A Quick Introduction to the Domain Name System

David Conrad

<david.conrad@nominum.com>

Chief Technology Officer

nominum

COMMUNICATION BY NAME™

Overview

- Introduction to the DNS
- DNS Components
- DNS Structure and Hierarchy
- The DNS in Context

The DNS is…

- The "Domain Name System"
  - Created in 1983 by Paul Mockapetris (RFCs 1034 and 1035), modified, updated, and enhanced by a myriad of subsequent RFCs
- What Internet users use to reference anything by name on the Internet
- The mechanism by which Internet software translates names to addresses and vice versa

nominum
COMMUNICATION BY NAME™

A Quick Digression: Names versus Addresses

- An address is how you get to an endpoint
  - Typically, hierarchical (for scaling):
    - 950 Charter Street, Redwood City CA, 94063
    - 204.152.187.11, +1-650-381-6003
- A “name” is how an endpoint is referenced
  - Typically, no structurally significant hierarchy
    - “David”, “Tokyo”, “itu.int”

The DNS is also…

- A lookup mechanism for translating objects into other objects
- A globally distributed, loosely coherent, scalable, reliable, dynamic database
- Comprised of three components
  - A "name space"
  - Servers making that name space available
  - Resolvers (clients) which query the servers about the name space

DNS as a Lookup Mechanism

- Users generally prefer names to numbers
- Computers prefer numbers to names
- DNS provides the mapping between the two
  – I have "x", give me "y"
- DNS is **NOT** a directory service
  – No way to search the database
    - No easy way to add this functionality

DNS as a Database

- Keys to the database are “domain names”
  - www.foo.com, 18.in-addr.arpa, 6.4.e164.arpa
- Over 100,000,000 domain names stored
- Each domain name contains one or more attributes
  - Known as “resource records”
- Each attribute individually retrievable

Global Distribution

- Data is maintained locally, but retrievable globally
  - No single computer has all DNS data
- DNS lookups can be performed by any device
- Remote DNS data is locally cachable to improve performance

Loose Coherency

- The database is always internally consistent
  - Each version of a subset of the database (a zone) has a serial number
    - The serial number is incremented on each database change
- Changes to the master copy of the database are replicated according to timing set by the zone administrator
- Cached data expires according to timeout set by zone administrator

nominum
COMMUNICATION BY NAME™

Scalability

- No limit to the size of the database
  - One server has over 20,000,000 names
    - Not a particularly good idea
- No limit to the number of queries
  - 24,000 queries per second handled easily
- Queries distributed among masters, slaves, and caches

nominum
COMMUNICATION BY NAME™

Reliability

- Data is replicated
  - Data from master is copied to multiple slaves
- Clients can query
  - Master server
  - Any of the copies at slave servers
- Clients will typically query local caches
- DNS protocols can use either UDP or TCP
  - If UDP, DNS protocol handles retransmission, sequencing, etc.

Dynamicity

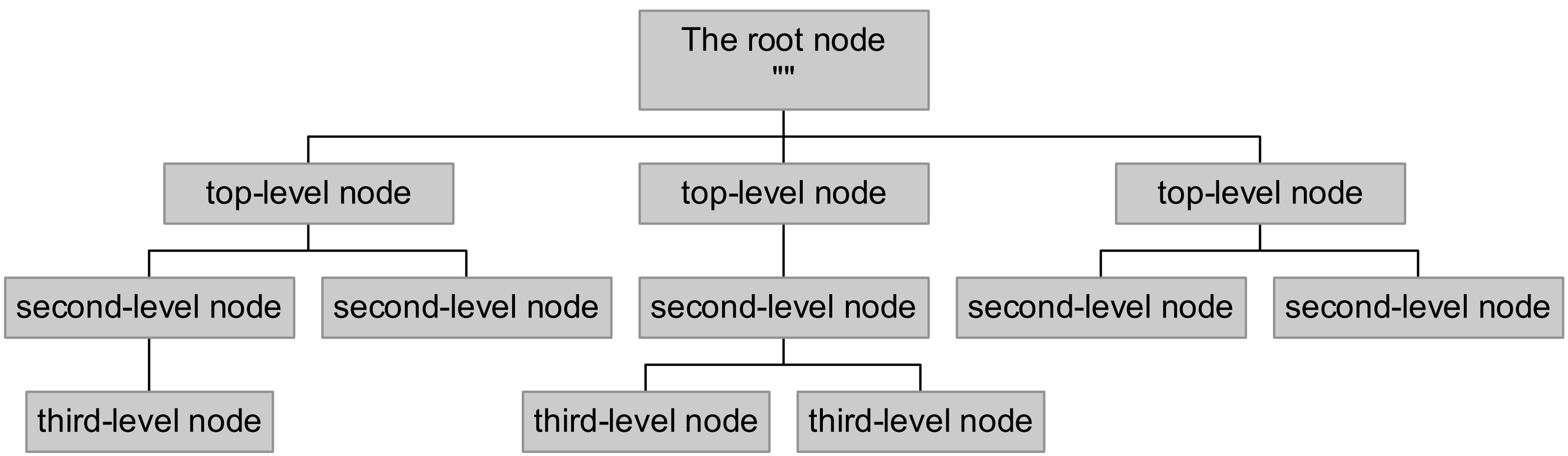
- Database can be updated dynamically
  - Add/delete/modify of any record
- Modification of the master database triggers replication
  - Only master can be dynamically updated
    - Creates a single point of failure

Overview

- Introduction to the DNS
- DNS Components
  - The name space
  - The servers
  - The resolvers
- DNS Structure and Hierarchy
- The DNS in Context

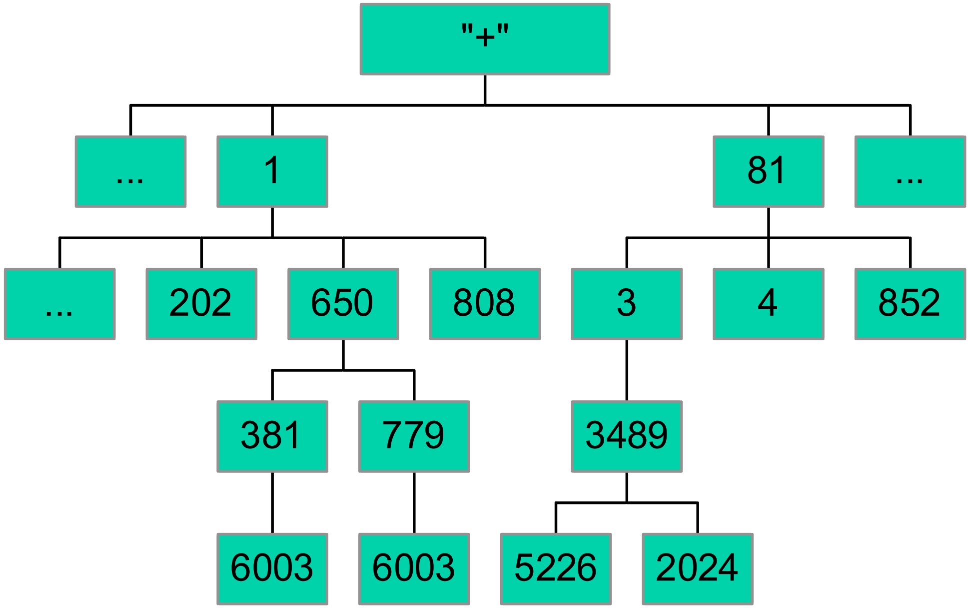
The Name Space

- The *name space* is the structure of the DNS database
  - An inverted tree with the root node at the top
- Each node has a label
  - The root node has a null label, written as “”



An Analogy – E.164

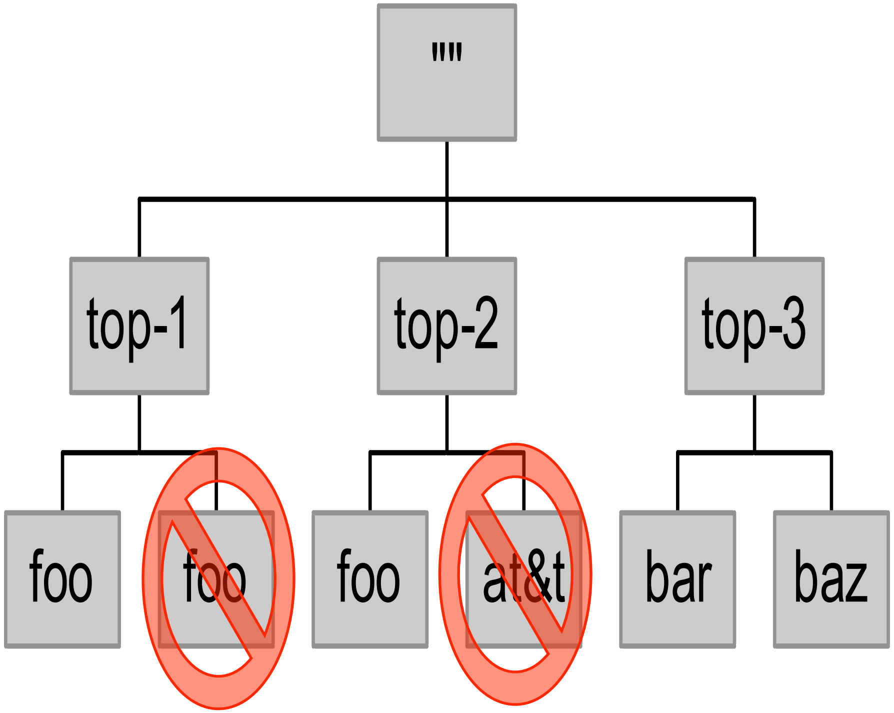
- Root node maintained by the ITU (call it “+”)
- Top level nodes = country codes (1, 81, etc)
- Second level nodes = regional codes (1-808, 81-3, etc.)



nominum
COMMUNICATION BY NAME™

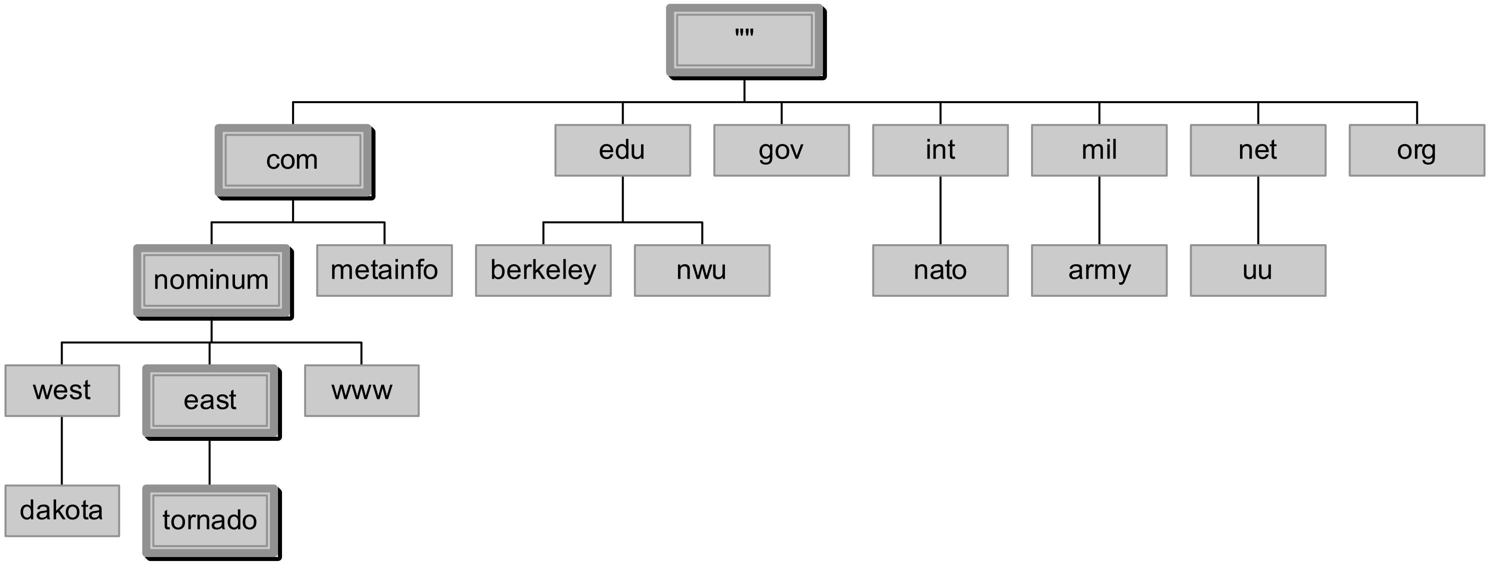
Labels

- Each node in the tree must have a label
  - A string of up to 63 8 bit bytes
- The DNS protocol makes NO limitation on what binary values are used in labels
  - RFCs 852 and 1123 define legal characters for "hostnames"
    - A-Z, 0-9, and "-" only with a-z and A-Z treated as the same
- Sibling nodes must have unique labels
- The null label is reserved for the root node



