Building Database Systems Infrastructure for Performance and Reliability

Keynote speaker for MinervaDB Athena 2020

Shiv Iyer

Founder and Principal @ MinervaDB Inc.

WebScale Database Infrastructure Operations Experts

Speaker Bio

- Name: Shiv Iyer
- Profession: Open Source Database Systems guy for last 17 years
- Technology focus: MySQL, MariaDB, PostgreSQL and Linux
- Passion: Building WebScale platforms for performance and reliability
- My current role: Founder and Principal of MinervaDB Inc.
- Contacts:
  - Email: shiv@minervadb.com
  - LinkedIn: https://www.linkedin.com/in/thewebscaledba/
  - Twitter: https://twitter.com/thewebscaledba
  - Facebook: https://www.facebook.com/DataOpsGeek

MinervaDB Athena 2020 Introduction

- Vendor neutral and independent Open Source Database Systems Infrastructure Operations **virtual conference** hosted by MinervaDB Inc. addressing topics like performance, scalability, architecture / internals, data SRE, automation / DevOps. and production engineering / real-life Database Ops. stories
- Technology focus
  - MySQL
  - MariaDB
  - PostgreSQL
  - ClickHouse
  - NoSQL
  - Linux

Speakers of MinervaDB Athena 2020

Shiv Iyer
(MinervaDB Inc.)

Roy
(Sequoia Capital)

Valerii Kravchuk
(MariaDB)

Peter Zaitsev
(Percona)

Karthik.P.R
(Mydbops)

Colin Charles
(Galera Cluster)

Krunal Bauskar
(Huawei)

René Cannaò
(ProxySQL)

Tanel Poder
(Gluent)

Federico Razzoli
(Vettabase)

Jean-François Gagné
(HubSpot)

MinervaDB Athena 2020 Talks Schedule

Speaker Name	Topic	When
Shiv Iyer	Building Database Infrastructure for Performance and Reliability	09:00 AM PST to 10:00 AM PST
Anandamoy Roychowdhary	Role of Venture Capital Companies in building Database Systems Companies	10:00 AM PST to 11:00 AM PST
Valerii Kravchuk	Dynamic Tracing for Finding and Solving MySQL Performance Problems on Linux	11:00 AM PST to 12:00 PM PST
Peter Zaitsev	MySQL 8 Observability	12:00 PM PST to 01:00 PM PST
Karthik.P.R	MySQL Shell for Database Engineers	01:00 PM PST to 02:00 PM PST
Rene Cannao	ProxySQL Talks	02:00 PM PST to 03:00 PM PST
Tanel Poder	Profiling Linux Operations for Performance and Troubleshooting	03:00 PM PST to 04:00 PM PST
Federico Razzoli	MariaDB, MySQL and Ansible: automating database infrastructures	04:00 PM PST to 05:00 PM PST
Krunal Bauskar	Growing Open Source Database Systems ecosystem on ARM	05:00 PM PST to 06:00 PM PST
Jean-Francois Gagne	Demystifying MySQL Replication Crash Safety	06:00 PM PST to 07:00 PM PST
Colin Charles	WebScale Database Infrastructure Operations with Galera Cluster 4	07:00 PM PST to 08:00 PM PST

About this talk

- This talk is about functional and technical components involved when you are building Database Infrastructure for an high traffic internet property, It covers mostly how I have worked with some of the customers in past and present for scaling their Database Systems for performance and reliability
- The slides of this talk are intentionally kept vendor neutral and independent. But, I have shared scripts and tools used in MySQL, MariaDB and PostgreSQL Infrastructure Operations for troubleshooting the performance
- Greatly this talk help in functionally building your Database Infrastructure for WebScale addressing Performance and Reliability

Changing times in Database Systems Operations

- Database Infrastructure powers digital economy globally:
  - Internet powered Database Systems aggregates and integrates suppliers and consumers globally with their heavily funded digital assets:
    - Uber, The world's larges transport provider owns no vehicles
    - Facebook is the world's largest media company creates no content
    - Alibaba, The world's largest retailer owns no inventory
    - Airbnb, The world's largest accommodation and hospitality services providers own no hotels or real-estate
    - PayPal: The world's largest multi-disciplinary financial intermediatory services provider without any formal banking infrastructure
    - According to IBM, The global populace daily creates more than 2.5 quintillion bytes of data

What this means to Database Systems geeks ?

