

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

The 2009 International Conference on Internet Computing
July 14, 2009

Organization of this talk

- 1 Introduction
 - Domain Setup
 - Hierarchy
- 2 Search
 - Hash Scheme
 - Space Division
 - Routing Table
- 3 Enhancements
 - Routing Issues I
 - Routing Issues II
 - Search Redundancy
- 4 Search
 - Distributing Search
- 5 Questions

Distributed
Search for
Multicast
Sessions

Piyush Harsh
& Richard
Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing
Search

Questions

mDNS is a multicast session directory architecture. It is

- DNS aware,
- hierarchial and
- scalable.

It allows for multicast session registration and makes them discoverable in real time.

A typical mDNS domain components

- Distributed Search for Multicast Sessions
- Piyush Harsh & Richard Newman
- Introduction
- Domain Setup Hierarchy
- Search
- Hash Scheme
- Space Division
- Routing Table
- Enhancements
- Routing Issues I
- Routing Issues II
- Search Redundancy
- Search
- Distributing Search
- Questions

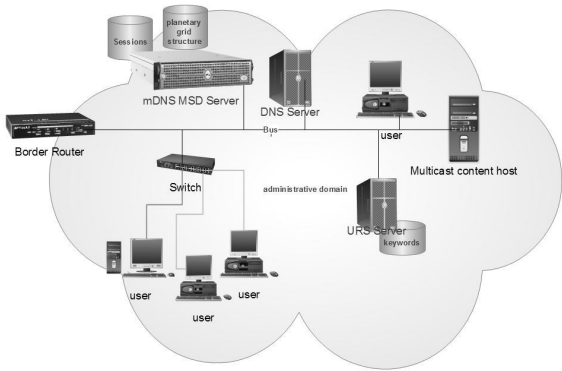


Figure: a typical mDNS domain setup

A typical mDNS domain hierarchy

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

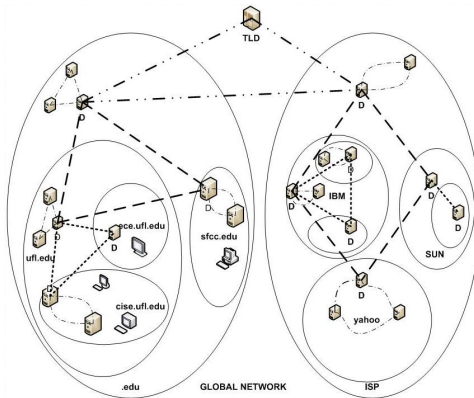
Introduction
Domain Setup Hierarchy

Search
Hash Scheme
Space Division
Routing Table

Enhancements
Routing Issues I
Routing Issues II
Search
Redundancy

Search
Distributing Search

Questions



Hierarchy buildup is due to inclusion of these parameters in MSD servers.

@MCAST {

}

A typical mDNS domain hierarchy

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

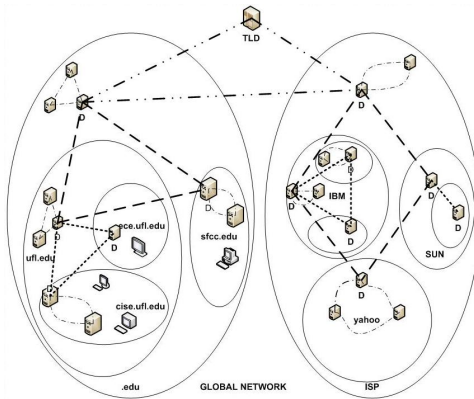
Introduction
Domain Setup Hierarchy

Search
Hash Scheme
Space Division
Routing Table

Enhancements
Routing Issues I
Routing Issues II
Search Redundancy

Search
Distributing Search

Questions



Hierarchy buildup is due to inclusion of these parameters in MSD servers.

