



DDR5 服务器平台解决方案内的 I3C 信号完整性挑战

主讲者：Azusena Lupercio, Juan Orozco, Nestor Hernandez

英特尔



Juan Orozco: 大家好，我是 Juan Orozco，我和两位同事 Azusena Lupercio 及 Nestor Hernandez 都任职于英特尔，将介绍 DDR5 服务器平台解决方案内的 I3C 信号完整性挑战。

Agenda

- Introduction
- DDR5 SPD server connectivity and bus characteristics
- I2C and MIPI I3C Retro-compatibility challenges
 - Non-dynamic pullup impact
 - Dynamic pullup on open-drain
- Buffer R_{ON} design implications
- Critical time margin calculation
 - Frequency and AC/DC parameters impact
- Non-monotonic signal behavior
- Slope reversal capability and timing improvement
- Summary

Juan Orozco: 今天简报在开场之后，会简介 DDR5 SPD 服务器联机与总线的特质，再说明 I2C 与 MIPI I3C 向下兼容的挑战，以及 I3C 控制器与目标装置缓冲对系统设计的意涵；并深入探讨临界时间余裕计算，以及不同 AC/DC 参数对操作频率有何影响。也会呈现我们所见的非单调性信号，以及为何目标装置需要斜率反转能力，才能够改善装置的时序余裕及逻辑适当功能。最后再做总结。



MIPI ALLIANCE DEVELOPERS CONFERENCE

Introduction

28-29
SEPTEMBER
2021

[MIPI.ORG/DEVCON](https://mipi.org/devcon)

MOBILE & BEYOND

© 2021 MIPI Alliance, Inc.

Introduction

- The MIPI I3C Improved Inter-Integrated Circuit interface is first introduced in a server application for the DDR5 DIMM Serial Presence Detect (SPD) function.
- Its implementation exceeds by far the bus capacitance/loading specification, which was defined for low capacitance Mobile/IoT applications.
- This presentation covers the interoperability challenges of the dynamic push-pull and open-drain operating modes on I3C BASIC server applications.
 - Covering an in-depth analysis of the implications of long PCB traces, multiple DIMM routing branches, several loads, to the electrical and timing parameters.

Juan Orozco: 进入正题，MIPI I3C 改良积体总线电路最先出现在服务器应用中，以满足 DDR5 双线内存模块的串行存在侦测（SPD）功能，当建置在服务器中，超越 MIPI I3C BASIC 规格内的总线电容，这项规格定义了高低总线电容、行动物联网、小型手持式装置应用等。今天会介绍到 DDR5 服务器应用内的互用性挑战，是因为 I3C BASIC 解决方案运作模式内的动态推挽与开漏特质，我们将深入分析，当主板具有多种电路设计与多项负载时，电路板线路很长会造成什么影响，而这些特质又会如何影响电子与时序参数。

Introduction cont'd

- I3C Communication Bus specification was released by MIPI Alliance in 2016, as an improved communication protocol compared to its predecessor I2C, but the implementation of I3C, in a Data Center (Server) application was materialized until 2020.
- The main enhancements in I3C adopted by the DDR SPD function are:
 - Higher bit rate: up to 12.5MHz, compared to 100KHz-1MHz I2C SPD in prior DDR generations (125x to 12.5x higher bit rate).
 - Better IO electrical interface: Push-pull driver vs Open Drain only.
 - In-band interrupts (IBI) support – Not supported in DDR5 now, but looking for support in the future (or in other Server use cases).
 - In band Common Command Codes (CCCs) – Direct or Broadcast.
 - Reduced interface power (1.0V IOs).

Juan Orozco: I3C 通信总线规格是由 MIPI Alliance 在 2016 年发表，做为改良前代 I2C 规格的通信协议，不过一直到首批 DDR5 服务器问世后，I3C 才在 2020 年首度出现在数据中心的服务器应用中。

DDR SPD 功能使用 I3C 之后，主要改良之处当然是提高比特率，理论上可达 12.5MHz，相对而言，先前 DDR 的 I2C SPD 功能比特率为 100KHz 至 1MHz，依据实际建置情况不同，比特率可提高 12.5 倍至 125 倍。

改良的 IO 电子接口具备推挽驱动器，而非开漏缓冲，因此更加适合，虽然 DDR5 SPD 目前并未使用频内中断 (IBI) 功能，但我们正在研究是否未来需支持，以符合其他服务器用途。

使用 CCCs 直接或广播时提供自目标装置或控制器装置的优点，也能降低接口能耗，目前为 1.0V IO。

Introduction cont'd

- The DDR5 SPD interface transitioned from I2C to MIPI I3C based on the following requirements for the next generation DDR DIMM technology:
 - Lower IO operating Voltage (as low as **1V** aligned to advanced process node)
 - DDR4 SPD IO voltage was **2.5V**
 - Higher interface bit rate (**400KHz** to **8-12.5MHz** in real applications) due to the increased number of devices per DIMM to be managed
 - DDR4 had **two** devices per DIMM vs **five** devices in DDR5
 - Considering 8 DIMMs per SPD segment, this is 16 vs 40 devices
 - Higher bit rate to reduce boot time (diminishing Memory Reference code execution time)

