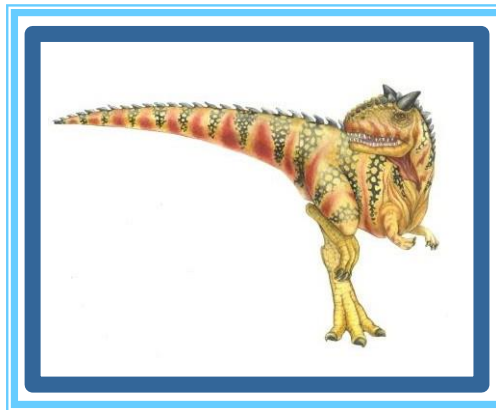
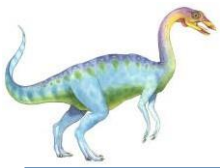


Chapter 9:

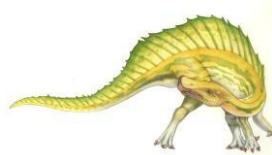
Distributed Systems

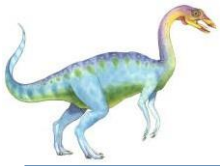




Chapter 9: Distributed Systems

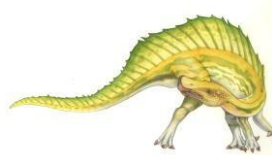
- Advantages of Distributed Systems ข้อดีของระบบแบบกระจาย
- Network Structure โครงสร้างเครือข่าย
- Communication Protocols โพรโทคอลการสื่อสาร
- Network and Distributed Operating Systems เครือข่ายและระบบปฏิบัติการแบบกระจาย
- Design Issues of Distributed Systems ปัญหาการออกแบบระบบแบบกระจาย

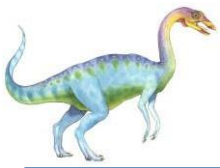




Chapter Objectives

- อธิบายข้อดีของระบบเครือข่ายและแบบกระจาย
- Explain the advantages of networked and distributed systems
- กำหนดบทบาทและประเภทของระบบแบบกระจายที่ใช้อยู่ในปัจจุบัน
- Define the roles and types of distributed systems in use today





Overview

ระบบแบบกระจายคือชุดของโหนดที่เชื่อมต่อกันอย่างหลวมๆ ซึ่งเชื่อมต่อกันด้วยเครือข่ายการสื่อสาร

- A **distributed system** is a collection of **loosely coupled nodes** interconnected by a communications network

โหนดต่างๆ ที่เรียกว่าโปรเซสเซอร์, คอมพิวเตอร์, เครื่องจักร, โฮสต์

- Nodes variously called **processors, computers, machines, hosts**

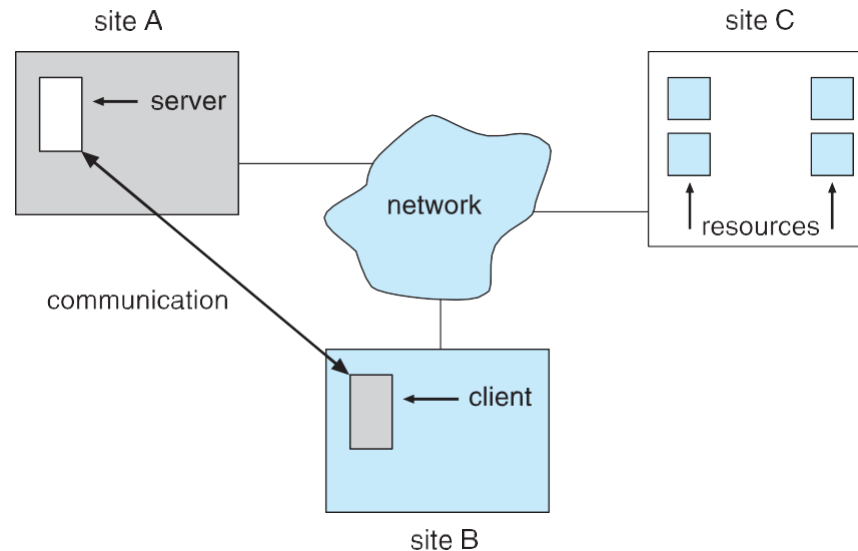
โหนดคือตำแหน่งของเครื่อง, โหนดหมายถึงระบบเฉพาะ

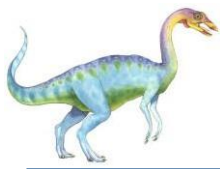
a specific system at a site

- **Site** is **location** of the machine, **node** refers to **specific system**

- Generally a **server** has a resource a **client** node at a **different site wants to use**

โดยทั่วไปเซิร์ฟเวอร์จะมีทรัพยากรที่โหนดไคลเอ็นต์ที่ไซต์อื่นต้องการใช้



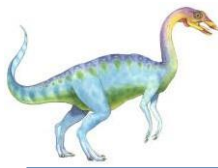


Overview (cont.)



