Centre de Calcul de l'Institut National de Physique Nucléaire et de Physique des Particules

Manage your hardware failures

(Mattieu Puel, Hepix Spring '17)

CCIN2P3

Manage your hardware failures in an (almost) automated workflow

(Mattieu Puel, Hepix Spring '17)

CCIN2P3

Preliminary notes

Straight to an example

Architecture concerns

Some figures

Evolutions

Conclusions

Are you concerned ?

May ring a bell if :

- your own people handles hardware failures (from detection to resolution)
- your interactions with support teams are time consuming :
  - filling in incidents into vendor ticketing systems
  - providing diags based on specific support requests
  - answering satisfaction polls
  - **even for the most simple case**

ONE word about DELL

DELL EMC TechDirect

Mattieu Puel

Dashboard | Services | Utilities | Reports | Documents | Info Center | Message Center 220

Technical Support
Self-Dispatch
APIs
SupportAssist
ProDeploy Client Suite
Certification & Authorization

Dashboar

Customize Your Dashboard

Quick

220 New Message(s)
Support Request
(Add Action)
(Add Action)
Manage Actions

Technical Support

Create and monitor your requests.

Visit Technical Support

Self-Dispatch

Create and monitor your Self-Dispatch requests, check warranties and access self-dispatch online training.

Visit Self-Dispatch

APIs

View and request APIs.

Visit APIs

Certification & Authorization

Schedule individual exams, manage company progress to achieve competencies.

Visit Certification & Authorization

Administrator Control Panel

Manage key capabilities such as Technician Users, Address Book, Join Requests as they relate to

https://techdirect.dell.com/Portal.30/Services.aspx

Message Center

Messages 220 | Alerts

L	Dispatch Processed This message is to alert you a recent parts dispatch request has been processed. Work Order #: SR946638571. Request date: 4/10/2017. Service Tag: 6Q4Y102. Reference / PO #: . DPS #: 7325440511. (033-111716)	4/10/2017 3:55 AM
L	Dispatch Processed This message is to alert you a recent parts dispatch request has been processed. Work Order #: SR937114664. Request date: 10/2/2016. Service Tag: 35LQMW1. Reference / PO #: . DPS #: 7321671247.	4/9/2017 9:29 PM
L	Dispatch Processed This message is to alert you a recent parts dispatch request has been processed. Work Order #: SR936436793. Request date: 9/21/2016. Service Tag: DP6T35J. Reference / PO #: . DPS #: 7321316311.	4/9/2017 9:28 PM
L	Dispatch Processed This message is to alert you a recent parts dispatch request has been processed. Work Order #: SR936436126. Request date: 9/21/2016. Service Tag: 5L5XR12. Reference / PO #: . DPS #: 7321312996.	4/9/2017 8:35 PM
L	Dispatch Denied This is to alert you that we were unable to process a recent parts dispatch WO8327278. Request Date 9/1/2016 Service Tag F1MZTB2 Reference / PO # Please review the details and resubmit your request online	4/9/2017 6:20 PM

Showing 1 to 5 of 220 entries

Mark all as read
Preferences

ONE word about **DELL**

DELL EMC TechDirect

Mattieu Puel

Dashboard | Services | Utilities | Reports | Documents | Info Center | Message Center 220

Technical Support
Self-Dispatch
APIs
SupportAssist
ProDeploy Client Suite
Certification & Authorization

Customize Your Dashboard

Quick

220 New Message(s)
Support Request
(Add Action)
(Add Action)
Manage Actions

Visit Technical Support

Self-Dispatch requests, check warranties and access self-dispatch online

Visit Self-Dispatch

APIs

View and request APIs.

Visit APIs

Certification & Authorization

Schedule individual exams, manage company progress to achieve competencies.

Visit Certification & Authorization

Administrator Control Panel

Manage key capabilities such as Technician Users, Address Book, Join Requests as they relate to

https://techdirect.dell.com/Portal.30/Services.aspx

Vendor SOAP APIs

Message Center

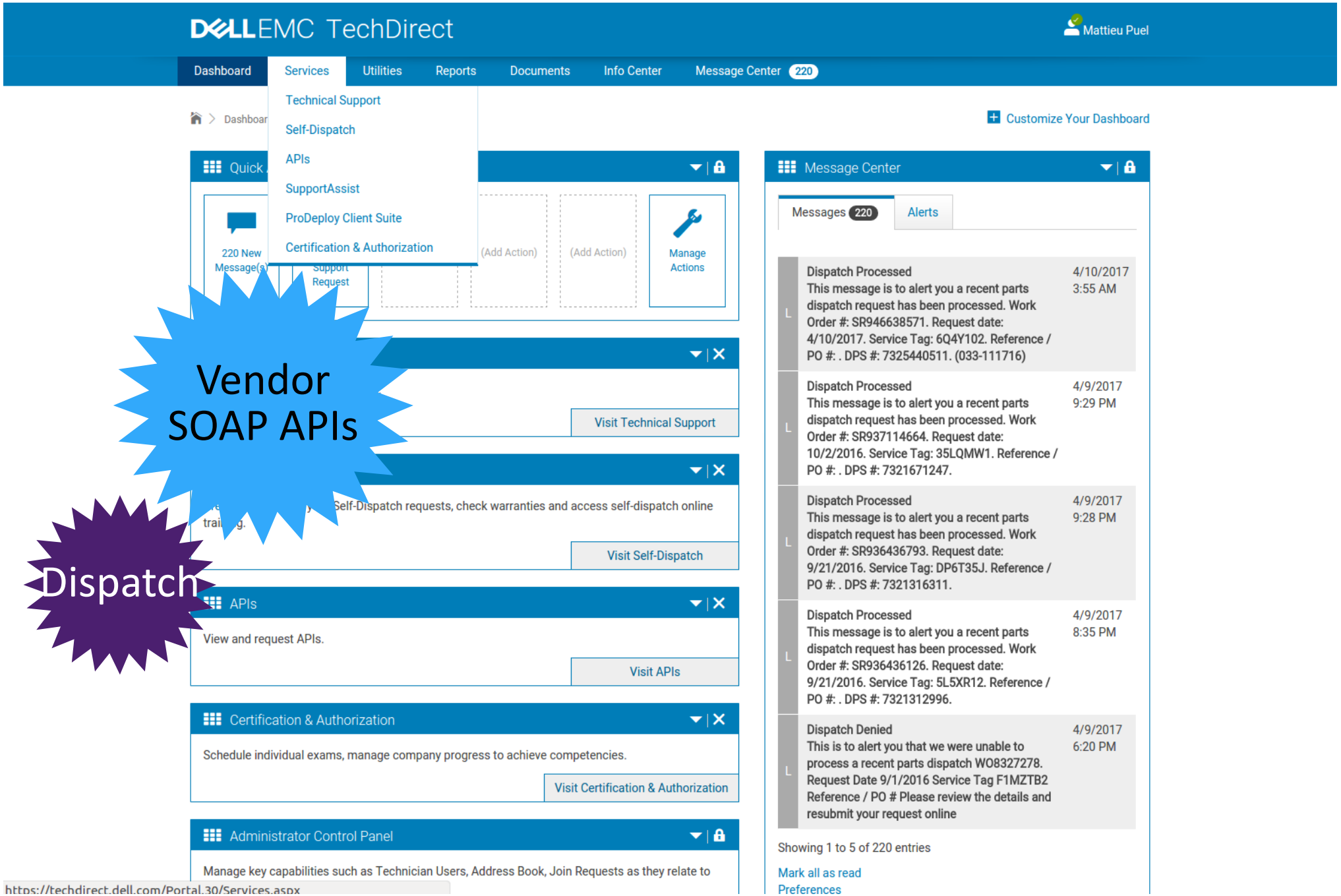
Messages 220 | Alerts

L	Dispatch Processed This message is to alert you a recent parts dispatch request has been processed. Work Order #: SR946638571. Request date: 4/10/2017. Service Tag: 6Q4Y102. Reference / PO #: . DPS #: 7325440511. (033-111716)	4/10/2017 3:55 AM
L	Dispatch Processed This message is to alert you a recent parts dispatch request has been processed. Work Order #: SR937114664. Request date: 10/2/2016. Service Tag: 35LQMW1. Reference / PO #: . DPS #: 7321671247.	4/9/2017 9:29 PM
L	Dispatch Processed This message is to alert you a recent parts dispatch request has been processed. Work Order #: SR936436793. Request date: 9/21/2016. Service Tag: DP6T35J. Reference / PO #: . DPS #: 7321316311.	4/9/2017 9:28 PM
L	Dispatch Processed This message is to alert you a recent parts dispatch request has been processed. Work Order #: SR936436126. Request date: 9/21/2016. Service Tag: 5L5XR12. Reference / PO #: . DPS #: 7321312996.	4/9/2017 8:35 PM
L	Dispatch Denied This is to alert you that we were unable to process a recent parts dispatch WO8327278. Request Date 9/1/2016 Service Tag F1MZTB2 Reference / PO # Please review the details and resubmit your request online	4/9/2017 6:20 PM

