Intersystem End-to-End QoS Provision in 4G Heterogeneous Networks

NINO KUBINIDZE, MAIRTIN O'DROMA, IVAN GANCHEV
Department of Electronic and Computer Engineering
University of Limerick
IRELAND

Abstract: -The paper presents an IP-based multi-service approach for end-to-end (E2E) Quality of Service (QoS) provision and discusses it in the context and perspective of the fourth generation (4G) networks. Particular emphasis is placed on distribution of functionalities between core, edge, multi-access networks and mobile host. Existing QoS mechanisms are briefly discussed and possible QoS mapping techniques between various wireless and fixed protocols, namely GPRS/UMTS and MPLS/DiffServ are proposed and evaluated. Intersystem E2E QoS support model provided across different network parts and suitable for deployment in 4G heterogeneous network environment is presented. Two options for E2E QoS provision are specified and consequent shifting of functionalities of the core, edge and multi-access networks are demonstrated and justified.

Key- Words: - 4G, QoS, MPLS/DiffServ, GPRS/UMTS, MuMAcWiN, all-IP, ABC, ABS

1. Introduction

Current networks evolve towards all-IP infrastructure, where networks and services converge into one common, cost-effective and scalable architecture. The essential and sufficient requirement for flexible E2E QoS provision in the future 4G networks is a seamless interoperability between fixed and mobile wireless networks, where a system is capable of providing reliable always best connection (ABC) and always best served (ABS) adaptable to the dynamic changes of the environment. Due to the importance of voice and multimedia applications considerable attention has to be paid to the QoS guarantees. Optimised schemes for E2E QoS provision should be proposed in order to reflect the dynamic behaviour of terminals moving across heterogeneous access networks and address variety of QoS features, such as bandwidth reservation, admission control, scalability, per-class service support etc.

2. 4G Integrated Network Architecture

4G networks paradigm is going to be service-oriented, therefore service creation platforms should be carefully designed in order to support scalable network structure with guaranteed levels of reliability and broadband multiple service availability [1]. 4G networks will be based on the integration/coupling of all existing access network technologies in order to provide ABS service to an ABC user on any mobile terminal, through any access network anytime, anywhere, anyhow. Currently there are two main approaches for coupling WLAN/WPAN with GPRS/UMTS access networks [2]: (i) *tight coupling*, where the WLAN/WPAN is connected to the GPRS/UMTS network as an alternative radio access network, i.e. the WLAN/WPAN router is connected directly to the serving GPRS support node (SGSN) and is treated by it as a radio network controller (RNC); (ii) *loose coupling*, where the WLAN/WPAN is connected to a gateway GPRS support node (GGSN) of the GPRS/UMTS network as a separate network, i.e. the WLAN/WPAN router is treated as a GGSN. Figure 1 shows a general architecture for tightly coupled multiple multi-access wireless networks (MuMAcWiNs[1]) interconnected via IP backbone (core network). Each multi-access network domain (A, B, C) consists of network sub-domains / cells, each providing different type of access, e.g. (E)GPRS, EDGE, UMTS, WPAN, WLAN etc. Mobile users may have an access to at least two different sub-domains and may move between them while using a service. The core network is QoS enabled IP-based backbone. The "edge network" includes the ingress and/or egress nodes of the core network and the GGSN/SGSN infrastructure. This type of network architecture is fully utilized, based on QoS and TE features; and applications sharing the same connection are provided with guaranteed

[1] *MuMAcWiN, Multiple Multi-Access Wireless Networks, means more than one MAcWiN, providing different (more than one) access technologies to a core network. Such access technologies are sufficiently different to merit being so distinguished (e.g. having different values for defining physical layer attributes, such as the modulation scheme). Which access technology is invoked by the user or the network will be determined dynamically on the basis of certain criteria etc., or by prior agreement (e.g. between the user, and network operator, and maybe also by service provider) seeking to meet certain criteria for the communications access, connection and information transfer.*

QoS levels. Vertical handover from one access network to another is to be provided in a seamless manner, without service / communication interruption and with the same level of security as GPRS/UMTS.

