

# Kubernetes Example Install Journal 3.3.0 GA install with 1.0.0 helm charts

Monday, 7 March 2022 4:44 PM

## Update 7 Mar 2022

- Updated labelling to use the Azure CLI API (az commands) as kubectl labeling of nodes does not hold between shutdown/startup.
- Pin Neo4j 3.5.30 since there are breaking changes in 3.5.31

## Objective

To take an existing CluedIn 3.2.5 environment standard setup that was manually installed in Azure Kubernetes Services (AKS) and fresh install 3.3.0 GA with latest cluedin-platform 1.0.0 helm charts reusing the node pools.

Kubectl, Azure Portal and Lens (<https://k8slens.dev/>) will be used.




**Note:** to perform an upgrade requires the careful following of CluedIn provided upgrade instructions (which are WIP as of the time of writing)

## Overview

- Tear down cluster while keeping what we can
- Delete what we know we don't need or will potentially confuse the process or us
- Use helm to deploy 3.3.0 with cluedin-platform 1.0.0 charts
- Create additional orgs as desired

## Journal

The charts have been split into

	cluedin-application	Deploys the applications for the CluedIn platform
	cluedin-infrastructure	Deploys infrastructure for CluedIn Platform
	cluedin-platform	Deploys all parts of the CluedIn platform

this is part of decoupling the application version from the charts and utilizing 3rd party subcharts where possible

deploying cluedin-platform will co-ordinate deploying infrastructure then application on top

Doco starting points:

<https://documentation.cluedin.net/release-notes>

<https://github.com/CluedIn-io/Charts>

<https://github.com/CluedIn-io/Charts/releases/tag/cluedin-platform-1.0.0>

## Prep Work

Shutdown existing cluster

```
kubectl scale deploy -n cluedin --replicas=0 --all
```

```
kubectl scale statefulset -n cluedin --replicas=0 --all
```

```
PS C:\Users\rudi> kubectl get pods -n cluedin
```

NAME	READY	STATUS	RESTARTS	AGE
cluedin-demo-bs-cluedin-91lw1	0/1	Completed	0	39h
cluedin-demo-prometheus-node-exporter-4xdmb	1/1	Running	0	41h
cluedin-demo-prometheus-node-exporter-7kxhc	1/1	Running	0	41h
cluedin-demo-prometheus-node-exporter-br66w	1/1	Running	0	41h
cluedin-demo-prometheus-node-exporter-cqznc	1/1	Running	0	41h
cluedin-demo-prometheus-node-exporter-hbbpf	1/1	Running	0	41h
cluedin-demo-prometheus-node-exporter-jnrr6	1/1	Running	0	41h

Delete volumes since we want a fresh install

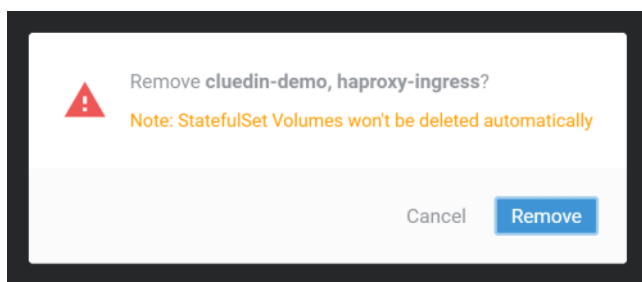
**note:** if we were upgrading we would follow those steps to ensure these are NOT deleted

Persistent Volume Claims				9 items
<input checked="" type="checkbox"/>	Name ▾	Namesp... ▾	Storage class ▾	Size ▾
<input checked="" type="checkbox"/>	cluedin-demo-neo4j-data	cluedin	<a href="#">default</a>	8Gi
<input checked="" type="checkbox"/>	cluedin-demo-openrefine-data	cluedin	<a href="#">default</a>	512Mi
<input checked="" type="checkbox"/>	cluedin-demo-redis-data	cluedin	<a href="#">default</a>	512Mi
<input checked="" type="checkbox"/>	cluedin-demo-sqlserver-backup	cluedin	<a href="#">default</a>	8Gi
<input checked="" type="checkbox"/>	cluedin-demo-sqlserver-data	cluedin	<a href="#">default</a>	8Gi
<input checked="" type="checkbox"/>	cluedin-demo-sqlserver-master	cluedin	<a href="#">default</a>	8Gi
<input checked="" type="checkbox"/>	cluedin-demo-sqlserver-transact	cluedin	<a href="#">default</a>	8Gi
<input checked="" type="checkbox"/>	data-cluedin-demo-rabbitmq-0	cluedin	<a href="#">default</a>	4Gi
<input checked="" type="checkbox"/>	elasticsearch-master-elasticsearch-mast...	cluedin	<a href="#">default</a>	4Gi

delete releases

we could keep the haproxy-ingress but I want the new one nevertheless as part of checking the haproxy-ingress integrated install process