Domain Names

- A *domain name* is the sequence of labels from a node to the root, separated by dots (“.”s), read left to right
  - The name space has a maximum depth of 127 levels
  - Domain names are limited to 255 characters in length
- A node’s domain name identifies its position in the name space



nominum COMMUNICATION BY NAME™

Subdomains

- One domain is a *subdomain* of another if its apex node is a descendant of the other's apex node
- More simply, one domain is a subdomain of another if its domain name ends in the other's domain name
  - So *sales.nominum.com* is a subdomain of
    - *nominum.com*
    - *com*
  - *nominum.com* is a subdomain of *com*

nominum
COMMUNICATION BY NAME™

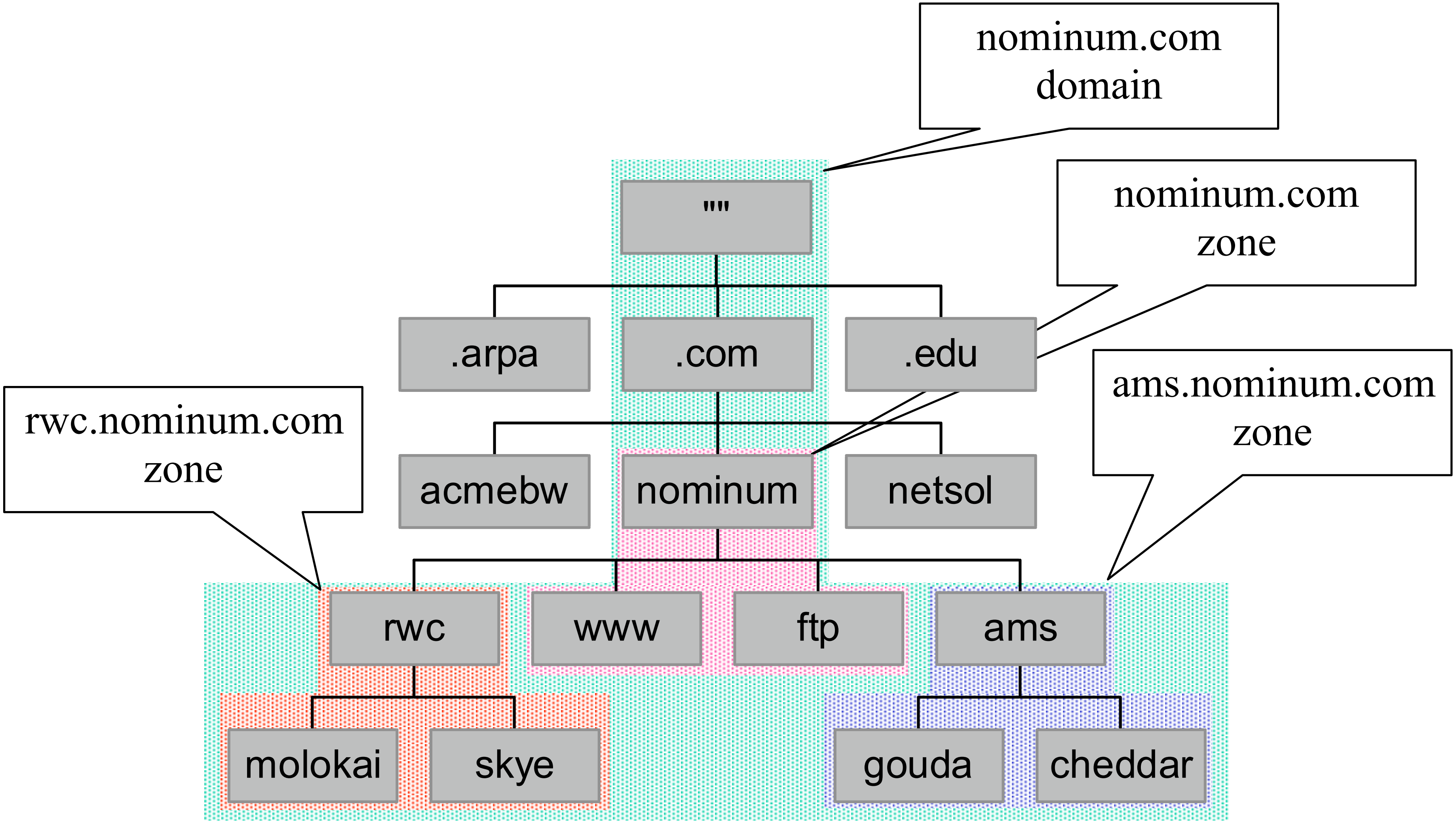
Delegation

- Administrators can create subdomains to group hosts
  - According to geography, organizational affiliation or any other criterion
- An administrator of a domain can delegate responsibility for managing a subdomain to someone else
  - But this isn't required
- The parent domain retains links to the delegated subdomain
  - The parent domain "remembers" who it delegated the subdomain to

Delegation Creates Zones

- Each time an administrator delegates a subdomain, a new unit of administration is created
  - The subdomain and its parent domain can now be administered independently
  - These units are called *zones*
  - The boundary between zones is a point of delegation in the name space
- Delegation is good: it is the key to scalability

Dividing a Domain into Zones



nominum COMMUNICATION BY NAME™

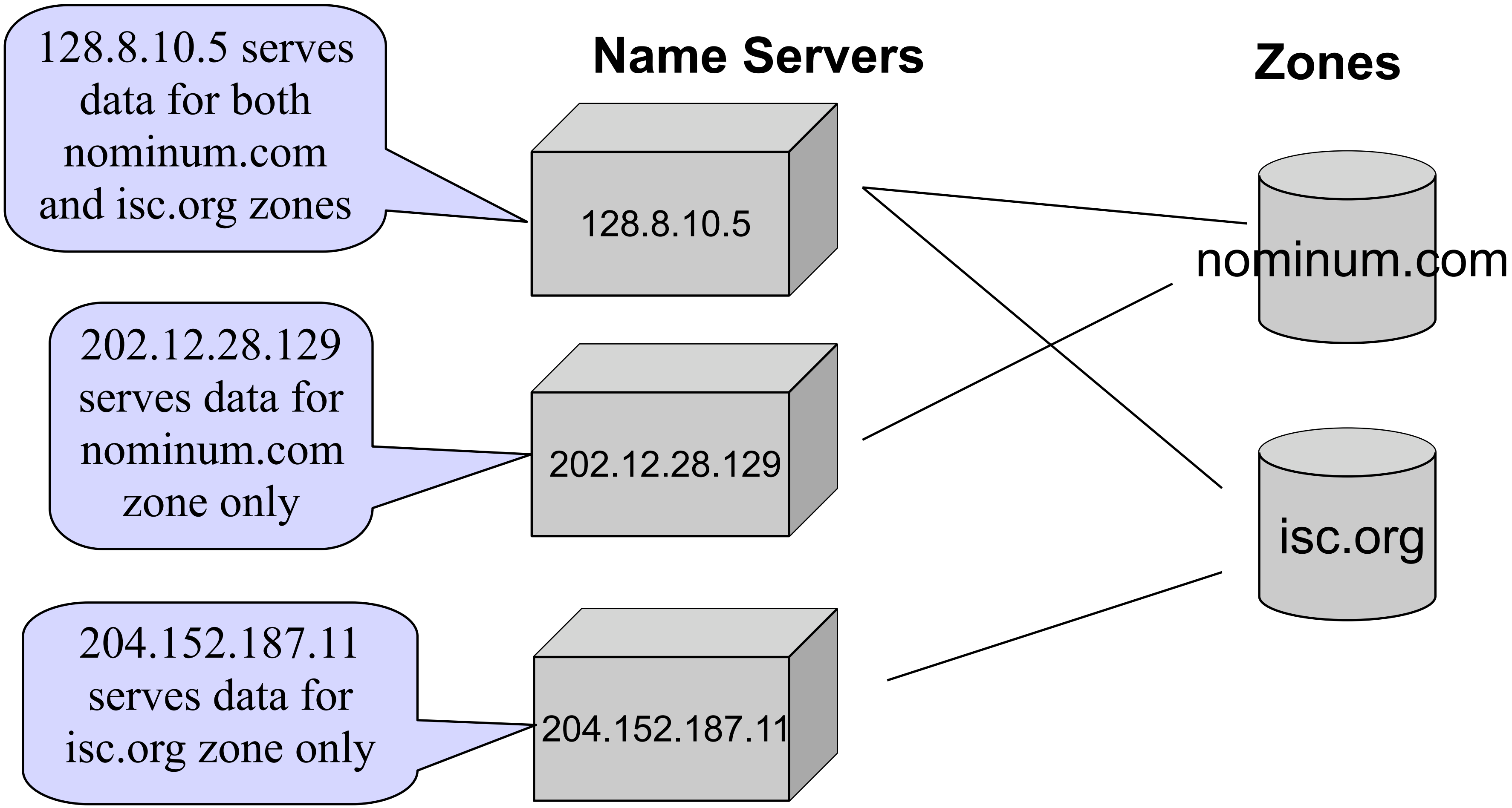
Overview

- Introduction to the DNS
- DNS Components
  - The name space
  - The servers
  - The resolvers
- DNS Structure and Hierarchy
- The DNS in Context

Name Servers

- Name servers store information about the name space in units called “zones”
  - The name servers that load a complete zone are said to “have authority for” or “be authoritative for” the zone
- Usually, more than one name server are authoritative for the same zone
  - This ensures redundancy and spreads the load
- Also, a single name server may be authoritative for many zones

Name Servers and Zones



Types of Name Servers

- Two main types of servers
  - Authoritative – maintains the data
    - Master – where the data is edited
    - Slave – where data is replicated to
  - Caching – stores data obtained from an authoritative server
  - The most common name server implementation (BIND) combines these two into a single process
- Other types exist…
- No special hardware necessary

Name Server Architecture

- You can think of a name server as part:
  - *database server,* answering queries about the parts of the name space it knows about (i.e., is authoritative for),
  - *cache,* temporarily storing data it learns from other name servers, and
  - *agent,* helping resolvers and other name servers find data that other name servers know about