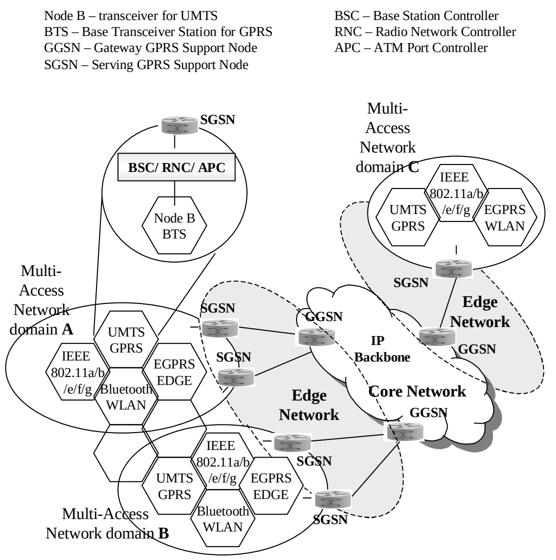

Fig.1. MuMAcWiN architecture interconnected via IP Backbone.

Functional requirements of 4G mobile terminals (MTs) include ability to support intelligent interface(s), so that users have a choice to initially pick (or further on switch to) the best service depending on the capabilities of the access networks available in the area, constraints of the terminal currently used, cost/performance preferences specified in users profiles, etc. Multi-access domains should have a set of IP functionality as in the core as well as at the edge of its integrated network architecture. The edge network provides the functionalities of Authentication, Authorization, and Accounting (AAA), policy management, mobility, QoS, billing, e.g. for dynamic control and provision of network resources and services.

QoS management as a part of intelligent control of the network (control plane, figure 2), together with functions like mobility, monitoring of the status of access resources and its information management, is accomplished independently of the IP-based transport network. This approach leaves space for further development of control functions without interfering with transport network. In order to provide communications services (charging, billing, authentication and security) in a seamless manner during the handover process from one access network technology to another, service support functions should be independent of transport network and control layers (Figure 2).

3. Intersystem QoS support across integrated MuMAcWiNs

With the evolutionary trend being towards all-IP-based 4G environment, E2E QoS support is becoming an important focus for research and development. What this E2E QoS means in effect is guaranteed bandwidth provision and per-class treatment of traffic flows. QoS schemes have already been adopted for IP networks (IntServ [3], DiffServ [4]) and proposed by IETF.

- *IP QoS for GPRS and UMTS*

IntServ approach introduces Guaranteed Load and Controlled Load Services to the best-effort-based IP networks and proves through it, it is possible to support real-time applications and per-class treatment and control [3]. Together with RSVP signaling it provides networks with per-flow packet classification, bandwidth reservations and explicit admission control, to ensure that each flow receives the service requested. Thus IntServ satisfied two main requirements of QoS, namely bandwidth guarantees and per-class treatment. IntServ has been proposed for deployment in GPRS/UMTS core networks, that is that RSVP signaling to be used between SGSN and GGSN for QoS enabled GPRS Tunneling Protocol (GTP) tunnel establishment [5, 6]. However complex signaling and per-flow state of IntServ create scalability problems and limit deployment of IntServ in the networks with larger number of nodes.

As a solution to scalability issues the DiffServ approach has been developed by IETF [4]. DiffServ introduces Classes of Service (CoS) e.g. aggregates, together with its specific features like Behaviour Aggregate (BA), Per Hop Behaviour (PHB), Ordered Aggregate (OA), Per Hop Scheduling Class (PSC), Differentiated Services Code Point (DSCP). DiffServ eliminates the need for per-flow state processing, scalability in large networks, and provides the required per-class QoS treatment to BAs. However DiffServ has one distinctive deficiency, it does not guarantee the required bandwidth for the traffic flows. IntServ/RSVP and DiffServ complement each other's limitations and may be used in combination [7] in order to accomplish different types of classification and traffic control at different network parts for E2E QoS provision. Traffic is identified and marked with particular DSCP at the edge network depending on IntServ parameters and RSVP reservations, and transmitted in the core transparently.

