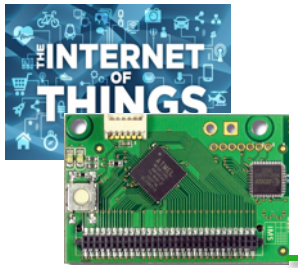
Image and Audio Challenges

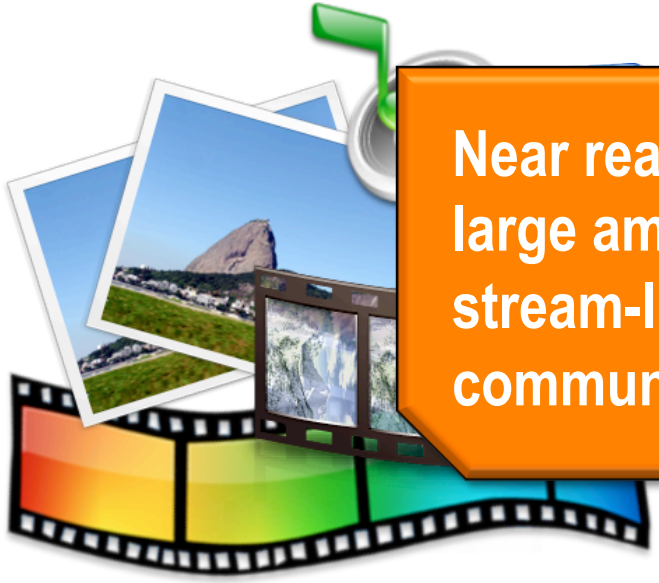
(for surveillance applications)

CARI 2014 tutorial - PART II
Gaston Berger University
October, 17th, 2014
Saint-Louis, Senegal

Prof. Congduc Pham
http://www.univ-pau.fr/~cpham
Université de Pau, France

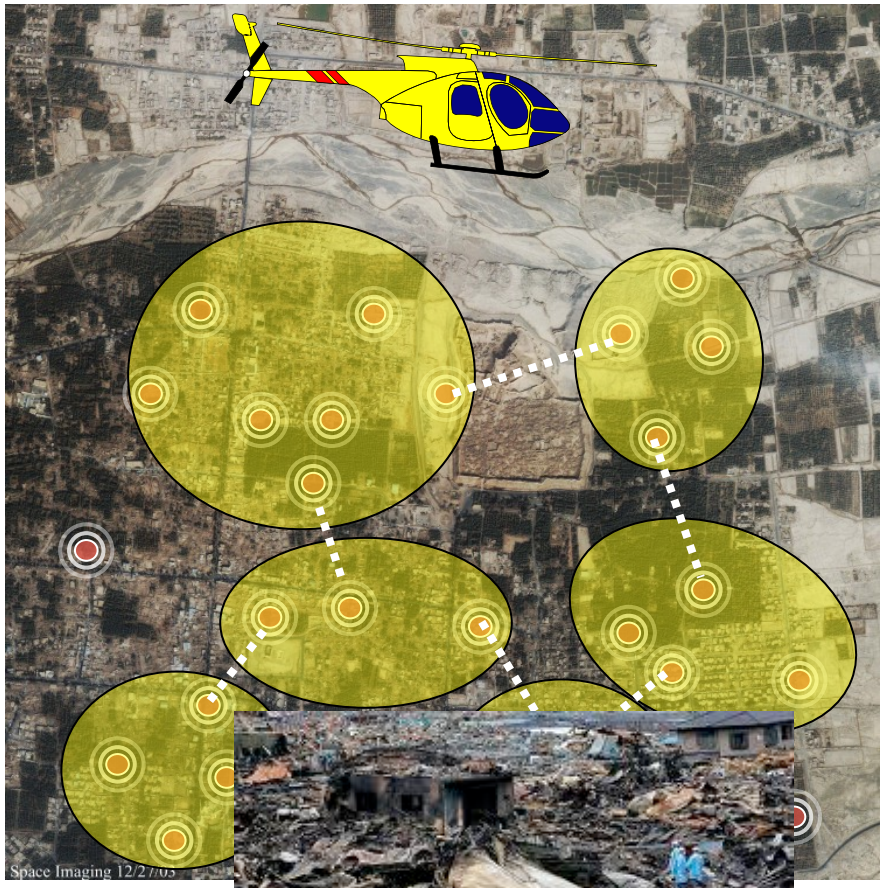


Towards multimedia information

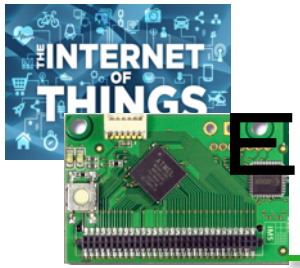


Near real-time constraints, large amount of data, stream-like communication,…

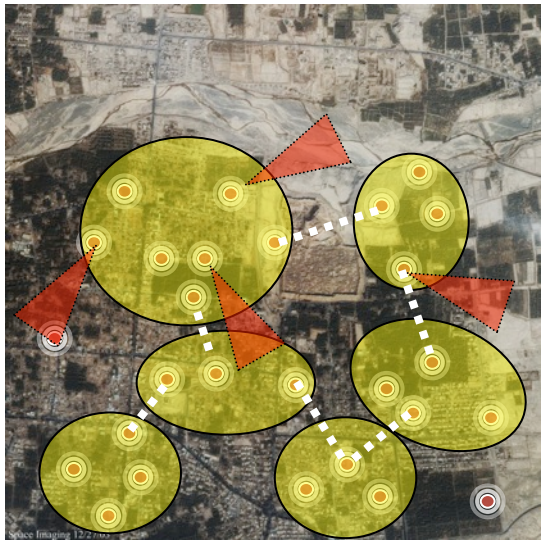
Mission-critical apps

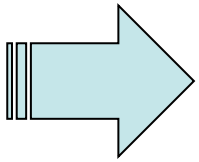


Disaster relief, Search & Rescue, Intrusion detection, …



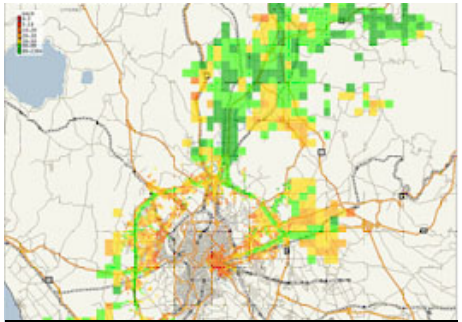
Ex: Situation-awareness





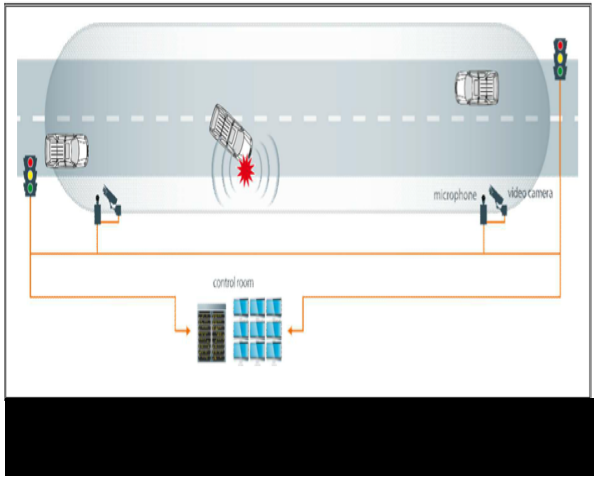
Collect data to improve the responsiveness of rescue operations

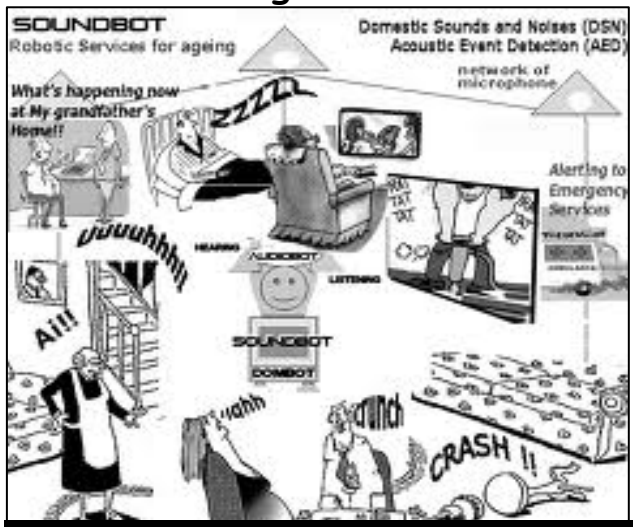
Exploiting Acoustic data



Management

efficiency

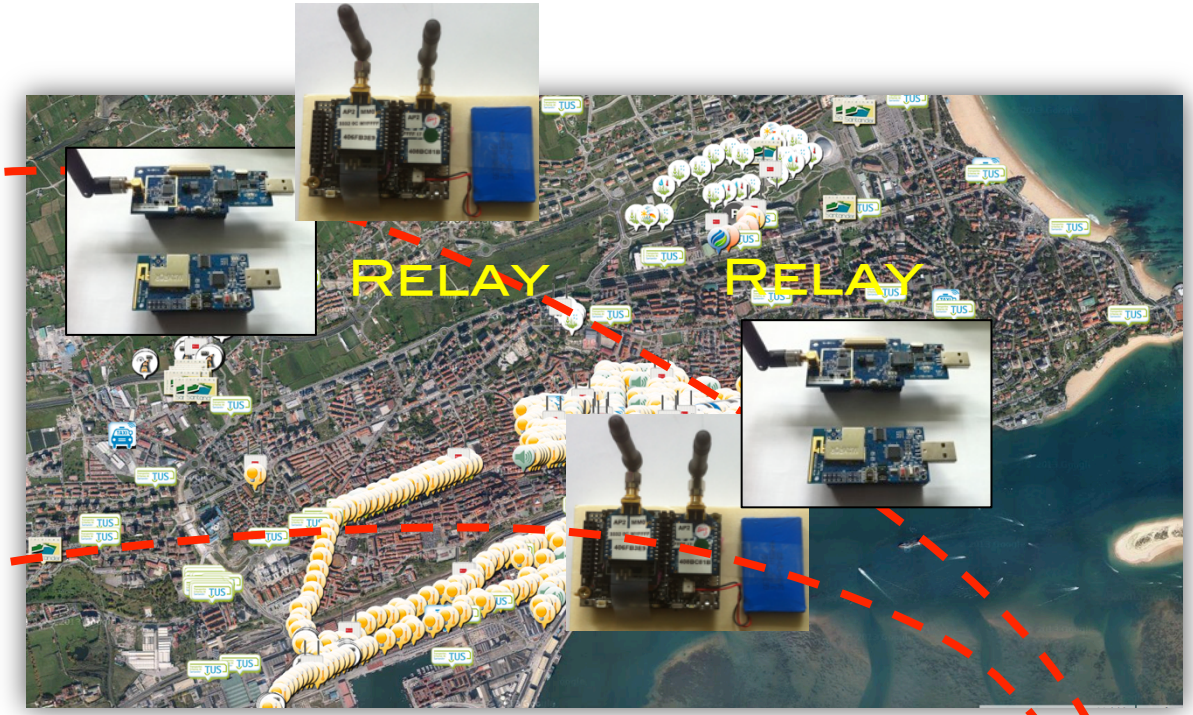




Surveillance

From EAR-IT slides

Use deployed low-resource IoT node to enhance acoustic services



Play/store received Audio data

SmartSantander test-bed

Santander's sensor network deployment

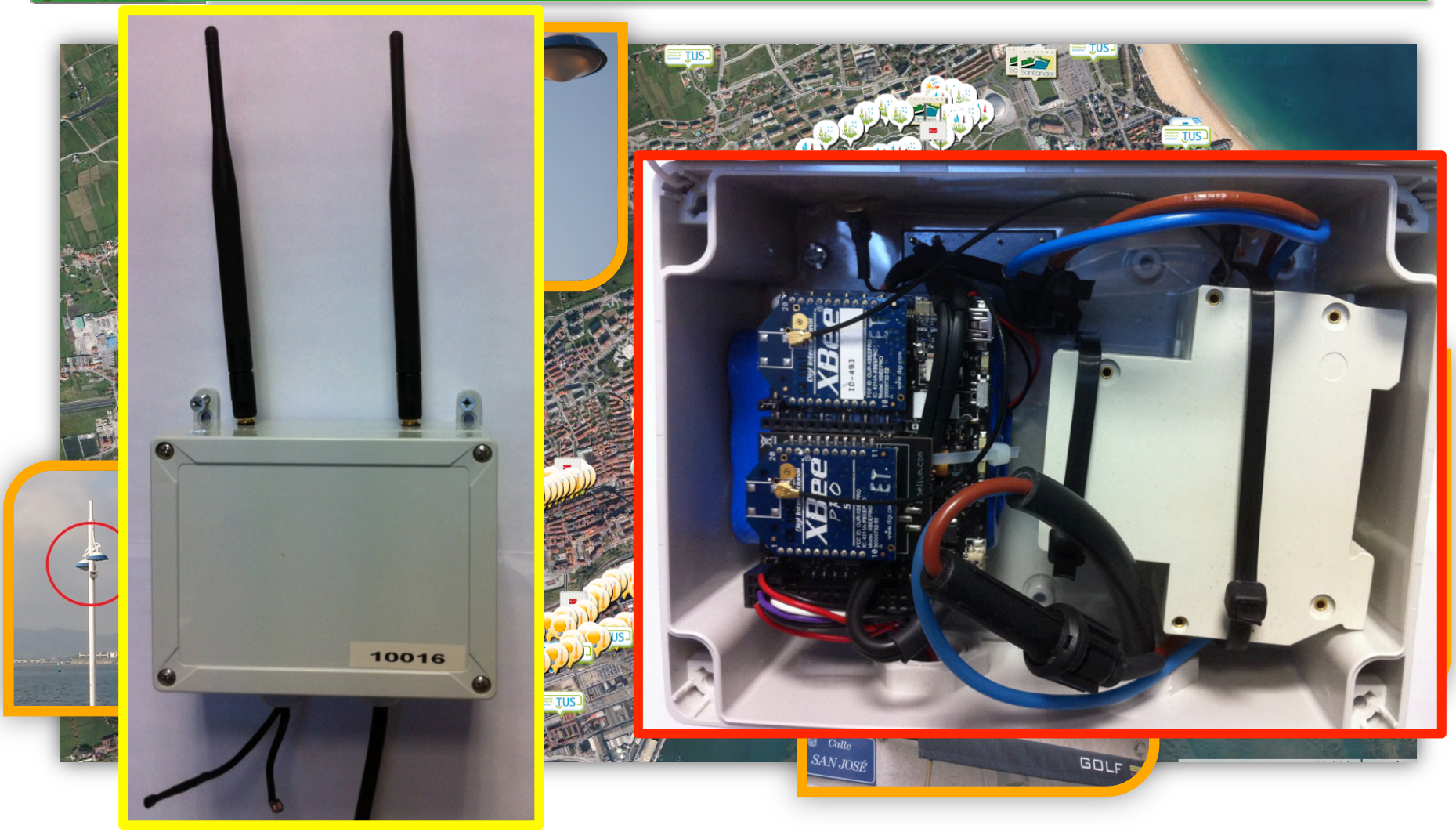
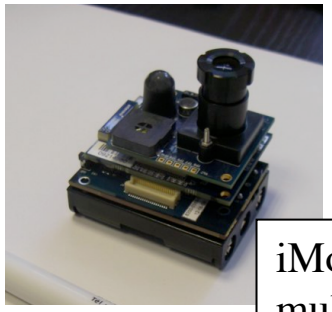
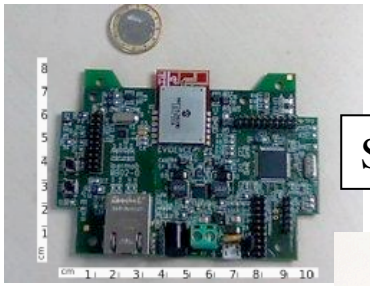


