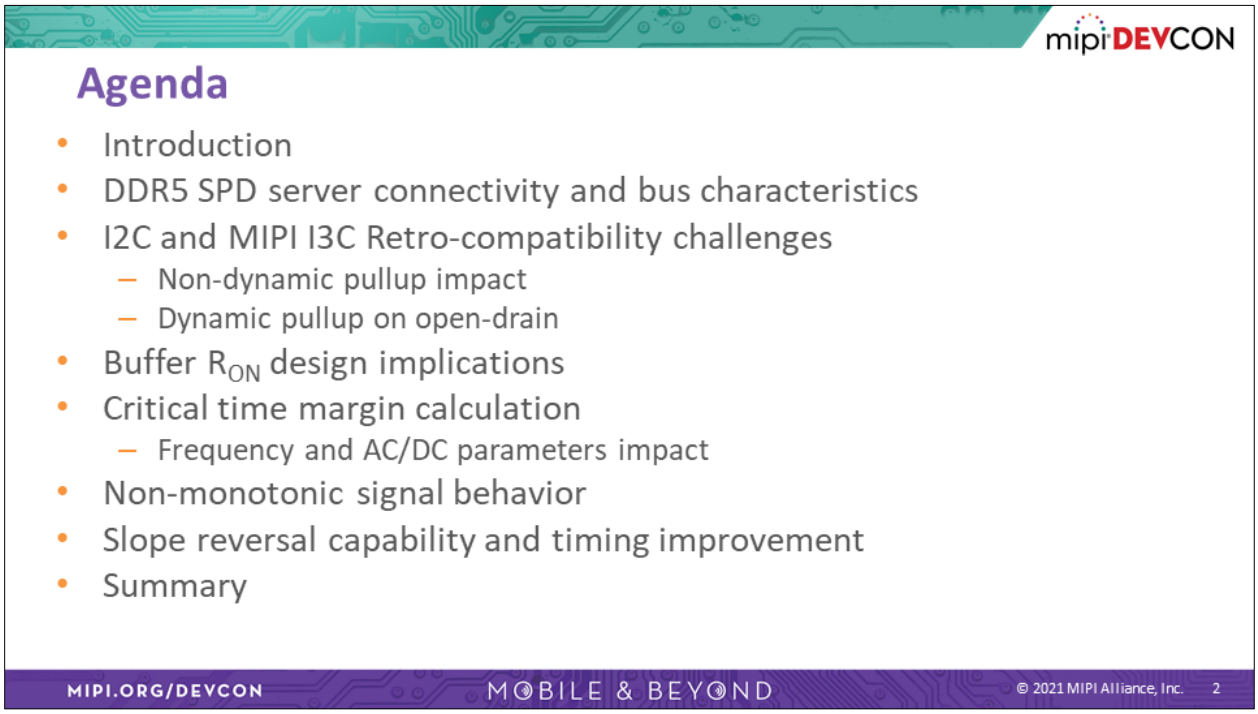
DDR5 伺服器平台解決方案內的 I3C 訊號完整性挑戰

主講者：Azusena Lupercio, Juan Orozco, Nestor Hernandez

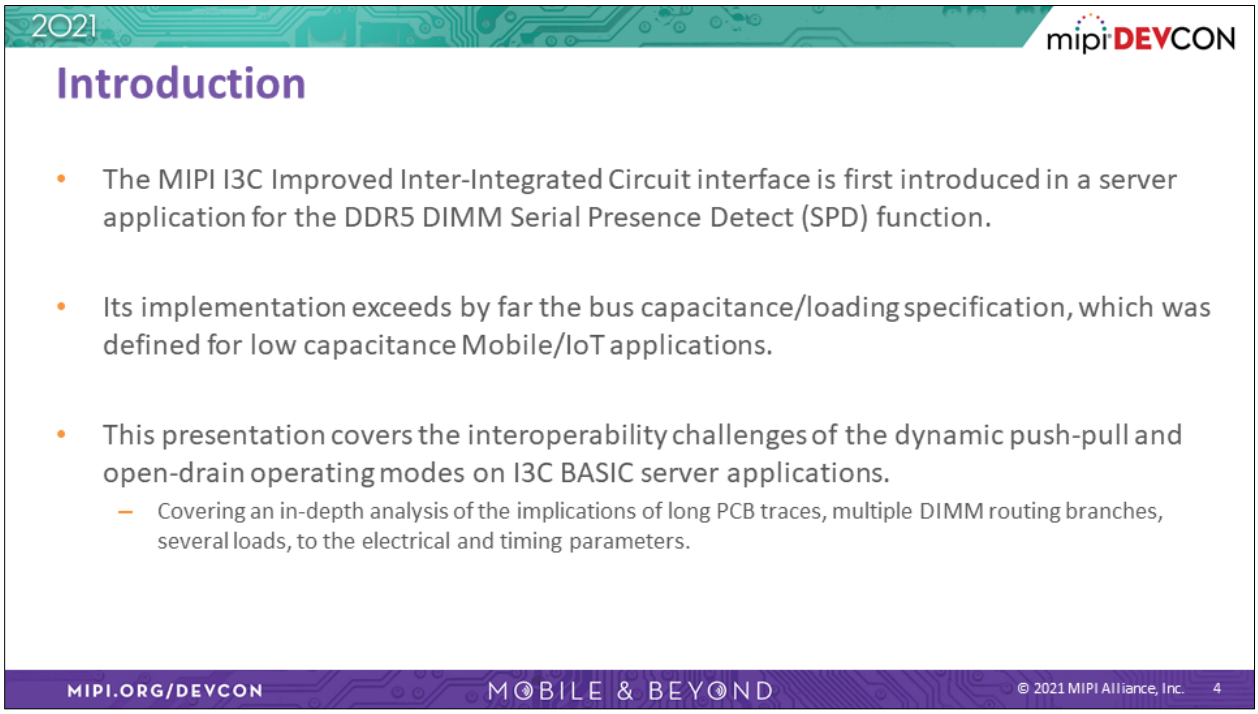
英特爾

Juan Orozco: 大家好，我是 Juan Orozco，我和兩位同事 Azusena Lupercio 及 Nestor Hernandez 都任職於英特爾，將介紹 DDR5 伺服器平台解決方案內的 I3C 訊號完整性挑戰。

