Bringing it together
Multimedia in IP

(Web view)

1

Multimedia Networking Applications

2

- **Fundamental characteristics:**
  - Typically delay sensitive
    - end-to-end delay
    - delay jitter
  - But loss tolerant: infrequent losses cause
  - ..., which are loss intolerant but delay tolerant.

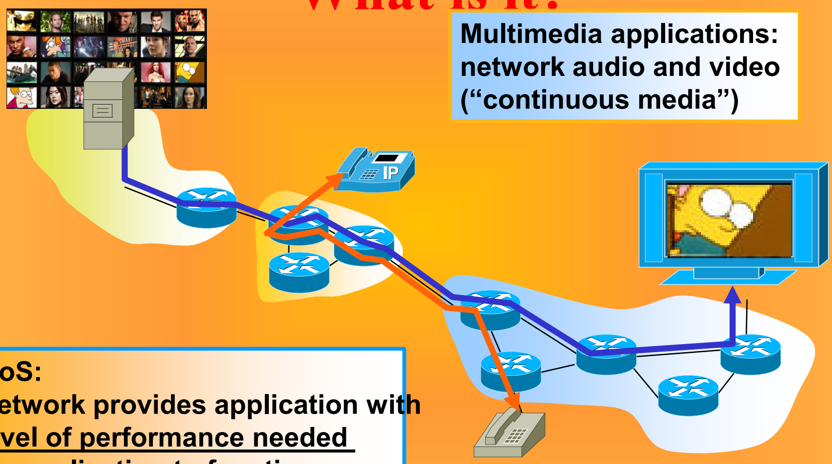
Jitter is the variability of packet delays within the same packet stream

- **Classes of multimedia applications:**
  - Streaming stored audio and video
  - Streaming live audio and video
  - Real-time interactive audio and video

2

Multimedia, Quality of Service: What is it?

Multimedia applications: network audio and video ("continuous media")



QoS:
Network provides application with <u>level of performance needed for application to function.</u>

3

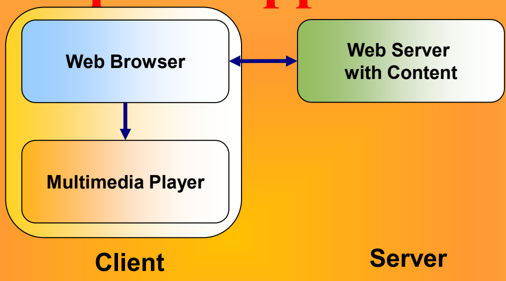
Internet Multimedia Support

- Integrated services philosophy.
  - Requires dedicated links/channels with QoS requirements.
- Differentiated services philosophy.
  - Fewer changes to Internet infrastructure.
- Best effort.
  - No major changes.
  - More bandwidth when needed.
  - Application-level control and distribution.

Would require QoS
Only possible in private networks or operator infrastructure

4

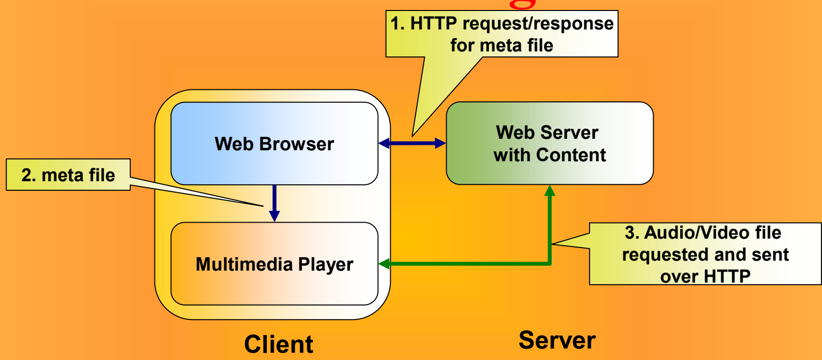
Internet Multimedia: Simplest Approach



- Audio or video stored in file.
- Files transferred as HTTP object (or using P2P).
  - Received in entirety at client as a file.
  - Then passed to default player in client.
- Audio&video is not streamed!
- No “pipelining”, long delays until playout!

5

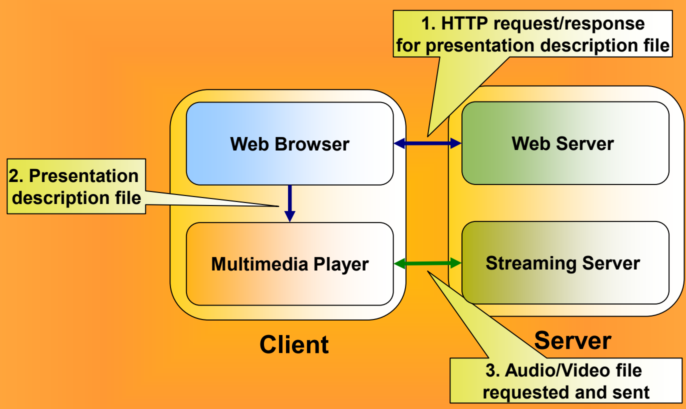
Internet Multimedia: Web Streaming



- Browser GETs metafile.
  - Content negotiation may happen.
- Browser launches player, passing metafile
- Player contacts server.
- Server streams audio/video to player.

6

Streaming from a streaming server

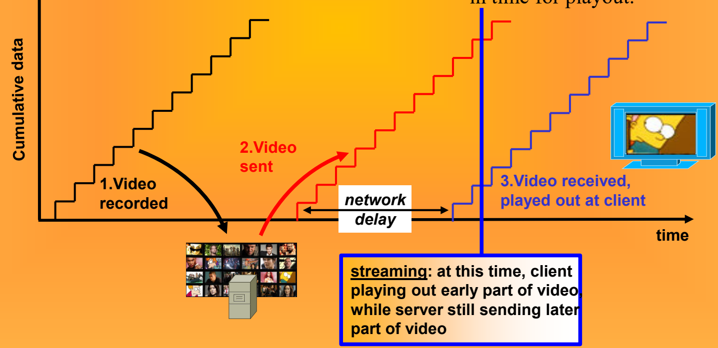


- This architecture allows for non-HTTP protocol between server and media player.
- Can use UDP or TCP transport.

7

Web Streaming: Stored Multimedia

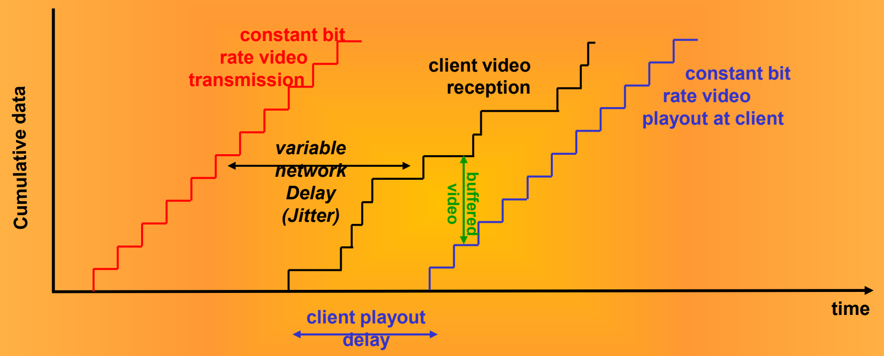
- Streaming:
  - Media stored at source is transmitted to client.
  - Client visualization begins before all data has arrived.
    - Timing constraint for transmitted data: in time for playout.



8

9

Streaming Multimedia: Client Buffering

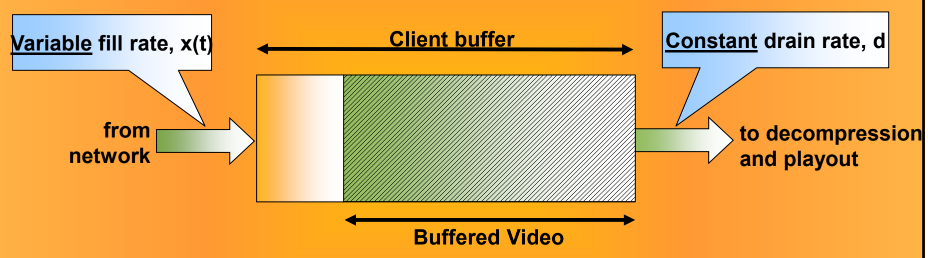


- Client-side buffering, playout delay compensate for network-added delay, delay jitter.
  - Size of buffer/playout delay are parameters that can be adjusted dynamically
  - For VoIP, delay between 2 packets is about 20ms (plus *jitter*).

9

10

Streaming Multimedia: Client Buffering



- Client-side buffering, playout delay compensate for network-added delay, delay jitter.
  - Size of buffer/playout delay are parameters that can be adjusted dynamically

10

Streaming Stored Multimedia

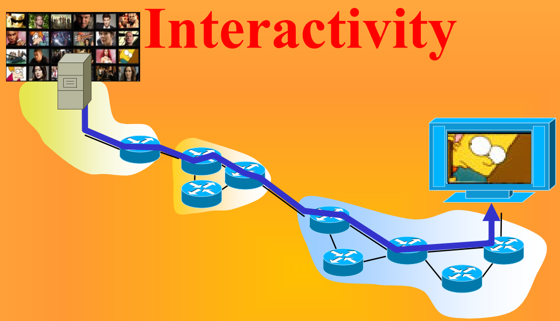
11

- Application-level streaming techniques for making the best out of best effort service:
  - Client side buffering.
  - Use of UDP versus TCP.
  - Multiple encodings of multimedia.
- Multimedia Player
  - Jitter removal,
  - Decompression,
  - Error concealment,
  - Graphical user interface with controls for interactivity.
- Network
  - Close to client content (multi-content) buffering for faster interactivity
  - Only viable in network operator proprietary services.

11

Streaming Stored Multimedia: Interactivity

12



- VCR-like functionality: client can pause, rewind, fast-foward, push slider bar.
  - 10 sec initial delay OK.
  - 1-2 sec until command effect OK.
  - Timing constraint for still-to-be transmitted data: in time for playout.

12

13

Streaming Live Multimedia

- Examples:
  - Internet TV/radio show.
  - Live sporting event.
- Streaming
  - Playback buffer.
  - Playback can lag tens of seconds after transmission.
  - Still have timing constraint.
- Interactivity
  - Fast forward impossible.
  - Rewind, pause possible!

13

14

Interactive Real-Time Multimedia

- Applications:
  - IP telephony, video conference, online-game multimedia actions, distributed interactive worlds.
- End-end delay requirements:
  - Audio: < 150 msec good, < 400 msec OK
    - Includes application-level (packetization) and network delays.
    - Higher delays noticeable, impair interactivity.
- Requires session initialization
  - Advertise its IP address, port number, encoding algorithms, required contents, available contents

14

15

UDP Streaming vs. TCP Streaming

- UDP
  - Server sends at rate appropriate for client .
    - Often send rate = encoding rate = constant rate.
    - Then, fill rate = constant rate - packet loss.
  - Short playout delay (2-5 seconds) to compensate for network delay jitter.
  - Error recover: time permitting.
- TCP
  - Send at maximum possible rate under TCP.
  - Fill rate fluctuates due to TCP congestion control.
  - Larger playout delay: smooth TCP delivery rate.
  - HTTP/TCP passes more easily through firewalls.

15

16

HTTP/TCP Streaming

- Multiple versions with distinct/complementary characteristics are generated for the same content
  - With different bitrates, resolutions, frame rates.
- Each version is divided into time segments.
  - e.g., two seconds.
- Each segment is provided on a web server and can be retrieved through standard HTTP GET requests.
- Examples of protocols:
  - MPEG's Dynamic Adaptive Streaming over HTTP (DASH).
    - Standard ISO/IEC 23009-1. YouTube's default.
  - Adobe HTTP Dynamic Streaming (HDS).
  - Apple HTTP Live Streaming (HLS).
  - Microsoft Smooth Streaming (MSS).