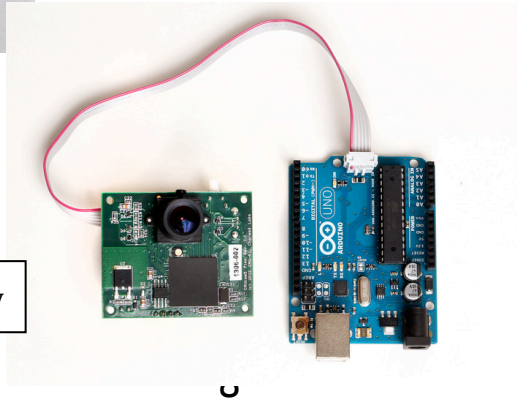
Image sensors



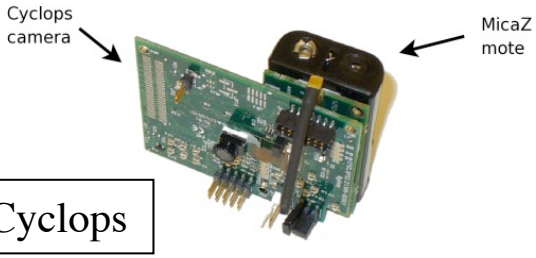
iMote2 with IMB400 multimedia board



Seedeye



Pixy



Cyclops

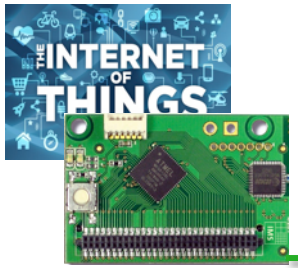
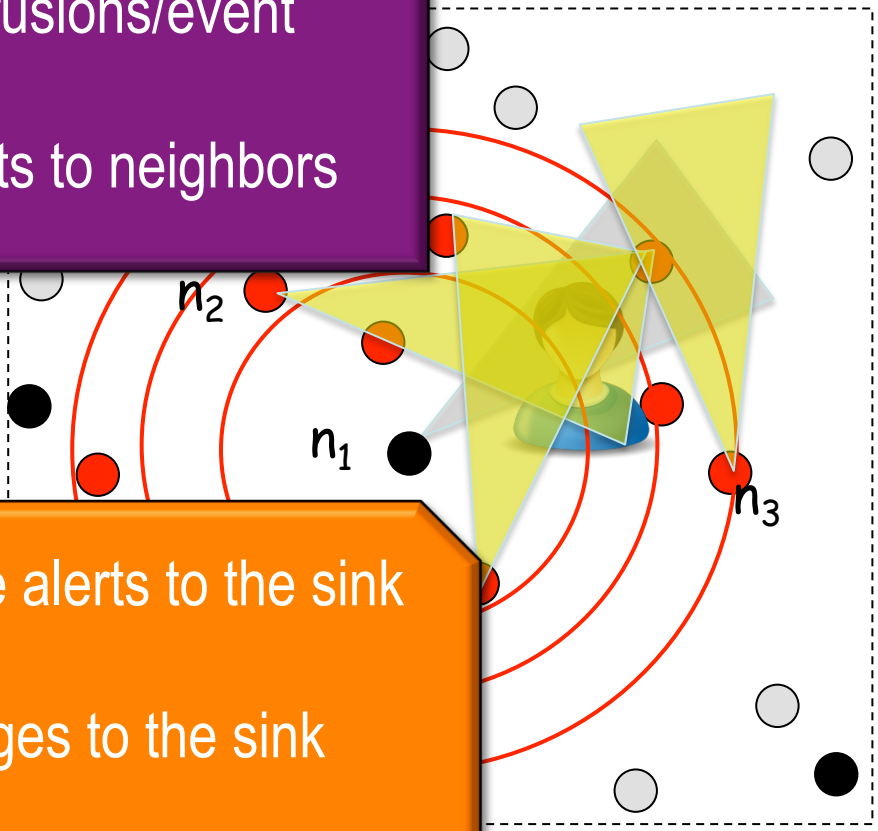


IMAGE SENSORS FOR SURVEILLANCE

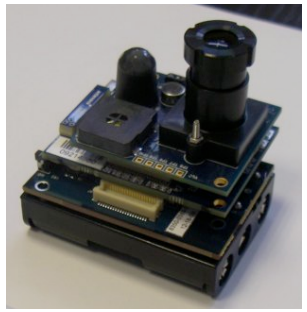
Periodically capture to detect intrusions/event

Send alerts to neighbors

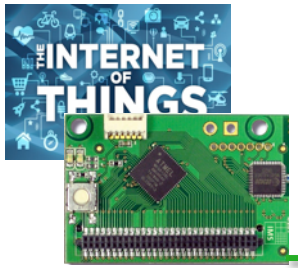


Propagate alerts to the sink

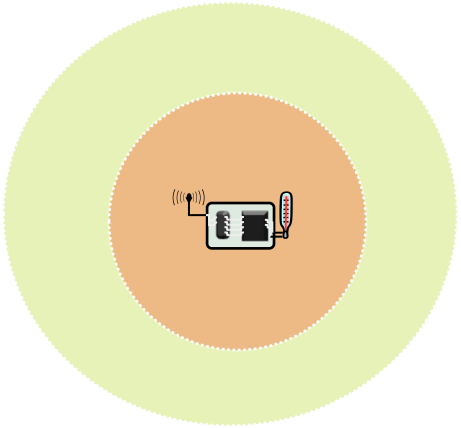
Send images to the sink



alerted node



Sensing range & coverage



vs

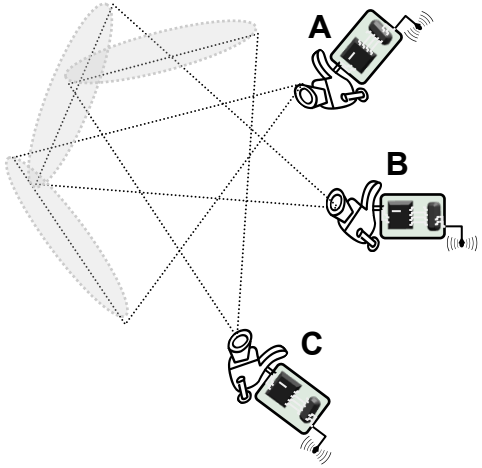
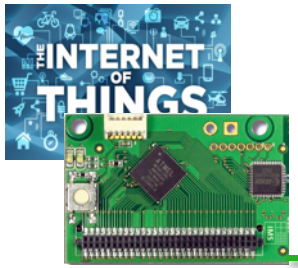


Image sensors capture scene with a Field of View ~ a cone

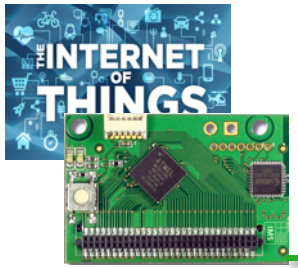
Image resolution, capture speed, rotation,…



Don't miss important events!

Whole understanding of the scene is wrong!!!

What is captured



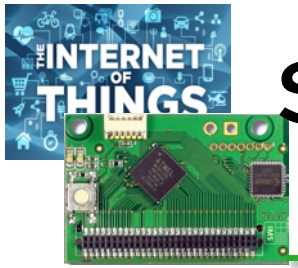
« High-quality » not necessarily good

333x358 16M colors, no light

167x180 16 colors, light

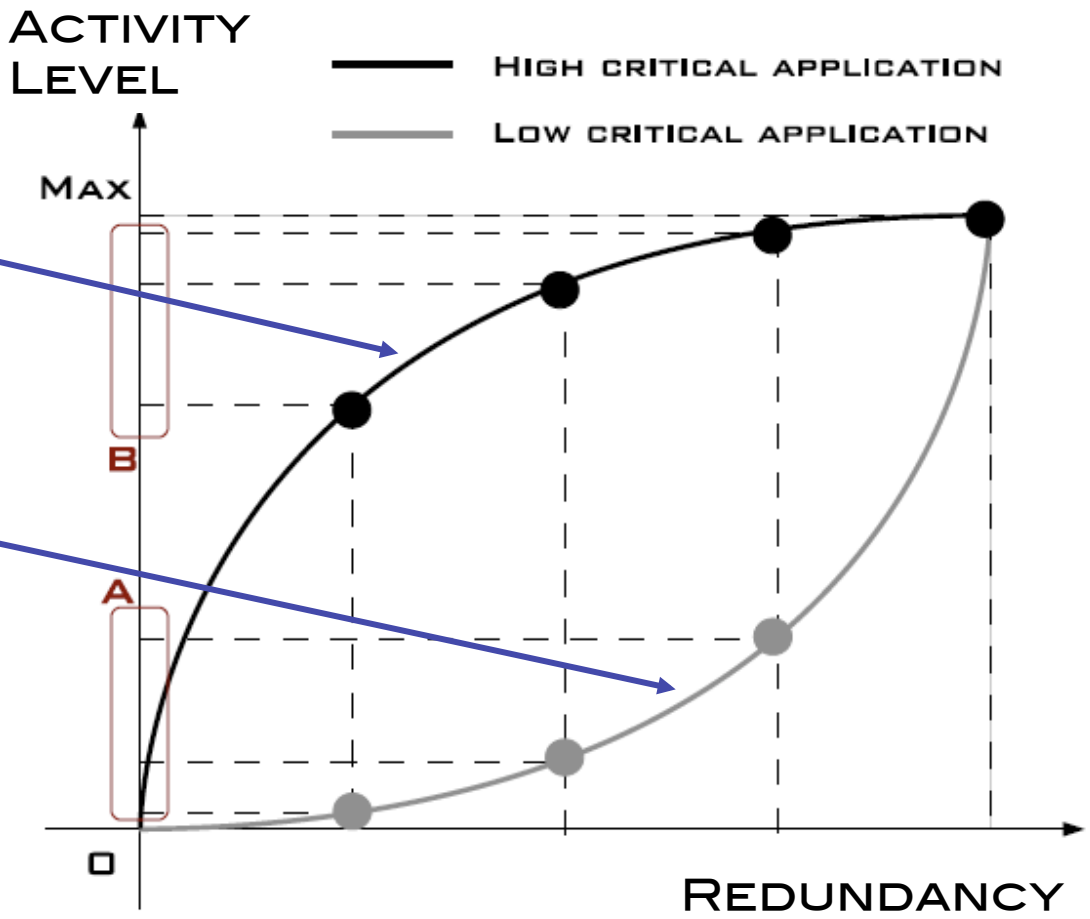
167x180 BW (2 colors), light

Keep in mind the goal of the application!



Schedule activity with criticality in mind

- Link the activity to redundancy level
- High criticality
  - Convex shape
  - Most projections of x are close to the max activity
- Low criticality
  - Concave shape
  - Most projections of x are close to the min activity
- Concave and convex shapes automatically define sentry nodes in the network



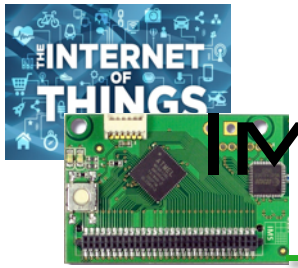
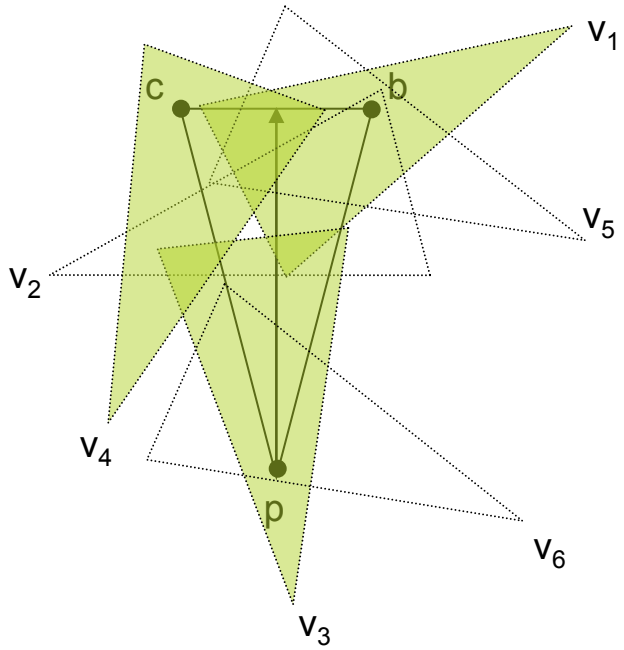
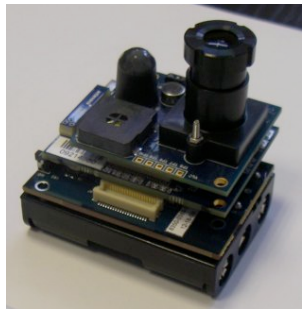


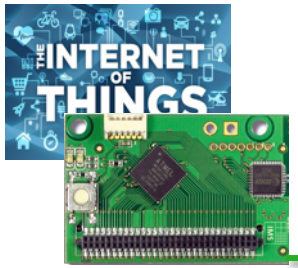
IMAGE SENSOR'S COVER SET

Co(V)= {
{V },
$\{V_1, V_3, V_4\}$,
$\{V_2, V_3, V_4\}$,
$\{V_3, V_4, V_5\}$,
$\{V_1, V_4, V_6\}$,
$\{V_2, V_4, V_6\}$,
$\{V_4, V_5, V_6\}$
}

|Co(V)| = 7







Proposed criticality model

- R^0 can vary in [0
- Behavior funct
  (BV) defines the
  capture speed
  according to R^0
- $R^0 < 0.5$
  - Concave shape

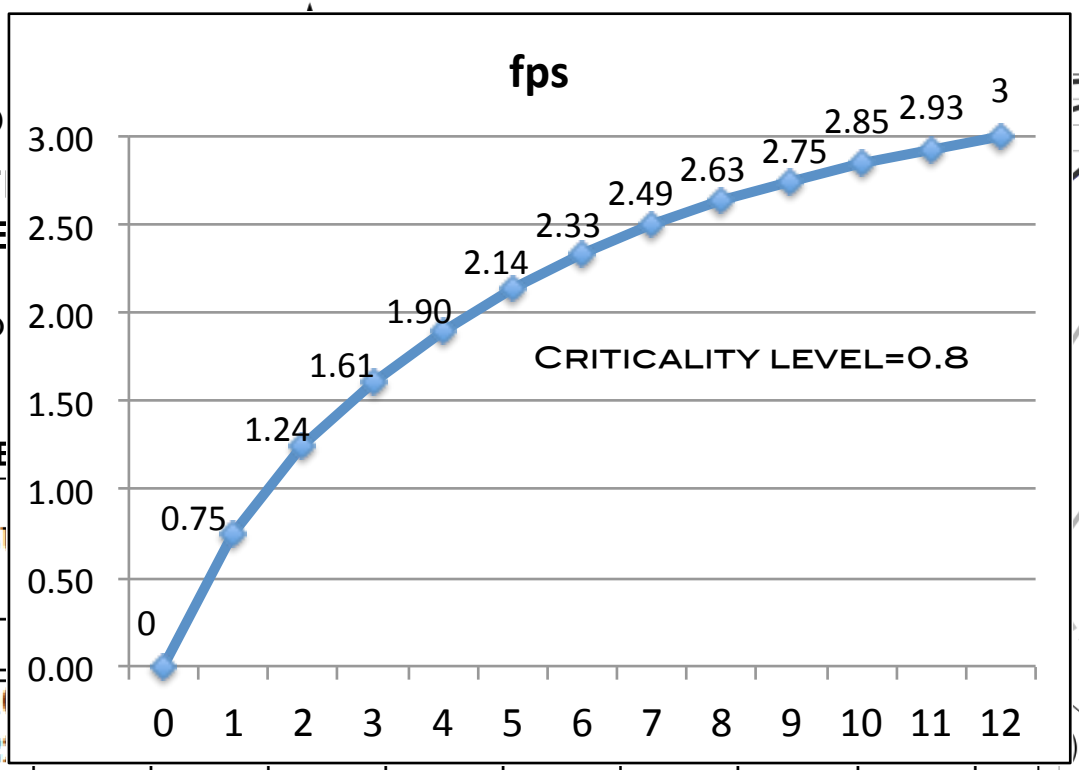
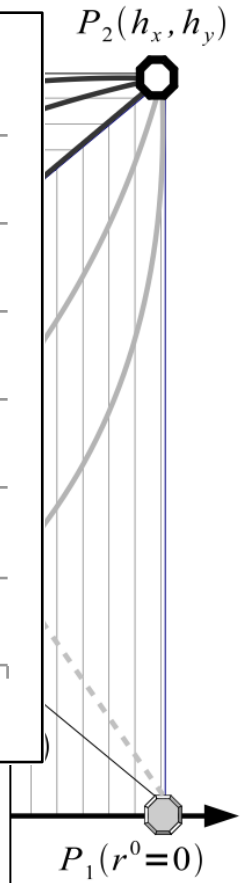
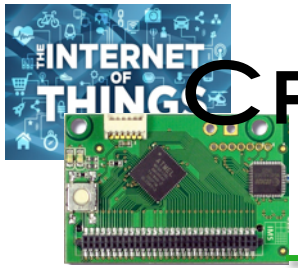