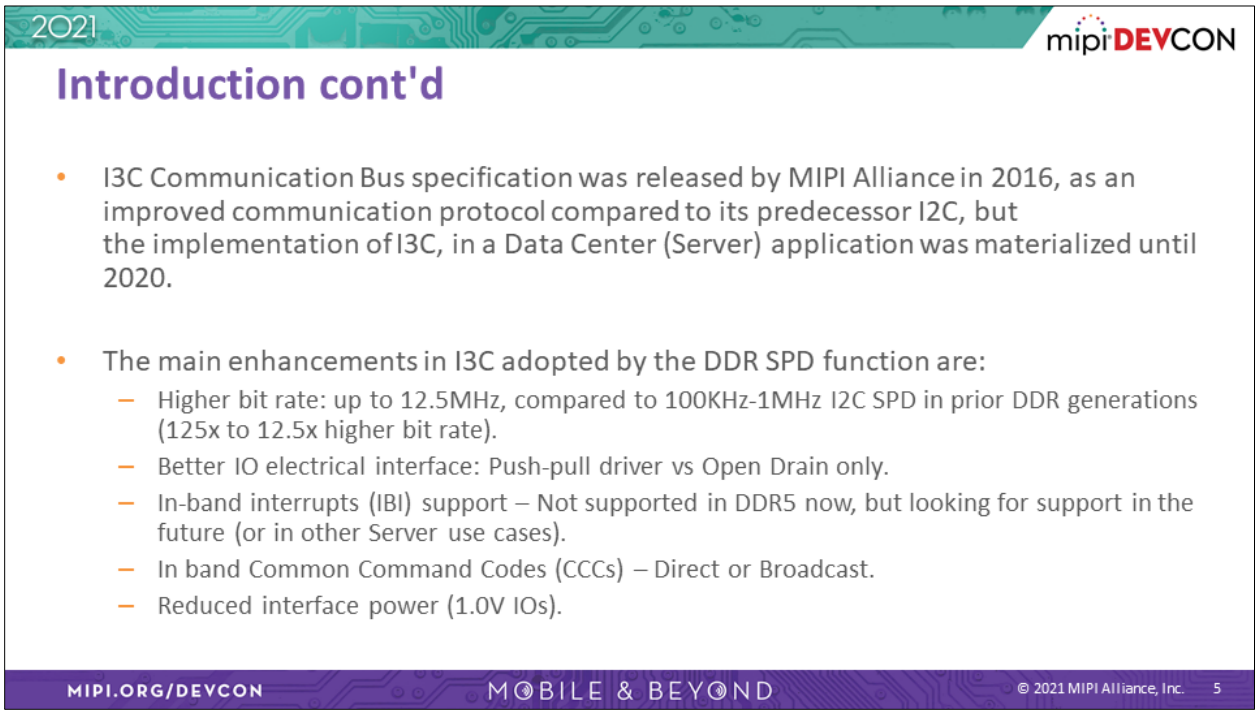


Juan Orozco: 今天簡報在開場之後，會簡介 DDR5 SPD 伺服器連線與匯流排的特質，再說明 I2C 與 MIPI I3C 向下相容的挑戰，以及 I3C 控制器與目標裝置緩衝對系統設計的意涵；並深入探討臨界時間餘裕計算，以及不同 AC/DC 參數對操作頻率有何影響。也會呈現我們所見的非單調性訊號，以及為何目標裝置需要斜率反轉能力，才能夠改善裝置的時序餘裕及邏輯適當功能。最後再做總結。


mipi
DEVCON
MIPI ALLIANCE DEVELOPERS CONFERENCE
28-29
SEPTEMBER
2021
Introduction
MIPI.ORG/DEVCON
MOBILE & BEYOND
© 2021 MIPI Alliance, Inc.



Juan Orozco: 進入正題，MIPI I3C 改良積體匯流排電路最先出現在伺服器應用中，以滿足 DDR5 雙線記憶體模組的串列存在偵測（SPD）功能，當建置在伺服器中，超越 MIPI I3C BASIC 規格內的匯流排電容，這項規格定義了高低匯流排電容、行動物聯網、小型手持式裝置應用等。今天會介紹到 DDR5 伺服器應用內的互用性挑戰，是因為 I3C BASIC 解決方案運作模式內的動態推挽與開漏特質，我們將深入分析，當主機板具有多種電路設計與多項負載時，電路板線路很長會造成什麼影響，而這些特質又會如何影響電子與時序參數。

