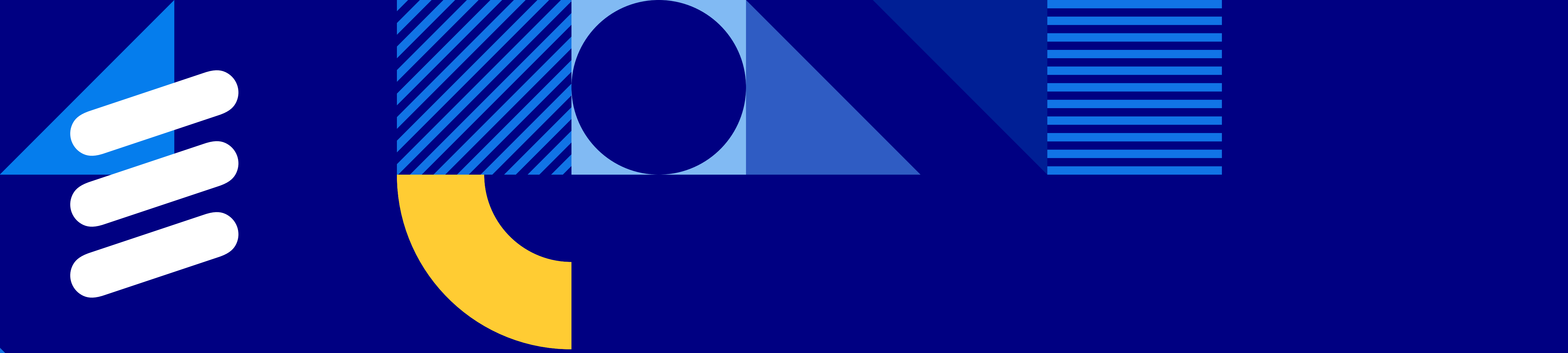
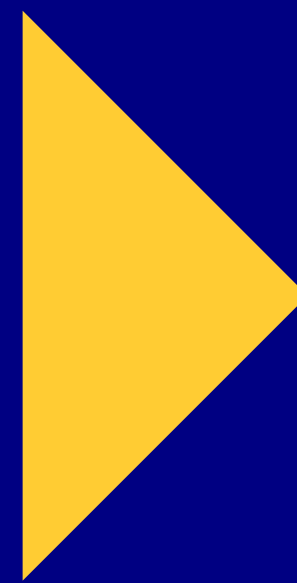




Ericsson Technology Review



#9, October 2026

Faster applications in 6G
with contention-based
buffer status reporting

Charting the future of innovation

Faster applications in 6G with contention-based buffer status reporting

Authors:
Jonas Pettersson, Kristofer Sandlund

Round-trip time can be a significant bottleneck for applications on 5G networks, affecting both application performance and user experience. Ericsson therefore proposes a 6G mechanism that reduces uplink scheduling delays, addressing one contributor to application-level latency. Simulation results with end-to-end protocols and detailed application models demonstrate a 33 percent increase in web-page download bitrate and a 51 percent reduction in latency for UL real-time video.



6G is expected to deliver higher performance across a wide range of services. This is particularly relevant for mobile broadband, which accounts for a large majority of the data carried by 5G networks and is expected to do so in 6G as well [1].

Many mobile broadband (MBB) applications require exchanges where the application must wait for a response before proceeding. For example, when a smartphone app requests data over HTTP (Hypertext Transfer Protocol), the request may be followed by multiple exchanges as the server provides the requested content and associated resources. Larger downloads are also broken into multiple segments that require timely transmission of acknowledgement packets in the uplink (UL) to avoid transport-level congestion control from reducing the connection rate. With all these messages traveling back and forth, round-trip time (RTT) can become a bottleneck.

This becomes particularly relevant with the higher link speeds expected in 6G. At such speeds, link capacity is less likely to be the limiting factor. This also applies to downlink (DL) heavy services, such as large web pages, where DL throughput has traditionally been assumed to be the limiting factor.

