

## **Team 13 DDS Project- Phase 3**

### **Experiments on operations developed in Phase 2 and Phase 3**

In phase 2, we implemented the following operations in Apache Spark:

- 1.) Geometric union
- 2.) Convex hull
- 3.) Farthest pair
- 4.) Closest pair
- 5.) Spatial Join query
- 6.) Spatial Range query

We ran these operations on single cluster and clusters of multiple nodes and noted down the per node CPU usage, per node Memory usage, Run time and the communication cost in the cluster. We also used different data sets.

#### **New Operation developed in Phase 3: Heat Map System Specification:**

##### **Master**

**Processor:** Intel Core i3 CPU M 370 @ 2.40Hz\*4

**Memory:** 2.9GB

**OS:** 64-bit Ubuntu

##### **Worker-1**

**Processor:** Intel Core i3 CPU @ 1.90GHz × 4

**Memory:** 5.7 GiB

**OS:** 64-bit Ubuntu

##### **Worker-2**

**Processor:** Intel Core i5 CPU @ 2.4GHz × 4

**Memory:** 5.88 GiB

**OS:** 64-bit Ubuntu

##### **Worker-3**

**Processor:** Intel Core i3 CPU M 370 @ 2.40Hz\*4

**Memory:** 2.9GB

**OS:** 64-bit Ubuntu

#### **Software Used For Measurement**

1. Glances

## 2. Dstat

### **DataSet Used**

- 1) **Big Data Set:** Arealm.csv
- 2) **Small Data Set:** Sample File of size 12kb

### **1. Geometric Union**

#### **(i.) Big Data Set**

#### **A. CPU Utilization (Percentage Increase)**

**Single Node:** Worker1- 55.27

**2 Nodes:** Worker 1 - 41.04 ;Worker 2- 49.6

**3 Nodes:** Worker 1- 38.26 ;Worker 2- 24.6 ; Worker 3- 18.2

**4 Nodes:** Worker 1- 31.45 ;Worker 2- 18.63 ;Worker 3-20.04 ; Worker 4 -15.26

#### **B. Memory Utilization (Percentage Increase)**

**Single Node:** Worker1-13.04

**2 Nodes:** Worker 1 - 6.23 Worker 2-9.8

**3 Nodes:** Worker 1- 5.1 Worker 2- 8.9 Worker 3-8.1

**4 Nodes:** Worker 1- 4.72 Worker 2- 6.8 Worker 3-6.2 ;Worker 4 -6.01

#### **C. Run Time (Seconds)**

**Single Node:** 403

**2 Nodes:** Job 0- 175 Job 1-187

**3 Nodes:** Job 0- 156 Job 1- 182

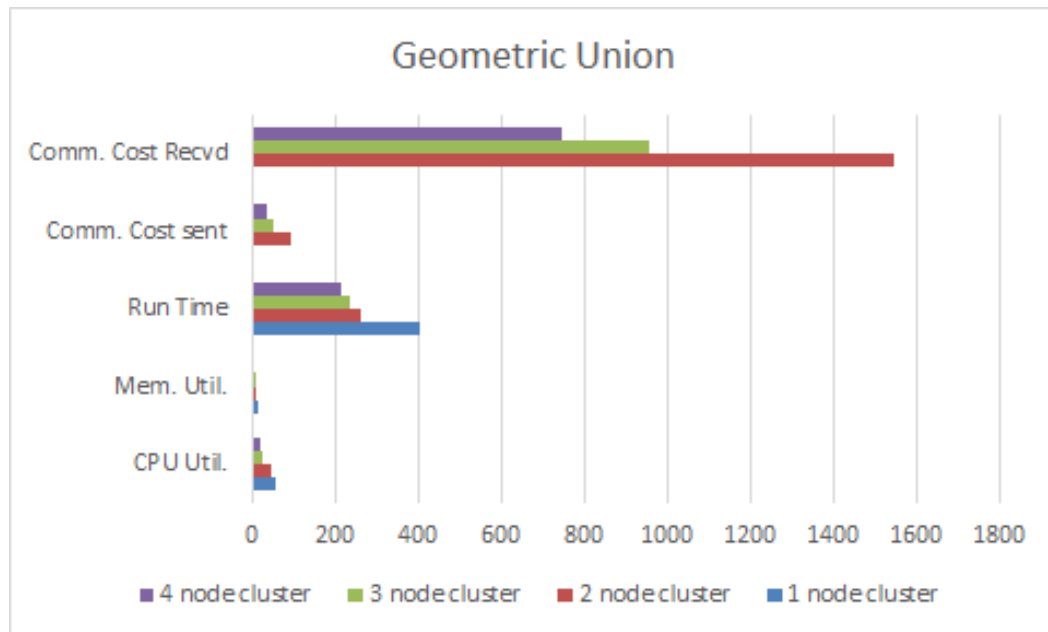
**4 Nodes:** Job 0- 148 Job 1-176

#### **D. Communication Cost**

**2 Nodes:** Worker1 Sent- 88.8 K Worker1 Received- 1573.4K  
Worker2 Sent- 97.3 K Worker2 Received- 1512 K

**3 Nodes:** Worker1 Sent- 53.5 K Worker1 Received- 971 K  
Worker2 Sent- 44.3 K Worker2 Received- 945.5 K  
Worker3 Sent- 59.5 K Worker3 Received- 958.06 K

