TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI

VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

──────── \* ───────

**BÁO CÁO**

MÔN: Phát triển phần mềm phân tán

Nhóm sinh viên thực hiện: **Trần Thị Phượng**

**Hà Bảo Khiêm**

Giáo viên hướng dẫn: TS. **Nguyễn Nhất Hải**

***Hà Nội, tháng 4 năm 2021***

# **MỤC LỤC**

[**MỤC LỤC** 2](#_Toc70453124)

[**PHẦN I:** 3](#_Toc70453125)

[**1.** **Tổng quan về hệ phân tán:** 3](#_Toc70453126)

[**1.1.** **Hệ phân tán là gì?** 4](#_Toc70453127)

[**1.2.** **Đặc điểm của hệ phân tán** 4](#_Toc70453128)

[**1.3.** **Tính chia sẻ tài nguyên.** 4](#_Toc70453129)

[**1.4.** **Tính co giãn** 5](#_Toc70453130)

[**1.5.** **Tính trong suốt** 5](#_Toc70453131)

[**1.6.** **Tính mở** 5](#_Toc70453132)

[**2.** **Các thành phần hệ phân tán** 6](#_Toc70453133)

[**PHẦN II:** 8](#_Toc70453134)

[**1.** **Phát biểu bài toán:** 8](#_Toc70453135)

[**2.** **Phân tích bài toán:** 9](#_Toc70453136)

[**a.** **Mục tiêu:** 9](#_Toc70453137)

[**b.** **Các yếu tố ảnh hưởng** 9](#_Toc70453138)

[**c.** **Giải quyết bài toán bằng Top-sis Method** 10](#_Toc70453160)

[**3.** **Demo sản phẩm.** 10](#_Toc70453161)

# **PHẦN I:**

**CƠ SỞ LÝ THUYẾT**

# **Tổng quan về hệ phân tán:**

Do nhu cầu tăng không ngừng việc chia sẻ tài nguyên và thông tin mà các hệ điều hành đã có từ trước không đáp ứng được. Trong quá trình triển khai ứng dụng Tin học vào đời sống, các mạng máy tính được phát triển không ngừng, các tài nguyên của các máy tính trong mạng (phần cứng, phần mềm) ngày càng được mở rộng và nâng cấp, giá trị các tài nguyên này càng tăng nhanh dẫn đến sự tăng trưởng vượt bậc nhu cầu chia sẻ tài nguyên và thông tin trong một hệ thống thống nhất.

Hệ điều hành tập trung và hệ điều hành mạng thuần túy không đáp ứng được nhu cầu đối với sự tăng trưởng đó. Việc giá các trạm làm việc giảm nhanh chóng: làm cho chúng được sử dụng phổ dụng hơn, số lượng và chất lượng các trạm làm việc cũng tăng không ngừng, từ đó làm tăng yêu cầu xử lý phân tán trên cơ sở việc kết nối mạng để triển khai hệ điều hành mạng tạo nên một cơ sở kỹ thuật hạ tầng (phần cứng, kết nối mạng, phần mềm) làm nền tảng phát triển hệ phân tán.

* 1. **Hệ phân tán là gì?**

Hệ thống phân tán có thể coi là 1 hệ thống tính toán với các thành phần tính toán được phân bố trên các vị trí địa lý khác nhau.

Là tập hợp các máy tính độc lập, không phụ thuộc lẫn nhau, kết nối với nhau bằng 1 hạ tầng truyền thông, có hạ tầng phần cứng và phần mềm khác nhau kết nối với nhau bằng các công nghệ mạng (mạng máy tính), có khả năng phối hợp và chia sẻ tài nguyên, thực hiện một nhiệm vụ chung và cung cấp các tài nguyên tính toán cho người sử dụng dưới dạng nhất định.

Một hệ phân tán cần thống nhất về giao diện cũng như cách thức truy cập dịch vụ. Người sử dụng không cần phải quan tâm tới các chi tiết của hệ thống.