Charts Releases							
Releases				2 Items		Namespace: cluedin ▾	Search Releases...
<input checked="" type="checkbox"/>	Name ▾	Namespace ▾	Chart ▾	Revision ▾	Version	App Version	Status ▾
<input checked="" type="checkbox"/>	cluedin-demo	cluedin	cluedin	7	3.2.5	3.2.5	Deployed
<input checked="" type="checkbox"/>	haproxy-ingress	cluedin	haproxy-ingress	1	0.13.6	v0.13.6	Deployed



update helm charts from the repos

```
PS C:\rudi.harris\k8s\cluedin\apac-demo> helm repo list
NAME      URL
haproxy-ingress https://haproxy-ingress.github.io/charts
cluedin    https://cluedin-io.github.io/Charts/

PS C:\rudi.harris\k8s\cluedin\apac-demo> helm repo update
Hang tight while we grab the latest from your chart repositories...
...Successfully got an update from the "cluedin" chart repository
...Successfully got an update from the "haproxy-ingress" chart repository
Update Complete. ✨Happy Helming!✨
```

Charts Releases				
Search Helm Charts...				
Name ▾	Description	Version	App Version	Repository ▾
cluedin	Deploys the CluedIn Application	3.2.5		cluedin
cluedin-application	Deploys the applications for the CluedIn platform	1.0.0		cluedin
cluedin-infrastructure	Deploys infrastructure for CluedIn Platform	1.0.0		cluedin
cluedin-platform	Deploys all parts of the CluedIn platform	1.0.0		cluedin
haproxy-ingress	Ingress controller for HAProxy loadbalancer	0.14.0-alpha.1		haproxy-ingress

new charts, including cluedin-platform, are here

**note:** I probably don't need haproxy-ingress anymore, but no harm in leaving it

follow the repo location

```
PS C:\rudi.harris\k8s\cluedin\apac-demo> helm repo list
NAME      URL
haproxy-ingress https://haproxy-ingress.github.io/charts
cluedin    https://cluedin-io.github.io/Charts/
```



## Charts

This repository contains a collection of CluedIn Helm charts that can be used to install the CluedIn application into a Kubernetes cluster.

Information on Helm and how to use it can be found here: [https://helm.sh/docs/intro/using\\_helm/](https://helm.sh/docs/intro/using_helm/)

Current charts:

- CluedIn - The whole cluedin application + all dependant services.

## Using charts from the CluedIn chart repository

Before you can install the CluedIn chart you need to add the chart repository to your local collection so that Helm knows where to retrieve the chart from.

Adding the `cluedin` chart repository to your local repository collection.

```
$ helm repo add cluedin https://cluedin-io.github.io/Charts/
```

List all repositories currently installed.

```
$ helm repo list
```

## Examples of Helm usage

Search repositories for charts called `cluedin` in the chart repository and display the versions.

let's go to the github repo

<https://github.com/CluedIn-io/Charts>

click the release info on the right for latest release

<https://github.com/CluedIn-io/Charts/releases/tag/cluedin-platform-1.0.0>

## Release cluedin-platform-1.0.0

Latest

dervalp released this 9 hours ago cluedin-platfo... 9c870e4

Initial Release!

### Features

- (cluedin-application) EngineRoom feature is now enabled by default
- (cluedin-infrastructure) Kube-Prometheus-Stack added as sub-chart (for engine-room)
- (cluedin-application) Custom Grafana charts can be added to the `dashboards` section (Remember to surround JSON with `{{\` `}}`)
- (cluedin-application) Metrics are > **enabled** < by default .. to *disable* set:

```
monitoring:
  enabled: false
```

- (cluedin-infrastructure) Metrics are > **enabled** < by default .. to *disable* (wont install services) set:

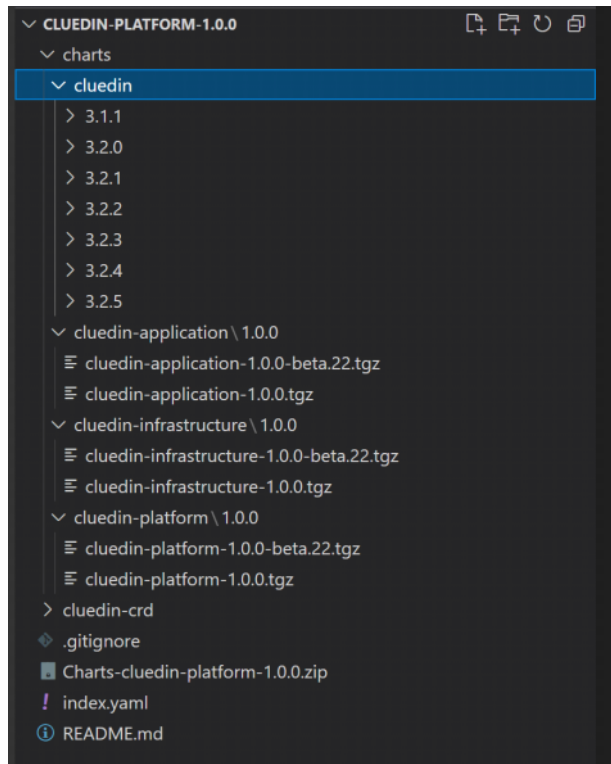
```
monitoring:
  enabled: false
```

```
rabbitmq:
```

if we want to inspect the charts we could download zip file from assets at the bottom and unzip to inspect

<https://github.com/CluedIn-io/Charts/archive/refs/tags/cluedin-platform-1.0.0.zip>

you would then untar the platform one potentially



## Build New Values Yaml

Let's use our existing 3.3.0 values yaml as a reference point to reimplement from the 3.2.5 charts to the cluedin-platform-1.0.0 charts