Showing 1 to 5 of 220 entries

Mark all as read
Preferences

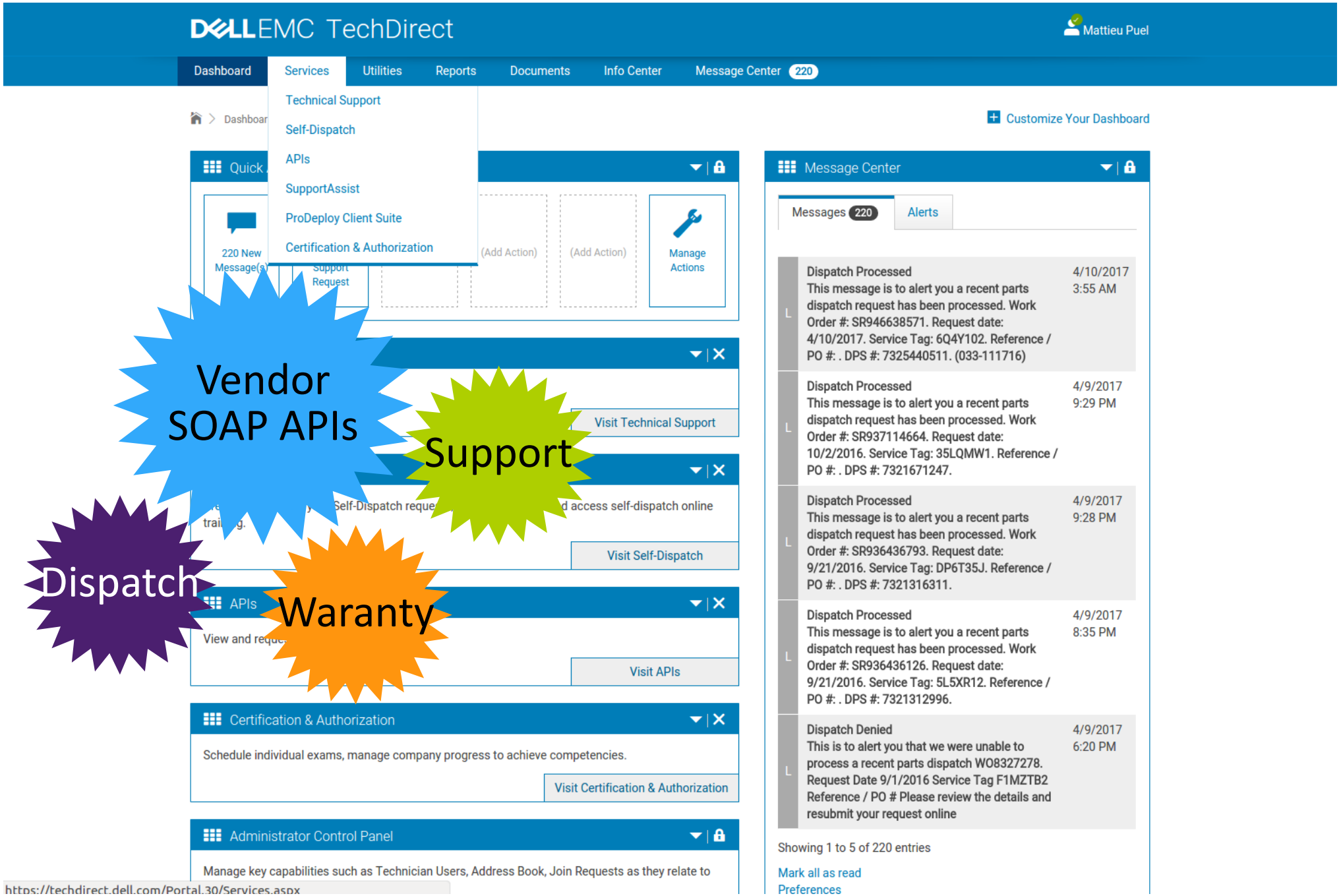
ONE word about DELL



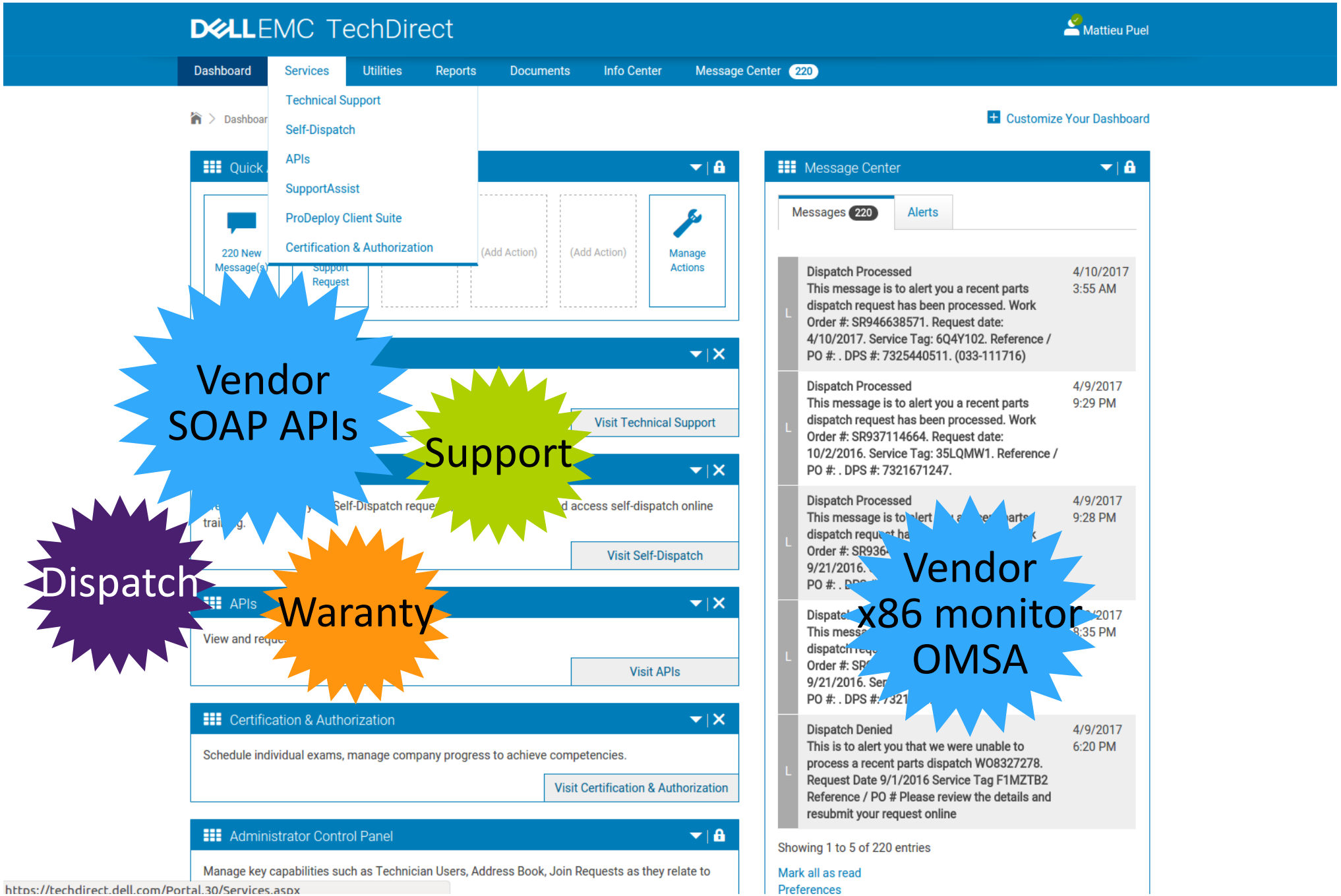
Vendor SOAP APIs

Dispatch

ONE word about DELL



ONE word about DELL



Agenda

Preliminary notes

Straight to an example

Architecture concerns

Some figures

Evolutions

Conclusions

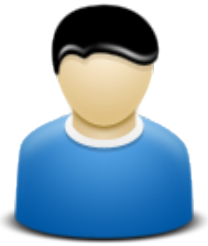
Straight to an example : a good old **DRIVE FAILURE**

nono2dtd

Vendor API

Redmine (ITS)

Dispatch validator



Sysadmin

Straight to an example : a good old **DRIVE FAILURE**

nono2dtd

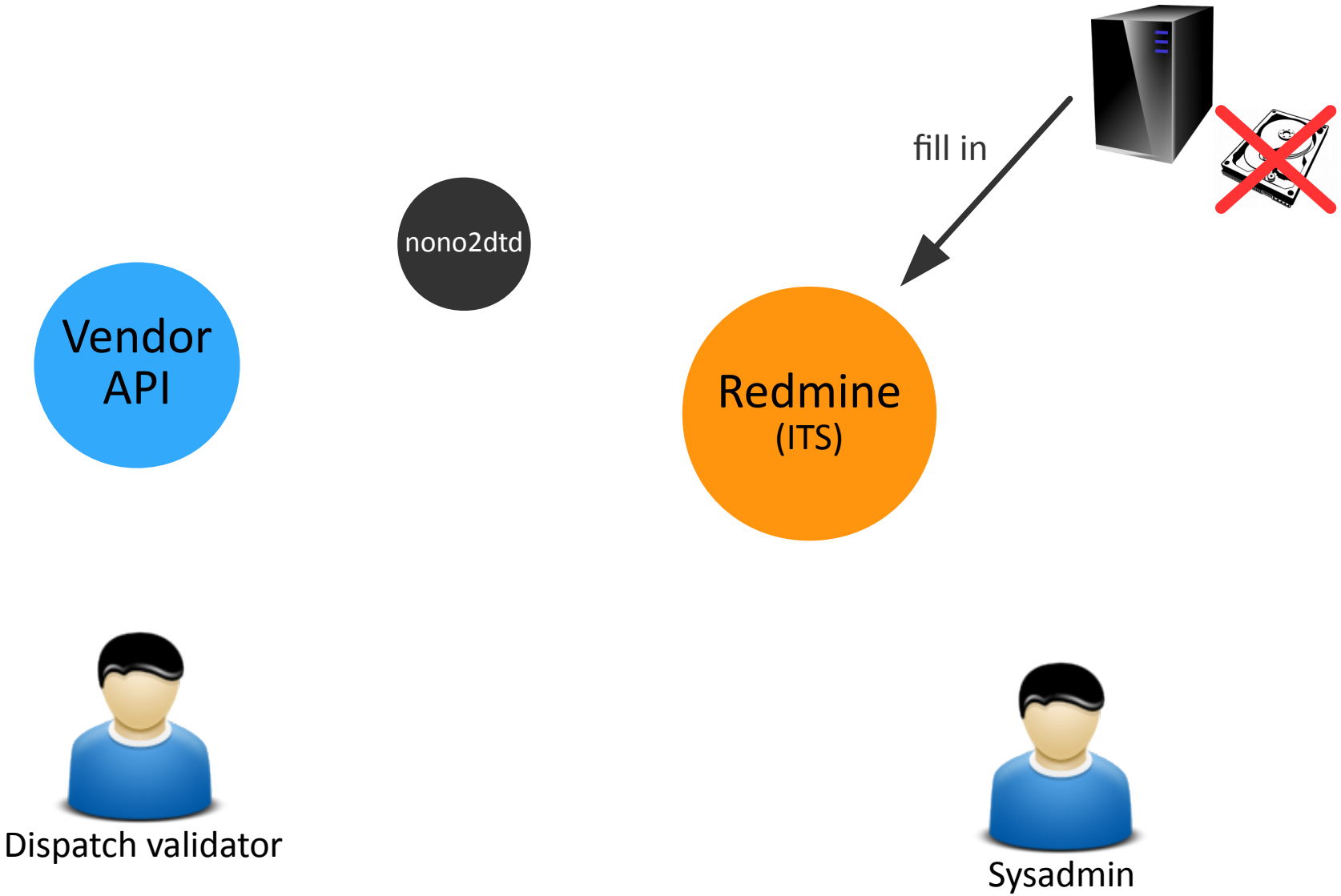
Vendor API

Redmine (ITS)

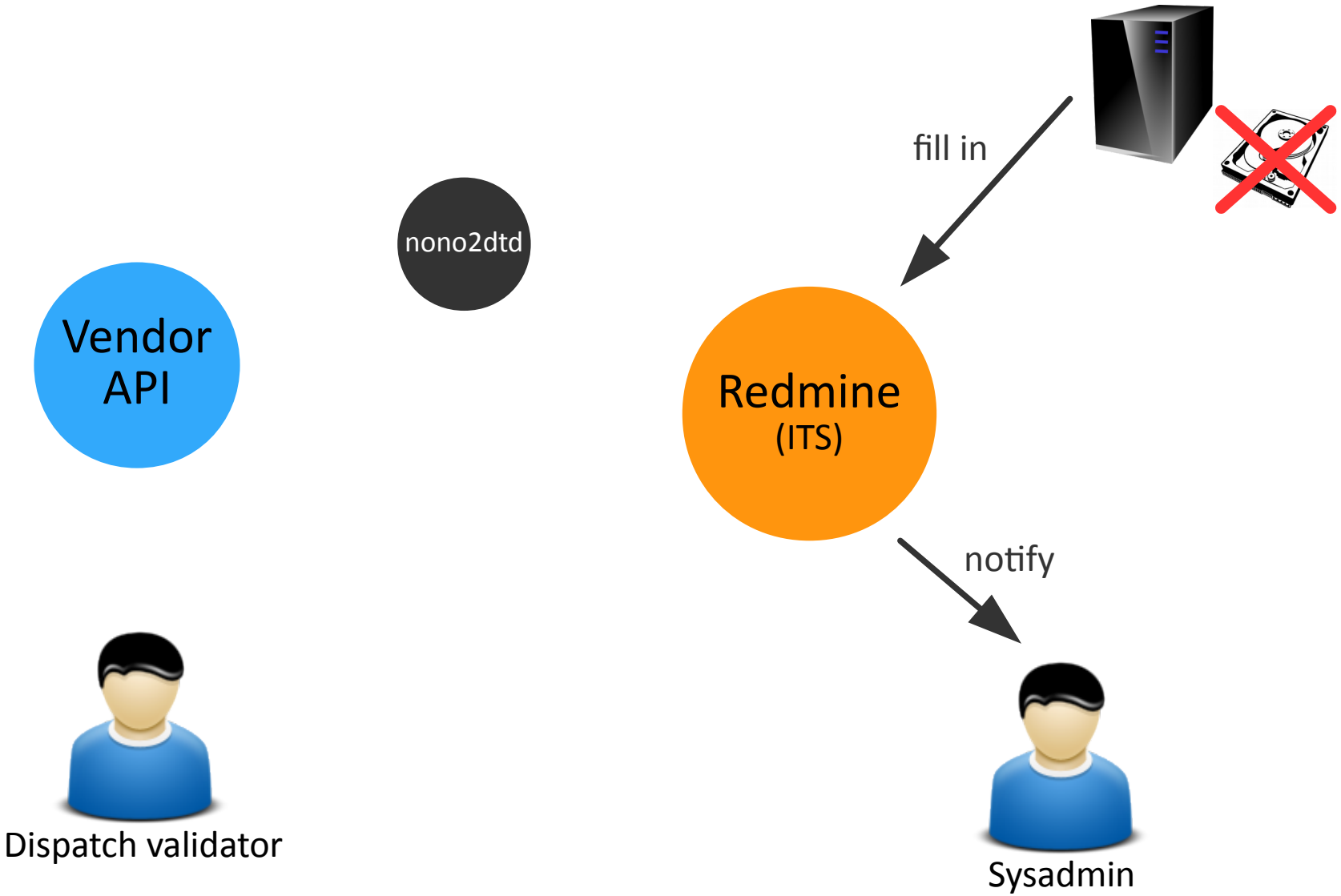
Dispatch validator

Sysadmin

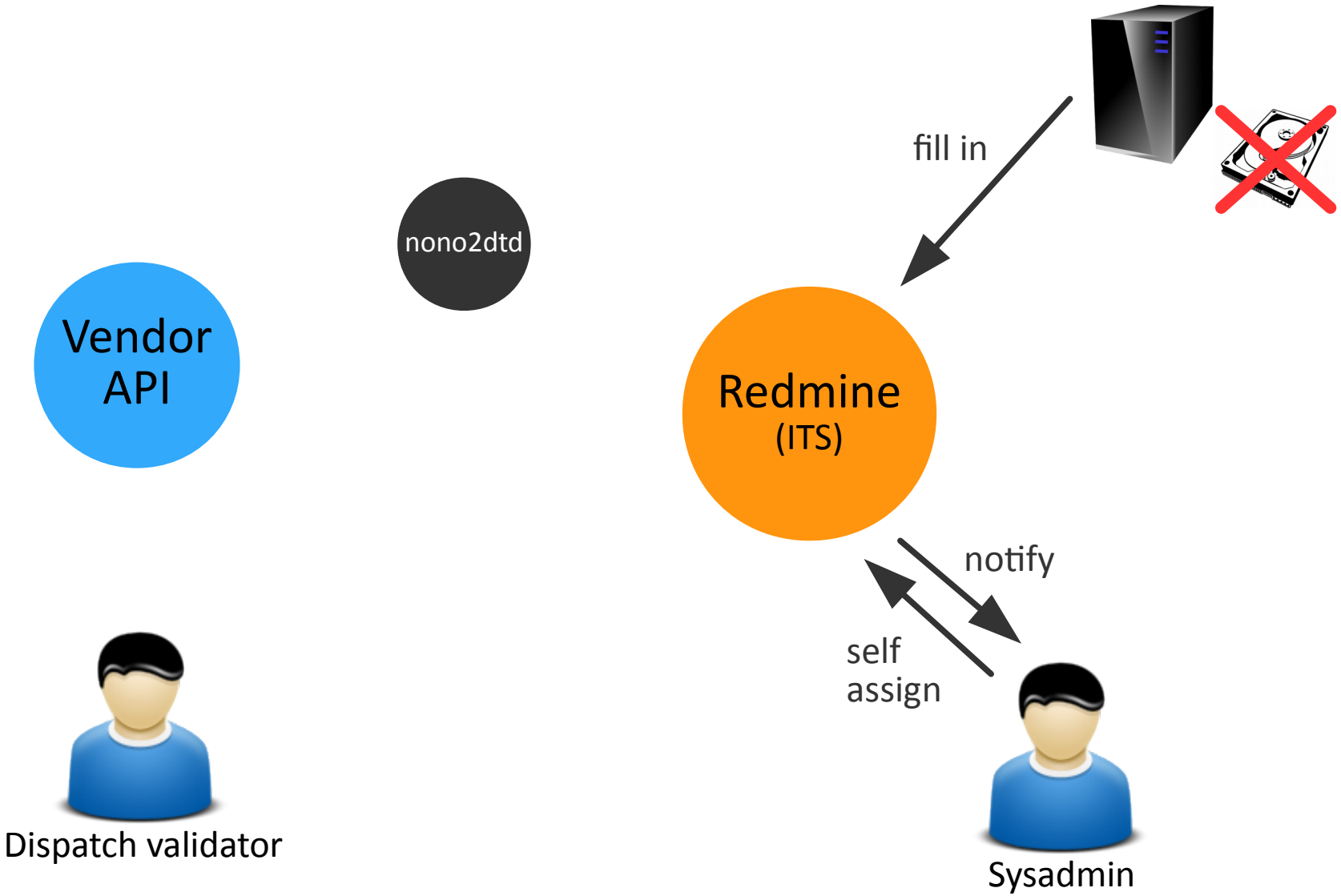
Straight to an example : a good old **DRIVE FAILURE**



