

# Project Report (CSE-230 System Administration)

**Title:** Implementing an Intelligent Personal Assistant on local network

**Project ID:** 5

## Team

Aditya Dwivedi, 2014128, [aditya14128@iiitd.ac.in](mailto:aditya14128@iiitd.ac.in)

Tarun Kumar Yadav, 2014110, [tarun14110@iiitd.ac.in](mailto:tarun14110@iiitd.ac.in)

## Objective

To familiarize with working of Docker (containers) and GPU computing by implementing an open source voice assistant “Sirius”.

## Architecture Diagram

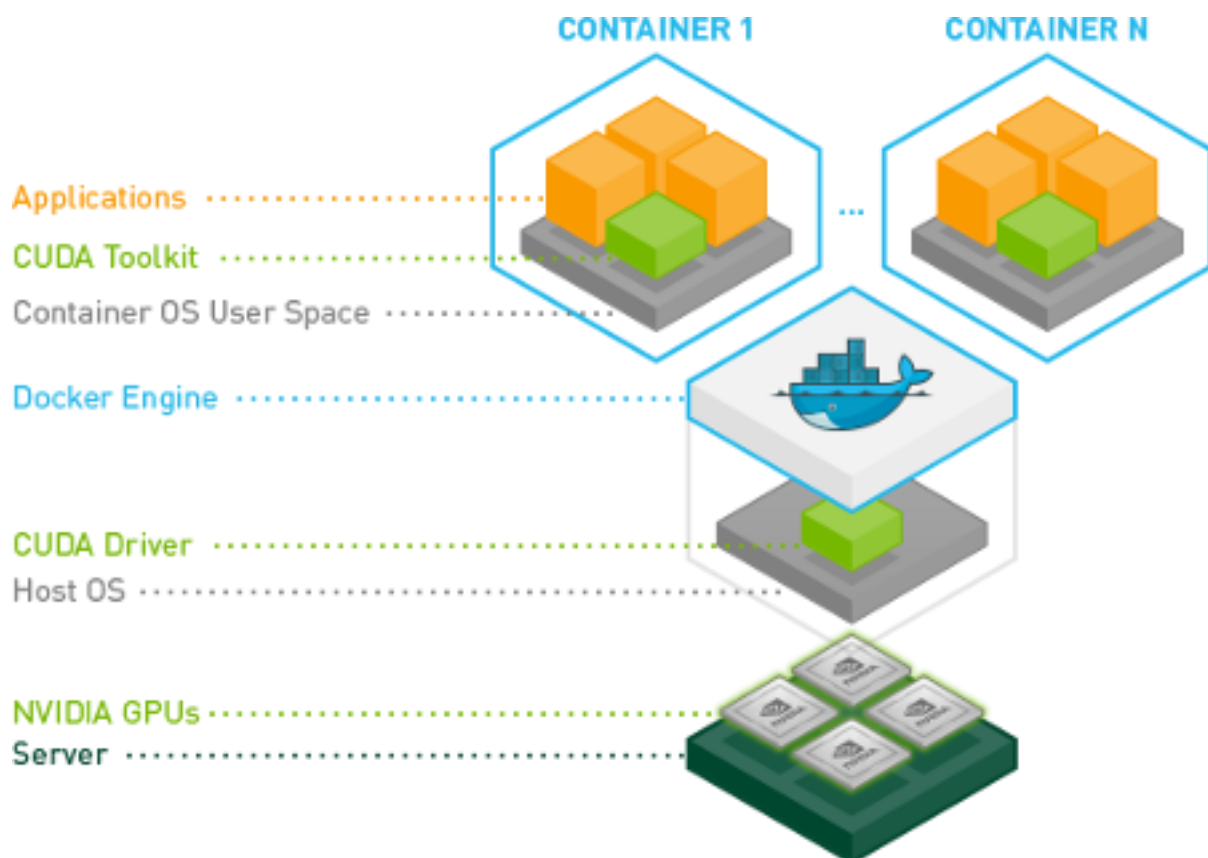
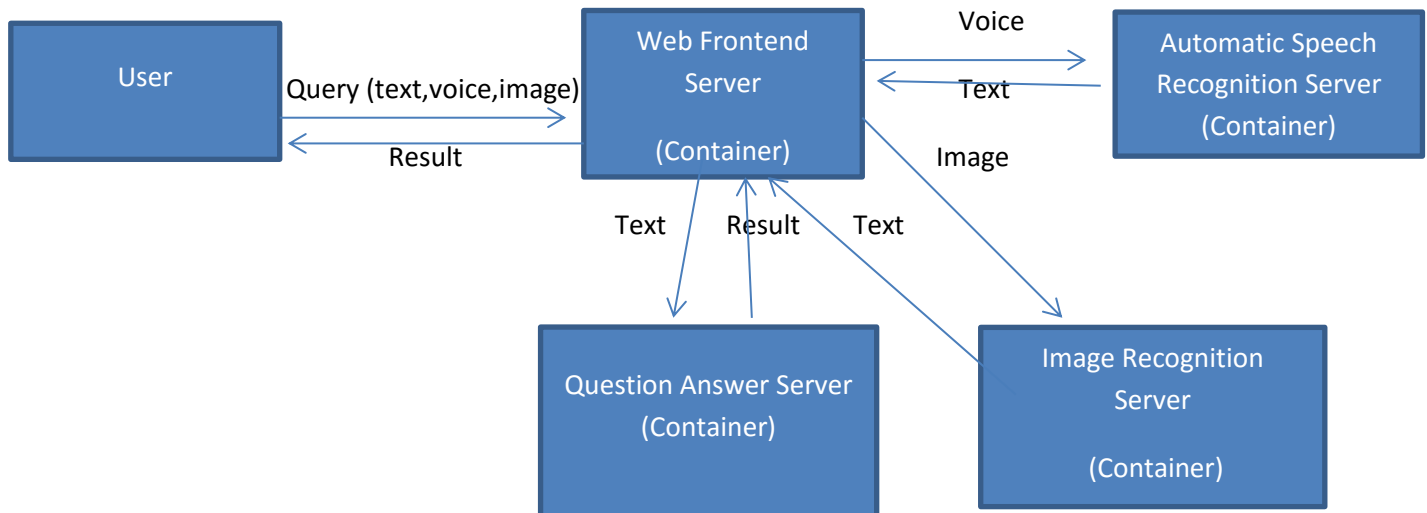


