# Interview Assignment for FiveTran: Technical Support Engineer

Database and tables were created in **Postgres SQL** for testing purposes. **PopSQL** client application was used to query database.

## DDL Create Tables

**Table - ManagerDetails**

| MANAGER_ID | NAME | DIRECT_REPORTS | ANNUAL_BUDGET |
|---|---|---|---|
| 203 | Ned Stark | 1 | 100000 |
| 202 | Walter White | 1 | 100000 |
| 201 | Leia Organa | 3 | 200000 |

```
Create TABLE ManagerDetails (
    MANAGER_ID INT PRIMARY KEY,
    "NAME" VARCHAR(50),
    DIRECT_REPORTS INT,
    ANNUAL_BUDGET INT
);
INSERT INTO ManagerDetails VALUES (203, 'Ned Stark', 1, 100000);
INSERT INTO ManagerDetails VALUES (202, 'Walter White', 1, 100000);
INSERT INTO ManagerDetails VALUES (201, 'Leia Organa', 3, 200000);
```

**Table - EmployeeSalary**

| EMP_ID | PROJECT | SALARY | Bonus |
|---|---|---|---|
| 101 | P1 | 8000 | 500 |
| 102 | P2 | 10000 | 1000 |
| 103 | P3 | 12000 | 0 |
| 104 | P1 | 20000 | 2000 |
| 105 | P2 | 15000 | 1000 |
| 106 | P0 | 25000 | 3000 |

```
CREATE TABLE EmployeeSalary (
    EMP_ID INT PRIMARY KEY,
    PROJECT VARCHAR(2),
    SALARY INT,
    BONUS INT
);
INSERT INTO EmployeeSalary VALUES (101, 'P1', 8000, 500);
INSERT INTO EmployeeSalary VALUES (102, 'P2', 10000, 1000);
INSERT INTO EmployeeSalary VALUES (103, 'P3', 12000, 0);
INSERT INTO EmployeeSalary VALUES (104, 'P1', 20000, 2000);
INSERT INTO EmployeeSalary VALUES (105, 'P2', 15000, 1000);
INSERT INTO EmployeeSalary VALUES (106, 'P0', 25000, 3000);
```

**Table - EmployeeDetails**

| EMP_ID | FULL_NAME | MANAGER_ID | JOIN_DATE | TEAM |
|--------|-----------|------------|-----------|------|
| 101 | John Snow | 203 | 1/31/2019 | AMER |
| 102 | Jesse Pinkman | 202 | 1/30/2020 | EMEA |
| 103 | Luke Skywalker | 201 | 11/27/2018 | APAC |
| 104 | Ned Stark | 201 | 7/10/2016 | EMEA |
| 105 | Walter White | 201 | 8/24/2017 | APAC |
| 106 | Leia Organa | | 7/10/2015 | AMER |

```
CREATE TABLE EmployeeDetails (
    EMP_ID INT PRIMARY KEY,
    FULL_NAME VARCHAR(50),
    MANAGER_ID INT,
    JOIN_DATE DATE,
    TEAM VARCHAR(10),
    FOREIGN KEY(MANAGER_ID) REFERENCES ManagerDetails(MANAGER_ID) ON DELETE SET NULL
);
INSERT INTO EmployeeDetails VALUES (101, 'John Snow', 203, '1/31/2019', 'AMER');
INSERT INTO EmployeeDetails VALUES (102, 'Jesse Pinkman', 202, '1/30/2020', 'EMEA');
INSERT INTO EmployeeDetails VALUES (103, 'Luke Skywalker', 201, '11/27/2018', 'APAC');
INSERT INTO EmployeeDetails VALUES (104, 'Ned Stark', 201, '7/10/2016', 'EMEA');
INSERT INTO EmployeeDetails VALUES (105, 'Walter White', 201, '8/24/2017', 'APAC');
INSERT INTO EmployeeDetails VALUES (106, 'Leia Organa', NULL, '7/15/2015', 'AMER');
```

# DML Query Challenges

1. You are approached by your boss and he asks you to provide him a list of all the employees who report to Walter White. Please provide the SQL query you would use to display this information.

```
SELECT
    e.emp_id,
    e.full_name,
    e.manager_id,
    e.join_date,
    e.team
FROM
    employeedetails as e
    LEFT OUTER JOIN managerdetails as m
    ON (e.manager_id = m.manager_id)
WHERE
    m.name = 'Walter White';
```

2. Write an SQL query to add Thomas Anderson with the Employee ID of 131 to the EmployeeDetails table. He joined the Oakland team on April 1, 2020, and his manager's name is Ned Stark.

```
INSERT INTO employeedetails
VALUES (
    131,
    'Thomas Anderson',
    NULL,
    '4/1/2020',
    'Oakland'
);

UPDATE employeedetails
SET manager_id = (
    SELECT m.manager_id
    FROM managerdetails as m
    WHERE m.name = 'Ned Stark'
    LIMIT 1
)
WHERE emp_id = 131;
```