**4 Nodes:** Worker1 Sent- 32.1 K Worker1 Received- 736 K  
Worker2 Sent- 38.4 K Worker2 Received- 778.8 K  
Worker3 Sent- 31.2 K Worker3 Received- 728.2 K  
Worker4 Sent- 43.1 K Worker4 Received- 745.7 K



## (ii.) Small Data Set

### A. CPU Utilization (Percentage Increase)

**Single Node:** Worker1- 44.5

**2 Nodes:** Worker 1 - 38.76 ;Worker 2-34.23

**3 Nodes:** Worker 1- 35.45 ; Worker 2- 24.5 ;Worker 3- 24.4

**4 Nodes:** Worker 1- 32.26 ; Worker 2- 20.6 ;Worker 3-18.36 ;Worker 4 - 15.63

### B. Memory Utilization (Percentage Increase)

**Single Node:** Worker1- 7.36

**2 Nodes:** Worker 1 -5.63 ;Worker 2-4.23

**3 Nodes:** Worker 1- 2.89 ;Worker 2- 3.2 ;Worker 3- 3.1

**4 Nodes:** Worker 1-1.69 ;Worker 2- 2.36 ;Worker 3- 2.3 ;Worker 4 -1.21

### C. Run Time

**Single Node:** 11.23 seconds

**2 Nodes:** Job0- 5.6 seconds Job1- 4.45 seconds

**3 Nodes:** Job0- 4.92 seconds Job1- 3.76 seconds

**4 Nodes:** Job0- 3.26 seconds Job1- 3.43 seconds

### D. Communication Cost

**2 Nodes:** Worker1 Sent- 23.6K Worker1 Received- 526.4 K

Worker2 Sent- 27.8K Worker2 Received- 514.6 K

**3 Nodes:** Worker1 Sent- 16.8K Worker1 Received- 387.3K

Worker2 Sent- 12.3K	Worker2 Received- 344.1K
Worker3 Sent- 15.7K	Worker3 Received- 362K

4 Nodes: Worker1 Sent- 12.9K	Worker1 Received- 236.36K
Worker2 Sent- 11.7K	Worker2 Received- 242.23K
Worker3 Sent- 15.1K	Worker3 Received- 256.23K
Worker4 Sent- 12.9K	Worker4 Received- 229.5K

