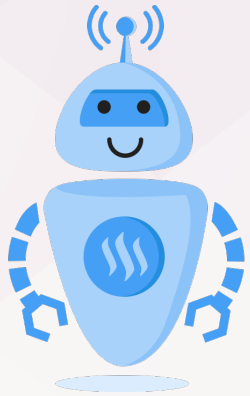
*ceph node*

CEPHALOCON APAC 2018

THE FUTURE OF STORAGE

22-23 March 2018 | BEIJING

# Q & A

Thanks

youji@ebay.com