Experimental Performance Evaluation of NB-IoT

Subho Shankar Basu\textsuperscript{1}, Ashish Kumar Sultania\textsuperscript{2}, Jeroen Famey\textsuperscript{2}, Jeroen Hoebeke\textsuperscript{1}

\textsuperscript{1}IDLab, Ghent University – imec, Belgium
\texttt{\{subho.basu, jeroen.hoebeke\}@ugent.be}

\textsuperscript{2}IDLab, University of Antwerp – imec, Belgium
\texttt{\{ashishkumar.sultania, jeroen.famey\}@uantwerpen.be}

Abstract—Narrowband Internet of Things (NB-IoT) is gaining prominence as a key Low Power Wide Area Network (LPWAN) technology for IoT applications. Since it operates on licensed frequency spectrum it can provide guarantees to applications demanding Quality of Service (QoS). NB-IoT has emerged as a competitive rival for other LPWAN technologies such as LoRa and Sigfox, which work in the unlicensed frequency spectrum and are vulnerable to interference. Therefore, NB-IoT is the trivial fit for industries and other business companies that demand guaranteed services. In this paper the different features of the NB-IoT technology have been studied on the commercial Orange network in Belgium using the ublox SARA-N210 module [1] as the user equipment (UE). We focused on the device and network performance in terms of setup times, signal quality, throughput, latency, and reliability and studied the network dynamicity on signal strength. These observations are then compared with the theoretical defined limits of NB-IoT.

I. INTRODUCTION

The future is moving towards a smart world where almost every day-to-day usable object will be connected and communicable leading to the concept of Internet of Things (IoT). Use case applications include smart metering (electricity, gas and water), smart cities, smart agriculture, e-health and many more. Different applications have different priorities for the technological parameters as per their requirements. Amongst those, the most common ones are scalability, throughput, latency, reliability, coverage, power consumption, security and cost. For example, smart metering applications prioritize coverage and lower power consumption rather than throughput, latency or reliability. On the other hand, e-health applications require high reliability, moderate throughput and minimal latencies to function in an emergency situation. Different communication technologies have different advantages and disadvantages, therefore it is difficult for a single communication technology to serve all these use cases requirements completely at the same time. Hence, the appropriate technology needs to be identified for a corresponding use case. Focusing on long range communication, there are different LPWAN technologies available today for use such as LoRa, Sigfox, e-MTC, and NB-IoT. In general, the telecommunication companies propose NB-IoT as a better candidate over other LPWAN technologies [2]. This paper focuses on the performance metrics of NB-IoT mainly in terms of uplink and downlink latencies, throughput, signal quality and how the network dynamically adapts to it, and device setup times.

The remainder of the paper is organized as follows. Section II gives an overview of NB-IoT and its working principles, followed by the theoretical performance limits derived from these principles in section III. Section IV refers to related works on NB-IoT, whereas section V explains the test setup and section VI presents the detailed evaluation of NB-IoT based on various metrics and tries to cross-refer the obtained results with the theoretical values of NB-IoT with possible analysis of the matches or deviations. Finally, section VII concludes the paper.