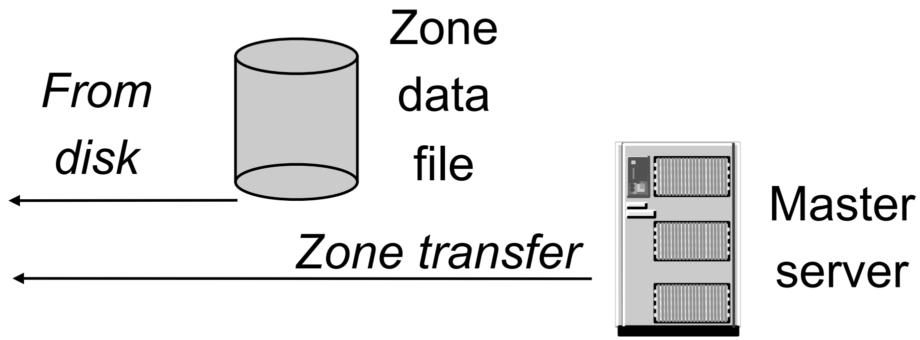
Name Server Architecture

Name Server Process

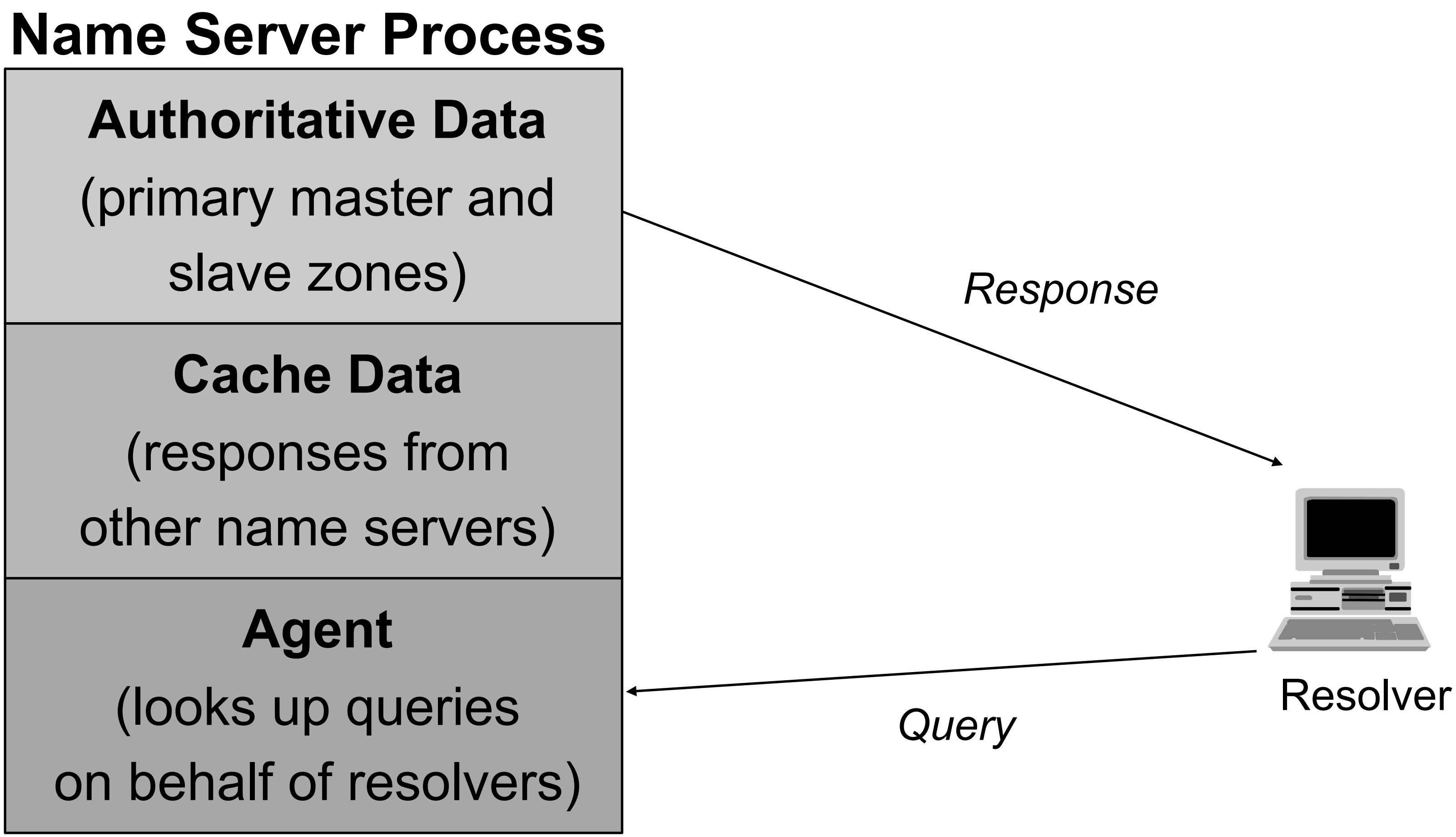
Authoritative Data
(primary master and slave zones)

Cache Data
(responses from other name servers)

Agent
(looks up queries on behalf of resolvers)

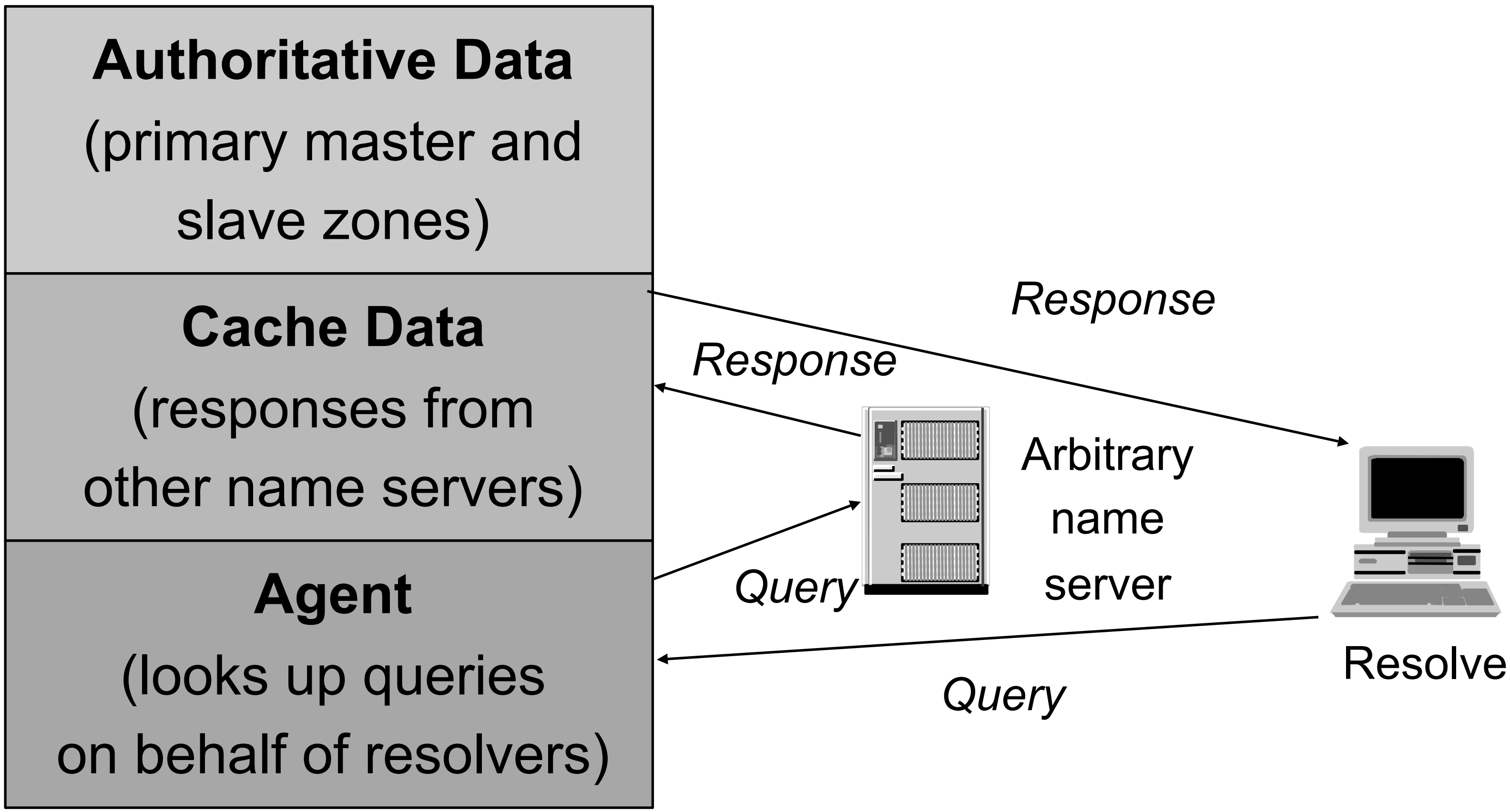


Authoritative Data

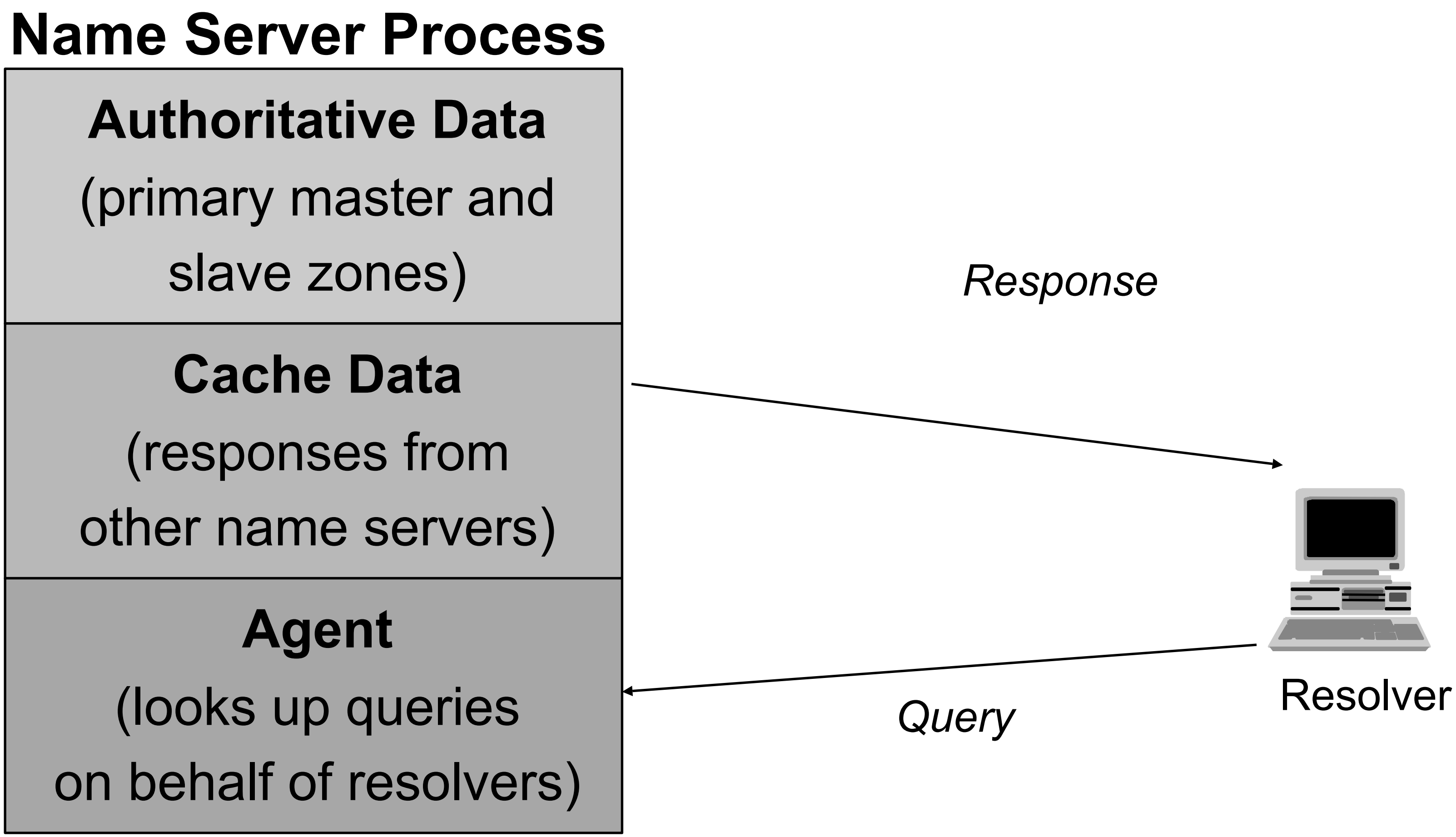


Using Other Name Servers

Name Server Process



Cached Data



nominum
COMMUNICATION BY NAME™

Overview

- Introduction to the DNS
- DNS Components
  - The name space
  - The servers
  - The resolvers
- DNS Structure and Hierarchy
- The DNS in Context

Name Resolution

- *Name resolution* is the process by which resolvers and name servers cooperate to find data in the name space
- To find information anywhere in the name space, a name server only needs the names and IP addresses of the name servers for the root zone (the "root name servers")
  - The root name servers know about the top-level zones and can tell name servers whom to contact for all TLDs

