Prerequisites

H.M.P. 3.0

Date : **Sep. 2007**
Version : **V 1.2 EN**

DOCUMENT DISCLAIMER

While every reasonable precaution has been taken in the preparation of this document, neither the author nor VOCALCOM development or support teams assumes responsibility for errors or omissions, or for damages resulting from the use of the information contained herein.
The information contained in this document is believed to be accurate. However, no guarantee is provided. Any trademarks referenced in this document are the property of their respective owners.

LEGAL DISCLAIMER

Table of Contents

1 Overview

This document describes the necessary prerequisites, requirements and implementation for HMP 3.0.

Dialogic® Host Media Processing (HMP) Software performs media processing tasks on general-purpose servers without requiring the use of specialized hardware. Dialogic HMP software provides media services for building flexible, scalable, and cost-effective next-generation media servers, converged telephony applications, gateways, and video portals. When installed on a system, Dialogic HMP software looks like a Dialogic® board with DM3 architecture to the customer application, but all media processing takes place on the host processor.

2 Vocalcom H.M.P. Solution Architecture

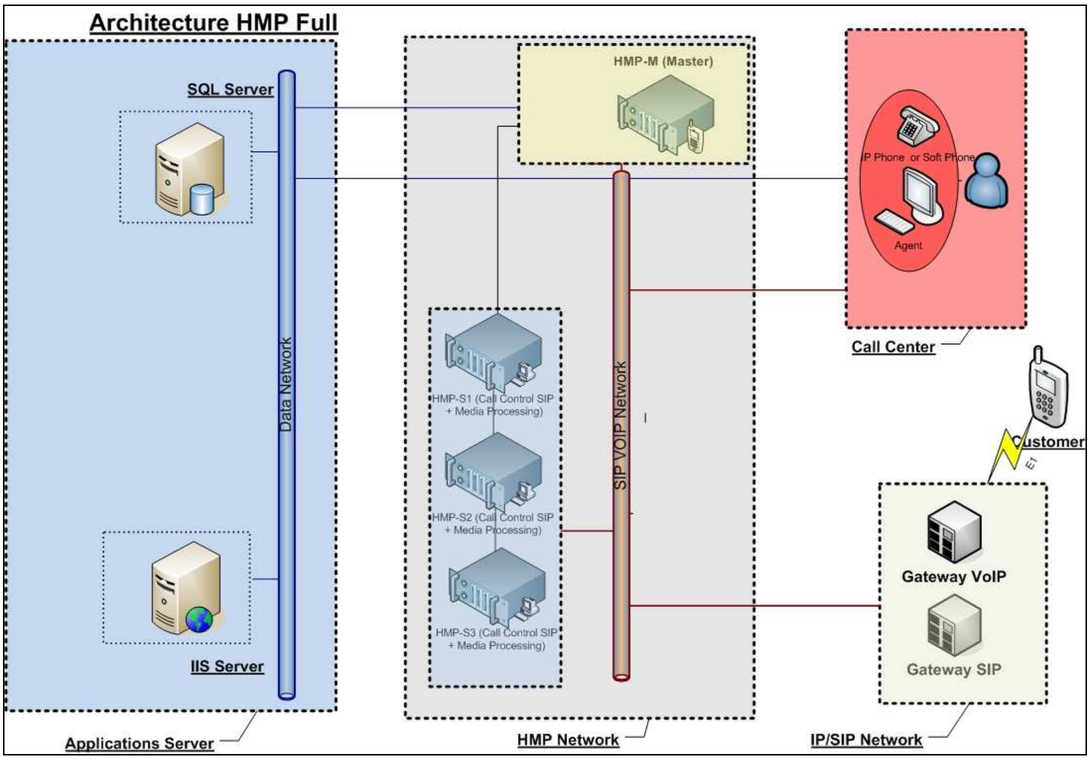
The Vocalcom H.M.P solution is built on a pure software architecture in order to manage the voice communication. The traditional ISDN (Integrated Services Digital Network) methods have been replaced by VOIP (Voice over Internet Protocol) technology that uses the SIP (Session Initiation Protocol) that is now the standard in multimedia telecommunications.