- *MPLS role in 4G*

A new QoS approach, namely MPLS protocol [8] has been reviewed in the context of next generation networks. MPLS originally is a TE and routing protocol, which combines simplified forwarding mechanism at the data link layer with network layer routing, thus supporting and/or being compatible with the various protocols on these two layers. MPLS brings connection-oriented paradigm into IP networks by reserving and maintaining Label Switched Paths (LSPs) for traffic transmission. QoS flavour is borrowed from DiffServ technology and accommodated by MPLS header in its Experimental Bits (EXP) field, thus supporting eight different levels or CoS. Two different ways of DSCP mapping into MPLS EXP field to serve LSPs are explained in the following section. The advantage of deploying MPLS together with DiffServ protocol relies on its ability to operate at the edge networks along with the core networks and process traffic at the layer 2 or 2.5, thus dramatically reducing traffic processing delay. Other advantages of MPLS are: better price/performance ratio of routing, improved scalability in the network layer and greater flexibility in the delivery of routing services [9, 10].

Using MPLS TE and QoS features in multi-access wireless networks, by incorporating MPLS functionality into Base Station Controllers (BSCs) and RNCs as well as in SGSN and GGSN will further extend E2E LSP setup, and consequently QoS and TE support. Chaskar et.al at Nokia Research Centre [11], propose a combination of MPLS and DiffServ to be used for advanced route setup and control between SGSN and GGSN with particular emphasis on LSP establishment with Label Distribution Protocol (LDP) signaling protocol to establish aggregate GTP tunnels, for better TE and QoS support.

- *UMTS/GPRS CoS encapsulation into MPLS/DiffServ CoS*

QoS of UMTS network defines service levels provided to subscribers on demand and outlines four different CoS:

– *Conversational*, used for conveying real-time traffic streams and multimedia applications with strict requirement of low delay and jitter;
– *Streaming*, used for audio/video traffic streams with stringent demands on low jitter;
– *Interactive*, used for data transmission, requiring low round trip delay;
– *Background*, equivalent to Internet's best effort service.

When UMTS/GPRS networks communicate through MPLS/DiffServ enabled IP backbone, traffic arriving from a UMTS/GPRS access domain to SGSN/GGSN (or from the core network to the access network) should undergo QoS parameters translation, e.g. mapping, so that a new system treated traffic flows according to the service requirements. Table 1a shows the actual mapping of UMTS CoS into MPLS/DiffServ CoS for the following set of conditions, compliant with [12]:

– LSP type: EXP-Inferred-PHB Service Class (PSC) LSP, (E-LSP);
– "PHB Mapping into Encapsulation (Encaps) Layer" mode: per signaled or pre-configured;
– "PHB → Encaps Mapping" type: PHB →EXP mapping;
– To be applied to: a set of BAs which share an ordering constraint (OAs).

MPLS EXP	DSCP	UMTS
000	BE	Out of order packet
001	AF21	Background
010	AF22	Interactive
011	AF23	Interactive
100	AF11	Streaming
101	AF12	Streaming
110	AF13	Streaming
111	EF	Conversational

Table 1a. Mapping of UMTS CoS into MPLS/DiffServ CoS for E-LSP

E-LSP serves a set of OAs for a given destination address (FEC) and supports eight or less BAs. EXP bits of the MPLS header are used by LERs and LSRs to determine PHB to be applied to the packets arriving from an access network or transiting an MPLS enabled domain. EXP field includes PSC and drop precedence priorities. E-LSPs are scalable due to smaller number of labels used, however per-LSP bandwidth reservation reduces the advantage of PSC-based granularity and bandwidth availability in queues for all PSCs.

Another type of LSP can serve a single FEC-OA pair, for which PSCs are explicitly signaled and read directly from the label value, regardless of EXP field [13]. EXP field shows only the drop precedence priority. Such LSP is called Label-Only-inferred-PSC LSP (L-LSP). Table 1b shows possible mapping of UMTS CoS into MPLS/DiffServ CoS, under the following set of conditions:

– LSP type: L-LSP;

– “PHB Mapping into Encapsulation (Encaps) Layer” mode: explicitly signalled;
– “PHB → Encaps Mapping” type: EXP/PSC → PHB mapping;
– To be applied to: a single FEC-OA pair.

MPLS EXP	PSC	PHB	UMTS
000	DF	DF	Background
-	-	-	-
-	-	-	-
000	CS *n*	CS *n*	-
001	AF *n*	AF*n*1	Streaming/Interactive
010	AF *n*	AF*n*2	Streaming/Interactive
011	AF *n*	AF1*n*3	Streaming/Interactive
000	EF	EF	Conversational

Table 1b. Mapping of UMTS CoS into MPLS/DiffServ CoS for L-LSP

CS field in the Table 1b indicates Class Selector and DF- Default Forwarding. In case of L-LSP establishment individual PSC of a FEC should be labeled, because label value contains scheduling information. This form of LSP provision is less scalable, however more granular in terms of bandwidth allocation for each PSC.