Name Resolution

- A DNS query has three parameters:
  - A domain name (e.g., *www.nominum.com*),
    - Remember, every node has a domain name!
  - A class (e.g., *IN*), and
  - A type (e.g., *A*)
- A name server receiving a query from a resolver looks for the answer in its authoritative data and its cache
  - If the answer isn't in the cache and the server isn't authoritative for the answer, the answer must be looked up

nominum
COMMUNICATION BY NAME™

The Resolution Process

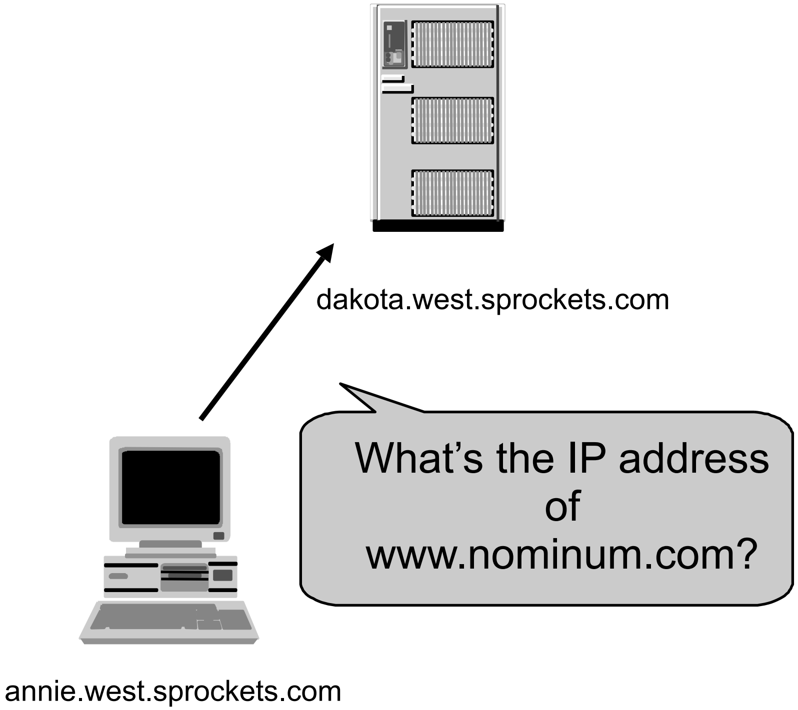
- Let's look at the resolution process step-by-step:

annie.west.sprockets.com

```
ping www.nominum.com.
```

The Resolution Process

- The workstation *annie* asks its configured name server, *dakota,* for *www.nominum.com's* address

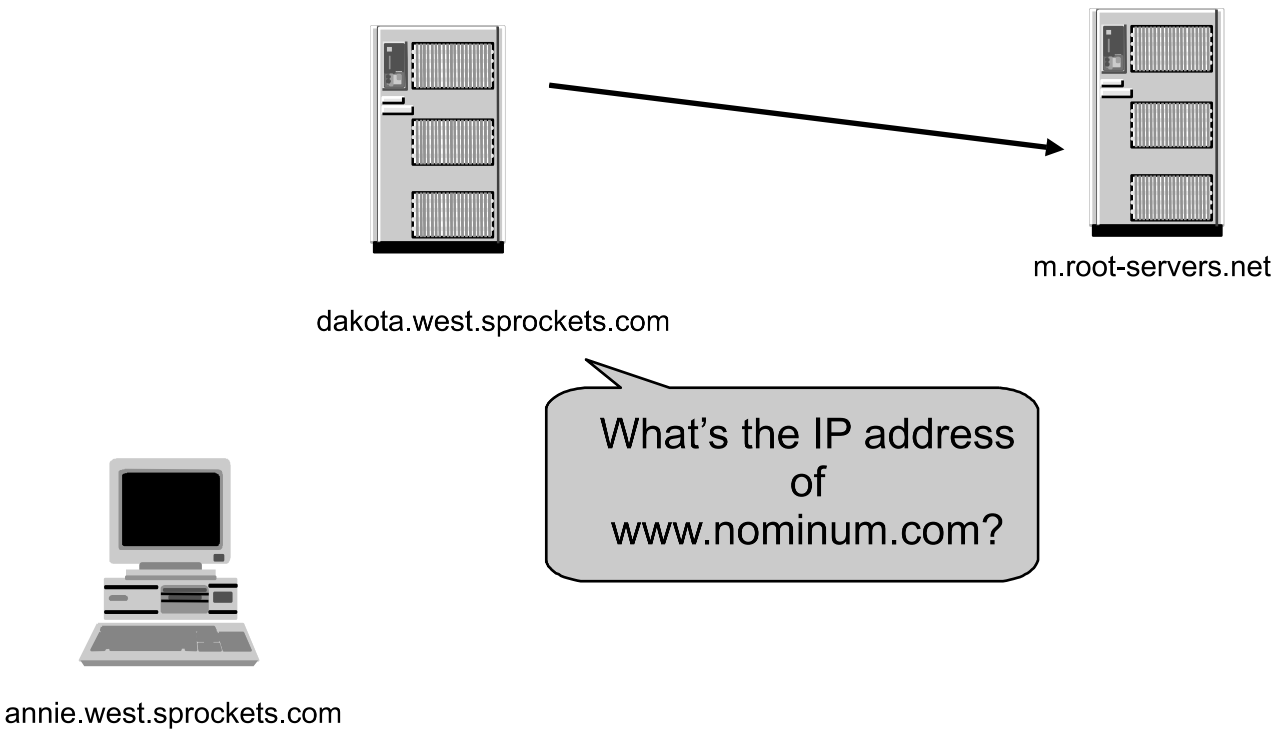


`ping www.nominum.com.`

nominum COMMUNICATION BY NAME™

The Resolution Process

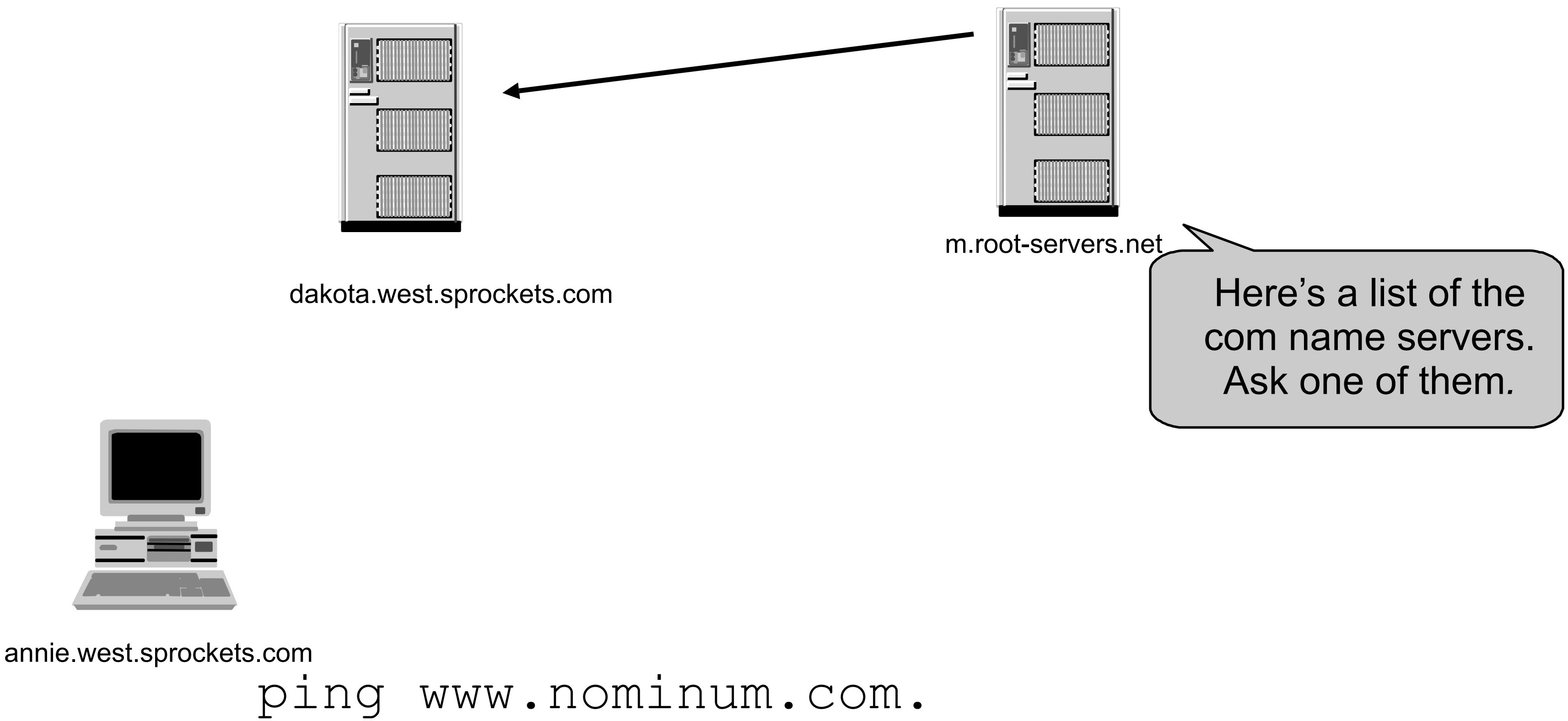
- The name server *dakota* asks a root name server, *m*, for *www.nominum.com's* address



`ping www.nominum.com.`

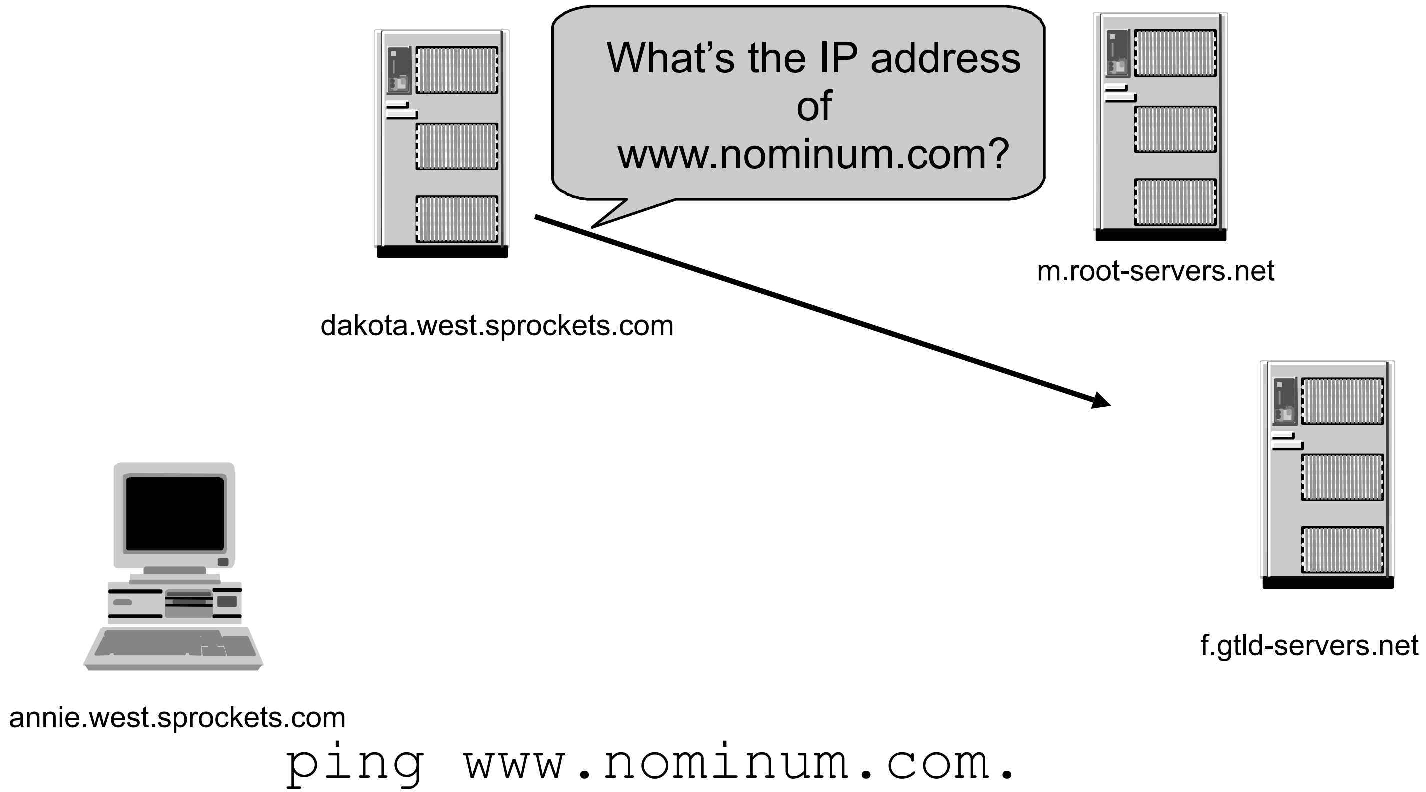
The Resolution Process

- The root server *m* refers *dakota* to the *com* name servers
- This type of response is called a “referral”