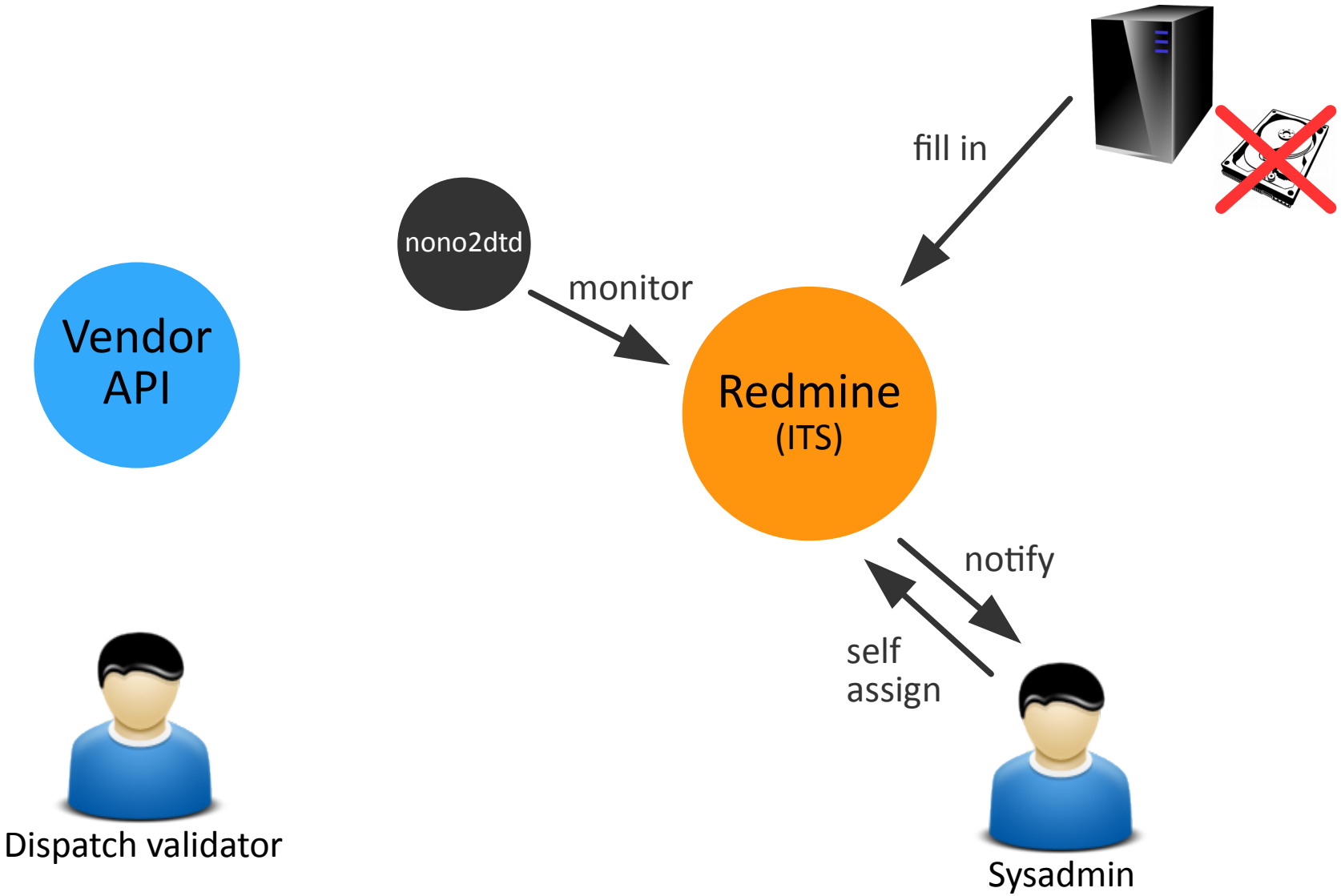
Straight to an example : a good old **DRIVE FAILURE**



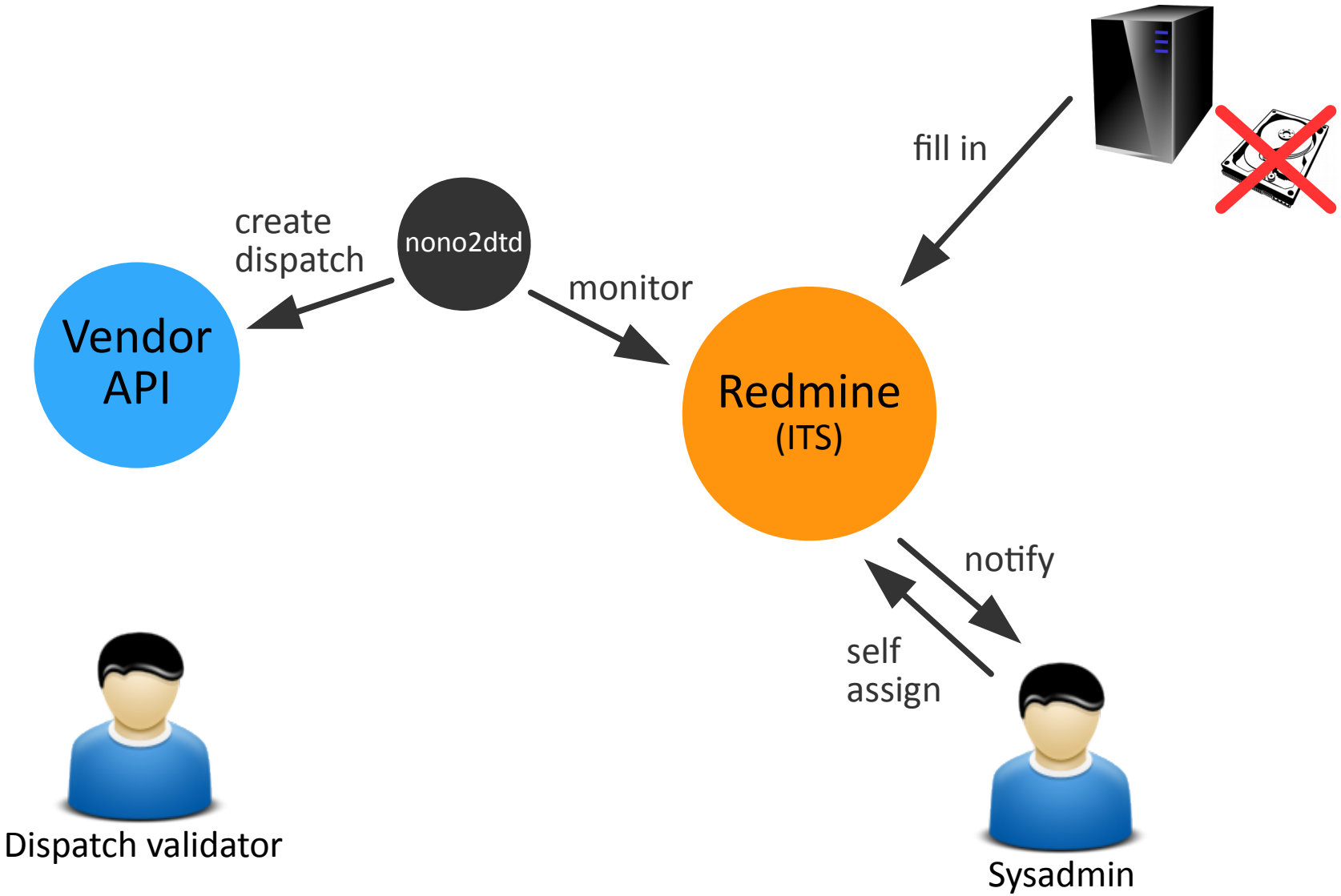
Straight to an example : a good old **DRIVE FAILURE**



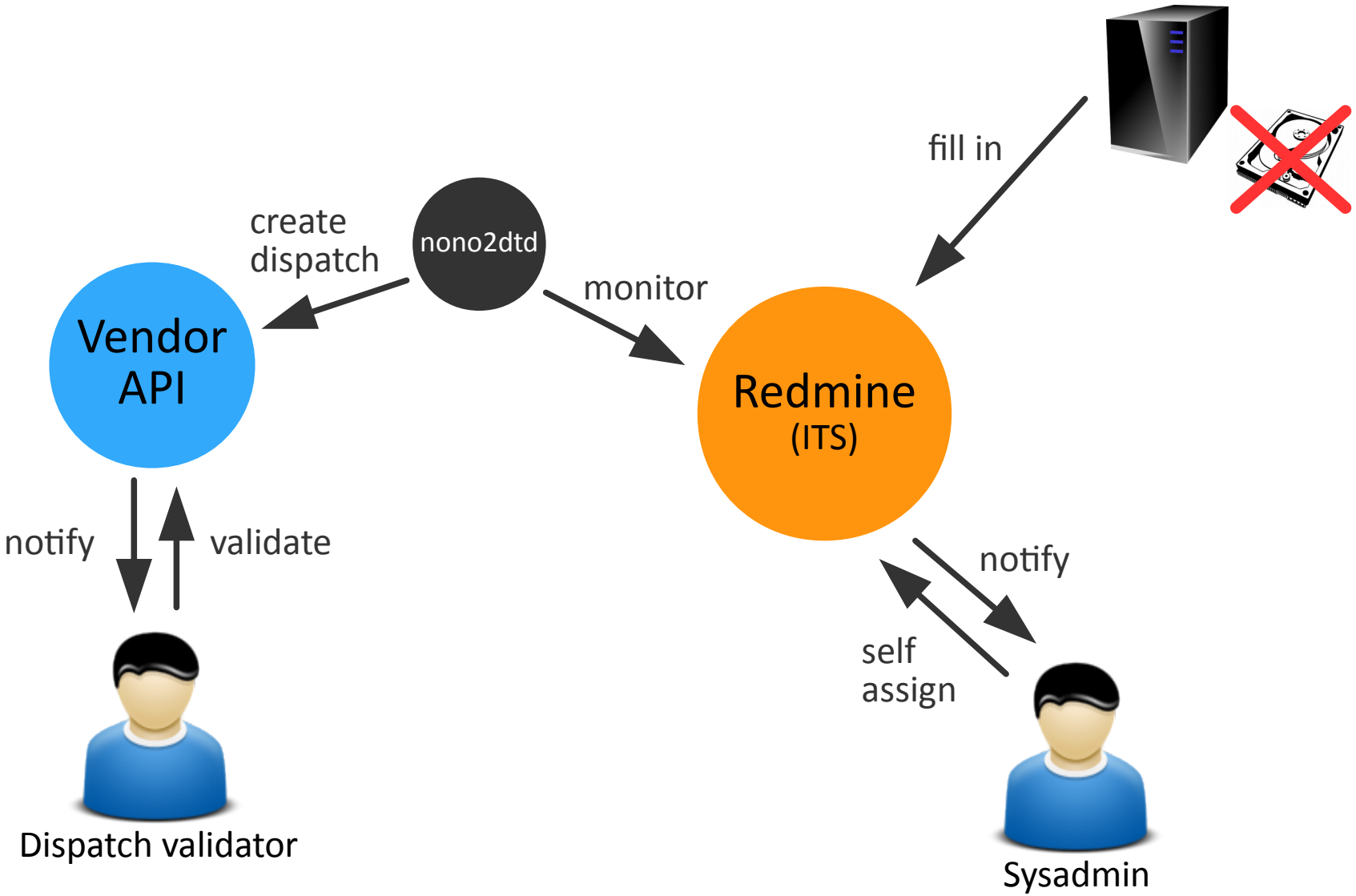
Straight to an example : a good old **DRIVE FAILURE**



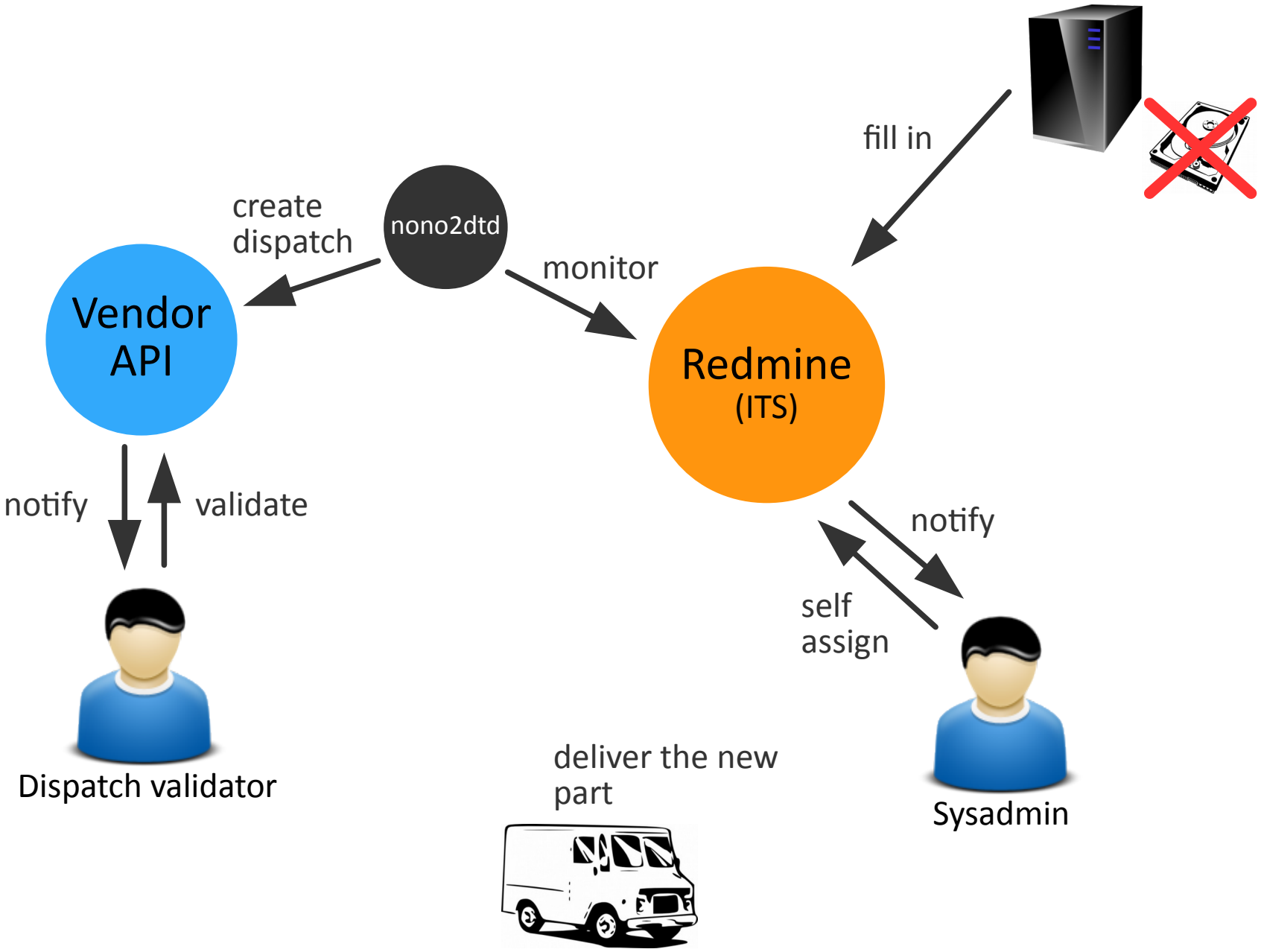
Straight to an example : a good old **DRIVE FAILURE**



