MinervaDB

Data SRE

#Monitoring #Sizing #Performance #Availability

MinervaDB

Speaker Bio

- Name: Shiv Iyer
- Occupation: Founder and Principal of MinervaDB
- Years of experience in Database Infrastructure Operations: 18 years

Technology focus:

- MySQL and InnoDB
- MariaDB
- PostgreSQL
- ClickHouse
- Building database infrastructure for performance, scalability and reliability

MinervaDB

Building Systems for Reliability

- Business impact if systems are not sized optimally
- How performance matters in data driven planet
- Proactive on database infrastructure health and performance
- High endurance database systems
- Database infrastructure operations performance visibility
- No single-point-of-failure in database infrastructure components
  - Redundant and distributed
  - self-healing and fault-tolerant
  - Multi-location backup retention
- Emergency outage troubleshooting checklist and run-books

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

MinervaDB

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**

Data Ops Performance Matters..

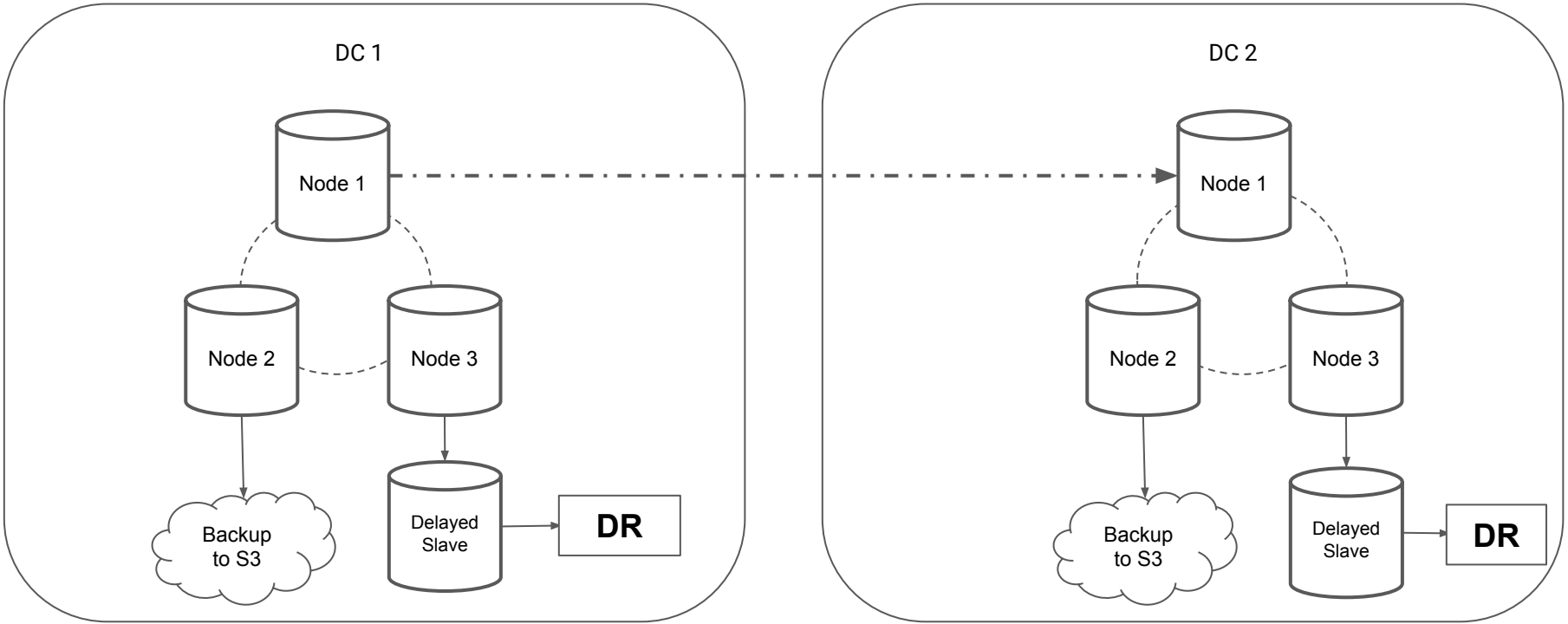
- Configuring Linux for Database Performance
- Optimal installation and configuration
  - Not all the Database Systems variables / configuration parameters benefit in overall performance by setting too high always
    - There are per instance and per session / thread variables in MySQL, MariaDB and PostgreSQL
- Optimal SQL:
  - No SELECT * FROM queries
  - Query data only needed, In some cases trimming columns in queries by 10% improves performance by 30%
  - Index when needed and remove unused indexes
- Data grows - Archive and partition data for query performance

Database Infrastructure Observability and Resilience - Monitoring performance by Response Time and Throughput

- Expensive SQLs by latency (response time) and throughput (system resource usage)
- SQL execution plan analysis, Data Access Path profiling and index usage:
  - Cost of Data Access Path
  - Monitoring indexes:
    - Missing indexes
    - Redundant / duplicate indexes
    - unused indexes
- Monitoring infrastructure usage by Database System
  - Expensive SQLs by CPU usage
  - Disk to Memory ratio analysis - It's great if you can fit entire DB in the memory
  - Distribute READ / WRITE for Disk I/O performance and reliability
- Connection handling and threads performance

Building Database Infrastructure Operations for availability and reliability - Distributed and Redundant

MinervaDB


DC 1
Node 1
Node 2
Node 3
Backup
to S3
Delayed
Slave
DR
DC 2
Node 1
Node 2
Node 3
Backup
to S3
Delayed
Slave
DR

MinervaDB

Data Ops. Checklist

- How to change system variables for performance and reliability ?
  - Guidelines, workflows and approval
  - Documentation and next steps
- Monitoring Data Ops. charts to troubleshoot performance proactively
  - Time-series query performance (latency) monitoring - **from peaks to off-peaks**
  - Time-series charts to measure throughput against latency - **from peaks to off-peaks**
- Troubleshooting performance using logs
  - Error log
  - Slow query log
  - Audit log
- Monitoring Replication
- Monitoring Backups

MinervaDB

Data Ops. run-book

- Changing system variables and configuration parameters
- Backup / DR Ops. automation scripts and validation process
- Interpreting the error log, report bugs and upgrades / migration
- Troubleshooting Data Ops. performance:
  - Query performance by latency / response-time
  - Top 5 queries by throughput / system resource usage
  - Scripts to monitor Index usage:
    - Missing indexes
    - Redundant and duplicate indexes
    - unused indexes
- Troubleshooting Replication infrastructure - performance and consistency
- Run-book to archive and purge the database

MinervaDB

How can you contact me ?

Email - shiv@minervadb.com

Twitter - https://twitter.com/thewebscaledba

LinkedIn - https://www.linkedin.com/in/thewebscaledba/

Phone / MinervaDB Toll Free - (844) 588-7287