**Splunk admin session**

Why Splunk?  
- Volume of data ->Data in tera even peta byte is common and using Splunk we can do easily compared to DB  
-Types of data -> we can have structured or non-structured data. We can do both in Splunk without changing schema.  
-Dynamics analytics use case ->here search happen simultaneously in Splunk.

Because of all these main challenges analytics tool comes into picture that’s were Splunk introduced , but Splunk initially introduced for log analysis But later on it upgraded with its feature to full-fledged analytic tool.

**Splunk Architecture:**

Single server environment ->All components will work under this and using webserver we can upload data and do analysis. But if it goes down, then no body cant access and if data grows, then it impact performance issues. **Its recommended when we have small amount of data or for testing.**  
Distributed environment 🡪 here we have multiple linux server. And those servers helps to replicate data. Since there are multiple nodes, we will get much much better performance.

**To install software:**

* You should use, Splunk.com
* Where you can create account and click free Splunk.
* And check for Splunk enterprise for full fledged instance.
* And install Splunk software depends on OS.

For linux, depends on flavors obuntu(.deb) ,redhat(.rpm) etc.. eg. Fedora linux used by redhat and release it with .rpm version. Similarly, if aws use fedora linux and release as dif version.

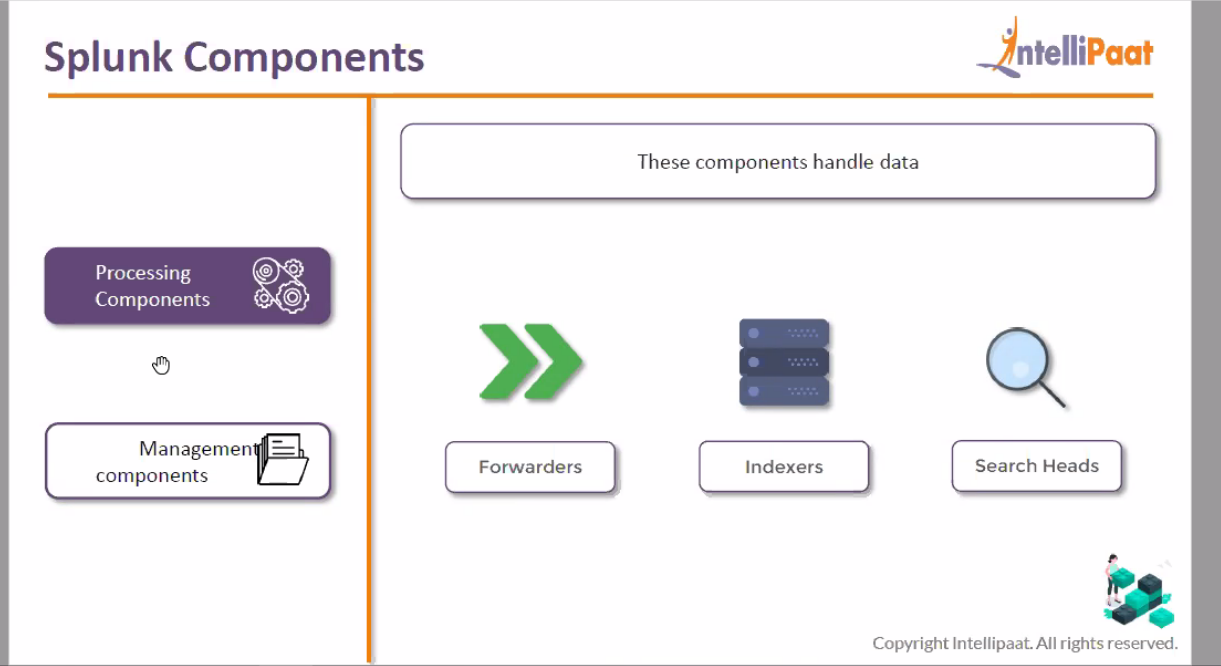
Fedora (contributed by multo ppl. Aws, redhat) so it has regular updates which will not good for pdicution. So we shd check for proper vendor and take proper version. So for productin,its good to have redhat or other vendor. For pslunk, we shd use .rpm. and obuntu is paid one(for this we shd use .deb)

Fedora ->fedora server would be preferable.

**Linux is must cheaper than windows and also security purpose.**

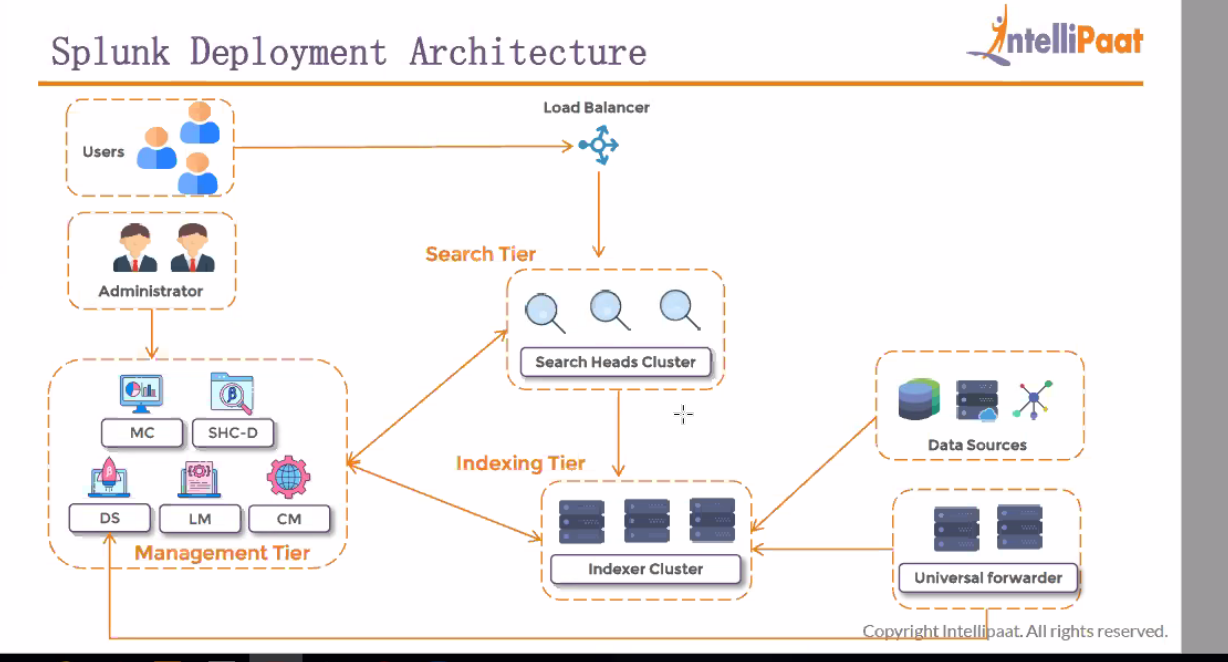
**One main advantage of using Splunk, all we can do in one. Like importing data,storing and analysis. Whereas if we take Hadoop for analysis dif and for store dif..**

Splunk **components:**

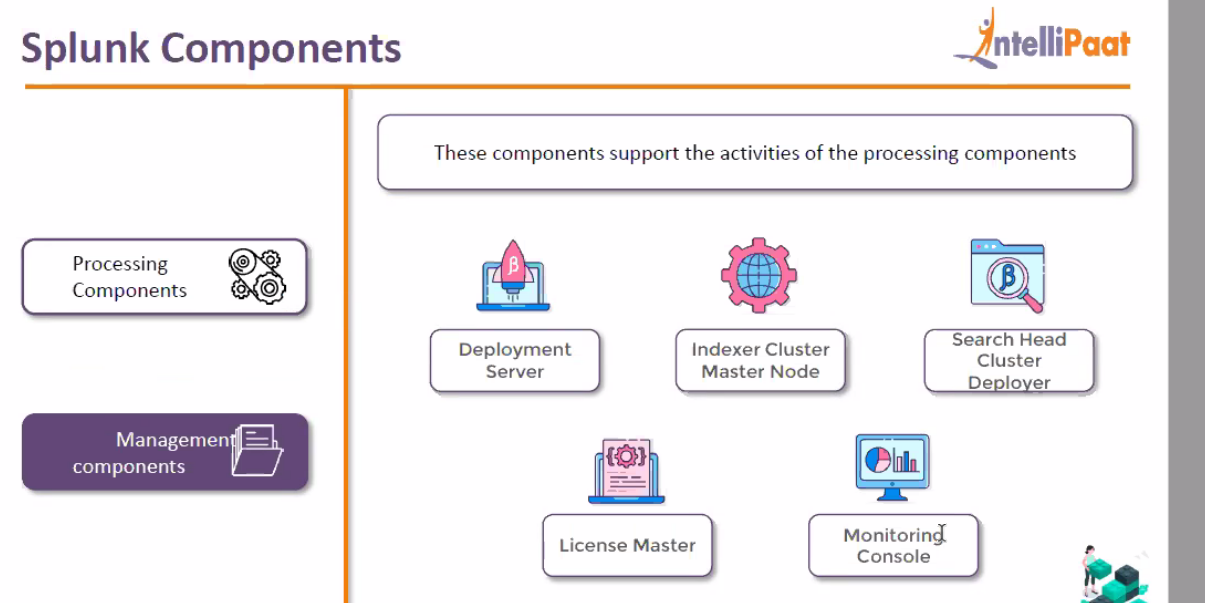
****

**You know detailed descripton of processing components such as forwarders, indexers and search heads.**

**Remember, processing happens in indexer and storing also happens at indexer. And results gets displayed on console by search head. Here serach head renders data from indexer and display on its console.**

****

**Here Managemetn tier is not resource hungry since it controls other component. So we install all DS,LM,CM ,MC and SHC-D all in one as mgmt. tier.**

****

**Forwarders managed by deployment server.**

**All logs which you want to collect to Splunk those configuration changes will be done in deployment server and will push those changes to deployment-clients(forwarders) using server class.**

**Index clustermaster node:** it mainly helps in manages index cluster and it also manages serach cluster as well. Here It monitor how replication happens at captain in search head and also how data getting replicate d in indexer.

**Search head cluster deployer: it helps to distribute apps or config changes across indexers and search heads.**

**License master: splunk is full fledged premium product. It will monitor hwo much data indexing and if ti exceeds license amount , it will trigger alert.**

**Monitoring console: How many users logged in to Splunk, how much CPU occupied by search head etc.. similarly we have few inbuilt alerts which we can enable if required,**

**Restart universal forwarder:**

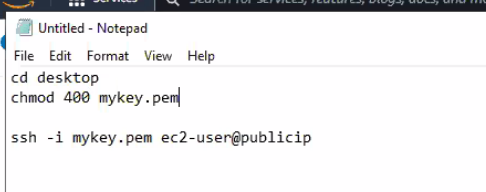
**, For version upgrade of forwarders, then we shd go for deplyemnt server but if upgrades on sale, u can go for individual one.**

**Or most of the case, u can script and push it across.**

**And most of the case, you can do restart of failed ones**

**Installation of linux instance:**

* **Launch instance**
* **If you are using moextem, there u use public ip, usre name as ec2-user and apply private key and launch linux instance by selecting ssh session.**
* **If its Mac os, then open terminal and navigate to directory where u kept .pem file. And run ssh command as below. Here we kept .pem(key pair file) in desktop.**

****

**Install splunk instance:**

**In EC2 instance, we get private address which is actual server ip address. And public address temp kind and which we use while launching instance.**

**Always keep in mind, be as splunk root user to install splunk and to isntall software always install inside a foler instead of installing in OS default folder.**

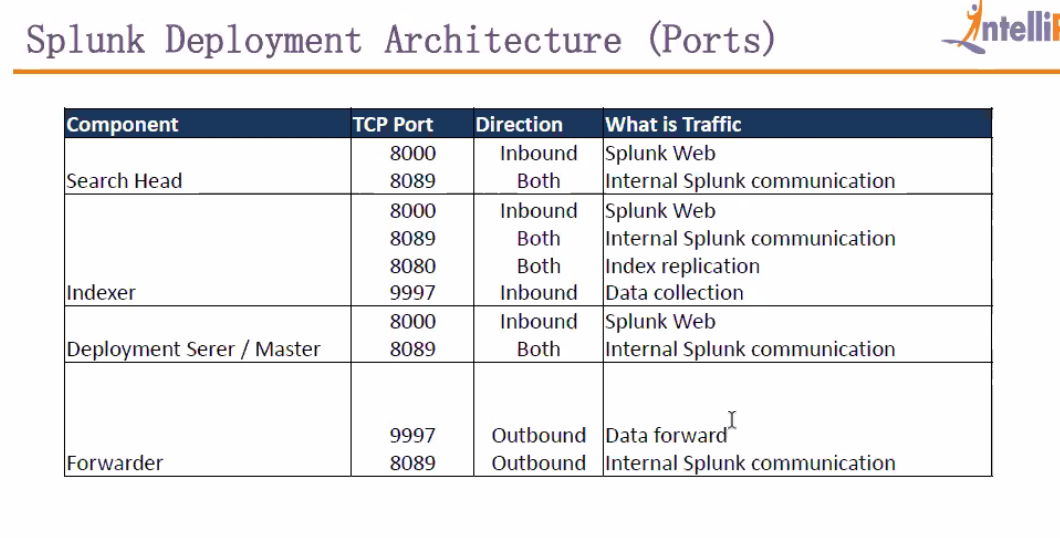
**So that if something happens, it will not crash OS.**

