Routing Proximity

Stefano Previdi - spreviddi@cisco.com

Agenda

- Introduction and motivation
- Proximity
- Standardization & Research

Introduction and Motivation Contexts

- Caching, replication, positioning and find the location of content and services are vital for optimization of network resources
  - Need to locate the nearest copy of a movie or the closest instance of a service among several available resources
- In different contexts:
  - Managed Content/Application Network
    - Service Routing (Service Node / Service Gateway)
    - Traditional CDNs
  - Peer to Peer Overlay
    - BitTorrent, eMule, ...

Introduction and Motivation

- Application Networking assumes no constraint in location of elements

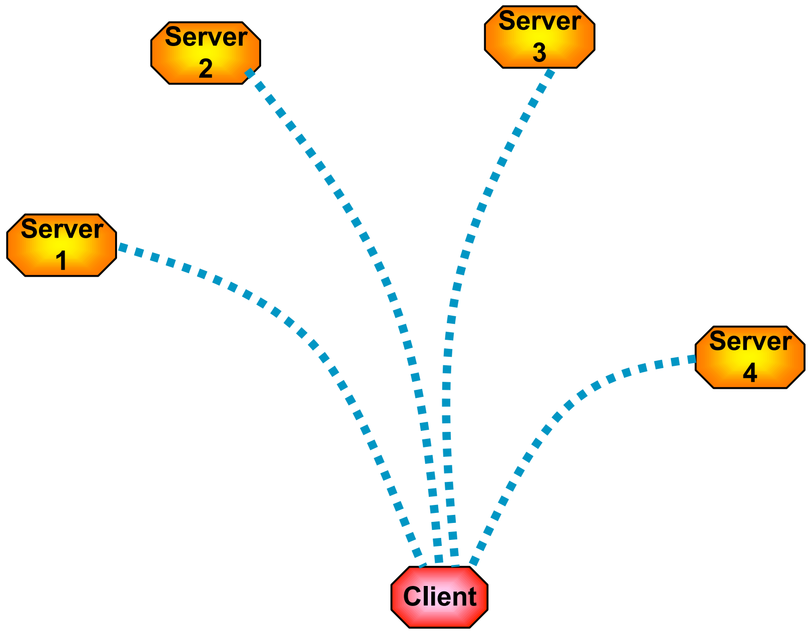
  Servers, caches, peers, hosts, ...

- Peer to peer applications are often observed to move a single piece of content back and forth across an inter-provider link multiple times, because the peer selection algorithm does not take account of the network layer topology.

  According to surveys, up to ~75% of currently used BW is consumed by p2p traffic

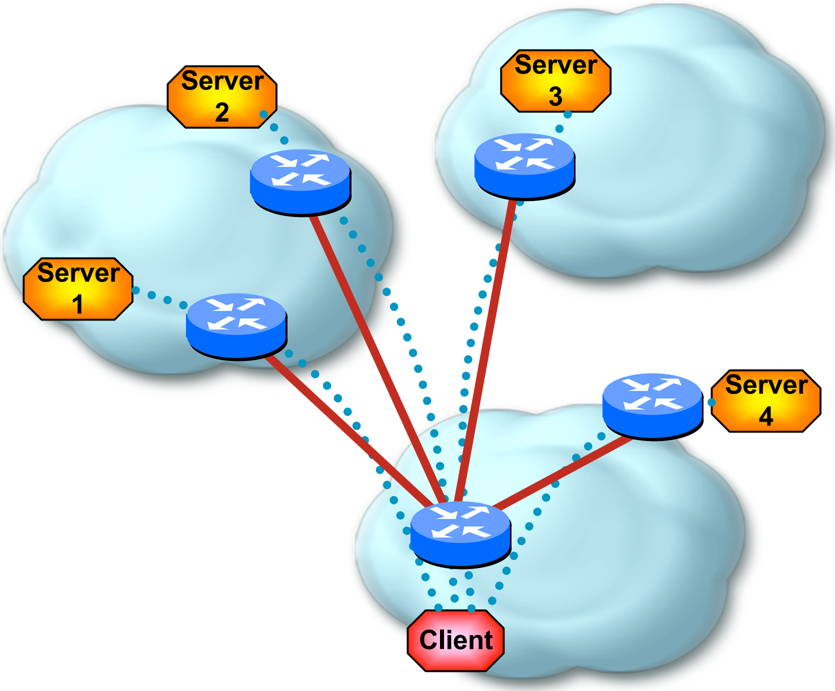
  Provisioning is dictated by low-revenue traffic