@MCAST { **PMCAST** }

A typical mDNS domain hierarchy

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction
Domain Setup Hierarchy

Search

Hash Scheme
Space Division
Routing Table

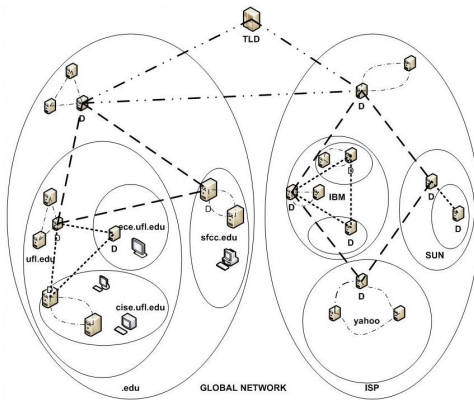
Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing Search

Questions



Hierarchy buildup is due to inclusion of these parameters in MSD servers.

@MCAST { PMCAST **CMCAST** }

A typical mDNS domain hierarchy

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction
Domain Setup Hierarchy

Search

Hash Scheme
Space Division
Routing Table

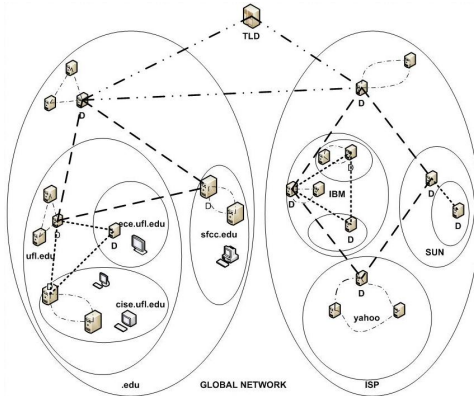
Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing Search

Questions



Hierarchy buildup is due to inclusion of these parameters in MSD servers.

@MCAST { PMCAST CMCAST **MSD-LOCAL-MCAST** }

A typical mDNS domain hierarchy

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction
Domain Setup Hierarchy

Search

Hash Scheme
Space Division
Routing Table

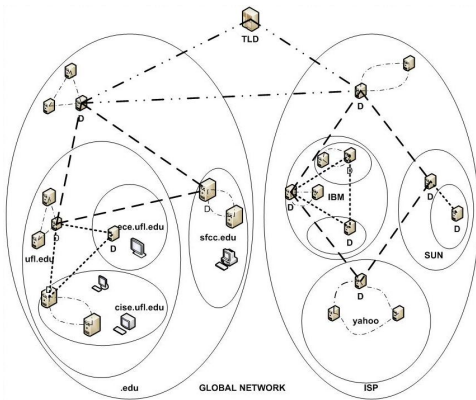
Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing Search

Questions



Hierarchy buildup is due to inclusion of these parameters in MSD servers.

@MCAST { PMCAST CMCAST MSD-LOCAL-MCAST **URS** }

Distributed
Search for
Multicast
Sessions

Piyush Harsh
& Richard
Newman

Introduction
Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing
Search

Questions

In a major design improvement from earlier proposal, we have optimized the session search algorithm in current version.

New algorithm is based on appropriate keyword routing that allows activation of a few MSD servers along the route compared to a broadcast storm approach deployed in earlier implementation.

- Each MSD server maintains global session information for the keyword set assigned to it.

- Each MSD server maintains global session information for the keyword set assigned to it.
- A session keyword is hashed using MD5.

- Each MSD server maintains global session information for the keyword set assigned to it.
- A session keyword is hashed using MD5.
- Keyword hash is used to route the session registration and search requests to appropriate server.

- Each MSD server maintains global session information for the keyword set assigned to it.
- A session keyword is hashed using MD5.
- Keyword hash is used to route the session registration and search requests to appropriate server.
- Each MSD server maintains a keyword routing table that facilitates correct request routing.

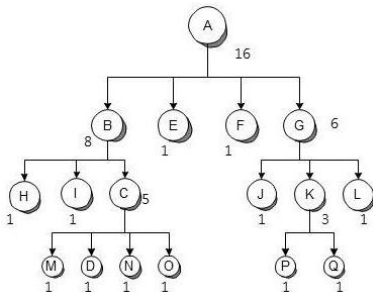


Figure: a sample hierarchy

- Each MSD server reports total count of domains below it including itself to its parent.

