

# **Media Services in the IMS: Evolution for Innovation**

**Eric Burger**

*Vice President  
CTO of Next-Generation Communications*

**Guy Redmill**

*EMEA Market Development Manager*



**cantata**  
TECHNOLOGY

# table of contents

<b>Abstract</b>	<b>3</b>
<b>Introduction</b>	<b>3</b>
<b>New Service Architecture</b> Figure : Representation of IMS Framework Architecture	<b>4</b> 4
<b>MRFC/P Decomposition</b> Figure 2: Architecture for Media-Rich Services	<b>5</b> 5
<b>Alternatives for the Mp Interface</b> H.248 SIP SIP + Markup: VoiceXML and MSCML	<b>7</b> 7 7 8
<b>Deployment Issues</b>	<b>9</b>
<b>Conclusion</b>	<b>9</b>
<b>Appendix 1: Configuration of IM Subsystem Entities</b>	<b>10</b>
<b>Appendix 2: Functional Architecture for the Provision of Service in the IMS</b>	<b>11</b>
<b>Key</b>	<b>11</b>

## abstract

*SIP is a critical building block for service delivery in next generation networks. It has a prominent role in the IP Multimedia Subsystem (IMS), a new framework for service delivery and management in next generation fixed and mobile networks. The media resource function (MRF) within the IMS is responsible for the delivery of media services. Evolution of the MRF from its original specification towards a SIP-orientated architecture is advocated. In commercial practice, while there are virtually no deployments of the MRF in the IMS as specified by 3GPP, an alternative MRF architecture, based on SIP, has been successfully deployed and is already generating revenue for service providers. Several models for the decomposed MRF are discussed, but true flexibility is only made possible by the adoption of SIP throughout the architecture and the elimination of H.248, which offers no benefits and many disadvantages.*

*An alternative MRF framework is proposed that offers an open, standardized approach to service creation via application-level protocols, including SIP and specialized markup languages, such as VoiceXML and MSCML. This paper examines the benefits of a SIP driven IMS architecture for controlling the media resource function and the inherent limitations of the alternative, H.248 driven framework for the MRF.*

## introduction

The Session Initiation Protocol or SIP is one of the driving technologies behind the telecom industry's shift from circuit switched TDM based infrastructure to the more flexible and web-centric packet network infrastructure of IP telephony. Over the last four years, SIP has been adopted by the vendor, development, and service provider communities as the protocol of choice for developing and deploying next generation IP services.

Its flexibility and openness have made it a critical building block for the creation of innovative enhanced services, such as voice and video messaging, conferencing and prepaid calling. SIP has become a cornerstone of "new" next generation architectures for delivering and controlling services. With growing adoption of packet communications, service providers are transitioning from a voice-centric orientation; the connection and delivery of voice calls; to a service-centric model focused on the rapid and efficient delivery of innovative, value-added services to enterprises and consumers. This shift has been apparent for some time. Service providers can no longer rely on basic voice transport services to drive revenue. Ovum Research predicts that service revenue will exceed basic call revenue by 2006. SIP is the critical glue within these new network topologies for developing and delivering generation upon generation of enhanced services.

Although there are many competing visions for how service delivery can be achieved in next-generation networks, there is one framework that is currently gaining considerable momentum. The IMS or IP Multimedia Subsystem was originally proposed for Third Generation (3G) mobile networks. However, "the IMS can be used for any mobile access technology"<sup>1</sup>. Moreover, because it is abstracted from the network access layer it "can be seen as the next generation service delivery platform framework" for any network<sup>2</sup>. The IMS is thus seen by many proponents as being the foundation of next-generation services, irrespective of the network access mode (*i.e. for wireless and wireline applications*). SIP is an integral part of this framework and is utilized throughout the architecture as proposed by several standards bodies.

However, there is one key exception to the adoption of SIP within the IMS. This occurs at the level of media services, where a different protocol (*H.248, sometimes known as MEGACO*) has been proposed for deployment. Media services are to be realized through a series of functional elements in the IMS, collectively defined as the Media Resource Function, or MRF. While this approach, as originally conceived,

---

<sup>1</sup> Magedanz et al, Fokus @ The Fraunhofer Institute.