- คณา ฟูโซมรพ ฟางน หดาคณได
นัชน - กวิ - นัชน
- Nodes may exist in a **client-server**, **peer-to-peer**, or **hybrid** configuration. โหนดอาจมีอยู่ในการกำหนดค่าไคลเอนต์-เซิร์ฟเวอร์, เพียร์ทูเพียร์, หรือไฮบริด
ในการกำหนดค่าไคลเอนต์-เซิร์ฟเวอร์ เซิร์ฟเวอร์มีทรัพยากรที่ไคลเอนต์ต้องการใช้
 - In client-server configuration, server has a resource that a client would like to use
ในการกำหนดค่าแบบเพียร์ทูเพียร์ แต่ละโหนดจะแบ่งความรับผิดชอบเท่ากันและสามารถทำหน้าที่เป็นทั้งไคลเอนต์และเซิร์ฟเวอร์ได้
 - In peer-to-peer configuration, each node shares equal responsibilities and can act as both clients and servers
 - Communication over a network occurs through **message passing**
การสื่อสารผ่านเครือข่ายเกิดขึ้นผ่านการส่งข้อความ
ฟังก์ชันระดับสูงกว่าทั้งหมดของระบบสแตนด์อโลนสามารถขยายให้ครอบคลุมระบบแบบกระจายได้
 - All higher-level functions of a standalone system can be expanded to encompass a distributed system





Reasons for Distributed Systems

การแบ่งปันทรัพยากร

□ Resource sharing

การแชร์ไฟล์หรือการพิมพ์ที่ไซต์ระยะไกล

- Sharing files or printing at remote sites
- การประมวลผลข้อมูลในฐานข้อมูลแบบกระจาย
- Processing information in a distributed database
- การใช้อุปกรณ์ฮาร์ดแวร์พิเศษระยะไกล เช่น หน่วยประมวลผลกราฟิก (GPU)
- Using remote specialized hardware devices such as **graphics processing units (GPUs)**

การเร่งความเร็วในการคำนวณ

□ Computation speedup

กระจายการคำนวณไปยังไซต์ต่างๆ เพื่อให้ทำงานพร้อมกัน

- Distribute subcomputations among various sites to run concurrently

- **Load balancing** – moving jobs to more lightly-loaded sites

ความน่าเชื่อถือ

□ Reliability

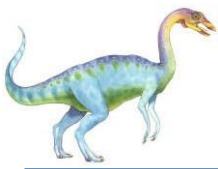
ตรวจจับและกู้คืนจากความล้มเหลวของไซต์ การถ่ายโอนฟังก์ชัน รวมไซต์ที่ล้มเหลวอีกครั้ง

- Detect and recover from site failure, function transfer, reintegrate failed site

ซึ่งงานไม่ซ้ำกัน

โหลดบาลานซ์ – ย้ายงานไปยังไซต์ที่มีการโหลดน้อยมากขึ้น





Network Structure

โครงสร้างเครือข่าย

เครือข่ายท้องถิ่น (LAN) – ออกแบบมาเพื่อครอบคลุมพื้นที่ทางภูมิศาสตร์ขนาดเล็ก

- **Local-Area Network (LAN)** – designed to cover small geographical area ประกอบด้วยคอมพิวเตอร์หลายเครื่อง (เวิร์กสเตชัน แล็ปท็อป อุปกรณ์พกพา) อุปกรณ์ต่อพ่วง (เครื่องพิมพ์ อาร์เรย์จัดเก็บข้อมูล) และเราเตอร์ที่ให้การเข้าถึงเครือข่ายอื่น

- Consists of multiple computers (workstations, laptops, mobile devices), peripherals (printers, storage arrays), and routers providing access to other networks

- Ethernet and/or Wireless (**WiFi**) most common way to construct LANs อิเทอร์เน็ตและ/หรือไร้สาย (**WiFi**) วิธีที่ใช้กันทั่วไปในการสร้าง LAN อิเทอร์เน็ตที่กำหนดโดยมาตรฐาน IEEE 802.3 ด้วยความเร็วโดยทั่วไปจะแตกต่างกันไปตั้งแต่ 10Mbps ถึงมากกว่า 10Gbps

- ▶ Ethernet defined by standard IEEE 802.3 with speeds typically varying from 10Mbps to over 10Gbps

- ▶ WiFi defined by standard IEEE 802.11 with speeds typically varying from 11Mbps to over 400Mbps.

