Scaling Dense NB-IoT Networks to the Max:
Performance Benefits of Early Data Transmission

Pascal Jörke, Tim Gebauer, Stefan Böcker and Christian Wietfeld
Communication Networks Institute
TU Dortmund University, Germany


Abstract—The growing number of IoT devices will lead to a massive number of users in communication networks. Therefore, this paper analyzes the scalability boundaries of NB-IoT networks with different transmission modes, called standard transmissions, Cellular IoT Optimization, and Early Data Transmission, using a novel detailed implementation of NB-IoT in the NS-3 LTE simulation framework LENA. The results show that Early Data Transmission clearly outperforms NB-IoT standard transmissions and Cellular IoT Optimization by providing up to 4:1 times less latency and 1:6 times longer battery lifetime, while only using one-fourth of the downlink spectrum. Further, a good scalability for up to 864,000 devices per day in a cell with an area of 4.91 km², or 176,000 devices per day and km² for all NB-IoT standard transmission, Cellular IoT Optimization, and Early Data Transmission scenarios is given. It is shown that the scalability is limited by downlink spectrum capacity for non-Early Data Transmission scenarios and Random Access windows for all scenarios. Doubling the number of Random Access windows improves the performance in highly scaled scenarios in terms of a larger packet delivery rate and fewer Random Access collisions. Still, the number of Random Access collisions for scenarios over 1,000,000 devices per day is very high, which indicates the necessity of a detailed optimization for the radio resource configuration when using new features like Early Data Transmission for optimal cellular performance. Still, the analysis results show a great positive impact of Early Data Transmission on the overall performance and is highly recommended to be used by default.

Index Terms—Scalability, NB-IoT, EDT, Early Data Transmission, NS-3, LENA

I. INTRODUCTION

WASTE bins that alarm the waste collection when they need to be emptied, environmental sensors that transmit air quality reports frequently, parking sensors that help to navigate to free parking lots. These are only a few use cases enabled by the upcoming Internet of Things (IoT). With the digitalization of everyday life, billions of sensors and actuators will be installed, transmitting and receiving small amounts of data (Fig. 1) [1]. To enable the rollout of these new applications, current communication technologies are faced with new challenges.

Taking these new challenges into account, the 5th mobile radio generation (5G) addresses the massive IoT area in particular and promises a very high subscriber density for different environmental scenarios. For this purpose, the ITU-R defines correlating requirements for massive IoT applications in the context of the massive Machine Type Communication (mMTC) [2]. With NB-IoT, a promising technology has been derived from LTE, designed for small and efficient data transmissions. With a typical bandwidth of 180 kHz, the goal of providing cell access to 1,000,000 devices per km² is very challenging. Besides enabling communication for a massive number of devices, NB-IoT is also designed for a long battery lifetime, deep indoor penetration for flexible device placement as well as a latency not exceeding 10 s.

Since NB-IoT is designed for small data transmissions, new features are added with every new release. With Cellular IoT Optimization (C-IoT Opt) and Early Data Transmission (EDT), new techniques have been added. Both features reduce the overall signaling overhead of single transmissions, which results in lower latency, power consumption, and spectral usage. Latter enables an improved user capacity per cell. Since C-IoT Opt only enables piggybacking data, but still requires a full NB-IoT Radio Resource Control (RRC) connection resume procedure, the improvement will be limited. With EDT, data can be transmitted without an active connection, drastically reducing signaling overhead.

For a detailed look into the scalability of NB-IoT networks with these different transmission modes, this paper is organized as follows: Section II briefly outlines previous works on EDT, NB-IoT scalability evaluations and simulation tools, while section III gives a short overview of techniques for reduced signaling overhead and more efficient transmission in terms of energy and frequency resources. Section IV introduces a novel, detailed NS-3 LENA extension for NB-IoT, which is used in this work. It is followed by an analysis of the NB-IoT performance in high-scaled networks in section V and finally, the results are concluded in section VI.

Fig. 1: Worldwide IoT growth leads to great challenges for all communication technologies [1]
II. RELATED WORK

As one of NB-IoT Release 15 main features, EDT was first proposed by Ericsson in [3] and later evaluated by [4] using an analytic framework. While the impact on the individual device performance is addressed, the results miss an analysis of network performance in future high-scaled IoT networks, which is discussed in this proposed work. Since EDT is a new feature, work related to EDT is very limited for now. Other related works address the scalability of NB-IoT networks, but are limited to the initial 3GPP Release 13.

