CENTER for TELECOMMUNICATIONS RESEARCH

KING'S College LONDON
University of London

Quality of Service (QoS)
3G and Beyond

Prof. Hamid Aghvami
Centre for Telecommunications Research
King's College London

Interworking/Integration

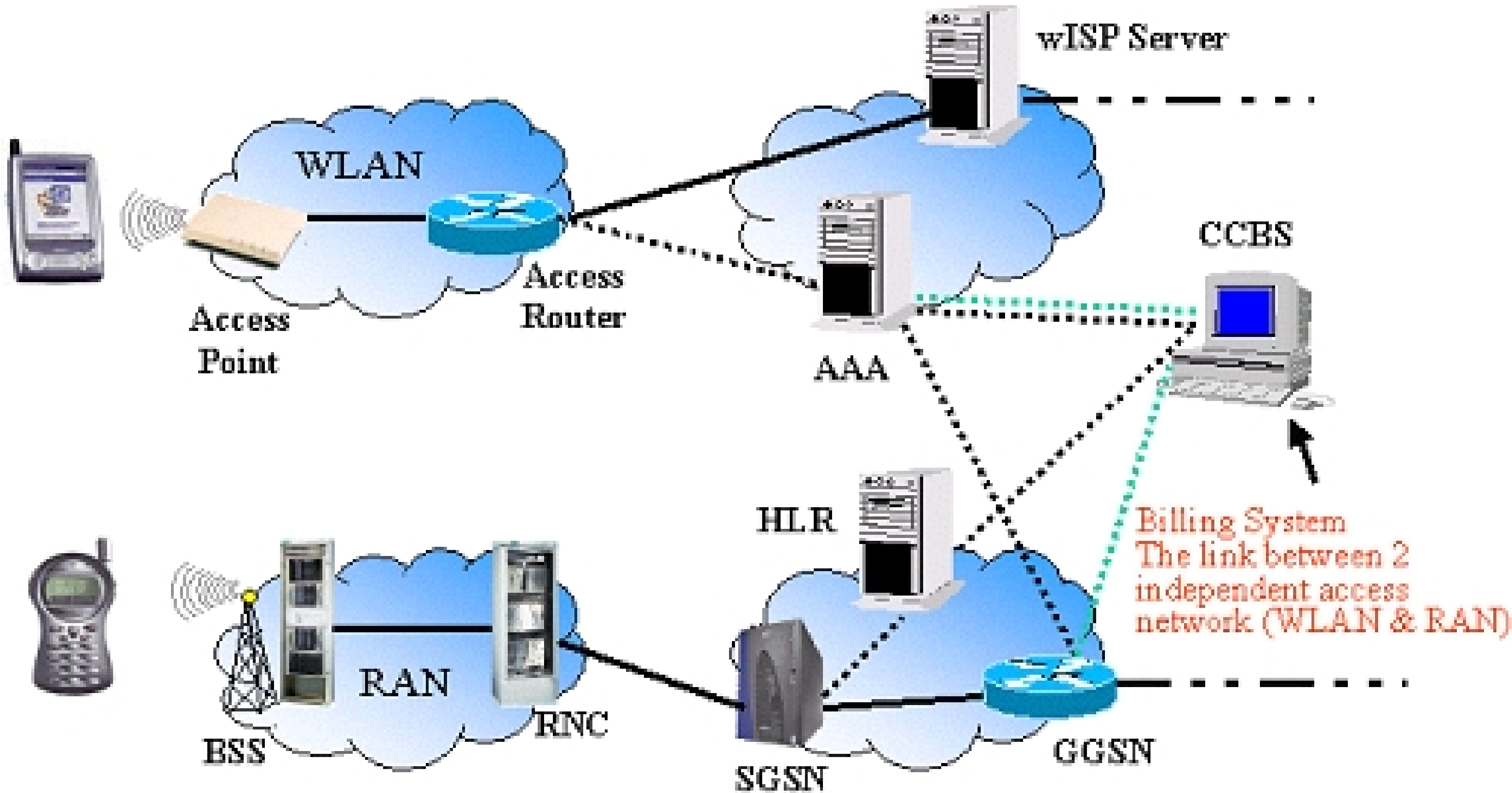
- For the design of next generation wireless networks two different approaches are currently being considered. They are:
  - Interworking with next generation Internet (tight coupling)
  - Integration within next generation Internet (very tight coupling)

- In the first approach, the access network and the core network use different IP protocols and mechanisms and only the core network is considered as a sub-network of the Internet.

- In the second approach, both the core and access networks use common IP based protocols and mechanisms and the access network is considered as a sub-network of the Internet.

C ENTER for
T ELECOMMUNICATIONS
R ESEARCH

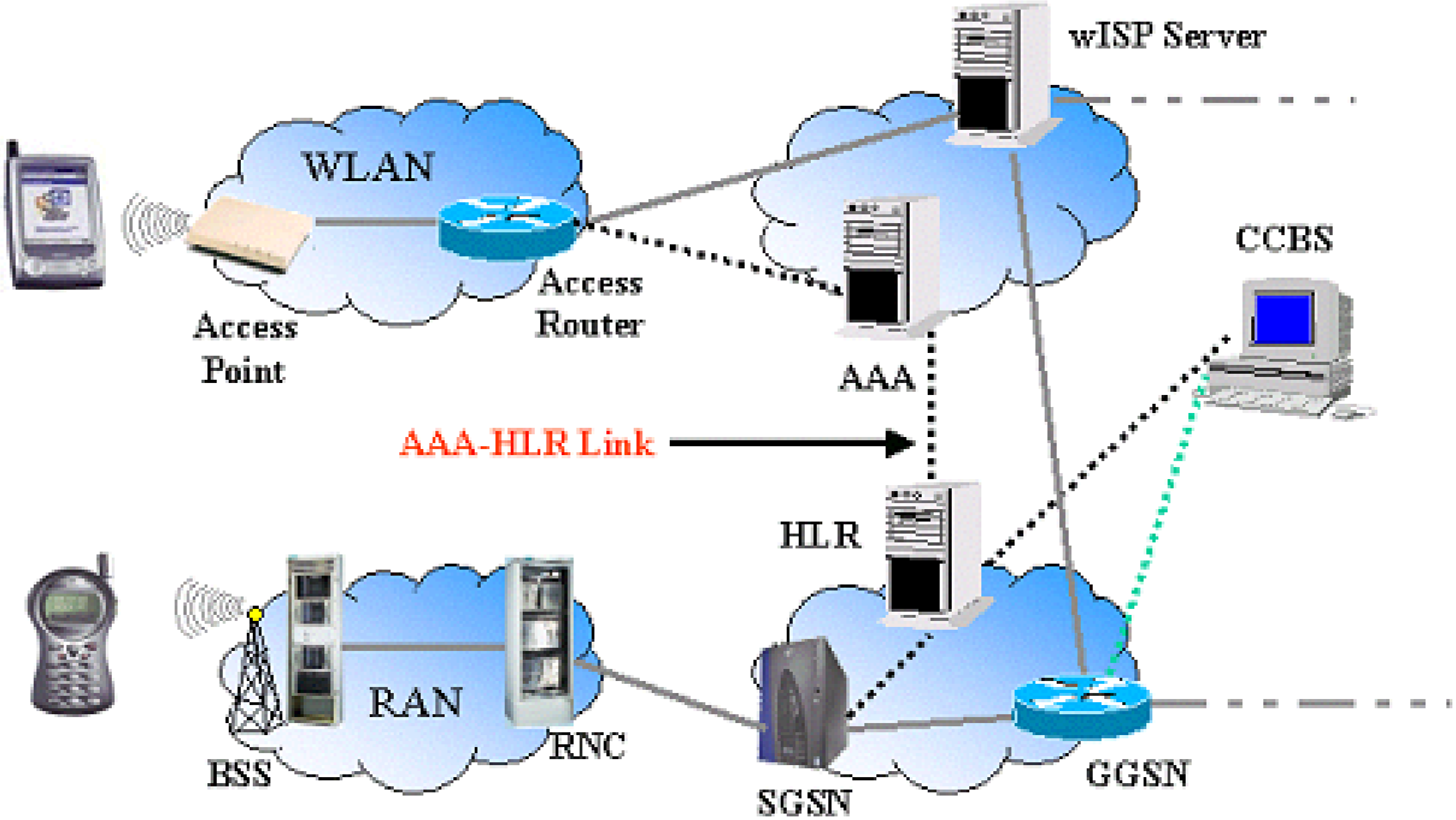
KING'S College LONDON
University of London