Many of the delays that contribute to RTT cannot be reduced by refining cellular standards. UL scheduling latency is an

Uplink scheduling — how it works today

- In cellular networks, the scheduling of uplink (UL) data is controlled by the base station (BS).
- The BS sends UL grants that specify which time-frequency resources to use and how many bits to transmit.
- A buffer status report can be included in a UL transmission and indicates how much data the user equipment (UE) has in its buffers in a compact format.
- When there is no ongoing UL data transfer, the BS has no immediate way of knowing that new data has arrived in the UE's buffer.
- In 5G, a scheduling request is used to notify the BS when a UE does not already have a grant for UL transmission.

exception. In cellular networks, the scheduling of UL data is controlled by the base station (BS), but the BS lacks immediate knowledge of the timing, quantity and nature of incoming user equipment (UE) buffer data. When new data arrives in an empty UE buffer, the BS must first be made aware of this before it can send a grant for a UL transmission. The first granted transmission and many

Terms and abbreviations

BS — Base Station | **BSR** — Buffer Status Report | **CB** — Contention-Based | **CB-BSR** — Contention-Based Buffer Status Reporting | **CB-PUSCH** — Contention-Based Physical Uplink Shared Channel | **CG** — Configured Grant | **CG-BSR** — Configured Grant Buffer Status Reporting | **CRC** — Cyclic Redundancy Check | **DL** — Downlink | **DMRS** — Demodulation Reference Signal | **MBB** — Mobile Broadband | **PRB** — Physical Resource Block | **PUSCH** — Physical Uplink Shared Channel | **RTT** — Round-Trip Time | **SR** — Scheduling Request | **UE** — User Equipment | **UL** — Uplink



subsequent transmissions will contain a buffer status report (BSR) indicating how much data the UE has in its buffers. Thus, it is the first data that arrives in an empty buffer that experiences this extra scheduling delay.

This scheduling delay is particularly relevant to MBB traffic because users tend to alternate between short bursts of activity and longer periods of inactivity. To keep the connection responsive, the network keeps a UE connected for some time after its last activity — typically 10 seconds — rather than releasing it immediately. As a result, the majority of UEs at any given time are connected without an ongoing data transfer, and although a burst of activity can generate many requests in a short period, this is offset by the longer idle stretches: each connected MBB user sends requests only sparsely — one to two per second on average. Yet when a burst begins, new data can arrive in the UE buffer repeatedly, making efficient notification of the BS particularly important.

5G uplink scheduling: limitations and challenges

Reducing UL scheduling latency in 5G is challenging because the network has limited information about newly arriving UL data. Except in niche cases where the traffic pattern is largely known to the BS, such as voice over IP, 5G UL scheduling relies on a single-bit scheduling request (SR) transmitted by the UE on a periodically available UL control channel resource. When the BS receives an SR, it responds with a small UL grant that carries a BSR and a small amount of data. Only after receiving the BSR can the BS schedule the remaining data in the UE buffer.

One possible approach would be to provide a large grant in response to an SR, allowing the UE to transmit its buffered

data without waiting for a BSR. However, the BS would have to estimate how much data is in the buffer. This would often result in the buffer being overestimated, causing the UE to transmit padding and creating unnecessary interference for other cells.

The 5G approach has three fundamental limitations. First, the BS initially knows only that the UE has data to transmit, not how much data is in the buffer. The BS cannot schedule all the buffered data until the first granted transmission containing the BSR has been received.

Second, all connected UEs require a dedicated periodic resource for transmitting an SR. Because MBB users send requests only sparsely, this results in inefficient use of radio resources. The large number of connected but inactive UEs also means that the SR period must be relatively long. A typical value in 5G is 20ms, while values of up to 80ms or even 160ms may be used under high load.

Third, the SR mechanism does not provide error protection. If SR resources are configured frequently —for example, once every millisecond — the false alarm probability becomes a problem. With a false alarm probability of up to 1 percent, 10 false alarms could be generated per second. Given the scarcity of requests — typically only one to two real requests per second, as noted above — false alarms could therefore occur 5 to 10 times more frequently than actual requests.