Introduction - Localization Example



- Movie “Spiderman 42” is available from different servers. Which one to select ? Based on which criteria ?
- Application layer goal: minimize delay, maximize user experience
- Network layer: reduce resources utilization while maintain best performance experience
- In this context: localization is intended as the ability to find/compute the exact location of an object

Introduction - Localization Example



- RTT measurement *may* give indication

  However, not efficient as solely based on one factor

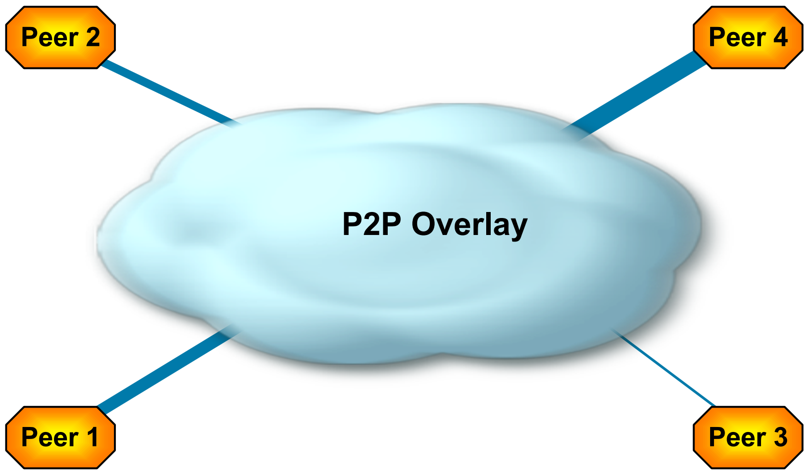
  Example: what if difference is insignificant ?

- Routing visibility gives exact view of traffic path

  Taking into account traffic direction

  Taking into account inter-AS connectivity and cost

Introduction - Localization Example



- Example:

  Peer to peer overlays, base their peer selection mechanisms on uplink capacity.

  Infrastructure is not considered.

Introduction - Localization

- The localization query in the most general form is: which targets in a given list are the closest to a particular source?

  E.g.: Tell me who is *closest* to 192.168.10.1 among:

  192.168.20.44
  192.168.43.32
  192.168.65.76

- Applications have implemented selection mechanisms taking into account application specific requirements

  Most of the time, network infrastructure not in selection process

Existing Solutions

- Several solutions exist trying to answer the localization question:
  - DNS based
  - RTT measurements
  - Manual configuration
  - GPS coordinates
  - Synthetic co-ordinate systems (create a virtual topology based on measurements)

Existing Solutions

- But each have its own set of issues:
  - Not always reliable
  - Do not reflect network/link resources availability
  - Not scalable when the large number of resources need to be tracked
  - Not responsive when routing topology changes
  - Does not address user mobility
  - Incentives not aligned between Application and Infrastructure
    - No easy way for provider to specify policy and steer traffic
    - Doesn’t allow flexible traffic management
- Little attention has been given to SP infrastructure
- Application layer lacks topology information
  - When selecting a given resource there’s no information about network resources going to be utilized

Agenda

- Introduction and motivation
- Cisco Proximity
- Standardization & Research

Goal

- Introduce localization/topology information in the selection process of any application component having to select a traffic source or destination
- Implement a Proximity component delivering localization services to application layer
- Define and implement a standard protocol for handling proximity requests
  - Between Proximity clients and servers

Proximity

- The localization query in the most general form is: which targets in a given list are the closest to a particular source?

