

# ERD Connection Traps

by Sophia



## WHAT'S COVERED

This tutorial explores connection traps that can be created in entity-relationship models in two parts:

1. Chasm Traps
2. Fan Traps

## 1. Chasm Traps

The two common connection traps in entity-relationship models include the chasm trap and fan trap. The **chasm trap** can occur when there are two one-to-many joins that converge on a single table, and the query includes aggregate data from both leaf tables. As a result, you can run into a situation where the results from multiple rows are returned from the tables when the query is being processed. This type of issue can occur due to the join between the leaf tables as those values would be multiplied potentially resulting in incorrect results.

A chasm trap can occur when the existence of a relationship is seen in the ERD but in reality, there isn't actually any connection.

➔ **EXAMPLE** Take, for example, a situation where you have a group that has students and the group works on projects. The relationship, in this case, would set it such that not all projects are worked on by students as they may be handled externally.

By having this type of model, if any project is not associated with a student, it cannot be associated with the group. So, if you have a situation where a group without students is assigned to a project (perhaps the students have not been assigned to the group yet), you could not actually create that link. To work around that scenario, you would have to create a link between the group and the project directly rather than depend on the middle relationship to include the student.



### TERM TO KNOW

#### Chasm Trap

A chasm trap can occur when there are two one-to-many joins that converge on a single table, and the query includes aggregate data from both leaf tables.

## 2. Fan Traps

The **fan trap** occurs when you have two one-to-many joins that follow a parent-child form. If you try to aggregate both measures simultaneously, you will probably get incorrect results. For example, in our PostgreSQL database, you have customer, invoice, and invoice\_line. If you attempted to aggregate the sum using the total from the invoice as well as the sum of the quantity \* the unit\_price at the same time. Let us see what would happen:

```
SELECT customer.customer_id, sum(total) as "SUM of Total", sum(quantity * unit_price) as "SUM of quantity times unit_price"
FROM customer, invoice, invoice_line
WHERE customer.customer_id = invoice.customer_id AND invoice.invoice_id = invoice_line.invoice_id
GROUP BY customer.customer_id;
```

## Query Results

Row count: 59

customer_id	SUM of Total	SUM of quantity * unit_price
29	338	37.62
54	338	37.62
4	366	39.62
34	356	39.62
51	344	38.62
52	338	37.62
10	338	37.62
35	338	37.62

You would have expected the sum of the values to be the same. The sum of the quantity \* unit\_price is calculated correctly, but the sum of the total is not calculated correctly. This is because you get a value for every invoice\_id that occurs in the invoice\_line which will inflate the results for each invoice\_id listed.

With an entity-relationship model, it can be a different scenario where you may have the relationship being ambiguous where you can have two or more one-to-many relationships that fan out from the same entity set. For example, if you have a scenario where the bank has a 1:M connection with a person and the bank also has a 1:M relationship with a clerk, you have two one-to-many relationships that come from the bank table. If you wanted to identify which clerk helped which person, you could not do that as they aren't linked in that way and you do not have the means to store that. If you have that situation, you could reconstruct the relationship to have the bank have a 1:M relationship with the clerk. Then the clerk could have a 1:M relationship with the person. This eliminates the fan trap.



### TERM TO KNOW

#### Fan Trap

The fan trap occurs when you have two one-to-many joins that follow a parent-child form. If you try to aggregate both measures simultaneously, you will probably get incorrect results.



### SUMMARY

In this tutorial, you learned about two common connection traps. The **chasm trap** can occur when there are two one-to-many joins that converge on a single table, and the query includes aggregate data from both leaf tables. Also, the **fan trap** can occur when you have two one-to-many joins that follow a parent-child form. Both forms of traps can produce incorrect results with your entity-relationship model.

Source: Authored by Vincent Tran



### TERMS TO KNOW

#### Chasm Trap

A chasm trap can occur when there are two one-to-many joins that converge on a single table, and the query includes aggregate data from both leaf tables.

#### Fan Trap

The fan trap occurs when you have two one-to-many joins that follow a parent-child form. If you try to aggregate both measures simultaneously, you will probably get incorrect results.