Juan Orozco: DDR5 SPD 接口从 I2C 转变为 MIPI I3C，是为因应下一代 DDR DIMM 技术的下列需求，第一，降低 IO 运作电压，最低可降至 1V，符合先进硅工艺节点，前一代 DDR4 SPD 电压为 2.5V IO，和这项先进硅工艺已不相容。

未来二至四年都需要 1.0V IO。接口比特率也较高，DDR4 使用二项 DIMM，而 DDR5 应用在五部装置中，装置数量增加，就需要更高比特率，若考虑到每项 SPD 最高可达 8 DIMM，DDR4 应用可能最高管理 16 部装置，DDR5 应用更高达 40 部装置，等于超过一倍。

实际建置时，I2C 是否可达 400KHz，受限于两项因素，即为信号开漏特质，若要突破 400KHz，必须提高时序及升压电容；其次，I2C MoCs [ph?]等标准化 I2C 装置受限于 400KHz，I3C 实际应用可达 8-12.5MHz，取决于接口上的对象数量。

提高比特率亦可缩短计算机开机时间，减少内存参考代码需花在整体 bin 复原流程的时间。



28-29
SEPTEMBER
2021

**DDR5 SPD Server connectivity and
bus characteristics**

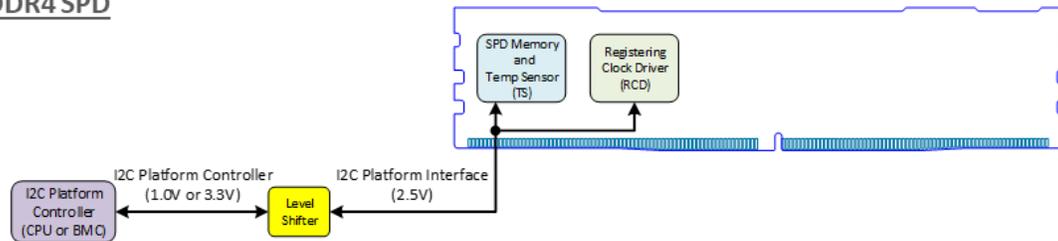
[MIPI.ORG/DEVCON](https://mipi.org/devcon)

MOBILE & BEYOND

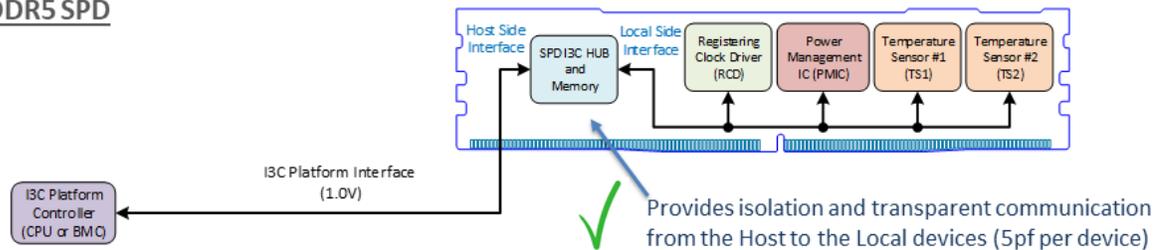
© 2021 MIPI Alliance, Inc.

DDR4 vs DDR5 SPD DIMM Connectivity

DDR4 SPD



DDR5 SPD

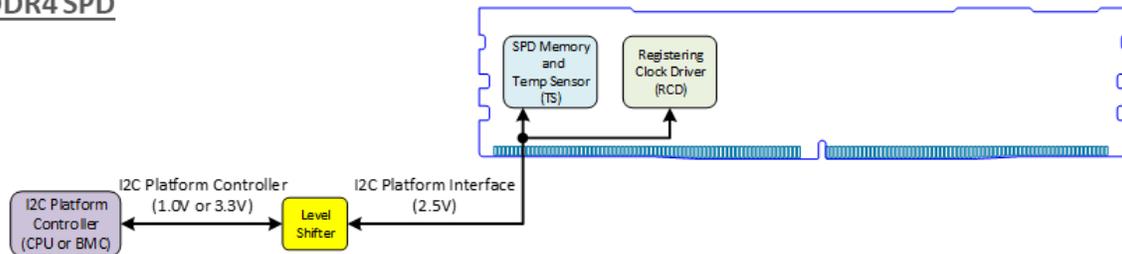


Juan Orozco: 请看这项 DDR5 SPD 在服务器主板的联机情况，以 DDR4 为例，一项是 SPD 内存及温度传感器，另一项是暂存频率驱动器，这两项装置与 I2C 接口平台互连时，是通过至主机 CPU 或主板控制器等目标控制器装置的 2.5V IO，由于各装置的运作电压为 1V 或 3.3V，中间需要电压电平转换器。

