ADVANCED RESEARCH COMPUTING (ARC) IN CANADA:

EMPOWERING DISCOVERY THROUGH DIGITAL INFRASTRUCTURE

Mohammad Alnabhan

AGENDA

- What is ARC?
- Why ARC is Critical for Research
- Canada's ARC Infrastructure
- How to Access ARC Systems
- Running Jobs on Clusters
- Using GPUs for Deep Learning
- Best Practices
- Conclusion

WHAT IS ADVANCED RESEARCH COMPUTING (ARC)?

- ARC uses powerful computers (clusters) to solve complex research problems.
- Enables data processing, simulations, modelling, machine learning, and artificial intelligence.
- Essential for fields like climate science, genomics, physics, and social science analytics.

WHY ARC MATTERS FOR RESEARCH

- Accelerates scientific discovery by shortening computation time dramatically.

- Supports multi-disciplinary research across Canada.

- Makes advanced data-intensive techniques possible for researchers in all fields.

CANADA'S ARC INFRASTRUCTURE

Niagara — University of Toronto (AI and Climate Modeling)

Cedar — Simon Fraser University (General purpose HPC)

Graham — University of Waterloo (Big data and analytics)

Narval — Calcul Québec (Large memory and GPU capabilities)

These systems are interconnected nationally via the Digital Research Alliance of Canada.

HOW TO ACCESS ARC RESOURCES

Apply for an account through the Compute Canada Database (CCDB).

Request compute and storage resources via the Resource Allocation Competitions or Rapid Access.

Connect remotely using SSH from your personal device to the cluster login nodes.

RUNNING JOBS ON CLUSTERS

Submit jobs using SLURM workload manager.

Job scheduling ensures fair sharing of resources among all users.

RUNNING JOBS ON CLUSTERS

Submit jobs using SLURM workload manager.

Job scheduling ensures fair sharing of resources among all users.

SAMPLE SLURM SCRIPT

```
#!/bin/bash
#SBATCH --job-name=test_job
#SBATCH --time=02:00:00
#SBATCH --gpus=1
#SBATCH --mem=16G

module load python/3.10
python train_model.py
```

USING GPUS FOR DEEP LEARNING

Many ARC systems include specialized GPU nodes.

Deep learning models (like BERT for fake news detection) can be trained much faster using GPUs.

Libraries supported: TensorFlow, PyTorch, HuggingFace Transformers.

BEST PRACTICES FOR USING ARC

Request only the resources you truly need.

Move large datasets to $SCRATCH space; keep minimal files in $HOME.

Monitor job progress using squeue and analyze job efficiency using seff.

Regularly back up critical results to external storage.

CONCLUSION

ARC provides the foundation for innovation and discovery in Canadian research.

Researchers from all disciplines are encouraged to leverage national HPC resources.

The Digital Research Alliance ensures equitable access for everyone.

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

THANK YOU!