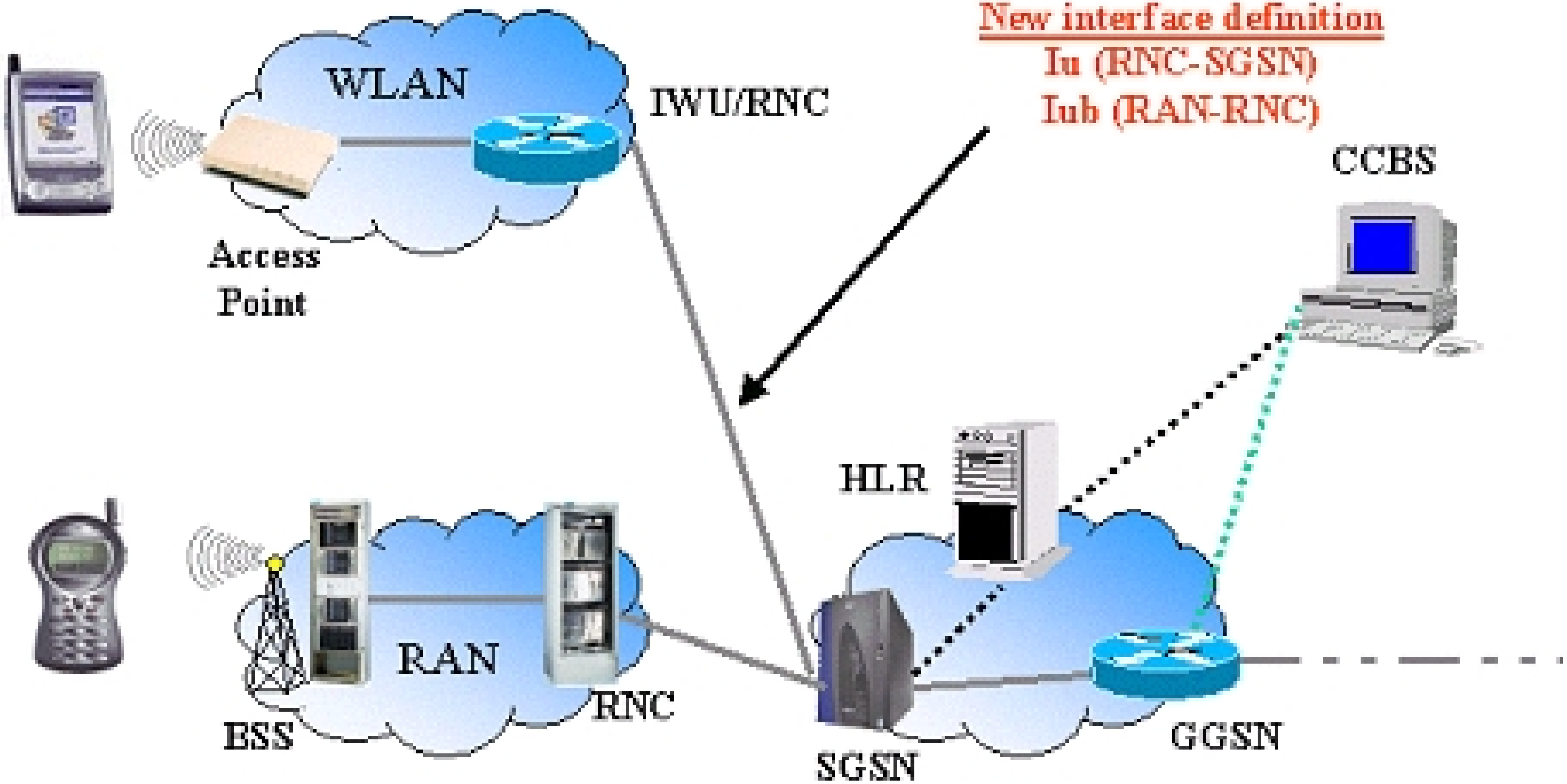
Open coupling

ENTER for
ELECOMMUNICATIONS
ESEARCH

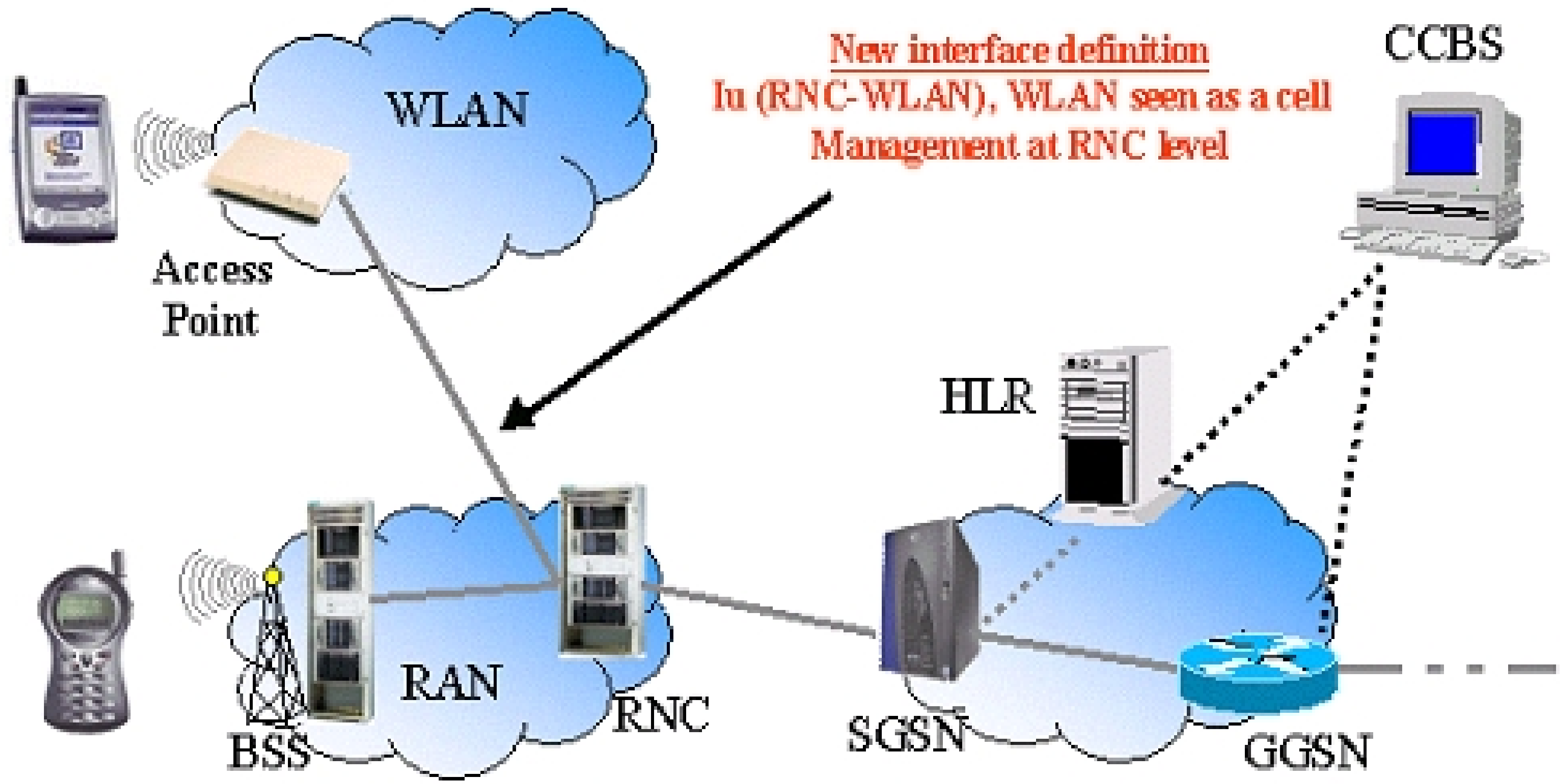
KING'S College LONDON
University of London



Loose coupling



Tight coupling



Very Tight coupling

Three main functionality groups of an access network are:

- Mobility management

- QoS

- Inter-layer interfaces, especially above the TCP/IP stack (towards application layers) and below (towards wireless link layers)

Mobility Management

- The mobility management of a wireless access network includes how to solve the problems of local seamless handover, idle mode and paging, and of course the routing capabilities within the access network that allow these to take place without constant address reassignment.

- All of these, especially the last, have major impacts on the internal architecture of the access network, in terms of location of functionality in different components.

QoS

• The problems are:

- How to identify the specific problems of mobility and radio access impacts on the provision of end-to-end QoS.

- What requirements on external networks might be needed to enable QoS mechanisms within the access network to operate.

- How to adapt QoS concepts for the fixed Internet to the mobile wireless environment.

Inter-Layer Interfaces

- How to solve the network layer problems that could take into account application layer requirements and could be implemented over real link layers.

UMTS Quality of Service

- In the context of UMTS networks, QoS refers to the collective effect of service performance that determines the degree of satisfaction of the end-user of the service.

- The QoS architecture consists of all the UMTS network entities and functions that participate in providing the end-user with the appropriate service quality.

End User QoS Requirements

- Only the QoS perceived by end-user matter.
- The number of user defined/controlled attributes has to be as small as possible.
- Derivation/definition of QoS attributes from the application requirements has to be simple.
- QoS attributes shall be able to support all applications that are used, a certain number of applications have the characteristic of asymmetric nature between two directions, uplink/downlink.
- QoS definitions have to be future proof.
- QoS has to be provided end-to-end.

attributes = parameters of a specific QoS class

General Requirements for QoS

- QoS concept should be capable of providing different levels of QoS by using UMTS specific control mechanism.
- All attributes have to have unambiguous meaning.
- QoS mechanism have to allow efficient use of radio capacity.
- Allow independent evolution of core and access networks.
- Allow evolution of UMTS network (i.e., eliminate or minimise the impact of evolution of transport technologies in the wireline world).
- All attribute combinations have to have unambiguous meaning.

CENTER for TELECOMMUNICATIONS RESEARCH

Technical Requirements for QoS

KING'S College LONDON University of London

- UMTS QoS control mechanisms shall provide QoS attribute control on a peer to peer basis between UE and 3G gateway node.
- UMTS QoS mechanisms shall provide a mapping between application requirements and UMTS services.
- UMTS QoS control mechanisms shall be able to efficiently interwork with current QoS schemes.
- UMTS shall provide a finite set of QoS definitions.
- Overhead and additional complexity of the QoS scheme should be kept reasonably low.
- QoS shall support efficient resource utilisation.
- QoS attributes are needed to support asymmetric beares.

- QoS behaviour should be dynamic, i.e., it should be possible to modify QoS attributes during an active session.

- Number of attributes should be kept reasonably low.

- User QoS requirements shall be satisfied by the system, including when change of SGSN within the core network occurs.

