the gridbus project

THE UNIVERSITY OF MELBOURNE

Content Delivery Networks: Overlay Networks for Scaling and Enhancing in the Web

Mukaddim Pathan and Rajkumar Buyya

Grid Computing and Distributed Systems (GRIDS) Laboratory
Dept. of Computer Science and Software Engineering
The University of Melbourne, Australia
www.gridbus.org

Tutorial@ADCOM 2008, Chennai, India

The University of Melbourne

- The City of Melbourne
  - World's most livable city
  - Cultural diversity is the key essence
  - Greater Melbourne population is 3.5 million
  - Attractive tourist attractions
  - The place to be!
- The University of Melbourne
  - Founded in 1853
  - Australia's top-ranked university in teaching and research

Aerial view of the University of Melbourne

GRIDS Lab @ Melbourne University

R & D

- Youngest and one of the rapidly growing research labs in our School/University:
  - Founded in 2002
  - Houses (20+) consisting of:
    - Research Fellows/PostDocs
    - Research Programmers
    - PhD candidates
    - Honours/Masters students
- Funding
  - National and International organizations
  - Australian Research Council & DEST
  - Many industries (Sun, StorageTek, Microsoft, IBM, Microsoft)
- University-wide collaboration:
  - Faculties of Science, Engineering, and Medicine
- Many national and international collaborations.
  - Academics
  - Industries
- Software:
  - Widely in academic and industrial users.
- Publication:
  - Gridbus research team produces over 20% of our Dept's research output.

Education

+ Community Services: e.g., IEEE TC for Scalable Computing

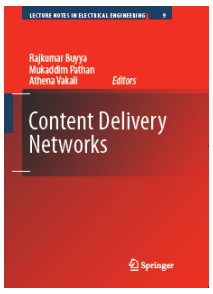
Knowledge Transfer Video Showcase

The Grids Lab and Next-Generation Grid Technologies

Available in YouTube !

Outline

- Part I: CDN Fundamentals
  - CDN Insights
  - CDN in Practice – Akamai Case Study
  - CDN Taxonomy
- Part II: CDN Modeling and Performance
  - CDN Pricing
  - CDN Performance
- Part III: Advanced CDN Platforms and Applications
  - Mobile Dynamic CDNs
  - Content Delivery for Community Networks
  - CDN Internetworking

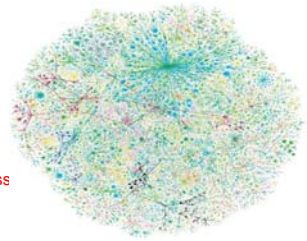


Part I

Content Delivery Networks: State of the art, insights and imperatives

Internet Dynamics

- With the rapid growth of Internet and Web
  - Services are competing each other for finite network and computing resources
  - High availability and responsiveness are keys to business Web sites
  - Large number of users are trying to simultaneously access the same Web site, causing "Flash Crowd"



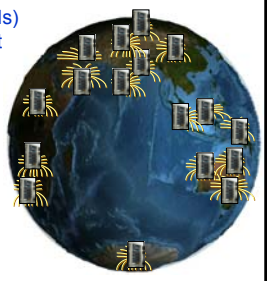
Centralized Web!

- Slow
  - content must traverse multiple backbones and long distances
- Unreliable
  - delivery may be prevented by congestion or backbone peering problems
- Not scalable
  - usage limited by bandwidth available at master site
- Inferior streaming quality
  - packet loss, congestion, and narrow pipes degrade stream quality

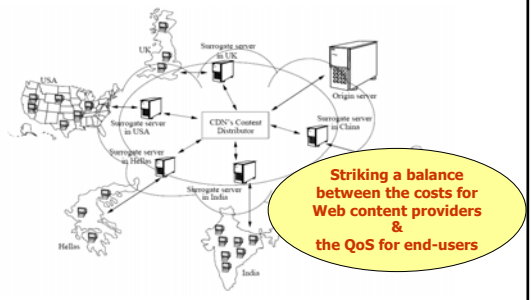
Source: Bruce Maggs, CCGrid 2001 Keynote

Content Delivery Network (CDN)

- Content Delivery Networks (CDNs) emerged as a solution to Internet service degradation
  - Moving content to the "edge" of the Internet, close to end-users
- Alternatives
  - Increased bandwidth, Web caching, Web pre-fetching
- CDN advantages
  - Reduced server loads
  - Distributed network traffic
  - Reduced latency

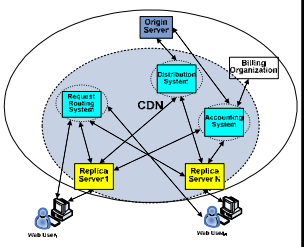


Abstract View



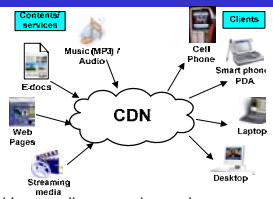
CDN Architecture

- Content delivery component
  - Origin server and a set of edge servers (surrogates) to replicate content
- Request-routing component
  - Direct user requests to edge servers
  - Interact with the distribution component to keep an up-to-date view of content
- Content distribution component
  - Moves content from the origin to edge servers and ensures consistency
- Accounting component
  - Maintains logs of client accesses and records usage of the servers
  - Assists in traffic reporting and usage-based billing



CDN Supported Content/Services

- Static content
  - Static HTML pages, images, documents, software patches
- Streaming media
  - Audio, real-time video
- User Generated Video (UGV)
- Content services
  - Directory, e-commerce, file transfer services
- Sources of content
  - Large enterprises, Web service providers, media companies, and news broadcasters
- Customers
  - Media and Internet advertisement companies, data centers, ISPs, online music retailers, mobile operators, consumer electronics manufacturers, and other carrier companies
- User interaction
  - Cell phone, smart phone/PDA, laptop, and desktop



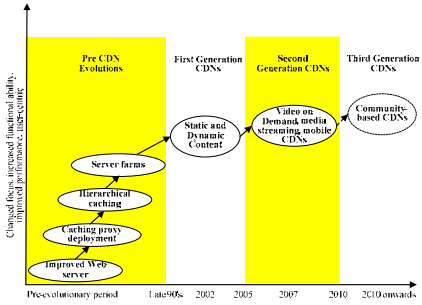
CDN Insights – Business Goals

- Scalability
  - Ability to expand in order to handle new and large amounts of data, users and transactions
  - Requires capabilities for dynamic provisioning and high quality content delivery, with low operational cost
  - Future trend: content providers as well as end-users will pay to get high quality content
- Security
  - Protection of content against unauthorized access and modification
  - Requires physical, network, software, data and procedural security
  - Future trend: reduce business interruption by combating against DDoS attacks and other malicious activities
- Reliability, Responsiveness and Performance
  - Service availability, handling possible outages and end-user experience
  - Requires a fault-tolerant network with appropriate load balancing
  - Future trend: distributed content location, cache consistency and routing mechanisms

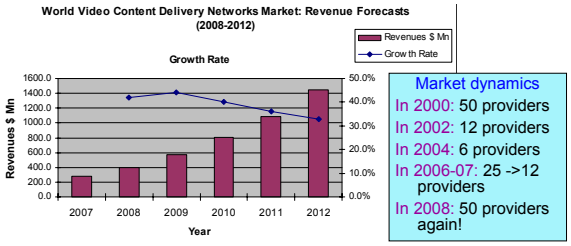
Comparison with Related Systems

Features	CDNs	Data Grids	Distributed Databases	P2P Networks
Category	A collection of networked computers spanning the Internet	Data intensive computing environment	Locally organized collection of data distributed across multiple physical locations	Information retrieval network formed by ad-hoc aggregation of resources
Constitution	Distribution of cache servers to the edge of the Internet	Formation of a VO of participating institutions	Federation or splitting of existing database(s)	Collaboration among peers
Main goal	Reducing Web latency during content delivery	Performance gain through data distribution by pre-staging, optimal source selection, and high speed data movement	Integration of existing databases and replication of database fragments in a transparent manner	File sharing among peers
Integrity	Integrity between caches	Integrity between data grid replicas	Integrity between multiple DBs	N/A
Consistency	Strong cache consistency between replicated content	Weak consistency between data grid replicas	Strong database consistency between distributed DBs	Weak consistency between cached content
Autonomy	None	Autonomous participants	Autonomous DDB sites	Autonomous peers
Operational Activities	Content caching	Seamless analysis, collaboration, and maintenance of data across organizational and regional Boundaries	Query processing, optimization, and Management	Locating or caching content, encrypting, retrieving, decrypting, and verifying content
Administration	Individual companies, proprietary in nature	Institutions who cooperate on some shared goals	Single authoritative entity	Self-interested end users/peers

CDN Evolutions



CDN Growth and Market Forecast



Market dynamics

In 2000: 50 providers
In 2002: 12 providers
In 2004: 6 providers
In 2006-07: 25 ->12 providers
In 2008: 50 providers again!

