LI Vendor Perspective

Upload Version 3

APNIC 50, Sep 2020
Jeff Brower, Signalogic CEO

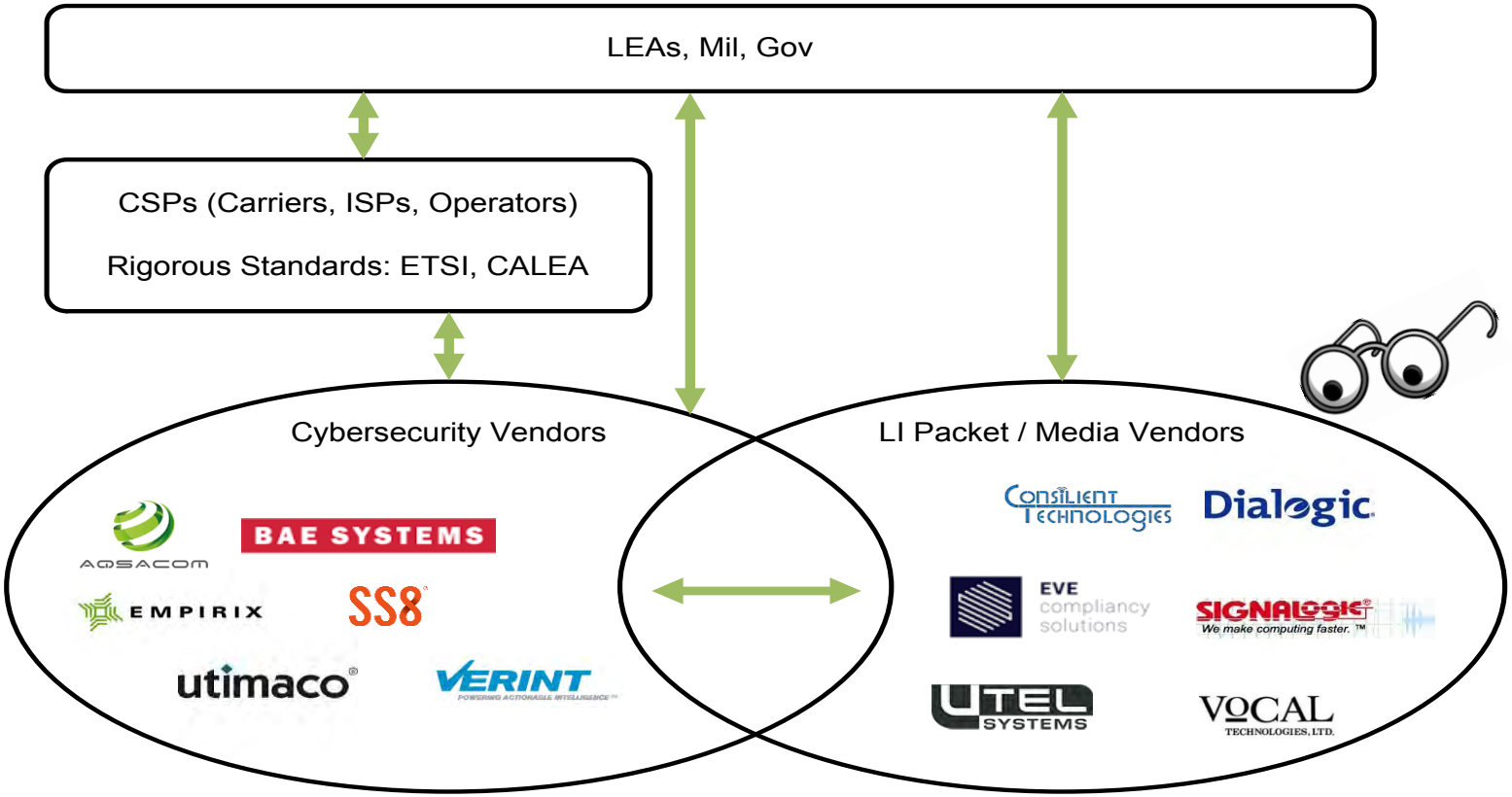
Contents

LI Vendor Perspective

- **LI Vendor Ecosystem**
  - Perspective of today's presentation
- **LEA Expectations**
  - Audio Quality
  - Capacity
  - Reliability
  - User defined signal processing
- **Challenges**
  - Stream alignment
  - Crazy packet send rates
  - Encapsulated media
- **What's Coming**
  - Edge computing (5G cloud)
  - ASR, diarization
  - Containerization, Kubernetes