Several approaches exist in 5G to address these limitations, but none provides a practical solution to all three problems across both low and high system loads. Extending the SR signal to carry more bits is one option, and a UE can have multiple SR configurations in 5G. However, this increases

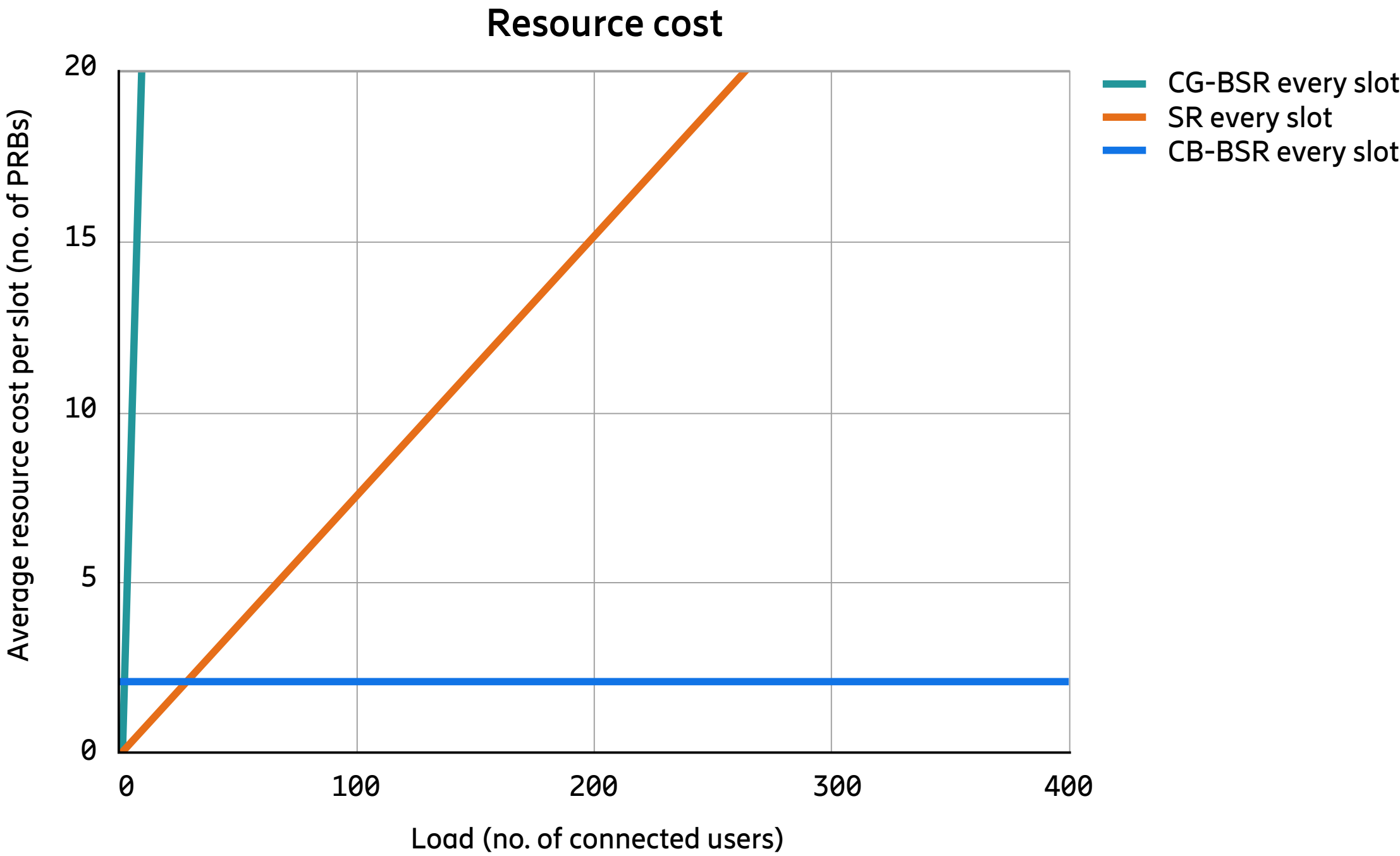


Figure 1: How the resource cost of frequent SR and CG scales with the number of users in the system

the resource cost per user. At the same resource cost, transmitting more bits also requires longer intervals between resource opportunities, increasing the waiting time before newly arrived data can be transmitted.

A configured grant (CG) for a PUSCH (physical uplink shared channel) transmission can carry a BSR and addresses the false alarm problem by protecting the transmission with a cyclic redundancy check (CRC). However, the amount of dedicated radio resources required for each connected UE is even greater than for a multi-bit SR.