Hash Space Distribution

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search Redundancy

Search

Distributing Search

Questions

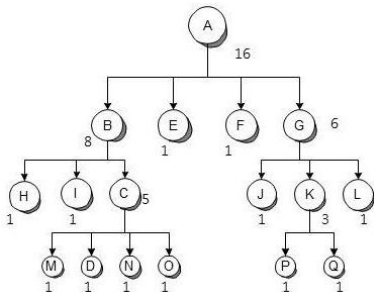


Figure: a sample hierarchy

- Each MSD server reports total count of domains below it including itself to its parent.
- In time, the root node knows the total count of mDNS domains in the hierarchy.

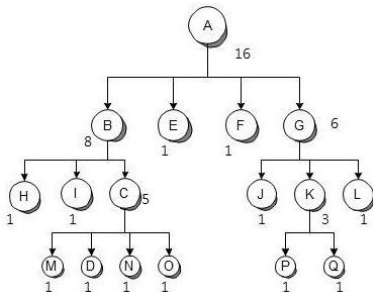


Figure: a sample hierarchy

- Each MSD server reports total count of domains below it including itself to its parent.
- In time, the root node knows the total count of mDNS domains in the hierarchy.
- The root node does hash space division using this knowledge.

Hash Space Distribution

Distributed
Search for
Multicast
Sessions

Piyush Harsh
& Richard
Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing
Search

Questions

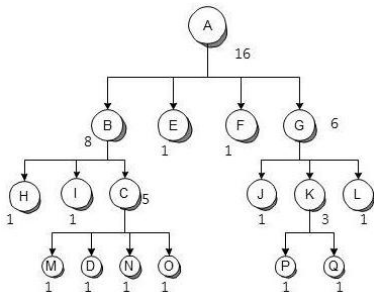


Figure: a sample hierarchy

- Each MSD server reports total count of domains below it including itself to its parent.
- In time, the root node knows the total count of mDNS domains in the hierarchy.
- The root node does hash space division using this knowledge.
- The space allotment then propagates top-down from root to leaves.

Hash Space Distribution

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search Redundancy

Search

Distributing Search

Questions

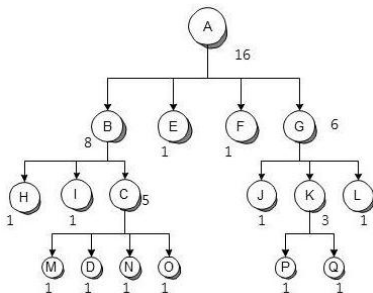


Figure: a sample hierarchy

- Each MSD server reports total count of domains below it including itself to its parent.
- In time, the root node knows the total count of mDNS domains in the hierarchy.
- The root node does hash space division using this knowledge.
- The space allotment then propagates top-down from root to leaves.
- Each node uses this space division to construct the forwarding table for correct routing of service requests.

Routing Table Construction

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing
Search

Questions

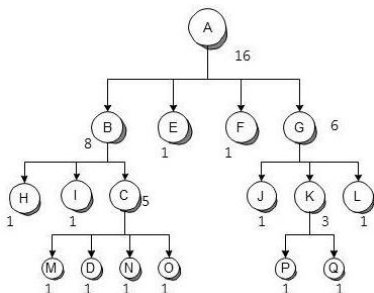


Figure: a sample hierarchy

1 Any request can be acted upon three possible ways.

- if associated keyword's hash falls in self managed hash range, then process locally.
- if keyword's hash lies in the range assigned to child domain, then propagate on CMAST.
- else forward on the PMCAST channel
- of course, care is taken to avoid request propagation looping!