LI Vendor Ecosystem

- **Approximate view only … lots of overlap**
- **Excluding gateway, SBC, router vendors [1]**



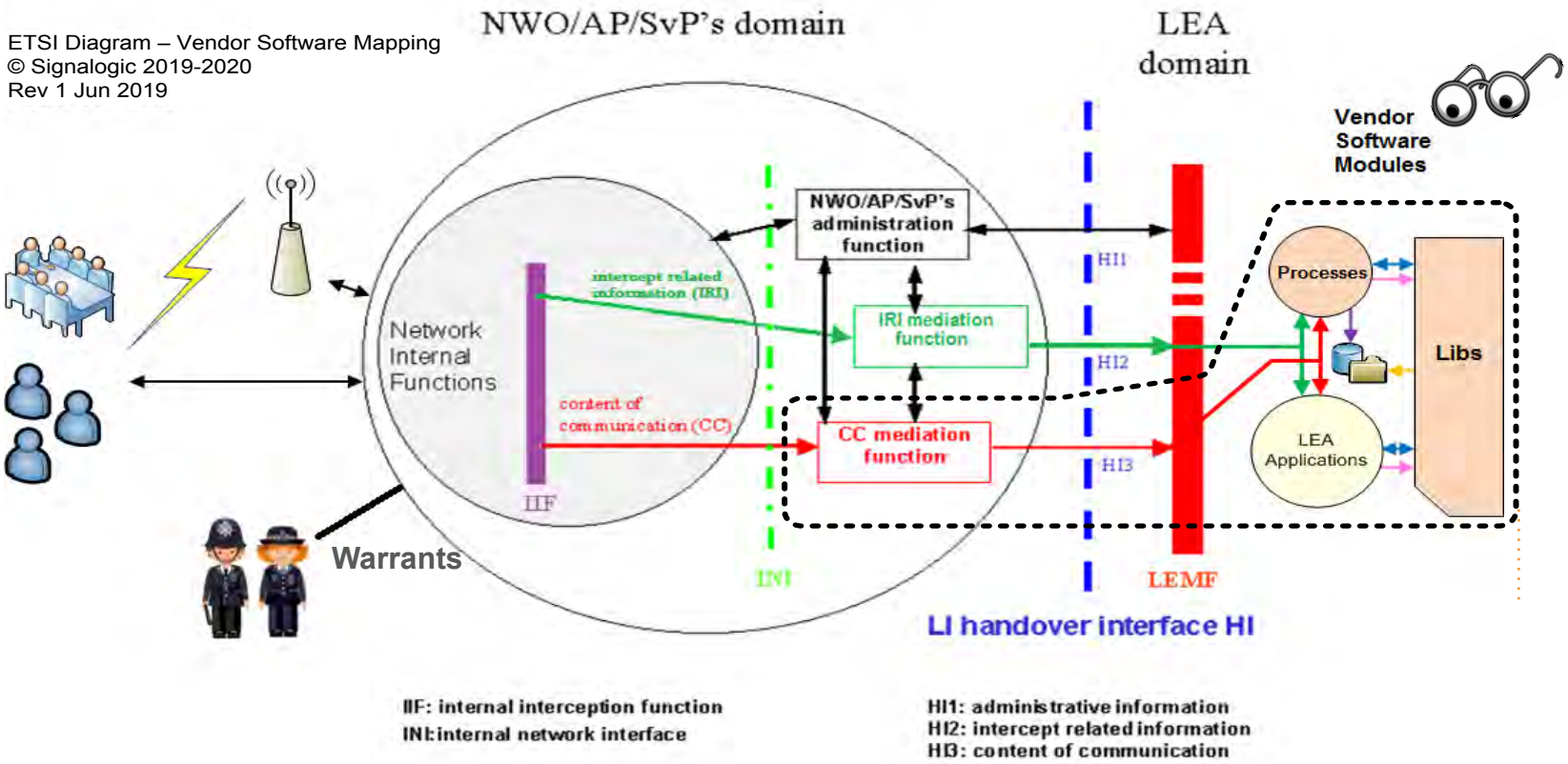
LI Vendor Ecosystem
© Signalogic 2020
Rev 1, Aug 2020

Business Relationships

[1] Cisco, Ericsson, Nokia, Juniper, Ribbon, Mavenir, MetaSwitch, etc.