The schematic bellow describes the solutions architecture, this is a full VOIP solution connected directly to a SIP carrier for call termination and origination therefore both calls to the agents and the carrier are on VOIP. As described in the schematic it is possible to add one or multiple SIP gateways if the technical environment can not accommodate a SIP carrier this way the system can interface with tradition ISDN (E1/T1).

Regardless of the technology used, **all of the solutions components must imperatively take into consideration the SIP protocol and manage the re-invite** notion that allows you to change the codec's or address / IP port of the current RTP call destination traffic.

The call control components are based on the recommendation described in the **rfc 3725 document** concerning the implementation of the « Third Party Control Call (3pcc) » in the SIP protocol. This document is available at the following address http://www.ietf.org/rfc/rfc3725.txt

2.1 Full VOIP Architecture Schematic



The agents that are connected to the system can use a soft phone or a hardphone (IP Phone) to manage their calls. For softphone configuration; reference paragraph describing minimal requirements for Agents PC.

Notice:
PC network components today do not allow proper QoS, to minimize quality risks and poor voice quality related to loss of transmission packets, **we recommend the use of a hardphone (IP Phone)** then a Softphone.

2.2 HMP Servers Role and Capacity

Each HMP server is capable of managing up to 500 ports. The number of maximum ports depends on the codec that is utilized. In the event where there is no voice compression required (G.711) the number of available configurable ports will be 500.

The table bellow describes the number of available port based on the codec's compression levels. For a call center environment the table bellow can advise you on the number of lines and agents you can configure by server.

HMP Server Capacity		
Codec RTP	Number of ports per server (IVR)	Number of Agents per server
G 711	500	200
G 729	200	80

For a configuration inferior or equal to 200 agents (G.711) the hardware architecture will consist of one unique HMP server. This server's role will be Master and Slave.

The number of servers depends on the number agents, a minimum of 3 servers are required as 1 of the servers in dedicated as a Master.

The table bellow describes the number of servers required based on the number of agents.

Number of Agents	Number of Master Servers	Number of Slave Servers
Up to 200 – Up to 80	1	0
Up to 400 – Up to 160	1	2
Up to 600 – Up to 240	1	3
Up to 800 – Up to 320	1	4

* The red figures correspond to the maximum number of agents if G.711 Codec is used.
* The black figures correspond to the maximum number of agents if G.729 Codec is used.

Note:
Regardless of the codec, the maximum number of agents per system is fixed at 800, this is due to for virtual memory allocation issue.

As described above, there are therefore two distinguished types of HMP servers.

- The HMP- Master Server
- The HMP(s) _ S1 Slave Server

The role of the Master Server:
- Manage the all incoming and outgoing calls.
- Repartition the load in accordance to the available resources on each HMP Slave server.

The HMP Slave server's role (Call control SIP / Media Processing) is to treat and manage the inbound and outbound call traffic. The HMP Slave server holds all the voice, recording and conferencing resources.

2.2.1 HMP Server Configuration

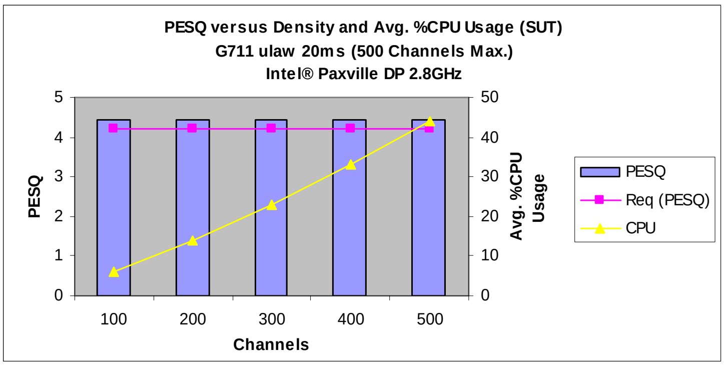
The hardware configuration detailed bellow is based on the Intel / Dialogic configuration. In conjunction with the resources that are software related, **it is imperative that the specifications bellow are respected.**