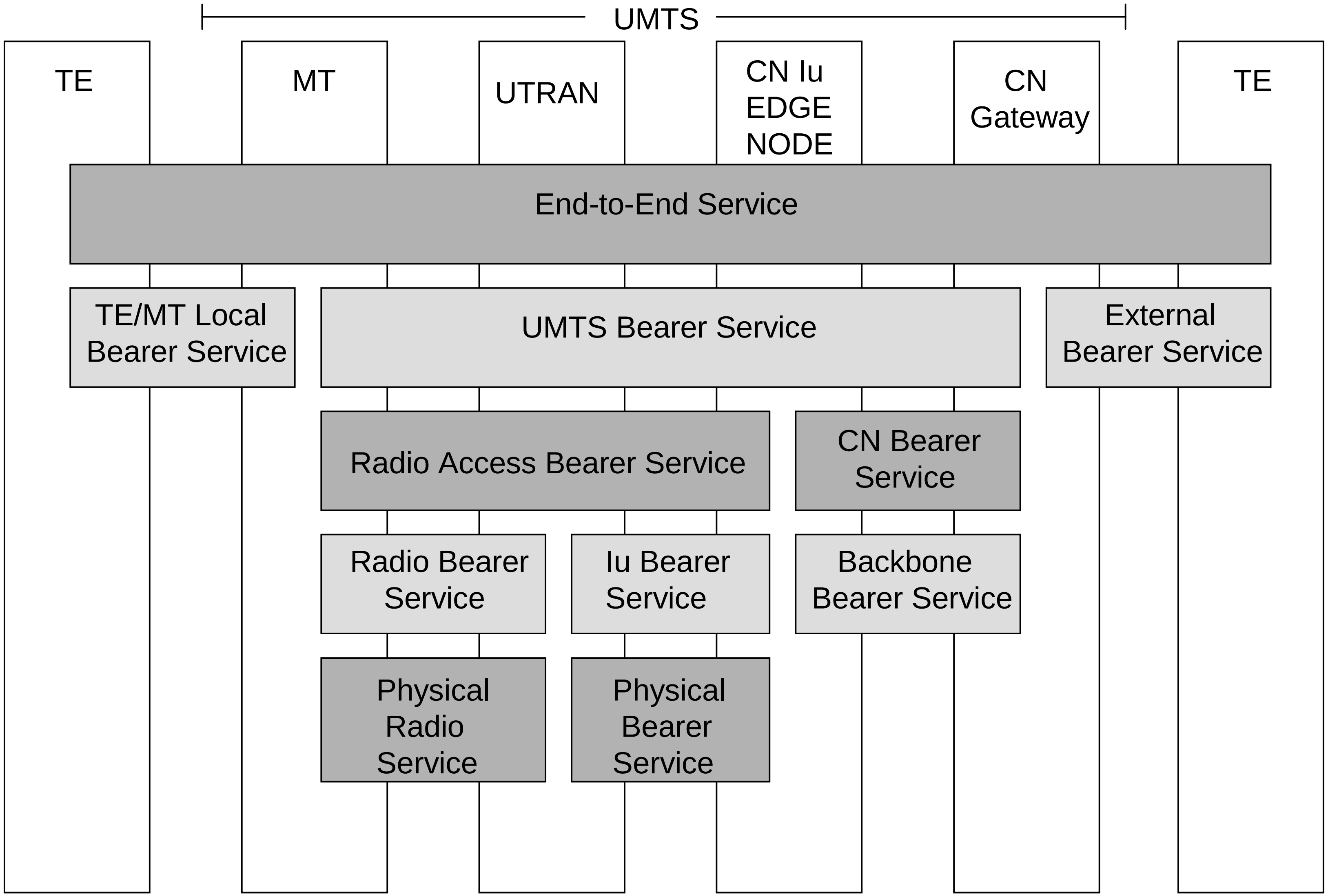
UMTS Bearer Service

- QoS support in UMTS is based on the concept of bearer service.
- A bearer is a logical connection between two end points with specific service capabilities.
- A bearer service includes all aspects to enable the provision of a contracted QoS (e.g., controlling, signalling, user plane transport, management functionalities).
- Each bearer service on a specific layer offers it’s individual services using services provided by the layers below.

CENTER for TELECOMMUNICATIONS RESEARCH

End-to-End QoS Architecture

KING'S College LONDON
University of London



- The End-to-End service on the application level uses the bearer services of the underlying networks.

- The End-to-End service used by the TE will be realised using a TE/MT Local Bearer Service, a UMTS Bearer Service, and as External Bearer Service.

- The UMTS Bearer Service, that the UMTS operator offers, provides the UMTS QoS.

- The UMTS Bearer Service consists of two parts:

    - Radio Access Bearer Service
    - Core Network Bearer Service

- The Radio Access Bearer Service provides confidential transport of signalling and user data between MT and CN Iu Edge Node with the QoS adequate to the negotiated UMTS Bearer Service.

- The Radio Access Bearer Service is realised by a Radio Bearer Service and an Iu-Bearer Service.

- The role of the Radio Bearer Service is to cover all the aspects of the radio interface transport. This bearer service uses the UTRA FDD/TDD.

- The Iu-Bearer Service together with the physical Bearer Service provides the transport between UTRAN and CN.

- The core Network Bearer Service uses a generic Backbone Network Service.

- The Backbone Bearer Service covers the layer 1/layer 2 functionality and is selected according to operator's choice in order to fulfil the QoS requirements of the core Network Bearer Service.

- The Backbone Network Service is not specific to UMTS but may reuse an existing standard.

QoS Management Functions

- The QoS Management Functions are needed to establish, modify and maintain a UMTS Bearer Service with a specific QoS.

- The QoS Management Functions are located in the control plane and the user plane.

- In the control plane:
  - Service manager
  - Translation function
  - Admission / Capability control
  - Subscription control

- In the user plane:
  - Mapping function
  - Classification function
  - Resource manager
  - Traffic conditioner

Service Manager

- Co-ordinates the functions of the control plane for establishing, modifying and maintaining the service it is responsible for.

- Provides all user plane QoS management functions with relevant attributes.

- Performs an attribute translation to request lower layer services.

- Interrogates other control functions to receive permission for service provision.

Translation Function

- Converts between UMTS bearer service attributes and QoS attributes of the external networks.

- The service manager may include a translation function to convert between its service attributes and the attributes of a lower layer service it is using.

Admission / Capability Control

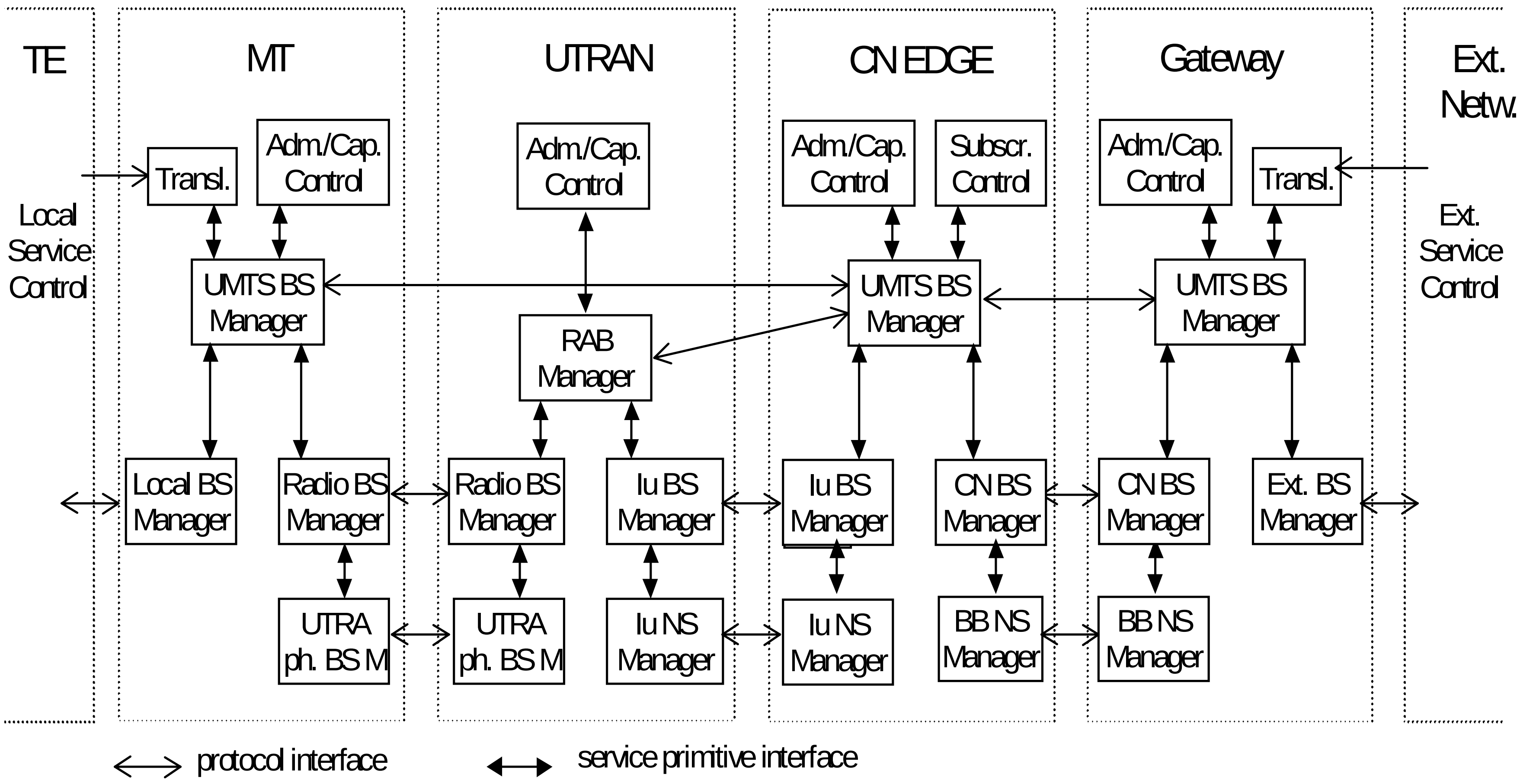
- Maintains information about all available resources of a network entity and about all resources allocated to UMTS bearer services.

- Determines whether the required resources, for each UMTS service request or modification, can be provided by this network entity and if yes, it reserves these resources.

- Checks also the capability of the network entity to provide the requested services.

Subscription Control

- Checks the administrative rights of the UMTS bearer service to use the requested service with the specified QoS attributes.



QoS management functions for UMTS bearer service in the control plane

Functions for UMTS bearer service in the user plane

- User plane QoS management functions should provide the QoS negotiated for a UMTS bearer service.

- Mapping function
  - provides each data unit with the specific marking required to receive the intended QoS by a bearer service.