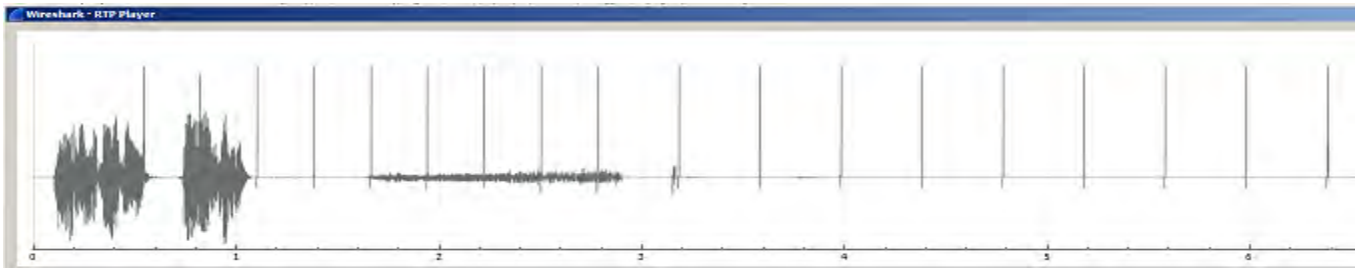
Packet / Media Perspective

- **ETSI LI Terminology: CC mediation (communication content), HI3 (Handover Interface port 3)**
- **Packet Handling**
  - Jitter buffer, packet repair, rate adjustment
- **Media**
  - Decoding (AMR, AMR-WB, EVS, more), stream alignment
- **Signal Processing**
  - Stream merging, conferencing, speech recognition

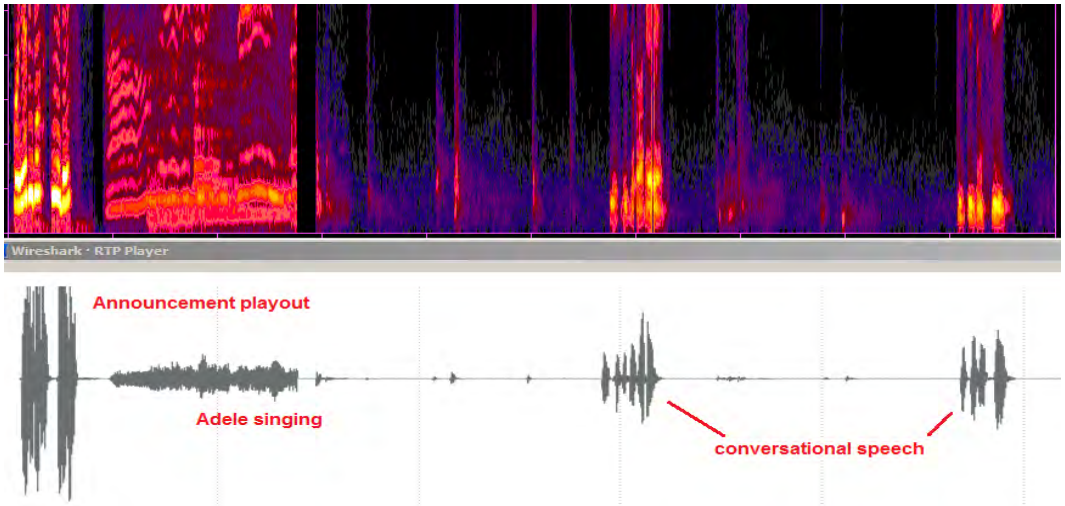


Audio Quality

- **LEAs are obsessive about audio quality**
  - "no sound left behind" – repair lost packets to extent possible, recover gaps due to slow packet send rates, deal with fast send rates by detecting silence
  - some use metronomes, duck quacks, whale sounds, etc to check timing and frequency integrity. Not kidding
- **We use a wide range of techniques to verify LEA test cases**
  - visual audio markers to verify timing, audio frame repair, etc
  - frequency domain analysis