Figure 1 illustrates how the resource cost of frequent SR and CG scales with the number of users in the system. As both approaches rely on dedicated resources, the resource cost increases linearly with the number of users. Changing the resource period affects the slope of this relationship but does not change its linear nature.

Designing 6G uplink scheduling for shared access

An ideal mechanism for buffer reporting would allow each UE to transmit granular buffer information immediately



and reliably in any UL slot, thereby avoiding all three of the limitations in 5G. The upper part of **Figure 2** illustrates how the current 5G SR solution compares with the desired solution for 6G. In effect, the desired solution would give each UE access to a CG that can carry a BSR in every slot – a solution that is not feasible in 5G because it would be too costly.

Since we know that MBB users send requests sparsely, sharing time/frequency resources and allowing UEs to contend for access is a natural way to achieve our goal. In this way, hundreds of UEs can transmit a BSR in the next slot at the same resource cost as a single 5G CG. In principle, this could be achieved in 5G by assigning the same CG to hundreds of UEs. In practice, however, the BS would have to

blindly search for transmissions from each UE on the shared time/frequency resources. This would require testing against hundreds of sequences, creating a substantial processing burden for the receiver. There is also no effective way to handle cases where the BS fails to detect a transmission altogether. On PUSCH, the BS explicitly schedules retransmissions, which works because the BS can detect that a UE has transmitted even when the transmission itself cannot be decoded. This approach is much less feasible when hundreds of UEs share the same resource, however.

In the wanted scheme, despite the low average utilization, transmissions are uncoordinated, meaning that multiple UEs will sometimes transmit on the same time/frequency resources. For the scheme to perform well, it must therefore have a high probability of resolving multiple transmissions that occur on the same time/frequency resources. The bottom section of Figure 2 illustrates how such a scheme works. All contention-based (CB) BSR resources can be accessed by any user at any time, and the channel design makes it possible to resolve multiple transmissions on the same time-frequency resources.

Contention-based buffer status reporting in 6G

CB-BSR is designed to address the limitations of 5G SR while making efficient use of radio resources. The basic idea is to replace the dedicated, periodically available SR resources with shared resources that any UE can use when it has data to report. Because UEs generate requests only occasionally, not every UE needs to have its own resource available at all times.

The BS allocates one or more CB-BSR resources, each typically consisting of one or two resource blocks and carrying

All contention-based BSR resources can be accessed by any user at any time.

a small amount of data (56 bits, for example). These resources can be accessed by any UE at any time. Because each resource is small, CB-BSR resources can be available in every UL slot without significantly reducing the resources available for PUSCH. The BS can also dynamically adjust the number of CB-BSR resources in each slot according to the current load.

When a UE has new data in an otherwise empty buffer, it selects a CB-BSR resource and transmits its BSR. Since the resources are shared, several UEs may choose the same resource in the same slot. This contention is intentional. The system is designed so that multiple transmissions can often be separated and decoded, while UEs whose transmissions are not successfully received can try again.

To allow multiple UEs to share the same resource, each CB-BSR resource is associated with a set of different demodulation reference signal (DMRS) sequences. Whenever a UE needs to transmit a BSR, it randomly selects one of the available sequences. The DMRS sequences allow the BS to distinguish between transmissions from different UEs and, in many cases, decode multiple CB-BSR transmissions on the same resource. The random selection also provides diversity across retransmissions, reducing the probability that