Operating System:
Windows server 2003 – Windows server 2003 R2 enterprise Edition

Hardware:
Dual Core CPU Support Intel **XEON 70xx**
Hyper-Threading Technology
4 GB DDRAM
HDD SAS

2.2.2 Example of CPU Load

- **Processor utilization**
  - Dual core Intel® Xeon® 7030 (2.8 GHz) running 500 channels of IVR at under 43% of CPU utilization
- **Densities Tested**
  - 500 Channels Voice Play/Record RTP G.711
  - 200 Channels Voice Play/Record with RTP G.729
  - 500 Conferees/server with RTP G.711
  - 240 Speech Integration with IVR and RTP G.711
  - 120 Channels H.263 Video Messaging
  - 120 Channels of Fax**



HMP r3.0 for Windows Doing Transcoding on a Dual Core Intel® Xeon® 7030 (2.8 GHz)

The gold of the PESQ (Perceptual Evaluation of Speech Quality) is to propose an objective model that is based on a psycho-acoustic human model. It performs a comparative calculation between two variables; the first is injected to an extremity and the same information is collected to the second extremity after having crossed and studied the systems. After the system has treated the PESQ send back the results based on a MOS scale. A 5 PESQ is equivalent to perfect quality.

The voice quality is verified in the following way:

- PESQ between 4 and 5 = Excellent – Good
- PESQ between 3 and 4 = Correct
- PESQ inferior to 3 = Poor

3 Network Equipments

Processing VOIP requires good quality bandwidth with a good response time and little latency, at the heart of the network your architecture needs to be sufficiently large (In terms of bandwidth).

The detailed requirements bellow are based on standard VOIP implemented solutions configurations.

3.1 Network Configuration

3.1.1 Switch

- All LAN Networking elements must be on switches, hubs are not recommended.
  The switches must have the ability to manage VLANs for transporting the voice packets

- In the event where the voice traffic is separated from the data, the VLAN can not surpass 500 IP Phones.

- In the event where the voice and data are on the same VLAN, each VLAN can not surpass 250 agent stations.

All equipments, between the Ethernet switches and the IP Phones must be configured on 100Mbps (Deactivate auto negotiation). The broadcast rate must be the least as possible, the recommended amount is 500/s at a maximum.

It is strongly suggested to have in place P.O.E switches that integrate the ports to the Ethernet standard of 802.3af in order to power the IP Phones.

3.1.2 Bandwidth

The amount of bandwidth that is required is based on the codec that will be used and on the sample packet that is defined. The rule of thumb is a fixed sample at 20ms (sampling rate)

The G.711 codec holds the following characteristics:

- It samples at a 8000hz (like the RNIS) therefore a sample of 8000 for 1s
- It used 8 binary elements per sample that represent 8000 octets/s or **64kbit/s**

The voice transmission using the standard G.711 codec therefore utilizes 64kbit per seconds. At the network transmission level this solution is not feasible because in real time it is not possible to loose 1s of voice conversation in case of a lost packet. The idea is therefore to send the voice packets more frequently, each packet will be sent at 10 to 40 ms (sampling rate).

To optimize the transmission, the value of the sampling rate takes into consideration the report: bits used / control bits which represents 50 packets per second for a sampling rate of 20 ms(1/0,2)

Each packet will then consist of 8000/50 = 160 octets of voice payload (useable data)

At the internet level, each packet is encapsulated in the RTP/UDP/IP that represents 40 octets of headers that are added to the VOIP charge (Voice Payload).

In G.711 the amount of packets that are handled for 20ms of codec contains: ((160+40)*8) * (50/1000)= 80 KBps

This will also be added to the Ethernet header (Ethernet L2 Overhead)

The size of the packet, including the header is therefore:

G711 20ms : ~ 90Kbps
G 729 20ms : ~ 35 Kbps

Note:

Regardless of the codec used the header represents a very important part of the size of the packet when the voice in send over a WAN. It is strongly recommended to provide routers that have the ability to compress the VOIP header of the packet in order to minimize the bandwidth usage.