Table 1: Capt

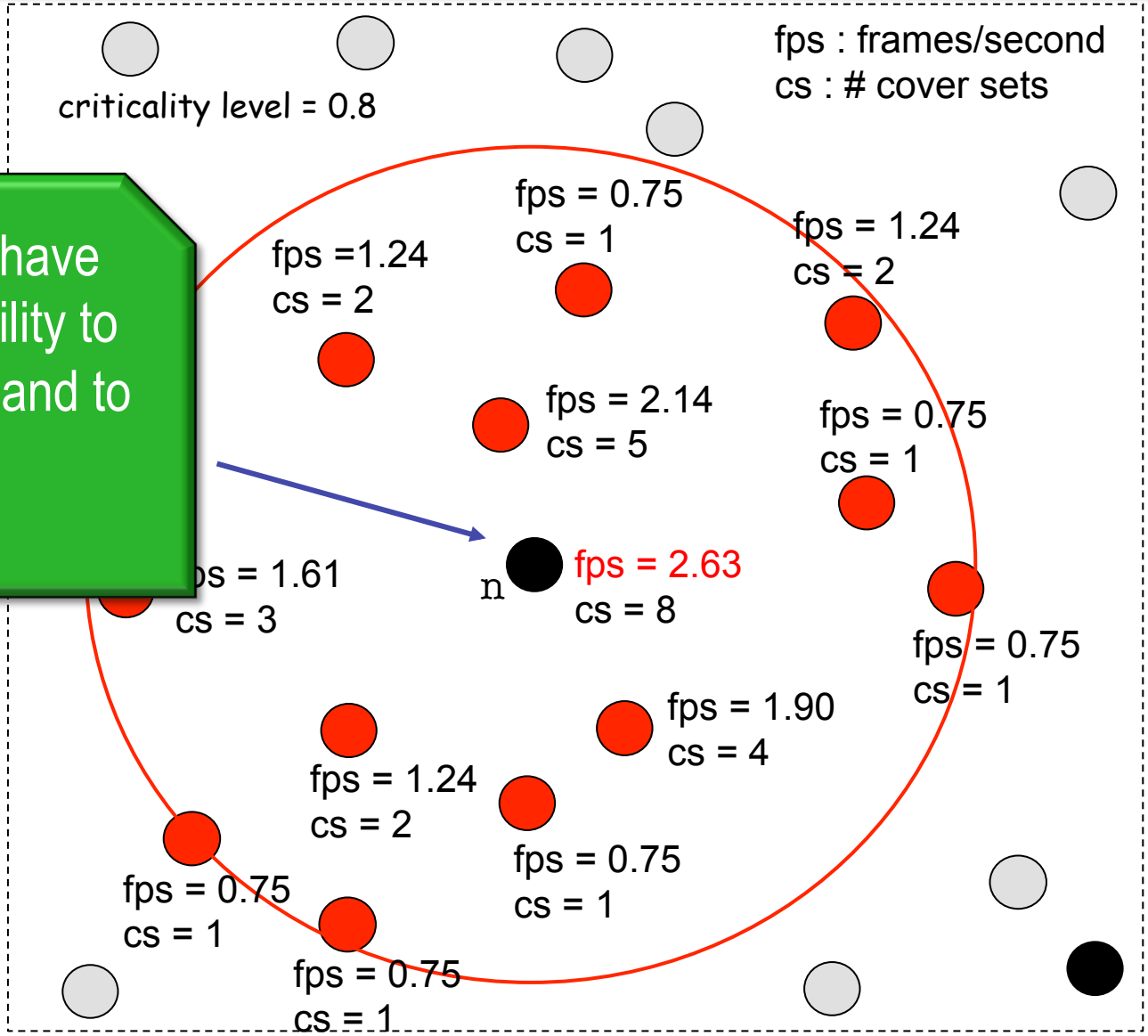
r^0	1	2	3	4								
0	0.01	0.02	0.05	0.1								
.1	0.03	0.08	0.14	0.2								
.4	0.17	0.35	0.55	0.75	0.97	1.20	1.45	1.72	2.00	2.31	2.64	3
.6	0.36	0.69	1.00	1.28	1.55	1.80	2.03	2.25	2.45	2.65	2.83	3
.8	0.75	1.24	1.61	1.90	2.14	2.33	2.49	2.63	2.75	2.85	2.93	3
1	1.48	1.95	2.25	2.46	2.62	2.74	2.83	2.90	2.95	2.98	2.99	3



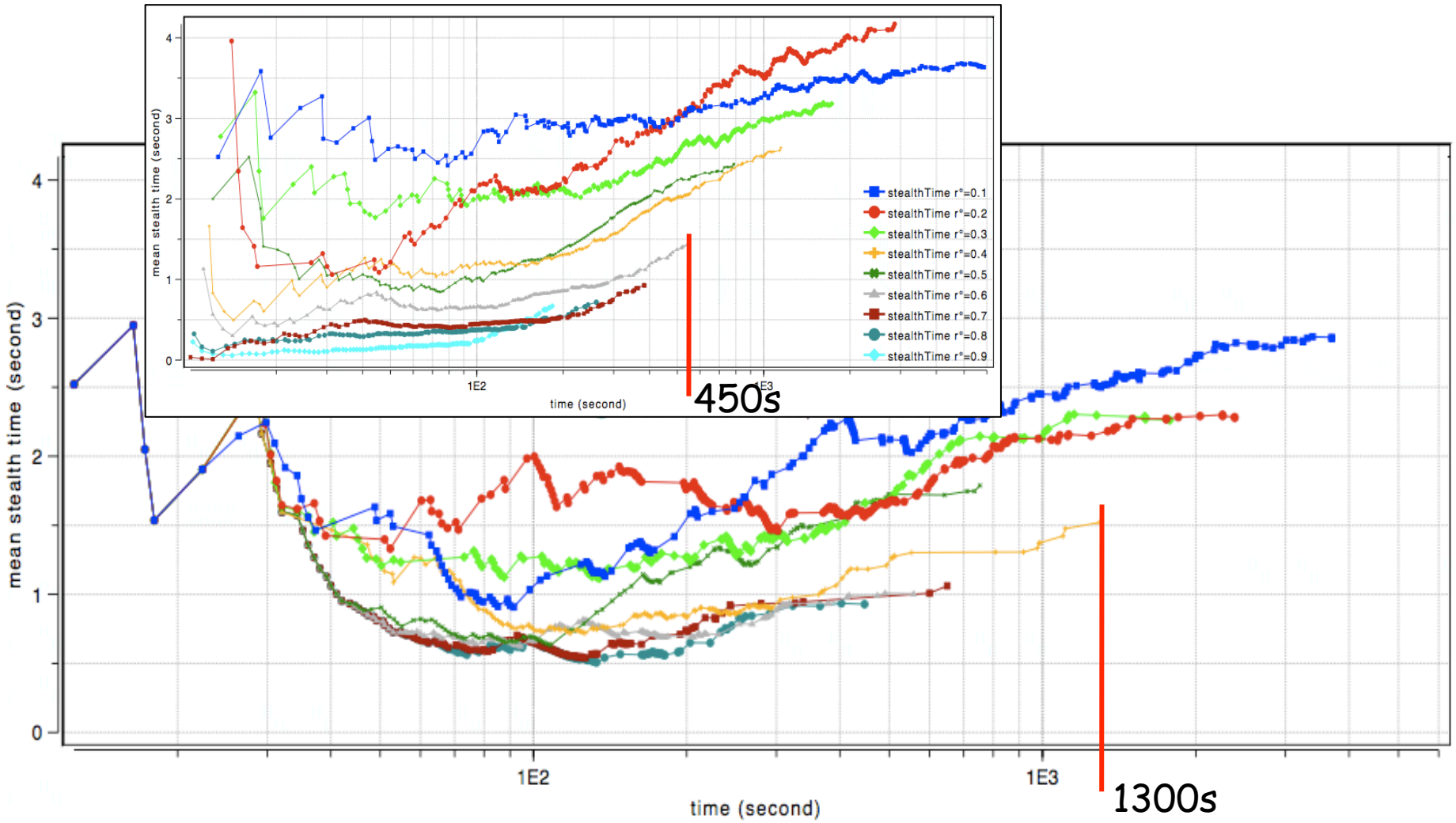


CRITICALITY-BASED ACTIVITY SCHEDULE

Sentry nodes have higher probability to detect events and to send alerts



MEAN STEALTH TIME UNDER CRITICALITY-BASED SCHEDULING



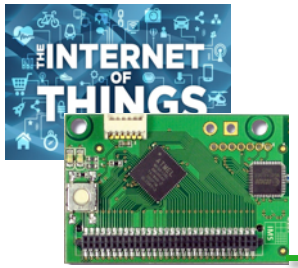


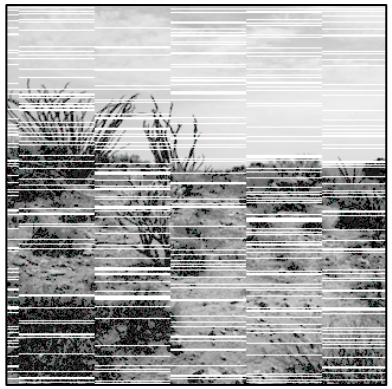
Image quality? Uncompressed BMP

1617 packets, 64 bytes payload, one Hop
Loss rate: 20%, No loss bursts (radio), No duty-cycling

Original 320x320
256 gray levels,
BMP 102400 bytes

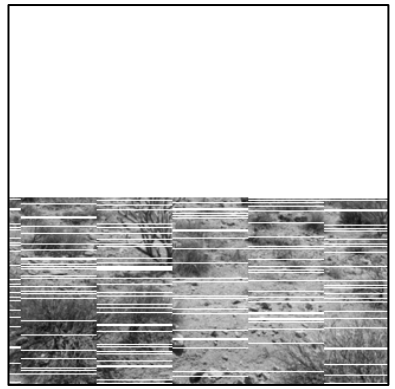
Max TX rate = 250 kps
(IEEE 802.15.4)

Minimum latency = 3.27s



1340 out of 1617
packets received

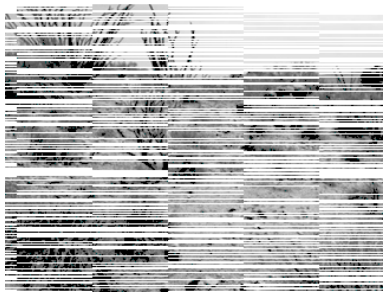
1303 out of 1617
packets received



674 out of 1617
packets received

With loss bursts (radio)

921 out of 1617
packets received



689 out of 1617
packets received

913 out of 1617
packets received

Cannot really use the compressed version of BMP using RLE.

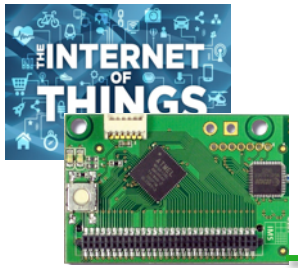


Image quality? Standard JPG

427 packets, 64 bytes payload, one Hop
Loss rate: 20%, No loss bursts (radio), No duty-cycling

Original 320x320
256 gray levels,
JPG 27303 bytes

Max TX rate = 250 kps
(IEEE 802.15.4)

Minimum latency = 0.87s

Encoding cost of JPEG2000 is too high for these devices.

348 out of 427
packets received

351 out of 427
packets received

9 out of 12 images
could not be decoded

349 out of 1617
packets received

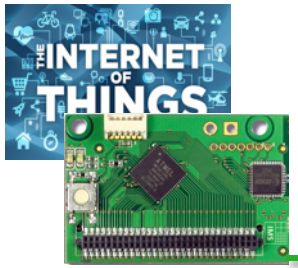
With loss bursts (radio)

258 out of 427
packets received

270 out of 427
packets received

8 out of 12 images
could not be decoded

269 out of 427
packets received



Improving image robustness

Original BMP 40000b

ORIGINAL 200X200
256 GRAY LEVELS,
ADJUSTABLE
IMAGE QUALITY: 6236B
(Q=20)

MAX TX RATE = 250 KPS
(IEEE 802.15.4)

MINIMUM LATENCY = 0.20S

Collaboration with CRAN laboratory, Nancy, France, for robust image encoding techniques for WSN.