16

17

User Control of Streaming Media: RTSP

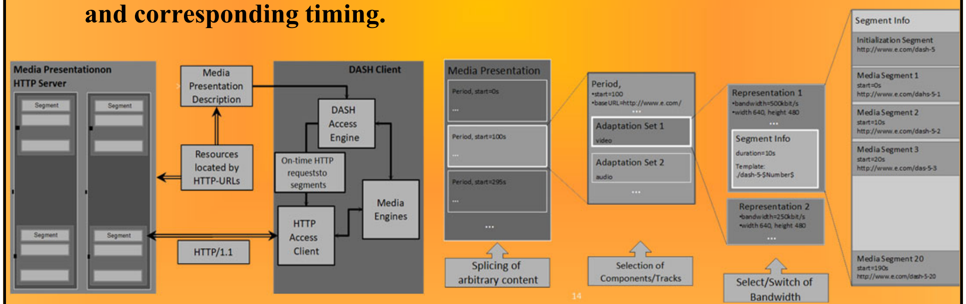
- RTSP (Real Time Streaming Protocol): RFC 2326
  - Client-server application layer protocol.
  - For user to control display: rewind, fast forward, pause, resume, repositioning, etc…
- Does not define how audio/video is encapsulated for streaming over network.
- Does not restrict how streamed media is transported.
  - Can be transported over UDP or TCP.
- Does not specify how the media player buffers audio/video.
- RTSP messages are also sent out-of-band:
  - RTSP control messages use different port numbers than the media stream: out-of-band
    - Port 554
  - The media stream is considered "in-band"

17

22

Dynamic Adaptive Streaming over HTTP (DASH)

- Developed to be an Open Standard Delivery Format.
  - MPEG DASH ISO/IEC 23009-1.
- Video streaming solution where pieces of video streams/files are requested with HTTP and spliced together by the client.
  - Client entirely controls delivery.
- Media Presentation Description (MPD) describes accessible Segments and corresponding timing.



22

WebRTC

23

- Peer-to-peer connections.
  - An instance allows an application to establish peer-to-peer communications with another instance in another browser, or to another endpoint implementing the required protocols.
- RTP Media.
  - Allow a web application to send and receive media stream over a peer-to-peer connection (discussed in a minute)
- Peer-to-peer Data
  - Allows a web application to send and receive generic application data over a peer-to-peer connection.
- Peer-to-peer DTMF.

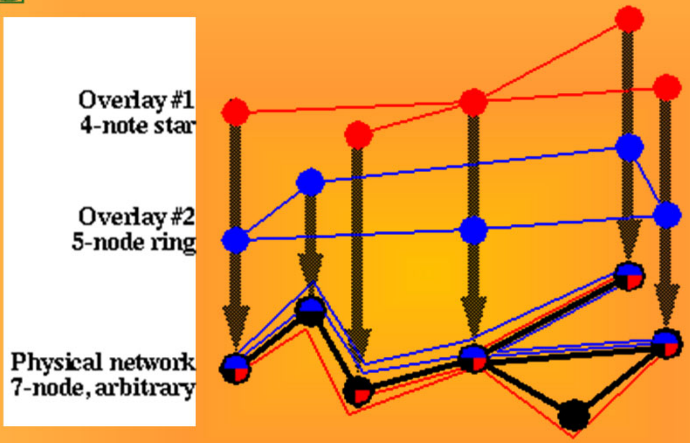
23

CDNs

Everyone in the same network ?

24

Recall: what is an Overlay ?



What is the topology of this network?

WHICH network??

www.isi.edu/xbone

26

Overlay Networks: Overview

- **Networks built using an existing network as substrate (*Virtual Networks)***

Internet

  - **Initially an overlay on the POTS (Plain Old Telephone System) network**
- **Overlays are a (quasi) structured virtual topology above the basic transport protocol level that facilitates deterministic search and guarantees convergence**
  - **Overlays could consist of routing software installed at selected sites, connected by encapsulation tunnels or direct links**
- **Examples of overlays:**
  - **MBone, 6Bone**
  - **P2P (Napster, FreeNet, Gnutella, Bittorrent)**
  - **Cooperating Caches**
  - **Server Farms**
  - **Content Distribution Networks (CDNs)**

27

Content Distribution Networks

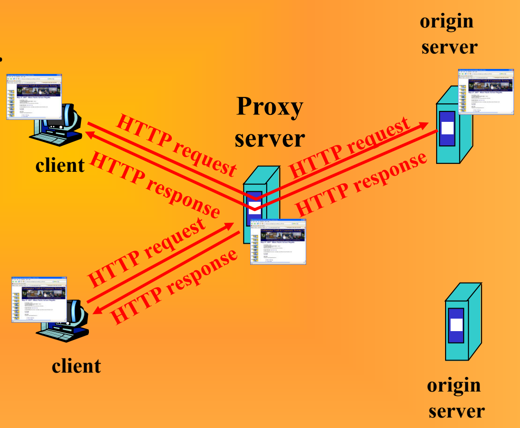
Client-Server and distribution models
Caching and load balancing

28

(recall FR): Web caches (proxy server)

Goal: satisfy client request without involving origin server

- user sets browser: Web accesses via proxy server
- browser sends all HTTP requests to proxy
  - object in cache: cache returns object
  - else proxy requests object from origin server, then returns object to client



30

More about Web caching

- Proxy server acts as both client and server
- typically proxy server is installed by ISP (university, company, residential ISP)

Why Web caching?

- reduce response time for client request
- reduce traffic on an institution's access link.

Conditional GET

- **Goal:** don't send object if cache has up-to-date cached version
- cache: specify date of cached copy in HTTP request
  `If-modified-since: <date>`
- server: response contains no object if cached copy is up-to-date:
  `HTTP/1.0 304 Not Modified`

cache **server**

HTTP request msg
`If-modified-since: <date>`

HTTP response
`HTTP/1.0 304 Not Modified`

object not modified

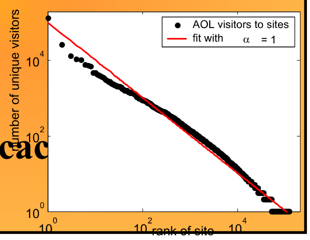
HTTP request msg
`If-modified-since: <date>`

HTTP response
`HTTP/1.0 200 OK`
`<data>`

object modified

Optimizing performance

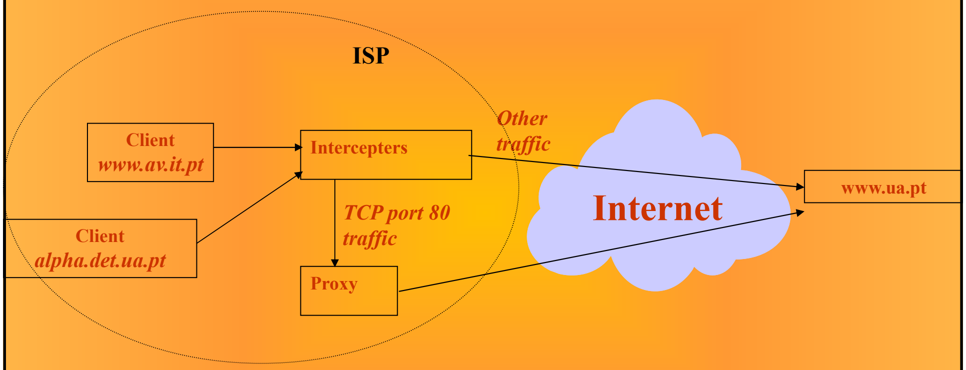
- **Where to cache content?**
  - **Popularity of Web objects is Zipf-like**
    - a few elements that score *very* high (the left tail in the diagrams)
    - a medium number of elements with middle-of-the-road scores (the middle part of the diagram)
    - a huge number of elements that score very low (the right tail in the diagram)
  - **Small number of sites cover large fraction of requests**



- **Given this observation, how should cac replacement work?**

37

Potential content network structure: Caching Proxies

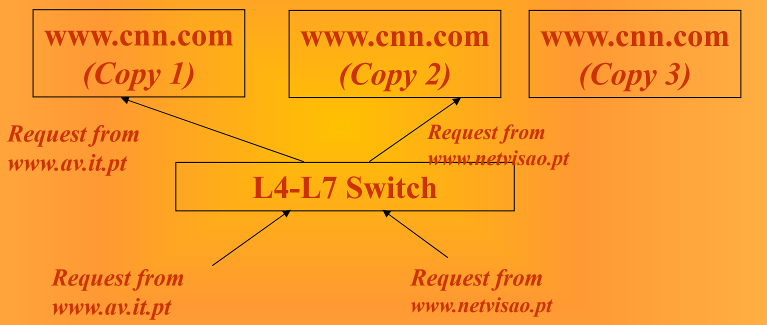


- **Mostly motivated by ISP business interests – reduction in bandwidth consumption of ISP from the Internet**
  - **Reduced network traffic**
  - **Reduced user perceived latency**

41

Potential content network structure : Server Farms

Simple solution to the content distribution problem:
deploy a large group of servers on site

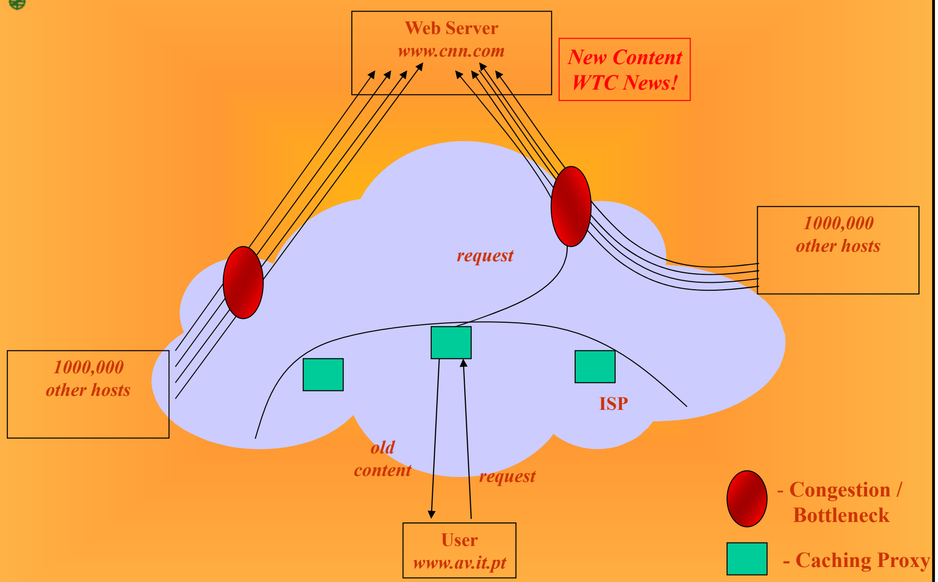


- Arbitrate client requests to servers using an "intelligent" L4-L7 switch
  - Quite used today

42

Flash Crowds:

Consider, On September 11, 2001



43

44

Why Not Web-only approaches for content networks?

- Integrating file caching in proxies
  - Optimized for 10KB objects
  - 10GB = 1.000.000 x 10KB
- Memory pressure
  - Disk access is 1000 times slower
  - Working sets do not fit in memory
- Waste of resources
  - More servers needed
  - Provisioning is a must

44

Problems with *Server farms and Caching proxies*

- Server farms do nothing about problems due to network congestion, or to improve latency issues due to the network
- Caching proxies serve only their clients, not all users on the Internet
- Content providers *(say, Web servers)* cannot rely on existence and *correct* implementation of caching proxies
- Accounting issues with caching proxies.
  For instance, *www.cnn.com* needs to know the number of hits to the webpage for advertisements displayed on the webpage

45

47

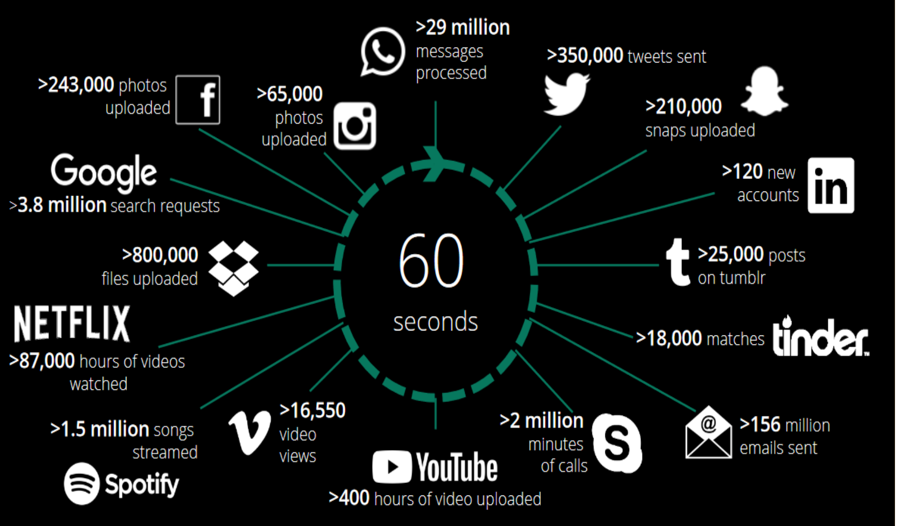
CDNs

47

So much happened in our digitalized world in 2017 – and we have the numbers behind it

Things that happened online in 2017 within 60 seconds

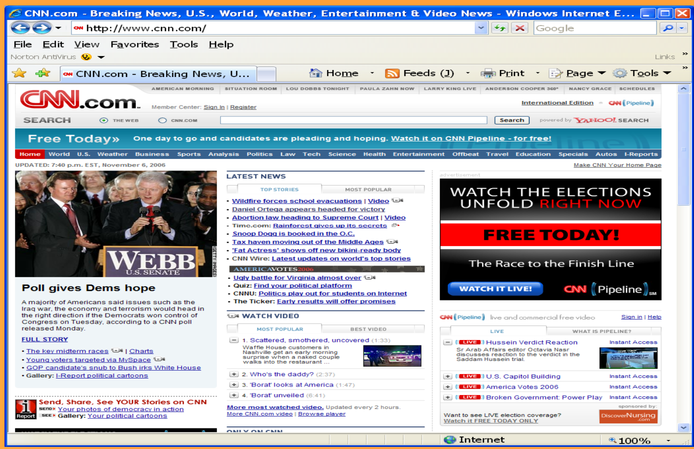
Lots of multimedia content!