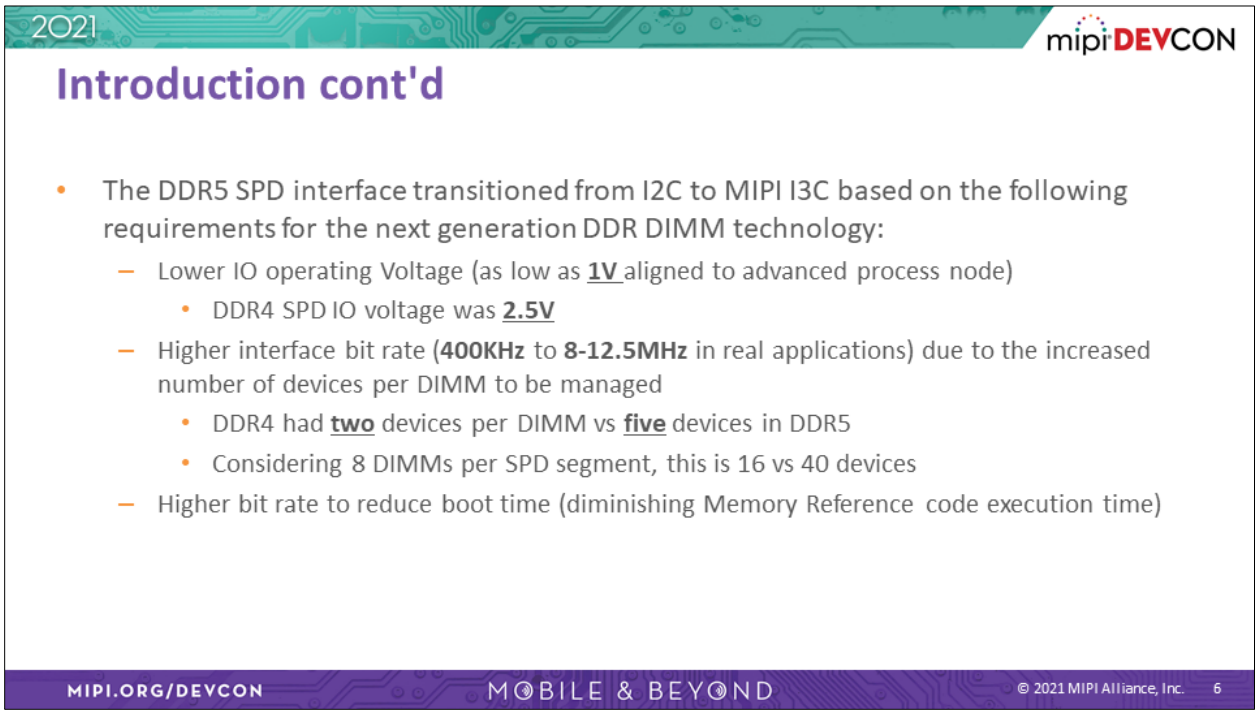


Juan Orozco: I3C 通訊匯流排規格是由 MIPI Alliance 在 2016 年發表，做為改良前代 I2C 規格的通訊協定，不過一直到首批 DDR5 伺服器問世後，I3C 才在 2020 年首度出現在資料中心的伺服器應用中。

DDR SPD 功能使用 I3C 之後，主要改良之處當然是提高位元率，理論上可達 12.5MHz，相對而言，先前 DDR 的 I2C SPD 功能位元率為 100KHz 至 1MHz，依據實際建置情況不同，位元率可提高 12.5 倍至 125 倍。

改良的 IO 電子介面具備推挽驅動器，而非開漏緩衝，因此更加適合，雖然 DDR5 SPD 目前並未使用頻內中斷（IBI）功能，但我們正在研究是否未來需支援，以符合其他伺服器用途。

使用 CCCs 直接或廣播時提供自目標裝置或控制器裝置的優點，也能降低介面能耗，目前為 1.0V IO。



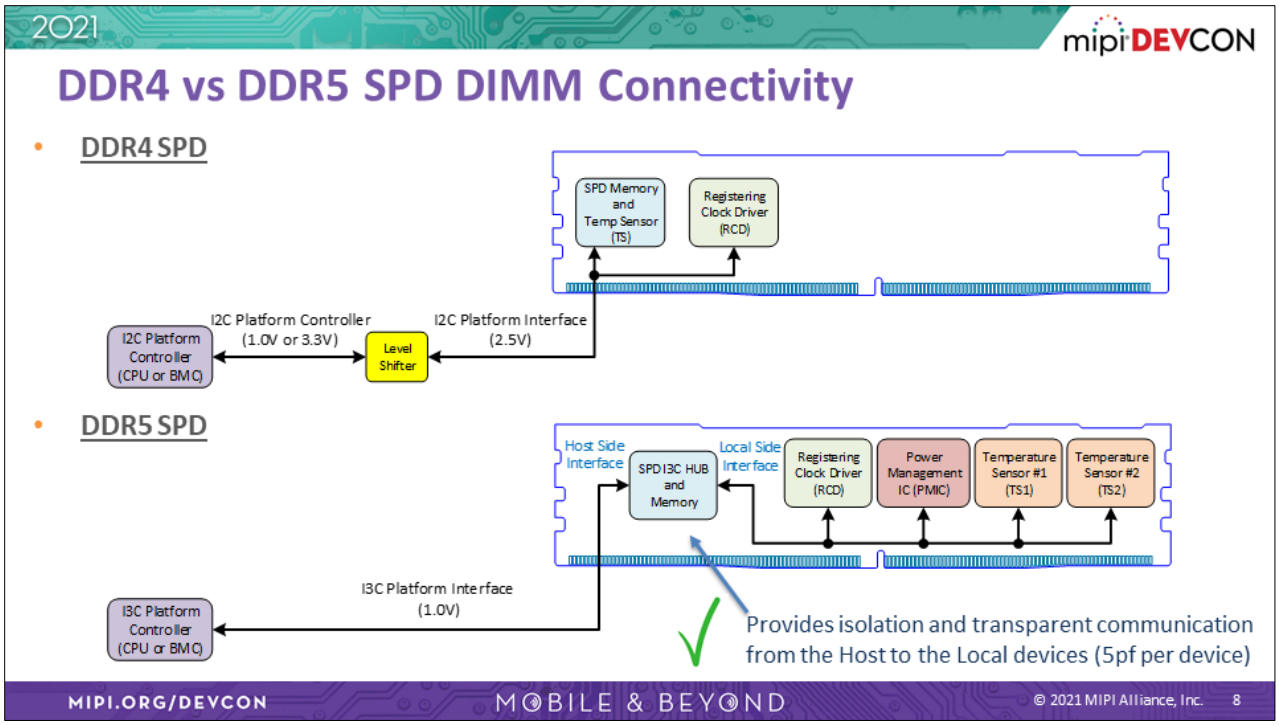
Juan Orozco: DDR5 SPD 介面從 I2C 轉變為 MIPI I3C，是為因應下一代 DDR DIMM 技術的下列需求，第一，降低 IO 運作電壓，最低可降至 1V，符合先進矽製程節點，前一代 DDR4 SPD 電壓為 2.5V IO，和這項先進矽製程已不相容。

未來二至四年都需要 1.0V IO。介面位元率也較高，DDR4 使用二項 DIMM，而 DDR5 應用在五部裝置中，裝置數量增加，就需要更高位元率，若考量到每項 SPD 最高可達 8 DIMM，DDR4 應用可能最高管理 16 部裝置，DDR5 應用更高達 40 部裝置，等於超過一倍。

實際建置時，I2C 是否可達 400KHz，受限於兩項因素，即為訊號開漏特質，若要突破 400KHz，必須提高時序及升壓電容；其次，12C MoCs [ph?]等標準化 I2C 裝置受限於 400KHz，I3C 實際應用可達 8-12.5MHz，取決於介面上的物件數量。