Routing Table Construction

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing
Search

Questions

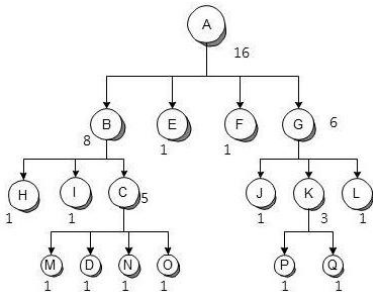


Figure: a sample hierarchy

- 1 Any request can be acted upon three possible ways.
 - if associated keyword's hash falls in self managed hash range, then process locally.
 - if keyword's hash lies in the range assigned to child domain, then propagate on CMAST.
 - else forward on the PMCAST channel
 - of course, care is taken to avoid request propagation looping!
- 2 Routing table contains significant bits, start - end of range and next channel.

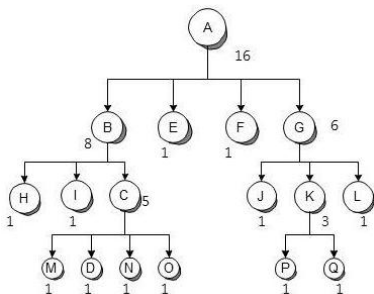


Figure: a sample hierarchy

start	end	node ID	channel
0000	0000	self	MSD-LOCAL-MCAST
0001	0001	H	CMCAST
0010	0010	I	CMCAST
0011	0111	C	CMCAST
*	*	-	PMCAST
Significant Bits: 4			

routing table maintained at node B.

Here the root node is not participating as a MSD server. If it were, the distribution would have been slightly different.

- Most state information is maintained using soft-state approach.

- Most state information is maintained using soft-state approach.
 - this allows for dynamic adaptability to changing topology
 - if topology change is very frequent, this may lead to instability

- Most state information is maintained using soft-state approach.
 - this allows for dynamic adaptability to changing topology
 - if topology change is very frequent, this may lead to instability
- Routing tables need to be updated when new domains are added or old domains go offline

- Most state information is maintained using soft-state approach.
 - this allows for dynamic adaptability to changing topology
 - if topology change is very frequent, this may lead to instability
- Routing tables need to be updated when new domains are added or old domains go offline
 - routing stability is essential for smooth operations
 - frequent routing updates may degrade the service quality offered by mDNS.

- Most state information is maintained using soft-state approach.
 - this allows for dynamic adaptability to changing topology
 - if topology change is very frequent, this may lead to instability
- Routing tables need to be updated when new domains are added or old domains go offline
 - routing stability is essential for smooth operations
 - frequent routing updates may degrade the service quality offered by mDNS.

- In order to improve routing stability, domain count reporting to parent nodes is governed by change variables α and β

Distributed
Search for
Multicast
Sessions

Piyush Harsh
& Richard
Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing
Search

Questions

- In order to improve routing stability, domain count reporting to parent nodes is governed by change variables α and β
 - typical values used are: $\alpha=0.4$ and $\beta=0.8$
 - if fractional change in node count is less than α then no action is taken

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing Search

Questions

- In order to improve routing stability, domain count reporting to parent nodes is governed by change variables α and β
 - typical values used are: $\alpha=0.4$ and $\beta=0.8$
 - if fractional change in node count is less than α then no action is taken
 - if change lies between α and β then hash reassignment is done for all child nodes
 - this may lead to routing table updates from that particular node to the leaves along that branch

- In order to improve routing stability, domain count reporting to parent nodes is governed by change variables α and β
 - typical values used are: $\alpha=0.4$ and $\beta=0.8$
 - if fractional change in node count is less than α then no action is taken
 - if change lies between α and β then hash reassignment is done for all child nodes
 - this may lead to routing table updates from that particular node to the leaves along that branch
 - if fractional change is more than β then the updated count is reported to the parent node.
 - this may lead to global route updates
 - a higher value of β is suggested to minimize this effect

- In order to improve routing stability, domain count reporting to parent nodes is governed by change variables α and β
 - typical values used are: $\alpha=0.4$ and $\beta=0.8$
 - if fractional change in node count is less than α then no action is taken
 - if change lies between α and β then hash reassignment is done for all child nodes
 - this may lead to routing table updates from that particular node to the leaves along that branch
 - if fractional change is more than β then the updated count is reported to the parent node.
 - this may lead to global route updates
 - a higher value of β is suggested to minimize this effect

As a precaution against intermittent MSD server failures, the registration request is duplicated and sent along different route by inverting all keyword hash bits. Therefore two copies of session details are maintained by two separate MSD servers in two different domains.

