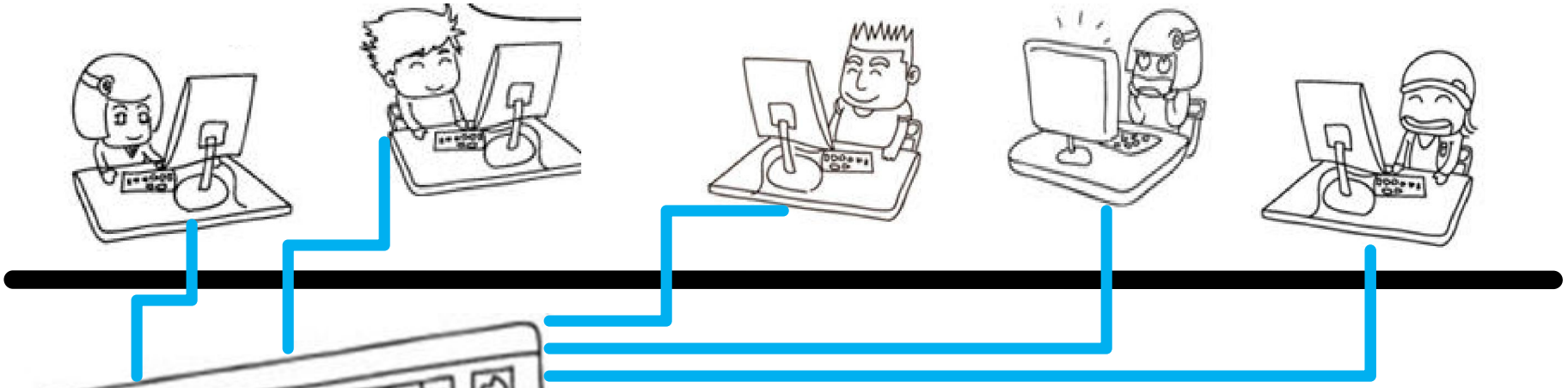
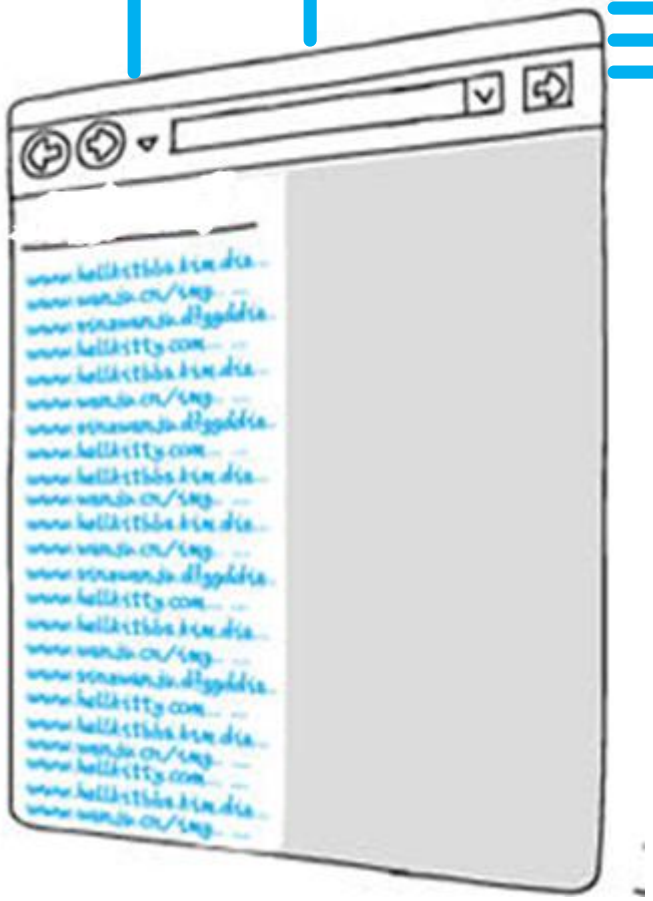
OS Hot Topics: Streaming Processing

Instructor: Dr. Liting Hu

Walmart
1 million transaction per hour
You Tube
100 hours of video per min
400 million tweets per day
The fire hose of social and business data

… data is an opportunity !