So, we have [values.3.2.5charts\\_3.3.0app.redacted.example.yaml](#) as an example chart for this exact environment as sized in [202202\\_sizing.xlsx](#)

method, let's create a new one based on internal CluedIn training and tailor it to match the refinements from the above reference (and thus provide the public with a good new example) - if we have any issues with the deploy we can use the command `helm template` to output what helm would have done in the case where `helm upgrade` doesn't work

pool labeling

the recommendation is to label all ~~nodes~~ node pools with cluedin.com based pooltypes and then reference these in the yaml

### Update!

Use az API - either on pool creation or you can use the following to update the existing pools

```
az aks nodepool update --cluster-name APAC-demo-k8s --name corepool --resource-group APAC-demo-k8s-rg --labels kubernetes.cluedin.com/pooltype=core
az aks nodepool update --cluster-name APAC-demo-k8s --name datapool --resource-group APAC-demo-k8s-rg --labels kubernetes.cluedin.com/pooltype=data
az aks nodepool update --cluster-name APAC-demo-k8s --name generalpool --resource-group APAC-demo-k8s-rg --labels kubernetes.cluedin.com/pooltype=general
az aks nodepool update --cluster-name APAC-demo-k8s --name processpool --resource-group APAC-demo-k8s-rg --labels kubernetes.cluedin.com/pooltype=process
```

*To be explicit about assignments you can also add nodeSelectors—example can be found in values-deploy.yaml inside the root of each chart. You will need to label each node accordingly:*

- `kubectl label node <node-name> -n cluedin kubernetes.cluedin.com/pooltype=data`*
- `kubectl label node <node-name> -n cluedin kubernetes.cluedin.com/pooltype=processing`*

so label them

```
PS C:\rudi.harris\k8s\cluedin\apac-demo> kubectl get nodes
NAME                                STATUS    ROLES    AGE   VERSION
aks-corepool-17590510-vmss00000b    Ready    agent    2d    v1.20.9
aks-datapool-17590510-vmss00000m    Ready    agent    2d    v1.20.9
aks-datapool-17590510-vmss00000n    Ready    agent    2d    v1.20.9
aks-generalpool-17590510-vmss00000m    Ready    agent    2d    v1.20.9
aks-generalpool-17590510-vmss00000n    Ready    agent    2d    v1.20.9
aks-processpool-17590510-vmss00000b    Ready    agent    2d    v1.20.9
```

e.g.



values yml copied and passed at the end for reference

delete old secrets with cluedin-demo- prefix

try to install using our new values yml

the output of helm can be inspected to ensure what we intended and what was parsed match