As a precaution against intermittent MSD server failures, the registration request is duplicated and sent along different route by inverting all keyword hash bits. Therefore two copies of session details are maintained by two separate MSD servers in two different domains.

In case one of these two servers is down, the session details can be retrieved from the backup server by sending the search request by inverting all the keyword hash bits along the alternate route.

Search Algorithm - By Example

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

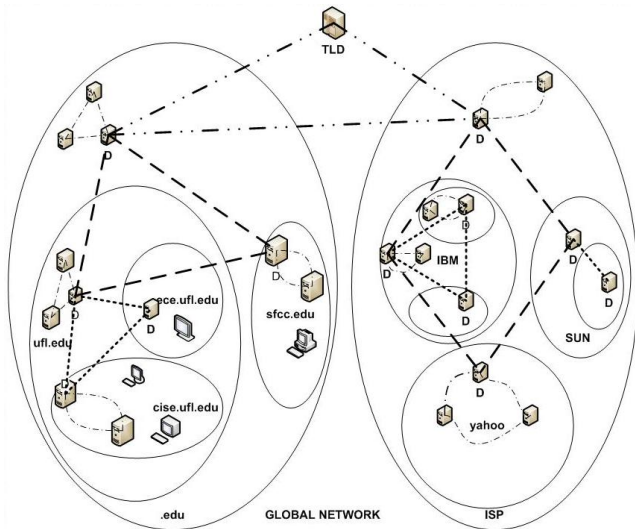
Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing
Search