In [5], the authors analyzed the performance of enhanced Machine Type Communication (eMTC) and NB-IoT in smart city applications, especially in the context of battery lifetime and power-saving features. Further [6] compares NB-IoT in anchor and non-anchor mode with eMTC for high connection density scenarios. [7] extensively compares NB-IoT user performance to other Low Power Wide Area Network (LPWAN) technologies like LoRaWAN, but does not take scalability into account. To analyse this, multiple attempts are present in the context of existing NB-IoT implementation for NS-3 LENA. First, there is an NB-IoT implementation draft linked to the official NS-3 repository at [8], but it does not go beyond the implementation of broadcast signals such as MIB-NB and SIB1-NB. Next, a relatively high-level approach for evaluating NB-IoTs energy consumption was implemented in [9]. A working Random Access (RA) implementation was evaluated in [10]. Although this work calculates the RA windows correctly, it does not consider the transmitted System Information Block 2 - Narrowband (SIB2-NB) and does not allow flexible configuration of the User Equipment (UE) Medium Access Control (MAC) layer. Further, the selected parameter set is not referenced or explained.

In the subsequent attempt [11], the authors implemented basic Narrowband Internet of Things (NB-IoT) scheduling capabilities into the evolved Node B (eNB) MAC layer, including Narrowband Physical Downlink Control Channel (NPDCCH), Narrowband Physical Downlink Shared Channel (NPDSCH), and Narrowband Physical Uplink Shared Channel (NPUSCH). It is essential to mention that this scheduling only logs its results and does not actually coordinate the transmission between the simulated devices.

III. FUNDAMENTALS OF NB-IoT

With 3GPP Release 13, NB-IoT has been introduced as a promising cellular IoT solution. With new power-saving techniques, such as extended Discontinuous Reception (eDRX) and Power Saving Mode (PSM) as well as small data transmission optimizations (Cellular-IoT Optimization, C-IoT Opt), NB-IoT is designed to enable cellular connections to battery-powered IoT devices. For flexible device placement, even in basements, NB-IoT also integrates coverage improvements. Besides a small transmission bandwidth, Coverage Enhancement (CE) is provided by using up to 2048 transmission repetitions for an indoor and deep indoor coverage [12] (ref. Fig. 2). While these repetitions enable a Maximum Coupling Loss of 164 dB, repetitions are redundant data and lower the overall network capacity as well as the device’s energy efficiency [13]. Therefore, NB-IoT includes features for a reduced signaling overhead detailed in the following sections.

A. Control Plane C-IoT EPS Optimization

In standard NB-IoT transmissions, a connection on the RRC layer has to be established, before user data can be transmitted (Fig. 3a). While this RRC Connection Resume procedure remains mandatory, with Control Plane C-IoT Optimization data can be piggybacked as a Non-Access Stratum (NAS) dedicated information transfer in the RRCConnectionSetup message (Msg4) in DL direction and the RRCConnectionSetupComplete message (Msg5) in UL direction [14]. In Fig. 3b it is shown that in this case a CoAP PUT uplink message with its mandatory NPDCCH message for the Downlink Control Information (DCI) transmission is included in Msg5 and thus reduces the overall number of messages.

B. Early Data Transmission

In Release 15, a new transmission mode is introduced. With Early Data Transmission (EDT), data now can be piggybacked on Msg3. While Msg3 was originally the RRC-ConnectionResumeRequest message, the UE transmits an EarlyDataRequest message as a Msg3 in EDT, which includes the piggybacked user data and can be as large as 1000 bits [15]. After a successful reception by the base station / evolved Node B (eNB), the Random Access (RA) procedure is terminated with an EarlyDataComplete message, which again can carry DL user data, such as an application acknowledgment, and therefore drastically decreases the message overhead of a user data transmission (Fig. 3c). When the eNB rejects the EarlyDataRequest message, the UE falls back to the standard RRC Connection Resume procedure, is moved to connected mode, and then transmits its user data.