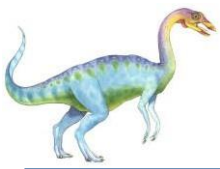
มาตรฐานทั้งสองมีการพัฒนาอย่างต่อเนื่อง

- ▶ Both standards constantly evolving

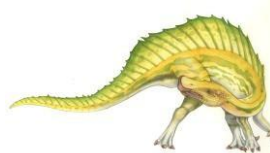
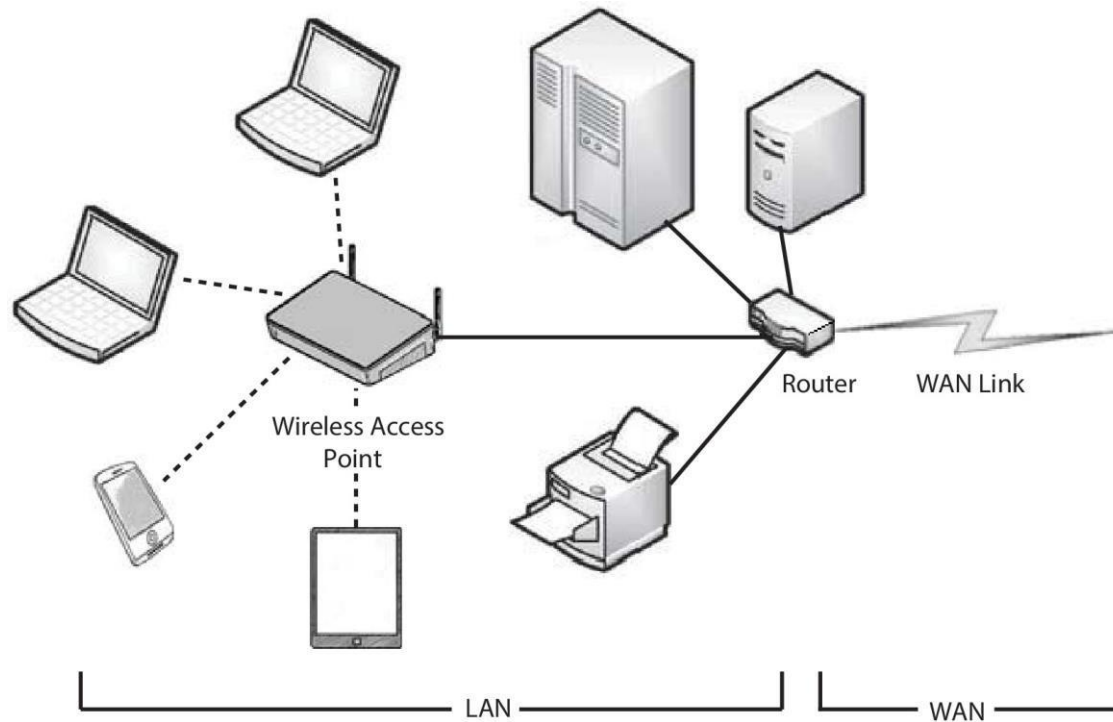
WiFi ที่กำหนดโดยมาตรฐาน IEEE 802.11 ด้วยความเร็วโดยทั่วไปจะแตกต่างกันไปตั้งแต่ 11Mbps ถึงมากกว่า 400Mbps

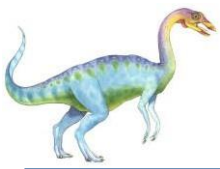
กันไปตั้งแต่ 11Mbps ถึงมากกว่า 400Mbps





Local-Area Network (LAN)





Network Structure (Cont.)

