

D4Science Infrastructure - VM Creation #11492

Please create new VM for hosting Natural Language Processing (NLP) web applications:
nlp.d4science.org

Mar 20, 2018 05:49 PM - Massimiliano Assante

Status:	Closed	Start date:	Mar 20, 2018
Priority:	Urgent	Due date:	Apr 10, 2018
Assignee:	_InfraScience Systems Engineer	% Done:	100%
Target version:	Parthenos		
Infrastructure:	Production	Applications list and version:	nlphub 1.0 available at http://ftp.d4science.org/apps/hub.war
RAM:	3	Applications context path:	hub
CPU:	4	External storage data disk:	
DiskSpace:	10	Smartgears auth connector:	
Smartgears:	Yes	data transfer enabled:	No
Smartgears Version:	LATEST	PostgreSQL database:	No
Application Instructions:	nothing particular	Number of VMs:	1
Smartgears scopes:	ParthenosVO/ParthenosLab		
Description <p>The host name should be "nlp.d4science.org", a standard Apache Tomcat 7 (Not Smartgeared) would suffice. A SSL certificate is required as the applications would have to run over https.</p> <p>The web applications deploy in the Apache Tomcat 7 would do a minimal work (the heavy part of the process is done in DataMiner) so just input and output mostly. Not sure how much RAM and CPU are necessary, keep it to the minimum acceptable.</p>			

History

#1 - Mar 20, 2018 05:53 PM - Andrea Dell'Amico

We take care of deploying war in the tomcat instance. No need for tomcat manager activated.

That should never happen in production.

#2 - Mar 20, 2018 05:56 PM - Andrea Dell'Amico

Also, please fill in reasonable values for ram, cpu and disk space.

#4 - Mar 20, 2018 06:07 PM - Andrea Dell'Amico

- RAM changed from 1 to 3
- CPU changed from 1 to 2
- DiskSpace changed from 5 to 10

Do you know the applications context names yet or do I have to reverse proxy /?

Who do must have access to the VM?

#5 - Mar 21, 2018 12:52 PM - Pasquale Pagano

- CPU changed from 2 to 4
- Smartgears changed from No to Yes

I am a bit confused.
First, this application must be a SmartGears application.

Second, the application has to be released as a gCube application. It will be on Nexus. The difference is that we wish and need to update without waiting for the release closure.

Third, each single exploitation of the application is not heavy but we hope to have several users accessing it. I increased the amount of CPUs for this reason.

#6 - Mar 21, 2018 01:55 PM - Massimiliano Assante

Pasquale Pagano wrote:

I am a bit confused.
First, this application must be a SmartGears application.

If we want to use SmartGears is ok but then some context must be able to be called without token, i'm sure this is possible

Second, the application has to be released as a gCube application. It will be on Nexus. The difference is that we wish and need to update without waiting for the release closure.

Applicable for NLPHub not for the others derived from my prortypte and developed by Clarin (Liner2, CSTLemma)

Third, each single exploitation of the application is not heavy but we hope to have several users accessing it. I increased the amount of CPUs for this reason.

#7 - Mar 21, 2018 02:02 PM - Andrea Dell'Amico

Pasquale Pagano wrote:

Third, each single exploitation of the application is not heavy but we hope to have several users accessing it. I increased the amount of CPUs for this reason.

what about load balancing it from the start? I guess that a new developed application will be load-balancer friendly.

#8 - Mar 21, 2018 02:03 PM - Andrea Dell'Amico

Massimiliano Assante wrote:

Applicable for NLPHub not for the others derived from my prortypte and developed by Clarin (Liner2, CSTLemma)

They will never have a public distribution?

#9 - Mar 21, 2018 02:38 PM - Massimiliano Assante

Andrea Dell'Amico wrote:

Pasquale Pagano wrote:

Third, each single exploitation of the application is not heavy but we hope to have several users accessing it. I increased the amount of CPUs for this reason.

what about load balancing it from the start? I guess that a new developed application will be load-balancer friendly.

There is not load balancing need nor of 4 vCPUs need for me. All this web applications do is to take a text, construct a call (string concatenation) and invoke DataMiner which does the heavy process job, if any. Then get back an XML, parse it and display it. 4 CPUs is nonsense to do this.

#10 - Mar 21, 2018 02:39 PM - Massimiliano Assante

Andrea Dell'Amico wrote:

Massimiliano Assante wrote:

Applicable for NLPHub not for the others derived from my prortypte and developed by Clarin (Liner2, CSTLemma)

They will never have a public distribution?

They may, but we can't assume this.

#11 - Apr 03, 2018 06:52 PM - Andrea Dell'Amico

As we discussed, all the applications should be installable from a public repository. I can create the VM in the meantime, but I don't see the advantage.

#12 - Apr 06, 2018 03:28 PM - Massimiliano Assante

- *Priority changed from High to Urgent*

The NLP Hub is going to be presented mid April and we need this VM up & running as soon as possible. For the time being it is going to be managed as dev (stable) machine so we're going to deploy the stable dev artifacts in there.

#13 - Apr 06, 2018 03:36 PM - Andrea Dell'Amico

I still need the application contexts and the scopes.

If you still want to install the applications manually, this one can only be a dev VM, with dev scopes and a name in the dev domain.

