Report

on a one-day workshop on

Mastering Data Systems

From PostgreSQL Internals to Distributed Architectures

Date: February 1, 2025

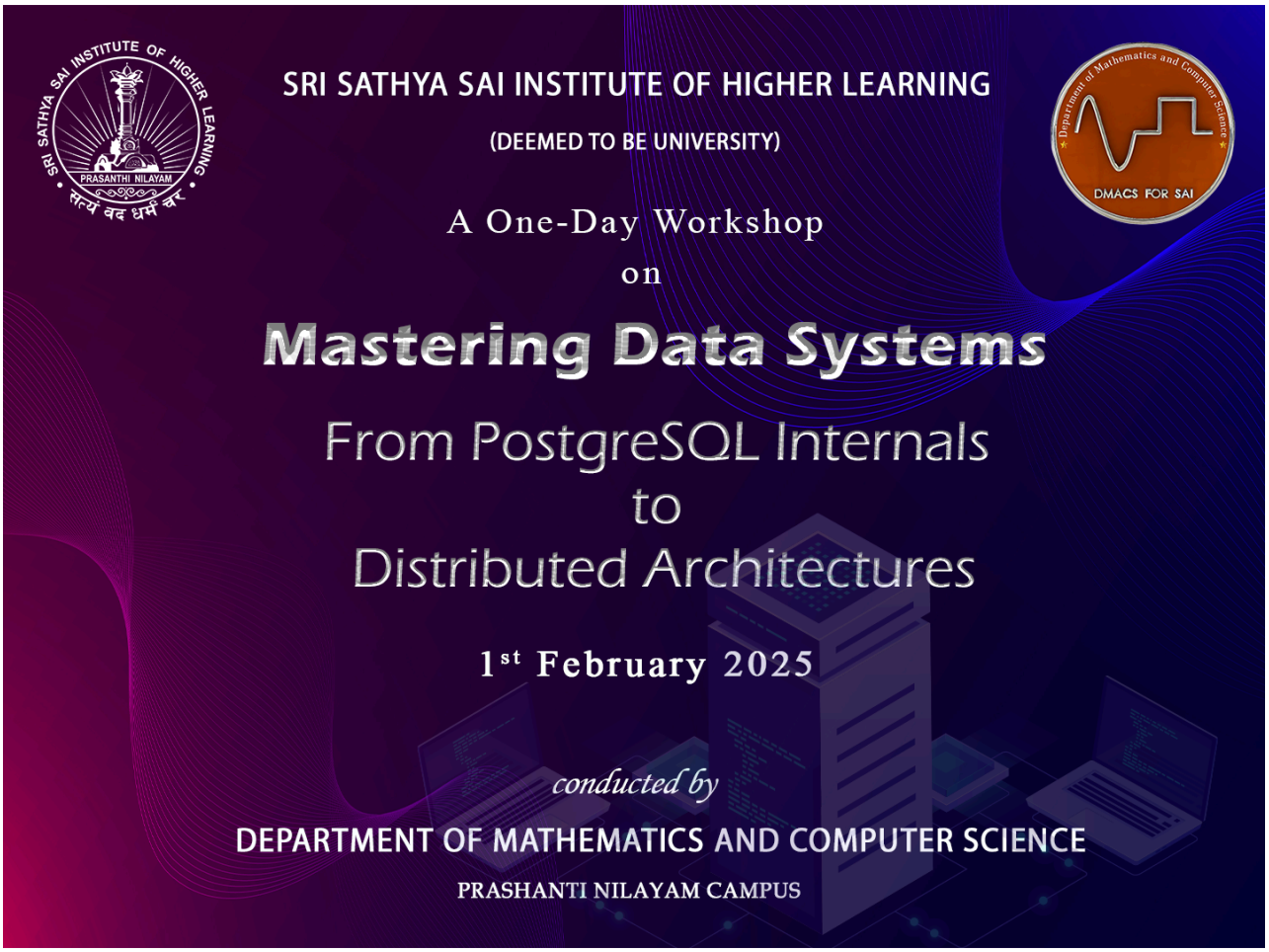
Venue: Multimedia Learning Center, Prasanthi Nilayam Campus

conducted by

Department of Mathematics and Computer Science

Sri Sathya Sai Institute of Higher Learning

Flyer


SRI SATHYA SAI INSTITUTE OF HIGHER LEARNING
PRASANTHI NILAYAM
सत्यं वद धर्मं चर
SRI SATHYA SAI INSTITUTE OF HIGHER LEARNING
(DEEMED TO BE UNIVERSITY)
Department of Mathematics and Computer Science
DMACS FOR SAI
A One-Day Workshop
on
Mastering Data Systems
From PostgreSQL Internals
to
Distributed Architectures
1st February 2025
conducted by
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
PRASHANTI NILAYAM CAMPUS