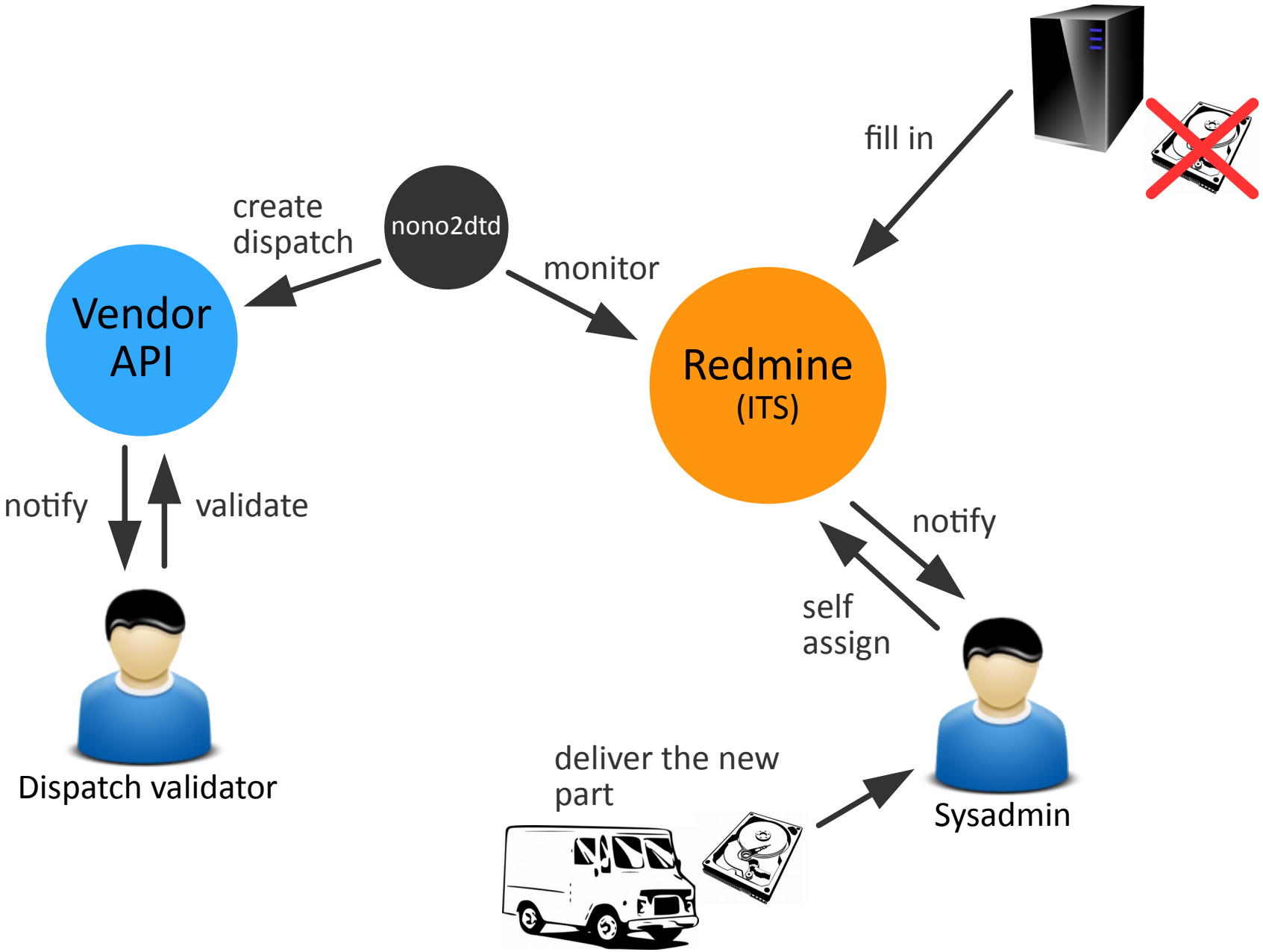
Straight to an example : a good old **DRIVE FAILURE**



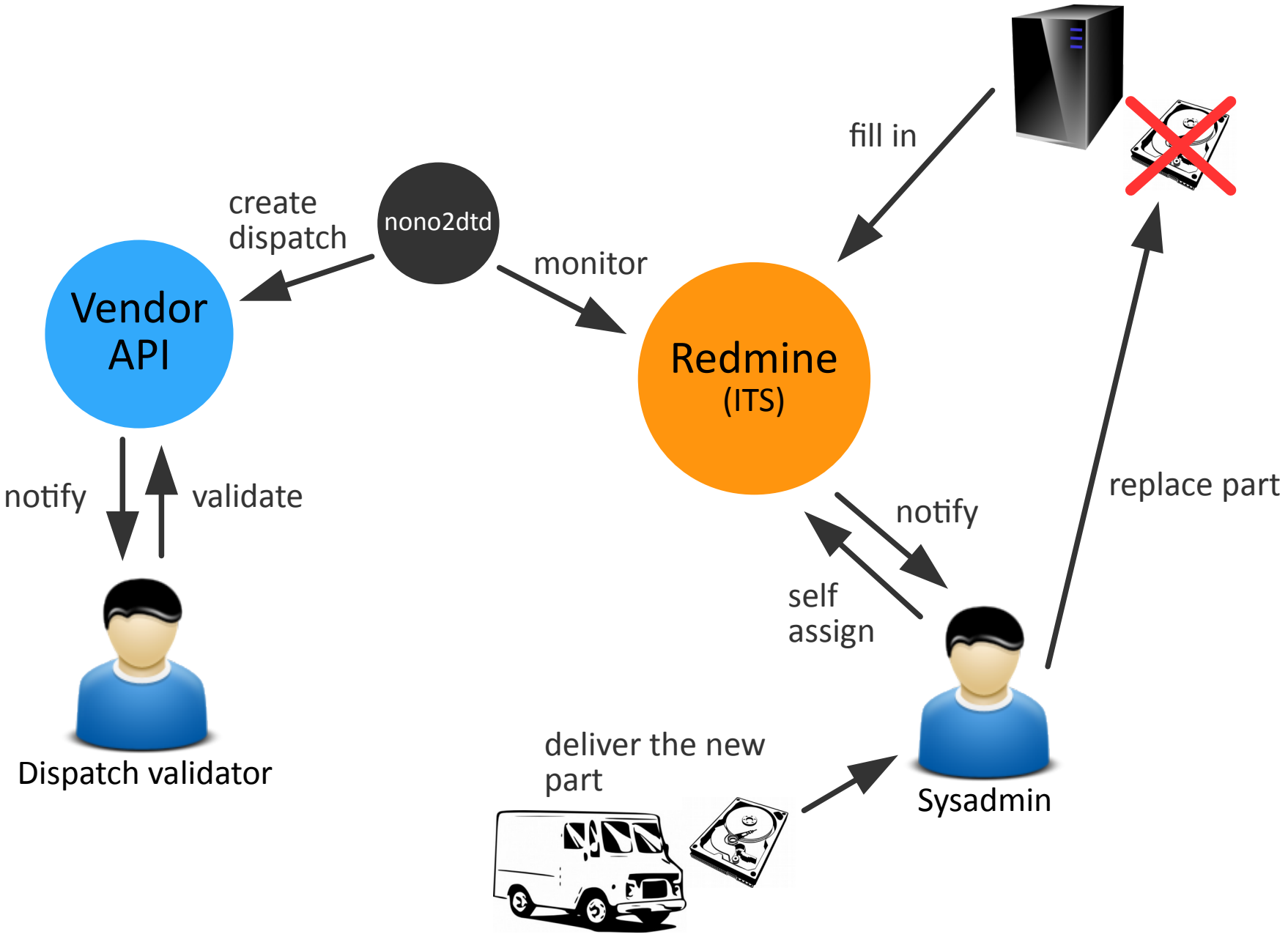
Straight to an example : a good old **DRIVE FAILURE**



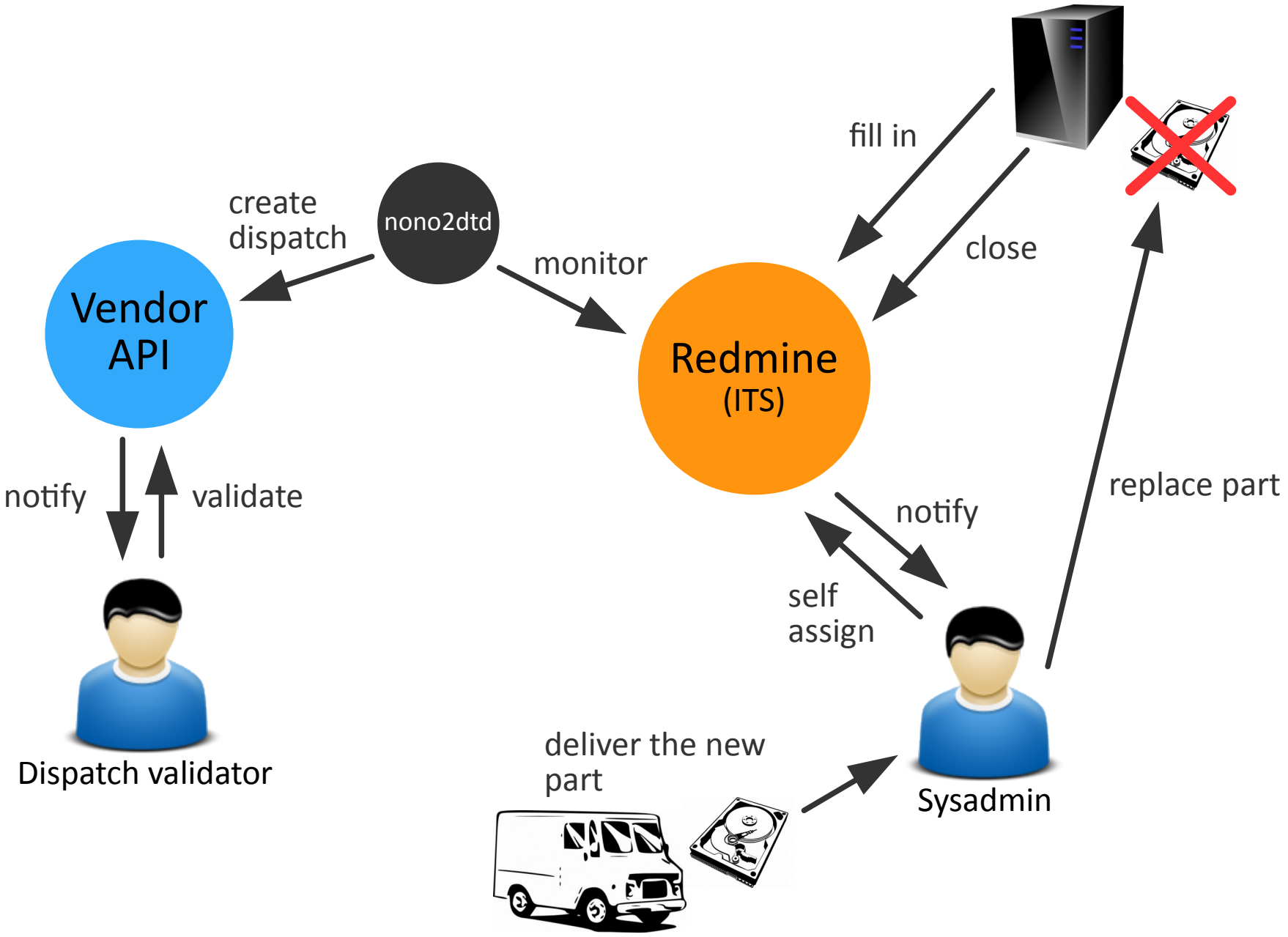
Straight to an example : a good old **DRIVE FAILURE**



Straight to an example : a good old **DRIVE FAILURE**



Straight to an example : a good old **DRIVE FAILURE**



1- I am notified by an email from the ITS

Courrier entrant - mattieu.puel@cc.in2p3.fr - Mozilla Thunderbird

Sujet **[Hardware incidents - Incident #17581] (New) disk issue on ccosndli0032**

Issue #17581 has been reported by LePetitRobot Nono.