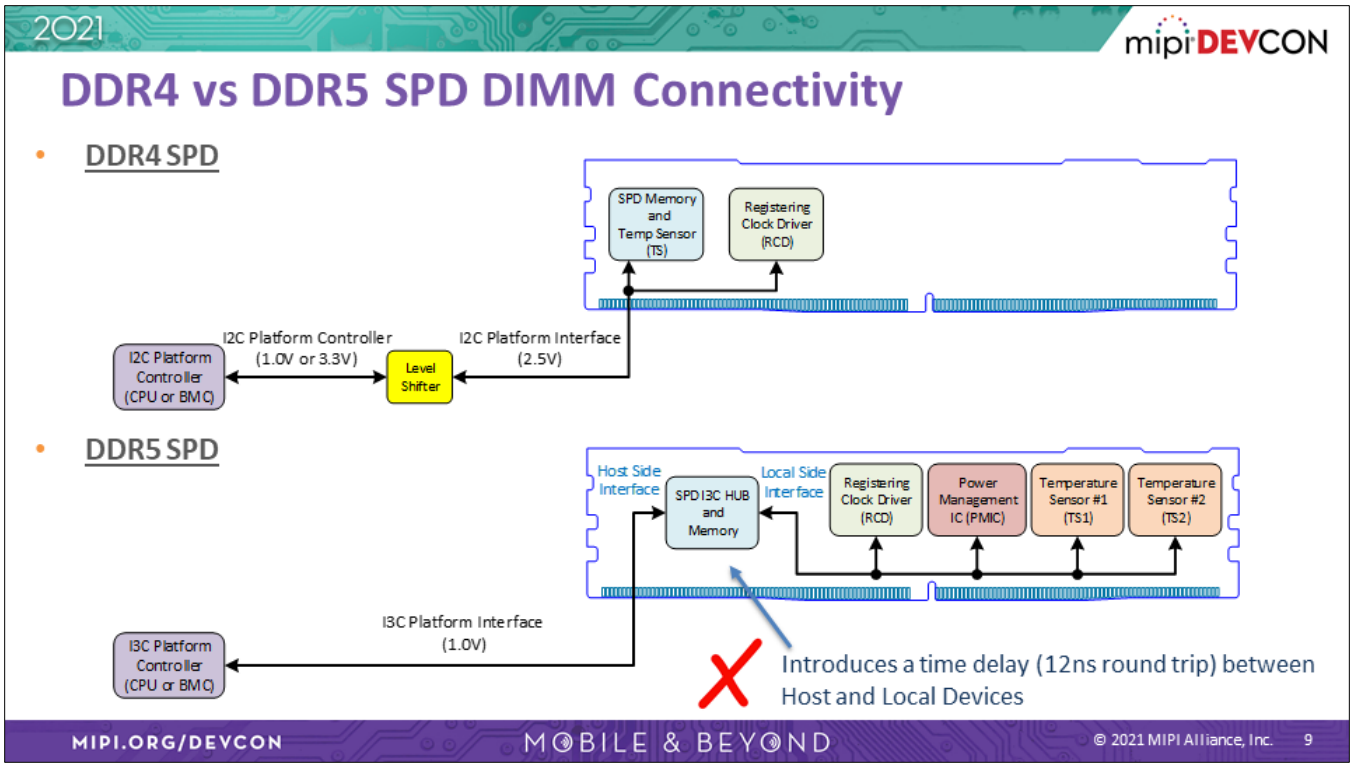
提高位元率亦可縮短電腦開機時間，減少記憶體參考代碼需花在整體 bin 復原流程的時間。


mipi
DEVCON
MIPI ALLIANCE DEVELOPERS CONFERENCE
28-29
SEPTEMBER
2021
DDR5 SPD Server connectivity and bus characteristics
MIPI.ORG/DEVCON
MOBILE & BEYOND
© 2021 MIPI Alliance, Inc.

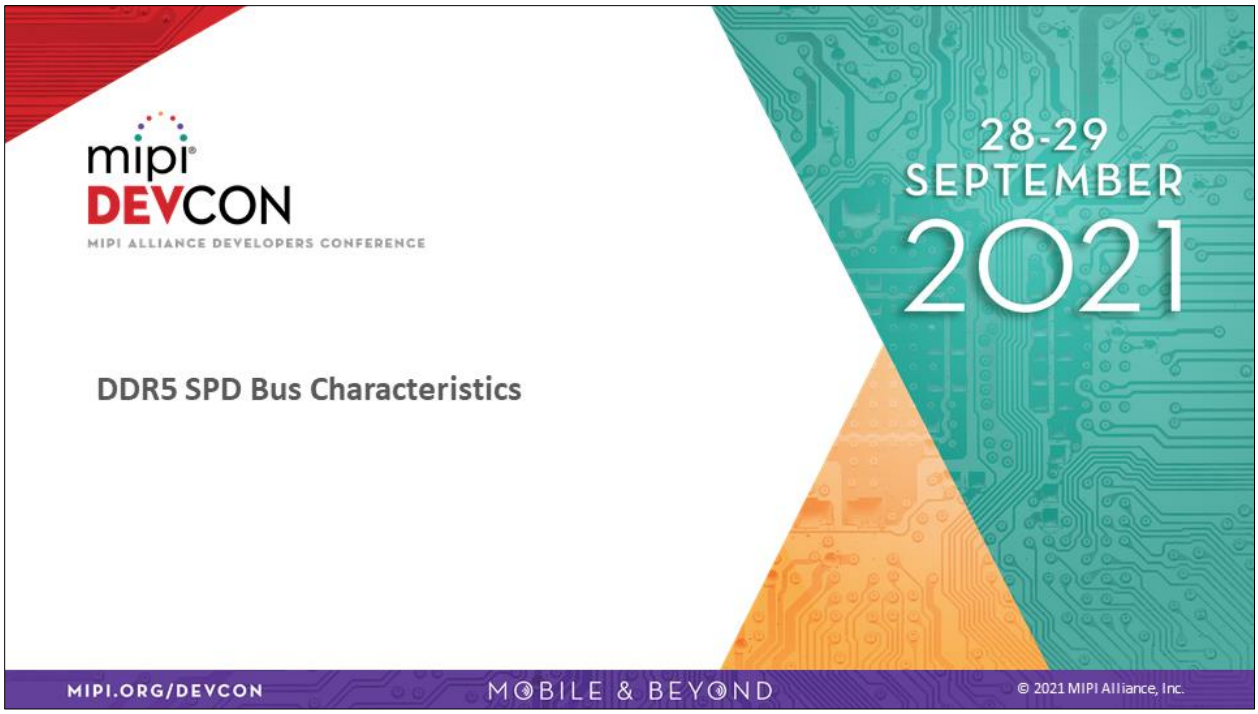


Juan Orozco: 請看這項 DDR5 SPD 在伺服器主機板的連線情況，以 DDR4 為例，一項是 SPD 記憶體及溫度感測器，另一項是暫存時脈驅動器，這兩項裝置與 I2C 介面平台互連時，是透過至主機 CPU 或主機板控制器等目標控制器裝置的 2.5V IO，由於各裝置的運作電壓為 1V 或 3.3V，中間需要電壓電平轉換器。