Wide-Area Network (WAN) – เชื่อมโยงไซต์ที่แยกออกจากกันทางภูมิศาสตร์

- **Wide-Area Network (WAN)** – links geographically separated sites

การเชื่อมต่อแบบจุดต่อจุดผ่านลิงก์

- Point-to-point connections via links

สายโทรศัพท์ สายเช่า (ข้อมูลเฉพาะ) สายเคเบิลอปติก ลิงค์ไมโครเวฟ คลื่นวิทยุ และช่องสัญญาณดาวเทียม

- ▶ Telephone lines, leased (dedicated data) lines, optical cable, microwave links, radio waves, and satellite channels

ใช้งานผ่านเราเตอร์เพื่อกำหนดเส้นทางการรับส่งข้อมูลจากเครือข่ายหนึ่งไปยังอีกเครือข่ายหนึ่ง

- Implemented via **routers** to direct traffic from one network to another

WAN อินเทอร์เน็ต (เวิลด์ไวด์เว็บ) ช่วยให้โฮสต์ทั่วโลกสามารถสื่อสารได้

- Internet (World Wide Web) WAN enables hosts world wide to communicate

ความเร็วแตกต่างกันไป

- Speeds vary

ผู้ให้บริการแบ็คโบนหลายรายมีความเร็วที่ 40-100Gbps

- ▶ Many backbone providers have speeds at 40-100Gbps

ผู้ให้บริการอินเทอร์เน็ตในพื้นที่ (ISP) อาจช้าลง

- ▶ Local **Internet Service Providers (ISPs)** may be slower

ลิงก์ WAN ได้รับการอัปเดตอย่างต่อเนื่อง

- ▶ WAN links constantly being upgraded

WAN และ LAN เชื่อมต่อระหว่างกัน คล้ายกับเครือข่ายโทรศัพท์มือถือ:

- WANs and LANs interconnect, similar to cell phone network:

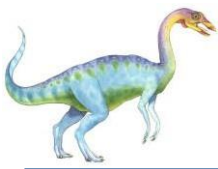
โทรศัพท์มือถือใช้คลื่นวิทยุไปยังเสาส่งสัญญาณ

- ▶ Cell phones use radio waves to cell towers

ทาวเวอร์เชื่อมต่อกับทาวเวอร์และฮับอื่นๆ

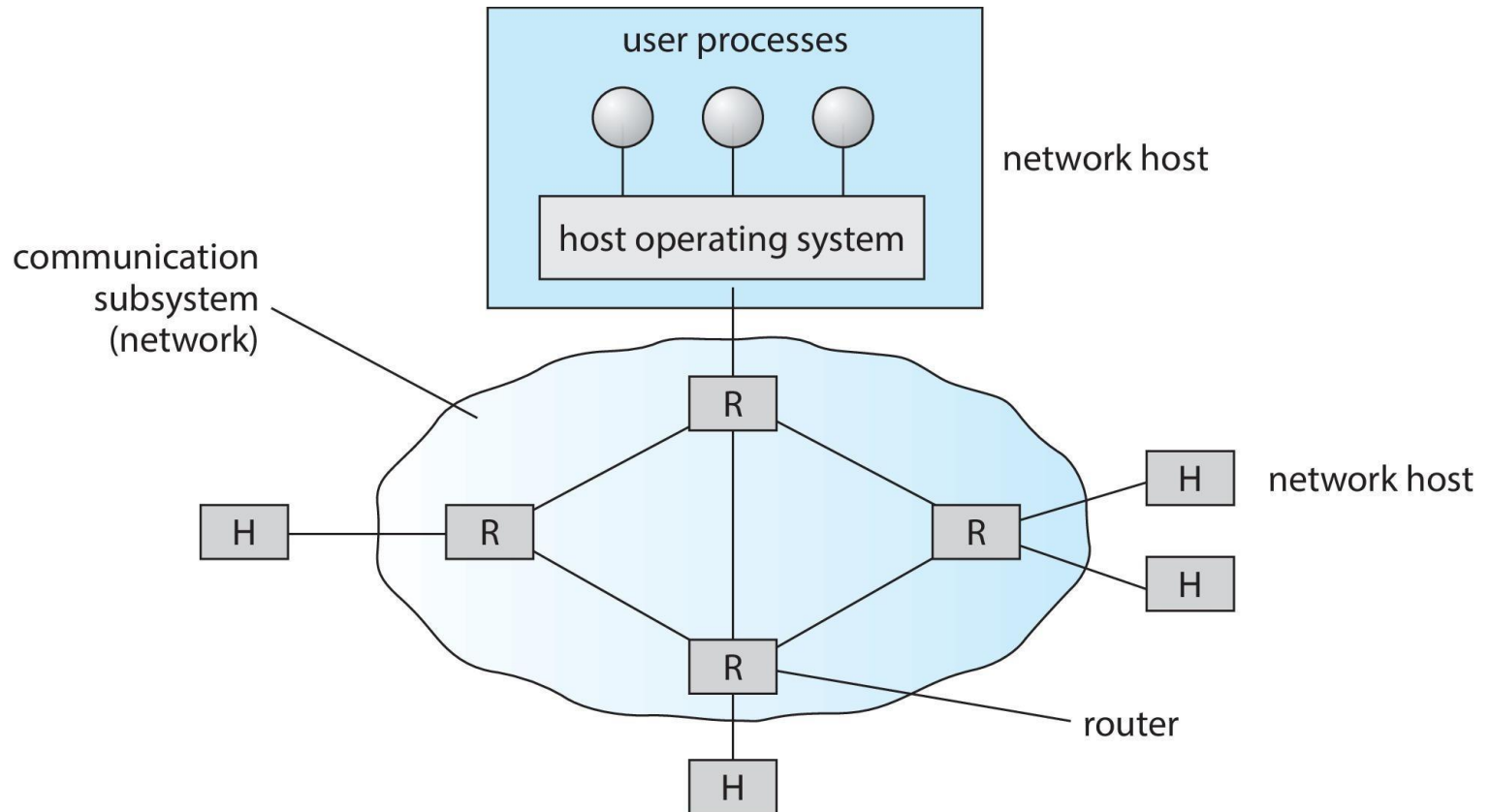
- ▶ Towers connect to other towers and hubs

