

# Niranda Perera

January 10, 1990

Last update on January 26, 2022

+1 812-558-8884 • [niranda@niranda.dev](mailto:niranda@niranda.dev) • [github.com/nirandaperera](https://github.com/nirandaperera) • [www.linkedin.com/in/niranda](https://www.linkedin.com/in/niranda)

## RESEARCH INTERESTS

---

High Performance Computing, Data Engineering, *Dataframes*, Distributed Programming, Distributed Deep Learning

## EXPERIENCE

---

### Digital Science Center, Indiana University

BLOOMINGTON, USA

#### Research Assistant

Aug 2018 – Present

Working on high performance computing, data engineering, and AI/ML

- Lead developer of **Cylon**, a **distributed high performance data engineering framework** based on **Apache Arrow**, with a **C++** backend & a **Python** frontend using **Cython**
- Developing a pandas-like *dataframe* API for Bulk Synchronous Parallel (**MPI**) environments to seamlessly integrate data engineering with distributed Deep Learning
- Developing *Cylon-Flow* execution environment that integrates Cylon with **Parsl**, **Dask**, & **Ray**
- Working on parallel DNN training in **PyTorch** (ex:**Pipedream**, **GPipe**)
- Developing **Twister2**, a **Java** based composable Big Data toolkit for cloud and high performance computing infrastructure

#### Teaching Assistant

Aug 2019 – Apr 2020

Engineering Cloud Computing (E516)

- Assisted students with their **Cloudmesh** projects
- Technologies covered: Cloud infrastructure (**Azure**, **AWS**, **GCP**, **OpenStack**), Container management (**Docker**, **Kubernetes**), REST APIs

### Voltron Data Inc. (FKA Ursa Computing Inc.)

REMOTE - GREENWOOD, INDIANA

#### Software Engineering Intern

May 2021 – Aug 2021

Working on **Apache Arrow C++ Compute API** development

- Developed to compute kernels, bug fixes, and compute utilities in C++ and Python
- Extensively worked on **Template Metaprogramming (TMP) & Object Oriented Programming (OOP) in C++**. Major contributions - *Hash Semi-Join Node*, *If-else Kernel*, *Bitmap Word Visitor API*
- Apache Arrow JIRAs worked on [can be accessed here](#).
- Apache Arrow Github Pull Requests [can be accessed here](#).

### Department of Computer Science and Engineering, University of Moratuwa

MORATUWA, SRI LANKA

#### Research Assistant

Jan 2017 – Jun 2018

Developing a Cloud Based, Real-Time Weather Modeling and Forecasting Framework for **Center for Flood Control and Water Management**, Sri Lanka.

- Worked on statistical and numerical weather prediction models (WRF, SHER, and HEC-HMS)
- Developed a **Python**-based distributed execution framework to run models in production using **Apache Airflow**, **Google Cloud Platform**, & **Docker**

### WSO2 Inc

COLOMBO, SRI LANKA

#### Senior Software Engineer

Apr 2016 – Dec 2016

Open source middleware stack using **OSGi** spec. Member of **WSO2 Data Analytics Server (DAS)**, an enterprise platform for batch, streaming, & predictive data analytics.

- Integrated **Apache Spark** into DAS
- Developed analytics solutions for WSO2 Enterprise Service Bus, API Manager, & Identity Server
- Experienced in OSGI architecture, Open Source Software development practices, and providing enterprise production & support

#### Software Engineer

Mar 2014 – Apr 2016

## EDUCATION

---

### Indiana University

BLOOMINGTON, USA

#### Doctor of Philosophy, Intelligent Systems Engineering

2018 – Present

**PhD Candidate** advised by Prof. **Geoffrey Fox** with a **GPA of 3.989** (completed cr. 82/90)

Major: **Computer Engineering**, Minor: **Cyber-Physical Systems** (Expected graduation Aug, 2022)

Dissertation topic: **Towards Scalable High Performance Data Engineering Systems**

- Proposing a generic model & operator patterns for distributed data engineering systems
- Developing *Cylon* and benchmarking with TPC-H TPCx-BB benchmark suites
- Developing *Cylon-Flow*, a hybrid execution environment that integrates Bulk Synchronous Parallel and Distributed Asynchronous execution models