II. WORKING PRINCIPLE OF NB-IoT

Narrowband IoT (NB-IoT) is a comparatively new technology for the Cellular Internet of Things (CIoT), targeting Low Power Wide Area Network (LPWAN) solutions. As the name says it works in a narrow channel bandwidth of just 180 KHz (1 resource block of an LTE channel, with 12 sub-carriers and each sub-carrier being separated by 15 KHz). It is introduced in the 3GPP cellular standards of Release 13 onwards and is a new branch out of LTE. Functioning in the licensed spectrum, it is less susceptible to channel interference and thus able to deliver high reliability. NB-IoT is not to be confused with the minimal architecture of LTE, as legacy LTE requires a minimum channel bandwidth of 1.4 MHz which consists of 6 Physical Resource Blocks (PRBs). Working on a single PRB (180 kHz) with much reduced complexity and power consumption, many of the legacy LTE control channels and signals (NPSS, NSSS, NPDCCH, NPDSCH, NPUSCH, NPRACH, etc.) have been redefined to fit into a single PRB of NB-IoT. So, essentially when a User Equipment (UE) attaches to a network, it searches for an NB-IoT channel assigned by the network operator. The network also differentiates between the NB-IoT UE and other legacy LTE UEs as they have separate architectures for PRB and treat the corresponding signals in different ways. But even though NB-IoT has a re-designed architecture, the essential elements which constitute the slot, subframe, radio frame, resource elements (REs) and their timing and scheduling remain the same. Therefore, NB-IoT can re-use the raw essentials of legacy LTE by updating the RAN at the eNodeBs. It uses the same Orthogonal Frequency Division Multiple Access (OFDMA) and Single Carrier Frequency Division Multiple Access (SC-FDMA) principles as LTE for downlink and uplink communication respectively, making it easy to co-exist with legacy LTE systems. NB-IoT reuses the basic features of LTE such as robust authentication, encryption and other security mechanisms and thus can be targeted towards mission critical IoT applications. NB-IoT can be deployed in three modes namely Guard-band (uses the LTE guard band in between adjacent channels), In-band (uses one resource block within an allocated LTE channel), and Stand-alone (uses some other frequency band from the
them have been taken up in Release 14 for Cat NB2.

connected mode, localization and VoLTE, although some of
that are not supported by NB-IoT Cat NB1 are roaming in
to power optimization and bandwidth constraints, the features
trade off between availability and energy conservation. Due
wide range of setup parameters which enables applications to
the network, NB-IoT has the possibility to choose from a
of the device of how long it needs to remain connected to
more energy. Thus depending on the application and the need
just after finishing its transmission, hence conserving even
( RA1 ) which allows the device to go to sleep even faster
of LTE also help NB-IoT devices to sleep for longer periods
of time or enter deep sleep to conserve more energy. LTE
Also introduced the concept of Release Assistance Indication
(RAI) which allows the device to go to sleep even faster
just after finishing its transmission, hence conserving even
more energy. Thus depending on the application and the need
of the device of how long it needs to remain connected to
the network, NB-IoT has the possibility to choose from a
wide range of setup parameters which enables applications to
trade off between availability and energy conservation. Due
to power optimization and bandwidth constraints, the features
that are not supported by NB-IoT Cat NB1 are roaming in
connected mode, localization and VoLTE, although some of
them have been taken up in Release 14 for Cat NB2.

III. NB-IoT THEORETICAL LIMITS FOR THROUGHPUT
AND LATENCY

This section details the theoretical concepts and provides
numerical figures on the achievable throughput and latency
on both downlink and uplink in NB-IoT. NB-IoT is a
centralized system where the entire control of the scheduling
takes place at the eNodeB. The eNodeB instructs the UE
the specific time and frequency when it should transmit the
message on the uplink or expect a downlink message from
the eNodeB. Based on the scheduling mechanism, one can
determine the theoretical limits of throughput and latency.
Among the central parameters that affect the latency and
throughput are the link quality leading to the use of a
particular MCS, the packet size in the MAC layer which
determines the Transmission Block Size (TBS) (and the
number of TBSs required for the transmission of a single
packet), the number of repetitions used, the deployment
mode, etc. All of these parameters are network configurable
and thus may vary from one operator to another. The normal
way how the network functions is as follows and shown in
Figure 1.

If a UE intends to transmit a packet, it sends a NPRACH
request to the eNodeB, the eNodeB receives the request
and responds the UE with a scheduling grant on the NPDCCH
channel mentioning the details of the time and the
frequency when the UE can transmit. The UE then transmits
an uplink packet on NPUSCH and the eNodeB responds
with an ACK/NACK on NPDCCH. For the downlink, the
eNodeB first pages the UE and this phase depends on
multiple network configurations involving eDRX, PSM, etc.
The eNodeB then sends a Downlink Control Information
(DCI) on NPDCCH which carries the scheduling information
indicating the time and frequency of NPDSCH. The eNodeB
then transmits the data on the NPDSCH and finally the UE
responds with an ACK/NACK on the NPUSCH channel. All
these channels have different configurations for single tone
and multi-tone operation and are operator controlled.

