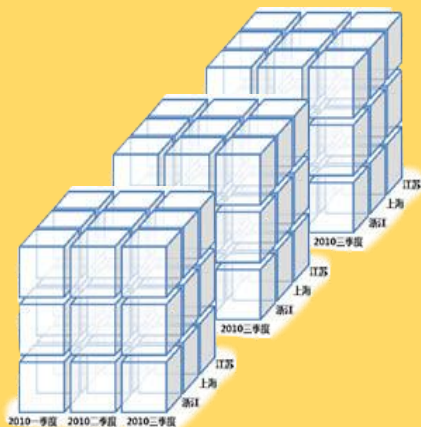




High Performance Multidimensional Analysis Solution

esProc SPL Base Application scenarios

Common Choices of OLAP Servers for Multidimensional Analysis



Pre-aggregation Solution

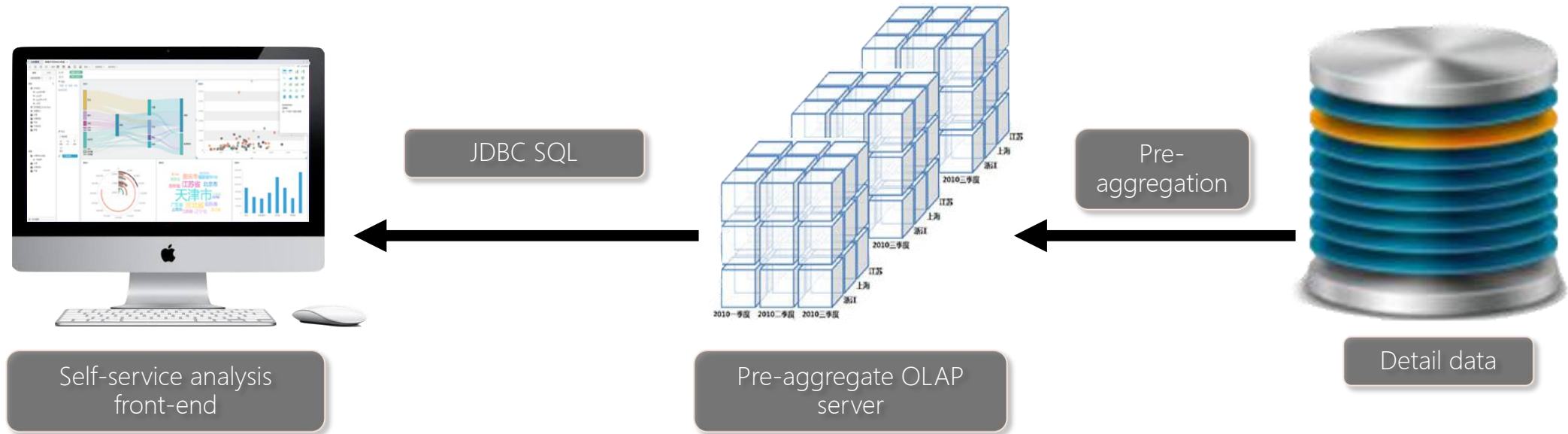
Pre-computed statistical results (cube)
Change the complexity from $O(n)$ to $O(1)$