Image Source: <https://cloud.githubusercontent.com/assets/3028125/11199468/c0e09f50-8c82-11e5-846d-1f5e6a410598.png>



## Hardware and Software Prerequisites

Hardware: System with NVidia GPU preferably server GPUs like tesla etc.

Software: Docker engine and NVidia CUDA.

## Links to packages, libraries used

Sirius voice assistant: <http://sirius.clarity-lab.org/downloads/#sirius>

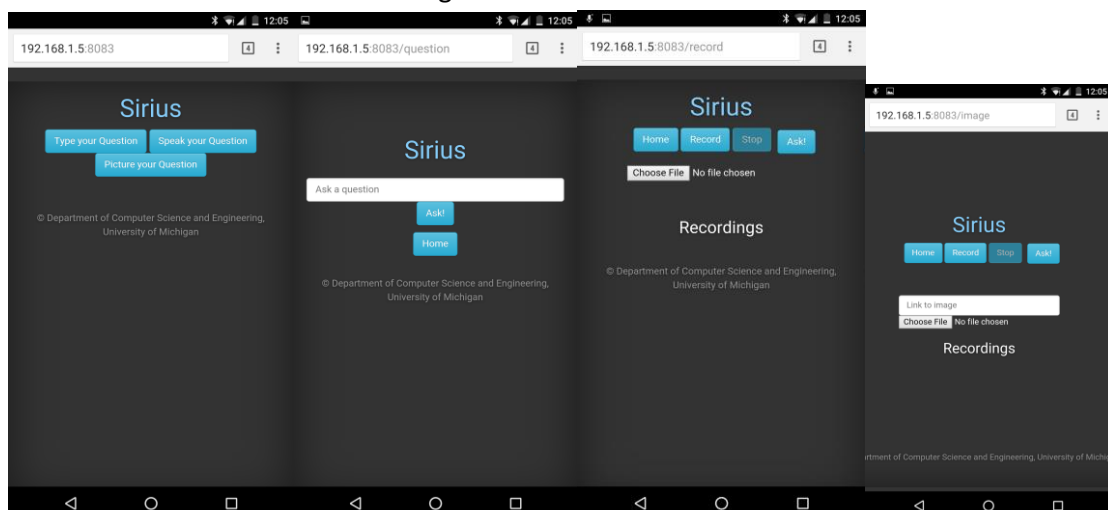
Docker engine: <https://get.docker.com>

NVidia CUDA: <https://developer.nvidia.com/cuda-toolkit-70>

NVidia Docker: <https://github.com/NVIDIA/nvidia-docker>

## Use cases

- Can be used as a system for portable devices especially wearables further giving users the ability to customize it as per their needs. E.g. Raspberry pi can be converted to a smart glass. Screenshots of web frontend running on mobile device



- Since the system is packed into Docker containers, portability and deployment of system becomes very easy.
- GPU implementation lowers the hardware requirement of servers (CPUs) as GPUs could be used to take off load from CPUs thereby enabling even a personal computer with a GPU to cater to needs of few people.

Screenshots of system being hosted on a laptop and resultant load on resources

The first screenshot shows the terminal output of the system startup. It includes commands to start the server and the application, and the resulting output showing the server starting on localhost:8081.

```
root@improvedadi:/opt/sirius/sirius-application/run-scripts# ./start-asr-server.sh
Starting server on localhost:8081, use <Ctrl-C> to stop
.../src/online2bin/online2-wav-nnet2-ltgen-faster --do-endpointing=false --only-neutree --config=net_a_gpu_online/conf/online_nnet2_decoding.conf --max-actives=7000 --beam=15.0 --lattice-beam=6.0 --acoustic-scale=0.1 --word-symbol-tables=gr4ph/words.txt nnet_a_gpu_online/smr_epoch2.mdl graph/HCLG.fst 'ark:echo utterance-ld1 utterance-ld1' scp:echo utterance-ld1 null ark:/dev/null
LOG (online2-wav-nnet2-ltgen-faster:ComputeDerivedVars):lvector-extractor.cc:180 Computing derived variables for lvector extractor
LOG (online2-wav-nnet2-ltgen-faster:ComputeDerivedVars):lvector-extractor.cc:231 Done.
```