Wireshark screen capture showing audio markers



Frequency domain analysis aligned with Wireshark time domain capture

Capacity

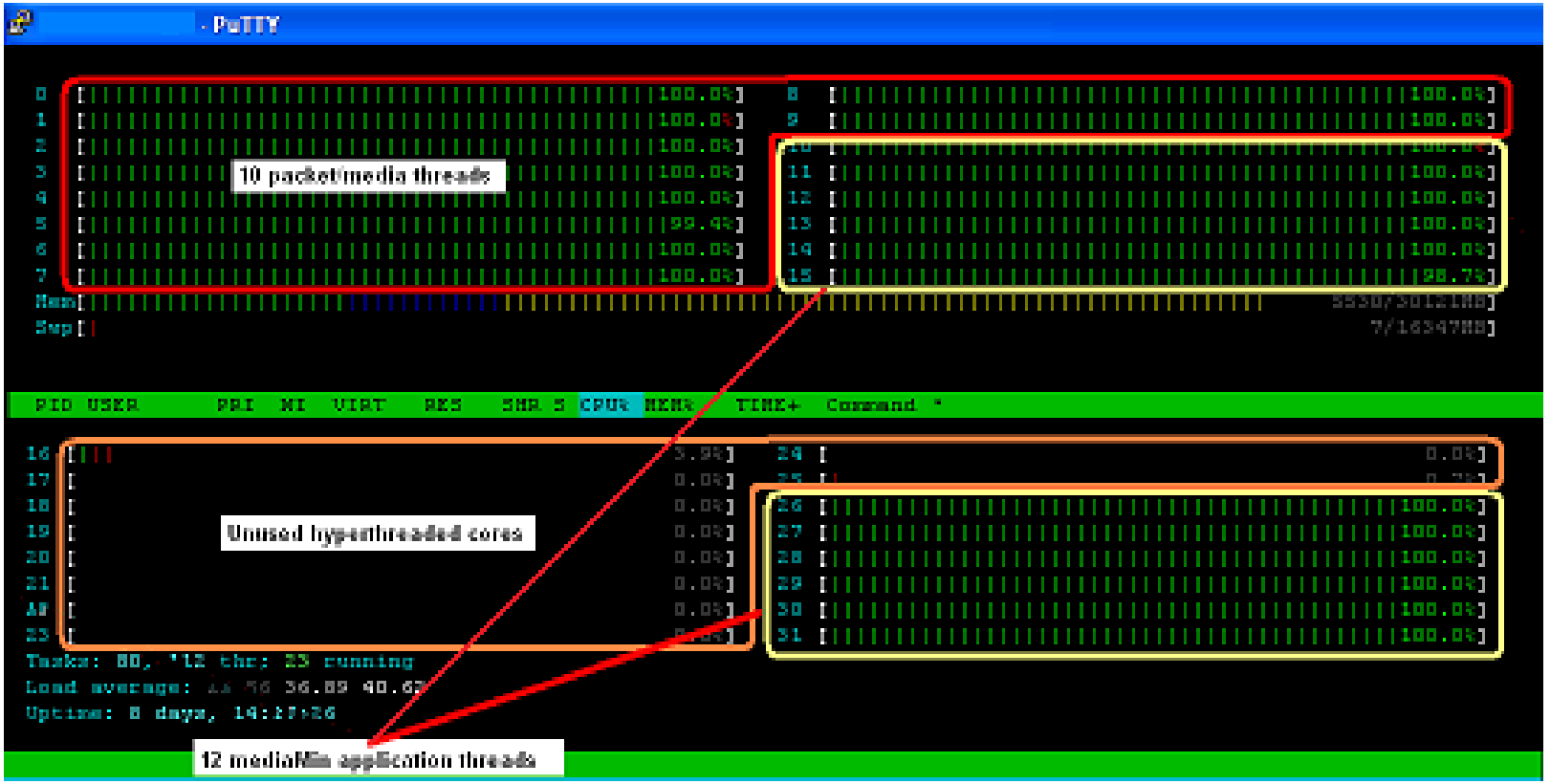
- **LEAs expect extreme per-box / per-VM performance**
  - LI vendor is allowed a specific number of cores, no exceptions
  - the telecom influence is strong – a long history of applications coded for high capacity, real-time performance
- **Linux makes it difficult**
  - not deterministic, not a small footprint RTOS
  - LEAs and carriers know that “software defined solutions” are not deterministic, but no excuses are allowed
  - we have several alarms to detect “thread preemption” – things that Linux housekeeping and other user applications may do to impede performance
- **DPDK[1] can help in some cases**
- **Others**
  - Texas Instruments exited the multicore CPU market in 2016, no longer an option
    - they had an effective solution, 64 cores on an x8 PCIe card, but were unable to embrace servers, open source, and modern development methods
    - they’re now an “analog company” facing existential pressure from US mergers and competition in China
  - GPUs are typically not helpful
    - no help with packet processing
    - only a small subset of media processing can be accelerated
    - each x86 core needs a dedicated GPU card PCIe lane to maintain max performance

[1] Data Plane Development Kit – refers to non-Linux x86 cores dedicated to packet processing

Capacity, cont.

