

**Illinois Institute of Technology**

**Advanced Operating System**  
**(CS-550)**

**PA-1**

**Submitted By:**

**Batkishig Dulamsurankhor (#A20543498)**

**Nitin Singh (#A20516824)**

**Vaishnavi Papudesi Babu (#A20527963)**

**Instructor: Professor Ioan Raicu**

**TA: Mr. Lan Nguyen**

**TA: Ms. Sonal Gaikwad**

**TA: Mr. Adarsh Agrawal**





1KB\_056492\_peer1.txt  
1KB\_461688\_peer1.txt  
1KB\_492521\_peer1.txt  
1KB\_075774\_peer1.txt  
1KB\_317623\_peer1.txt  
1KB\_416917\_peer1.txt  
1KB\_360183\_peer1.txt  
1KB\_480946\_peer1.txt  
1KB\_273317\_peer1.txt  
1KB\_299009\_peer1.txt  
1KB\_324866\_peer1.txt  
1KB\_440553\_peer1.txt  
1KB\_129622\_peer1.txt  
1KB\_413501\_peer1.txt  
1KB\_413539\_peer1.txt  
1KB\_344764\_peer1.txt  
1KB\_211007\_peer1.txt  
1KB\_362162\_peer1.txt  
1KB\_280291\_peer1.txt  
1KB\_429237\_peer1.txt  
1KB\_441600\_peer1.txt  
1KB\_306787\_peer1.txt  
1KB\_243917\_peer1.txt  
1KB\_082592\_peer1.txt  
1KB\_471355\_peer1.txt  
1KB\_408880\_peer1.txt  
1KB\_140317\_peer1.txt  
1KB\_088106\_peer1.txt  
1KB\_285571\_peer1.txt  
1KB\_104620\_peer1.txt  
1KB\_082220\_peer1.txt  
1KB\_138795\_peer1.txt  
1KB\_147596\_peer1.txt  
1KB\_029247\_peer1.txt  
1KB\_010214\_peer1.txt



## Search files

Peer started!

[Server]: Waiting for a peer ...

[Client]: What do you want to do?

[0] - Register

[1] - Search a file

[2] - Obtain a file

1

[[Client]: File name:z1.txt

[Client]: Connected!

[Client]: peer1 peer1.mynetwork

[Client]: What do you want to do?

[0] - Register

[1] - Search a file

[2] - Obtain a file

1

[[Client]: File name:1KB\_130040\_peer1.txt

[Client]: Connected!

[Client]: peer1 peer1.mynetwork

[Client]: What do you want to do?

[0] - Register

[1] - Search a file

[2] - Obtain a file

1

[Client]: File name:



*Do a weak scaling scalability study to measure search time of 10K requests per peer, on 1 node and 2 nodes. Report the average and standard deviation. Plot your data in figures graphically.*

Each file request attempt time stamps are captured in the logs,

file - weak\_scale\_search\_1\_node.log

**1 node: mean and standard deviation**



<b>Standard Deviation</b>	<b>s = 1.178265</b>
Variance	$s^2 = 1.3883083$
Count	n = 10000
Mean	$\bar{x} = 5.0757$
Sum of Squares	SS = 13881.695

Solution

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{\frac{SS}{n - 1}}$$

$$s = \sqrt{\frac{13881.695}{10000 - 1}}$$

$$s = \sqrt{\frac{13881.695}{9999}}$$

$$s = \sqrt{1.3883083}$$



Each file request attempt time stamps are captured in the logs,  
file - weak\_scale\_search\_2\_nodes.log

**2 node: mean and standard deviation**



<b>Standard Deviation</b>	<b>s = 2.1986762</b>
Variance	$s^2 = 4.8341772$
Count	n = 10000
Mean	$\bar{x} = 11.7175$
Sum of Squares	SS = 48336.938