NB-IoT defines MCS values from 0 to 12 which a UE
or the network chooses from based on the Received Signal
Strength Indicator (RSSI) at a certain instant of time. NB-
IoT also has the notion of Resource Unit (RU) as a basic
unit for NPUSCH allocation, and this has 2 different formats
with multiple possible configurations within each format.
The format referred to in this paper is format 1 operating
at 15 kHz on single tone for normal uplink data, and format
2 operating at 15 kHz on single tone for ACK/NACK. Hence
a RU in format 1 takes 8 ms while a RU in format 2 requires
2 ms [3].

A. Throughput bounds

The TBS tables [4] for both uplink and downlink map
the MCS and the RUs to determine the number of bits
(TBS) that can be transmitted as a single block in the MAC
layer. From the tables the throughput at which the bits are
sent can be calculated as the ratio of the TBS to the time
taken by the corresponding number of RUs. In case of
uplink, the maximum throughput with MCS 10 (MCS 11
and 12 are not supported for uplink) evaluates to (1000 /
6 / 8) bits/millisecond = 20.8 kbps on a single-tone link.
Correspondingly, the minimum throughput evaluates to 3.2
kbps at MCS 0. For downlink, the maximum throughput
evaluates to (680 / 3 / 1) bits/millisecond = 226.7 kbps and
the minimum to 25.6 kbps on a multi-tone link. However,
we consider the case of a single tone link for downlink,
the theoretical maximum expected throughput for which
evaluates to (226.7 / 12) kbps = 18.9 kbps.

B. Latency bounds

The latency bounds of the network can be derived from
the scheduling information as shown in Figure 1. Essentially,
the latency of a packet transmission depends on multiple
factors. Firstly it depends on the packet size; higher the
packet size, higher the transmission time due to the higher
number of TBSs used to transmit the packet. Next, it depends
on the TBS which is dependent on the link quality at the
time of transmission. Depending on the link quality, the
UE chooses the appropriate MCS and then a corresponding
TBS is chosen from the possible ones within that MCS
from the table. Based on the chosen TBS the corresponding
required number of RUs can be found out and since the
transmission time for a RU is known, the latency can be
calculated. Finally the latency also depends on the number
of repetitions being used for the transmission, NB-IoT uses
up to 128 repetitions for uplink and 2048 for downlink,
the differences being due to the unbalanced link budgets in uplink and downlink [4]. Lastly, referring to the scheduling information, the net latency for a packet transmission can be found. For a single tone uplink transmission using 15 kHz, the minimum possible latency is determined by the transmission of a single RU with no repetition at the best signal quality. Thus it evaluates to \((1 + 8 + 8 + 3 + 1)\) ms = 21 ms, since 1 RU takes 8 ms. The maximum latency is thus calculated as the transmission time of 10 RUs (maximum possible in the TBS table) with the maximum number of repetitions at the worst signal quality. It evaluates to \((2048 + 65 + 10240 + 3 + 2048)\) ms = 14404 ms. Similarly for downlink multi-tone the minimum and maximum possible latencies evaluate to 20 ms and 22788 ms respectively. Apart from this, the latency of a packet transmission also depends on the network deployment mode (in/guard band, standalone) as they have different link budgets hence leading to varying number of repetitions. Also it depends on the eDRX and PSM modes [5], which can be configured on the UE as a measure of conserving power.