Release "cluedin-demo" does not exist. Installing it now.

```
NAME: cluedin-demo
LAST DEPLOYED: Wed Feb 23 00:29:37 2022
NAMESPACE: cluedin
STATUS: deployed
REVISION: 1
NOTES:
```



```

===== \_/_/ =====
====[ Mode : [INSTALL]
====[ Chart : cluedin-platform [v1.0.0]
====[ Namespace : cluedin
====[ Domain : *.k1.cluedin.me
====[ TLS : [ENABLED] Cert Secret: ingress-cert-secret
=====
====[ Global Image Tag : 3.3.0
====[ Global Registry : docker.io
====[ Global Pull Policy : IfNotPresent
====[ Global ImagePullSecret : .
====[ Global ServiceAccount : cluedin-serviceaccount
====[ Global Environment : Production
====[ Global Strategy : RollingUpdate
====[ Global UploadLimit : 1gb
=====
====[ CluedIn Server : docker.io/cluedin/cluedin-server:3.3.0
====[ CluedIn Nuget Install : docker.io/cluedin/nuget-installer:3.3.0
====[ CluedIn Controller : docker.io/cluedin/controller:3.3.0
====[ Annotation : docker.io/cluedin/cluedin-micro-annotation:3.3.0
====[ Prepare : docker.io/cluedin/cluedin-micro-clean:3.3.0
====[ DataSource : docker.io/cluedin/cluedin-micro-datasource:3.3.0
====[ Submitter : docker.io/cluedin/cluedin-micro-submitter-node:3.3.0
====[ GQL : docker.io/cluedin/cluedin-ui-gql:3.3.0
====[ UI : docker.io/cluedin/ui:3.3.0
====[ WebAPI : docker.io/cluedin/app:3.3.0
====[ InitWait : docker.io/groundnuty/k8s-wait-for:v1.3
====[ InitCleanUp : docker.io/bitnami/kubect1:latest
====[ InitSQL : docker.io/cluedin/sqlserver-init:3.3.0
====[ InitNeo4J : docker.io/cluedin/neo4j-init:3.3.0
=====
====[ SQL Server : mcr.microsoft.com/mssql/server:2017-latest
====[ OpenRefine : docker.io/cluedin/openrefine:3.3.0
====[ Elasticsearch : docker.elastic.co/elasticsearch/elasticsearch:7.8.0
====[ Neo4J : neo4j:3.5
====[ RabbitMQ : docker.io/bitnami/rabbitmq:3.9.11-debian-10-r28
====[ Redis : docker.io/bitnami/redis:6.2.6-debian-10-r81
====[ AlertManager : quay.io/prometheus/alertmanager:v0.23.0
====[ Prometheus : quay.io/prometheus/prometheus:v2.31.1
====[ Grafana : grafana/grafana:8.2.3
====[ Prometheus Operator : quay.io/prometheus-operator/prometheus-operator:v0.52.0
=====
====[ Cluedin : [ENABLED]
====[ Roles : Main ☒ - Processing ☒ - Crawling ☐
====[ Connection (mssql) : AuthenticationStore (DataStore.Db.Authentication) - User:
clientUser
====[ Connection (mssql) : AuditLog (DataStore.Db.AuditLog) - User: clientUser
====[ Connection (mssql) : BlobStorage (DataStore.Db.BlobStorage) - User: clientUser
====[ Connection (mssql) : ConfigurationStore (DataStore.Db.Configuration) - User:
clientUser
====[ Connection (mssql) : CluedInEntities (DataStore.Db.OpenCommunication) - User:
clientUser
====[ Connection (mssql) : TokenStore (DataStore.Db.TokenStore) - User: clientUser
====[ Connection (mssql) : Training (DataStore.Db.Training) - User: clientUser
====[ Connection (mssql) : ExternalSearch (DataStore.Db.ExternalSearch) - User: clientUser
====[ Connection (mssql) : ML-Logging (DataStore.Db.ML-Logging) - User: clientUser
====[ Connection (mssql) : Metrics (DataStore.Db.Metrics) - User: clientUser
====[ Connection (mssql) : WebAppStore (DataStore.Db.WebApp) - User: clientUser
====[ Connection (rabbitmq) : MessageBus () - User: cluedin
====[ Connection (rabbitmq) : MessageBusManagement () - User: cluedin
====[ Connection (rabbitmq) : SignalRScaleOut () - User: cluedin
====[ Connection (neo4j) : GRAPHSTORE_READ () - User: neo4j
====[ Connection (neo4j) : GRAPHSTORE_WRITE () - User: neo4j
====[ Connection (redis) : CACHESTORE () - User: redis
====[ Connection (redis) : ETAGSTORE () - User: redis
====[ Connection (redis) : JOBSTORE () - User: redis
====[ Connection (redis) : DATAPROTECTIONPERSISTENCE () - User: redis
====[ Connection (elasticsearch) : SEARCHSTORE () - User: elastic
=====
====[ SQL SERVER : [ENABLED]
====[ Host Name : cluedin-sqlserver
====[ Host Port : 1433
====[ Timeout : 15
====[ Encrypted ? : false
====[ Extra Parameters : MultipleActiveResultSets=True;Max Pool Size=200;Pooling=True;
====[ User : clientUser (cluedin-sqlserver-clientuser-secret)
====[ User : sa (cluedin-sqlserver-secret)
=====
====[ OpenRefine : [ENABLED]
====[ Host Name : cluedin-openrefine
====[ Host Port : 3333
=====
====[ Redis : [ENABLED]
====[ Host Name : cluedin-redis-master
====[ Host Port : 6379
====[ Encrypted ? : false
====[ User Secret : cluedin-redis (redis-password)
=====
====[ Neo4J : [ENABLED]
====[ Host Name : cluedin-neo4j
====[ Host HTTP Port : 7474
====[ Host BOLT Port : 7687
====[ User : neo4j
====[ User Secret : cluedin-neo4j-secrets (neo4j-password)
=====
====[ Elasticsearch : [ENABLED]
====[ Host Name : cluedin-elasticsearch
====[ Host Port : 9200
====[ User Secret : elasticsearch-credentials (password)
=====
====[ RabbitMQ : [ENABLED]
====[ Host Name : cluedin-rabbitmq
====[ Host Port : 5672
====[ User : cluedin
====[ User Secret : cluedin-rabbitmq (rabbitmq-password)
=====
====[ Bootstrap Org ? : [YES => cluedin]
====[ SSO : [DISABLED]
=====

```

watching in lens - it's coming up

Pods		31 items	Namespace: cluedin		Search Pods...
<input type="checkbox"/> Name	Containers	Restarts	Node	Age	Status
<input type="checkbox"/> cluedin-grafana-7b45668d44-t2dcd	■■■	0	aks-datapool-17590510-vmss00	10s	Pending
<input type="checkbox"/> cluedin-server-7b67c75c8c-jzdv8	■■■■■▲	0	aks-processpool-17590510-vmss00	10s	Pending
<input type="checkbox"/> cluedin-server-processing-74996cdc59-x5zzx	■■■■■■■	0	aks-processpool-17590510-vmss00	10s	Pending
<input type="checkbox"/> cluedin-cert-manager-6555bf977-hkwfb	■	0	aks-datapool-17590510-vmss00	11s	Pending
<input type="checkbox"/> cluedin-haproxy-ingress-ddfc9cf9-k2l67	■	0	aks-datapool-17590510-vmss00	11s	Pending
<input type="checkbox"/> cluedin-annotation-85576b49bb-qlc9g	■■■■■	0	aks-datapool-17590510-vmss00	12s	Pending
<input type="checkbox"/> cluedin-cert-manager-cainjector-556ddb47d6-grrzd	■	0	aks-datapool-17590510-vmss00	12s	Running
<input type="checkbox"/> cluedin-cert-manager-webhook-59ddc48bf-frlxn	■	0	aks-datapool-17590510-vmss00	12s	Pending
<input type="checkbox"/> cluedin-controller-5c95967ff4-zcmkb	■■	0	aks-processpool-17590510-vmss00	12s	Pending
<input type="checkbox"/> cluedin-datasource-7644786b48-n7c2r	■■■	0	aks-processpool-17590510-vmss00	12s	Pending
<input type="checkbox"/> cluedin-elasticsearch-0	■■	0	aks-datapool-17590510-vmss00	12s	Pending
<input type="checkbox"/> cluedin-gql-5fc4f5d858-lmj7x	■■■■■■■	0	aks-generalpool-17590510-vmss00	12s	Pending
<input type="checkbox"/> cluedin-kube-state-metrics-659ffc4664-mctxn	■	0	aks-datapool-17590510-vmss00	12s	Pending
<input type="checkbox"/> cluedin-neo4j-core-0	■■■	0	aks-datapool-17590510-vmss00	12s	Pending
<input type="checkbox"/> cluedin-openrefine-55f89987ff-f2ln4	■	0	aks-processpool-17590510-vmss00	12s	Pending