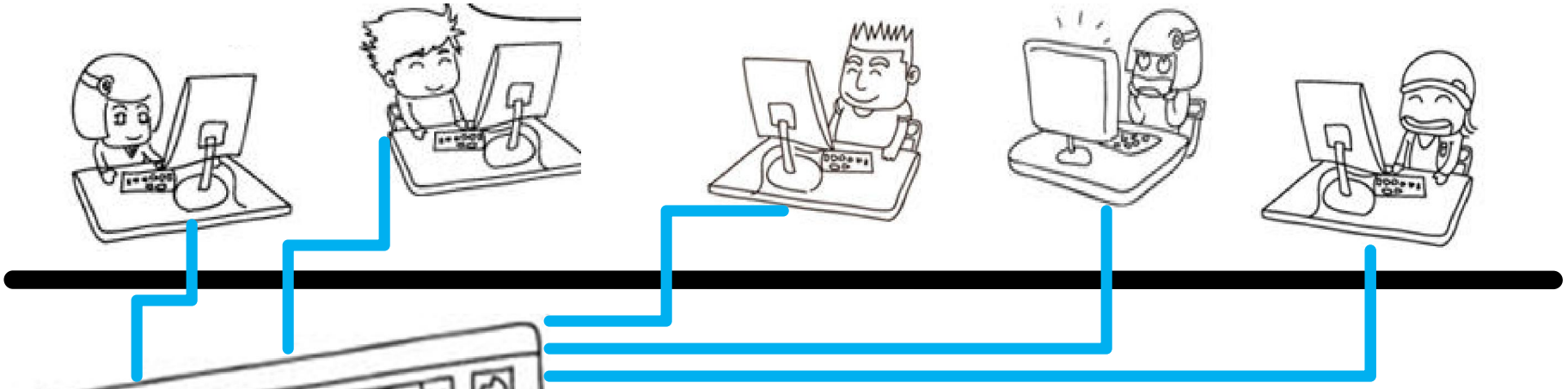


#TopK

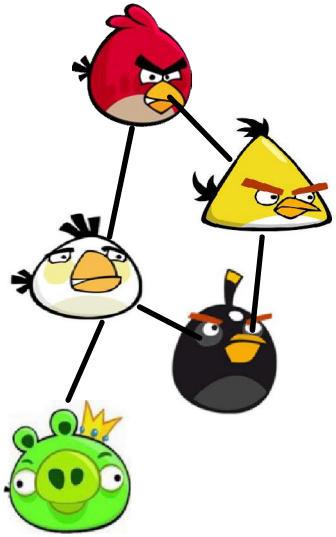
Jobs

Micro-promotion

Coupons
→Boost Sales



Jobs



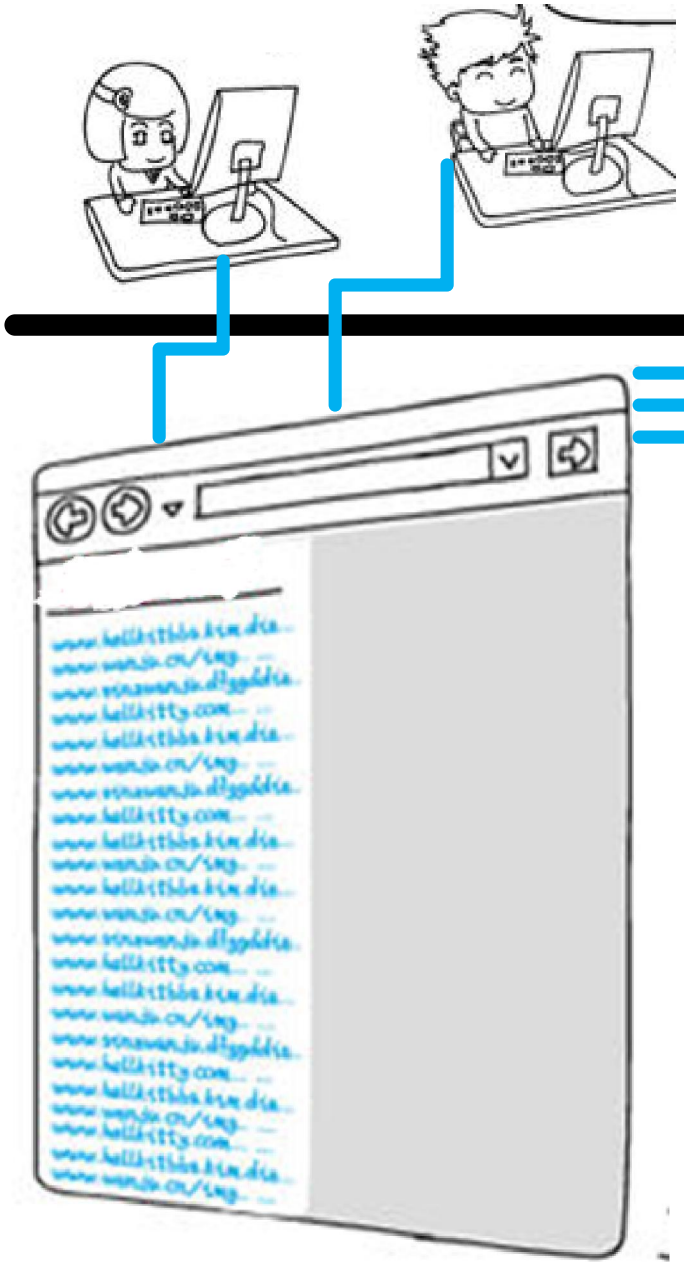
Micro-promotion

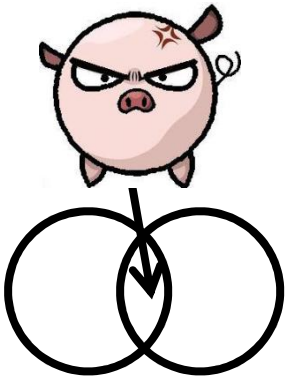
Product-bundling

Coupons