<sup>2</sup> IBID

has theoretical merits, real-world implementation experience has highlighted its limitations. This paper will explore the issues surrounding the SIP-based model for control of the media server resource function versus the alternative H.248 model and illustrate why the SIP model is the more appropriate model moving forward.

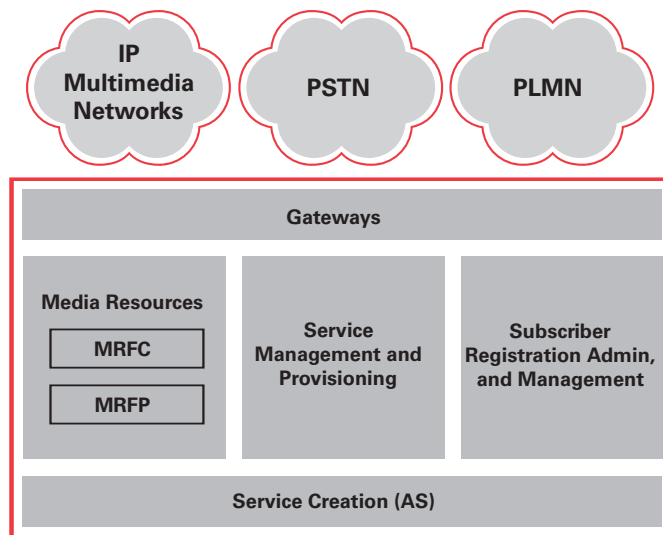
A second key issue concerns the architecture of the Media Resource Function itself. The initial specification mandated further functional separation into two separate units, a media resource processor function and a media resource function controller, mirroring the disaggregation of the Media Gateway and Media Gateway Controller. H.248's role within this architecture was for the control of functions between the two sub elements. However, it has become clear that there are a number of alternative options for the MRF architecture that are also valid and we shall explore these in subsequent sections.

## **new service architectures**

The evolution of standards for ensuring compatibility and consistency in capabilities in mobile networks is now largely the responsibility of two bodies, the 3GPP and its North American and Asian counterpart, the 3GPP2. The two groups collaborated to ensure that the foundation of future service delivery will reside in a new platform, known as the IP Multimedia Subsystem or IMS. This defines a complete architecture for distributed enhanced service delivery appropriate to the enhanced transport capabilities planned for third generation mobile networks—offering greater bandwidth across the radio interface to the user and permitting a wider range of value added services.

The IMS is itself decomposed into a number of functional elements, each of which is clearly defined and each of which presents specific interfaces to related components. Further, the architecture makes explicit the distinction between service delivery and service provisioning. Thus, a further architectural framework is defined to encompass devices associated with service creation and management. A simplified representation of the IMS framework is provided in Figure 1. A fuller picture is provided in the appendices.

**Figure 1: Representation of IMS Framework Architecture**



The service delivery process requires the connection of user equipment (via media streams) to media processing equipment; this activity takes place under the direction of the service-provisioning layer. This duality is important because it retains flexibility and may permit considerable innovation in terms of future service offerings. It is important to remember that the IMS framework is not designed to offer a range of specific services, but rather to provide a platform for any conceivable multimedia service to be delivered in both current and future networks. The key component of the service-provisioning plane is the Application Server (AS). This device offers value-added services to subscribers and can host single or multiple services, depending upon implementation.

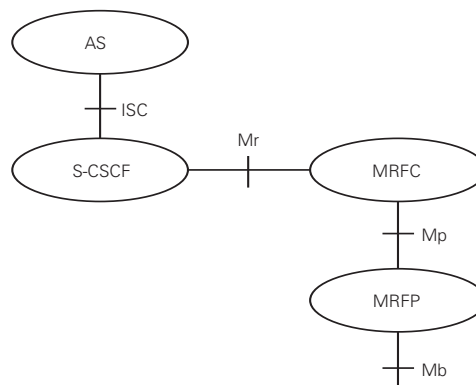
The Application Server interacts with the Media Resource Function (MRF) to ensure delivery of multimedia services. In fact, media functionality within the IMS is most commonly realized through the inclusion of dedicated Media Server platforms, the purpose of which is to provide unique and diverse services through common and shared media resource processing capabilities.