*If you are not admin, then ask admin to create mount folder with size of 30GB.*

*Why we are setting min space is 50MB. Splunk minimum free space requirement is 5GB. Since its free trial, while installingit occupies 4.5 GB. So to make it work, we are keeping min space as 50MB. But in production it shd have 5000MB.*

*And that comes by default. And if tis about to film then we shd create volume and apply to disk using lvcreate and fdisk commands.*

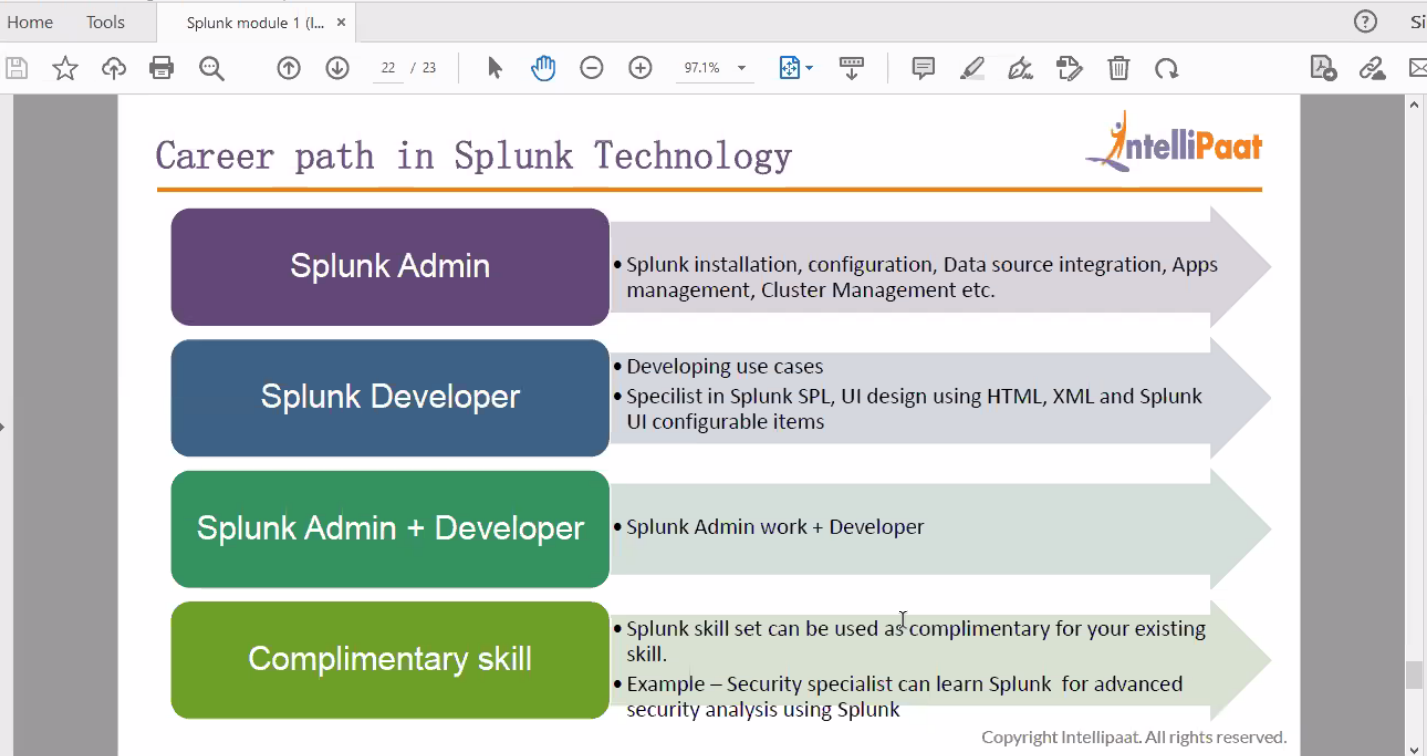
***Ports:***

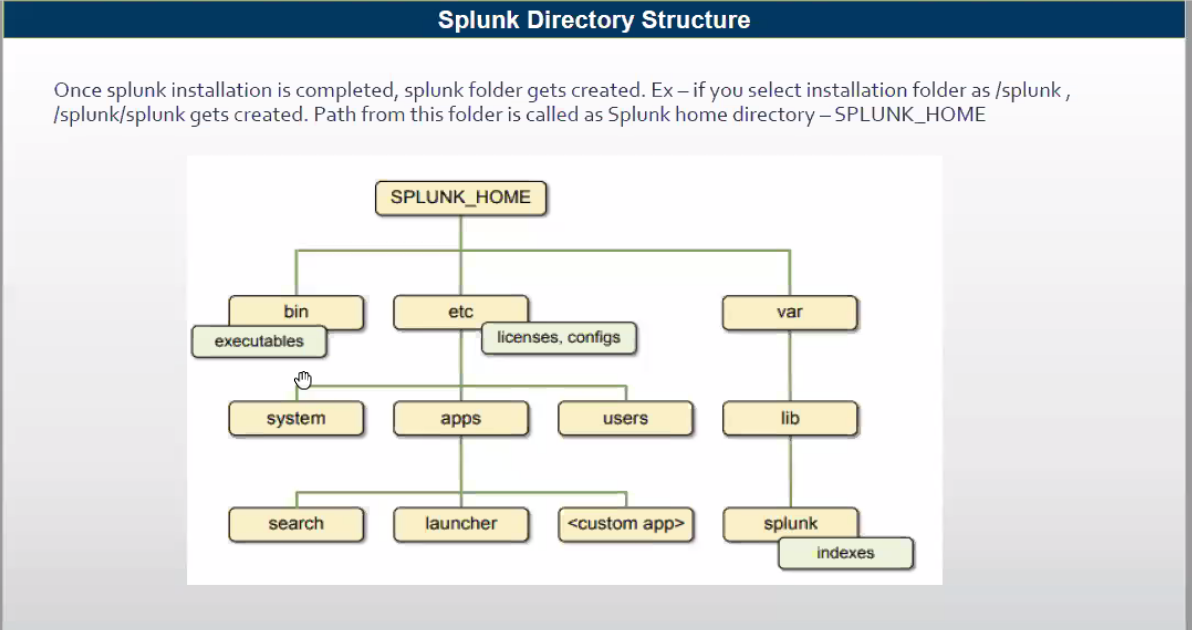
******

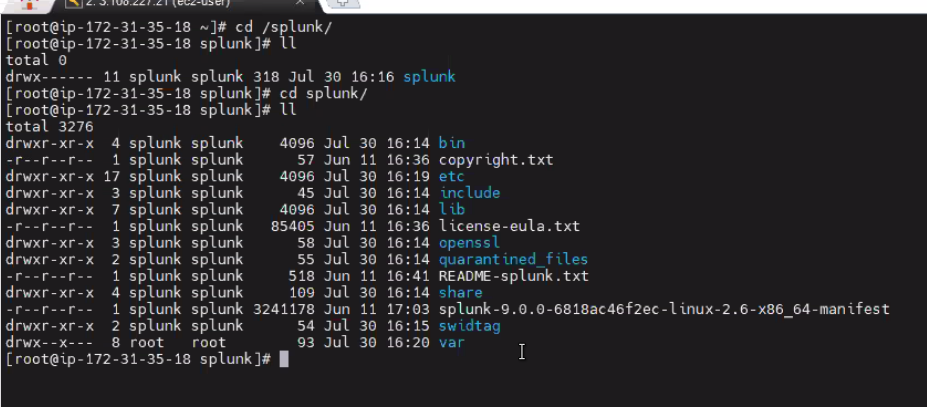
***2 types of flow we have***

***Inbound->data comes from user to search head***

***Outbound->data comes from indexer to search head***

******

******

******

***Default folder is Splunk and generally we call it as SPLUNK\_HOME***

***And inside that we have some default folders***

***Bin:***

***->here we have*** *hell no of executables. But all areused internaly by splunk. But only we use splunk execuatle file to apply config changes including start and stop splunk server.*

***Etc***

*->apps  
 ->here we will get info about apps. If u see apps in homepage all will get dsplayed here.*

*In phone, if u buy it will come with default android os and default apps. And if u need to install, u can install from play store or if u want u can create as well.  
similarly, in splunk if u want u can install from splunkbase.com or if u want u can create on ur own. And these apps gets stored under /splunk/splunk/etc <here splunk new mount folder where we installed splunk>*

*🡪system  
 default->here we have default setup/config setup. If any changes need, we shd copy frm here and paste it in local and do modification. And once done, restart splunk.  
 local ->here it has internal SSL config infomaton, disk freespacce info, And* ***files in /local folder has highest precedence. And it has customized config files.***

*->users  
-if splunk used in multi persons, then we create users. Those users foldrs are created under user folder and their metadata and user preferences are saved under their folder.*

***Var:***

* */lib/log  
  here we have log files relates to splunk. Here we get various logs. In wich important one is splunkd.log where we get to know if any transfer is nto happenng or any server is not startng.or any search head or indexer not starting.. those we can get to knw from splunkd log.*

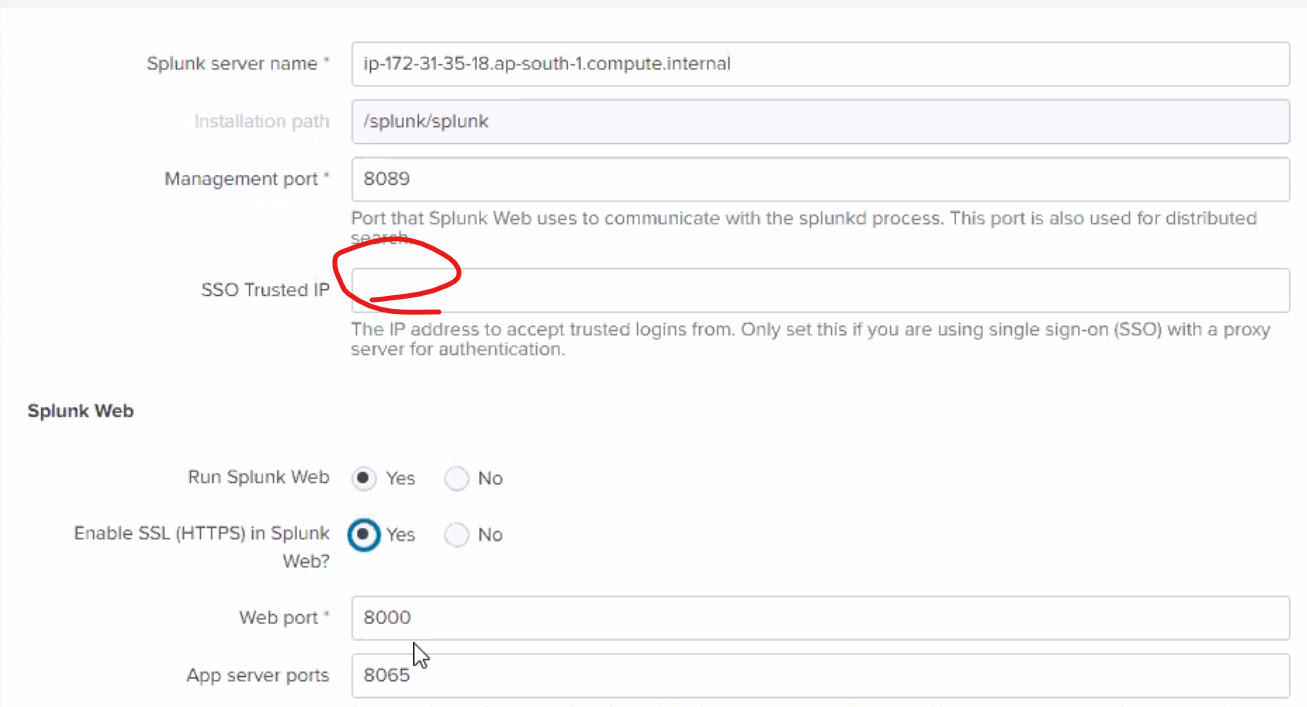
***31st July-22:***

***Preferences-*** *to change default home app and to change tiemzone  
Time zone will not change timing of raw data and it jsu change for splunk UI*