General database Solution

Using Relational Database as Backstage
OLAP drag SQL query results

Pre-aggregate OLAP Server Architecture



Pre-calculated statistical results



Complexity changes from $O(n)$ to $O(1)$



Return results in seconds

300 meters

250 meters



Tens of thousands of TB in full pre-aggregation of dozens of dimensions



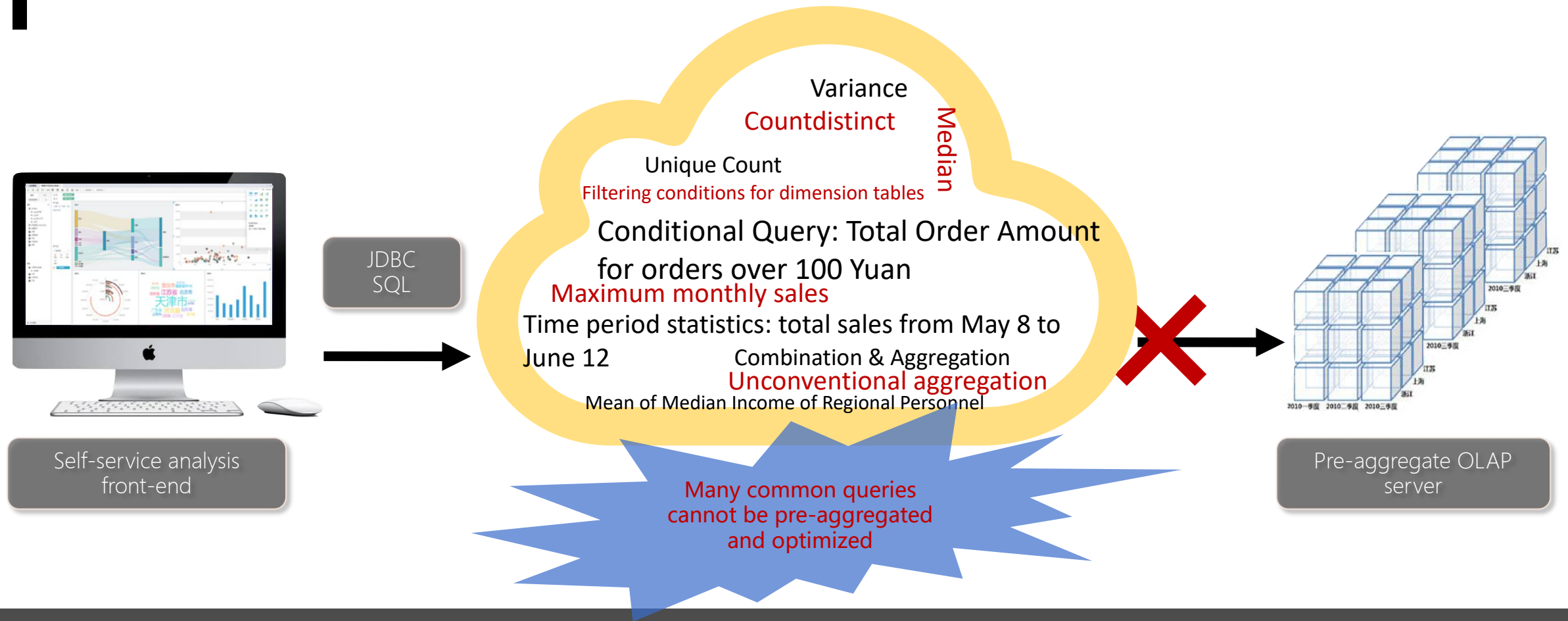
10,000 TB need to be stored in
10,000 1TB hard disks



10,000 1TB hard drives stacked
together will be 250 meters high

Pre-aggregate OLAP Server Problem: Huge Storage Space

Pre-aggregate OLAP Server Problem: Functional Blind Zone



Conditional Query
Conditions for Dimensional Tables

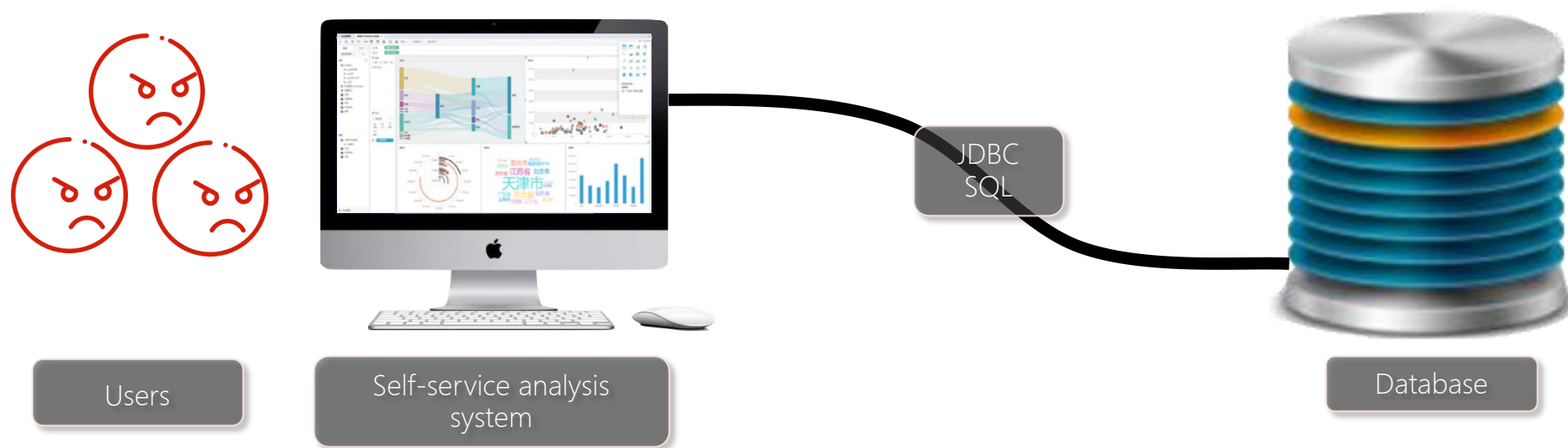


Combination & Aggregation
Unconventional aggregation



Time period statistics

General database solution problems: poor performance, long waiting time for users



