# Roshan R Nair

roshannairwork@gmail.com roshanrn.github.io Roshan R Nair

## Education

**Birla Institute of Technology and Science, Pilani, (Goa Campus)** (India) Aug. 2015 to Aug. 2020
M.Sc. (Hons.) Mathematics (Class Rank: 9/45) , CGPA 8.79/10
B.E. (Hons.) Electrical and Electronics Engineering (Class Rank: 15/123) (Dual Degree Program)
*(Completed with 212 units = ∼ 352 ECTS)*

**Final Year Projects (Practice School) at Samsung Semiconductor India Research (SSIR)**
(6 month stay for each degree) Advised by Arun George (Samsung) and Vishak G (Samsung)

***I/O signature and workload identification in a userspace framework.*** Jan. 2020 to June 2020
Explored using a classifier chains model to perform multi-workload identification in a datacenter environment. This work led to a primary inventor patent application.

***Statistical Modelling and DL for NAND Memory access patterns and I/O workloads.*** July 2019 to Dec. 2019
Explored the use of various statistical modeling and deep-learning approaches to extract access patterns of various block-based workloads.

## Publications

**Towards Efficient Flash Caches with Emerging NVMe Flexible Data Placement SSDs.*** Michael Allison, Arun George, Javier Gonzalez, Daniel Helmick, Vikash Kumar, ***Roshan R Nair***, and Vivek Shah.
**EuroSys 2025** ***Acceptance Rate: 12.4%*** *(Alphabetical Order)

**Numerical solution of a nonlinear fractional model for hepatitis C by using Haar wavelets**. Amit Setia, Pushpendu Ghosh, and ***Roshan R Nair.***
AIP Conference Proceedings (Vol. 1975, No. 1). AIP Publishing, 2018.

## Patents

**Method and system for performing replication recovery operation in a disaggregated distributed storage systems**, (US18/090,023) *Primary Inventor* (Granted)

**Device and method for data replication in disaggregated distributed storage system**, (US18/089,986) *Primary Inventor* (Granted)

**Method and system for distributing and managing io in a disaggregated storage architecture**, (US18/089,163) *Co-Inventor* (Grant Pending)

**Methods and systems for identifying multiple workloads in a heterogeneous environment**, (US17/935,124) *Primary Inventor* (Grant Pending)

## Industry Contributions

**The Promise of NVMe Flexible Data Placement in Data Center Sustainability**.
***Roshan R Nair*** and Arun George.
Presented at SNIA Storage Developer Conference, 2024
Presented to OCP Sustainability Initiative - Invited Guest Speaker, 2024

**Getting Started with Flexible Data Placement (FDP)** Samsung White Paper.
Arun George, Vikash Kumar, ***Roshan R Nair***, Vivek Shah, et. al. White Paper, 2024

**A Brief History of Data Placement Technologies** Technical Blog Post.
***Roshan R Nair.*** Published in Samsung's Technical Blog, 2024

## Industry Research Experience

**Samsung Semiconductor India Research (SSIR) - Staff Engineer (SDE 3)** ***Aug. 2020 to Present***

*Currently affiliated with Samsung's Global Open-ecoSystem Team (GOST)*

**Project: CXL Memory Pooling for AI applications** Feb. 2025 to Present

*Co-researchers: Arun George*

This project aims to create a memory pooling solution based on features provided by the Compute Express Link (CXL) specification to cater to AI inference and training use-cases.

- Investing dynamic capacity device (DCD) based multi-headed CXL devices to create a memory pooling solution. This involves studying the current kernel, qemu, and ndctl patches around DCD.
- Investigating VectorDB DiskANN bulk insert and search workloads to showcase the memory pooling solution.

**Project: SSD Write-Amplification Modeling Framework** Aug. 2024 to Feb. 2025

*Co-researchers: Arun George, Shrihari ES*

This project involves designing and developing a Python-based framework to estimate and model SSD internals, such as write amplification, garbage collection, and read-modify-writes. This framework enables modeling SSD configurations and features that are unavailable in the market.

