Operator Placement with QoS Constraints for Distributed Stream Processing

Yuanqiang Huang, Zhongzhi Luan, Rong He, Depei Qian, Sino-German Joint Software Institute,
Beijing Key Laboratory of Computer Network Beihang University Beijing, China

In Network and Service Management (CNSM), 2011 7th International Conference on (pp. 1-7). IEEE.

Motivation

- Challenge:
  - Operator placement (in-network) :
    - To achieve an optimal resource allocation.
    - An optimization problem with QoS (Quality of Service) constraints: throughput and end-to-end delay.
    - Getting a global optimization is a NP-hard problem.

Motivation

- Solution:
  1. Formalize the operator placement problem
     - with network usage as the optimization objective and constraints.
  2. Propose a concept of Optimization Power
     - describe the host's capacity to reach a global optimal solution as soon as possible.
     - Consider QoS metrics : throughput and end-to-end delay
  3. Propose a corresponding Optimization Power-based heuristic algorithm for operator placement.

Application Model

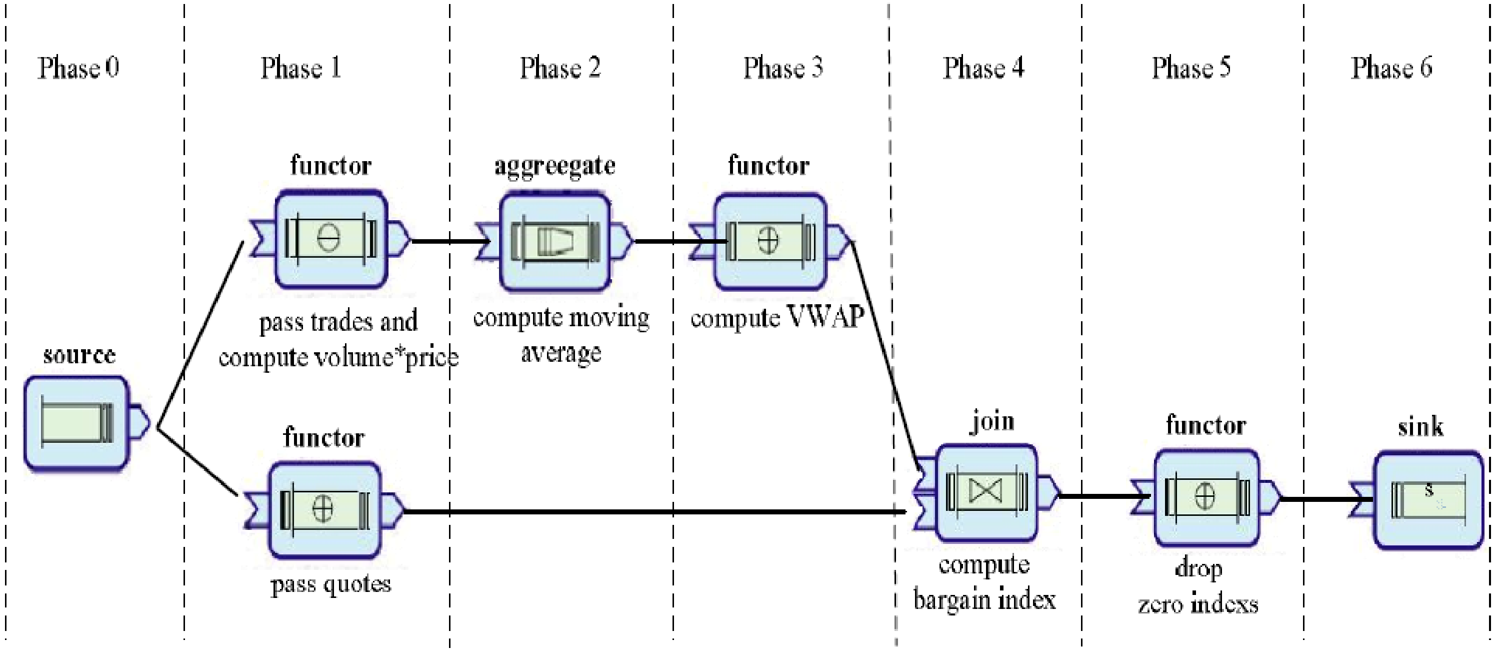


Figure 1. Application of finical analysis using distributed stream processing

Heuristic Operator Placement - ***Optimization Power***

- ***Optimization Power*** need to consider:
  1. network delay between upstream-downstream hosts.