The second screenshot shows the network status of the system. It displays the status of the 'lo' and 'wlan1' interfaces, including IP addresses, netmasks, and packet statistics.

```
lo
Link encap:Local Loopback
Inet addr:127.0.0.1 Mask:255.0.0.0
Inet6 addr: ::1/128 Scope:Host
UP LOOPBACK RUNNING MTU:65536 Metric:1
RX packets:179 errors:0 dropped:0 overruns:0 frame:0
TX packets:179 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueue:len:0
RX bytes:12483 (12.4 KB) TX bytes:12483 (12.4 KB)

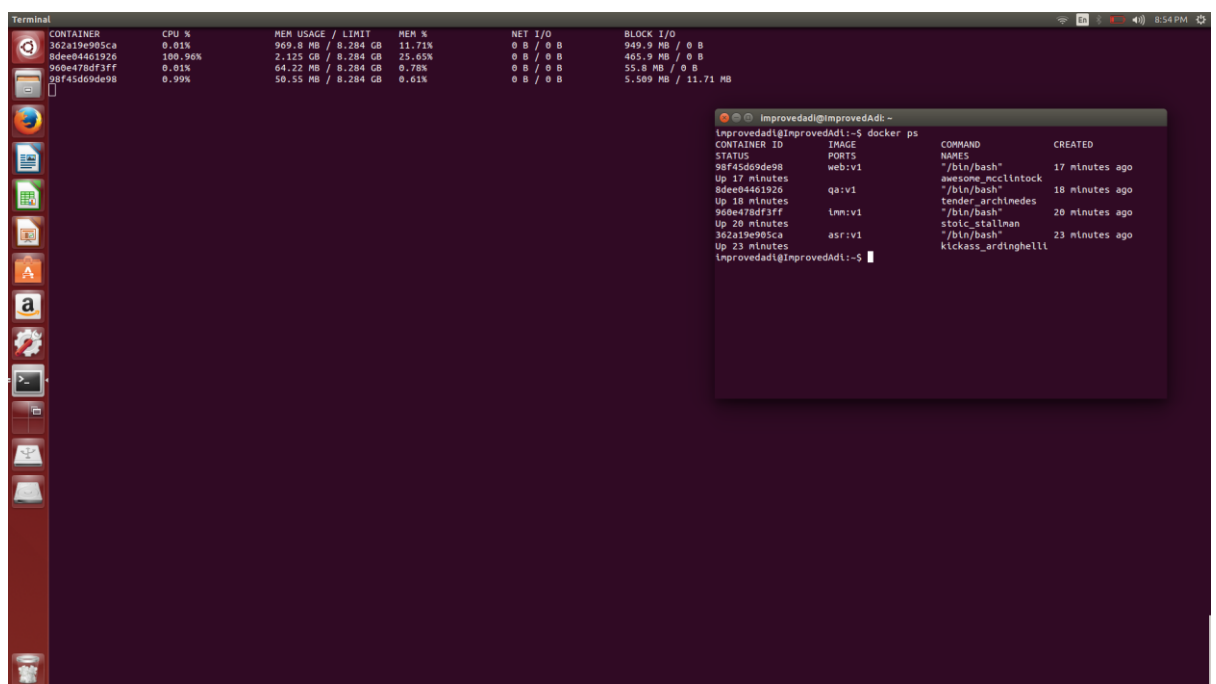
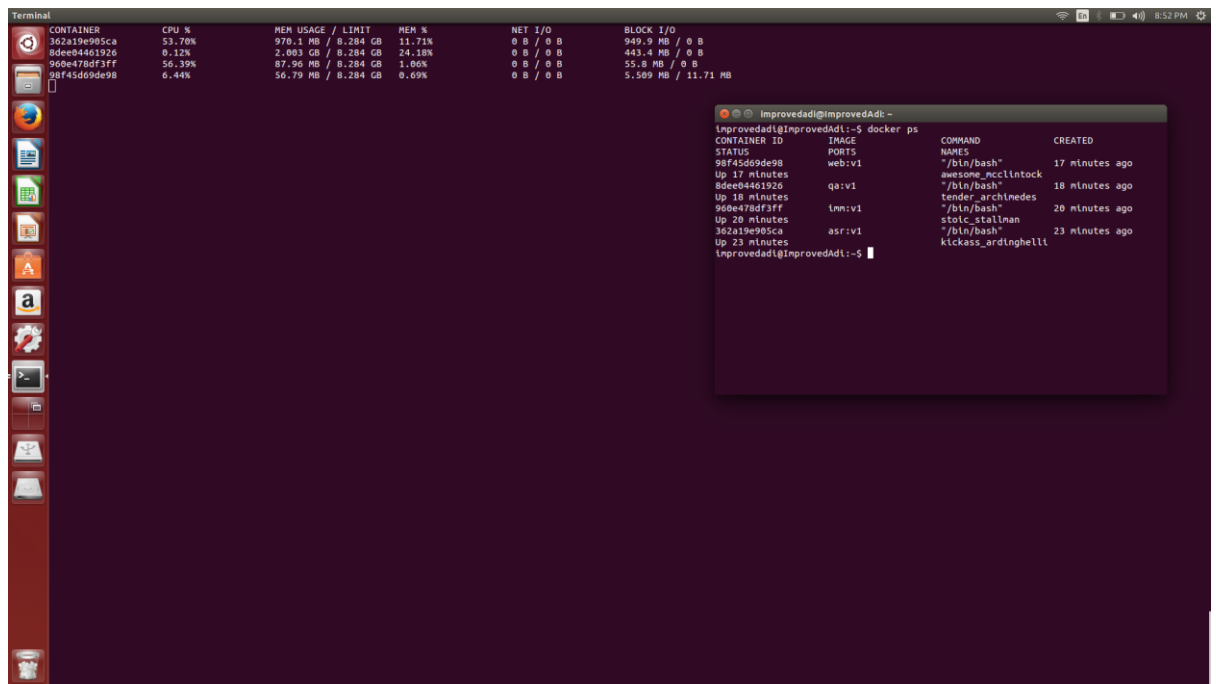
wlan1
Link encap:Ethernet HWaddr 18:c:f:5e:a1:1d:d5
Inet addr:192.168.1.5 Bcast:192.168.1.255 Mask:255.255.255.0
Inet6 addr: fe80::1acf:5eff:fe01:1dd5/64 Scope:Link
UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
RX packets:127 errors:0 dropped:0 overruns:0 frame:0
TX packets:146 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueue:len:1000
RX bytes:12340 (12.3 KB) TX bytes:18407 (18.4 KB)
```