- Classification function
  - assigns data units to the established services of a Mobile Terminal (MT) to the related QoS attributes if the MT has multiple UMTS bearer services established.

- Resource Manager
  - Distributes the available resources between all services sharing the same resource by means of, for example, scheduling, bandwidth management and power control for the radio bearer.

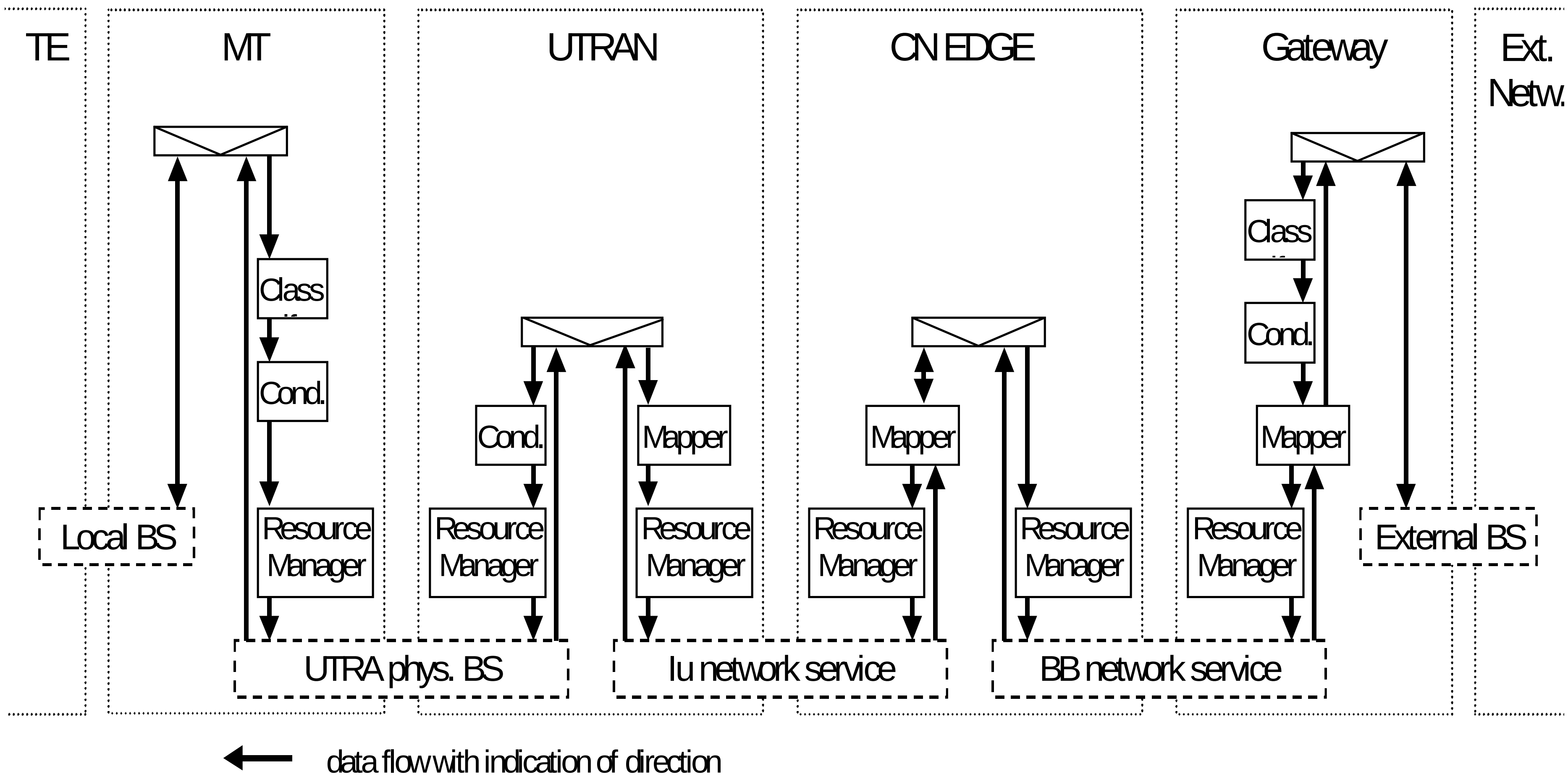
- Traffic Conditioner

  - Provides conformance between the negotiated QoS for a service and the data unit traffic.

  - Traffic conditioning is performed by policing or by traffic shaping.

  - The policing function compares the data unit traffic with the related QoS attributes. Data units not matching the relevant attributes will be dropped or marked as not matching.

  - the traffic shaper forms the data unit traffic according to the QoS of the service.



QoS management functions for the UMTS bearer service in the user plane

Traffic class	Conversational class Conversational RT	Streaming class streaming RT	Interactive class Interactive best effort	Background Background best effort
Fundamental characteristics	- Preserve time relation (variation) between information entities of the stream Conversational pattern (stringent and low delay)	- Preserve time relation (variation) between information entities of the stream	- Request response pattern - Preserve payload content	- Destination is not expecting the data within a certain time - Preserve payload content
Example of the application	- Voice	- Streaming video	- Web browsing	- Background download of emails

UMTS QoS classes

Value ranges for UMTS Bearer Service Attributes

Traffic class	Conversational class	Streaming class	Interactive class	Background class
Maximum bitrate (kbps)	< 2 048 (1) (2)	< 2 048 (1) (2)	< 2 048 – overhead (2) (3)	< 2 048 – overhead (2) (3)
Delivery order	Yes/No	Yes/No	Yes/No	Yes/No
Maximum SDU size (octets)	<=1 500 or 1 502 (4)	<=1 500 or 1 502 (4)	<=1 500 or 1 502 (4)	<=1 500 or 1 502 (4)
SDU format information	(5)	(5)		
Delivery of erroneous SDUs	Yes/No/-(6)	Yes/No/-(6)	Yes/No/-(6)	Yes/No/-(6)
Residual BER	$5*10^{-2}$, 10^{-2}, $5*10^{-3}$ 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}	$5*10^{-2}$, 10^{-2}, $5*10^{-3}$ 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}	$4*10^{-3}$, 10^{-5}, $6*10^{-8}$ (7)	$4*10^{-3}$, 10^{-5}, $6*10^{-8}$ (7)
SDU error ratio	10^{-2}, $7*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}	10^{-1}, 10^{-2}, $7*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}	10^{-3}, 10^{-4}, 10^{-6}	10^{-3}, 10^{-4}, 10^{-6}
Transfer delay (ms)	100 – maximum value	280 (8) – maximum value		
Guaranteed bit rate (kbps)	< 2 048 (1) (2)	< 2 048 (1) (2)		
Traffic handling priority			1,2,3	
Allocation/Retention priority	1,2,3	1,2,3	1,2,3	1,2,3
Source statistic descriptor	Speech/unknown	Speech/unknown		

Value ranges for Radio Access Bearer Service Attributes

Traffic class	Conversational class	Streaming class	Interactive class	Background class
Maximum bitrate (kbps)	< 2 048 (1) (2)	< 2 048 (1) (2)	< 2 048 – overhead (2) (3)	< 2 048 – overhead (2) (3)
Delivery order	Yes/No	Yes/No	Yes/No	Yes/No
Maximum SDU size (octets)	<=1 500 or 1 502 (4)	<=1 500 or 1 502 (4)	<=1 500 or 1 502 (4)	<=1 500 or 1 502 (4)
SDU format information	(5)	(5)		
Delivery of erroneous SDUs	Yes/No/-	Yes/No/-	Yes/No/-	Yes/No/-
Residual BER	$5*10^{-2}$, 10^{-2}, $5*10^{-3}$ 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}	$5*10^{-2}$, 10^{-2}, $5*10^{-3}$ 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}	$4*10^{-3}$, 10^{-5}, $6*10^{-8}$ (6)	$4*10^{-3}$, 10^{-5}, $6*10^{-8}$ (6)
SDU error ratio	10^{-2}, $7*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}	10^{-1}, 10^{-2}, $7*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}	10^{-3}, 10^{-4}, 10^{-6}	10^{-3}, 10^{-4}, 10^{-6}
Transfer delay (ms)	80 – maximum value	250 – maximum value		
Guaranteed bit rate (kbps)	< 2 048 (1) (2)	< 2 048 (1) (2)		
Traffic handling priority			1,2,3	
Allocation/Retention priority	1,2,3	1,2,3	1,2,3	1,2,3
Source statistic descriptor	Speech/unknown	Speech/unknown		

Definition

- Per flow- A flow is defined as an individual, uni-directional, data stream between two applications (sender and receiver), uniquely identified by 5 parameters (transport protocol, source address, source port number, destination address, and destination port number).

- per Aggregate- An aggregate is simply two or more flows. Typically the flows will have something in common (e.g., any one or more of the parameters, a label or a priority number, or perhaps some authentication information).

QoS mechanism and protocols for IP networks are generally divided into two types:

- Resource reservation (Integrated Services): The network resources are reserved according to an application QoS request, and subject to resource management policy (per flow).

- prioritization (differentiated Services): The network traffic is classified into a fix set of categories and allocated the network resources to them according to resource management policy criteria (per aggregate).

Integrated Services (Insert) / Reservation Protocol (RSVP)

- The IntServ model specifies the characteristics of a traffic flow (service type), quantifies its resource requirements, reserves the resources within the network (using RSVP signalling) and provides the traffic control needed to classify and process the traffic flow in order to ensure the requested QoS (using a classifier, a packet scheduler and admission control).

- The RSVP is a signalling protocol for setting up and control of resource reservations within a network.

- The IntServ / RSVP offers two types of services:

  - Guaranteed (G) Services; it offers a strict mathematical assurance of both throughput and queuing delay. It closely emulates a dedicated virtual circuit. Hence, it offers the highest level of QoS per flow.

  - Control Load (CL) Services; it is equivalent the "best effort" services under unloaded conditions. Hence, their performance is better than that of the best effort services.

