



# HyperQueries:

Dynamic Distributed Query Processing on  
the Internet

---

Alfons Kemper **Christian Wiesner**

Universität Passau

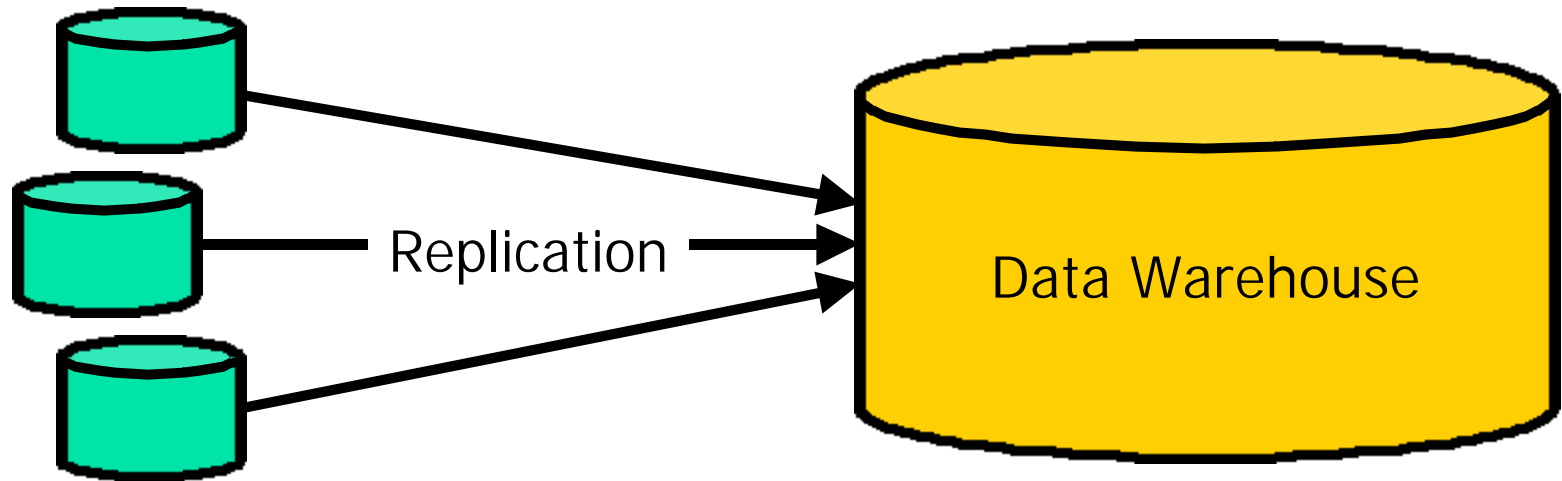


# Outline

---

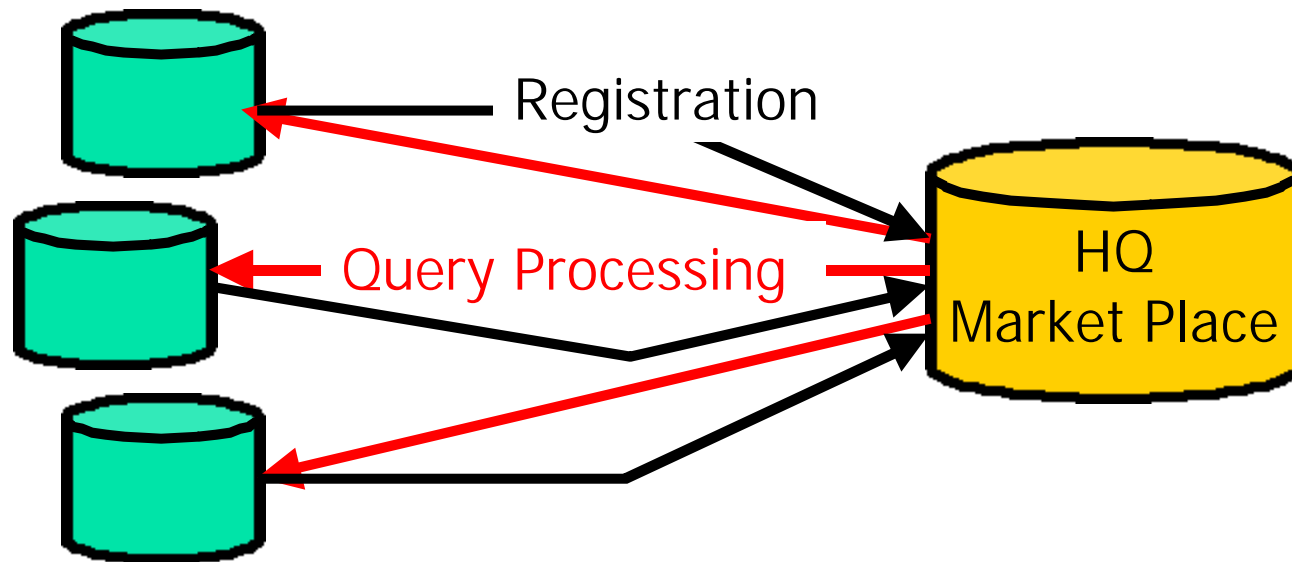
- Motivation
- Hyperlinks & HyperQueries
- Execution of HyperQueries
- Optimization Issues
- Related Work
- Conclusions

# Motivation: E-Market Places (1)



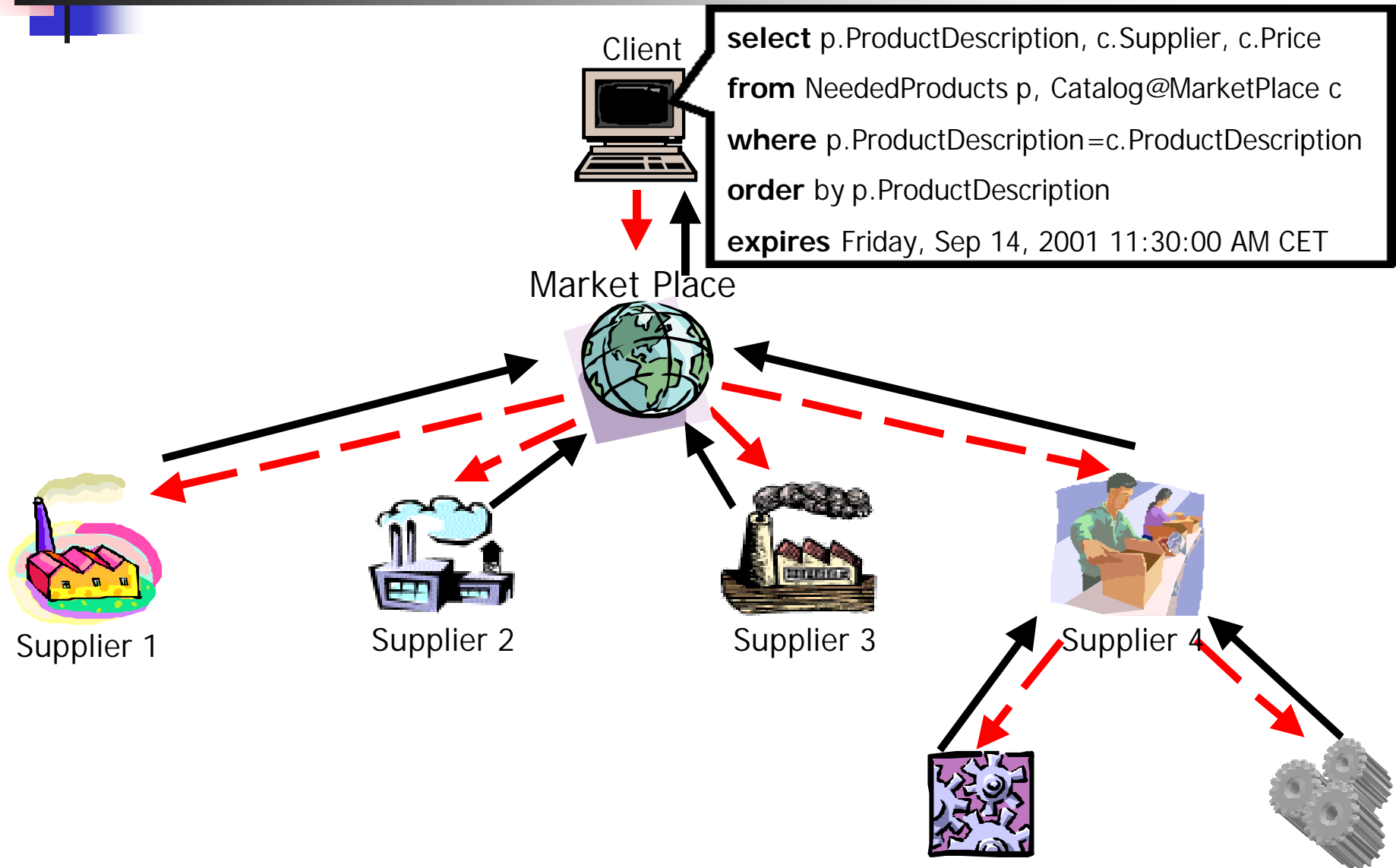
- Traditional techniques
  - Replicate data in a central data warehouse
  - Severe problems
    - Security and privacy violations
    - Coherence problems
    - Schema integration problems
    - Fixed query operators

## Motivation: E-Market Places (2)

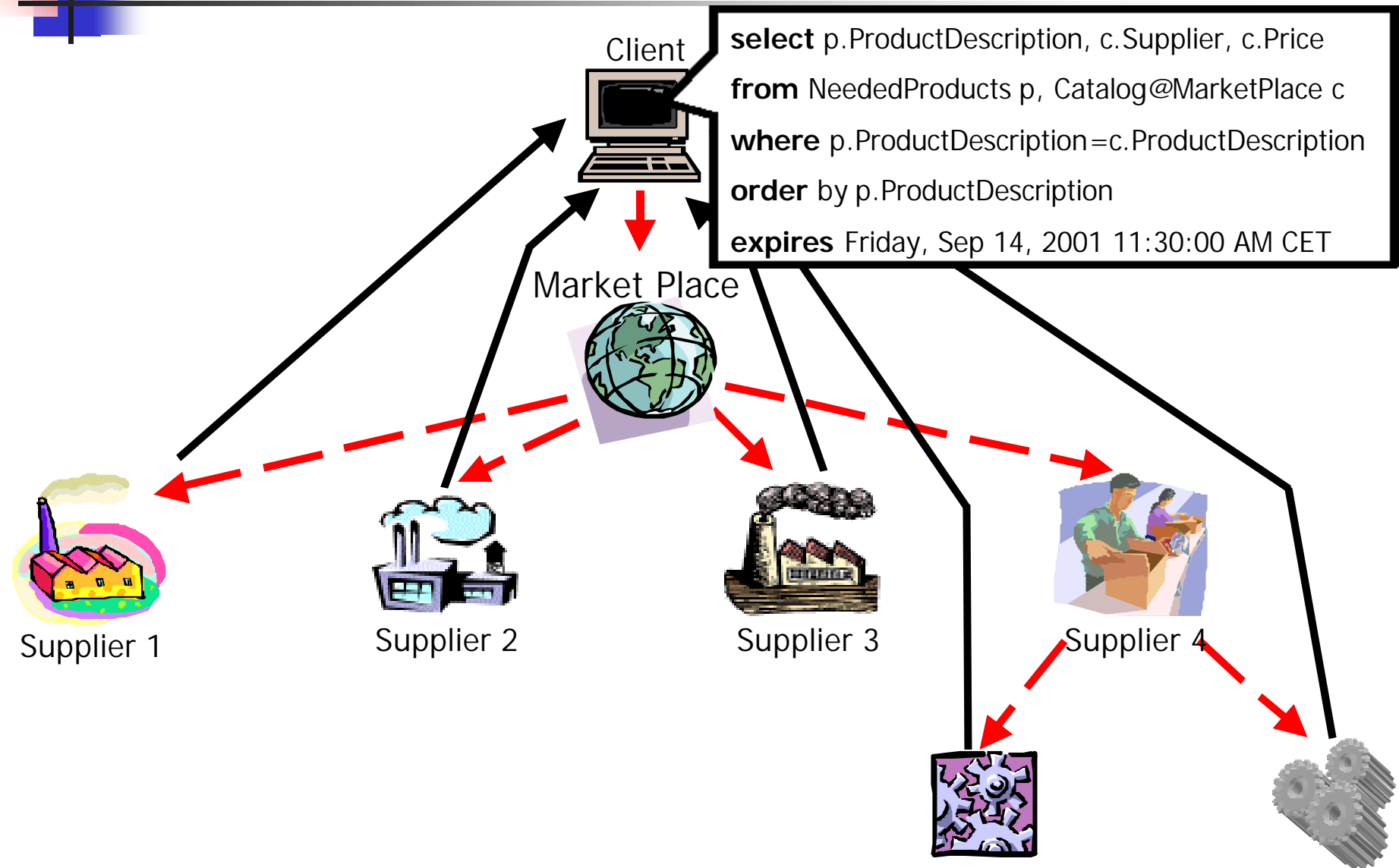


- HyperQuery framework
  - Market place is intermediary
  - Distribute query processing along allocation schema of data
  - Objects & queries flow through the WWW
  - Scalability

# Hierarchical Flow of Objects



# Broadcast-like Flow of Objects





# Hyperlinks

---

- Embed hyperlinks as *virtual attributes* in databases
- Hyperlinks refer to HyperQueries

# Hyperlinks - Example

virtual attribute

ProductDescription	Supplier	Price
Battery, 12V 32A	Supplier 1	hq://supplier1.com/Electrical/Price?ProdID=CB1232
Battery 12V 55A	Supplier 1	hq://supplier1.com/Electrical/Price?ProdID=CB1255
Tires 175/65TR14	Supplier 2	hq://supplier2.com?ProdKey=175_65TR14
Spare		234

hq  
supplier1.com  
Electrical/Price  
?ProdID=CB1232

HyperQuery Protocol  
DNS  
HyperQuery ID  
Object-Specific Parameter

!Currency=USD

Global Parameter





# HyperQueries

---

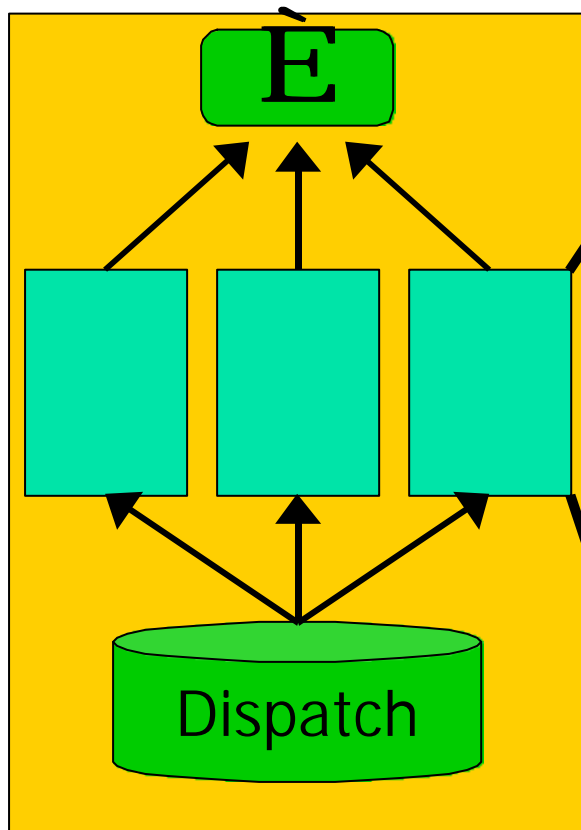
- Sub-plans at remote hosts
- Virtual table **HyperQueryInputStream**
- SQL Dialect

Electrical/Price@Supplier1.com:

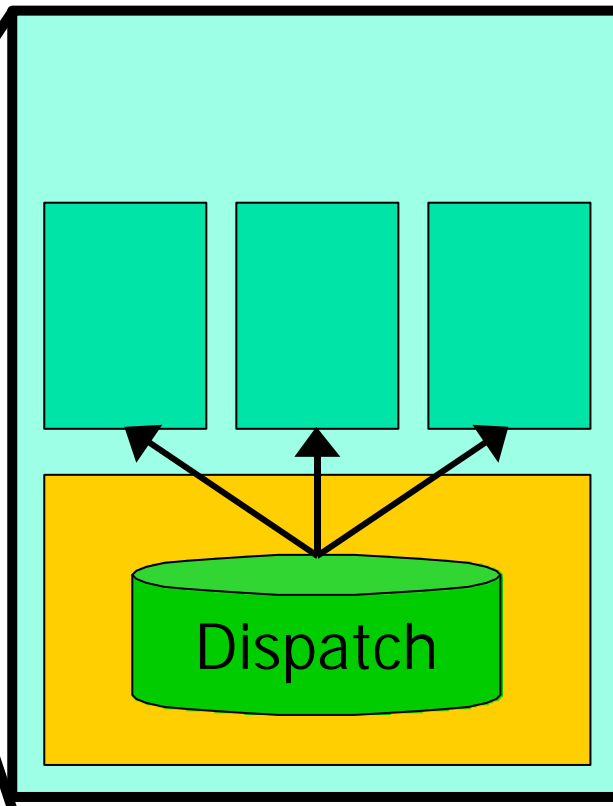
```
select h.*, p.Price as Price  
from HyperQueryInputStream h, Products p  
where h.ProdID = p.ProdID
```

# Templates for Sub-Plans

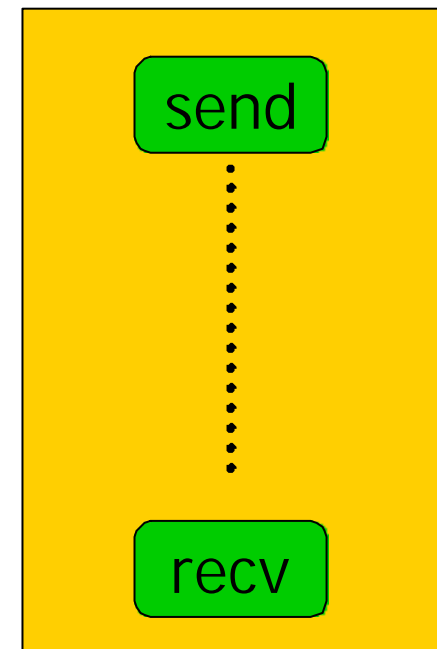
Nesting



Sequencing



Inner



# Example

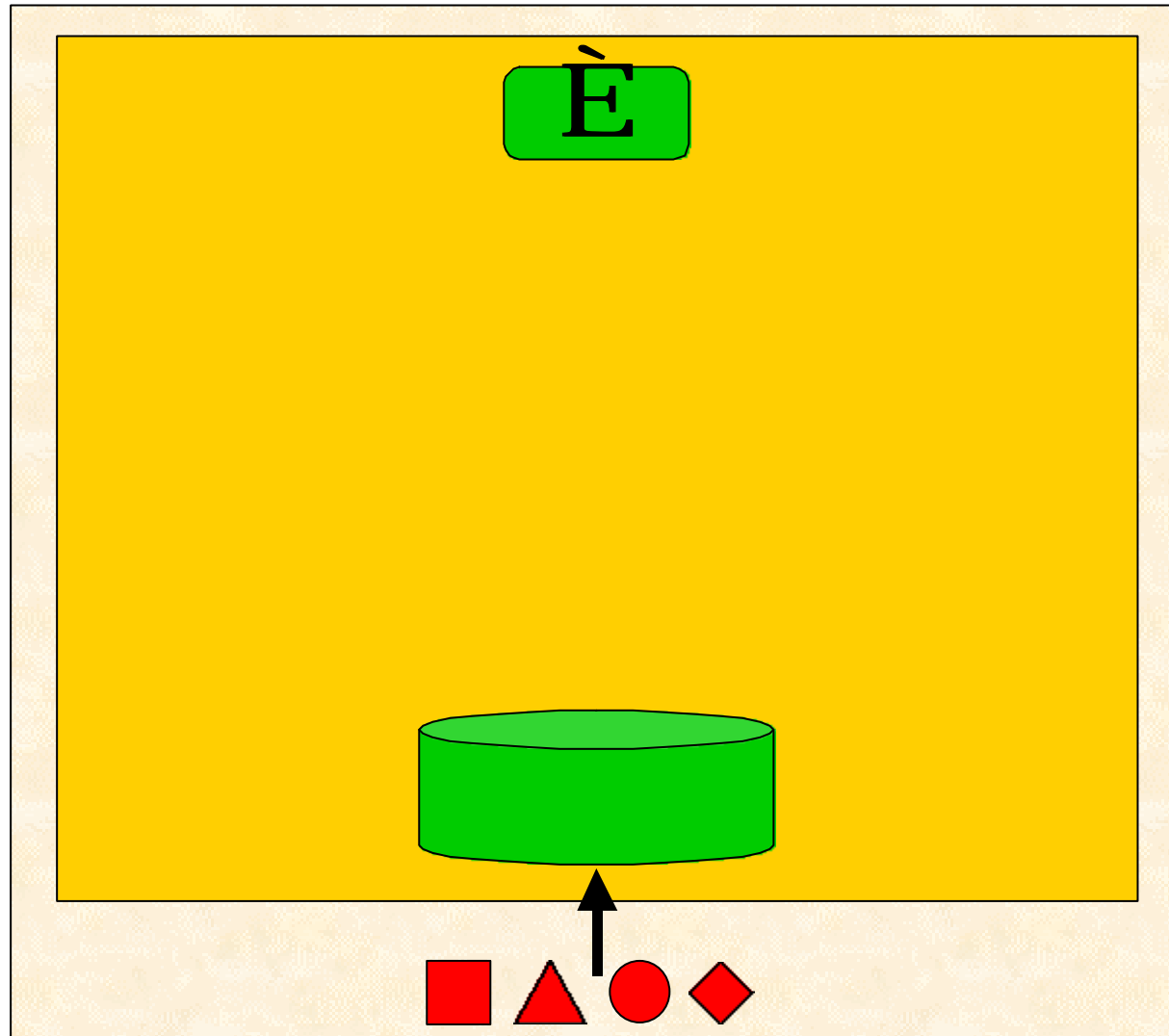
- Query

```
select p.ProductDescription, c.Supplier, c.Price
from NeededProducts p, Catalog@MarketPlace c
where p.ProductDescription=c.ProductDescription
order by p.ProductDescription
expires Friday, Sep 14, 2001 11:30:00 AM CET
```

