
# Surf Host Media Processor Specification Guide

Version: 1.09
Date: August 2015
By: *Oleg Fomenko*

Copyright © 2005-2015, SURF Communication Solutions Ltd.

This document contains confidential and proprietary information of SURF Communication Solutions Ltd., henceforth referred to as “SURF.” All rights reserved. No part of this documentation may be reproduced in any form or by any means or used to make any derivative work (such as translation, transformation, or adaptation) without written permission from SURF Communication Solutions.

SURF reserves the right to revise this document, and to make changes therein, from time to time without providing notification of such revision or change.

This document contains descriptive information regarding the subject matter herein, and is not an offer to purchase or license any products or services of, or from SURF. SURF expressly disclaims any and all representations or warranties, expressed or implied herein, including but not limited to warranties of merchantability or fitness for a particular purpose. The licensing or sale of any product or service by or of SURF shall only be made in accordance with, and subject to the terms of, an agreement for the relevant product or service, to be signed by both the customer and SURF or its authorized agent or representative. Consequently, SURF shall carry no liability to any such customer based on this document, the information contained herein, or the omission of any other information.

Trademarks

The name “SURF Communication Solutions” is a registered trademark of SURF Communication Solutions Ltd.

Any other trademarks, trade names, service marks or service names owned or registered by any other company and used herein are the property of their respective owners.

SURF Communication Solutions, Ltd.

5 Hamada St., Yoqneam Industrial Park

YOQNEAM – 20667, Israel

Tel: +972 (0)73 714-0700

Fax: +972 (0)4 959-4055

| | |
|---|---|
| Web site | http://www.surfsolutions.com/ |
| Email | info@surfsolutions.com |

**Document Origins**

| | Name | Position | Date |
|---|---|---|---|
| Written By | Oleg Fomenko | AVP Signal Processing | 11/05/2014 |
| Inspected By | Avi Fisher | CEO | 11/05/2014 |
| Approved By | | | |

# Table of Contents

# 1. About this guide

## 1.1 Abstract

This document details the specification of the technical infrastructure, services and sizing capabilities supported by the Surf Host Media Processor package (referred to as SURF HMP). SURF HMP is designed to simultaneously support all types of media processing requirements including Voice and Video.

This specification document provides high level information regarding the supported features available and in Roadmap. This document does not provide information regarding the methods which should be applied in order to use specific capabilities. For this purpose the user is encouraged to read reference documents [1], [2] and [3].

**Please refer to the HMP package 1.0.9 Release Notes for availability of asterisk (*) marked items in this guide**.

## 1.2 Reference Documents

[1] MediaProcessor_API_spec.pdf: the detailed specification of the SURF HMP package API.

[2] MediaProcessor_getting_started.pdf: the getting started guide of the above package.

[3] MediaProcessor package109 Release Notes.pdf: the release notes guide of above package.

## 1.3 System requirements

### 1.3.1 Voice and Audio functionality requirements

Host requirements: Any Intel 64bit platform.

OS requirements: Any 64 bit Linux distribution, with Kernel 2.6.32 and up.

It should be noted that the amount of concurrent transcoding/processed streams/channels/ conferences and participants per conference is highly dependent on the following elements:

- The number of Intel processing cores in the server under use
- The Linux distribution under use
- The Kernel under use
- The Network Interface Controller (NIC) under use (1Gb/s and up)
- Whether or not the software package was compiled and tested on the same Linux distribution based platform
- Enabling of Hyper-threading

### 1.3.2 Video functionality requirements

Host requirements: Any INTEL 4th generation (Haswell) or 5th generation (Broadwell) core i3/i5/i7 CPU with GPU (Graphic Processing Unit) acceleration - and above.
[Support for specific features in this spec guide depend of course on the generation of the processor used]

OS requirements: Linux, CentOS 7.1 (CentOS 7-1503) build, 64 bit - and above.