3.1.3 Quality of Service - QoS

Quality of service is the ability to process a specific condition under good conditions, a type of traffic in terms of availability, transmission delays, waist and amount of packets lost.

To guarantee a good voice quality on VOIP, the equipments must imperatively manage the QoS notion **from end to end.**

In a LAN environment, the switches and the IP Phones must have the ability to manage QoS level 2. The equipments compatibility with **QoS level 2** is defined as « 802.1p/Q ».

In a WAN environment, the routers must imperatively manage the **QoS level 3**.

The compatibility of the equipments with QoS level 3 are defined as « ToS, DIFFSERV, DSCP ».

The equipments technical notices indicate their compatibility

Reminder:
If a softphone is used the QoS can not be guaranteed, it is therefore strongly recommended to use a hardphone.

3.1.4 Latency

Managing the transmission delays is an essential element in order to benefit from a true conversation quality transmissions and minimizing the **perception of echo.**

The amount of time it takes for the network to process the voice depends on a number of factors:

- The transmission load on each link
- The amount of elements on the network
- The time it takes to process each element
- The information and propagation delay
- The amount of time to translate codec and the voiced packets.

The figures bellow (extracted from the UIT-T G114 recommendations) are provided to indicate the precise class of quality and interactivity in function to delay of transmission in a conversation. These figures consist of the total delay when treating a call.

.

Class	Delay	Comments
1	0 to 150 ms	Acceptable for most communications
2	150 to 300 ms	Acceptable for communications with minimal interaction
3	300 to 700 ms	Practically a half duplex communication
4	> 700 ms	Unusable without good practice of Half duplex

We generally consider that the superior limit « acceptable » for a telephone conversation is between 150 and 200 ms per transmission. By considering the time it takes to treat the conversation and the deliver delay.

Note:
In the actual version, Host Media Processing introduces a latency time by means of transmission of 120 ms, therefore (see jitter paragraph)

3.1.5 Loss of Packets

When the buffers of the different IP network elements are congested, they automatically release the bandwidth by eliminating certain proportion of packets that come in.

If there is no mechanism in place to recuperate these lost IP packets this will results in ruptures on the conversation level causing crackling in the voice communications.

The loss rate per packet must generally be < 2%

3.1.6 Jitter

Jitter is a variance caused by a delay in transmission, it measures the temporal variation between the moment when two packets arrive and their effective arrival.

This irregularity in the packets behavior is caused by multiple reasons:

- Encapsulated packets in supported protocols
- The network load in any given time

To compensate for jitter, we use jitter buffers to control the irregularity of the packets, this gives the process time to cross the system (See latency graph). Their size must therefore be carefully defined and adapted in a dynamic way to the conditions of the network.

On a LAN, acceptable jitter is 20 ms, even though it is not possible to obtain these results on a WAN, let's not forget that the delay (variable) must be as small as possible in order to guarantee a good voice quality on VOIP. When the jitter becomes more apparent, the conversations are very poor in quality.

3.1.7 Echo Cancellation

Echo cancellation consists of eliminating echo that is present on a voice transmission in order to increase the quality of the call. The echo cancellation is often necessary because the compression techniques used for voice and the way packets are treated generate a resonance.

To be efficient the echo cancellation must always be compared to and as close as possible to a IP/TDM environment, in our architecture it must be managed at the gateway level.

The size of the buffer must be properly defined in accordance with the total time send / receive between the gateway and the client in order to guarantee a good and proper echo cancellation.
A value of 128 ms is a minimum.

Echo cancellation enhances the quality of the communication and also reduces the bandwidth used thanks to the silence suppression technique.

Note:
Echo perception is much more amplified when heterogeneous networks are combined. This is the case for wireless telephony mobile networks connect to the general networks. In essence when an outbound call is sent to a wireless mobile network, the echo phenomenon can be very minimal. (Reference standard G168.de l'UIT-T related to echo cancellation).

3.1.8 Headset

In VOIP environments no elements must be neglected, a good quality VOIP headset is necessary for optimal clarity, comfort and quality.

End of Document