Data is the most value asset of digital economy and the database infrastructure scales business globally without compromising performance and reliability. This means, When you are building a Data Powered commerce the entire Database Infrastructure stakeholders are accountable for:

- ✓ Capacity Planning and Sizing
- ✓ Performance Optimization and Tuning
- ✓ Building Self-Healing and Fault-Tolerant Database Systems Infrastructure
- ✓ Data Privacy / Audits / Statutory Compliance
- ✓ Scalable (Horizontally and Vertically) Scalable Database Systems Platforms
- ✓ Accommodating Data – Transactional, Columnar and Eventually Consistent

Business impact on Capacity Planning and Sizing - Too big or too small is an challenging situation

Trouble of generous sizing

- Super confident stakeholders on system's endurance:
  - These systems has high frequency reliability issues or even expensive outages
  - Multi-purpose strategy:
    - Accommodating production, backup and archive data in the same infrastructure
  - Most often monitoring systems are not managed and alerts are ignored
- Reactive database performance management
  - Often SQL performance testing ignored
  - Indexes are often created everywhere, duplicated and most will be unused:
    - More indexes is a different problem to solve
  - SQL performance bottleneck also triggers extensive disk operations
    - More data to scan, process and housekeep - Higher cost of Data Ops.
- Will eventually lead to more expensive infrastructure procurement for **DATA**

How it all starts ?

- Single-Point-of-Failure (SPoF) Database Instances scaled vertically (heavy investments on processing power)
- Missing Capacity Planning / Sizing Documentation – **Too big or small**
- DR Plan – What will be missing in most of the times ?
  - How to Backup and Recovery checklist
  - Backup quality validation
  - Rollback plan, In case of a data corruption scenario
- Generic Monitoring Systems for Operating Systems and Database Systems Infrastructure – Everyone troubleshoot Database Systems
- Missing Database Systems Anomaly Documentation

Why no complaints about Database Systems Performance early days ?

- **ALL DATABASE SYSTEMS ARE OPTIMAL TILL DATA VOLUME GROW SIGNIFICANTLY**
- Database Systems are highly tolerant, They usually perform well even when SQLs are expensive and scale vertically very well
- In early days, Some data loss is acceptable
- Mostly no DBA so everyone is accountable for Database Systems Performance
- When more application code / SQLs gets in the complexity of transaction handling spikes significantly …. The first (early) sign of things are not cool, This will be ignored by many startups

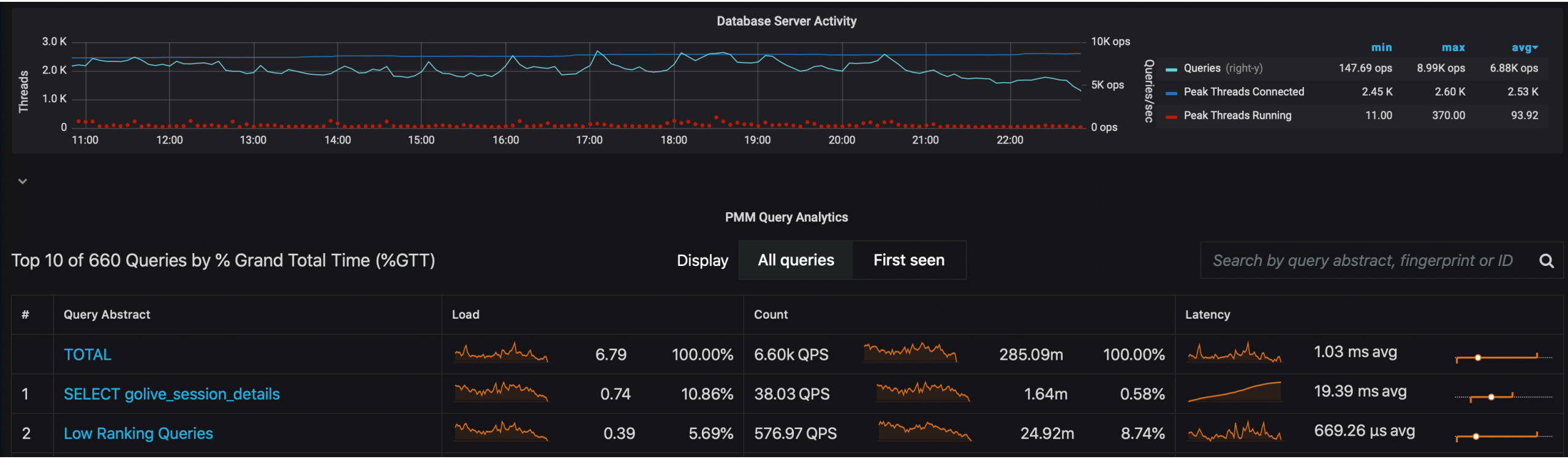
How complex is troubleshooting without DBA ecosystem?