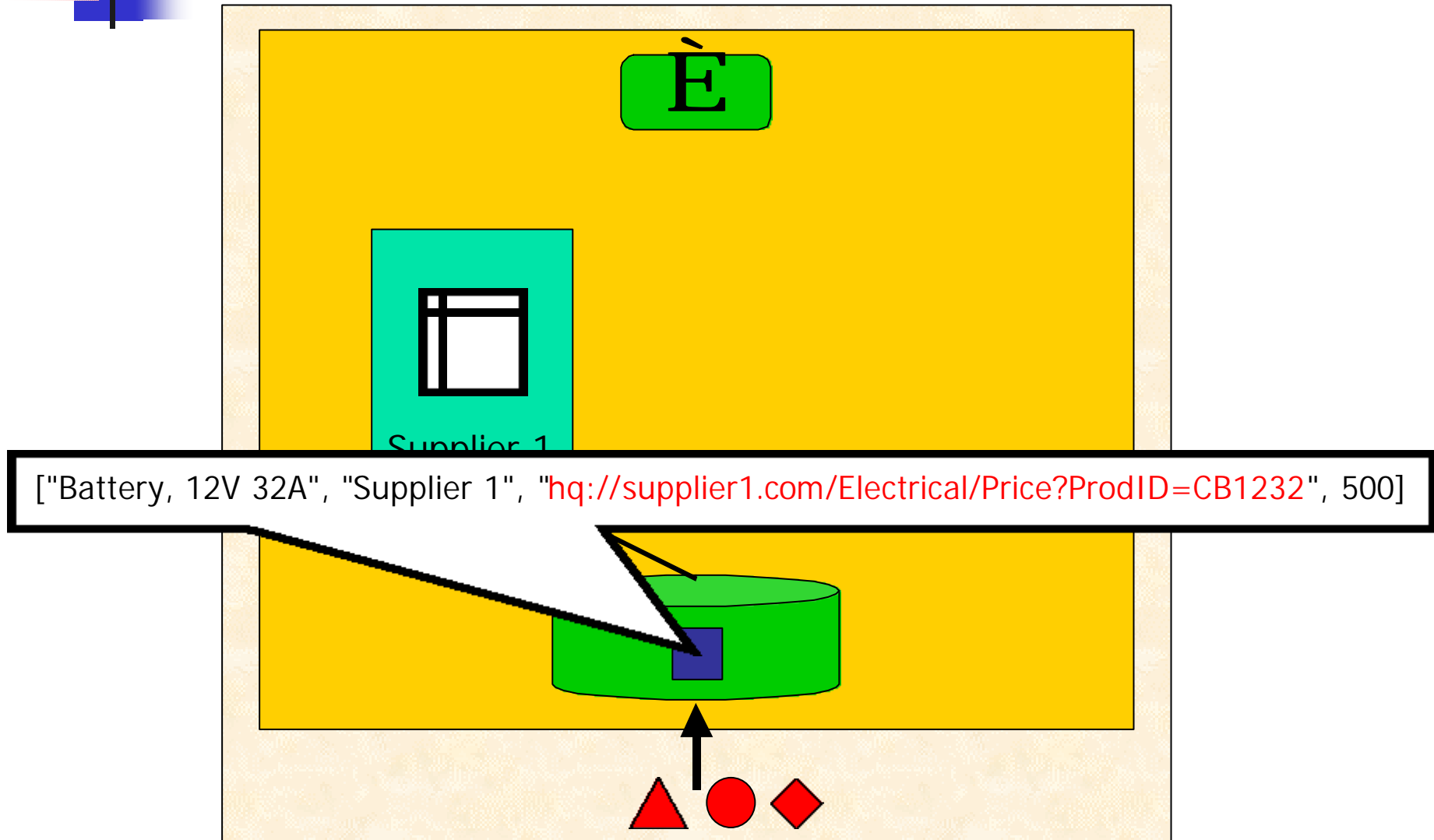
- NeededProducts@Client:

ProductDescription	Quantity	
Battery, 12V 32A	500	■
Battery, 12V 55A	750	▲
Tires 175/65TR14	1000	●
Spark Plug VX	8000	◆