4. Intersystem E2E QoS vision

Most of the current long distance and international networks are ATM based and at the same time modern IP networks provided by ISPs are IP routed. Hence the main question is whether or not an ATM switch or an IP router should be MPLS enabled. All-IP notion has already outlined the trend of future 4G networks. However among the issues raised by this all-IP drive the cost of the reconfiguration of existing networks is an important one. For this reason Figure 2 is constructed in such a manner that both connection-oriented and connectionless networks (and their protocols) to be taken into account in the perspective of emerged wireless (both fixed and mobile) technologies. Figure 2 visualizes 4G from E2E QoS perspective and emphasizes differentiation between core network, edge network, access network and end user (mobile host). Future core network will predominantly be IP based, however integrated networks may also be based on ATM and Frame Relay, due to the cost of network renovation. The edge in the block diagram includes both edges of core and access networks, therefore SGSN and GGSN, gateway WLAN routers, MPLS ingress and egress LERs all are parts of this section. The purpose of the multi-access wireless network block is to unite potential wireless access technologies together with wired networks, namely GSM, GPRS, EGPRS, EDGE, UMTS, WPAN (Bluetooth, 802.15), Ethernet, WLAN (HyperLAN/2, 802.11), WMAN (802.16), satellite networks.

QoS issues arise in most of the layers of the protocol architecture, however the complexity of QoS provision and assurance are mostly concentrated on the link layer, which deals with transit traffic treatment across neighbouring cells, and the network layer, which takes responsibility on mobility / location management. The control and management plane is also involved in controlling for the provision of guaranteed QoS levels, seamless handover (both horizontal and vertical) and reconfiguration of multi-access interfaces. Existing core networks are usually formed by different domains representing different networks owned by various operators.

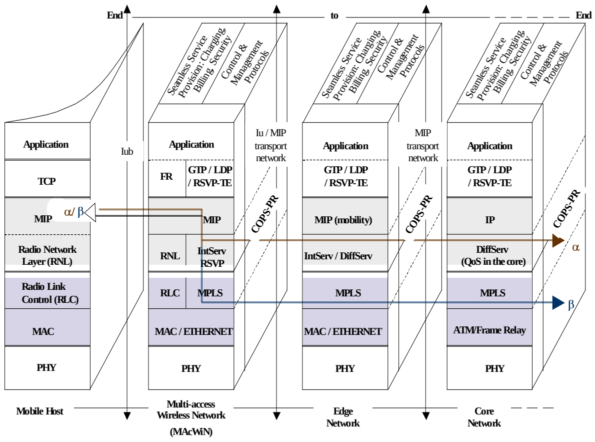

Fig.2. Intersystem E2E QoS vision

The core (current DiffServ enabled IP Backbones) is managed per aggregate of the network resources, and is over-provisioned, providing guaranteed bandwidth even at peak times. The requirement of the next generation network architecture though dictates a need for efficient techniques to eliminate unnecessary waste of network resources, at the same time maintaining the required guarantees of QoS, reliability, service availability and minimum management and intervention.

- *‘Option α‘ for E2E QoS*

Option α, (red dashed line on Figure 2), represents one way of E2E QoS support based on IP QoS protocols, IntServ and DiffServ. Due to all-IP-based structure of future networks, DiffServ protocol may be used in the core, as it provides and satisfies most of the requirements of 4G, such as CoS differentiation, independence from underlying link layer technology, scalability, relative simplicity, and has no signaling requirements. The issue of bandwidth guarantees can be handled by IntServ-enabled routers residing at the edge network. Appropriate information on reservations per user-service will be provided by QoS Broker (QB) entity and encoded in DSCP header of IP packets, to be signalled by RSVP across the network. DSCP-marked IP packets will be identified and treated accordingly by the core routers. IntServ/DiffServ deployment in multi-access wireless networks is a novel approach, targeting further extension of E2E QoS notion. For example, RSVP signaling function mapped into the Core Network Line Cards of RNCs, as well as into SGSNs and GGSNs will be able to provide mechanisms for the connections establishment and signaling information exchange between RNC and GGSN, and RNCs themselves. In this case the notion of the edge network and especially its functionality would move towards multi-access network domains (Figure 3a).