new volumes created

Persistent Volume Claims			9 items		Namespace: cluedin		Search Persistent
<input type="checkbox"/>	Name	Namespace	Storage class	Size	Pods	Age	Status
<input type="checkbox"/>	cluedin-elasticsearch-cluedin-elasticsearch-0	cluedin	default	4Gi	cluedin-elasticsearch-0	60s	Bound
<input type="checkbox"/>	cluedin-openrefine-data	cluedin	default	512Mi	cluedin-openrefine-55f89987ff-f2ln4	61s	Bound
<input type="checkbox"/>	cluedin-sqlserver-backup	cluedin	default	8Gi	cluedin-sqlserver-64ff77dd95-bmp9k	61s	Bound
<input type="checkbox"/>	cluedin-sqlserver-data	cluedin	default	8Gi	cluedin-sqlserver-64ff77dd95-bmp9k	61s	Bound
<input type="checkbox"/>	cluedin-sqlserver-master	cluedin	default	8Gi	cluedin-sqlserver-64ff77dd95-bmp9k	61s	Bound
<input type="checkbox"/>	cluedin-sqlserver-translog	cluedin	default	8Gi	cluedin-sqlserver-64ff77dd95-bmp9k	61s	Bound
<input type="checkbox"/>	data-cluedin-rabbitmq-0	cluedin	default	4Gi	cluedin-rabbitmq-0	60s	Bound
<input type="checkbox"/>	datadir-cluedin-neo4j-core-0	cluedin	default	8Gi	cluedin-neo4j-core-0	60s	Bound
<input type="checkbox"/>	redis-data-cluedin-redis-master-0	cluedin	default	512Mi	cluedin-redis-master-0	60s	Bound

ingresses created

		Services	Endpoints	Ingresses	Network Policies	Port Forwarding
Ingresses		3 items		Namespace: cluedin		Search Ingresses...
<input type="checkbox"/> Name	Namespace	LoadBalancers	Rules	Age		
<input type="checkbox"/> cluedin-clean	cluedin		https://clean.k1.cluedin.me/ → cluedin-openrefine:3333	112s		
<input type="checkbox"/> cluedin-ui	cluedin		https://*k1.cluedin.me/ → cluedin-ui:80 https://*k1.cluedin.me/graphql → cluedin-gql:8888 https://*k1.cluedin.me/upload → cluedin-gql:8888 https://*k1.cluedin.me/logout → cluedin-gql:8888 https://*k1.cluedin.me/app → cluedin-webapi:80 https://*k1.cluedin.me/auth → cluedin-server:9001 https://*k1.cluedin.me/api → cluedin-server:9000 https://*k1.cluedin.me/public → cluedin-server:9007 https://*k1.cluedin.me/webhooks → cluedin-server:9006	112s		
<input type="checkbox"/> cluedin-ui-grafana	cluedin		https://*k1.cluedin.me/charts → cluedin-grafana:service	112s		

check public ip

```
PS C:\rudi.harris\k8s\cluedin\apac-demo> kubectl get services -n cluedin | FindStr haproxy
cluedin-haproxy-ingress      LoadBalancer    10.0.90.9      20.193.9.158
80:30025/TCP,443:30007/TCP    3m2s
cluedin-haproxy-ingress-metrics ClusterIP        10.0.15.39    <none>
9101/TCP,10254/TCP           3m2s
cluedin-haproxy-ingress-stats ClusterIP        10.0.146.28   <none>
3m2s                          1936/TCP
```

update dns k1 and \*.k1 A records for cluedin.me

all pods up



Pods		33 items		Namespace: cluedin		Search Pods...
<input type="checkbox"/>	Name	Containers	Restarts	Node	Age	Status
<input type="checkbox"/>	prometheus-cluedin-prometheus-0	<div><div></div><div></div><div></div></div>	0	<a href="#">aks-datapool-17590510-vmss00</a>	6m46s	Running
<input type="checkbox"/>	alertmanager-cluedin-alertmanager-0	<div><div></div><div></div></div>	0	<a href="#">aks-datapool-17590510-vmss00</a>	6m47s	Running
<input type="checkbox"/>	cluedin-grafana-7b45668d44-t2dcd	<div><div></div><div></div><div></div></div>	0	<a href="#">aks-datapool-17590510-vmss00</a>	7m6s	Running
<input type="checkbox"/>	cluedin-server-7b67c75c8c-jzdvr	<div><div></div><div></div><div></div><div></div><div></div></div>	0	<a href="#">aks-processpool-17590510-vmss00</a>	7m6s	Running
<input type="checkbox"/>	cluedin-server-processing-74996cdc59-x5zzx	<div><div></div><div></div><div></div><div></div><div></div></div>	0	<a href="#">aks-processpool-17590510-vmss00</a>	7m6s	Running
<input type="checkbox"/>	cluedin-cert-manager-6555bff977-hkwfb	<div><div></div></div>	0	<a href="#">aks-datapool-17590510-vmss00</a>	7m7s	Running
<input type="checkbox"/>	cluedin-haproxy-ingress-dddfc9cf9-k2l67	<div><div></div></div>	0	<a href="#">aks-datapool-17590510-vmss00</a>	7m7s	Running
<input type="checkbox"/>	cluedin-annotation-85576b49bb-qlc9n	<div><div></div><div></div><div></div></div>	0	<a href="#">aks-datapool-17590510-vmss00</a>	7m8s	Running