## **2. CONVEX HULL**

### **(i.) Big Data Set**

#### **A. CPU Utilization (Percentage Increase)**

**Single Node:** Worker1- 42.2%

**2 Nodes:** Worker 1 - 37% ; Worker 2- 40%

**3 Nodes:** Worker 1- 32% ; Worker 2- 31% ; Worker 3- 21%

**4 Nodes:** Worker 1- 21% ; Worker 2- 19% ; Worker 3- 15% ; Worker 4 - 19%

#### **B. Memory Utilization (Percentage Increase)**

**Single Node:** Worker1- 14.1%

**2 Nodes:** Worker 1 - 7% ; Worker 2- 6%

**3 Nodes:** Worker 1- 6% ; Worker 2- 4.5% ; Worker 3- 3.2%

**4 Nodes:** Worker 1- 4.2% ; Worker 2- 4% ; Worker 3- 3.3% ; Worker 4 - 3.7%

#### **C. Run Time**

**Single Node:** 17.24s

**2 Nodes:** Job0- 9.8s Job 1- 6.9s

**3 Nodes:** Job0- 7.9s Job1- 6.3s

**4 Nodes:** Job0- 7.2s Job 1- 6s

#### **D. Communication Cost**

**2 Nodes:** Worker1Sent- 42k Worker1Received- 621k

Worker2Sent- 58k Worker2Received- 511k

**3 Nodes:** Worker1Sent- 31k Worker1Received- 318k

Worker2Sent- 38k Worker2Received- 480k

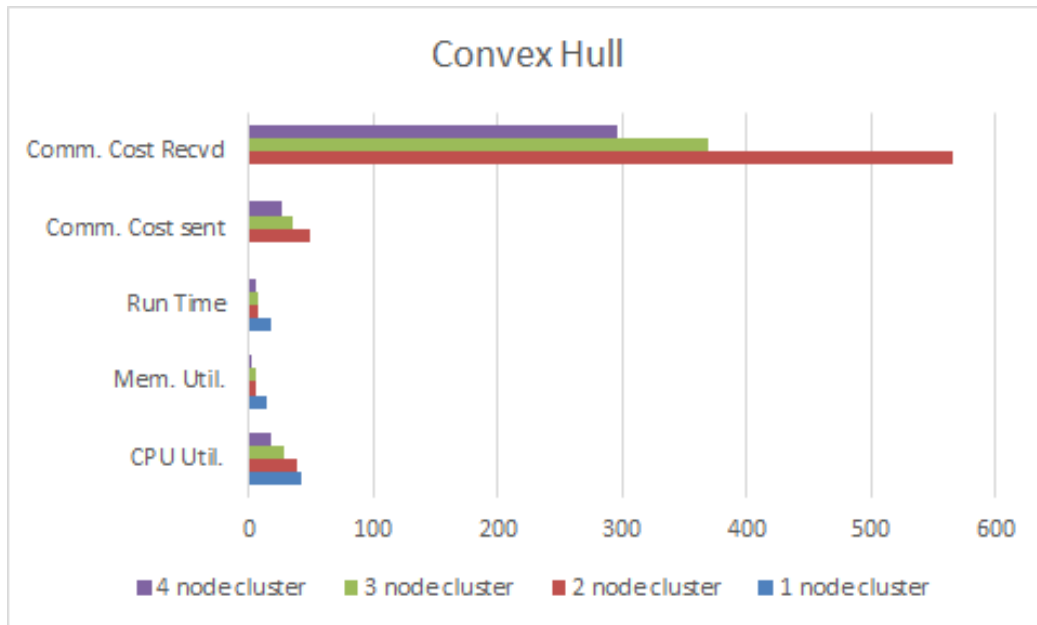
Worker3Sent- 29k Worker3Received- 311k

**4 Nodes:** Worker1 Sent- 31k Worker1 Received- 319k

Worker2 Sent- 27k Worker2 Received- 250k

Worker3 Sent- 18k  
Worker4 Sent- 30k

Worker3 Received- 218k  
Worker4 Received- 400k



## (ii.) Small Data Set

### A. CPU Utilization (Percentage Increase)

**Single Node:** Worker1- 38%

**2 Nodes:** Worker 1 - 31% ;Worker 2- 36%

**3 Nodes:** Worker 1- 24% ;Worker 2- 24.5% ;Worker 3- 34.1%

**4 Nodes:** Worker 1- 21% ;Worker 2- 23% ;Worker 3- 18% ;Worker 4 - 27%

### B. Memory Utilization (Percentage Increase)

**Single Node:** Worker1- 11.4%

**2 Nodes:** Worker 1 - 9.7% ;Worker 2- 8%

**3 Nodes:** Worker 1- 3.4% ;Worker 2- 8.1% ;Worker 3- 7.1%

**4 Nodes:** Worker 1- 6.2% ;Worker 2- 3.4% ;Worker 3- 5% ;Worker 4 - 5%

### C. Run Time

**Single Node:** 5.3s

**2 Nodes:** Job0- 4.3s Job1- 3.7s

**3 Nodes:** Job0- 3.54s Job1- 3.2s

**4 Nodes:** Job0- 3.1s Job1- 2.8s

### D. Communication Cost

**2 Nodes:** Worker1Sent- 26k

Worker1Received- 211k

Worker2Sent- 15k

Worker2Received- 180k

**3 Nodes:** Worker1Sent- 17k

Worker1Received- 126k

	Worker2Sent- 15k	Worker2Received- 149k
	Worker3Sent- 9k	Worker3Received- 170k
<b>4 Nodes:</b>	Worker1Sent- 12k	Worker1Received- 86k
	Worker2Sent- 8k	Worker2Received- 73k
	Worker3Sent- 18k	Worker3Received- 72k
	Worker4Sent- 6k	Worker4Received- 117k

### **3. FARTHEST PAIR**

#### **(i.) Big Data Set**

##### **A. CPU Utilization (Percentage Increase)**

**Single Node:** Worker1- 48.7

**2 Nodes:** Worker 1 - 32.5; Worker 2- 31.2

**3 Nodes:** Worker 1- 27.2; Worker 2- 27.8; Worker 3- 17.6

**4 Nodes:** Worker 1- 22.7; Worker 2- 24.4; Worker 3- 13.5 Worker 4 - 14.6

##### **B. Memory Utilization (Percentage Increase)**

**Single Node:** Worker1- 13.3

**2 Nodes:** Worker 1 - 9.8; Worker 2- 6.5

**3 Nodes:** Worker 1- 7.3; Worker 2- 4.5; Worker 3- 5.6

**4 Nodes:** Worker 1- 6.7; Worker 2- 4.1; Worker 3- 4.2;  
Worker 4 - 3.9

##### **C. Run Time (seconds)**

**Single Node:** 21.76

**2 Nodes:** Job0- 15.57 Job1- 8.89

**3 Nodes:** Job0- 6.3 Job1- 6.8

**4 Nodes:** Job0- 5.57 Job1- 6.2

##### **D. Communication Cost**

**2 Nodes:** Worker1 Sent- 40k Worker1 Received- 730k

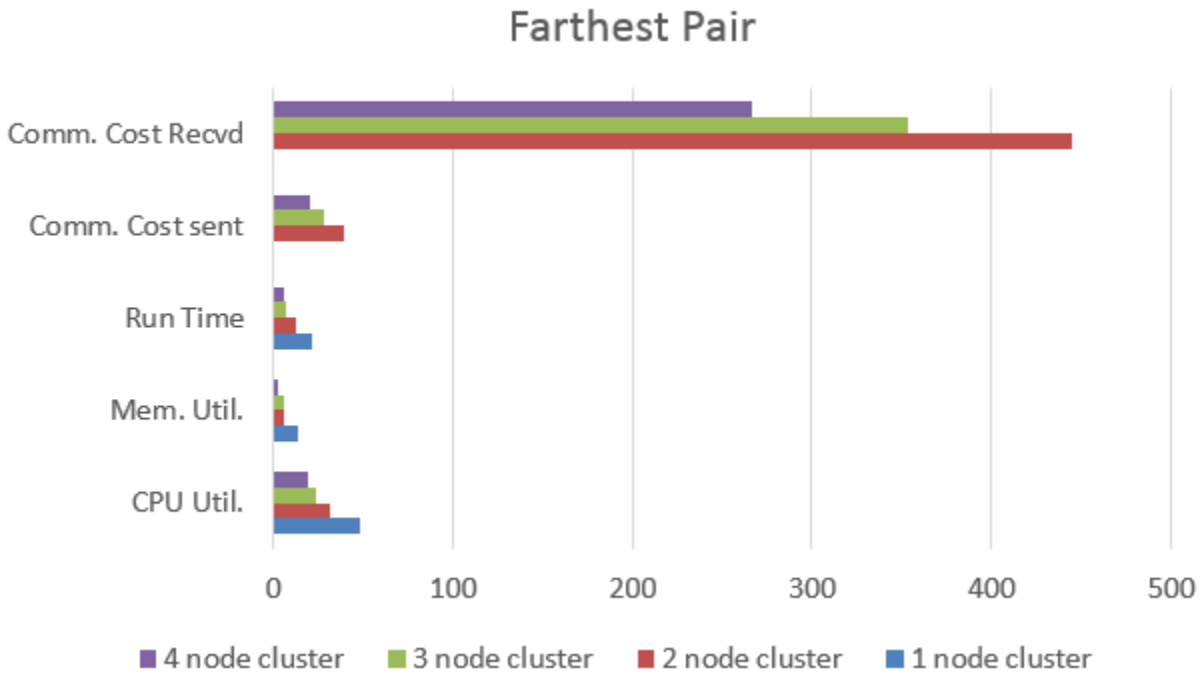
Worker2 Sent- 39k Worker2 Received- 260k

