

uRLLC业务时延分析及低时延网络部署探讨

Analysis of Latency of uRLLC Service and discussion of Low Latency Network Deployment

张斌¹,张鹏²,薛超粤²(1. 中讯邮电咨询设计院有限公司,北京 100048;2. 中国联合网络通信集团有限公司,北京 100033)
Zhang Bin¹,Zhang Peng²,Xue Chaoyue²(1. China Information Technology Designing & Consulting Institute Co.,Ltd.,Beijing 100048, China;2. China United Network Communications Group Co.,Ltd.,Beijing 100033,China)

摘要:

uRLLC技术因低时延、高可靠的特性,广泛应用于工业控制、自动驾驶、车联网等场景。为了降低空口时延,3GPP在R15及R16标准中,从优化调度和减少时间间隔等方面提出了一系列的增强技术。这些低时延技术可以不同程度地降低5G网络的空口时延,但是都无法克服TDD制式的等待时延,相对应的FDD制式在低时延方面具有天然的优势。在实际网络中,除了空口时延,影响用户端到端时延的因素还有传输时延及业务平台时延等方面,因此,打造低时延网络是一个系统性工程。

Abstract:

Due to the characteristics of low latency and high reliability, uRLLC technology is widely used in industrial control, automatic driving, vehicle networking and other scenarios. In order to reduce the air interface delay, 3GPP proposes a series of enhancement technologies in R15 and R16 standards, such as optimizing scheduling and reducing waiting time interval. These low latency technologies can reduce the air interface delay of 5G network in varying degrees, but they can not overcome the waiting delay of TDD system. On the contrary, the FDD system has a natural advantage in low delay. In the actual network, except the air interface delay, the factors that affect the user end-to-end delay are transmission delay and service platform delay. Therefore, building a low latency network is a systematic project.

Keywords:

Low latency; Ultra-reliable; Grant free transmission; Downlink preemption

引用格式:张斌,张鹏,薛超粤. uRLLC业务时延分析及低时延网络部署探讨[J]. 邮电设计技术,2022(5):37-41.

1 概述

5G标准设立之初就同时考虑了2B、2C 2种应用场景,uRLLC具有超低时延、超高可靠等特性,主要为满足2B应用场景而生的,是运营商切入垂直行业的突破口,uRLLC技术可以广泛应用于工业控制、自动驾驶、车联网、远程医疗等时延敏感且网络稳定性要求较高的场景。

uRLLC技术分为低时延技术和高可靠性技术2个部分,本文主要对低时延技术进行研究,首先分析影

关键词:

低时延;高可靠性;上行免调度;下行抢占

doi:10.12045/j.issn.1007-3043.2022.05.008

文章编号:1007-3043(2022)05-0037-05

中图分类号:TN929.5

文献标识码:A

开放科学(资源服务)标识码(OSID):



响5G空口时延的关键因素,然后介绍uRLLC为降低空口时延所采用的技术手段,并对各项技术降低时延的效果进行分析,在此基础上,提出uRLLC低时延网络的部署建议。

2 5G空口时延分析

2.1 影响空口时延因素分析

5G空口时延分为用户面时延及控制面时延,其中用户面时延为数据从基站PDCP层到终端PDCP层的时延(下行)或数据从终端PDCP层到基站PDCP层的时延(上行);控制面时延为终端在网络中注册或者状态转换(如从空闲态到连接态)所花费的时间。用户

收稿日期:2022-03-02

面时延对2B业务影响更大,因此本文主要关注用户面时延。

用户面时延主要受传输时隙间隔、资源调度时间、重传时间、终端和基站的处理时间等因素的影响。

在TDD制式,上下行数据必须等待在相应的上下行时隙才能进行数据的发送,以5G现网2.5ms双周期的帧结构为例,在上行方向最大等待时间为4个Tslot,在下行方向最大等待时间为2个Tslot,因此TDD制式上下行时隙倒换间隔为影响空口时延的关键因素,如图1所示。