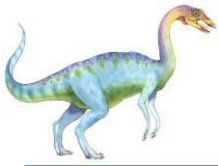




Wide-Area Network (WAN)

Diagram





Naming and Name Resolution

ระบบคอมพิวเตอร์แต่ละระบบในเครือข่ายมีชื่อไม่ซ้ำกัน

- Each computer system in the network has a unique name
- แต่ละกระบวนการในระบบที่กำหนดคิชื่อเฉพาะ (process-id)
- Each process in a given system has a unique name (process-id)
- ระบบกระบวนการบนระบบระยะไกลโดย
- Identify processes on remote systems by

<host-name, identifier> pair คู่ <ชื่อโฮสต์ ตัวระบุ>

- Domain name system (DNS) – specifies the naming structure of the hosts, as well as name to address resolution (Internet)

ระบบชื่อโดเมน

(DNS) – ระบบ

โครงสร้างการตั้งชื่อ

โฮสต์ ตลอดจนการ

แก้ไขชื่อไปยังที่อยู่

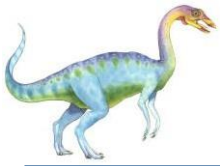
(อินเทอร์เน็ต)

```
/**
 * Usage: java DNSLookup <IP name>
 * i.e. java DNSLookup www.wiley.com
 */
public class DNSLookup {
    public static void main(String[] args) {
        InetAddress hostAddress;

        try {
            hostAddress = InetAddress.getByName(args[0]);
            System.out.println(hostAddress.getHostAddress());
        }
        catch (UnknownHostException uhe) {
            System.err.println("Unknown host: " + args[0]);
        }
    }
}
```

Figure 19.4 Java program illustrating a DNS lookup.





การสื่อสารผ่านโมเดล OSI (7 Layers)

- Communication via OSI model (7 Layers)

โปรโตคอลการขนส่ง:

- Transport Protocols:

TCP (โปรโตคอลควบคุมการส่ง)

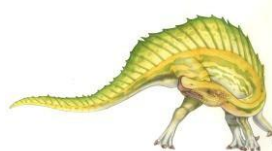
- TCP (Transmission Control Protocol)

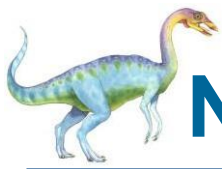
▶ Reliable and Connection-oriented เชื่อถือได้และเน้นการเชื่อมต่อ

UDP (โปรโตคอลเดตาแกรมผู้ใช้)

- UDP (User Datagram Protocol)

▶ Unreliable and Connectionless ไม่น่าเชื่อถือและไม่มีการเชื่อมต่อ



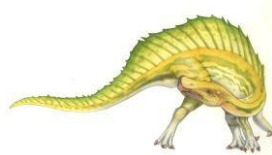


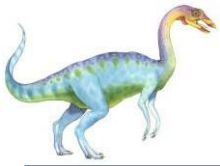
Network-oriented Operating Systems



- สองประเภทหลัก
- Two main types
- ระบบปฏิบัติการเครือข่าย
- **Network Operating Systems**
 - ผู้ใช้ตระหนักถึงความหลากหลายของเครื่องจักร
 - **Users are aware** of multiplicity of machines
- ระบบปฏิบัติการแบบกระจาย
- **Distributed Operating Systems**
 - ผู้ใช้ไม่ทราบถึงความหลากหลายของเครื่องจักร
 - **Users not aware** of multiplicity of machines

multiplicity : ความมากมายหลากหลายรูปแบบ





Network Operating Systems



ผู้ใช้ตระหนักถึงความหลากหลายของเครื่องจักร

- **Users are aware of multiplicity** of machines

การเข้าถึงทรัพยากรของเครื่องต่างๆ ทำได้อย่างชัดเจนโดย:

- **Access to resources of various machines** is done explicitly by:

การเข้าสู่ระบบระยะไกลในเครื่องระยะไกลที่เหมาะสม (ssh)

- Remote logging into the appropriate remote machine (ssh)

▶ `ssh kristen.cs.yale.edu`

การถ่ายโอนข้อมูลจากเครื่องระยะไกลไปยังเครื่องท้องถิ่นผ่านกลไก File Transfer Protocol (FTP)

- Transferring data from remote machines to local machines, via the File Transfer Protocol (FTP) mechanism

อัปโหลด ดาวน์โหลด เข้าถึง หรือแชร์ไฟล์ผ่านที่เก็บข้อมูลบนคลาวด์

- Upload, download, access, or share files through cloud storage

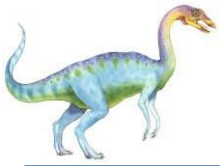
ผู้ใช้งานต้องเปลี่ยนกระบวนทัศน์ – สร้างเซสชัน, ออกคำสั่งตามเครือข่าย, ใช้เว็บเบราว์เซอร์

- **Users must change paradigms – establish a session, give network-based commands, use a web browser**

ยากขึ้นสำหรับผู้ใช้งาน

- **More difficult for users**





Distributed Operating Systems

ผู้ใช้ไม่ทราบถึงความหลากหลายของเครื่องจักร

- **Users not aware** of multiplicity of machines

การเข้าถึงทรัพยากรระยะไกลคล้ายกับการเข้าถึงทรัพยากรในที่อื่น

- **Access to remote resources similar to access to local resources**

- **Data Migration** – transfer data by transferring entire file, or transferring only those portions of the file necessary for the immediate task

การย้ายการคำนวณ – ถ่ายโอนการคำนวณข้ามระบบ แทนที่จะเป็นข้อมูล

- **Computation Migration** – transfer the computation, rather than the data, across the system

ผ่านการเรียกขั้นตอนระยะไกล (RPC)

- Via remote procedure calls (RPCs)

ผ่านทางระบบส่งข้อความ

- Via messaging system

migration : การเคลื่อนย้ายจากที่หนึ่งไปยังอีกที่หนึ่ง





Distributed-Operating Systems (Cont.)

การย้ายกระบวนการ — ดำเนินการกระบวนการทั้งหมดหรือบางส่วนที่ไซต์ต่างๆ

- **Process Migration** — execute an entire process, or parts of it, at different sites

โหลดบาลานซ์ — กระจายกระบวนการต่างๆ ทั่วทั้งเครือข่ายแม่กระทั่งเวิร์กโหลด

- **Load balancing** — distribute processes across network to even the workload

เร่งความเร็วในการคำนวณ — กระบวนการย่อยสามารถทำงานพร้อมกันบนไซต์ต่างๆ

- **Computation speedup** — subprocesses can run concurrently on different sites

การตั้งค่าฮาร์ดแวร์ — การดำเนินการตามกระบวนการอาจต้องใช้โปรเซสเซอร์พิเศษ

- **Hardware preference** — process execution may require specialized processor

การตั้งค่าซอฟต์แวร์ — ซอฟต์แวร์ที่จำเป็นอาจมีให้บริการเฉพาะบางไซต์เท่านั้น

- **Software preference** — required software may be available at only a particular site

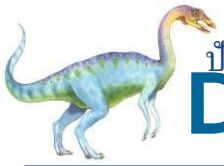
การเข้าถึงข้อมูล — เรียกใช้กระบวนการจากระยะไกล แทนที่จะถ่ายโอนข้อมูลทั้งหมดภายในเครื่อง

