Labeling Automation

PSSA Pillar

This document is part of **PSSA Tooling pillar**. This automation is expected to contribute to the following two Operational OKRs

- Improve operations efficiency by 30% - Tickets per Ops Engineer
- Reduce ticket volume per host by 40%

Problem Statement:

Recognizing patterns in tickets is crucial for identifying and organizing them into appropriate categories. However, manual categorization is tedious and nearly impossible given the number of categories and the complexity of the JQL queries involved. Without a proper identification & categorisation of issues, we miss out on significant insights that could be unlocked through accurate data analysis and trend identification.

Issue Categorization:

We analyzed six months of ticket data for VMIDP, VMIFM, SCDP and SCCP queues, based on summaries, identified recurring themes to determine which issues required dedicated categories. By grouping similar issues, we could identify recurring patterns that may indicate systemic problems or opportunities for improvement. This analysis also helped us focus on priorities by determining which noise reduction efforts were worth the effort.

Once we had all categories and labels sorted out, we integrated them into our automation code. These labels now serve as the foundation for our Jira & Retrospective dashboards with trend analysis, enabling better decision-making and visibility into the data.

Pattern identification & Ticket Trend Analysis:

We derive weekly data on categories through labelling, enabling us to conduct trend analysis. This process helps us identify emerging patterns or recurring issues that require our attention.

- PSSA team perform pattern identification exercise on weekly basis
- It is done based on above labelling method & categorization which helps us to easily figure out if there is ticket influx in particular category.
- Week on week comparison dashboard also help us to compare trends in past weeks/months
- PSSA team also has deep dived into each category tickets to perform root cause analysis to fix the issue and reduce ticket count.
- Further they also check for the opportunity to create run book.
- Based on above, team raises SCCP, SCDP, VMIDP, or VMIFM ticket with Service team who will further investigate reported issue & apply fixes.
- JIRA Ticket alerts are getting labeled for respective identified issue & JQL along with other investigation details will be added in ticket raised with Service team.

Ticket categorization process using label assignment:

1. A manually triggered python automation fetches previous week's tickets on pre-defined summary category. Date & Time window considered for this ticket collection is set between Sunday 00:00 AM to Saturday 11:59 PM UTC for previous week.
2. Perform Ticket Categorization based on summary keywords searches patterns through an automation & apply Category Label to each resolved ticket
   Example:
   - ***vmidp-canary-issue***
   - ***sccp-latency-alarm***

3. If labelling automation in Step 1 doesn't find a matching summary pattern, then it will apply ***others*** label (as in ***vmidp-others***)

4. ***"Others"*** category tickets will be reviewed manually by the appropriate team in their weekly retrospective call to add appropriate existing category label based on the issue

If we do not find pre-existing category for an issue & if such issue has multiple occurrences, then we will create a new category label & assign the ticket to new category label.

Wherever there is lack of pattern emerging those tickets will remain in the ***"Others"*** labeled category. PSSA engineers keep reviewing this category of tickets for any pattern of issues that may emerge.

Data Reporting and Service Follow-ups :

1. Data Reporting using dashboard: Human triggered automation is setup to integrate with JIRA to report tickets based on issue categories with the following:

- Load the Teams config file which specifies the Team Name, Jira Queues and their respective Categorizes and the Filter Patterns .
- A filter pattern is what the ticket is matched against. Generally it is a list of Jira fields to match like Summary, Labels, Description, Components, etc.
- Fetch all tickets for the queue from Jira SD.
- Take each ticket one by one attempting to match against the categories in the config file, taking the first match to allocate the ticket to.
- If a ticket is not matched to any category, its put into the final `others` bucket.
- Once all tickets are Categorized, upload the labels to Jira for all categories and their respective categories.

- Reporting tables are created for weekly data & WOW trend for each category of issues based on Severity of ticket as Sev-1/2 and Sev-3/4/5 as separate tabs.

2. Ticket Trend Analysis & Response Action:

- Based on above labelling method & categorization, any aberrations or abnormalities or trend strains in a particular category for the previous week & WOW trends are analysed to identify the issues that are different than they are labelled for initially and shows a new sub-trend which may require a new categorisation.
- For any unusual trend or in others category or even before lodging into others category, a deep investigation to identify the root cause or issue source is performed
- Service tickets are created wherever it merits based on the investigation outcome and for those that shall be dealt by ops are fixed immediately to arrest or reduce the trend. However if it requires a service level fix then it will be reported to the service team with a ticket.
- Once we raise a ticket with the service team, add Service Ticket number as “label” to the tickets relevant to reported issue.
  Example: SCCP-39936
- Update Operations Review Dashboard with this ticket information which will be discussed with Service team during weekly Ops Service call.

