Log Ingestor

A Logs Ingestion and Analysis System for Distributed Systems

Project - IT214 - Database Management System
Prof. P M Jat

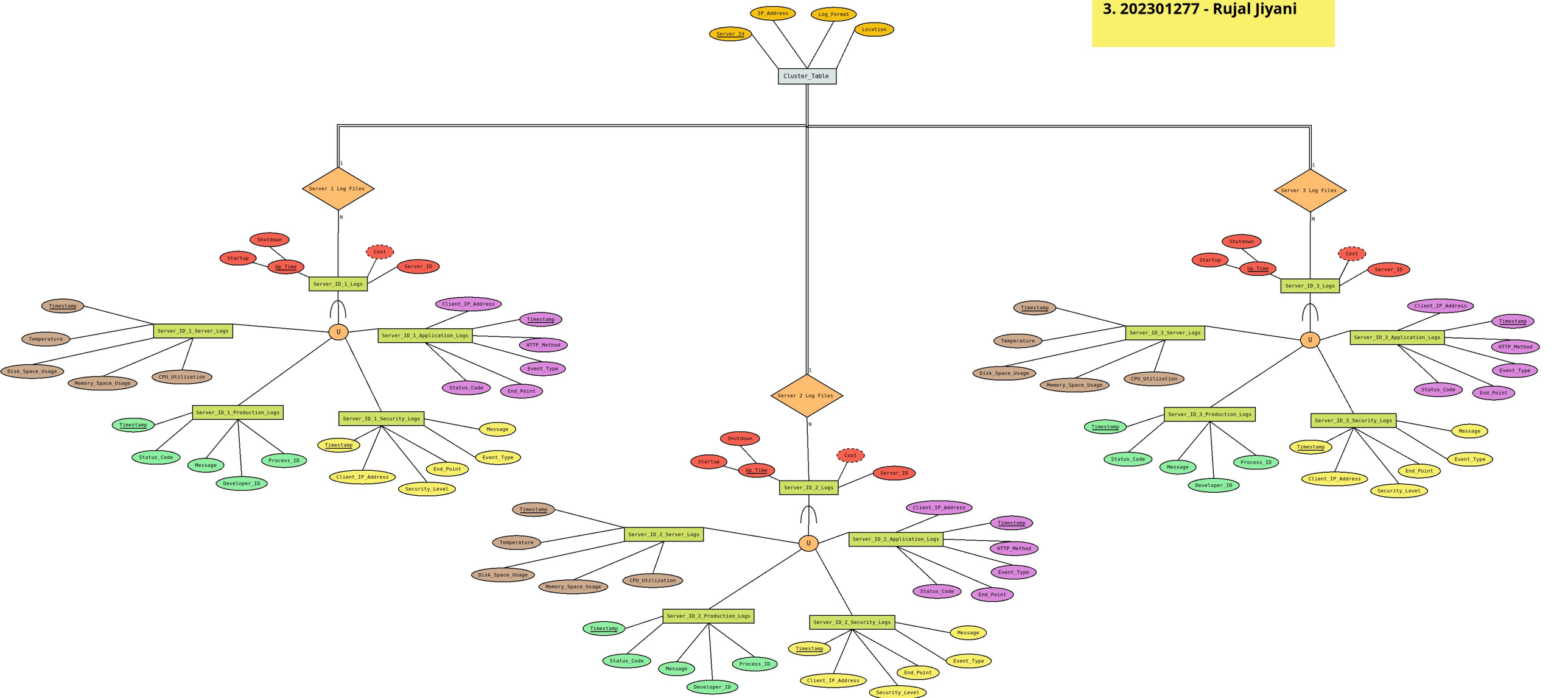
Group Members

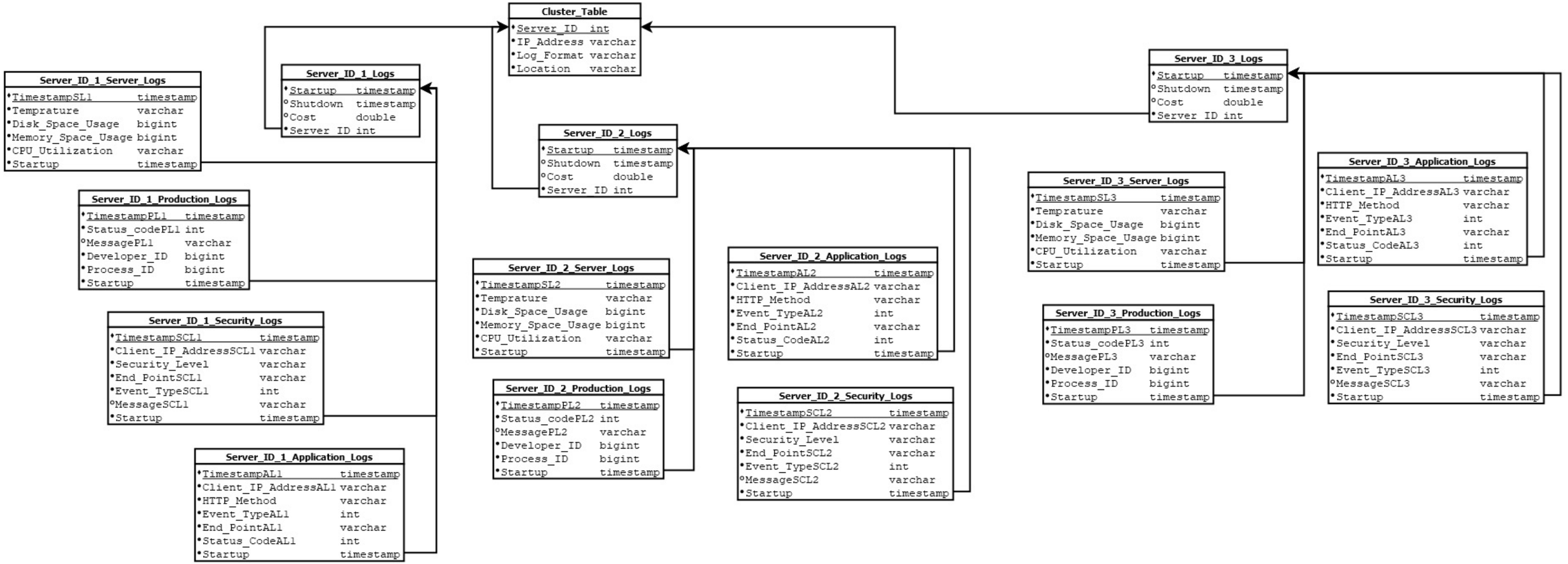
1. Dhruv Jain - 202301272
2. Jevik Rakholiya - 202301276
3. Rujal Jiyani - 202301277

G5-T4

LOG INGESTOR - Entity Relationship Diagram

1. 202301276 - Jevik Rakholiya
2. 202301272 - Dhruv Jain
3. 202301277 - Rujal Jiyani




Cluster_Table
Server_ID int
IP_Address varchar
Log_Format varchar
Location varchar
Server_ID_1_Server_Logs
TimestampSL1 timestamp
Temprature varchar
Disk_Space_Usage bigint
Memory_Space_Usage bigint
CPU_Utilization varchar
Startup timestamp
Server_ID_1_Logs
Startup timestamp
Shutdown timestamp
Cost double
Server_ID int
Server_ID_2_Logs
Startup timestamp
Shutdown timestamp
Cost double
Server_ID int
Server_ID_3_Logs
Startup timestamp
Shutdown timestamp
Cost double
Server_ID int
Server_ID_1_Production_Logs
TimestampPL1 timestamp
Status_codePL1 int
MessagePL1 varchar
Developer_ID bigint
Process_ID bigint
Startup timestamp
Server_ID_1_Security_Logs
TimestampSCL1 timestamp
Client_IP_AddressSCL1 varchar
Security_Level varchar
End_PointSCL1 varchar
Event_TypeSCL1 int
MessageSCL1 varchar
Startup timestamp
Server_ID_1_Application_Logs
TimestampAL1 timestamp
Client_IP_AddressAL1 varchar
HTTP_Method varchar
Event_TypeAL1 int
End_PointAL1 varchar
Status_CodeAL1 int
Startup timestamp
Server_ID_2_Server_Logs
TimestampSL2 timestamp
Temprature varchar
Disk_Space_Usage bigint
Memory_Space_Usage bigint