C. Random Access Procedure

While most NB-IoT signaling is scheduled and therefore collision-free, all devices have to send a Random Access preamble at first in predefined Random Access windows in the Narrowband Physical Random Access Channel (NPRACH). In NB-IoT, for each CE class, distinct Random Access windows are defined, considering different channel conditions of end devices. For this work, the network configuration, including the Random Access configuration, is derived from live NB-IoT networks in Dortmund, Germany, using an Ettus USRP B210 SDR with srsRAN [16] and is given in Table I.
The NB-IoT Radio Resource configuration has to be well-balanced between RA windows and available spectrum for uplink data transmissions. An insufficient number of RA windows leads to high congestion in the remaining RA occasions, ending in many collisions and only a few successful devices being able to resume their RRC connection and transmit data. In this case, the RA windows are overloaded, while the rest of the available spectrum is unused. On the other hand, an oversized RA allocates more spectrum, which then is unavailable for data transmissions, but may be unused. Additionally, in highly scaled networks too many devices will establish a connection to the network and will overload the available downlink (DL) and uplink (UL) spectrum. Using the configuration in Table I, 4.5% of the uplink spectrum is allocated to RA windows and is therefore not available for scheduled data transmissions (Fig. 4). Later the results will show if this configuration is sufficient.

Besides the RA in UL, in the DL spectrum, broadcasts allocate spectrum as well, which is not available for DL transmissions. In NB-IoT networks, the following signals need to be considered:

- Narrowband Primary Synchronization Signal (NPSS)
- Narrowband Secondary Synchronization Signal (NSSS)
- Master Information Block (MIB)
- System Information Block 1 - Narrowband (SIB1-NB)

The performance comparison of IoT application protocols in future NB-IoT is of great importance since it especially the SIB2-NB is of great importance since it carries the Radio Resource configurations (ref. Tab. I) of the

<table>
<thead>
<tr>
<th>Parameter</th>
<th>CE0</th>
<th>CE1</th>
<th>CE2</th>
</tr>
</thead>
<tbody>
<tr>
<td>rtp-Thresholds</td>
<td>-</td>
<td>-116 dBm</td>
<td>-128 dBm</td>
</tr>
<tr>
<td>nprrach-InfoList-r13</td>
<td>ms250</td>
<td>ms40</td>
<td>ms2560</td>
</tr>
<tr>
<td>nprrach-StartTime-r13</td>
<td>ms256</td>
<td>ms256</td>
<td>ms256</td>
</tr>
<tr>
<td>nprrach-SubcarrierOffset-r13</td>
<td>n36</td>
<td>n24</td>
<td>n12</td>
</tr>
<tr>
<td>nprrach-NumSubcarriers-r13</td>
<td>n12</td>
<td>n12</td>
<td>n12</td>
</tr>
<tr>
<td>nprrach-SubcarrierMSG3</td>
<td>twoThird</td>
<td>twoThird</td>
<td>twoThird</td>
</tr>
<tr>
<td>RangeStart-r13</td>
<td>n10</td>
<td>n10</td>
<td>n10</td>
</tr>
<tr>
<td>nRepRepetitionsPer-PreambleAttempt-r13</td>
<td>n1</td>
<td>n8</td>
<td>n32</td>
</tr>
<tr>
<td>npdch-NumRepetitions-RA-r13</td>
<td>r8</td>
<td>r64</td>
<td>r512</td>
</tr>
<tr>
<td>npdch-StartSF-CSS-RA-r13</td>
<td>v2</td>
<td>v1dot5</td>
<td>v4</td>
</tr>
<tr>
<td>npdch-Offset-RA-r13</td>
<td>zero</td>
<td>zero</td>
<td>zero</td>
</tr>
</tbody>
</table>

Fig. 3: Control message overhead for different types of NB-IoT data transmissions

Fig. 4: Allocated NB-IoT uplink spectrum

Fig. 5: Allocated NB-IoT downlink spectrum

IV. INTRODUCING NS-3 LENA-NB

To perform a fine-grained scalability analysis of NB-IoT and its signaling features, we implemented an NB-IoT simulator based on the NS-3 LTE framework LENA. Due to the official NS-3 implementation of NB-IoT, having only the MIB-NB and SIB1-NB on the eNB side implemented yet, LENA-NB was completely new implemented using LENA as a base. Fig. 6 gives an overview of the different modifications made for LENA-NB. First, the adapted transmission of control information of the MIB-NB and the SIB-NBs were implemented, making it possible to simulate the acquisition of NB-IoT-specific control information and the decision process at the UE. Especially the SIB2-NB is of great importance since it carries the Radio Resource configurations (ref. Tab. I) of the