49

50

CDNs Target environment?



Most Web files are small (1KB ~ 100KB)

(initially….)

50

Motivation

51

- IP based networks
- Web based applications have become the norm for corporate internal networks and many business-to-business interactions
- Large acceptance and explosive growth
  - Serious performance problems
  - Degraded user experience

  For a large set of applications, including VIDEO access
- Improving the performance of networked applications
  - Use many sites at different points within the network
    - Stand alone servers
    - Routers

51

CDNs basics

- What is a CDN?
  - A network of servers delivering content on behalf of an origin site
    - A number of CDN companies well established now
      - E.g. Akamai, Digital Island, Speedera, CDN77, Cloudfare, Stackpat
    - Many companies are exploring CDNs
      - Avoid congested portions of the Internet
- Consist of
  - Edge servers deployed at several ISP (Internet Service Provider) access locations and network exchange points
- Large-file service with no custom client, no custom server, no prepositioning
- Improve the response time of an Internet site
  - Offloading the delivery of bandwidth-intensive objects, such as images and video clips
- Intelligent Internet infrastructure that improves the performance and scalability of distributed applications by moving the bulk of their *computation* to servers located at the edge of the network
  - Applications are logically split into two components
    - Executed at an edge server close to the user
    - Executed on a traditional application server

52

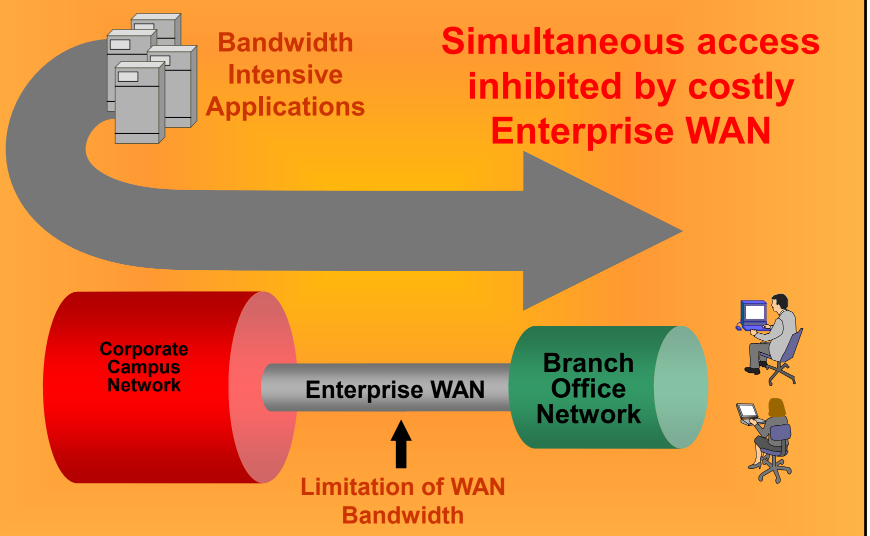
53

CDN Generations

- First generation (early 90ies)
  - Accelerate the performance of web sites
  - Support increasing volumes of traffic
    - Key disruption event: 9/11
    - Akamai technologies created
- Second generation (early 2000ies)
  - Support high volumes of multimedia traffic
  - Audio/video intensive networks
    - All ISPs developed/used CDNs
- Third generation (2010+)
  - Cloud computing
    - Amazon cloud (2008)
  - UGC (user generated content)
  - P2P and interactivity
    - AT&T distributed data centers (2011)
  - Mobile support, and device adapted content

53

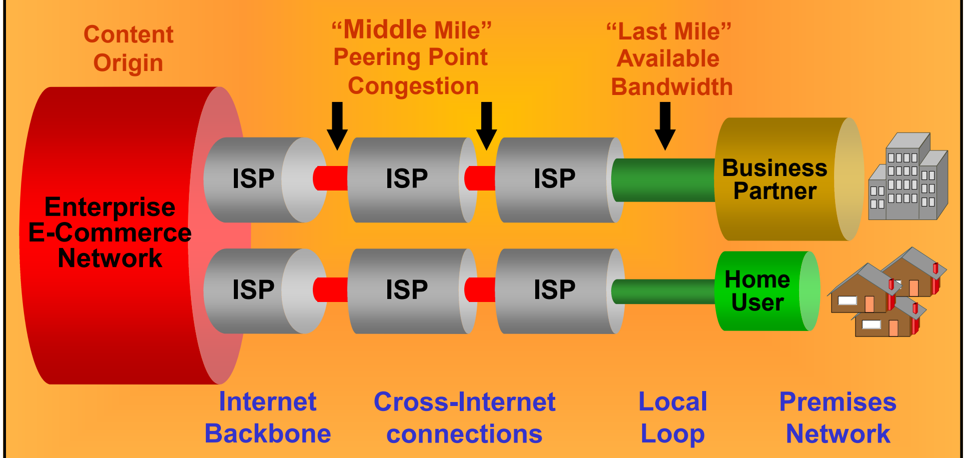
Internal Enterprise Applications



54

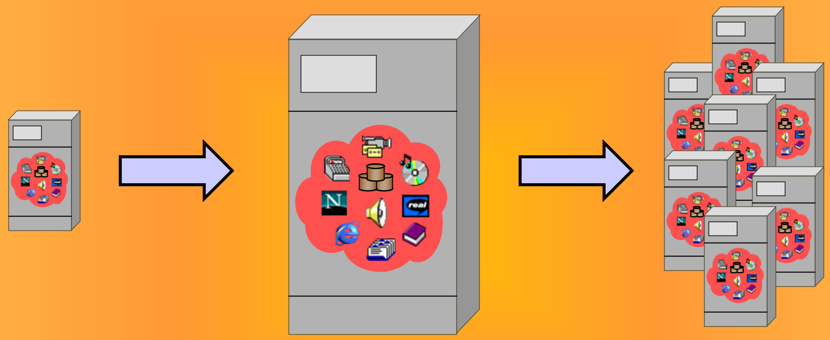
External Enterprise Applications

E-Commerce applications limited by congestion points within the Internet



55

Content Scaling



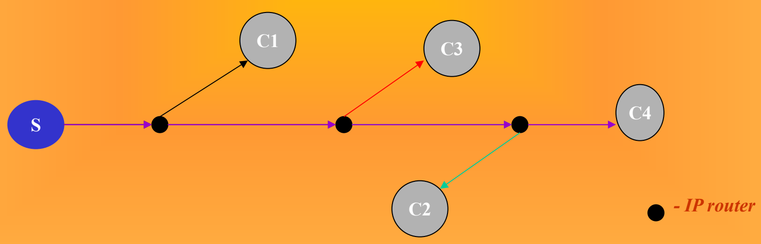
- Need to scale content to handle numerous clients
  - One can only scale 'vertically' to a point

Multiple servers/locations introduces new issues

56

Early Motivations for Content Networks (1st generation)

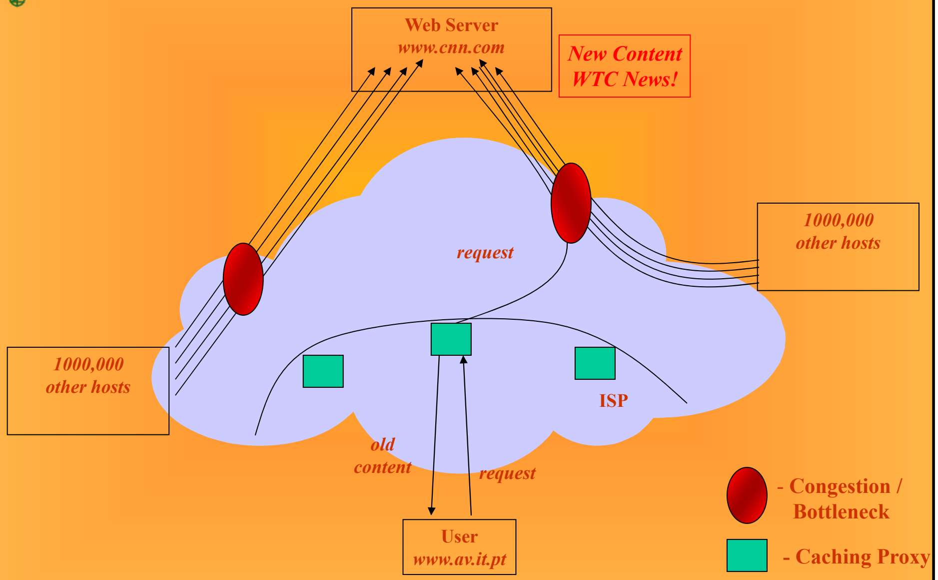
- More hops between client and Web server => more congestion!
- Same data flowing repeatedly over links between clients and Web server
- Origin server is bottleneck as number of users grows
- Flash Crowds *(for instance, Sept. 11)*
  - *The Content Distribution Problem:* Arrange a rendezvous between a content source at the origin server *(www…com)* and a content sink *(users)*



57

Flash Crowds:

Consider, On September 11, 2001



59

60

Flash crowd solution: CDNs..

What is a CDN?

A network of servers delivering content on behalf of an origin site

Large-file service with

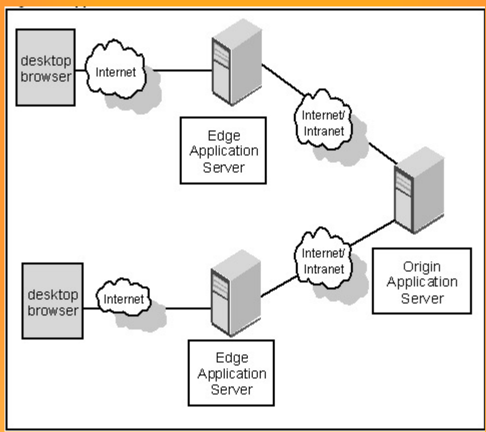
- No custom client
- No custom server
- No prepositioning
- No rehosting
- No manual provisoning

60

61

Model

- Application offload (1st generation concern)



61

62

Content distribution networks

- Client attempts to access the main server site for an application
- It is redirected to one of the other sites
- Each site caches information
  - Avoid going to the main server to get the information/application
- Access a closelly located site
  - Avoid congestion on the path to the main server
- Set of sites used to improve the performance of web-based applications collectivelly
  - Content distribution network

62

Inside a CDN

- Servers are deployed in clusters for reliability
  - Some may be offline
    - Could be due to failure
    - Also could be "suspended" (e.g., to save power or for upgrade)
- Could be multiple clusters per location (e.g., in multiple racks)
- Server locations
  - Well-connected points of presence (PoPs)
  - Inside of ISPs

63

Advantages

64

- Better scalability
- Higher availability
- Improved response time from a centrally managed solution
- Nodes constituting the distribution network are designed to be
  - Self-configuring
  - Self-managing
  - Self-diagnosing
  - Self-healing

to ensure easy management and operational convenience

64

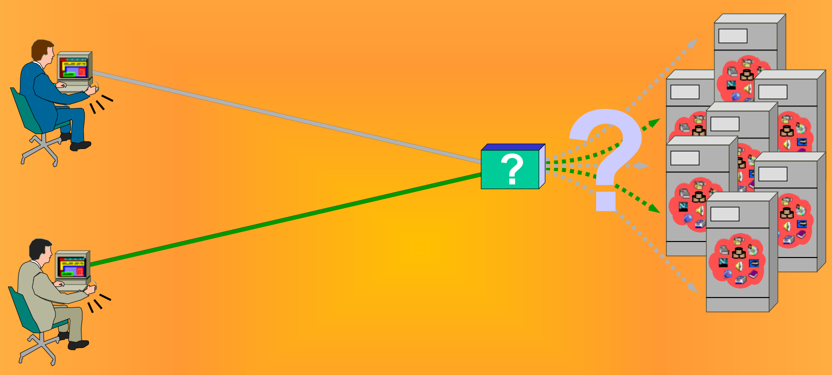
Challenges

- Keep consistency among the enterprise data hosted by the offloaded applications
- Share session state among edge and origin application servers
- Distribution, configuration, and management
- Develop programming models consistent with current industry standards such as J2EE
- Application security.