**3 Nodes:** Worker1 Sent- 25k Worker1 Received- 340k

Worker2 Sent- 26k Worker2 Received- 401k

Worker3 Sent- 35k Worker3 Received- 320k

**4 Nodes:** Worker1 Sent- 15k      Worker1 Received- 310k  
                  Worker2 Sent- 16k      Worker2 Received- 340k  
                  Worker3 Sent- 32k      Worker3 Received- 198k  
                  Worker4 Sent- 19k      Worker4 Received- 219k



## (ii.) Small Data Set

### A. CPU Utilization (Percentage Increase)

**Single Node:** Worker1- 39

**2 Nodes:** Worker 1 - 36.3; Worker 2- 34.5

**3 Nodes:** Worker 1- 23.4; Worker 2- 23.45; Worker 3- 24.17

**4 Nodes:** Worker 1- 18.7; Worker 2- 12.8%; Worker 3- 20.3% ; Worker 4 - 15.6%

### B. Memory Utilization (Percentage Increase)

**Single Node:** Worker1- 12

**2 Nodes:** Worker 1 - 9; Worker 2- 8

**3 Nodes:** Worker 1- 2.78; Worker 2- 6.43; Worker3- 6.4

**4 Nodes:** Worker 1- 3.45; Worker 2- 3.56; Worker 3- 5.7 ;  
                  Worker 4 - 2.4

### C. Run Time

**Single Node:** 2.53s

**2 Nodes:** Job0- 1.93s                      Job1- 3.13s

**3 Nodes:** Job0- 1.13s                      Job1- 1.19s

**4 Nodes:** Job0- 0.92s                      Job1- 0.89s

### D. Communication Cost

**2 Nodes:**     Worker1Sent- 14k                      Worker1Received- 86k  
                 Worker2Sent- 29k                      Worker2Received- 136k

**3 Nodes:**     Worker1Sent- 15k                      Worker1Received- 90k  
                 Worker2Sent- 13k                      Worker2Received- 100k  
                 Worker3Sent- 15k                      Worker3Received- 89k

**4 Nodes:**     Worker1Sent- 6.3k                      Worker1Received- 89k  
                 Worker2Sent- 4.5k                      Worker2Received- 74k  
                 Worker3Sent- 11k                      Worker3Received- 42k  
                 Worker4Sent- 9k                      Worker4Received- 36k