# 2. Detailed specification

## 2.1 Audio/Voice Processor

### 2.1.1 Supported codecs

- G.711 A/U law
- G.722
- G.723.1
- G.726
- G.729AB
- Linear 16 bit
- NB-AMR (all standard rates) – Bandwidth efficient and octet aligned
- WB-AMR (all standard rates) – Bandwidth efficient and octet aligned
- OPUS – Mono/ Stereo, 8-48kHz sampling rates
- Roadmap Codecs: EVS* (Project Dependent, Check for Availability)

### 2.1.2 Audio/Voice processing

- Rate conversion from any to any rate
- Conversion mono->stereo/stereo->mono

### 2.1.3 Supported voice stream processing

- RTP input → RTP output using a single voice tool.
- The input and output RTP can contain any arbitrary codec from the above specified list. The tool performs the appropriate processing and transcoding.
- Full duplex requires the configuration of 2 voice tools.

### 2.1.4 Tonal event handling

- Supported tones: DTMF's, user defines tones and tone sequences
- RFC2833/4733 detection and reporting to application via API, with optional suppression.
- In-band tonal even detection and reporting to application via API, with optional suppression.
- Generation of in-band tone events to channel based on API commands

- Generation of RFC2833/4733 tone events to channel based on API commands
- Tonal event relay from in-band in the input to RTC2833 in the output
- Tonal event relay from RFC2833/4733 in the input to in-band in the output

### 2.1.5 RTP frame duration

- According to specific codec restrictions, from 5ms and up to 40ms

### 2.1.6 IP/UDP

- IPv4, IPv6*
- UDP port setting on a per-stream basis

### 2.1.7 VAD/CNG/PLC

- G.729AB, NB and WB-AMR have a built-in supported mechanism for VAD/CNG/PLC.
- For G.711:
    - VAD/CNG: compliant with G.711 Annex II
    - PLC: compliant with G.711 Annex I

### 2.1.8 RTP/RTCP encapsulation

- RTP/RTCP according to RFC 3550, 3551, 3389
- For NB-AMR and WB-AMR: RFC 4867

### 2.1.9 SRTP Encryption

- SRTP Support for the media layer; TLS cryptographic support for the signaling layer
- AES Encryption Functions support (Counter mode, F8)
- Advanced Encryption Standard (AES) support – up to 256bit key
- SHA-1 Authentication, up-to 160bit

### 2.1.10 Jitter buffer

- Up to 300ms history per stream
- Multiple modes of operation, configured on a per-stream basis:
    - Adaptive
    - Fixed
    - Short-run-adaptation: Fixed with recovery mechanism in case of clock skew/drift
- Configurable initial jitter buffer delay

### 2.1.11 Automatic Gain Control (AGC)

- Configurable minimum/maximum range
- Configurable gain step level
- Configurable energy averaging window
- Configurable maximum applied gain
- AGC can be configured separately on a per-stream basis

### 2.1.12 Output stream duplication

- Each output IP/UDP/RTP stream can be sent to up to 16 different destinations

### 2.1.13 File Playing

- File is read from a file system and streamed into a voice tool
- Any local/remote file system supported by Linux can be used, for example: ext4, tmpfs, NFS, etc.
- Play lists are supported – several files can be configured to be streamed one after another
- The following container types are supported: WAV, AVI, 3GP and MP4
- Codecs supported for file playing: G.711, Wide Band AMR and Narrow Band AMR

### 2.1.14 Voice Conferencing

- Voice conferencing supports unlimited number of conference participant (limited only by machine resources).
- Narrow band/Wide band conference (8000/16000 voice sampling rate)
- Different types of participants
    - Regular – this participant can talk, the decision if the participant is heard by the others or not is taken according to internal algorithm and participant speaking energy.
    - Always dominant – these participants are always heard by the others.
    - Listener only – this participant can hear the conference, but does not have the right to talk.
    - Audio Codecs (such as OPUS) can enter a conference via an Audio Tool
- Configurable "hangover" period, during this period a participant configured as dominant which does not speak still remains dominant.
- Whisper functionality; one participant can "whisper" to another participant so that all others will not hear that.