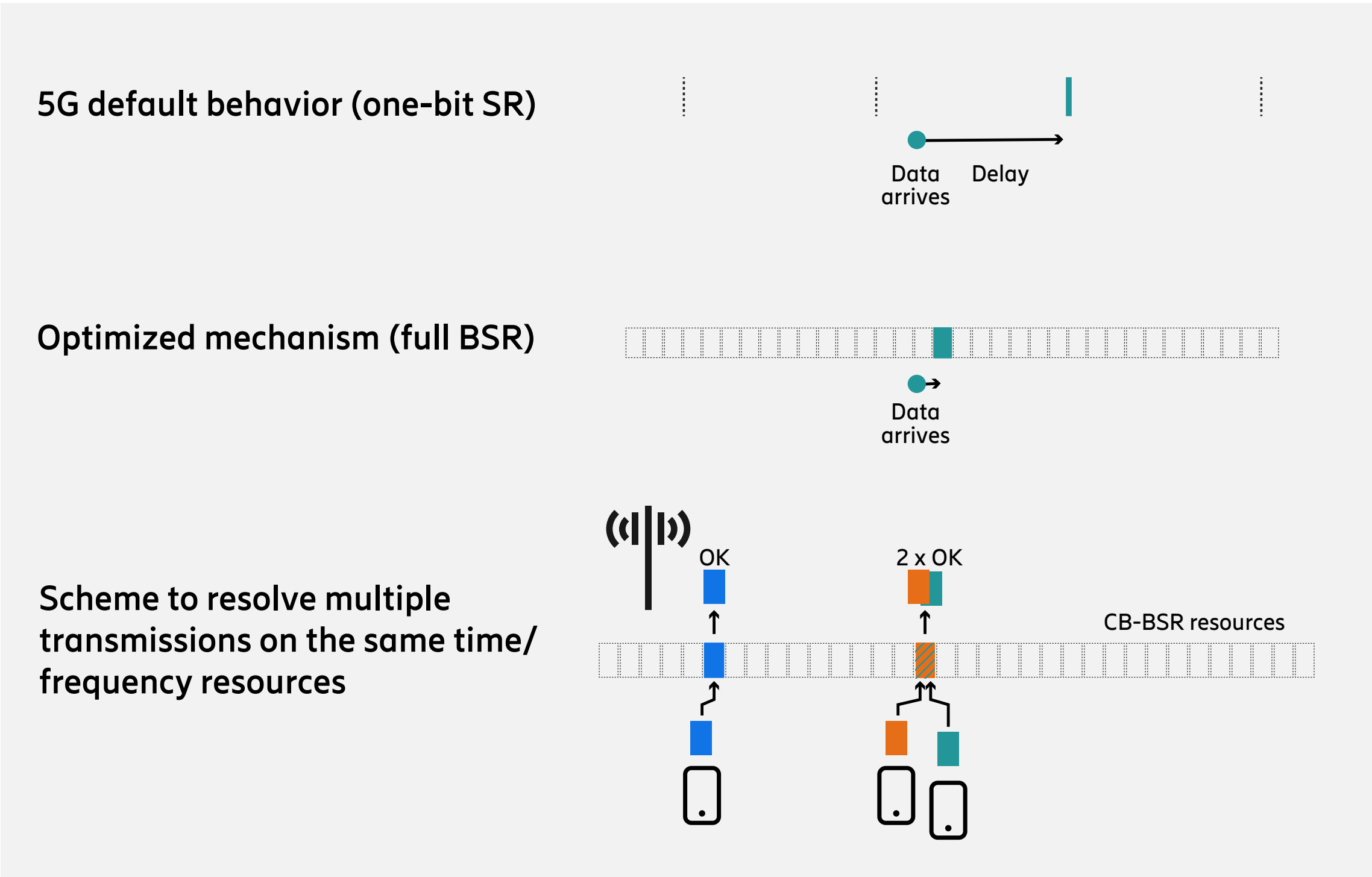


Figure 2: 5G default behavior, an optimized mechanism for buffer reporting and a scheme to resolve multiple transmissions on the same time/frequency resources



the same UEs will collide again when they reattempt on the same time slot, thereby alleviating the need for a backoff timer.

There is, however, a practical limit to the number of sequences that the BS receiver can process for each resource. With the proposed DMRS scheme, efficient receiver processing is possible for approximately 10 sequences, allowing up to 10 transmissions to be resolved on the same resource. When two UEs use the same time/frequency resources but select different DMRS sequences, there is a high probability that both transmissions can be decoded. The small payload allows for robust modulation and channel coding to be used. If two UEs select the same DMRS sequence, however, there is a high risk that both transmissions will be lost.

A reattempt mechanism is therefore needed to handle CB-BSR transmissions that are not successfully received. When a CB-BSR is successfully received, the UE is normally scheduled quickly, allowing the resulting UL grant to serve as an implicit acknowledgement. If the BS fails to detect or decode the CB-BSR, no grant is sent. The UE uses a timer that is started with each CB-BSR attempt and can make another attempt after the timer expires.

The simulated system includes a mix of users following different realistic traffic models.

The CB-BSR payload contains both the BSR and a UE identifier, allowing the BS to determine which UE sent the transmission. The transmission is also protected by a CRC. This effectively eliminates the false detections that can occur with the single-bit SR mechanism.