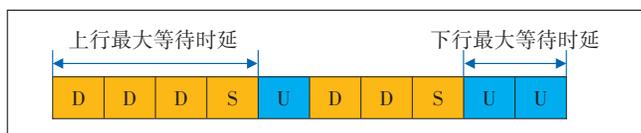


图1 TDD制式下上下行的最大等待时延

在传统的业务调度方式中,终端发送数据都需要经历调度请求、调度授权、数据传输及重传等过程,调度时延和数据重传时延也是影响空口的重要因素,如图2所示。

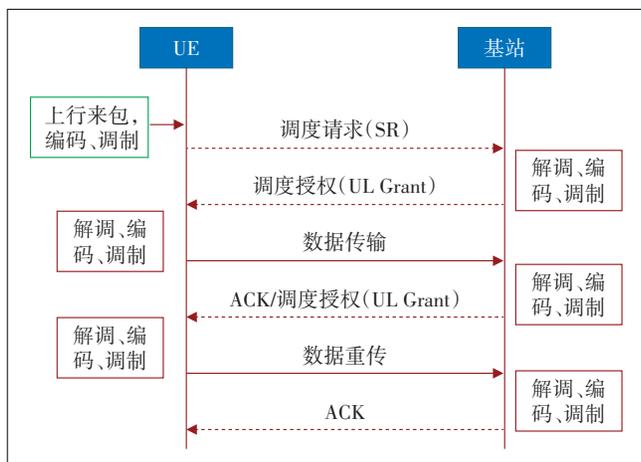


图2 上行数据发送流程

此外,终端和基站的处理时间也是影响空口时延的因素,但处理时间主要受限于基站和终端的自身处理能力,不在本文讨论范围。

2.2 uRLLC业务时延要求

根据国际电联IMT-2020标准的要求,uRLLC业务空口时延为1ms;3GPP对uRLLC业务的用户面空口时延要求为上行0.5ms、下行0.5ms,控制面时延为10ms。上述标准为极限情况下的时延要求,3GPP根据业务需求的不同,给出智能电网、工业自动化,自动驾驶等业务的时延与可靠性要求(见表1)。

表1 uRLLC业务典型应用需求评估

应用场景	业务名称	可靠性要求	时延要求
配电应用	配电网故障及断电管理	99.9999%	端到端时延为5ms; 空口时延为2~3ms
	差动保护	99.999%	端到端时延为15ms; 空口时延为6~7ms
工业自动化	运动控制	99.9999%	端到端时延为2ms; 空口时延为1ms
交通运输	远程驾驶	99.999%	端到端时延为5ms; 空口时延为3ms
	智能运输系统(ITS)	99.999%	端到端时延为10ms; 空口时延为7ms

为了达到uRLLC业务的低时延要求,根据2.1节分析结果,除了终端和基站提升自身处理能力外,通信侧也需要在优化资源调度、减少传输时间间隔、降低重传次数等方面做出改进。

3 uRLLC低延时技术及效果分析

为了满足uRLLC的低时延特性,3GPP在R15及R16标准中提出了一系列的增强技术,主要包括智能预调度、上行免调度、下行抢占机制、Mini-slot技术等优化调度技术,特殊帧结构、更大子载波间隔等减少时间间隔技术,以及提高可靠性、减少重传次数等相关技术。本章将对这些技术做简要的介绍,并重点分析这些技术的引入能够带来的降时延效果。

3.1 智能预调度及上行免调度技术

在正常的上行数据发送流程中,UE需要发送SR请求基站,基站发送上行授权(UL Grant)信息给终端后,终端才可以发送上行数据,为了减少SR请求及调度时间,3GPP提出了智能预调度及上行免调度技术。

智能预调度技术的原理是在上下行信息交互过程中,下行数据传输总是伴随着上行信息的发送,智能预调度功能使得基站在发送下行数据时主动触发上行预调度,加快上行信息的发送,从而降低整体数据传输的时延,如图3所示。智能预调度虽然节省SR请求时间,但是仍然需要解析DCI(Scheduling Response),与完全免调度的理想还有差距。