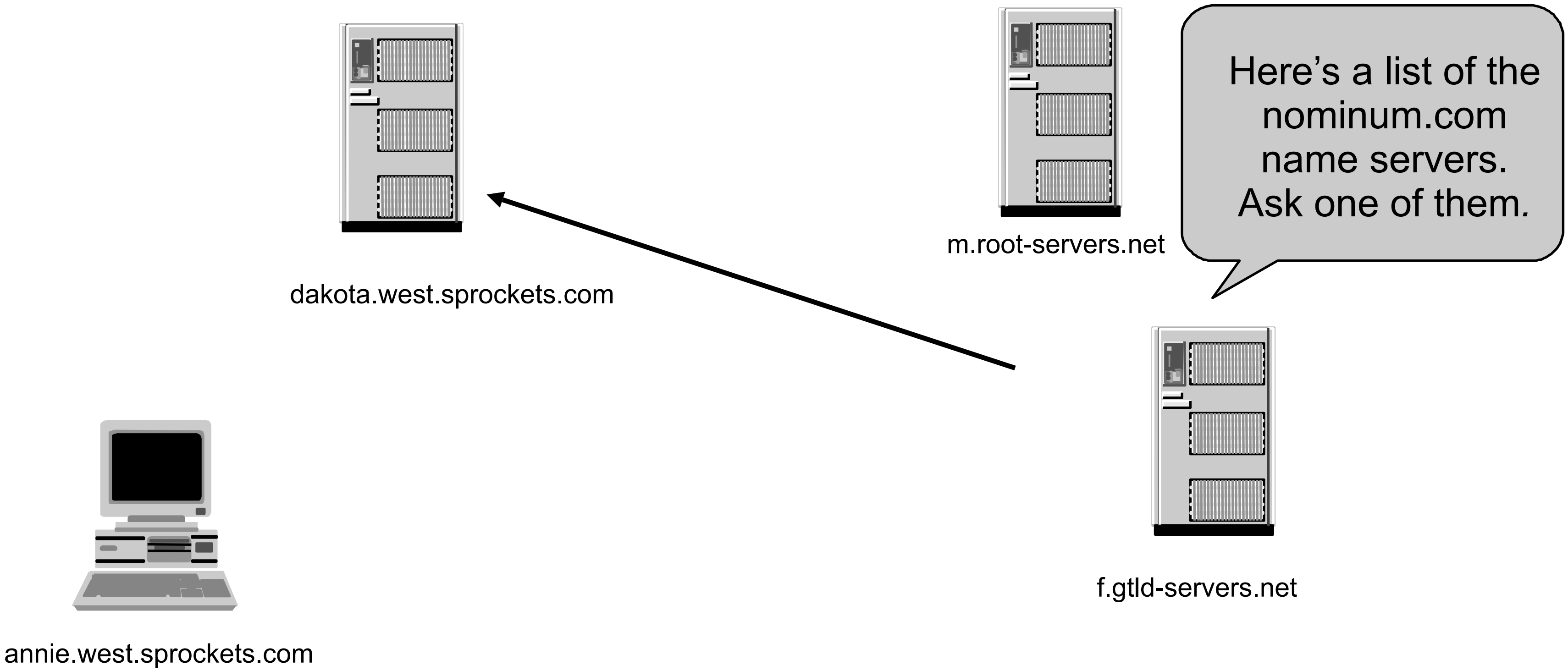
The Resolution Process

- The name server *dakota* asks a *com* name server, *f*, for *www.nominum.com's* address



The Resolution Process

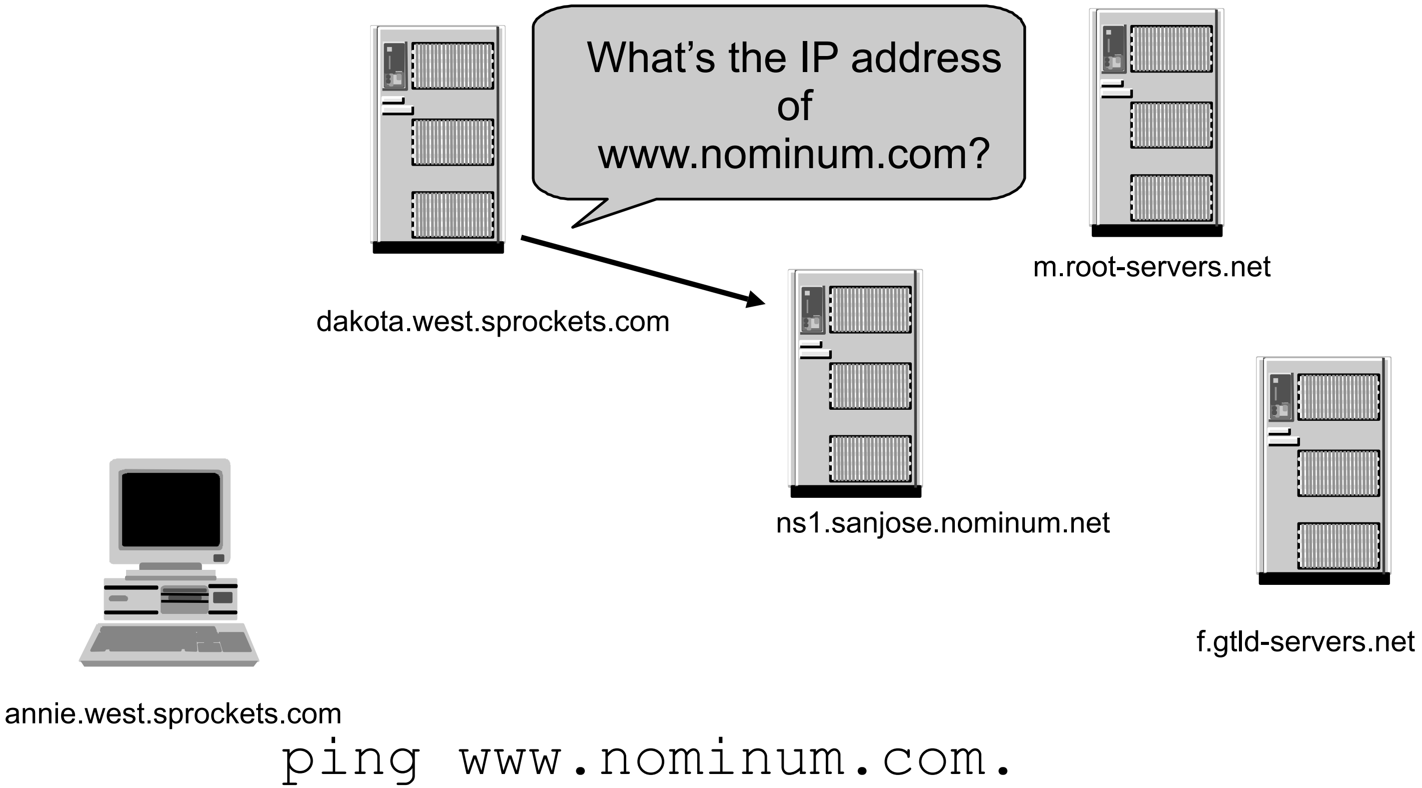
- The *com* name server *f* refers *dakota* to the *nominum.com* name servers



`ping www.nominum.com.`

The Resolution Process

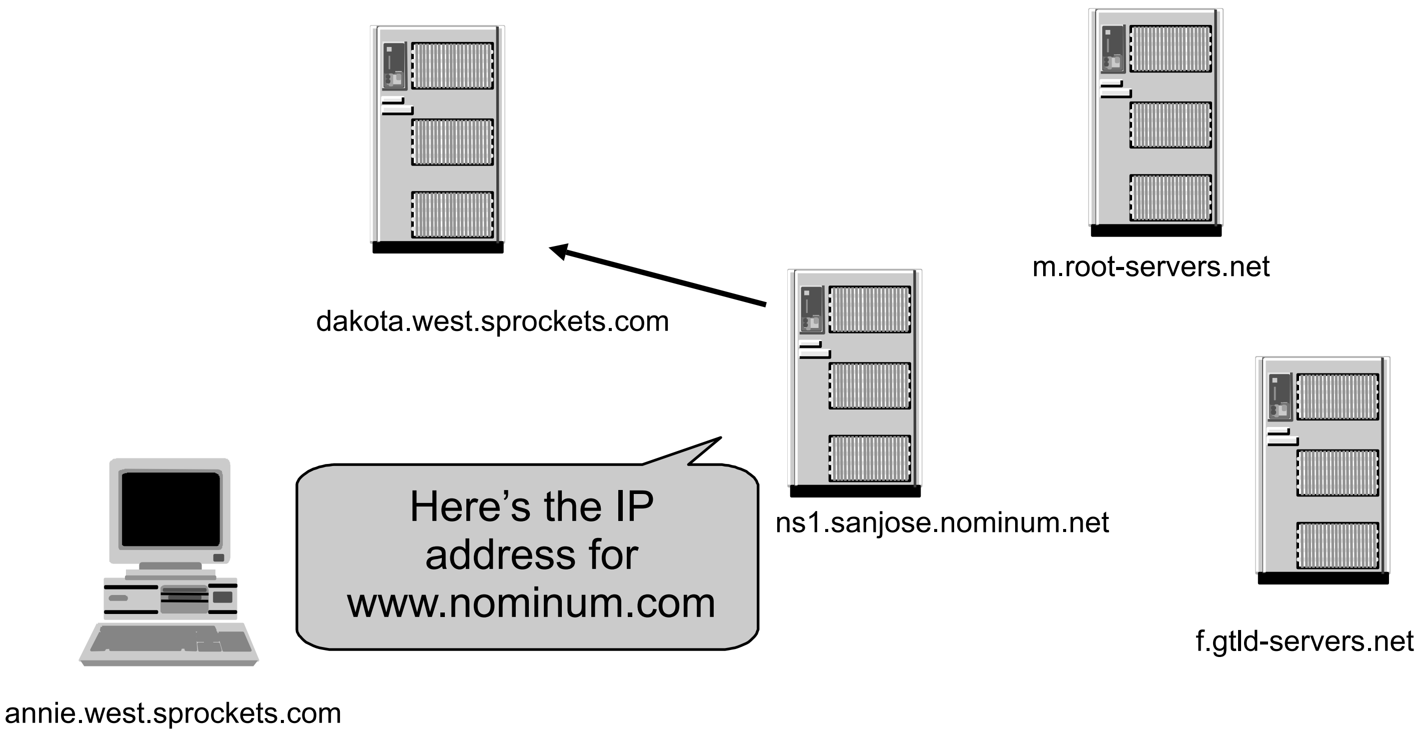
- The name server *dakota* asks an *nominum.com* name server, *ns1.sanjose*, for *www.nominum.com's* address



nominum COMMUNICATION BY NAME™

The Resolution Process

- The *nominum.com* name server *ns1.sanjose* responds with *www.nominum.com's* address