---

Solution

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

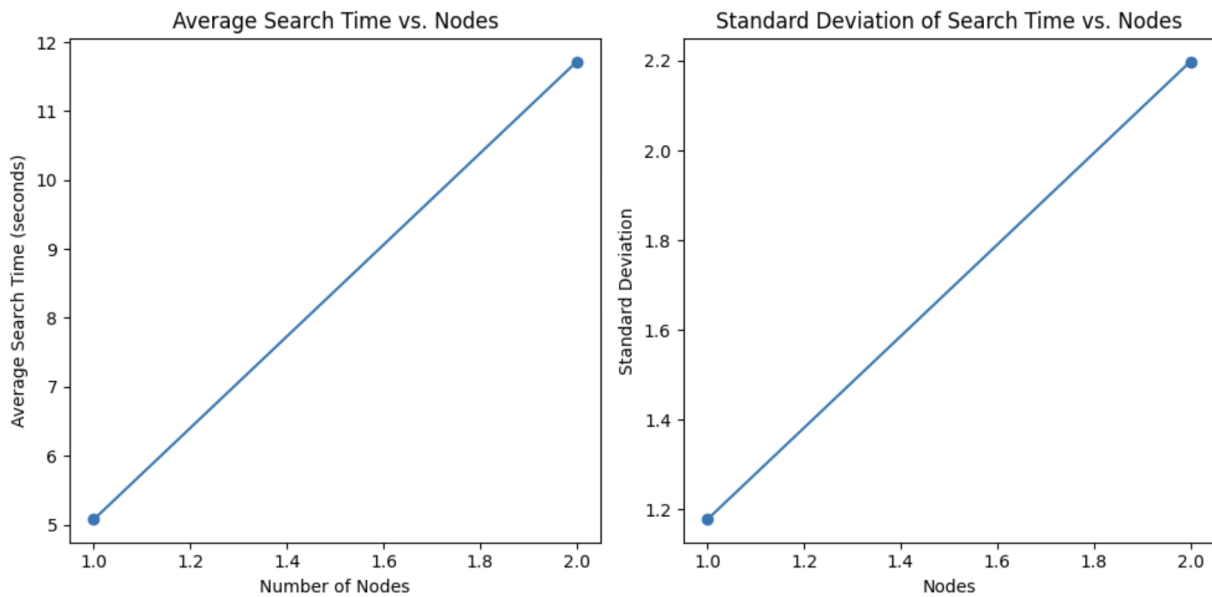
$$s = \sqrt{\frac{SS}{n - 1}}$$

$$s = \sqrt{\frac{48336.938}{10000 - 1}}$$

$$s = \sqrt{\frac{48336.938}{9999}}$$







*Do a strong scaling scalability study that measures the search and transfer time of 10K small files (1KB), on 1 node and 2 nodes. Repeat the study on 1K medium files (1MB). Repeat the study on 8 large files (1GB). Report the average and standard deviation. Plot your data in figures graphically.*

1 Node - Strong scale - 10K 1KB files test performed from peer2 to the files available in peer1:  
Log file: strong\_scale\_search\_1KB\_node1.log



<b>Standard Deviation</b>	<b><math>s = 2.0141041</math></b>
Variance	$s^2 = 4.0566154$
Count	$n = 10000$
Mean	$\bar{x} = 6.8095$
Sum of Squares	$SS = 40562.097$

---

Solution

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{\frac{SS}{n - 1}}$$

$$s = \sqrt{\frac{40562.097}{10000 - 1}}$$

$$s = \sqrt{\frac{40562.097}{9999}}$$



1 Node - Strong scale - 1K 1MB files test performed from peer2 to the files available in peer1:  
File name: strong\_scale\_search\_1MB\_node1.log



<b>Standard Deviation</b>	<b><math>s = 15.118464</math></b>
Variance	$s^2 = 228.56794$
Count	$n = 1000$
Mean	$\bar{x} = 30.925$
Sum of Squares	$SS = 228339.38$

---

Solution

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{\frac{SS}{n - 1}}$$

$$s = \sqrt{\frac{228339.38}{1000 - 1}}$$

$$s = \sqrt{\frac{228339.38}{999}}$$

$$s = \sqrt{228.56795}$$

$$s = 15.118464$$



1 Node - Strong scale - 8 1GB files test performed from peer2 to the files available in peer1:  
Filename: strong\_scale\_search\_1GB\_node1.log



<b>Standard Deviation</b>	<b>s = 686.582</b>
Variance	$s^2 = 471394.84$
Count	n = 8
Mean	$\bar{x} = 8905.625$
Sum of Squares	SS = 3299763.9

---

Solution

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{\frac{SS}{n - 1}}$$

$$s = \sqrt{\frac{3299763.9}{8 - 1}}$$

$$s = \sqrt{\frac{3299763.9}{7}}$$



2 Nodes - Strong scale - 10K 1KB files test performed from peer2 to the files available in peer1:  
Filename: strong\_scale\_search\_1KB\_nodes\_2\_peer1.log



<b>Standard Deviation</b>	<b>s = 6.8668427</b>
Variance	$s^2 = 47.153529$
Count	n = 9979
Mean	$\bar{x} = 18.677623$
Sum of Squares	SS = 470497.91

---

Solution

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{\frac{SS}{n - 1}}$$

$$s = \sqrt{\frac{470497.91}{9979 - 1}}$$

$$s = \sqrt{\frac{470497.91}{9978}}$$





2 Nodes - Strong scale - 10K 1KB files test performed from peer1 to the files available in peer2:  
Filename: strong\_scale\_search\_1KB\_nodes\_2\_peer2.log



<b>Standard Deviation</b>	<b>s = 6.8915257</b>
Variance	$s^2 = 47.493127$
Count	n = 10000
Mean	$\bar{x} = 18.8326$
Sum of Squares	SS = 474883.77

Solution

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{\frac{SS}{n - 1}}$$

$$s = \sqrt{\frac{474883.77}{10000 - 1}}$$

$$s = \sqrt{\frac{474883.77}{9999}}$$



2 Nodes - Strong scale - 1K 1MB files test performed from peer1 to the files available in peer1:  
Filename: strong\_scale\_search\_1MB\_nodes\_2\_peer1.log