上行免调度技术是智能预调度技术的升级,其原理是基站为终端预先配置可进行免调度的PUSCH资源,终端在发送数据前不再需要发送SR请求,而是在预先配置的PUSCH资源上进行上行传输,在终端收到去激活信息前,将会一直使用该资源进行传输,上行免调度的业务流程如图3所示。

此处分别以TDD 2.5ms双周期制式和FDD制式

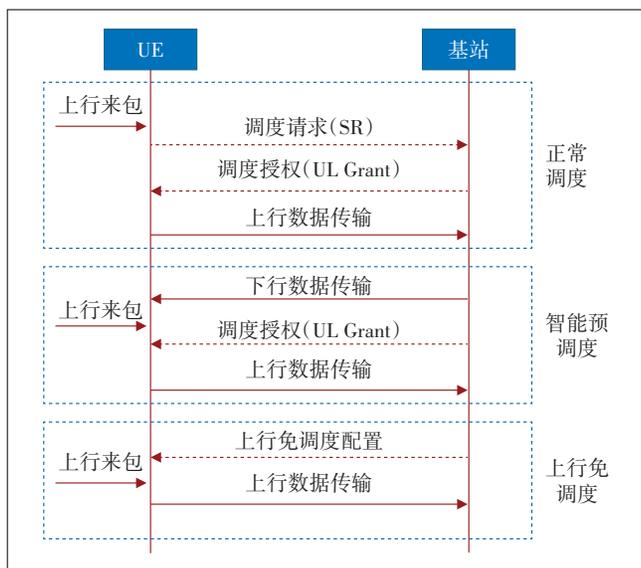


图3 3种调度模式业务流程对比

为例进行时延分析,并假设基站的SR周期足够小,上行来包后可以在下一个上行时隙及时发送SR请求,同时假设基站和终端有快速的编解码能力,收到SR请求后可以在下一个下行时隙及时发送UL Grant信息(DCI),UE收到DCI后,可以在下一个上行时隙及时发

送上行数据。

图4对比了TDD制式下3种调度模式的时延,由图4可知,对于TDD 2.5 ms双周期制式,上行免调度技术相比正常调度流程可以减少4~10个Tslot时延,但仍然有0~4个Tslot的等待及调度时延;智能预调度技术可以减少0~6个Tslot时延,但仍然有2~7个Tslot的等待及调度时延。

图5对比了FDD制式下3种调度模式的时延,由图5可知,对于FDD制式,上行免调度技术相比正常调度流程可以减少2个Tslot时延,且没有等待及调度时延;智能预调度技术可以减少1个Tslot时延,但存在1个Tslot的等待及调度时延。

3.2 下行抢占机制

对于eMBB业务和uRLLC业务共享频谱的场景,为了降低uRLLC业务时延,3GPP在R16中定义了下行抢占机制,突发的uRLLC业务可以抢占已经在传输的eMBB业务的时频资源进行传输。

下行抢占调度原理如图6所示,uRLLC业务一般数据量较小,可以通过打孔方式占用eMBB业务的部分资源模块,即在一个时隙同时发送uRLLC和eMBB

	D	D	D	S	U	D	D	S	U	U	D	D	D	S	U	
正常调度	上行				SR	DCI			T							
		上行			SR	DCI			T							
			上行		SR	DCI			T							
				上行	SR	DCI			T							
					上行		上行		SR		DCI					T
							上行		SR		DCI					T
								上行	SR		DCI					T
									上行	SR	DCI					T
										上行	SR	DCI				T
											上行	SR				SR
上行免调度	上行				T											
		上行			T											
			上行		T											
				上行	T											
					上行/T		上行									
							上行		T							
								上行	T							
									上行/T							
										上行/T						
											上行/T					
智能预调度	上行	DCI			T											
		上行	DCI		T											
			上行	DCI	T											
				上行	DCI											
					上行	DCI			T							
						上行	DCI		T							
							上行	DCI	T							
								上行	DCI	T						
										上行	DCI					T
											上行	DCI				T