***Messages –*** *its good practice to check for it and take action as soon as logged in*

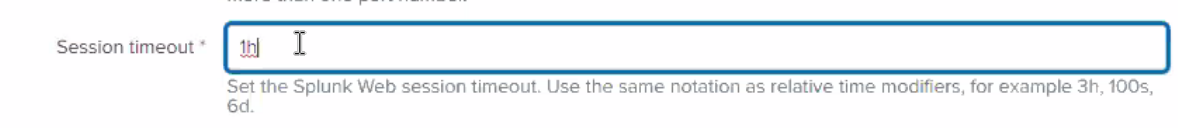
***Settings-***  *left side u have add data,explore data and monitoring console.*

*And along with it we have 5 sections.  
Knowledge ->this we covered under developer part  
Data  
Distributed environment  
User and authentication  
System*

*Under* ***server settings:  
You can change server name, port and index pausing as well.  
***

***If you are using Single sign on, then u can update trusted ip’s there if u really want to implement those in splunk***

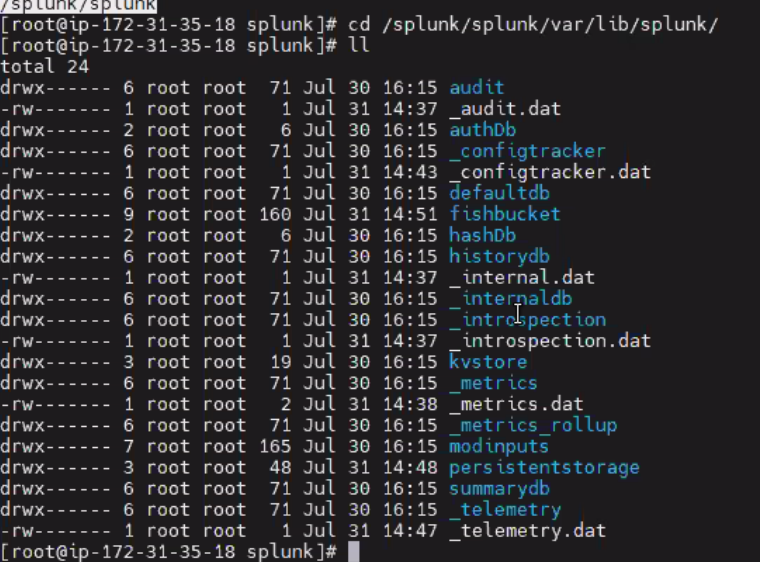
***App server port->mostly we will not open. Only we use webport often.***

******

***As per current policy, we need to set this up.***

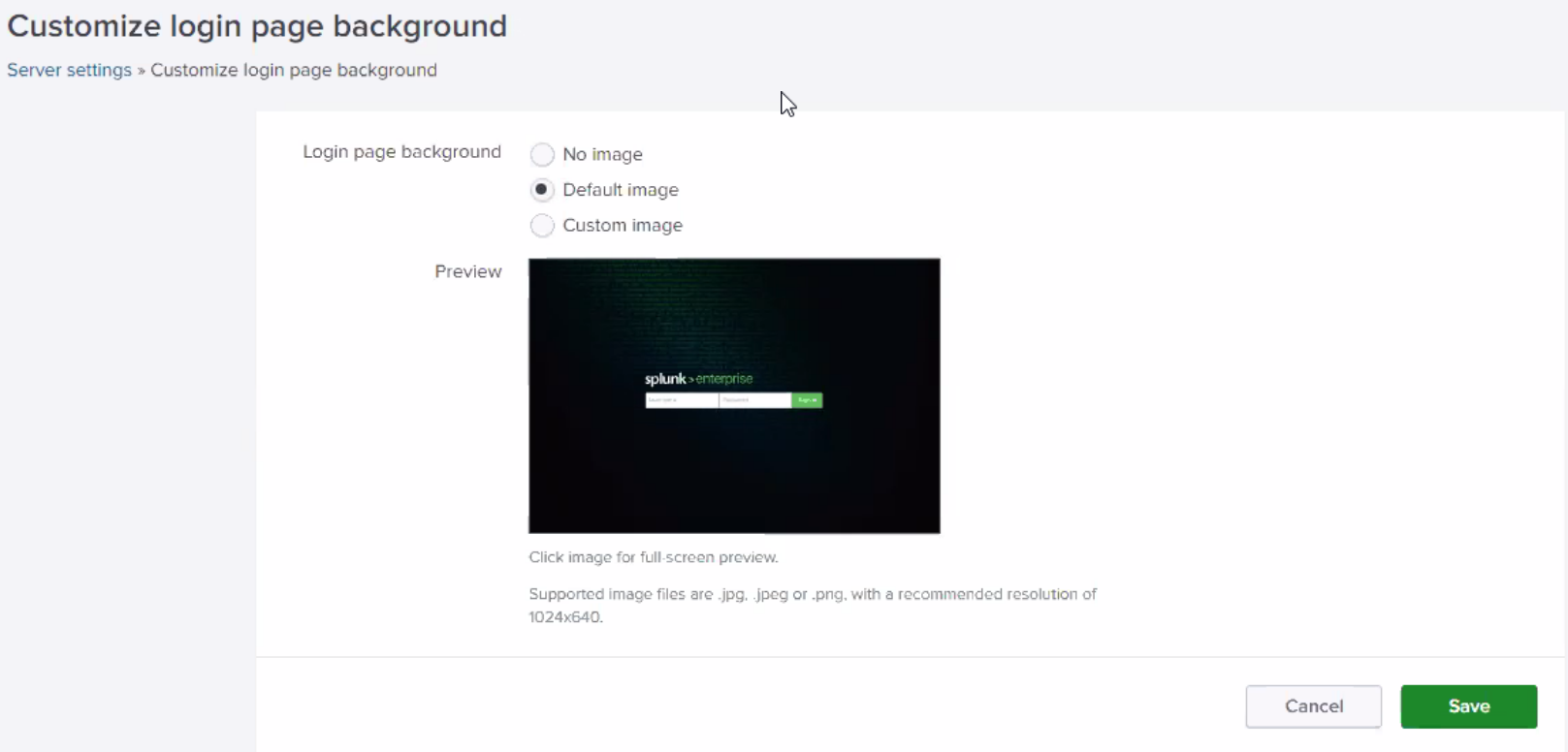
******

***Here, we have our indexer, search head everything resides on same. So here to specifically identify index with hostname, u can change here. And if you want to pause free diskspace then u can specify here. Similarly, u can update under server.conf as well using stanza named [diskUsage] and with info as minFreeSpace = 50***

******

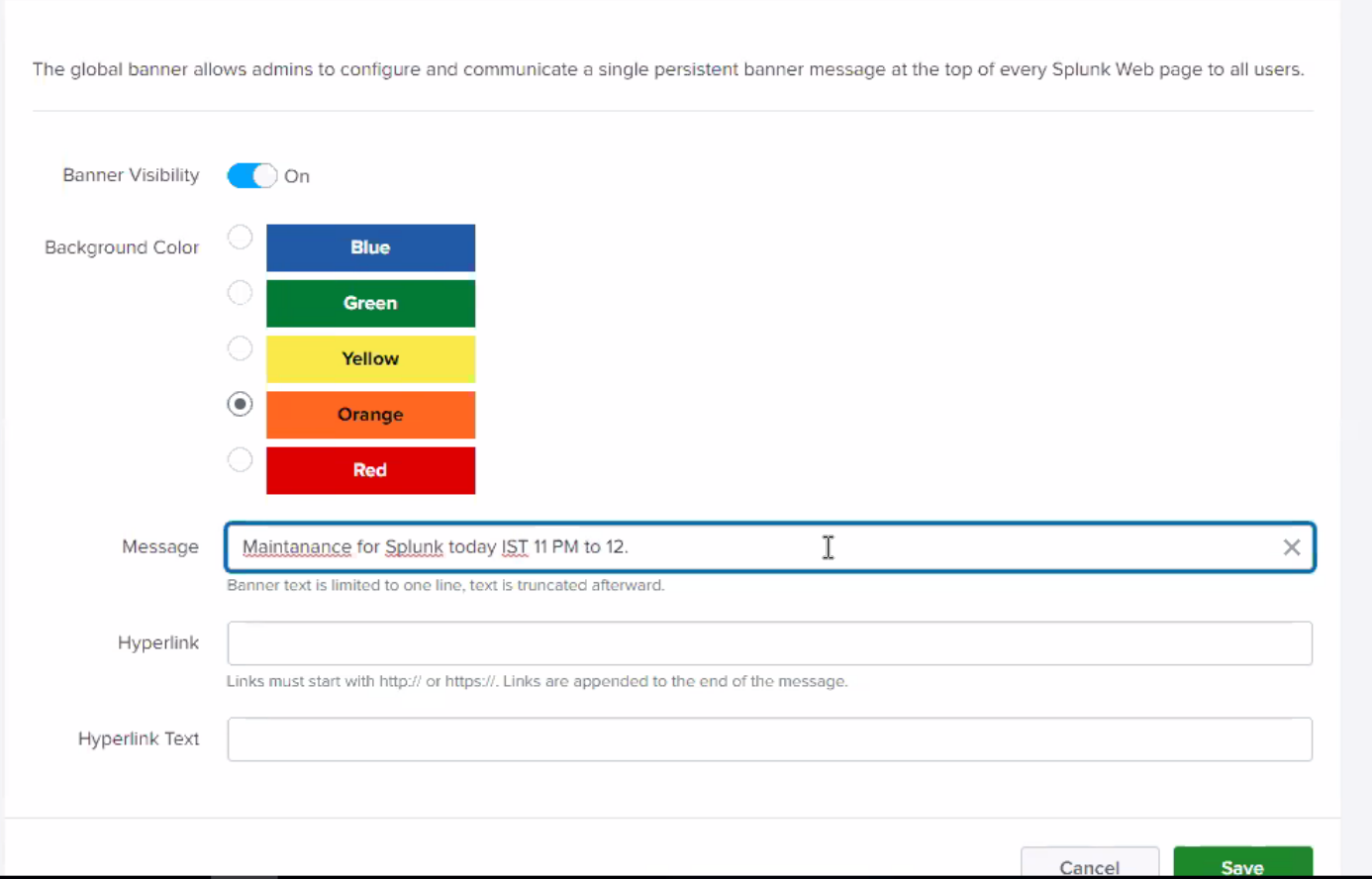
***This is path where default index gets generated.***

***To change custom background:***

******

***Here image it shd be .jpg,.png or .jpeg and pixel size shd be 1024\*640 and restart now***

***Global banner:***

******

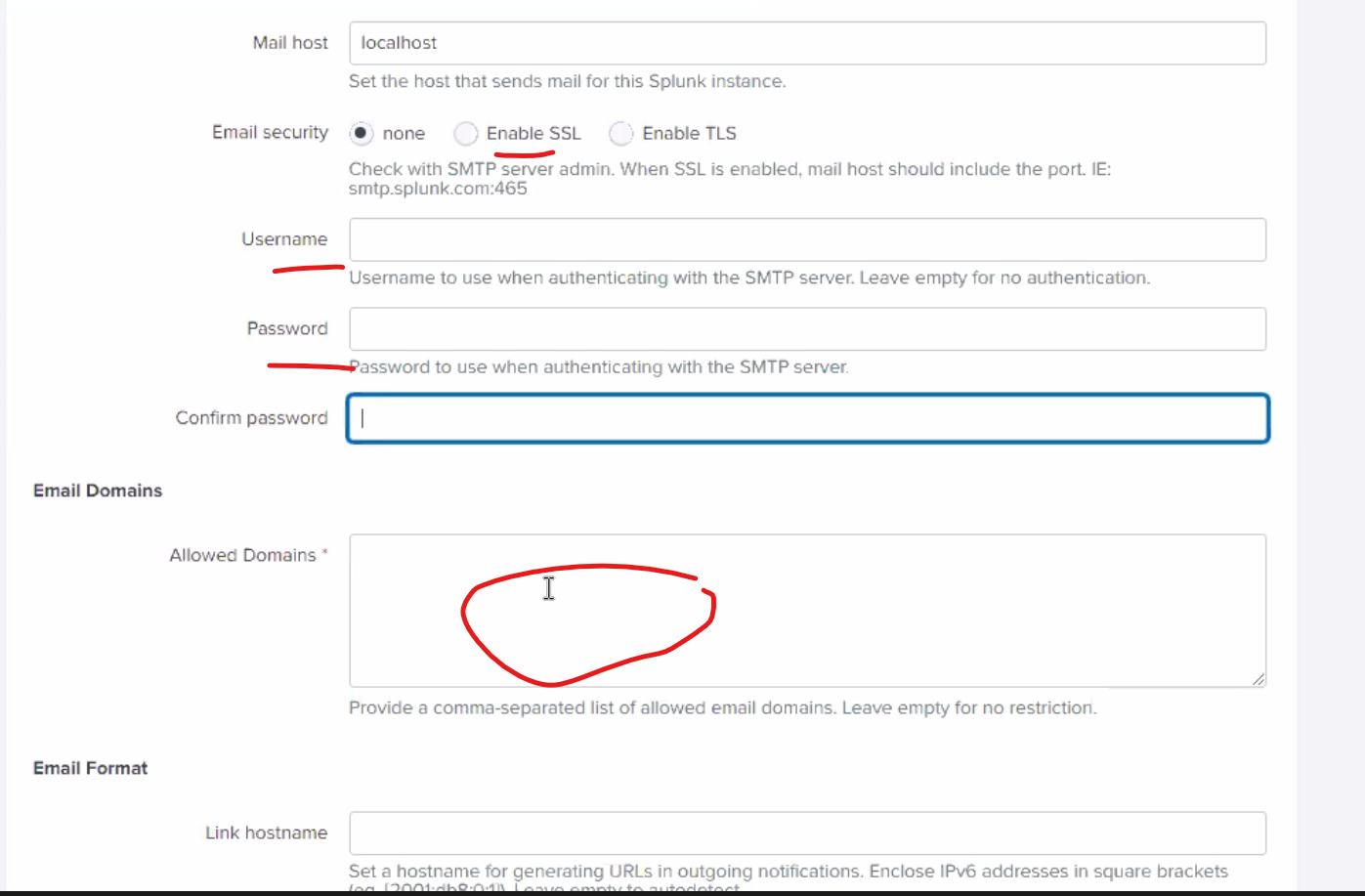
***If you save ike this, as soon as u loggedin, u will get message on top.***