- There is active research into general frameworks to be used to support distributed applications, as well as prototyping the ideas for specific application instances

65

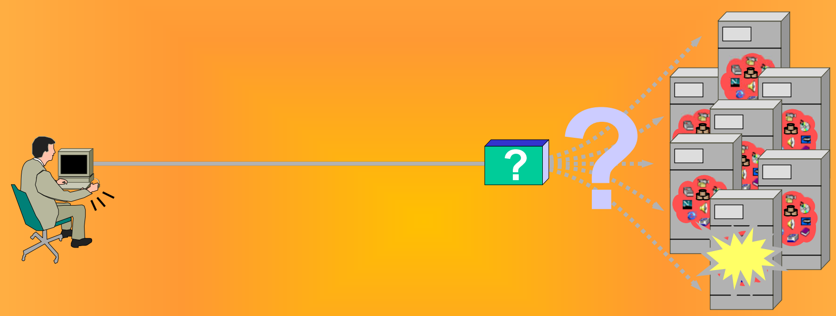
Load-Sharing Content



- Handle requests fairly amongst servers/sites
- Easily add servers/sites to content service
- Adjust connections based on server/site load

66

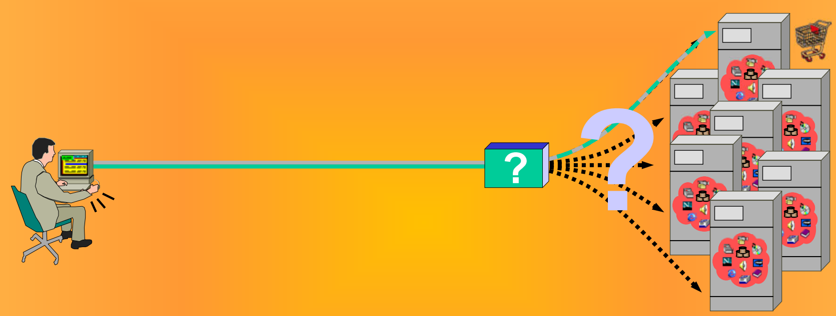
Content Availability with multiple servers?



- Synchronize content amongst servers/sites
- Avoid faulty servers/sites
- Faulty servers/sites includes invalid/dated content

67

Persistence with multiple servers?



- Handle applications which use 'state'
  - Need to learn client ID to satisfy state requirement
  - Need to maintain state for period of time - variable

68

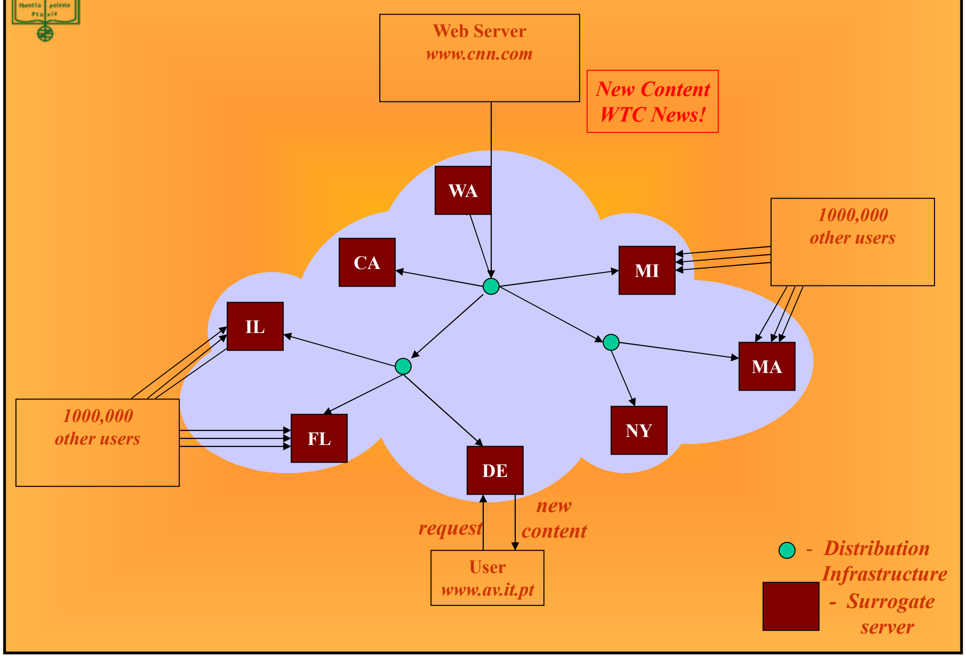
69

Outline

- Overall context
- Challenges
- Potential alternatives?
- **Architecture**

69

CDN, On September 11, 2001



70

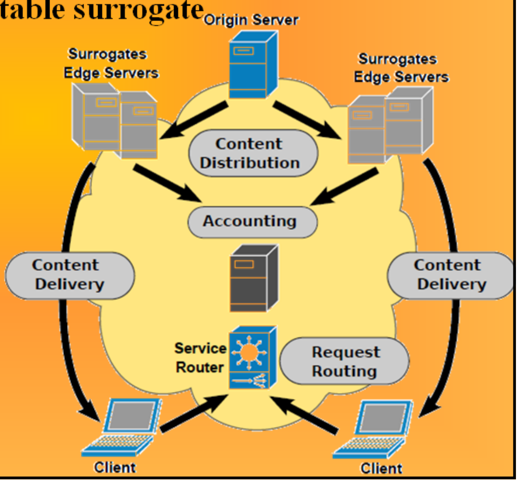
With CDNs

- Overlay network to distribute content from origin servers to users
  - Avoids large amounts of same data repeatedly traversing potentially congested links on the Internet
  - Reduces Web server load
  - Reduces user perceived latency
  - Tries to route around congested networks
- CDN is not a cache!
  - Caches are used by ISPs to reduce bandwidth consumption, CDNs are used by content providers to improve quality of service to end users
  - Caches are reactive, CDNs are proactive
  - Caching proxies cater to their users (web clients) and not to content providers (web servers), CDNs cater to the content providers (web servers) and clients
  - CDNs give control over the content to the content providers, caching proxies do not

71

CDN Components

- *Content Delivery Infrastructure:* Delivering content from producer to clients by surrogates
- *Request Routing Infrastructure:* Steering or directing content request from a client to a suitable surrogate
- *Distribution Infrastructure:* Moving or replicating content from content source (origin server, content provider) to surrogates
- *Accounting Infrastructure:* Logging and reporting of distribution and delivery activities



72

Mapping clients to servers

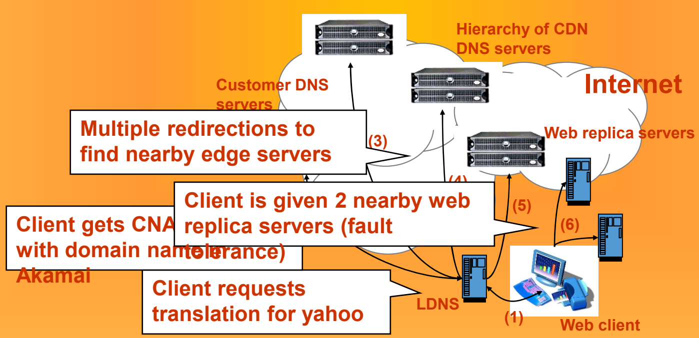
- **CDNs need a way to send clients to the "best" server**
  - The best server can change over time
  - And this depends on client location, network conditions, server load, …
  - What existing technology can we use for this?

- **DNS-based redirection**
  - Clients request www.foo.com
  - DNS server directs client to one or more IPs based on request IP
  - Use short TTL to limit the effect of caching

73

DNS Redirection

- **Web client's request redirected to 'close' by server**
  - Client gets web site's DNS CNAME entry with domain name in CDN network
  - Hierarchy of CDN's DNS servers direct client to 2 nearby servers



74

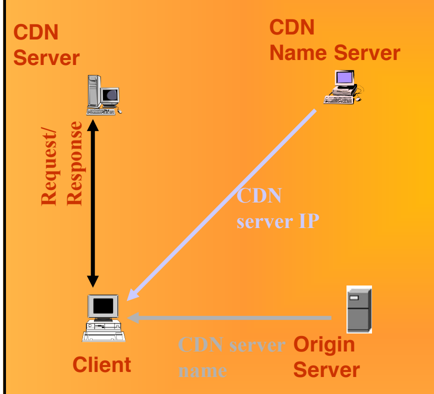
DNS Redirection Considerations

- **Advantages**
  - Uses existing, scalable DNS infrastructure
  - URLs can stay essentially the same
- **Limitations**
  - DNS servers see only the DNS server IP
    - Assumes that client and DNS server are close. Is this accurate?
  - Content owner must give up control
  - Unicast addresses can limit reliability

75

76

What other CDN techniques are being used?



- DNS redirection (DR)
  - Full-site delivery
  - Partial-site delivery
- URL rewriting
- Hybrid scheme
  - URL rewriting + DNS redirection
- Manual hyperlink selection
- HTTP redirection
- Layer 4 switching
- Layer 7 switching
- Anycast

76

77

Offloading a portal

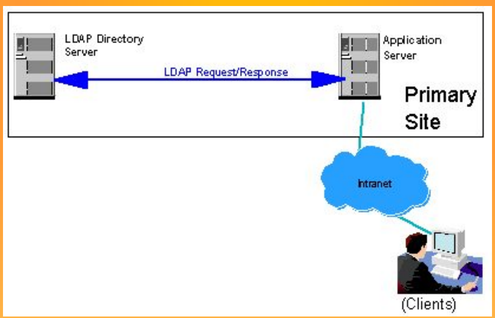
- Portal servers allow users to access content and applications from a single access point
  - Users can create persistent, customized views of applications and content chosen from the set of applications and content by the portal administrators
- Portal server pages are personalized
- Often include dynamic content
- Significant amount of computation required for page assembly
  - Application offload

77

78

Offloading an Enterprise directory

- E.g. a common e-Workplace tool
- The employee data is often stored in a central LDAP directory
  - Separate web-based application providing the interface to the directory

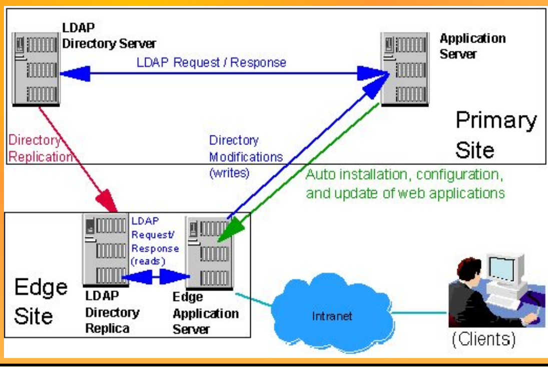


78

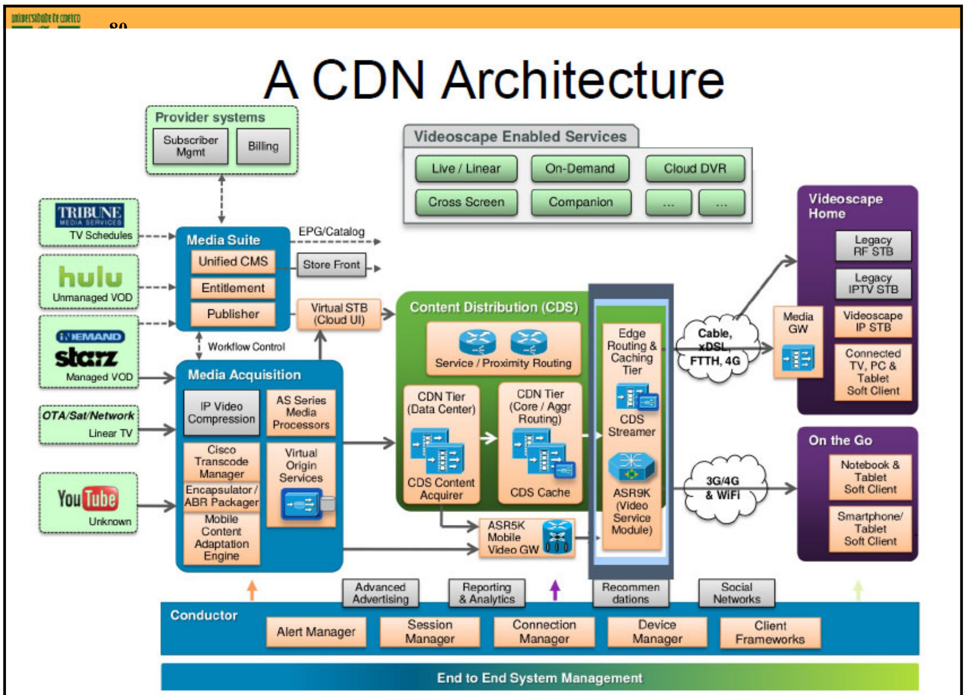
Offloading an Enterprise directory

79

- **Centralized directory**
  - **Convenient to manage**
  - **Performance for clients accessing the directory from remote sites can be poor**
    - E.g. transcontinental network connections suffer from a long delay
- **Offloaded version of the application**



79



80

81

Interconnecting (two) Large Networks

How to interconnect PSTN and ISPs

81

82

What is VoIP?