图4 TDD制式下3种调度模式时延对比

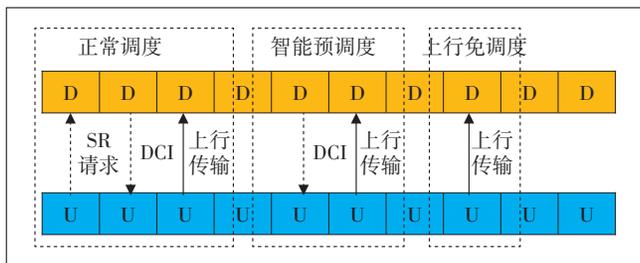


图5 FDD制式下3种调度模式时延对比

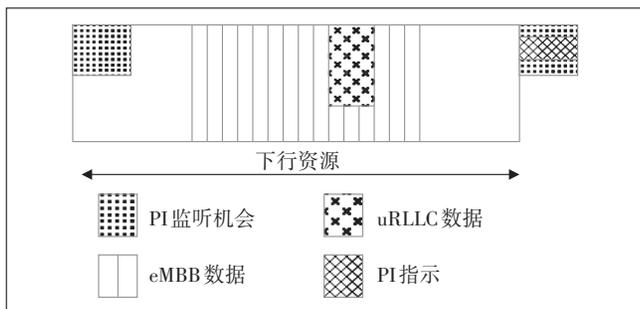


图6 下行抢占机制资源分配

业务。这种下行抢占机制会对 eMBB 业务产生一定影响,为了降低这种影响,3GPP 引入了下行抢占指示机制,即通过组播发送 PI 指示信息,通知终端被抢占的资源,终端可以通过 HARQ 机制,要求基站重新发送受影响的数据。

图7给出了 TDD 及 FDD 模式下下行抢占机制节省时间,由图7可知,对于 TDD 2.5 ms 双周期制式,下行抢占调度机制相比正常调度流程可以最多减少3个 Tslot 的时延,对于 FDD 制式,下行抢占调度机制相比正常调度流程可以减少1个 Tslot 的时延。

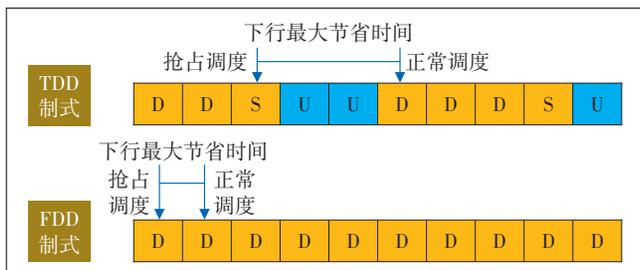


图7 TDD 及 FDD 模式下下行抢占机制节省时间

3.3 Mini-slot 技术

Mini-slot 继承了 LTE 中减小传输时间间隔(sub-slot)的设计理念,将最小的调度单元由时隙变为符号,实现符号级别的调度,可以减少数据发送的等待时间,适应 uRLLC 业务的低时延小数据量的特点。mini-slot 长度为 2、4、7 个符号,基于时隙的调度和基于 mini-slot 的调度如图8所示。

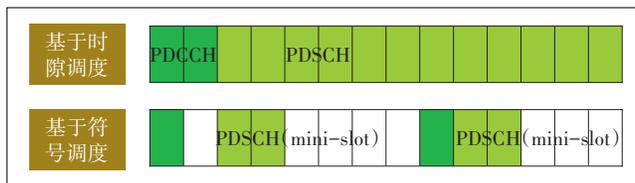


图8 基于时隙调度与基于符号调度对比

如图9所示,本文以下行2符号调度方案为例来计算节省时间,相比时隙调度,对于 TDD 2.5 ms 双周期制式,根据来包时间与上下行时隙的对应关系,下行最大可以降低 $2+6/7T_{slot}$ 的时延,对于 FDD 制式,最大可以节省 $6/7T_{slot}$ 的时延。