## 2.2 Video Processor

Video capabilities include video decoding, processing, encoding and RTP streaming

### 2.2.1 Supported codecs

- H.264 Baseline/Main and High profiles – Up to 4K resolution (3840x2160)
- VP8 – Up to 720p resolution (1280x720)
- Roadmap Codecs: H.265, VP9*

### 2.2.2 Video Processing

- Video resizing
- Frame rate modification
- Mixing (creating composite layout from several inputs)
- Text overlay*
- Static image insertion*

### 2.2.3 RTP Streaming

- RTP and RTCP support
- RFC 6184 support (RTP payload format for H.264 video)
- Draft-ietf-payload-vp8-01 (RTP payload format for VP8 video)

### 2.2.4 SRTP Encryption

- SRTP Support for the media layer; TLS cryptographic support for the signaling layer
- AES Encryption Functions support (Counter mode, F8)
- Advanced Encryption Standard (AES) support – up to 256bit key

### 2.2.5 Video Transcoding

- Low latency transcoding ( < 1 frame duration)
- High to low resolutions supported (SQCIF up to Ultra-HD 4K )
- High frame rates supported (limited only by processor's resources)
- Ultra High density transcoding (see table in section 3.4)

### 2.2.6 File Playing*

- File is read from a file system and streamed into a video decoder tool
- Any local/remote file system supported by Linux can be used, for example: ext4, tmpfs, NFS, etc.
- Play lists are supported – several files can be configured to be streamed in sequence
- Container types supported: mp4, 3gp, avi, h264, ivf, mkv and more

### 2.2.7 File Recording*

- Created video stream can be saved into a file
- Any local/remote file system supported by Linux can be used, for example: ext4, tmpfs, NFS, etc.
- Container types supported: mp4, 3gp, avi, h264, ivf, mkv and more

### 2.2.8 Video Mixing and Conferencing*

- Various user defined layouts
- Dynamic layout change
- High to low resolutions supported (SQCIF up to Ultra-HD 4K )
- High frame rates supported (limited only by processor's resources)
- Ultra high density concurrent conferencing (see table in section 3.5)
- Ultra high amount of concurrent participants (see table in section 3.6)

# 3. Performance data

The following section provides detailed performance data for the SURF HMP product. It should be noted that this performance is highly dependent on the following parameters:

1. Input/output codecs in use
2. Media processing application (such as IP←→IP transcoding, file→transcoding→IP, and more)
3. The processor type used - including existence of SSE and/or AVX and other SIMD extended instruction sets
4. Number of processors in the system
5. Existence of Hyper-threading support
6. Linux Kernel version and configuration
7. Linux operating system configuration

## 3.1 Voice Packet to Packet trans-coding

This scenario covers trans-coding from packet to packet: stream is received via RTP, decoded, encoded and sent to the destination via RTP

**I7-3770 3.4Ghz performance:**

| Input/ Output codecs | Application | Processor Type | Number of processors /total cores | OS Linux distribution | Number of concurrent processed streams | Average CPU utilization |
|---|---|---|---|---|---|---|
| G.711⇔G.729 | Full duplex transcoding | I7-3770 3.4Ghz, with Hyper-threading, SSE4.2, AVX | 1/4 | MINT | 1500 full duplex (1500 A→B + 1500 B→A) | 89% |
| G.711⇔NB-AMR (rate 12.2) | Full duplex transcoding | I7-3770 3.4Ghz, with Hyper-threading, SSE4.2, AVX | 1/4 | MINT | 1300 full duplex | 96% |
| G.729⇔NB-AMR | Full duplex transcoding | I7-3770 3.4Ghz, with Hyper-threading, SSE4.2, AVX | 1/4 | MINT | 700 full duplex | 88% |
| G.711⇔WB-AMR | Full duplex transcoding | I7-3770 3.4Ghz, with Hyper-threading, SSE4.2, AVX | 1/4 | MINT | 360 full duplex | 92% |
| NB-AMR⇔WB-AMR | Full duplex transcoding | I7-3770 3.4Ghz, with Hyper-threading, SSE4.2, AVX | 1/4 | MINT | 310 full duplex | 92% |