- Many CDN providers will not make it 24 months from now!!
- The market is expected to achieve revenues over a third of a billion in 2008 and cross the billion dollar mark in 2011

Source: Dan Rayburn, Streaming Media West, 20 May 2008

State of the Art – Commercial CDNs

and so on...

State of the Art – Commercial CDNs (Cont'd)

CDN Name	Service Type	Coverage	Products/Solutions
Akamai www.akamai.com	Provides CDN service, including streaming	Covers 85% of the market. 40,000 servers in 900 networks in 70 countries. It handles 20% of total Internet traffic today	Edge Platform for handling static as well as dynamic content, Edge Control for managing applications, and Network Operations Control Center (NOCC)
EdgeStream www.edgestream.com	Provides disrupted video streaming applications over the public Internet	Provides video streaming over consumer cable or ADSL modem connections around the globe, even over paths that have 20 router hops between server and end user	EdgeStream video on-demand and IPTV Streaming software for video streaming
Limelight Networks www.limelightnetworks.com	Provides distributed on-demand and live delivery of video, music, games and download	Edge servers located in 72 locations around the world	Limelight ContentEdge for distributed content delivery via HTTP, Limelight MediaEdge Streaming for distributed video and music delivery via streaming, and Limelight Custom CDN for custom distributed delivery solutions
Mirror Image www.mirror-image.com	Provides content delivery, streaming media, Web computing and reporting services	Edge servers located in 22 countries	Global Content Caching, Extensible Rules Engine (XRE), Video On-Demand, and Live Webcasting

State of the Art – Academic CDNs

CDN Name	Description	Service Type	Implementation & Testing	Availability
CoDeeN www.codeen.cs.princeton.edu	CoDeeN is an academic testbed CDN built on top of PlanetLab	Provides caching of content and redirection of HTTP Requests	Implemented in C/C++ and tested on Linux (2.4/2.6) and MacOS (10.2/10.3)	N/A
Coral www.coralcdn.org	Coral is a free P2P CDN. It is hosted on PlanetLab	Provides content replication in proportion to the content's popularity	Implemented in C++ and tested on Linux, OpenBSD, FreeBSD, and Mac OS X	No official release yet. Coral is a Free software, licensed under GPLv2 (GNU General Public License)
Globule www.globule.org	Globule is an open source collaborative CDN	Provides replication of content, monitoring of servers and redirecting client requests to available replicas	Implemented using PHP scripting, C/C++ and tested on Unix/Linux and Windows	Globule is open source, subject to a BSD-style license and the Apache software license for the packaged Apache HTTP server

CoDeeN CORAL

Emerging CDN Practices

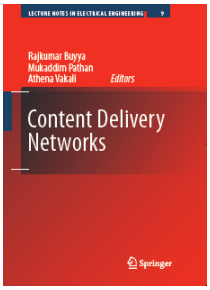
- Unified content network
  - Integration of "Content Delivery/Distribution" and "Content Services" domains
  - Content Service Network (CSN) as a "service" distribution channel for value added services
- Dynamic content delivery
  - On-demand content generation using Web applications based on end-user request specifications (e.g. scripts, animations, DHTML, XML)
  - Edge computing, context-aware data caching, content replication, content blind data caching, as well as proprietary solutions (e.g. EdgeSuite, IBM WebSphere edge services)
- Web services hosting
  - Use of XML parsing, Java serialization, reflection copy, clone copy
  - Application Delivery Network (ADN) to host .NET and J2EE applications
  - Capacity Provisioning Network (CPN) for trading cache capacities
- Service-Oriented Architecture
  - Content management is expected to be motivated by user preferences
  - User personalization based on data mining

Research Directions

- Load balancing and content replication in cooperative domain
  - Request locality
  - Integration of replication and caching
- Deployment of market mechanisms
  - Economics models based on an SOA
- Adaptive CDN for media streaming
  - P2P approach for collaborative media streaming
- Mobile dynamic CDN
  - High variability in demand due to user mobility
- Content distribution through internetworking/peering/brokering
  - CDN cooperation for global coverage with high quality performance

Outline

- Part I: CDN Fundamentals
  - CDN Insights
  - CDN in Practice – Akamai Case Study
  - CDN Taxonomy
- Part II: CDN Modeling and Performance
  - CDN Pricing
  - CDN Performance
- Part III: Advanced CDN Platforms and Applications
  - Mobile Dynamic CDNs
  - Content Delivery for Community Networks
  - CDN Internetworking



What is Akamai?

- Research began at MIT in 1995
- Company founded in August 1998
- The largest CDN provider to date
  - 40000 servers in 70 countries
  - Handles 20% of today's Internet traffic !
- The market share leader (approx. 85%)
- Provides managed edge services to content providers
- Operates a global network of servers deployed at the edge of Internet
- Charged with highly complex, proprietary mathematical algorithms and patented technologies
- Partner with third-party application developers via open APIs

Content Delivery – The Akamai Way

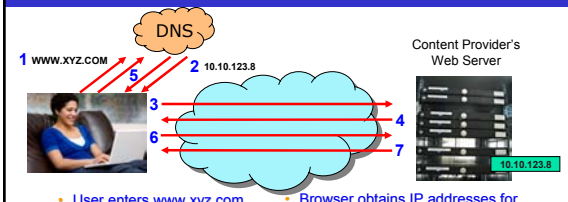
- What is the technology behind Akamai?
- How Akamai has improved the content delivery performance in the Internet?
- What are the design principles?
- How does it handle failures within its world-wide distributed network?

HTML Title Page for www.xyz.com with Embedded Objects

```
<html>
<head>
<title>Welcome to xyz.com!</title>
</head>
<body>
<img src="http://www.xyz.com/jpgs/navbar1.jpg">
<img src="http://www.xyz.com/jpgs/navbar1.jpg">
<h1>Welcome to our Web site!</h1>
<a href="page2.html">Click here to enter</a>
</body>
</html>
```

Source: Bruce Maggs, Akamai Technologies

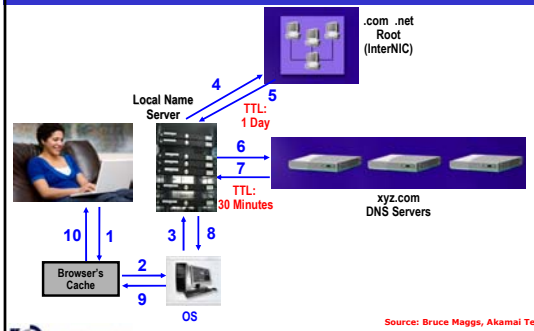
Downloading www.xyz.com - before Akamai



- User enters www.xyz.com
- Browser requests IP address for www.xyz.com
- DNS returns IP address
- Browser requests HTML
- Content provider's web server returns HTML
- Browser obtains IP addresses for hostnames listed in URLs of objects embedded on page
- Browser requests embedded objects
- Content provider's web server returns embedded objects

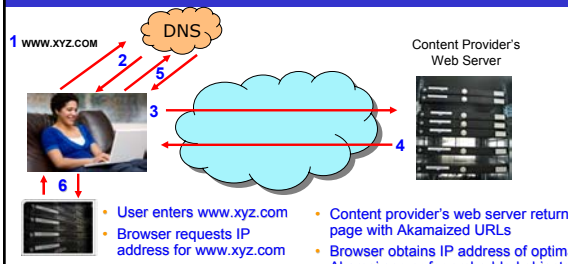
Source: Bruce Maggs, Akamai Technologies

DNS Resolution



Source: Bruce Maggs, Akamai Technologies

Downloading www.xyz.com - The Akamai Way



- User enters www.xyz.com
- Browser requests IP address for www.xyz.com
- DNS returns IP address
- Browser requests HTML
- Content provider's web server returns page with Akamaized URLs
- Browser obtains IP address of optimal Akamai server for embedded objects
- Browser obtains objects from optimal Akamai server