- Additional System Information (SI) messages

These broadcast messages allocate approx. 30% of the downlink spectrum without any data transmission, leaving only 70% for control information and data transfer (Fig. 5). Compared to the 4.5% allocated spectrum in UL direction, the downlink can be considered a bottleneck in terms of spectrum. Therefore, this work will focus on spectrum usage in DL direction.
individual CE levels and EDT parameters. The simulated UE then calculates its corresponding CE level based on the received SIB2-NB and the measured Reference Signal Received Power (RSRP). Furthermore, a detailed random access procedure was implemented. The simulated UE uses the previously determined information to perform correct random access at the right random access occasion. The eNB was extended to recognize the CE level of the UE based on the selected random access parameters and to perform collision detection. At the moment, we assume that the collisions end in destructive interference and that none of the colliding UEs completes the random access successfully. Another major part of LENA-NB involves a proper NB-IoT cross-subframe scheduler, which has been newly implemented. Due to the low-cost devices in NB-IoT, certain intervals must be maintained between, for example, UL and DL messages to give the UE enough time to switch the transceiver. Accordingly, an NB-IoT scheduler must schedule across multiple subframes and not just on a subframe basis. Furthermore, NB-IoT uses so-called search spaces where the UE listens for scheduling information. The implemented scheduler further uses an empirical modulation and coding scheme to use the best possible transmission configuration, e.g., subframes, Transport Block Size (TBS), repetitions, for downlink and uplink transmissions.

Fig. 6: Overview over the different modifications made in LENA for LENA-NB

A detailed energy model has been implemented in LENA-NB to simulate and analyze the much-discussed energy consumption of NB-IoT devices. For this purpose, the functionality of the simulated NB-IoT devices was first extended to include the connection resume procedure together with the eDRX and PSM modes, into which the UE also dynamically switches. With empirically measured values of the individual power states, LENA-NB can thus also simulate the energy consumption most accurately. Furthermore, the features C-IoT-Opt and EDT presented in section III were implemented to enable a scalability analysis of the upcoming NB-IoT releases.

Finally, LENA-NB contains many profiler optimizations to simulate large numbers of UEs with NS-3 LENA, which is limited to a few devices and short simulation times.

The source code to LENA-NB is available in [17].

V. EVALUATION AND RESULTS

With growing numbers of IoT devices, the communication networks face the challenges of highly scaled networks. For a detailed scalability analysis, we defined different scenarios with growing numbers of devices in different coverage conditions. All simulations are performed in NB-IoT networks with a diameter of 2.5 km, which is the average cell size in the city of Dortmund, Germany and is derived from the Dortmung mobile radio register [18]. Table II lists all relevant simulation parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation time</td>
<td>15 min</td>
</tr>
<tr>
<td>Cell diameter</td>
<td>2.5 km</td>
</tr>
<tr>
<td>Cell area</td>
<td>4.91km²</td>
</tr>
<tr>
<td>Channel Model</td>
<td>Winner+ (UMaNLOS)</td>
</tr>
<tr>
<td>Base station height</td>
<td>50 m</td>
</tr>
<tr>
<td>Device height</td>
<td>1.5 m</td>
</tr>
<tr>
<td>UDP data size (UL direction)</td>
<td>49 Bytes</td>
</tr>
<tr>
<td>Transmission interval</td>
<td>24 hours</td>
</tr>
</tbody>
</table>

TABLE II: NS-3 LENA-NB simulation parameters

While the payload data size is considered fix with 49 Bytes at UDP layer (32 Bytes 5G mMTC payload + 4 Bytes CoAP header + 13 Bytes DTLS header), the devices are distributed equally over three different coverage conditions: 1/3 of all devices are considered to be placed outdoors, while 1/3 are indoor and 1/3 are deep indoor (basement) devices. In respect of indoor and deep indoor placement, the device’s path loss is supplemented with additional building entry losses (15.4 dB for indoor and 20.9 dB for deep indoor), which are derived from [19]. Note that a payload data size, which exceeds the 1000 bits TBS limit, leads to a fallback to standard transmissions or C-IoT-Opt, and thus EDT will not be used.

