The MIDI Wire Protocol Packetization

draft-lazzaro-avt-mwpp-midi-nmp-00.txt

John Lazzaro and John Wawrzynek
CS Division
UC Berkeley
{lazzaro,johnw}@cs.berkeley.edu

Outline

```
Slide 1:  Network Musical Performance
Slide 2:  MWPP
Slide 3:  MWPP temporal controversies
Slide 4:  MWPP resiliency controversies
Slide 5:  MWPP MPEG controversies
Slide 6:  The road ahead
```

Network Musical Performance

Musicians located at different places who interact over a network to perform as if they were in the same room.

Isn't Latency An Issue?

Measured Network Delay (from UCB)

Stanford (40 miles): 2.1 ms net delay → 2.4 ft separation
Caltech (375 miles): 14.2 ms net delay → 16 ft separation

Audio Output Hardware Delay

2-4 ms (Linux, Mac OS X)

Handling Lost and Late Packets

A 50ms receiver buffer won't work!

- Send the physical gestures that musicians make.
- Receiver immediately turns these gestures into audio.
- Semantic recovery from late and lost packets.

The Payoff

Artifacts produced by imperfect networks sound like mistakes made by imperfect musicians, not codec gurgling.

How to Create Gestural Data?

Easy way: musicians use MIDI controllers and softsynths.

The MIDI Wire Protocol

- Instruments have a **MIDI Out** jack (5-pin DIN).
- Serial wire protocol: a stream of MIDI commands
- 1-3 byte commands: 320μs/byte

- Commands encode gestures: `NoteOn, NoteOff, ...`
- No timestamps: execute on receipt.

- Very compact: fast piano playing → 300 bps.
- Very fragile: a lost `NoteOff` → a hung note.

MIDI RTP Packetization

- Handles late packets gracefully
- Handles lost packets w/o retransmission
- Multicast compatible, uses standard RTCP only
- Prototyped in `sfront`, presented at `NOSSDAV`

```
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|X|  CC   |M|     PT      |        Sequence Number        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Timestamp                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             SSRC                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|R|    LEN    |          MIDI Command Payload ...             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Recovery Journal ...                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

MWPP and Time

```
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|X|  CC   |M|     PT      |        Sequence Number        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Timestamp                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             SSRC                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|R|    LEN    |          MIDI Command Payload ...             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Recovery Journal ...                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

`When Local MIDI Command Happens, the Sender ...`

- Synthesizes sound locally.
- Sends RTP packet with new command.
- Timestamp reflects local time.

`When an RTP packet arrives, the Receiver ...`

- Determines if packet is late or on-time.
- Always execute on-time packets.
- Follow semantic rules for late packets.

Open Issues

- Semantic rules are normative: are they optimal?
- Late/ontime algorithm is **not** normative: should it be?
- MIDI System unsupported (following MPEG lead).

MWPP and Resiliency

One Lost NoteOff → Note Sounds Forever

```
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|X|  CC   |M|     PT      |        Sequence Number        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Timestamp                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             SSRC                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|R|    LEN    |          MIDI Command Payload ...             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Recovery Journal ...                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Recovery Journal

- Minimal session history since a **checkpoint packet**.
- For recovery from arbitrary packet loss since checkpoint.
- Updated journal sent with every packet.

What Bounds the Journal Size?

- Not a trace log; only most recent history needed.
- RTCP "last packet received" advances checkpoint.
- 4 kbs payload bandwidth for complex piano work.

Open Issues

- Normative recovery semantics: are they correct?
- No RTCP changes, for simplicity: best tradeoff?
- Strong recovery approach too paranoid?

MWPP and MPEG 4 Structured Audio

MWPP Designed for MP4-SA

- MWPP recovery/time semantics MP4-SA compatible.
- Normatively includes MP4-SA execution semantics.
- **sfront** prototype uses MP4-SA.

Use Not Limited to MP4-SA

- MWPP semantics compatible with all softsynths.
- Most softsynths do not use MP4-SA.

Dual Support

- RFC-generic for MP4-SA softsynths.
- RTP for other softsynths.

Open Issue

- Is "RFC-generic" usage in I-D correct?

The Road Ahead

- Detailed review by the network-savvy.
- Detailed review by the MIDI-savvy.
- Standards track? WG item?

Why MWPP Now?

- NMP is "in the air."
- Most implementors not IETFers.
- Home-grown TCP- or UDP-based.
- Congestion? Multicast? Security?

Early IETF action may help.