Data models

DB1: Prod{[Cat],Supp}, St, Wa, OL, Cl

DB2: Prod{[Cat],Supp, [**St**]}, Wa, OL, Cl

DB3: **St**{Prod{[Cat],Supp}}, Wa, OL, Cl

DB4: St, Wa, **OL**{Prod{[Cat],Supp}}, Cl

DB5: Prod{[Cat],Supp, [**OL**]}, St, Wa, Cl

Queries

1. Filter queries

Q1 The stock of a given ID product in a given warehouse:

```
SELECT S.quantity, S.location
FROM Stock S
WHERE S.IDP = $IDP AND S.IDW = $IDW;
```

Q2 Names and prices of product from a given brand (take “Apple” as example):

```
SELECT P.name, P.price
FROM Product P
WHERE P.brand = $brand;
```

Q3 Product ID and quantity from order lines ordered at a given date:

```
SELECT O.IDP, O.quantity
FROM OrderLine O
WHERE O.date = $date;
```

2. Join queries

Q4 Stock (list of product names, as well as their quantity) from a given warehouse;

```
SELECT P.name, S.quantity
```

```
FROM Stock S JOIN Product P ON S.IDP = P.IDP
WHERE S.IDW = $IDW;
```

Q5 Distribution of “Apple” brand products (name & price) in warehouses (IDW & quantity);

```
SELECT P.name, P.price, S.IDW, S.quantity
FROM Product P JOIN Stock S ON P.IDP = S.IDP
WHERE P.brand = "Apple";
```

3. Aggregation queries

Q6 The 100 most ordered product names and price (sum of quantities).

```
SELECT P.name, P.price, OL.NB
FROM Product P JOIN (
SELECT O.IDP, SUM(O.quantity) AS NB
FROM OrderLine O
GROUP BY O.IDP
) OL ON P.IDP = C.IDP
ORDER BY OL.NB DESC
LIMIT 1;
```

Q7 Name and price of the product most ordered by customer no. 125;

```
SELECT P.name, P.price, OL.NB
FROM Product P JOIN (
SELECT O.IDP, SUM(O.quantity) AS NB
FROM OrderLine O
WHERE O.idClient = 125
GROUP BY C.IDP
) OL ON P.IDP = OL.IDP
ORDER BY OL.NB DESC
LIMIT 1;
```

DB1 : Prod{[Cat],Supp}, St, Wa, OL, Cl

	Sharding	C1		C2			
		S1	O1	Loops	S2	O2	#msgs
R1.1	S(#idW)	1	1				
R1.2	S(#idP)	1	1				
R2.1	P(#brand)	1	50				
R2.2	P(#idP)	1000	50				
R3.1	O(#idC)	1000	10 958 904				
R3.2	O(#idP)	1000	10 958 904				
R4.1	S(#idW), P(#idP)	1	10^5	10^5	1	1	
R4.2	S(#idP), P(#idP)	1000	10^5	10^5	1	1	
R5.1	S(#idP), P(#brand)	1	50	50	1	200	
R5.2	S(#idP), P(#idP)	1000	50	50	1	200	

C1 = #S1 * size S1 + #O1 * size O1
C2 = #S2 * size S2 + #O2 * size O2

Vt = C1 + loops * C2

CommunicationTime = Vt * Ct (Ct is bandwidth speed)

DB2 : Prod{[Cat],Supp, [St]}, Wa, OL, Cl

	Sharding	C1		C2			
		S1	O1	Loops	S2	O2	#msgs
R1.1	P(#idP)	1 (152B)	1 (112B)				
R1.2							
R2.1	P(#brand)	1 (204B)	50 (112B)				
R2.2							
R3.1	O(#idC)	1000 (72B)	10 958 904 (40B)				
R3.2							
R4.1	P(#idP)	1000 (132B)	10^5 (40B)				
R4.2							
R5.1	P(#brand)	1 (244 B)	50 (8112B)				
R5.2							

DB3: St{Prod{[Cat],Supp}}, Wa, OL, Cl

	Sharding	C1		C2			
		S1	O1	Loops	S2	O2	#msgs
R1.1	S(#idW)	1 (152B)	1 (112B)				
R1.2							
R2.1	S(#idW)	1000 (204B)	50*200 (112B)				
R2.2							
R3.1	O(#idC)	1000 (72B)	10 958 904 (40B)				
R3.2							
R4.1	S(#idW)	1 (132B)	10^5 (40B)				
R4.2							
R5.1	S(#idW)	1000 (244 B)	50*200 (132B)				
R5.2							