## Manually Create Additional Orgs

```
kubectrl run -i --tty --rm debug --image=alpine --restart=Never -- sh
```

```
apk --no-cache add curl jq
```

```
export org=cluedin
```

```
export password=<redacted>
```

create_org.sh	<pre>#!/bin/sh  URL=https://app.&lt;your cluster FQDN here&gt;/auth/api/account/new echo \$URL  organizationName=\$org emailDomain=\$org.com username=admin@\$emailDomain  # it seems we do not require the bearer token # removed --header "Authorization: Bearer \$TOKEN" \  curl -v -k --location -g --request POST \$URL \   --header 'Accept: application/json' \   --header 'Content-Type: application/x-www-form-urlencoded' \   --data-urlencode email=\$username \   --data-urlencode username=\$username \   --data-urlencode password=\$password \   --data-urlencode confirmPassword=\$password \   --data-urlencode grant_type=password \   --data-urlencode applicationSubDomain=\$org \   --data-urlencode organizationName=\$organizationName \   --data-urlencode emailDomain=\$emailDomain \   --data-urlencode allowEmailDomainSignup=true \     jq  URLHost=`echo \$URL   sed -e 's ^[/]*[//] ' -e 's /.*\$  '` LoginHost=`echo \$URLHost   sed -e 's app \${org} '` URLSchema=`echo \$URL   cut -d: -f 1` SignInURL=\$URLSchema://\${LoginHost}/signin  echo echo \$SignInURL echo \$username echo \$password</pre>
---------------	--

```
{
  "Active": true,
  "ApplicationSubDomain": "cluedin",
  "CustomerId": null,
  "ExternalAuthenticationId": null,
  "EmailDomainName": "cluedin.com",
  "Id": "719047d7-3488-413b-a780-8b08de94057f",
  "IsEmailDomainSignupActivated": true,
  "LastLoginTime": "2022-02-22T14:51:53.2361614Z",
  "Name": "cluedin",
  "Plan": "6749e7a5-da15-47e2-918b-6a385e288865",
  "PlanEndDate": "2022-03-24T14:51:53.2361615Z",
  "PlanIsActive": true,
  "PlanStartDate": "2022-02-22T14:51:53.236162Z",
  "RefreshTokenLifeTime": 14400,
  "SubscriptionId": null
}
```

## Successful Result - 3.3.0 GA



# Sign in

Welcome back, you are signing into the **cluedin** account.


Email

Password

SIGN IN

[I forgot my password](#)

← → ↻ cluedin.k1.cluedin.me/home/




CluedIn

- Integrations
- Governance
- Preparation
- Engine Room
- Management
- Consume
- Administration


All | Search

All 🔍


### Integrations




Import Data From File




Import Data From Database




Import Data From Endpoint




Import Data From Connector




Build Integration



View All Data sources




0  
Integrations




0  
Connectors

### Governance




2  
Number of entities




2  
Number of used entity types

### Management



12



277

Let's ingest some mock data

<https://mockaroo.com/>

a 1000 entries will do

### Name\*

User friendly name to help you find

MOCK\_DATA.csv

### Group\* ?

Data is generally grouped by department domains (e.g. Financial)

rudi-test-mockdata



## MOCK\_DATA.csv

Unfinished Mapping

Created by admin@cluedin.com a few seconds ago

Preview

Map

Prepare

Validate

Process

id	first_name	last_name	email
1	Micky	Reeve	mreeve0
2	Joelie	Conachy	jconachy
3	Chryste	Kirdsch	ckirdsch;
4	Rosemarie	Laphorn	rlaphorn
5	Eli	Sherrington	esherring
6	Margarette	Ellerbeck	mellerbe
7	Jessamyn	O'Moylane	jomoylar

### Create mapping

✓ Choose mapping type

2

This step will create auto mapping which will decide the type to create when the data is processed. In CluedIn we call this:

### Choose Entity Type

Select from the list or enter name to create new entity type

Person

[View more details on Person](#)