→Boost Sales

Recommendations





Jobs

Micro-promotion

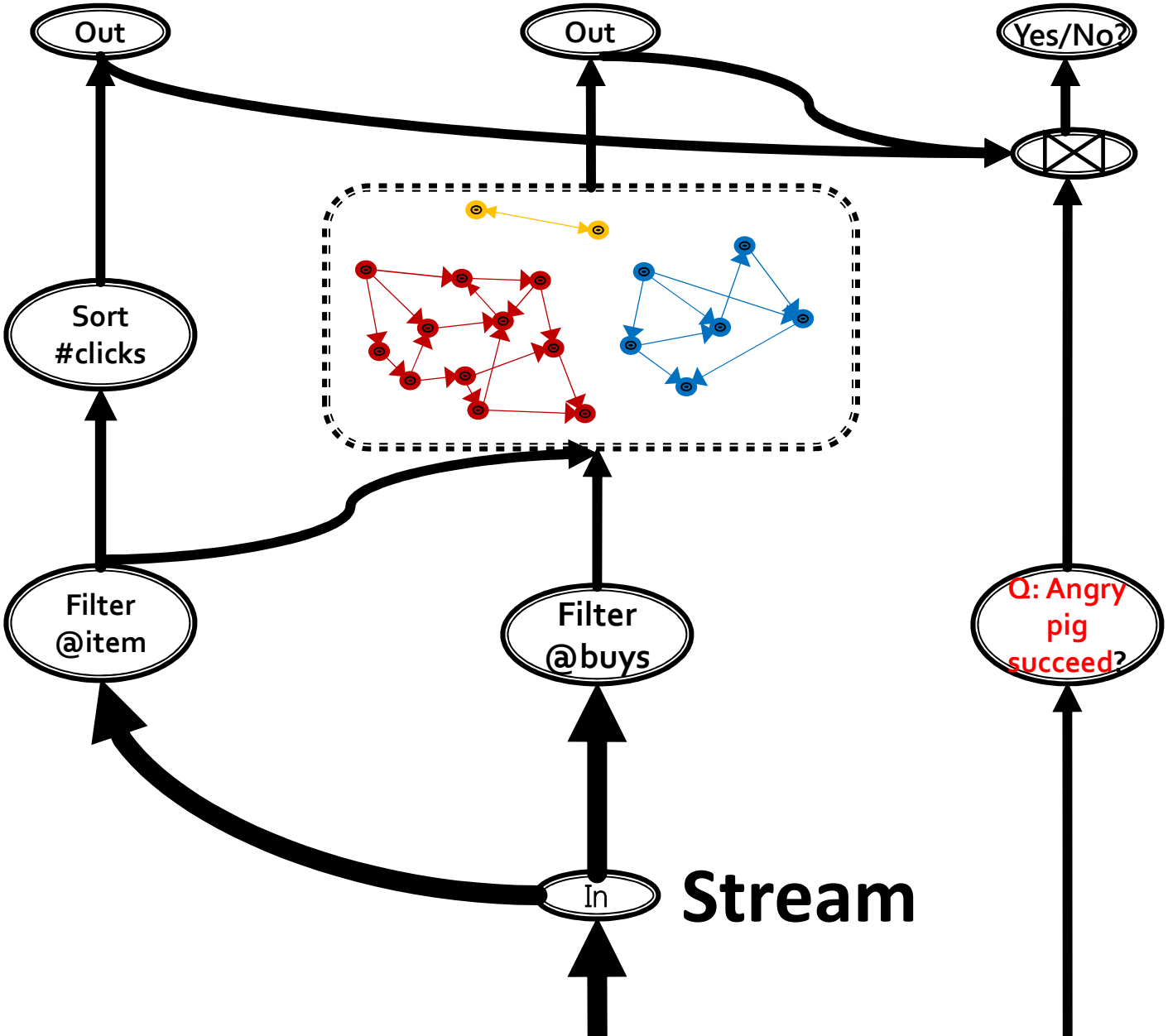
Product-bundling

Sales-prediction

Coupons
→Boost Sales

Recommendations

Predictions


Out
Out
Yes/No?
Sort
#clicks
Filter
@item
Filter
@buys
Q: Angry
pig
succeed?
In
Stream