## 4. CLOSEST PAIR

### (i.) Big Data Set

#### A. CPU Utilization (Percentage Increase)

**Single Node:** Worker1- 43.7%

**2 Nodes:** Worker 1 - 41%     ;Worker 2- 42.6%

**3 Nodes:** Worker 1- 41.7%   ;Worker 2- 32% ;Worker 3- 40%

**4 Nodes:** Worker 1- 34%     ;Worker 2- 28% ;Worker 3- 27%   ;Worker 4 - 29%

#### B. Memory Utilization (Percentage Increase)

**Single Node:** Worker1- 11%

**2 Nodes:** Worker 1 - 7.2%                      ;Worker 2- 8.7%

**3 Nodes:** Worker 1- 8%   ;Worker 2- 3.3%   ;Worker 3- 6.4%

**4 Nodes:** Worker 1- 4%   ;Worker 2- 3%   ;Worker 3- 4%   ;Worker 4 - 5%

### C. Run Time

**Single Node:** 16.47s

**2 Nodes:** Job0- 10.2s                      Job1- 11.9s

**3 Nodes:** Job0- 7.64s                      Job1- 10.9s

**4 Nodes:** Job0- 6.4s                      Job1- 7.3s

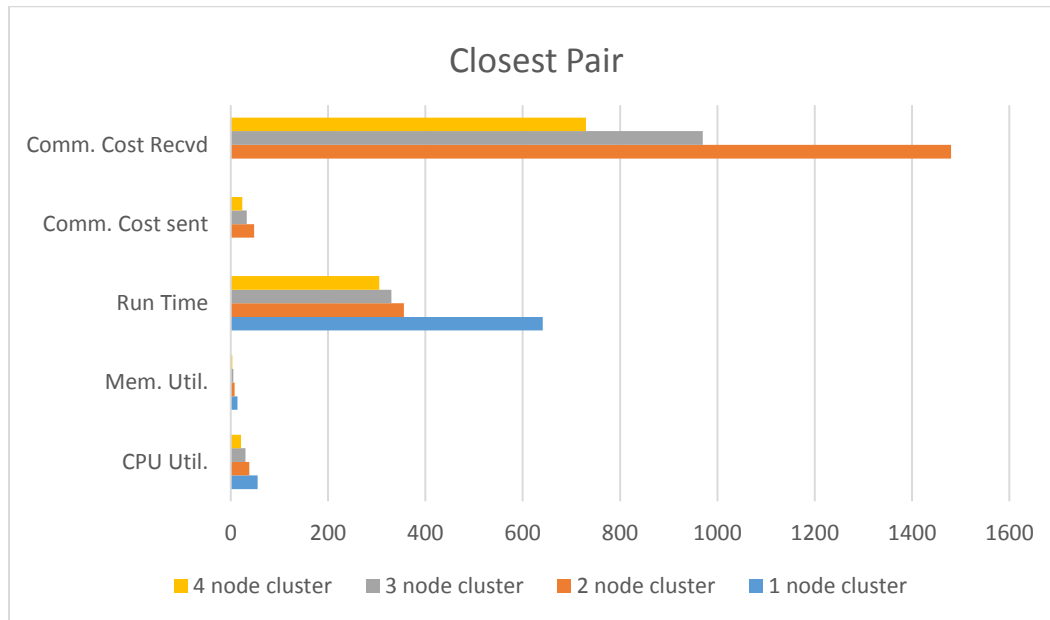
### D. Communication Cost

**2 Nodes:**     Worker1Sent- 35k                      Worker1Received- 698k  
                 Worker2Sent- 37k                      Worker2Received- 239k



**3 Nodes:**      Worker1Sent- 19k                      Worker1Received- 310k  
                      Worker2Sent- 23k                      Worker2Received- 381k  
                      Worker3Sent- 29k                      Worker3Received- 280k

**4 Nodes:**      Worker1Sent- 12k                      Worker1Received- 298k  
                      Worker2Sent- 13k                      Worker2Received- 325k  
                      Worker3Sent- 31k                      Worker3Received- 218k  
                      Worker4Sent- 16k                      Worker4Received- 239k



## (ii.) Small Data Set

### A. CPU Utilization (Percentage Increase)

**Single Node:** Worker1- 36%

**2 Nodes:** Worker 1 - 35.3% ;Worker 2- 32%

**3 Nodes:** Worker 1- 23.3% ;Worker 2- 27.1% ;Worker 3- 32.1%

**4 Nodes:** Worker 1- 18.2% ;Worker 2- 13.8% ; Worker 3- 21%  
 ;Worker 4 - 14.6%

### B. Memory Utilization (Percentage Increase)

**Single Node:** Worker1- 9%

**2 Nodes:** Worker 1 - 9% ;Worker 2- 8%

**3 Nodes:** Worker 1- 3.1% ;Worker 2- 8% ;Worker 3- 5.4%

**4 Nodes:** Worker 1- 6% ;Worker 2- 3.4% ;Worker 3- 5.7% ;Worker 4 - 2.4%

### C. Run Time

**Single Node:** 1.67s

<b>2 Nodes:</b>	Job0- 1.48s	Job1- 1.39s
<b>3 Nodes:</b>	Job0- 1.1s	Job1- 1.09s
<b>4 Nodes:</b>	Job0- 0.89s	Job1- 0.92s

#### **D. Communication Cost**

<b>2 Nodes:</b>	Worker1Sent- 12k	Worker1Received- 80k
	Worker2Sent- 27k	Worker2Received- 141k
<b>3 Nodes:</b>	Worker1Sent- 10k	Worker1Received- 88k
	Worker2Sent- 14k	Worker2Received- 110k
	Worker3Sent- 18k	Worker3Received- 100k
<b>4 Nodes:</b>	Worker1Sent- 7k	Worker1Received- 97k
	Worker2Sent- 4k	Worker2Received- 68k
	Worker3Sent- 13k	Worker3Received- 32k
	Worker4Sent- 8k	Worker4Received- 39k