IV. Related work

Previous works have studied the capabilities of NB-IoT and its performance. The authors in [6] have evaluated the extended coverage of NB-IoT with the help of a simulation model in different deployment modes, showing that NB-IoT can reach a Maximum Coupling Loss (MCL) of 164 dB which is 20 dB more than legacy LTE. Authors in [7] have studied the effects of channel estimation quality and coherence time of the channel on the coverage of NB-IoT from an analytical perspective. They have also validated the results with a testbed implementation using Software Defined Radio (SDR), but this mostly highlights the physical layer performance with a subset of the NB-IoT MAC and RRC functionalities. [8] studies LTE-CatM and NB-IoT on a comparative basis and shows how NB-IoT provides an even deeper indoor coverage compared to LTE-CatM. It involves a simulated approach based on values calibrated from a deployed network and shows how the network behaves and adapts itself to different MCL values as the device moves further away from the eNodeB. The step pattern of the latency and throughput observed also matches the pattern from our tests and validates the fact about the dynamic network adaptation. [9] performs a comparative study of the different available technologies for a smart-grid application and shows from a simulated approach how NB-IoT suits to the needs of such an application. [10] studies the effect of interference on a partially deployed NB-IoT network in in-band mode and addresses possible solutions to mitigate them. [11] extends the LTE module of ns-3 to implement the NB-IoT architecture and studies the enhanced network coverage due to repetitions, but also highlights the trade-off with reduced system efficiency as compared to legacy LTE.

However, most of these works are either analytical or simulation based following some mathematical model and based on certain assumptions of the different parameters. In contrast, this paper focuses on evaluating the performance of NB-IoT in a real deployed network with real devices and validates it against the theoretical values. The tests also involve using an attenuator to study the effects of signal quality on the performance metrics. To the best of our knowledge, it is one of the initial efforts to evaluate the performance of NB-IoT in a real deployed network.

V. Setup environment

To perform the tests, the NB-IoT network from Orange, Belgium has been used. So all results presented represent the performance of the Orange network and might vary for other operators, as different operators have different algorithms for channel scheduling, link quality adaptation, etc. The test setup is shown in Figure 2.

Our setup consists of a client-server implementation with the client application written in Python running on a laptop and the server application as a Click modular router implementation running on a Linux server. Both the client and the server are time synchronized with a Network Time Protocol (NTP) server. The OCTA-Connect module with the embedded Sara N210 modem from ublox and an NB-IoT enabled SIM card from Orange has been used as the user equipment (UE) device, connected to the laptop over a USB to TTL serial cable and controlled using AT commands. For the tests, the device initially sends a configuration which determines the test conditions involving uplink and downlink intervals, packet sizes, total test time, reboot and registration parameters, etc. which initializes the test setup before the actual test execution. The Click implementation on the server reads this configuration and sets it up for the test accordingly. Additionally, a quadAtten RF attenuator from octoScope has been used to attenuate the signals between the device and the eNodeB to test different link quality parameters.

VI. Evaluation

This section studies the network and the device behaviour for NB-IoT on a real deployed network. The network performance measurements involve multiple parameters that
include link quality, latency, throughput, packet size, inter-packet intervals, packet error rates, etc. and their dependencies on one another. The device side performance studies the capabilities and limitations of the SARA-N210 module involving buffer size, handling simultaneous uplink and downlink with a single RF chain, device initialization times, getting and setting different attributes and timers, etc. We discuss more about these in the next subsections.

A. Link quality with attenuation

First, we study the link quality for increasing levels of attenuation from 0 dB up to the value at which the device loses connection with the eNodeB. The RSSI in dBm, and Signal to Noise Ratio (SNR) in dB have been plotted against the Reference Signal Received Power (RSRP) in dBm as shown in the Figure 3.

It can be seen that the device just loses connection at an RSRP of -141 dBm. As the device transmission power is 23 dBm, the MCL comes out to be (23 - (-141)) dBm = 164 dBm, which indeed is the theoretical value for NB-IoT (20 dB more as compared to GPRS with an MCL of 144 dB). The RSSI value corresponding to an RSRP of -113 dBm and below becomes static. This is a device limitation of the SARA-N210 module which cannot measure RSSI values below that threshold. However, a gradual increase in the RSSI value is noticeable with decreasing attenuation. The same characteristics have been noticed for SNR as well although we see some small peaks and falls in between. This is due to the fact that with decreases in the signal strength, the eNodeB adapts the transmission power with a certain threshold, for which the SNR has a sudden peak and then falls back again with increasing attenuation. So the normal characteristics that has been mentioned in the NB-IoT specification can be verified as well from this plot.

