

# DS 5110 Big Data Systems Term Project

Team 1: Harold Haugen, George Shoriz, Zach Lisman, Dan Anthony

11.19.2024

# Prior Work (Steps #1-3)

- ▶ 1: Step Function
  - ▶ Facilitates orchestration of FMI lambda functions
  - ▶ Measure execution time and cost
- ▶ 2: Rendezvous Server
  - ▶ Establishes communication between AWS lambda functions
- ▶ 3: AI for Astronomy Inference: Run Inference.py file
  - ▶ resized\_images
- ▶ Additional: Review Machine Learning

# Recent Work

- ▶ Review fork of GitHub Repo with edits from Mills
- ▶ Integration of Cosmic AI with AWS/Lambda
  - Make edits to inference.py and Plot\_Redshift.py for when it is run
    - Passes in rank & world\_size
    - When run, it calls engine function which calls inference function
  - Inference function
    - Reduction operation
      - ▶ Pass in function type (sum)
      - ▶ Works with execution info (as doubles)
    - Added arguments for Rank & World\_Size

The screenshot shows a GitHub repository interface. At the top, it indicates the branch is 'main' and shows commit status: 'This branch is 4 commits ahead of, 20 commits behind UVA-MLSys/AI-for-Astronomy:main'. Below this is a search bar and a file explorer on the left. The file explorer shows a directory structure: 'code' > 'Inference' > 'inference.py'. The right pane shows a table of commit history for the 'inference.py' file.

Name	Last commit message	Last commit date
..		
.ipynb_checkpoints	astronomy model and image overview...	last month
fmlib	changes based on aws lambda env	3 weeks ago
Inference Step by Step Instructi...	astronomy model and image overview...	last month
__init__.py	changes based on aws lambda env	3 weeks ago
inference.py	changes based on aws lambda env	2 weeks ago
resized_inference.pt	data is resized and inference and meas...	last month

```
execution_info = {
    'total_time': total_time,
    'execution_time': total_time / num_batches, # Average execution time per batch
    'num_batches': num_batches, # Number of batches
    'batch_size': batch_size, # Batch size
    'device': device, # Selected device
    'throughput_bps': total_data_bits / total_time, # Throughput in bits per second
    'sample_persec': num_samples / total_time
}

plt_rdshft.err_calculate(redshift_analysis, real_redshift, execution_info,
                        plot_to_save_path) #invoke for calculating statistical prediction evaluation metrics

reduce_res = comm.reduce(total_time, 0, fmi.func(fmi.op.sum), fmi.types(fmi.datatypes.double))
```

```
#Iterate over data for predicting the redshift and invoke the evaluation modules
def inference(model, data_loader, real_redshift, plot_to_save_path, device, batch_size, rank, world_size):

    parser.add_argument('--rank', type=int, **environ_or_required('RANK', required=False))
    parser.add_argument('--world_size', type=int, **environ_or_required('WORLD_SIZE', required=False))
```

# Recent Work

## ► Copied/created our own:

- State Machine (*copied IAM roles*)
- S3 Bucket (*and copied files*)

## ► First Trial for Experiments

- Ran Step Function
  - Output to Cosmic AI
- Changed Inference.py 'Bucket =' to point to our S3 Bucket
- Try Different World Sizes
  - World Size 1: ~6 second
  - World Size 6: ~6 Second
  - World Size 16 & 128: Did not finish

Cosmic\_AI\_Team\_One

Standard

Nov 18, 2024, 19:28:57 (UTC-05:00)

Active

Bucket name

Info

Team\_One\_S3\_Bucket

Bucket name must be unique within the global namespace and follow the bucket naming rules.

[See rules for bucket naming](#)

<a href="#">Cosmic_Trial_TOne_Mon_903</a>	<div><div>✖</div>Failed</div>	Nov 18, 2024, 21:03:58	Nov 18, 2024, 21:19:12	00:15:13.800
<a href="#">Cosmic_Trial_TOne_Mon_850</a>	<div><div>✔</div>Succeeded</div>	Nov 18, 2024, 20:50:14	Nov 18, 2024, 20:50:29	00:00:15.237
<a href="#">Cosmic_Trial_TOne_Mon_836</a>	<div><div>✔</div>Succeeded</div>	Nov 18, 2024, 20:36:18	Nov 18, 2024, 20:36:34	00:00:15.674
<a href="#">Cosmic_Trial_TOne_Mon_828</a>	<div><div>✔</div>Succeeded</div>	Nov 18, 2024, 20:28:30	Nov 18, 2024, 20:28:47	00:00:17.074

```
ResultsReduce-01.json  Results-01.json  Results-06.json  ResultsReduce-06.json
File Edit View File Edit View
{
  "total_cpu_time (seconds)": 5.936356256999996,
  "total_cpu_memory (MB)": 14320.256568,
  "execution_time (seconds/batch)": 2.968178128499998,
  "num_batches": 2,
  "batch_size": 512,
  "device": "cpu",
  "throughput_bps": 28317006.716330588,
  "sample_persec": 172.4963859425596,
  "cpu_info": {
    "processor": "x86_64",
    "architecture": [
      "64bit",
      "ELF"
    ],
    "machine": "x86_64",
    "system": "Linux",
    "platform": "Linux-5.10.227-239.884.amzn2.x86_64"
  },
  "ram_info (GB)": 10.455680847167969,
  "avg_profile": "<FunctionEventAvg key=Total self_cpu_time=0.000us input_shapes= cpu_memory_usage=143202568 self_cpu_memory (MB)": 10.510336
}

{
  "total_cpu_time (seconds)": 5.814478928000014,
  "total_cpu_memory (MB)": 14320.256568,
  "execution_time (seconds/batch)": 2.907239464000007,
  "num_batches": 2,
  "batch_size": 512,
  "device": "cpu",
  "throughput_bps": 28910559.67036082,
  "sample_persec": 176.11208376194455,
  "cpu_info": {
    "processor": "x86_64",
    "architecture": [
      "64bit",
      "ELF"
    ],
    "machine": "x86_64",
    "system": "Linux",
    "platform": "Linux-5.10.227-239.884.amzn2.x86_64-x86_64-with-glibc2.35"
  },
  "ram_info (GB)": 10.455680847167969,
  "avg_profile": "<FunctionEventAvg key=Total self_cpu_time=2.229s cpu_time=375.273us self_cuda_time=0.000us input_shapes= cpu_memory_usage=14320256568 cuda_memory_usage=0",
  "self_cpu_memory (MB)": 10.510336
}
```

# Future Work (Deliverables)

- ▶ Try restarting rendezvous server and review video from last week
- ▶ Experiment with partitioning and run inference with varied parameters and visualize results via charts or graphs
- ▶ Record 5-10 Minute video presenting term project
  - Introduction
  - Data
  - Experimental Design
  - Beyond original specification
  - Results
  - Testing
  - Conclusions (Systems & ML)
- ▶ Update README to reflect presentation and final outcomes
- ▶ Upload any files to GitHub
  - Code
  - Presentations/Updates