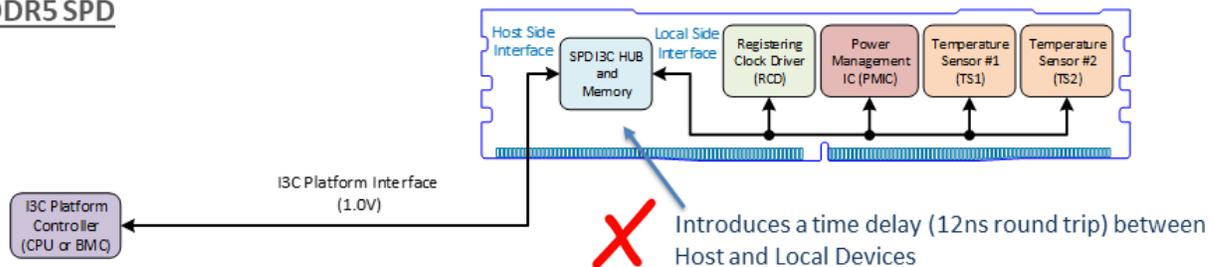
反观 DDR5 SPD，CPU 或 DMC 等平台控制器直接通过 1V I3C IO，与 DIMM 接口相连，DIMM 内共有五部装置，其中一项为 SPD I3C HUB，在主机 ID 接口与 DIMM 上的接口之间提供电子隔离，让平台上的控制器只会看到一部装置的电子负载，而非同时五部装置的电子装置，如此最能够降低整体电容。

DDR4 vs DDR5 SPD DIMM Connectivity

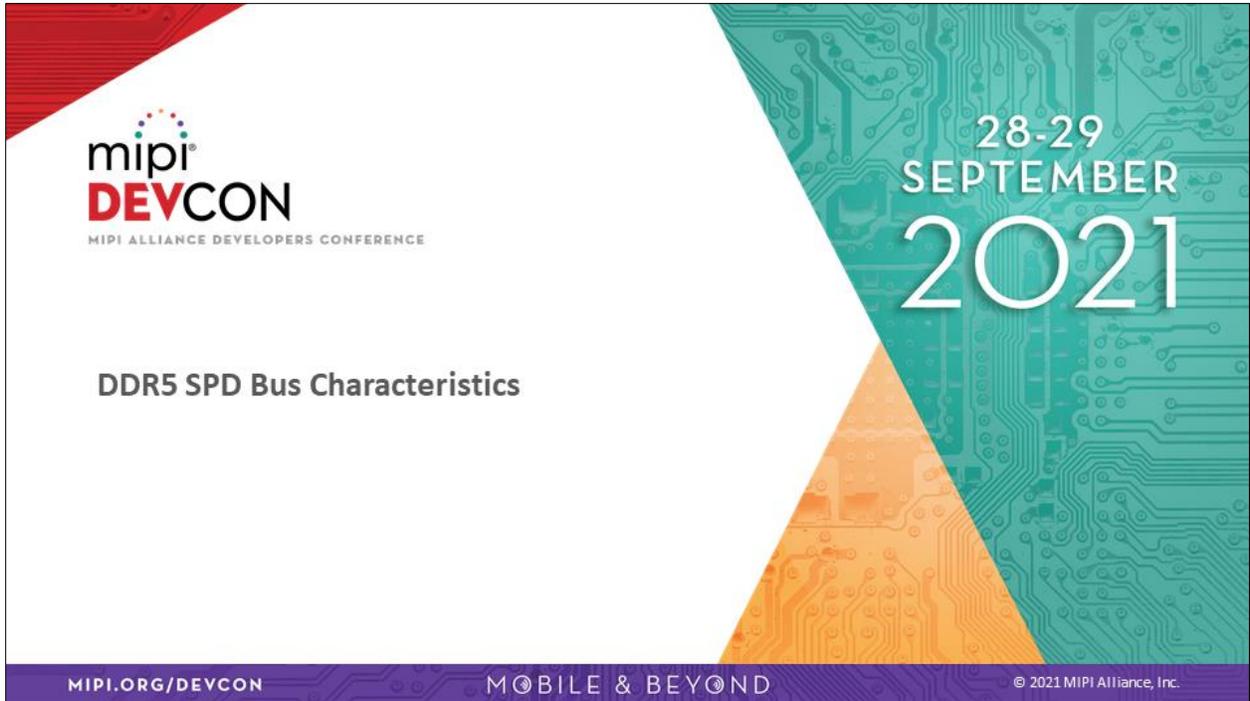
- DDR4 SPD



- DDR5 SPD



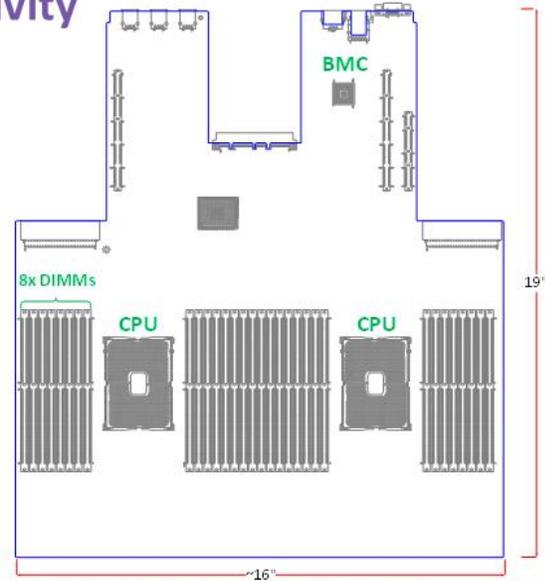
Juan Orozco: 这种方式的缺点在于，SPD HUB 会造成单向六纳秒的时间延迟，等同于控制器与目标装置来回共 12 纳秒，这是总线电容必须牺牲之处，在分析控制器驱动时序时，就必须考虑 12 纳秒的延迟，接下来请 Nestor 继续说明。