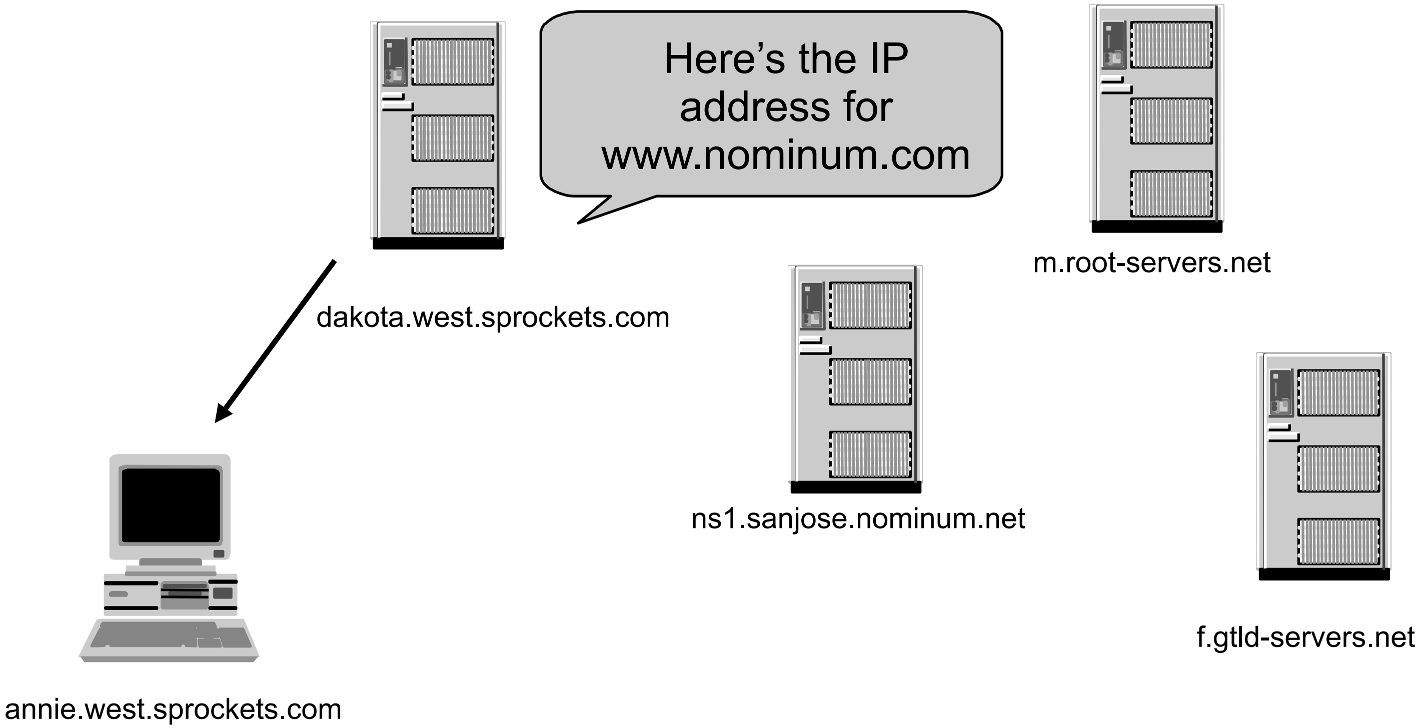


`ping www.nominum.com.`

nominum COMMUNICATION BY NAME™

The Resolution Process

- The name server *dakota* responds to *annie* with *www.nominum.com's* address




```
ping www.nominum.com.
```

Resolution Process (Caching)

- After the previous query, the name server *dakota* now knows:
  - The names and IP addresses of the *com* name servers
  - The names and IP addresses of the *nominum.com* name servers
  - The IP address of *www.nominum.com*
- Let's look at the resolution process again

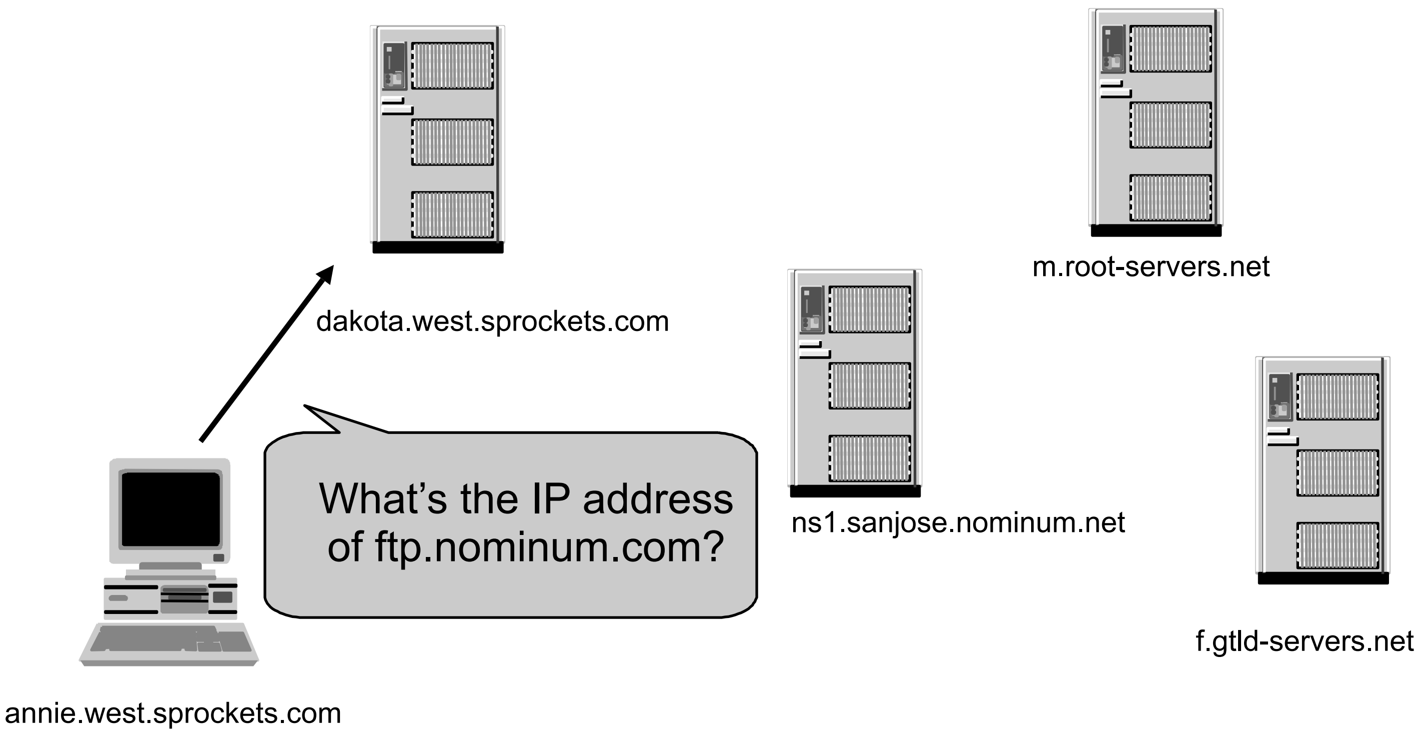
annie.west.sprockets.com

`ping` **`ftp`**`.nominum.com.`

nominum COMMUNICATION BY NAME™

Resolution Process (Caching)

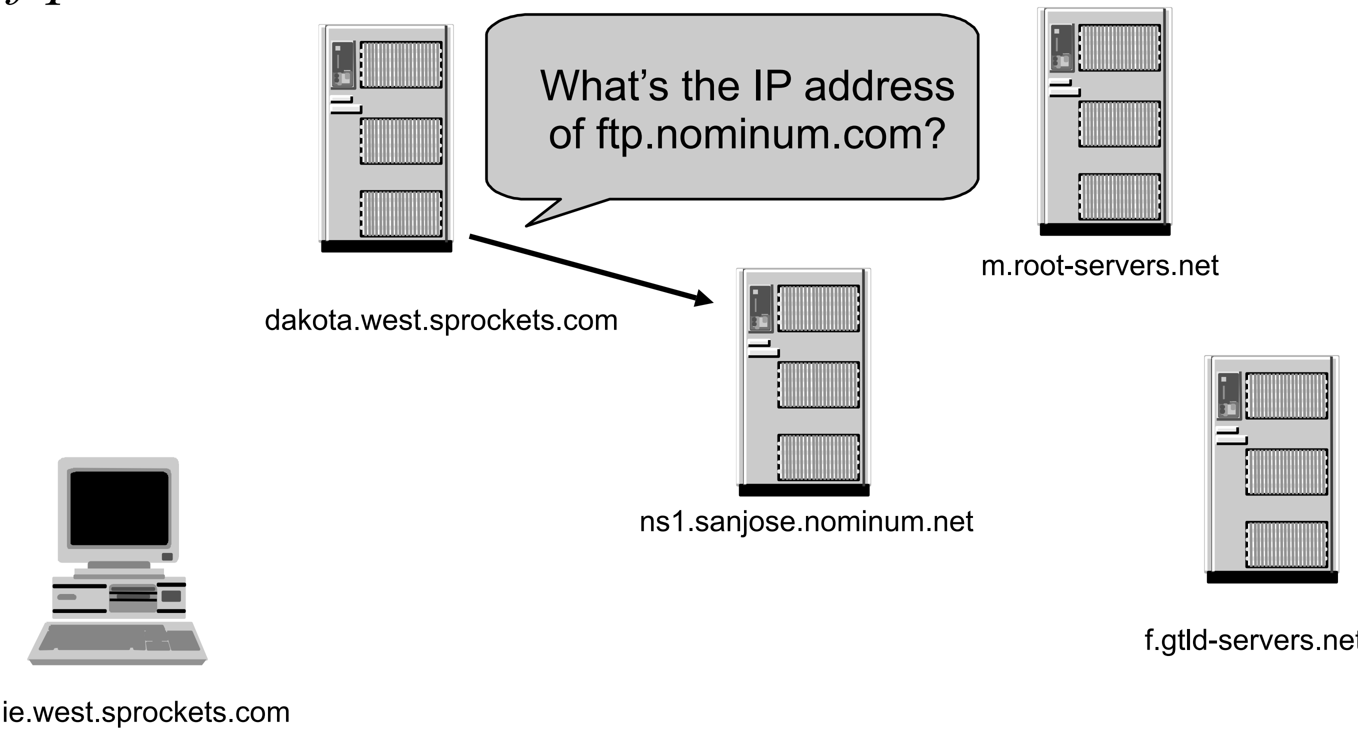
- The workstation *annie* asks its configured name server, *dakota,* for *ftp.nominum.com's* address




```
ping ftp.nominum.com.
```

Resolution Process (Caching)

- *dakota* has cached an NS record indicating *ns1.sanjose* is an *nominum.com* name server, so it asks it for *ftp.nominum.com's* address

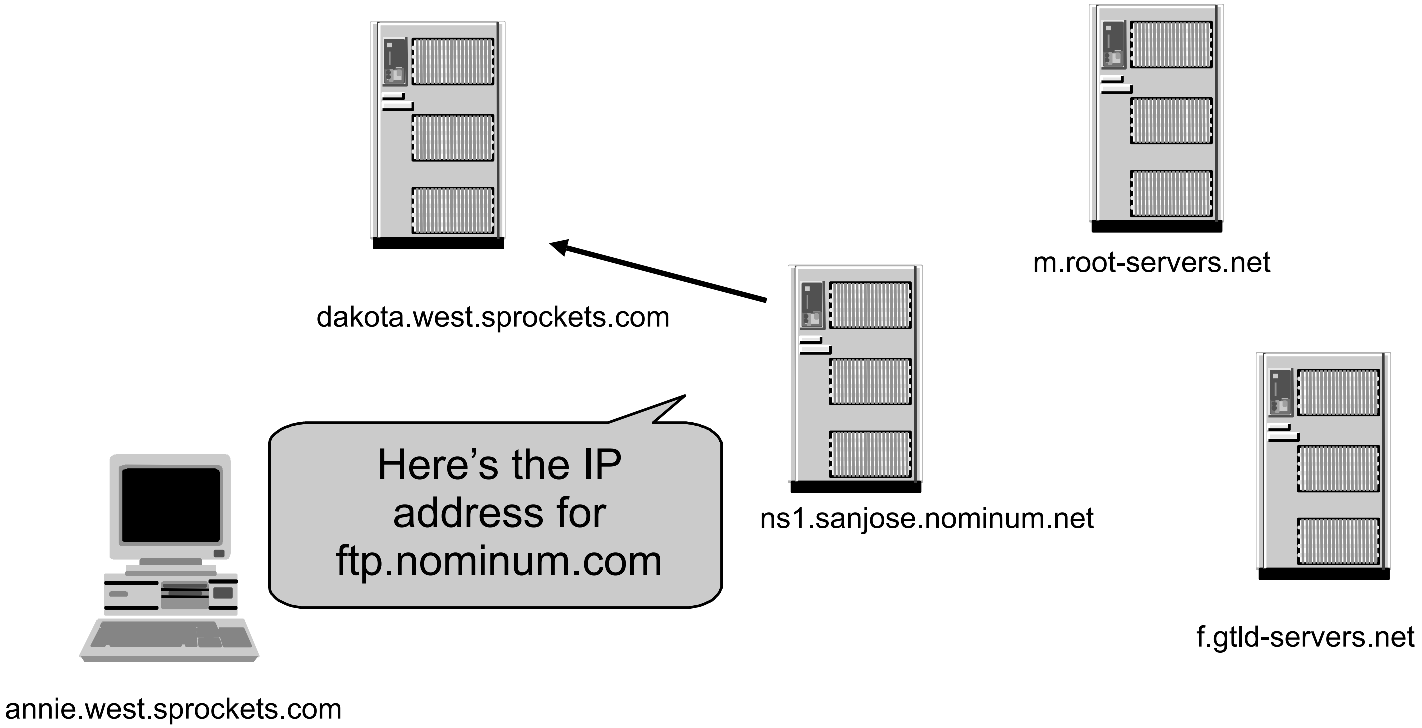



```
ping ftp.nominum.com.
```


nominum COMMUNICATION BY NAME™

Resolution Process (Caching)