Q=50 S=11045b 142pkts

PSNR=25.1661

Q=40 S=9701b 123pkts

PSNR=24.2231

Q=30 S=8100b 101pkts

PSNR=23.2264

Q=20 S=6236b 76pkts

PSNR=22.1293

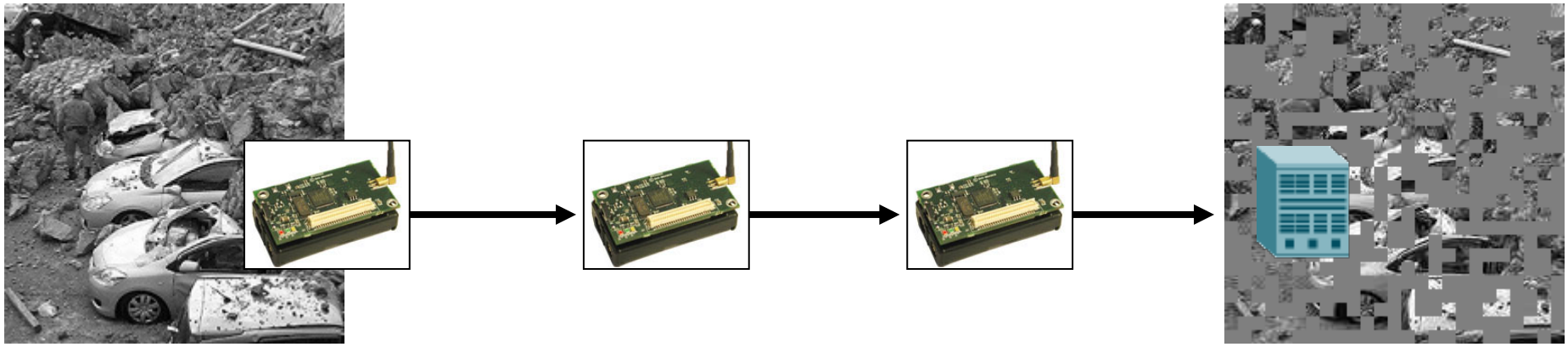
Q=15 S=5188b 63pkts

PSNR=21.4475

Q=10 S=3868b 47pkts

PSNR=20.5255

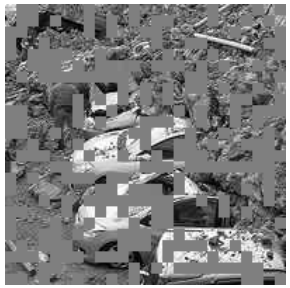
Robust to packet losses

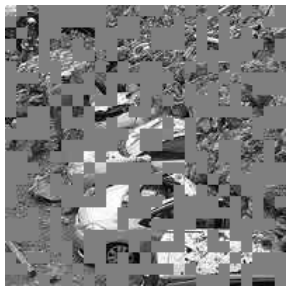


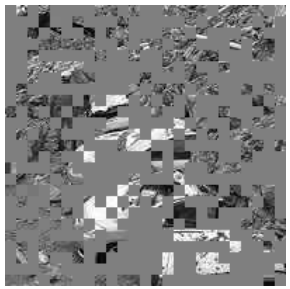

10%

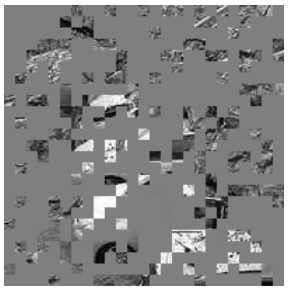

20%

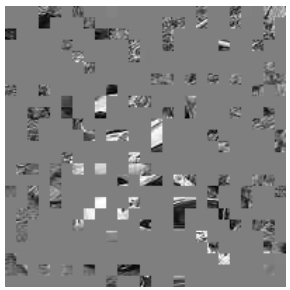

30%

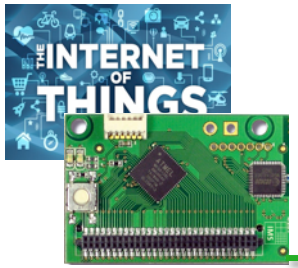

40%


50%


60%

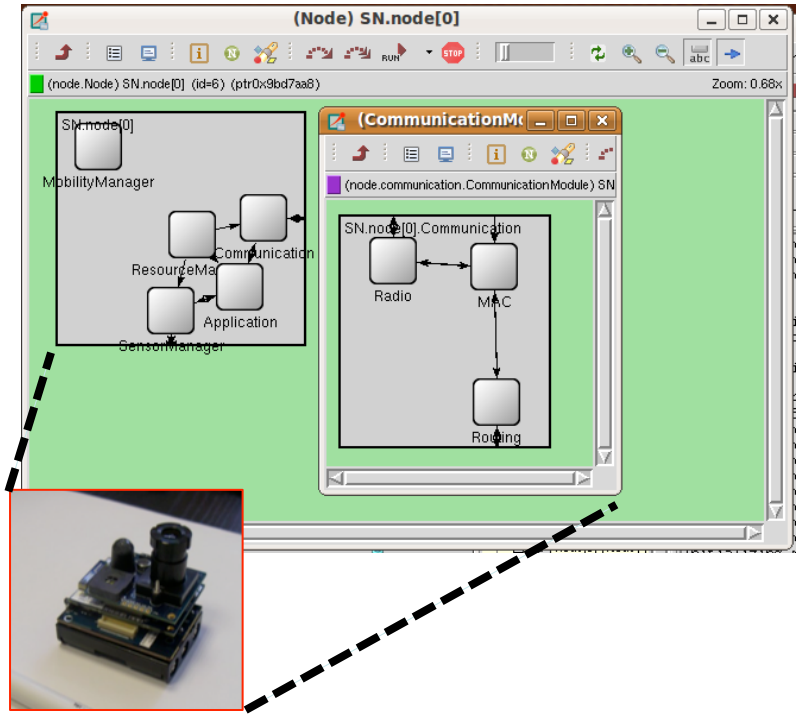

70%

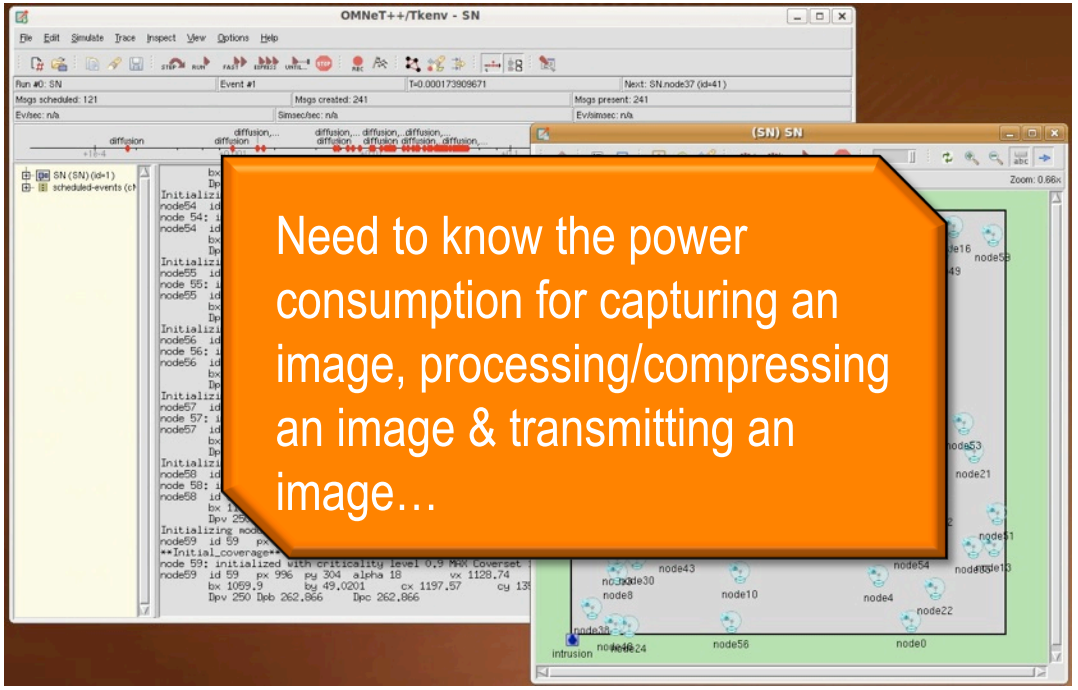

80%



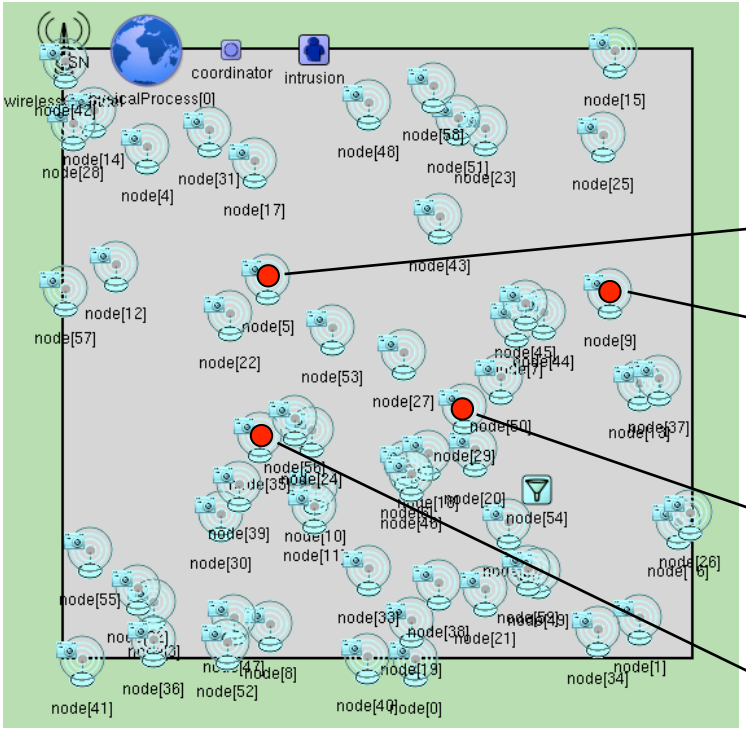
Multimedia sensor simulation model

- Advanced propagation and radio models
- Layered, flexible architecture





Intrusion detection scenario



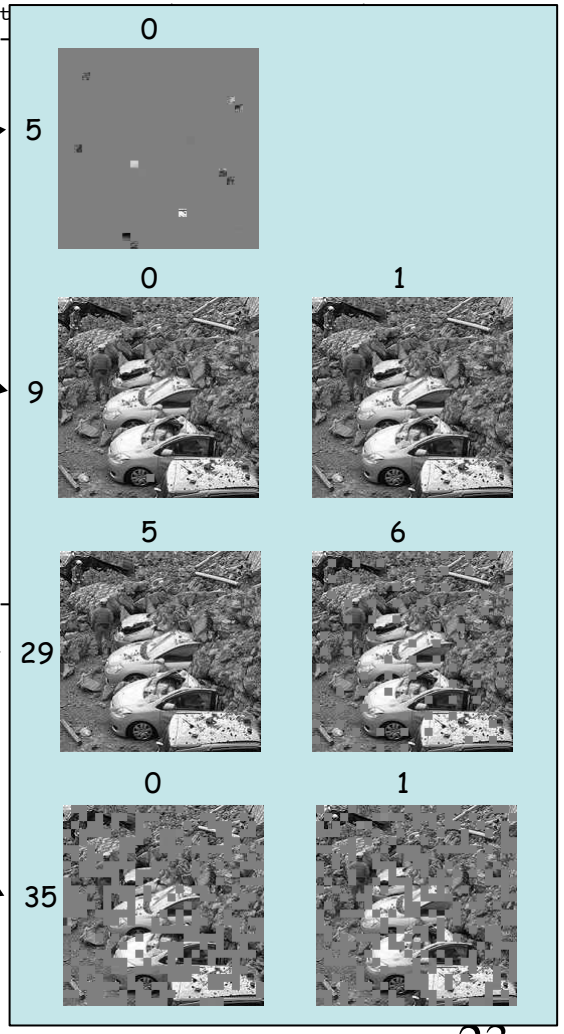
60 image sensor nodes
75mx75m
1 sink (node 54)

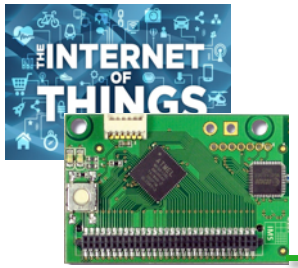
Application:Image sent

	Images	Packets	by coverset
node=2	1	206	0
node=5	4	824	0
node=9	2	412	2
node=10	6	1236	6
node=12	1	206	0
node=15	2	412	2
node=17	1	206	0
node=19	3	618	0
node=22	4	824	0
node=23	2	412	0
node=24	6	1236	0
node=26	1	206	1
node=27	6	1236	0
node=29	7	1442	6
node=33	6	1236	6
node=35	12	2472	0
node=37	5	1030	0
node=40	8	1648	3
node=46	2	412	2
node=48	2	412	0
node=50	2	412	2

Application:Image displayed

	all	complete	truncated
index=-1	39	21	18
index=5	1	0	1
index=9	2	1	1
index=10	6	3	3
index=23	2	0	2
index=24	3	0	3
index=27	4	4	0
index=29	7	6	1
index=33	3	3	0
index=35	4	0	4
index=37	5	3	2
index=50	2	1	1



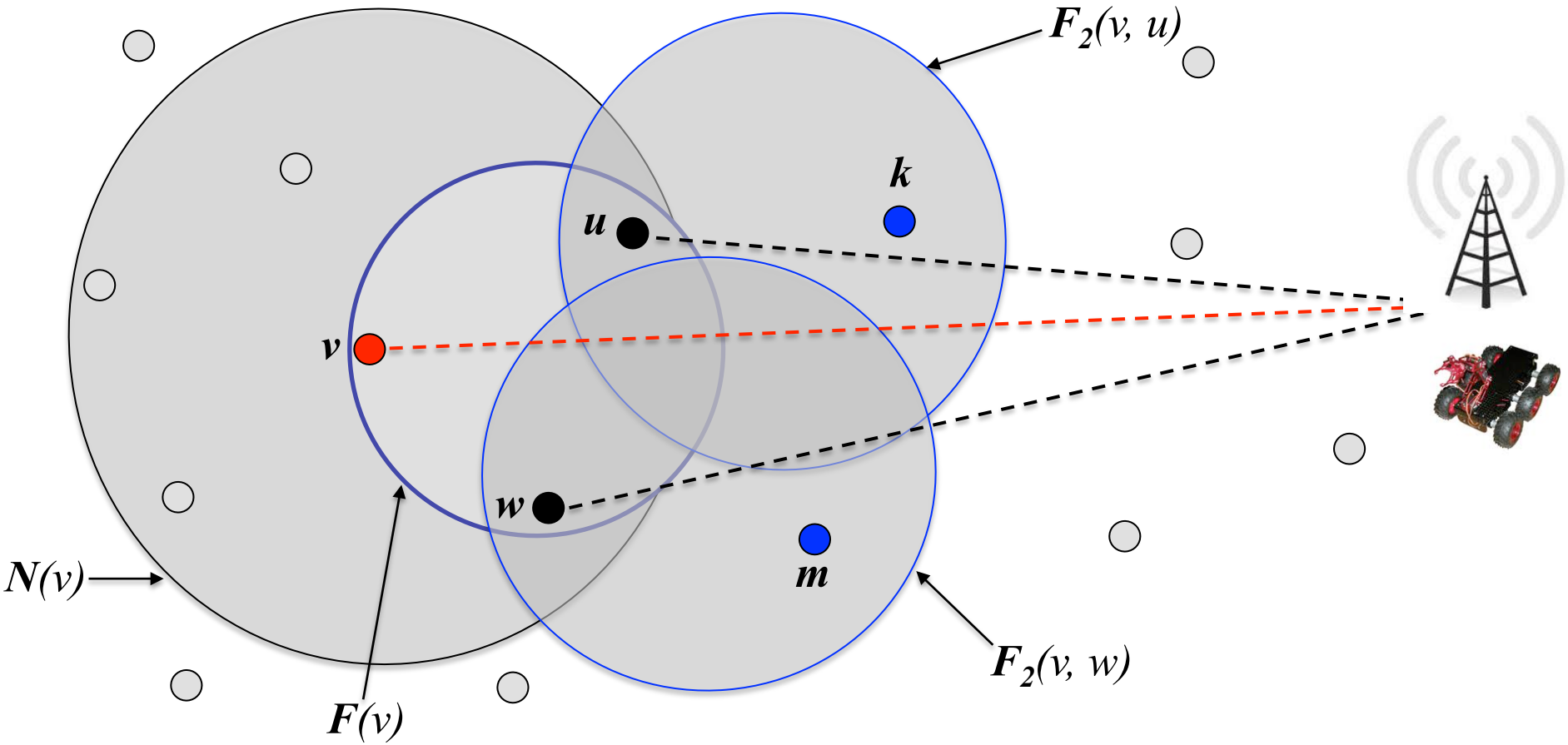


Routing issues

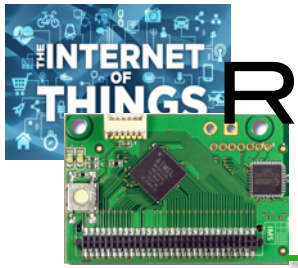
- Images generate large amount of data transported in MANY PACKETS
- Images from multiple sources to the sink can SATURATE THE RADIO CHANNEL
- MULTIPATH ROUTING can be used for reliabiity, load-balancing, mitigating congestion thus packet losses
- As more images need to be sent, a HIGH NUMBER OF PATHS towards the sink is desirable

How to find the best next hop?