Nestor Hernandez: 感谢 Juan 开场，接下来介绍 SPD 总线特质。

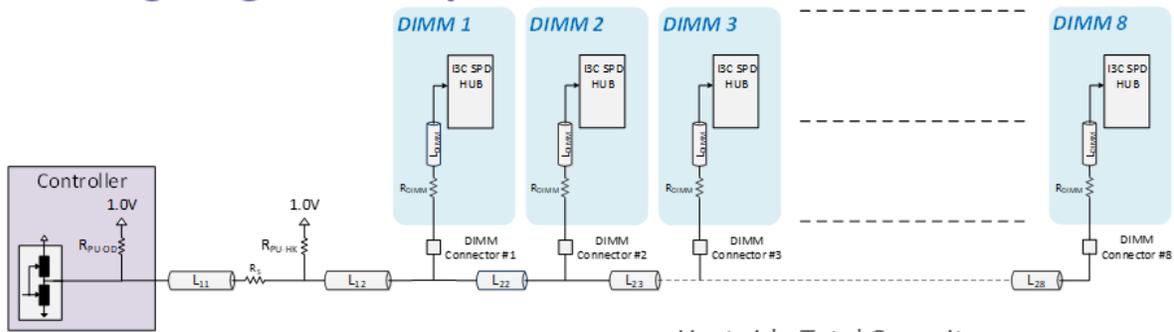
DDR5 SPD Platform Connectivity

- Host side Server PCB routing
 - Total length could exceed 50"
 - Server motherboards are BIG: up to 16"x 19"
 - BMC controller is located at the North side
 - DIMMs are located at the South side
 - CPU to DIMM SPD routing has lower priority than high speed IO routing (DDR5, PCIe G4/G5, etc.)



Nestor Hernandez: 如图表所示，主要问题在于如何配置零组件，这是一大难题，因为 CPU 和单元相近，但 BMC 通常较远，若电路板尺寸为 16 乘 19 吋，线路长度很容易就长达 50 吋，由于 SPD 非优先事项，线路会先连结 PCIe、DDR 及其他零组件，这是我们要克服的第一项挑战。

Routing length and capacitance



- Host side Server PCB routing
 - From Controller to DIMMs ($L_{11}+L_{12}+L_{2x}$): $\sim 25''$
 - DIMM routing (L_{DIMM}): 3.5"/DIMM, 28" 8x DIMMs
 - Total PCB trace length: $\sim 53''$
- Host side Total Capacitance
 - Each device apport 5pF
 - 1 CPU + 8 HUBs = 45pF
 - PCB routing is $\sim 3\text{pF/in}$
 - $53'' * 3\text{pF} = \sim 159\text{pF}$
 - **Totaling:**
 - Dev (45pF) + PCB (53pF) = **$\sim 204\text{pF}$**

Nestor Hernandez: 下一项挑战为实际链接，因为有许多分支，8 项 DIMM 代表线路非常长，若所有 DIMM 和线路的负载加总之后，电容大约有 200 Bus，而驱动器必须随时管理，因此下一项问题是，这些电容总和该如何相容？



28-29
SEPTEMBER
2021

I2C and MIPI I3C Retro-compatibility Challenges

[MIPI.ORG/DEVCON](https://mipi.org/devcon)

MOBILE & BEYOND

© 2021 MIPI Alliance, Inc.

I2C and MIPI I3C Retro-compatibility Challenges

- There are three operating modes supported by the I3C protocol:
 - I2C mode with Open-Drain(OD) buffer class.
 - I3C mode with Open-Drain buffer class.
 - I3C mode with Push-Pull(PP) buffer class.
- The OD class requires a pullup to set a stable “Logic-high”.
 - The pullup is set accordingly with the total capacitance on the bus.
 - High capacitance busses requires a “**Strong pullup**”
 - Strong pullup guarantees rise time specification to pass.
- The PP class requires a High-Keeper pullup.
 - A “Weak pullup” is required to the target device with low current can pull SDA signal low within a minimum low period.
 - Weak pull-up lessens the voltage levels disturbances

MIPI I3C Basic Spec requires Dynamic pull-up control to switch between “strong pull-up” and “weak pull-up” to optimize open-drain and push-pull timing requirements.