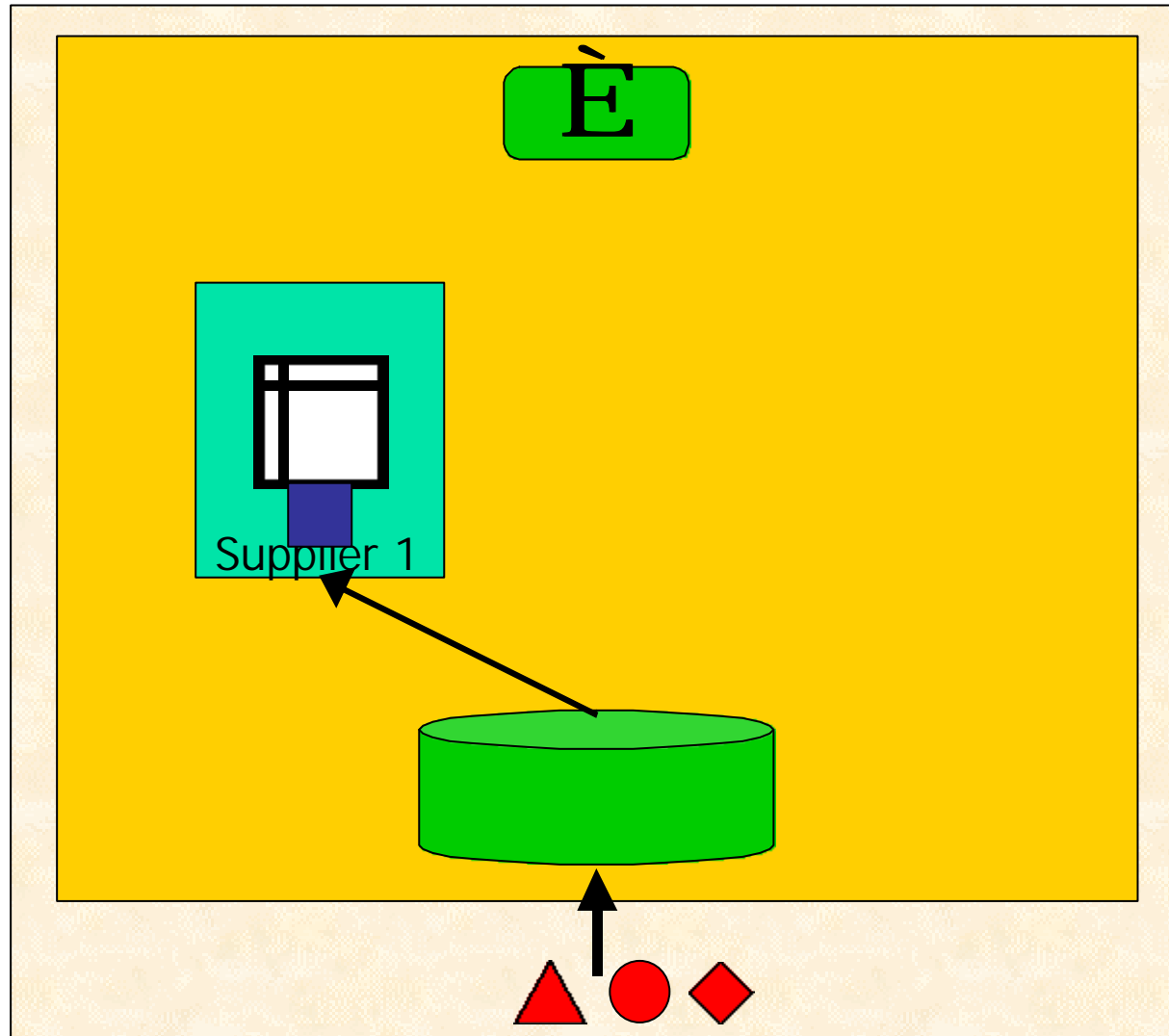
# HyperQuery Execution



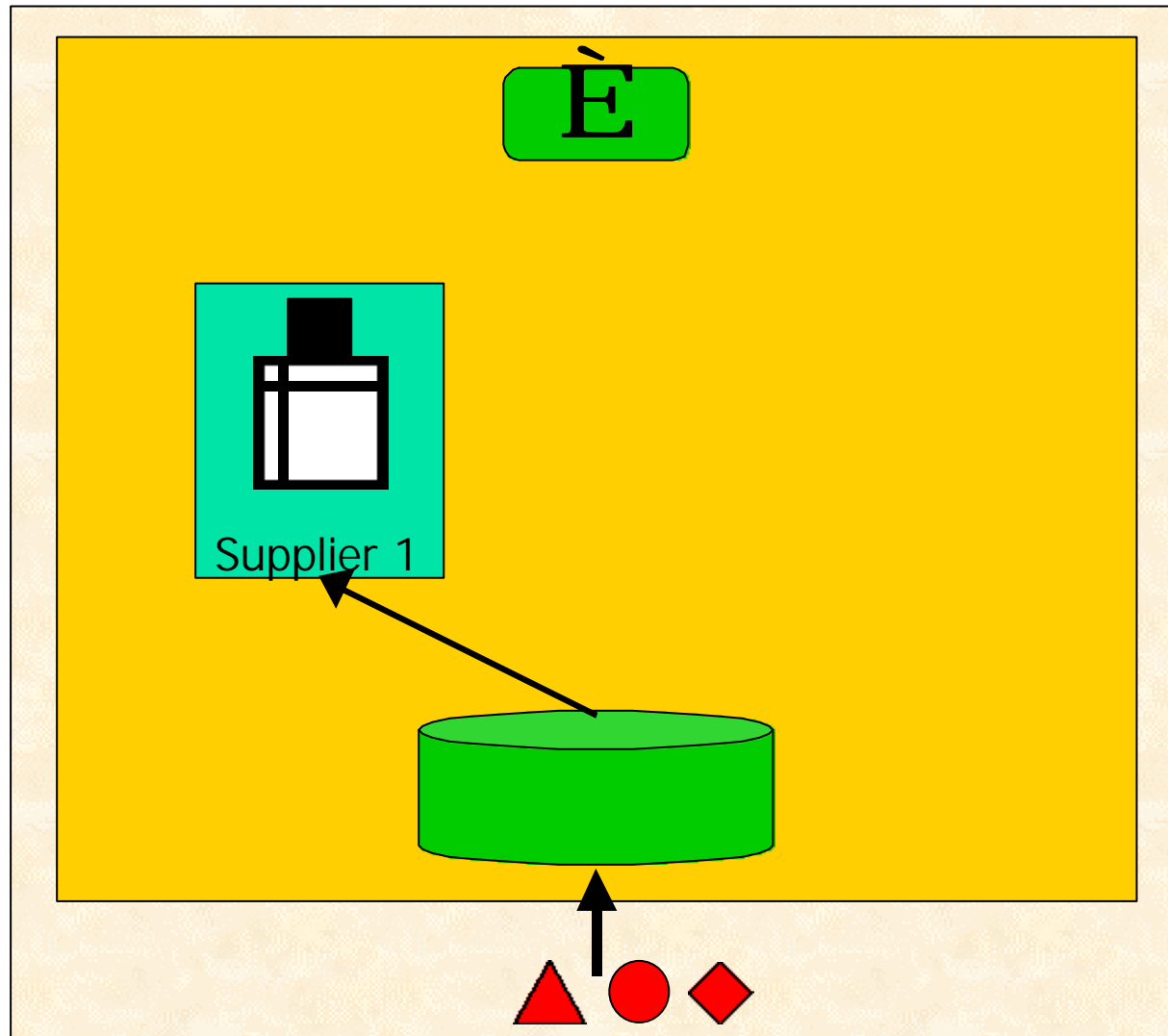
# HyperQuery Execution



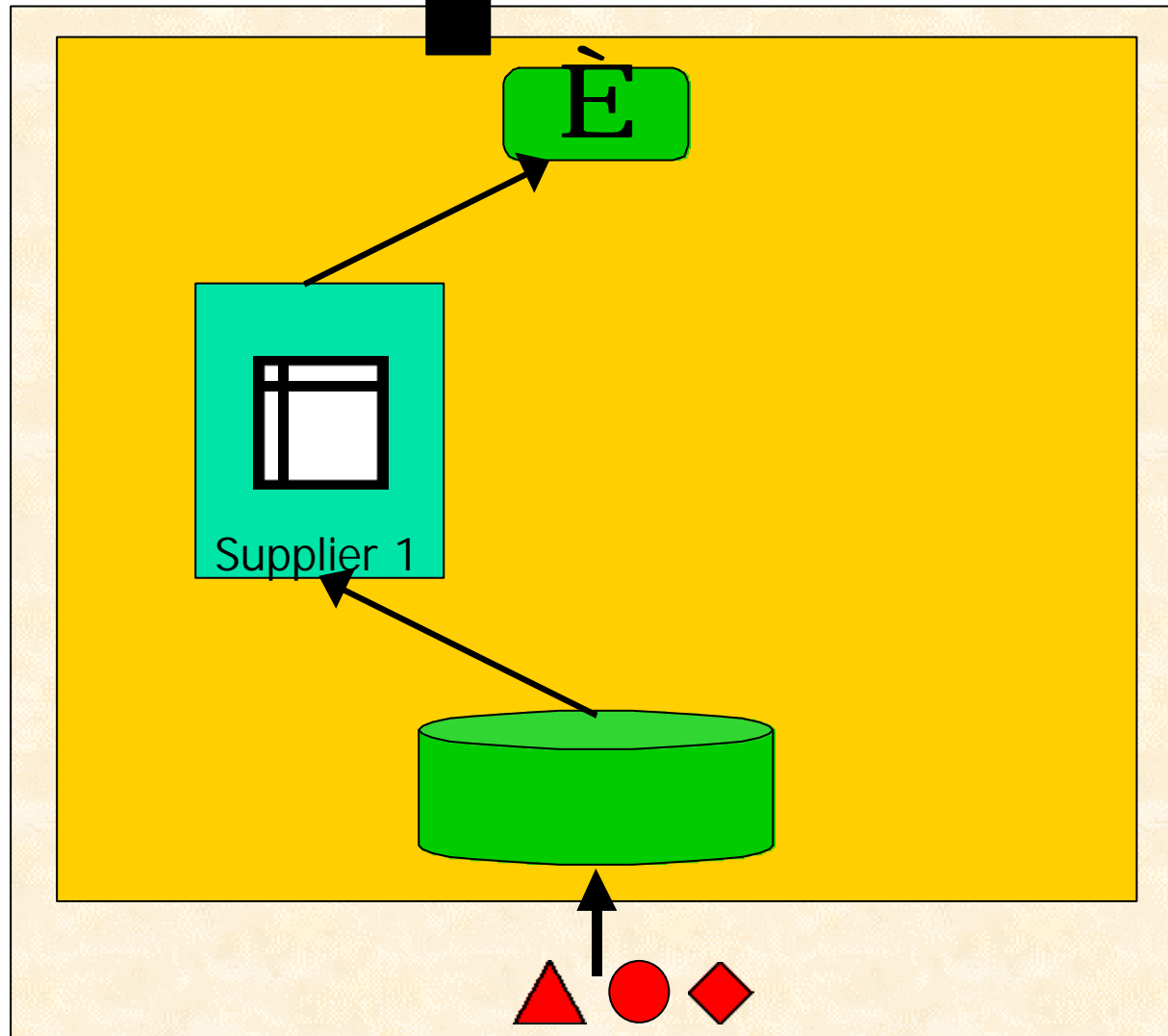
# HyperQuery Execution



# HyperQuery Execution

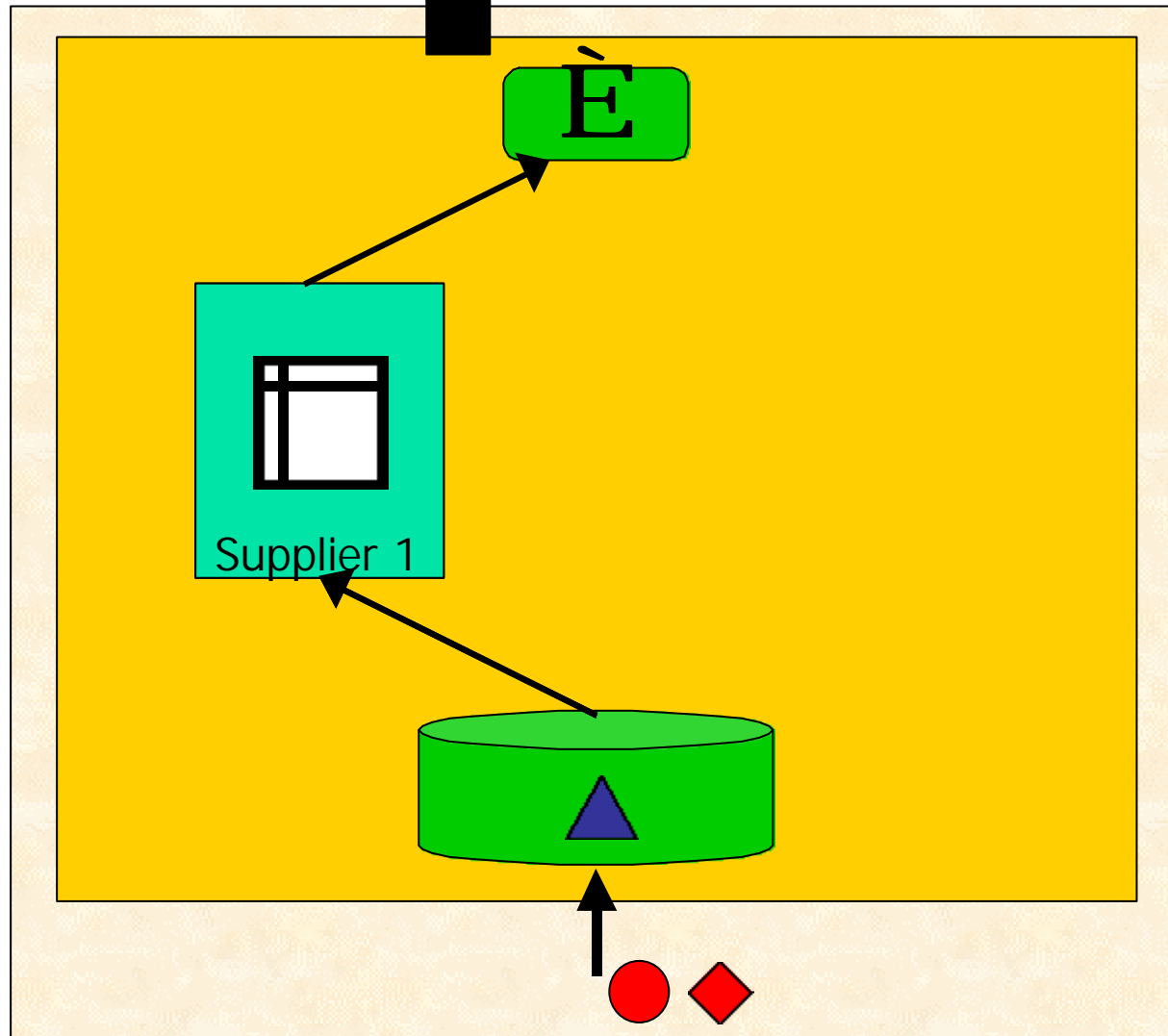


# HyperQuery Execution

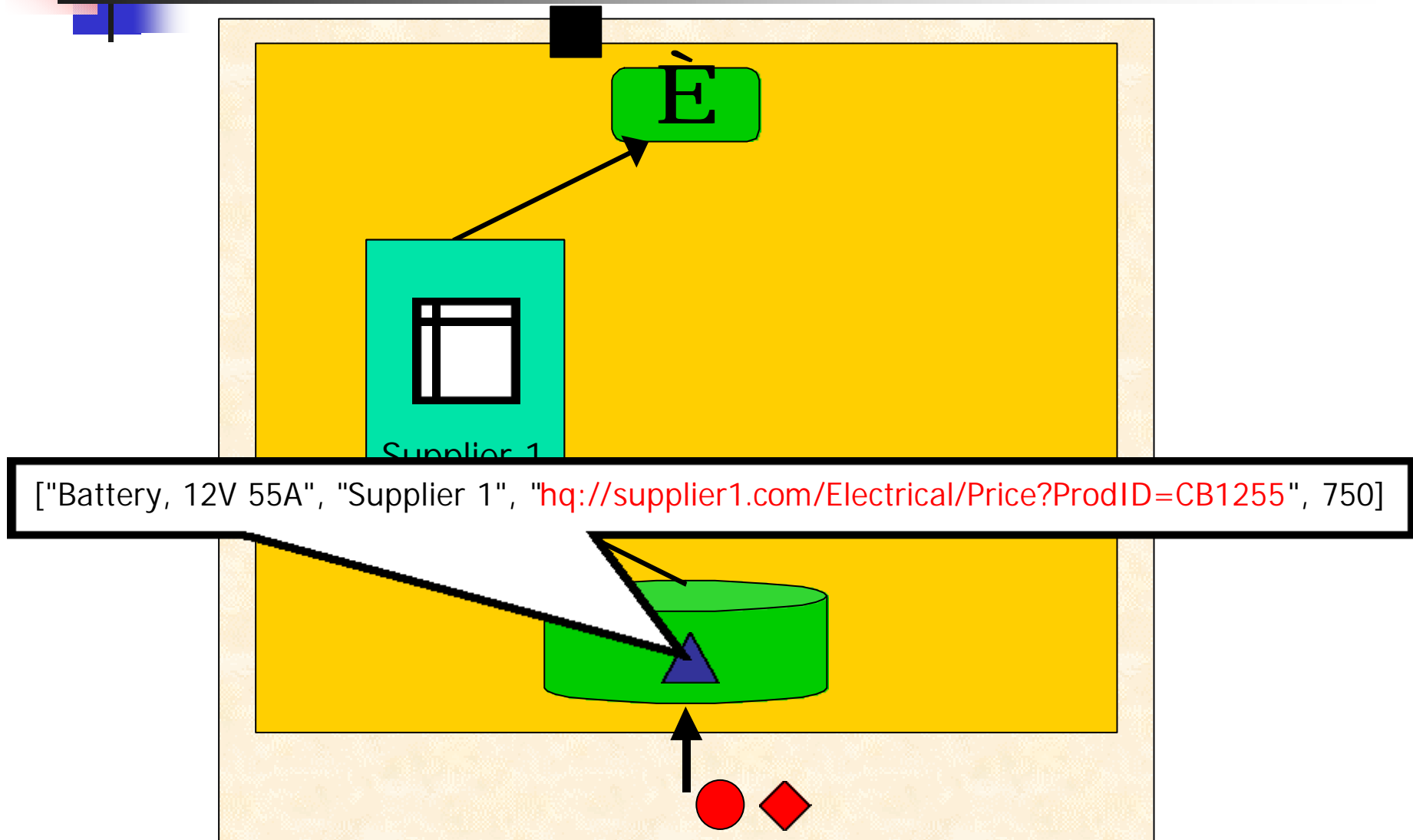




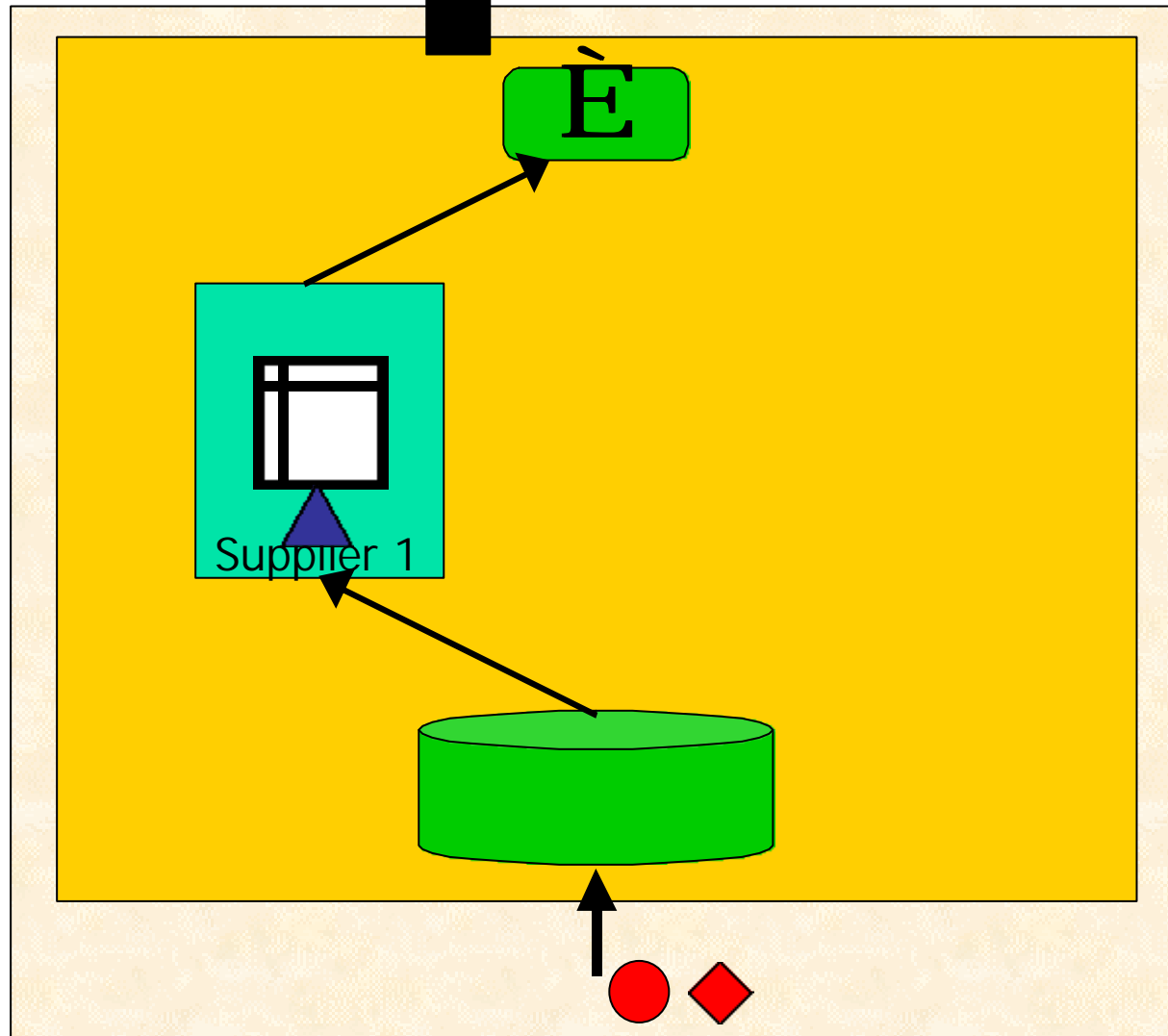
# HyperQuery Execution



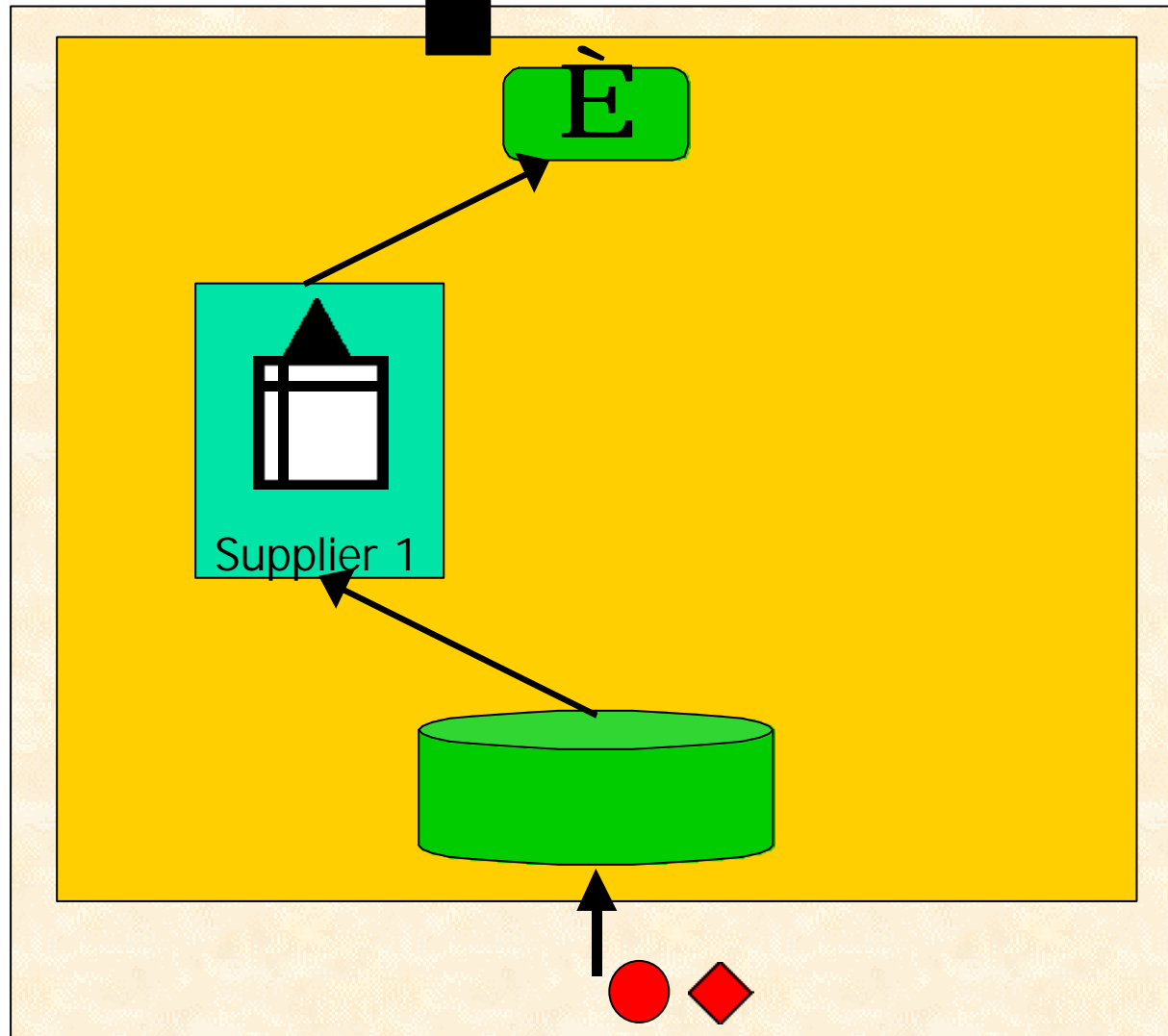
# HyperQuery Execution



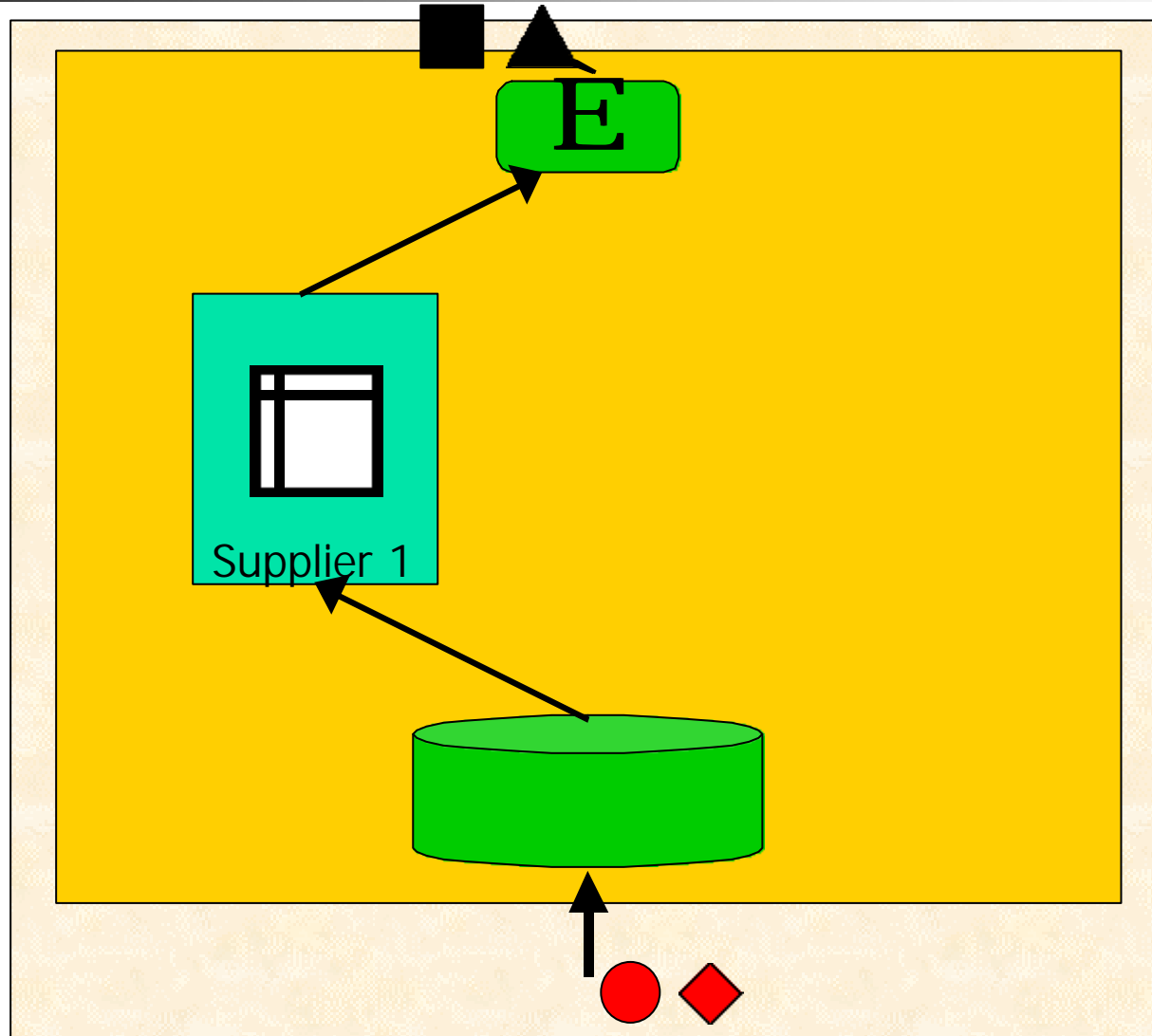
