

Unit-5

Naming

By Prashant Gautam

BCA 6th Semester

Outlines

- Name, Identifiers and Address
- Structured Naming
- Attribute-Based Naming
- Case Study: The Global Name service

Names, Identifiers, and Addresses

- A name in a distributed system is a string of bits or characters that is used to refer to an entity.
- An entity in a distributed system can be practically anything.
 - Typical examples include resources such as hosts, printers, disks, and files.
 - Other well-known examples of entities that are often explicitly named are processes, users, mailboxes, newsgroups, Web pages, graphical windows, messages, network connections, and so on.

For example, a resource such as a printer offers an interface containing operations for printing a document, requesting the status of a print job, and the like. Furthermore, an entity such as a network connection may provide operations for sending and receiving data, setting quality-of-service parameters, requesting the status, and so forth.

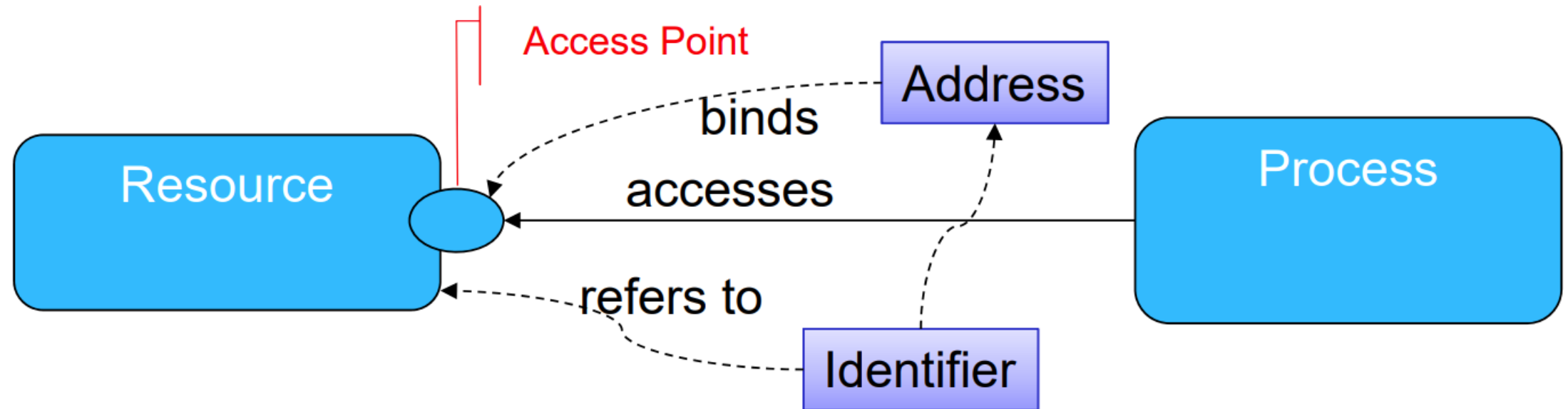
- To operate on an entity, it is necessary to access it, for which we need an access point.
- An access point is yet another, but special, kind of entity in a distributed system.
- The name of an access point is called an address.
- The address of an access point of an entity is also simply called an address of that entity.
- An entity can offer more than one access point.

- As a comparison, a telephone can be viewed as an access point of a person, whereas the telephone number corresponds to an address. Indeed, many people nowadays have several telephone numbers, each number corresponding to a point where they can be reached.

In a distributed system, a typical example of an access point is a host running a specific server, with its address formed by the combination of, for example, an IP address and port number (i.e., the server's transport-level address).

- An entity may change its access points in the course of time.
- For example: when a mobile computer moves to another location, it is often assigned a different IP address than the one it had before. Likewise, when a person moves to another city or country, it is often necessary to change telephone numbers as well.
- In a similar fashion, changing jobs or Internet Service Providers, means changing your e-mail address.

- Identifier: a name that uniquely identifies an entity
 - the identifier is unique and refers to only one entity
- Address: the name of an access point, the location of an entity