  E.g.: Tell me who is *closest* to 192.168.10.1 among:

  192.168.20.44
  192.168.43.32
  192.168.65.76

- *Closest* may have different meaning in different contexts

  Routing cost ? Makes sense as traffic is routed/forwarding according to routing costs/topology

  State information (e.g.:RTT)? Makes sense as throughput may be affected by RTT

  Policy ? Makes sense as each SP may have how own preferences rules

  Others?

- Applications have implemented selection mechanisms taking into account application specific requirements

Cisco Proximity - Components

- Cisco Proximity Technology provides localization information in order to improve/guide selection mechanisms
  - Implemented through a ranking system
- Leverage Routing databases (ISIS/OSPF/BGP) in order to compute proximity/localization services
- Request/Reply based
  - Simple, lightweight protocol
  - KISS principle for scalability
  - a la DNS
- May be extended to other Proximity sources
  - NMS: state/performance Information
  - Application Specific information: DHT specifics

Cisco Proximity - Why Routing Layer ?

- Interface to routing layer is well known: routing adjacencies. No need to define anything new
- OSPF/ISIS and BGP Data has good proximity/localization properties
  - Includes protocol extensions/features for policy management
    - BGP Communities
    - Route Tags
    - ...
- Routing Proximity can be based on
  - IP addresses, Prefixes, AS numbers, BGP Communities, ...
  - Any routing information type

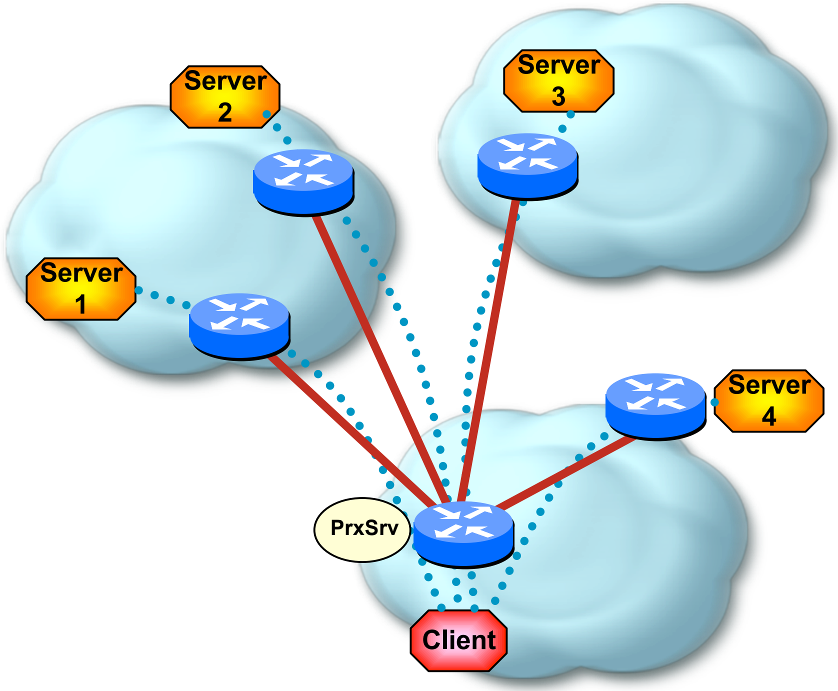
Cisco Proximity - Why Routing Layer ?

- Consistency between
  - Routing/forwarding decisions
  - Proximity decisions
- Use same topology information as used for routing
  - Not derived from third parties
- Information is kept up to date by routing layer: adapt to network events
- Leverage existing (and future) IGP/BGP protocol enhancements
  - Examples:
    - RFC 5029 - ISIS Link Attribute
    - RFC 5130 - ISIS Route Tags
    - draft-ietf-isis-genapp - ISIS Generic Application TLV
    - BGP Community / Extended Community
    - ...