Row-wise storage
Poor query performance of wide table



Low compression rate
Read slowly for large amount of data



No pre-association mechanism
Lots of JOINS are slow to compute

Breakthrough: esProc as OLAP Server



Solve the problems of pre-aggregate OLAP engine



OLAP Speed Up



Low hardware requirements

esProc: Using Multithread Parallel Mechanism



OLAP Query Based on Detailed Data

Detailed data do not need to be pre-aggregated according to various dimensions, and will not generate a large amount of cube data.



Multithread Parallel Accelerates the Access of Data Files

For multi-CPU (multi-core) servers, multi-thread parallel computing can effectively improve access to detailed data and computing speed.

esProc: Compressed column storage

1
2
3

Column storage

There are many fields in data file, and only a few participate in OLAP calculation. The column storage method can effectively improve the speed of data reading.

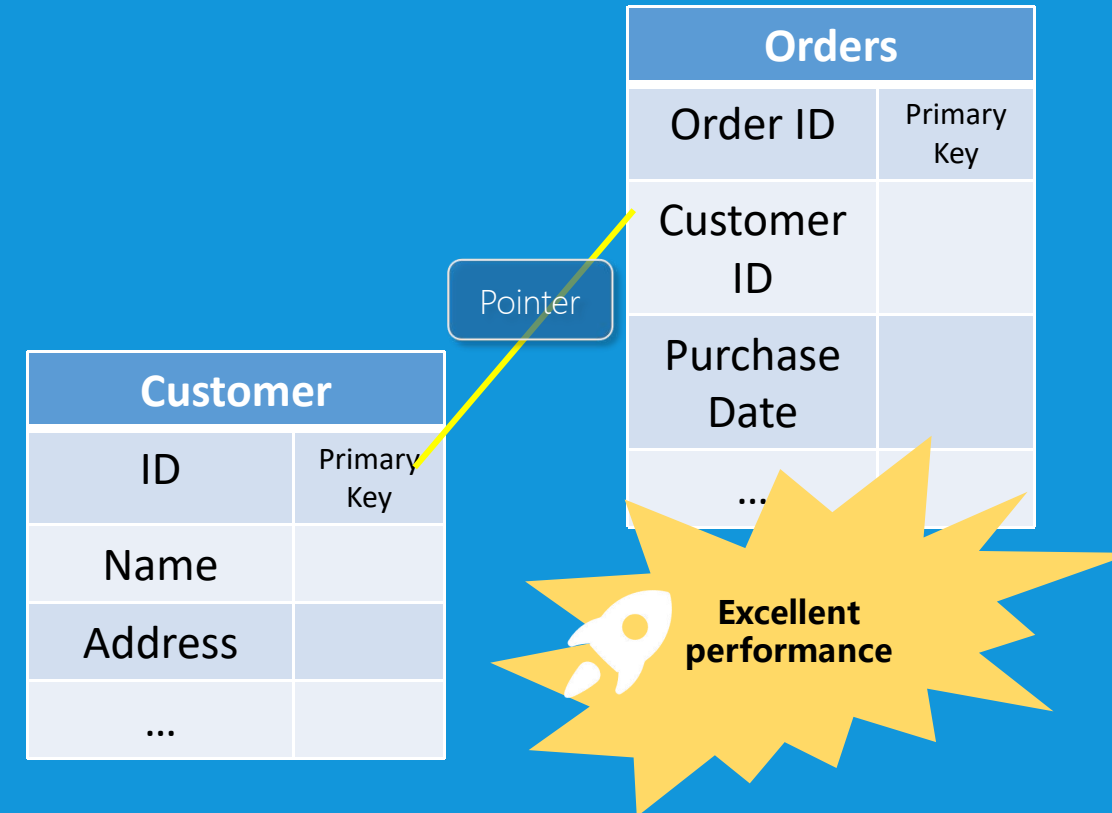
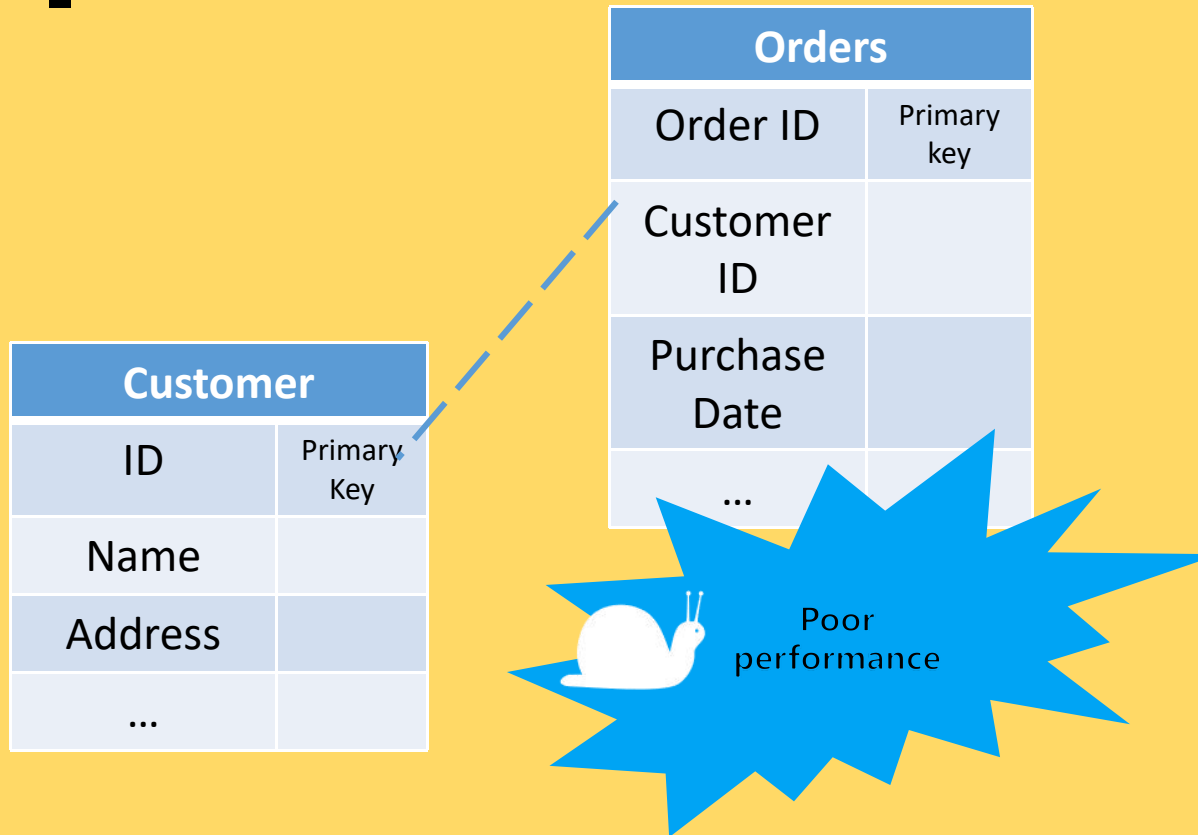