Incident #17581: disk issue on ccosndli0032

- Author: LePetitRobot Nono
- Status: New
- Priority: Urgent
- Assigned To:
- Category: disk
- Target version:
- Machine: ccosndli0032
- S/N: 6Q4Y102
- Model: PowerEdge R720
- P/N: CN0F617N726223AO008DA03
- SR#:
- Vendor: Dell Inc.
- Usage: openstack
- PieceID: 0:0:1:1
- Location: Service 6/14
- Dispatch number:

ISSUES ON disk/0:0:1:1:

```
6Q4Y102      278.88    GB Online/Non-Critical Yes   6SJ7KKS4            CN0F617N726223AO008DA03
```

FULL REPORT

```
processor check
-------------------------------
processor id: 0
    CPU1 Ok/Present
processor id: 1
    CPU2 Ok/Present

bmc check
-------------------------------
bmc id: 1
    Hardware Log: Ok

memory check
-------------------------------
memory id: 0
    DIMM A1 Ok 16384 MB
memory id: 1
    DIMM A2 Ok 16384 MB
memory id: 10
    DIMM B5 Ok 16384 MB
memory id: 11
    DIMM B6 Ok 16384 MB
memory id: 2
    DIMM A3 Ok 16384 MB
memory id: 3
    DIMM A4 Ok 16384 MB
memory id: 4
    DIMM A5 Ok 16384 MB
memory id: 5
    DIMM A6 Ok 16384 MB
memory id: 6
    DIMM B1 Ok 16384 MB
memory id: 7
    DIMM B2 Ok 16384 MB
memory id: 8
    DIMM B3 Ok 16384 MB
memory id: 9
    DIMM B4 Ok 16384 MB
```

2- I assign to myself the incident

Home | My page | Projects | Scrum statistics | Help | Logged

Groupe Système CCIN2P3 » Hardware incidents

Search:

Overview | Activity | Issues | New issue | Files | Settings

Issues

Filters

Options

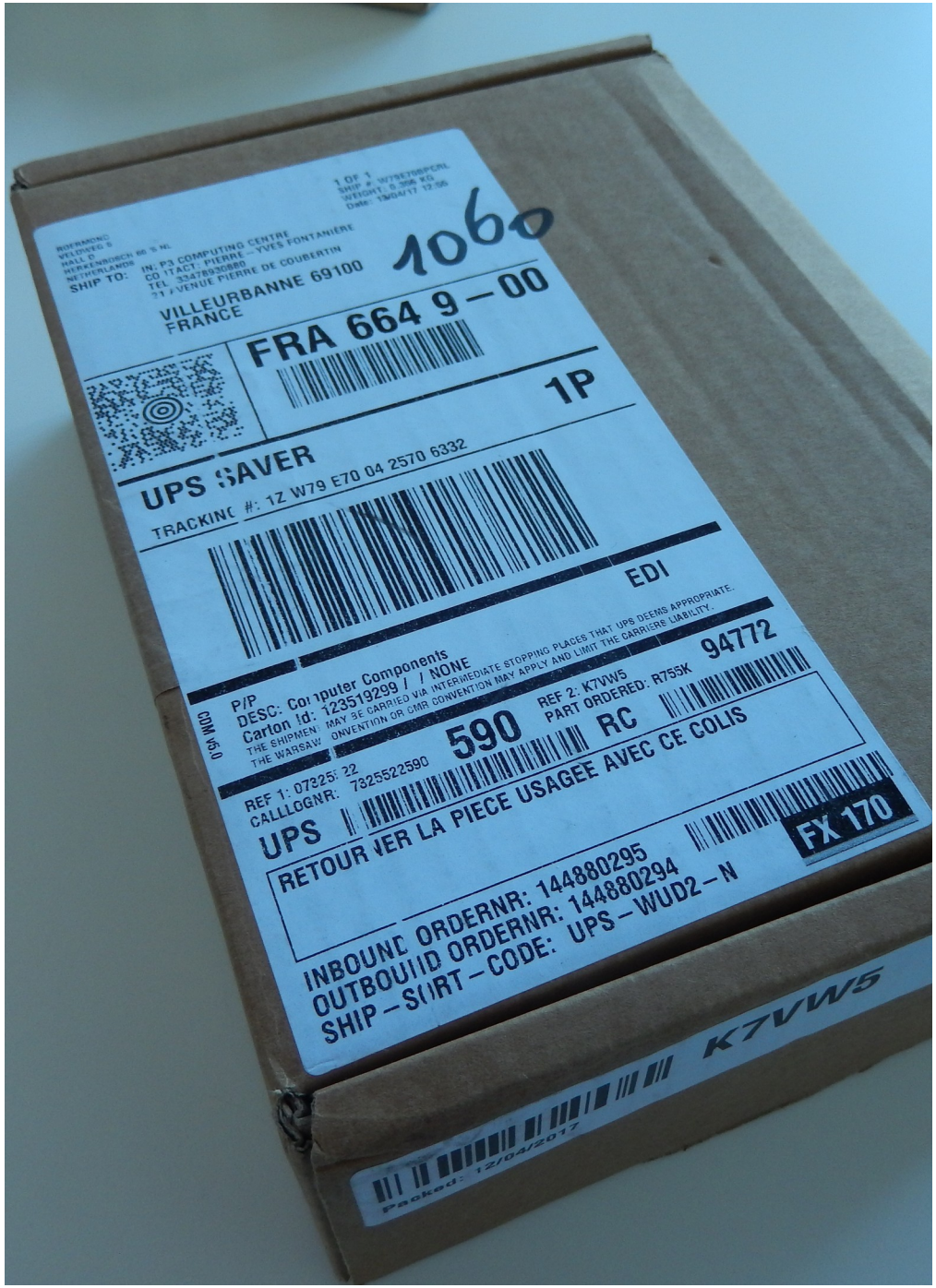
Apply | Clear | Save

#	Tracker	Status	Priority	Subject	Assigned To	Updated	Category
17602	Incident	New	Urgent	disk issue on cchmdli007		04/12/2017 12:01 AM	disk
17581	Incident	Closed	Urgent	disk issue on ccosndli0032	Mattieu	04/11/2017 12:56 PM	disk
17570	Incident	New	Urgent	bmc issue on ccstest03		04/07/2017 08:45 AM	bmc
17540	Incident	Closed	Urgent	disk issue on cchmdli008	Pierre-Yves	04/07/2017 11:54 AM	disk
17536	Incident	Closed	Normal	memory issue on ccwgpge005		03:13 PM	memory
17535	Incident	Closed	Urgent	disk issue on ccosndli0051		02:43 PM	disk
17522	Incident	Closed	Urgent	disk issue on ccdcacli042		10:06 AM	disk
17518	Incident	Closed	Urgent	openmanage issue on ccdtli82		01:20 PM	openmanage
17517	Incident	Closed	Urgent	openmanage issue on ccdtli83		01:17 PM	openmanage
17503	Incident	Closed	Urgent	power issue on cctest35		01:49 PM	power
17502	Incident	Closed	Urgent	disk issue on ccdcacli040	Wern	10:21 AM	disk
17501	Incident	Closed	Urgent	disk issue on ccdcacli201	Wern	01:11 AM	disk
17463	Incident	Resolved	Normal	cmos battery low sur ccwsge0119 Poweredge C6220 GJRKC5J	Niu	02:46 PM	
17388	Incident	Closed	Urgent	disk issue on ccdcacli027	Babel Jahson	03/30/2017 08:54 AM	disk
17346	Incident	Closed	Urgent	disk issue on ccxrdli009	Hamar Vanessa	03/28/2017 02:54 PM	disk
17345	Incident	Closed	Urgent	disk issue on ccxrdli046	Hamar Vanessa	03/29/2017 02:06 PM	disk

Edit
Status
Tracker
Priority
Assigned To
Category
% Done
Unwatch
Log time
Copy
Delete

<< me >>
Babel Jahson
Eloto Christelle
Ferrand Rémi
Fontanière Pierre-Yves
Fournials Nicolas
Hamar Vanessa
Khouder Ahmed
LePetitRobot Nono
Niu Xiaomei
Puel Mattieu
Wernli Fabien
nobody

3- Xmas present for me ! I relate this to it's original issue



My page | Projects | Scrum statistics | Help

Groupe Système CCIN2P3 » Hardware incidents

Search:

Overview | Activity | Issues | New issue | Files | Settings

Incident #17602

Edit | Log time | Unwatch

disk issue on cchmdli007

Added by LePetitRobot Nono 2 days ago. Updated less than a minute ago.

Status:	New	**Start date:**	04/12/2017
Priority:	Urgent	**Due date:**	
Assigned To:	Fontanière Pierre-Yves	**% Done:**	0%
Category:	disk	**Estimated time:**	30.00 hours
Target version:	-	**Spent time:**	-
Machine:	cchmdli007	**Vendor:**	Dell Inc.
S/N:	GP6T35J	**Usage:**	hpss
Model:	PowerEdge R510	**PieceID:**	1:0:0:9
P/N:	TH0R755K21233172009FA02	**Location:**	Rittal 00/28
SR#:	SR946847557	**Dispatch number:**	7325522590

Description

ISSUES ON disk/1:0:0:9:

```
BVXT35J      1,862.50  GB Online/Non-Critical Yes   9WM6LXQ2          TH0R755K21233172009FA02
```

FULL REPORT

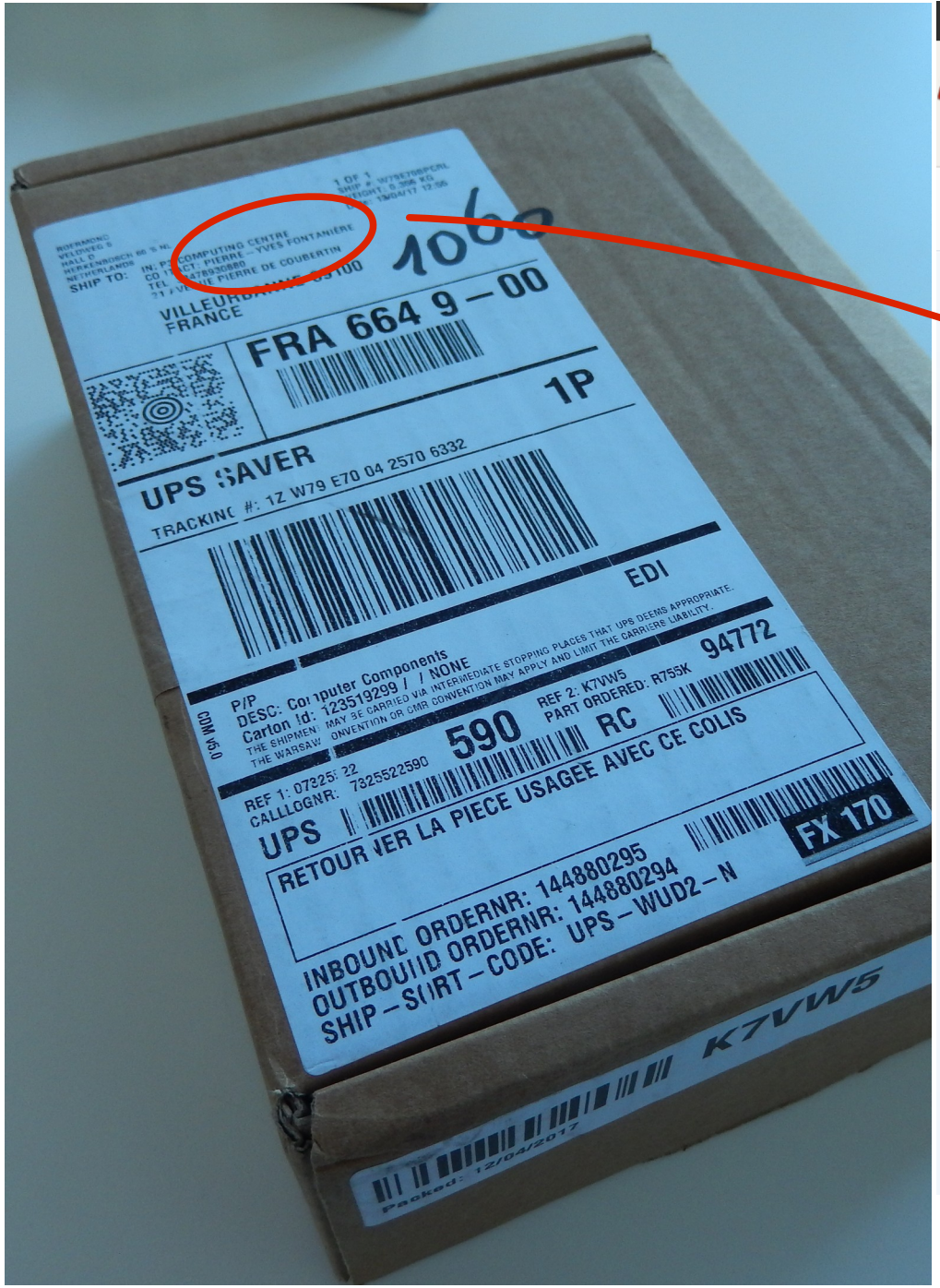
```
processor check
-------------------------------
processor id: 0
    CPU1 Ok/Present

bmc check
-------------------------------
bmc id: 1
    Hardware Log: Ok

memory check
-------------------------------
memory id: 0
    DIMM_A1 Ok 2048 MB
memory id: 1
    DIMM_A2 Ok 2048 MB

controller check
-------------------------------
```

3- Xmas present for me ! I relate this to it's original issue



My page | Projects | Scrum statistics | Help

Groupe Système CCIN2P3 » Hardware incidents

Search:

Overview | Activity | Issues | New issue | Files | Settings

Incident #17602

Edit | Log time | Unwatch

disk issue on cchmdli007

Added by LePetitRobot Nono 2 days ago. Updated less than a minute ago.