Differentiated Services (DiffServ)

- The DiffServ architecture is based on a simple model where packets entering a DS-capable network are classified into one of a small number of aggregated flows .
- Packets belonging to one aggregated flow are identified by a DiffServ code point (DSCP) inserted in their IP headers. This is known as behaviour aggregate (BA) classification.
- DS behaviour aggregate- A collection of packets based on the content of packet headers according to defined rule.
- At each DiffServ router, packets are subjected to a “per-hop behaviour” (PHB) according to their DSCP.
- Per-Hop-Behavior (PHB)- The externally observed forwarding behaviour applied at a DS-enable router to a DS behaviour aggregate.

- They are currently two standard per hop behaviour defined to represent two service levels.

  - Expedited Forwarding (EF)- It has a single DSCP. EF minimises delay and jitter and offers the highest level of aggregate QoS.

  - Assured Forwarding (AF)- It has four classes and three dropping mechanisms per class (a total of twelve code point). It offers lower levels of aggregate QoS than that of EF.

- A DS network has two DS border routers and many DS interior routers. A DS border router connects one DS network to a router in another network. It acts as ingress and engress router handling traffic as enters and leaves the DS network. A DS interior router is one that is not a DS border router.

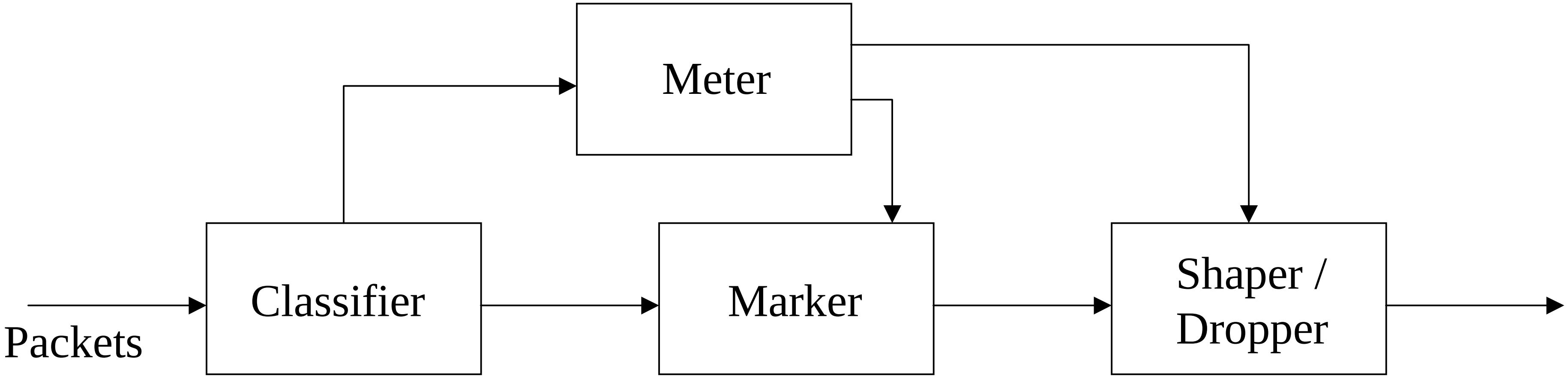
- Packets are classified and marked to receive a particular per-hop forwarding behaviour on nodes along their path.

- Sophisticated classification, marking, policing, and shaping operations need only be implemented at network boundaries or hosts.

- Network resources are allocated to traffic streams by service provisioning policies which govern how traffic is marked and conditioned upon entry to a DS – capable network, and how traffic forwarded within that network.

- PHBs are implemented in routers by means of some kind of buffer management and packet scheduling mechanism.

- A PHB is selected at a router by identifying the DSCP in a received packet and using a code point - PHB mapping look up table.



Entities of DS-compliant border router

Classifier – Selects packets based on content of packet headers according to define Rules.
Meter – Measures the temporal properties (e.g., rate and burst size).
Marker – Sets the DSCP in a packet based on defined rules.
Shaper – Delays packets within a traffic stream to conform to some traffic profile.
Dropper – Discards packets based on specific rules.

IntServ vs. DiffServ

		IntServ	DiffServ
General	Granularity of Service Differentiation	Per-Flow	Per-Class
	State in Routers	Per-Flow	Per-Class
	Scalability	~ number of flows	~ number of Classes
	Coordination	End-to-End	Per-Hop
	Network management	Close to VC Networks	Close to today's IP Networks
Mobility	Path Change	Re-negotiation of QoS (RSVP) – Until the Cross-Over-Router	Re-negotiation with BB for the new path
		Actions Performed Locally (smaller delay)	BB decisions may introduce larger delays

IntServ Versus DiffServ

IntServ

- Advantage
  - offers the best level of QoS possible
- Disadvantage
  - complex – high overhead
  - not scalable

DiffServ

- Advantage
  - simple and low overhead
  - scalable
- Disadvantage
  - does not guarantee QoS per user or per application
  - less efficient use of network resources

Design Approach

For the design of next generation wireless networks, two different approaches are being currently considered:

1. Inter-working with next generation Internet (tight coupling), it is also applied to the existing wireless networks

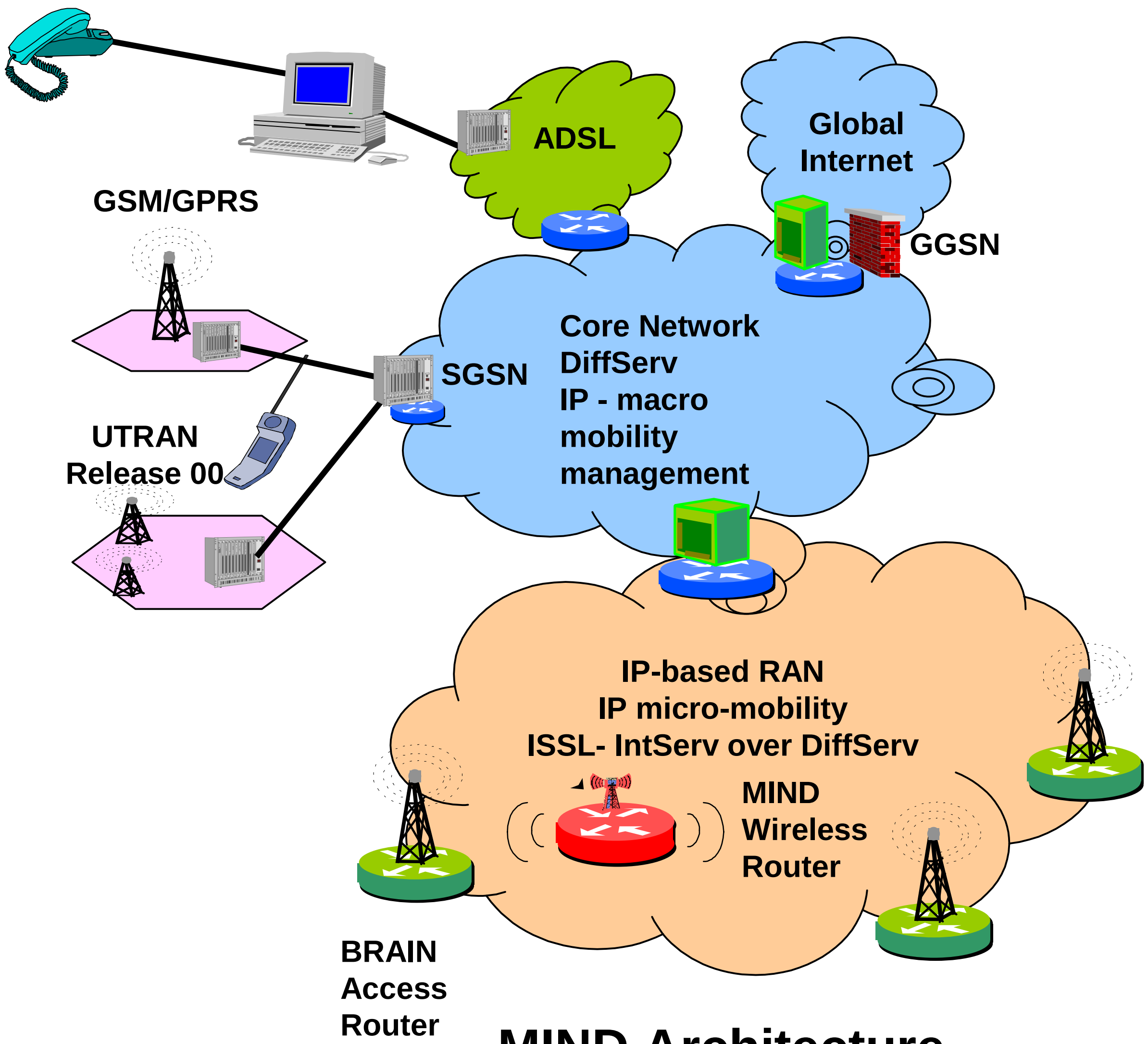
2. Integration within next generation Internet (very tight coupling-plug-in)

First Approach

- To ensure the establishment, maintenance and termination of end-to-end QoS requires a proper mapping among traffic classes and their attributes of two QoS domains (wireless and Internet networks). In addition, inter-operability of the QoS management functions of both domains at the cotrol plane is needed.

Second Approach

- To support end-to-end QoS, new mechanisms and protocols are required to replace or enhance the existing IETF proposed QoS models.



MIND Architecture

Interworking Approach Example

- The example given here is the interoperability issue between UMTS and a specific prototype implementation of the next generation of Internet (The RCL architecture).

- The Resource Control Layer (RCL) architecture was developed within the framework of the IST project, AQUILA (IST-1999-1007).