## MRFC/P decomposition

The 3GPP R99 standards proposed a decomposition of the MRF into two separate elements: the Media Resource Function Controller (MRFC) and the Media Resource Function Processor (MRFP). Figure 2 illustrates this relationship, which has been maintained in subsequent revisions of the architecture. The ISC reference point is crucial to service delivery as it connects the media resource function to the CSCF and thus to the service provisioning plane. SIP has been defined as the signaling interface of choice for this interface.

The MRFC/MRFP division was prompted, in part, by a desire to draw on work in the PacketCable and IPCC groups based on creating a common architecture for IP communications equipment. This activity had led to a logical and physical separation of media access from media control. This approach was realized with the definition of the Media Gateway (MG), to provide TDM/IP access and conversion, and the Media Gateway Controller (MGC), to effectively drive the MG. The interface from the MGC to control the access device, the MG, was chosen as H.248 (or "MEGACO"), a joint initiative between the ITU-T and the IETF. However, neither the PacketCable nor IPCC groups promote decomposition of the MRF into two entities. In these architectures, there is a single media server.

**Figure 2: Architecture for Media-Rich Services<sup>3</sup>**



<sup>3</sup> 3GPP TS 23.228 V6.7.0

Thus, although it might have appeared sensible to the 3GPP to replicate the MGC/MG logical and physical distinction and extend this into the MRFC and MRFP, several factors militate against this. Firstly, there simply isn't the gross distinction between control processing and media processing that is evident in the MGC/MG decomposition. In access and core networks, media (call) sessions need to be routed to appropriate onward destinations and different media formats need to be rendered compatible. The MGC/MG division is perfect for this task as there are many potential destinations for media and many potential sources of media streams. However, within the IMS itself, the routing/formatting tasks will likely have been dealt with by the time the media arrives at the IMS (having been resolved in the routing process); hence, there is no need to replicate this functionality to the same extent as, for example, at the edge of the packet domain. The MGC/MG division as originally conceived is thus too heavyweight for deployment at this level of the network which by definition serves only a proportion of available traffic at any one time. Secondly, one of the tasks of the MGC is to co-ordinate switching functions in the access network. This is not applicable to the MRF.

In fact, a number of different instantiations of the MRFC/MRFP architecture are possible and have their own proponents. For example, one alternative rests on the premise that much of the activity that might be required in the MRFC is actually carried out in the AS. The billing, roaming, and routing functionality already exists in the AS and CSCF and hence replication in the MRFC is unnecessary. The original purpose of the MRFC was conference control (now recognized as an AS function), roaming control (now recognized as a CSCF function), and media control (*now recognized as integrated into the MRFP*). Many of the 3GPP documents reflect this fact by leaving the role of the MRFC unspecified, except to say that the functionality occurs in the AS. Under this architecture, the MRFC is redundant as a discrete entity, with its functionality removed to the AS.

Another alternative limits the function of the MRFC to resource location and management, with other intended functions integrated into the CSCF. A third architecture has the MRFC embedded into the MRFP. Indeed, there are sound reasons why each is valid. It is important to recognize that no vendor today offers an H.248-based MRFC device in the open market.

Given that variations on the theme of MRFC/MRFP are possible, it is appropriate to consider the interfaces used by the MRF itself, that is, the Mr and Mp interfaces. SIP has been selected as the signaling protocol for the Mr interface (between S-CSCF and MRFC), suggesting that further conversion between the MRFC and MRFP to H.248 for the Mp interface is superfluous. For completeness, we will consider options for the Mp interface between the MRFC and the MRFP. In the following sections we will explore issues with H.248 and examine the alternatives for this interface.

Factors that we will consider include:

- Technical fit of the protocol to the needs of application developers
- Technical fit of the protocol to the needs of service providers
- Commercial availability of platforms offering the protocol

# alternatives for the Mp interface

## H.248

Among the possible alternatives for the Mp interface is H.248, which has the virtue of being currently specified by the 3GPP for the MG-MGC interface. For the MG-MGC interface, H.248 is clearly the best choice for the transport-oriented nature of the MG. However, as a master/slave protocol, H.248 is limited for use in the enhanced services environment, which is the domain of the MRF. In fact, its use was originally mandated when the primary service envisioned was conferencing, whereas today we need to consider a wide range of 3GPP service offerings and the possibility that yet more will emerge. The distinction between access and transport with master/slave protocols; and applications and services, with client/server protocols, is critically important for the development and deployment of enhanced services. Indeed, a fully-interoperable H.248-based MRFP supporting conferencing is not currently available in the marketplace, despite this being the rationale behind the MRFC/MRFP separation.