- **We use htop and other tools to scrutinize x86 core usage**
  - hyperthreading must be disabled
  - stream groups must not split across cores

htop screen capture showing
- packet/media threads
- application threads
- disabled hyperthread cores



Reliability

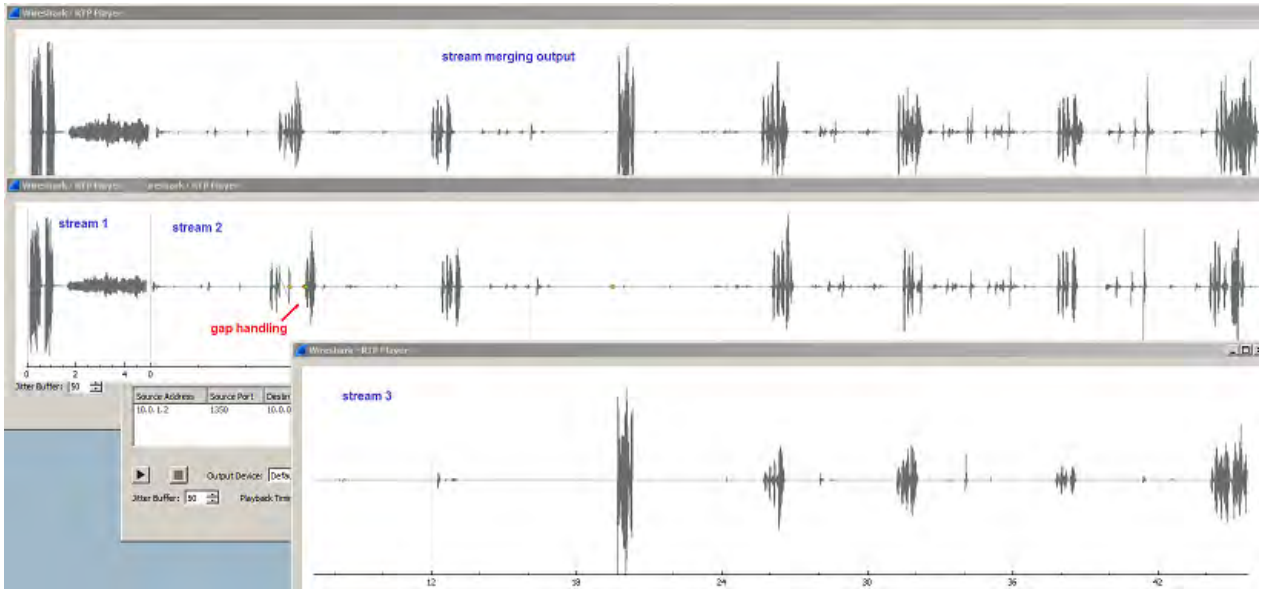
- **LEAs are also obsessive about reliability**
  - extremely long calls are common. Every code and packet data wrap that could occur must be tested
  - as with capacity, the telecom influence pervades. “5 9s” up time is a minimum
- **LEAs run stress tests for weeks at a time**
  - we run stress tests for 2+ months
  - tests include pcaps with artificial wraps, 10x packet push rates, deliberate thread preemptions, more
  - tests run at max per-box / per-VM capacity ratings

Application Specific Signal Processing

- **Common for mil/gov guys to ask for specific signal processing. Some examples:**
  - "deduplication" due to multiple intercepts of the same end point (with different latencies)
  - removing room echo / reverb
  - AGC
  - separating overlapped talkers / conversations
- **Less common for LEAs, but it happens**
- **These typically have a substantial impact on performance**
- **With speech recognition, these needs increase**
  - training is ultra sensitive to small changes in audio characteristics
  - production systems are trained with wide variety of "augmentations", including background noise and babble, loud and quiet speech, frequency warping, etc.
  - "preprocessing" to normalize speech input decreases reliance on augmentation training and increases accuracy

Challenges

- **Stream alignment**
  - when merging intercepted streams, correct time alignment must be maintained between all endpoints
- **Crazy packet send rates**
  - slow, fast, variable. We've seen rates up to 15% slow/fast
  - media playout servers are particularly bad offenders
- **ETSI encapsulated packet format**
  - intercept packet rate may be very different than original audio RTP packet rate