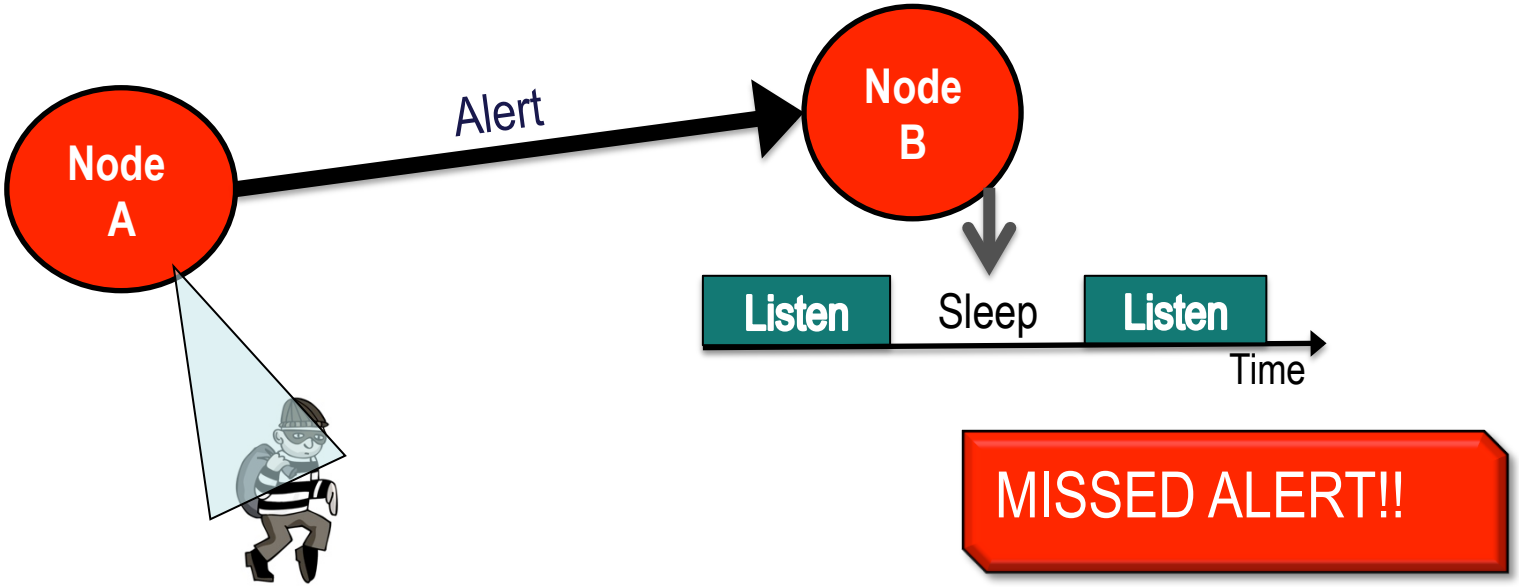
$$\mathbf{F_2}(v,u) = \left\{k | d(k, Sink) < d(u, Sink), u \in F(v), k \in N(u)\right\}$$



$$\mathbf{F}(v) = \left\{u | d(u, Sink) < d(v, Sink), u \in N(v)\right\}$$

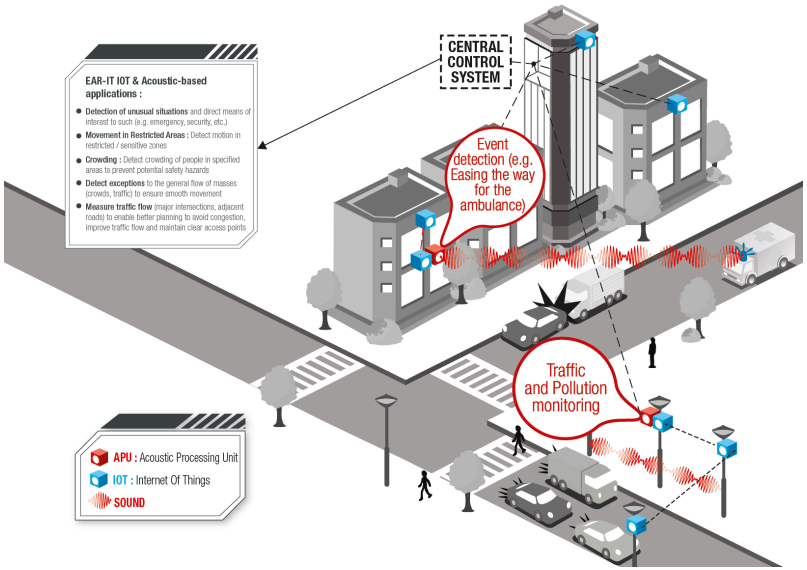


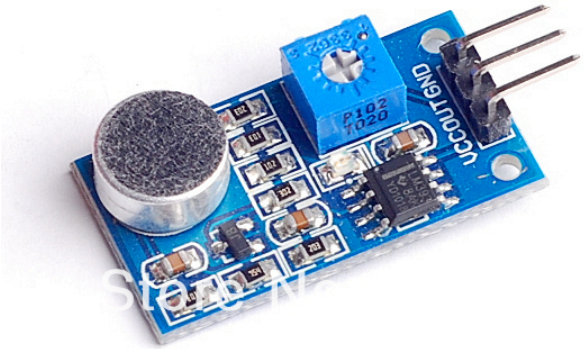
Radio/MAC duty-cycling Issues



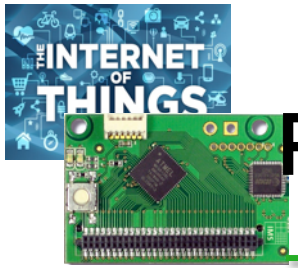
- Radio & MAC layer activities represent a large part of energy consumption
- Most of operation modes imply duty-cycling behavior

Acoustic sensing

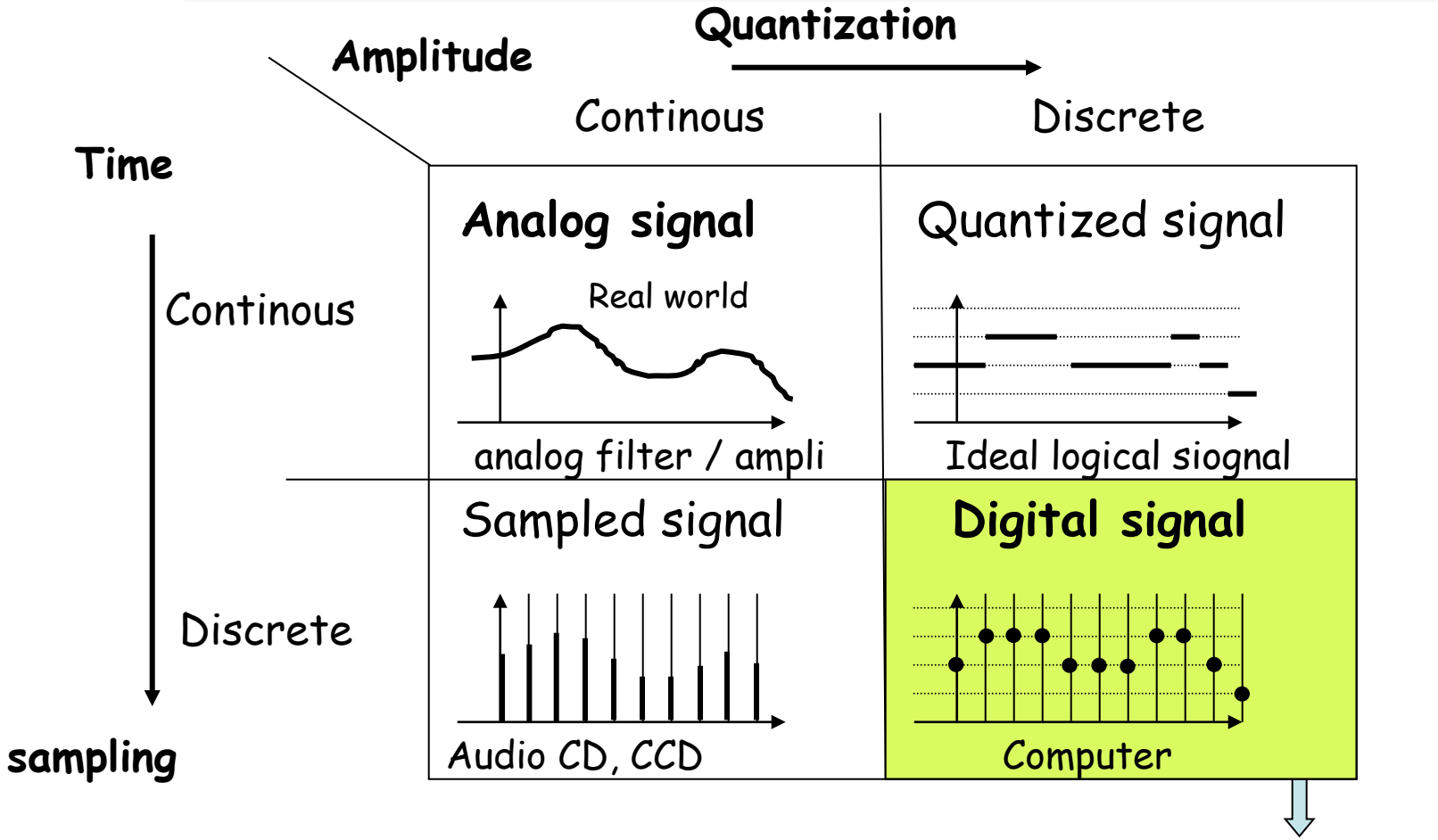




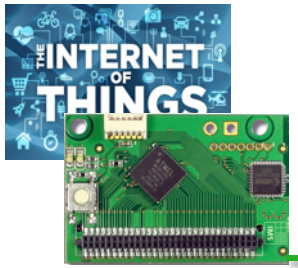
EAR-IT



REVIEW OF DIGITAL AUDIO



Only in this case can we associate an integer value to the signal



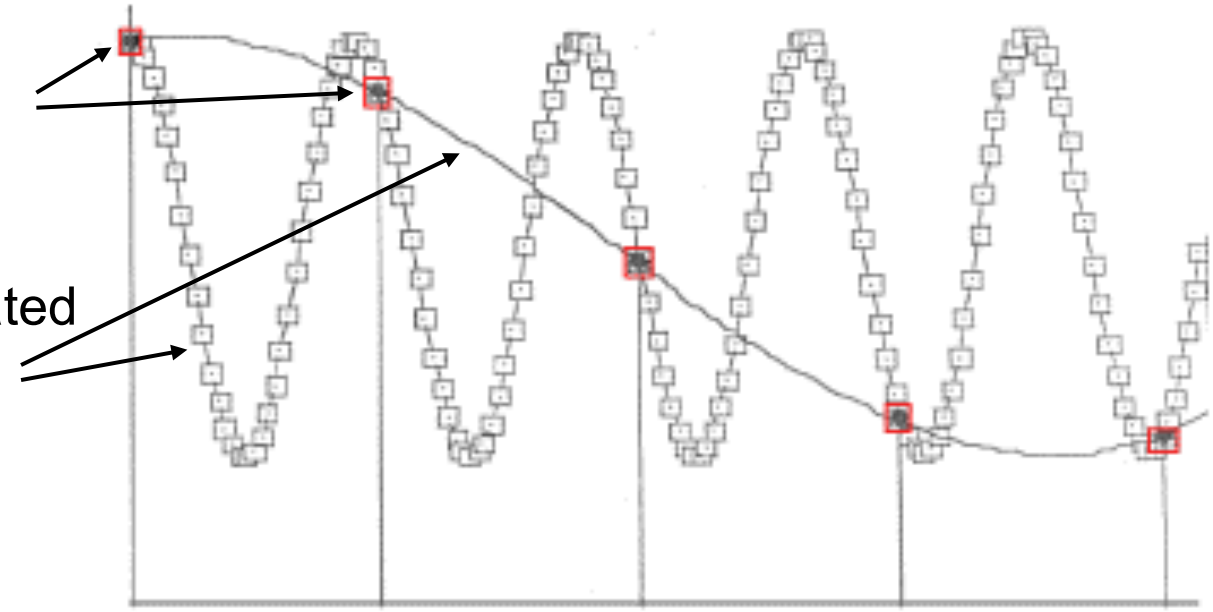
Sampling: Shannon's theorem

Shannon's theorem: Fe > 2 x Fmax(Signal)

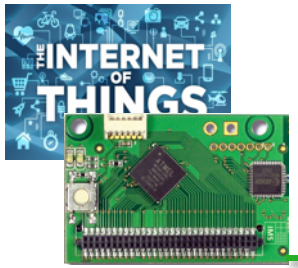
Example :

Samples

Extrapolated signal ?



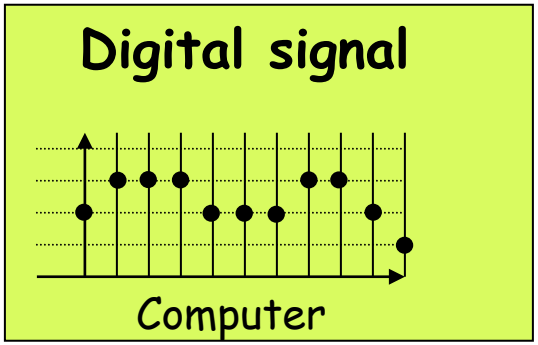
An incorrectly sampled signal will not be reconstituted

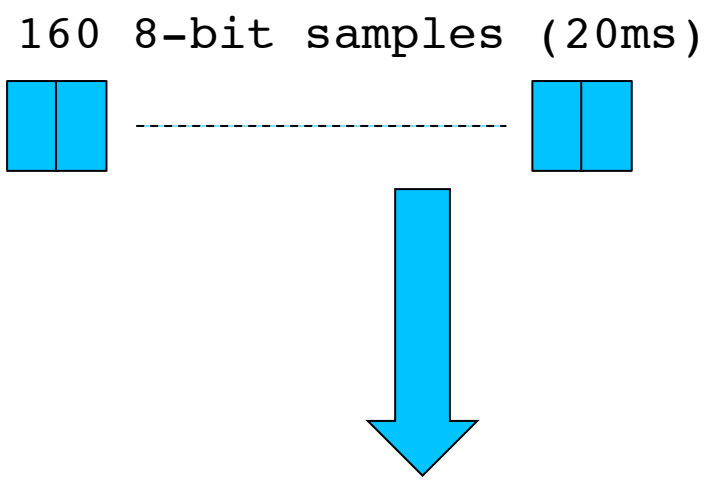


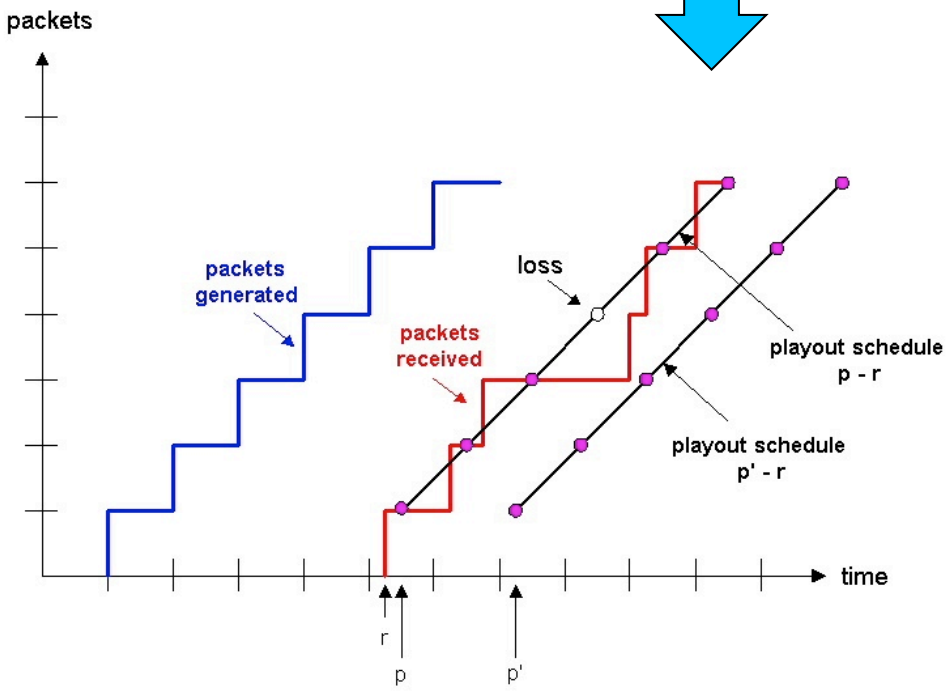
Narrow-band audio

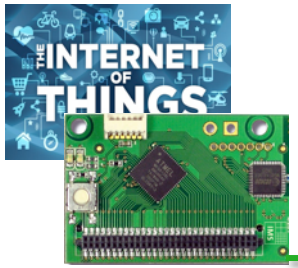
- Sampling rate up to 8kHz
- 1 sample every 1/8000s (125us)
- Sample coded on 8 bits
- Raw throughput of 64kbps
- So-called Pulse Code Modulation (PCM) used in most wired telephony systems
- With 4kHz sampling rate, can reduce to 32kbps