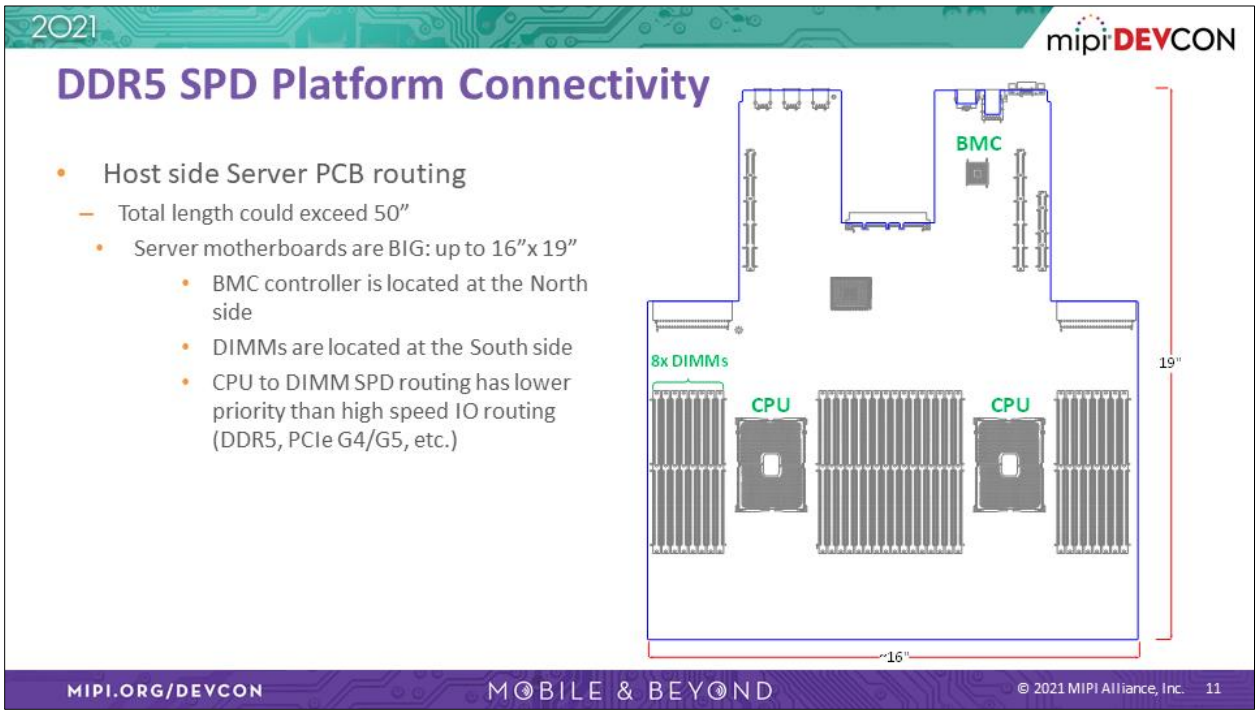
反觀 DDR5 SPD，CPU 或 DMC 等平台控制器直接透過 1V I3C IO，與 DIMM 介面相連，DIMM 內共有五部裝置，其中一項為 SPD I3C HUB，在主機 ID 介面與 DIMM 上的介面之間提供電子隔離，讓平台上的控制器只會看到一部裝置的電子負載，而非同時五部裝置的電子裝置，如此最能夠降低整體電容。



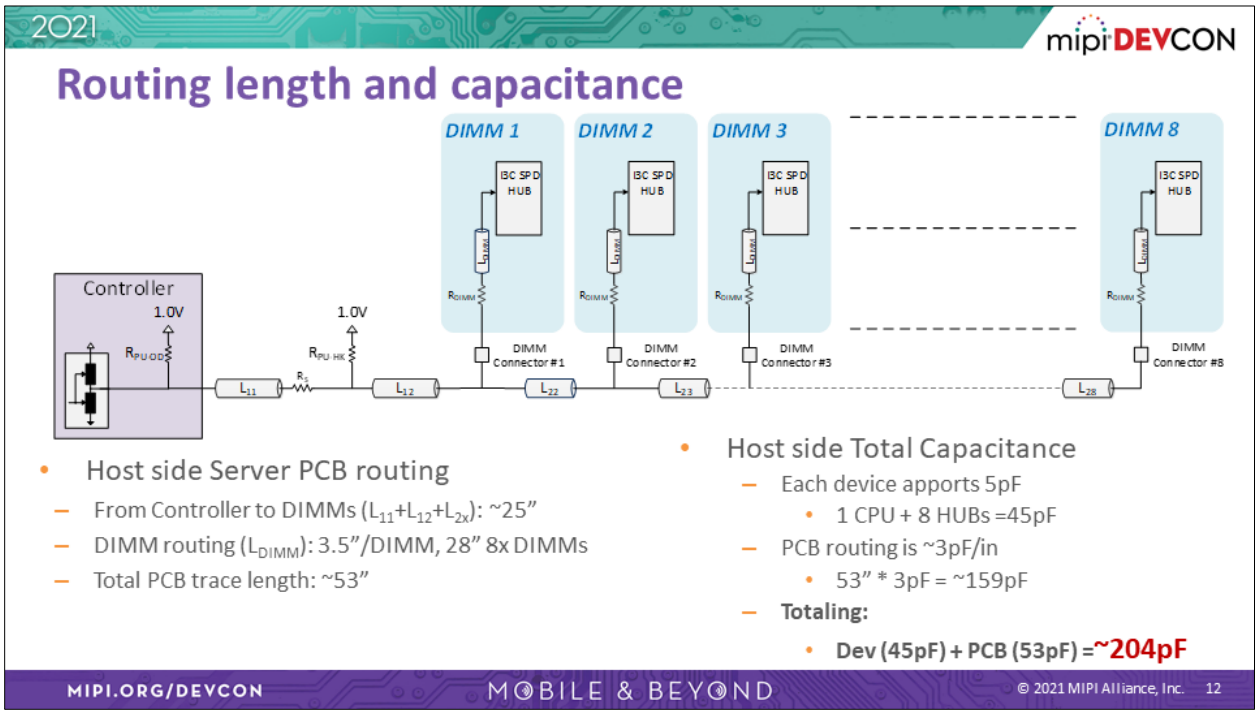
Juan Orozco: 這種方式的缺點在於，SPD HUB 會造成單向六奈秒的時間延遲，等同於控制器與目標裝置來回共 12 奈秒，這是匯流排電容必須犧牲之處，在分析控制器驅動時序時，就必須考量 12 奈秒的延遲，接下來請 Nestor 繼續說明。