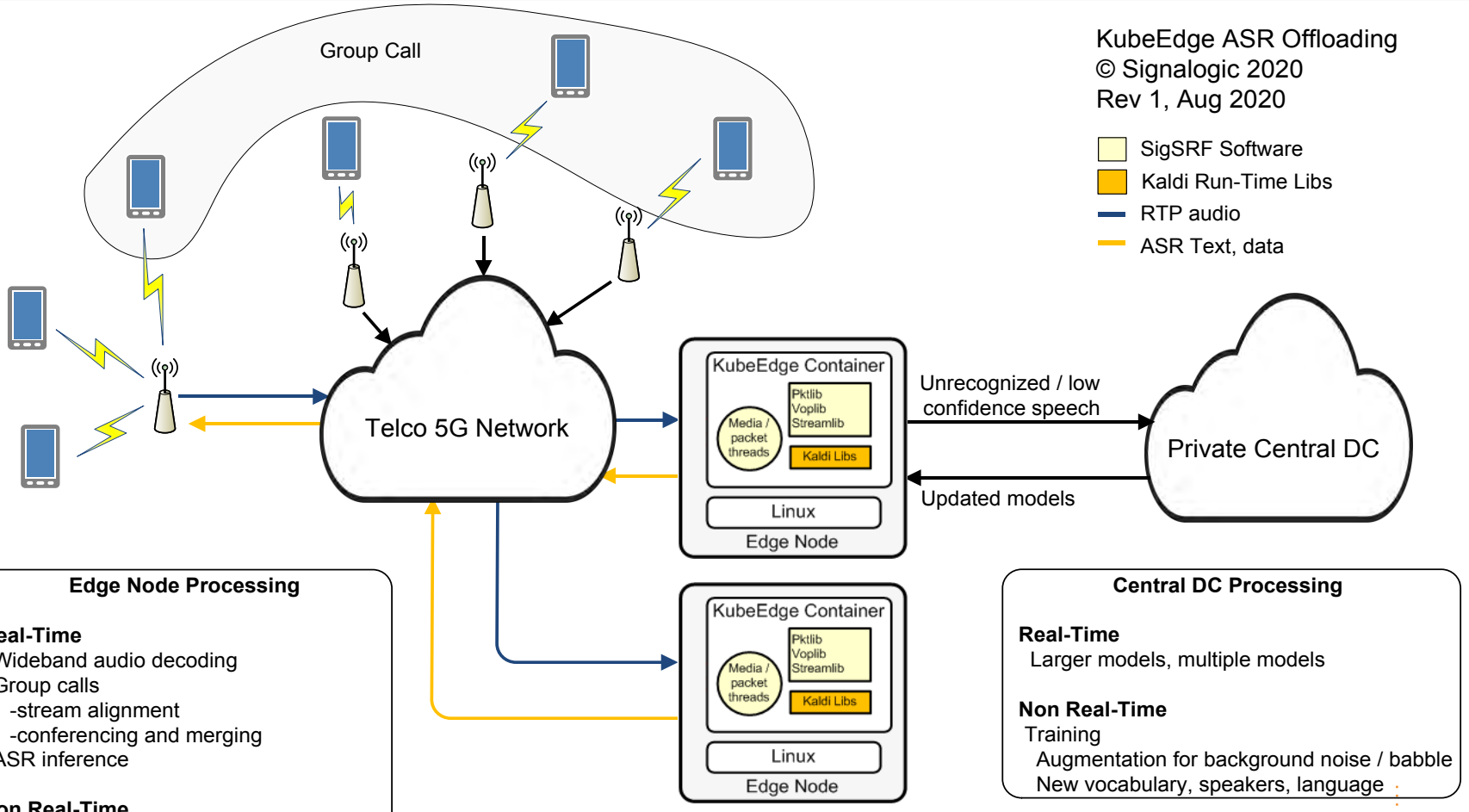
Multiple Wireshark captures showing stream merging of 3 intercepted endpoints

What's Coming

- **Edge Computing**
  - enabled by 5G performance, reduced latency
  - decouple from big tech ("hyperscaler") cloud when needed -- e.g. privacy
- **ASR (Automatic Speech Recognition)**
  - can be done in real-time, but substantially less capacity
  - cannot yet be done in real-time: individual speaker identification and transcription, known as "diarization"
  - potential to greatly reduce LEA workload, accurately alert on "conversations of interest"
  - open source accuracy only a few % WER[2] above proprietary code bases
- **Containerization**
  - easier to scale and deploy
  - allow CICD[1], for example improving ASR accuracy with "on the fly" training based on collected data

