20 Concerns that system needs to be addressed.

Config, Adaptor, Bean, logging, java - implicit I think?, xml, json, concurrency, memory, math, io, graphics, gui, Text, networking, fileio, sql, nosql, spring, console, persistence, debug, benchmark, performance - same as benchmark?, Bash

8 + 2 Word list

Concurrency

Graphics

Gui

Io

Math

Security

Sound

Xml

**Analysis part 1**

***Chukwa-0.4.0 me, nb and original comparision :***

In Chukwa-0.4.0 there are 3 major clusters named scr, build and contrib, all these 3 major clusters are depended or categorized as IO component where as in NB- it is categorized as security components, but in ME – they are majority of the clusters are not depended or they don’t match with any category.

Clusters are mainly dependent on Graphics, networking, SQL in old, where as in NB they are math and IO, ME – they are IO, GUI

For example – metric sub-components like abstract matrics context is graphics component in old , in NB it is security where as in ME it is GUI.

Another example is data loader sub-components like dataLoadFactory and socket factor are catergorixed as IO dependent clusters in OLD where as in NB they are SQL dependent and in ME they are IO dependent.

Similarly , configuration cluster with sub-cluster object are graphics dependent in Old where as in NB they are math dependent and in ME they are GUI dependent

(*Optional)* Heatmap , swimlane which are Graphics clustered in old , in NB and ME they are GUI.

Do u think recovered architecture is accurate – it varies example – in OLD where the basic concern of topic is IO, where as in NB it is security but in ME most of them wont match

Example – QueueToWriterConnector , BackfillingLoader – which are sub-clusters of Backfilling was IO components where as in NB they are security this doesn’t seem accurate because these are writer and loader components which will be more accurate catergorizing to IO or GUI component than security. Same way in ME they are no-match they dont dependent on anything.

Visualisation sub-clusters like heatmap and swimlane are graphic clusters where as NB and ME they are GUI , which actually seems accurate to be categorized or depenedent as GUI than graphics as they are visualization cluster.

Log4J cluster is catergorized as Sound component in ME, which doesn’t make any sense but in NB they are security and in OLD they are IO. It can acceptable to some extend they are of no-match to any cluster as I have not chosen logging topic but in all 3 cases they are categorized to wrong cluster.

Rank three result in term of there accuracy, - OLD ranked 1st, ME ranked 2nd and NB ranked 3rd

Criteria used for this is – in OLD most of the components are categorized properly to some extent like Client Finalizer, chukwaAgentController which are catergorized as networking components looks accurate when compared to NB where it is catergorized as security and in ME it is no-match as my chosen topic doesn’t contain network category.

***Chukwa-0.6.0 me, nb and original comparision***

There are only 2 major clusters in Chukwa-0.6.0 they are src and contrib, build is merged with src as collectors are explicitly removed from 0.6.0 version.

Similar to previous version Majority of the clusters in OLD are IO and graphics where as in NB it is security but in ME most of them no-match and few are IO