CPU_Utilization varchar
Startup timestamp
Server_ID_2_Application_Logs
TimestampAL2 timestamp
Client_IP_AddressAL2 varchar
HTTP_Method varchar
Event_TypeAL2 int
End_PointAL2 varchar
Status_CodeAL2 int
Startup timestamp
Server_ID_2_Production_Logs
TimestampPL2 timestamp
Status_codePL2 int
MessagePL2 varchar
Developer_ID bigint
Process_ID bigint
Startup timestamp
Server_ID_2_Security_Logs
TimestampSCL2 timestamp
Client_IP_AddressSCL2 varchar
Security_Level varchar
End_PointSCL2 varchar
Event_TypeSCL2 int
MessageSCL2 varchar
Startup timestamp
Server_ID_3_Server_Logs
TimestampSL3 timestamp
Temprature varchar
Disk_Space_Usage bigint
Memory_Space_Usage bigint
CPU_Utilization varchar
Startup timestamp
Server_ID_3_Application_Logs
TimestampAL3 timestamp
Client_IP_AddressAL3 varchar
HTTP_Method varchar
Event_TypeAL3 int
End_PointAL3 varchar
Status_CodeAL3 int
Startup timestamp
Server_ID_3_Production_Logs
TimestampPL3 timestamp
Status_codePL3 int
MessagePL3 varchar
Developer_ID bigint
Process_ID bigint
Startup timestamp
Server_ID_3_Security_Logs
TimestampSCL3 timestamp
Client_IP_AddressSCL3 varchar
Security_Level varchar
End_PointSCL3 varchar
Event_TypeSCL3 int
MessageSCL3 varchar
Startup timestamp

Projected and Minimal Functional Dependency

Cluster_Table(Server_ID, IP_Address, Log_Format, Location)

Projected Functional Dependency Set :

- Server_ID → IP_Address, Log_Format, Location

Minimal Functional Dependency Set :

- Server_ID → IP_Address
- Server_ID → Log_Format
- Server_ID → Location

Server_ID_X_Logs(Startup, Shutdown, Cost, Server_ID) (X = 1,2,3,..)