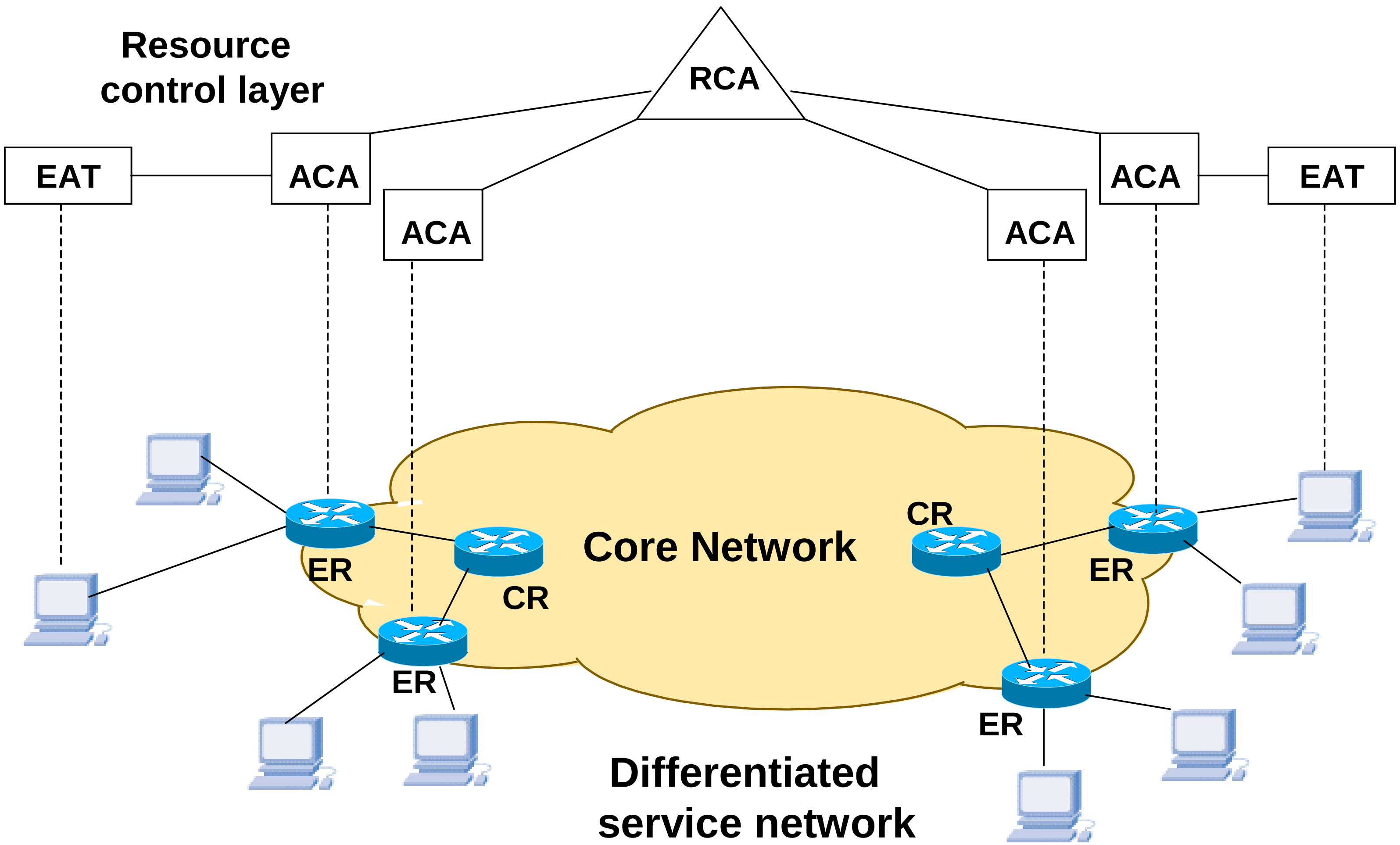
- The RCL architecture is based on the Diffserv model of the IETF.

- It provides a set of five traffic classes, which extend the EF and AF per-hop behaviours (PHBs) to further incorporate admission control and traffic engineering mechanisms.

ENTER for
ELECOMMUNICATIONS
ESEARCH

The RCL architecture and main components

King's College London
University of London



RCA: Resource Control Agent
ACA: Admission Control Agent
EAT: End-user Application Toolkit

ER : Edge Router
CR : Core Router

<u>RCL traffic classes</u> (IST AQUILA project)

	Mapping onto
Premium constant bit rate (PCBR)	Conversational
Premium variable bit rate (PVBR)	Streaming
Premium multimedia (PMN) OR Premium mission critical (PMC)	Interactive
Premium mission critical (PMC) OR Best effort (BE)	Background

CENTER for TELECOMMUNICATIONS RESEARCH

KING'S College LONDON
University of London

Mapping of UMTS QoS classes to RCL network services

Characteristics	UMTS classes	RCL architecture NS	
	Conversational	**PCBR**	
Maximum bit rate (kbps)	< 2048	Maximum per flow 200kb/s	
Maximum packet size (bytes)	≤ 1500 or 1502	< 256	
Packet error ratio	10^{-2}, $7*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}	$< 10^{-8}$	
Transfer delay (ms)	100 Maximum	150 Maximum	
	Streaming	**PVBR**	
Maximum bit rate (kbps)	< 2048	Maximum per flow 1000kb/s	
Maximum packet size (bytes)	≤ 1500 or 1502	< 1000	
Packet error ratio	10^{-1}, 10^{-2}, $7*10^{-3}$, 10^{-3}, 10^{-4}, 10^{-5}	$< 10^{-6}$	
Transfer delay (ms)	250 Maximum	250 Maximum	
	Interactive	**PMM**	**PMC**
Maximum bit rate (kbps)	< 2048 - overhead	Maximum per flow 250kb/s	Maximum per flow 50kb/s
Maximum packet size (bytes)	≤ 1500 or 1502	< 1500	< 1500
Packet error ratio	10^{-3}, 10^{-4}, 10^{-6}	$< 10^{-3}$	$< 10^{-4}$
Traffic handling priority	1, 2, 3	1	2, 3
	Background	**PMC**	**BE**
Maximum bit rate (kbps)	< 2048 - overhead	Max per flow 50kb/s	
Maximum packet size (bytes)	≤ 1500 or 1502	< 1500	
Packet error ratio	10^{-3}, 10^{-4}, 10^{-6}	$< 10^{-4}$	

Problems

- Mutual transformation of some service attributes is not possible.

- Interoperability of the QoS management functions of both domains at the control plane is difficult.

Core-Centric vs. Radio Access Centric Approach

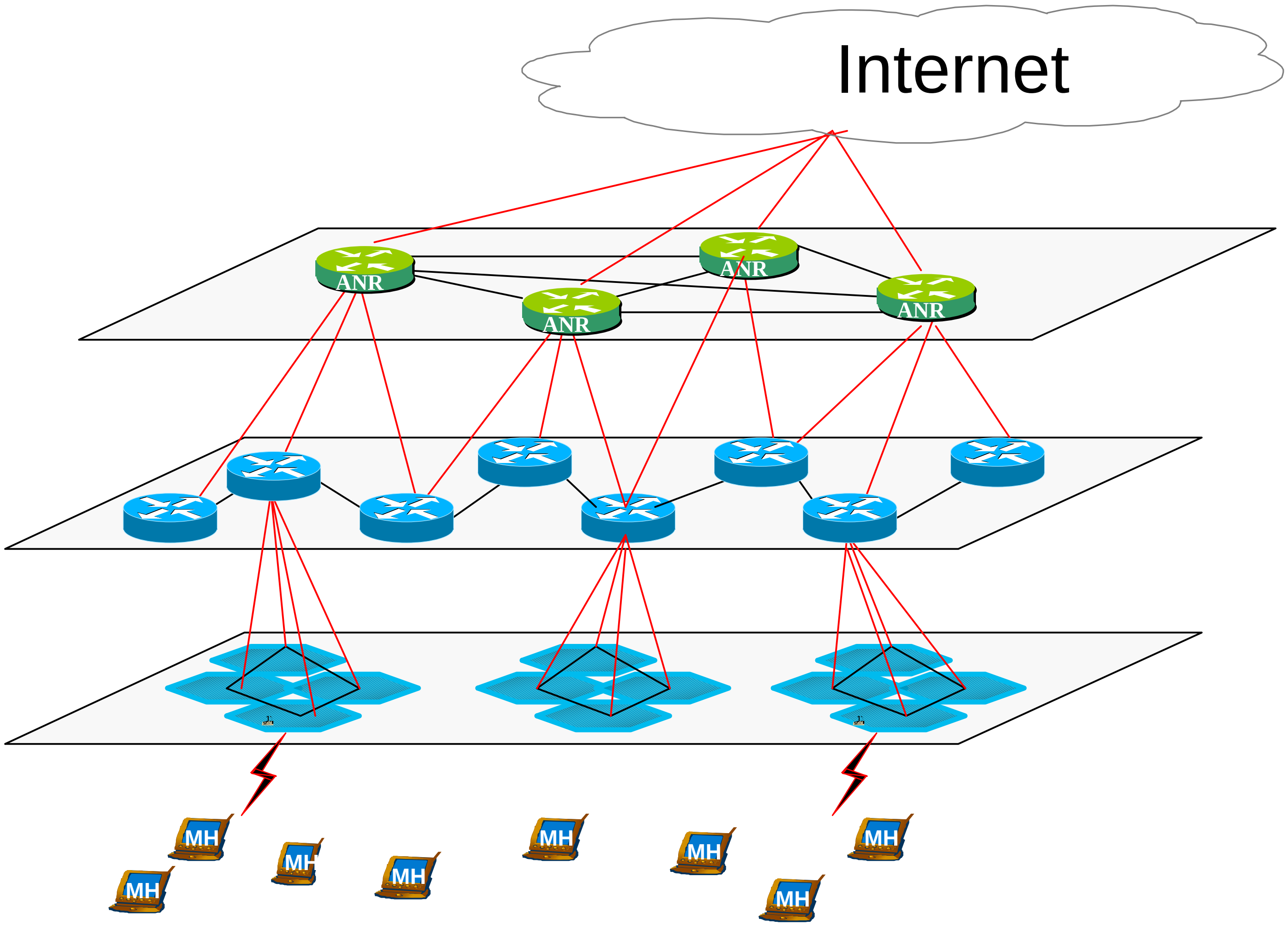
- Separation of the traffic classes in Core-IP from the traffic classes of Radio Access

  - To support heterogeneous Mobile/Wireless environment

  - To use a minimum number of traffic classes possible in order to simplify the Control Plane of the Core-IP