# HyperQuery Execution



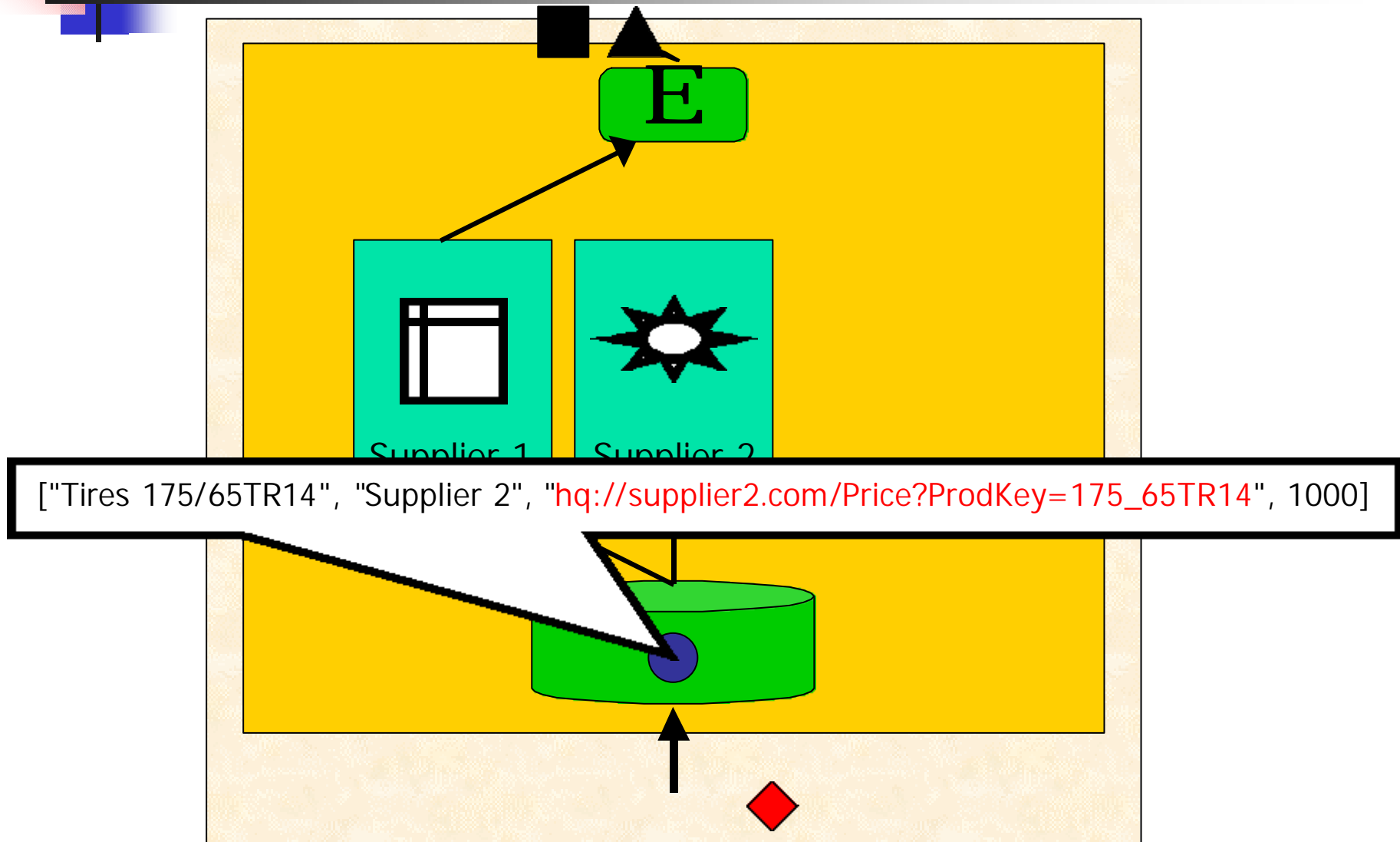
# HyperQuery Execution



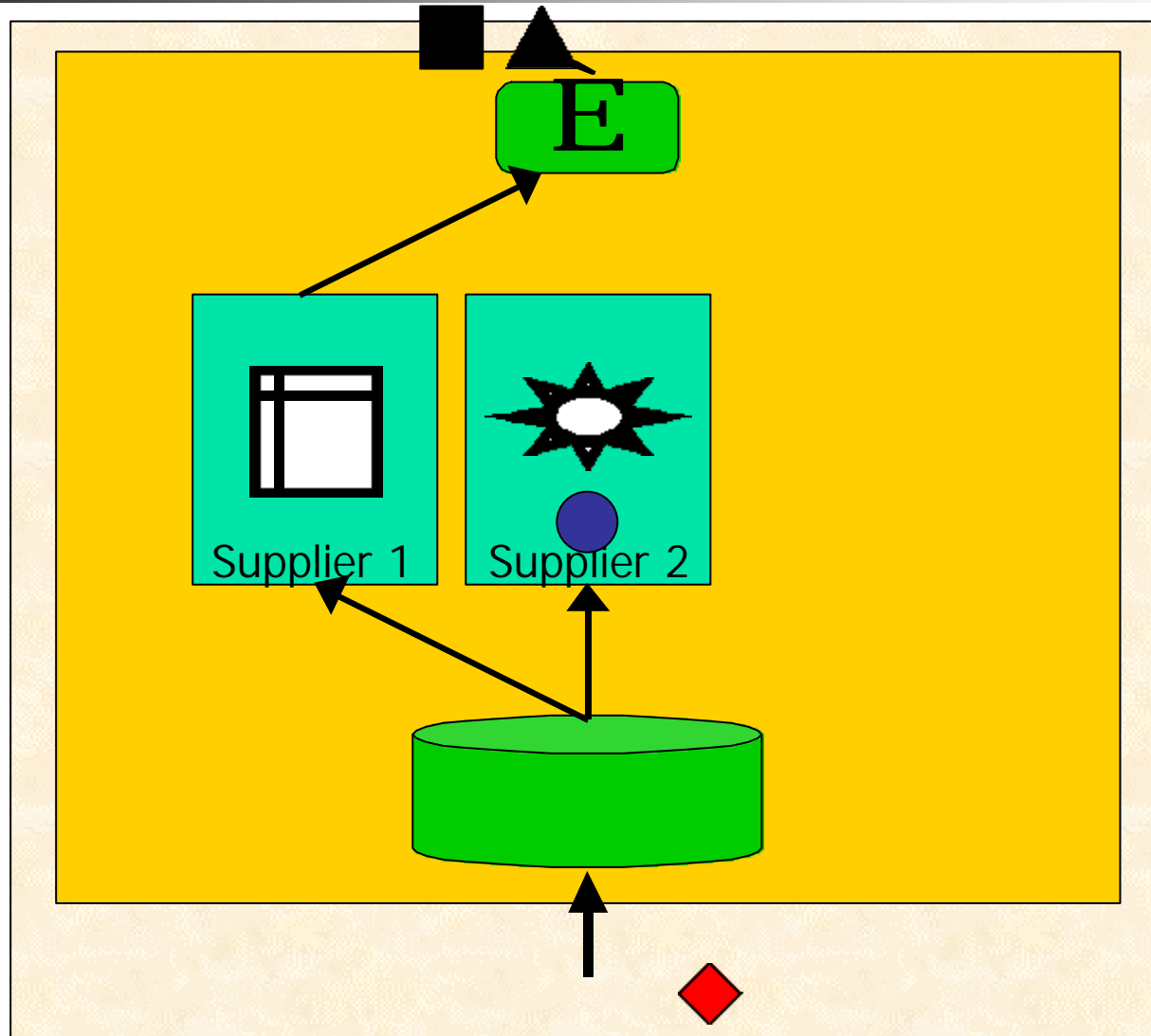
# HyperQuery Execution



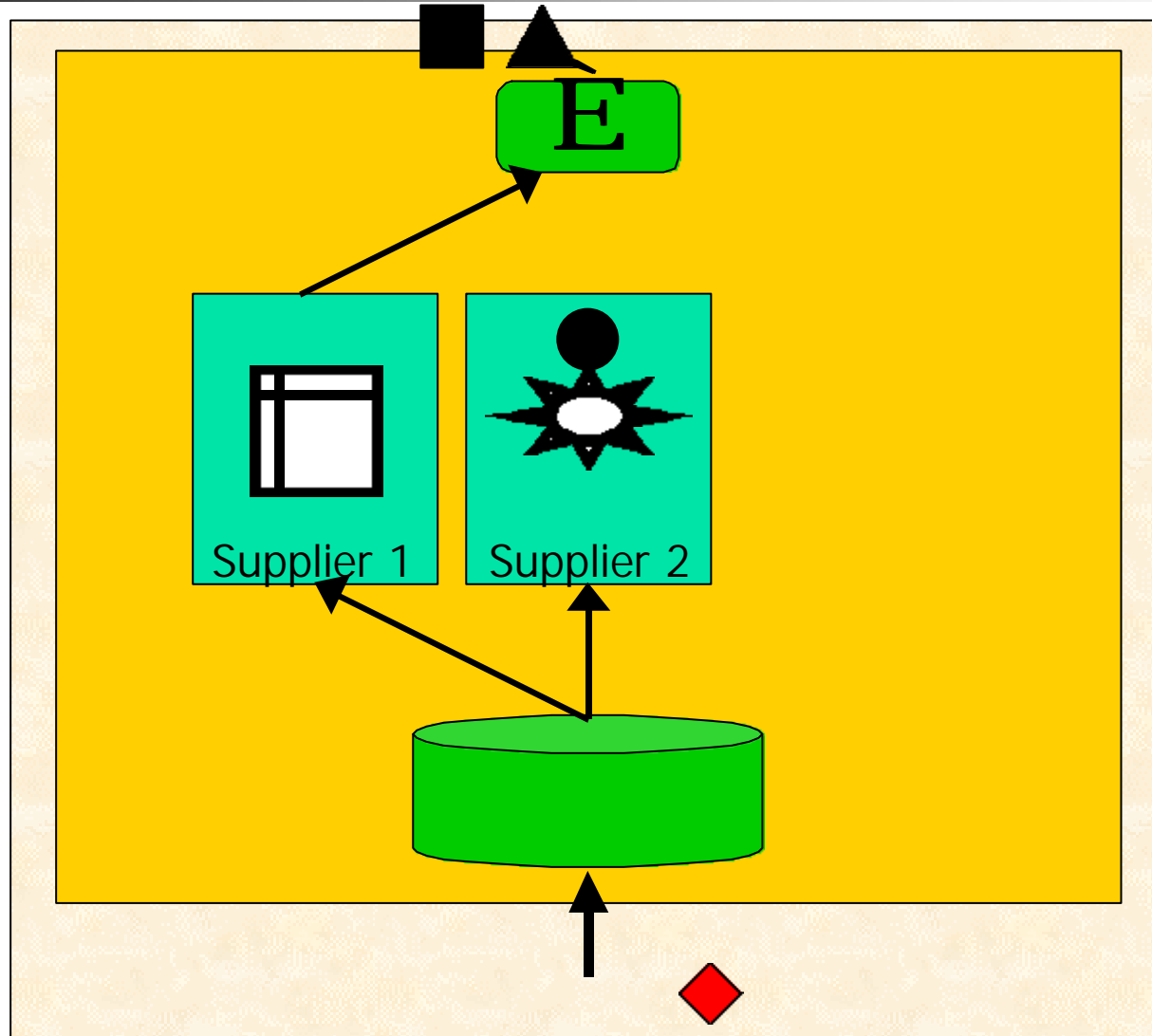
# HyperQuery Execution



# HyperQuery Execution

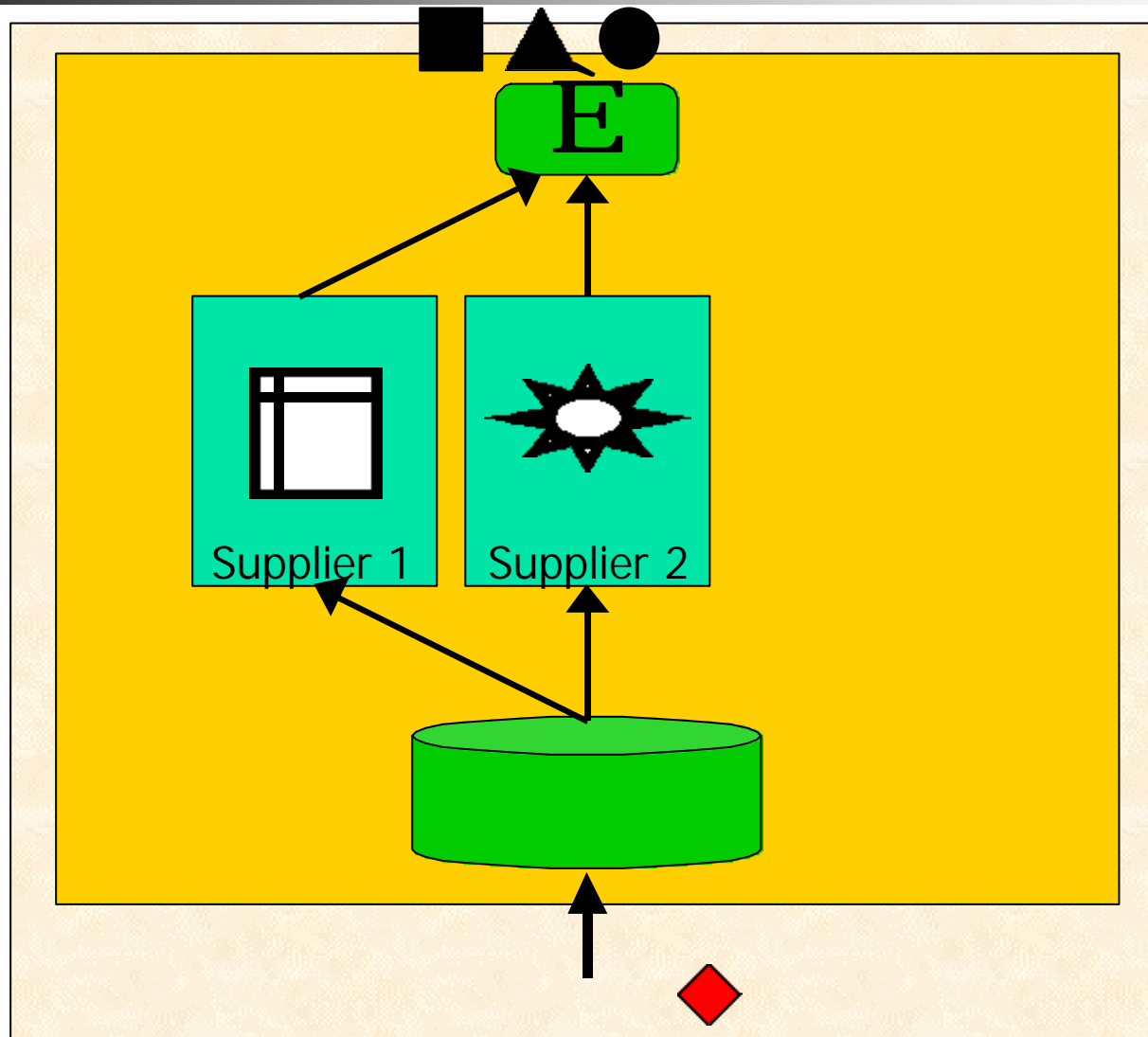


# HyperQuery Execution

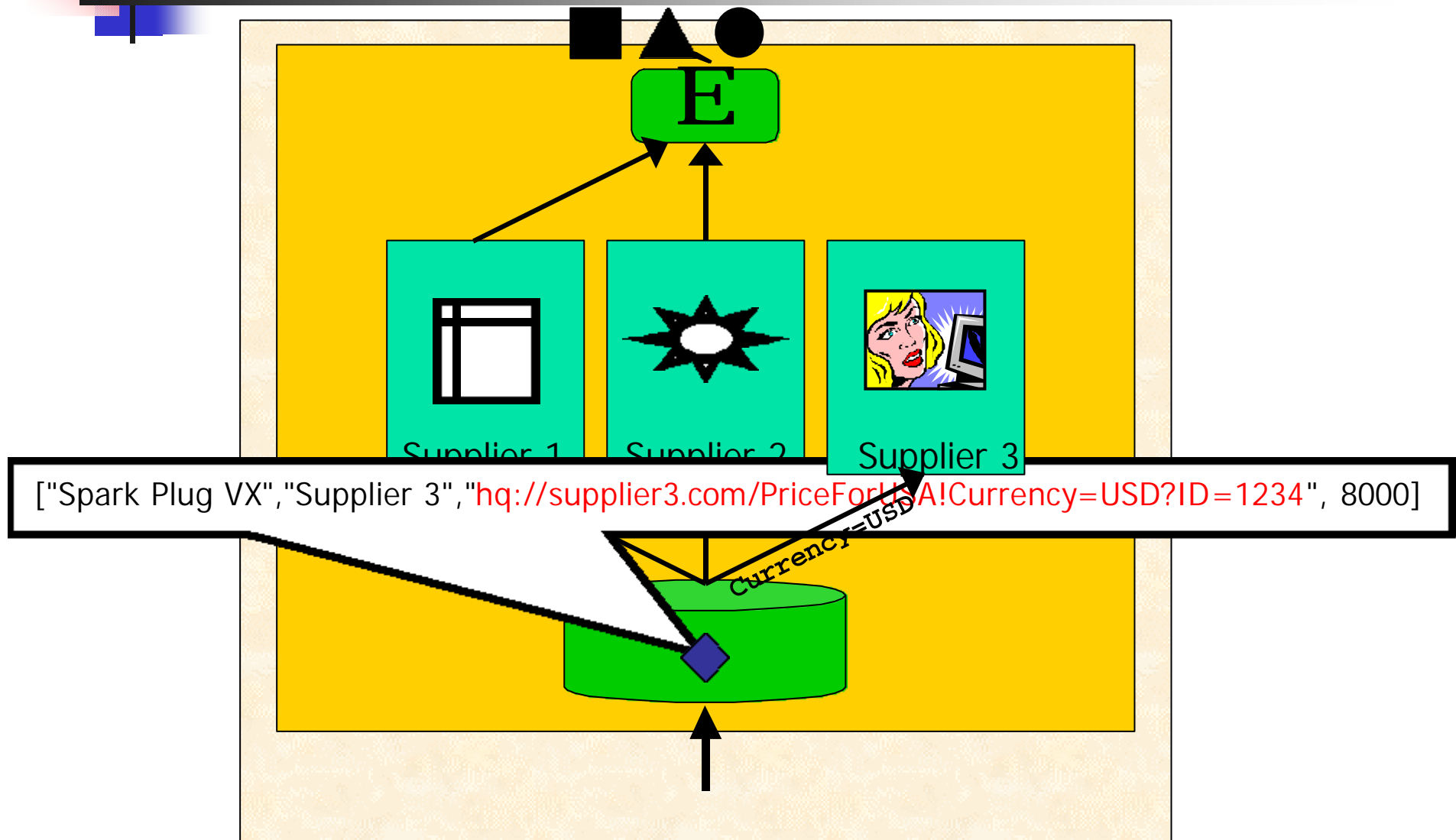




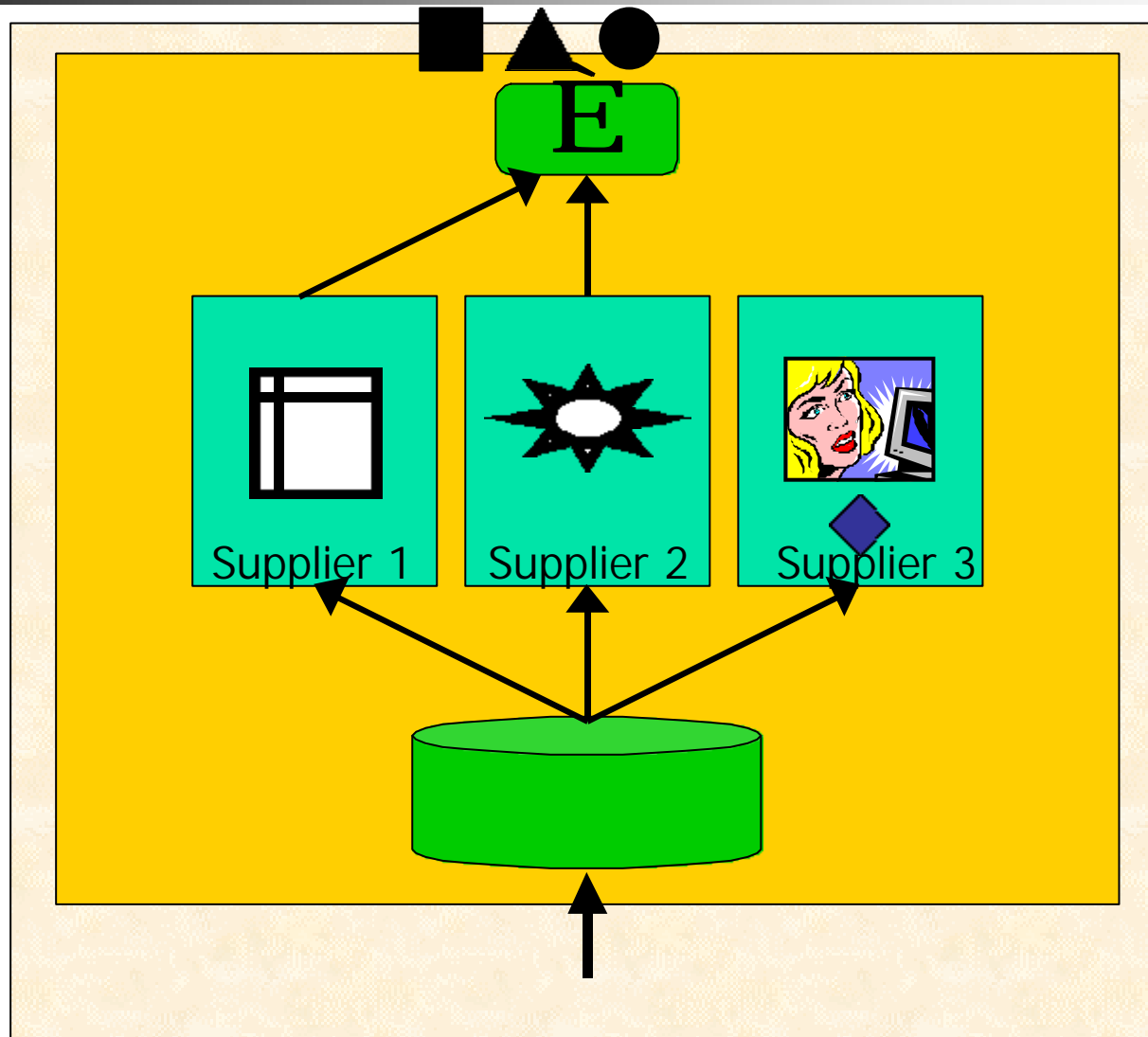
# HyperQuery Execution



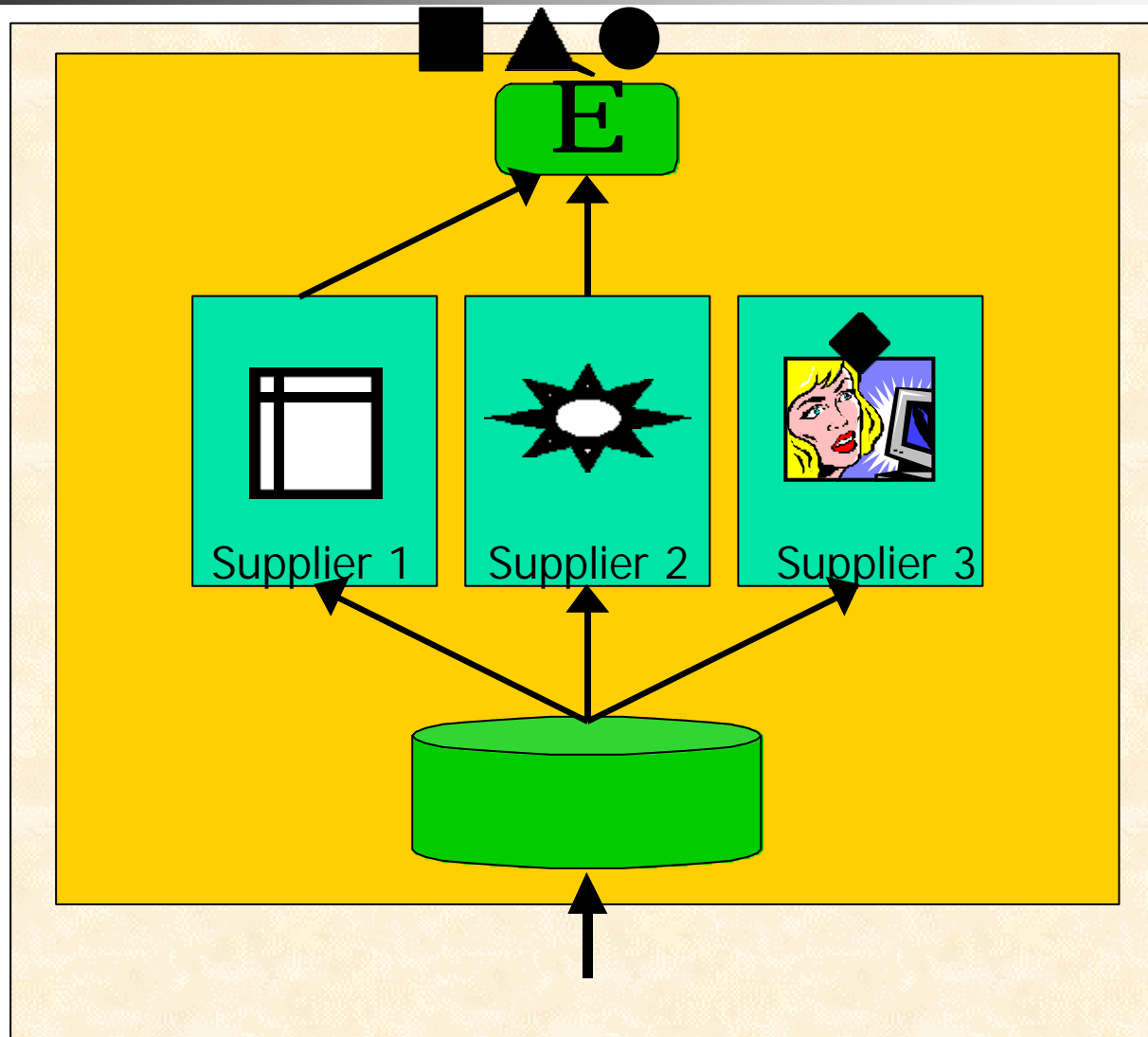
# HyperQuery Execution



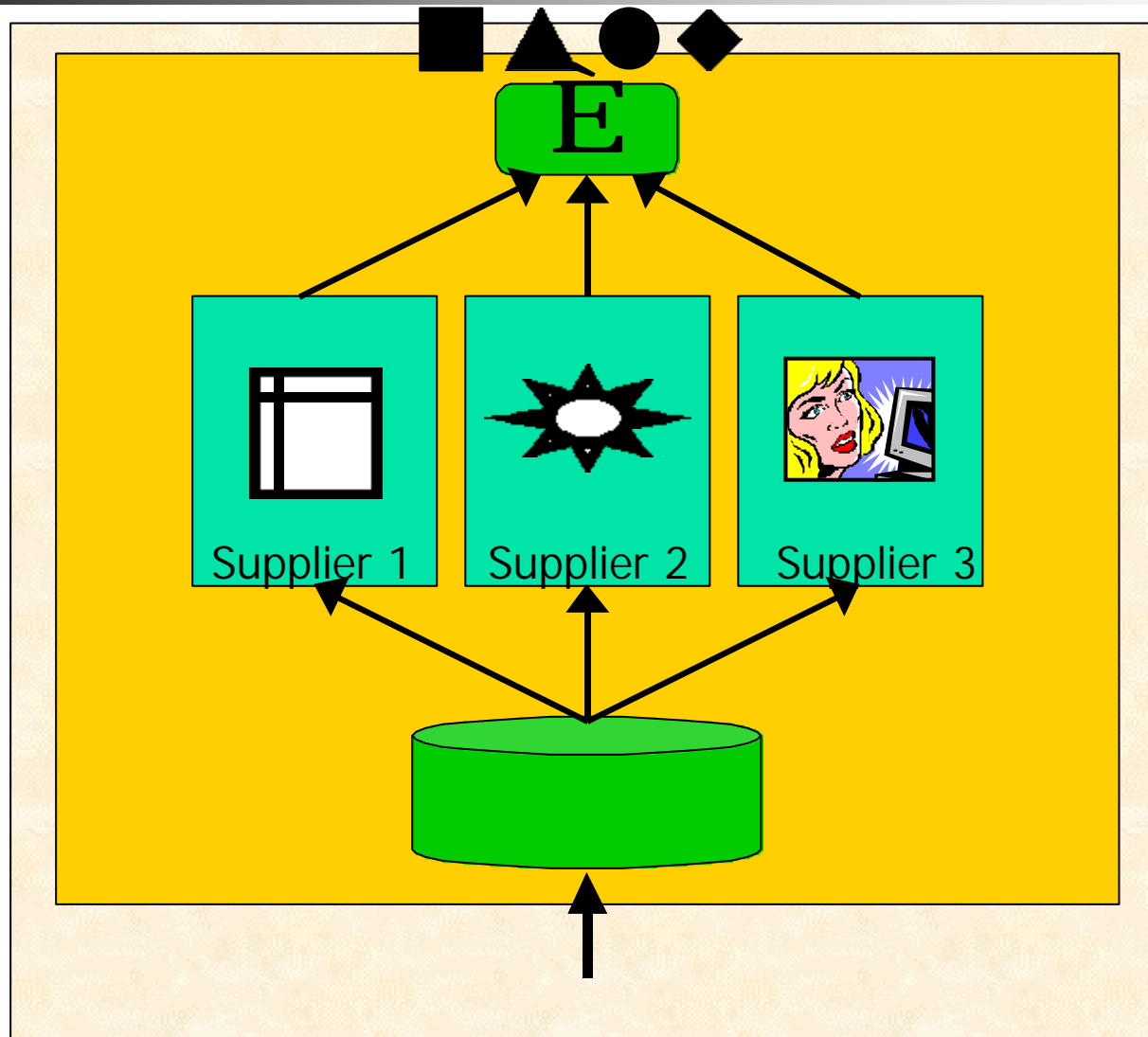