***Email settings:***

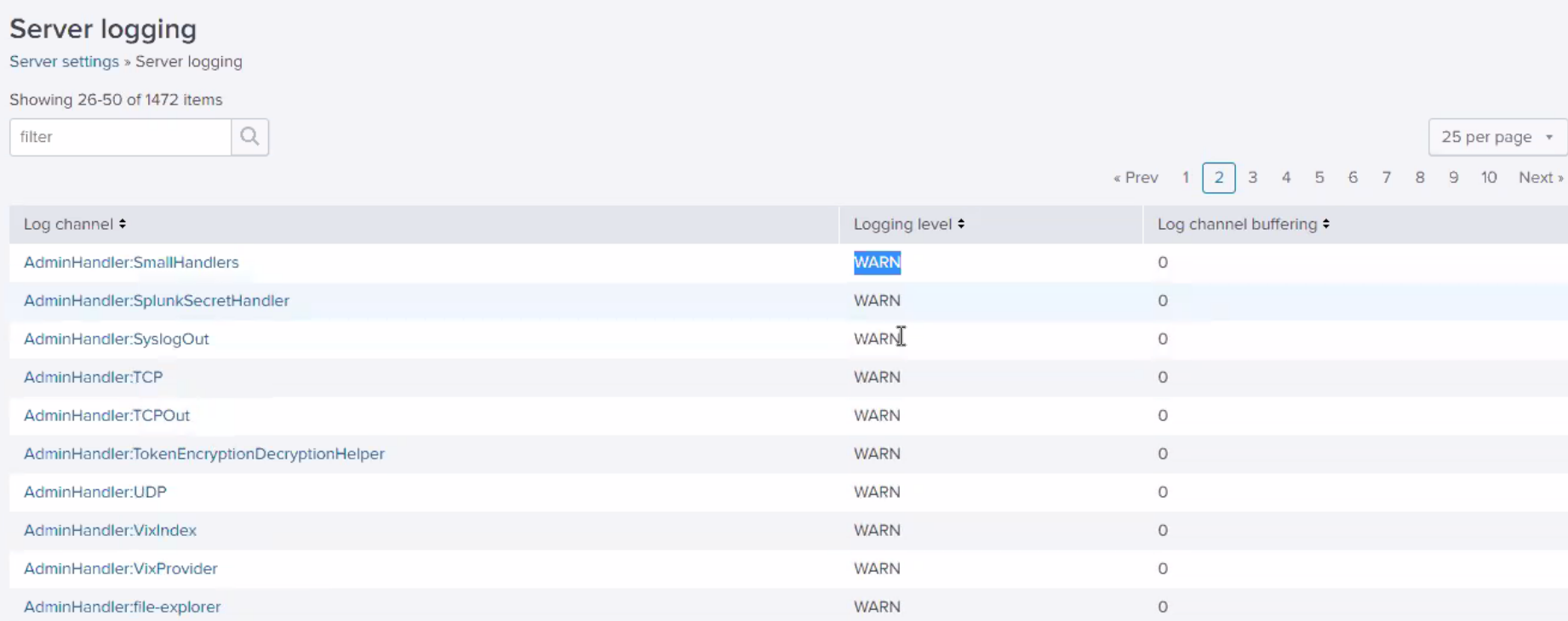
***Even in alert, if u setup alert it will not deliver because splunk is not aware of email server so its not sure to accept email request.***

***So here, we shd define email server by checking email team and get information for smtp email server information and put ip address, and their credentials.***

***You can also restrict it***

******

***Server logging:***

******

***Here we can see log channel and its logging level. And mostly we wont change log level. And if suppose we want to restrict log level, u can change it.***

***But we cant add it***

******

***Here we can change default time range. By default it takes 24 hours. If u want u can change it.***

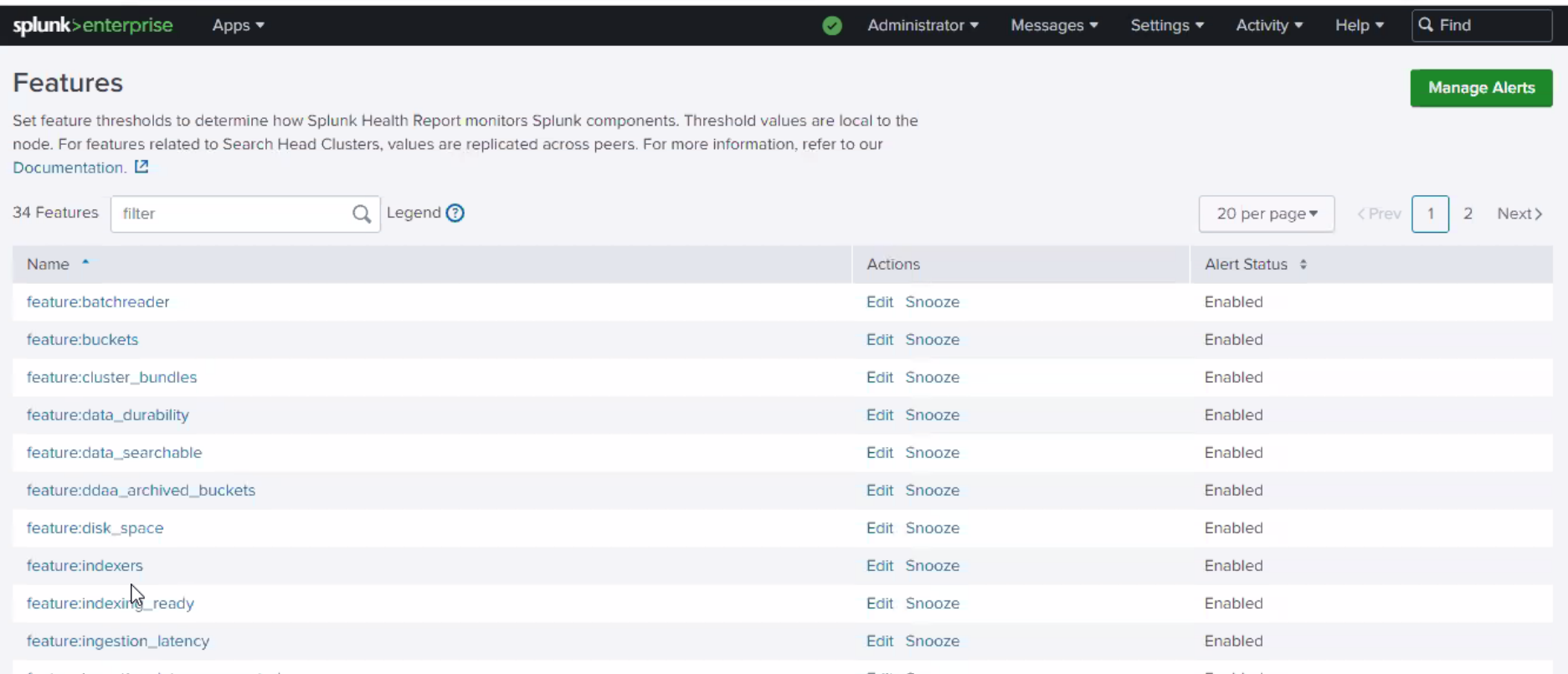
***Relative concurrency for schedule seach means no of searches scheduler can run. Means we have 2 types of serach one is live search and other is scheduled or report. Acc to admin decision, we mostly will not change but if lot of developer running more, then we shd change it,***

***And relative concurrency for summarization searches-> this limit for summarization search but it has least precedence than schedule search.***

***Server controls:***

***This helps to restart splunk from GUI***

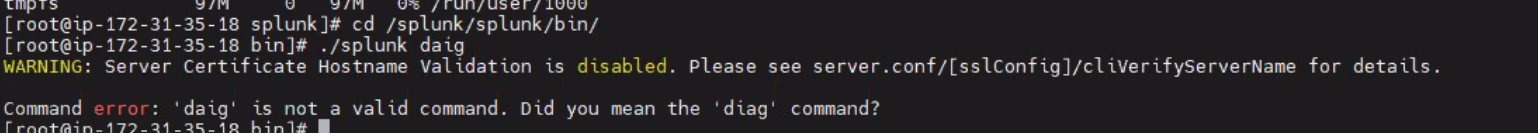
***Helath manager:***

******

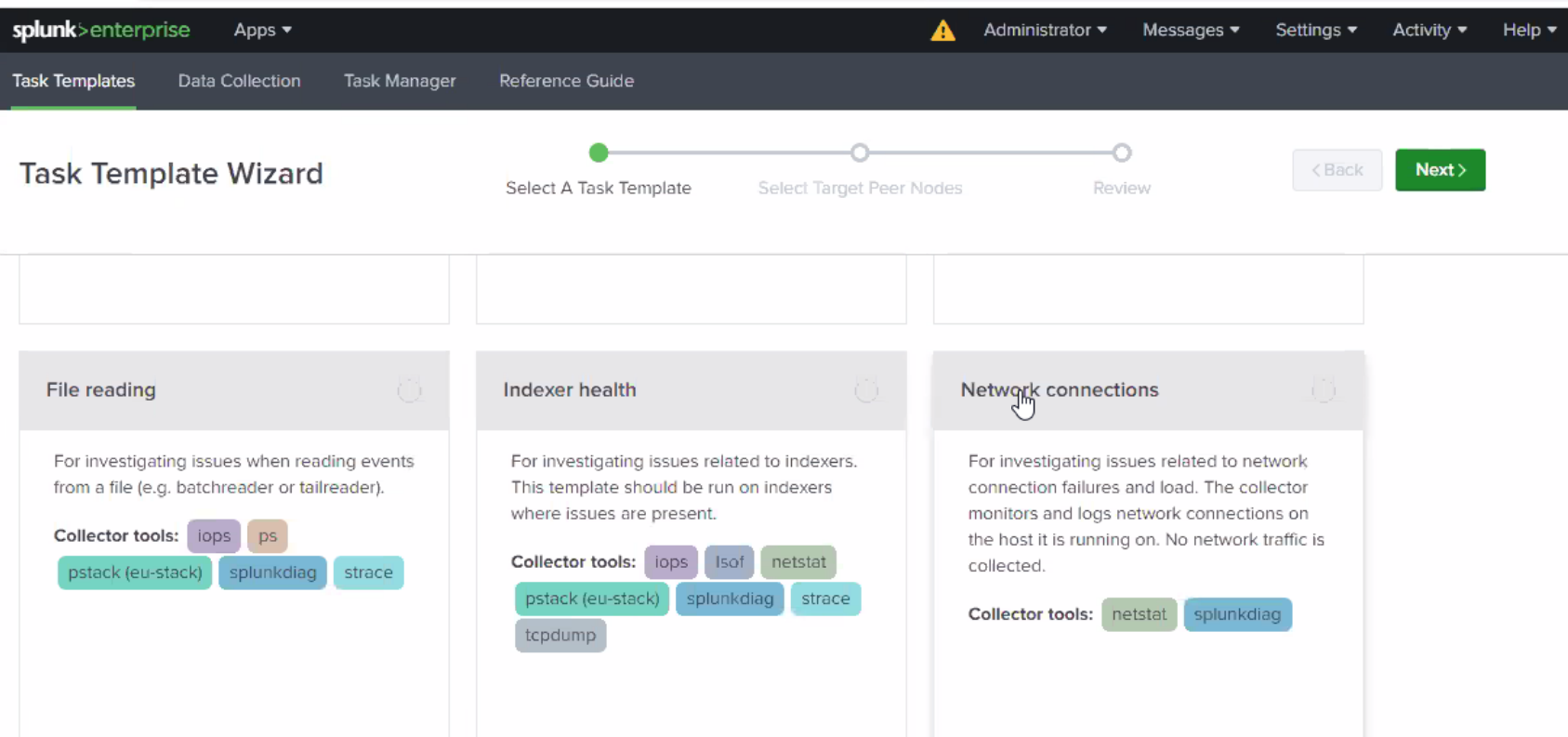
***Whenever there is some issues, by default alert trigger messages and those are comes from above one alerts.***