Together, these design principles address the three main limitations of 5G SR. Shared resources improve resource efficiency by allowing many UEs to use the same resources. The use of multiple DMRS sequences allows multiple transmissions to be resolved on the same resource, while reattempts provide a way to recover from collisions that cannot be resolved. CRC protection avoids the false detections associated with SR. CB-BSR can therefore provide many of the properties of the ideal buffer-reporting mechanism described in the previous section, while allowing buffer information to reach the BS with much lower scheduling delay.

The development of CB-BSR for 6G was driven by a detailed analysis of 5G traffic, applications and protocols to identify opportunities for improving the performance experienced by smartphone users. The objective was to develop a resource-efficient solution without adding substantial implementation complexity. Ericsson has taken a leading role in introducing the concept into 6G standardization [2]. It is supported by a coalition of more than 10 companies and is a leading candidate for 6G UL scheduling.

Solution performance results

Evaluating a mechanism like CB-BSR requires more than a link-level analysis of the radio channel. Because its benefit is ultimately experienced at the application level, evaluating CB-BSR requires a simulator that models realistic, time-dynamic traffic and follows its behavior end to end. For this

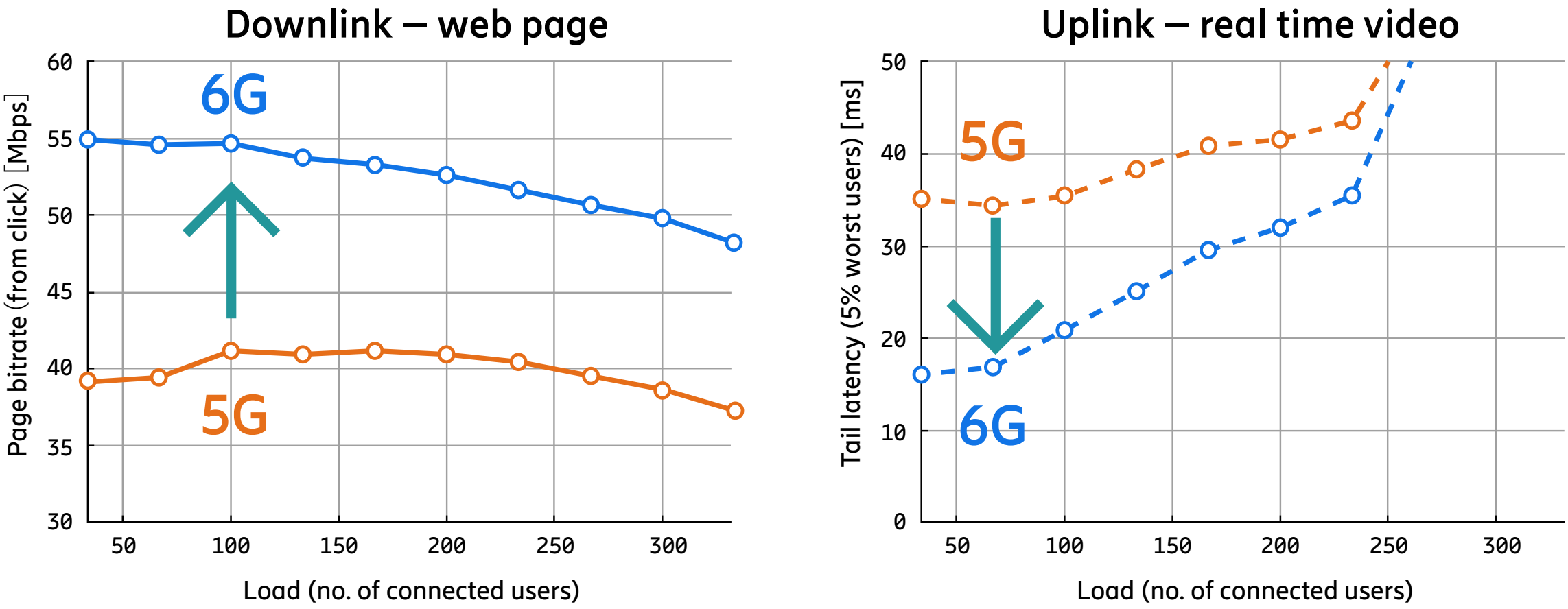


Figure 3: Solution performance results for web-page downlink (left) and real-time video uplink (right)

reason, we used a simulator with detailed models spanning every layer from radio propagation up to the application itself, for the UE, the BS and the internet servers involved in each exchange.