3 Classes of naming systems

1. FLAT NAMING / UNSTRUCTURED NAMING
2. STRUCTURED NAMING
3. ATTRIBUTE-BASED NAMING

FLAT NAMING / UNSTRUCTURED NAMING

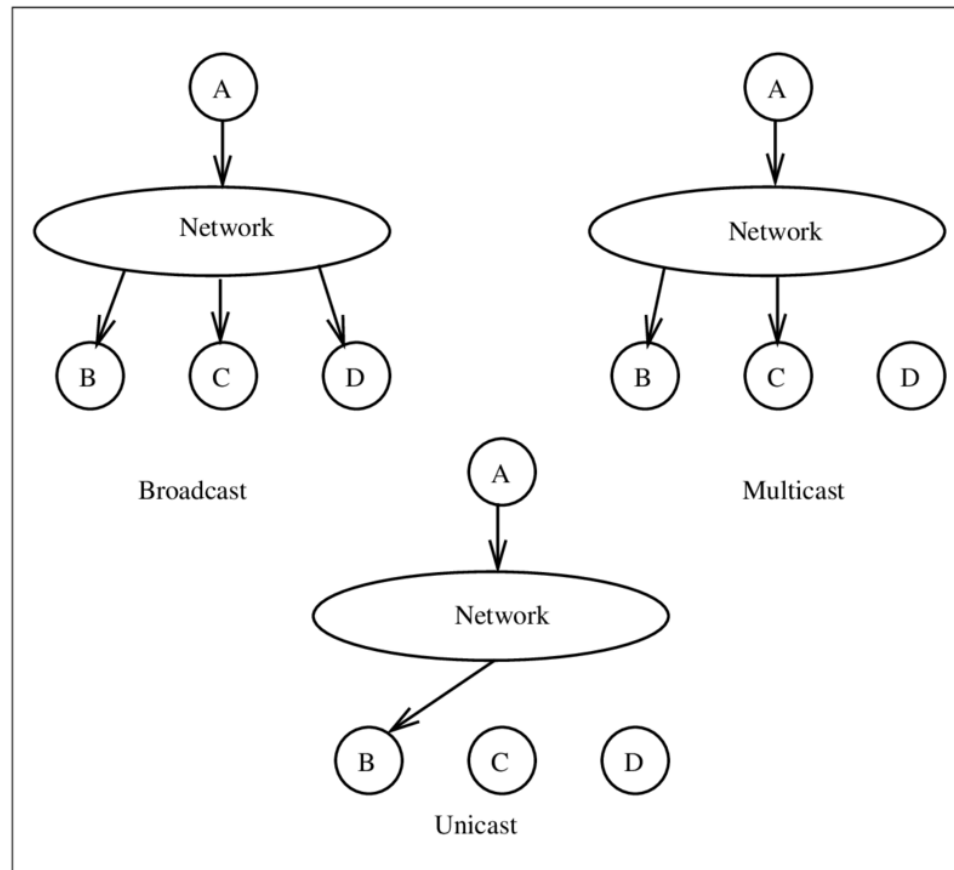
- Identifiers are convenient to uniquely represent entities.
- In many cases, identifiers are simply random bit strings. which we conveniently refer to as unstructured, or flat names.
- An important property of such a name is that it does not contain any information whatsoever on how to locate the access point of its associated entity.
- We will study these aspects:
 - How flat names can be resolved ?
 - How we can locate an entity when given only its identifier ?

FLAT NAMING / UNSTRUCTURED NAMING

- Simple Solution for locating an entity
 - Broadcasting and Multicasting
 - Forwarding Pointers
- Both solutions are applicable only to local-area networks.
- But they often do the job well, making their simplicity particularly attractive.

Broadcasting and Multicasting

BROADCAST VERSUS MULTICAST



BROADCAST

A method of transferring a message to all recipients simultaneously

Packets are transmitted to all the connected devices in the network

There is no need for group management

Less secure

More traffic

Slower

MULTICAST

A group communication where data transmission is addressed to a group of destination computers simultaneously

Packets are transmitted to some of the devices in the network

Requires group management

More secure

Less traffic

Faster

Visit www.PEDIAA.com

- Consider a distributed system built on a computer network: that offers efficient broadcasting facilities.
- Typically, such facilities are offered by local-area networks and LAN wireless in which all machines are connected to a single cable or equivalent.
- Locating an entity in such an environment is simple: a message containing the identifier of the entity is broadcast to each machine and each machine is requested to check whether it has that entity.
- Only the machines that can offer an access point for the entity send a reply message containing the address of that access point.

Example:

- a machine broadcasts a packet on the local network asking who is the owner of a given IP address.
- When the message arrives at a machine, the receiver checks whether it should listen to the requested IP address.
- If so, it sends a reply packet containing, for example, its Ethernet address.

Problems with Broadcasting:

- Broadcasting becomes inefficient when the network grows.
- Not only is network bandwidth wasted by request messages, but, more seriously, too many hosts maybe interrupted by requests they cannot answer.
- One possible solution is to switch to multicasting, by which only a restricted group of hosts receives the request.
- For example, Ethernet networks support data-link level multicasting directly in hardware.

- Multicasting can also be used to locate entities in point-to-point networks.
- For example, the Internet supports network-level multicasting by allowing hosts to join a specific multicast group.
- Such groups are identified by a multicast address.
- When a host sends a message to a multicast address, the network layer provides a best-effort service to deliver that message to all group members.