Nestor Hernandez: 感謝 Juan 開場，接下來介紹 SPD 匯流排特質。

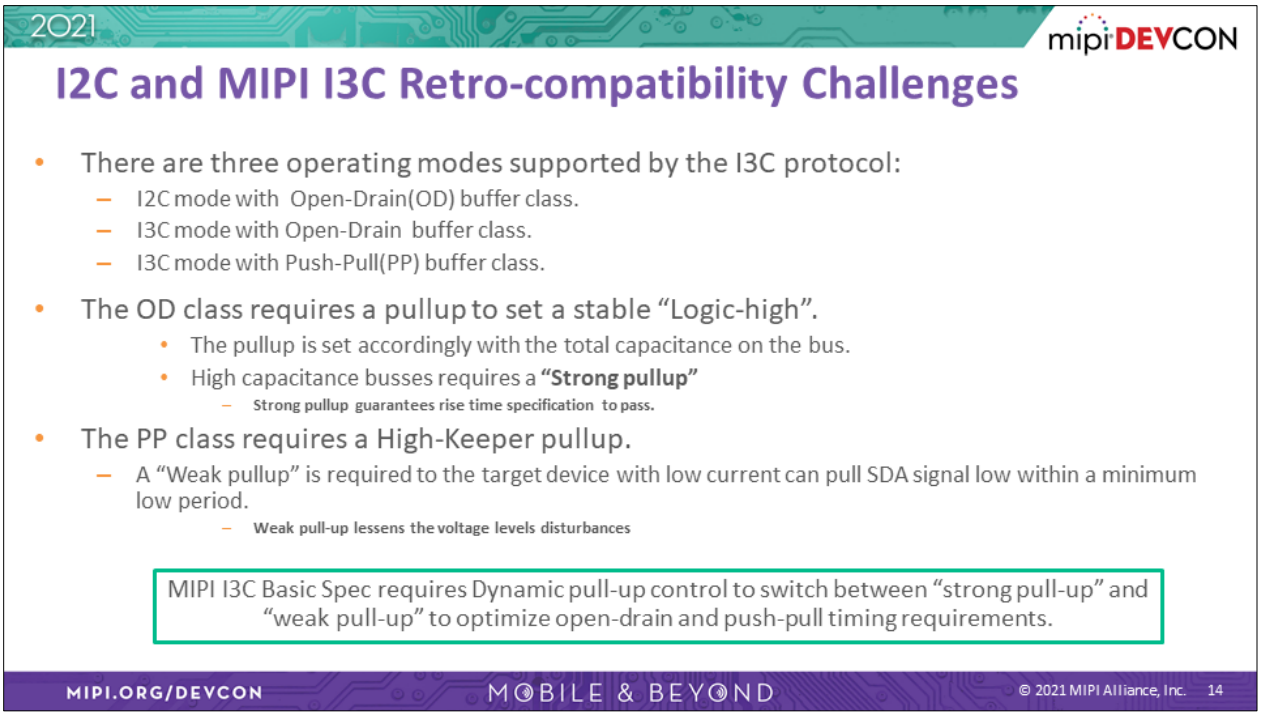


Nestor Hernandez: 如圖表所示，主要問題在於如何配置零組件，這是一大難題，因為 CPU 和單元相近，但 BMC 通常較遠，若電路板尺寸為 16 乘 19 吋，線路長度很容易就長達 50 吋，由於 SPD 非優先事項，線路會先連結 PCIe、DDR 及其他零組件，這是我們要克服的第一項挑戰。

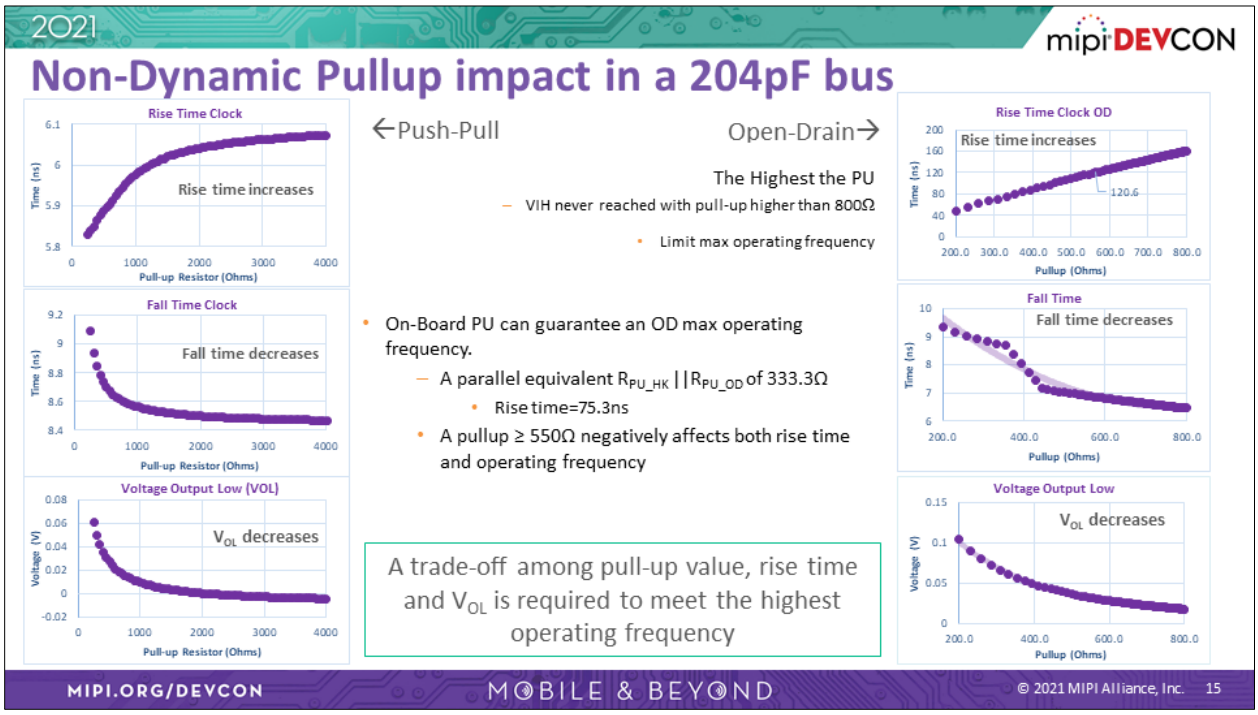


Nestor Hernandez: 下一項挑戰為實際連結，因為有許多分支，8 項 DIMM 代表線路非常長，若所有 DIMM 和線路的負載加總之後，電容大約有 200 Bus，而驅動器必須隨時管理，因此下一項問題是，這些電容總和該如何相容?


mipi
DEVCON
MIPI ALLIANCE DEVELOPERS CONFERENCE
28-29
SEPTEMBER
2021
I2C and MIPI I3C Retro-compatibility Challenges
MIPI.ORG/DEVCON
MOBILE & BEYOND
© 2021 MIPI Alliance, Inc.