- We believe DBA is a function than a individual / team, When everyone is DBA then recording with accountability of who did what is almost impossible. In such cases and extreme scenarios, We have rebuilt entire Database Infrastructure with fresh installation. This process is extremely complex and super expensive
- Database Systems Observability Infrastructure is completely different compare to Operating System or Network Operations Monitoring, Especially when you want to deep drill and find out queries by Response Time

How challenging is managing fast growing DB

- Capacity Planning and Sizing
- Performance
- Scalability (both horizontal and vertical)
- Building No-Single-Point-of-Failure Database Systems Infrastructure
- Zero Data Loss DR / Database SRE
- Proactive Troubleshooting: Time-Series Data of Performance Metrics
- Building Data Privacy / Security and Auditing Policies
- Hiring expert DBA staff is tough and retaining top talent is even more complex

Time-series Performance Monitoring Of Database Systems Infrastructure helps us to troubleshoot proactively



Considering Performance and Reliability

- Optimal Performance even when the Database Systems are distributed
- Proactive on database infrastructure health and performance
- Building Database Infrastructure for Performance and Reliability
- Can you rollback if there is data corruption beyond repair
- No single-point-of-failure in database infrastructure components
  - Redundant and distributed
  - self-healing and fault-tolerant
  - Multi-location backup retention

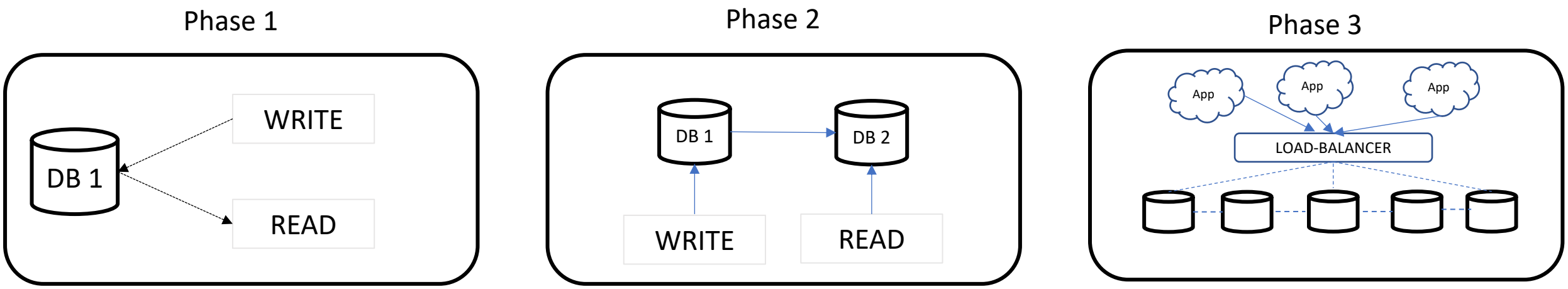
Database Infrastructure Operations Metrics expected to recorded for troubleshooting proactively

- Record slow queries
- Monitoring unused indexes
- Tables with full table scans
- Active sessions and queries triggered from them
- Monitoring deadlocks and queries
- **Time-series data of infrastructure usage from DB**