Compressed storage

With a compression ratio of about 10 times, the hard disk space occupied by data files becomes smaller and the reading speed is greatly improved.



esProc: Pre-association accelerates JOIN



Database JOIN:

Temporarily compute Join at query time



esProc JOIN

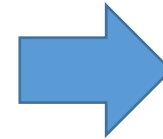
Pre-compute Joins

Stored in memory or hard disk in various ways

There is no need to compute Joins when querying

esProc: Optimization for Complex SQL

```
Select F1, F2, F3,  
  (select FF1 from TABLE1 WHERE...) AS F4,  
  (select min(FF1) from TABLE2 GROUP BY...) AS F5,  
From  
  (select FFF1 from TABLE3 WHERE...) T1  
Left join  
  (select FFFF1 from TABLE4 WHERE...) T2  
On T1.FFF2=T2.FFFF2  
WHERE  
T1.FFF2 in (select min(FFFF1) from TABLE5 GROUP  
BY...)
```



Converting Subqueries to JOIN Computing

Stripping Multi-Layer Nested SQL to Single Layer Computing

• • •



SQL multi-layer nesting with subqueries
dragged by OLAP front-end



OLAP engine can only optimize
simple SQL



esProc optimizes this type of SQL in
depth

esProc: Provide innovative pre-aggregation capabilities



Partial pre-aggregation to effectively balance the contradiction between space and time



Give a feasible method of pre-aggregation for time-period statistics

- The End -

THANK YOU