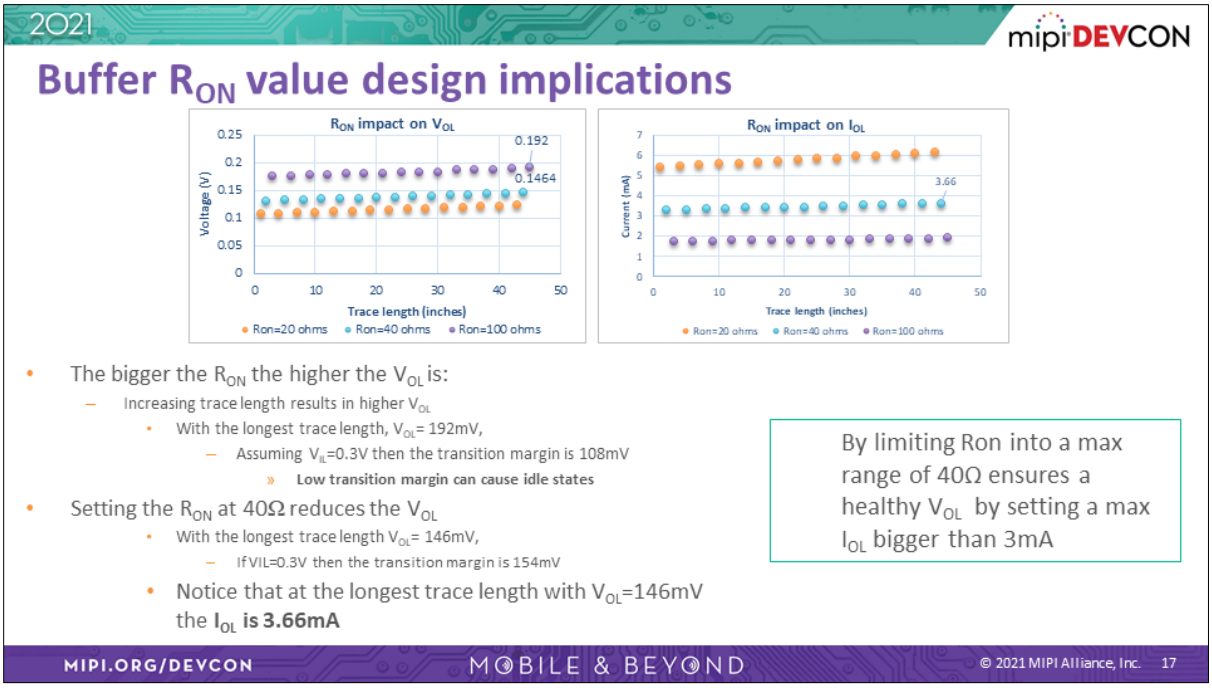


Nestor Hernandez: 第一項障礙是，由於 I3C 必須與 I2C 相容，我們必須深究 I2C 的需求為何，I2C 必須為開漏緩衝級，需要強拉，開漏需要強拉，必須鎖定在低高阻。I2C 需要推挽緩衝，因此又需要弱拉，才能實際移動低階，因此 MIPI I2C 基本規格內需要動態上拉電阻，以切換開漏所需的強拉，以及新 I2C 模型內推挽所需的弱拉。



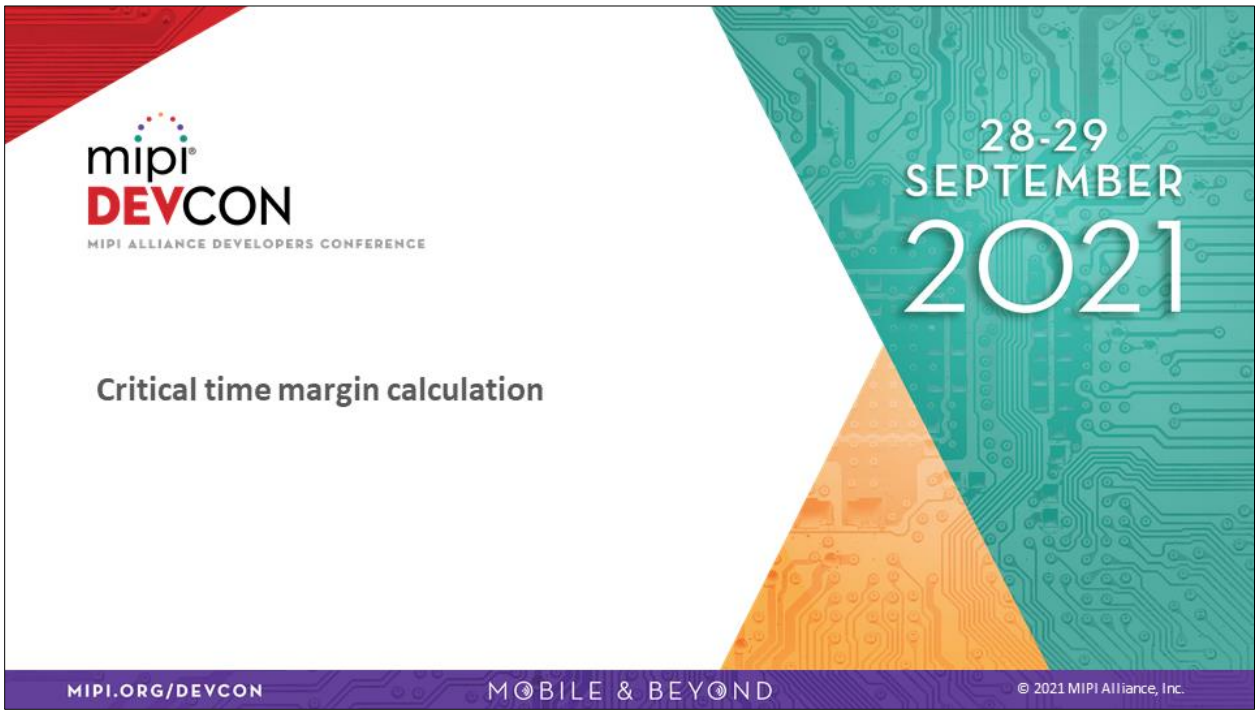
Nestor Hernandez: 運用這三項數據，要掌握上升時間不容易，其次是下降時間及 V_{OL}，右側開漏的電阻數值較低，左側的推挽需要較高電阻數值，若對照這兩項要求，會找到交集點，通常交焦點大約 303 歐姆，上升時間大約 75 奈秒，再加上推挽，就能找到符合最高運作效率所需的 V_{OL} 安全數值。