Status:	New	Start date:	04/12/2017
Priority:	Urgent	Due date:	
Assigned To:	Fontanière Pierre-Yves	% Done:	0%
Category:	disk	Estimated time:	30.00 hours
Target version:	-	Spent time:	-
Machine:	cchmdli007	Vendor:	Dell Inc.
S/N:	GP6T35J	Usage:	hpss
Model:	PowerEdge R510	PieceID:	1:0:0:9
P/N:	TH0R755K21233172009FA02	Location:	Rittal 00/28
SR#:	SR946847557	Dispatch number:	7325522590

Description

ISSUES ON disk/1:0:0:9:

```
BVXT35J      1,862.50  GB Online/Non-Critical Yes   9WM6LXQ2          TH0R755K21233172009FA02
```

FULL REPORT

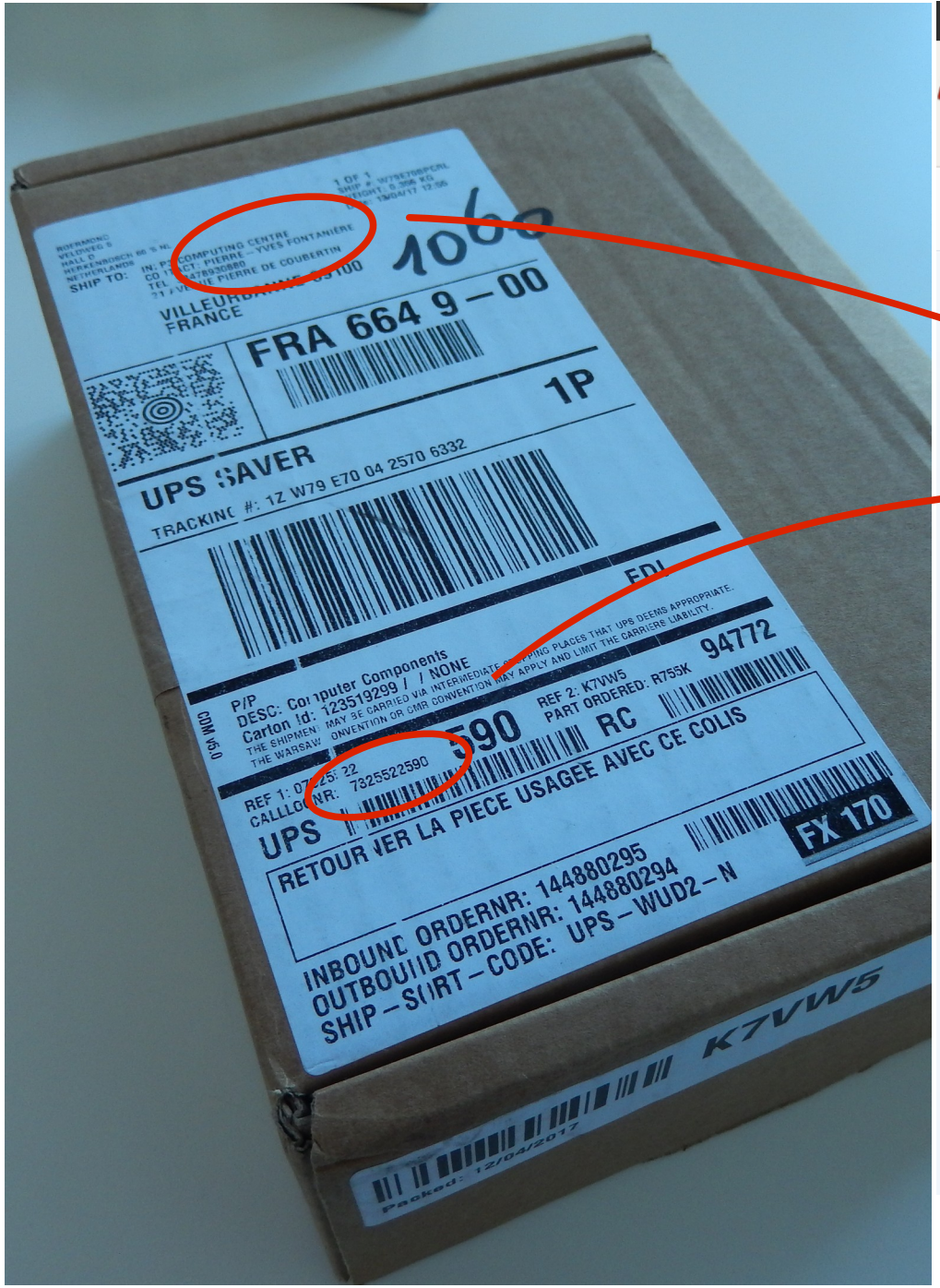
```
processor check
-------------------------------
processor id: 0
    CPU1 Ok/Present

bmc check
-------------------------------
bmc id: 1
    Hardware Log: Ok

memory check
-------------------------------
memory id: 0
    DIMM_A1 Ok 2048 MB
memory id: 1
    DIMM_A2 Ok 2048 MB

controller check
-------------------------------
```

3- Xmas present for me ! I relate this to it's original issue



My page | Projects | Scrum statistics | Help

Groupe Système CCIN2P3 » Hardware incidents

Search:

Overview | Activity | Issues | New issue | Files | Settings

Incident #17602

Edit | Log time | Unwatch

disk issue on cchmdli007

Added by LePetitRobot Nono 2 days ago. Updated less than a minute ago.

Status:	New	Start date:	04/12/2017
Priority:	Urgent	Due date:	
Assigned To:	Fontanière Pierre-Yves	% Done:	0%
Category:	disk	Estimated time:	30.00 hours
Target version:	-	Spent time:	-
Machine:	cchmdli007	Vendor:	Dell Inc.
S/N:	GP6T35J	Usage:	hpss
Model:	PowerEdge R510	PieceID:	1:0:0:9
P/N:	TH0R755K21233172009FA02	Location:	Rittal 00/26
SR#:	SR946847557	Dispatch number:	7325522590

Description

ISSUES ON disk/1:0:0:9:

```
BVXT35J      1,862.50  GB Online/Non-Critical Yes   9WM6LXQ2          TH0R755K21233172009FA02
```

FULL REPORT

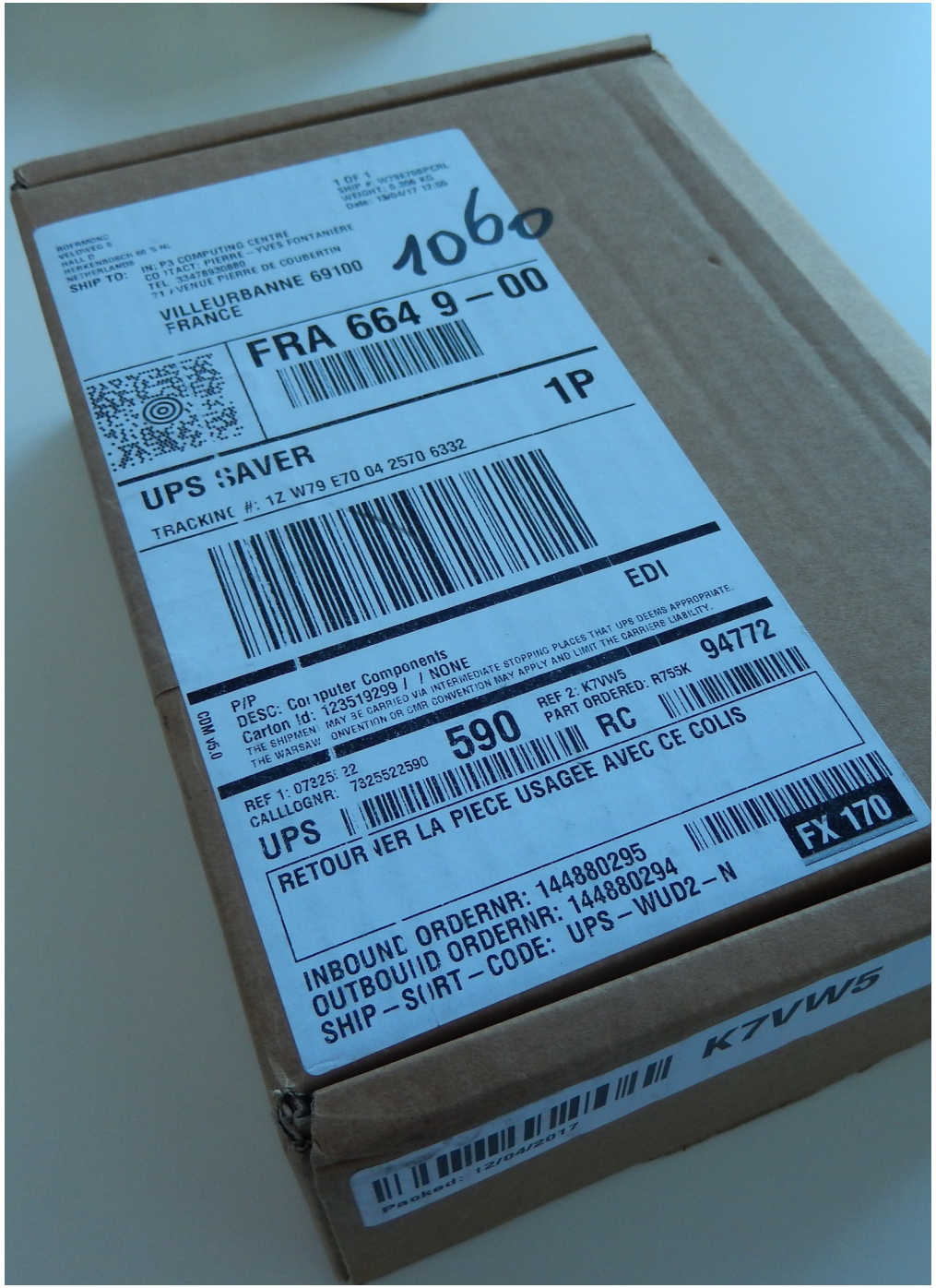
```
processor check
--------------------------------
processor id: 0
    CPU1 Ok/Present

bmc check
--------------------------------
bmc id: 1
    Hardware Log: Ok

memory check
--------------------------------
memory id: 0
    DIMM_A1 Ok 2048 MB
memory id: 1
    DIMM_A2 Ok 2048 MB

controller check
--------------------------------
```

3- Xmas present for me ! I relate this to it's original issue



My page | Projects | Scrum statistics | Help

Groupe Système CCIN2P3 » Hardware incidents

Search:

Overview | Activity | Issues | New issue | Files | Settings

Incident #17602

Edit | Log time | Unwatch

disk issue on cchmdli007

Added by LePetitRobot Nono 2 days ago. Updated less than a minute ago.

Status:	New	Start date:	04/12/2017
Priority:	Urgent	Due date:	
Assigned To:	Fontanière Pierre-Yves	% Done:	0%
Category:	disk	Estimated time:	30.00 hours
Target version:	-	Spent time:	-
Machine:	cchmdli007	Vendor:	Dell Inc.
S/N:	GP6T35J	Usage:	hpss
Model:	PowerEdge R510	PieceID:	1:0:0:9
P/N:	TH0R755K21233172009FA02	Location:	Rittal 00/28
SR#:	SR946847557	Dispatch number:	[illegible]

Description

ISSUES ON disk/1:0:0:9:

```
BVXT35J      1,862.50  GB Online/Non-Critical Yes   9WM6LXQ2          TH0R755K21233172009FA02
```

FULL REPORT

```
processor check
-------------------------------
processor id: 0
    CPU1 Ok/Present

bmc check
-------------------------------
bmc id: 1
    Hardware Log: Ok

memory check
-------------------------------
memory id: 0
    DIMM_A1 Ok 2048 MB
memory id: 1
    DIMM_A2 Ok 2048 MB

controller check
-------------------------------
```

I said **almost**

5- Once the drive rebuild ends, Redmine notifies the closure

Courrier entrant - mattieu.puel@cc.in2p3.fr - Mozilla Thunderbird

Courrier entrant - mat... | Agenda

Relever | Écrire | Chat | Adresses | Étiquette | Filtre rapide | Rechercher <Ctrl+K>

mattieu.pu...c.in2p3.fr
- Courrier entrant (3)
- Brouillons
- Envoyés
- Corbeille

mattieu.p...gmail.com
- Courrier entrant (1)
- Corbeille
- Accusés de réception
- Déplacement
- [Gmail]
  - Brouillons
  - Message...yés (2)
  - Tous les...s (297)
  - Spam
  - Corbeille
  - Important (16)
  - Suivis
- Personnel
- Professionnel

Dossiers locaux

Non lus | Suivis | Contacts | Étiquettes | Pièces jointes | 7 messages | osndli0032

Filtrer les messages selon : Expéditeur | Destinataires | Sujet | Corps du message

Sujet	Expéditeur	Date
[NAGIOS] OPENSTACK RECOVERY : Check compute on ccosndli0032 is OK	Nagios User	31/03/2017 16:37
[Hardware incidents - Incident #17581] (New) disk issue on ccosndli0032	redmine@cc.in2p3.fr	09/04/2017 10:26
[Hardware incidents - Incident #17581] (Closed) disk issue on ccosndli0032	redmine@cc.in2p3.fr	12:56

Répondre | Transférer | Archiver | Indésirable | Supprimer | Autres

De redmine@cc.in2p3.fr

Sujet **[Hardware incidents - Incident #17581] (Closed) disk issue on ccosndli0032** 12:56

Issue #17581 has been updated by LePetitRobot Nono.

- **Status** changed from *New* to *Closed*
- **% Done** changed from 0 to 100

I clearly think this piece of hardware is fine...

Incident #17581: disk issue on ccosndli0032

- Author: LePetitRobot Nono
- Status: Closed
- Priority: Urgent
- Assigned To: Puel Mattieu
- Category: disk
- Target version:
- Machine: ccosndli0032
- S/N: 6Q4Y102
- Model: PowerEdge R720
- P/N: CN0F617N726223AO008DA03
- SR#: SR946638571
- Vendor: Dell Inc.
- Usage: openstack
- PieceID: 0:0:1:1
- Location: Service 6/14
- Dispatch number:

ISSUES ON disk/0:0:1:1:

```
6Q4Y102      278.88    GB Online/Non-Critical Yes   6SJ7KKS4            CN0F617N726223AO008DA03
```

FULL REPORT

```
processor check
--------------------------------
processor id: 0
    CPU1 Ok/Present
processor id: 1
    CPU2 Ok/Present

bmc check
--------------------------------
bmc id: 1
    Hardware Log: Ok

memory check
--------------------------------
memory id: 0
    DIMM_A1 Ok 16384 MB
memory id: 1
    DIMM_A2 Ok 16384 MB
memory id: 10
    DIMM_B5 Ok 16384 MB
memory id: 11
    DIMM_B6 Ok 16384 MB
memory id: 2
    DIMM_A3 Ok 16384 MB
memory id: 3
    DIMM_A4 Ok 16384 MB
memory id: 4
    DIMM_A5 Ok 16384 MB
```

Non lus : 3 | Total : 3062 | Panneau « Aujourd'hui »

