QoS Provision for Mobile Access via GPRS

Presented by:
Shaily Verma
Overview

- Introduction to GPRS
- QoS provision
  - Purpose
  - Approach
- SSF simulation platform
- Implementation
- Simulation results
- Next Steps
Introduction to GPRS:

GPRS is GSM's extension for packet oriented data transmission.

Information Flow
**Purpose**

- Multimedia applications need a wide range of QoS to be supported by the GPRS network.
- Adaptation and control strategies to ensure QoS have not been defined in GPRS.
- Aims to look at call admission and link adaptation schemes for both real time/non real time traffic in GPRS.
Where does QoS come into the picture?

MS

BS

SGSN

GGSN

PDN

MS
QoS negotiations occur at the time of PDP (Packet Data Protocol) context activation/modification.
As per GSM 3.60, the QoS profile consists of the following parameters:

1. service precedence (priority)
2. reliability
3. delay
4. user data throughput
   - peak
   - mean
QoS profile

- **Service precedence (priority)**
  - Indicates the relative priority of maintaining the service.
    - **High priority** (class 1): Service commitments will be maintained ahead of all other precedence levels.
    - **Normal priority** (class 2)
    - **Low priority** (class 3)
Reliability

The reliability class defines probability of data loss, data delivered out of sequence, duplicate data delivery and corrupted data.

<table>
<thead>
<tr>
<th>Reliability Class</th>
<th>GTP Mode</th>
<th>LLC Frame Mode</th>
<th>LLC Data Protection</th>
<th>RLC Block Mode</th>
<th>Traffic Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Acknowledged</td>
<td>Acknowledged</td>
<td>Protected</td>
<td>Acknowledged</td>
<td>Non real-time traffic, error-sensitive application that cannot cope with data loss.</td>
</tr>
<tr>
<td>2</td>
<td>Unacknowledged</td>
<td>Acknowledged</td>
<td>Protected</td>
<td>Acknowledged</td>
<td>Non real-time traffic, error-sensitive application that can cope with infrequent data loss.</td>
</tr>
<tr>
<td>3</td>
<td>Unacknowledged</td>
<td>Unacknowledged</td>
<td>Protected</td>
<td>Acknowledged</td>
<td>Non real-time traffic, error-sensitive application that can cope with data loss, GMM/SM, and SMS.</td>
</tr>
<tr>
<td>4</td>
<td>Unacknowledged</td>
<td>Unacknowledged</td>
<td>Protected</td>
<td>Unacknowledged</td>
<td>Real-time traffic, error-sensitive application that can cope with data loss.</td>
</tr>
<tr>
<td>5</td>
<td>Unacknowledged</td>
<td>Unacknowledged</td>
<td>Unprotected</td>
<td>Unacknowledged</td>
<td>Real-time traffic, error non-sensitive application that can cope with data loss.</td>
</tr>
</tbody>
</table>

NOTE: For real-time traffic, the QoS profile also requires appropriate settings for delay and throughput.
Delay

- Defines the maximum values for the mean delay and 95- percentile delay.
- This includes the radio channel access delay, radio channel scheduling delay, the radio channel and GPRS network transit delay.

<table>
<thead>
<tr>
<th>Delay Class</th>
<th>SDU size: 128 octets</th>
<th>SDU size: 1024 octets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Transfer Delay (sec)</td>
<td>95 percentile Delay (sec)</td>
</tr>
<tr>
<td>1. (Predictive)</td>
<td>&lt; 0.5</td>
<td>&lt; 1.5</td>
</tr>
<tr>
<td>2. (Predictive)</td>
<td>&lt; 5</td>
<td>&lt; 25</td>
</tr>
<tr>
<td>3. (Predictive)</td>
<td>&lt; 50</td>
<td>&lt; 250</td>
</tr>
<tr>
<td>4. (Best Effort)</td>
<td>&lt; Unspecified</td>
<td></td>
</tr>
</tbody>
</table>
Throughput

- Indicates the user data throughput requested by the user.
- Defined by two negotiable parameters:
  - Peak
  - Mean
- Various classes defined according to peak and mean throughput requirements.
Approach

Intends to see how QoS can be delivered for different traffic types through:

- call admission and control.
- radio resource management.
- link adaptation (changing code rate according to the current C/I).