Building Database Infrastructure for Reliability

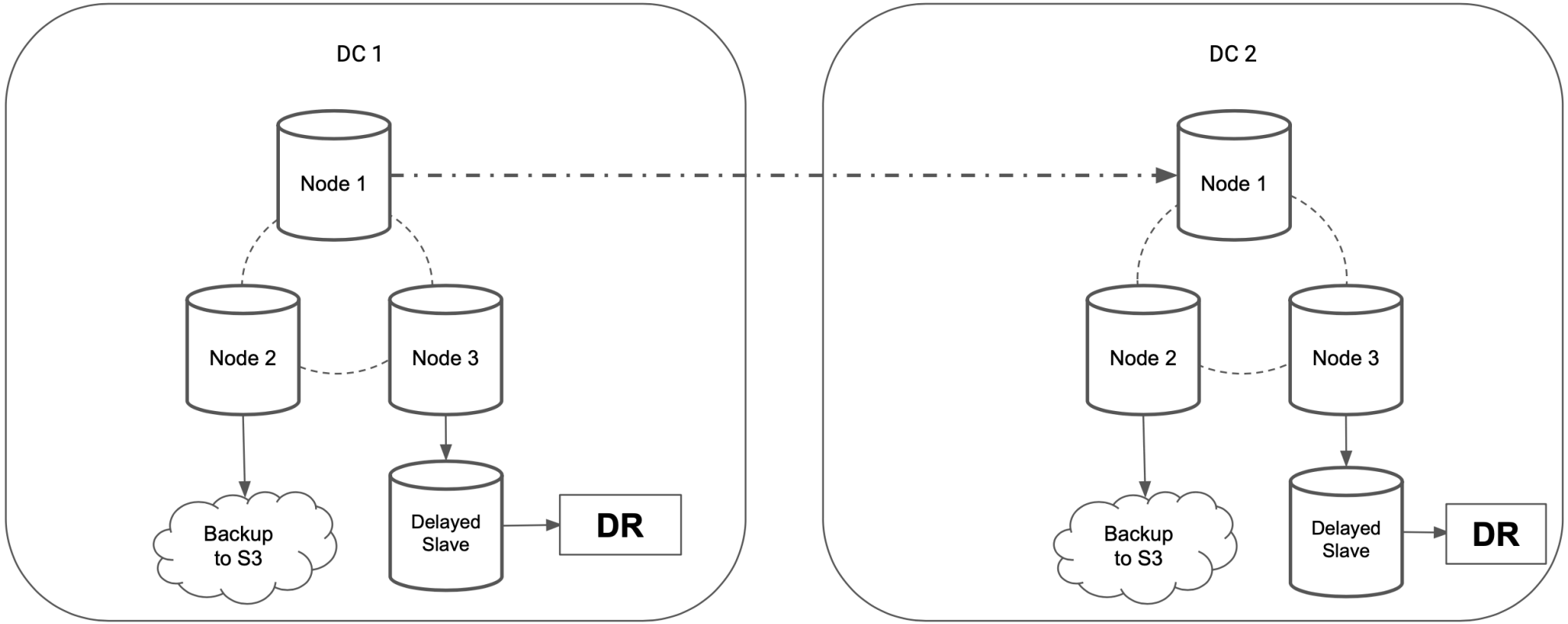
- What is the cost of Database Systems outage ?
  - ⁉️ Customer churn
  - 💰 Revenue loss
  - ✓ SLA to customers and partners
  - 🔧 Cost of Database Infrastructure Maintenance Operations
    - 1️⃣ Have you planned for downtime ? If yes, Do you have maintenance steps audited ?
    - 2️⃣ Last known good condition backup is recoverable for rollback when needed ?
    - 3️⃣ Do you have a failover Database Infrastructure available ?
    - 4️⃣ Do you have a Database Systems Support provider 24*7 ?
- Is your Database Infrastructure Globally Distributed ?
- DR Policies for both local and global Database Instances

How it grows from a Single Server DB Infra. ?



- ✓ Phase 1 – It's all about building Database Systems Operations for performance, Vertically scaled with huge infrastructure investments
- ✓ Phase 2 - READ : WRITE splitting for performance and reliability
- ✓ Phase 3 – Heavily distributed and replicated (in some cases across Data Centers or even cities / geographies), Need to address replication latency and load-balancing for ROI. Proactive detection and troubleshooting of replication infrastructure corruption scenarios

Multi-DC Database Infrastructure for Performance and Reliability – Operationally complex



Transaction level awareness

- The following are major reasons why troubleshooting larger Database Systems is complex, resource intensive and super expensive:
  - Blame it on Data Volume
  - Longer Data Retention in Primary Database Infrastructure
    - Do business really need data older than 12 months in Primary Database Infrastructure ?
    - How much data we can afford to move on soft archive infra. , The data is available on-demand
    - Performance SLA on queries
      - Primary Database Systems Infrastructure
      - Soft Archive Database Systems Infrastructure
  - Transaction Audit to isolate expensive and rogue SQLs
  - Index Life Cycle Management – Who is approving index creation and what is the retirement plans for indexes

Indexing – Do we have access to regularly audit index usage ?