Nestor Hernandez: 第一项障碍是，由于 I3C 必须与 I2C 兼容，我们必须深究 I2C 的需求为何，I2C 必须为开漏缓冲级，需要强拉，开漏需要强拉，必须锁定在低高阻。I2C 需要推挽缓冲，因此又需要弱拉，才能实际移动低阶，因此 MIPI I2C 基本规格内需要动态上拉电阻，以切换开漏所需的强拉，以及新 I2C 模型内推挽所需的弱拉。

Non-Dynamic Pullup impact in a 204pF bus

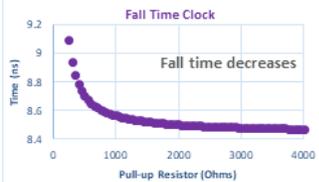
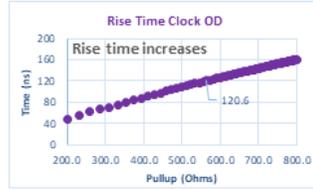


← Push-Pull

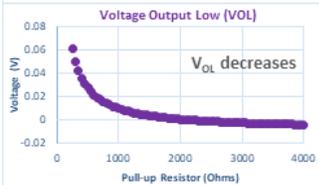
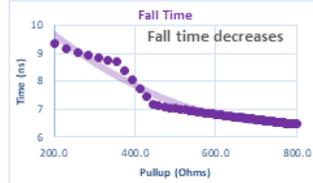
Open-Drain →

The Highest the PU

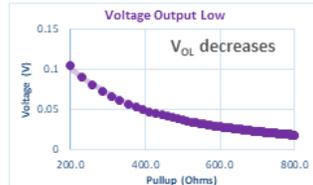
- VIH never reached with pull-up higher than 800Ω
- Limit max operating frequency



- On-Board PU can guarantee an OD max operating frequency.
 - A parallel equivalent $R_{PU_HK} || R_{PU_OD}$ of 333.3Ω
 - Rise time=75.3ns
 - A pullup $\geq 550\Omega$ negatively affects both rise time and operating frequency



A trade-off among pull-up value, rise time and V_{OL} is required to meet the highest operating frequency



Nestor Hernandez: 运用这三项数据，要掌握上升时间不容易，其次是下降时间及 V_{OL} ，右侧开漏的电阻数值较低，左侧的推挽需要较高电阻数值，若对照这两项要求，会找到交集点，通常交点大约 303 欧姆，上升时间大约 75 纳秒，再加上推挽，就能找到符合最高运作效率所需的 V_{OL} 安全数值。



28-29
SEPTEMBER
2021

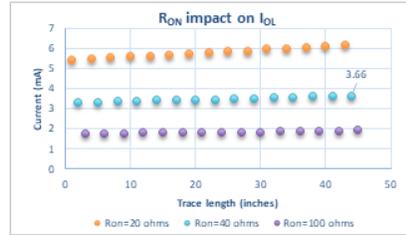
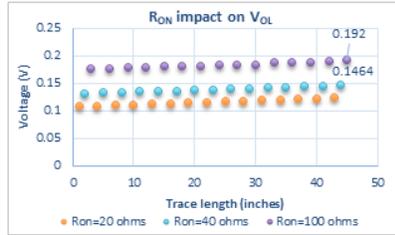
Buffer R_{ON} design implications

[MIPI.ORG/DEVCON](https://mipi.org/devcon)

MOBILE & BEYOND

© 2021 MIPI Alliance, Inc.