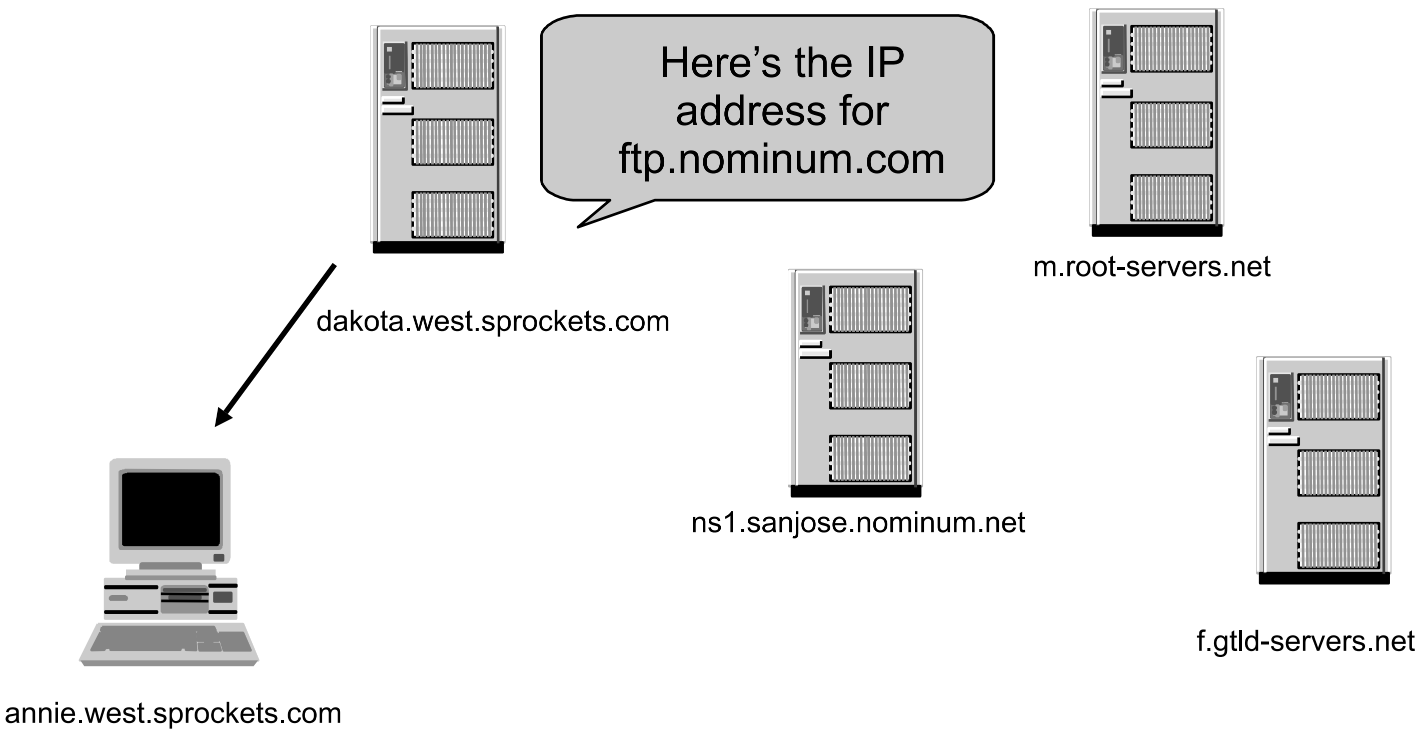
- The *nominum.com* name server *ns1.sanjose* responds with *ftp.nominum.com's* address



`ping ftp.nominum.com.`

Resolution Process (Caching)

- The name server *dakota* responds to *annie* with *ftp.nominum.com's* address




```
ping ftp.nominum.com.
```


nominum COMMUNICATION BY NAME™

What can be Resolved?

- Any name in the name space
- Class
  - Internet (IN), Chaos (CH), Hesiod (HS)
- Type
  - Address (A, AAAA, A6)
  - Pointer (PTR, NAPTR)
  - Aliases (CNAME, DNAME)
  - Security related (TSIG, SIG, NXT, KEY)
  - Etc.

Overview

- Introduction to the DNS
- DNS Components
- DNS Structure and Hierarchy
- The DNS in Context

DNS Structure and Hierarchy

- The DNS imposes no constraints on how the DNS hierarchy is implemented except:
  - A single root
  - The label restrictions
- If a site is not connected to the Internet, it can use any domain hierarchy it chooses
  - Can make up whatever TLDs you want
- Connecting to the Internet implies use of the existing DNS hierarchy

Top-level Domain (TLD) Structure

- In 1983 (RFC 881), the idea was to have TLDs correspond to network service providers
  - e.g., ARPA, DDN, CSNET, etc.
    - Bad idea: if your network changes, your email address changes
- By 1984 (RFC 920), functional domains was established
  - "The motivation is to provide an organization name that is free of undesirable semantics."
  - e.g., GOV for Government, COM for commercial, EDU for education, etc.
- RFC 920 also provided for
  - Provided for country domains
  - Provided for "Multiorganizations"

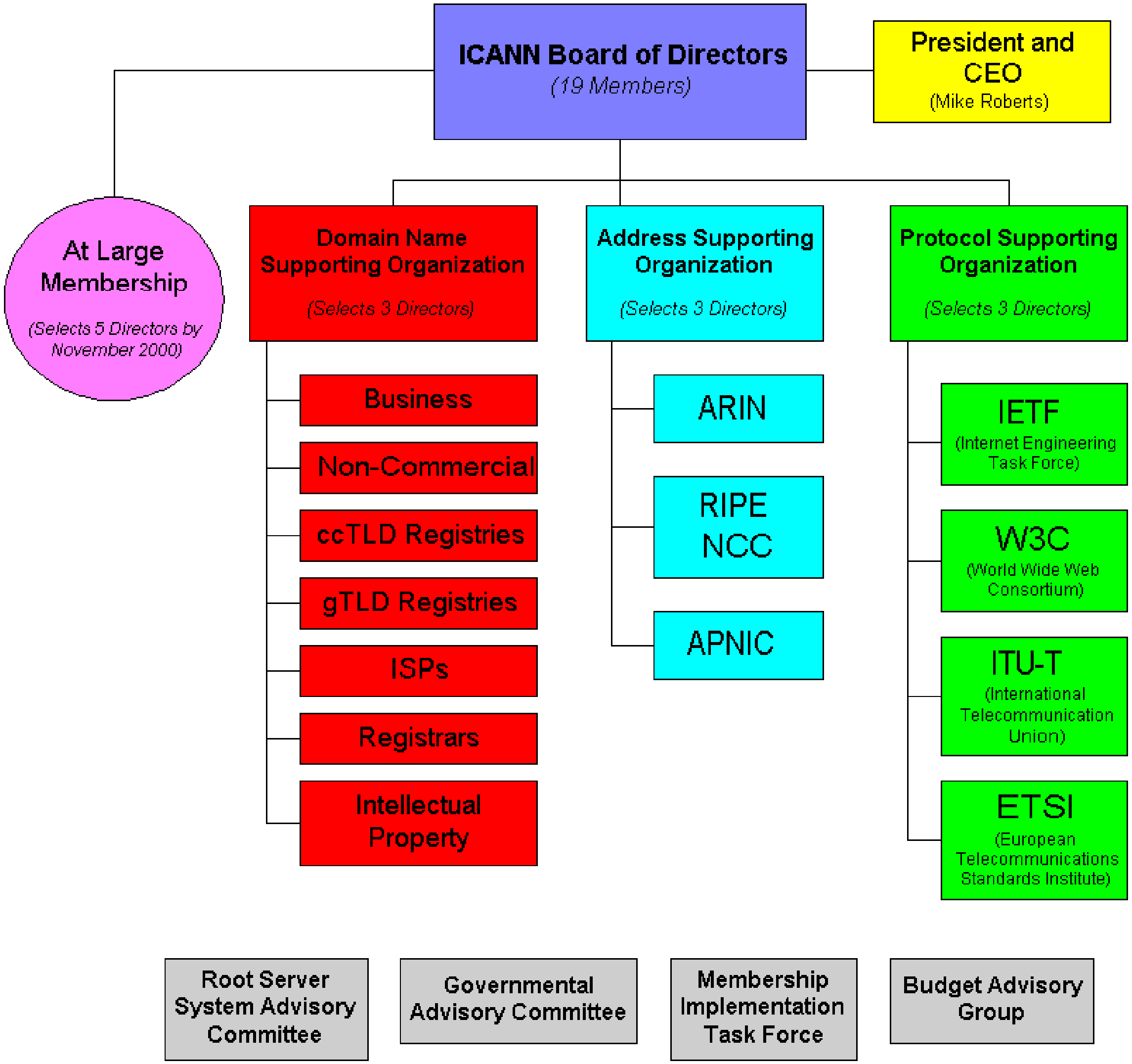
The Domain Name Wars

- In 1996,the US National Science Foundation permitted Network Solutions to charge a usage fee for the allocation and registration of domain names
  - To compensate for the explosive growth the Internet was facing at the time
- The resultant controversy caused the US Government (Dept. of Commerce) to take a much more active role
  - Official governmental policy (the White Paper) on Internet resource administration created
- That policy resulted in the creation of ICANN

Internet Corporation for Assigned Names and Numbers

- California non-profit, operating in Marina Del Rey, California, USA
- Consists of:
  - A set of Support Organizations
    - Address Support Organization, Domain Name Support Organization, Protocol Support Organization
  - A board of 19 members
    - 9 elected by public membership
    - 3 each by each of the SOs
    - 1 President/CEO
  - A set of committees
    - Governmental Advisory Committee, Addressing Ad Hoc Committee, etc. that advise the board

ICANN Organizational Chart



ICANN’s Role

- To oversee administer Internet resources including
  - Addresses
    - Delegating blocks of addresses to the regional registries
  - Protocol identifiers and parameters
    - Allocating port numbers, OIDs, etc.
  - Names
    - Administration of the root zone file
    - Oversight of the operation of the root name servers

The Internet Root

- The DNS protocol assumes a consistent name space
  - This consistency is enforced by the constraint of a **SINGLE** root for the Internet domain name space
    - There is no assumption on how that single root is created
- ICANN oversees modification of the zone file that makes up the Internet DNS root

Multiple Roots?

- The single root is often seen as a single point of control for the entire Internet
  - Edit control of the root zone file implies the ability to control the entire tree
- Multiple root solutions have often been proposed
  - Unless coordinated, inconsistencies will almost certainly result
    - The answer you get depends on where you ask
      - This would be "bad".