For all scenarios, 15 minutes of simulation time are simulated, but only the intermediate 5 minutes are evaluated in the following. The first 5 minutes produce no significant results, since devices at the beginning are scheduled in an empty cell and experience very good transmission conditions. After 5 minutes, new devices will find ongoing transmissions of previous devices, which enables a more realistic situation and produces significant results. Since devices that have started transmissions within the intermediate 5 minutes of the simulation may not complete their transmissions in this intermediate time slot, additional 5 minutes are simulated with more new transmissions to keep the channels busy and letting the intermediate devices complete their transmissions.

In our simulations scenarios with different number of devices are simulated, ranging from 9 to 18,000 devices per cell within the 15 minutes simulation time, which results in a total of 864 devices per cell on a daily base in small scenarios.
A. User-relevant Performance Results

Fig. 7a shows the results for the average end to end latency and Packet Delivery Rate (PDR) results using the NB-IoT Radio Resource configuration given in Tab. I, which is called from now on Standard RA windows, as defined in Fig. 4. For configurations without EDT, an increasing latency can be noted starting at approx. 9,000 devices per cell. Since all devices use the same shared channels, one device needs to wait for other devices to pause or finish their transmissions, before it can be scheduled. Since the signalling overhead for devices without C-IoT Opt and EDT is larger (ref. Fig. 3), the latency for these devices rises faster than for other configurations, since more messages occupy more spectrum. While C-IoT Opt can only slightly lower the average latency, EDT clearly outperforms all other configurations with an up to 4.1 times lower latency. In high user density scenarios, the results show a minor drop in PDR just below 1.000.000 devices for scenarios without EDT and a major drop for the 1.7 mio devices scenarios for all configurations. The PDR results show, that the scalability in the given configurations is limited to approx. 500.000 devices for non-EDT scenarios and approx. 1.000.000 devices in EDT scenarios, limited by spectral resources. Taking the different control message overhead of these configurations into account (ref. Fig. 3), these results clearly meet the authors expectations.

Besides latency, the power consumption is simulated as well and is presented in Fig. 7c. While latency experiences a rather steep rise, the overall power consumption increases as well, but much less steep. When devices have to wait for their turn to be scheduled, they can remain in a low power mode to save energy. The overall high power states, especially when transmitting data in uplink direction, remain the same in amount and length, independent on the number of devices. Using a linear battery model and a 5 Wh battery, the estimated battery lifetime is presented in Fig. 7c as well. In most NB-IoT use cases latency is unimportant, since NB-IoT is usually not used for real-time applications. But power consumption and battery lifetime are significant for those use cases that use batteries as an energy source. Comparing the results of the different configurations, EDT clearly outperforms non-EDT configurations with 1.6 times the battery lifetime. From a user-view, EDT is highly recommended for an overall better NB-IoT performance.

B. Cell-relevant Performance Results

While the latency in Fig. 7a rises for all scenarios up to approx. 1.000.000 devices, the latency for non-EDT cells with 1.728.000 devices does not increase further, which indicates a cell capacity overload. The results of the PDR (ref. Fig. 7a)
confirm, that devices in the largest scenario experience massive packet drops. In the next step, we evaluated the spectrum usage in the downlink direction (Fig. 8a).

As described in section III-C the DL broadcast messages already occupy 30% of the downlink spectrum, which is not available for user data. In addition, the increasing number of devices in the simulated scenarios results in increasing spectrum usage. Since devices without C-IoT Opt and EDT provide the most signaling overhead, these scenarios also result in the highest spectrum usage of up to 90%. Note that a 100% spectrum usage is not possible, since timing offsets for device scheduling are limited to predefined values. Again, the results meet the expectations, since spectrum usage decreases with C-IoT Opt by 14% and with EDT by 73%, which indicates the massive spectral efficiency improvement possible with EDT. While transmitting the same user data, EDT uses only 1/4 of the DL spectrum. Although much DL spectrum is still available in the EDT high-density scenarios, the PDR drops in EDT scenarios with 1.728.000 devices as well as with other configurations. While the DL spectrum is the bottleneck for scalability of NB-IoT standard transmissions and NB-IoT C-IoT Opt transmissions, the reason for EDT packet drops can be found in the number of RA collisions (Fig. 8c). Since all scenarios rely on the same NPRACH configuration, the results for the RA collisions are mostly equally for standard transmissions, C-IoT Opt and EDT. Still, the results show a massive increase of RA collisions in the largest scenarios, which indicates overloaded RA windows, refusing most devices successful cell access.