Provide QoS by abstracting the functionalities from LLC, RLC, MAC and BSSGP protocols.
### UMTS traffic classes

<table>
<thead>
<tr>
<th>Error tolerant</th>
<th>Conversational voice and video</th>
<th>Voice messaging</th>
<th>Streaming audio and video</th>
<th>Fax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error intolerant</td>
<td>Telnet, interactive games</td>
<td>E-commerce, WWW browsing,</td>
<td>FTP, still image, paging</td>
<td>E-mail arrival notification</td>
</tr>
</tbody>
</table>

- **Conversational** (delay $\ll 1$ sec)
- **Interactive** (delay approx 1 sec)
- **Streaming** (delay $<10$ sec)
- **Background** (delay $>10$ sec)
“All bits are not created equal!”

Have two main traffic classes in GPRS:

- GPRS Conversational class/real time (rt)
  - have absolute QoS requirements.
  - if the negotiated QoS cannot be met, the MS is rejected

- GPRS non-conversational class/non real time(nrt)
  - PDF (Packet data flows) adjust to the available bandwidth.
  - the relative QoS between different PDFs is kept.
Approach

- **Call Admission and Control (CAC)**
  - CAC criteria
    - RT: check bandwidth requirement and also available bandwidth
    - NRT: check nrt queue and buffers available.
  - Maintain separate queues for rt and nrt packets.
  - Schedule packets on the basis of priority and delay class.

- **QoS profile in the BSS**
Simulation: Top Level

Transmission Plane Implementation
Simulation Platform: JSSF
(JAVA SCALABLE SIMULATION FRAMEWORK)

- SSF provides a single, unified interface for discrete-event simulation.
- Makes it possible to build models that are
  - Efficient
  - Scalable
  - able to utilize parallel processor resources.
  - Object-oriented
    - utilize and extend the framework
    - maximize the potential for direct reuse of code
The SSF syntax comprises five base class interfaces:

- **Entity**: Object that can own processes and channel endpoints, and can be aligned with other Entities
- **Event**: sent on channels to processes in Entities
- **Process**: sends events on outChannels.
- **inChannel**: pipe leading to a process carrying events
- **outChannel**: pipe going out from a process carrying events
Scope of the simulator

Simulation scenario currently restricted to:
- one GSM carrier
- one cell
- One BSS incorporating functionalities of BS and SGSN
- use of all coding schemes
- mean C/I can be varied as also variance
- concentrates on uplink performance
- User can specify the number of MS
Scope of the simulator

- Traffic Models:
  - Mobitex (mean packet length 30 bytes)
  - Exponentially distributed packets (mean 200 bytes)

- Channel Model
  - TU3
  - RT and NRT packet separation
  - Delay class and priority based selection at the BS
  - RT packets are unacknowledged (no retransmissions)
  - NRT packets are acknowledged (selective ARQ)
Implementation:

- **Physical Layer:**
  - C/I value is generated from a lognormal distribution with user specified parameters.
  - Emulated with a set of graphs of the BLER vs. C/I for the 4 channel coding schemes
  - As per BLER, RLC block is dropped/transmitted.

- **LLC**
  - sole functionality is the segmentation and reassembly of the LLC frames.
Implementation:

- **RLC:**
  - Acknowledged/Unacknowledged data transfer mode
  - BEC with sliding window ARQ mechanism
  - All four coding schemes are supported

- **MAC:**
  - Uses slotted ALOHA model
  - Includes capture model
  - Channel assignment by the BSS
  - Logical channels are implemented
Uplink Model Description

1. PDP context negotiation done
2. Contend for channel using slotted ALOHA
3. Contention winner sent PUA (Packet Uplink Assignment)
4. USF (uplink state flag) /channel allocation sent
5. Data blocks sent to BS:
   - RT: if delay class 1
   - NRT: if delay class 2, 3, 4
6. ACK sent to NRT with retransmission bitmap and channels
Initially say out of the 7TS, one is allocated for nrt (on which 7 users can be mux.) and 6 TS are allocated for rt users.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>av_sl &lt; P/S</td>
<td>Allocate</td>
</tr>
<tr>
<td>Lq &gt; 7</td>
<td>Reject request</td>
</tr>
<tr>
<td>Snrt &lt; Snrt_max</td>
<td>Add to queue</td>
</tr>
</tbody>
</table>

Peak throughput = P
Throughput realizable per time slot = S = 5Kbps
Lq = Length of nrt queue
Available slots = av_sl
Slots for nrt = Snrt
Max. # of slots for nrt = Snrt_max
Algorithm for Resource allocation
(Channel request)

MS \( \xrightarrow{\text{Channel \_request[MS \_id]}} \) BSS + SGSN

Add to rt Q \( \xrightarrow{\text{Rt/nrt?}} \) Add to nrt Q

Real time(rt) Non real time(nrt)