Source: Bruce Maggs, Akamai Technologies

Content Delivery Using Akamai

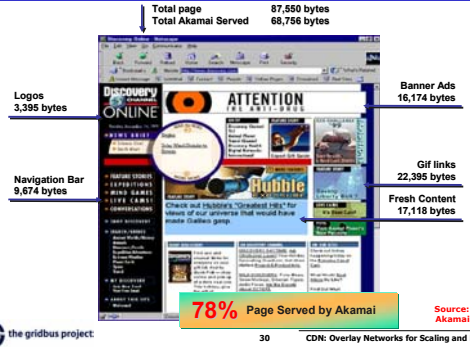
Embedded URLs are Converted to ARLs

```
<html>
<head>
<title>Welcome to xyz.com!</title>
</head>
<body>
<img src=" http://www.xyz.com/logos/logo.gif">
<img src=" http://www.xyz.com/jpgs/navbar1.jpg">
<h1>Welcome to our Web site!</h1>
<a href="page2.html">Click here to enter</a>
</body>
</html>
```

ak

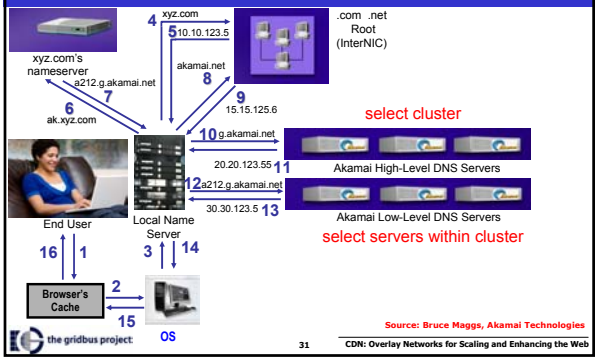
Source: Bruce Maggs, Akamai Technologies

Typical Page Content



Source: Bruce Maggs, Akamai Technologies

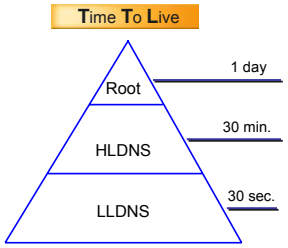
Akamai DNS Resolution



Source: Bruce Maggs, Akamai Technologies

DNS Maps and Time-To-Live

- Map creation is based on measurements of:
  - Internet congestion
  - System loads
  - User demands
  - Server status
- Maps are constantly recalculated:
  - Every few minutes for HLDNS
  - Every few seconds for LLDNS



TTL of DNS responses gets shorter further down the hierarchy

Source: Bruce Maggs, Akamai Technologies

Engineering Methodology

- C programming language (gcc)
- Reliance on open-source code
- Automated unit and system builds and tests
- Staged rollout to production
- Independent release management
- Burn-in on “invisible” system

Source: Bruce Maggs, Akamai Technologies

Design Principles

- No single point of failure
- Minimal human intervention
- Decentralized organization
- Fail-over at multiple scales, redundancy
- Sophisticated algorithms
- Multiple, disjoint, reporting systems
- Backwards compatibility
- Secure and authenticated communications

Source: Bruce Maggs, Akamai Technologies

Varieties of Failures

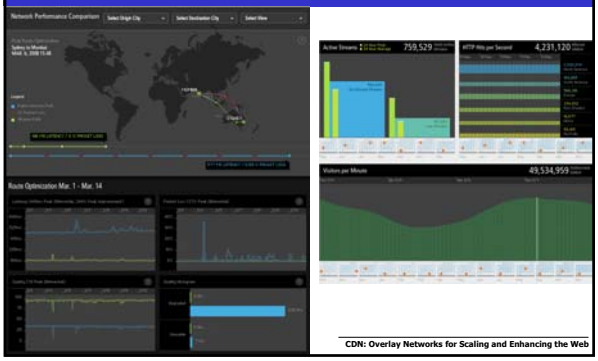
- Network
  - Congestion at public and private peering points
  - Misconfigured routers and switches
  - Inaccessible networks
- Hardware
  - Memory SIMMS jumping out of their sockets
  - Network cards screwed down but not in slot
  - Switches configured to drop broadcasts
- Software
  - Third-party software
  - Problems with streaming clients/servers
- Misperceptions
  - Personal firewalls
  - Reporting Tools
  - Customer-side problems
  - Third-party measurements
- Attacks
  - Distributed Denial of Services (DDoS)

Source: Bruce Maggs, Akamai Technologies

Real Time Web Monitor

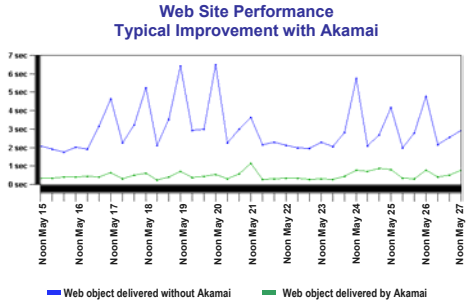
Network Traffic | Client Perceived Latency | Attacks

Network Performance and Akamai Visualization



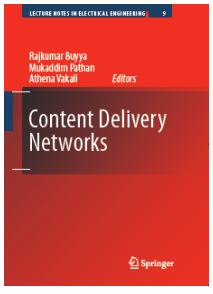
Network Metrics

Keynote™ Results

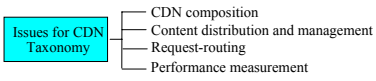


Outline

- Part I: CDN Fundamentals
  - CDN Insights
  - CDN in Practice – Akamai Case Study
  - CDN Taxonomy
- Part II: CDN Modeling and Performance
  - CDN Pricing
  - CDN Performance
- Part III: Advanced CDN Platforms and Applications
  - Mobile Dynamic CDNs
  - Content Delivery for Community Networks
  - CDN Internetworking

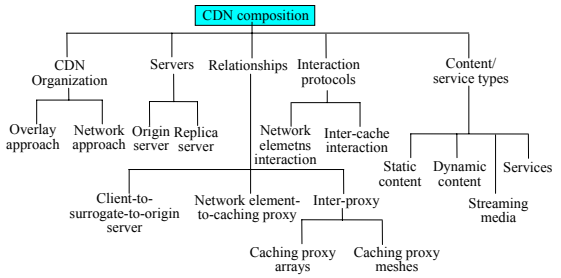


CDN Taxonomy



- Core issues
  - What is required for a harmonious CDN composition?
  - How to perform effective content distribution and management?
  - How to route client requests to appropriate CDN node?
  - How to measure a CDN's performance?
- Additional issues
  - How to handle wide-area failures in a CDN?
  - How to ensure security in a wide-area CDN system?
  - How to achieve wide-area application hosting?

CDN Composition

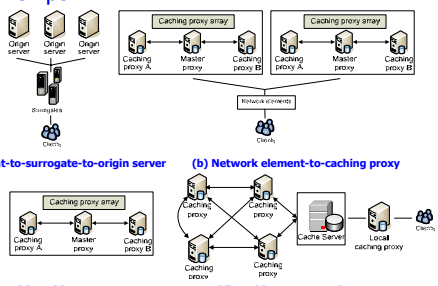


CDN Composition (Cont'd)

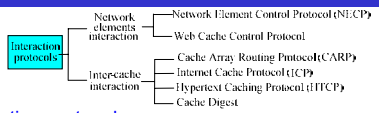
- CDN organization
  - Overlay approach
    - Application-specific servers and caches at several places in the network to handle content delivery
    - Network components play no active role for content delivery
    - Used by most commercial CDN provider
  - Network approach
    - Network components are equipped with code for identifying specific application types and for request forwarding based on predefined policies
    - Used by some CDNs (e.g. Akamai), in addition to the overlay approach
- Servers
  - Origin server
  - Replica server

CDN Composition (Cont'd)

- Relationships



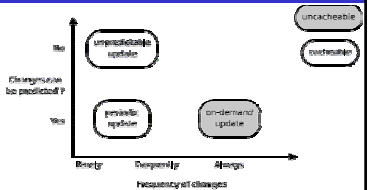
CDN Composition (Cont'd)



- Interaction protocols
  - **NECP:** Lightweight protocol for signaling between servers and network elements for load balancing
  - **WCCP:** Specifies interaction between one or more routers and one or more Web-caches
  - **CARP:** Distributed caching protocol
  - **ICP:** Lightweight message format used for inter-cache communication
  - **HTCP:** Protocol for discovering HTTP caches, cached data, managing sets of HTTP caches and monitoring cache activity
  - **Cache Digest:** Exchange protocol and data format

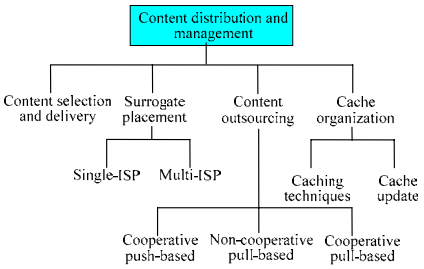
CDN Composition (Cont'd)

- Content/service types
  - **Static content:** Frequency of change is low (e.g. static HTML pages, embedded images, executables, PDF documents, software patches, audio and/or video files)