**Xeon E5-2450 v2 @ 2.5Ghz performance:**

| Input/output codecs | Application | Processor type | Number of processors /total cores | OS Linux distribution | Number of concurrent processed streams | Average CPU utilization |
|---|---|---|---|---|---|---|
| G.711⇔G.729 | Full duplex transcoding | Xeon E5-2450 v2 @ 2.5Ghz | 2/16 | MINT | 3200 full duplex (3200 A→B + 3200 B→A) | 90% |
| G.711⇔NB-AMR (rate 12.2) | Full duplex transcoding | Xeon E5-2450 v2 @ 2.5Ghz | 2/16 | MINT | 3000 full duplex | 90% |
| G.729⇔WB-AMR | Full duplex transcoding | Xeon E5-2450 v2 @ 2.5Ghz | 2/16 | MINT | 1300 full duplex | 96% |
| G.711⇔WB-AMR | Full duplex transcoding | Xeon E5-2450 v2 @ 2.5Ghz | 2/16 | MINT | 1600 full duplex | 93% |
| G.729⇔NB-AMR | Full duplex transcoding | Xeon E5-2450 v2 @ 2.5Ghz | 2/16 | MINT | 2200 full duplex | 93% |
| NB-AMR⇔WB-AMR | Full duplex transcoding | Xeon E5-2450 v2 @ 2.5Ghz | 2/16 | MINT | 1300 full duplex | 96% |

## 3.2 Voice File playing performance measurement

In this scenario a file is read using "file_reader" tool and passes to "voice_p2p" tool. After that the voice tool trans-codes the received stream and send it to the destination via RTP.

The density measurements of file reading functionality heavily depends on the type of the file system used and the hardware that stands behind I/O operations.

For example, with SSD disk much higher density can be reached than with mechanical hard drive. The highest density requires usage of some type of "ram disk": tmpfs or ramfs. Using tmpfs is recommended since this file system is newer and in general more effective than ramfs.

SURF HMP is designed to reach maximal utilization of the machine resources and therefore file I/O almost does not affect number of tools that the CPU is able to handle. However, file I/O can be a bottle neck in some cases.

In the following table there are density measurements using different file systems and file I/O hardware. These numbers are maximal capacity that could be reached without harming the voice quality.

| Input/output codecs | Application | Processor type | Number of processors /total threads | OS Linux distribution | Number of concurrent processed streams | Average CPU utilization | File system type/ HW used |
|---|---|---|---|---|---|---|---|
| G.711⇔G.711 | File reading and transcoding (half duplex) | Intel core i7-2600 CPU@3.4 Ghz | 1/8 | MINT | 2300 | 31% | Ext4/ mechanical hard drive |
| G.711⇔G.711 | File reading and transcoding (half duplex) | Intel core i7-2600 CPU@3.4 Ghz | 1/8 | MINT | 5000 | 79% | Ext4/SSD drive |
| G.711⇔G.711 | File reading and transcoding (half duplex) | Intel core i7-2600 CPU@3.4 Ghz | 1/8 | MINT | 6200 | 96% | tmpfs/ DDR |

## 3.3 Voice Opus ⇔ linear trans-coding

This testing scenario includes full duplex packet to packet trans-coding. Each channel contains both Opus to linear and linear to Opus trans-coders. All the streams are received from RTP and sent using RTP to the network. Sampling rate on the linear side is always 8000Hz. Opus streams have the following parameters: Mono, Complexity=0, Frame Duration = 20ms.