- Largest indexes by size
- List of most often used indexes
- Unused indexes
- Indexes causing maximum disk I/O
- In MySQL 8 Histograms can greatly influence optimizer to choose cost efficient and optimal query plan. **THE BIGGEST ADVANTAGE OF USING MYSQL 8 HISTOGRAMS COMPARED TO INDEX IS THEY ARE LESS EXPENSIVE, THERE WILL BE NO CHANGE TO HISTOGRAM WHEN YOU MANIPULATE THE DATA**
- **In PostgreSQL 13 you can vacuum indexes in parallel, This means much better PostgreSQL Maintenance Operations for performance**

Troubleshooting Performance with slow logs

- Most preferred Database Systems Performance Troubleshooting method is slow logs for following reasons:
  - If you are looking out for epicenter of Database Systems Performance Bottleneck from the queries causing it then it's slow log is a right pick.
  - When you are interpreting performance by "Response Time" then from slow log you can list top queries by latency. MySQL, MariaDB and PostgreSQL have implemented this feature very gracefully:
    - **MySQL and MariaDB** – slow_query_log
    - **PostgreSQL** – log_min_duration_statement

Note: We strongly recommend to configure slow log carefully, There is no fun in logging more information than you really want to interpret, More evidence collection is also sometimes root cause of performance bottleneck due to the extra load for disk I/O and also larger file to annotate is expensive.

Transaction Audit for Performance and Compliance

- What happens if Data Ops. of Internet companies are not serious about information privacy and security:
  - Data theft, identity robbery, digital terrorism, ransomware, expect more worst beyond all these
  - Digital economy and related infrastructure will collapse
  - Connected world will become unreliable
- Who can help us here to scale secured digital infrastructure for global equality and transparency ?
  - Database SRE, DBAs, Database Architects / Engineers, Database Infrastructure Security Experts, SREs, Production Engineers, Infrastructure Ops. Engineers

Secured Database Infrastructure Operations

- Password management:
  - Strong password rule enforcement
  - Password policies and retention audits
  - How are passwords maintained in multi-tenant Database Applications ?
  - Data Obfuscation for privacy protection
- Data Access Policies and Profiles / Roles
- Transaction Audit:
  - All the successful and failed transactions should be recorded to even session level
  - PLEASE DO NOT CONSIDER REPLACING AUDIT LOGS WITH SLOW QUERY LOGS

Data Reliability

- DR Functional Architecture Review and Recommendations (audit and update policy with dates)
- Backup policy
  - ✓ Daily – How often ?
  - ✓ Weekly
  - ✓ Monthly
  - ✓ Logical Backup – scripts, frequency and audit log report
  - ✓ Physical Backup - scripts, schedule and validation report
  - ✓ Backup validation – date –time, restoration infrastructure details and time to restore / recover
  - ✓ Backup retention policy – on-premises, remote and cloud
  - ✓ Backup archiving and purging
  - ✓ Recent database outage and detailed report of restoration scripts used including time to restore

Complexity with Data Infrastructure of future

- Heterogeneous Data Platforms
  - Transaction Processing Systems
  - Eventually Consistent Systems
  - Columnar Stores exclusively for SORT / SEARCH intensive queries
  - New SQL addressing best of both eventually consistent and ACID
- Heterogenous Database Infrastructure Operations
  - On-premises
  - Cloud
  - Bare metal
  - Containers
- Data Ops @ Scale – AI to fuel performance troubleshooting of planet-scale database infrastructure intuitively

Thank you !

We are hiring @ MinervaDB, Please contact me on shiv@minervadb.com to know more about opportunities in MinervaDB and you can also visit our Jobs page here - https://minervadb.xyz/careers/