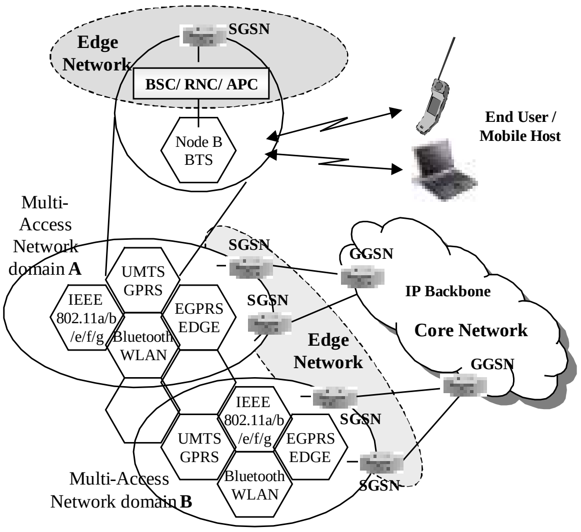

Fig.3a. Migration of edge network with regard to MuMAcWiN

Access network requirements are more dynamic in terms of per user-service management. The QB is a typical entity for access network management and control of points of attachment, i.e. access routers for various wireless domains. QB receives information on policy and Service Level Agreements (SLA) from a mobile host (MH) entering a network domain and takes responsibility for management and monitoring of resources in the access network and edge gateways as well as user requests according to user profile.

MH in the proposed scheme is a multi-mode device, capable of automatically connecting to different access networks including fixed, WLAN, GPRS, satellite networks etc. MH has minimal intervention into the scheme of E2E QoS support and maintenance, and therefore does not fall within the scope of this paper.

- *'Option β' for E2E QoS*

Relatively new and optimised approach would be to deploy MPLS functionality in the core IP routers (blue dashed line on Figure 2) and benefit from DiffServ QoS through MPLS, for more efficient CoS support in application-specific backbone. MPLS, with RSVP-TE as a transport label distribution protocol, brings in additional features such as:

- TE, for better utilization of available bandwidth;
- Constraint-based routing, for congestion avoidance and more flexible and efficient network usage for operators;
- Virtual private LAN Service (VPLS);
- Ethernet Point-to-Point VPNs.

Interoperability test operations and interworking of various implementations are being accomplished regularly by MPLS/Frame Relay Forum [14] including a number of vendors. Recent results in regard to MPLS DiffServ / TE backbone showed successful PHB treatment and mapping between EXP bits and PHBs for AF, EF and BE traffics in Label Switched Routers (LSRs). RSVP-TE and OSPF-TE basic tests were run without signaling or routing issues.

In MPLS enabled networking environment notions of core, edge and access networks depend on whether the nodes support MPLS-specific functionality or not. In this case "edge" is formed by routers (ingress or egress with respect to MPLS domain), which are connected to nodes that either do not support MPLS features or reside outside of the MPLS network. These types of routers are called Label Edge Routers.

In the proposed architecture (Figure 3b) the edge network is shifted to the area, which includes RNC, BSC and APC only. SGSN and GGSN retain their own functions and additionally run MPLS, therefore are also capable of accomplishing LSR functions. Incorporation of MPLS features in multi-access network domains, particularly inside the controllers, is a novel approach, proposed to further extend the idea of E2E QoS and TE. In this scenario edge, or an ingress interface, communicates SLAs with the users and stores the information in a label switching table, which is associated with the user profile entries. On the arrival of a packet LER reads the information from the table and treats the packet according to its SLA requirements. The major issues in the 4G multi-access wireless networks, namely QoS and mobility management, vertical and horizontal handovers and related signaling, and inefficient location management schemes leading to connection interruption are discussed in [15, 16] in the context of MPLS. Lightweight signaling with LDP and/or RSVP-TE, OSPF-TE and loosely explicit LSP setup may reduce or eliminate problems like heavy and complicated handover and location management in wireless domains.