  ➔ In general, **smaller network delay** ➔ smaller network usage.

  2. host resource capacity ➔ processing delays of operators

  ➔ application's **end-to-end delay**

  3. expected time needed by an operator *Oi* to process a tuple on node n_i can be estimated:

$$\forall o_i, n_j \quad d_p(o_i, n_j) = \frac{er_{cpu}^{o_i} / rr_{cpu}^{n_j}}{1 - Rate_{in}^{o_i} \cdot er_{cpu}^{o_i} / rr_{cpu}^{n_j}} = \frac{er_{cpu}^{o_i}}{rr_{cpu}^{n_j} - Rate_{in}^{o_i} \cdot er_{cpu}^{o_i}}$$

Heuristic Operator Placement - ***Optimization Power***

- *Optimization Power* (*OP*) :
  - measure the appropriateness of node n_k for hosting operator o
  - calculated by :

residual CPU capacity on node n_k

maximal delay allowed

maximal network delay when choosing n_k to host o

$$OP^{o}_{n_k} = \left(\frac{rr^{n_k}_{cpu}}{MAX_{nd}(o,n_k)}\right)^{(1/SUM_{nu}(o,n_k))} \cdot (q^{max}_{d} - d_p(o,n_k) - MAX_{nd}(o,n_k))$$

maximal network delay when choosing n_k to host o

Increased network usage when choosing n_k to host o

processing delay when choosing n_k to host o

Heuristic Operator Placement - *Algorithm*

- relies on:
  - Resource Discovery Service (RDS) to discover potential hosts that can satisfy resource requirements for processing operators.
  - Network Coordinate Service (NCS) to estimate network delay between any pair of nodes using Euclidean Distance between their given network coordinates.