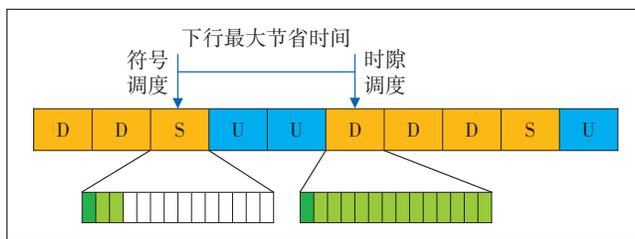


图9 TDD 制式基于符号调度的时间节省

3.4 特殊帧结构及更大的子载波间隔

5G TDD 具有灵活的帧结构,除了典型的 2.5 ms 双周期、5 ms 单周期外,为了满足 uRLLC 业务的低时延特性要求,还可以考虑 1 ms 单周期的帧结构,如图10所示。

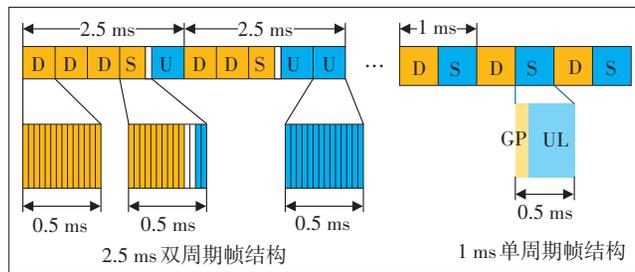


图10 2.5 ms 双周期与 1 ms 单周期帧结构

相对于 2.5 ms 双周期的帧结构,1 ms 单周期的帧结构可以明显减少等待时延,根据来包时间与上下行时隙的对应关系,上行最多可以减少3个 Tslot 的时延,下行最多可以减少1个 Tslot 的时延。

虽然 1 ms 单周期帧结构可以降低等待时延,但是也有明显缺点,主要是 GP 符号占比达到 7% 的,对频谱效率影响较大,同时对于上下行非对称业务,也会影响某一方向的数据传输速率。

相对 LTE 固定的 15 kHz 的子载波间隔, NR 具有不同类型的子载波间隔,如图11所示。时隙长度随着子载波间隔的增大而减少,60 kHz 子载波间隔对应的

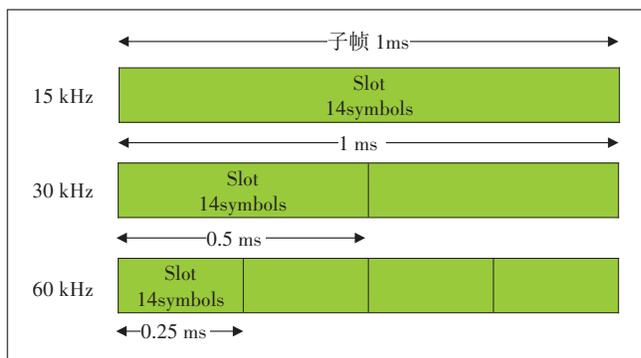


图 11 TDD 制式基于符号调度的时间节省

时隙长度 0.25 ms, 相对 30 kHz 减少一半。

时隙长度的减少可以带来时延的降低,但是由于符号长度较短,循环前缀(CP)较短,抵抗多径干扰的能力较弱,在某些场景和频段中会存在性能风险。

3.5 降低 HARQ 重传次数

根据 HARQ 机制,数据发送方收到 NACK 反馈或重传指示后,需要重新发送数据,直到收到 ACK 信息才认为本次数据发送成功,数据的重传必然带来时延的增加。此处以 FDD 上行为例来分析重传带来的时延,且假设终端和基站的处理速率足够快,即接收方收到数据后的第 1 个 slot 里反馈 ACK/NACK/重传指令,由图 12 可知,在此模式下,重传会带来 2 个 Tslot 的时延;如果终端和基站的处理时间为 1 个 Tslot,即接收方收到数据后的第 2 个 slot 里反馈 ACK/NACK/重传指令,则重传会带来 4 个 Tslot 的时延。