Requirements

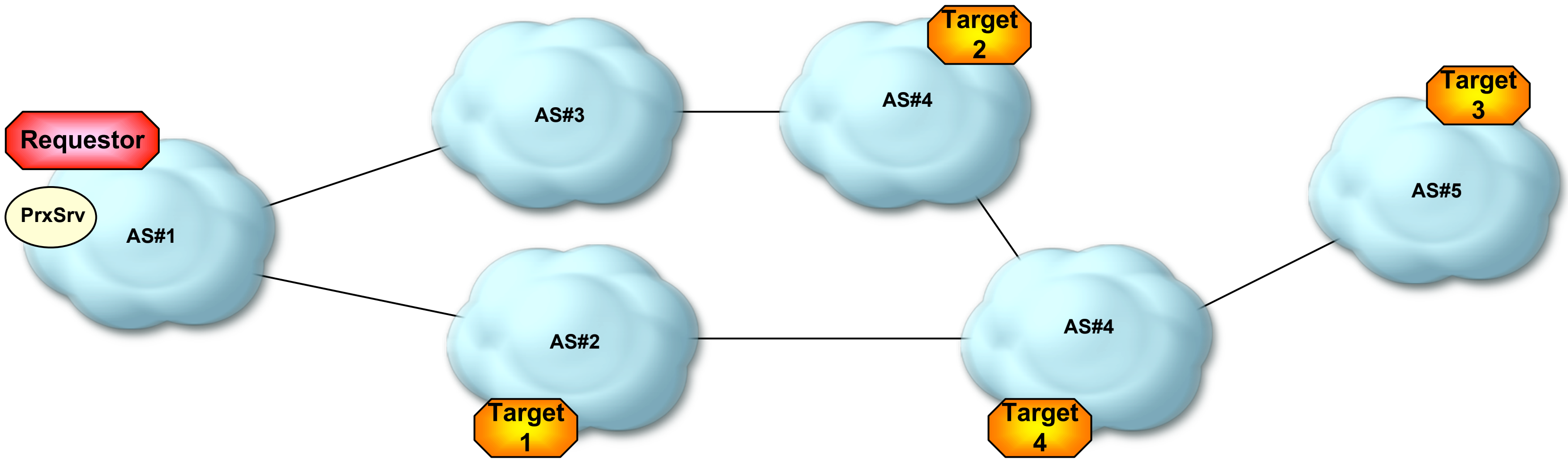
- Improve current schemes
  - Both at SP and application level
- Take into consideration different scenarios
  - Example: Content pre-positioning Vs. content download
- Preserve confidentiality/security across layers
  - Do not disclose any routing/topology information
  - Do not disclose any application specific information
- Application agnostic
  - Proximity is not specific to any given application
- Ensure interoperability
  - Signaling to be standardized

Cisco Proximity



- Example:
  - Client found content replicated in different peer/servers
  - Client sends Proximity Request to Proximity Server, get reply and start selection process taking into account Proximity Rankings
- Proximity service returns a ranked list

Cisco Proximity - Example



- Improve selection mechanism in order to take into account the AS's traversed by the traffic

  Proximity component to deliver best ranking

- Ranking based on Inter-AS routing scheme

  Including deployed BGP policies

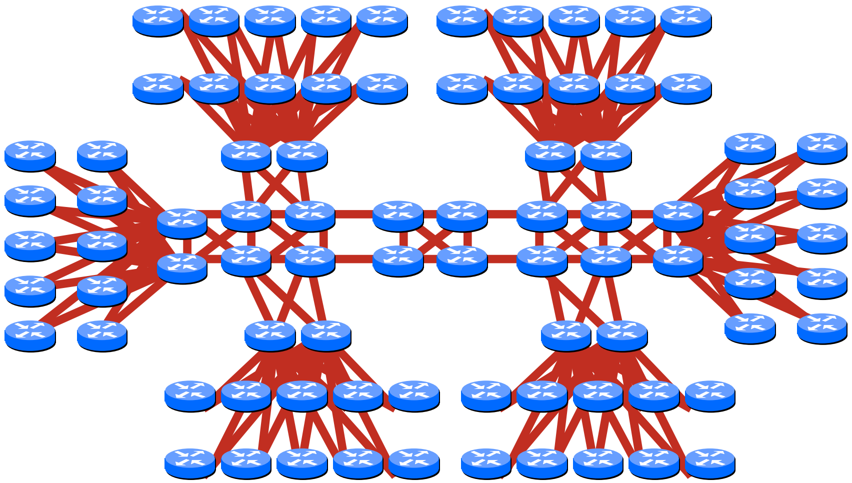
Proximity and network topology visibility