[1] Continuous Integration, Continuous Deployment
[2] WER = Word Error Rate

What's Coming: Edge Computing



Edge Node Processing

Real-Time
- Wideband audio decoding
- Group calls
  - -stream alignment
  - -conferencing and merging
- ASR inference

Non Real-Time
- Diarization – not possible in real-time yet

Challenges
- Varying latencies between endpoints
- Background noise
- Overlapping conversations
- Avoid using GPUs when possible

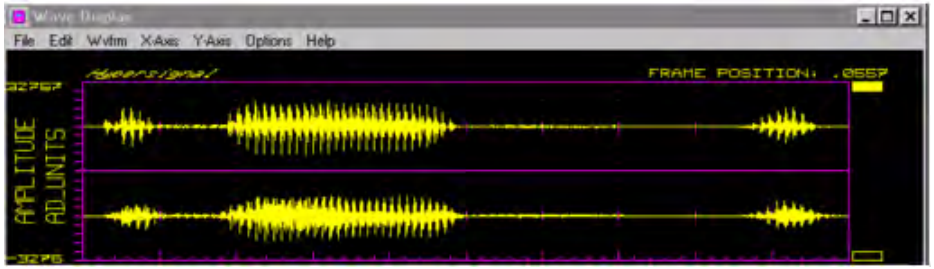
Central DC Processing

Real-Time
- Larger models, multiple models

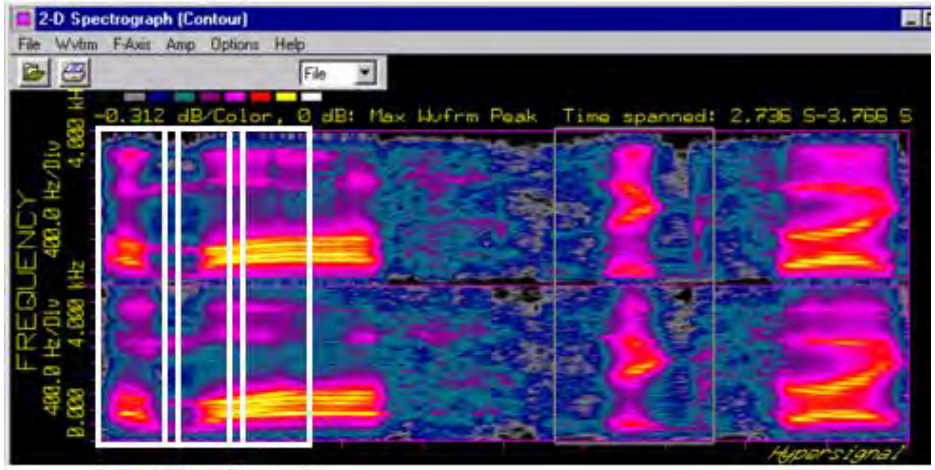
Non Real-Time
- Training
  - Augmentation for background noise / babble
  - New vocabulary, speakers, language

What's Coming: ASR

- **Deep Learning Architecture**
  - combines previous generation xMM[1] technology with DNNs (Deep Neural Networks)
  - relies on extensive training and "augmentation" methods
  - Kaldi open source is basis for Alexa, Google Home, and Cortana
- **Frequency domain "images"**
  - formed by sliding FFT analysis of incoming time series data. Each FFT frame output is similar to cochlea in human ears
  - groups of FFT frames form images
  - successive images are called "TDNN" (time delayed DNN), similar to series of CNNs[2]