AUDIO STREAMING PRINCIPLE



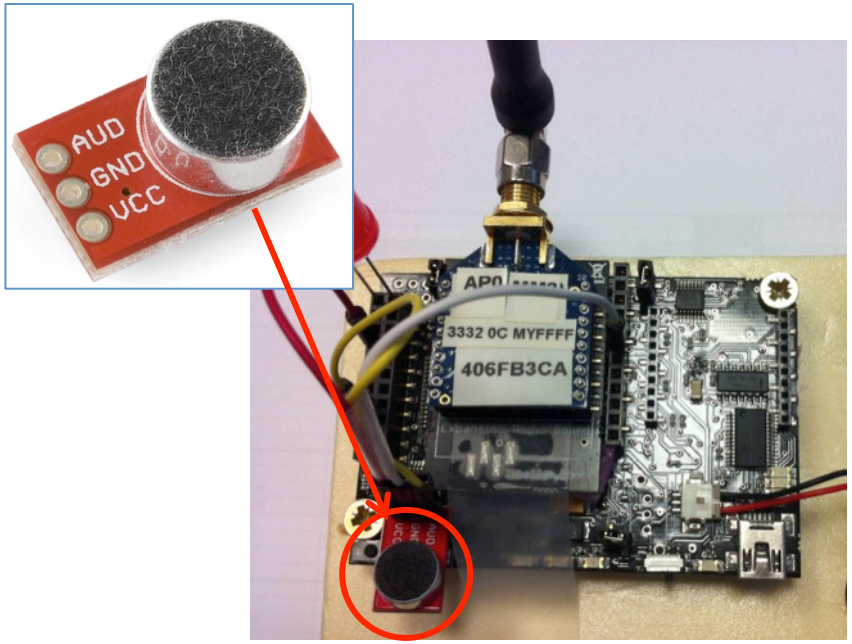


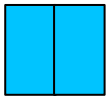




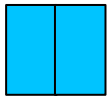
Practical audio

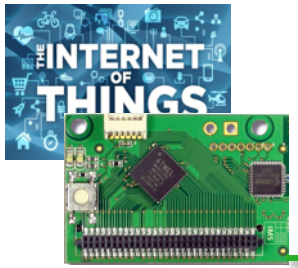
- Electret mic with amplifier on ADC input pin
- Convert from 10-bit to 8-bit sample
- 8Khz sampling gives 64000bps
- 4Khz sampling gives 32000bps





```
100 8-bit samples (12.5ms or 25ms)
-----------------------------------------------------------
```

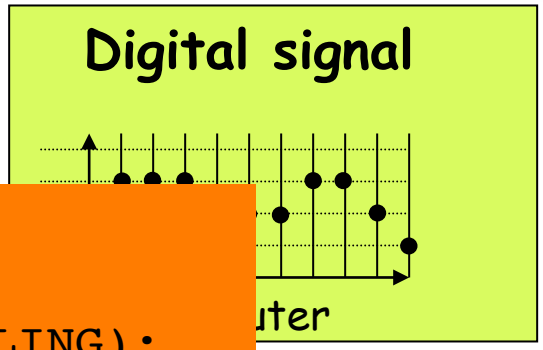




Simple program

```
#define TIMING_SAMPLING 125 // 8000Hz
#define CAPTURE_DURATION 15000000UL // in us 15s
#define SAMPLE_COUNT_CAPTURE CAPTURE_DURATION/TIMING_SAMPLING

void setup() {
  Timer1.initialize(TIMING_SAMPLING);
}

void ca

  sampl

  if (s

    val
    val

    // write on UART1
    ser
  } els

    sto

    Timer1.detachInterrupt();
  }
```

```
void setup() {
  Timer1.initialize(TIMING_SAMPLING);
}
```



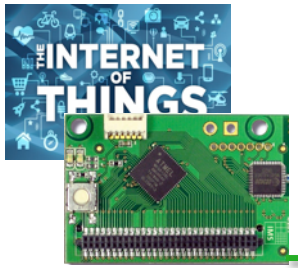
of sound pressure level,

Or any other way to put the sample into

```
val = analogRead(ANALOG2) ; // read analog value
val8bit = ((val >> 2) ) ; // convert into 8 bit
```

```
// we have to wake up
if (millis() - lastSleepTime > 15000 && !capturingAudio) {
                sampleCount=0L;
                lastWakeupTime = millis();
                capturingAudio = true;
                    Timer1.attachInterrupt(callback);
}
```

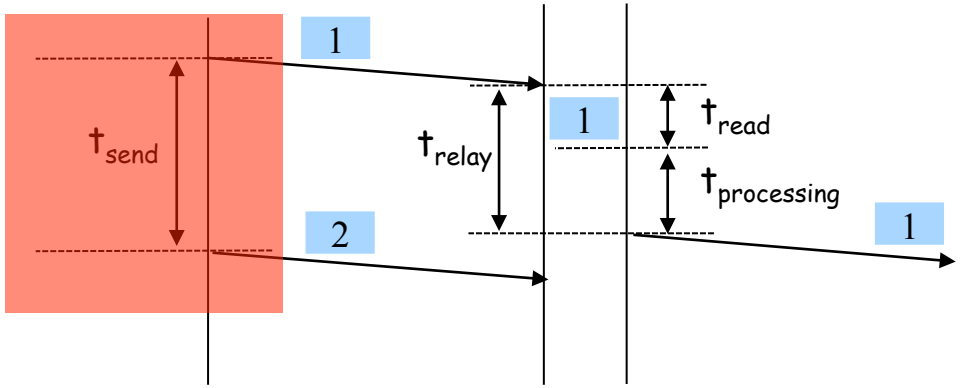
```
                Timer1.attachInterrupt(callback);
    }
}
```

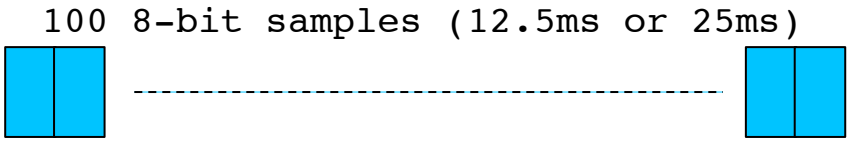


Communication performances

- Application level performances depends on OS, API, hardware architecture
- Usually much lower than radio performances
- What are minimum latencies & max. throughput?
  - For sending?
  - For receiving?
  - For relaying?

Sending performances





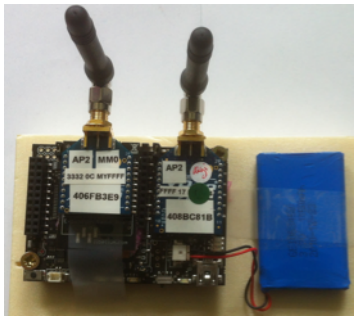
Traffic Generator

```
void loop() {
        T0;
        L0=T0;
        ...
        T1;
        send(buf);
        T2;
        ...
}
```

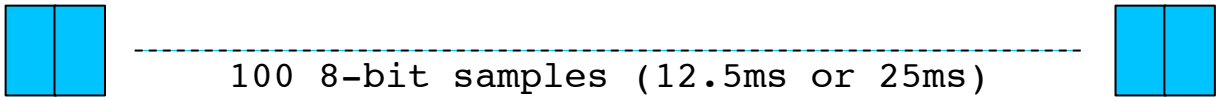
« Time in send() » is T2-T1
« Time between 2 pkt generation » is T0-L0
Time resolution is millisecond
Minimum data manipulation

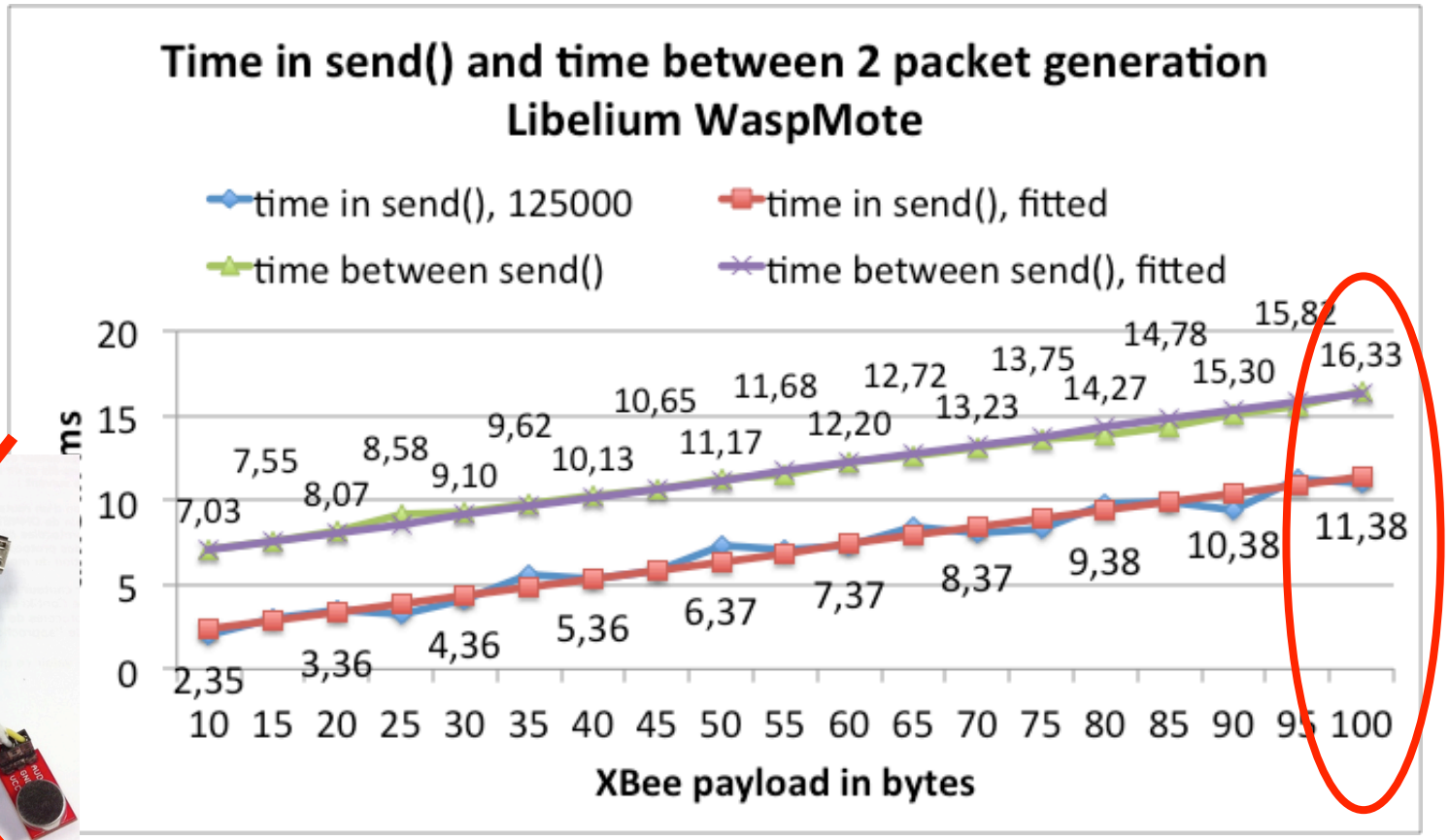
Measure the time in various part of API `send()` when possible.

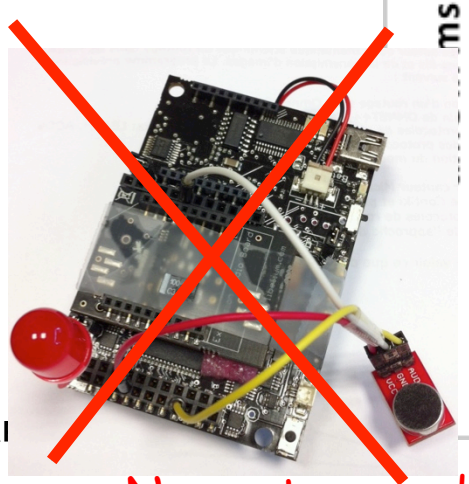
IoT node sending performance



Libelium WaspMote



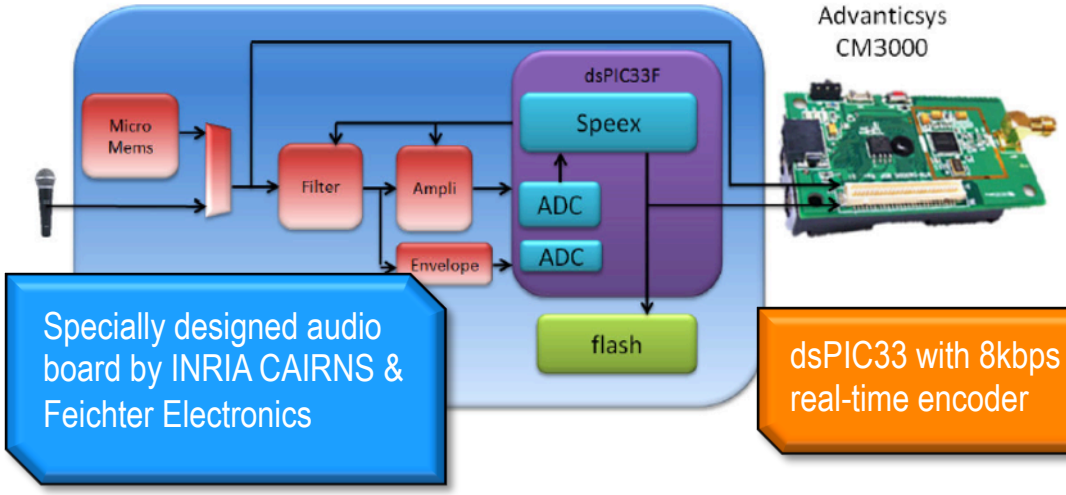


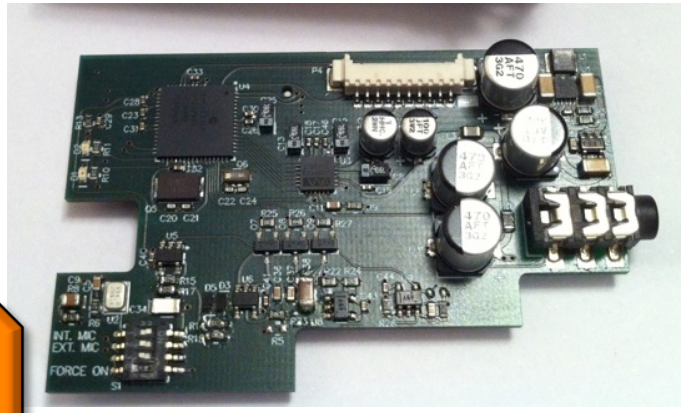


No capture and transmission at the same time if using only mote ucontroller!