- **VoIP is not a protocol!**
  - **VoIP is a set of protocols and equipments that allow coding, transport and routing of audio calls (multimedia) through IP networks**
    - Both data (media) and signaling have to be tackled
    - Audio streams are coded in digital environment and encapsulated in IP for transport in the network.
- **Examples of VoIP inclusion (required interoperation)**
  - **PSTN → VoIP → PSTN**
  - **VoIP Native → PSTN**
  - **VoIP Native → VoIP Native**

82

83

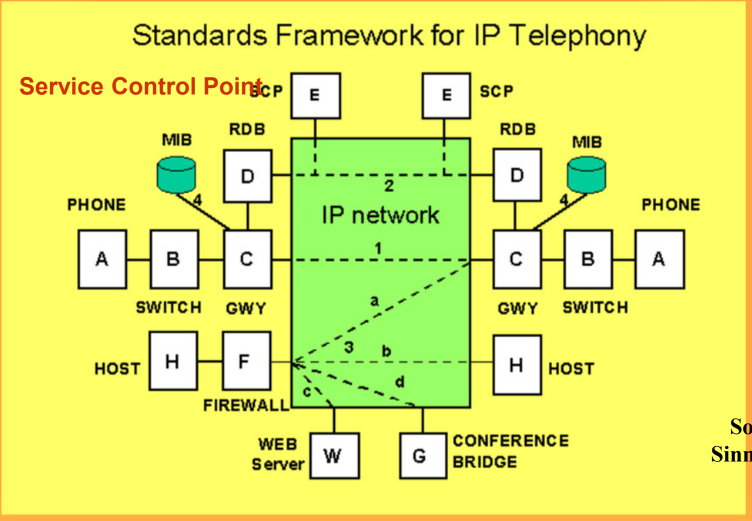
VoIP advantages

- Cost reduction
  - Do not need to pay for PSTN circuits for call transport (user side) / consolidate infrastructure (provider side)
  - Bandwidth reduction
    - Distributed nature of VoIP
    - Operation costs reduction – voice and data traffic both in the same network
- 'Open' standards and interoperability between operators
  - Does not depend on proprietary solutions
- Integration of voice and data networks
  - Considered as 'just another IP application'
  - Two major approaches: ITU-T (early on) and IETF (current)
  - As long as the quality is similar to the PSTN network, companies can easily invest in new services and applications

83

84

Voice over IP Framework



Source: Sinnreich et al

Most challenges are associated with control plane.

84

85

Different levels of VoIP problem

1. The transport level
   - How to transport multimedia information.

   Covers also content, but mostly RTP (and associated protocols)
2. The session control
   - How to signal a VoIP session.

   Covers also application protocols, but we talk mostly about SIP and H.323 or RTSP (web)
3. The gateway control
   - How to signal interface entities between Internet and POTS.

   Mostly Megaco

85

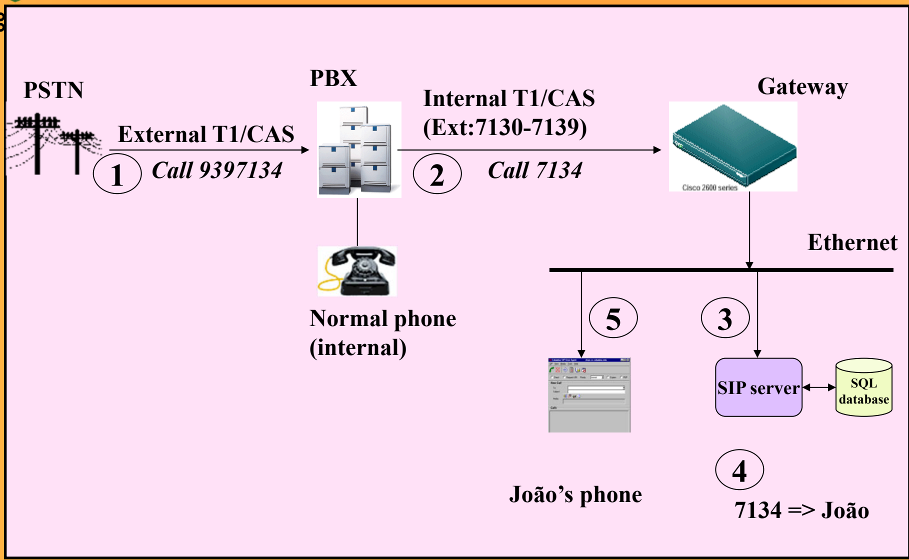
87

PSTN interoperation

- Translate audio (PCMU/PCMA)
- Translate signalling (PRI/T1,ISUP)
  - Different signalling
  - Advanced capabilities in SIP are not mapped in PSTN
- Translate identifiers (phone numbers)
- Determine transition points

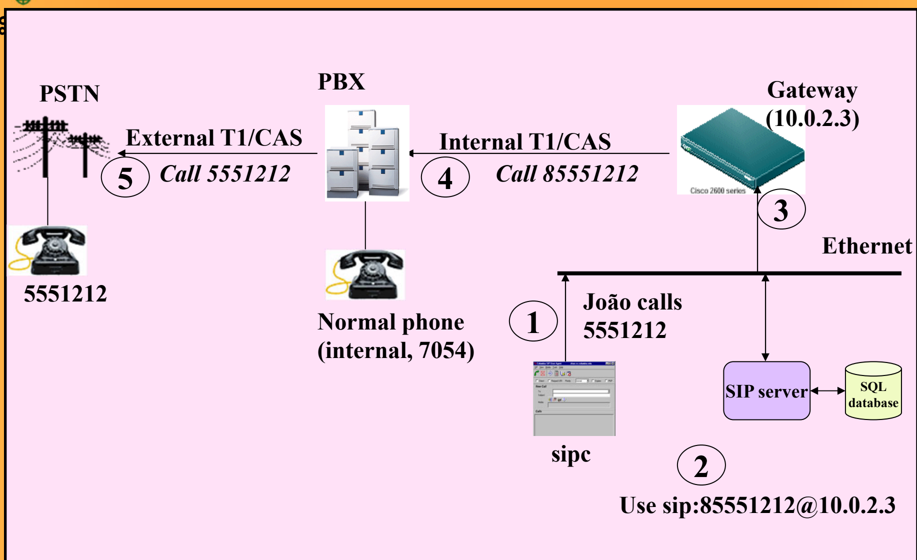
87

PSTN calls to IP



88

IP calls to PSTN



89

90

ISPs and PSTN

- Having VoIP (specially voice) sessions connecting to old-style phone networks implies:
  1. Interconnecting voice signalling
  2. Interconnecting data (voice)
     - Typically this is set by routing tables in both sides
  3. Linking both inteconnection actions
  4. Selecting where to do each one of these

90

91

What about REAL interoperation?

- Signaling boxes between the data and circuit systems must be interconnected
  - Multiple interconnection points may exist
- Systems must select best interconnection points
  - This implies best routing solution
    - And this is mixed routing – both in data and circuit systems
  - Interoperation points may be different for the data and control planes
- Different types of boxes may exist (interoperation of data/control/both)

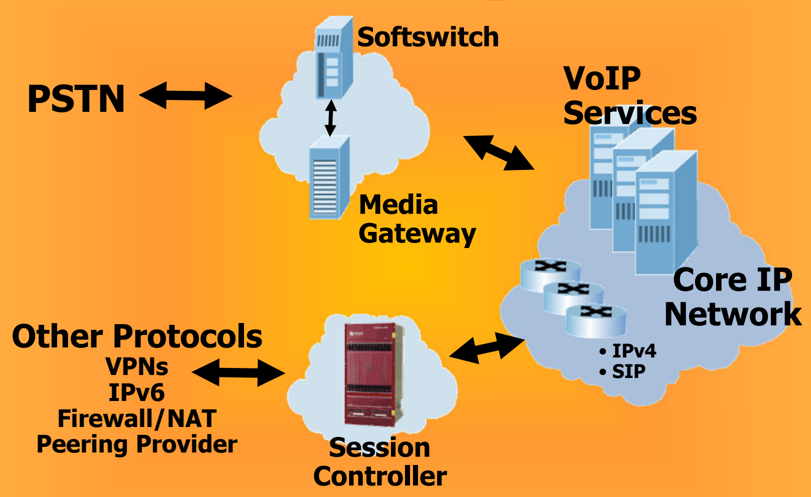
91

VoIP and PSTN Interoperability in Large Scalable Scenarios

- Requires an application programming interface and a corresponding protocol for controlling VoIP Gateways from external call control elements.
- Signaling must be inter-operable between PSTN and VoIP.
- Protocols:
  - Media Gateway Controller Protocol (MGCP) - RFC 2705
  - MGCP evolution/successor → H.248/Megaco (RFC 3015) → H.248.1/Gateway Control Protocol (RFC 3525)
    - These are control plane signaling only.
  - SIGTRAN (Signaling Transport) is the standard telephony protocol used to transport Signaling System 7 (SS7) signals over the Internet.
    - Stream Control Transmission Protocol (SCTP) – RFC 3286
      - Is an IP transport designed for transporting signaling information over an IP network.
      - Reliable transport protocol with support for framing of individual message boundaries.

92

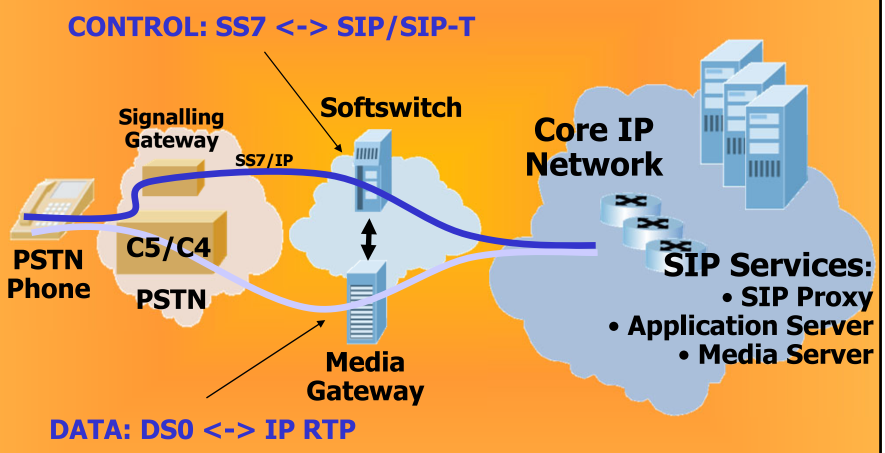
Network interoperation



Technologically addressed by media gateways and softswitches (PSTN ⇦⇨ IP) and session controllers (IP ⇦⇨ IP), and associated VoIP data plane protocols

93

PSTN/IP Interworking



Softswitches are associated with Media Gateways (often colocated in the same box), to bridge both signalling and data.

94

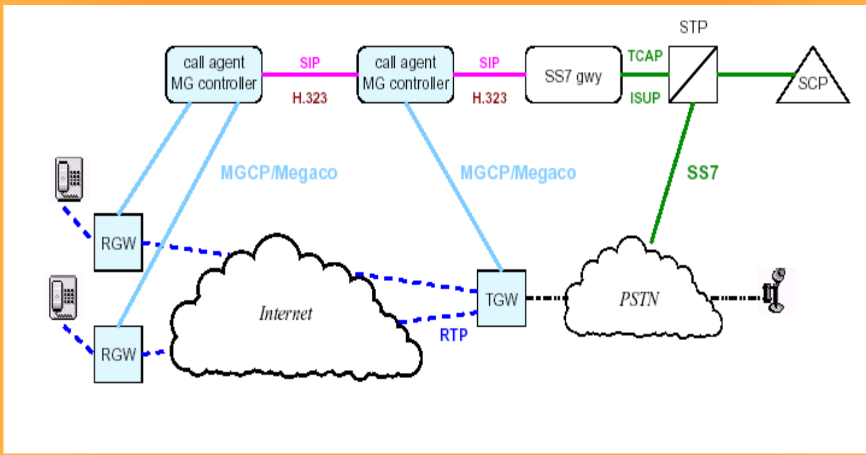
MGCP e Megaco

- Media Gateway Controller Protocol (RFC 2705)
- Controls phone Gateways resorting to *external* control elements, the media gateway controllers (MGC) a.k.a. call agents
  - Gateways: Eg: RGW (residential gateway): physical interconnection between VoIP networks and phone interfaces at homes
  - The call control “intelligence” is outside the gateways, and is controlled by external elements
  - master-slave philosophy
- Objective: scalable gateway infrastructure between PSTN and IP networks
- MGCP Successor: H.248/Megaco
- These are control plane signaling ONLY.

