Adaptive Cache Architecture for Multimedia Content Delivery

HCL Technologies

Engineering and R&D Services - Networking & Telecom Practice

Author: Deepak Kataria
Co-Authors: Mukta Agarwal, Saurabh Chattopadhyay, Sachin Gupta
Topics

- Introduction
- Data Traffic Analysis & Growth
- Content Service Providers
- Adaptive Cache Architecture – Addressing Content Delivery Challenges
- Summary
- References
Content Delivery Trends in Service Providers

- Trend towards an across the board media-centric network that is targeted for *any-content* at *any-time* over *any-device* type of service with robustness, security, convenience and highest quality of experience (QoE)

- Convergence of Telecom, IT, Internet and Media coupled with competitive pressures from non-traditional players forcing service providers to readjust their cost structures down to a new maximum efficiency of scale

- Rich digital media content ingest, movement and delivery coupled with asset management, DRM, and security causing fundamental shifts in network architectures from tiered centralized architectures to edge based distributed architectures

- Service providers looking at creative ways to monetize their investments
  - Offering high margin, high value service bundling for top line growth
  - Looking to reduce network cost (both CAPEX and OPEX) to shore up bottom line
  - Targeting Ad-based revenue and revenue-sharing models with content owners
  - Providing media management & distribution services for content owners & aggregators
  - Leveraging location awareness capabilities and higher throughputs to provide rich media services

- LTE timelines are accelerating, fueled by the auction of the 700 MHz analog TV spectrum in the US and 2.6 GHz band (IMT extension) in Europe
Topics

- Introduction
- Data Traffic Analysis & Growth
- Content Service Providers
- Adaptive Cache Architecture – Addressing Content Delivery Challenges
- Summary
- References
Mobile and Fixed Traffic

Mobile data takes off
Mobile PC main driver

Fixed traffic growth driven by
Internet and IPTV

Fixed traffic ~100 times mobile traffic
Both mobile and fixed networks traffic will grow over the next few years
According to Strategy Analytics, mobile phones - excluding wireless PC cards and other cellular enabled devices - will generate almost 2.2 Exabytes of data traffic; 84% will come from web browsing. This volume of data is equivalent to watching over 120,000 years of DVD quality video.

**Data Volumes per Subscriber & Different Access Methods**

FTTH access methods “pump” a lot of data into core networks
Traffic Distribution


- While developed regions will account for just 25% of the cellular user population by 2015, they will generate 65% of total global wireless network traffic“, according to Dr Mark Heath, the report’s co-author. “This is due to a higher proportion of advanced handsets and the earlier deployment of more advanced cellular technologies, such as LTE, which have higher throughput.”

- Historically, voice telephony has dominated wireless network traffic, and popular data services, such as SMS, consume a tiny amount of network resource. However, the take-up of USB modems and a broad range of data services on smart-phones will increase wireless network traffic. Average wireless network traffic per cellular user (for all voice and data services) in developed regions will increase to eight times its 2008 level by 2015, rising from 56MB per month to 455MB per month.

- While voice traffic will continue to increase, as a result of on-going fixed-mobile substitution, data traffic will rise at significantly faster and come to dominate wireless network traffic. By 2015, data will account for 94% of total wireless network traffic in developed regions.

- A number of uncertainties could result in traffic levels differing from our base-case forecasts. An upside forecast derived by Analysis Mason indicates that by 2015, traffic per customer in developed regions will grow to almost 30 times its level in 2008.

Source: Paper from Continuous Computing, Published in October 8, 2009

Few users generate majority of traffic
Typical Telco Network Architecture for Video

Example of Network used by Telcos

Source: Cisco
Typical MSO Network Architecture

Example of Network used by MSOs
Tiered Content Delivery Infrastructure

Master Headends

Regional Headends

Local Headends

IP Stream

IP Stream

IP Stream

Example of Network used by CDNs

Content Aggregation

Add Local Content & Advertising

On-Demand

Gigabit Ethernet

FTTH

ADSL2+

Gigabit Ethernet

FTTH

ADSL2+
3GPP MBMS Architecture for Mobile TV

Example of Network used by Mobile Operators
DVB-H High Level Architecture

High Level Architecture

Content Provider → Contribution Network → DVR-H stream generator

Conditional Access

1st level sites – Broadcast Sites

2nd level sites – GAP Fillers Sites

Distribution Network

Cooperation Platform

Billing, Authentication, CRM, ...