* 1. **Đặc điểm của hệ phân tán**

Hệ phân tán có 4 đặc điểm đặc trưng sau:

* Chia sẻ tài nguyên
* Tính trong suốt
* Tính mở
* Tính co giãn
  1. **Tính chia sẻ tài nguyên.**

Các tài nguyên trên máy tính được quản lý bởi chương trình quản lý tài nguyên. Chương trình quản lý tài nguyên có khả năng nhận các yêu cầu do các chương trình khác gửi đến, chuyển các yêu cầu này thành các yêu cầu truy cập tài nguyên vật lý rồi nhận trả lời từ tài nguyên vật lý và cung cấp ngược lại cho các chương trình

Lợi ích của việc chia sẻ tài nguyên: Tiết kiệm chi phí đầu tư => số lượng các thiết bị ngoại vi đầu tư cho máy tính giảm => giảm suất đầu tư trên từng người sử dụng Việc cho phép NSD kết nối các tài nguyên ở xa và các máy khác nhau làm tăng khả năng sẵn sàng của hệ thống

Nhược điểm: Chương trình có những kết nối mạng => có lỗ hổng bảo mật => giảm mức độ an toàn bảo mật của hệ thống. Khi quá trình chia sẻ thông tin kéo dài việc theo dõi tất cả thông tin được phép chia sẻ có thể tìm ra những thông tin ẩn, từ đó có những thông tin liên quan đến tính riêng tư có thể bị lộ.

* 1. **Tính co giãn**

Tính co giãn của hệ phán tán được thể hiện qua : Qui mô: Khi số lượng người dùng và tài nguyên thay đổi hệ phân tán phải thích nghi được. Không gian địa lý: Khi vùng địa lý có tài nguyên và người sử dụng thay đổi. Tổ chức: Khi tổ chức kết cấu của hệ thay đổi

* 1. **Tính trong suốt**

Tính trong suốt là khả năng cung cấp một khung cảnh lôgic của hệ thống cho người dùng, độc lập với hạ tầng vật lý. Hệ thống luôn là duy nhất đối với người dùng song nó sẽ che giấu được tính phân tán của hệ phân tán phía dưới.

Trong suốt được xem xét dưới nhiều góc độ khác nhau:

Trong suốt truy cập (Access): Che giấu sự khác nhau trong biểu diễn dữ liệu và cách thức truy cập tài nguyên.

* Trong suốt vị trí (Location): Che giấu vị trí của tài nguyên, vị trí của TN không bị người sử dụng nhìn thấy.
* Trong suốt di trú (Migration): Che giấu việc tài nguyên chuyển đến địa điểm khác.
* Trong suốt về việc chuyển địa điểm (Relocation): Che giấu việc tài nguyên chuyển đến địa điểm khác ngay trong khi đang được sử dụng.
* Trong suốt sao lưu (Replication): Che giấu việc dữ liệu được cung cấp từ nhiều bản sao khác nhau (thường được sử dụng rỗng rãi trong các hệ thống phân tán để tăng hiệu năng và tính sẵn sàng của hệ thống)
* Trong suốt tương tranh (Concurency): Che giấu việc tài nguyên được truy cập đồng thời bởi nhiều người sử dụng.
* Trong suốt sự cố (Failure): Che giấu lỗi và quá trình phục hồi của tài nguyên
* Trong suốt bền vững (Persistence): Che giấu việc tài nguyên/dữ liệu được lưu trữ bền vững (disk) hoặc không (RAM)

Việc đảm bảo tính trong suốt là 1 trong những yêu cầu chắc chắn phải thực hiên để đảm bảo định nghĩa của hệ thống phân tán. Tuy nhiên để có được tính trong suốt ở mức độ tuyệt đối sẽ kéo theo chi phí về TN rất cao. Do đó không phải lúc nào cũng hướng tới trong suốt tuyệt đối => Cần xem xét trường hợp nào cần trong suốt đến đâu để tiết kiệm chi phí.