Schedule

Time	Event
8:30 – 8:45	*Invocation Vedam* by Students and *Welcome Address* by ***Prof. Raghunatha Sarma***, The Director, PSN Campus, SSSIHL, and former Head, DMACS, SSSIHL.
8:45 – 9:45	Talk by ***Prof. Ravi Mukkamala***, Chair and Professor of Computer Science, Old Dominion University, and Adjunct Faculty, DMACS, SSSIHL titled "*Distributed and Decentralized Database Systems: Concepts and Challenges*".
9:45 – 10:45	Talk by ***Mr. Hari Kiran***, Founder, OpenSourceDB titled "*All you need is PostgreSQL*".
10:45 – 11:15	Break – Photo Session followed by Snacks
11:15 – 12:15	Talk by ***Mr. Hari Kiran*** continued.
12:15 – 1:30	Lunch
1:30 – 2:30	Talk by ***Mr. Suresh Chandra Dash***, Senior Solution Architect, Fujitsu titled "*Exploring PostgreSQL Internals: From Structure to Query Execution*".
2:30 – 3:30	Talk by ***Mr. R K N Sai Krishna***, Senior Software Engineer, Teradata, and Alumnus, DMACS, SSSIHL titled "*Unlocking Data Freedom: The Power of Apache Iceberg Open Table Format*".
3:30 – 3:40	*Vote of Thanks* by ***Dr. M S Srinath***, Assistant Professor, DMACS, SSSIHL, and The Convener and Coordinator of the workshop.
3:40 – 4:00	Snacks

Participants: The workshop's target audience comprised Master of Technology (Computer Science) students from the Department of Mathematics and Computer Science (DMACS) at Sri Sathya Sai Institute of Higher Learning (SSSIHL). The discussion focused on the internal workings of databases, aligning with the participants' academic background and interests.

Proceedings

The day's proceedings began with *Vedam* chanting by all the participants. Prof. Raghunatha Sarma R (Director, Prasanthi Nilayam Campus, SSSIHL, and former Head, DMACS, SSSIHL) then delivered the welcome address, highlighting the uniqueness of Sri Sathya Sai Institute of Higher Learning.

The welcome address was followed by a talk by Prof. Ravi Mukkamala (Professor and Chair, Department of Computer Science, Old Dominion University, and Adjunct Faculty, DMACS, SSSIHL) titled *"Distributed and Decentralized Database Systems: Concepts and Challenges."* In his talk, he covered various aspects of distributed and decentralized databases, including big data, distributed file systems, and partitioning. He also discussed different types of databases beyond relational databases, such as key-value stores, NoSQL databases, time-series databases, and graph databases.

Next, Mr. Hari Kiran (Founder, OpenSourceDB) addressed the gathering with a talk titled *"All You Need is PostgreSQL."* In the first part, he discussed the FOSS and COSS ecosystem, while the second part focused on PostgreSQL's internal architecture. He elaborated on various concepts, including different processes and their functionalities, the data dictionary layout, key features, replication types, and replication modes.