Testing was done using the following CPU:

**Intel i7-2600 CPU @ 3.40GHz  4 cores hyper-threading**

| Opus sampling rate | Opus bit-rate (for both directions) | Number of concurrent processed streams |
|---|---|---|
| 8000Hz | 12Kbps | 485 full duplex |
| 16000Hz | 20Kbps | 500 full duplex |
| 24000Hz | 26Kbps | 460 full duplex |
| 48000Hz | 60Kbps | 460 full duplex |

## 3.4 Video trans-coding performance

These measurements were done on an Intel i7-5650U CPU with encoder settings optimized for maximal speed, using the H.264 codec.

| Video Resolution | Application | Processor Type | Number of processors /total cores | OS Linux distribution | Number of concurrent channels | Average CPU utilization |
|---|---|---|---|---|---|---|
| Ultra-HD, 4K, 60 frames/sec | H.264 Video transcoding | I7-5650U, 2.2Ghz, | 2/4 | CENT OS | 1 | 10% typ. |
| Full HD, 1080p, 60 frames/sec | H.264 Video transcoding | I7-5650U, 2.2Ghz, | 2/4 | CENT OS | 6 | 10% typ. |
| Full HD, 1080p, 30 frames/sec | H.264 Video transcoding | I7-5650U, 2.2Ghz, | 2/4 | CENT OS | 11 | 10% typ. |
| HD, 720p, 30 frames/sec | H.264 Video transcoding | I7-5650U, 2.2Ghz, | 2/4 | CENT OS | 21 | 10% typ. |
| Wide-VGA, 30 frames/sec | H.264 Video transcoding | I7-5650U, 2.2Ghz, | 2/4 | CENT OS | 61 | 10% typ. |

## 3.5 Video Max Concurrent Conferences

These measurements were done on an Intel i7-5557U CPU with encoder settings optimized for maximal speed, using the H.264 codec.

| **Single INTEL i7 5557U CPU** | | | | |
|---|---|---|---|---|
| **Encoders/ Decoders Mix** | **Max Concurrent Conferences** | | | **Non-Viewable Participants** |
| | **4 Part/conf** | **9 Part/conf** | **16 Part/conf** | |
| **Enc**: 4K/60fps<br>**Dec**: 1080p/60 | **1** | **0** | **0** | Unlimited* |
| **Enc**: 1080p/60; 720p/60<br>**Dec**: 720p/60 | **2** | **1** | **0** | |
| **Enc**: 1080p/30; 720p/30<br>**Dec**: 720p/30 | **4** | **2** | **1** | |
| **Enc**: 720p/30; WVGA/30<br>**Dec**: WVGA/30 | **7** | **4** | **2** | |
| **Enc**: WVGA/30 (852x480)<br>**Dec**: WVGA/30 | **13** | **7** | **4** | |

*** Number of non-viewable participants watching the conference and participating in voice is limited only by the network BW availability**

## 3.6 Video Max Participants per Conference

These measurements were done on an Intel i7-5557U CPU with encoder settings optimized for maximal speed, using the H.264 codec.

**MAX. PARTICIPANTS on SINGLE CONFERENCE**

| INTEL i7 5557U CPU | | |
|---|---|---|
| **Single Conference Encoder** | **Max Viewable Participants/Conf.** | **Decoder Resolution** |
| 4K/60 | 4 | 1080p/60 |
| 1080p/60 | 11 | 1080p/60 |
| | 22 | 1080p/30 |
| | 45 | 720p/60 |
| | 91 | 720p/30 |
| | 160 | WVGA30 |
| 1080p/30 | 22 | 1080p/30 |
| | 45 | 720p/60 |
| | 91 | 720p/30 |
| | 160 | WVGA30 |
| 720p/60 | 45 | 720p/60 |
| | 91 | 720p/30 |
| | 160 | WVGA30 |
| 720p/30 | 91 | 720p/30 |
| | 160 | WVGA30 |
| WVGA/30 | 160 | WVGA30 |