The third screenshot shows the resource usage of the system. It displays a table with columns for CPU %, MEM USAGE / LIMIT, MEM %, NET I/O, and BLOCK I/O. The table shows the usage of the '36219e905ca' container, which is using 644.86% CPU and 2.867 GB of memory.

CONTAINER	CPU %	MEM USAGE / LIMIT	MEM %	NET I/O	BLOCK I/O
36219e905ca	644.86%	2.867 GB / 8.284 GB	34.61%	0 B / 0 B	949.9 MB / 0 B
8dee04461926	0.14%	1.936 GB / 8.284 GB	23.36%	0 B / 0 B	358.6 MB / 0 B
960e478df3ff	0.02%	8.544 MB / 8.284 GB	0.10%	0 B / 0 B	606.2 KB / 0 B
28f45d69de98	1.23%	37.87 MB / 8.284 GB	0.46%	0 B / 0 B	5.419 MB / 4.776 MB

The fourth screenshot shows the output of the 'docker ps' command, which lists the status, image, ports, command, and creation time of the containers. The table shows that the '36219e905ca' container is running and has been created 17 minutes ago.

```
Improvedadi@ImprovedAdi:~$ docker ps
CONTAINER ID        IMAGE               COMMAND             CREATED
STATUS            PORTS              NAMES
28f45d69de98       web:vi             "/bin/bash"        17 minutes ago
Up 17 minutes
8dee04461926       qa:vi              "/bin/bash"        18 minutes ago
Up 18 minutes
960e478df3ff       tmn:vi             "/bin/bash"        20 minutes ago
Up 20 minutes
36219e905ca        asr:vi             "/bin/bash"        23 minutes ago
Up 23 minutes
Improvedadi@ImprovedAdi:~$
```



- Can be used to address queries related to specific domain like managing a library system by providing a custom database.
- Can be used to serve the needs of blind people especially making use of automatic speech recognition feature.
- Image matching can be used to implement biometric authentication.

## Unfinished Tasks

- Customizable database

- GPU performance monitoring

## Citation/Reference

Sirius Team: Johann Hauswald, Michael A. Laurenzano, Yunqi Zhang, Cheng Li, Austin Rovinski, Arjun Khurana, Ron Dreslinski, Trevor Mudge, Vinicius Petrucci, Lingjia Tang, and Jason Mars. Sirius: An Open End-to-End Voice and Vision Personal Assistant and Its Implications for Future Warehouse Scale Computers. In *Proceedings of the Twentieth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, ASPLOS '15, New York, NY, USA, 2015. ACM. Acceptance Rate: 17%

Base image used for system: <https://github.com/danielchalef/sirius-docker>