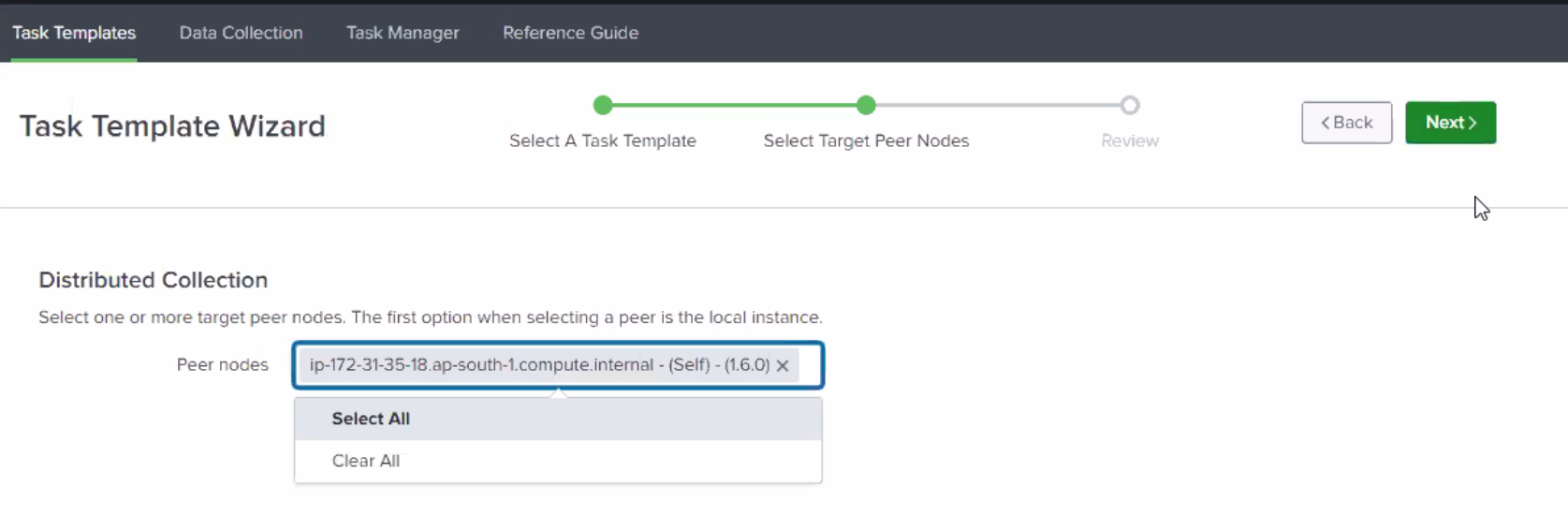
***Rapid diag:***

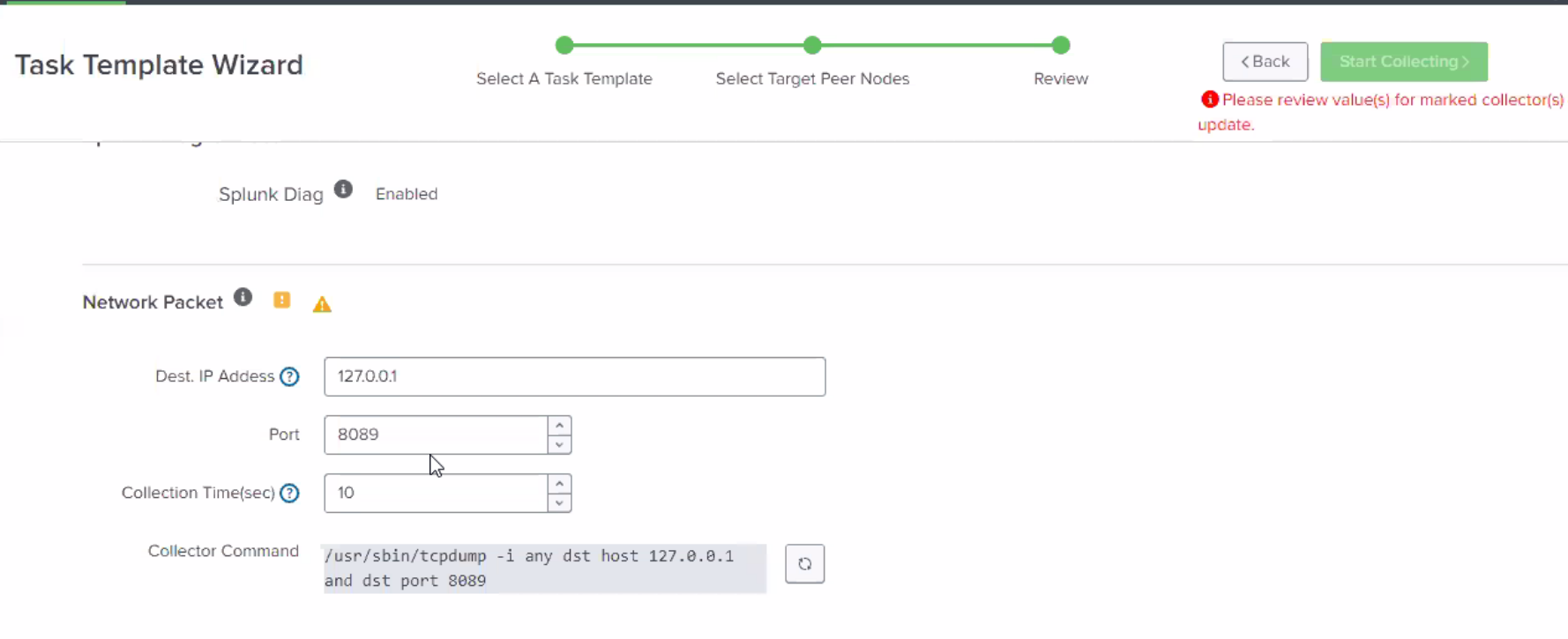
***If we go to splunk support team for any issues, forst they will ask to submit report diag.***

******

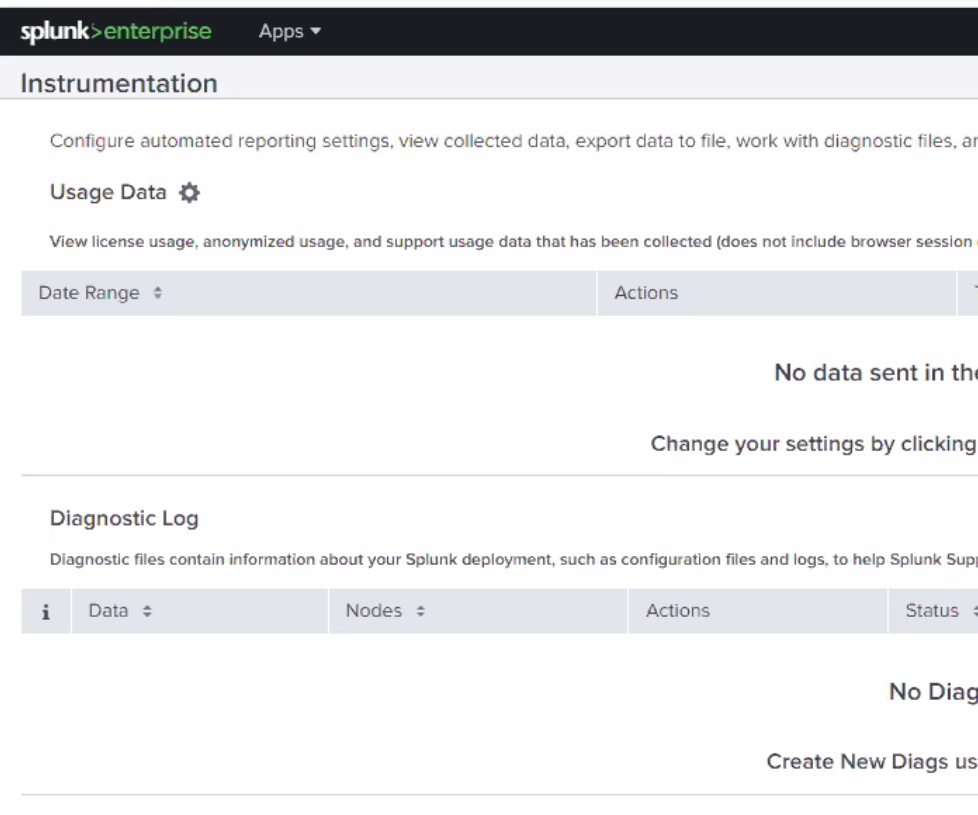
***It will generate .tgz file which will have log file etc. and if we submit to splunk support team and they will monitor it and analyze it,***