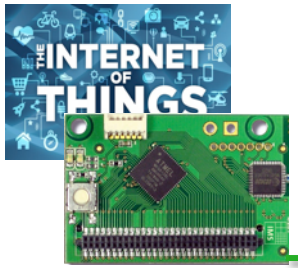
DEVELOPMENT OF AUDIO BOARD

- USE DEDICATED AUDIO BOARD FOR SAMPLING/STORING/ENCODING

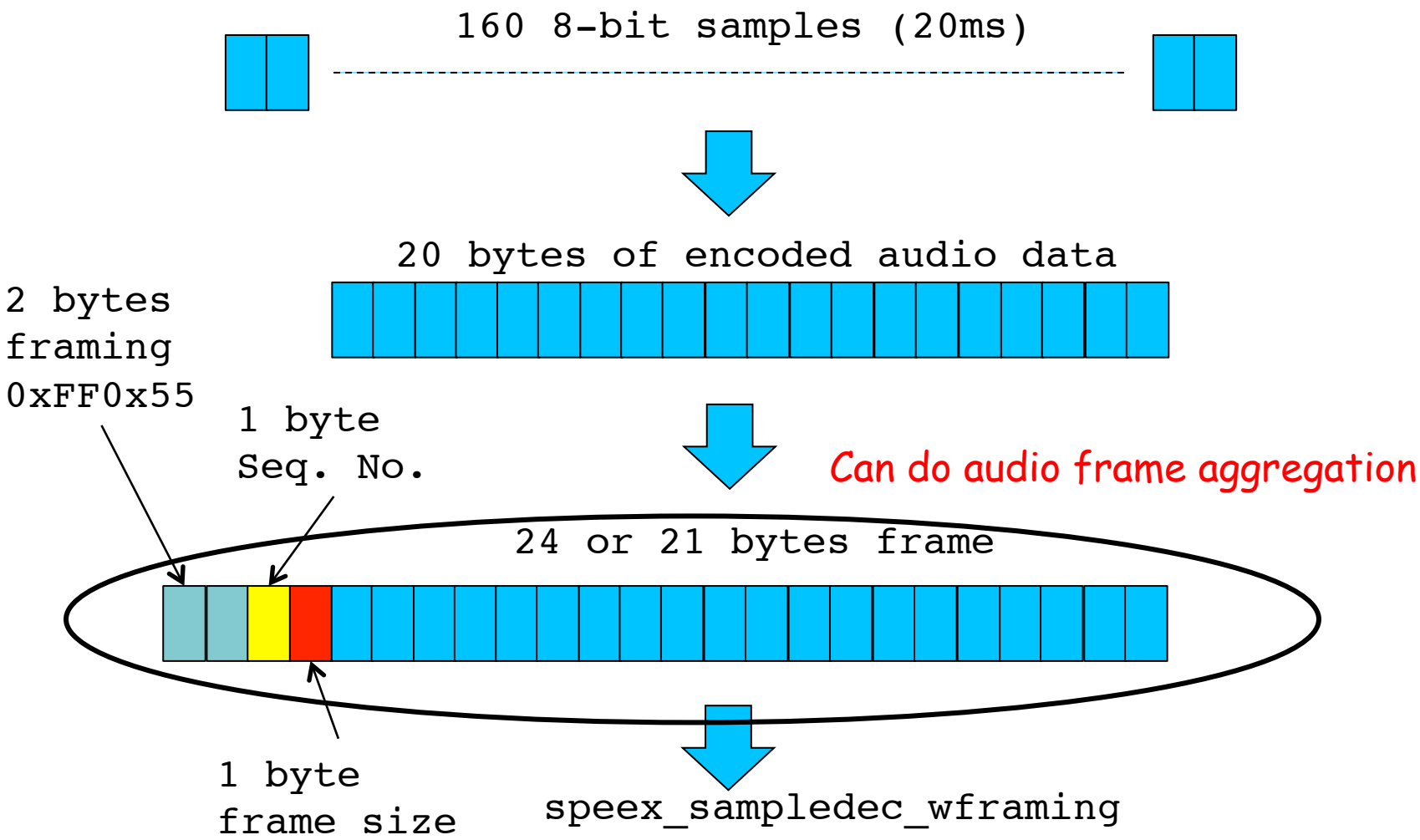


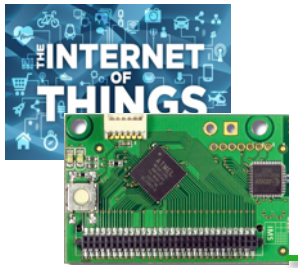


- ENCODING SCHEME IS SPEEX AT 8KBPS
- DESIGNED FOR MULTI-PLATFORM MOTES
- DEVELOPED FOR EAR-IT PROJECT



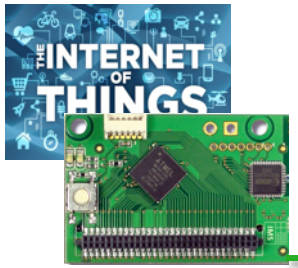
SPEEX AT 8KBPS





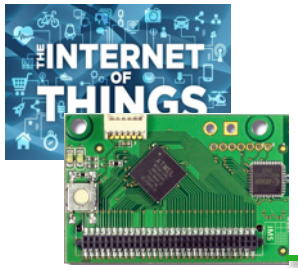
Summary of audio constraints

Codec	Minimum sending rate
Raw 4KHz 8KHz	 100 bytes every 25ms 100 bytes every 12.5ms
Speex 8000bps A1 A2 A3 A4	 24 bytes every 20ms 48 bytes every 40ms 72 bytes every 60ms 96 bytes every 80ms

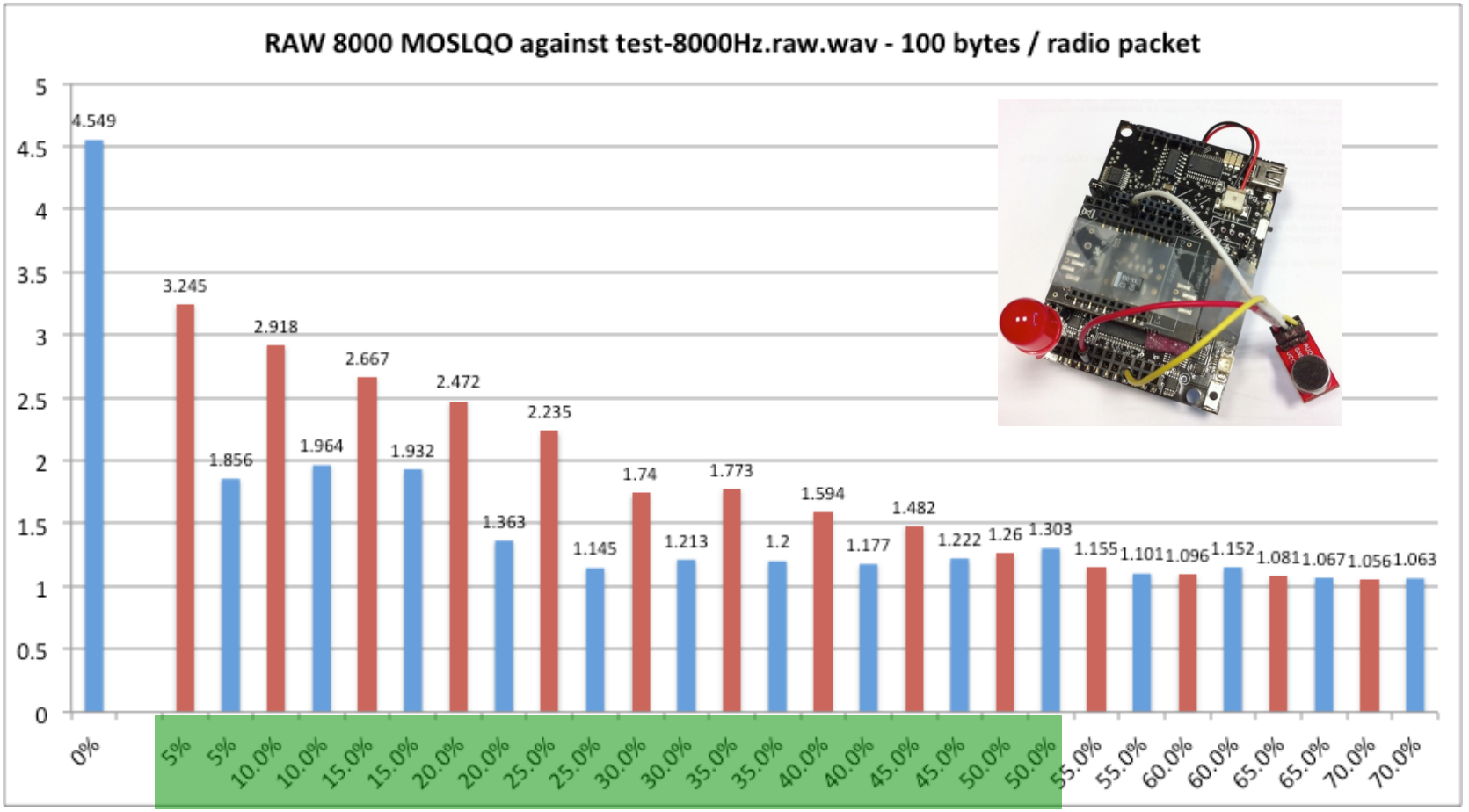


Audio quality: PESQ & MOS (1)

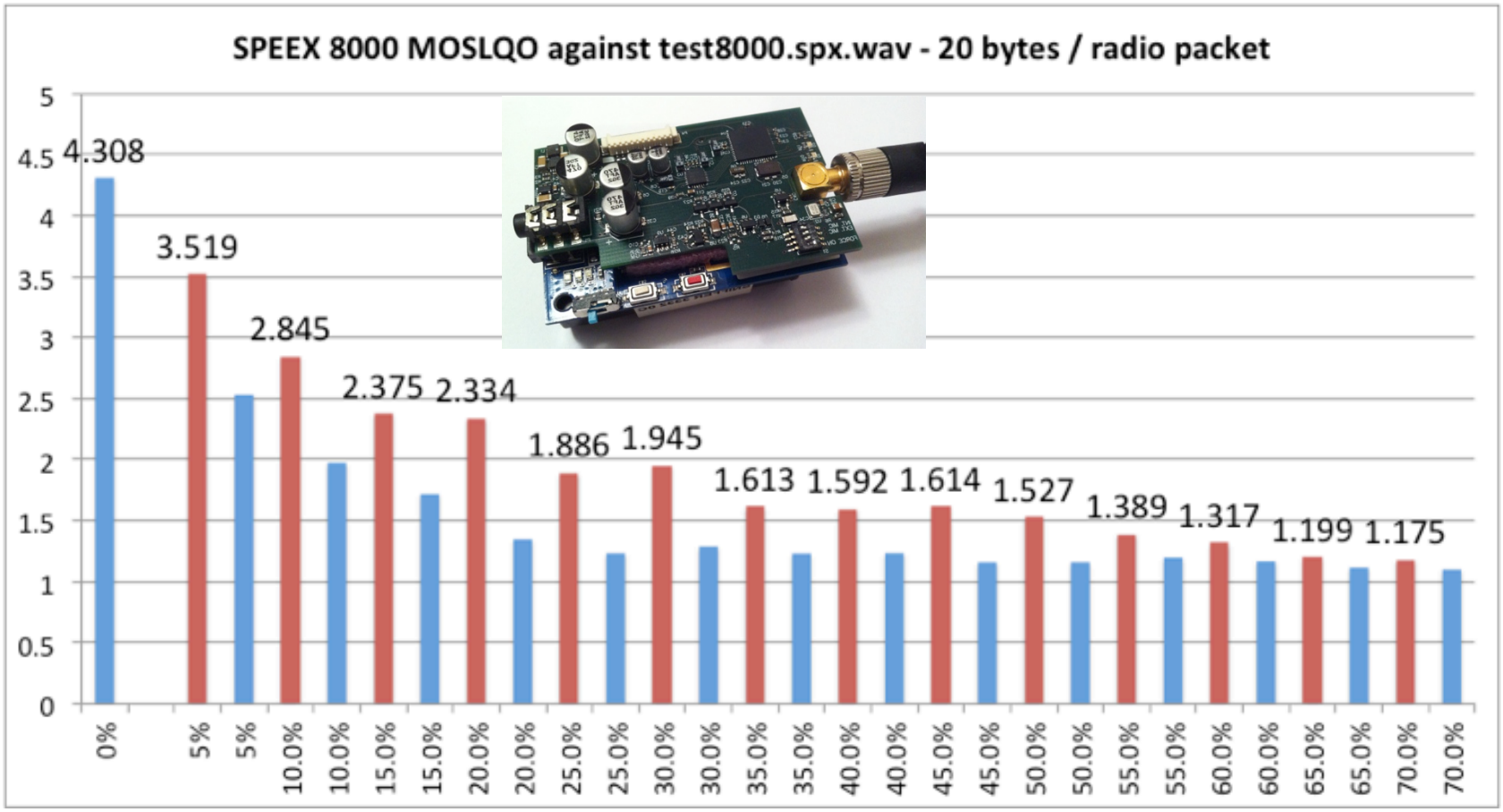
- ITU-T P.862 Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs.
- We can use ITU-T PESQ tool to determine the MOS value for loss-free encoded audio (codec2, speex, ...). MOS-LQO values greater than 2.6 are considered good.