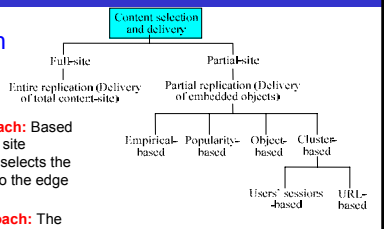
  - **Dynamic content:** Personalized for user or created on-demand (e.g. animations, scripts, and DHTML)
  - **Streaming media:** Live or on-demand telecast (e.g. audio and/or video on-demand, movie files and music clips)
  - **Content services:** Internet infrastructure services (e.g. directory, Web storage, file transfer, and e-commerce services)

Content Distribution and Management



Content Distribution and Management (Cont'd)

- Content selection and delivery



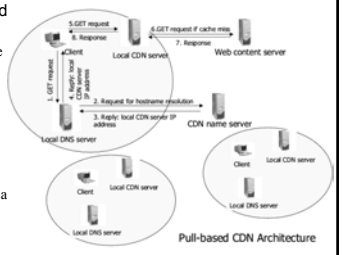
  - **Empirical-based approach:** Based on some heuristics, Web site administrator empirically selects the content to be replicated to the edge servers
  - **Popularity-based approach:** The most popular objects are replicated to the surrogates
  - **Object-based approach:** Greedy approach to replicate content to the surrogate servers in units of objects
  - **Cluster-based approach:** Web content is grouped based on either correlation or access frequency and is replicated in units of content clusters

Content Distribution and Management (Cont'd)

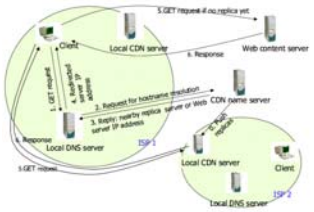
- Surrogate placement strategies
  - **Center placement problem:** For the placement of a given number of centers, minimize the maximum distance between a node and the nearest center
  - **Greedy and topology-informed placement:** suboptimal algorithms take into account the existing information from CDN, such as workload patterns and the network topology
  - **Hotspot replica placement:** Places replicas near the clients generating greatest load
  - **Tree-based replica placement:** Based on the assumption that the underlying topologies are trees
  - **Scalable replica placement:** Generates replicas on demand and organizes them into an application-level multicast tree
- Optimal placement
  - Single-ISP and multi-ISP approach

Content Distribution and Management (Cont'd)

- Content outsourcing
  - Non-cooperative pull-based
    - Upon cache miss, surrogate servers pull content from the origin server
    - Used by most CDN providers
  - Cooperative pull-based
    - Surrogate servers cooperate with each other to get the requested content in case of a cache miss
    - Only used by Coral CDN, making use of DHTs

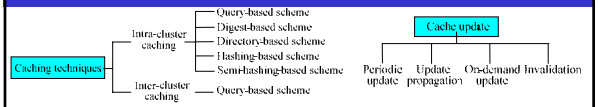


Content Distribution and Management (Cont'd)



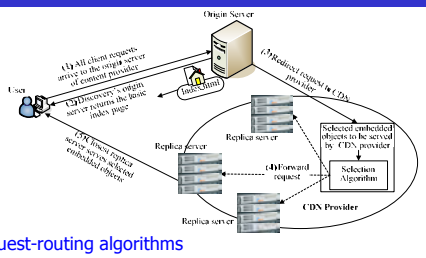
- Content outsourcing (Cont'd)
  - Cooperative push-based
    - Pre-fetching of content to the surrogates
    - Content is pushed to the surrogate servers from the origin, and surrogate servers cooperate to reduce replication and update cost

Content Distribution and Management (Cont'd)



- Cache organization and management
  - Caching techniques
    - **Query-based:** Query broadcast to other CDN servers
    - **Digest-based:** Each CDN server maintains a digest of content held by the other cooperating surrogates
    - **Directory-based:** A centralized server keeps content information of all the cooperating surrogates inside a cluster
    - **Hashing-based:** A designated CDN server holds a content based on content's URL, IP addresses of the CDN servers, and the hashing function
    - **Semi-hashing-based:** A local CDN server allocates a certain portion of its disk space to cache the most popular content for its local users and the remaining portion to cooperate with other CDN servers via a hashing function
  - Cache Update

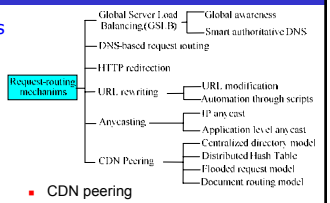
Request-Routing



- Request-routing algorithms
  - **Adaptive**
    - Consider the current system condition to select a cache server for content delivery
  - **Non-adaptive**
    - Use of heuristics to select a cache server rather than to consider current system condition

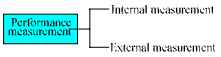
Request-Routing (Cont'd)

- Request-routing mechanisms
  - GSLB
    - Global awareness and smart authoritative DNS
  - DNS-dispatching
    - Modified DNS server
  - HTTP redirection
    - Information of replica server sets in HTTP headers
  - URL rewriting
    - Rewrite dynamically generated pages' URL links
  - Anycasting
    - IP Anycasting and Application-level anycasting



- CDN peering
  - Centralized directory model: Centralized directory to publish content
  - Distributed hash table: Indexing through hash keys
  - Flooded request model: Broadcast queries
  - Document routing model: Authoritative peer for referral

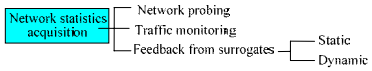
Performance Measurement



- Performance evaluation metrics
  - Cache hit ratio
    - Ratio of the number of cached content versus total content requested
  - Reserved bandwidth
    - Measure of the bandwidth used by the origin server
  - Latency
    - User perceived response time
  - Surrogate server utilization
    - Fraction of time during which the surrogate servers remain busy
  - Reliability
    - Packet-loss measurements

Performance Measurement (Cont'd)

- Network statistics acquisition



- Measurement-based study and simulations
  - Real-time testbed
    - PlanetLab
  - Simulators
    - CDNSim, CSIM, NS-2, OMNET++ and so on…

CDN Composition Taxonomy Mapping

	CDN Name and Type	CDN Organization	Servers	Relationships	Interaction Protocols	Content/Service Types
Commercial CDNs	Akamai	Network and overlay approach	Origin and replica servers	Client-to-surrogate-to-origin server, Network element-to-caching proxy, Inter-proxy	Network elements interaction, inter-cache interaction	Static content, dynamic content, streaming media, and services (network monitoring, geographic targeting)
	Edge Stream	Network approach	N/A	N/A	Network elements interaction	Video streaming, video hosting services
	Limelight Networks	Overlay approach	Origin and replica servers	Client-to-surrogate-to-origin server, Network element-to-caching proxy	Network elements interaction	Static content, streaming media
	Mirror Image	Network and Overlay approach	Origin and replica servers	Client-to-surrogate-to-origin server, Network element-to-caching proxy	Network elements interaction	Static content, streaming media, Web computing and reporting services

CDN Composition Taxonomy Mapping (Cont'd)

	CDN Name and Type	CDN Organization	Servers	Relationships	Interaction Protocols	Content/Service Types
Academic CDNs	CoDeeN	Overlay approach with "open" proxies	Origin and replica/ proxy (forward, reverse, redirector) Servers	Client-to-surrogate-to-origin server, Network element-to-caching proxy, inter-proxy	Network elements interaction, inter-cache interaction	*Participating* users receive better performance to *most* sites; only provides static content
	Coral	Overlay approach with an underlying indexing infrastructure	Origin and replica/ (cooperative) proxy cache servers	Client-to-surrogate-to-origin server, Network element-to-caching proxy, inter-proxy	Network elements interaction, inter-cache Interaction	Most *users* receive better performance to *participating* sites; only provides static content
	Globule	Overlay approach with end user nodes	Origin, replica, backup and/or redirector servers	Client-to-surrogate-to-origin server, Network element-to-caching proxy, inter-node	Network elements interaction, inter-cache interaction	A Web site's performance and availability is improved; provides static content and monitoring services