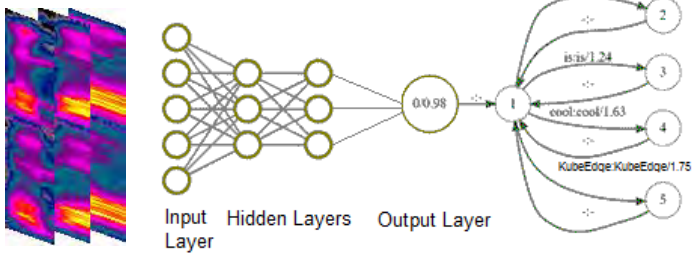


Sliding FFT → **time domain (time series)**



frequency domain

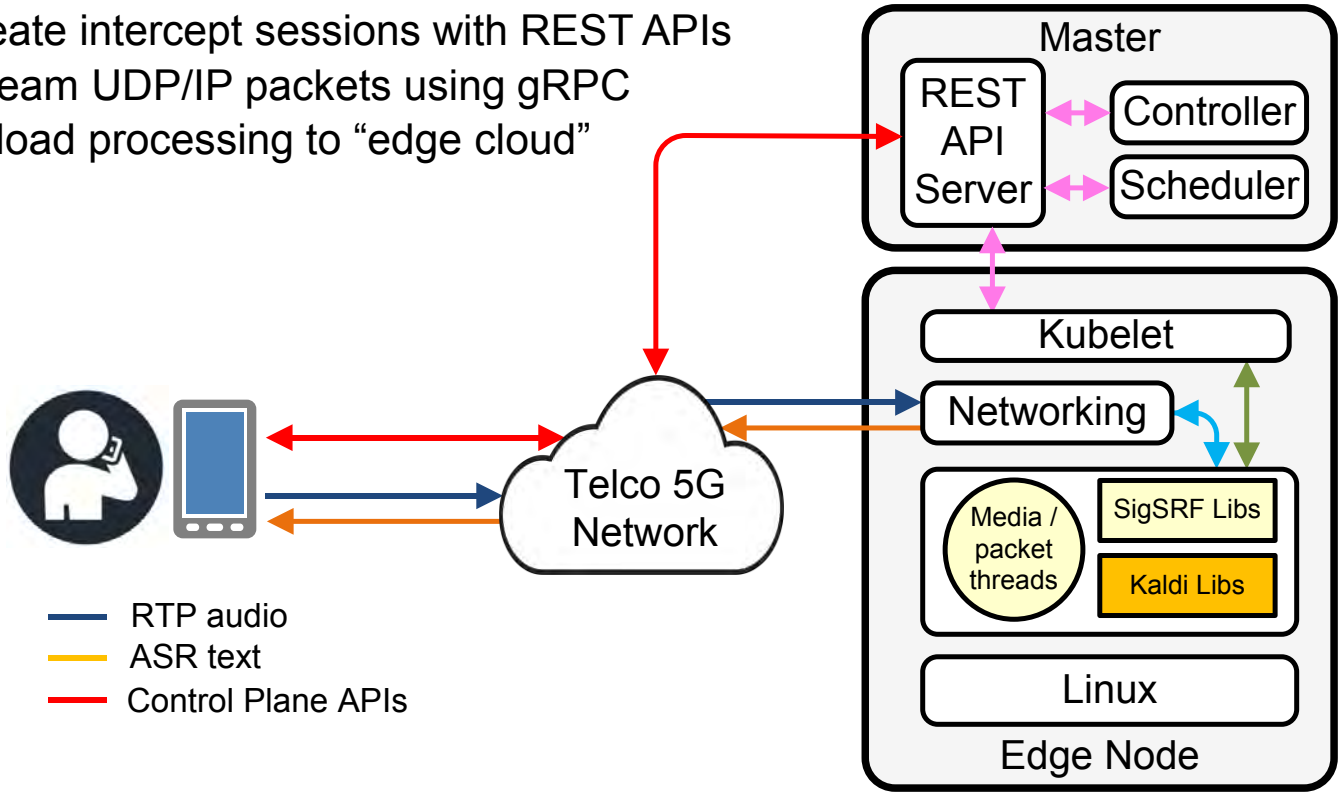
DNN Input Layers (ILn)



[1] Hidden Markov Model, Gaussian Mixed Model, [2] Convolutional Neural Network

What's Coming: Containers and Kubernetes

- **Packet + media + ASR inside container**
  - minimum 2 x86 cores, 32 GB mem, 1 TB HDD can handle 32 sessions
  - a session is wideband decode (e.g. EVS), jitter buffer, stream merging up to 8 stream groups, G711 pcap output, wideband wav file output
  - scales up linearly with more cores
- **Field apps**
  - create intercept sessions with REST APIs
  - stream UDP/IP packets using gRPC
  - offload processing to "edge cloud"



Thanks !

- **Questions or comments, e-mail me at jbrower (at) signalogic (dot) com**
- **For deployment references, possibly I can tell you under NDA, with customer permission**
- **If you need certain pcap test cases, possibly we can help. We have 100s, but we don't publish them**
- **If you have pcaps your system can't handle, or you suspect audio quality could be better, you can try our demo or send and let us try**