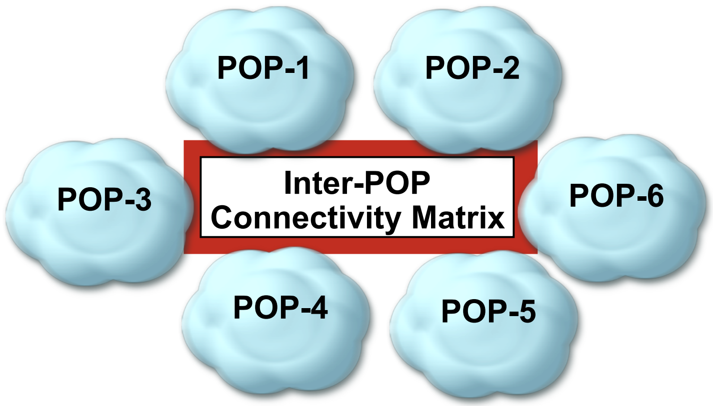
- Routing Layer to deliver localization hint
- IGP and BGP databases have most of needed info
  - Topology information about: AS, PoP, Prefix
  - Transit information: link and node traffic will traverse from/to a given address
  - Traffic direction
  - Route preferences
  - Policies (BGP)
- Confidentiality between layer is preserved
  - Req/Res based protocol between layer entities
  - Application IDs are not disclosed
  - Topology is not disclosed

Proximity and network topology visibility

- Proximity computation to aggregate topology information
- Example: aggregate all network topology into a Inter-POP connectivity scheme
  - Intra-POP topology may have limited usefulness
  - Most important is impact each POP may have on backbone/transit core
- Current deployments allows to correlate BGP communities to POPs or regions

Proximity and network topology visibility





- Proximity Server is part of routing layer

  Collects routing databases

- For Proximity efficiency, performance and scalability topology is aggregated in order to cope with application requirements

  More interested into coarse granularity

Routing Proximity - Ranking System

- Ranking based system - Preference list
- Ranking criteria may vary
  - Expressed in the Proximity Request
- Example: Proximity Client requests ranking on given IP (host) addresses
- Extendible to other routing layer attributes:
  - Prefixes, AS numbers, Communities, ...

Routing Proximity - Ranking System

- Routing computation requirements for localization are not the same as normal routing calculation:
  - Routing: shortest path to destination computed at source
  - Application: shortest path from source computed at destination
- Algorithms are similar but not the same
  - Forward-SPF, Reverse-SPF, MultiRoot-SPF, Modified BGP PS, ...
  - Combination of all above...
  - And Maybe More...
- Scaling implementation through pre-computation of coarse topology

Routing Proximity - Ranking System

- Confidentiality
  - Application request ranking (e.g.: IP addresses)
  - Does not deliver application internal details
  - Proximity Server delivers ranked list of IP addresses
  - Service Provider needs no to disclose any topology or policy information