Buffer R_{ON} value design implications

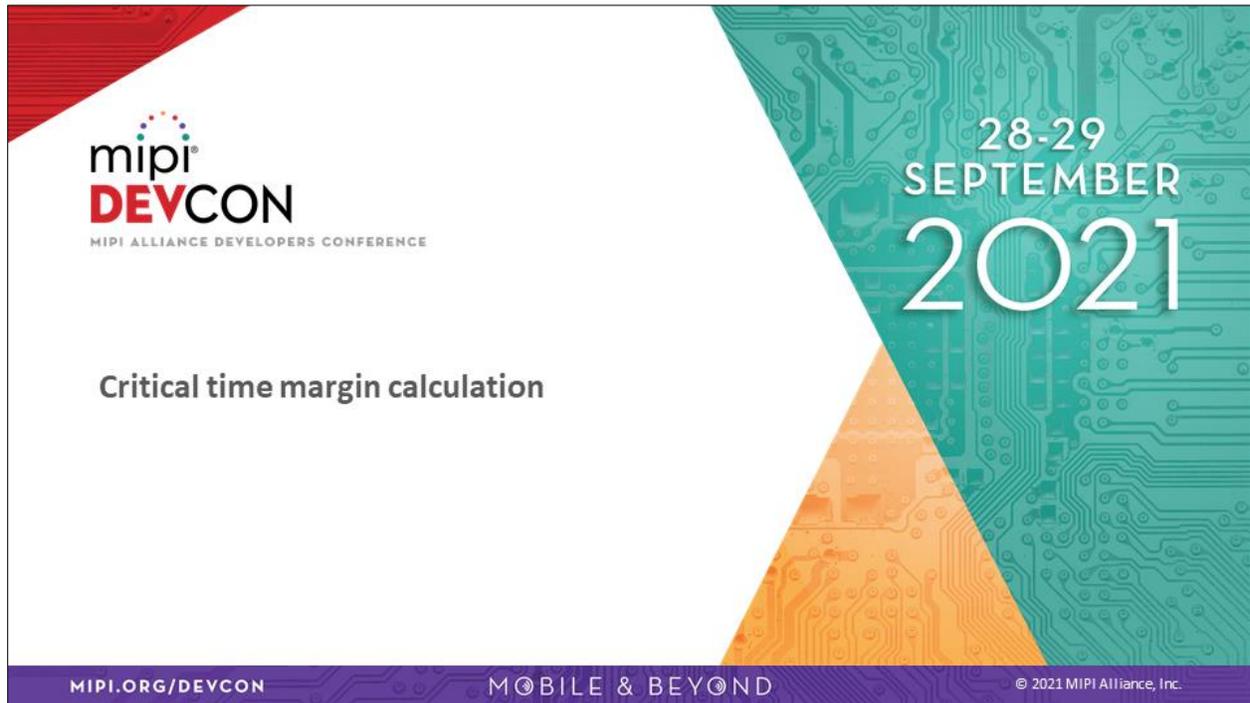


- The bigger the R_{ON} the higher the V_{OL} is:
 - Increasing trace length results in higher V_{OL}
 - With the longest trace length, $V_{OL}=192\text{mV}$,
 - Assuming $V_{IL}=0.3\text{V}$ then the transition margin is 108mV
 - Low transition margin can cause idle states
 - Setting the R_{ON} at $40\ \Omega$ reduces the V_{OL}
 - With the longest trace length $V_{OL}=146\text{mV}$,
 - If $V_{IL}=0.3\text{V}$ then the transition margin is 154mV
 - Notice that at the longest trace length with $V_{OL}=146\text{mV}$ the I_{OL} is 3.66mA

By limiting R_{ON} into a max range of $40\ \Omega$ ensures a healthy V_{OL} by setting a max I_{OL} bigger than 3mA

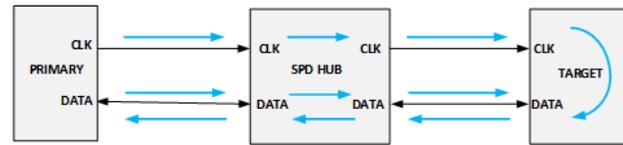
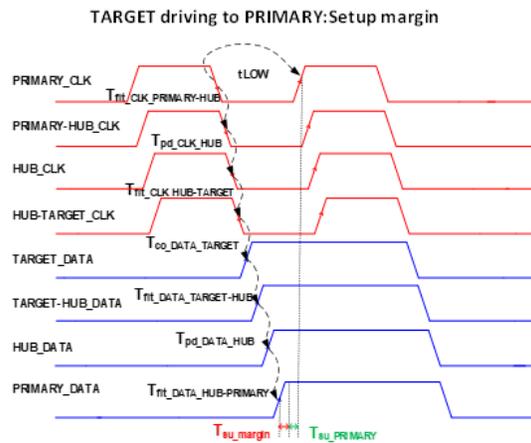
Nestor Hernandez: 驱动器必须驱动，总线必须为强驱动器，依据走线长度，会需要具备 40 欧姆的强驱动器，当线路达到 50 英寸左右，仍然能够驱动电容，但会面临一些问题，若最长线路无法使用 V_{OL} ，就会出现约 146mV；若 V_{OL} 线路负载为 300，则会有 150mV。各位要注意到使用的电流为 3.60 mA，因此 R_{ON} 必须设定上限为 40 欧姆，以达到适当的 V_{OL} ，让电流最高不会超过 3 mA，这将是 I3C 设计缓冲的主要特质。

接下来可进入实际联机计算。



Azusena Lupercio: 谢谢 Nestor, 接下来讨论临界时间余裕的计算。

Critical time margin calculation



TARGET driving to PRIMARY: Setup margin

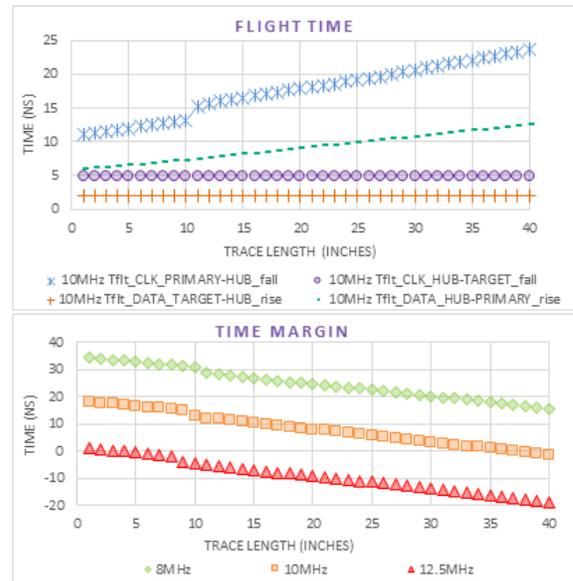
$$T_{su_mar} = t_{LOW} - (T_{flt_CLK_PRIMARY \rightarrow HUB_fall} + T_{pd_CLK_HUB} + T_{flt_CLK_HUB \rightarrow TARGET_fall} + T_{co_DATA_TARGET} + T_{flt_DATA_TARGET \rightarrow HUB_rise} + T_{pd_DATA_HUB} + T_{flt_DATA_HUB \rightarrow PRIMARY_rise}) - T_{su_PRIMARY_max}$$