# HyperQuery Execution



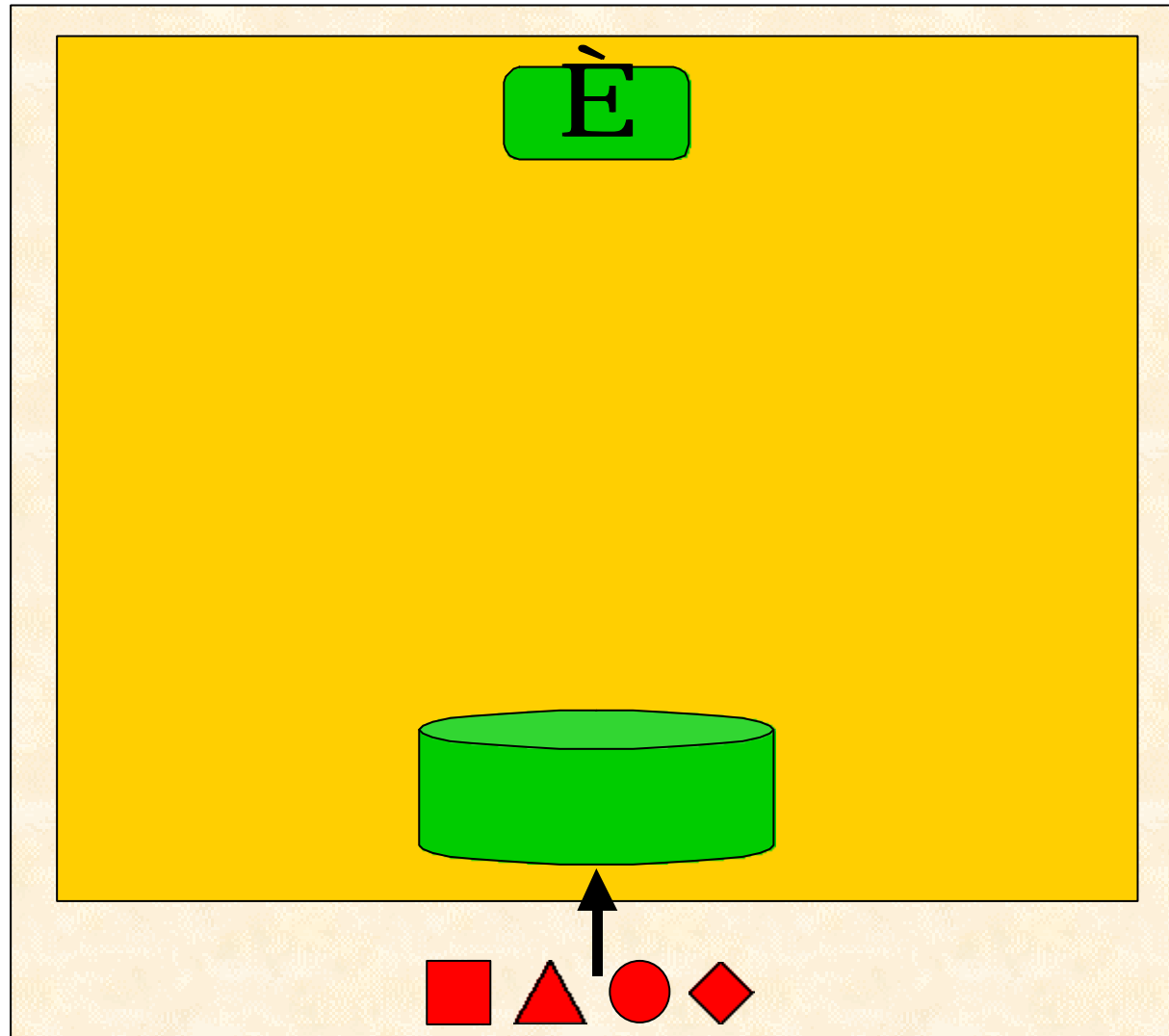
# HyperQuery Execution



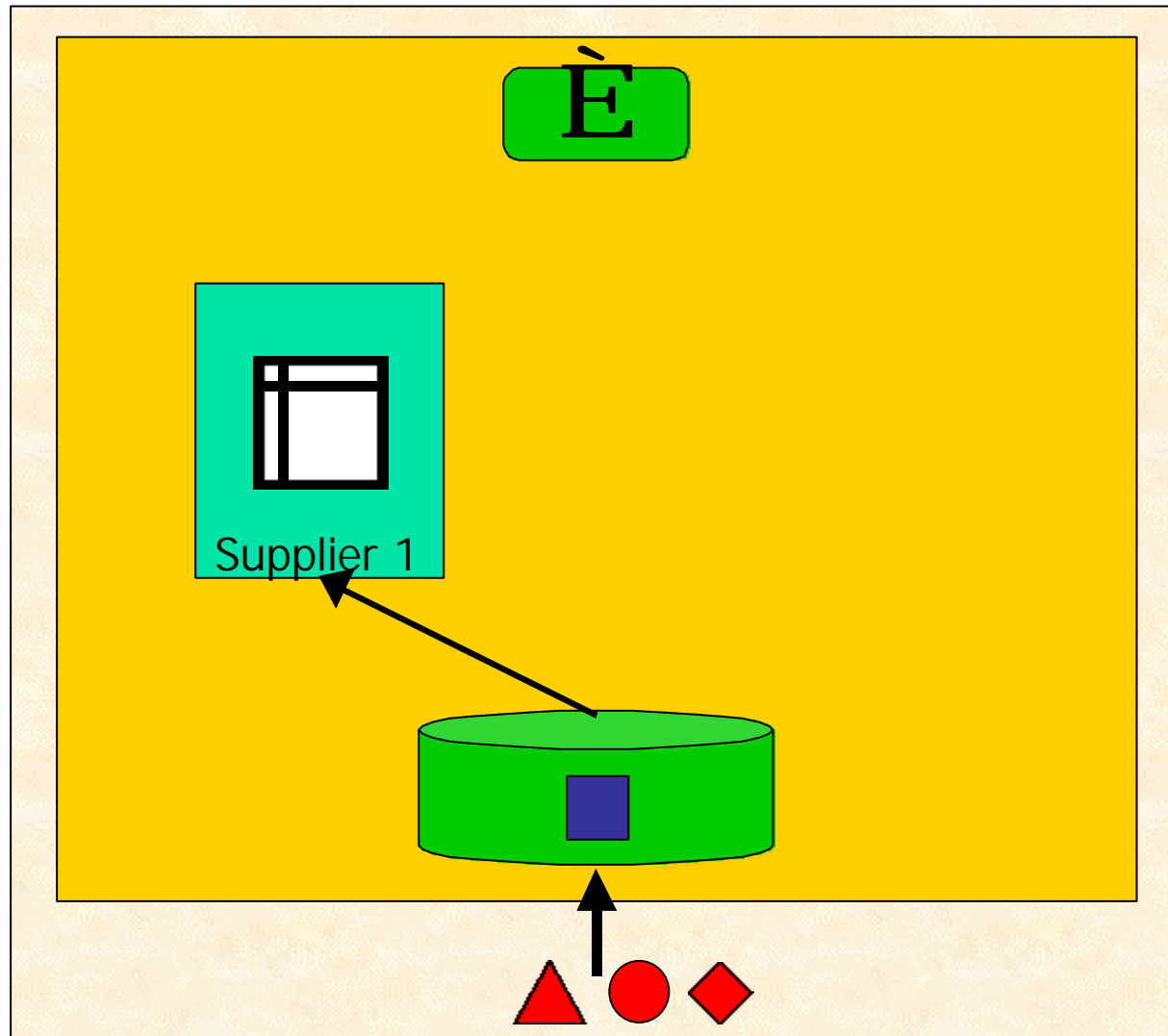
# HyperQuery Execution



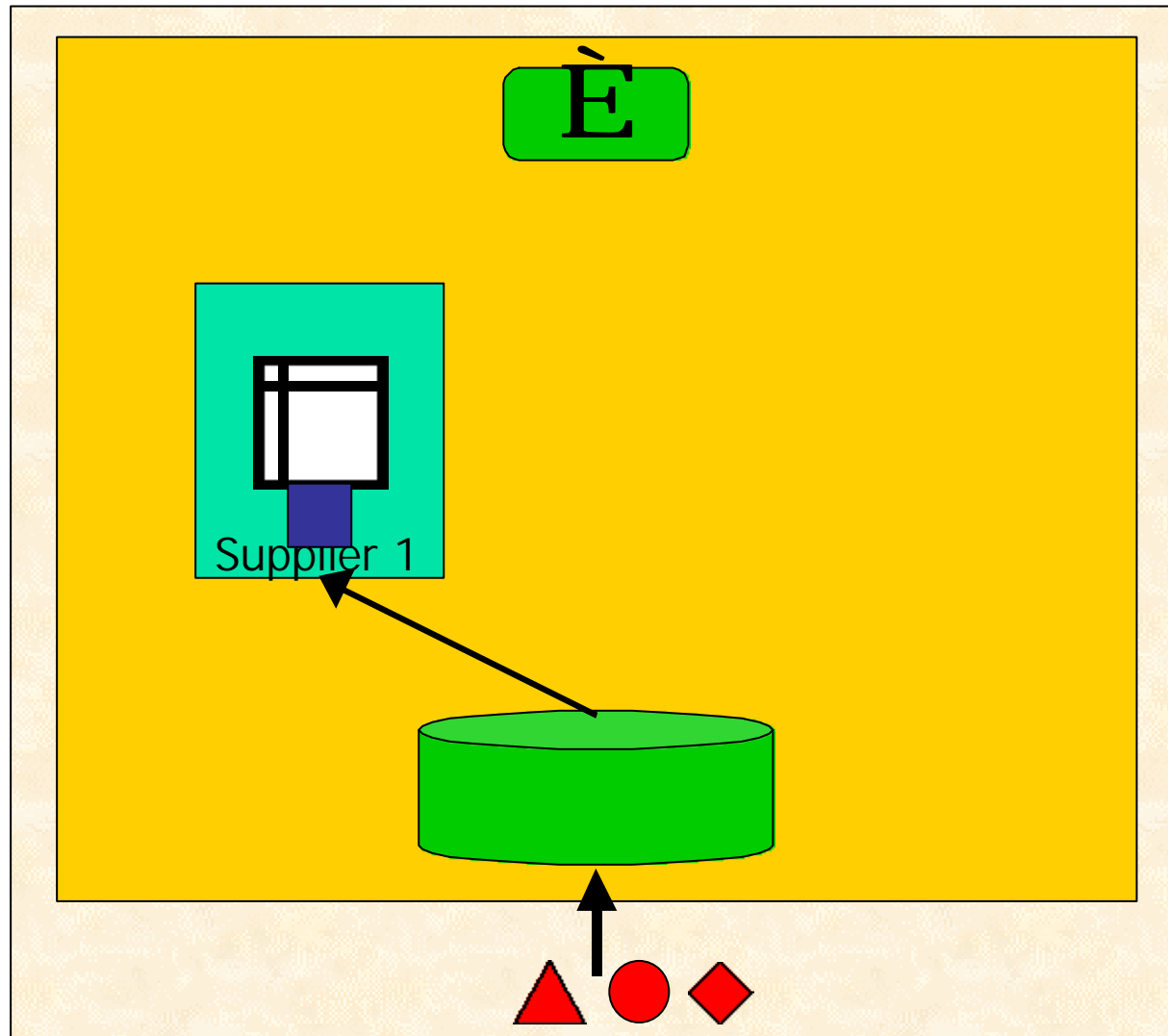
# Real HyperQuery Execution



# Real HyperQuery Execution

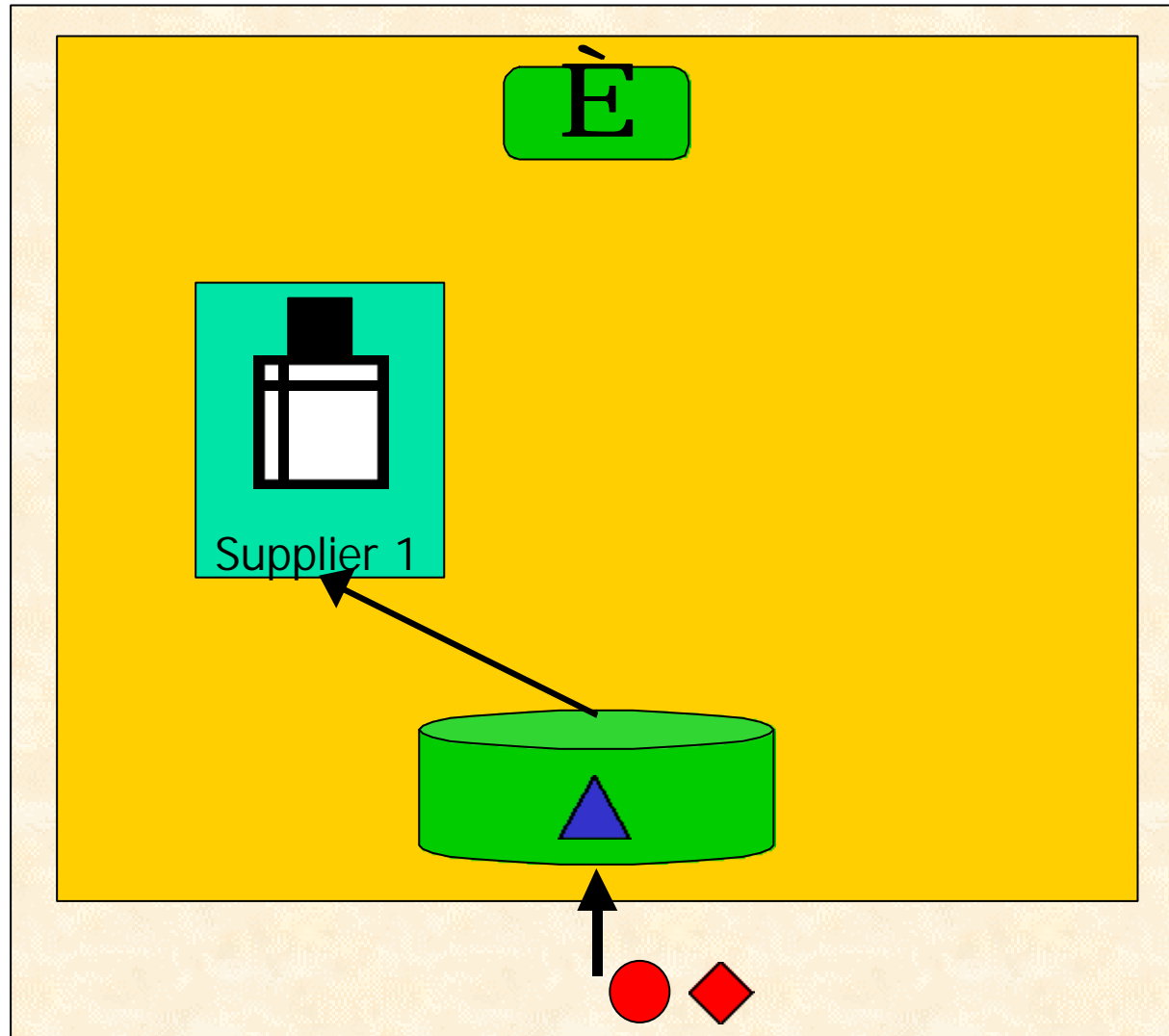


# Real HyperQuery Execution

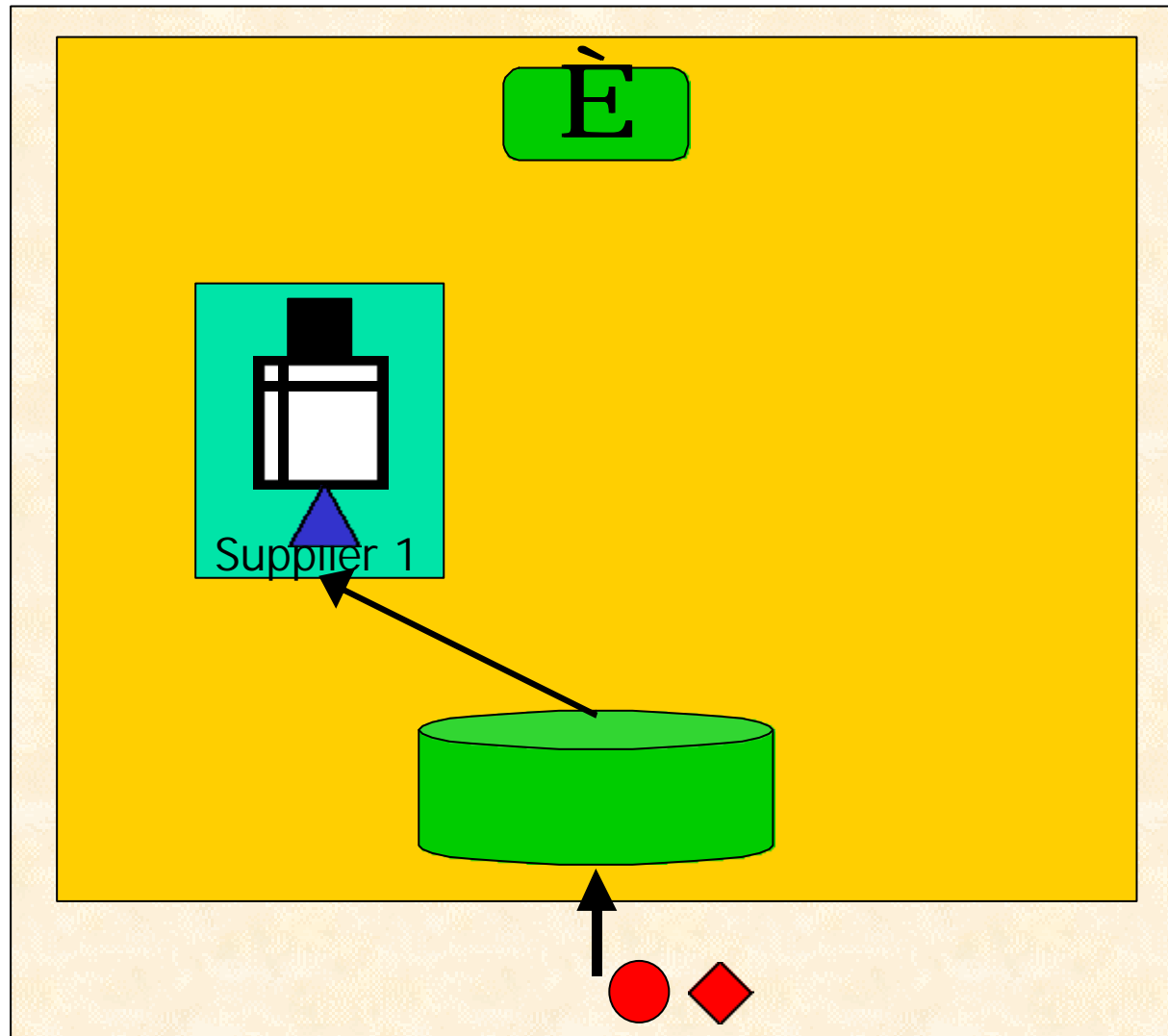




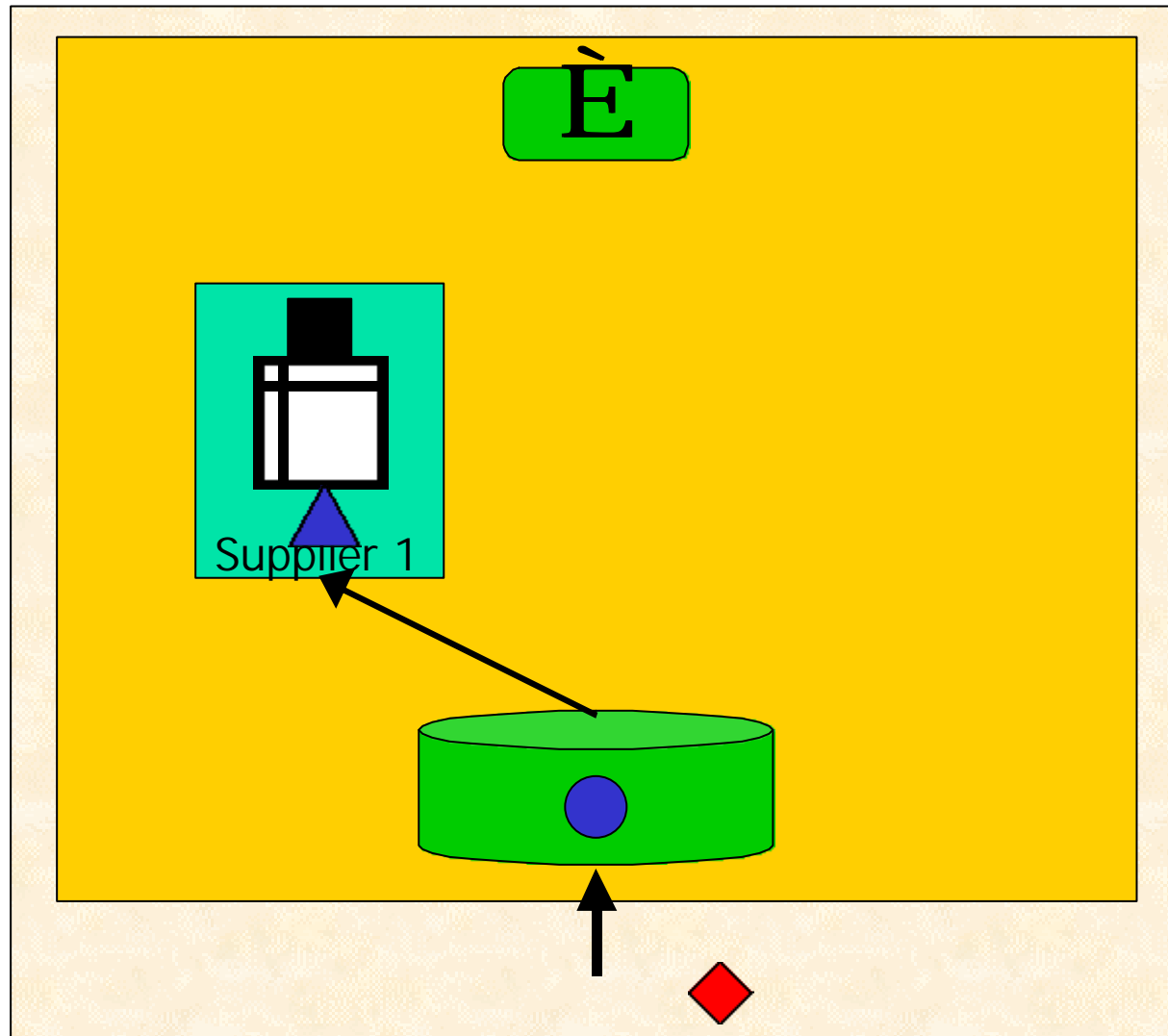
# Real HyperQuery Execution



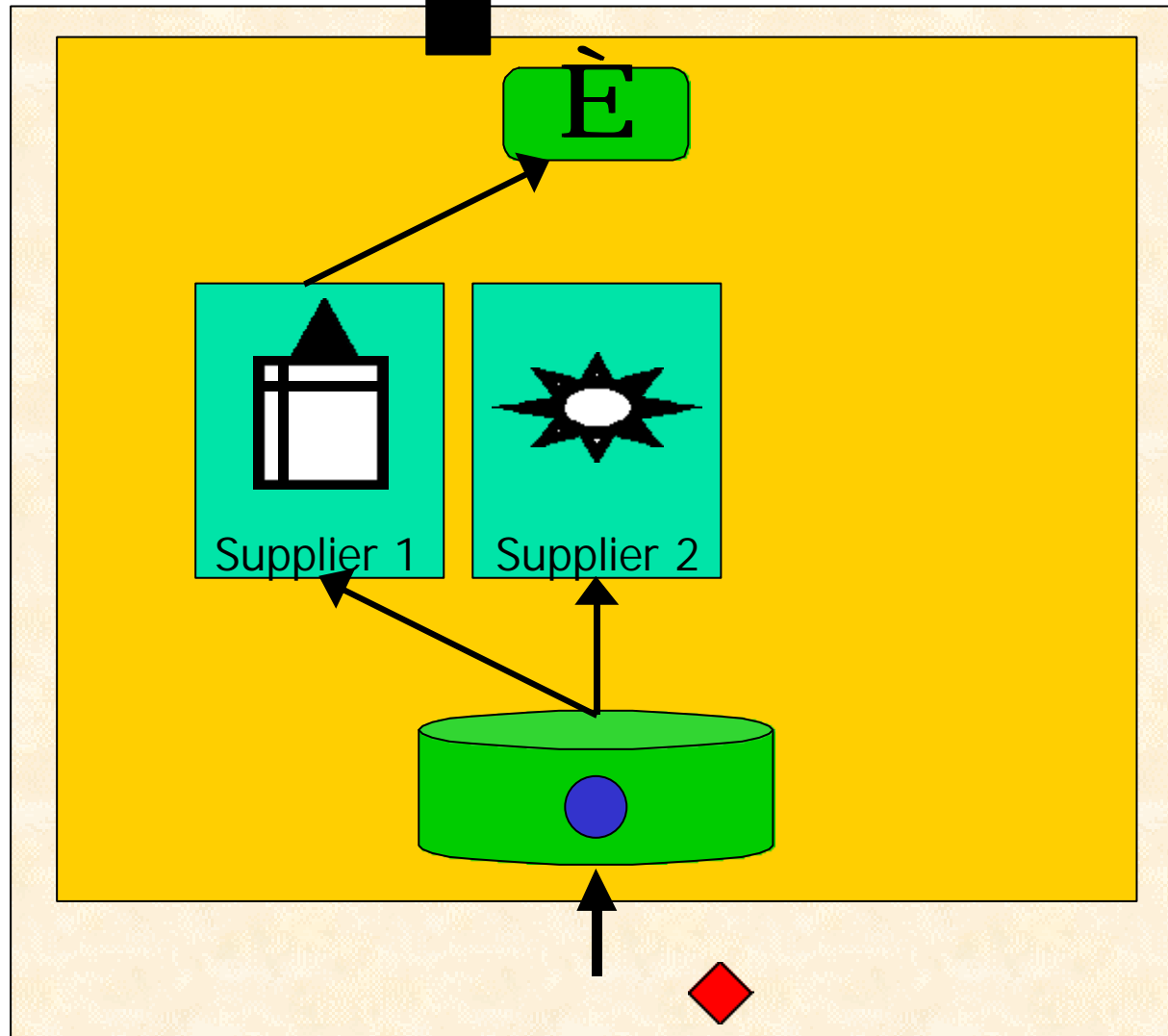
# Real HyperQuery Execution



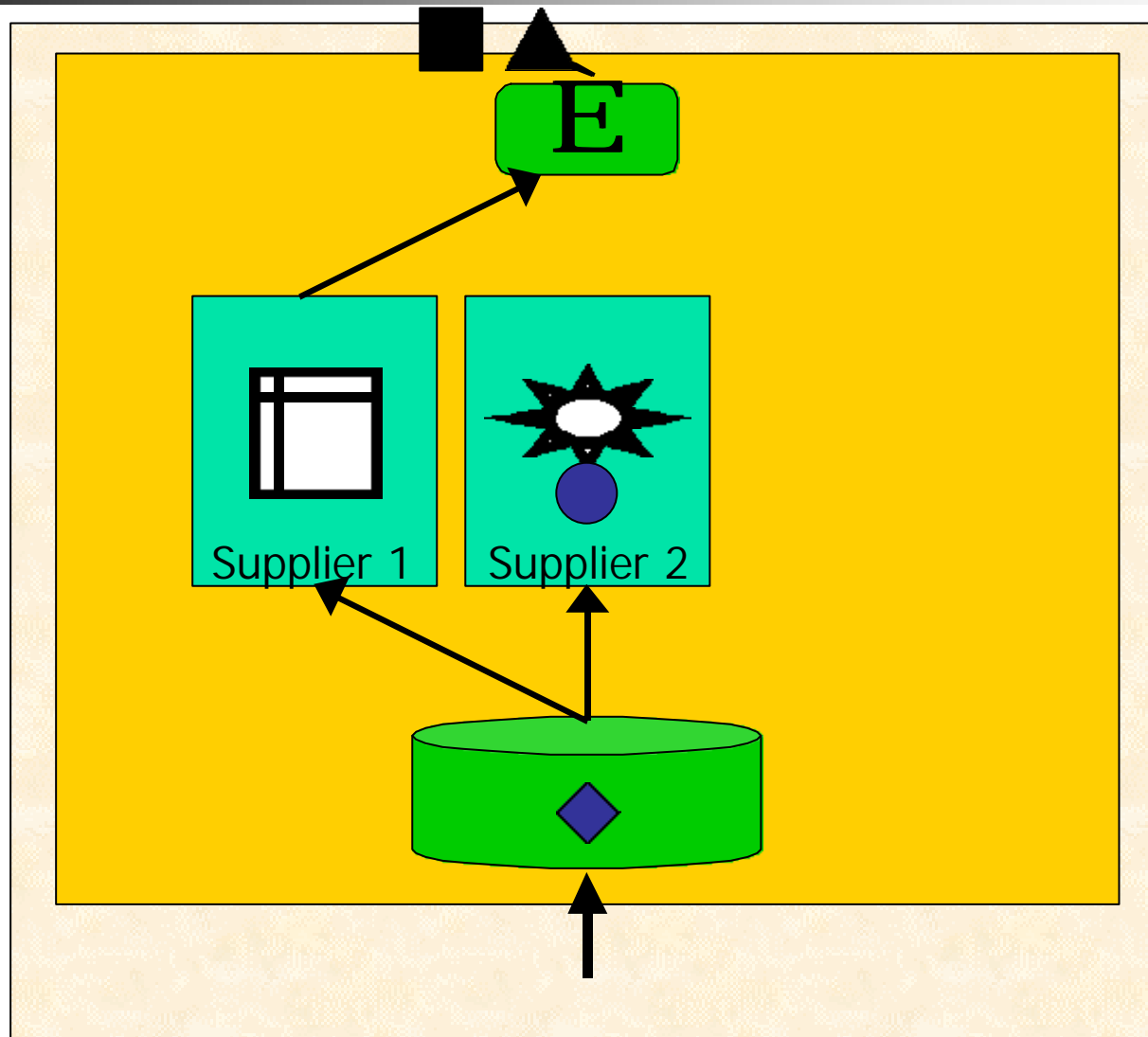