B. Latency with link quality

In this subsection we evaluate the effects of the link quality on latency with different packet sizes. The one-way latency has been measured as the time difference between the packet leaving a host and arriving at the other host, for both uplink and downlink. The latencies against increasing attenuation for uplink and downlink with packet sizes of 8, 64 and 512 bytes are shown in Figures 4 and 5 respectively.

As seen from the figures, three zones can be identified where significant changes are observable for both the uplink and downlink latencies. This is due to the same fact that, when reducing signal strengths, the network adapts itself and assigns lower and hence slower Modulation Coding Scheme (MCS) values to the device which results in an increase in latency for increased attenuation. In Zone 1, with high RSSI signal strengths, the latencies for both uplink and downlink are around 300 ms and around 1 second for RSRP values around -110 dBm. Then it goes up to 5 seconds in Zone 2 with RSRP values around -135 dBm, and in Zone 3 with very poor signal strength on the edge of the network reach it goes up to 15 or 10 seconds in uplink and downlink respectively depending on the MCS and TBS used at that signal strength, repetitions used, etc. in the network. In general it is noticed that the uplink latency is slightly higher than the downlink latency. This also conforms to the fact that NB-IoT uses OFDMA for downlink and SC-FDMA for uplink. As OFDMA is faster than SC-FDMA, the latencies for uplink are higher than the latencies for downlink under the same environmental conditions. It is also important to note that none of the packets were lost even in low signal strength scenarios which testifies the strong reliability of NB-IoT. A correlation is observable between the link quality and the latency, which verifies that the network adapts itself to lower the MCS value to send data slower and also applies more repetitions at low signal strengths.

C. Latency with varying packet sizes

NB-IoT functions in a way that the data symbols are mapped to Resource Elements (REs) and a set of REs form a Resource Block (RB). Depending on the signal strength, the operator assigns the MCS and the number of REs used for forming a block that can be used to transmit some data as a unit chunk of bits and this is known as the Transmission
Block. So with varying signal strengths, the TBS (in bits) varies and the operator takes the values from a mapping table as defined in the specification. So depending on the signal strength and the packet size that needs to be transmitted, there can be differences in performance in the network. Figures 6 and 7 show the latency against varying packet sizes for three different signal strengths for uplink and downlink respectively.

As is noticeable from the figures, the uplink latencies are in general greater than the downlink latencies because of NB-IoT using SC-FDMA in uplink and OFDMA in downlink. A significant increase in latency is visible for both 30 dB and 60 dB attenuation as compared to zero attenuation for both uplink and downlink. This shows that the channel conditions are quite stable for good signal strengths and as the device moves away from the eNodeB it switches to a lower MCS value with higher number of repetitions leading to a higher latency.

However, in all the tests performed, there are certain deviations from the expected values. This has been noted multiple times and it is presumed that the real time network performance changes during different times of the day. These random fluctuations might be network dependent or planned network downtime or maintenance during off-peak hours.

D. Throughput with link quality

As the latencies of the packets vary with varying link quality, so does the throughput. This section studies the achievable application throughput for both uplink and downlink for NB-IoT. The results are shown in Figures 8 and 9 respectively. The tests were performed so as to keep the buffer of the devices always non-empty, i.e. having some packets to transmit, but not overflowing it. All packets were noticed to be received on the receiver end, which again shows the high reliability provided by NB-IoT.

Theoretically, NB-IoT has a peak throughput of 20 kbps for single tone mode. The obtained results show the values fall within this range with the uplink throughput reaching a maximum of 11 kbits/sec and downlink 17 kbps under good signal strength conditions. The three zones, as have been seen during the latency tests, are also visible here due to the network changing the MCS and repetitions. Downlink latencies are lower compared to uplink latencies. Therefore, the achievable downlink throughput is higher than the uplink throughput.

E. Throughput with varying packet sizes

As the throughput has been measured considering only the application layer data, it would make more sense to send packets of larger payload for more efficient data transmission and hence to obtain maximum throughput. As the SARA-N210 module supports a maximum of 512 bytes of application payload that can be transmitted in a UDP packet, tests were performed from application payload sizes from 8 bytes up to 512 bytes. This is a device constraint and not a network constraint. However, it limits our test cases. The tests were performed with the device buffer having some packets to transmit every time but not overflowing it. All packets were noticed to have been received on the receiver end. Figure 10 shows the throughput characteristics with varying packet sizes for both uplink and downlink.