Heuristic Operator Placement - *Algorithm*

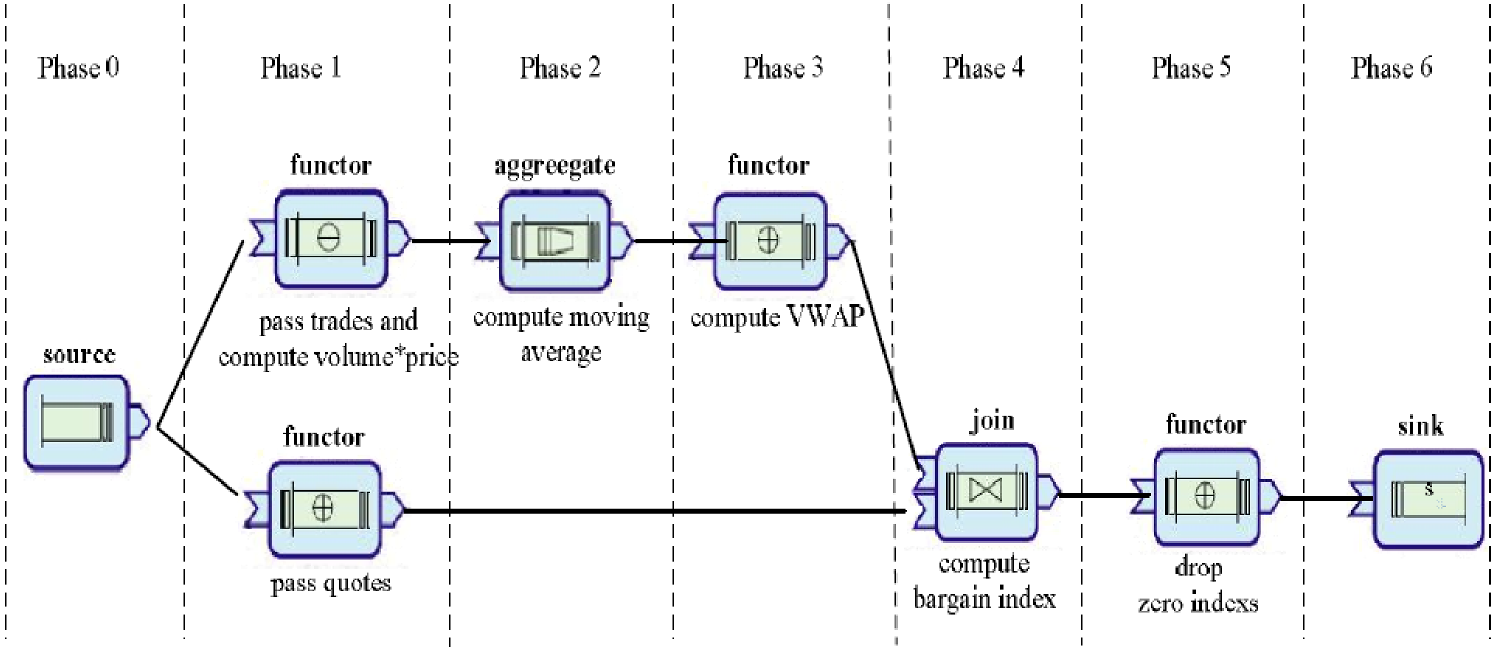


Figure 1. Application of finical analysis using distributed stream processing

Evaluation - *Experimental Settings*

- Use a trace data from PlanetLab network platform, which includes:
  - a span of 10 months (July 2007--April 2008) collecting for network delay of every PlanetLab node pair.
  - the total number of nodes is more than 240 and the total number of records is over 110,000.
- Generate network coordinate for every PlanetLab node by using Vivaldi algorithm.
- Since the data of bandwidth between node pair is not provided in the trace file, we used the BRITE [4] to simulate the bandwidths.
- Bandwidth distribution is based on exponential model with the value range of [10KBps, 10MBps].
- Adopt Zipf distribution model for resource distribution of nodes.
- In our experiments, we considered 3 types of important node resource: CPU speed, memory size and disk size. Each resource is assigned a value in the range of [2000, 20000].

Evaluation - *Experimental Settings*

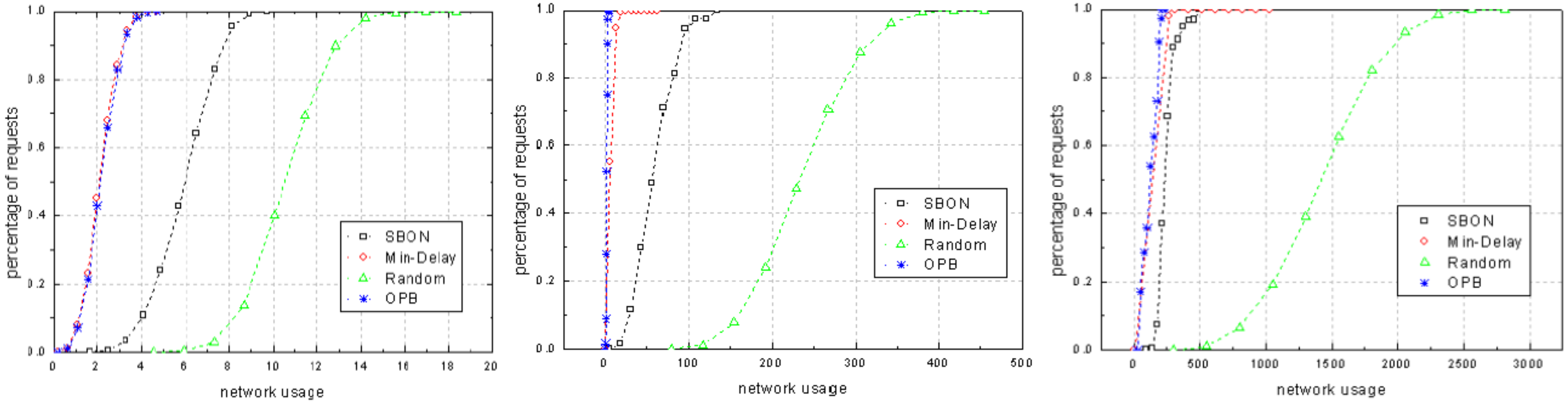
- Application consists of 10 operators including:
  - 2 sources and1 sink (fixed hosts)
  - 7 intermediate operators
  - every non-sink operator can have 1 to 3 downstream operators.
- By default, source's stream output rate is 5 tuples per second. Selectivity of all intermediate operators is set to 1.0, and the average size of tuple is 10 bytes.
- Define two adjustable factors ftp and fd to control application's throughput and end-to-end delay objectives respectively.
- ftp is for controlling throughput objective. The stream output rate of the sources is $5 \cdot ftp$ tuples per second. In same phase, half of intermediate operators set their selectivity to ftp , and the other half set to $1/ ftp$.