* 1. **Tính mở**

1 hệ thống có tính mở là hệ thống cho phép các thành phần được sản xuất bởi các nhà sản xuất khác nhau và có thể thay thế lẫn nhau, đồng thời cũng có khả năng cho phép thành phần mới bổ sung vào hệ thống. Hệ phân tán mở cung cấp các dịch vụ theo các đặc tả về cú pháp và ngữ nghĩa của các dịch vụ hay còn gọi là giao diện.

Trong giao diện có 2 thành phần : Thành phần cài đặt giao diện: chịu trách nhiệm cung cấp dịch vụ cho thành phần khác Thành phần sử dụng giao diện: sử dụng dịch vụ do các thành phần các cung cấp. => Điều kiện để 2 thành phần này có thể tương tác, phối hợp nhau là chúng cài đặt và sử dụng cùng 1 giao diện.

Để cài đặt và sử dụng 1 giao diện cần có các điều kiện sau:

* Đầy đủ. Nếu giao diện không quy định đầy đủ để các thành phần có thể sử dụng và cài đặt thì khi sử dụng và cài đặt các thành phân sẽ tự bổ sung vào những thành phần của giao diện cho đầy đủ => người cài đặt và người sử dụng bổ sung các kiểu khác nhau => không giao tiếp được
* Trung lập: độc lập, không phụ thuộc vào công nghệ, nền tảng, hạ tầng nào cả. Nó chi định nghĩa chung tương tác giữa 2 thành phần

Sử dụng ngôn ngữ giao diện gọi là IDL (Interface Definition Language) để đảm bảo tính trung lập của giao diện cũng như thuận tiện hơn trong việc vaildate xem giao diện có đầy đủ hay không

1. **Các thành phần hệ phân tán**

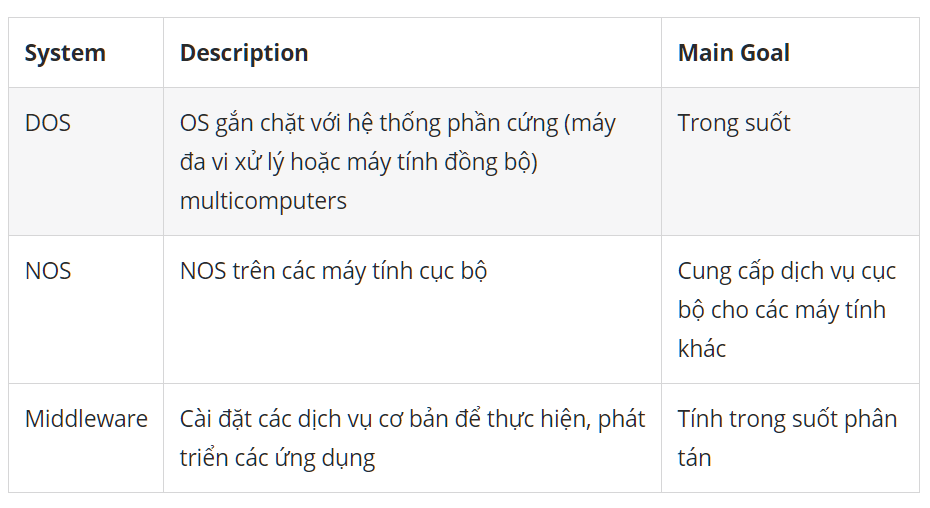
*Phần cứng*

Hệ thống có thể đơn vi xử lý hoặc đa vi xử lý. Hiện nay là các thiết bị đa vi xử lý được sử dụng rộng rãi