Sort rt Q as per priority and delay class. Allocate USFs

Sort nrt Q as per priority and delay class. Allocate USFs
# GPRS Coding Schemes

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Code Rate</th>
<th>Payload (bits/block)</th>
<th>Data rate (Kb/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS 1</td>
<td>1/2</td>
<td>181</td>
<td>9.05</td>
</tr>
<tr>
<td>CS 2</td>
<td>2/3</td>
<td>268</td>
<td>13.4</td>
</tr>
<tr>
<td>CS 3</td>
<td>3/4</td>
<td>312</td>
<td>15.6</td>
</tr>
<tr>
<td>CS 4</td>
<td>1</td>
<td>428</td>
<td>21.4</td>
</tr>
</tbody>
</table>
**Optimizing Channel Capacity**

**Link Adaptation:**
Use higher coding schemes (less coding, more payload) when the radio conditions are good:

- **RT packets:** use CS3
- **NRT packets:**
  - use CS1 at $C/I < 10\text{dB}$
  - use CS3 at $10<C/I < 18\text{dB}$
  - use CS4 at $C/I \geq 18\text{dB}$
Sequence of messages on the MS side

(1) Packet from higher layer
(2) PacketEvent
(3) If ACK, remove head
(4) Sends ACTIVATE_PDP_REQUEST(QoS)
(5) Sends ACTIVATE_PDP_ACCEPT/REJECT
(6) Sends ACTIVATE_PDP_ACCEPT/REJECT
(7) ACTIVATE_PDP_ACCEPT/REJECT from BS with negotiated QoS
(8) Change PDP state from inactive to active
(9) RLCBlock from BS containing USF free
(10) USF free from BS
(11) Sends PACKET_CHANNEL_REQUEST
(12) Send RLCBlockEvent(Data_Block) to BS
(13) Receive ACK/NACK for data block
(14) Send USF=MS_ID
(15) Status 1->2
(16) Start T3164
(17) Stop T3164. Start T3166
(18) Stop T3166.
Sequence of messages on the BS side

1. Listen_to_PDP_Uplink
2. Listen_to_Uplink
3. Contend and Capture
4. ConnectionRtSet
5. ConnectionNrtSet
6. ControlMessagesACK_NACKset
7. Sink
8. to_BEM
9. in_BEM
10. Block_Error_Mgr
11. Downlink_Sender
12. RLCBlock containing channel request
13. Receive Data_Block from MS
14. ConnectionEvent
15. Sort the requests based on priority.
16. Send RLCBLOCK with USF=USF_FREE in TS0
17. Send the PUA in TS0. Set uplink_sent to 1.
18. In other TS check if uplink_sent =1 in ConnectionRtSet/ ConnectionNrtSet
19. Send USF=MS_ID
20. Check for error
21. ControlMessagesACK_NACKset
22. Send ACK/NACK for data block from ControlMessagesACK_NACKset
23. Send ACTIVATE_PDP_REQUEST(QoS requested)
24. Sends ACTIVATE_PDP_ACCEPT/REJECT (QoS negotiated)
25. (14) ConnectionEvent
26. (13) ConnectionRtSet
27. (12) RLCBlock containing channel request
28. (11) Send ACK/NACK for data block from ControlMessagesACK_NACKset
29. (4) Sends ACTIVATE_PDP_REQUEST
30. (5) Sends ACTIVATE_PDP_REQUEST(QoS requested)
31. (6) Sends ACTIVATE_PDP_ACCEPT/REJECT (QoS negotiated)
Total load per TS vs. total throughput per TS
10 MS (3 RT, 7 NRT), exp. distributed packet (mean 200 bytes)
Throughput vs. load
10 MS (3 RT, 7 NRT), exp. distributed packet (mean 200 bytes)
Blocking and dropping

- Packets can be lost in two ways:
  - Blocking
  - Dropping

<table>
<thead>
<tr>
<th>Traffic</th>
<th>Blocking</th>
<th>Dropping</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>&lt;1%</td>
<td>0%</td>
</tr>
<tr>
<td>NRT</td>
<td>&lt;1%</td>
<td>&lt; 1%</td>
</tr>
</tbody>
</table>
Next Steps

- Get BLER vs. C/I curves for the 8 coding schemes of EDGE and study the link adaptation gains of EDGE vs. GPRS.
- Investigate throughput and delay gains for other traffic models like railway and web through link adaptation.