#14 - Apr 07, 2018 10:57 AM - Pasquale Pagano

- *Smartgears Version set to LATEST*

- *Smartgears scopes updated*

#15 - Apr 09, 2018 09:54 AM - Massimiliano Assante

- *Due date changed from Mar 30, 2018 to Apr 10, 2018*

- *Application Instructions updated*

- *Applications list and version updated*

- *Applications context path updated*

The VM is needed urgently by tomorrow at last, it is going to work on production on the scope indicate by Lino, you can download the artifact from here: <http://ftp.d4science.org/apps/hub.war>

It should respond to this URL <https://nlp.d4science.org/hub>

Applications context path is /hub

#16 - Apr 09, 2018 02:26 PM - Andrea Dell'Amico

- *Status changed from New to In Progress*

hostname and IP address are going to be nlp.d4science.org 146.48.122.59

#17 - Apr 09, 2018 02:58 PM - Andrea Dell'Amico

There is no artifact:

```
root@ftp:/data/ftp/d4science/www# ls -la nlp
total 4
drwxr-xr-x  2 d4science d4science    6 2018-04-09 09:31 .
drwxr-xr-x 18 d4science d4science 4096 2018-04-09 09:31 ..
root@ftp:/data/ftp/d4science/www# ls -la nlp-hub/
total 4
drwxr-xr-x  2 d4science d4science    6 2018-04-09 09:36 .
drwxr-xr-x 18 d4science d4science 4096 2018-04-09 09:31 ..
```

#18 - Apr 09, 2018 03:03 PM - Andrea Dell'Amico

- *% Done changed from 0 to 70*

@massimiliano.assante@isti.cnr.it let me know when it appears, all the other components have been installed.

#19 - Apr 09, 2018 03:21 PM - Massimiliano Assante

you're looking in the wrong place and wrong name:

it is downloadable from <http://ftp.d4science.org/apps/hub.war>

```
massi-laptop:~ massi$ wget http://ftp.d4science.org/apps/hub.war
```

```
--2018-04-09 15:19:36--  http://ftp.d4science.org/apps/hub.war
Resolving ftp.d4science.org... 146.48.122.130
Connecting to ftp.d4science.org|146.48.122.130|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 3513210 (3.3M) [application/java-archive]
Saving to: `hub.war'
```

```
100%[=====>] 3,513,210    429K/s   in 9.8s
```

```
2018-04-09 15:19:46 (350 KB/s) - `hub.war' saved [3513210/3513210]
```

```
massi-laptop:~ massi$
```

#20 - Apr 09, 2018 03:42 PM - Andrea Dell'Amico

The answer when I run the playbook was:

```
*****
fatal: [nlp.d4science.org]: FAILED! => {"changed": false, "dest": "/home/gcube/downloads/nlp.war", "msg": "Request failed", "response": "HTTP Error 404: Not Found", "state": "absent", "status_code": 404, "url": "http://ftp.d4science.org/apps/nlp.war"}
```

So, the app wasn't there or the file was not readable.

#21 - Apr 09, 2018 03:48 PM - Massimiliano Assante

Andrea Dell'Amico wrote:

The answer when I run the playbook was:

```
*****
fatal: [nlp.d4science.org]: FAILED! => {"changed": false, "dest": "/home/gcube/downloads/nlp.war", "msg": "Request failed", "response": "HTTP Error 404: Not Found", "state": "absent", "status_code": 404, "url": "http://ftp.d4science.org/apps/nlp.war"}
```

So, the app wasn't there or the file was not readable.

Nope, you used the name nlp.war instead of hub.war before

```
"url": "http://ftp.d4science.org/apps/nlp.war"}
```

#22 - Apr 09, 2018 03:50 PM - Andrea Dell'Amico

Ops. Yes.

#23 - Apr 09, 2018 03:56 PM - Andrea Dell'Amico

- Status changed from In Progress to Feedback

- % Done changed from 70 to 100

The app has been deployed.

#24 - Apr 09, 2018 04:00 PM - Massimiliano Assante

```
This site can't be reached
nlp.d4science.org's server IP address could not be found.
Go to http://www.d4science.org/
Search Google for nlp d4science org
ERR_NAME_NOT_RESOLVED
```

#25 - Apr 09, 2018 04:02 PM - Massimiliano Assante

but it works using the VPN (I'm at home) so I suppose it is a matter of minutes to make it working from my home as well.

#26 - Apr 09, 2018 04:05 PM - Andrea Dell'Amico

Yes, the DNS servers have problems updating zones to the outside since some weeks.

#27 - Apr 09, 2018 04:21 PM - Massimiliano Assante

There are a number of contexts that should be enabled (the one corresponding to the servlets of this webapp)

please see 3 url-pattern below

```
<servlet-mapping>
  <servlet-name>NLPServlet</servlet-name>
  <url-pattern>/nlphub-servlet</url-pattern>
</servlet-mapping>

<servlet-mapping>
  <servlet-name>NLPUploader</servlet-name>
  <url-pattern>/nlphub-uploader-servlet</url-pattern>
</servlet-mapping>

  <servlet-mapping>
    <servlet-name>NLPMapper</servlet-name>
    <url-pattern>/nlphub-mapper-servlet</url-pattern>
  </servlet-mapping>
</web-app>
```

#28 - Apr 09, 2018 04:25 PM - Andrea Dell'Amico

Done.

#29 - Apr 10, 2018 10:37 AM - Massimiliano Assante

Thanks, it works now. When is the SSL certificate (non self signed i mean) going to be available ?

#30 - Apr 10, 2018 10:48 AM - Massimiliano Assante

The Nlphub on production is available at <http://nlp.d4science.org/hub> , soon over HTTPS too.

#31 - Apr 10, 2018 12:24 PM - Andrea Dell'Amico

Massimiliano Assante wrote:

Thanks, it works now. When is the SSL certificate (non self signed i mean) going to be available ?

I really don't know. We are almost always over quota, so we'll have to wait several days. I cannot estimate it, it all depends by how many certificates are going to be renewed in the next days.

#32 - May 15, 2018 06:12 PM - Andrea Dell'Amico

- *Status changed from Feedback to Closed*

The certificate has been emitted on April 12th.