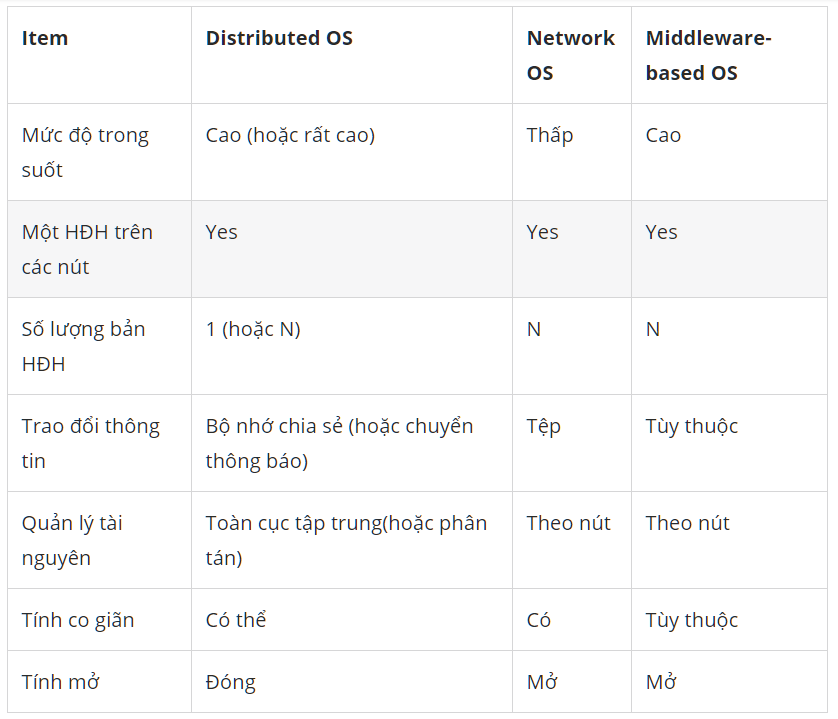
Trong các thiết bị tính toán sử dụng đa vi xử lý, thông thường các bộ vi xử lý kết nối với bộ nhớ thông qua 1 trục thông tin chung của cả hệ thống. Trong các hệ thống xử lý tính toán với hiệu năng rất lớn có nhiều CPU và module bộ nhớ thì việc sử dụng 1 trục chung cho tất cả các CPU và CPU bộ nhớ dẫn đến việc khi 1 cặp CPU bộ nhớ sử dụng trục thì những cặp khác không sử dụng được => thời gian chờ dài => hệ thống có nhiều CPU và module bộ nhớ sẽ sử dụng kiên trúc khác như: kiến trúc có 1 switch, kiến trúc có hệ thống chuyển mạch với tốc độ truyền cao, tuy nhiên đòi hỏi chi phí chế tạo cao.

Với các tổ chức nhỏ hơn khi người ta có nhu cầu về thiết bị tính toán có hiệu năng cao đồng thời không có đủ khả năng tài chính để có thể mua những super computer thì người ta chọn phương án hiệu quả hơn là mua nhiều các máy tính mini computer hay các super computer với hiệu năng thấp hơn, giá thành nhỏ hơn và tìm cách làm cho hiệu năng tổng của các máy tính đó xấp xỉ bằng tổng hiệu năng của tất cả các máy tính con => gọi là các máy tính bó (hay còn goi là computer classter) Computer classter dùng cho nhiều mục đích khác nhau như dùng cho các máy chủ tính toán, các máy chủ hosting, ... Trong các hệ thống computer classter để có thể kết nối với nhau thông thường các máy tính phải tương đối giống nhau => gọi là tập hợp các máy tính đồng nhất Trường hợp các máy tính sử dụng cho các mục đích khác nhau nhưng lại muốn kết nối với nhau => gọi là tập hợp các máy tính không đồng nhất