Agenda

Preliminary notes

Straight to an example

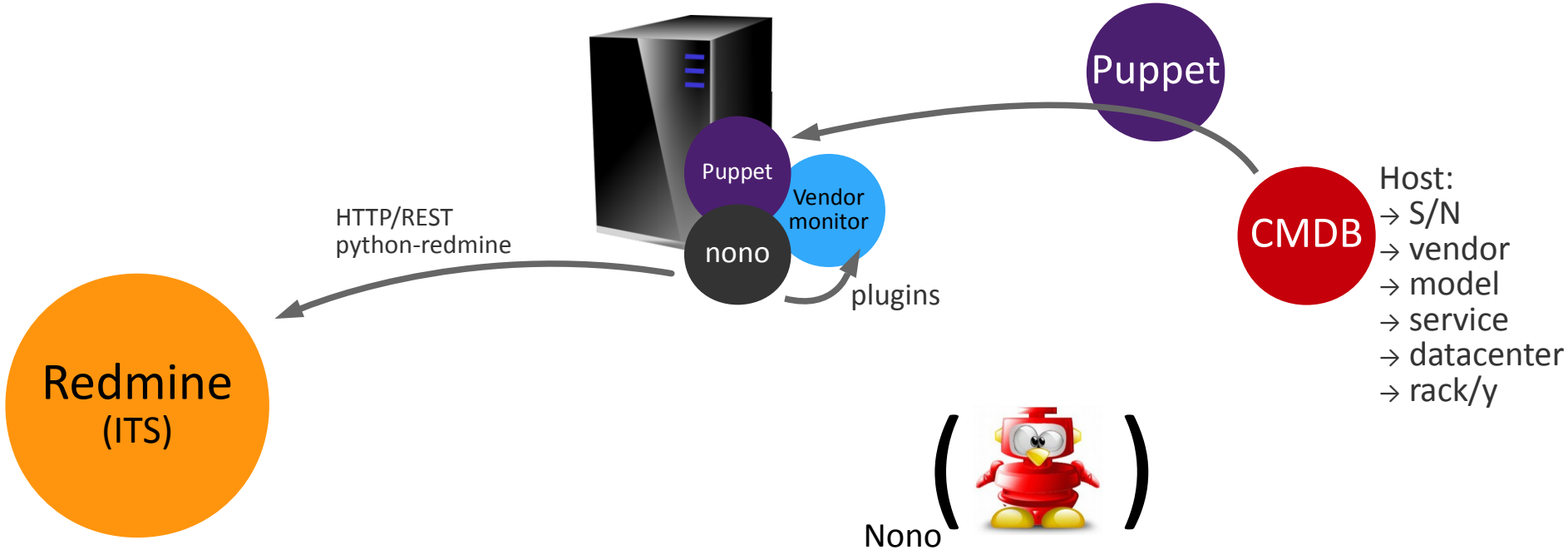
Architecture concerns

Some figures

Evolutions

Conclusions

A closer look at the architecture : **The incident detection**



( )

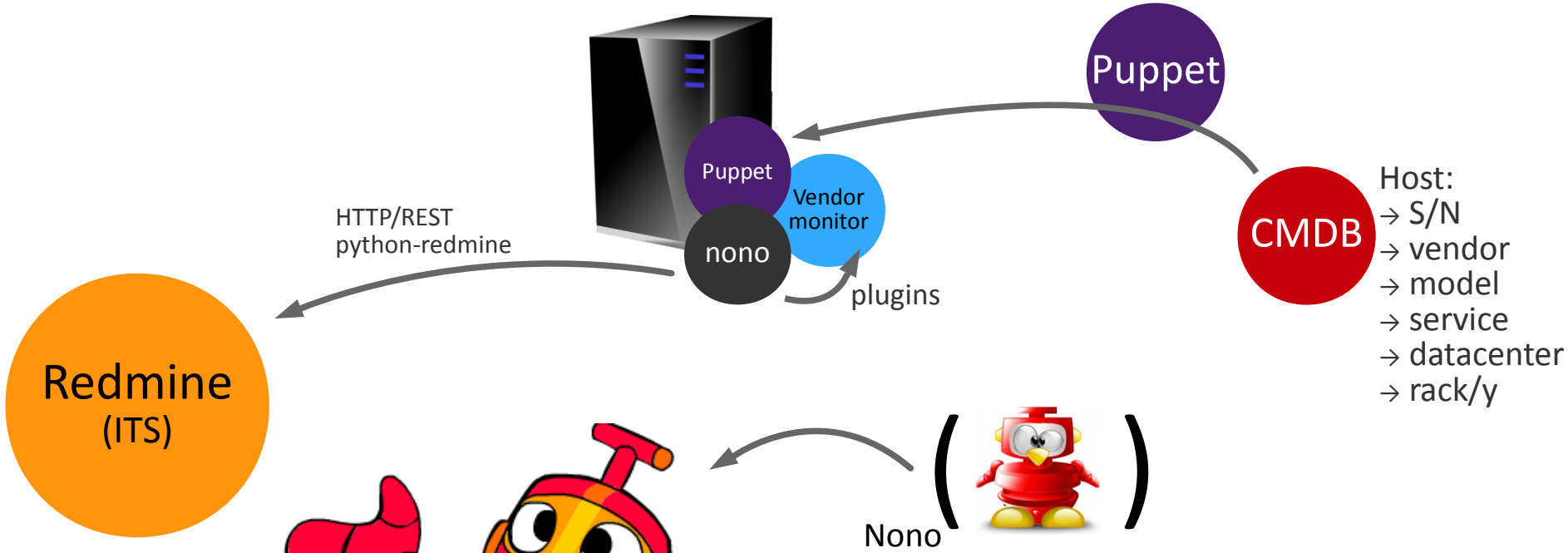
Nono

- queries hardware status through plugins (disk, memory, controller...)
- generates a diagnostic file to be provided to the vendor when necessary
- is responsible for tracking failures in the ITS (open/close)

An issue has the following metadata :

```
category            disk
serial number       6Q4Y102
vendor              DELL Inc.
model               PowerEdge R720
part number         CN0F617N726223AO008DA03
service             openstack
datacenter          Villeurbanne2
location            C12/32
service request
dispatch number
```

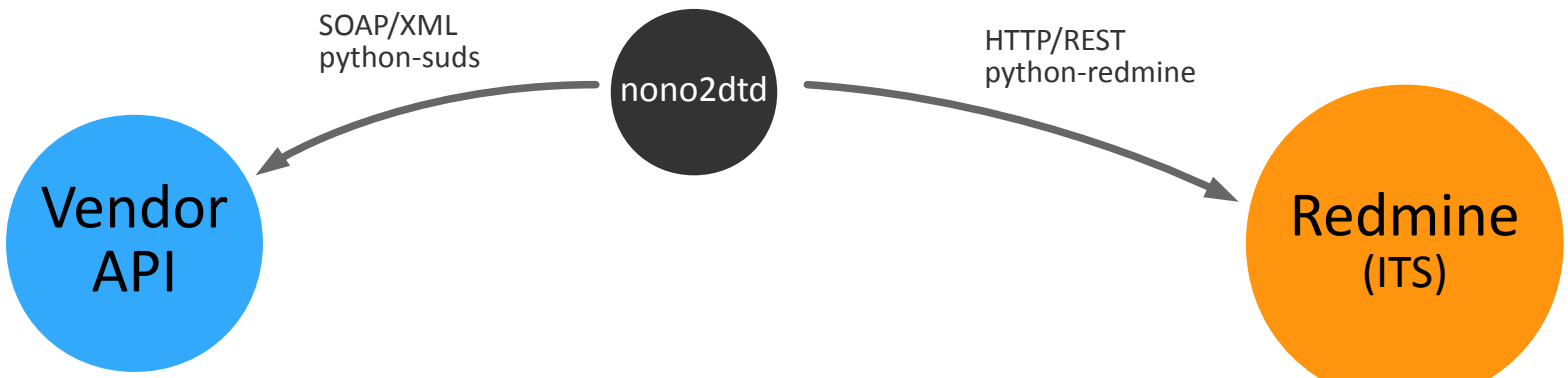
A closer look at the architecture : **The incident detection**



Nono

- queries hardware status through plugins (disk, memory, controller...)
- generates a diagnostic file to be provided to the vendor when necessary
- is responsible for tracking failures in the ITS (open/close)

A closer look at the architecture : **The dispatch**



SOAP API calls :

```
list_dispatch()
list_parts(SN)
list_parts(model)
create_dispatch(PN, SN, diag, location, contact)
get_dispatch_status(SR)
```

()

Nono2dtd

- queries redmine for incidents that are assigned, dispatchable and with no SR specified
- creates the dispatch for explicit technician through the dispatch API
- retrieves the SR or denial reason and report it to redmine
- reports the DN to redmine once set (after human validation)

This requires a little thought though

About technical details :

- do you use in-band or out-of-band host information ?
- do you send fully introspected vendor diagnostics or targeted information ?

About workflows :

- how do you manage issue assignement : self-assign ? round robin it ?
- how do you handle named deliveries and front desk interaction ?

About the diagnostics :

- what probes will you rely on ? (memory, batteries, BMCs, drives, enclosures, fans, PSUs, CPUs)
- how to deal with flapping alarms or false positives (batteries, PSUs) ?

Drawbacks & limitations

Specific to *this* vendor :

- usage of self-dispatch requires vendor certifications
- usage of APIs requires enrolment program
- not much documented
- XML/SOAP
- unable to integrate the worker nodes (~1k hosts)

In general :

- need to deal with the support for all appropriate diags
- you can’t manage some components this way (DIMMs, PSUs...)

Agenda

Preliminary notes

Straight to an example

Architecture concerns

Some figures

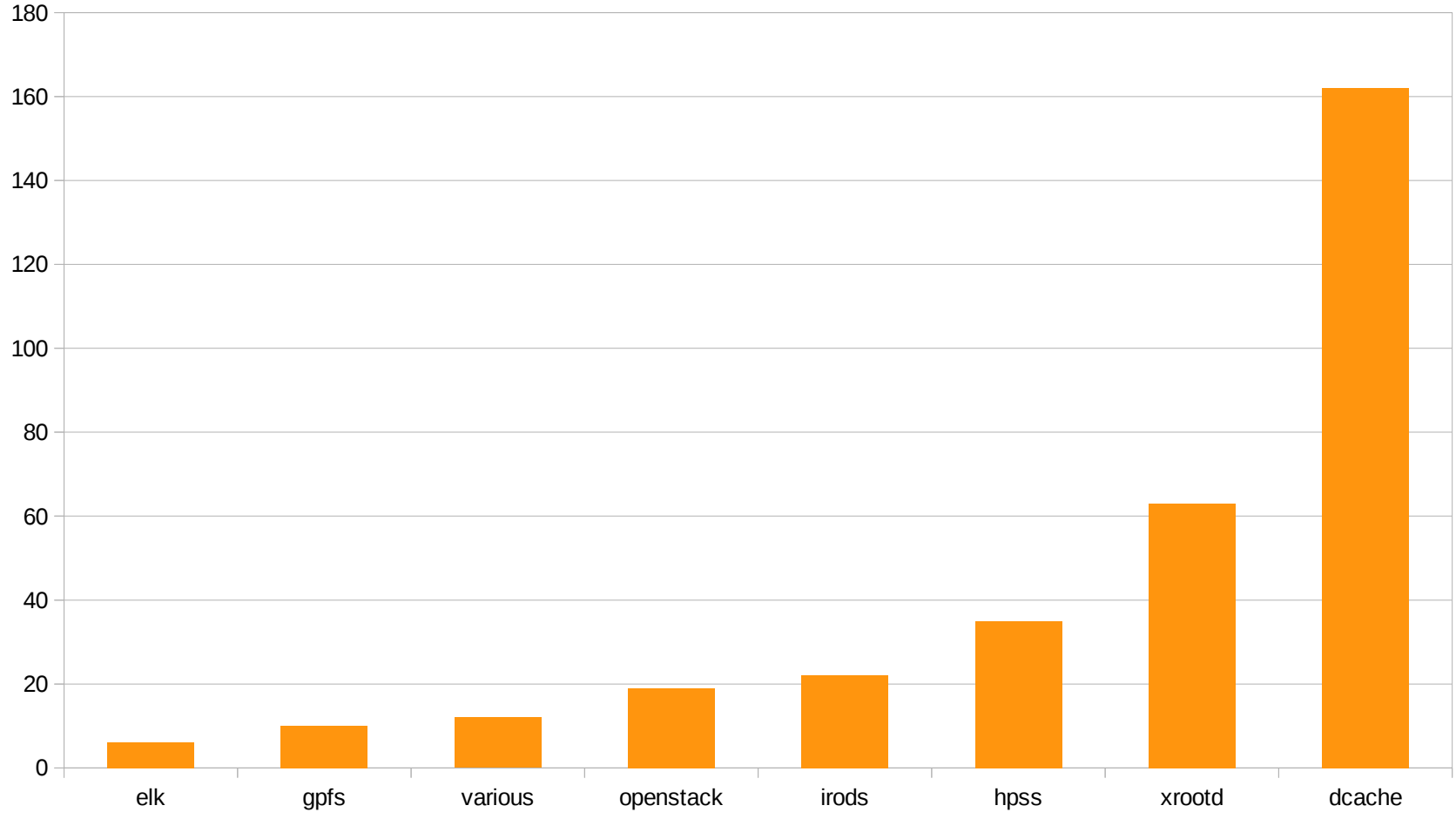
Evolutions

Conclusions

Draw whatever **chart** you like !

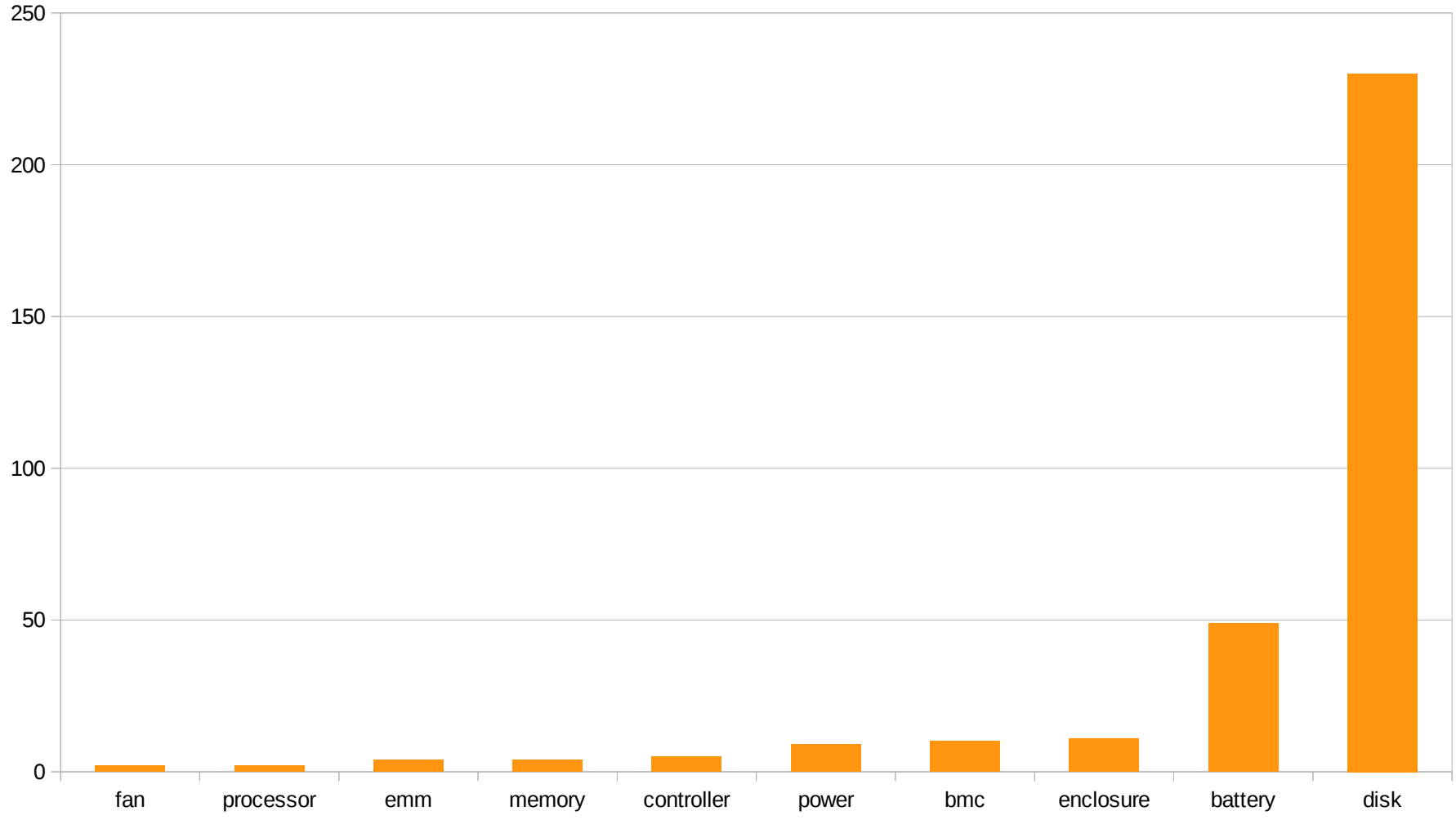
Following charts are based on around 420 days of data (329 incidents). This is a **cured** subset.