Projected Functional Dependency Set :

- Startup → Shutdown, Cost, Server_ID

Minimal Functional Dependency Set :

- Startup → Shutdown
- Startup → Cost
- Startup → Server_ID

Server_ID_X_Server_Logs(TimestampSLX, Temperature, Disk_Space_Usage, Memory_Space_Usage, CPU_Utilization, Startup)

Projected Functional Dependency Set :

- TimestampSLX → Temperature, Disk_Space_Usage, Memory_Space_Usage, CPU_Utilization, Startup

Minimal Functional Dependency Set :

- TimestampSLX → Temperature
- TimestampSLX → Disk_Space_Usage
- TimestampSLX → Memory_Space_Usage
- TimestampSLX → CPU_Utilization
- TimestampSLX → Startup

Server_ID_X_Production_Logs(TimestampPLX, Status_codePLX, MessagePLX, Developer_ID, Process_ID, Startup)

Projected Functional Dependency Set :

- TimestampPLX → Status_codePLX, MessagePLX, Developer_ID, Process_ID, Startup

Minimal Functional Dependency Set :

- TimestampPLX → Status_codePLX
- TimestampPLX → MessagePLX
- TimestampPLX → Developer_ID
- TimestampPLX → Process_ID
- TimestampPLX → Startup

Server_ID_X_Security_Logs(TimestampSCLX, Client_IP_AddressSCLX, Security_Level, End_PointSCLX, Event_TypeSCLX, MessageSCLX, Startup)

Projected Functional Dependency Set :

- TimestampSCLX → Client_IP_AddressSCLX, Security_Level, End_PointSCLX, Event_TypeSCLX, MessageSCLX, Startup

Minimal Functional Dependency Set :

- TimestampSCLX → Client_IP_AddressSCLX
- TimestampSCLX → Security_Level
- TimestampSCLX → End_PointSCLX
- TimestampSCLX → Event_TypeSCLX
- TimestampSCLX → MessageSCLX
- TimestampSCLX → Startup

Server_ID_X_Application_Logs(TimestampALX, Client_IP_AddressALX, HTTP_Method, Event_TypeALX, End_PointALX, Status_CodeALX, Startup)

Projected Functional Dependency Set :

- TimestampALX → Client_IP_AddressALX, HTTP_Method, Event_TypeALX, End_PointALX, Status_CodeALX, Startup

Minimal Functional Dependency Set :

- TimestampALX → Client_IP_AddressALX
- TimestampALX → HTTP_Method
- TimestampALX → Event_TypeALX
- TimestampALX → End_PointALX
- TimestampALX → Status_CodeALX
- TimestampALX → Startup

BCNF Proof for Relational Schema

To prove that the given relational schema is in Boyce-Codd Normal Form (BCNF), we verify that for every relation, all non-trivial functional dependencies (FDs) have a superkey as their left-hand side (LHS).

Cluster_Table(Server_ID, IP_Address, Log_Format, Location)

- Primary Key: Server_ID
- FDs: Server_ID → IP_Address, Log_Format, Location

Since Server_ID is a candidate key, BCNF condition is satisfied.

Server_ID_X_Logs(Startup, Shutdown, Cost, Server_ID) (X = 1,2,3,..)

- Composite Primary Key: Startup
- FDs: Startup → Server_ID, Shutdown, Cost

LHS is a candidate key, so BCNF condition is satisfied.

Server_ID_X_Server_Logs(TimestampSLX, Temperature, Disk_Space_Usage, Memory_Space_Usage, CPU_Utilization, Startup)

- Primary Key: TimestampSLX
- FDs: TimestampSLX → All other attributes

TimestampSLX is a candidate key, satisfying BCNF.

Server_ID_X_Production_Logs(TimestampPLX, Status_codePLX, MessagePLX, Developer_ID, Process_ID, Startup)

- Primary Key: TimestampPLX
- FDs: TimestampPLX → All other attributes

BCNF condition is satisfied.

Server_ID_X_Security_Logs(TimestampSCLX, Client_IP_AddressSCLX, Security_Level, End_PointSCLX, Event_TypeSCLX, MessageSCLX, Startup)

- Primary Key: TimestampSCLX
- FDs: TimestampSCLX → All other attributes

BCNF condition is satisfied.

Server_ID_X_Application_Logs(TimestampALX, Client_IP_AddressALX, HTTP_Method, Event_TypeALX, End_PointALX, Status_CodeALX, Startup)

- Primary Key: TimestampALX
- FDs: TimestampALX → All other attributes

TimestampALX is a candidate key, so BCNF is satisfied.

Conclusion:

All relations in the schema have functional dependencies where the left-hand side is a superkey. Therefore, the schema is in Boyce-Codd Normal Form (BCNF).