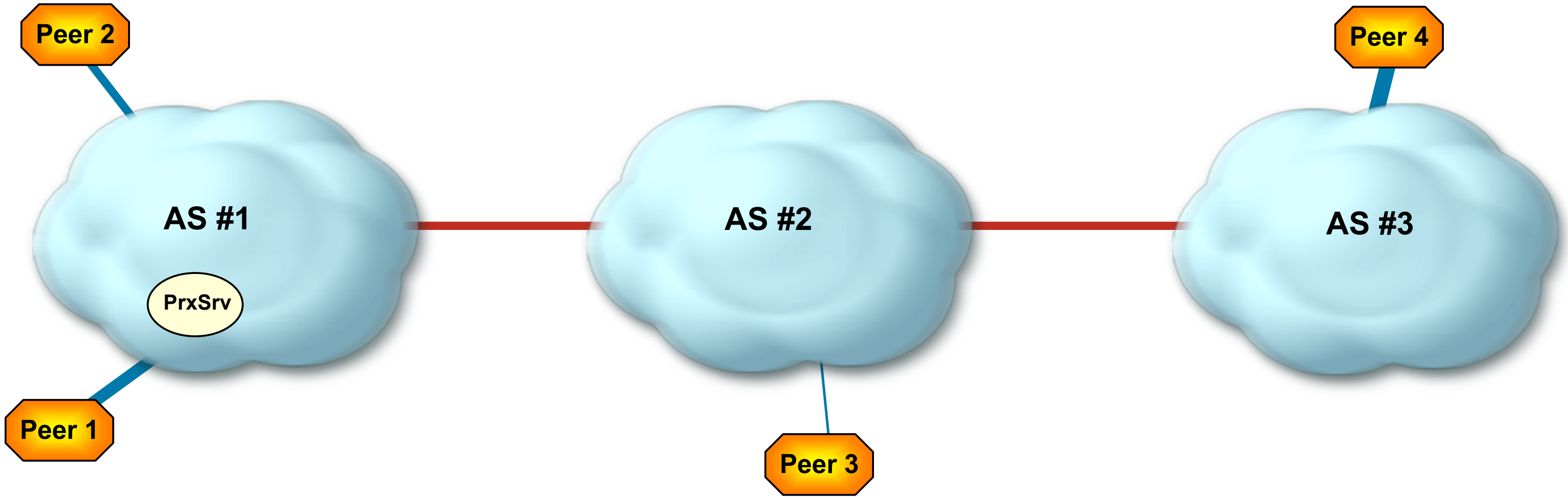
Proximity Service Advertisement

- Proximity Service Discovery
  - Need to propagate Proximity Service address
- Re-use existing well known methods
  - DNS
  - Anycast
  - Application specific
  - ...
- Not a big deal...

Routing Proximity - Signalling

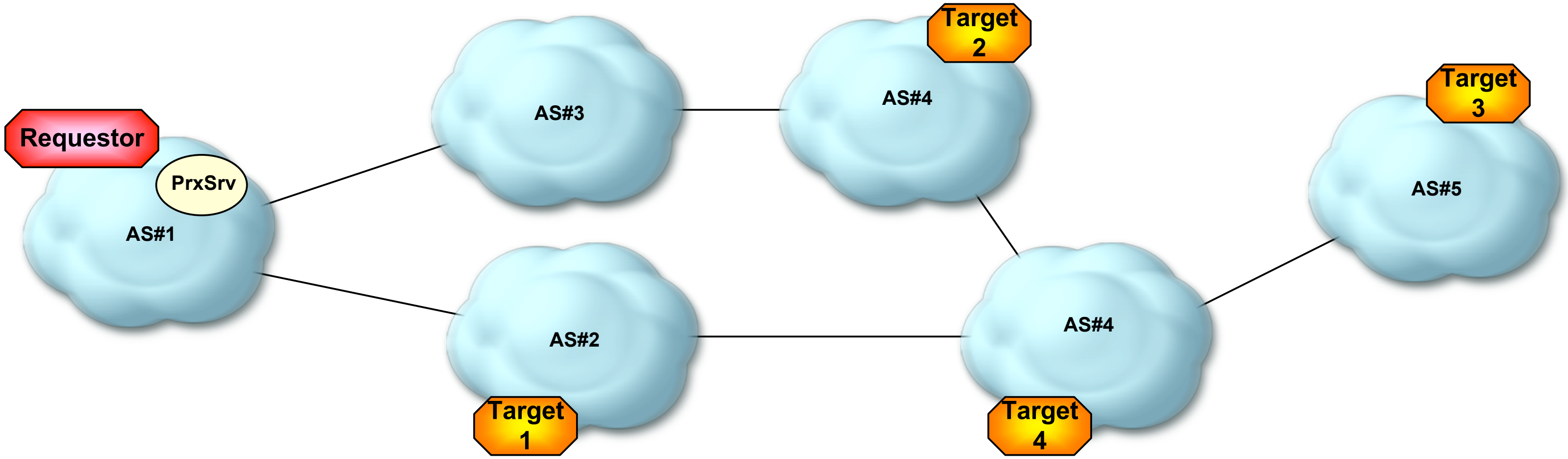
- Once known to the application (Proximity Consumer) request/response protocol is used
  - Between Proximity clients and server
- Need to formalize/standardize protocol
  - Again: KISS principle
  - Protocol must scale to thousands requests per second
- Routing Proximity is to be based on
  - IP addresses, Prefixes, AS numbers, BGP Communities, ... any routing information type
- Protocol extensibility and backward compatibility
  - Protocol objects and signaling messages will probably ‘grow’ over time