Issues by service :



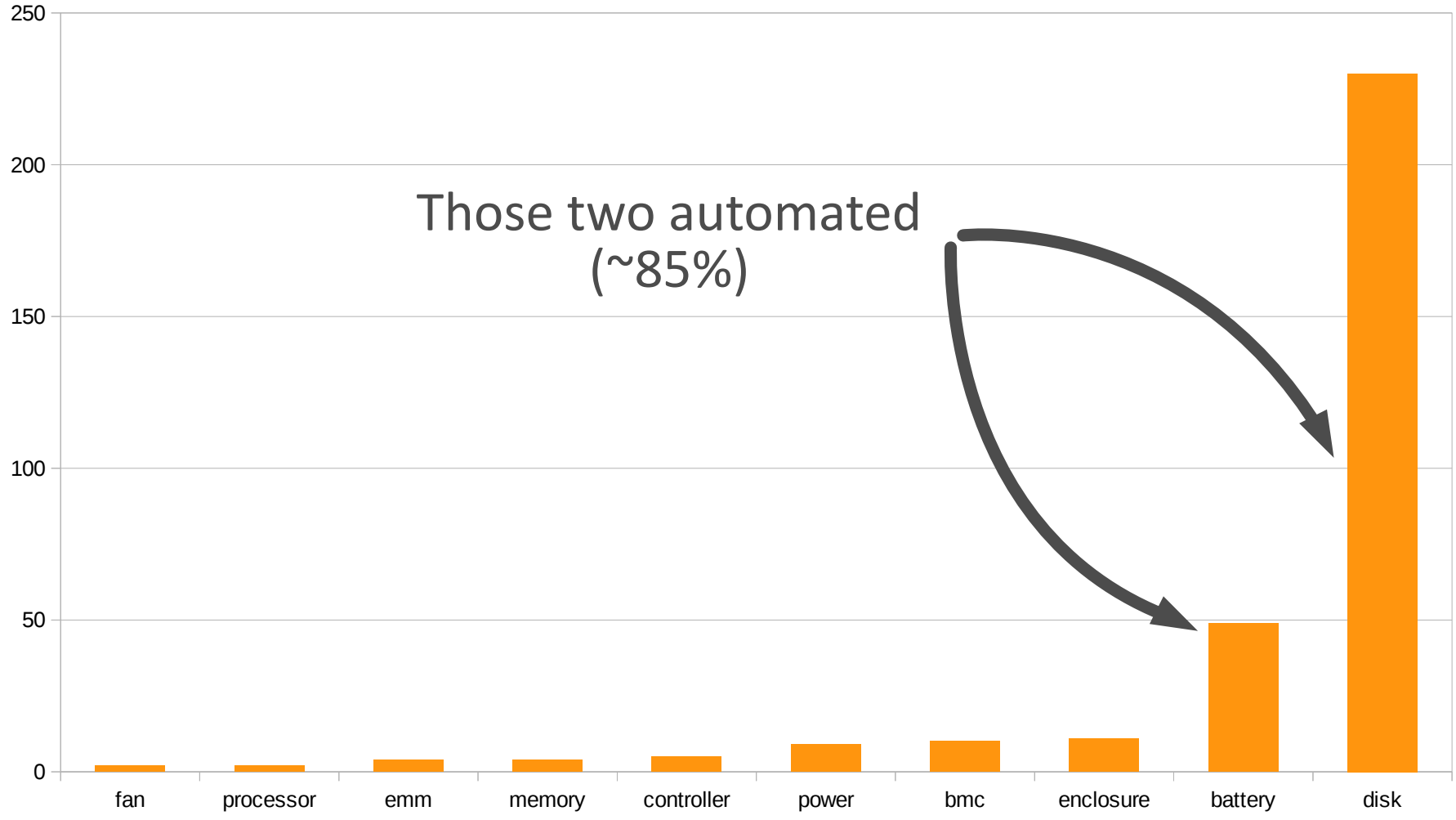
Draw whatever **chart** you like !

Issues by category :



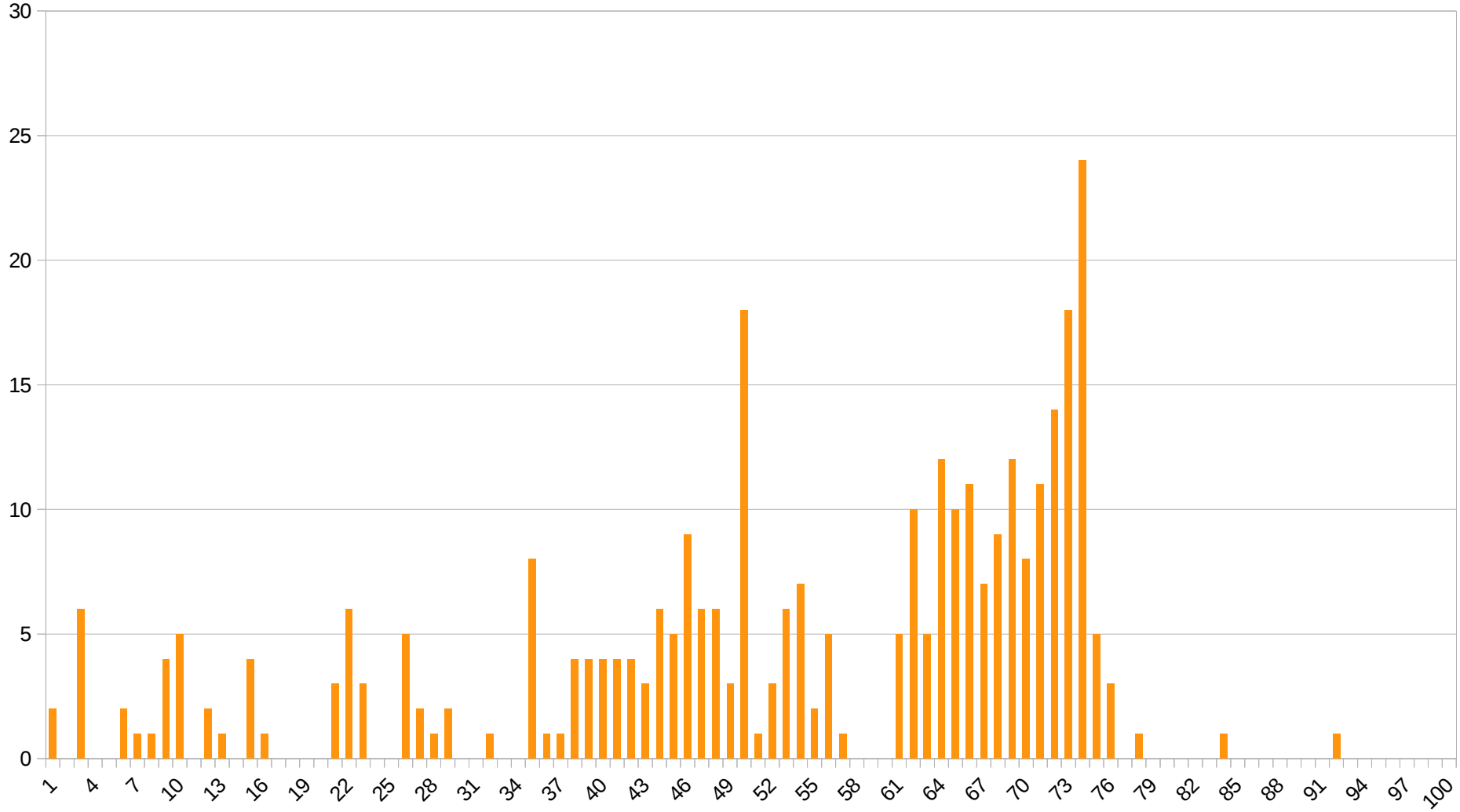
Draw whatever **chart** you like !

Issues by category :



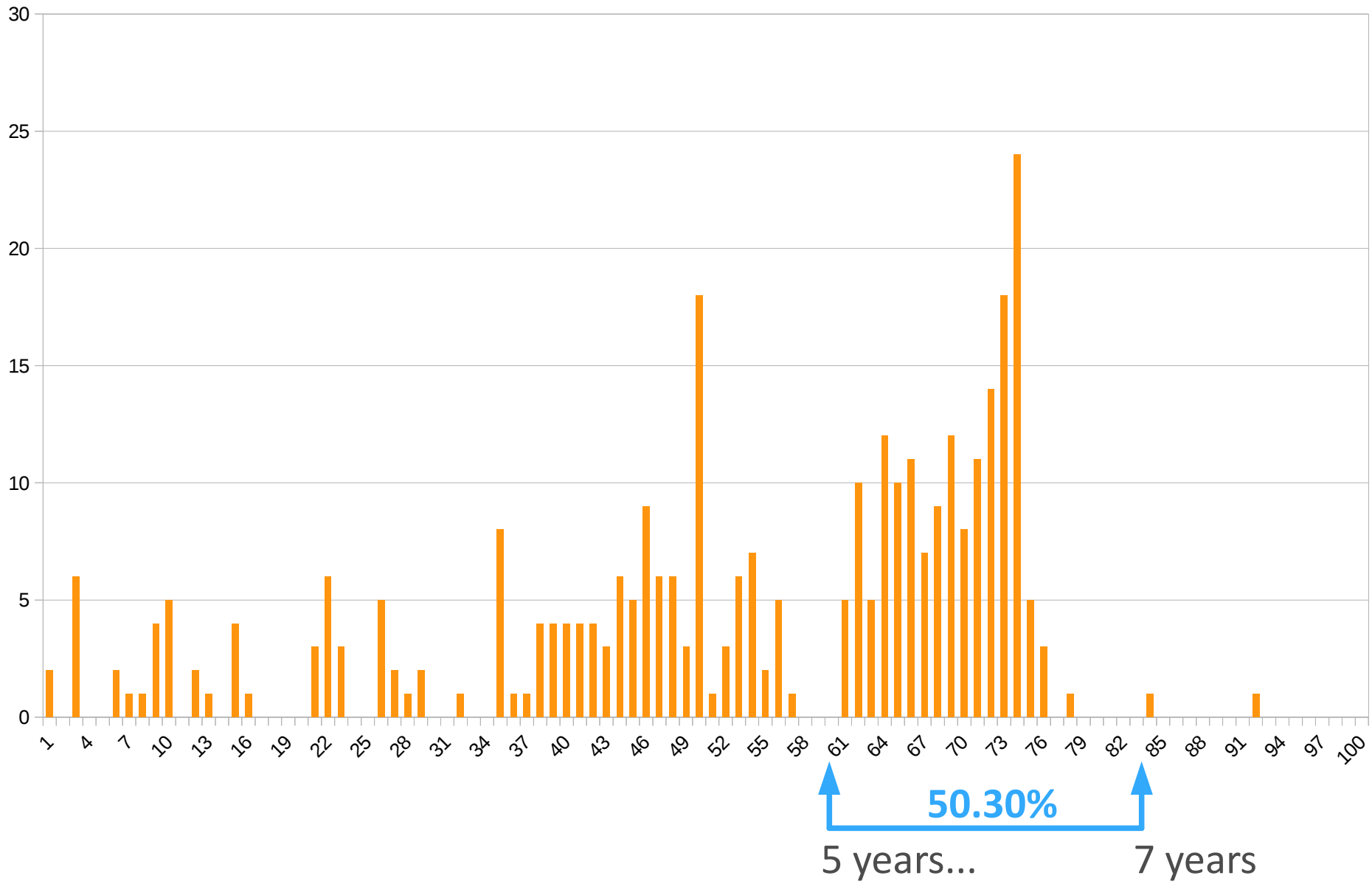
Draw whatever **chart** you like !

Issues over the host lifetime



Draw whatever **chart** you like !

Issues over the host lifetime



Agenda

Preliminary notes

Straight to an example

Architecture concerns

Some figures

Evolutions

Conclusions

Envisaged evolutions

- use dispatch API for incidents requiring human diagnostic (PSUs...)
- integrating non Linux/OMSA boxes like storage arrays
- integrating worker nodes
- enrich logs to let ELK do the graphs

Agenda

Preliminary notes

Straight to an example

Architecture concerns

Some figures

Evolutions

Conclusions

Key conclusions : how does this optimize sysadmins time ?

- avoid vendor support traditional portals
- no interaction with vendor staff
- rough estimation is 80% time saved using self-dispatch (720s / case)
- rough estimation is 95% time saved using APIs (195s / case)
- **TOTAL : 54h** saved for the 242 automated cases last year
- virtuous side effects :
  - full internal incident tracking
  - a saner environment where parts are replaced faster
  - clarifies your processes and cases allocations

Van még kérdése?

(google-translate© courtesy...)