# Real HyperQuery Execution



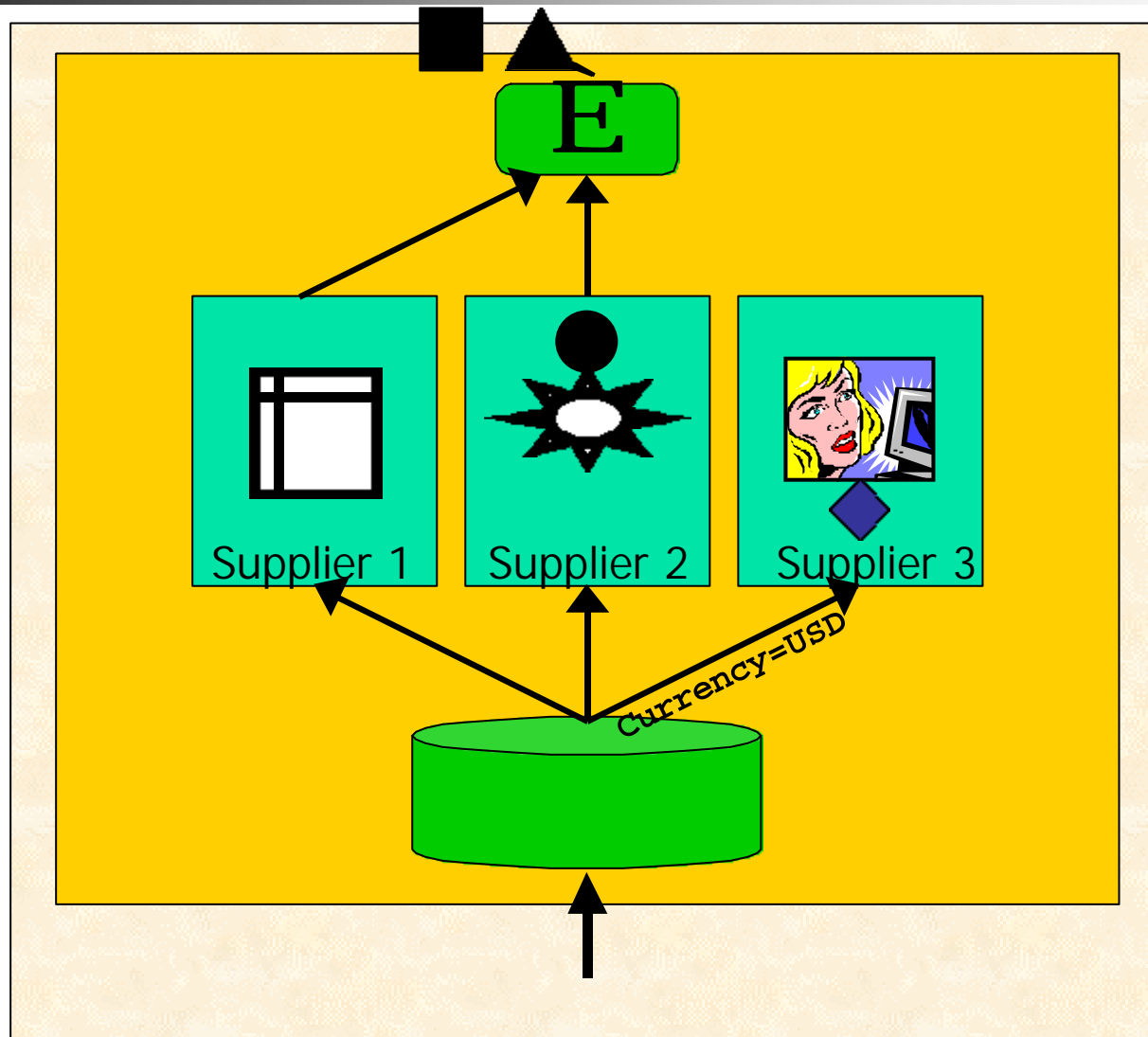
# Real HyperQuery Execution



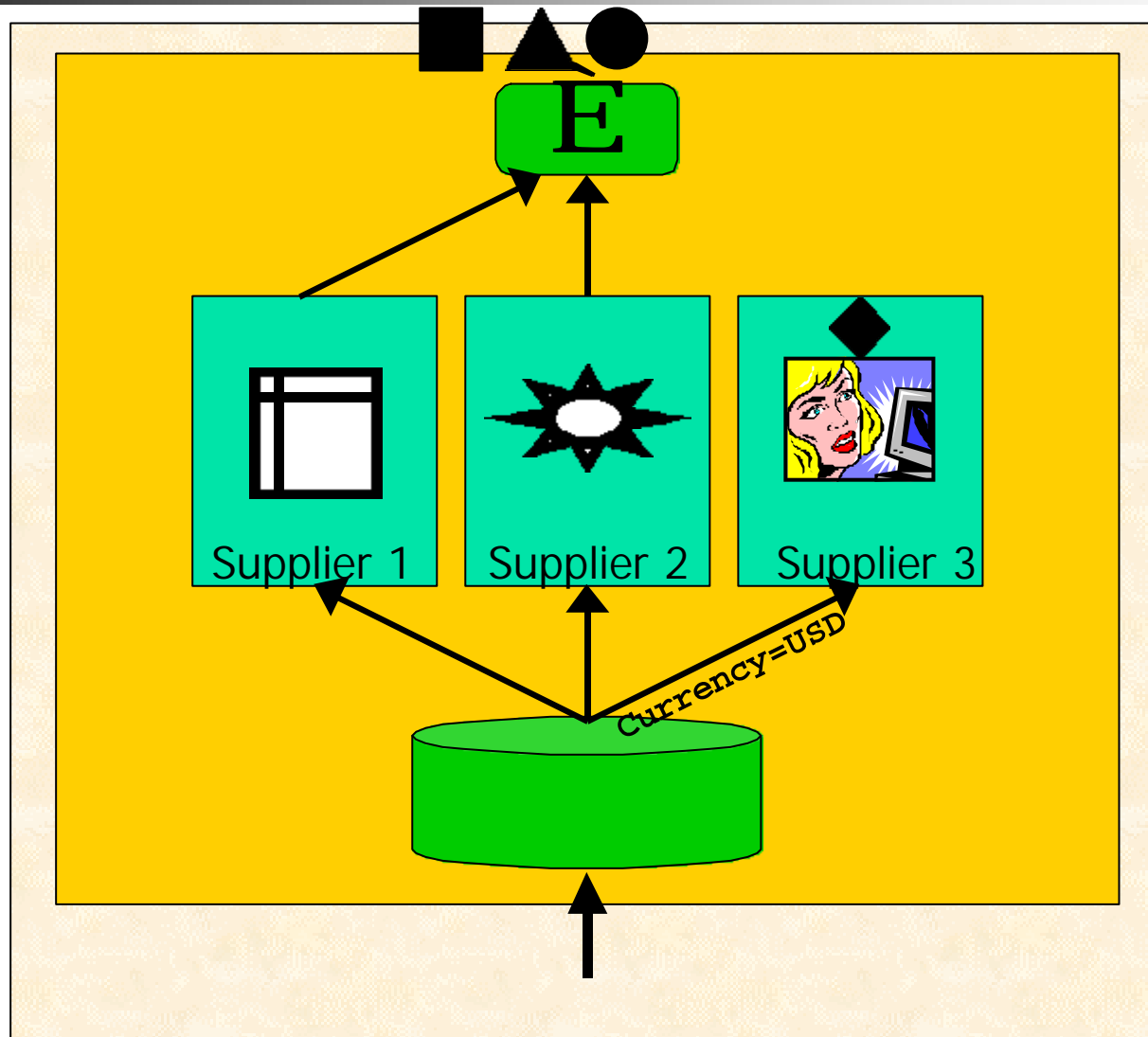
# Real HyperQuery Execution



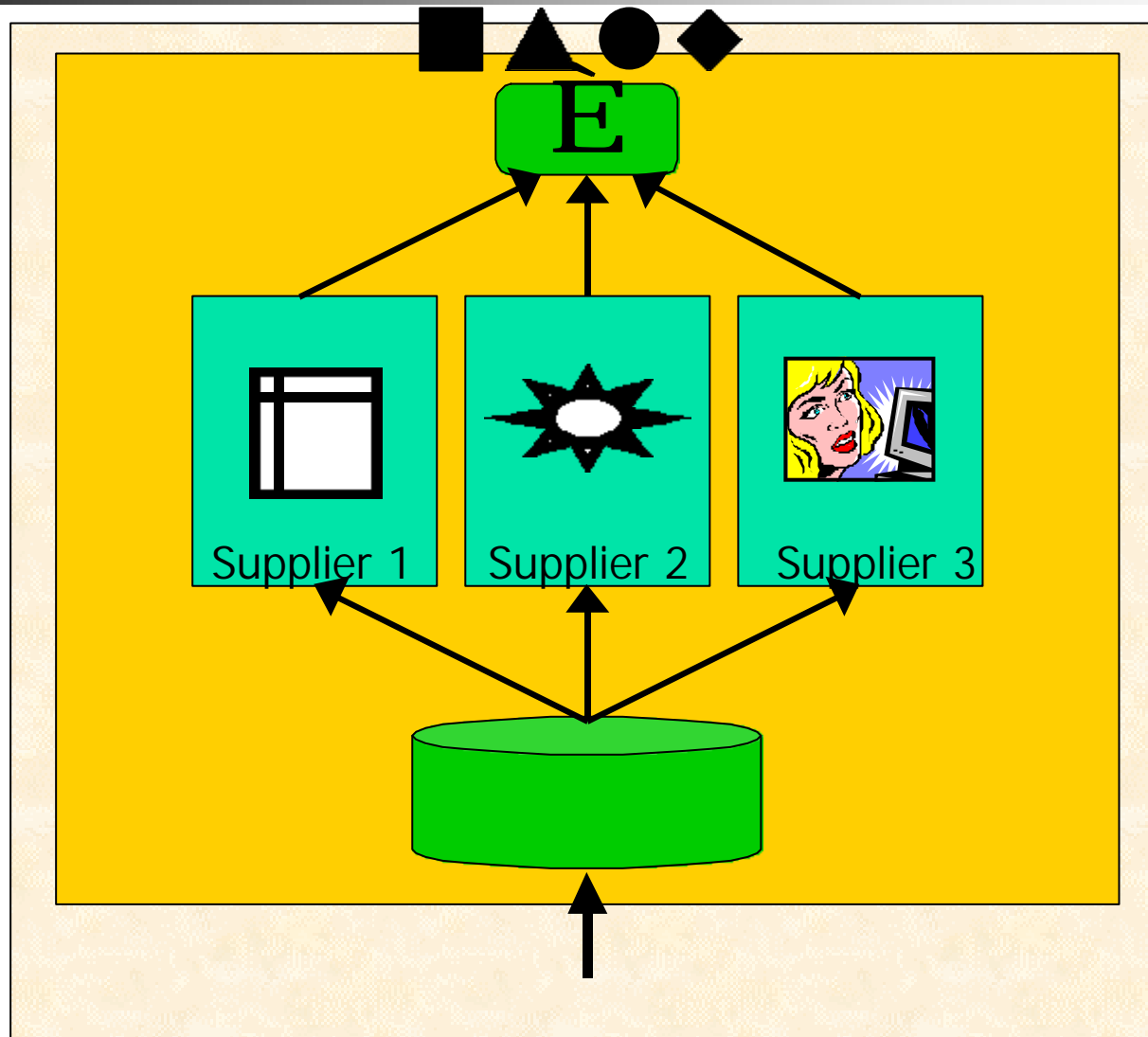
# Real HyperQuery Execution



# Real HyperQuery Execution

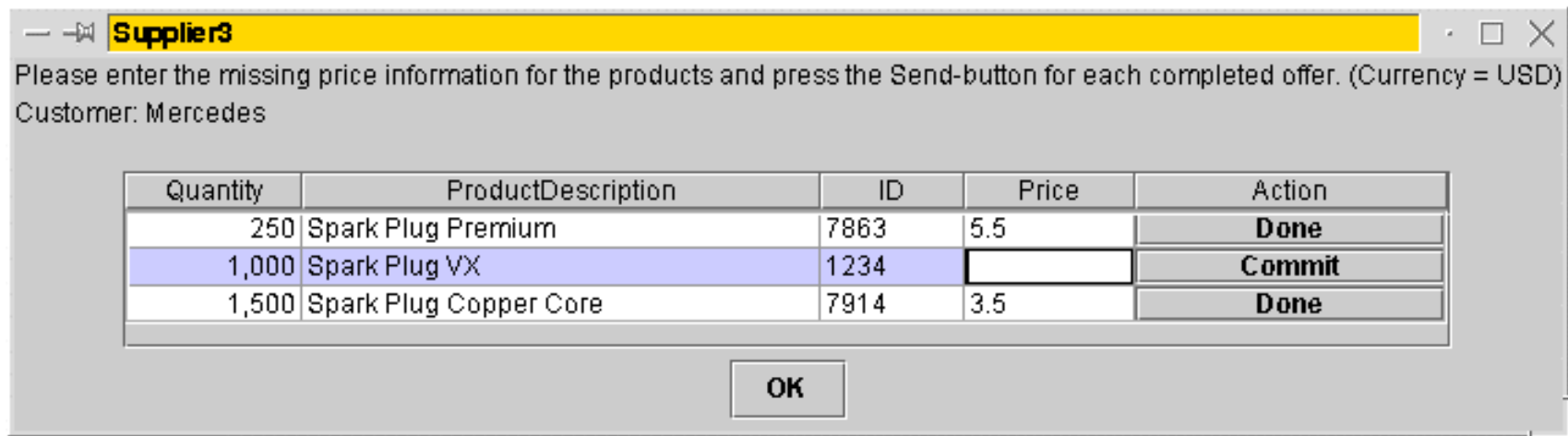


# Real HyperQuery Execution





# Origin of Data



Supplier3

Please enter the missing price information for the products and press the Send-button for each completed offer. (Currency = USD)  
Customer: Mercedes