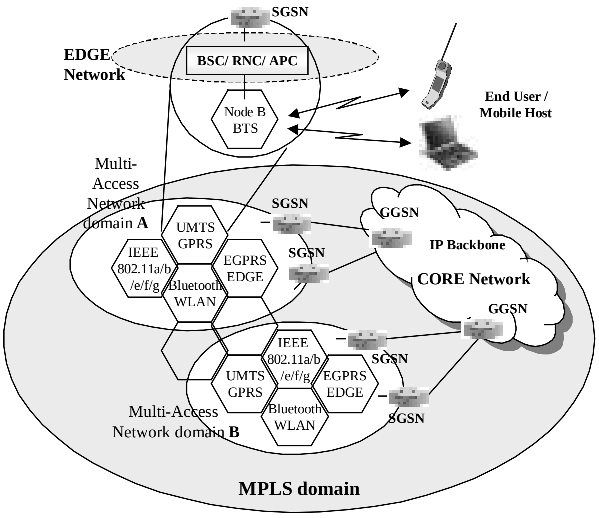

Fig.3b. MPLS core network and edge network formation

5. Conclusion

QoS is an important issue that should be addressed to provide acceptable and predictable CoS to the end user. In order to support the requirements of real-time and multimedia applications 4G networks should be able to provide unified QoS across heterogeneous network environment. The paper presented an all-IP based MuMAcWiNs tightly coupled architecture and outlined the provision of E2E QoS support. Appropriate QoS support protocols for fixed and mobile wireless networks have been reviewed and GPRS/UMTS mechanisms for CoS encapsulation into MPLS header field have been discussed in detail. Intersystem E2E QoS vision has been proposed in terms of layered protocol architecture blocks with distinctive differentiation of network functionalities in the core, edge, multi-access networks and mobile host. Migration of the functionalities of these network parts, invoked by deployment of the two different QoS schemes has been demonstrated and justified.

Acknowledgements

Authors would like to acknowledge the support of University of Limerick and ANWIRE project (www.anwire.org), which is funded by the European Community under the contract IST-2001-38835, within the framework of which this work has been produced.

References:

[1] Nokia "Next generation network architecture",www.nokia.com/nokia/0,8764,43318,00.html
[2] ETSI, "Requirements and Architectures for Interworking between HIPERLAN/2 and 3rd Generation Cellular Systems," Tech. Rep. TR 101 957, Aug. 2001.
[3] Braden R., et al. "Integrated Services in the Internet Architecture: an Overview," RFC 1633, Jun. 1994.
[4] Grossman D. "New Terminology and Clarifications for DiffServ", RFC3260, Apr. 2002.
[5] Puuskari M., "Quality of Service Framework in GPRS and Evolution towards UMTS", 3rd EPMCC, March 1999.
[6] Priggouris G., et. al., "Supporting IP QoS in the General Packet Radio Service", IEEE Network, pp. 8-17, Oct. 2000.
[7] Bernet Y., et. al., "A Framework for Integrated Services Operation over Diffserv Networks" RFC 2998, IETF, Nov. 2000.
[8] Rosen E., et.al., "MPLS Architecture" RFC 3031, IETF, Jan. 2001.
[9] Callon R. et. al., "A Framework for MPLS", Internet Draft IETF, 1999, at draft-ietf-mpls-framework-05.txt.
[10] Isoyama K. et. al., "Policy Framework QoS Information Model for MPLS", Internet Draft IETF, 2000, at draft-isoyama-policy-mpls-info-model-00.txt.
[11] Chaskar H., et. al. "MPLS and DiffServ for UMTS QoS in GPRS Core Network Architecture" ISOC Conference INET, Jun. 2001.
[12] Le Faucher F., et. al., "MPLS Support of Differentiated Services" RFC 3270, IETF, May 2002.
[13] Fineberg V., "QoS Support in MPLS Networks" MPLS Forum, White paper, May 2003.
[14] MPLS World Congress "New Revenue Streams With MPLS Service Differentiation" Public Interoperability Event, Test Plan and Results, EANTC, Feb. 2004.
[15] Kubinidze N., M. O'Droma, I. Ganchev, "On Deployment of MPLS Functionality in Next-Generation Wireless Networks," 2nd ANWIRE Workshop on Reconfigurability, Sep. 2003.
[16] Kubinidze N., M. O'Droma, I. Ganchev, "Integrated algorithm of HMIPv6 and MPLS for NS to support next generation wireless networks" Conference QoS 2004, The IEE, March 2004.