- The aim is to study the feasibility of deploying a small number of traffic classes in the Core-IP

Hierarchical Partially Mesh Topology



Simulations based on Strict Tree Topology

Simulation Scenarios

- DiffServ scenarios
  - DS [EF + 3AF] WRR
  - DS [EF + 2AF] WRR
  - DS [EF + AF] PQ
  - DS [EF + AF] WRR
- Best-effort scenarios
  - BE (Drop-Tail)
  - BE (RED)

QoS Requirements

App.	*QoS Requirement*
VoIP	Packet avg. delay < 80 msec
	Packet drops < 1 %
Video	Packet avg. delay < 150 msec
	Packet drops < 1 %
HTTP	Avg. download time < 10 sec
FTP	No QoS requirements (Background Traffic)

*Other important parameters measured :*VoIP jitter (mean and standard deviation), VoIP outages length, FTP download time

Over-provisioning:

TM A (VoIP 15 %) ⇒ **11.8%-39.6%** over-provisioning
TM B (VoIP 30 %) ⇒ **16.3%-35.4%** over-provisioning
TM C (VoIP 45 %) ⇒ **24.9%-36.2%** over-provisioning

Traffic Matrix:

As the percentage of VoIP traffic increases from 15% to 45% the bandwidth required to meet the QoS constraints **decreases**

AF Classes:

Excess bw needed in order to deploy DiffServ with EF and only 1 AF class (instead of EF and 3 AF classes) :

- Traffic Matrix A ⇒ **24.9%** over-provisioning
- Traffic Matrix B ⇒ **16.4%** over-provisioning
- Traffic Matrix C ⇒ **5%** over-provisioning

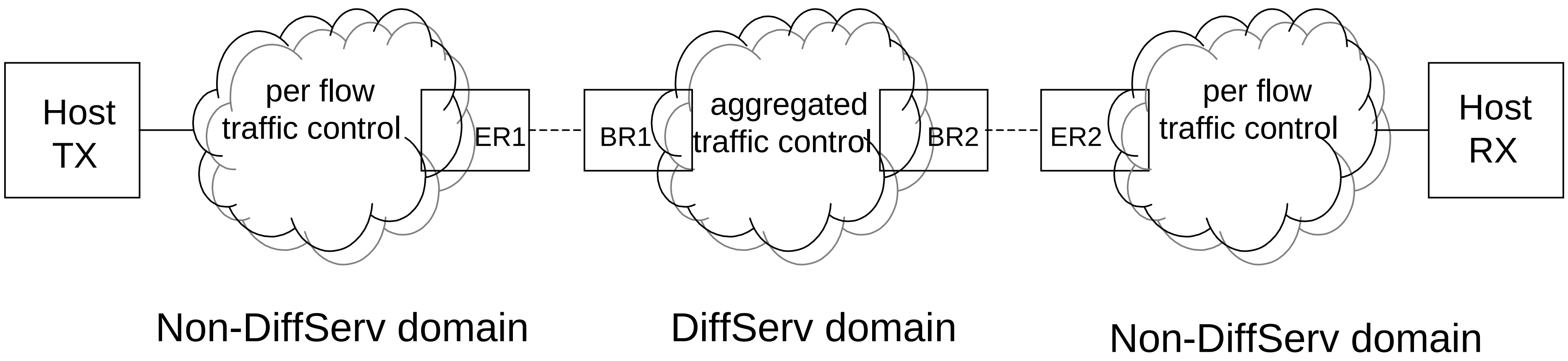
IntServ over DiffServ

- IntServ/RSVP and DiffServ can be viewed as complementary models and together can support end-to-end QoS of large variety of applications such as IP-telephony, video on demand and various non-real time applications.

- IntServ enables hosts to request per flow, quantifiable resources, along end-to-end data path, and to obtain feedback from the network elements regarding these requests.

- DiffServ enables scalability across large networks.

- There is a variety of architectures in which these models work together to provide end-to-end QoS across multiple service providers.



Network Configuration Example-IETF ISSLL Architecture

Main benefits of using IntServ over DiffServ:

- The use of the DiffServ model provides scalability.

- The use of explicit and dynamic admission control of the IntServ model ensures that network resources are optimally used.

QoS Routing

- QoS routing aims to extend the current Internet routing models of intra-and inter-domain routing to support QoS.

- Current Internet routing protocols are based on single parameter (metric) optimisation e.g., hop count (shortest path routing). Alternate paths with acceptable but non-optimal cost can not be used to route traffic.

- Therefore, if the best existing path cannot admit a new flow, the associated traffic cannot be forwarded even if an adequate alternate path exist.

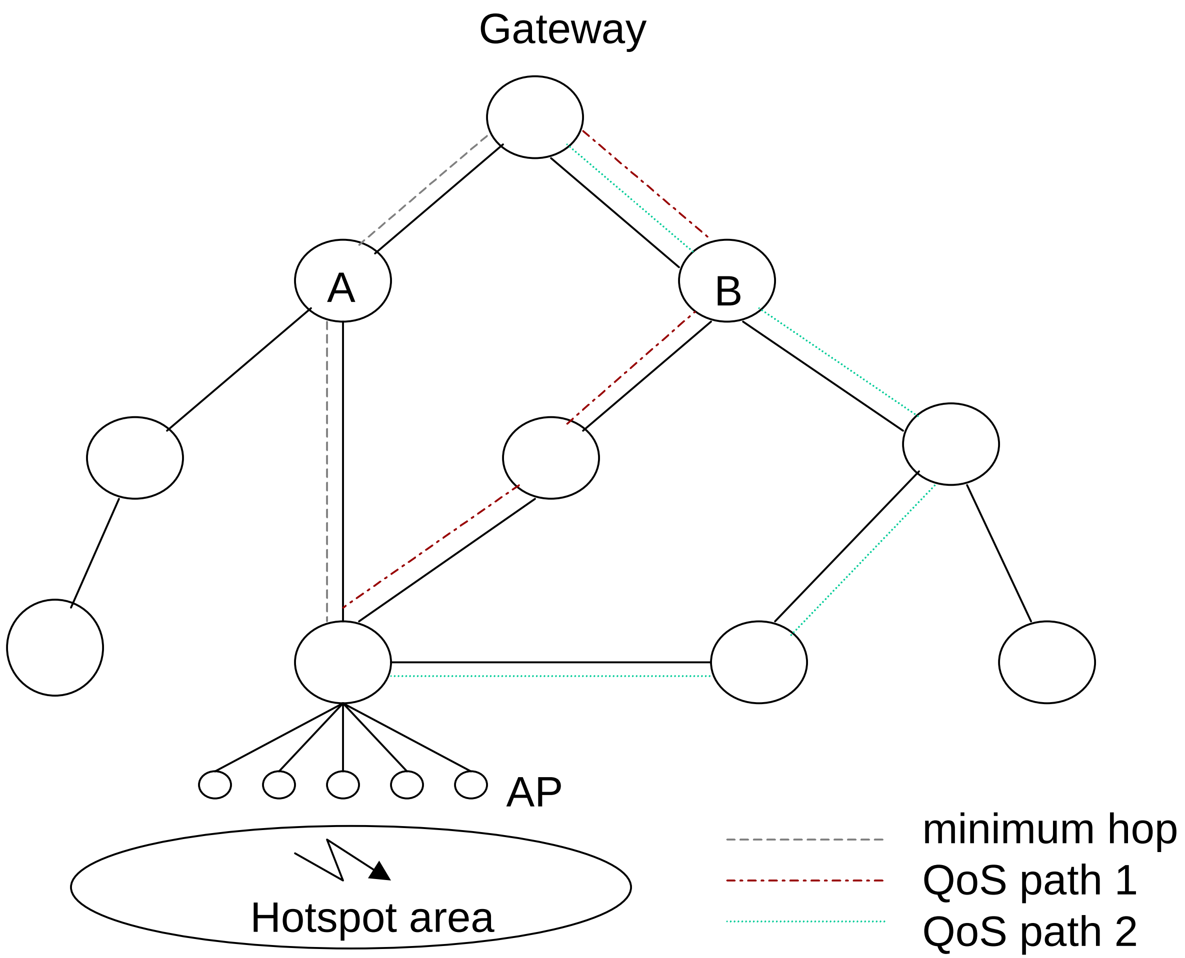
• QoS routing considers more than one metric e.g., hop count, available path bandwidth and/or path delay.

• QoS routing allows the routing of traffic using alternate paths, rather than just the shortest path.

• The objectives of alternate path routing are:

- To select paths that have a good chance of accommodating the request QoS.
- To distribute load among paths in the network.

• QoS routing is usually used in conjunction with some form of resource reservation mechanism such as RSVP.



QoS routing scenario under a multi-technology pure IP mobile/wireless network with hot-spot areas

Challenges in deployment of QoS routing:

- QoS routing capability deployment should have minimum possible impact on the existing routing infrastructure.

- The computational cost of QoS routing should be kept at a level comparable to those of existing routing algorithms.

- To minimize the amount of additional update traffic (overhead) due to more frequent changes to link metrics without adversely affecting the performance of path selection.

- Interaction between QoS routing and mobility.