### **5. SPATIAL RANGE QUERY**

#### **(i.) Big Data Set**

##### **A. CPU Utilization (Percentage Increase)**

Single Node: Worker1- 43

2 Nodes: Worker 1 - 39 ; Worker 2- 28

3 Nodes: Worker 1- 26 ;Worker 2- 25 ;Worker 3- 33

4 Nodes: Worker 1- 24 ;Worker 2- 22 ;Worker 3- 26 ;Worker 4 - 24

##### **B. Memory Utilization (Percentage Increase)**

Single Node: Worker1- 9.4

2 Nodes: Worker 1 - 7.1 ;Worker 2- 8

3 Nodes: Worker 1- 6 ;Worker 2- 3.3 ;Worker 3- 3.2

4 Nodes: Worker 1- 5.5 ;Worker 2- 2.7 ;Worker 3- 3.0 ;Worker 4 - 2.9

##### **C. Run Time (Seconds)**

Single Node: 20.96

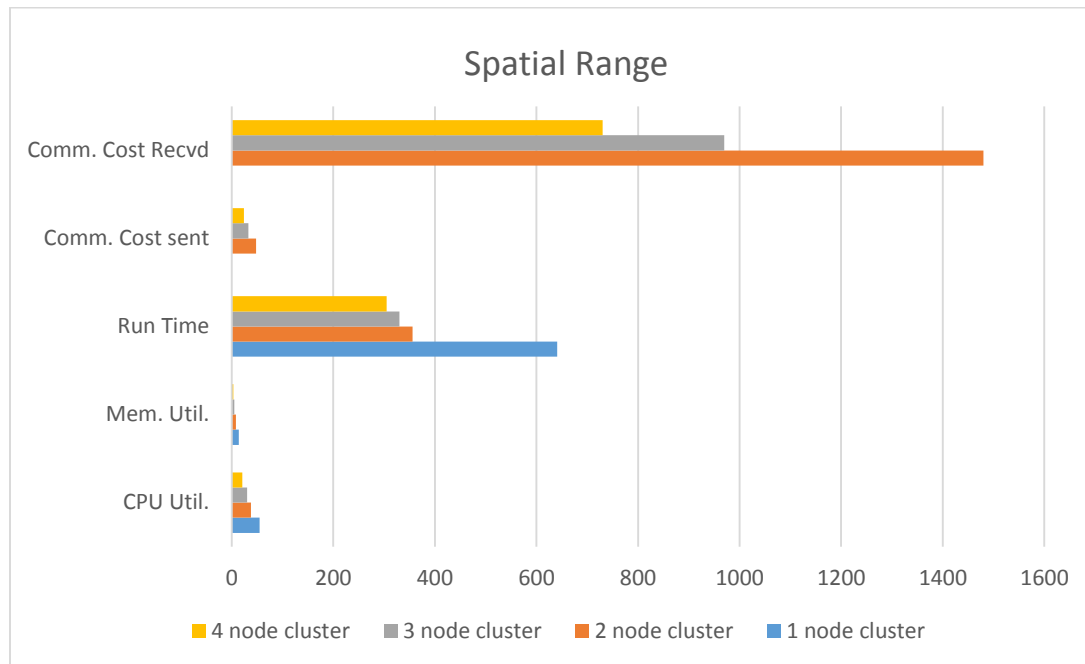
<b>2 Nodes:</b>	Job0- 9.8	Job1- 8.1
<b>3 Nodes:</b>	Job0- 4.94	Job1- 4.56
<b>4 Nodes:</b>	Job0- 3.67	Job1- 3.21

#### **D. Communication Cost**

<b>2 Nodes:</b>	Worker1Sent- 48.8K	Worker1Received- 370.8K
	Worker2Sent- 47.2K	Worker2Received- 367.9K
<b>3 Nodes:</b>	Worker1Sent- 32.5K	Worker1Received- 247.2K

Worker2Sent-	29.8K	Worker2Received-	238.5K
Worker3Sent-	33.3K	Worker3Received-	262.4K

4 Nodes:	Worker1Sent-	24.4 K	Worker1Received-	185.4K
	Worker2Sent-	22.7K	Worker2Received-	162.7K
	Worker3Sent-	19.9K	Worker3Received-	191.2K
	Worker4Sent-	26.7K	Worker4Received-	178.2K