- A multicast address can be used as a general location service for multiple entities.
- For example, consider an organization where each employee has his or her own mobile computer.
- When such a computer connects to the locally available network. it is dynamically assigned an IP address.
- In addition, it joins a specific multicast group. When a process wants to locate computer A, it sends a "where is A?" request to the multicast group.
- If A is connected, it responds with its current IP address.

Assignment-VI (Deadline: Feb 28)

- Explain the merits and demerits of broadcasting and multiple casting in locating entities.

Forwarding Pointers

- The principle is simple: when an entity moves from A to B, it leaves behind in A a reference to its new location at B.
- Advantages:
 - simplicity: as soon as an entity has been located, for example by using a traditional naming service, a client can look up the current address by following the chain of forwarding pointer
- Drawbacks
 - a chain for a highly mobile entity can become so long that locating that entity is prohibitively expensive.
 - the vulnerability to broken links

Explore more on Forward pointers

STRUCTURED NAMING

- Flat names:
 - good for machines, but are generally not very convenient for humans to use.
- structured names:
 - composed from simple, human-readable names.

In this section, we concentrate on structured names and the way that these names are resolved to addresses.

Name Spaces

- Names are commonly organized into what is called a name space.
- Name spaces for structured names can be represented as a labeled, directed graph with two types of nodes.

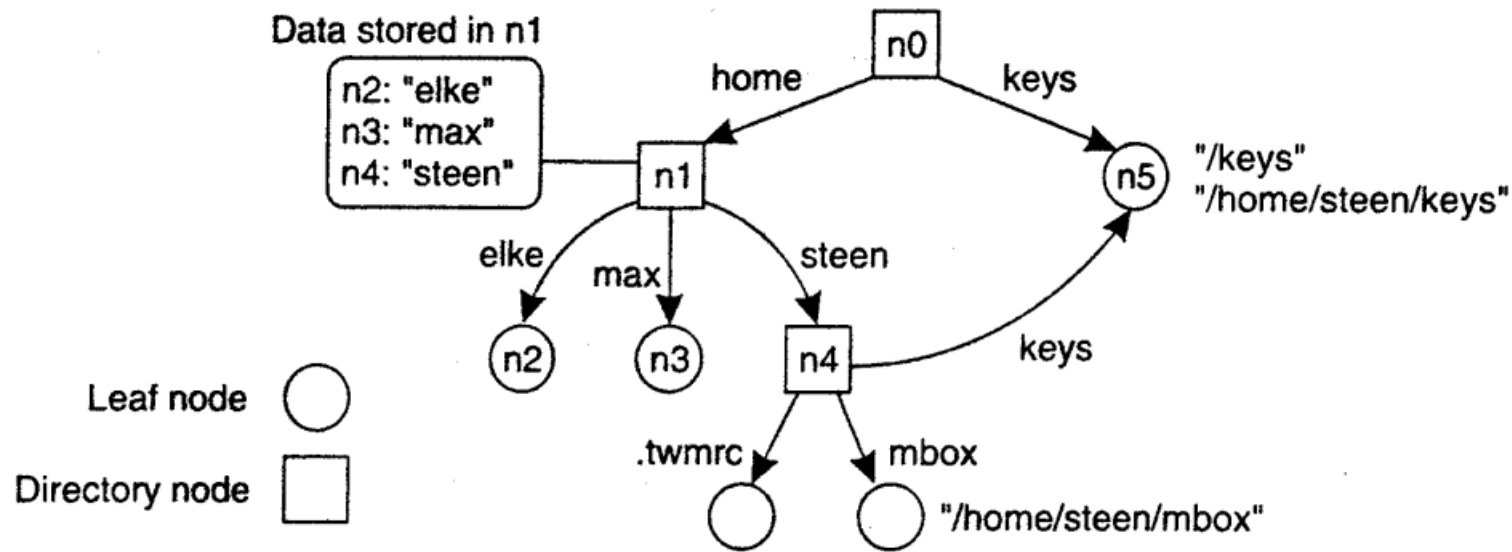


Figure 5-9. A general naming graph with a single root node.

Name Spaces

- **A leaf node** represents a named entity and has the property that it has no outgoing edges.
- **A leaf node** generally stores information on the entity it is representing—for example, its address—so that a client can access it.
- Alternatively, it can store the state of that entity, such as in the case of file systems 'in which a leaf node actually contains the complete file it is representing.
- **A directory node** has a number of outgoing edges, each labeled with a name
- **A directory node** stores a table in which an outgoing edge is represented as a pair (edge label, node identifier). Such a table is called a directory table.