*Phần mềm*



*So sánh các phần mềm của hệ phân tán:*



# **PHẦN II:**

**PHÁT TRIỂN PHẦN MỀM PHÂN TÁN BÀI TOÁN HỖ TRỢ TUYỂN ỨNG VIÊN TRONG CÔNG TY.**

# **Phát biểu bài toán:**

Phần mềm quản trị tuyển dụng với mục đích tối ưu hóa quy trình tuyển dụng,tiết kiệm thời gian và chi phí . Nó giúp cho nhà tuyển dụng và người quản lý cùng nhau giải quyết câu chuyện

(1) Đăng tin tuyển dụng,

(2) Thu thập – Sắp xếp,

(3) Sàng lọc,

(4) Đánh giá,

(5) Lựa chọn ứng viên, và đảm bảo đưa ứng viên đi qua quá trình tuyển dụng trên một hệ thống duy nhất Ban quản lý việc tuyển dụng có thể quản lý, theo dõi tốt quá trình tuyển dụng ( có thể xem thông tin về việc tuyển dụng tại bất kỳ thời điểm nào). Đánh giá được chất lượng của quá trình tuyển dụng

# **Phân tích bài toán:**

* 1. **Mục tiêu:**

Đưa ra được phương án tối ưu nhất để công ty có thể tuyển ứng viên.

* 1. **Các yếu tố ảnh hưởng**

Biến quyết định:

Ứng viên

Biến môi trường:

Trình độ học vấn

Lương

Kinh nghiệm làm việc

Chỉ tiêu...

*Các thông tin chung trong CSDL:*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Tên | Địa chỉ  A1 | Trình độ học vấn  A2 | Lương  A3 | Thái độ  A4 | Khả năng làm việc  A5 | Chỉ tiêu | Vị trí ứng tuyển |

*Các thông tin riêng của ứng viên*

|  |  |  |  |
| --- | --- | --- | --- |
| Địa chỉ  B1 | Trình độ học vấn  B2 | Lương đề xuất  B3 | Ngành học  B4 |

Các thành phần của bảng quyết định bao gồm:

1. Sự phù hợp về trình độ học vấn (Trọng số 0.3)

C1=1 nếu B2>=2A2

C1=0.5 nếu 2A2>B2>=A2

C1=0 nếu B2<A2

1. Lương đề xuất (Trọng số 0.1)

C2=1 nếu B3<=A3  
C2=0.5 nếu A3<B3<2B3  
C2=0 nếu B3>2A3

1. Sự phù hợp về địa điểm (Trọng số 0.1)

C3 = 1 với

|A3 - B3| = 0

|A3 - B3| = 1

|A3 - B3| = 2

Với miền bắc ứng giá trị = 1, miền trung ứng giá trị = 2, miền nam ứng giá trị = 3

Ngoài ra C3=0

1. Sự phù hợp về thái độ (Trọng số 0.2)

C4=1 nếu “Thái độ tốt”  
C4=0 nếu “Thái độ không tốt”

1. Sự phù hợp về khả năng làm việc: (Trọng số 0.3)

C5=1 nếu “Khả năng làm việc tốt”  
C5=0.5 nếu “Khả năng làm việc khá”  
C5=0 nếu “Khả năng làm việc trung bình”

* 1. **Giải quyết bài toán bằng Top-sis Method**
* Bước 1: chuẩn hoá, đưa các giá trị về rij thuộc [0,1]
* Bước 2: tính giá trị theo trọng số vij = rij \* wj
* Bước 3: tính các giải pháp lý tưởng A\* = (v1\*,v2\*,…,vn\*), với vj\* là giá trị tốt nhất của Xj A– = (v1–,v2–,…,vn–), với vj– là giá trị xấu nhất của Xj
* Bước 4: tính khoảng cách so với A\*, A– Si\* = (Σj (vij – vj\*)2)1/2, Si– = (Σj (vij – vj–)2)1/2
* Bước 5: tính độ tương tự: Ci\* = Si– / (Si\*+Si–)

# **Demo sản phẩm.**