******

******

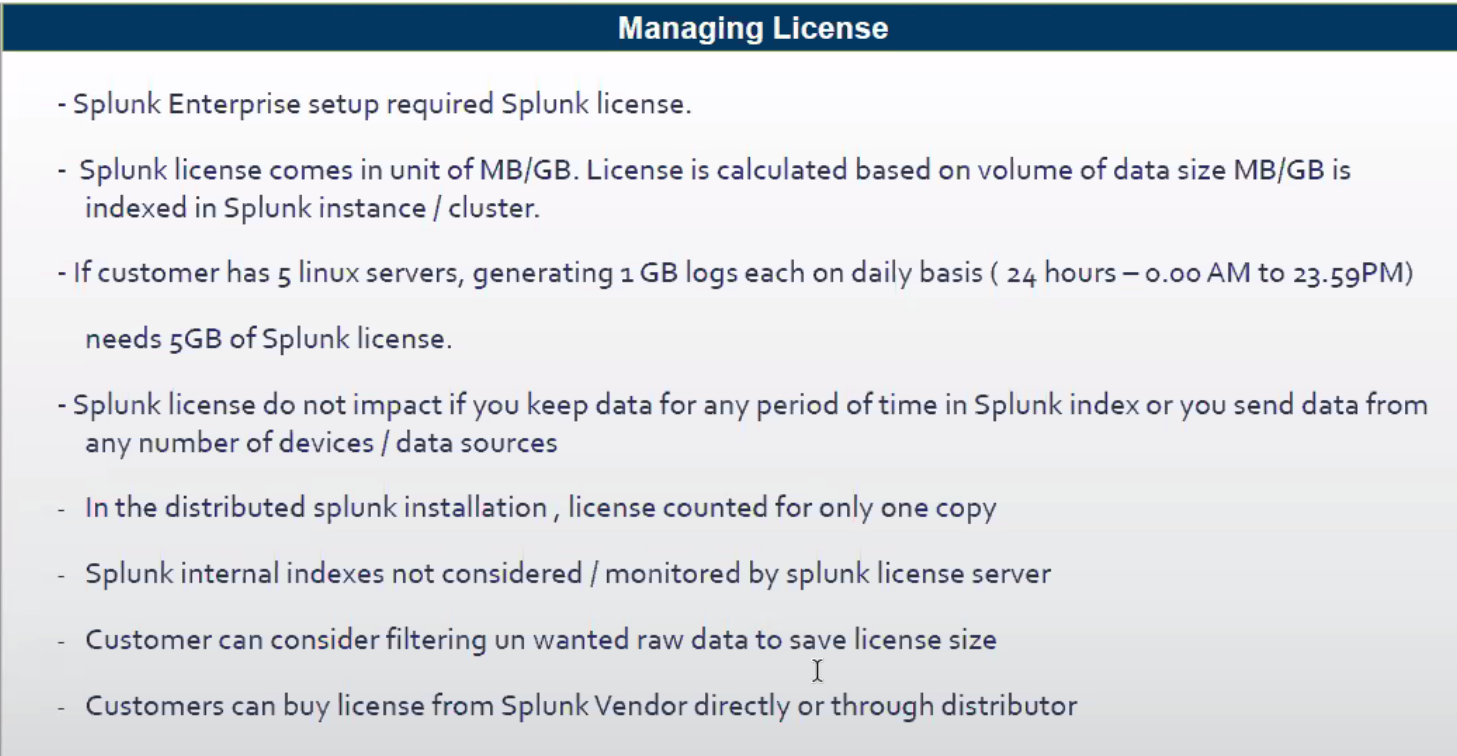
******

***Once u start collect, it wil generate .tgz file and that u can submit to support team.***

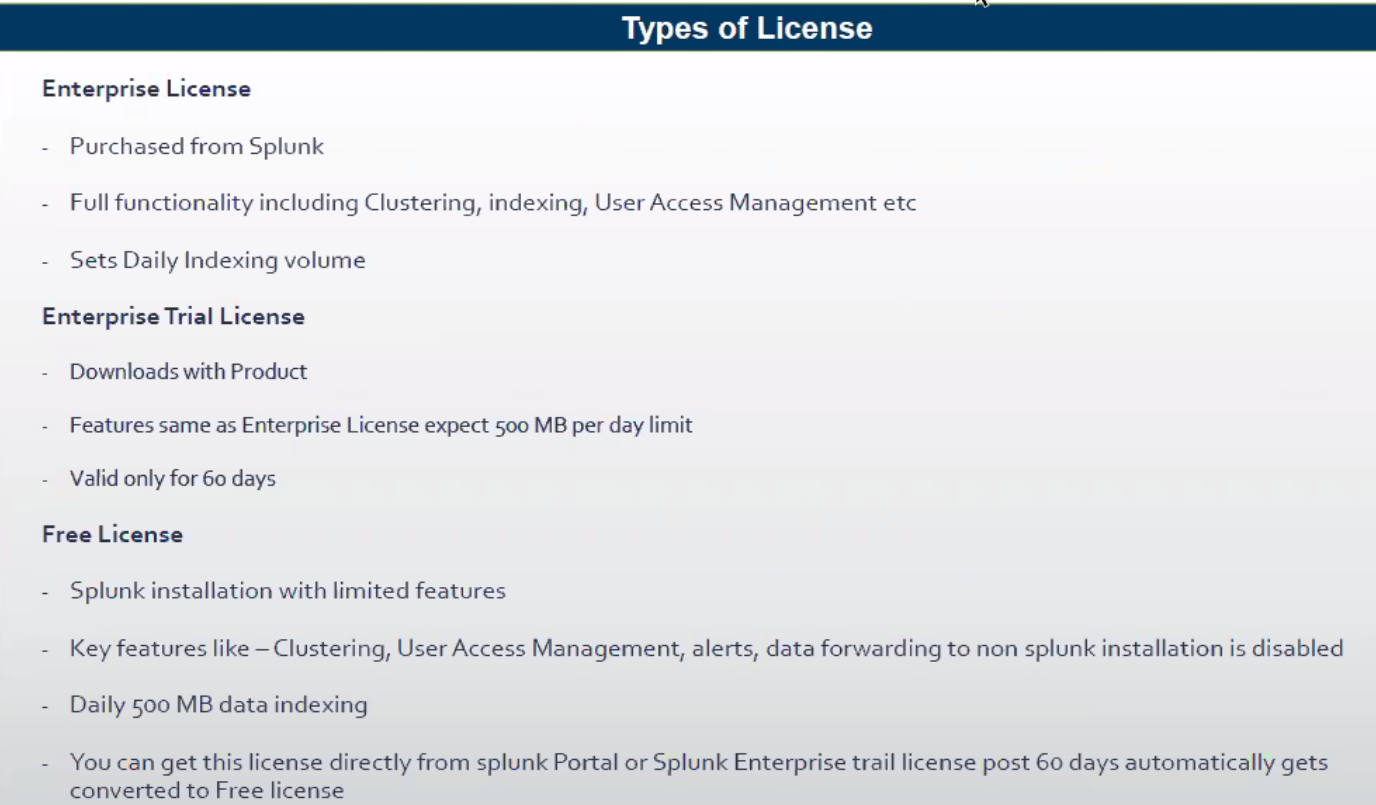
******

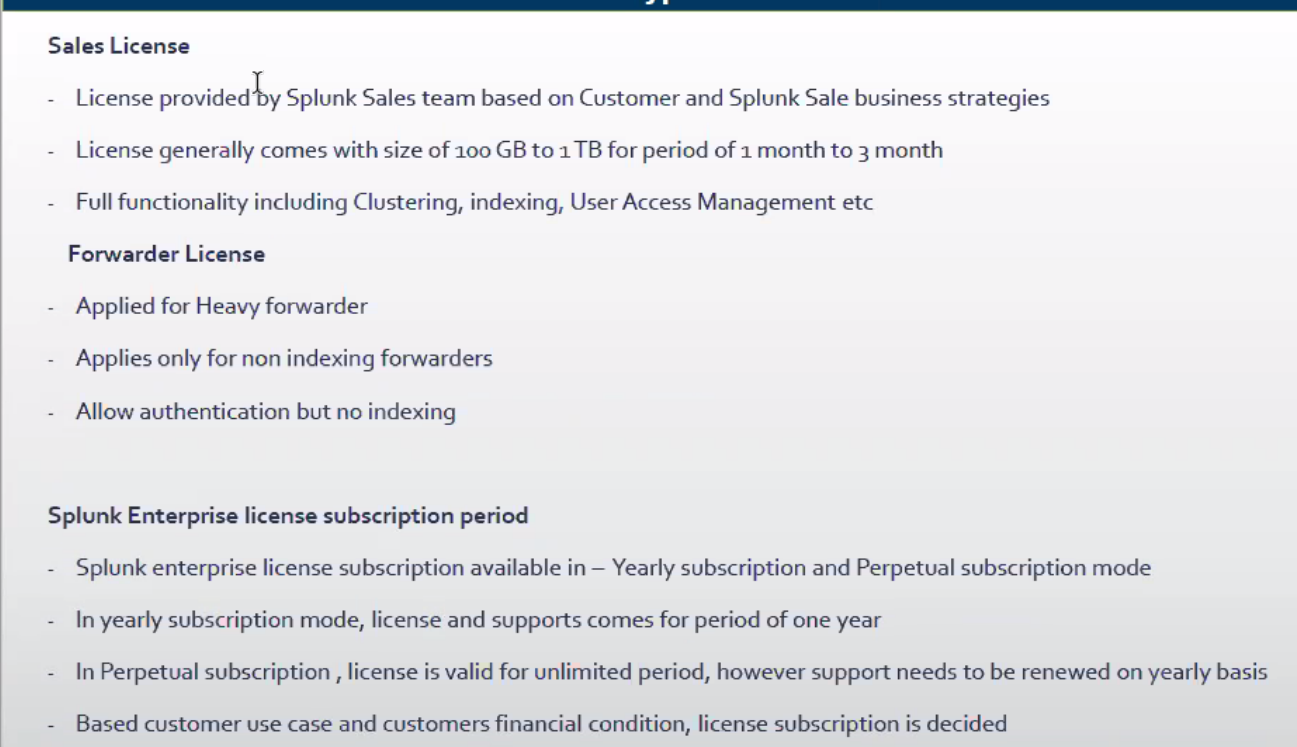
***Here instumentatin, where it will show diags and if u want u can create it***

***License management:***

******

***Internal index -> its getting created as a part of splunk installation. Eg. \_audit,\_internal.  
Within splunk, we have indexer role and whch has small small indexes.   
we have logs under /splunk/splunk/var/log/splunk and by default thes default indexes collects these logs internally and display on console.  
  
to check the same, index=\_internal u can use.  
  
And splunk license will not charge for internal index. But if we generate any external index, its payable.***

******

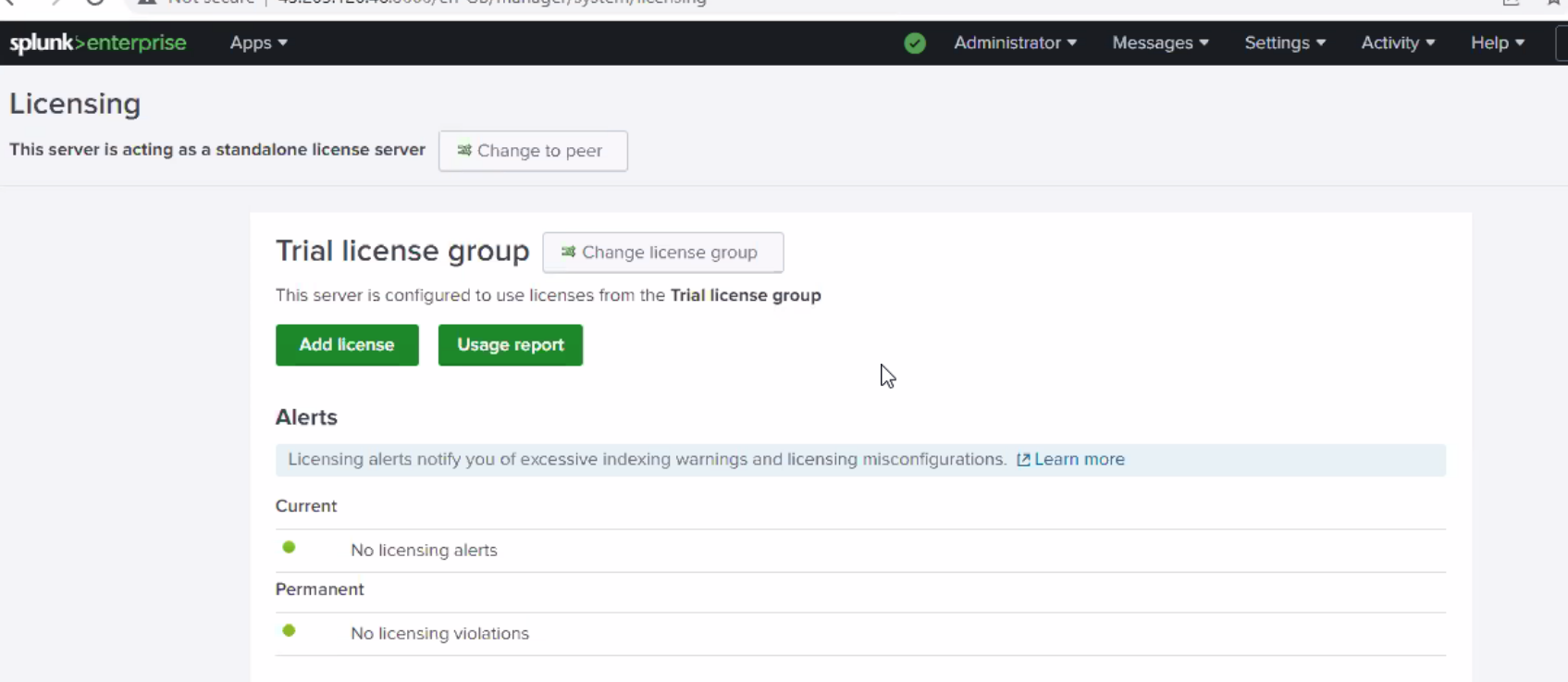
******

***To purchase sales license, we shd contact splunk sales team via splunk.com where we have connect option and call will get scheduled. After connect if theya re ok, they will provide licesne for 1 month or for 3 month.***

***Forwarder license comes by default along with other licenses so splunk will not allow component as its if it has forwarder license. It insists forwarder to have other license eg. Enterprise or sales along with it. It allow authentication but it shd not index. And also it has condition that***

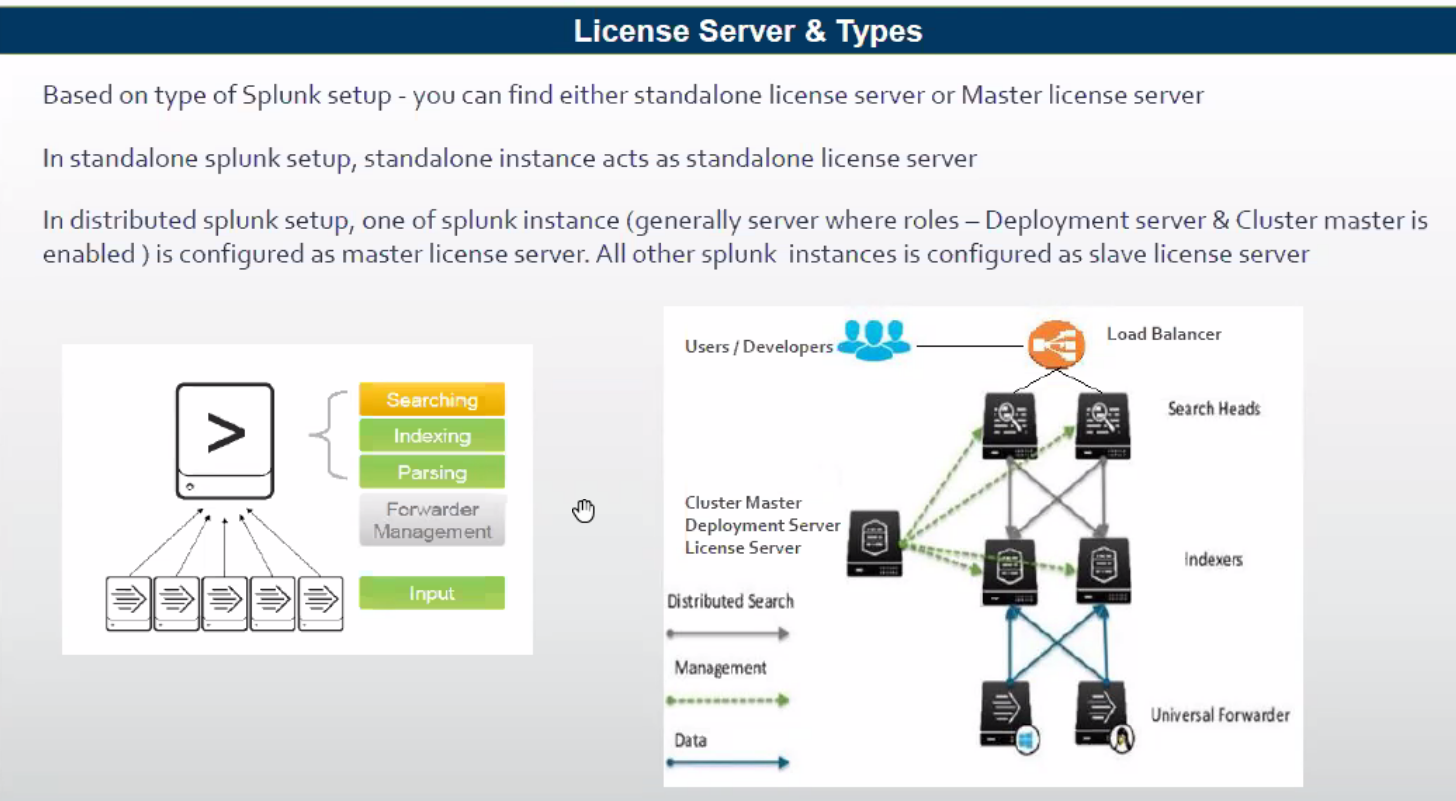
***In perpetual->license valid minimum for 15 years. But support needs to be renewed yearly basis.***