<b>Standard Deviation</b>	<b>s = 22.437087</b>
Variance	$s^2 = 503.42286$
Count	n = 1000
Mean	$\bar{x} = 40.242$
Sum of Squares	SS = 502919.44

---

Solution

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{\frac{SS}{n - 1}}$$

$$s = \sqrt{\frac{502919.44}{1000 - 1}}$$

$$s = \sqrt{\frac{502919.44}{999}}$$



2 Nodes - Strong scale - 1K 1MB files test performed from peer1 to the files available in peer2:  
Filename: strong\_scale\_search\_1MB\_nodes\_2\_peer2.log



<b>Standard Deviation</b>	<b>s = 22.158029</b>
Variance	s <sup>2</sup> = 490.97825
Count	n = 1000
Mean	$\bar{x}$ = 39.223
Sum of Squares	SS = 490487.27

---

Solution

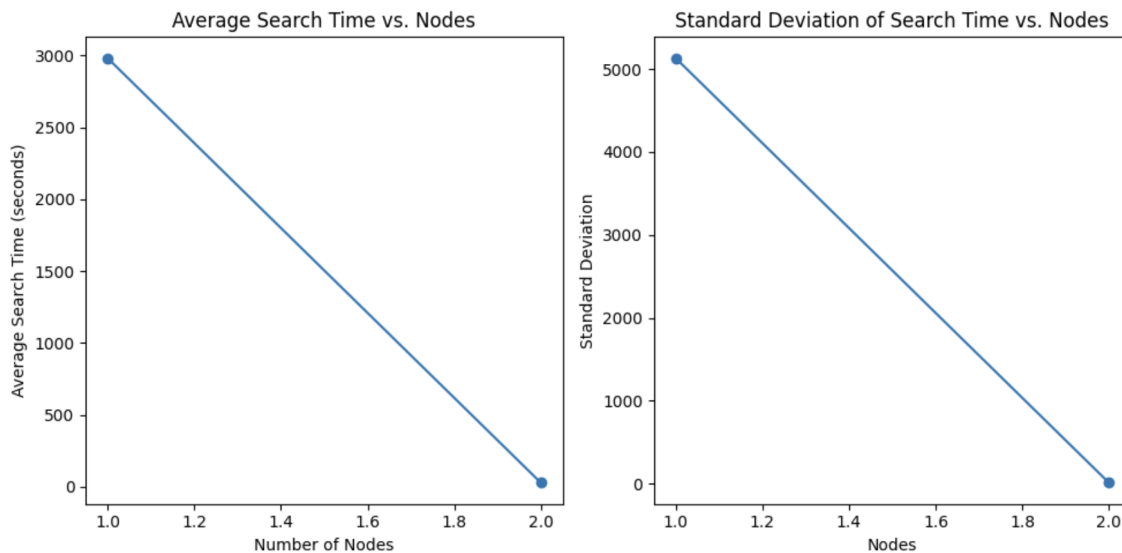
$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

$$s = \sqrt{\frac{SS}{n - 1}}$$

$$s = \sqrt{\frac{490487.27}{1000 - 1}}$$

$$s = \sqrt{\frac{490487.27}{999}}$$





*Can you deduce that your P2P centralized system is scalable up to 2 nodes? Does it scale well for some file sizes, but not for others? Based on the data you have so far, what would happen if you had 1K peers with small, medium, and large files? What would happen if you had 1 billion peers?*

i) Yes, scaling up to two nodes is doable. Scalability, however, is influenced by a variety of elements, such as the system's architecture, design, and implementation. Every peer (node) in a typical P2P system communicates directly with every other peer on the network. As a result, if your system is designed to run on just two nodes, it can be regarded as scalable because it can effectively manage a relatively small number of nodes.

A p2p system, on the other hand, is typically designed to function with additional p2p nodes to distribute the load and give fault tolerance. Therefore, it might be a little inaccurate to refer to it as "scalable" with only 2 p2p nodes.

ii) The size of the files shared can also affect scalability. We encountered this issue while working on a P2P system. Some p2p systems work well with small files, while others struggle with larger files because of bandwidth and latency limitations.

iii) Managing 1,000 p2p connections can be very resource-intensive, and it would require powerful algorithms to exchange data efficiently.

The performance of the system would depend on the capabilities and design of the system.



Small p2p files can be widely distributed among peers, particularly if the system supports simultaneous downloading and uploading.

Larger p2p files may have issues with bandwidth, storage and latency, but medium-sized p2p files might also work well.

iv) If your P2P system is centralized, one billion peers is a sizable number of peers. You'll need to control connections, monitor your resources, and make sure your data transmission and reception processes are effective. All of this would require a significant amount of infrastructure, bandwidth, and processing power.

At such a size, centralized systems could experience congestion and single points of failure, so you might want to think about decentralization or a hybrid strategy. To effectively manage a large number of peers, you might also want to think about how centralized or decentralized your system is.