Azusena Lupercio: 为了找出高总线电容系统的正确操作频率，必须检查目标装置驱动至首要装置时的设定时间余裕，可通过下列算式计算出来，算出信号周期内较低的部分后，再减去整条路径内所有延迟的参数，包括从首要装置到枢纽装置的频率飞行时间传播延迟、枢纽的频率传播延迟、从枢纽到目标的频率传播延迟、目标的数据 TCO、目标到枢纽的数据飞行时间、枢纽的数据传播延迟、枢纽到首要的数据飞行时间，以及首要装置的设定时间。

Critical time margin calculation

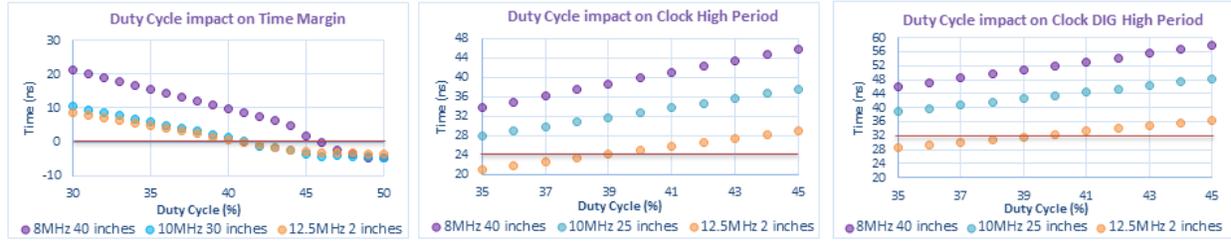
Frequency	10	MHz
Duty cycle	35	%
tLOW	65	ns
tflt_CLK_PRIMARY-HUB_fall	19.1	ns
tpd_CLK_HUB	6	ns
tflt_CLK_HUB-TARGET_fall	4.7	ns
tco_DATA_TARGET	12	ns
tflt_DATA_TARGET-HUB_rise	1.972	ns
tpd_DATA_HUB	6	ns
tflt_DATA_HUB-PRIMARY_rise	9.8	ns
tsu_PRIMARY_max	3	ns
Setup Margin	2.253	ns

- The longer trace length the biggest flight time
- Inner device propagation delay plays a significant role in defining the operating frequency.
 - The highest the Tpd the bigger the time margin reduction.
- Increasing tLOW provides extra timing margin.



Azusena Lupercio: 各位心中有了这些算式之后，就能厘清该检查哪些重要参数，以找出各项操作频率，例如取决于走线长度的飞行时间、各项装置的内部传播延迟、tLOW 等，只要走线长度愈长，飞行时间就会增加，因此走线长度若很长，飞行时间就会很长；传播延迟愈高，时间余裕就会缩短，但增加 tLOW 可提供额外时间余裕，下方图表内共有三种操作频率：12.5MHz、10 MHz、8 MHz，红色是 12.5 MHz，即使走线长度很短，时间余裕也很短，代表信号传播的时间不足；10 MHz 的解决方案空间约有 35 吋；8 MHz 的高电容总线解决方案空间足够，可大于 40 吋。

Frequency and AC/DC parameters impact



- Increasing duty cycle reduces t_{LOW} , thus reducing the Time Margin.
- When reducing the Duty Cycle the t_{HIGH} and t_{DIG_HIGH} are affected.
 - Small Duty Cycle can produce a NOT PASS on t_{HIGH}/t_{DIG_HIGH} .

A correct selection of Duty Cycle provides extra time margin to complete the setup transaction, granting higher operating frequency.

From MIPI I3C Spec t_{HIGH} min 24ns, t_{DIG_HIGH} min 32ns

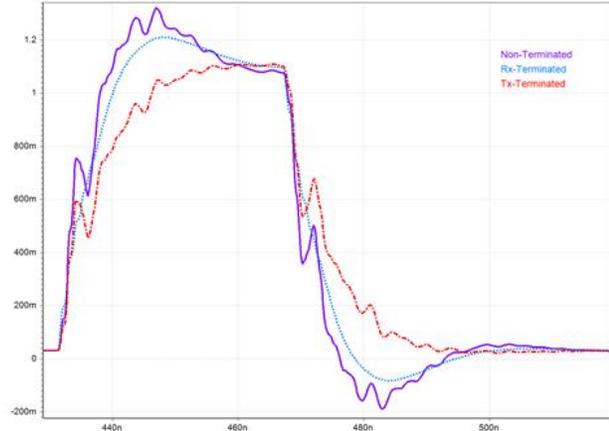
Azusena Lupercio: 了解 t_{LOW} 为重要参数之后, 可增加另外三项新参数, 例如工作周期、 t_{HIGH} 、 t_{DIG_HIGH} , 工作周期变长, 会缩短 t_{LOW} , 也会减少时间余裕; 若缩短工作周期, t_{HIGH} 和 t_{DIG_HIGH} 会受影响, 换言之, 工作周期短无法符合 t_{HIGH} 和 t_{DIG_HIGH} 规格, t_{HIGH} 至少需 24 纳秒, t_{DIG_HIGH} 至少需 32 纳秒, 若选择适当工作周期, 不仅能提供额外时间余裕, 可完成设定作业, 也能确保操作频率更高。