3. Time Saved from Manual Assessment:

Before automation, the dashboard required about **~ 4 hours per week** to compile all the statistics, such as weekly created and resolved tickets, categorized by severity, as well as week-on-week trend analysis based on categories. This was a manual process performed by engineers every week. With this Labelling & ticket categorization automation, manual effort of creating dashboard is eliminated, allowing us to generate reports instantly. Prior to the automation, engineers had to spend time dividing all Jira tickets into different categories manually to understand trends. With the automation in place, this ticket categorization & trend analysis is streamlined.

By automating this process, we not only save time but also gain more actionable insights into the issues, enables us to perform trend analysis quickly. In conclusion, the Retro Dashboard Automation brings benefits and direct time savings are measurable, with **16-20 hours saved per month** with labelling & categorization.

4. Improvements Achieved:

The automated dashboard helps identify trends and patterns that would have otherwise delayed to report or go unnoticed.
Combined with the faster identification and resolution of alerts, we expect an overall increase in efficiency, productivity.

Example: SCCP-39936 - Multiple High CPU utilization alarm

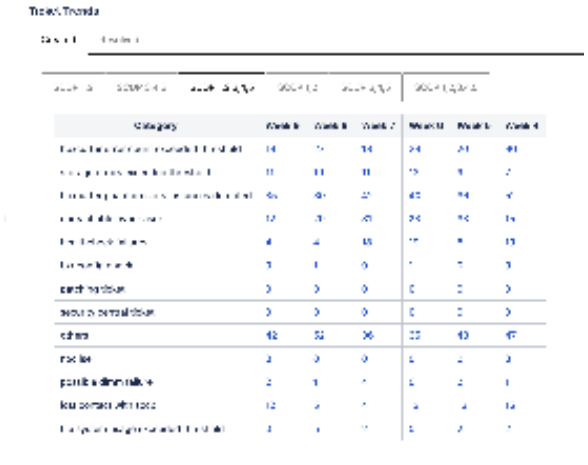
During our weekly trend analysis, we identified one particular alert **High CPU Utilization Alarm** showing an upward trend. This alerted the team to a potential issue, prompting us to escalate it to the Service team. As a result, the problem was investigated and eventually fixed, reducing the frequency of that specific alert.

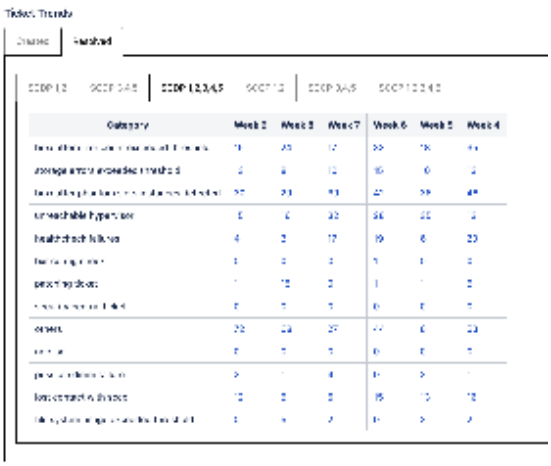
Using Labeling Automation and Ticket Categorization, we have reported various issues to Service team. Some of them are listed below.

- SCCP-39936 - Multiple High CPU utilization alarm
- SCCP-39681 - Alarm rule for BM instances stuck in terminating
- C3E-9211 - Disk Space Issue in Multiple Regions (VMIDP)
- C3E-9329: Automation Fails to Create Jira Tickets Due to Integration Issue (VMIFM)
- VMIDP-9710 - Secondary NIC down for HV shape: HV.Standard.A1-WB* series
- VMIDP-9868 - Stuck Terminating Instances
- VMIDP-9594 - [OKE] Volume Stuck in detach state.
- SCDP-24921 : Analysis for boxcutter-hypervisor-01365.node.ad1.me-abudhabi-1

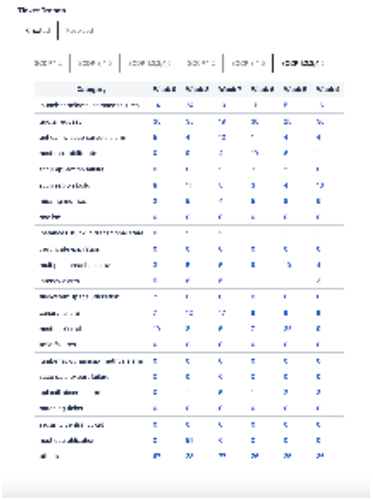
WOW Trend Analysis:

SCDP - Created WOW Trend Analysis **SCDP - Resolved WOW Trend Analysis**

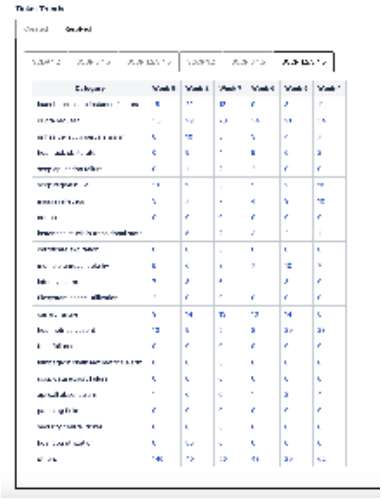




SCCP - Created WOW Trend Analysis



SCCP - Resolved WOW Trend Analysis



Weekly Trend:

VMIDP (Sev-1/2)

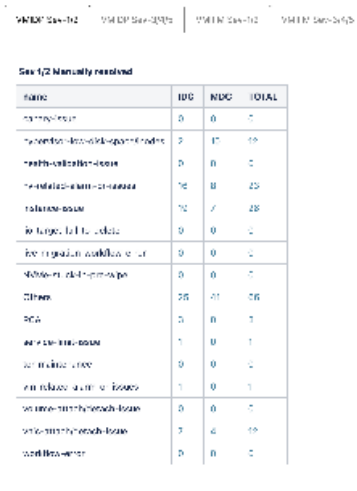
VMIDP (Sev-3/4/5)

VMIFM (Sev-1/2)

VMIFM (Sev-3/4/5)

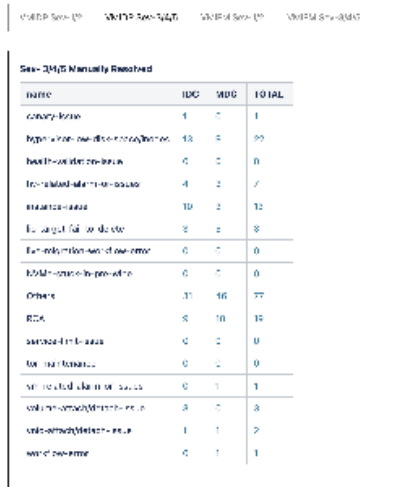
CIMM (Sev-1/2)

CIMM (Sev-3/4/5)



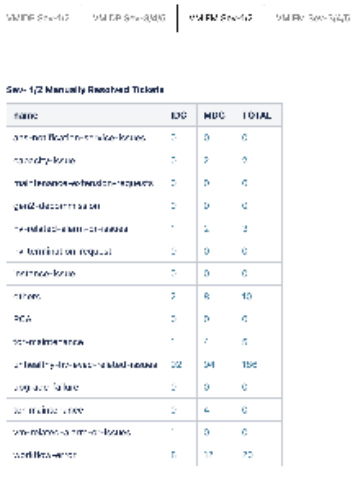
Sev-1/2 Manually resolved

Name	IDC	MDC	TOTAL
[illegible]	0	0	0
[illegible]	0	17	17
[illegible]	0	0	0
[illegible]	16	8	24
[illegible]	12	7	19
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	0	0	0
Others	26	21	26
RCA	0	0	0
[illegible]	1	0	1
[illegible]	0	0	0
[illegible]	1	0	1
[illegible]	0	0	0
[illegible]	3	2	10
[illegible]	0	0	0



Sev-3/4/5 Manually Resolved

Name	IDC	MDC	TOTAL
[illegible]	1	0	1
[illegible]	13	6	19
[illegible]	0	0	0
[illegible]	4	3	7
[illegible]	10	2	12
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	0	0	0
Others	31	16	77
RCA	8	10	18
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	0	1	1
[illegible]	3	0	3
[illegible]	1	1	2
[illegible]	0	1	1



Sev-1/2 Manually Resolved Tickets

Name	IDC	MDC	TOTAL
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	1	2	3
[illegible]	0	0	0
[illegible]	7	3	10
RCA	0	0	0
[illegible]	1	7	8
[illegible]	92	94	186
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	1	0	1
[illegible]	0	17	20

Sev-3/4/5 Manually Resolved Tickets

Name	IDC	MDC	TOTAL
[illegible]	0	0	0
[illegible]	0	1	1
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	0	4	4
[illegible]	0	0	0
[illegible]	0	0	0
[illegible]	3	1	4
RCA	0	1	1
[illegible]	0	0	0
[illegible]	28	44	72
[illegible]	0	1	1
[illegible]	0	0	0
[illegible]	1	9	10
[illegible]	5	8	13