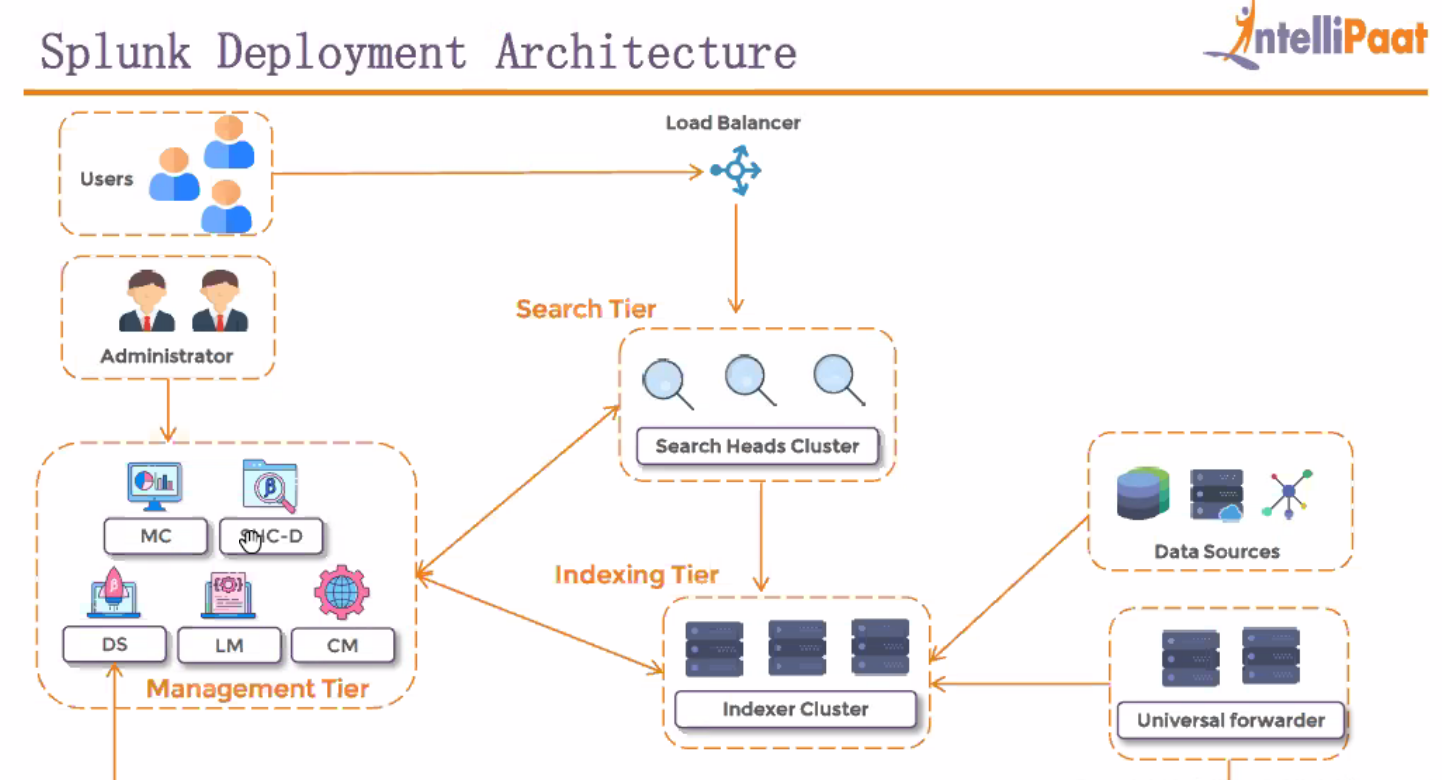
***License server:***

******

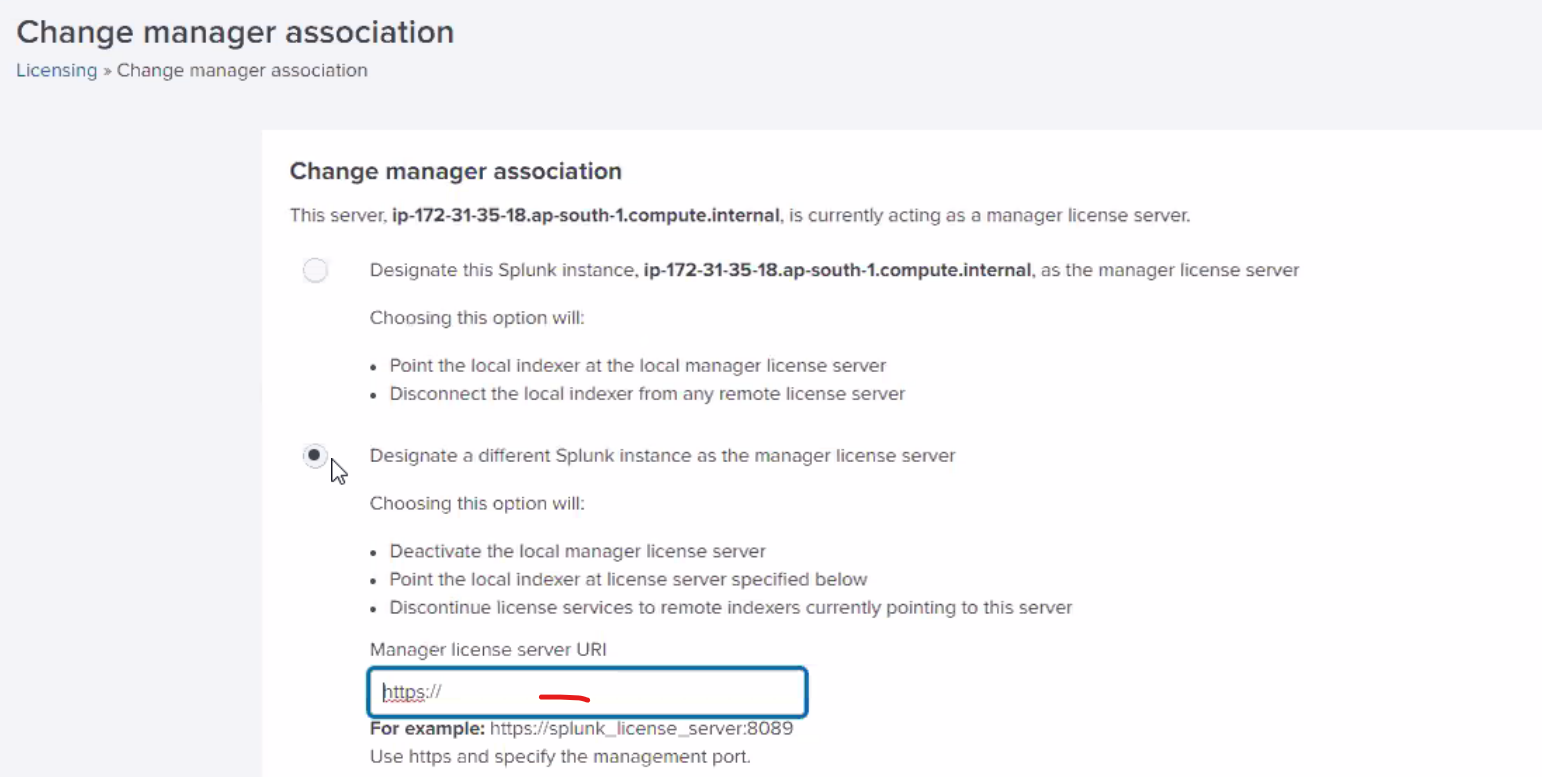
***If we are using standalone instance, then we can use this inbuilt standalone license server. And we can purchase license of 5GB and they will give xml file and we can upload and install it***

***But in distributed environments, we shd not have license on each server in such case, we shd have centralized license server which has master licnese and slave licnese applied to other components as shown below.***

******

******

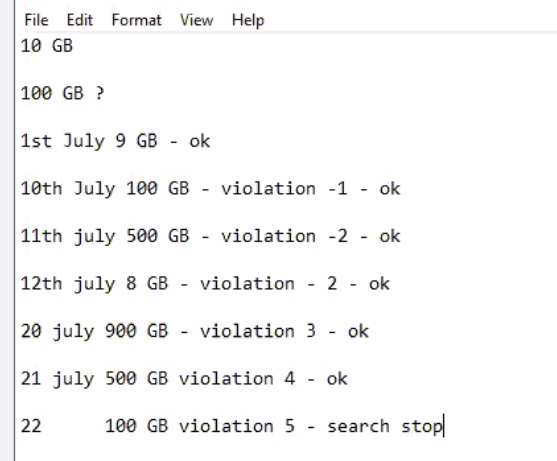
***Here if u see,LM under mgmt.***

******

***Here we have to make this change for all serach heads and indexers and fil IP of centralize license server.***

***What will happen if we dump 100GB data but we have 10GB license. So inthis case, Splunk marks this as Violation as 1 and all will be ok.***

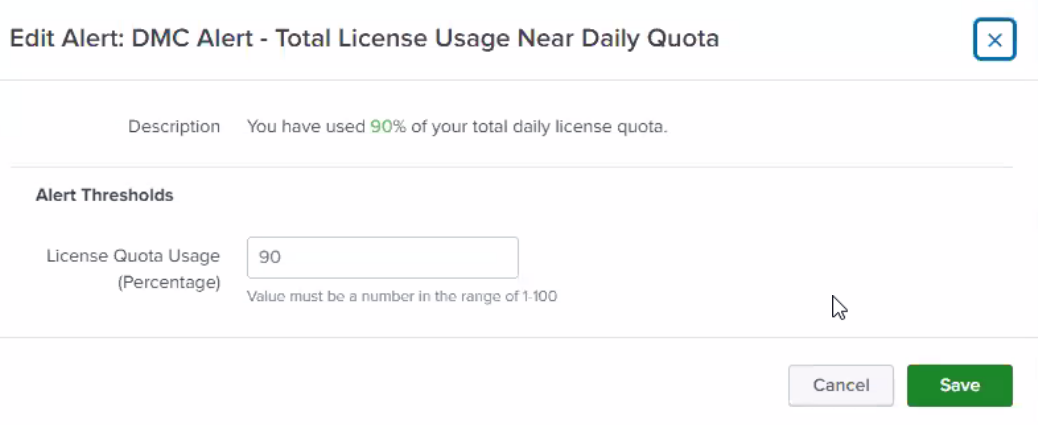
***And again if we dump more data, it will send one more violation and it will be ok and data will be searching.***