Example



- Example: Peer1 found 4 copies of a given movie residing in Peer 2, 3 and 4. Which one to select ?
  - Application-wise peer4 is preferred due to its uplink capacity
  - Topology-wise peer2 is preferred because residing in same AS
- Requestor sends a Proximity-Req to Proximity Server with:
  - Requestor IP address, list of targets IP addresses
- Proximity component runs ranking algorithm leveraging topology databases
  - From routing protocols as deployed in the backbone
- If policy is to avoid transit traffic, Proximity returns ranked list:
  - peer2, peer3, peer4

Example



- Inter-AS example
- Proximity computation to inspect AS connectivity
- What is the best proximity strategy ?

  Strict shortest AS-PATH ?

  Evaluate peering agreements and preference rules ?

  Monetary cost (in exchange points) ?

  Proximity policy/strategy is user defined

    Partially represented in BGP database

Routing Proximity Architecture - Summary

- Proximity Servers
  - Network infrastructure, Routing Layer
  - Simple interface with consumer (request/reply messages)
  - Data, Algorithms and Topologies remain inside infrastructure
- Proximity Consumers
  - Application clients, video/content distribution engines, cache engines, peer-to-peer clients
  - Commercial product of Content Distribution already integrate Proximity API
  - Application agnostic
- Messaging
  - Consumer ---> Provider: Proximity Request (list of IP addresses)
  - Provider ---> Consumer: Proximity Reply (ranked list of IP addresses)

Routing Proximity Architecture - Summary

- Proximity Servers: different implementation options
- Dedicated Proximity Server
  - Could implement additional components such as application routing protocols and inter-operate with application level
  - Require mechanisms to collect routing databases
- Embedded in router platform
  - Always available everywhere...
  - Embedded routing scheme
- Current implementation uses a dedicated server
  - Service Gateway
- Current implementation is done in the framework/context of Service Routing

Agenda

- Introduction and motivation
- Cisco Proximity
- Standardization & Research

IETF/IRTF

- Localization issues are being addressed by
  - IETF Working Group: ALTO (Application Layer Traffic Optimization)
  - IRTF Research Group: P2PRG
- IETF scope is to standardize the necessary part of the technology that will allow interoperability and consistent deployments
- IRTF scope is to conduct research activity in the peer to peer networking space including localization technologies

Context - IETF

- From draft-marocco-alto-problem-statement-02

  A significant part of the Internet traffic today is generated by peer-to-peer applications used for file sharing, real-time communications and live media streaming.

  Such applications often deal with large amounts of data in direct peer-to-peer connections, but they usually have little knowledge of the underlying topology, both at the overlay layer and the network layer.

  As a result, they may choose their peers based on measurements and statistics which, in some specific situations, often lead to suboptimal choices.