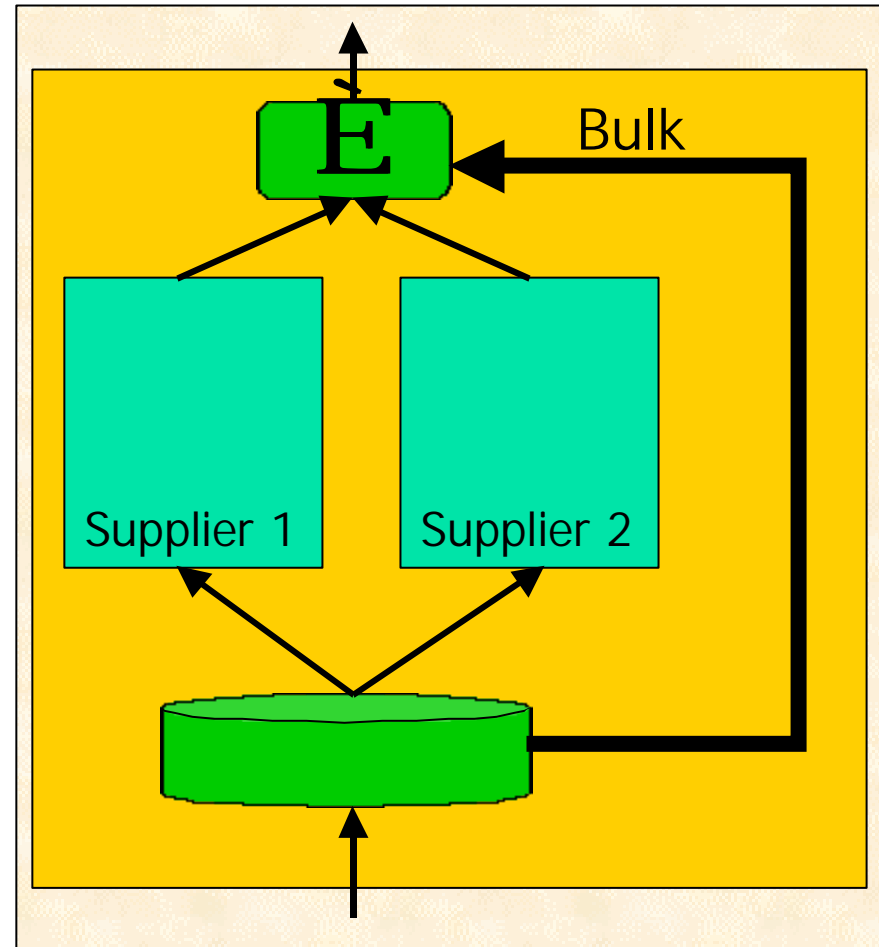
Quantity	ProductDescription	ID	Price	Action
250	Spark Plug Premium	7863	5.5	Done
1,000	Spark Plug VX	1234		Commit
1,500	Spark Plug Copper Core	7914	3.5	Done

OK

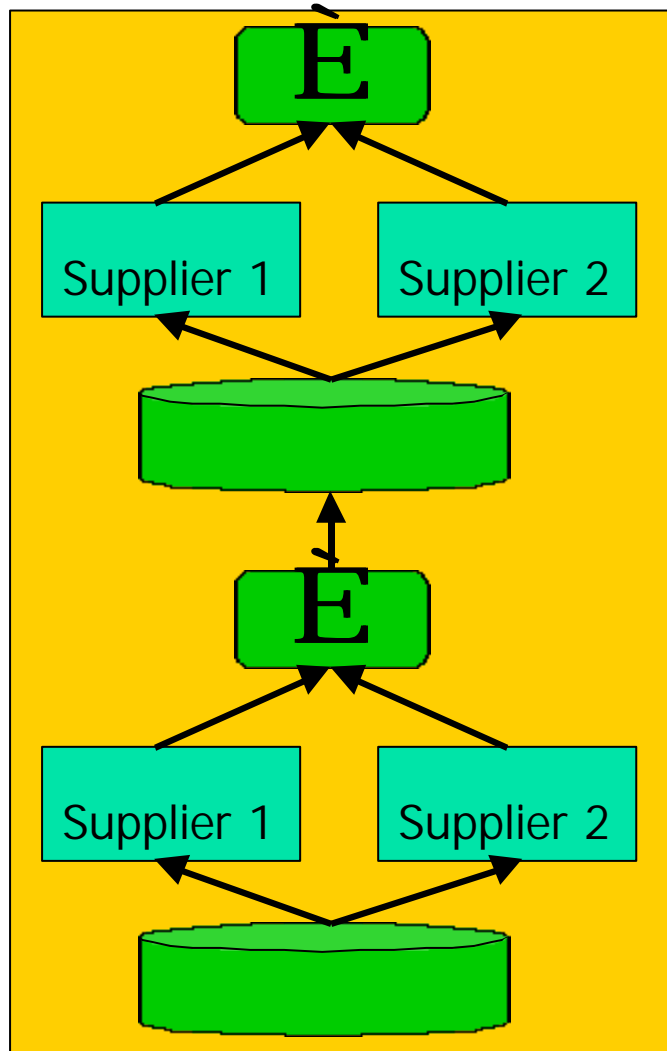
- Human interaction

# Optimization: Bulk Bypassing

- Strip off bulky attributes
- Bypass Sub-plans
- Re-merge
- Reduce network traffic



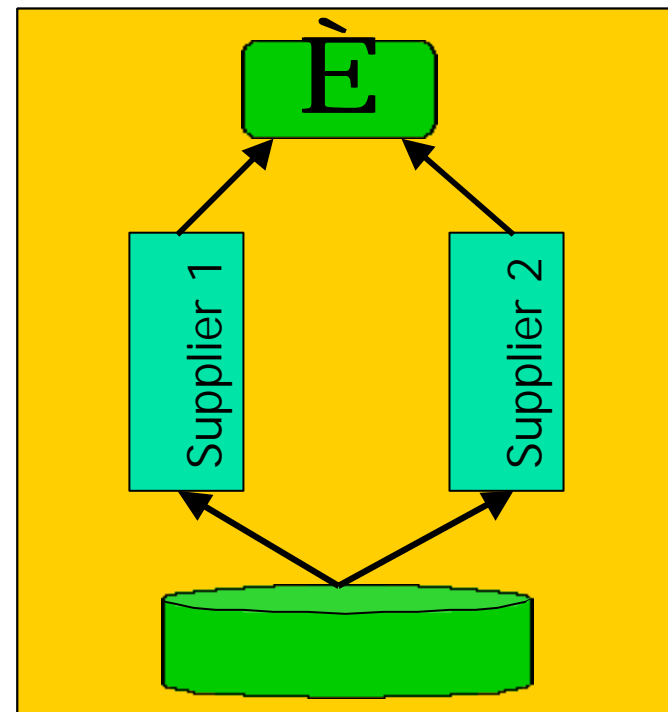
# Optimization: Multiple Virtual Attributes