******

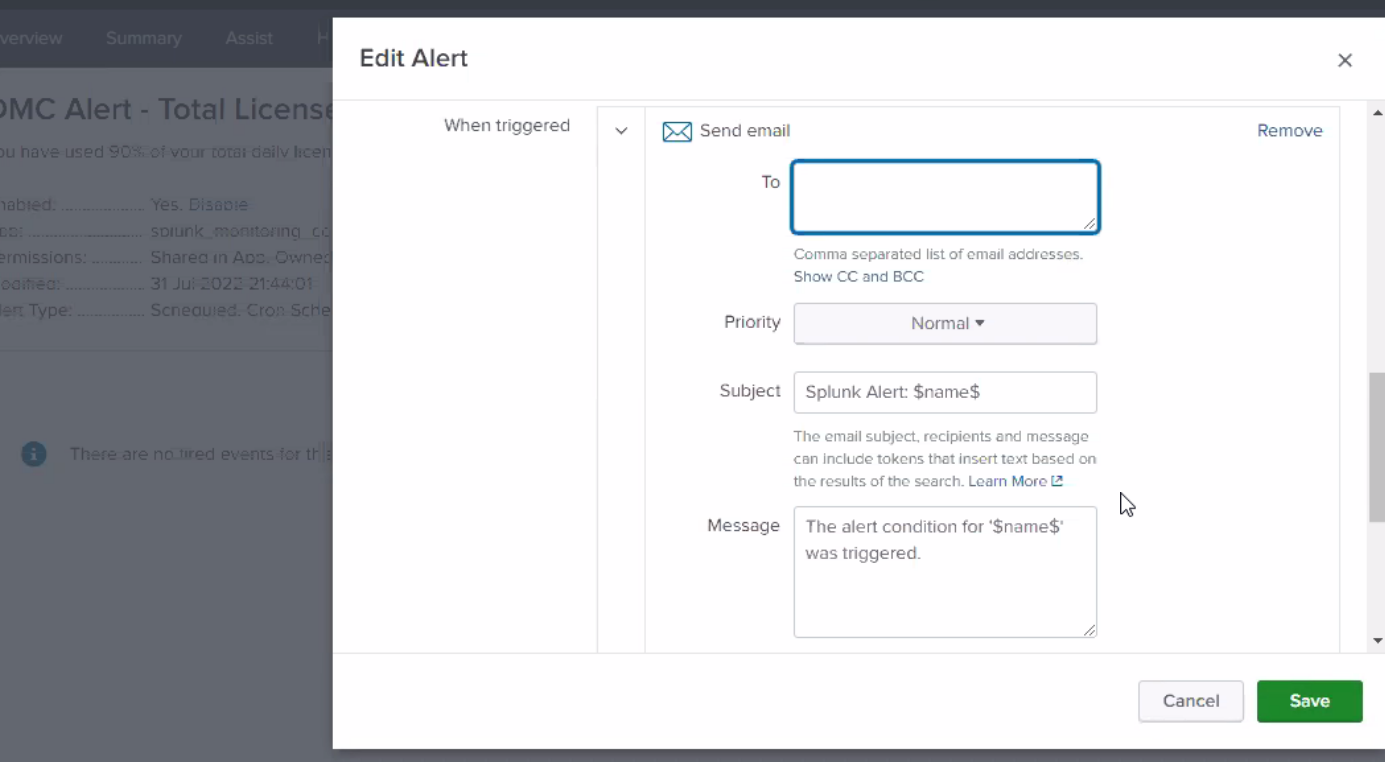
***Here after violation index continues but search will stop. So typically we have 4 viloation to take action***

***If you want u can set or enable alerts under monitoring console as well. And it also send messages to take action.***

***Violation will reset every month. But if we willnot take any action until 5 then it will not reset. But if its with 4, then it will reset every month.***

***Monitorng console>settings->alert setup->  
***

***If u want u can enable it open it and edit it and add action and save it***

******

***In distributed env, where we have centralized license where we shd apply licnese. And in other components search ehad, indexers ->settngs->licensing->change to peer->2nd option and give centraiize server ip address and save it***

***And here in searh head andindexer, for licensing it will consider only one copy. And it wil not consider other copies for license.***

***For eg. If we have 10 GB data comes from 2 forwarders to index cluster it will take only 20GB as part of licenseing and t will consider copy which is replicating.***

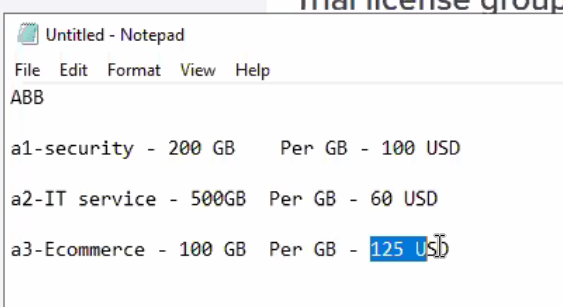
***How to calculate license:***

***First we need to get sales license for eg. 100 GB for 30 days.***

* ***First we need to list out usecases.***

***Take diagrm from session***

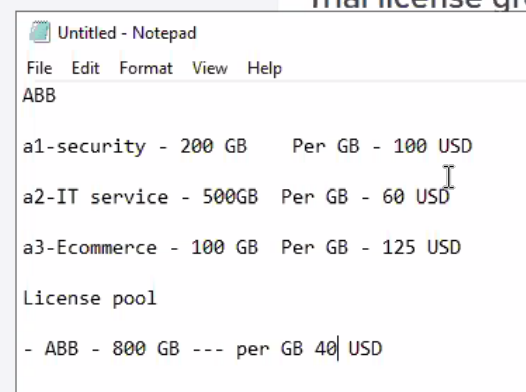
***Licnse pooling***

******

***As per splunk, if we purchse more volume then cost will reduce.***

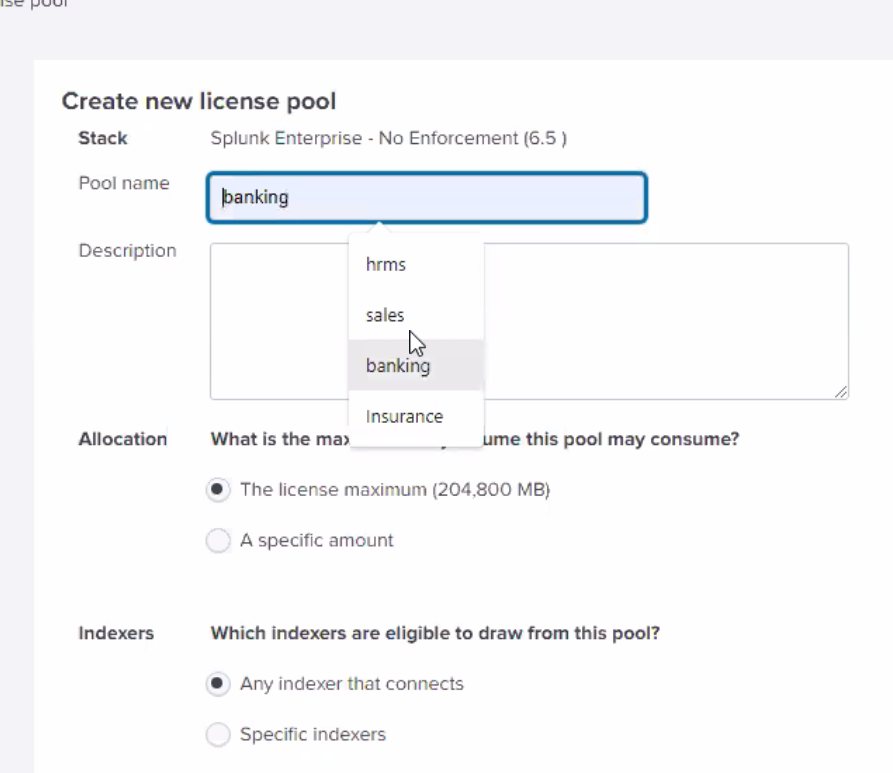
***So in big organization, what they will do is,   
they want to have centralized license server but here in ths case , we have challenge in terms of cost and managint license. For this license pool option came into picture***

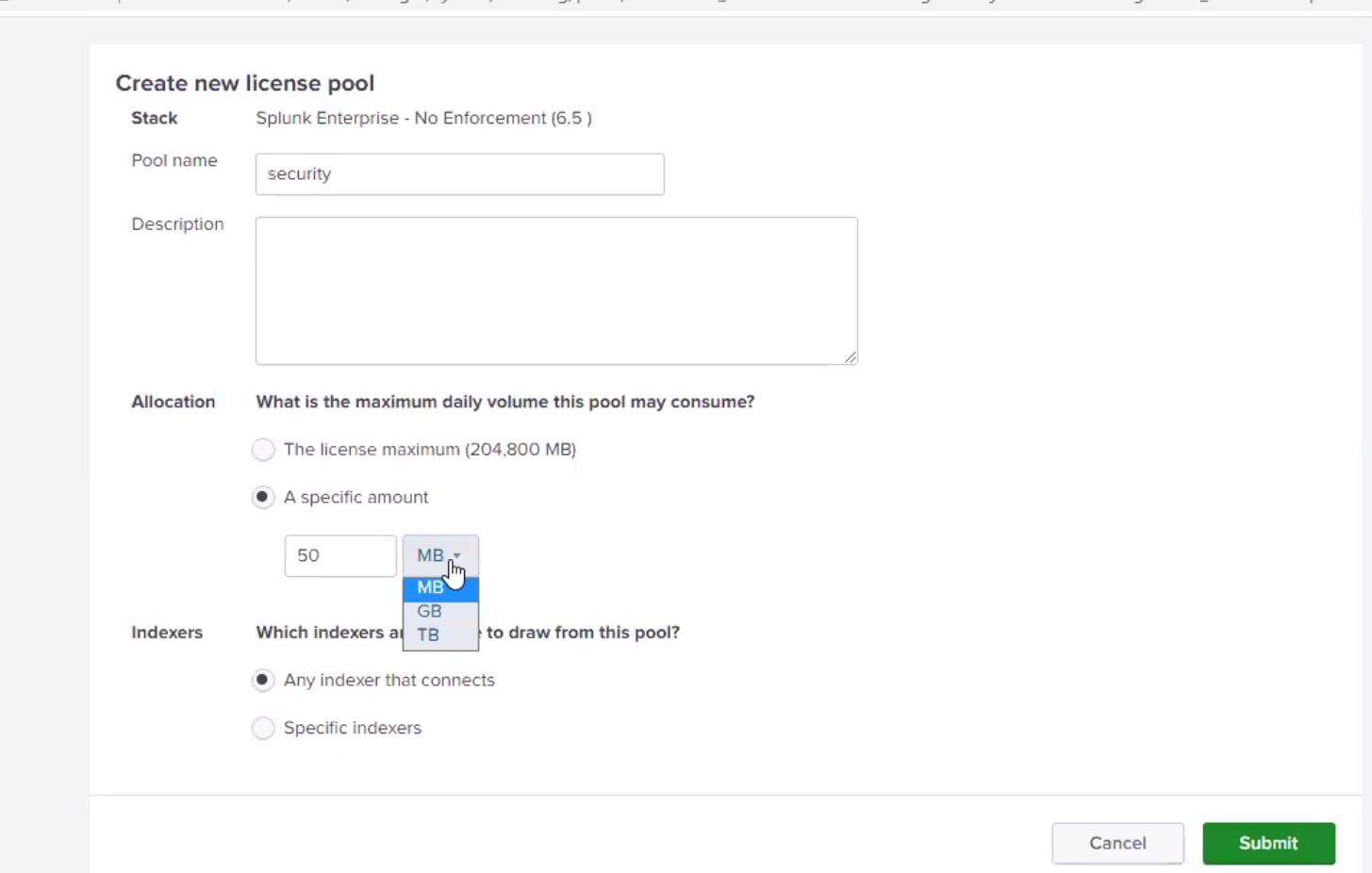
***So, here if u take ABB is company, which takes responsbilty of all 3 and disgtribute to all 3.***

******

***But here, if ABB purchase 800 GB it will cpme in single XML file, then how will u give this to 3 companies.***

***For eg, here u imagine u purchased 800 GB, apply it to centralized license master. And u can create pool to disgtirbute specific amputn as specified to all 3 companies. If we give same file, they might use extra license.***

******

******

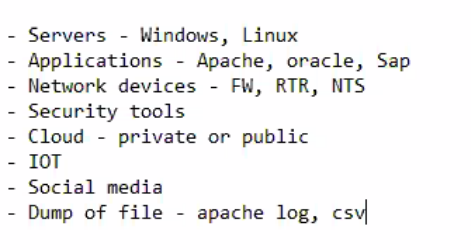
***General note: if we have 100 GB data but after filterationo we are indexing 50 GB then these 50GB replicated among indexers.***

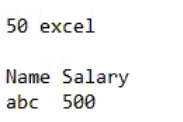
***Search factor => it shd be equal to replication factor. But as per theoritcally we shd kep less. But for better performance we shd keep equal to replication factor.***

***Data onboarding:***

***NTS – Network switches***

***Typical source data***

******

******

***You have 50 excel file and each has 1000 entries with name and salary and u need to push specific salary data.***

***How will you do?***

***In manyul way, we open each exel and find sum of all sal and do for all and finally we will do consolidate sum.***

***But in splum, we can push all excel in one go and in spluk we can cal using sum()***

***Sourcetype->to check log format.***

***You can check pre-trained source type.***

***Note: we have an app using that, we can track how much data comes from which source and to which index.***

***Data forwarding are mostly real time. But results delay it may occur in milliseconds.***

***Cp -pr filename des dire  
-pr->  
ping.eu ->to check current IP address of ours.***

***run script  
check for logs  
upload those logs using monitor option***

***try to track the dta***

***killercoda***