Content Distribution and Management Taxonomy Mapping

CDN Name	Content Selection and Delivery	Surrogate Placement	Content Outsourcing	Cache Organization
Akamai	Content selection • Full and partial-site delivery Content Clustering • Users' sessions based	Multi-ISP approach; Hotspot placement by allocating more servers to sites experiencing high load	Non-cooperative pull-based	Caching technique • Intra and inter-cluster caching Cache update • Update propagation • On-demand
Edge Stream	Content selection • Partial-site delivery Content Clustering N/A	Single-ISP approach	Non-cooperative pull-based	Caching technique • Inter-cluster caching Cache update N/A
Limelight Networks	Content selection • Partial-site delivery Content Clustering N/A	Multi-ISP approach	Non-cooperative pull-based	Caching technique • Intra-cluster caching Cache update • On-demand
Mirror Image	Content selection • Partial-site delivery Content Clustering • URL based	Multi-ISP approach; Center placement following a concentrated "Superstore" architecture	Non-cooperative pull-based	Caching technique • Intra-cluster caching Cache update • On-demand
CoDeeN	Content selection • Partial-site delivery Content Clustering N/A	Multi-ISP approach; Topology-informed replica placement	Cooperative pull-based	Caching technique • Intra and inter-cluster caching Cache update • On-demand
Coral	Content selection • Full and partial-site delivery Content Clustering • Users' sessions based	Multi-ISP approach; Tree-based replica placement	Cooperative pull-based	Caching technique • Intra and inter-cluster caching Cache update • Cache invalidation
Globule	Content selection • Full and partial-site delivery Content Clustering – N/A	Single-ISP approach; Best replica placement strategy is dynamically selected through regular evaluation of different strategies	Cooperative pull-based	Caching technique • Intra and inter-cluster caching Cache update • Adaptive cache update

Request-Routing Taxonomy Mapping

CDN Name	Request-routing Technique
Akamai	• Adaptive request-routing algorithms which takes into account server load and various network metrics • Combination of DNS-based request-routing and URL rewriting
EdgeStream	HTTP redirection
Limelight Networks	DNS-based request-routing
Mirror Image	Global Server Load Balancing (GSLB) • Global awareness • Smart authoritative DNS
CoDeeN	• Request-routing algorithm takes into account request locality, system load, reliability, and proximity information. • HTTP redirection.
Coral	• Request-routing algorithms with improved locality by exploiting on-the-fly network measurement and storing topology hints • DNS-based request-routing
Globule	• Adaptive request-routing algorithms considering AS-based proximity • Single-tier DNS-based request-routing

Performance Measurement Taxonomy Mapping

CDN Name	Performance Measurement
Akamai	Internal measurement • Network probing • Traffic monitoring (proactive) External measurement • Performed by a third party (Giga Information group)
EdgeStream	Internal measurement • Traffic monitoring through Real Time Performance Monitoring Service (RPMS)
Limelight Networks	N/A
Mirror Image	Internal measurement • Network probing • Traffic monitoring and reporting
CoDeeN	Internal measurement • Local traffic and system monitoring
Coral	Internal measurement • Traffic monitoring • Liveness checking of a proxy via UDP RPC
Globule	Internal measurement • Traffic monitoring • Monitoring of server availability by the redirectors

Additional Issues for CDN Taxonomy

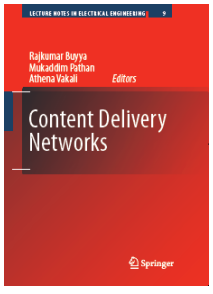
- Failure handling
  - Clustering, mirroring and multihoming
  - Distributed monitoring service
- Security
  - Intrusion detection, handling DDoS and protocol attacks
- Web application hosting
  - Usage-based content and application delivery
  - Akamai Edge Computing Infrastructure (ECI), active cache, ACDN

Part II

Content Delivery Networks:
Economic Models and Performance

Outline

- Part I: CDN Fundamentals
  - CDN Insights
  - CDN in Practice – Akamai Case Study
  - CDN Taxonomy
- Part II: CDN Modeling and Performance
  - CDN Pricing
  - CDN Performance
- Part III: Advanced CDN Platforms and Applications
  - Mobile Dynamic CDNs
  - Content Delivery for Community Networks
  - CDN Internetworking



Pricing Dynamics

- Content providers should continue to derive value and CDN providers simultaneously should have incentives to manage and deploy infrastructure
  - A market mechanism is needed to price the service
- Pricing based on aggregate usage
  - Simplest pricing structure which involves a content provider committing to a certain level of usage, e.g. $50TBs/month and a CDN determining a price per GB delivered based on this commitment
  - Volume/bulk discount is usual, $0.5/GB for a traffic commitment of $40-50TBs and $0.15/GB for commitment over 100TBs
  - Penalty if monthly commitment exceeds to induce commitment accuracy
- Percentile-based pricing
  - Periodic sampling of the bandwidth usage by a subscribing content provider, computing 95th percentile of usage at the end of the month and charging a price per Mbps based on it

Pricing Dynamics (Cont'd)

- Congestion pricing in networks
  - Focus is on the interaction between pricing and QoS by studying the trade-off between congestion cost and capacity cost
  - Pricing plays a key role to achieve desired QoS, when capacity can not be easily increased
    - Increase in pricing encourages users to shape their traffic and control the demand for network services, which in turn reduce network congestions
  - Congestion reduction is not a CDN's goal
    - Congestion pricing research do not readily transfer to CDN domain
- Economics of content delivery
  - Usage based pricing should entail volume discounts when subscribing content providers have similar traffic burstiness level
    - Volume discount can be sub-optimal for highly heterogeneous Bursty traffic
  - Profitability from a percentile-based pricing is substantially higher than usage-based billing

CDN Pricing Models

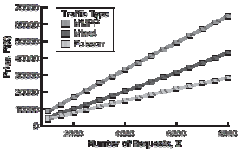
- Determining a CDN's optimal pricing policy
  - Determine a content provider's expected surplus from self-provisioning and that from delivering content through a CDN
    - Content provider chooses the option with higher expected surplus
  - Based on the content provider's subscription decision, determine the CDN's optimal pricing policy, which maximizes expected profit
- Self provisioning by content provider
  - Expected surplus, $U_{self} = V - C(I) - c*L(I)$
    - V is the benefit from responding to I requests/unit time
    - C(I) is the cost for maintaining infrastructure, which is concave in I because of economies of scale
    - c is the cost of each lost request and $L(I)$ is the number of lost requests
- Provisioning through a CDN
  - Expected surplus, $U_{CDN} = V + b(I)*\lambda - C_0 - P(X)$
    - b(N) is the benefit per request from faster delivery through a set of N CDN servers
    - C_0 is the cost of outsourcing content delivery
    - P(X) is the usage-based price the CDN changes the content provider

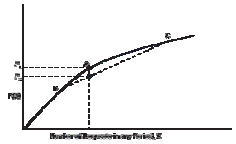
CDN Pricing Models (Cont'd)

- CDN's profit function, Π = CDN's expected revenue – CDN's cost
  - First term is the sum of revenues from all content providers and the second term is modeled to be quadratic over the man volume of traffic handled by CDN
  - A CDN does not know the outsourcing cost of individual content provider, but knows the cost distribution across content providers
- Simulation is used to compute optimal infrastructure, associated expected surplus under self-provisioning and CDN provisioning for a given CDN price function
- Optimal pricing for Poisson and Bursty traffic
  - 1000 content providers with arrival rate from a Pareto distribution [1000, 8000]
  - Poisson traffic: all content providers have Poisson traffic
  - Bursty traffic: all content providers have Bursty traffic
    - Modeled as a Markov Modulated Poisson Process (MMPP)
  - Mixed traffic: 500 CPs have Poisson and 500 CPs have MMPP traffic
  - Cost of serving 233 requests/min is $804 per month
  - Content provider's outsourcing cost is drawn from a Uniform [0, 30000]

Experimental Results

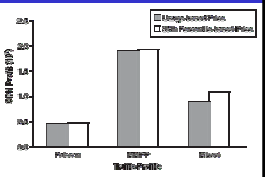
- Optimal prices are higher under Bursty traffic
  - Price (MMPP) > Price (Mixed) > Price (Poisson)
  - CDN's value proposition to content providers in terms of avoiding lost requests is enhanced in the presence of Bursty traffic
    - It leads to a price increase by a CDN
  - Optimal price under mixed traffic is convex and involves a volume tax rather than a volume discount
- Traditional usage-based pricing is inefficient under Bursty traffic
  - Penalize content providers with low burstiness or content providers with high volume





Experimental Results (Cont'd)

- Percentile-based pricing is more profitable for a CDN
  - Focus is on quadratic price function with simplify numerical computation
  - CDN profit is higher for mixed traffic
  - CDN profit is almost same for Poisson and MMPP traffic
    - Mean request rate is fixed
    - Mean-based pricing can be converted to a percentile-based pricing policy
  - With mixed traffic, a CDN is permitted to provide volume discounts and at the same time, charge a higher price with greater traffic burstiness
  - Drawbacks
    - Complicated billing and lack of standardization



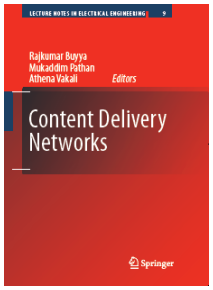
- Heterogeneous burstiness across content providers
  - Percentile-based pricing is preferable
- Usage-based pricing is usable for similar burstiness