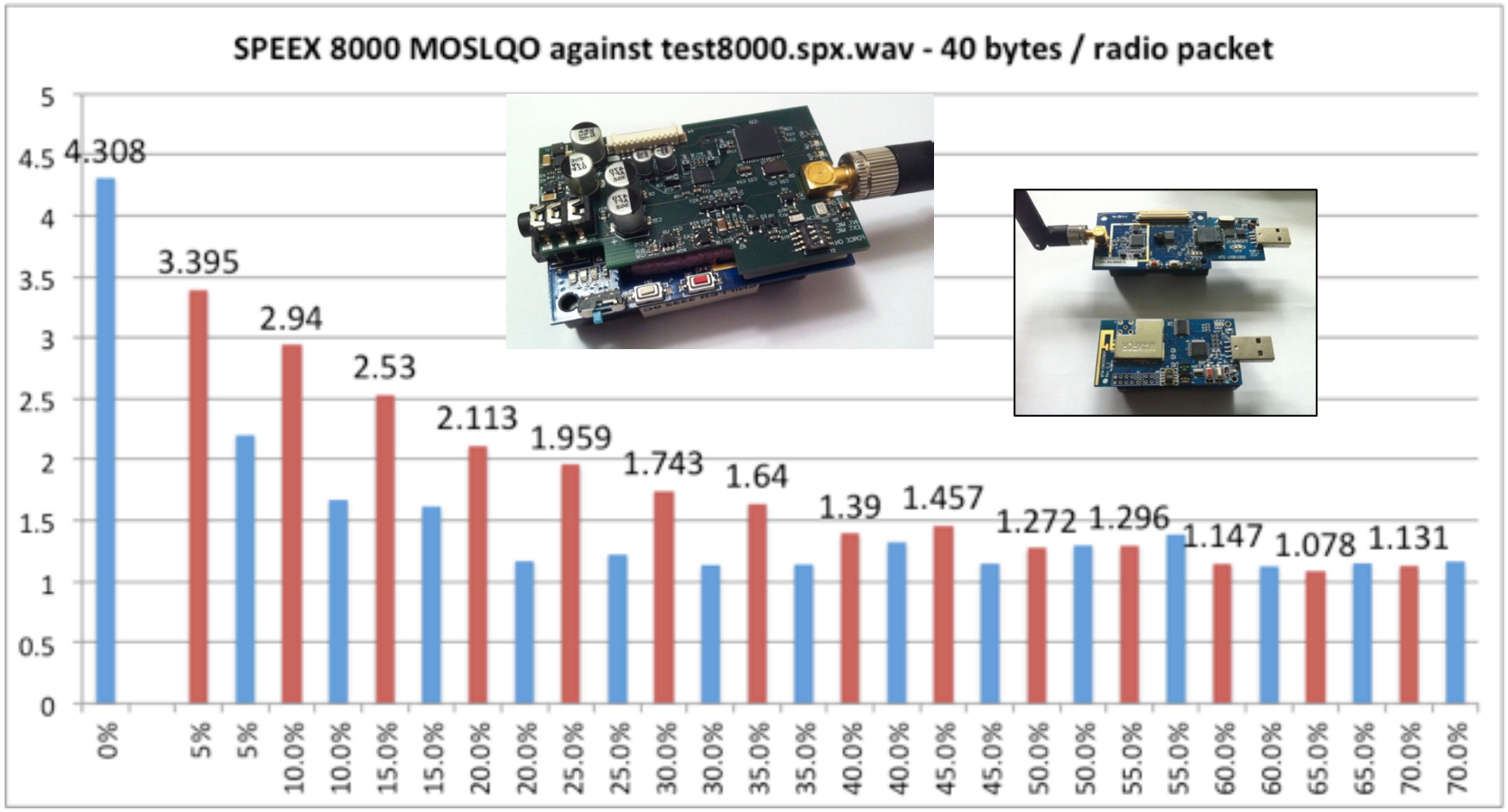
Test8000.raw



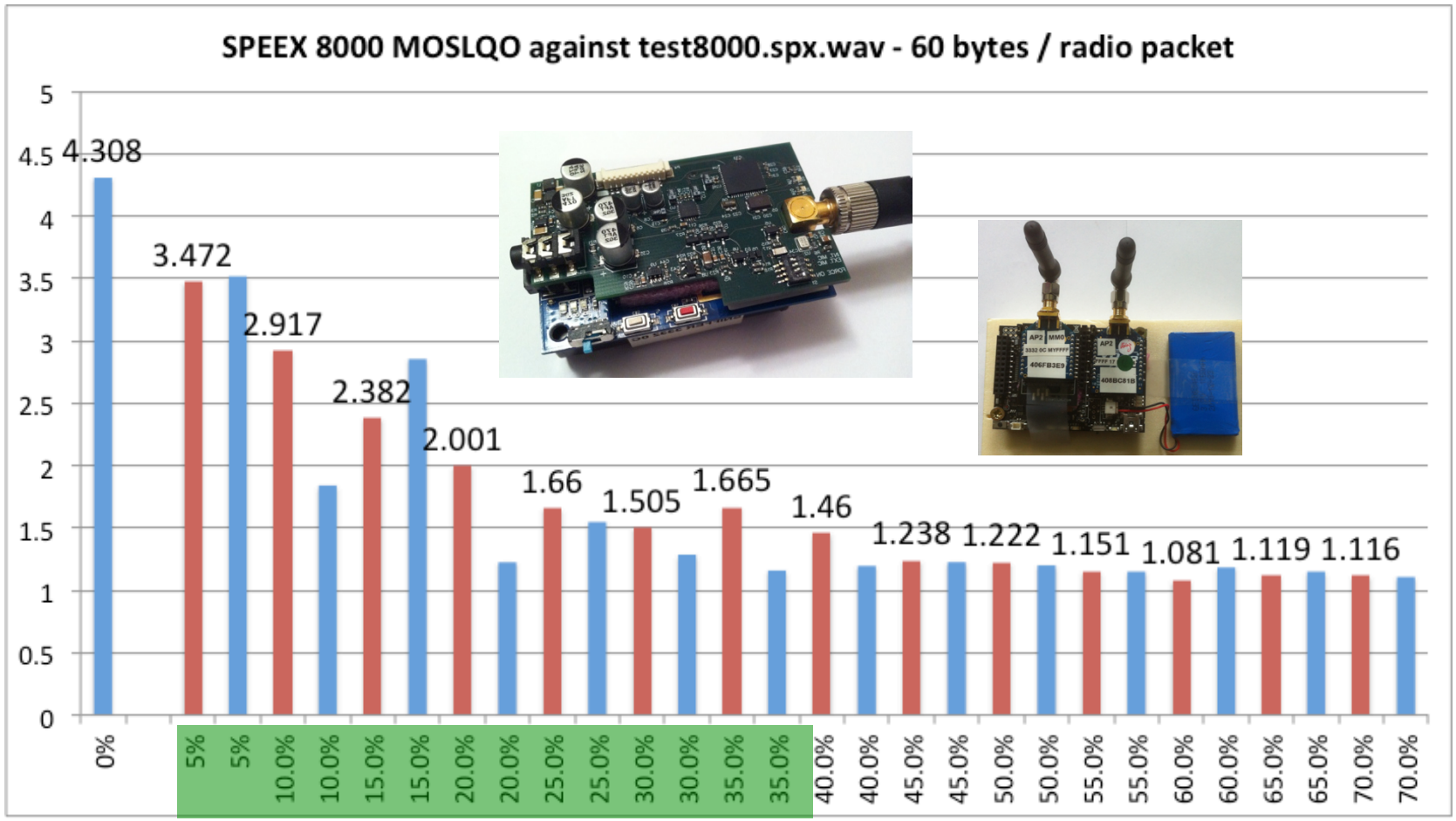
Test8000.spx, 20B/pkt (A1)



Test800.spx, 40B/pkt (A2)



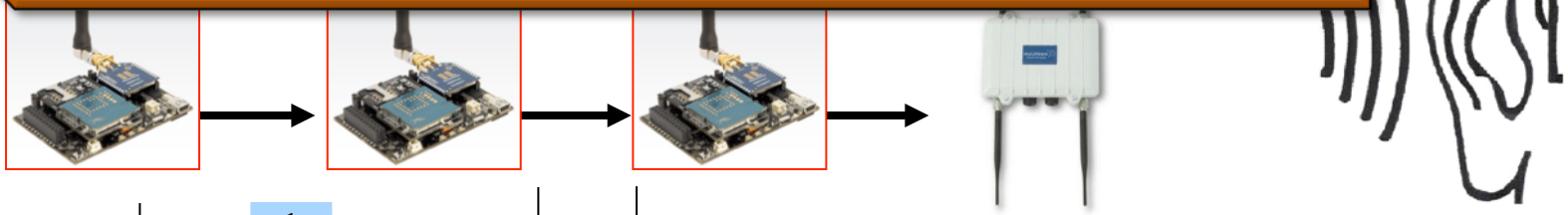
Test800.spx, 60B/pkt (A3)

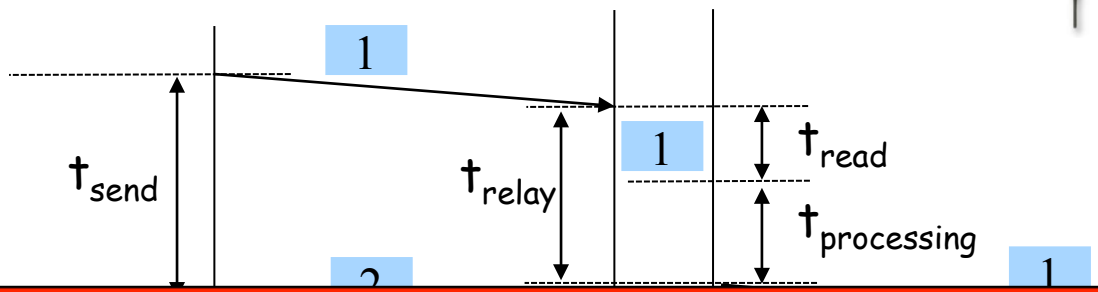


Multi-Hop Packet Forwarding

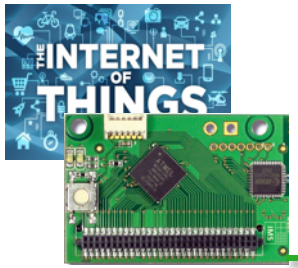
Multi-hop is very costly (routing) and generates lot's of packet losses!

In data-intensive applications, a lot of packets will be transmitted, usually at high transmission rate!

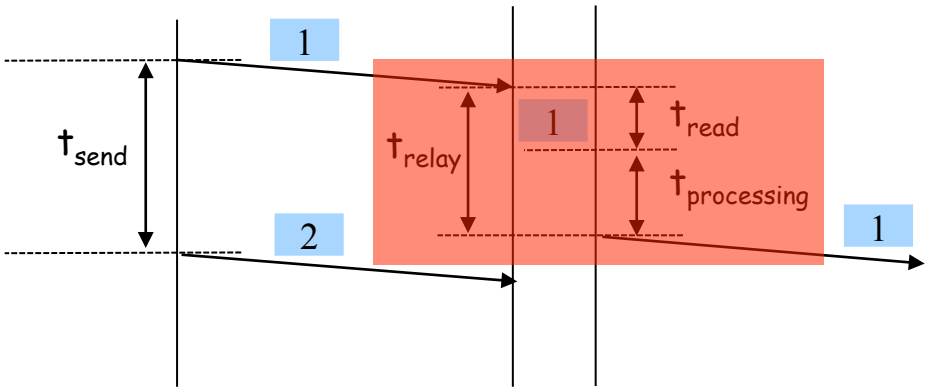




What level of performances can we expect?

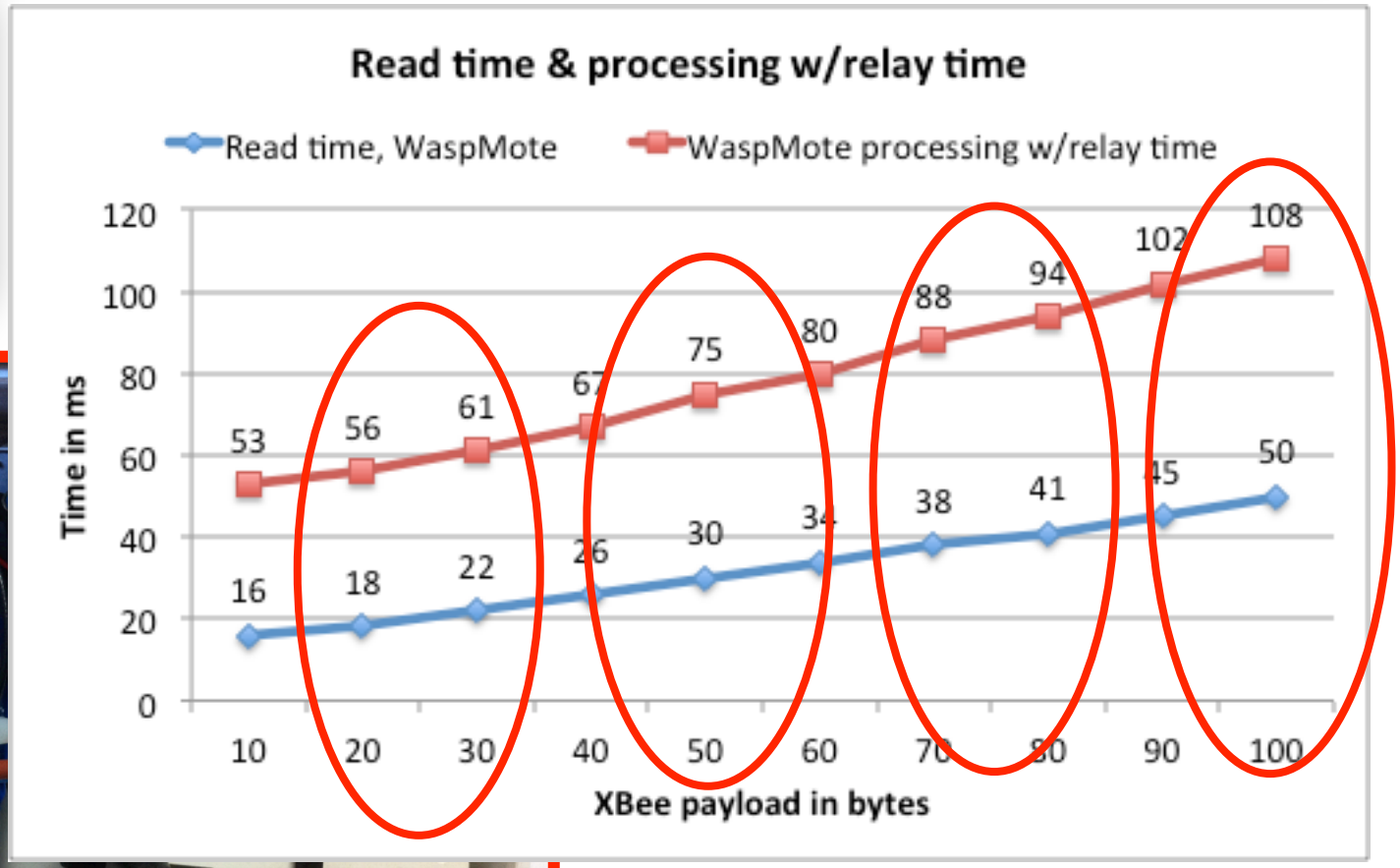


Relay performances



- Relaying are usually done at application-level (even OS level is considered app-level for the mote)
- Relaying means:
  - Read the packet in memory
  - Send the packet to next hop

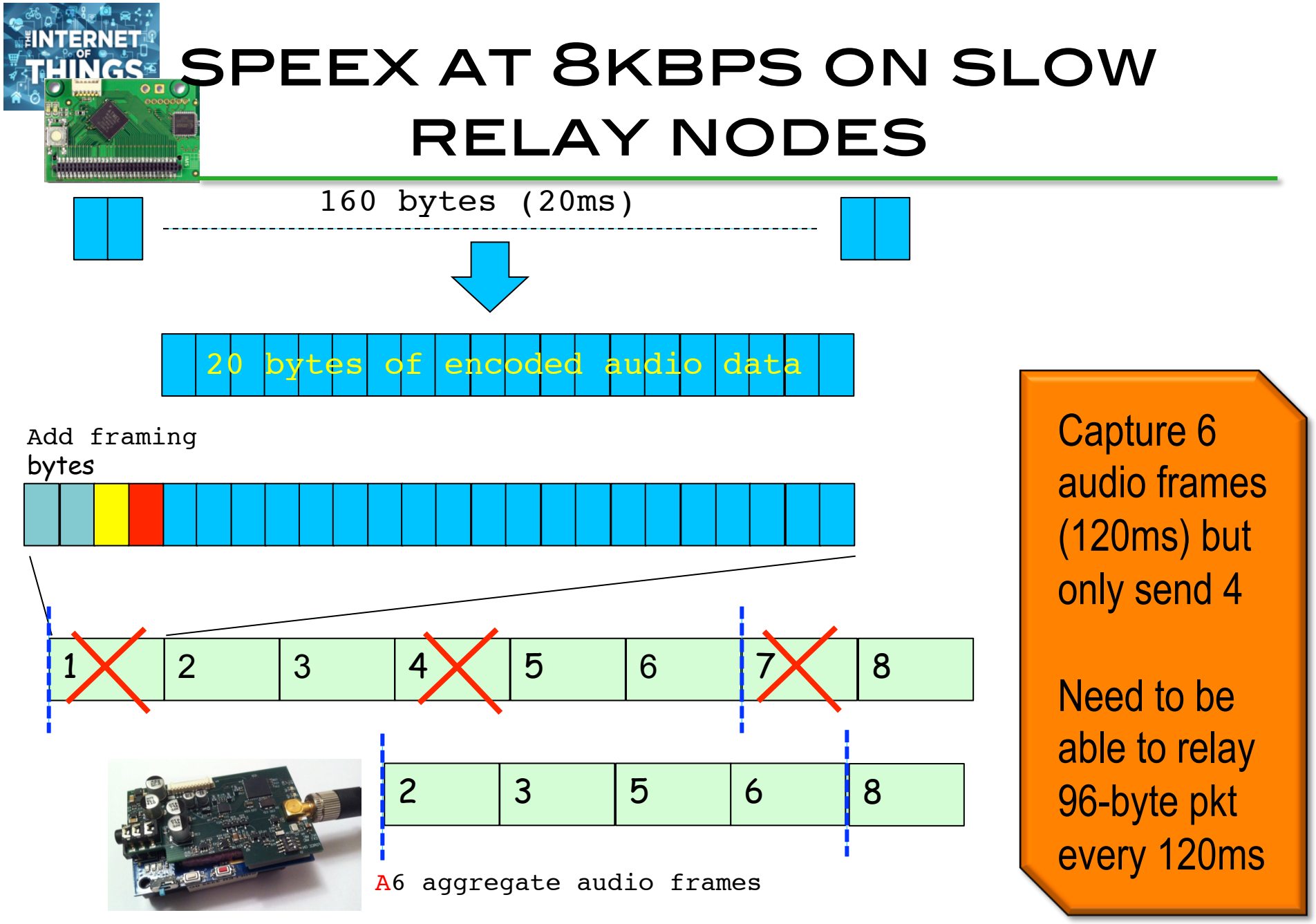
SANTANDER'S LIMITATIONS



SmartSantander's IoT node uses 38400 baud rate for communication between XBee radio and host ucontroller

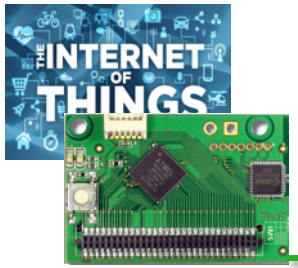
Needs to discard audio frame at the source to increase the time window

SPEEX AT 8KBPS ON SLOW RELAY NODES



Capture 6 audio frames (120ms) but only send 4

Need to be able to relay 96-byte pkt every 120ms



Conclusions

- Sensor Networks can provide large scale awareness to setup the foundation for ambient intelligence to offer new services for smart societies
- Hot topics are multimedia information for enhanced situation-awareness
- Dealing with images and audio is challenging and needs careful design and studies
- Testbed & real experimentations are needed to highlight realistic issues