- Investigated various functionalities of the SSD host interface and flash translation layers and simulated their behaviour. Collected detailed metrics to enable better WAF analysis using the modeling framework.
- Utilized the write-amplification model to investigate the effects of various flash cache designs and block trace workloads. Assessed the suitability of using QLC-based SSDs for these existing classes of workloads. The framework is planned to be open-sourced for the research community to use.

**Project: Storage Sustainability - Improve SSD Lifetime and Power Efficiency**

*Co-researchers: Dr. Vivek Shah, Arun George* July 2024 to Feb. 2025

The project studied the impact of various host configurables, such as data placement, large indirection unit (IU), NVMe power states, etc., on SSD power consumption, lifetime, and data center sustainability.

- Researched the role of NVMe FDP in improving SSD lifetime and power efficiency, thereby reducing operational and embodied carbon emissions in large-scale deployments.
- Investigated the impact of unaligned writes on power efficiency with large indirection unit drives (QLC SSDs).

**Project: NVMe Flexible Data Placement (FDP) Eco-system** Sep. 2023 to Feb. 2025

*Co-researchers: Arun George, Dr. Vivek Shah and Dr. Javier Gonzalez*

This project involved researching NVMe FDP and its integration into various eco-systems, focusing on reducing SSD write amplification.

- Flash Caches: Researched the benefits and trade-offs of using NVMe FDP with flash caches, such as CacheLib and Kangaroo. This work resulted in up-streaming FDP support to CacheLib and a research paper.
- Mutli-tenancy and data grouping: Researched the use of NVMe FDP to improve SSD write amplification for multi-tenant deployments.

**Project: Distributed Block Storage** Aug. 2020 to Sep. 2023

*Co-researchers: Arun George and Vishak G*

Worked on Samsung's in-house disaggregated (controller-storage) distributed block storage solution, leveraging SPDK and NVMe-oF. I was involved in the research and development of various modules.

- Load Balancing: Investigated different policies and simulated various cluster configurations to analyze the load distribution on controller and storage nodes. This work led to the design of an optimized NAND and resource-aware load-balancing policy, which is used in the solution.

- Data Redundancy (data replication and recovery): Investigated, designed, and implemented an N-way synchronous replication scheme and its recovery mechanism.
- Performance Analysis: Developed an in-house framework to identify performance deficiencies, utilizing tools such as perf and flamegraphs. This work resulted in an internal benchmarking white paper that compared Ceph with the in-house solution in terms of architecture, design, and KPIs including bandwidth, IOPS, and latency.
- Module Lead (design and development): State machine replication module using a c++ raft library, flexible node scaling and failure management, synchronous replication, etc.

## Teaching Experience

**Birla Institute of Technology and Science, Pilani** *(Goa Campus)*

**TA,** EEE F244: Microelectronic Circuits - Conducted and graded weekly lab sessions. Spring 2019
**TA,** MATH F312: Ordinary Differential Equations - Conducted tutorial sessions. Fall 2018
**TA,** MATH F214: Elementary Real Analysis - Conducted tutorial sessions and graded quizzes. Fall 2017

## Additional Experience And Awards

**Employee of the Month,** Samsung Semiconductor: August 2024 and March 2021
**Professional Software Competency (Jan. 2024),** Samsung Semiconductor India: Coding Competency Certification held by < 20% of eligible software developers.
**INSPIRE Scholarship, (Jan 2017)** Scholarship rolled out by the Department of Science and Technology, Government of India to top students pursuing natural science degrees.
**Chief Coordinator** The Literary and Debating Club, BITS Pilani (August 2016 - July 2017)
**Editor, College Magazine BITS, Pilani** (August 2016 - August 2017)

## Languages

English: Full Fluency - IELTS: 8.5 Band Score CEFR Level C2