Visionary Thoughts

- CDNs face a non-trivial pricing problem
  - Percentile-based pricing is more profitable, but usage-based pricing is preferential as the first is perceived as form of peak pricing
- It is expected that a transparent pricing plan coupled with competitive pressure will ultimately decrease the usage of percentile-based pricing in the next few years
- CDN pricing research has focused a monopolized market
  - Impact of competition on the pricing policies chosen by the CDNs could be investigated
  - Impact of P2P and hybrid CDNs on pricing are yet to be understood
- Another possibility is to provide content providers options to choose from usage-based or percentile-based pricing policies
  - It is not clear whether it is desirable
- CDNs require to devise a longer term solution
  - "Pay-as-you-go" as used in Cloud and utility computing, Storage Delivery Networks, e.g. Amazon S3 and EC2, could emerge in CDN domain

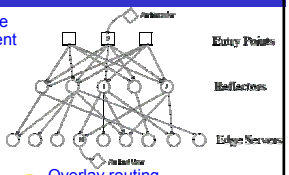
Outline

- Part I: CDN Fundamentals
  - CDN Insights
  - CDN in Practice – Akamai Case Study
  - CDN Taxonomy
- Part II: CDN Modeling and Performance
  - CDN Pricing
  - CDN Performance
- Part III: Advanced CDN Platforms and Applications
  - Mobile Dynamic CDNs
  - Content Delivery for Community Networks
  - CDN Internetworking



Transport Systems

- The types of content transported by the CDN system is varied and have different QoS requirements
  - Transporting live streaming content has different requirements as compared to transporting dynamic Web content
- Transport system for live streaming
  - Aim is to transmit live streams to enhance quality and minimize distortions
  - Optimization techniques
    - Sophisticated encoding for lose recovery
    - Pre-bursting for furst stream startup and pre-fetching for fast downloads
- Transport system for Web and online applications
  - Aim is to optimize end user response time
  - Non-http IP-based applications
    - Virtual Private Network (VPN), Voice-over-IP (VoIP)



- Overlay routing
  - Application-specific enhancement to meet end-to-end requirements
  - Aim to find a "better path" through the Internet between source and destination
  - Chooses the best overlay path (direct or indirect)
  - Enhances end-to-end availability and end user performance

Empirical Evaluation of Global Overlay Routing

- Aim is to study the performance and availability benefits of routing overlays on the commercial Internet
  - A global subset of Akamai CDN is used for data collection
  - Address the problem of picking optimum overlay paths between edge servers situated near end-users and origin servers situated in the Internet core
  - Investigate both performance characterized by round-trip latency as well as path availability
- Summary of outcomes
  - Randomly picking a small number of redundant paths (3 for Europe and North America, and 5 for Asia) achieves availability gains that approach the optimal
  - For reasonable probing intervals (10 mins) and redundancy (2 paths), over 90% of the source-destination pairs outside Asia have latency improvements within 10% of the ideal
    - Paths that originate or end in Asia require 3 paths to reach the same performance level
  - Overlay routing has high level of persistence over long periods of time
    - Indicating relatively infrequent network probing and measurements can provide optimal performance for almost all source-destination pairs

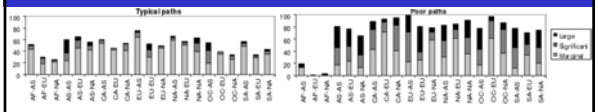
Experimental Setup

- Measurement platform
  - 1100 cluster locations distributed across many different kinds of ISPs in 77 countries, 630 cities and 6 continents
  - 15% of the clusters are located at the core, and the rest are at the edge
- Data collection
  - Each cluster run a task that sends ICMP echo request of size 64bytes in every 2 minutes to each node in the core set (rate: < 10req/sec)
  - 10 seconds timeout to deal no-response
  - Probing each path: 3780, total: 652 million
  - Modified standard all-pairs shortest-path algorithm is used
  - A path is unavailable if three or more consecutive pings are lost
  - Unavailability period: starting time for first lost ping to the ending time of last ping

Continent (Mnemonic)	Edge set	Core set
Africa (AF)	6	0
Asia (AS)	124	11
Central America (CA)	13	0
Europe (EU)	154	30
North America (NA)	624	110
Oceania (OC)	33	0
South America (SA)	38	0

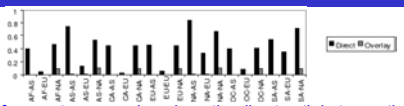
- Evaluation
  - Measurements from China are filtered out, since their failure characteristics are remarkably different from others due to the firewall policies by Chinese government

Latency Reduction



- Evaluation of what fraction of the poorly-connected source-destination pairs derive marginal, significant, or a large benefit from overlay routing
- Poorly-connected source-destination pairs see at least marginal benefits in over 80% of the samples
  - 67% of the samples see significant or large benefits
  - Some categories do deviate from the observation
    - E.g. poorly-connected source-destination pairs with destinations in Africa do not derive much help from an overlay

Availability Gains in Overlays



- Evaluation of percentage samples where the direct path between the source and destination failed for each category
- The failure percentage of the direct paths ranges from 0.03% to 0.83%
- Asia has the poorest availability, as nine of the ten categories with the largest failure percentage have an endpoint in Asia
- In the presence of overlay routing, the failure percentage goes down by 0.3% - 0.5% for most categories
  - Indirect path help mask failures of the direct path
- A small number of paths give rise to 30% of failures
  - 3% of the direct paths caused 30% of the failures, and 10% of the direct paths give rise to 50% of the failures
  - Reasons for the failures are specific chronic trouble rather than random, transient failures or short-lived congestions

Achieving Benefits in a Practical Design

- In practice, there is a bit time delay to construct overlay paths and only a small number of indirect paths could be considered
- Stability of optimal paths
  - The best paths between source to destination are quite persistent and do not change unless latency variations of all paths between them
  - Latency variations of the paths over time cause a significant reordering of the best paths between source and destination, causing reordering of optimal paths
- Performance gains of a predictive overlay
  - Overlays designed for high performance show reduced availability as compared to the ideal situation
  - This is because, better performing paths are typically constrained to share a small set of common links, leading to less path diversity and greater vulnerability
    - All shared links can simultaneously fail
- Persistence
  - There is a high degree of consistency in the relative performance of alternative paths between a source-destination pair, for most pairs
  - There is a small number of paths with high short term variations

Visionary Thoughts

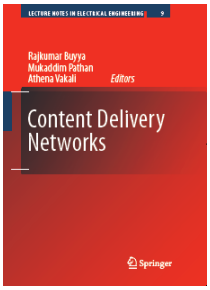
- Overlay routing optimizations become more and more prevalent
- Careful path selection is very important to build overlays for performance gains
- CDNs should provide even higher level of performance with little or no downtime
  - Due to recent e-commerce technology evolutions
- CDNs should deliver content in a scalable fashion even during flash crowds, without loss of availability or performance
- It is required to meet novel and more stringent availability and performance requirements to support the next-generation of Internet applications
  - Examples include applications such Virtual Private Networks (VPNs), Voice-over-IP (VoIP), which are highly latency sensitive

Part III

Content Delivery Networks:
Advanced Platforms and Applications

Outline

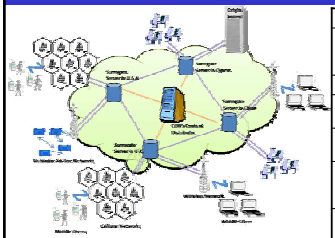
- Part I: CDN Fundamentals
  - CDN Insights
  - CDN in Practice – Akamai Case Study
  - CDN Taxonomy
- Part II: CDN Modeling and Performance
  - CDN Pricing
  - CDN Performance
- Part III: Advanced CDN Platforms and Applications
  - Mobile Dynamic CDNs
  - Content Delivery for Community Networks
  - CDN Internetworking



Mobile Environments and CDNs

- Mobile wireless networks offer a wide-range of dynamic and interactive services, such as GPS navigation information, mobile TV, vehicular traffic information, and location-oriented services
  - Represents a fundamentally different information medium than traditional Web
  - Techniques to efficiently disseminate content is required to reduce latency and to minimize network traffic
    - CDNs may offer a cost effective and scalable solution
- Mobile CDNs
  - Overlay network of surrogate servers to deliver content in mobile environments
- Content dissemination in mobile environments in an efficient and cost-effective manner is challenging
  - Mobile node constraints (limited storage, processing power, input capability) due to the portable size of mobile devices
  - Frequent network disconnections due to the navigational behavior of mobile users result in network fragmentation
  - Geo-location oriented services according to user mobility
  - Monitoring mechanism to obtain the real time status of the users