## (ii.) Small Data Set

### A. CPU Utilization (Percentage Increase)

Single Node: Worker1- 40.5

2 Nodes: Worker1- 37.5      Worker2- 25

3 Nodes: Worker1- 25.2 ;Worker2- 26 ;Worker3- 27.5

4 Nodes: Worker1- 22 ;Worker2- 23.7 ;Worker3- 24 ;Worker4- 21.8

### B. Memory Utilization (Percentage Increase)

Single Node: Worker1- 7.9

2 Nodes: Worker1- 5.6 ;Worker2- 6.4

3 Nodes: Worker1- 2.8 ;Worker2- 4 ;Worker3- 3.3

4 Nodes: Worker1- 2.5 ;Worker2- 3.1 ;Worker3- 2.9 ;Worker4- 3.2

### C. Run Time (Seconds)

Single Node: 3.2

2 Nodes: Job0- 1.7      Job1- 1.5

3 Nodes: Job0- 0.9      Job1- 0.8

4 Nodes: Job0- 0.54          Job1- 0.4

#### **D. Communication Cost**

2 Nodes:    Worker1Sent- 21.1K          Worker1Received- 170.2K

             Worker2Sent- 23.5K          Worker2Received- 165.5K

3 Nodes:    Worker1Sent- 14K          Worker1Received- 113.6K

             Worker2Sent- 13.7K          Worker2Received- 118.2K

             Worker3Sent- 12.5K          Worker3Received- 121K

4 Nodes: Worker1Sent- 10.55K          Worker1Received- 85.K

             Worker2Sent- 9.78K          Worker2Received- 92.5K

             Worker3Sent- 10.21K          Worker3Received- 97.8K

             Worker4Sent- 8.9K          Worker4Received- 72.1K

### **6. SPATIAL JOIN QUERY**

#### **(i.) Big Data Set**

##### **A. CPU Utilization (Percentage Increase)**

Single Node: Worker1-55%

2 Nodes: Worker 1 - 30%    ;Worker 2- 43%

3 Nodes: Worker 1- 36%    ;Worker 2- 26.5%    ;Worker 3- 37%

4 Nodes: Worker 1- 23%    ;Worker 2- 24%          ;Worker 3- 25%    ;Worker 4 - 11%

##### **B. Memory Utilization (Percentage Increase)**

Single Node: Worker1- 15%

2 Nodes: Worker 1 - 8.5%          ;Worker 2- 12%

3 Nodes: Worker 1- 7.9%          ;Worker 2- 4%          ;Worker 3- 9%

4 Nodes: Worker 1- 3.5%          ;Worker 2- 3.2%          ;Worker 3- 4.5%    ;Worker 4 - 4.9%

##### **C. Run Time**

Single Node: 720.88 seconds

2 Nodes: Job0- 430.01 seconds          Job1- 425.67 seconds

3 Nodes: Job0- 390 seconds                  Job1- 400 seconds

4 Nodes: Job0- 362.05 seconds          Job1- 378 seconds

#### **D. Communication Cost**

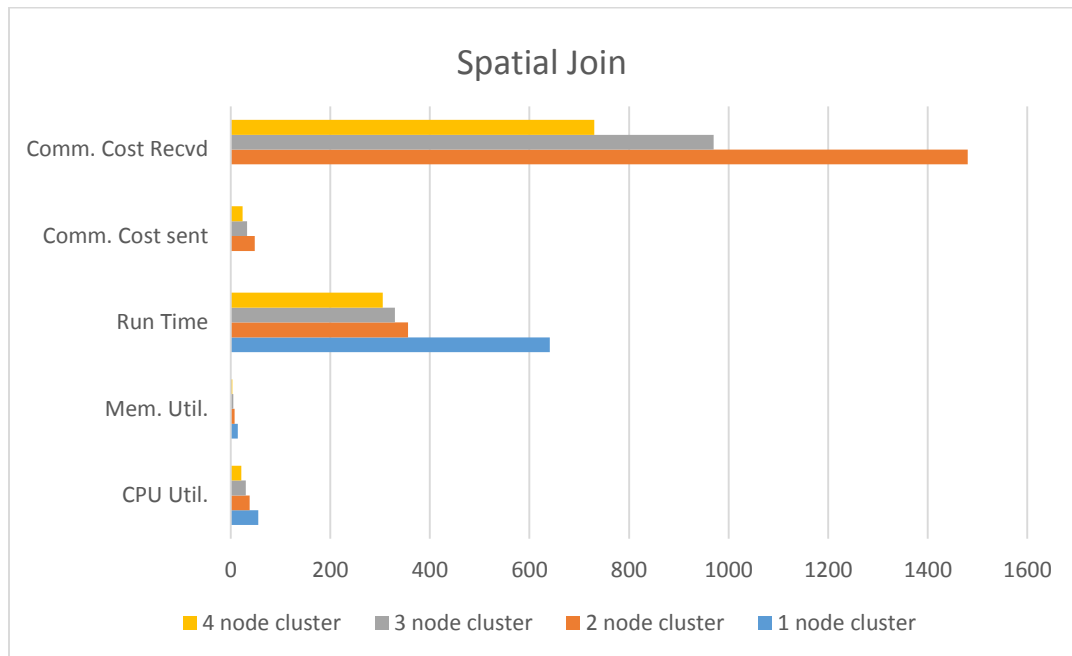
2 Nodes:          Worker1Sent- 78.8 K          Worker1Received- 1472.4K

             Worker2Sent- 77.3 K          Worker2Received- 1470 K

3 Nodes:          Worker1Sent- 50.5 K          Worker1Received- 981 K

Worker2Sent- 48.3 K	Worker2Received- 975.5 K
Worker3Sent- 51.5 K	Worker3Received- 978.06 K

4 Nodes:	Worker1Sent- 35.1 K	Worker1Received- 740 K
	Worker2Sent- 39.4 K	Worker2Received- 730.8 K
	Worker3Sent- 34.2 K	Worker3Received- 736.2 K
	Worker4Sent- 40.1 K	Worker4Received- 729.7 K