### Choose vocabulary

Select from the list or enter name to create new vocabulary

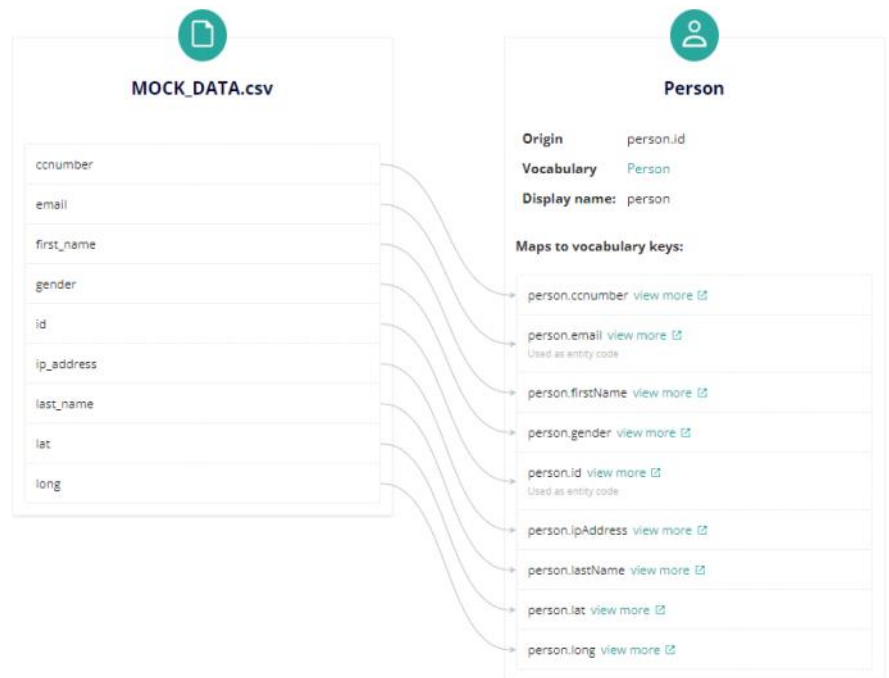
Person


[View more details on Person](#)



## MOCK\_DATA.CSV

Not Processed Created by  admin@cluedin.com a minute ago

view **Map** Prepare Validate Process



< Back  Integrations / ... / rudi-test-mockdata - MOCK\_DATA.csv / MOCK\_DATA.csv

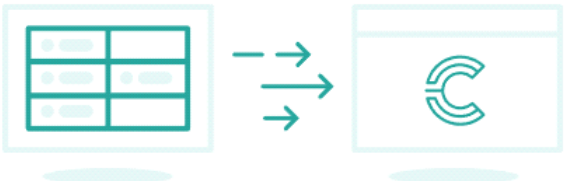
**MOCK\_DATA.csv**  
Not Processed Created by  admin@cluedin.com

Preview Map Prepare Validate

**Process**

Source

**Are you sure you want to process?**



**You are about to push Clues into CluedIn fabric.**

This means the data will now be processed and could be linked, merged and used to enhance your data fabric.

If you are not satisfied with the data yet, please create a 'Prepare' project to increase the quality.

**1000 clues will be created.**

**CONFIRM**

Stats	Created
Success 1000 	a few seconds ago

create rabbitmq  
the password is stored in the cluster secrets



Username:  \*

Password:  \*

secrets

```
rabbitmq-password
<redacted>
```

set graphs to large and last ten minutes so we can see a summary of the activity



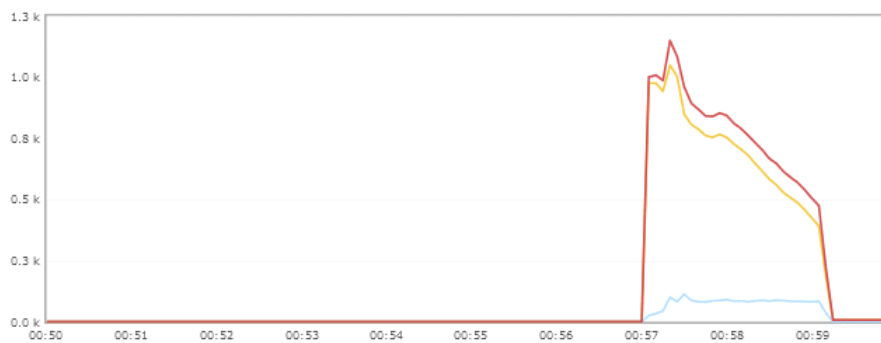
RabbitMQ 3.9.11 Erlang 24.2

Overview Connections Channels Exchanges Queues Admin

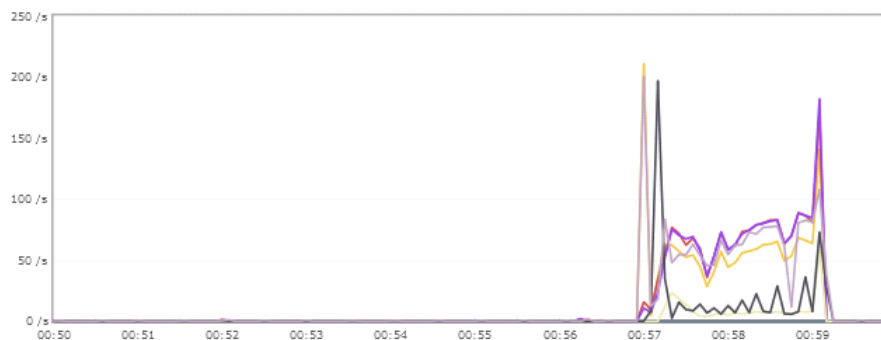
## Overview

▼ Totals

Queued messages last ten minutes ?



Message rates last ten minutes ?