Mobile CDNs

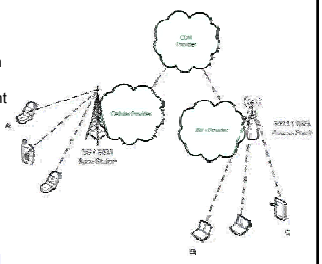


- Components
  - A set of surrogate servers, a network infrastructure, a network status monitoring mechanism, a cache manager, a content location manager and an accounting mechanism

Features	Typical CDN	Mobile CDN
Content type	Static, dynamic, streaming	Static, dynamic, streaming
Users location	Fixed	Mobile
Surrogate location	Fixed	Fixed
Surrogate topology	Close to ISPs	Close to Base Station (BS)
Replica maintenance cost	Medium	High
services	Application services	User and geo-location oriented services
Content outsourcing policy	Cooperative/ uncooperative pull-based scheme	Cooperative push-based scheme

Mobile CDNs Under Centralized Wireless Infrastructure

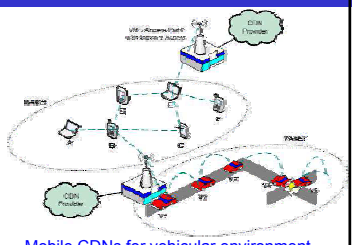
- User communicates with a central authority
  - For cellular network, it is a 3G and/or GSM enabled base station
  - For Wi-Fi network, it is an IEEE802.11 enabled Access Point
- Although it provides a good framework for mobile CDNs, it has limited scalability and resilience
- Careful and systematic provisioning is required to provide required communications coverage and user QoS



Centralized wireless network infrastructure

Mobile CDNs Under Ad-Hoc Wireless Infrastructure

- Characteristics of Mobile Ad-Hoc Network (MANET)
  - Self-organized and self-configurable with Ad-hoc and P2P connectivity
  - Dynamic network topology due to the unpredictable random mobility pattern of participating nodes
  - Mobile node constraints
- Distinct features of Vehicular Ad-Hoc Network (VANET)
  - More dynamic network topology
  - No constraint in terms of processing power of uptime duration
  - Predictable node movement



- Mobile CDNs for vehicular environment
  - Dissemination of traffic-related information
  - Location-oriented services
  - Dynamic and streaming content provisioning
  - Distributed gaming

Implementation and Experimentation Perspective

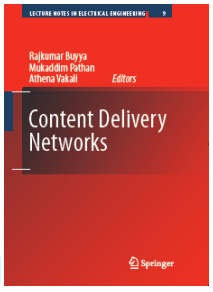
- Simulation testbed
  - ORBIT, a wireless network testbed and PlanetLab, a large scale testbed could be used for real-time experiments
- Simulation software
  - CDNSim has the potential to be extended for mobile CDNs simulations
  - Scope for extension
    - Realistic mobility traces
    - Support of wireless environments
    - Support for mobile resource-limited nodes
- Research issues to be addressed
  - Content placement technique
    - Taking user demand pattern and variation into account, future demand could be predicted
  - Disseminating dynamic content
    - A scalable and robust approach is required to alleviate the high sensitivity to delays
  - Disseminating mobile streaming media
    - Improvement is possible by CDN servers using state-of-the-art compression techniques

Visionary Thoughts

- Information dissemination in mobile wireless environment is an interesting and challenging problem
  - Since traditional CDNs do not take user mobility into account, mobile CDNs could be used for accelerating information dissemination in mobile wireless environments
- IBM WebExpress, specifically, client/intercept/server wireless computational model could be used for mobile CDNs
- Mobile computing and mobile networking are still evolving
  - There are lots of potential for further research
  - Mobile CDNs research should focus on minimizing the communication latency during mobile users Web access and reducing the network traffic over wireless links
- Only one mobile CDN, Ortiva Wireless, is existing
  - An industry specific rise for mobile CDNs could be predicted

Outline

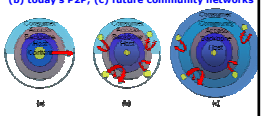
- Part I: CDN Fundamentals
  - CDN Insights
  - CDN in Practice – Akamai Case Study
  - CDN Taxonomy
- Part II: CDN Modeling and Performance
  - CDN Pricing
  - CDN Performance
- Part III: Advanced CDN Platforms and Applications
  - Mobile Dynamic CDNs
  - Content Delivery for Community Networks
  - CDN Internetworking



Content Delivery and Community Networks

- A network community is a group of people whose communication and collaboration over networks strengthens and facilitates their shared identity and goals
  - It is mainly driven by the common "people" or the average users
- Community-based CDNs with users as content producers
  - Decentralization in content delivery must be combined with self-configuring, self-organizing, self-managing and self-adapting solutions at all technical layers to minimize the need for human intervention
- Content delivery and usage is special in community networks context
  - Autonomic network and overlay solutions are needed to establish and maintain an appropriate CDN over physical community networks
  - Arbitrary and complex content services, e.g. content adaptation, transcoding, indexing, and storage, are required
- Spontaneous community networks model is quite successful
  - Dharamsala Wireless-Mesh community network came to life in February 2005, following the deregulation for outdoor use of WiFi in India

Content delivery paths: (a) traditional; (b) today's P2P; (c) future community networks



Architectural Framework

- Community networks are expected to play a central role in the intermediate future since they provide basic connectivity
- CDN implementation as overlay networks are captured through the delivery infrastructure layer in the architecture
- Content services networks consist of a set of services for handling multimedia content
- On top of the architectural framework is the QoE level, which reflects the actual experience of the end user
  - QoE is a function of different QoS parameters at network, system and application level, without a direct translation between QoS parameters and QoE

- Cross layer issues
  - relevant for the scenario of content delivery in community networks, e.g. correlation between QoS parameters at different layers, dependency between layers

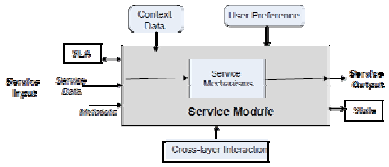
Description of Layers

- Community networks
  - Owned by individual users or groups of users sharing distributed resources in a relatively small geographical area
  - Technology usage for connectivity: xDSL, Powerline, FTTH for fixed nodes connected to an ISP; WiMAX, MBWA, 3G/UMTS/HSDPA for nodes with wireless access to an ISP; WiFi and Bluetooth for mobile nodes and home networks
- Delivery infrastructures
  - Logical infrastructure built on top of the community network with the specific purpose of enabling access to content services
  - P2P overlay network, content delivery and caching strategy for file sharing, live audio/video streaming, online gaming
- Content services networks
  - Infrastructure to provide a whole range of services to enhance content experience through easy navigation and personalized adaptation according to user needs
  - Service-Oriented architecture to provide value added services as infrastructure services
- Quality of Experience
  - Assessment of end-users perceived experience
  - Human Visual Senses (HVS) is crucial for accuracy and enhancement

Content Service Framework

- An End-to-End (E2E) infrastructure is required to provide seamless communication
  - Development of highly interactive applications within community networks
- More general Content Networks (CNs) are evolving which integrate overlay structures and content services, e.g. content management
  - Designing an autonomic CDN architecture to address not only actual content delivery, but also content management orchestration, services functionalities and communications for faster production and easier access
- Different services are to be placed into the overall service framework in order to make them a part of the content network infrastructure
  - Service Level Agreement (SLA) should be in place to access QoS and QoE
- A mechanism for cross-layer interaction is required for information exchange and retrieval from underlying layer and system components

Content Service Framework (Cont'd)



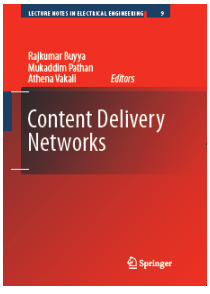
- The content service framework provides the context within which the services are placed
  - Service description and its representation within the service registry is crucial
  - Services can use other services through this service registry via interfaces
- A content service framework allows dynamic and automatic composition of content services and opens up new business opportunities for brokerage services

Visionary Thoughts

- Seamless handovers in a heterogeneous community network environment is challenging and it should be properly handled
- Community network should be coupled with techniques for user misbehavior detection and traffic anomaly detection
- Combination of P2P and classical CDNs could be beneficial
  - Recent technology trends clearly indicate that neighborhood and home networks will be connected to core CDN
- There is the need for supporting infrastructure services that provides more user freedom for interact and share content
  - How content is delivered, e.g. over a video streaming service
  - What to do in case of insufficient resources, i.e. what kind of adaptation strategy should be applied
  - What kind of incentive mechanisms should be used

Outline

- Part I: CDN Fundamentals
  - CDN Insights
  - CDN in Practice – Akamai Case Study
  - CDN Taxonomy
- Part II: CDN Modeling and Performance
  - CDN Pricing
  - CDN Performance
- Part III: Advanced CDN Platforms and Applications
  - Mobile Dynamic CDNs
  - Content Delivery for Community Networks
  - CDN Internetworking