- **Data access** — run process remotely, rather than transfer all data locally

พิจารณาเว็บไซต์เว็บ

- Consider the World Wide Web

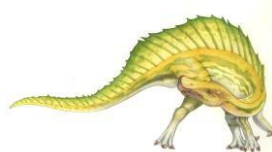


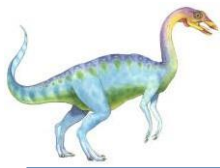


Design Issues of Distributed Systems

เราตรวจสอบคำถามการออกแบบสามข้อ:

- We investigate three design questions:
 - ความทนทาน – ระบบแบบกระจายสามารถทนต่อความล้มเหลวได้หรือไม่ = strong enough
 - **Robustness** – Can the distributed system withstand failures?
 - **Transparency** – Can the distributed system be transparent to the user both in terms of where files are stored and user mobility? ความโปร่งใส – ระบบแบบกระจายสามารถโปร่งใสต่อผู้ใช้ทั้งในแง่ของตำแหน่งที่เก็บไฟล์และความคล่องตัวของผู้ใช้ได้หรือไม่?
 - **Scalability** – Can the distributed system be scalable to allow addition of more computation power, storage, or users? ความสามารถในการปรับขนาด – ระบบแบบกระจายสามารถปรับขนาดเพื่อเพิ่มพลังการคำนวณ พื้นที่เก็บข้อมูล หรือผู้ใช้ได้มากขึ้นหรือไม่





Robustness

ความทนทาน

ความล้มเหลวของฮาร์ดแวร์อาจรวมถึงความล้มเหลวของลิงก์ ความล้มเหลวของไซต์ และการสูญเสียข้อความ

- Hardware failures can include failure of a link, failure of a site, and loss of a message.

ระบบที่ทนทานต่อข้อผิดพลาดสามารถทนต่อความล้มเหลวได้ในระดับหนึ่ง

- A fault-tolerant system can tolerate a certain level of failure

ระดับความทนทานต่อข้อผิดพลาดขึ้นอยู่กับ การออกแบบระบบและข้อบกพร่องเฉพาะ

- Degree of fault tolerance depends on design of system and the specific fault

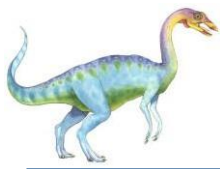
ยิ่งทนต่อข้อผิดพลาดได้มากเท่าไรก็ยิ่งดีเท่านั้น!

- The more fault tolerance, the better!

เกี่ยวข้องกับการตรวจหาความล้มเหลว การกำหนดค่าใหม่ และการกู้คืน

- Involves failure detection, reconfiguration, and recovery





Failure Detection

การตรวจจับความล้มเหลว

การตรวจจับความล้มเหลวของฮาร์ดแวร์เป็นเรื่องยาก

- ❑ Detecting hardware failure is difficult

ในการตรวจจับความล้มเหลวของลิงก์ สามารถใช้โปรโตคอลฮาร์ดบีทได้

- ❑ To detect a link failure, a heartbeat protocol can be used

สมมติว่าไซต์ A และ ไซต์ B ได้สร้างลิงก์แล้ว

- ❑ Assume Site A and Site B have established a link

ในช่วงเวลาที่กำหนด แต่ละ ไซต์จะแลกเปลี่ยนข้อความ I-am-up เพื่อระบุว่าไซต์เหล่านั้นพร้อมใช้งานแล้ว

- ❑ At fixed intervals, each site will exchange an I-am-up message indicating that they are up and running

หาก ไซต์ A ไม่ได้รับข้อความภายในช่วงเวลาที่กำหนด จะถือว่า (a) ไซต์อื่นไม่พร้อมใช้งาน หรือ (b) ข้อความสูญหาย

- ❑ If Site A does not receive a message within the fixed interval, it assumes either (a) the other site is not up or (b) the message was lost.

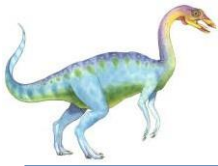
ขณะนี้ ไซต์ A สามารถส่งข้อความ Are-you-up? ข้อความถึง ไซต์ B

- ❑ Site A can now send an Are-you-up? message to Site B

หาก ไซต์ A ไม่ได้รับการตอบกลับ ก็สามารถส่งข้อความซ้ำหรือลองใช้เส้นทางอื่นไปยัง ไซต์ B

- ❑ If Site A does not receive a reply, it can repeat the message or try an alternate route to Site B





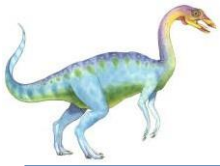
Failure Detection (Cont.)