95

MGCP Architecture

Objective: VoIP large scalable implementations



RGW: Residential Gateway
TGW: Trunk Gateway

96

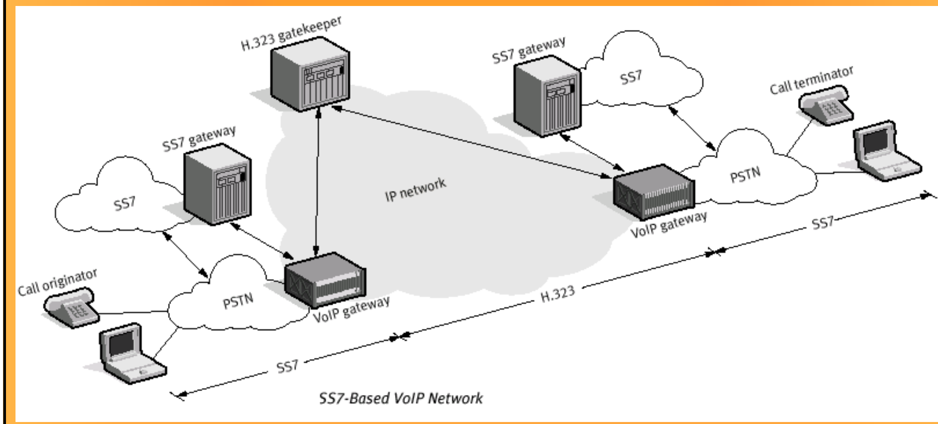
MGCP/H.248 Elements

97

- **Media Gateway Controller (MGC)**
  - **Controls the parts of the call state that pertain to connection control for media channels in a MG.**
- **Media Gateway (MG: RGW/TGW)**
  - **Converts media provided in one type of network to the format required in another type of network.**
  - **MG could terminate bearer channels from a switched circuit network (e.g., DS0s) and media streams from a packet network (e.g., RTP streams in an IP network).**
- **Signaling Gateway (SG)**
  - **Responsible for transferring signaling messages (e.g., SS7 messages) to different protocols and transports.**
    - Signaling Transport (SIGTRAN)
    - e.g., SS7 to SIGTRAN (SCTP/IP).

97

VoIP and SS7



98

Operator Networks and the FMBC

The arrival of common services

100

101

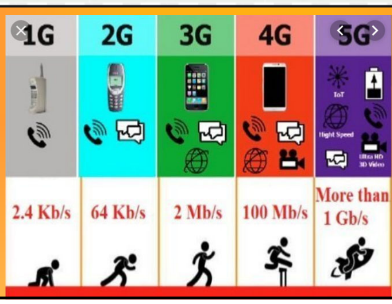
Rationale for the class

- Relate with the evolution of mobile networks
- Perceive the technologies underlying the integration of services, coming from the mobile networks, and integrating multimedia communications
- Understand the interworking of signal and data between networks
- Follow up the development of the FMC and FMBC concepts, as evolution of IP-based services

101

Wide Communication technologies: Cellular

Comparison	2G	3G	4G	5G
Introduced in year	1993	2001	2009	2018
Technology	GSM	WCDMA	LTE, WiMAX	MIMO, mm Waves
Access system	TDMA, CDMA	CDMA	CDMA	OFDM, BDMA
Switching type	Circuit switching for voice and packet switching for data	Packet switching except for air interference	Packet switching	Packet switching
Internet service	Narrowband	Broadband	Ultra broadband	Wireless World Wide Web
Bandwidth	25 MHz	25 MHz	100 MHz	30 GHz to 300 GHz
Advantage	Multimedia features (SMS, MMS), internet access and SIM introduced	High security, international roaming	Speed, high speed handoffs, global mobility	Extremely high speeds, low latency
Applications	Voice calls, short messages	Video conferencing, mobile TV, GPS	High speed applications, mobile TV, wearable devices	High resolution video streaming, remote control of vehicles, robots, and medical procedures



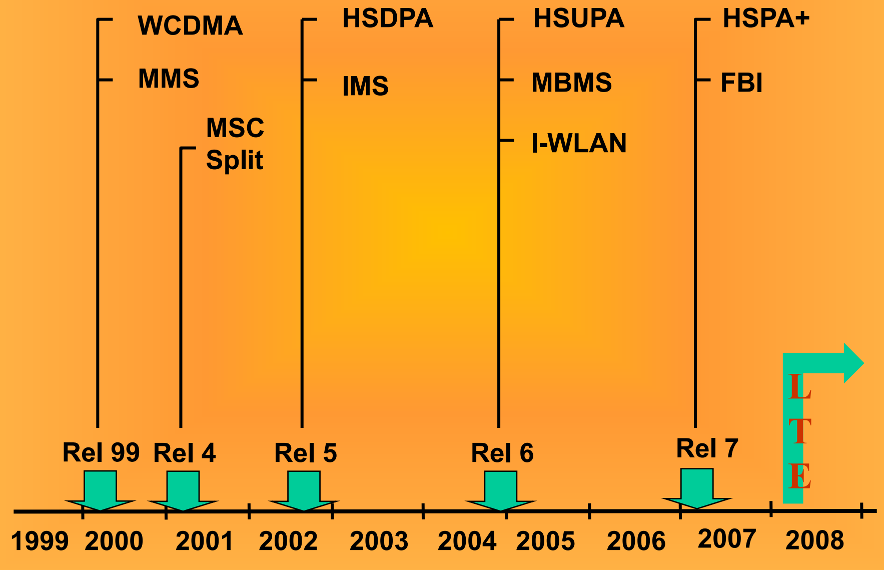
102

Early cellular systems

- 1G: analog systems (450-900 MHz)
  - Signalling: FSK
  - Share of medium: FDMA
  - NMT (Europe), AMPS (US)
- 2G: digital systems (900, 1800, 1900 MHz)
  - Share of medium : TDMA/CDMA
  - Circuit switching
  - GSM (Europe), IS-136 (US), PDC (Japan)
- 2.5G: extensions for packet switching
  - Digital: GSM → GPRS
  - Analog: AMPS → CDPD
- 3G: networks for data applications
  - High rates, data, Internet
  - Share of medium : TDMA/CDMA/CDMA
  - IMT-2000 (Europe: UMTS)

103

3GPP Release initial timeline: UMTS is phased



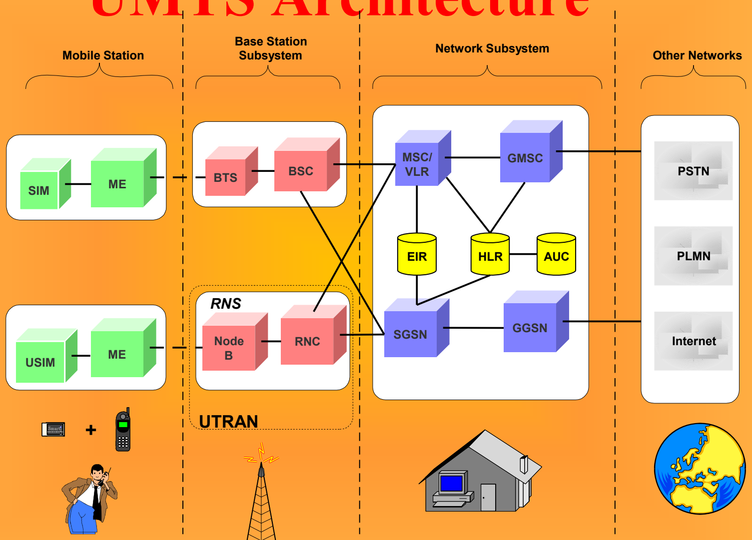
104

UMTS: first universal celular data system

- 3G system
- Oriented to generalized service diffusion and its future users trends
  - Combines cellular, wireless, paging, etc. functions
- "multimedia everywhere"
- Developed as an evolution path of 2.5G systems
  - Progressive evolution (GPRS-EDGE-UMTS)
- Direction towards IP networking, and relying of data services

105

UMTS Architecture

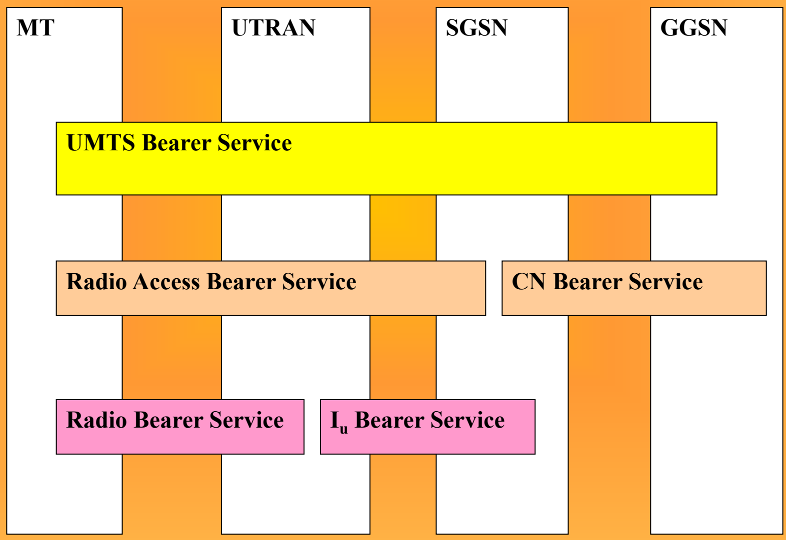


UMTS takes advantage of the existing GSM and GPRS networks: circuit and data paths
The main difference comes with the new radio interface.

106

108

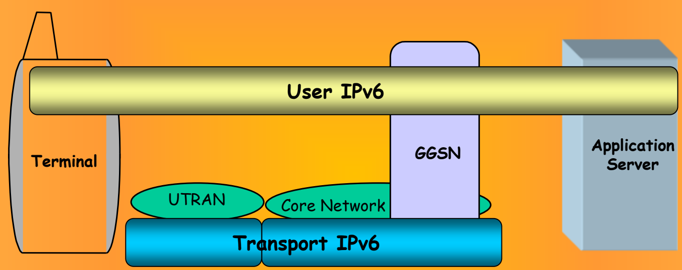
UMTS global architecture



108

109

User plane vs transport plane

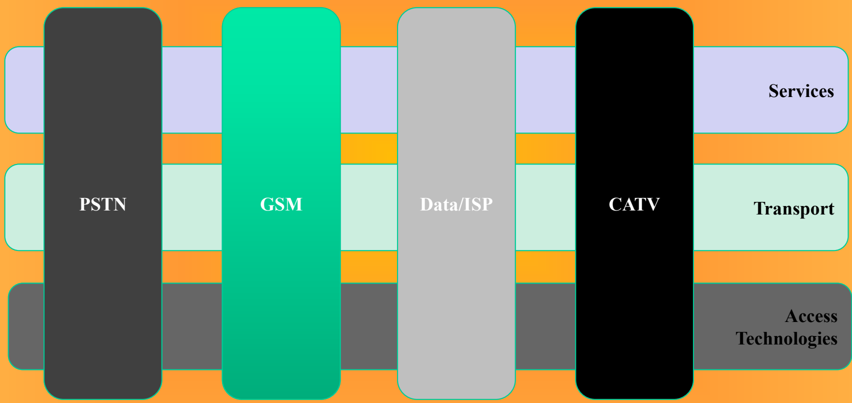


- User and transport planes are completely independent, i.e. the transport plane can run on a different IP version than the user plane
- UTRAN and Core Network transport can also run on different IP versions

109

110

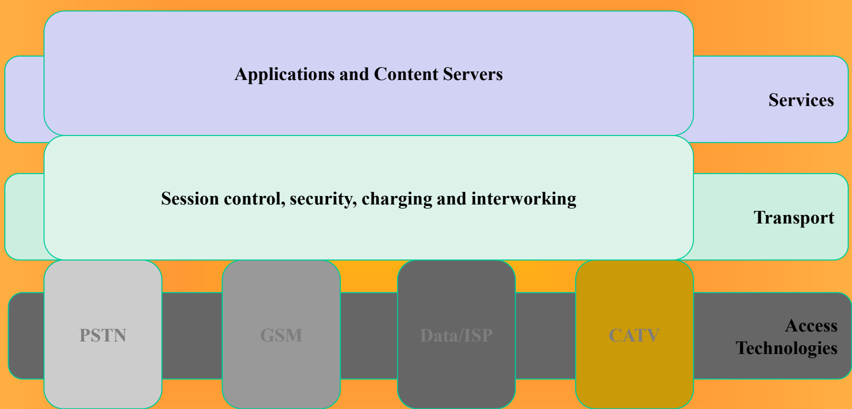
Legacy Telecom Model



110

111

3GPP/TI-SPAN Telecom Model



The topology is moving from a vertical to a horizontal layered architecture with a common IP-based connectivity network for all accesses, mobile, fixed, and data.
Relying in IP allows technology independence.

