



# 1722 AVTP Audio Format

---

Girault Jones  
Ashley Butterworth  
Apple Inc.

# What is AVTP Audio Format

---

- Simplest format for transmitting audio
- Higher channel counts than IEC 61883
- Samples can be integers or floating point
- All packets within a stream are the same size
- Permits any 802.1Qav compliant packet rate

# The Sample Types

---

- Floating Point (IEEE 754-2008)
  - 32-bit float
- Linear Integer
  - 4-byte (1 to 32 bits)
  - 3-byte (1 to 24 bits) “packed”
  - 2-byte (1 to 16 bits)

# Packet Format

---

- Time stamp
  - There is one in every packet
  - Always corresponds to first sample in packet
  - Every packet must have a valid timestamp
- Payload-specific header flags
  - 10 bits indicating channels per frame
  - 2 bits for synchronized event markers
  - 4 bits reserved for future use
- Interleaved or non-interleaved sample data

# Sample Format

---

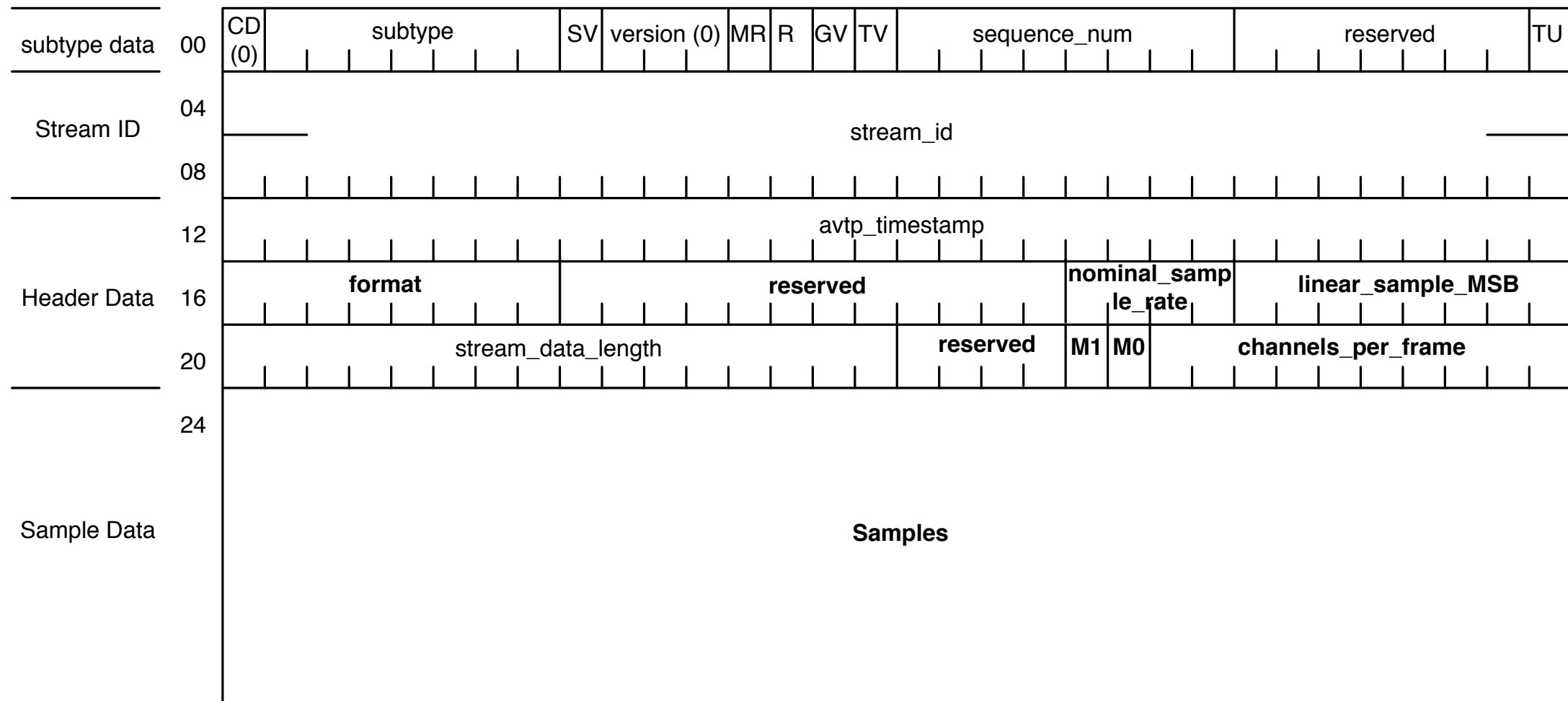
- Network byte order (big-endian)
- Linear integers are LSB zero padded
  - allows easy conversion between various bit widths

# Interoperability

---

- All end stations that support AVTP audio format are required to implement interleaved, 4-byte linear integer format.
- All end stations that support AVTP audio format are required to implement 6 samples per packet 48k sample rate
- Enumeration protocol (such as 1722.1) will fully describe an audio stream.
- However, header contains enough info for an end station to determine the payload format.