Enhanced service applications today typically leverage the SIP protocol. Developers think in terms of application level constructs, such as SIP dialogues (*call legs*) and service requests (*announcements, IVR, conference mixing, video, etc*). Few applications or developers operate at the level of media resources as such (RTP streams, DTMF detectors, players, mixers, etc). The target for H.248 is the low-level, transport oriented control of in-band signaling DSP resources, not application level constructs and service resources. Consequently, if one chooses H.248 as the Mp interface to the MRFP, an intermediary is required to convey application information and to convert it into something intelligible to media servers.

One proposed intermediary is the MRFC. However, although the MRFC can perform the task of translating between SIP and H.248, it is a commercial fact that most multi-function media servers already offer SIP and Markup-language (*e.g. XML*) interfaces, making this function redundant. In addition to adding network complexity and cost for service providers, it also limits the variety of services that can be offered by a service provider.

One of the most powerful advances in media server technology is the use of VoiceXML to specify IVR interactions. Unfortunately, the processing model of VoiceXML and H.248 are wholly incompatible. For example, a VoiceXML script can terminate a session or start a new call, neither of which activities fits into the master/slave model of H.248. This means that only applications of limited utility or of lower quality can be offered using H.248 as the Mp interface.

There are virtually no commercial, carrier-grade media servers with generally available H.248-controlled media services (MRFP). Without a vibrant market for these devices, service providers are effectively locked-in to a limited set of platforms. Moreover, from a technical point of view, the added complexity of unnecessary network elements that, in fact, remove value from the network, clearly makes H.248 an unsuitable alternative for the Mp interface.

## SIP

The *Basic Network Media Services with SIP* standard, also known as the IETF "netann" Convention, describes how an AS can invoke basic services at a multifunction media server. Cantata created the convention in 2000 simultaneously to the invention of the first SIP-controlled media server.

Netann is suitable for simple MRF functions, such as playing announcements (often needed by the CSCF) or simple 3-way or 6-way conferences. The model for netann is the Send-to-Resource function of CS-4 (ITU-T Q.1241). From the CSCF's or AS's perspective, they "forward" the call to the MRF, where the user part of the destination address informs the MRF what function to implement.

Being so simple, the netann interface enables the rapid development and deployment of basic enhanced services. Furthermore, the model introduces the concept of using call-level constructs, such as call legs and application-level resources, to the AS. As indicated above, this enables the delivery of more rich and robust applications in a much shorter time period with a higher quality than using device control-oriented protocols.

SIP has clearly demonstrated its market dominance in the telecommunications arena through its adoption by 3GPP as the application-level signaling standard. In addition, it is the standard adopted by new entrants such as Microsoft and AOL. All of the major media server vendors that offer any SIP interface use the netann interface.

### **SIP + Markup: VoiceXML & MSCML**

As we previously described, applications are typically written using constructs such as SIP dialogues (call legs) and service requests (“play announcement” “perform IVR” “mix streams”, etc). In fact, far from having to shoehorn this model into something that is basically incompatible (H.248 or MOML/MSML), there is an existing suite of protocols that directly implement these requirements, namely SIP, SIP with VoiceXML (for IVR), and SIP with MSCML (for conference control).

SIP provides routing and session establishment capabilities for applications, using SDP Offer/Answer (RFC3264) mechanisms. Additionally, one can directly use simple announcements and simple add-a-leg conference mixing services using SIP alone, following the conventions of the “Network Announcements with SIP” standard. Moreover, using the same standard, one can also launch a VoiceXML session to allow for more sophisticated media services to be realized.

By strictly adhering to the VoiceXML standard, the Cantata SnowShore IP Media Server™ allows full flexibility and interoperation with the largest number of applications from application vendors. In particular, dialogue results are always returned using HTTP. In contrast, there are implementations that return results using SIP for transport. However, such implementations violate the VoiceXML model. One might try to abuse SIP by attempting to re-implement HTTP over SIP or, worse, modify the VoiceXML state machine. However, this means that service providers will be locked-in to a proprietary platform with a limited number of application partners, if any.