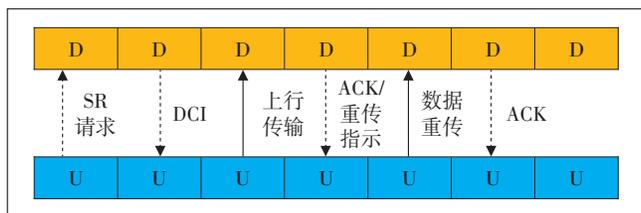


图 12 数据重传引起的时延

对于 TDD 模式,因为涉及到上下行时隙倒换,只有在对应的时隙才能发送上下行数据,因此重传带来的时延更大。

在现网中,可以采用通过完善信号覆盖质量、降低干扰水平、降低编码率、物理层重复、PDCP 层重复、Multi-TRP 传输等手段来降低 HARQ 重传次数。

4 uRLLC 低时延网络部署建议

从网络制式上看,uRLLC 低时延技术可以不同程度降低 5G 网络的空口时延,但是对于 TDD 制式,无论

何种低时延技术都无法克服上下行时隙倒换带来的等待时延,FDD 制式无上下行等待时延,上层数据到达后可以及时发送,在低时延方面具有天然的优势,因此对于时延要求严格的远程控制、自动驾驶类应用可以优先考虑在 FDD 制式上部署。

从低时延技术上看,上行免调度技术和下行抢占机制可以带来的时延增益较大,且对现有网络及终端影响较小,可以考虑优先引入;Mini-slot 技术带来的时延增益有限,且需要终端的支持,可以考虑在特定的场景部署;特殊帧结构及更大的子载波间隔会对频谱效率及抗干扰方面产生影响,需要谨慎评估;网络的低时延和高可靠性是相辅相成的,因此,降低数据重传次数,提升可靠性对降低 uRLLC 业务时延来说也非常关键。

5 结束语

本文主要从空口时延的角度对 uRLLC 低时延技术进行了理论分析,并给出了无线侧的部署建议。但是,在实际网络中,除了空口时延,影响用户端到端时延的因素还有传输时延及业务平台时延等方面,因此,打造低时延网络是一个系统性工程,需要无线、传输、核心网、业务平台等环节的共同努力。垂直行业应用已经风起云涌,专网时代已经到来,未来对 uRLLC 技术需求会越来越多,如何高质量构建低时延、高可靠性网络需要业界不断探索。

参考文献:

- [1] 梁辉,韩潇,李福昌. 5G uRLLC 无线网络部署方案分析[J]. 邮电设计技术,2020(3):15-18.
- [2] 杨一帆,姚键,吴祖辉,等. 面向 5G uRLLC 业务的时延分析与无线超低时延技术研究[J]. 电信工程技术与标准化,2021,34(2):10-15.
- [3] 黄韬,李鹏翔. uRLLC 关键技术和网络适应性分析[J]. 移动通信,2020,44(2):25-29.
- [4] 梁健生,陈晓冬. uRLLC 关键技术研究与时延分析[J]. 移动通信,2020,44(2):35-39+44.
- [5] Study on physical layer enhancements for NR ultra-reliable and low latency case:3GPP TR 38.824[S/OL]. [2021-12-09]. ftp://3gpp.org/.

作者简介:

张斌,毕业于北京邮电大学,高级工程师,硕士,主要从事移动通信新技术研究、无线网络规划设计工作;张鹏,毕业于吉林大学,高级工程师,硕士,主要从事移动网络规划、移动通信新技术研究等工作;薛超粤,毕业于西安邮电大学,工程师,主要从事无线网络规划与优化、数字化平台架构设计、项目管理等工作。