The figure shows the steady increase in the throughput for increasing packet sizes for both uplink and downlink. The difference in the throughput between them is more distinguishable at higher packet sizes because of the fact that a single RU in uplink requires 8 ms while a subframe on downlink requires 1 ms.
TABLE FOR DEVICE AND NETWORK SETUP PARAMETERS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baud rate 115200</th>
<th>Baud rate 115200 (device initialized)</th>
<th>Baud rate 9600</th>
<th>Baud rate 9600 (device initialized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialize argument</td>
<td>0.0027</td>
<td>0.0028</td>
<td>0.0035</td>
<td>0.0024</td>
</tr>
<tr>
<td>Open serial port</td>
<td>0.0036</td>
<td>0.0038</td>
<td>0.0041</td>
<td>0.0037</td>
</tr>
<tr>
<td>Reboot Module</td>
<td>4.0055</td>
<td>n.a.</td>
<td>4.0548</td>
<td>n.a.</td>
</tr>
<tr>
<td>Turn on RF circuit</td>
<td>1.6126</td>
<td>n.a.</td>
<td>1.5814</td>
<td>n.a.</td>
</tr>
<tr>
<td>Set different UE parameters</td>
<td>0.0637</td>
<td>n.a.</td>
<td>0.1598</td>
<td>n.a.</td>
</tr>
<tr>
<td>Module Initialize</td>
<td>5.6818</td>
<td>0.0066</td>
<td>5.80</td>
<td>0.0061</td>
</tr>
<tr>
<td>Network Registration wait</td>
<td>16.4457</td>
<td>0.0091</td>
<td>15</td>
<td>0.0450</td>
</tr>
<tr>
<td>Get different configuration parameters</td>
<td>0.1221</td>
<td>0.1232</td>
<td>0.5254</td>
<td>0.5085</td>
</tr>
<tr>
<td>Total Initialize SARA module</td>
<td>22.4118</td>
<td>0.2939</td>
<td>22.2268</td>
<td>1.1648</td>
</tr>
</tbody>
</table>

F. Connection setup and initialization parameters

The last test performed, aims to study and analyze how long it takes for a device to become operational, from the initialization of the module, over connecting to the network and getting or setting other device and network parameters for different baud rate settings. The results are shown in Table I. These numbers will help the application running on the device to understand and make use of the setup latencies from the device and the network side.

The tests have been performed with different baud rates to see if it has a major impact on the device setup times. In general it is observable that for a lower baud rate, the setup times are slightly slower although it does not have a significant effect on the total time for the device which is around 22 seconds for a fresh reboot until it can start sending or receiving packets to/from the network.

VII. Conclusion

From the extensive evaluations that have been done for NB-IoT in this paper, it turns out that the technology is quite robust in terms of availability (licensed spectrum without interference), coverage (both outdoor and deep indoor), throughput (more than 10kbps for both uplink and downlink), high reliability (no packet loss even in low signal quality) and acceptable latencies (depending on signal strength, but less than a second for moderate scenarios). On top, considering the scalability (’50K devices per cell), security (cellular LTE based), low power consumption (’10 years of battery life), low cost of devices (half-duplex and simpler LTE architecture with 1 resource block), low deployment cost of the network (update RAN of legacy LTE), and ease of use (plug-and-play), it can be well said that NB-IoT is indeed a good competitor of other LPWAN technologies that operate in unlicensed spectrum and offers IoT application providers a broader choice.

VIII. Acknowledgment

This research was funded by the ICON project MAGICiaN in collaboration with imec and with project support from VLAIO and Innoviris. Project partners are imec, Orange Belgium, Televic Healthcare, Restore and Citymesh. This work has been partly funded by the IDEAL-IoT (Intelligent DEnse And Longe range IoT networks) SBO project, funded by the FWO-V (Fund for Scientific Research-Flanders) under grant agreement S004017N.

REFERENCES