While the design goal of VoiceXML is sufficient for describing IVR interactions, it is not suitable for conference control. MSCML, an IETF proposal that is in use by over 15 application developers and offered by more than three media server vendors, enables control of advanced conferences at the level of the application developer. Specifically, the application developer uses SIP-level constructs for conference-level requests, such as recording the conference, recording a leg, doing basic IVR on a leg, forming sub-conferences or sidebars, forming coaching or whisper-in-your-ear arrangements, etc. It is particularly important to remember that MSCML does all of this at the application layer, not at the media layer. This means that application developers can deliver more features and functionality in a shorter time period, with fewer bugs. This directly results in enhanced revenues and fewer customer complaints for the service provider.

The fact that there are multiple independent implementations of AS and SIP+MSCML-controlled MRFs provides validation of MSCML as a likely future standard. Choosing SIP+MSCML means the service provider does not get locked-in to a single vendor protocol implementation. The service provider is able to source their MRF using the technology that is suitable for their needs, such as full-custom proprietary hardware as well as Cantata’s proven, carrier-class open-platform implementations. Note that of all of the carrier-class media server vendors, only Cantata has delivered on the promise of doubling performance every 18 months. Contrast this to equipment that uses transport protocols such as H.248 or transport-oriented, proprietary protocols where there have effectively been no performance improvements for over five years. This is an example of the value of truly open, multi-vendor standards and the competitive marketplace.

## deployment issues

In addition to discussion of the merits of SIP/Markup language control of the Mp interface as opposed to the currently-specified H.248, it is also appropriate to consider how different vendors are approaching the challenge of deploying the MRF, given the current position of the standards.

One major vendor has chosen to embed the MRFC into the CSCF, while retaining H.248 as the interface to the MRFP. By adopting this approach, they have effectively prohibited operators from selecting both independent MRFC and MRFP devices. Another vendor has taken a different tack and has integrated the MRFC into the AS. In their architecture, there is a direct interface into the MRFP, for which SIP has been chosen. Under the current standards, there are issues with both implementations as they stand, although each offers value. Clearly, the current position is not only unsatisfactory from a technical standpoint, but also does not meet the needs of operators and vendors. However, these issues can be readily resolved if the Mp interface is migrated towards SIP.

By standardizing on SIP, deployment flexibility—and vendor choice—is secured. For example, the relationship between the AS/CSCF/MRFC/MRFP can be determined by the operator. For small-scale IMS deployments, it may be determined that no MRFC is required. Alternatively, as deployments grow and multiple MRFPs are required, the operator may choose to insert a MRFC. With SIP, this becomes a straightforward matter. In this case, the MRFC assumes the role of a SIP proxy, locating resources in different MRFPs as required by the AS.

There is an interesting corollary of this in that although the MRFC can add value if there is a single AS, things can be more problematic where there are multiple AS devices from different vendors using a single MRFP. In such cases, integration can be problematic. Such challenges are minimized with consistent use of SIP throughout the MRF architecture.

SIP then, allows operators to choose the deployment option that is most appropriate to their goals and desires, and promotes true vendor independence. They can simply select components from whichever vendor is most suitable to their needs.

## conclusion

As in any technology adoption cycle, communications protocols undergo a transition from concept to reality as they move from laboratory trials to field deployment. Protocols and technologies can evolve based on real-world experience of changing circumstances. Although H.248 has its merits, these are far outweighed by the advantages of adopting a consistent use of SIP throughout the IMS for media service delivery and reducing the complexity inherent in the current framework model. H.248 limits the scope of the IMS to support future multimedia services—services that are vital to operator and service provider revenue, but many of which are as yet unimagined. It also limits deployment options and choice in vendor selection. It is this need to allow maximum future creativity that demands that the IMS becomes as flexible as possible with regard to service delivery. This is a view endorsed by a number of companies active in this field—Openwave, Ubiquity as well as Cantata Technology.