Service Platform

Core Network

UMTS Network

UTRA Network

Mobile TV Network Deployed in Europe
Example of Network used by Terrestrial Broadcasters
Satellite Distribution

Example of Network used by Satellite Operators
# Content Service Providers Capability Matrix: As on today

<table>
<thead>
<tr>
<th>Content Service Provider</th>
<th>Content</th>
<th>Network</th>
<th>Distribution</th>
<th>3 - Screen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telco</td>
<td>Leadership</td>
<td>Challenged</td>
<td>Challenged</td>
<td>Leadership</td>
</tr>
<tr>
<td>MSO</td>
<td>Challenged</td>
<td>Leadership</td>
<td>Leadership</td>
<td>Challenged</td>
</tr>
<tr>
<td>CDN</td>
<td>Challenged</td>
<td>Leadership</td>
<td>Challenged</td>
<td>Leadership</td>
</tr>
<tr>
<td>Mobile</td>
<td>Challenged</td>
<td>Leadership</td>
<td>Leadership</td>
<td>Challenged</td>
</tr>
<tr>
<td>Terrestrial</td>
<td>Challenged</td>
<td>Leadership</td>
<td>Leadership</td>
<td>Challenged</td>
</tr>
<tr>
<td>Satellite</td>
<td>Challenged</td>
<td>Leadership</td>
<td>Leadership</td>
<td>Challenged</td>
</tr>
</tbody>
</table>
Unified Content Centric Network View

Digital Content Data Center / Head End Content Server

Super Head End Live Broadcast & VoD Server

Internet Service Provider

Video Hub Office / CDC_Regional

Video Subscriber Office / CDC_Local

Service Provider Core

Internet

Hybrid Fiber/Coax

PC

HDTV

Mobile TV

[1] Data Modification / Analysis functions
[2] Security (Parental Ctrl) / Content based behaviour
[3] VSO is owned by Telco
[4] CDC Regional/Local works with SP Infrastructure, owned by CDN

This is the Network Everybody is Going After for 3-Screens
VoD workloads are primarily different in that they are read dominated, involve large data transfers with isochronous delivery requirements and have a Zipf-Mandelbrot distribution.

Need for an innovative network transport technology which is non-disruptive and works to augment performance of existing VoD distribution networks and offers the capability for managed peer-assisted deployment.

Leverage the network transport technology for the design of VoD server so as to decrease the server load while improving client transfer speed and affording scalability which is demand-driven.

Achieve high aggregate throughput and reliable performance from disk-array based VoD servers, through the design of high performance streaming engines and controller designs that are optimized for VoD workloads.

Provide for improved efficiency and enhanced interactivity by making use of the large total storage and server capacity of the participating peer network through virtualization of network, storage and server resources.
Topics

- Introduction
- Data Traffic Analysis & Growth
- Content Service Providers
- Adaptive Cache Architecture – Addressing Content Delivery Challenges
- Summary
- References
Solution Offering with Control, Service Enablement & Orchestration Capabilities

Service Enablement and Management Mediation

Adaptive Cache positioned in Edge Network

Local Concurrency of Service Delivery Points

Policies based on Popularity & Long Tail nature of Content

* TR 144 like Multi service Requirements
* TR 161 like Serviceability over Ethernet Access Aggregation

* TR 148 like Enablement of IPTV QoE
* TR 173 like Enablement of ADSL2+ IPTV

Spatial & Temporal Distribution of Content Request

Content Orientation

Bandwidth Mgmt

Congestion Control

Content Access Orientation

Content Cache Orientation

More Popular

Transport Storage Costly

SLA Associated Delivery

Optimum & Weighted Distribution of Media

Content Evolve Temporal Evolution

Content Cache Evolution

Content Cache Position

In-Netwerk Control & Elasticity Plane

Management & Governance Of End to End Services

Potential Elements to run Adaptive Cache as extended architecture

Fast Ingestion of content for immediate availability

Data Offloading to Wire-line Edge

MTOSI or SOAP/XML or TR-069 LIKE Mediation

Content access from adjacent localities to maximize distributed concurrency

Usage Analytics driving Network Policies, adapting to ambient conditions

Source: Whitepaper, Adaptive Caching Architecture for Multimedia Content Delivery, Networking & Telecom Practice, ERS, HCL
Solution Offering in Infrastructure & Virtualization Environment

Adaptation Policies based on localized knowledge using Interoperable Control Plane of Routers

PnP Enablement & Policies for P2P Media Caching

Network Resource Pool for efficient utilization

Local Concurrency, along with maximized distributed concurrency

Offline / Near-RealTime Usage & Cache Effectiveness Analytics

Automated control for trade off among Storage & Transport Cost and Bandwidth

Real Time Usage Analytics running along Adaptive Cache Instances

Any Content, Any Time, Any Where

Source: Whitepaper, Adaptive Caching Architecture for Multimedia Content Delivery, Networking & Telecom Practice, ERS, HCL
Future Internet Capabilities of the Architecture - Mapping Solution Architecture Against MANA Positioning

