



INSTITUTO POLITÉCNICO NACIONAL

UNIDAD PROFESIONAL INTERDISCIPLINARIA DE INGENIERÍA EN TECNOLOGÍAS AVANZADAS

Ingeniería Telemática Bases de Datos Distribuidas Tarea 3.

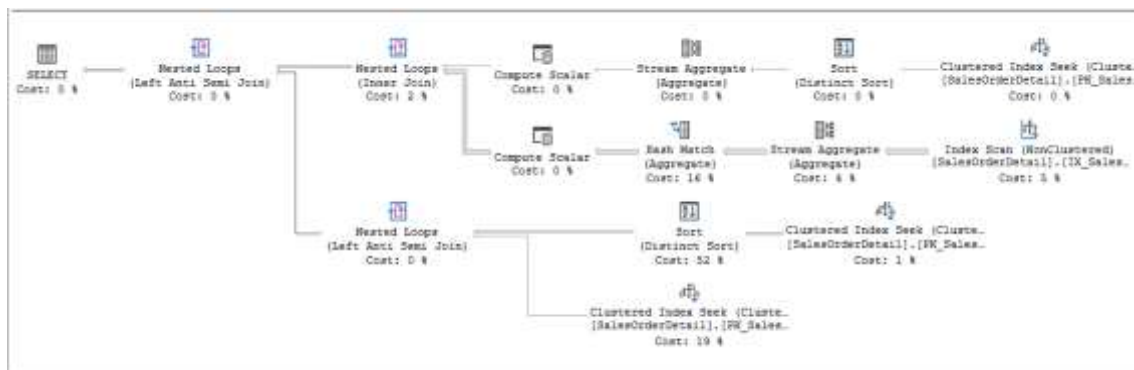
- **Docente:** De la Cruz Sosa Carlos
- **Alumnos:** De Jesús Lucio Oscar Daniel
Pérez Iturbe Carolina
- **Grupo:** 3TM3
- **Fecha de entrega:** 04 de junio 2025
- **Ciclo escolar:** 2025-2



Analizar el plan de ejecución de la consulta 3 de la práctica PLANES DE EJECUCIÓN con la siguiente solución:

```
WITH ProductsOrder43676 AS (  
    SELECT productid  
    FROM AdventureWorks2022.sales.SalesOrderDetail  
    WHERE salesorderid = 43676  
),  
CandidateOrders AS (  
    SELECT salesorderid, productid  
    FROM AdventureWorks2022.sales.SalesOrderDetail  
    WHERE salesorderid <> 43676  
)  
SELECT salesorderid  
FROM CandidateOrders  
GROUP BY salesorderid  
HAVING COUNT(DISTINCT productid) >= (SELECT COUNT(DISTINCT productid)  
    FROM ProductsOrder43676)  
AND NOT EXISTS (  
    SELECT productid  
    FROM ProductsOrder43676  
EXCEPT  
    SELECT productid  
    FROM CandidateOrders c  
    WHERE c.salesorderid = CandidateOrders.salesorderid  
);
```

Plan de ejecución de la consulta:



La consulta realiza varias operaciones de agregación, ordenamiento, uniones internas y externas (semi joins).

El operador más costoso es un **Sort (Distinct Sort)** con un **52% del costo total**.

Hay múltiples accesos a la tabla SalesOrderDetail, mediante **Clustered Index Seek**, los cuales indican acceso eficiente a la tabla usando la clave primaria e **Index Scan** el cual se ocupa por un índice no-clustered e implica que no se pudo usar un índice más selectivo.

También se puede notar que se están haciendo operaciones de agregación sobre los resultados intermedios, posiblemente para eliminar duplicados o calcular métricas. Y que se usa un Hash Match para realizar agregaciones, que es eficiente para grandes volúmenes de datos no ordenados.

Se nota que Sort es el paso que tiene el **mayor costo (52%)**, lo cual sugiere que la eliminación de duplicados o preparación para una operación de agregación está siendo muy costosa.

Hay varios Nested Loops, incluyendo **Left Anti Semi Join**, lo cual indica que se están haciendo comparaciones para excluir ciertos resultados (por ejemplo, "WHERE NOT EXISTS").

También se usa un Inner Join, aunque con un costo bajo.