3. Write a single query to display all employee names and manager names.

```
SELECT e.full_name as employee_name, m.name as manager_name
FROM
    employeedetails as e
    LEFT OUTER JOIN managerdetails as m
    ON (e.manager_id = m.manager_id);
```

4. Write a SQL query to calculate the average salary for all employees where the manager is Leia Organa.

```
SELECT AVG(s.salary) as Average_Salary_For_Employees_Under_Leia_Organa
FROM
    employeedetails as e
    INNER JOIN managerdetails as m
    ON (e.manager_id = m.manager_id)
    INNER JOIN employeesalary as s
    ON (e.emp_id = s.emp_id)
WHERE m.name = 'Leia Organa';
```

5. Write a SQL query to calculate the median total compensation per region where bonuses are always part of the compensation package

```
SELECT
  e.team,
  percentile_cont(0.5) WITHIN GROUP
  (ORDER BY (s.salary + s.bonus)) as median_total_compensation
FROM
    employeedetails as e
    INNER JOIN employeesalary as s
    ON (e.emp_id = s.emp_id)
GROUP BY e.team;
```

6. Write a SQL query to show the percentage of Leia Organa's annual budget that is used for employee compensation.

   **Note:** I assumed Leia Organa's compensation is not included in her annual budget.

```
SELECT
  SUM(1.0*(COALESCE(s.salary,0) + COALESCE(s.bonus, 0))/m.annual_budget)
FROM
    employeedetails as e
    LEFT OUTER JOIN managerdetails as m
    ON (e.manager_id = m.manager_id)
    INNER JOIN employeesalary as s
    ON (e.emp_id = s.emp_id)
WHERE m.name = 'Leia Organa';
```

# Database Architecture

Answer the following using fivetran docs, knowledge base, and internet searches. Respond to teach as if you were responding to a customer through a ticketing system.

1. "My company is looking to add a Postgres connector. We are unsure if we should choose WAL or XMIN replication. What is your suggestion?"

```
Hi John Doe,

In most scenarios, we recommend the WAL method. You should only consider
XMIN for updating records when WAL is not available.

Currently, you can use WAL method with either pgoutput or test_decode plugins.
You can see which plugins are supported for your Postgres instance here:
https://fivetran.com/docs/databases/postgresql#supportedservices.

Read more about WAL logical replciation and XMIN here:
https://fivetran.com/docs/databases/postgresql#updatingdata

Please reach out if you have further questions. I'm happy to help.

Best Regards,
Zaw Mai
```

2. "I've added a new Salesforce field, but I don't see it listed in the schema tab. Can you investigate why?"

```
Hi John Doe,

This can be caused one or more of the following:

1. Make sure fivetran has the right access permission for objects in Saleforce.
2. Review the synced field is not a formula or compound field types.
3. Review the table/field name from saleforce matches table/field name in fivetran.
4. Make sure there are no duplciate column fields or table with the same name.

Read more about syncing Salesforce fields here:
https://support.fivetran.com/hc/en-us/articles/360052347794

Please note you can sync formula fields but there's no gurantee of accurate
udpate because formula fields don't update when the source formula changes.
Read more here: https://fivetran.com/docs/applications/salesforce/formula

Please reach out if you have further questions. I'm happy to help.

Best Regards,
Zaw Mai
```

3. "What ETL/ELT steps does Fivetran take to move my data from my source database to my destination warehouse?"

Hi John Doe,

Here's general steps overview of our ETL/ELT process of moving the source data
to the destination warehouse.

1.  Fivetran secures an encrypted connection to the one or more source
    databases by using a pull or push connectors. Read more about connectors
    here: https://support.fivetran.com/hc/en-us/articles/1500003152601#connect

2.  Then, our core engine normalizes, cleans, and de-duplicates the data
    records to provide a optimal data format for the destination warehouse.

3.  Next, the formatted data is written to a temporary data store, waiting
    to be processed within 24 hours.

4.  Finally, Fivetran encrypts and writes the data to the destination data
    warehouse chunk by chunks. A success or a failed status will be sent via
    email or you can check on the fivetran dashboard.

Read more about Fivetran's ETL/ELT high-level process overview here:
https://fivetran.com/docs/getting-started/architecture

Each data source and destinatation warehouse requires application-specific
prequisites before Fivetran can start syncing, replicating, and updating data.
Read more about setting up your source and destination applications here:
https://fivetran.com/docs/applications

Please reach out if you have further questions. I'm happy to help.

Best Regards,
Zaw Mai

# Troubleshooting

## Challenge 1 - Logs / Monitoring

You are tasked to analyze the following log file.

```
 1   2018-10-11   200   192.168.11.1   test
 2   2019-10-04   201   192.168.11.1   test
 4   2018-08-11   302   212.168.11.1   stage
 7   2019-02-11   400   192.168.11.1   ci
10   2017-10-11   403   192.167.11.1   prod
11   2019-10-11   500   192.168.1.1    test
```

Using any method you like. Write the specific step by step instructions to execute the following tasks:

- Extract all IP addresses into a single column.
- Count unique IP addresses

**Answers below:**

Used bash scripts. Tested on Ubuntu 20.04 linux distro.

```
#!/bin/bash

# extract all IP Addresses into a single column
awk '{print $4}' log.txt

# count unique IP addresses
awk '{print $4}' log.txt | uniq -u | wc -l
```

# Challenge 2 - API GET

API GET challenge, have the candidate provide us with the results of a GET request.