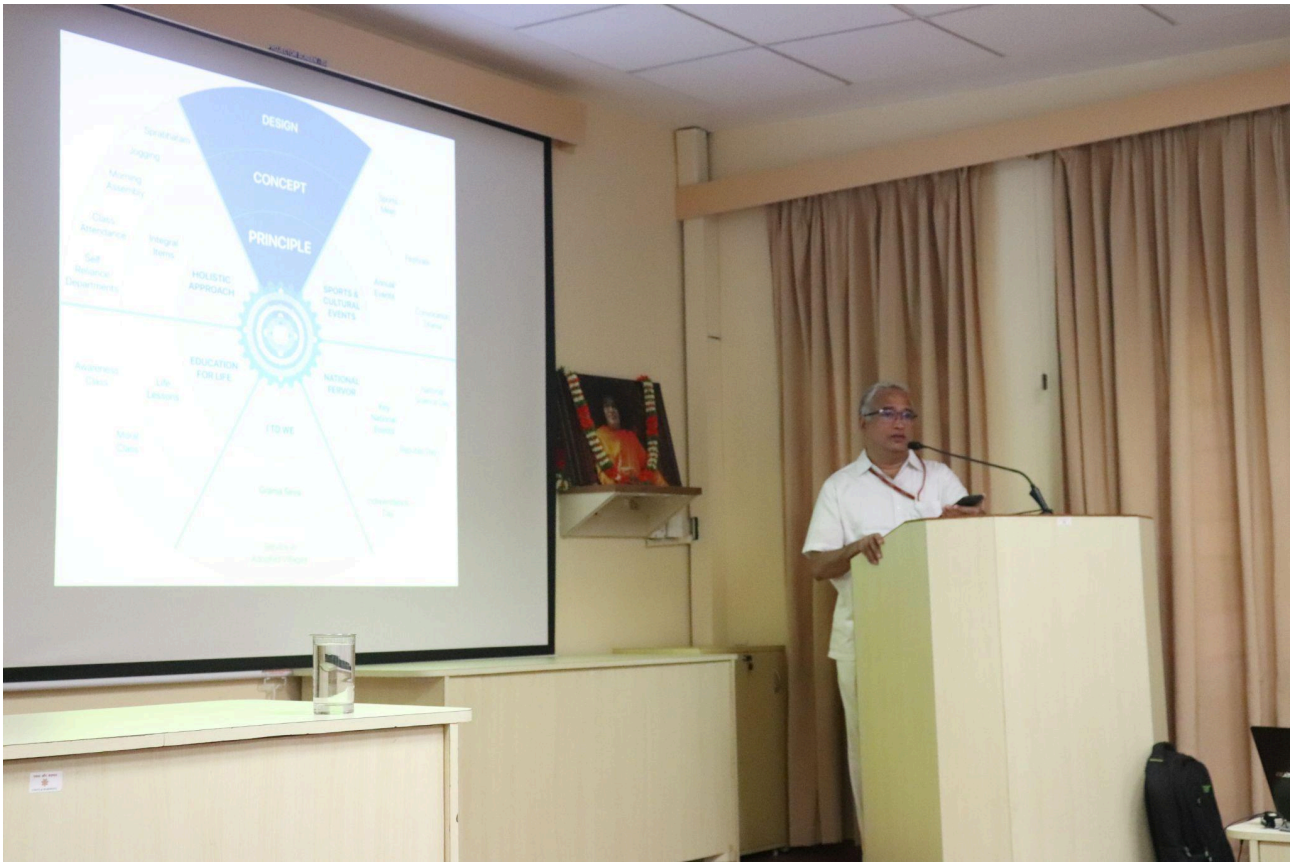
The afternoon session began with a talk by Mr. Suresh Dash (Senior Solution Architect, Fujitsu) titled *"Exploring PostgreSQL Internals: From Structure to Query Execution."* He provided insights into the internal structure and layout of PostgreSQL and elaborated on the various stages of query processing, including parsing, analyzing, rewriting, planning, and execution.

This was followed by a talk by Mr. R. K. N. Sai Krishna (Senior Staff Software Engineer, Teradata) titled *"Unlocking Data Freedom: The Power of Apache Iceberg Open Table Format."* He discussed current technological limitations and the need for data lakes, open file formats, and the Apache Iceberg open table format in a distributed environment.

The workshop concluded with a vote of thanks by Dr. Srinath M S (Assistant Professor, DMACS, SSSIHL), the Convener and Coordinator of the workshop.

Photos

Mastering Data Systems
From PostgreSQL Internals
to
Distributed Architectures
1st February 2025
DEPARTMENT OF MATHEMATICS
AND
COMPUTER SCIENCE
PRASHANTI NILAYAM CAMPUS
SRI SATHYA SAI INSTITUTE OF HIGHER LEARNING


DESIGN
CONCEPT
PRINCIPLE
HOLISTIC APPROACH
EDUCATION FOR LIFE
SPORTS & CULTURAL EVENTS
NATIONAL SERVICE
I TO WE

Distributed File Systems
• A distributed file system stores data across a large collection of machines, but provides single file-system view
• Highly scalable distributed file system for large data-intensive applications.
• E.g., 10K nodes, 100 million files, 10 PB
• Provides redundant storage of massive amounts of data on cheap and unreliable computers
• Files are replicated to handle hardware failure
• Detect failures and recovers from them

o Signal

Background - Data Lakes
• Definition:
• Data lakes are storage repositories that can hold vast amounts of raw data in its native format.
• Characteristics:
• Schema on Read: Data is stored in its raw form, and schema is applied when reading.
• Diverse Data Types: Can handle structured, semi-structured, and unstructured data.
• Scalability: Highly scalable and cost-effective.
• Challenges:
• Data Quality: Lack of data governance and quality control.
• Performance: Slower query performance due to unorganized data.
• Complexity: Requires advanced skills for managing and querying data.
GPS Map Camera
Puttaparthi, Andhra Pradesh, India
Vidyagiri, 5r77+32m Prasanthi Nilayam, Puttaparthi, Andhra Pradesh 515134, India
Lat 14.162188° Long 77.812746°
01/02/2025 03:22 PM GMT +05:30
Google

GPS Map Camera
Sri Sathya Sai Institute of Higher Learning, India
Prasanthi Nilayam, Puttaparthi, Andhra Pradesh 515134, India
Lat 14.162188° Long 77.812724°
01/02/25 08:56 AM GMT +05:30
Google

— Jai Sai Ram —