Questions



Search Algorithm - By Example

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

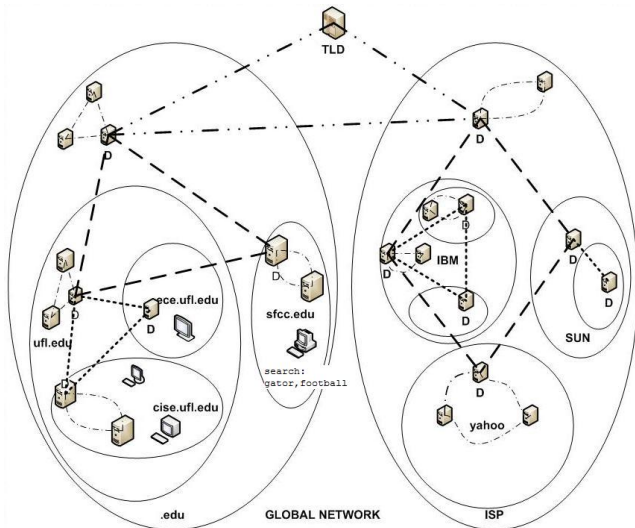
Enhancements

Routing Issues I
Routing Issues II
Search Redundancy

Search

Distributing Search

Questions



Search Algorithm - By Example

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

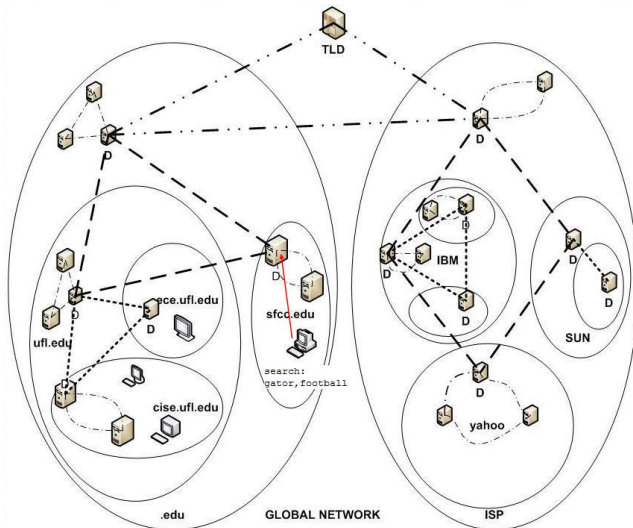
Enhancements

Routing Issues I
Routing Issues II
Search Redundancy

Search

Distributing Search

Questions



Search Algorithm - By Example

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

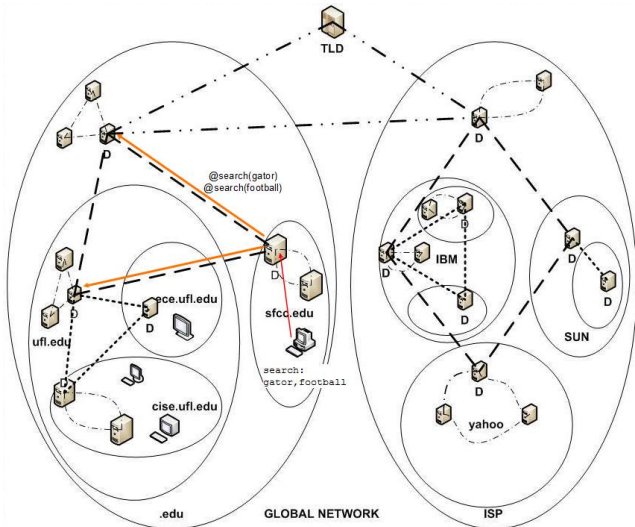
Enhancements

Routing Issues I
Routing Issues II
Search Redundancy

Search

Distributing Search

Questions



Search Algorithm - By Example

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

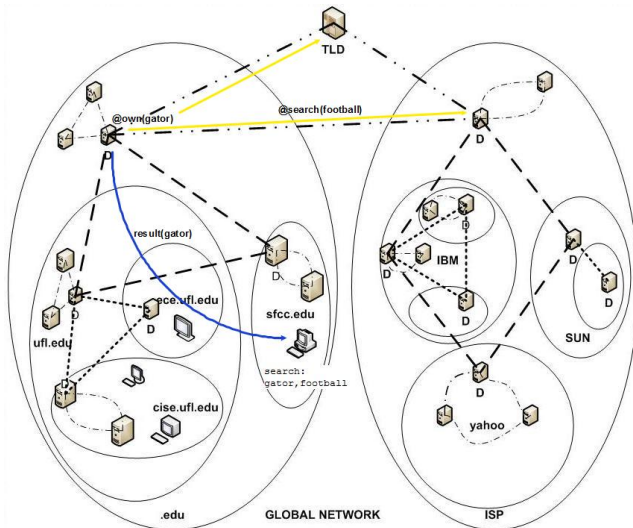
Introduction
Domain Setup
Hierarchy

Search
Hash Scheme
Space Division
Routing Table

Enhancements
Routing Issues I
Routing Issues II
Search Redundancy

Search
Distributing Search

Questions



Search Algorithm - By Example

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup Hierarchy

Search

Hash Scheme
Space Division
Routing Table

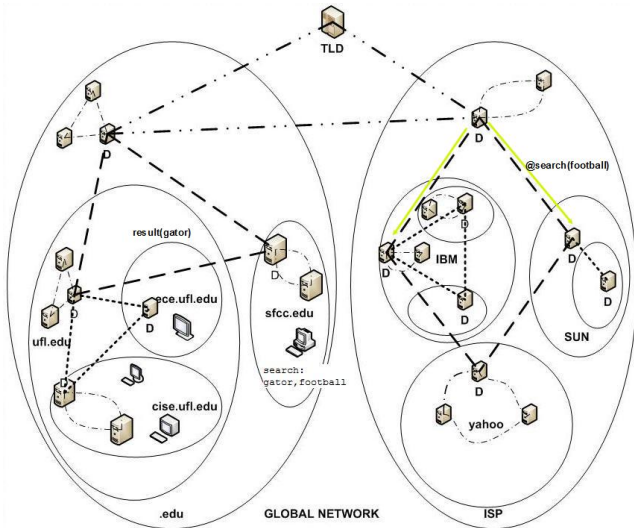
Enhancements

Routing Issues I