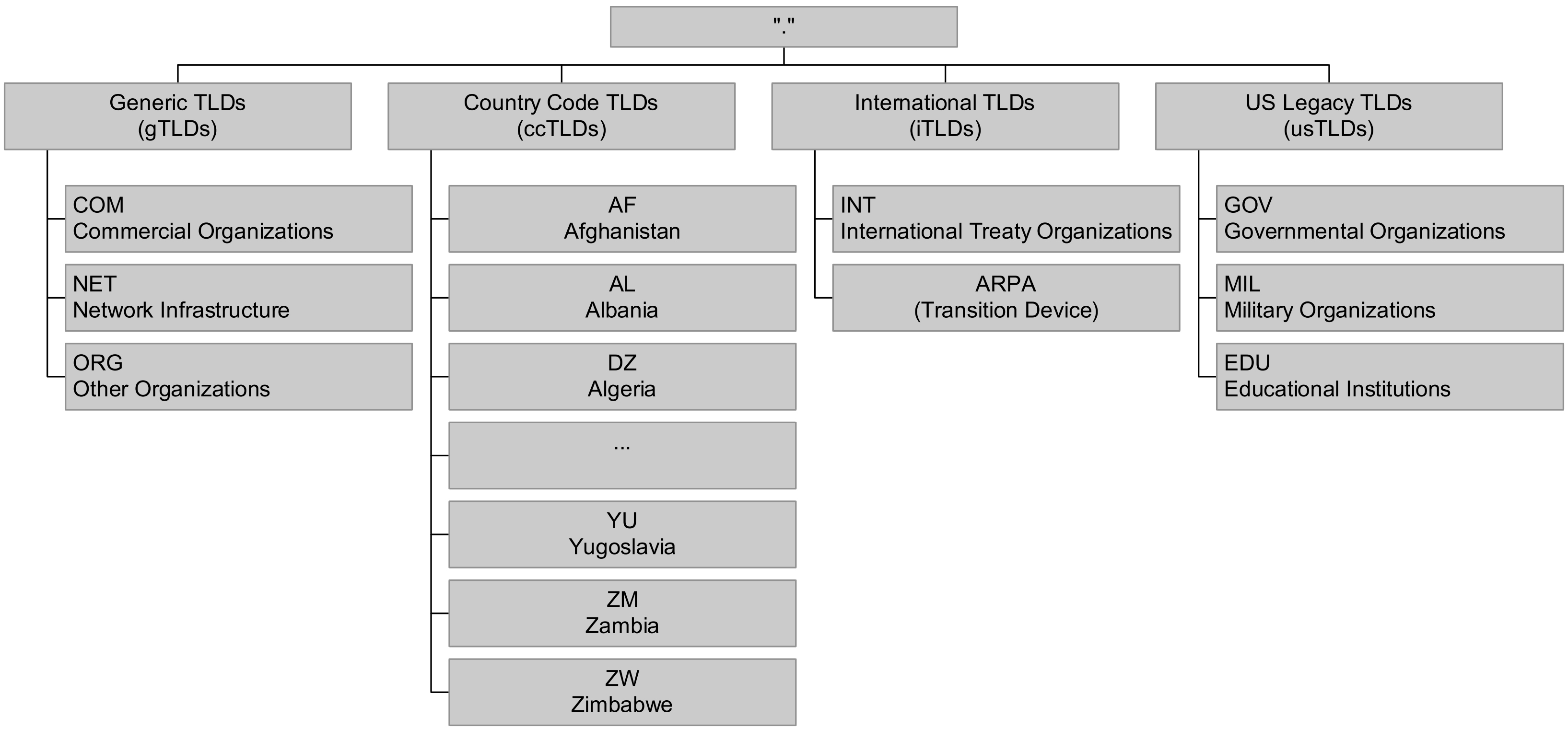
The Root Nameservers

- Modification of the root zone file is pointless unless that zone file is published
- The root zone file is published on 13 servers, “A” through “M”, around the Internet
  - Location of root nameserver is a function of network topology
- Root name server operations currently provided by volunteer efforts by a very diverse set of organizations
  - Volunteer nature will change soon

Root Name Server Operators

Nameserver	Operated by:
A	Verisign (US East Coast)
B	University of S. California –Information Sciences Institute (US West Coast)
C	PSI (US East Coast)
D	University of Maryland (US East Coast)
E	NASA (Ames) (US West Coast)
F	Internet Software Consortium (US West Coast)
G	U. S. Dept. of Defense (ARL) (US East Coast)
H	U. S. Dept. of Defense (DISA) (US East Coast)
I	KTH (SE)
J	Verisign (US East Coast)
K	RIPE-NCC (UK)
L	ICANN (US West Coast)
M	WIDE (JP)

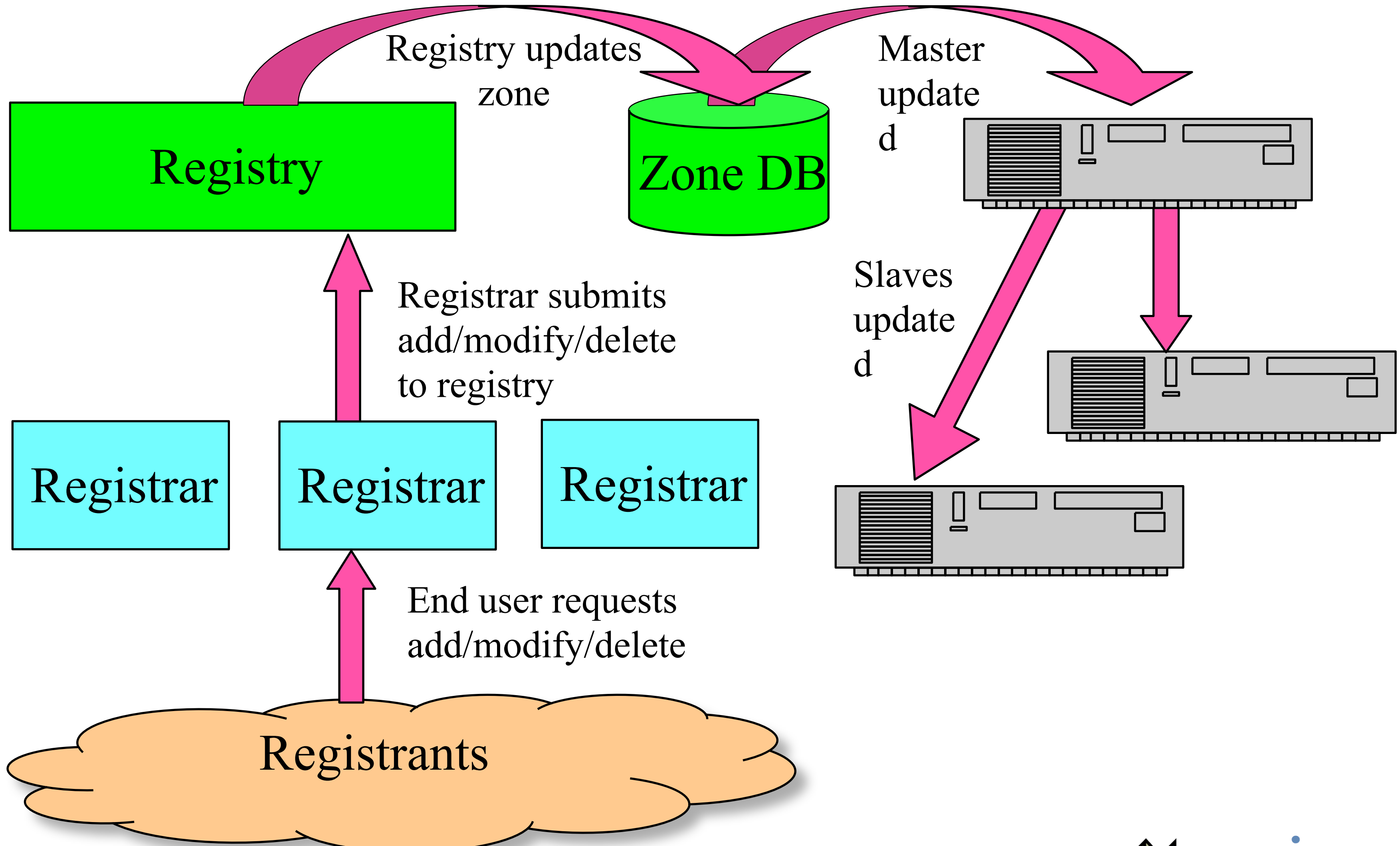
The Current TLDs



Registries, Registrars, and Registrants

- The Domain Wars resulted in a codification of roles in the operation of a domain name space
- Registry
  - the name space's database
  - the organization which has edit control of that database
    - Including dispute resolution, policy control, etc.
  - The organization which runs the authoritative name servers for that name space
- Registrar
  - the agent which submits change requests to the registry on behalf of the registrant
- Registrant
  - The entity which makes use of the domain name

Registries, Registrars, and Registrants



nominum
COMMUNICATION BY NAME™

The "Generic" Top-Level Domains (gTLDs)

- .COM, .NET, and .ORG
  - By far the largest top level domains on the Internet today
    - .COM has approx. 20,000,000 names
  - Essentially no restriction on what can be registered
- Network Solutions (now Verisign) received the contract for the registry for .COM, .NET, and .ORG
  - also a registrar for these TLDs

New Top Level Domains

- Recently, ICANN created 7 new top level domains:
  - .aero, .biz, .coop, .info, .museum, .name, .pro
    - Some are chartered (.aero, .coop, .museum, .name, .pro)
    - Some are generic (.biz, .info)
  - Expect these new TLDs to show up around 2Q01
- Many people unhappy with the process by which these new TLDs were created
  - Expect continued "discussion"

nominum
COMMUNICATION BY NAME™

Country Code Top-Level Domains

- With RFC 920, the concept of domains delegated on the basis of nations was recognized
- Conveniently, ISO has a list of “official” country code abbreviations
  - ISO-3166
- IANA has also used Universal Postal Codes
  - (e.g., .GG for Guernsey)
- Key consideration is to use lists other organizations define to avoid getting into political battles over what is or is not a valid ccTLD

ccTLD Organization

- How each country top-level domain is organized is up to the country
  - Some, like Australia's au, follow the functional definitions
    - *com.au*, *edu.au*, etc.
  - Others, like Great Britain's uk and Japan's jp, divide the domain functionally but use their own abbreviations
    - *ac.uk*, *co.uk*, *ne.jp*, *ad.jp*, etc.
  - A few, like the United State's us, are largely geographical
    - *co.us*, *md.us*, etc.
  - Canada uses organizational scope
    - *bnr.ca* has national scope, *risq.qc.ca* has Quebec scope
  - Some are flat, that is, no hierarchy

.arpa

- Now, Address and Routing Parameter Area
  - Was Advanced Research Projects Administration
    - US Dept. of Defense network, precursor to the Internet
- Used for infrastructure domains
  - IPv4 reverse (address to name) lookups
  - IPv6 reverse lookups
  - E.164
- Only .arpa is hardwired into the DNS sysem
  - DNS resolver software has it explicitly

Other TLDs

- .GOV – used by US Governmental organizations
  - E.g., state.gov, doj.gov, whitehouse.gov, etc.
- .MIL – used by the US Military
  - E.g., af.mil, army.mil, etc.
- .EDU – used for Educational institutions
  - Higher learning, not only US-based ones
  - E.g., harvard.edu, unu.edu, utoronto.edu
- .INT – international treaty organizations
  - E.g., itu.int, nato.int, wipo.int

Overview

- Introduction to the DNS
- DNS Components
- DNS Hierarchy
- The DNS in Context

Load concerns

- DNS can handle the load
  - DNS Root Servers get approximately 3000 queries per second (down from 8000 qps)
    - Empirical proofs (DDoS attacks) show root name servers can handle 50,000 queries per second
      - Limitation is network bandwidth, not the DNS protocol
  - in-addr.arpa zone, which translates numbers to names, gets about 2000 queries per second
    - Current closest analog to e164.arpa

Performance concerns

- DNS is a very lightweight protocol
  – Simple query – response
- Any performance limitations are the result of network limitations
  – Speed of light
  – Network congestion
  – Switching/forwarding latencies

Security Concerns

- Base DNS protocol (RFC 1034, 1035) is insecure
  - "Spoof" attacks are possible
- DNS Security Enhancements (DNSSEC, RFC 2565) remedies this flaw
  - But creates new ones
    - DoS attacks
    - Amplification attacks
    - Operational considerations
- DNSSEC strongly discourages large flat zones
  - Hierarchy (delegation) is good

Technically Speaking…

- ENUM is technically non-challenging
  – Intelligent delegation model will permit unlimited scaling
  – Performance considerations at the feet of service providers
  – Security concerns can be addressed by DNSSEC

Questions?

?