The results presented in Figure 3 are based on a 20MHz FDD (frequency division duplex) carrier in the 1.8GHz frequency band, with an inter-site distance of 500m. The simulated system includes a mix of users following different realistic traffic models, each of which stresses the CB-BSR solution in a different way, from short bursts of web browsing to continuous UL video streams. The proposed 6G CB-BSR mechanism is compared against a typical 5G SR configuration with a 20ms period, evaluated across multiple load points corresponding to an average of 33 to 333 MBB UEs connected per cell.

The graph on the left side of Figure 3 shows the results for web-page download, a traffic pattern that is representative of many other smartphone applications that communicate with servers over HTTP. Despite the fact that the large majority of the data in this traffic flows in the DL, faster UL scheduling of the accompanying requests and acknowledgements is enough to increase the experienced bitrate by 33 percent. This illustrates a point made earlier in the article: as link speeds increase in 6G, it is often RTT rather than raw throughput that limits application performance, even for DL-heavy services.

The graph on the right side of Figure 3 shows the impact on UL real-time video, a service whose dependence on UL scheduling is more direct. Here, CB-BSR reduces tail latency by 51 percent compared with 5G SR. Tail latency is

measured in two steps: we first determine the latency of each user’s slowest one percent of packets, and then look at the 5 percent of users with the highest values. Across all the evaluated load points, from lightly to heavily loaded cells, CB-BSR consistently outperforms the 5G SR baseline, confirming that the gains hold as the number of connected users grows rather than being limited to a single operating point.

These specific figures depend on the assumptions used in the simulation, including good backhaul and ample DL capacity, and should not be read as universal constants. They do, however, illustrate the order of magnitude of improvement that CB-BSR can deliver by removing the scheduling-request bottleneck inherent in 5G, and they demonstrate that this improvement comes without the linear growth in dedicated radio resources that alternative approaches would require. Research also shows that similar gains can be achieved in other scenarios, such as time division duplex (TDD) deployments [2].

Conclusion

As 6G pushes link speeds higher, the bottlenecks that limit application performance are shifting away from raw capacity and toward the delays inherent in how networks coordinate access to the radio interface. Uplink (UL) scheduling latency is one such delay, and unlike many of the other contributors to round-trip time, it can be addressed through improvements to the radio-access standard itself. It matters especially to new and evolved services such as immersive XR, physical AI and time-critical autonomous mobility, where responsiveness rather than throughput defines the experience – but it matters just as much to mobile broadband services too, as the results in this article show.

Contention-based buffer status reporting shows that this delay can be reduced substantially, without the linear growth in dedicated resources that would be needed to achieve the same effect with 5G-style mechanisms. The simulation results we have presented in this article – a 33 percent increase in web-page download bitrate and a 51 percent reduction in tail latency for UL real-time video – demonstrate that this is not a marginal gain but a meaningful improvement to the experience of everyday smartphone applications, achieved through more efficient use of existing radio resources rather than additional spectrum or hardware. Further work in the coming standardization phases will determine how these performance gains can be incorporated into commercial 6G networks.



The authors



Jonas Pettersson is a master researcher at Ericsson Research who joined the company in 1997. His research interests are radio resource management and how it interacts with applications to achieve high end-to-end performance. Pettersson holds an M.Sc. in engineering physics from Umeå University in Sweden.



Kristofer Sandlund joined Ericsson in 2005 and is currently a principal researcher at Ericsson Research. At present, his work focuses on end-to-end system simulators and research related to radio resource management and network protocols. Sandlund holds an M.Sc. in computer communications from Luleå University of Technology, Sweden.

References

1. Ericsson Mobility Report, June 2026 ↗
2. 3GPP, TSG-RAN WG2 #134, System simulation for contention based BSR, May 18-22, 2026 ↗



Further reading

- Ericsson, 6G explained ↗
- 6G RAN – key building blocks for new 6G radio access networks ↗
- Five research insights on 6G XR ↗
- 6G standardization: The technology realization step begins ↗
- Ericsson, Technology standards ↗