### 3.3GA Values Yaml Example

See <https://github.com/CluedIn-io/ImpTools/blob/master/k8s/pwsh/values.3.3.0.cluedin-platform-1.0.0.yml>

```
# based on migrating https://github.com/CluedIn-io/ImpTools/blob/master/k8s/pwsh/values.3.2.5charts
3.3.0app.redacted.example.yml from old 3.2.5 charts to
# platform based charts
# rha@cluedin.com Feb 2022

global:
  image:
    tag: "3.3.0" # what version of cluedin application we install
    pullPolicy: "IfNotPresent" # I like to not pull newer images unless I
really want them now
  ingress: # enabled HTTPS - I have my certs in the secret name as per 3.2.5
instructions - cert reused
    forceHttps: true
    annotations:
      ingress.kubernetes.io/ssl-redirect: "true"
    tls:
      secretName: "ingress-cert-secret"
  dns:
    hostname: "k1.cluedin.me"
```

```

# all sorts of create object type stuff can be done here
#platform:
# we could define docker-registry-key here if need be
# my cluster already has this defined

application:
  bootstrap:
    organization:
      name: cluedin
      email: admin@cluedin.com
      password: "<redacted>"
  cluedin:
    # this node selector is how we bind the pods to particular node pools
    nodeSelector:
      kubernetes.cluedin.com/pooltype: process
    roles:
      main:
        count: 1
        # all these resource sections were copy and pasted from the older
        # values yml and intended as required
        resources:
          limits:
            cpu: "2"
            memory: "10Gi"
          requests:
            cpu: "1.5"
            memory: "4Gi"
        processing:
          count: 1
          resources:
            limits:
              cpu: "4"
              memory: "8Gi"
            requests:
              cpu: "4"
              memory: "4Gi"
        crawling:
          count: 0
          resources:
            limits:
              cpu: "2"
              memory: "12Gi"
            requests:
              cpu: "2"
              memory: "8Gi"
  webapi:
    nodeSelector:
      kubernetes.cluedin.com/pooltype: general
    resources:
      limits:
        cpu: "0.5"
        memory: "512Mi"
      requests:
        cpu: "0.25"
        memory: "256Mi"
# note: this used to be called clean in 3.2.5 and below
  prepare:
    nodeSelector:
      kubernetes.cluedin.com/pooltype: core
    resources:
      limits:
        cpu: 1000m
        memory: 512Mi
      requests:
        cpu: 250m
        memory: 128Mi
  annotations:
    nodeSelector:
      kubernetes.cluedin.com/pooltype: general
    resources:
      limits:
        cpu: 1000m
        memory: 1000Mi
      requests:
        cpu: 512m
        memory: 512Mi
  datasource:
    nodeSelector:
      kubernetes.cluedin.com/pooltype: process
    resources:
      limits:
        cpu: 1
        memory: 4Gi
      requests:
        cpu: 125m
        memory: 2Gi
  submitter:
    nodeSelector:
      kubernetes.cluedin.com/pooltype: general

```

```

resources:
  limits:
    cpu: 1000m
    memory: 1Gi
  requests:
    cpu: 1000m
    memory: 1Gi
gql:
  nodeSelector:
    kubernetes.cluedin.com/pooltype: general
  resources:
    limits:
      cpu: 1000m
      memory: 1Gi
    requests:
      cpu: 200m
      memory: 64Mi
ui:
  nodeSelector:
    kubernetes.cluedin.com/pooltype: general
  resources:
    limits:
      cpu: 750m
      memory: 512Mi
    requests:
      cpu: 250m
      memory: 256Mi
infrastructure:
  elasticsearch:
    nodeSelector:
      kubernetes.cluedin.com/pooltype: data
    resources:
      limits:
        cpu: 5
        memory: 8Gi
  volumeClaimTemplate:
    resources:
      requests:
        # all the storage sizes were copied from particular places in the
        # original yml and these were put based on understanding the 3rd party
        # charts and internal samples
        storage: "4Gi"
neo4j:
  # there is an issue with 3.5.31
  # so pinning to this one 3 Mar 2022
  imageTag: "3.5.30"
  nodeSelector:
    kubernetes.cluedin.com/pooltype: data
  resources:
    limits:
      cpu: "7"
      memory: "8Gi"
    requests:
      cpu: "7"
      memory: "8Gi"
  core:
    persistentVolume:
      size: "8Gi"
# note: some of these used to be in sqlserver
mssql:
  nodeSelector:
    kubernetes.cluedin.com/pooltype: data
  resources:
    limits:
      cpu: "3"
      memory: "8Gi"
    requests:
      cpu: "3"
      memory: "8Gi"
  persistence:
    dataSize: "8Gi"
    transactionLogSize: "8Gi"
    backupSize: "8Gi"
    masterSize: "8Gi"
redis:
  master:
    nodeSelector:
      kubernetes.cluedin.com/pooltype: general
    resources:
      limits:
        cpu: "1"
        memory: "512Mi"
      requests:
        cpu: "0.1"
        memory: "128Mi"
    persistence:
      size: "512Mi"
rabbitmq:

```

```
nodeSelector:
  kubernetes.cluedin.com/pooltype: process
resources:
  limits:
    cpu: "2"
    memory: "4Gi"
  requests:
    cpu: "1"
    memory: "4Gi"
persistence:
  size: "4Gi"
openrefine:
  nodeSelector:
    kubernetes.cluedin.com/pooltype: core
```