Evaluation - *Experimental Settings*

- fd is the other factor for end-to-end delay objective. Let l denotes the maximal delay between the source hosts and the sink host.
- So we set the application's end-to-end delay threshold to $fd \cdot l$, l is unchanged during simulation since the positions of sources and sink are fixed beforehand.
- Also implemented three alternative operator placement algorithms for comparison:
  - i) SBON algorithm proposed assigns optimal virtual network coordinate for every operator based on Force-Energy theory, and then perform the k-nearest neighbor search (we set k=10) for each operator in the node space to find a host which has enough resource among these k neighbors.
  - ii) MIN-DELAY algorithm does a global search in node space for every operator to find a host which can introduce the minimal delay which is the sum of total processing delay on hosts and network delays from the current operator to the source and sink.
  - iii) RANDOM algorithm assigns a random host for every operator. For all the algorithms, when no eligible node which can meet application's SLOs is found, placement fails.

Evaluation - *Results and Analysis -*

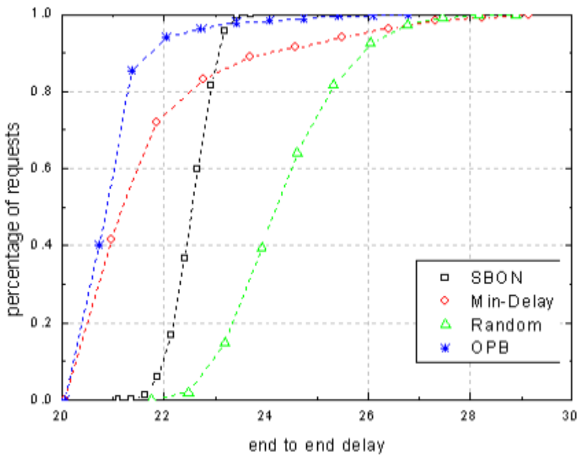
Comparision of Cumulative Percentage Distribution of 5000 placements for network usage and end-to-end delay with different value of f_{tp}



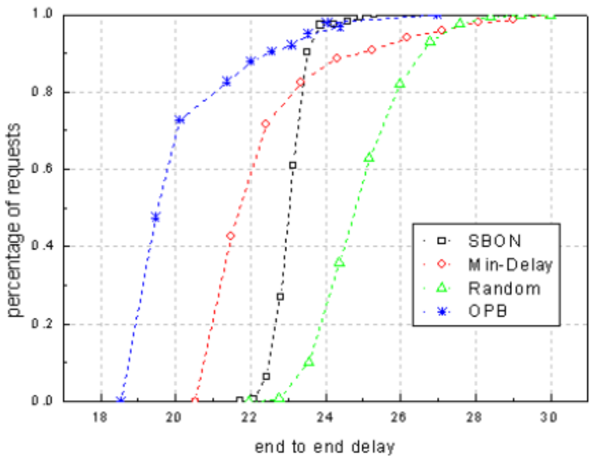
(a) $\mathbf{f_{tp}} = 1$ (b) $\mathbf{f_{tp}} = 5$ (c) $\mathbf{f_{tp}} = 10$

Evaluation - *Results and Analysis* - *end-to-end delay*

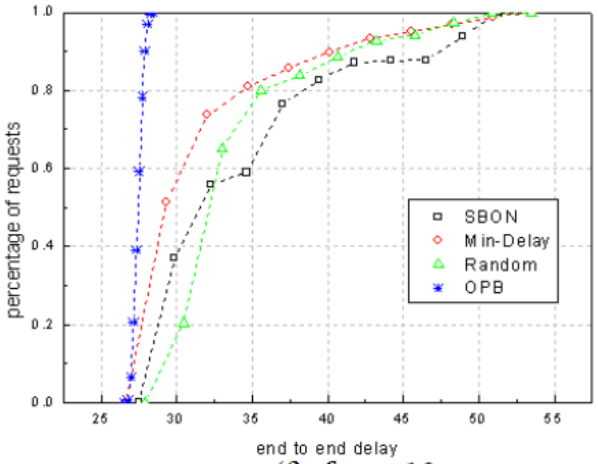
Comparision of Cumulative Percentage Distribution of 5000 placements for network usage and end-to-end delay with different value of f_{tp}



(d) $\mathbf{f_{tp}} = 1$



(e) $\mathbf{f_{tp}} = 5$



(f) $\mathbf{f_{tp}} = 10$

Conclusions

1. Formalize the operator placement problem
   - with optimizing network usage and meeting constraints.
2. Propose a concept of Optimization Power:
   - make the local optimal solution closer to the global one.
   - Consider QoS metrics : throughput and end-to-end delay
3. Propose a corresponding Optimization Power-based heuristic algorithm for operator placement.
4. Experimental results show that OPB has performance advantage compared to some other operator placement algorithms.