To verify the RA windows as a bottleneck, the Radio Resource Configuration from Table I has been modified to double the number of RA windows. In Fig. 7b, the latency in the largest scenarios is now much higher, since more UEs perform successful channel access and therefore increase the scheduling competition in the DL spectrum. It can be noted, that the PDR is clearly increased with doubled RA windows, meaning that more devices can transmit their data successfully. Still, the PDR is low, which still indicates an overloaded spectrum. Taking the spectrum usage into account (Fig. 8b), it is confirmed that the DL spectrum is overloaded by the number of UEs. Although the number of RA windows has been doubled, the number of RA collisions in the largest scenario is still very high: Each collision of RA preambles leads to retransmissions of these preambles, which occupy more spectrum and increase the probability of RA collisions even more. Though, the number of RA collisions in scenarios just below 1.000.000 devices has been reduced by a factor of 2.4, which is a significant improvement.

All in all, the simulation results show good scalability up to 864.000 devices in a cell with an area of 4.91km², or 176.000 devices per km² for all NB-IoT standard, C-IoT Opt and EDT scenarios. While the standard and C-IoT Opt scenarios are limited by RA windows and available DL spectrum, EDT is only limited by the RA windows in the largest scenarios.
Since RA windows are configurable, we have doubled the RA windows, which improved the overall performance in scenarios with high numbers of devices. For optimal performance of each cell configuration, a Radio Resource configuration optimization has to be performed, which will provide the best Radio Resource configuration concerning different NB-IoT features such as C-IoT Opt and EDT.

VI. CONCLUSION

With the growing number of IoT devices, communication networks need to provide connections to millions of devices. With NB-IoT, a promising solution of cellular IoT networks has been introduced to face these challenges. Introduced in 3GPP Release 13, it is still an object of improvement and is optimized with new features such as Early Data Transmission, which drastically reduces the signaling overhead of small data transmissions. In this work, a simulative analysis has been performed to compare the scalability boundaries of standard NB-IoT transmissions with optimization features such as Cellular-IoT Optimization and Early Data Transmission, using a novel, detailed implementation of NB-IoT into NS-3 LTE framework LENA. The results have shown that while the improvement using Cellular IoT Optimization was small, Early Data Transmission provided up to 4.1 times lower end to end latency and 1.6 times longer battery lifetime compared to NB-IoT standard transmission and Cellular IoT Optimization. In terms of spectrum usage, Early Data Transmission used 73% less downlink spectrum but was limited by the Random Access windows. For users and network operators, it is highly recommended to use the Early Data Transmission feature of upcoming NB-IoT releases, since it clearly improved the user and spectral performance of NB-IoT transmissions. Further, the simulation results have shown a good scalability for up to 864,000 devices per day in a cell with an area of 4.91km², or 176,000 devices per day per km² for all NB-IoT standard, Cellular IoT Optimization and Early Data Transmission scenarios.

In a second simulation run, the number of Random Access windows has been doubled for all scenarios, which led to an improved Packet Delivery Rate and fewer Random Access collisions. Since the number of Random Access collisions in scenarios with over 1,000,000 devices per day is still high an optimal Radio Resource configuration must be determined in future work. Besides the optimization of Random Access resources, new NB-IoT releases enable transmissions in non-anchor carriers and thus increase the available bandwidth for NB-IoT devices, which will increase the overall cell capacity and will be evaluated in a future work.

ACKNOWLEDGMENT

This work has been carried out in the course of the PuLS project, funded by the Federal Ministry of Transport and Digital Infrastructure (BMVI) under grant agreement no. 03EMF0203B and the 5shine project, funded by means of the Federal State NRW by the Ministry for Economic Affairs, Innovation, Digitalization and Energy (MWIDE) under the Funding ID 005-2108-0073, and was supported by the Collaborative Research Center SFB 876 “Providing Information by Resource-Constrained Analysis”, project A4.

REFERENCES