⊗

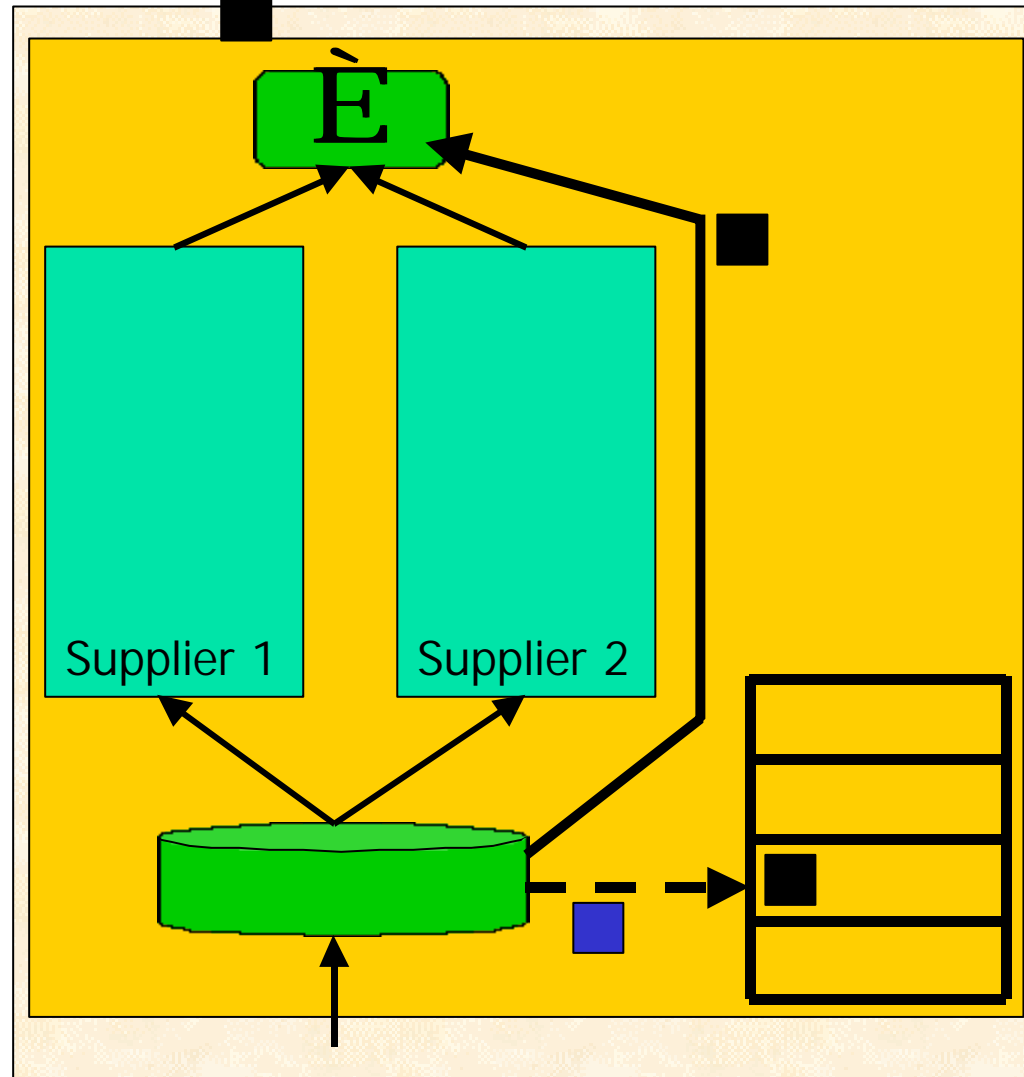
A v a i l a b i l

- Parallelize requests

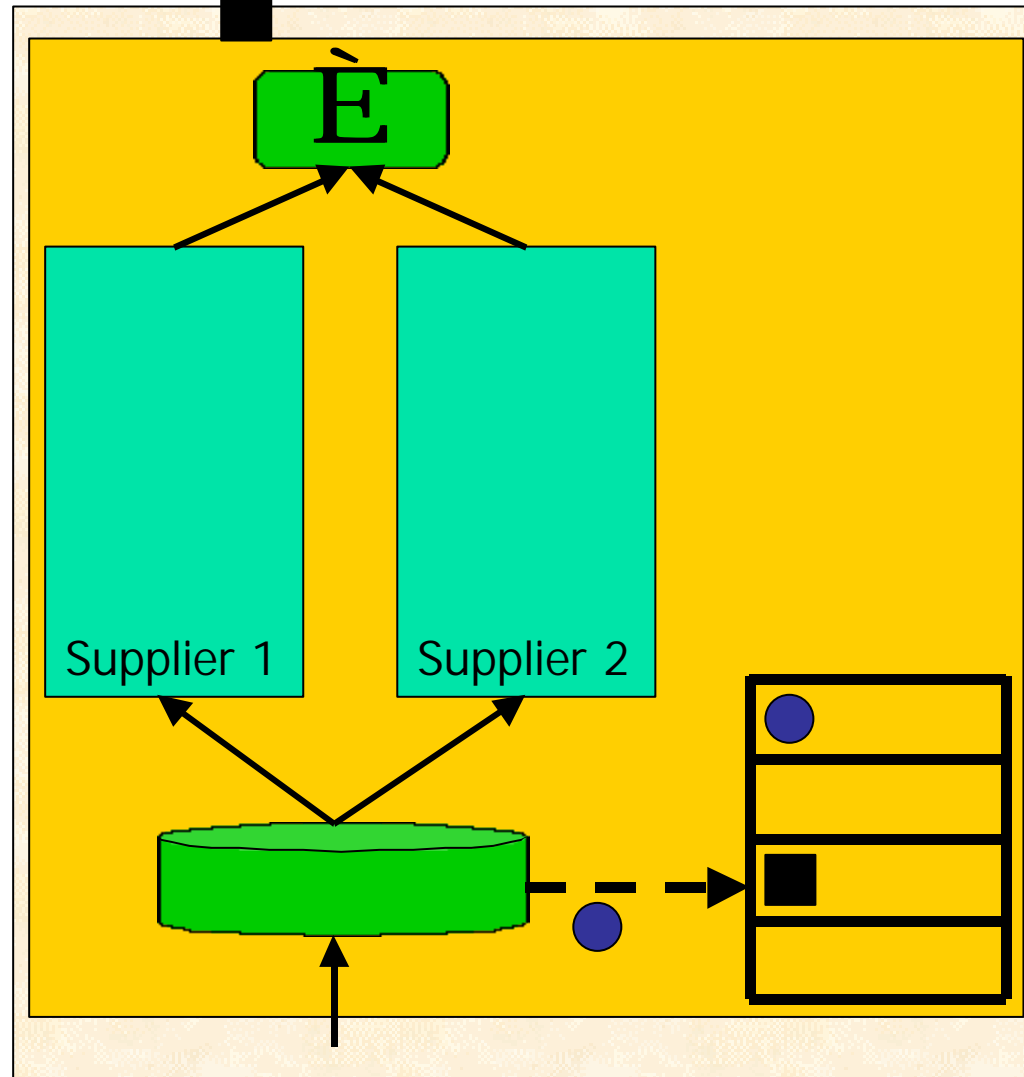


Availability & Price

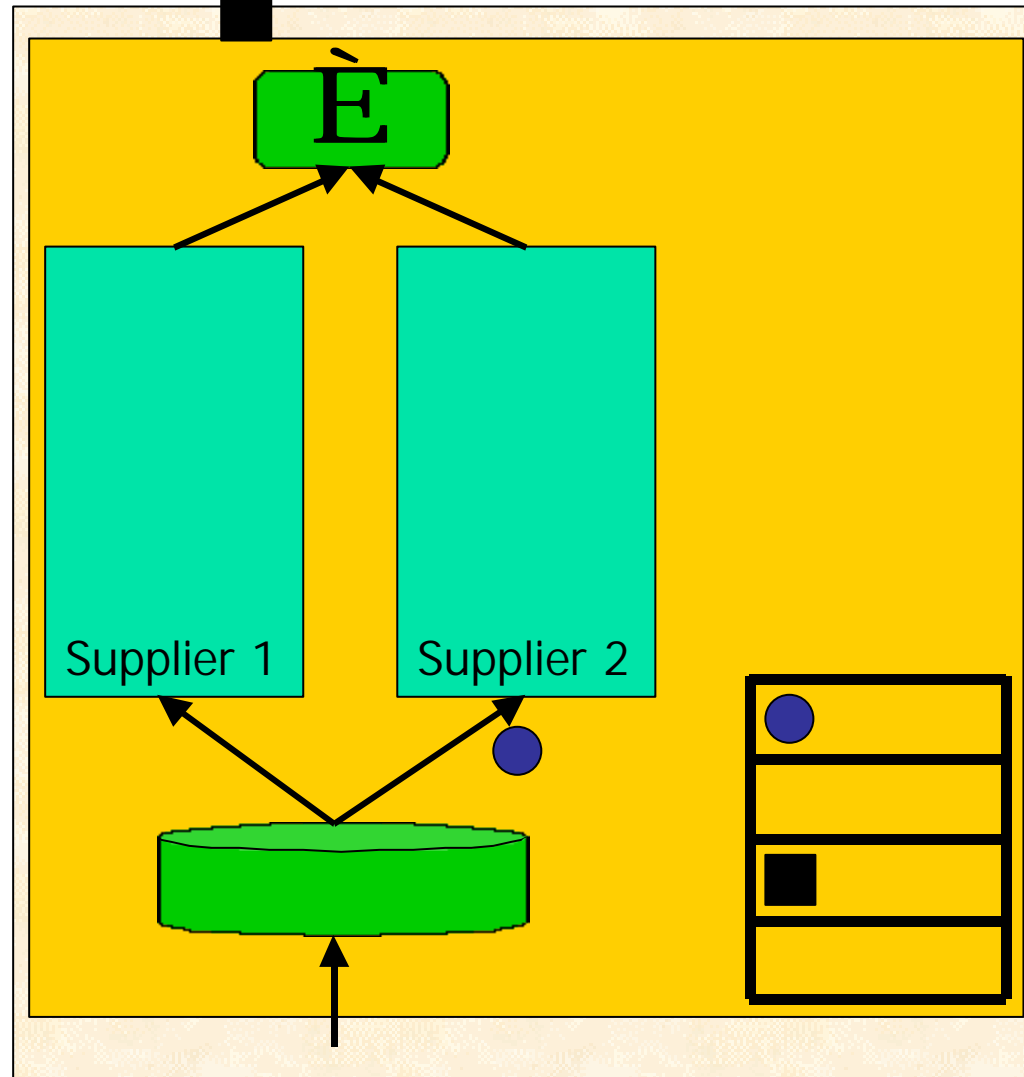
# Optimization: Caching



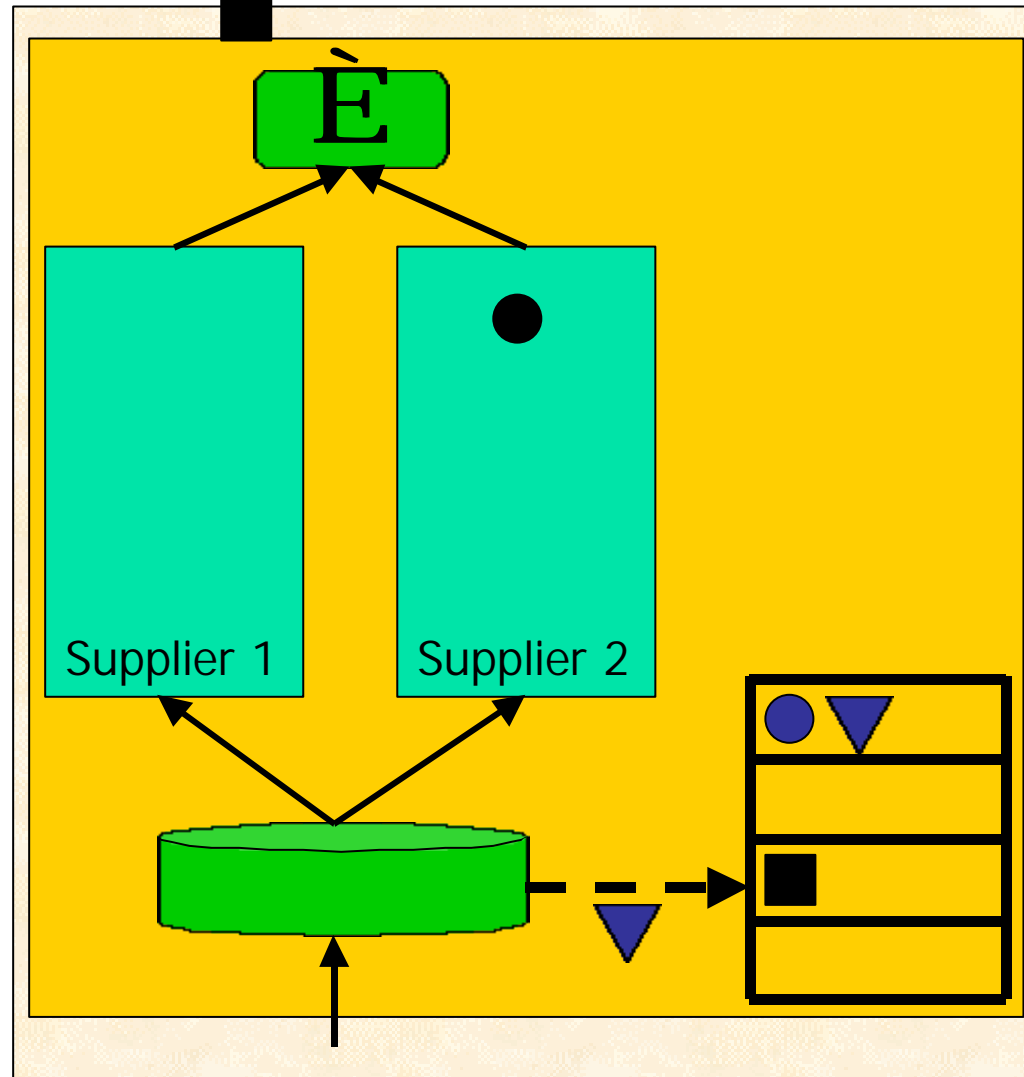
# Optimization: Caching



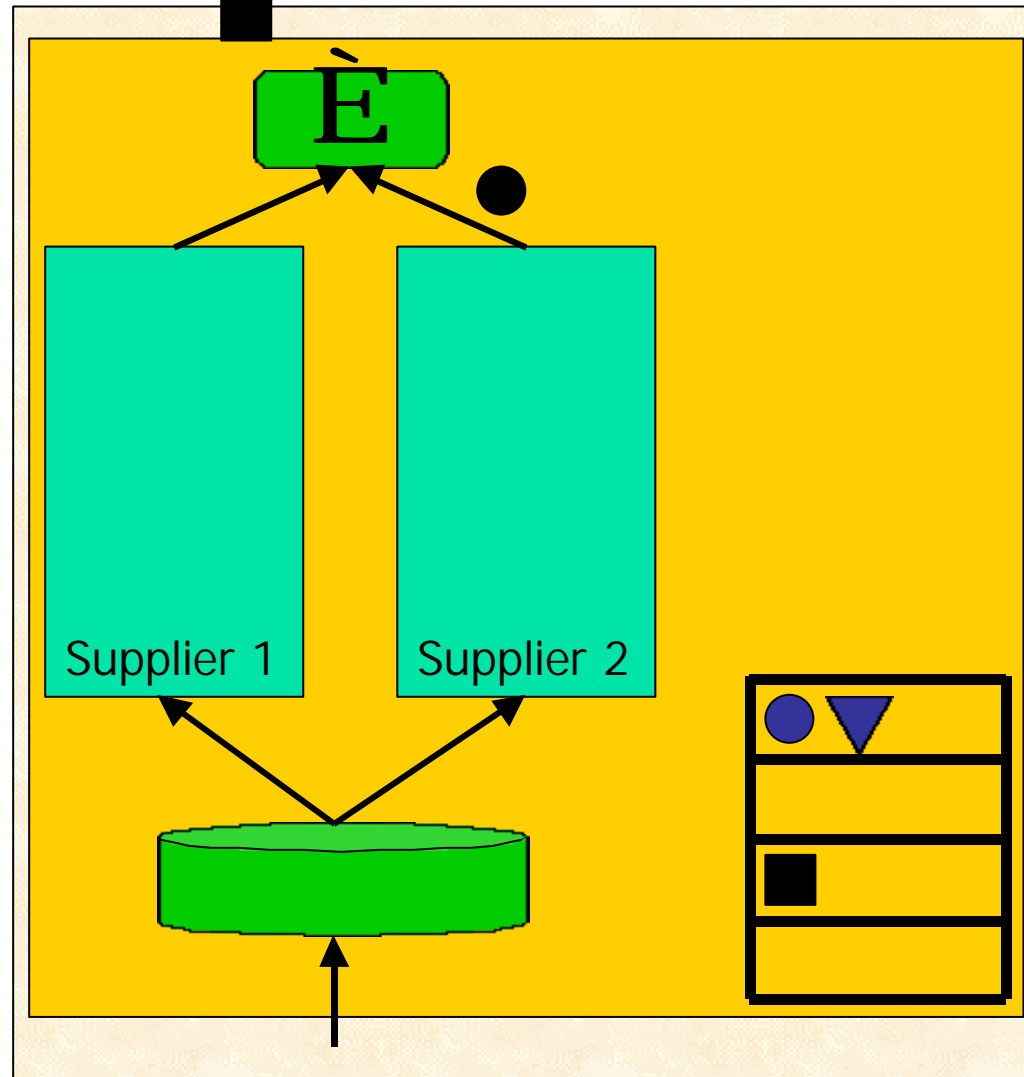
# Optimization: Caching



# Optimization: Caching

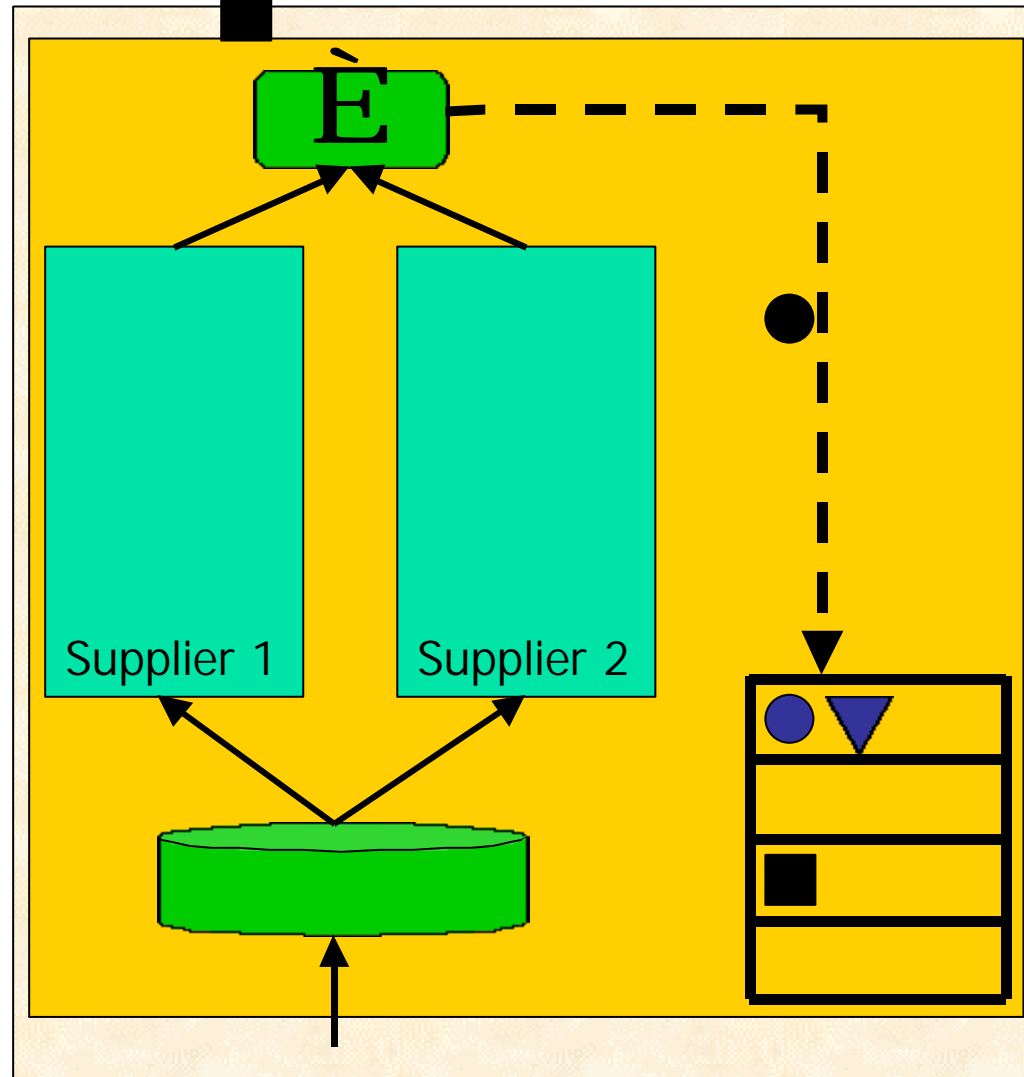


# Optimization: Caching

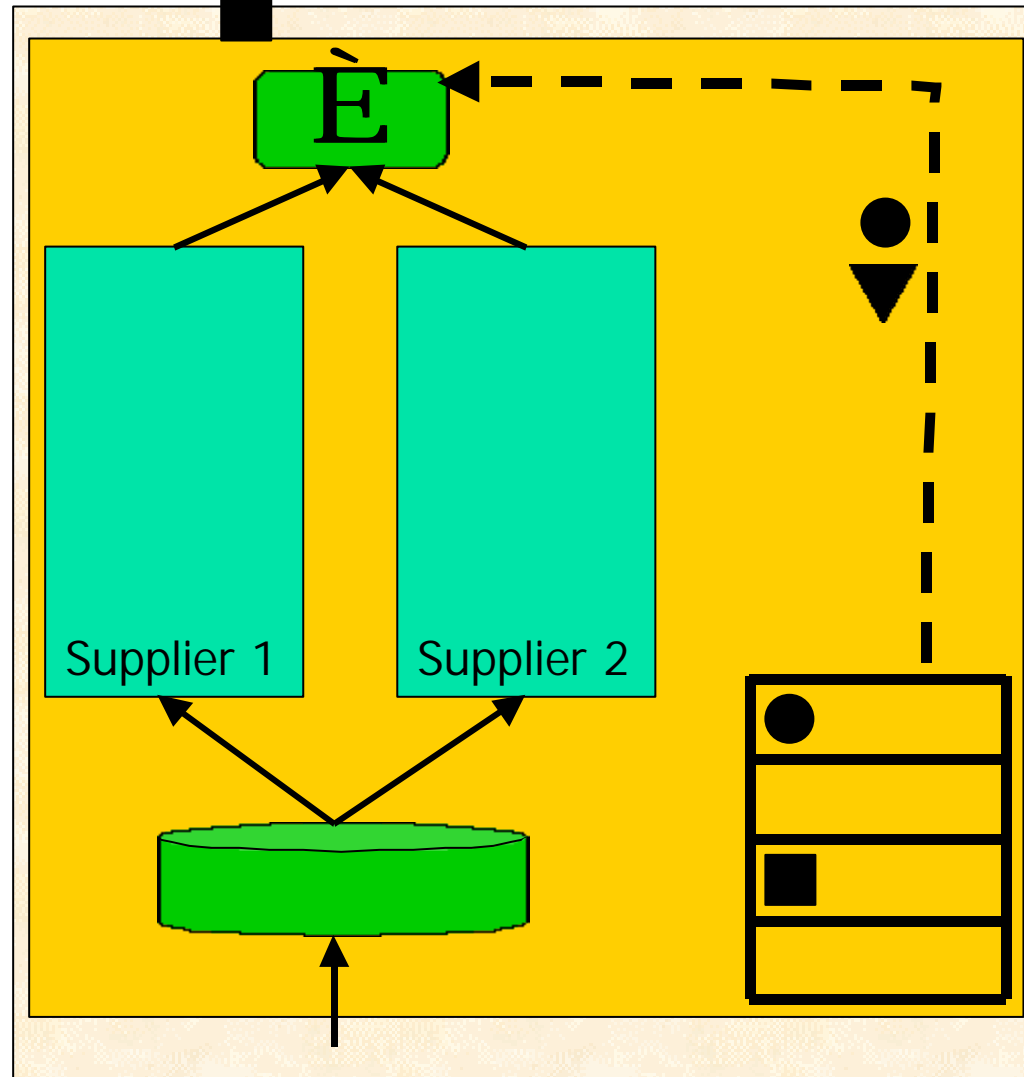




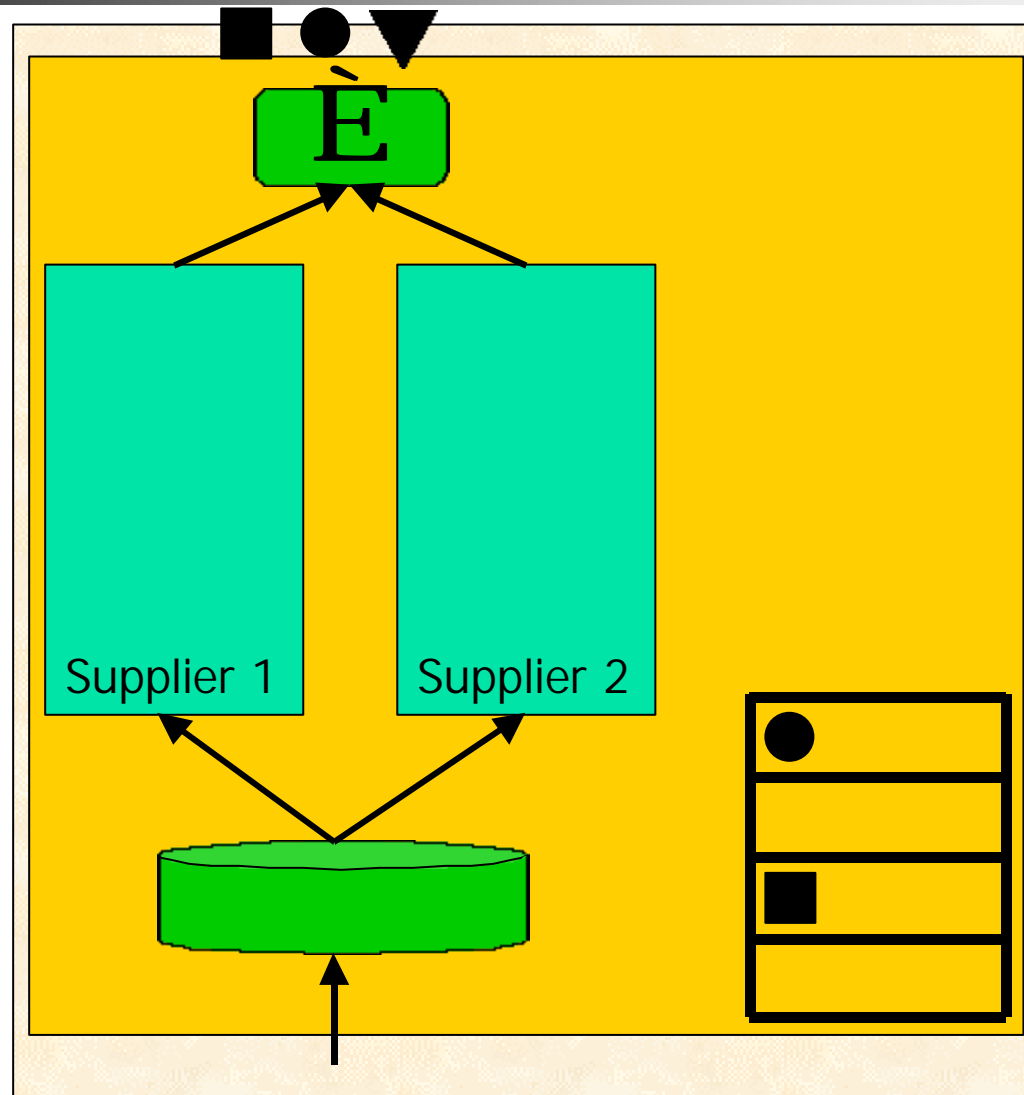
# Optimization: Caching



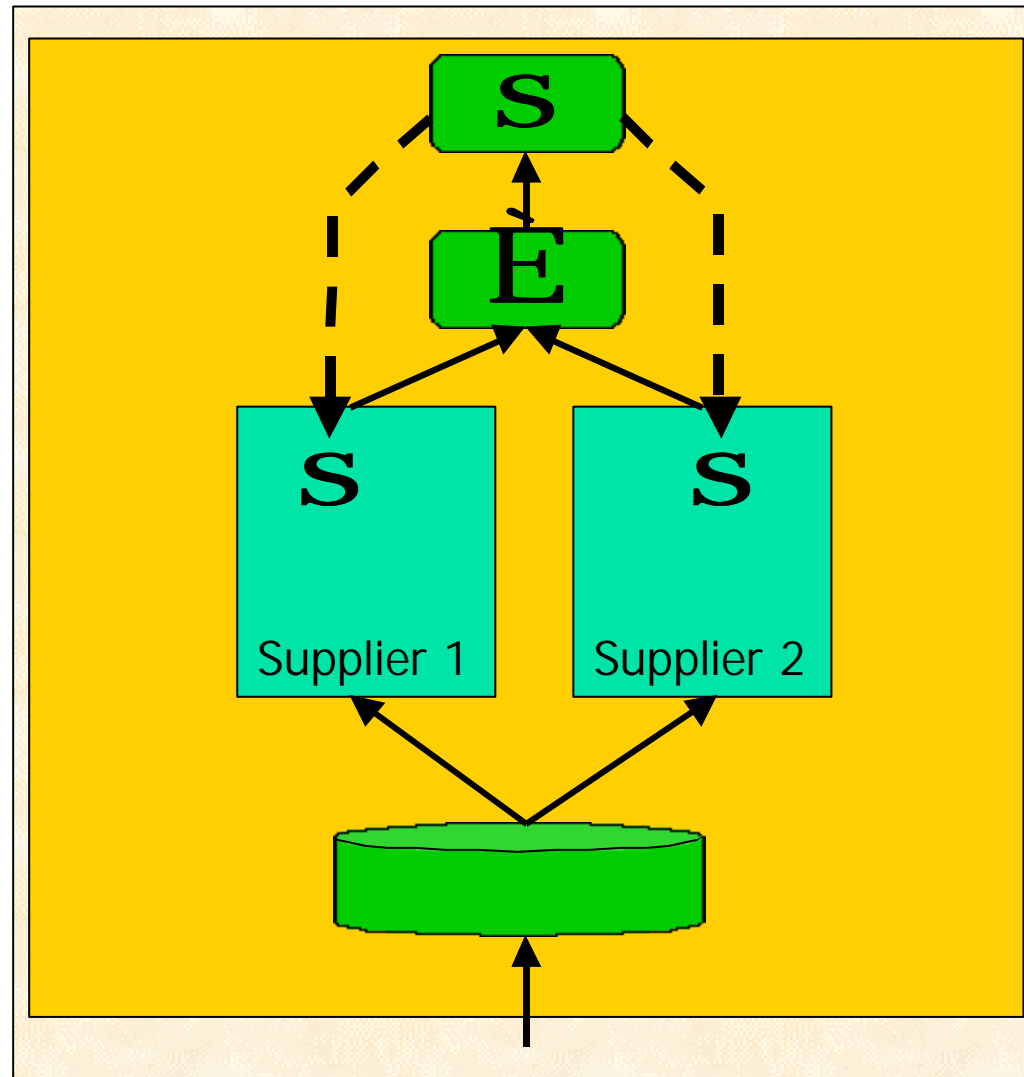
# Optimization: Caching



# Optimization: Caching



# Optimization: Predicate Migration





## Related Work

---

- Stonebraker et.al.: Quel as a Datatype, SIGMOD 1984
- Braumandl et.al.: ObjectGlobe: Ubiquitous query processing on the Internet, VLDBJ 2001
- Levy, Srivastava, Kirk: Data Model and Query Evaluation in Global Information Systems, JIIS 5(2), 1995
- Roth and Schwarz: Don't Scrap It, Wrap It! A Wrapper Architecture for Legacy Data Sources, VLDB 1997
- Stonebraker, Hellerstein: Content Integration for E-Business, SIGMOD 2001 (Cohera)
- Lucie Xyleme: Xyleme: A dynamic warehouse for XML Data of the Web, IEEE Data Engineering Bulletin 24(2), 2001
- Draper, HaLevy, Weld: The Nimble XML Data Integration System, ICDE 2001



# Conclusions

---

- Sub-plans "sitting behind" hyperlinks
- Market place is intermediary
- Embed hyperlinks at the market place
- Provide HyperQueries at remote sites
- Data integration at remote sites
- Optimization techniques