# Frame Format



# “Gateway Info”

---

- format
  - Sample format (float/linear integer)
- nominal\_sample\_rate
  - Predefined sample rate or unspecified
- linear\_sample\_MSB
  - Number of MSBs that are used
    - 0 = 256
    - Only applies to integer formats



# Formats

---

- 1 - Interleaved Float
- 2 - Non-interleaved Float
- 3 - Interleaved 32 bit int
- 4 - Non-interleaved 32 bit int
- 5 - Interleaved 24 bit packed int
- 6 - Non-interleaved 24 bit packed int
- 7 - Interleaved 16 bit int
- 8 - Non-interleaved int

# nominal\_sample\_rate

---

- 0 = Not specified
- 1-3 = 8, 16, 32
- 4-8 = 11.025, 22.05, 44.1, 88.2, 176.4
- 9-11 = 48, 96, 192
- 15 = Extended

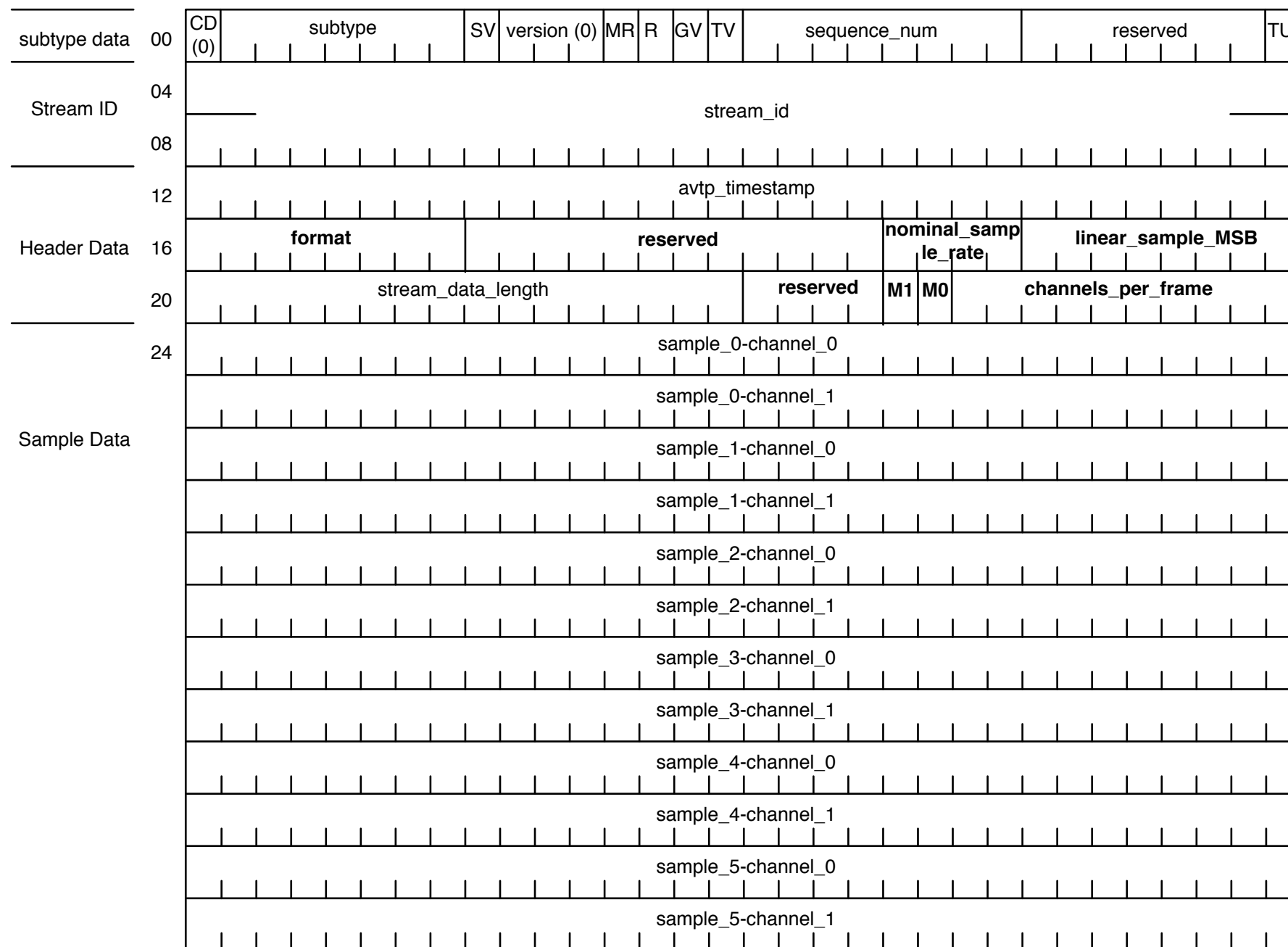
# Protocol Specific Header

---

- M1, M0
  - Marker bits for “synchronized” events
- channels\_per\_frame
  - The number of audio channels represented in the frame

# Frame Format Example

- 4-byte integer, stereo, interleaved, 6 samples/frame



# Encoded Audio

---

- Currently not supported
- But do we want to support it?