หากไซต์ A ไม่ได้รับการตอบกลับจากไซต์ B ในท้ายที่สุด จะสรุปได้ว่ามีความล้มเหลวบางประเภทเกิดขึ้น

= finally

- If Site A does not ultimately receive a reply from Site B, it concludes some type of failure has occurred
- ประเภทของความล้มเหลว:
 - Site B is down ไซต์ B ล่ม
 - The direct link between A and B is down การเชื่อมโยงโดยตรงระหว่าง A และ B ไม่ทำงาน
 - The alternate link from A to B is down ลิงก์สำรองจาก A ไป B ใช้งานไม่ได้
 - The message has been lost ข้อความหายไป
- However, Site A cannot determine exactly **why** the failure has occurred อย่างไรก็ตาม ไซต์ A ไม่สามารถระบุได้อย่างแน่ชัดว่าเหตุใดจึงเกิดความล้มเหลว





Reconfiguration and Recovery

เมื่อไซต์ A พิจารณาว่ามีความล้มเหลวเกิดขึ้น จะต้องกำหนดค่าระบบใหม่:

- When Site A determines a failure has occurred, it must reconfigure the system:

หากการเชื่อมโยงจาก A ไปยัง B ล้มเหลว จะต้องเผยแพร่ไปยังทุกไซต์ในระบบ

- If the link from A to B has failed, this must be broadcast to every site in the system

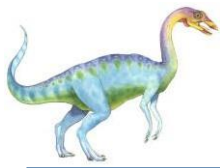
หากไซต์ล้มเหลว ไซต์อื่นๆ ทุกแห่งจะต้องได้รับแจ้งด้วยว่าบริการที่นำเสนอโดยไซต์ที่ล้มเหลวนั้นไม่มีให้บริการอีกต่อไป

- If a site has failed, every other site must also be notified indicating that the services offered by the failed site are no longer available

เมื่อลิงก์หรือไซต์พร้อมใช้งานอีกครั้ง ข้อมูลนี้จะต้องเผยแพร่ไปยังไซต์อื่นทั้งหมดอีกครั้ง

- When the link or the site becomes available again, this information must again be broadcast to all other sites





Transparency

ความโปร่งใส

ระบบแบบกระจายควรปรากฏเป็นระบบรวมศูนย์ทั่วไปสำหรับผู้ใช้งาน

- The distributed system should appear as a conventional, centralized system to the user

ส่วนติดต่อผู้ใช้ไม่ควรแยกความแตกต่างระหว่างทรัพยากรภายในและระยะไกล

- User interface should not distinguish between local and remote resources

ตัวอย่าง: NFS

- ▶ Example: NFS

ความคล่องตัวของผู้ใช้ทำให้ผู้ใช้สามารถเข้าสู่ระบบเครื่องใดก็ได้ในสภาพแวดล้อมและดูสภาพแวดล้อมของตนได้

- User mobility allows users to log into any machine in the environment and see his/her environment

ตัวอย่าง: LDAP พร้อมการจำลองเสมือนเดสก์ท็อป

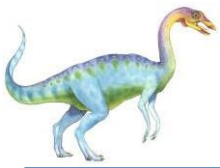
- ▶ Example: LDAP plus desktop virtualization

หมายถึง 11.5.1 5.9.7

NFS: Network File System (เป็นกลไกในการเก็บไฟล์ข้อมูลในระบบเครือข่ายเน็ตเวิร์ก)

LDAP: Lightweight Directory Access Protocol (ข้อตกลงที่ใช้ในการเข้าถึง ไดเรกทอรีในระบบเครือข่ายเน็ตเวิร์ก)





Scalability

ความสามารถในการขยายขนาด



เมื่อความต้องการเพิ่มขึ้น ระบบควรยอมรับการเพิ่มทรัพยากรใหม่ได้อย่างง่ายดายเพื่อรองรับความต้องการที่เพิ่มขึ้น

- As demands increase, the system should easily accept the addition of new resources to accommodate the increased demand

ตอบสนองอย่างสง่างามต่อภาระที่เพิ่มขึ้น

- Reacts gracefully to increased load

- Adding more resources may generate additional indirect load on other resources if not careful

- Data **compression** or **deduplication** can cut down on storage and network resources used

Data deduplication

the number of other resources
not increase but there may be
additional load

Many cloud storage providers use compression & deduplication to cut down on the amount of storage used

การบีบอัดข้อมูลหรือการจัดข้อมูลซ้ำซ้อนสามารถลดพื้นที่จัดเก็บข้อมูลและทรัพยากรเครือข่ายที่ใช้ได้



End of Chapter 9