SIP is critical to flexible service delivery in next-generation networks. The IMS model provides a framework architecture for realizing current and next-generation services and includes detailed specifications to cater for a rich array of multimedia services. However, the Media Resource Function of the IMS needs to reflect the

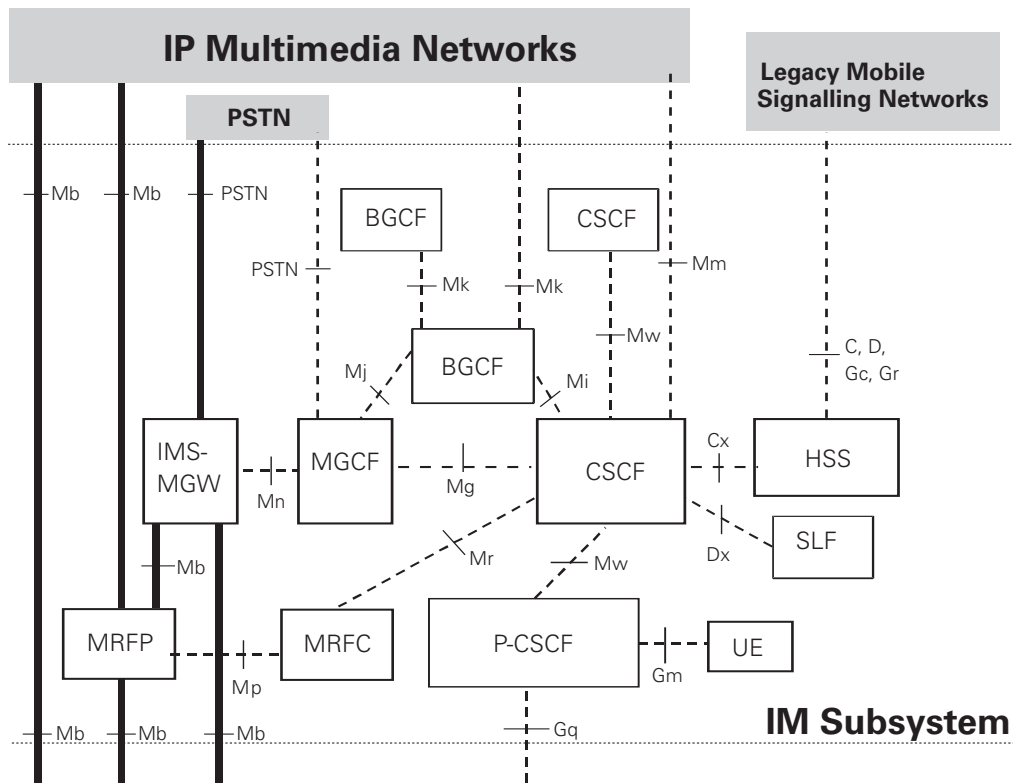
lessons learned over the past four years of field experience. As currently envisaged it is too rigid; too complex; includes the potential for duplication in the MRFC functions and distorts the simplicity promoted by the use elsewhere of SIP as a control protocol for media services. This is borne out by the multiple instantiations of the MRF that have emerged from different vendors. Further, it severely limits application programming models and approaches, which may have serious repercussions for future innovation and hence revenue generation. The adoption of SIP as the Mp interface will promote application flexibility and simplicity. This will also allow the MRFC to be deployed in the operator's chosen configuration.

In fact, the use of SIP as the Mp interface offers a further benefit in that the MRFC can act as a SIP proxy, allowing operators unlimited choices in both the range of application servers that they deploy and in the number of MRFP devices that are installed. Overall, the MRF can use the Mr interface, which is SIP, with additional instructions being encoded as appropriate in VoiceXML and MSCML. This is consistent with the overall IMS architecture, which derives much of its strength from use of the SIP protocol.

## appendix 1: configuration of IM subsystem entities<sup>4</sup>

### Legend:

- Bold Lines: Interfaces Supporting User Traffic
- Dashed Lines: Interfaces Supporting Only Signalling