Limitations of Current CDNs

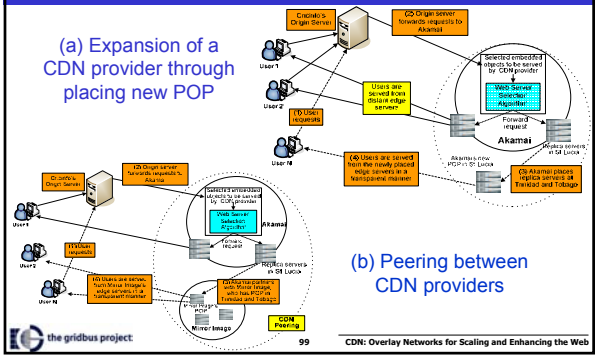
- Commercial CDNs sign Service Level Agreements (SLAs) with customers
  - Objective is to provide global competitive services satisfying QoS requirements
    - Over-provision or harness external resources on-demand
- A "Flash Crowd" or "SlashDot Effect" causes adverse business impact
  - Very often CDNs can not provide QoS to end-users requests
  - SLA violation to end-up costing the provider
- Existing CDNs are proprietary in nature
  - Each has expensive closed delivery network
  - No simultaneous usage of the resources of multiple CDNs
- CDN expenses to have led to consolidation in the market

Possible Solution

- Ad-hoc or planned peering arrangements between CDNs by leveraging existing infrastructures
  - Allows a CDN to "scale-out" to meet both flash crowds and anticipated increases in demand
  - Avoids expense of running a global CDN
- Achieves economics of scale, in terms of cost effectiveness and performance for both providers and end-users
  - Economy-based replication strategy for on-demand content placement
  - Avoids SLA violation
  - Assists to provide competitive service to catch-up the market

CDN Peering Scenario

(a) Expansion of a CDN provider through placing new POP

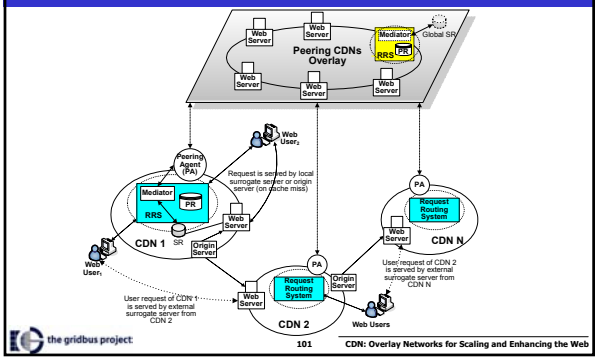


(b) Peering between CDN providers

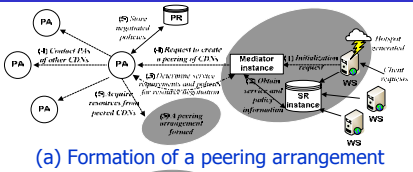
Peering CDNs

- A virtualized constellation between CDNs by leveraging existing infrastructures
  - Members are multiple semi-independent autonomous entities who come together to share resources and collaborate on common goals
  - Ad-hoc (short-term) or planned (long-term) peering arrangements
- Initiator of a peering negotiation
  - Primary CDN
- Participants
  - Peering CDNs or Peers
- Roles are fluid
  - Any participant can be acting as a primary or a peer

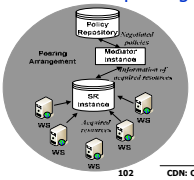
Peering CDNs: Abstract View



Peering Arrangement Formation

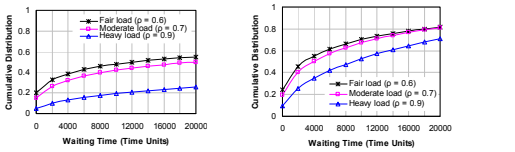


(a) Formation of a peering arrangement

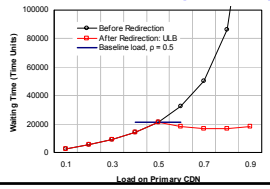


(b) A formed peering arrangement

Performance Gain Through Peering

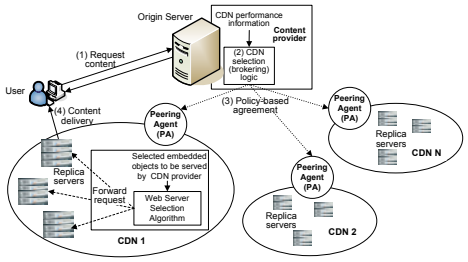


Cumulative distribution of waiting time for the primary CDN



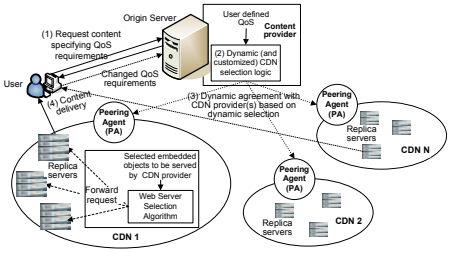
Impact of request-redirection on the waiting time of the primary CDN

New Models for CDN Peering



(a) Brokering-based approach

New Models for CDN Peering (Cont'd)

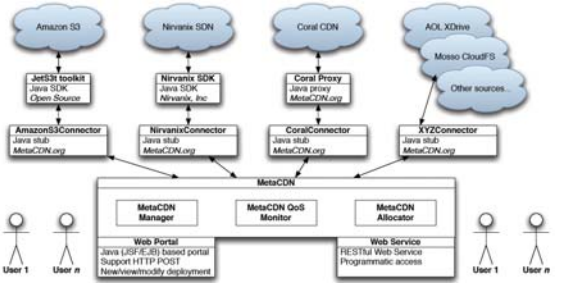


(b) QoS-driven (customized) brokering-based approach

Comparison of CDN Models

Features	Typical CDN models		Advanced models for CDN peering		
	Conventional CDNs	P2P-based CDNs	Peering CDNs	Brokering-based	QoS-driven (customized) brokering-based
Nature of Content Delivery	Based on Web server Collaboration	Based on peering and content availability	Based on CDN Internetworking/ Peering	Based on CDN Performance	Based on user defined QoS (Customized)
Responsibility for effective content delivery	CDN Provider	Peers/Users	Primary CDN Provider	Content Provider	Content Provider
Entities in agreement	CDN-Content Provider	No real Agreement (Self-interested users)	CDN-Content Provider, CDN-CDN	CDN-Content Provider	CDN-Content Provider
Agreement nature	Static	N/A	Short-term or long-term	Policy-based	Dynamic
Scalability	Limited	High	High	High	High
Cooperation with external CDNs	No	No	Yes	Yes	Yes
Cooperation between CDNs	No	No	Yes	No, CDNs work in Parallel	No, CDNs work in parallel
Cooperation between users	No	Yes	No	No	No

MetaCDN.org (Broberg et al, GRIDS Lab): Content Delivery Network over Clouds



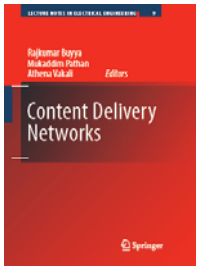
Visionary Thoughts

- Challenges in implementing peering CDNs
  - Legal/copyright issues, global reach, consolidation in CDN market, challenges in brokering-based CDN peering, challenges in P2P-based CDN peering, and lack of incentives for cooperation
- Technical issues for peering CDNs
  - Load distribution tp ensure reduced server load, less bandwidth consumption and improved content delivery performance
- Coordination mechanisms to ensure effectiveness, and to allow scalability and growth of cooperative CDNs
- Service and policy management for value-added services as infrastructure services, policies to support SLA negotiation and autonomous policy negotiation for time-critical short-term VO
- Pricing of contents and services to ensure maximum profit for providers in a competitive environment, yet maintain equilibrium of supply and demand

Tutorial Summary

- CDNs overcome Internet service degradation by offering infrastructure and mechanisms to delivery content and services in a scalable manner, and enhancing users' Web experience
- This tutorial builds on academic and industrial research and developments and case studies by prominent CDN researchers around the world
  - It identifies potential research directions and technologies that drive future innovations
  - It provides in-depth analysis and complete understanding of the current and future trends in the content distribution landscape
- Content networking is an emerging research topic as the CDN landscape is dynamic
  - Ongoing changes include introduction of P2P based CDNs, integration of cloud computing, evolution of storage delivery networks...

Thanks for your attention!



- R. Buyya, M. Pathan, A. Vakali (Eds.), Springer, Germany, 2008 http://www.gridbus.org/cdn/book/
- More information at: http://www.gridbus.org/cdn