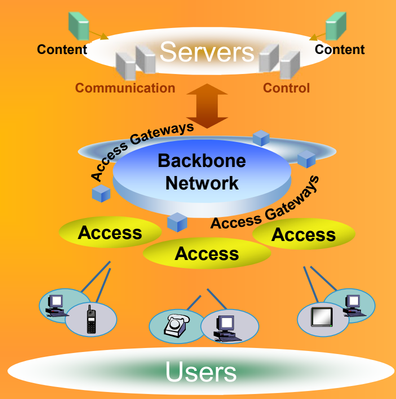
111

112

UMTS from release 5 on: IMS: IP Multimedia Subsystem

© Rui L. Aguiar (ruilaa@det.ua.pt) - Uni. Aveiro

- **Same Core network**
- **Same User on different accesses**
- **Same Services**
- **Can use WLAN, ADSL, LAN, UTRAN (GPRS) etc. as accesses in ONE system**
- **Can have several devices and move between them**
- Addresses circuit-oriented session concepts!



112

IP Multimedia Subsystem (IMS)

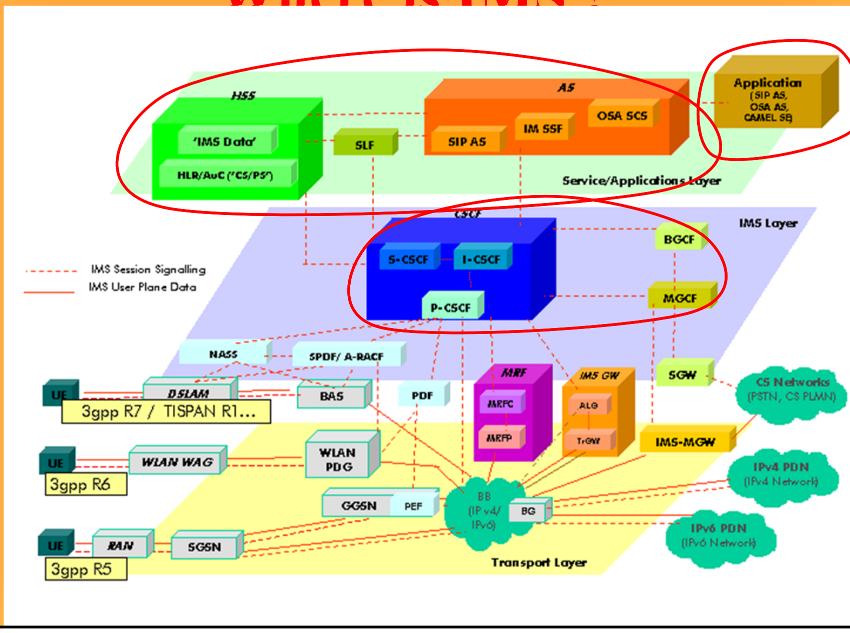
- **New IP-based mobile core network for 3G evolution**
- **Uses 3GPP variant of SIP & other IP protocols**
- **"Intelligent Network" over IP?**
- **New services drive IMS deployment**
  - **Push-to-Talk, FMC, IP Centrex**
- **PTT (PoC) & UMA FMC specs already turned over to 3GPP**
- **Developed by 3GPP for GSM-to-3G evolution**
  - **Defined in release 5; fully specified in release 6**

114

115

Where is IMS ?

3GPP IMS - Eng. de Serviços - Rui Aguiar/Diogo Gomes <dgomes@ua.pt>



115

116

x-CSCF

- Serving - CSCF
  - Controls the user's SIP Session
  - very few per domain
  - Located in the home domain
  - Is a SIP registrar (and proxy)
- Proxy-CSCF
  - IMS contact point for the user's SIP signaling
  - Several in a domain
  - Located in the visited domain
  - Terminals must know this proxy (e.g. DHCP used)
  - Compresses and decompresses SIP messages
  - Secures SIP messages
  - Assures correctness of SIP messages
- Interrogating – CSCF
  - domain's contact point for inter-domain SIP signaling
  - one or more per domain
  - In case there are more than one S-CSCFs in the domain, locates which S-CSCF is serving a user

116

118

SIP Protocol

- Defined in IETF RFC 3261
  - "… an application-layer control (signaling) protocol for creating, modifying, and terminating sessions with one or more participants. These sessions include Internet telephone calls, multimedia distribution, and multimedia conferences."
- In IMS, SIP is modified to include extra functionality and support a specific set of functions only
  - SIP is to the Internet what SS#7 is to telephony
- At the core of IMS there are several SIP proxies:
  - I-CSCF, S-CSCF, P-CSCF
  - The Call Session Control function (CSCF) is the heart of the IMS architecture
  - The main functions of the CSCF:
    - provide session control for terminals and applications using the IMS network
    - secure routing of the SIP messages,
    - subsequent monitoring of the SIP sessions and communicating with the policy architecture to support media authorization.
    - responsibility for interacting with the HSS.

118

119

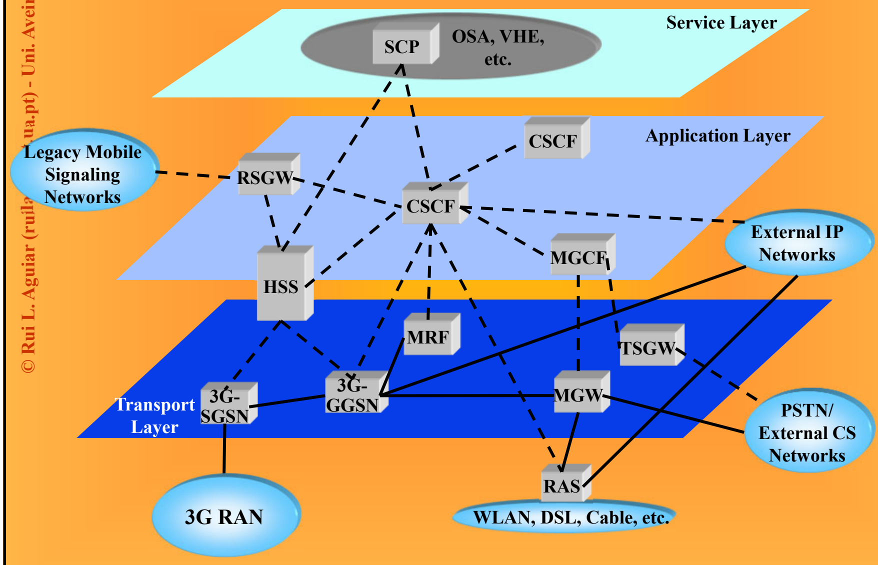
IMS Identity and User Profiles

- IMS uses SIP identity: SIP URIs
  - e.g. sip:ruilaa@ua.pt
  - Opposed to phone numbers
  - A user is uniquely identified in the HSS by his IMPI (Private User Identity).
  - IMPI is a unique global identity defined by the Home operator
  - used only in the process of registration
- to establish communication with a user IMPU (Public User Identity) is necessary.
  - Every user has one or more IMPUs.
  - Each IMPI can have several IMPUs
  - Users can classify their public identities: business, family, friends, …
    - E.g. sip:ruilaa@ua.pt, sip:steve.jobs@left.apple.com

119

120

UMTS Layered Approach for the PS services



120

121

Implications: what can be achieved with this add-on

- Any Device
- Any Access Technology
- Any Where

ALWAYS BEST CONNECTED

- One Network, multiple access technologies
- Common Session Control
- Generic Application Servers
- Single set of services that apply network wide
- Consistent user experience
- Operational efficiency
- New services/applications

121

Fixed Mobile (Broadcast) Convergence - FM(B)C

- One customer service
  - Handles mobile and fixed calls
  - Any network — mobile, WiFi, Broadband Cable…
  - Avoid mobile charging when in-building
- Single (customer) number with common suite of services
  - One voice mailbox, one phone directory…
  - Mobile, fixed, conference room
- New services? Irrespective of location, access technology or terminal device
  - Potentially gradative provision

Slide 122

122

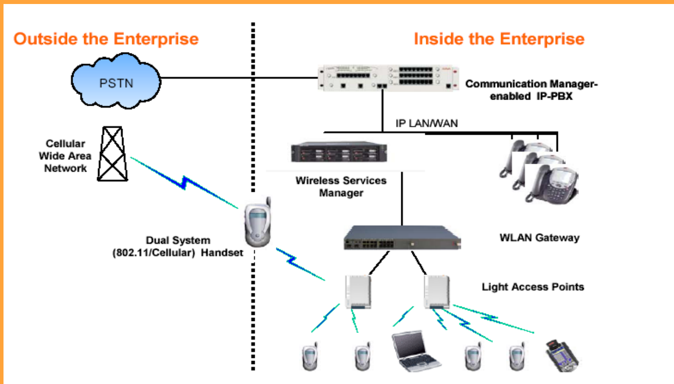
Implementing the FMBC concepts

- IP Multimedia Subsystem (IMS)
  - 3G vision of future IP-based mobile communications
- IP-PBX or softswitch with mobile network interface
  - Centered in company internal communications
- Wireless "fixed" line services
  - New (not FMBC) in developing nations, mobile, no handoffs
- Unlicensed Mobile Access (UMA)
  - voice & data services over WiFi

Slide 127

127

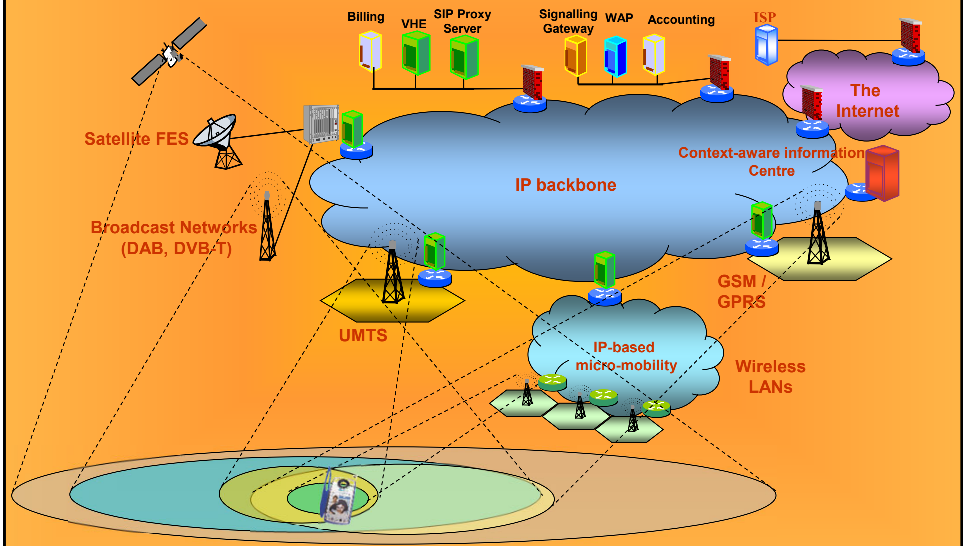
IP-PBX/Softswitch & Mobile Gateway



- IP-PBX or softswitch is in charge -> directed to company contacts
- Service hands off to mobile network when out of WLAN range
- Mobile network only used when necessary
  - "Tunnel" through the mobile network
  - Voice or text messages generated when necessary to connect enterprise calls to remote users