Azusena Lupercio: 非单调性信号行为

Non-monotonic signal behavior

- Termination effect on transmission lines
 - Non-terminated circuit:
 - Signal bounces back and forth between the driver and the receiver.
 - Tx-terminated circuit:
 - Reduces drive strength
 - Increases propagation delay
 - Limits buffer capabilities
 - Rx-terminated circuit:
 - Reduces bouncing effect
 - Increases propagation delay



Azusena Lupercio: 非单调性信号可通过决定性效果排除，若无终接线路，信号会不断在驱动器和接收器之间来回，由于信号存在非单调性，传输器的终接线路降低驱动器强度、增加传输延、限制缓冲能力；这项信号无单调性存在，传输器的终接线路减少反弹效果，但频率传输延迟很高，蓝线代表无单调性存在。

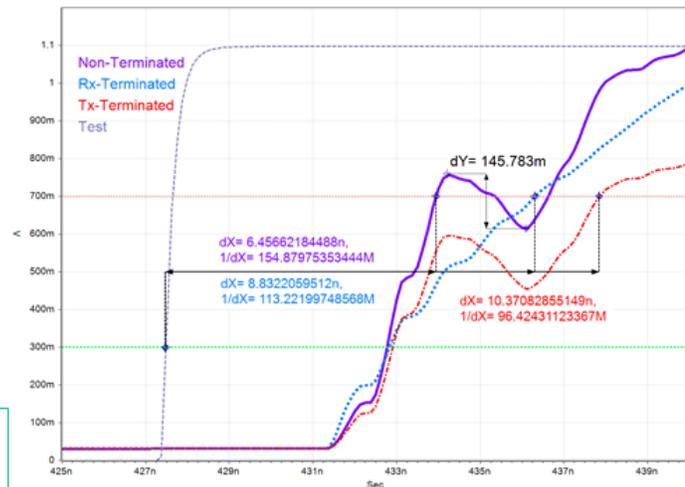


Azusena Lupercio: 斜率反转能力与时序改善

Slope reversal capability and timing improvement

- With the non-deterministic loading of an unterminated bus, there can be reflections on the bus causing slope reversal on the Rx signal.
- By sampling at the first threshold is possible to filter Non-Monotonicity's; Schmidt triggered inputs
 - Non-terminated VS Rx-Terminated: **Improves 2.3ns**
 - Non-terminated VS Tx-Terminated: **Improves 3.92ns**

Slope reversal capability provides additional time margin that improves operating frequency and prevent false logic states.



Azusena Lupercio: 由于无终接线路总线的非决定性负载，可能反射在总线上，造成接收器信号斜率反转，若在第一门坎取样，有可能过滤非单调性或施密特触发器输入，若比较终接与无终接接收器，飞行时间可改善约 2.3 纳秒；若比较无终接线路和终接传输器，飞行时间可改善约 3.9 纳秒，斜率反转能力提供额外时间余裕，可改变操作频率，也避免错误逻辑状态。



Azusena Lupercio: 总结

Summary

- I3C Applications in Server systems (such as DDR5 SPD) are dealing with higher Bus capacitance than the max limit assumptions in MIPI spec (for 12.5MHz capable buses).
- Higher Bus capacitance applications can be mitigated by using good Buffer Drive strength, strong open-drain class pull-up, and HUB isolation circuits.
- A dynamic pullup operation allows to drive the interoperability challenges between the open-drain and push-pull operating modes; by enabling higher operating frequencies on both modes and limiting critical parameters to meet latest specification.
- Strong buffers tend to increase signal energy reflections, specially in complex topologies resulting with slope reversal conditions at Devices' Inputs.
- Schmitt trigger capable inputs are required in order to mitigate slope reversal conditions when dealing with high bus capacitance and strong buffers

Azusena Lupercio: 在 DDR5 SPD 等服务器系统内，I3C 需处理的总线电容，大于 I3C MIPI 规格针对 12.5MHz 总线的高低限制，应用总线电容较高时，缓和途径包括缓冲驱动器强度、强开漏、上拉电阻、HUB 隔离电路，动态推挽操作可改善开漏与推挽操作模式之间的互用性，提高两种模式的操作频率，并限制重要参数，以符合最新规格。强大缓冲会增加信号能量反射，尤其在复杂组态中，导致装置输入出现斜率反转情况。需要施密特触发器输入，以缓和斜率反转情况，也同时处理高总线电容及缓冲。



Azusena Lupercio: 感谢大家聆听，接下来欢迎各位提问。