Source of background image: Position Paper - Management and Service-aware Networking Architectures (MANA) for Future Internet
Summary

- Rich media content will be the dominant traffic in future networks
- On-demand delivery modality will be key – any content, any time, anywhere to any device using any access network - the so-called “any-to-any” mode
- Mechanisms to optimize the delivery of rich media content for “any-to-any” mode will be of paramount importance
- Adaptive caching architectures that leverage a judicious mix of tier-ing (hierarchical) and peering approaches, that address the spatial and temporal distribution of content demand, that provide for fast ingestion of content for immediate availability, that maximizes the local concurrency of service delivery points, that leverages access of content from adjacent localities in the network to maximize distributed concurrency, that take into account the popularity and long tail nature of content, that utilizes usage analytics to drive network policies that adapt to ambient network conditions will be needed to deliver on the promise of any-to-any communication
- Architectural Positioning and Distribution that leverage content replication & availability, network topology information, concomitant transport cost information, and gathered analytics data to intelligently stage transfers, while maximizing the local concurrency of service delivery points and minimizing the total cost
- Architecture of Distributed Governance that enable and orchestrate the plane of End to End Content Centric Service Delivery and Management
References

Thank You
Lab Deployment Diagram

- **Support for various media types**
- **Interoperability with External Cache (ICP Stack)**
- **Automated Failover of Seed/Content Servers and Cache Sibling/Parents**
- **HTTP Streaming**
- **IP Backbone**
- **Users**
- **Adaptive Cache**
- **Cache / ICP Stack**
- **Collaborative Cache Hierarchy**
- **Aggregate Media Streaming**
- **Preemptive link monitoring**
- **Replicated Content Servers**
- **Fractional Distribution of Large Media for optimum use of Transport & Storage**
Current Feature Support

- **Core Cache**
  - Local Cache with ICP (RFC 2186) Interfaces
  - Interoperability with externally hosted ICP stack
  - Collaborative Cache Support
  - Caching Policies, Cache Clearance Policies and Cache Movement policies
  - Rule Engines for distinguishing Cacheable and Non-Cacheable Content parts

- **Security**
  - AA for Local Access
  - AA for Content Access
  - AA Wrapper over Management and Service Enablement Interfaces

- **Seed Server Load Balancing**
  - Core Engine with blend of Transparent and Non Transparent Load Balancing
  - URL based, Content based, Location based Switching

- **Media Streaming**
  - Support for HTTP Streaming
  - Support for diverse Media Types
  - Aggregated Streaming

- **Service Enablement & Management**
  - Web 2.0 based Administration
  - CLI / SSH console
  - Real-time Service Statistics Collection Policies
  - Near Real-time Reporting policies
  - Auto Detection of Failure and Failover
  - Pre-emptive Monitoring and Diagnostics

- **Quality & Experience of service**
  - Configurable Traffic Prioritization & Controlling Policies
  - Traffic Rate Controlling policies
  - Media Stream Adaptation

- **High Performance through SW High Availability & Stack Optimizations**
**Solution Roadmap (Internal)**

**M1 Jan 10**
- Milestone 1
  - Experimentation
  - Lab Setup of Experimentation Demonstration
  - Feasibility Analysis of Data Grid based Architecture

**M2 Mar 10**
- Milestone 2
  - Requirements Definitions / Feature List
  - Feature Compliance Analysis
  - Low Level Design
  - Whitepaper Presentation

**M3 / R0.9 Jul 10**
- M3 / Release 0.9
  - Implementation
  - Unit Testing
  - Roadmap Requirement Analysis
  - Integration Testing
  - Features
    - Basic Adaptive Caching
    - Load Balancing over Content Servers
    - Service Enablement & Management
    - QoS & EoS
    - Basic Self – Management Capabilities

**M4 / R1.0 Sep 10**
- M4 / Release 1.0
  - Release Trials in Simulated Lab Environment
  - Market Facing Documentation
  - Industry Briefing
  - Demo Showcasing
Key Areas in Broadband Forum’s Vision

Source: http://www.broadband-forum.org/about/mission.php

ACS: auto-configuration server
AN/MSE: access node/multiservice edge
AP: access point
BNG: broadcast network gateway
BS: base station
BSC/RNC: base station controller, radio network controller
CPE: customer premises equipment
CWMP: customer premises equipment wide area network management protocol
DQ: data quality management
GPON: Gigabit capable passive optical network
L2CP: Layer 2 control protocol
O&NM: operations and network management
OSS: operations support system
PCC: policy and charging control
QoE: quality of experience
RG: residential gateway
SOHO: small office, home office
STB: set-top box
TDM: time division multiplexing
VPN: virtual private network
P4P Field Test Deployment

Source: www.openp4p.net – Field Test Deployment Diagram