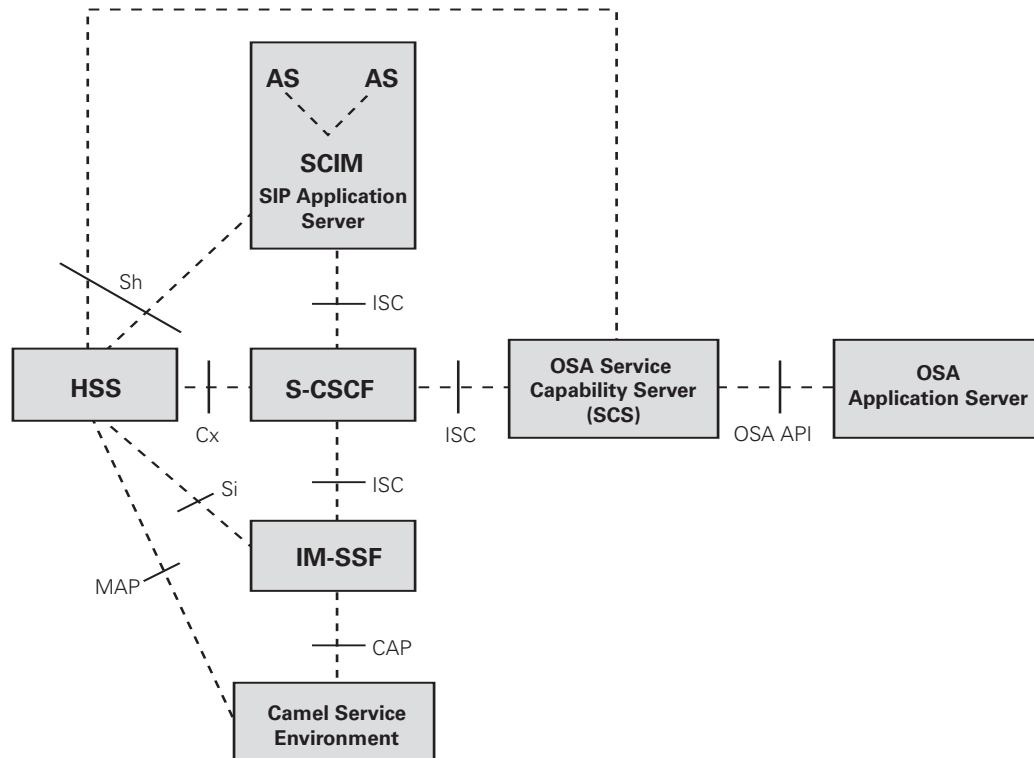


<sup>4</sup> From: 3GPP TS 23.002 V6.4.0

## appendix 2: functional architecture for the provision of service in the IMS<sup>5</sup>

### Legend:

Bold Lines: Interfaces Supporting User Traffic  
 Dashed Lines: Interfaces Supporting Only Signalling



### key

<b>3GPP:</b>	Third-Generation Partnership Project	<b>MRF:</b>	Media Resource Function
<b>3GPP2:</b>	Third-Generation Partnership Project 2	<b>MRFC:</b>	Media Resource Function Controller
<b>AS:</b>	Application Server	<b>MRFP:</b>	Media Resource Function Processor
<b>BGCF:</b>	Breakout Gateway Control Function	<b>MSCML:</b>	Media Server Control Markup Language
<b>CAP:</b>	Camel Application Part	<b>OSA:</b>	Open Service Access
<b>CSCF:</b>	Call Session Control Function	<b>P-CSCF:</b>	Proxy Call Session Control Function
<b>HSS:</b>	Home Subscriber Service	<b>PSTN:</b>	Public Switched Telephone Network
<b>IM-SSF:</b>	IP Multimedia Subsystem Service Switching Function	<b>SCIM:</b>	Service Capability Application Manager
<b>IMS:</b>	IP Multimedia Subsystem	<b>S-CSCF:</b>	Serving Call Session Control Function
<b>IMS-MGW:</b>	IP Multimedia Subsystem Media Gateway Function	<b>SIP:</b>	Session Initiation Protocol
<b>MGCF:</b>	Media Gateway Control Function	<b>SLF:</b>	Subscription Location Function
		<b>UE:</b>	User Equipment
		<b>VoiceXML:</b>	Voice Extensible Markup Language

<sup>5</sup> IBID



**cantata**  
TECHNOLOGY

**Corporate Headquarters**

410 First Avenue  
Needham, MA 02494  
USA

**Tel:** +1 (781) 449-4100

**Fax:** +1 (781) 449-9009

**Email:** [info@cantata.com](mailto:info@cantata.com)

Cantata Technology maintains multiple locations worldwide in North America, Asia and Europe.

**[www.cantata.com](http://www.cantata.com)**