Different System Goals

All systems

- Fairness - giving each process a fair share of the CPU
- Policy enforcement - seeing that stated policy is carried out
- Balance - keeping all parts of the system busy

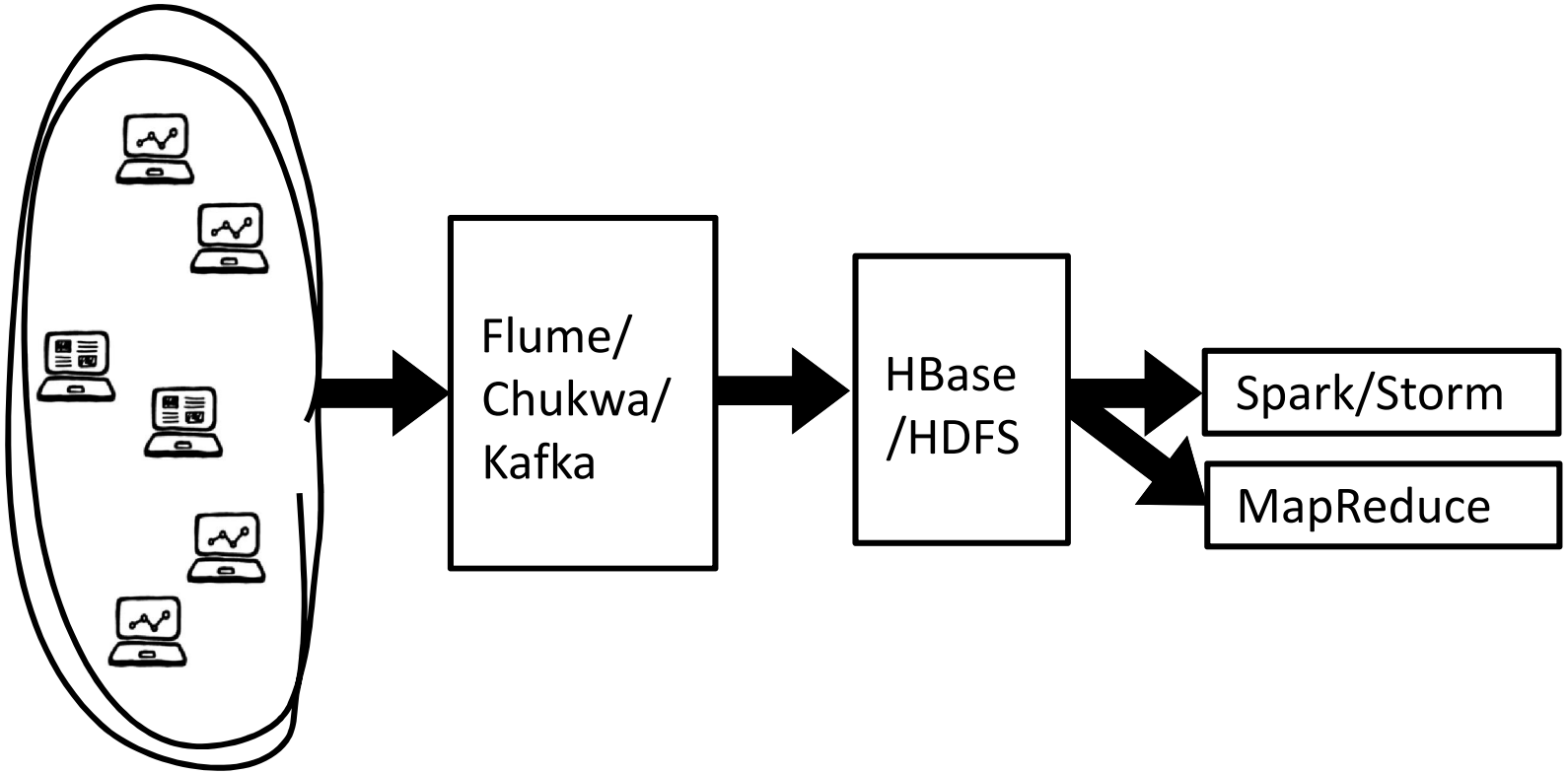
Batch systems

- Throughput - maximize jobs per hour
- Turnaround time - minimize time between submission and termination
- CPU utilization - keep the CPU busy all the time

Interactive systems

- Response time - respond to requests quickly
- Proportionality - meet users' expectations

Real-time systems

- Meeting deadlines - avoid losing data
- Predictability - avoid quality degradation in multimedia systems

Data Analytics Workflow