## PUBLICATIONS

- [1] V. Abeykoon, S. Kamburugamuve, C. Widanage, N. Perera, A. Uyar, T. A. Kanewala, G. von Laszewski, and G. Fox, "Hptmt parallel operators for high performance data science & data engineering," *arXiv preprint arXiv:2108.06001*, 2021.
- [2] N. Perera, V. Abeykoon, C. Widanage, S. Kamburugamuve, T. A. Kanewala, P. Wickramasinghe, A. Uyar, H. Maithree, D. Lenadora, and G. Fox, "A fast, scalable, universal approach for distributed data reductions," *arXiv preprint arXiv:2010.14596*, 2020.
- [3] V. Abeykoon, N. Perera, C. Widanage, S. Kamburugamuve, T. A. Kanewala, H. Maithree, P. Wickramasinghe, A. Uyar, and G. Fox, "Data engineering for hpc with python," in *2020 IEEE/ACM 9th Workshop on Python for High-Performance and Scientific Computing (PyHPC)*, IEEE, 2020, pp. 13–21.
- [4] C. Widanage, N. Perera, V. Abeykoon, S. Kamburugamuve, T. A. Kanewala, H. Maithree, P. Wickramasinghe, A. Uyar, G. Gunduz, and G. Fox, "High performance data engineering everywhere," in *2020 IEEE International Conference on Smart Data Services (SMDS)*, IEEE, 2020, pp. 122–132.
- [5] P. Wickramasinghe, N. Perera, S. Kamburugamuve, K. Govindarajan, V. Abeykoon, C. Widanage, A. Uyar, G. Gunduz, S. Akkas, and G. Fox, "High-performance iterative dataflow abstractions in twister2: Tset," *Concurrency and Computation: Practice and Experience*, e5998, 2020.
- [6] P. Wickramasinghe, S. Kamburugamuve, K. Govindarajan, V. Abeykoon, C. Widanage, N. Perera, A. Uyar, G. Gunduz, S. Akkas, and G. Fox, "Twister2: Tset high-performance iterative dataflow," in *2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, IEEE, 2019, pp. 55–60.
- [7] A. Uyar, G. Gunduz, S. Kamburugamuve, P. Wickramasinghe, C. Widanage, K. Govindarajan, N. Perera, V. Abeykoon, S. Akkas, and G. Fox, "Twister2 cross-platform resource scheduler for big data,"

## PRESENTATIONS

**IEEE BigData 2020 - IWBDP Workshop:** Presenting *A Fast, Scalable, Universal Approach For Distributed Data Aggregations* ([Video Link](#)), Dec 2020

**ApacheCon @Home 2020 Conference:** Presenting **Cylon** to Apache Community ([Video Link](#)), Sep 2020

## SKILLS

**Programming Languages:** C/C++, Python, Java, Cython

**Data Engineering:** SQL, Apache Spark, Apache Hadoop, NumPy, Python Pandas, Rapids CuDF

**Technologies:** OpenMPI, UCX, Ray-Project, Dask, Parsl, Docker, Kubernetes, Apache Airflow, REST

**Deep Learning:** PyTorch, Tensorflow

**Hardware Languages:** x86 Assembly with AVX/2/512, Verilog (moderate fluency)

## HONORS, AWARDS, AND ACHIEVEMENTS

**Hiking and travel:** Achieved a height of 16,200ft (4950m) on foot in Junarghali Pass, Roopkund Trek, Uttarkund, India in June, 2016.

**Basketball, WSO2:** Captained the team which emerged *champions* at the Mercantile Services Basketball Association League Tournament 2016 - Division E.

**Dean's List, Faculty of Engineering, UoM:** During the Semesters 1 and 8

**Basketball, UoM:** *Colorsman* during the years 2010, 2011, 2013. Emerged *champions* at the Inter University Basketball Championship 2011.

**G.C.E. Advanced Level Examination, Sri Lanka:** Ranked Island 21<sup>st</sup>, Colombo District 9<sup>th</sup> with a Z-Score of 2.9127 from Physical Science

**Nalanda College, Colombo:** Awards for the *Most Outstanding Student of the Year*, *Best Result – Science Section*, *Best Result – Physical Science & Best Student in General Knowledge* in the year 2008

**Dept. Electronic & Telecommunication, UoM:** *Student Representative* of the 2009 Batch for the year 2010/11

**Electronic Club, UoM:** *Treasurer* for the year 2013 & Chief Organizer, Sri Lanka Robotics Challenge 2012

**English Literary Association, UoM:** *Vice President* for the year 2011/2012

**English Debating Team, UoM:** Captained the Team B in 2012

**IEEE XTREME Programming Competition:** Participant in 2011, 2012 & 2013

**CIMA Global Business Challenge:** *Country Runner-Up*, Team Unorthodox in 2012