## (ii.) Small Data Set

### A. CPU Utilization (Percentage Increase)

Single Node: Worker1- 47%

2 Nodes: Worker 1 - 35% ;Worker 2- 22%

3 Nodes: Worker 1- 30% ;Worker 2- 28% ;Worker 3- 23%

4 Nodes: Worker 1- 20% ;Worker 2- 5% ;Worker 3- 8% ;Worker 4 - 15%

### B. Memory Utilization (Percentage Increase)

Single Node: Worker1- 8%

2 Nodes: Worker 1 - 3% ;Worker 2- 4%

3 Nodes: Worker 1- 2.3% ;Worker 2- 2.4% ;Worker 3- 4%

4 Nodes: Worker 1- 1.5% ;Worker 2- 2.3% ;Worker 3- 1.5% ;Worker 4 - 2.1%

### C. Run Time

Single Node: 9.7 seconds

2 Nodes: Job0- 4.6 seconds Job1- 3.78 seconds

3 Nodes: Job0- 4.1 seconds Job1- 3.2 seconds

4 Nodes: Job0- 3.61 seconds

Job1- 3.06 seconds

#### **D. Communication Cost**

2 Nodes: Worker1Sent- 25K Worker1Received- 506.4 K

Worker2Sent- 26K Worker2Received- 504.6 K

3 Nodes: Worker1Sent- 17K Worker1Received- 337.3K

Worker2Sent- 16.3K Worker2Received- 334.1K

Worker3Sent- 14.7K Worker3Received- 332K

4 Nodes: Worker1Sent- 13.9K Worker1Received- 256.1K

Worker2Sent- 12.7K Worker2Received- 253K

Worker3Sent- 14.1K Worker3Received- 259K

Worker4Sent- 13K Worker4Received- 249.5K

### **7. HEAT MAP**

#### **(i.) Big Data Set**

##### **A. CPU Utilization (Percentage Increase)**

Single Node: Worker1- 55%

2 Nodes: Worker 1 - 37% ;Worker 2- 39%

3 Nodes: Worker 1- 33% ;Worker 2- 25.9% ;Worker 3- 37%

4 Nodes: Worker 1- 12.9% ;Worker 2- 23% ;Worker 3- 13.6% ;Worker 4 - 26%

##### **B. Memory Utilization (Percentage Increase)**

Single Node: Worker1- 14.11%

2 Nodes: Worker 1 - 4.5% ;Worker 2- 14%

3 Nodes: Worker 1- 9% ;Worker 2- 2.5% ;Worker 3- 4.4%

4 Nodes: Worker 1- 3.7% ;Worker 2- 2.9% ;Worker 3- 6.7% ;Worker 4 - 3.1%

##### **C. Run Time**

Single Node: 641 seconds

2 Nodes: Job0- 356.19 seconds

Job1- 378.54 seconds

3 Nodes: Job0- 341.10 seconds

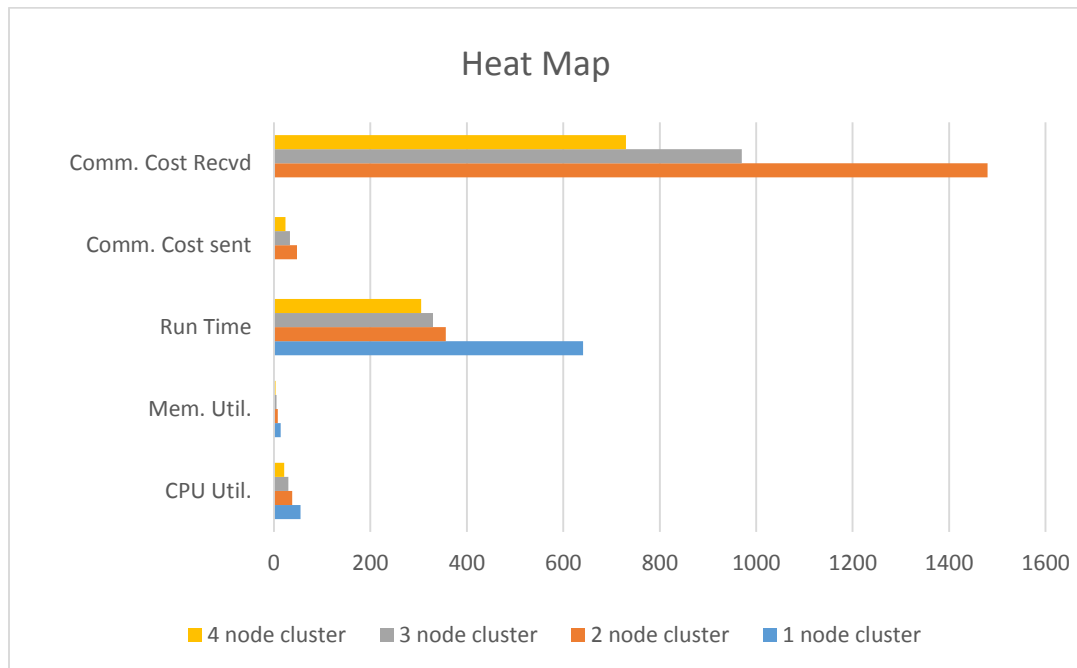
Job1- 323.77 seconds

4 Nodes: Job0- 310 seconds

Job1- 302 seconds

## D. Communication Cost

2 Nodes:	Worker1Sent- 49.2K Worker2Sent- 48.5K	Worker1Received- 1484.8K Worker2Received- 1396.5K
3 Nodes:	Worker1Sent- 32.8K Worker2Sent- 34.5K Worker3Sent- 33.3K	Worker1Received- 988.8K Worker2Received- 970.5K Worker3Received- 957.8K
4 Nodes:	Worker1Sent- 24.7K Worker2Sent- 22.8K Worker3Sent- 25.6K Worker4Sent- 24.2K	Worker1Received- 743.7K Worker2Received- 727.2K Worker3Received- 785K Worker4Received- 712.2K



## (ii.) Small Data Set

### A. CPU Utilization (Percentage Increase)

Single Node: Worker1- 50%

2 Nodes: Worker 1 - 34% ;Worker 2- 33%

3 Nodes: Worker 1- 30% ;Worker 2- 28% ;Worker 3- 32%

4 Nodes: Worker 1- 24% ;Worker 2- 31% ;Worker 3- 15% ;Worker 4 - 21%

### B. Memory Utilization (Percentage Increase)

Single Node: Worker1- 6.7%

2 Nodes: Worker 1 - 5.5% ;Worker 2- 4.2%

3 Nodes: Worker 1- 2.3% ;Worker 2- 3.4% ;Worker 3- 4%  
4 Nodes: Worker 1- 1.3% ;Worker 2- 2.1% ;Worker 3- 0.5% ;Worker 4 - 2.4%

### **C. Run Time**

Single Node: 9.2 seconds

2 Nodes: Job0- 4.2 seconds Job1- 3.7 seconds

3 Nodes: Job0- 4.1 seconds Job1- 3.2 seconds

4 Nodes: Job0- 3.05 seconds Job1- 2.98 seconds

### **D. Communication Cost**

2 Nodes: Worker1Sent- 27K Worker1Received- 506.7K

Worker2Sent- 25.4K Worker2Received- 545.5K

3 Nodes: Worker1Sent- 19.1K Worker1Received- 337.7K

Worker2Sent- 18.5K Worker2Received- 323.2K

Worker3Sent- 17.1K Worker3Received- 345.5K

4 Nodes: Worker1Sent- 13.3K Worker1Received- 247.9K

Worker2Sent- 15.1K Worker2Received- 256.2K

Worker3Sent- 11.2K Worker3Received- 221.6K

Worker4Sent- 14K Worker4Received- 233K