Conclusions

For the Interworking Approach

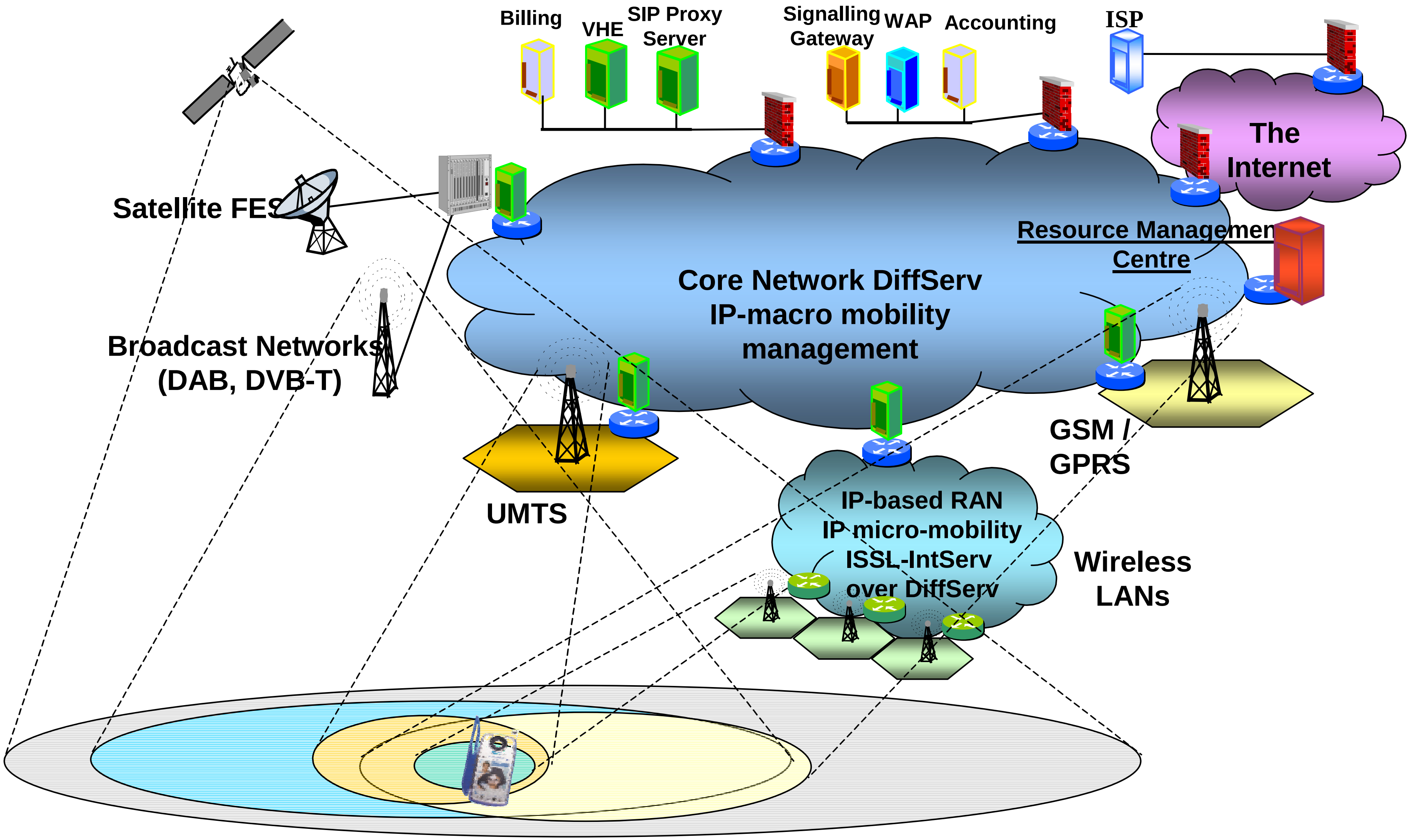
- A simple Diffserv with only two service classes combined with some reasonable over provisioning seems a good candidate.

For the Integration Approach

- QoS routing combined with IntServ over DiffServ with mobility support and its associated intra-and inter-domain signalling could be a candidate.

- Alternatively, one can ignore IETF and go for a radical revolutionary solution.

BRAIN QoS Base Line Architecture


Billing
VHE
SIP Proxy Server
Signalling Gateway
WAP
Accounting
ISP
The Internet
Satellite FES
Resource Management Centre
Core Network DiffServ
IP-macro mobility
management
Broadcast Networks
(DAB, DVB-T)
UMTS
GSM / GPRS
IP-based RAN
IP micro-mobility
ISSL-IntServ
over DiffServ
Wireless LANs

BRAIN Network key components

- Mobile Host (MH): IP hosts with one or more IP addresses and a single interface and possibly more than one simultaneous radio link with different ARs.

- Access Router (AR): IP routers which are located at the edge of the access network and offer IP connectivity to MHs. They have multiple wireless and wired interfaces.

- Anchor Point (ANP): They are located "inside" the access network at different selected positions. ANPs own and allocate IP addresses, authenticate users, maintain user records, and tunnel packets towards MHs. The AR terminates tunnels from ANPs and forward packets to/from MHs.

- Mobility Gateway (MG): They are standard border routers, distributing traffic within the access network to the correct ANPs. They hide any access network-specific routing functionality.

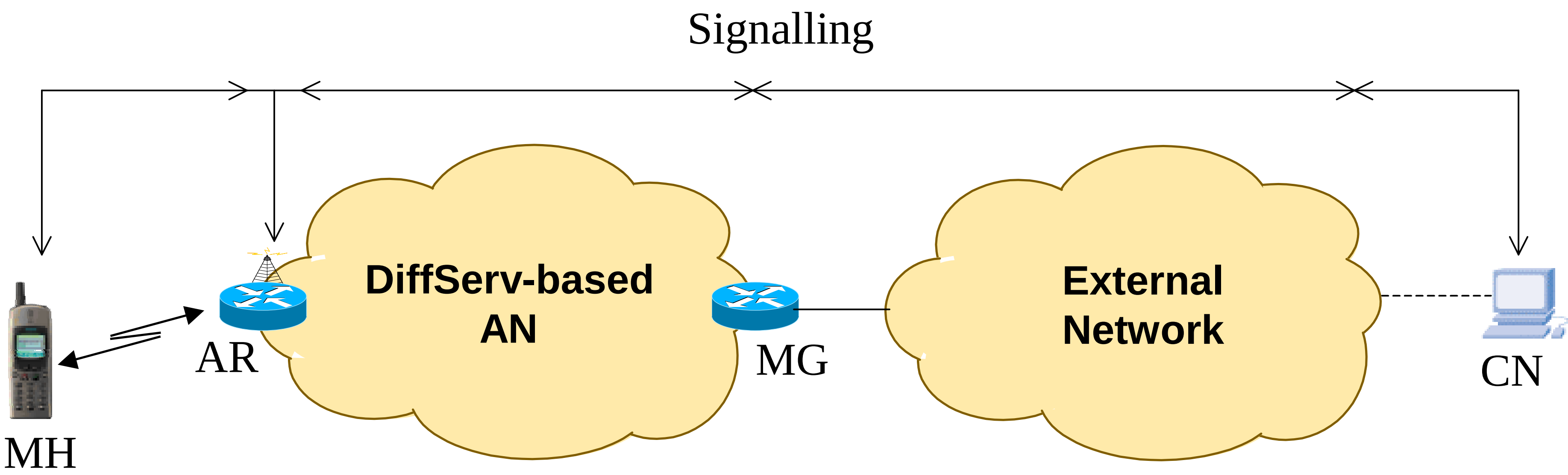
- Internal Node (IN): IP routers responsible of maintaining routing information both for fixed network and MH reachability.

BRAIN QoS Base Line Architecture

- The architecture is based on the work done in the IETF ISSLL working group using the IntServ over DiffServ model.

- The proposed architecture was chosen to address the need for an IP-based scalable QoS architecture that allows flexible support for micro-mobility in a wireless network.

- The fundemental design criteria were to use the existing IETF protocols and architectures where possible and to add new extensions if needed.



QoS BRAIN Architecture

MH	Mobile Host
CN	Corresponding Node
AR	Access Router
MG	Mobility Gate
AN	Access Network

- The architecture uses the IntServ/RSVP as the signalling method for explicit resource reservations.

- IntServ/RSVP messages are mapped to DiffServ forwarding classes at the edge of the access network and forwarded according to standard DiffServ operation within the access network.

- The access network between the AR and MG is viewed as a single transparent link to the RSVP messages.

Packet transfers

- MH-originated packet transfers

1. The MH sends a RSVP PATH message.
2. This message arrives at the AR which stores information about the reservation request and forwards the message to the MG.
3. The MG stores a similar state and forwards the message to the external network.
4. When the corresponding RESV message arrives from the external network to the MG, it checks for resource availability. If resources are available, the MG will forward the message to the AR, otherwise the MG will cancel the reservation according to standard RSVP processing.
5. When the AR receives the RESV message, it will perform similar operations and if resources are available, it can record the RSVP parameters-to-DSCP mapping to be used if the MH is not marking the upstream packets. The mapping information can be available at the AR for direct use, or the AR can request the SLA for the MH from an external entity like a Bandwidth Broker.

- CN-originated packet transfers

1. The CN sends a RSVP PATH message that will go through the MG and AR and stores a reservation state in these nodes.
2. The MH processes the arrived reservation request, calculate the required resources and responds with a RESV message.
3. If resources are available at the AR and MG, the RESV message will eventually reach the CN and the resources within the access network have been reserved.
4. When the data packets are arrived at the MG, they are marked with appropriate DSCP.

Some BRAIN Specific Enhancements

- QoS context transfer

- When MH changes its AR, the state information about the MH's QoS requirements needs to be transferred to the new AR. The exchange is triggered by handover indications received from the link layer.

- QoS Context Transfer enables the exchange of network layer parameters between network nodes involved in a handover without requiring any signalling from the MH itself.

• Bandwidth Broker (BB)

- Bandwidth brokers make the admission control decisions, on behalf of AR/MG nodes.

- A BB has global state knowledge, because it receives each admission control request, and so can make a correct admission control decision.

- Bandwidth Brokers may be a centralised unit or distributed between adjacent edge routers.

- Distributed BB approach enables possible mobility patterns to be taken into account, while removing the "single point of failure" that a centralised BB represents.

- Hop-to-hop call admission coupled with current micro-mobility mechanisms

- Reservation-based QoS implicitly assumes that the route taken by a traffic stream across a network is reasonably stable for the duration of a reservation.

- However, the routes in the mobile environment can be dynamic, changing every time the MH changes its AR.

- In order to improve the behaviour of reservation-based QoS in the micro-mobile environment, the QoS and micro-mobility mechanisms should be coupled to ensure the reservations are installed as soon as possible, after a mobility event such as handover.

- The loosely coupled approach uses mobility events to trigger the generation of RSVP messages, which distribute the QoS information along new paths across the network. The RSVP message can be triggered as soon as the new routing information has been installed in the network.