Routing Issues II
Search Redundancy

Search

Distributing Search

Questions



Search Algorithm - By Example

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

Introduction

Domain Setup Hierarchy

Search

Hash Scheme
Space Division
Routing Table

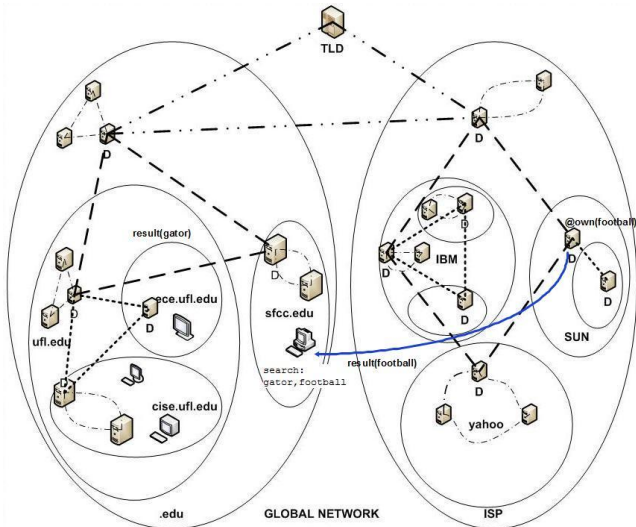
Enhancements

Routing Issues I
Routing Issues II
Search Redundancy

Search

Distributing Search

Questions



Search Algorithm - By Example

Distributed Search for Multicast Sessions

Piyush Harsh & Richard Newman

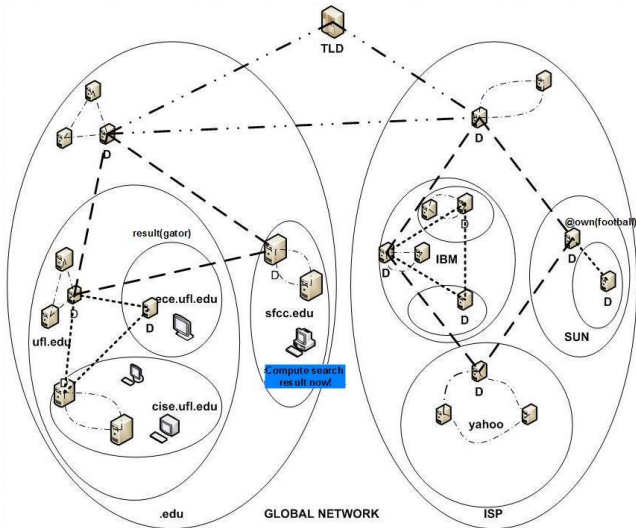
Introduction
Domain Setup Hierarchy

Search
Hash Scheme
Space Division
Routing Table

Enhancements
Routing Issues I
Routing Issues II
Search Redundancy

Search
Distributing Search

Questions



Distributed
Search for
Multicast
Sessions

Piyush Harsh
& Richard
Newman

Introduction

Domain Setup
Hierarchy

Search

Hash Scheme
Space Division
Routing Table

Enhancements

Routing Issues I
Routing Issues II
Search
Redundancy

Search

Distributing
Search

Questions

Any Questions?

Piyush Harsh

E-313A
CSE Building
University of Florida

email: pharsh@cise.ufl.edu

Richard Newman

E-346
CSE Building
University of Florida

email: nemo@cise.ufl.edu

Software Link

<http://www.cons.cise.ufl.edu/mdns/>