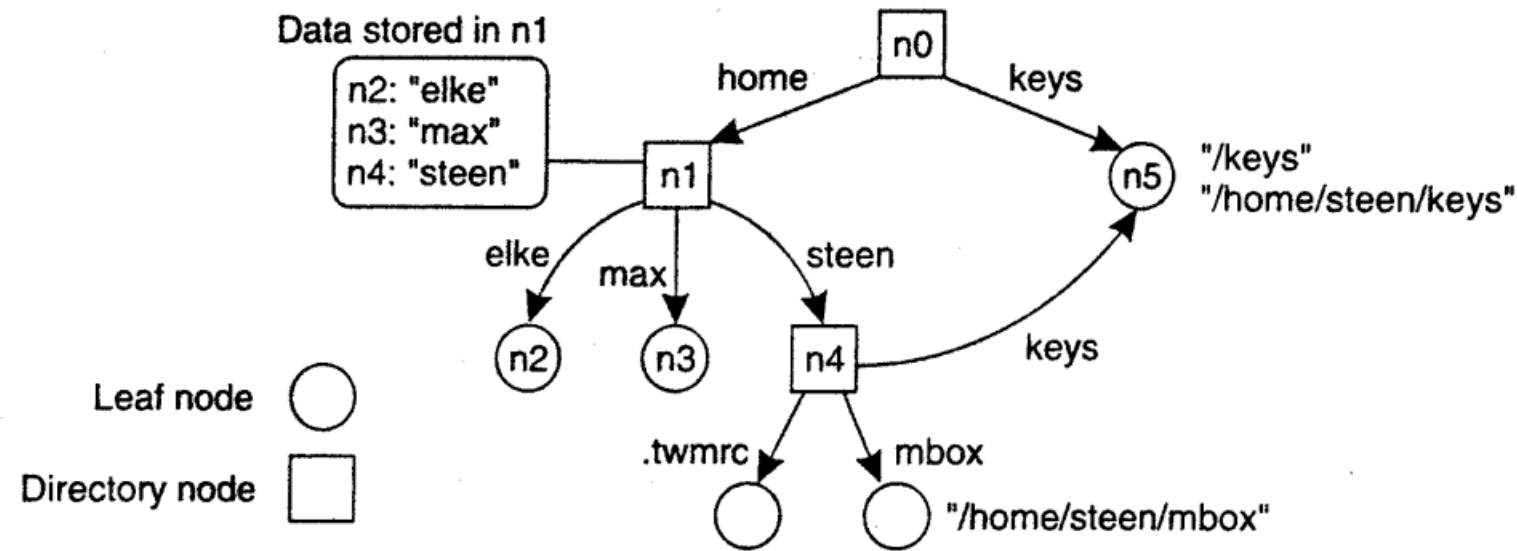


Figure 5-9. A general naming graph with a single root node.

Name Resolution

- Name spaces offer a convenient mechanism for storing and retrieving information about entities by means of names.
- More generally, given a path name, it should be possible to look up any information stored in the node referred to by that name.
- The process of looking up a name is called **name resolution**.

DNS

- We use to remember “human-readable” machine name
 - we have the name hierarchy
 - E.g., www.facebook.com
- But machines in Internet use IP address
 - E.g., 31.13.84.33
 - Application communication use IP addresses and ports
- DNS
 - Mapping from the domain name hierarchy to IP addresses

Information in records of DNS namespace

| Type of record | Associated entity | Description |
|----------------|-------------------|---|
| SOA | Zone | Holds information on the represented zone |
| A | Host | Contains an IP address of the host this node represents |
| MX | Domain | Refers to a mail server to handle mail addressed to this node |
| SRV | Domain | Refers to a server handling a specific service |
| NS | Zone | Refers to a name server that implements the represented zone |
| CNAME | Node | Symbolic link with the primary name of the represented node |
| PTR | Host | Contains the canonical name of a host |
| HINFO | Host | Holds information on the host this node represents |
| TXT | Any kind | Contains any entity-specific information considered useful |

Source: Andrew S. Tanenbaum and Maarten van Steen, Distributed Systems – Principles and Paradigms, 2nd Edition, 2007, Prentice-Hall

ATTRIBUTE-BASED NAMING

- A tuple (attribute, value) can be used to describe a property
 - E.g., („country“, “Austria“), („language“, „German“)
- A set of tuples (attribute, value) can be used to describe an entity

AustriaInfo

| Attribute | Value |
|-------------|---------|
| CountryName | Austria |
| Language | German |
| MemberofEU | Yes |
| Capital | Vienna |

- Employ (attribute, value) tuples for describing entities
 - Why flat and structured naming are not enough?
- Also called directory services.
- Name Resolution
 - Usually based on querying mechanism
 - Querying usually deal with the whole space
- Implementation
 - LDAP (Lightweight Directory Access Protocol)
 - RDF (Resource Description Framework): Framework for semantics

Assignment-VII (Deadline: Mar 3)

- Explore more about LDAP, how it works, considering real world example.
- Explore more about RDF, how it works, considering real world example.