IETF ALTO Requirements

- Improve application level performance/experience
- Optimize/rationalize resources utilization in SP infrastructure
  - Example: reduce impact of peer to peer application
- Ensure security and confidentiality of layer-specific information
- Allow different “flavors” of localization techniques
- Standardize what’s needed to be standardized
  - Keep implementation-specific aspects outside scope

Components and Aspects

- Different technical aspects
  - Core infrastructure
  - Transit (inter-SP) traffic
  - Edge congestion
  - Uplink capacity
- Other non-technical aspects
  - Confidentiality: SP doesn't want to publish any form of his topology
  - Applications: reluctant to disclose sensitive information such as content and/or user identifiers (in any form)
  - Inter-working between layers

Components and Aspects

- Different components requires different solutions
  - Service Routing Vs. Traditional CDN Vs. Tracker-based p2p Vs. DHT-based p2p, ...
- KISS methodology
  - Keep It Simple Straightforward
- Address what can be solved now and at limited cost

ALTO Protocol

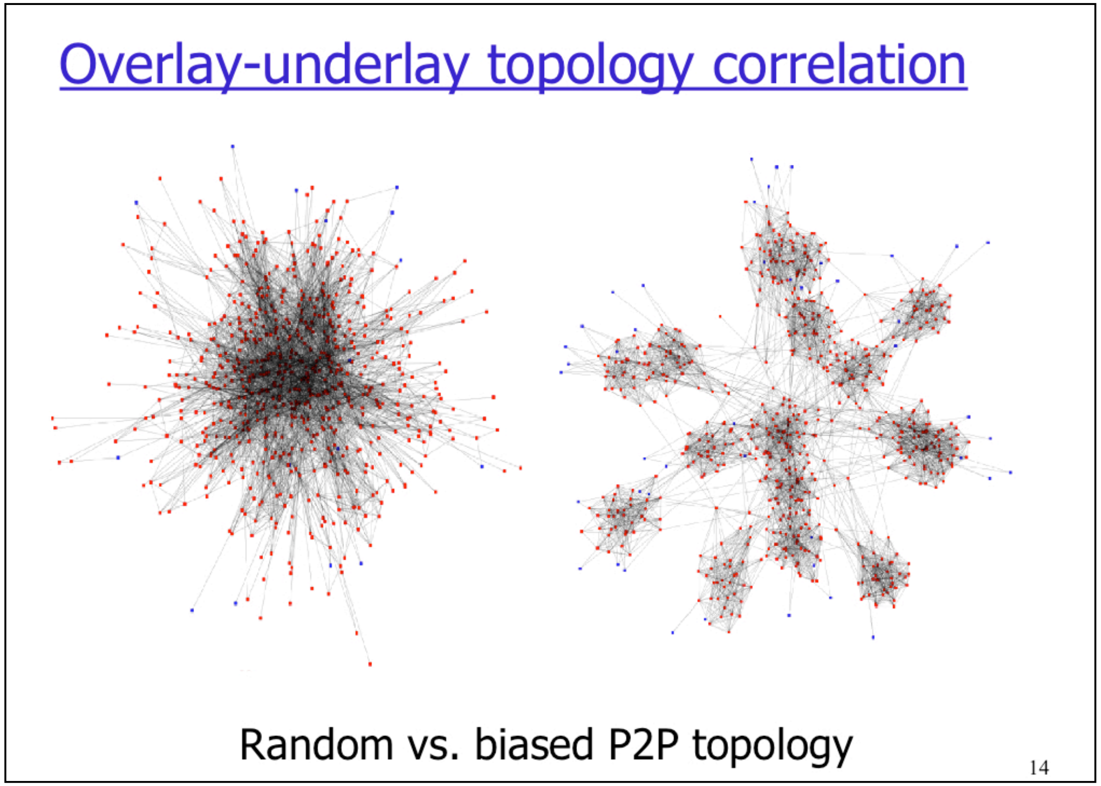
- ALTO to standardize a client/server protocol allowing application entities to request localization (proximity) services to an ALTO Service entity
- Request/Response model fits well. Example:

  ALTO Request: I’m IP4 and want to know among IP1, IP2, IP3 who is closer to me

  ALTO Response: IP3, IP1, IP2 (in order of smaller distance to IP4)
- Cisco Proximity architecture and implementation fits the model

Proximity/Localization: does it help ?

- Current peer selection mechanisms are close to random-based



- Source:

    Vinay Aggarwal, Anja Feldmann, Christian Scheideler.
    Can ISPs and P2P systems co-operate for improved performance? ACM SIGCOMM Computer Communications Review, Volume 37, Number 3, July 2007.