128

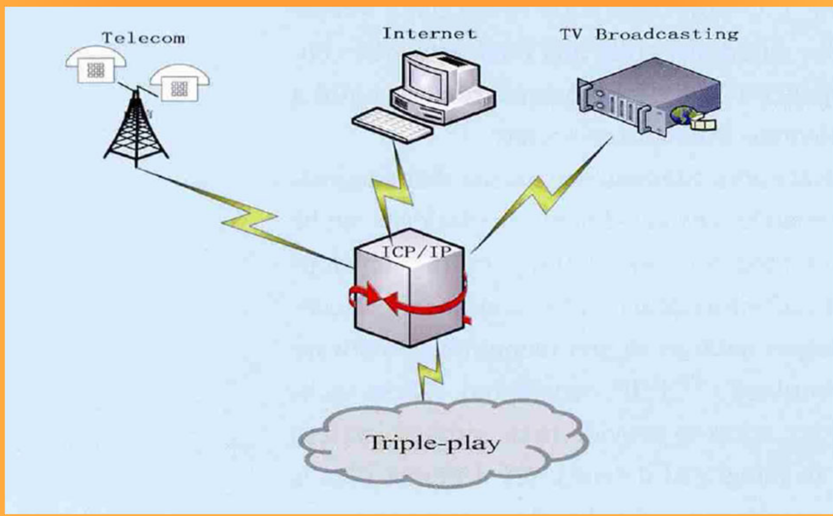
Inter-Working



130

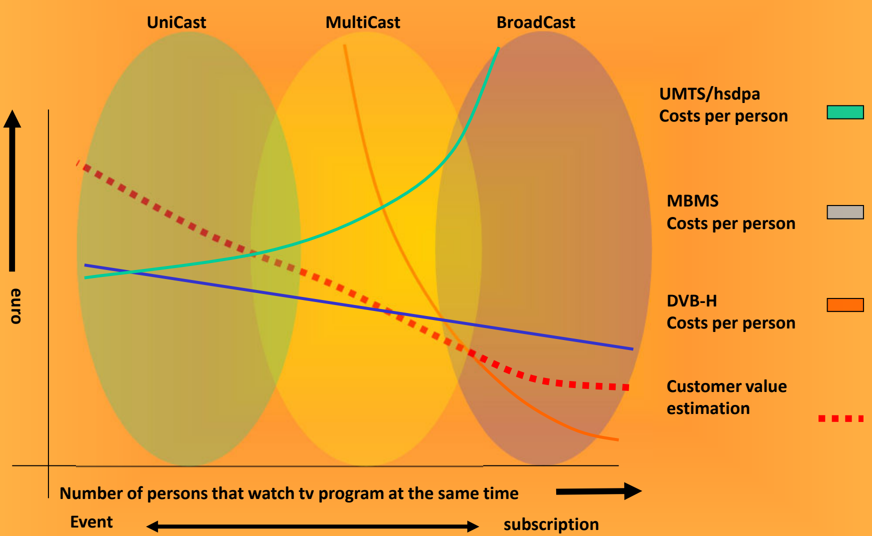
131

New service (FBC model): triple play



131

The broadcast problem: when to use it



Delivery cost vs. value for customer mobile bearers

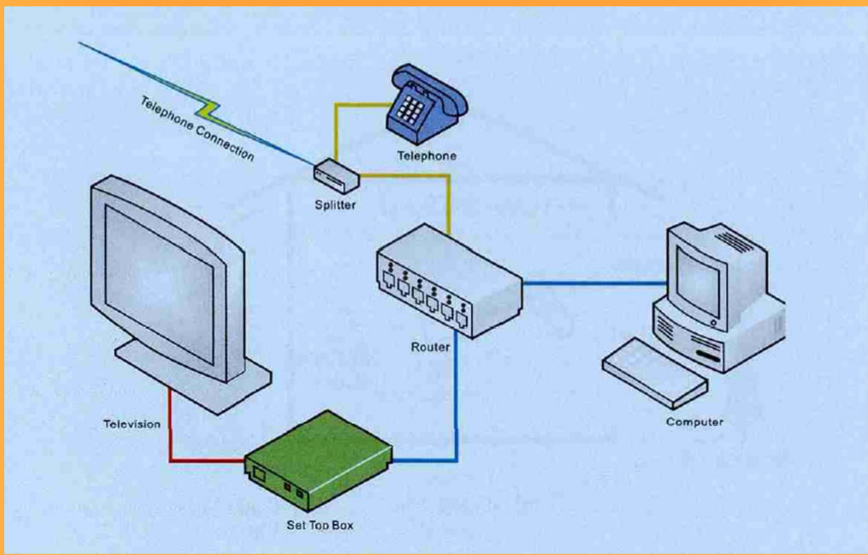
132

132

133

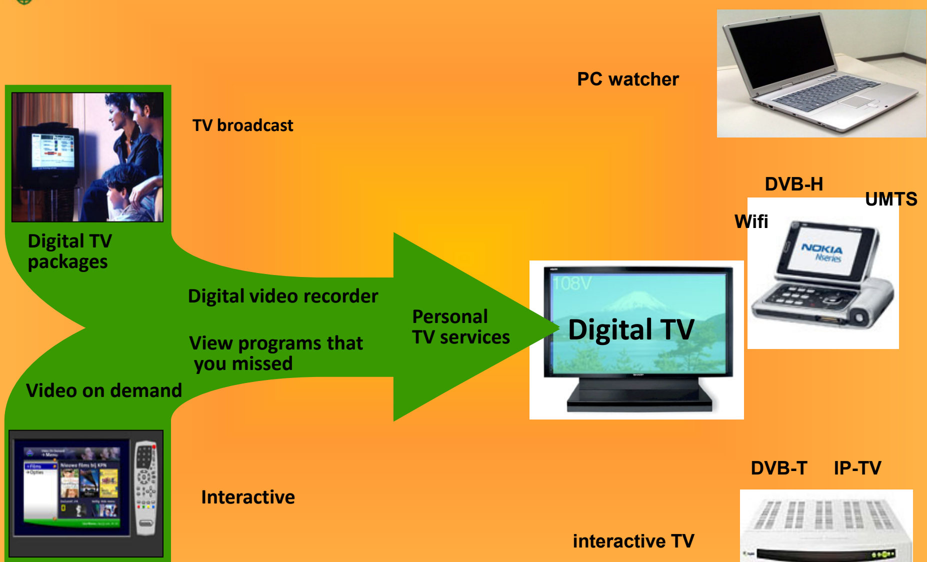
New methods of TV delivery

© Rui L. Aguiar (ruilaa@det.ua.pt) - Uni. Aveiro

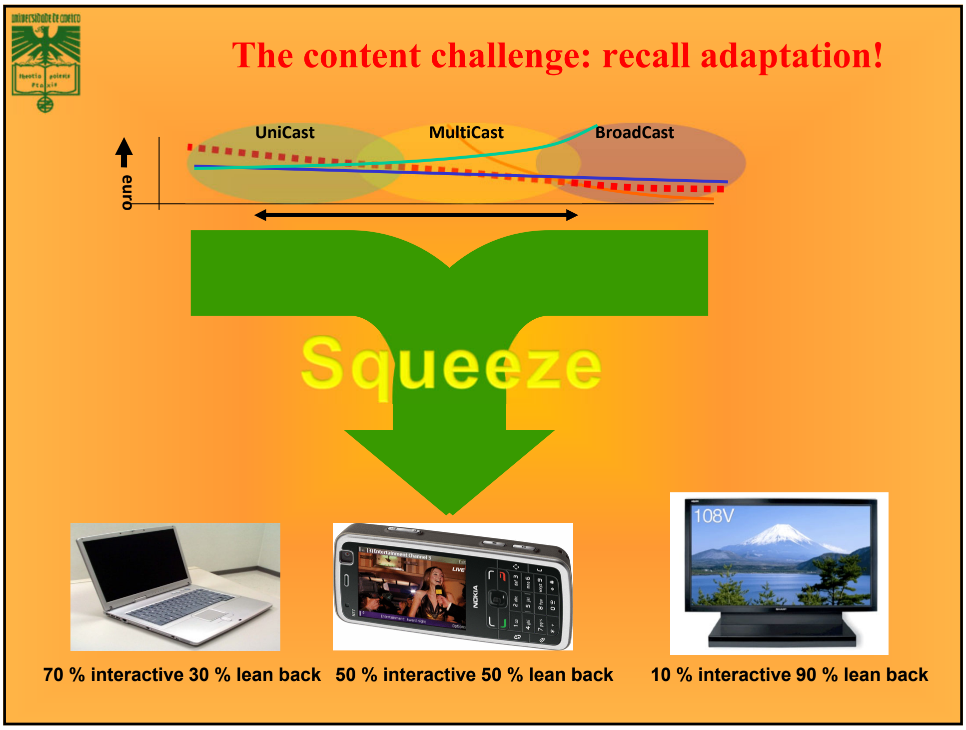


133

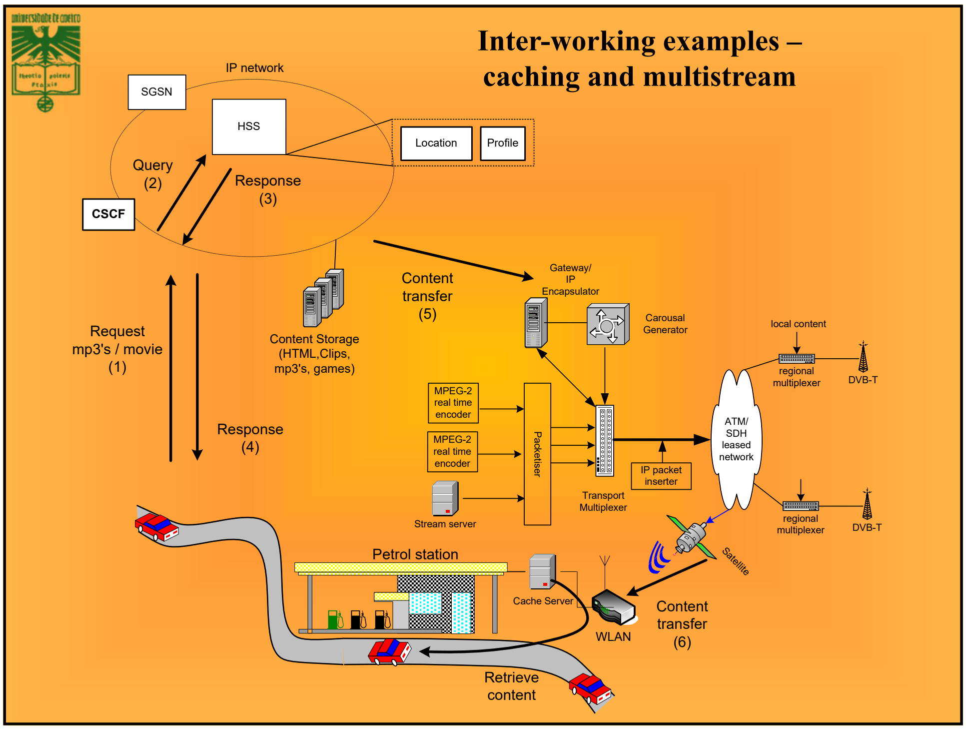
Potential of IP services for TV: Arriving of the multi- screen



134


The content challenge: recall adaptation!
UniCast
MultiCast
BroadCast
euro
Squeeze
108V
70 % interactive 30 % lean back
50 % interactive 50 % lean back
10 % interactive 90 % lean back

135

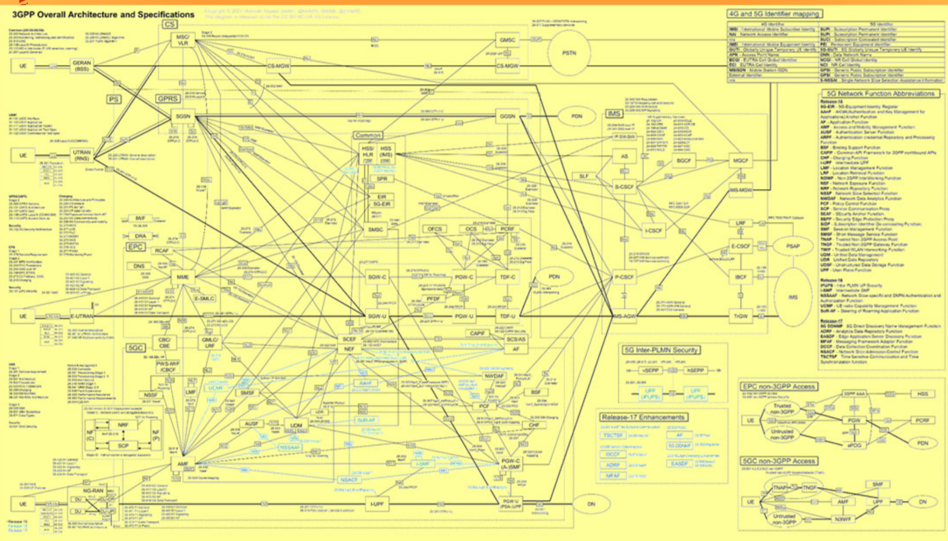

Inter-working examples –
caching and multistream
IP network
SGSN
HSS
Location
Profile
Query
(2)
Response
(3)
CSCF
Content transfer
(5)
Gateway/
IP
Encapsulator
Carousal Generator
Request
mp3's / movie
(1)
Content Storage
(HTML,Clips,
mp3's, games)
local content
regional multiplexer
DVB-T
MPEG-2 real time encoder
MPEG-2 real time encoder
Packetiser
ATM/
SDH
leased
network
IP packet inserter
Response
(4)
Stream server
Transport Multiplexer
regional multiplexer
DVB-T
Satellite
Petrol station
Cache Server
WLAN
Content transfer
(6)
Retrieve content

136

137

Example: current mobile network

(without IP/MPLS/CDN/DWDM…)



137

138

Conclusions

© Rui L. Aguiar (ruilaa@det.ua.pt) - Uni. Aveiro

- Overall large comercial networks are very complex
  - "Transport"/"ISP" is Multidimensional
  - Multistakeholder
  - Very large
  - Limited by costs
- No overall architecture vision easy to grasp exist, not even in comercial operators
  - Different departments handle different aspects

138