1. Sign up for a free Airtable account: https://airtable.com/
2. Navigate to the Airtable API documentation: https://airtable.com/api
3. Provide a working API call to pull all names from the project tracker table using curl headers for authentication.

**Answers below:**

Used "curl" command-line tool to make API GET request.
Used "jq" command-line tool to parse json response body. Additional guide here.
Tested on Ubuntu 20.04 linux distro.
Tested on Airtable's prebuilt Project Tracker template.

```
curl https://api.airtable.com/v0/app9S4NyrBMAWPLHM/Design%20projects?fields%5B%5D=Name -
```

**Results:**

```
Lemon headband
Coffee packaging
Convertible 3000 laptop
RITI media lab logo
New Door brand identity
Premier utility bike
HGH injection device
CubePad
CMCA brand identity
Locax notebook computer
B11 bike saddle
Hand hygiene system
EngineerU brand identity
Flapper brand identity
443 Huntington brand identity
Gotham City Parks brand identity
```

# Challenge 3 - Networking

1. Name a command to verify IP to IP connectivity. Give an example of using this in your command line.
   - "ping" is the CLI tool to verify IP to IP connectivity.

     ```
     ping -c4 -4 google.com
     ```

2. Explain the difference between ping and telnet.
   - Both can be used on the command-line.
   - Ping checkes if two machines are connected. It uses the Internet Control Message Protocol (ICMP) to get an "echo" or a reply from the remote machine. Ping works on the OSI Layer 3: Network/Internet layer.
   - Telnet is a server application tool to talk to remote comptuers. It uses Transmission Control Protocol/Internet Protocol (TCP/IP) to send and recieves commands via TCP ports (usually 23), and make remote user session. Telnet works on the OSI Layer 4: Transport Layer.
3. What would it mean if a ping to 1.1.1.1 is successful, but telnet 1.1.1.1 80 fails?
   - Ping: your local machine can reach the machine on network IP address 1.1.1.1.
   - Telnet: your local machine can send server application commands on TCP Port 80 with machine network IP address of 1.1.1.1.

**References:**

https://www.quora.com/What-is-the-difference-between-Telnet-and-Ping
https://www.ibm.com/docs/en/zos/2.2.0?topic=internets-tcpip-tcp-udp-ip-protocols
https://www.geeksforgeeks.org/layers-of-osi-model/

# Communications

## Challenge 1

We received the following message from a customer on a new ticket and the ticket has been assigned to you:

```
Hello,

All ten of our connectors are failing to sync.  Please help!

Thanks,
John Doe
```

1. What troubleshooting steps would you take with a general question like this?
   - If I have access to their account, I would take a quick look at how their connecters are configured.
   - If all ten of the connectors are failing, then I will try to find non-application specific causes for sync failures before diving into application-specific setup causes.
   - Then, I will search through the FAQs and Knowledge Base to see if there's common "gotchas" of sync failuer to review. I will relay the troubleshooting guide to the customer.
   - When the customer fails to troubleshoot on his or her own, I will go through the log file to look through error messages for possible lead/hint. This requires more "hand holding."
   - After troubleshooting is successful, I will reach out to the Customer Success team or personally add the learning to our public knowledge base.

2. Please write your first response to this issue below.

```
Hi John Doe,

Please make sure the followings configured for all ten of your connectors:
1.  User Access is granted to Fivetran for data sources.
2.  Source table names, sheet urls, or sftp/ftp urls are correct.
3.  If using authentication, make sure your user/password, ports, and
    authentication method are correct.
4.  Double check Fivetran supports your data instance for application, database,
    and files here:
    - https://fivetran.com/docs/files
    - https://fivetran.com/docs/databases
    - https://fivetran.com/docs/databases

I understand this must be frustrating for all ten connectors to fail sync.
Please reach out if you have further questions. I'm happy to help.

Best Regards,
Zaw Mai
```

# Challenge 2

Assume that you provided a workaround for a bug. The customer is not satisfied with the workaround and indicates that our engineering team needs to re-prioritize the issue and provide an immediate fix since it is affecting their business. However, our engineering is guaranteeing a fix will be deployed at the end of the quarter. Please provide your response to the customer.

```
Hi John Doe,

I sincerely would like to apologize that the workaround fix did not satisfy you.
I know this is inconvenient for your business operations and I can only imagine
the negative impact this has causesd.

Our engineering team is working on an official fix to be deployed at the end of
the quarter. Unforunately, we can't allocate more resource for a hotfix at this
time.

I've reported the situation to my manager. In the meantime, please keep
using the workaround and I will personally give you an udpate when the next
release is out.

I know this isn't what you want to hear, and again I want apologize the
inconvenience caused to your business. Thank you for understanding and
your support of Fivetran.

Sincerely,
Zaw Mai
```