mipi
DEVCON
MIPI ALLIANCE DEVELOPERS CONFERENCE
28-29
SEPTEMBER
2021
Buffer R_{ON} design implications
MIPI.ORG/DEVCON
MOBILE & BEYOND
© 2021 MIPI Alliance, Inc.

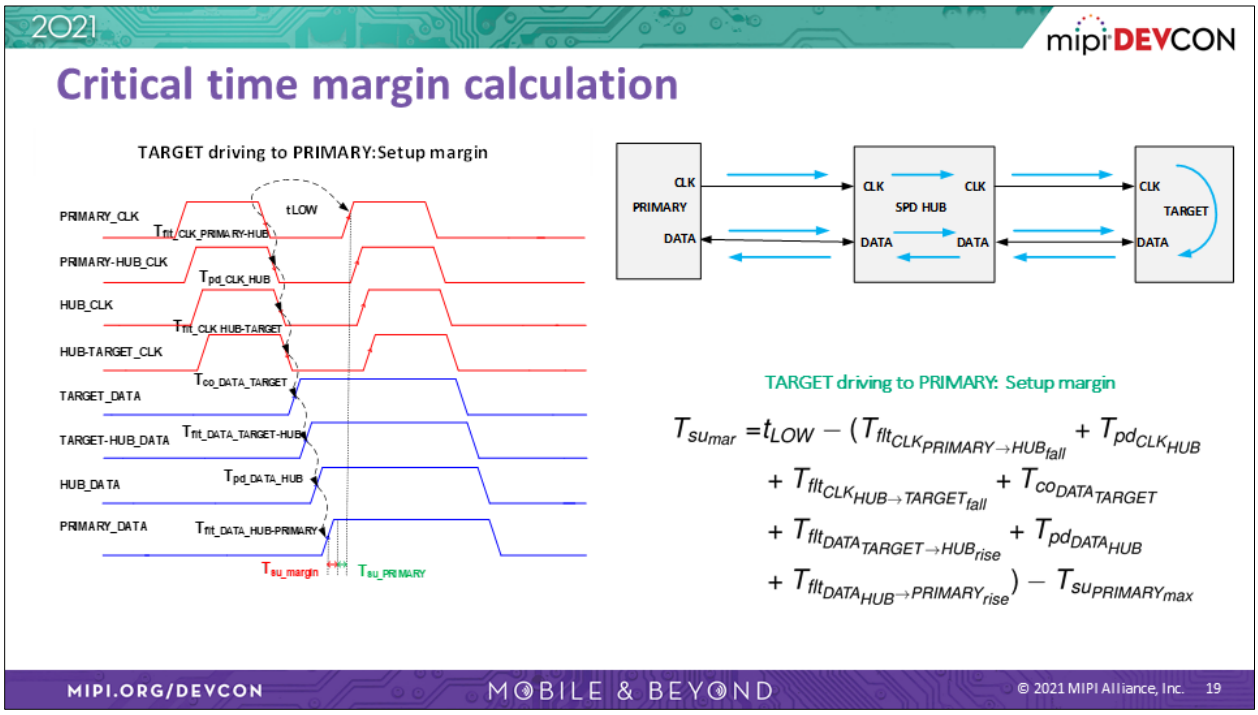


Nestor Hernandez: 驅動器必須驅動，匯流排必須為強驅動器，依據走線長度，會需要具備 40 歐姆的強驅動器，當線路達到 50 英吋左右，仍然能夠驅動電容，但會面臨一些問題，若最長線路無法使用 V_{OL}，就會出現約 146MV；若 V_{OL} 線路負載為 300，則會有 150MV。各位要注意到使用的電流為 3.60 mA，因此 Ron 必須設定上限為 40 歐姆，以達到適當的 V_{OL}，讓電流最高不會超過 3 mA，這將是 I3C 設計緩衝的主要特質。

接下來可進入實際連線計算。



Azusena Lupercio: 謝謝 Nestor，接下來討論臨界時間餘裕的計算。

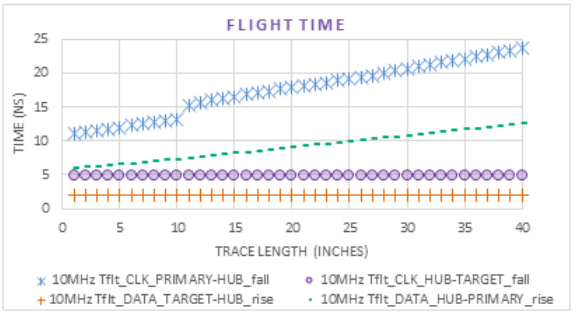


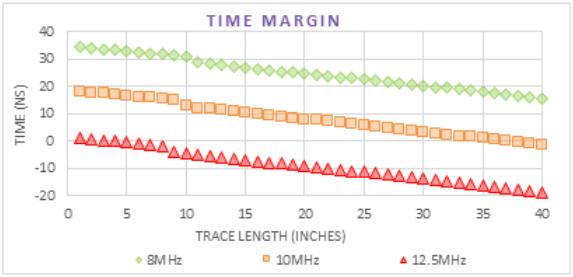
Azusena Lupercio: 為了找出高匯流排電容系統的正確操作頻率，必須檢查目標裝置驅動至首要裝置時的設定時間餘裕，可透過下列算式計算出來，算出訊號周期內較低的部分後，再減去整條路徑內所有延遲的參數，包括從首要裝置到樞紐裝置的時脈飛行時間傳播延遲、樞紐的時脈傳播延遲、從樞紐到目標的時脈傳播延遲、目標的資料 TCO、目標到樞紐的資料飛行時間、樞紐的資料傳播延遲、樞紐到首要的資料飛行時間，以及首要裝置的設定時間。

Critical time margin calculation

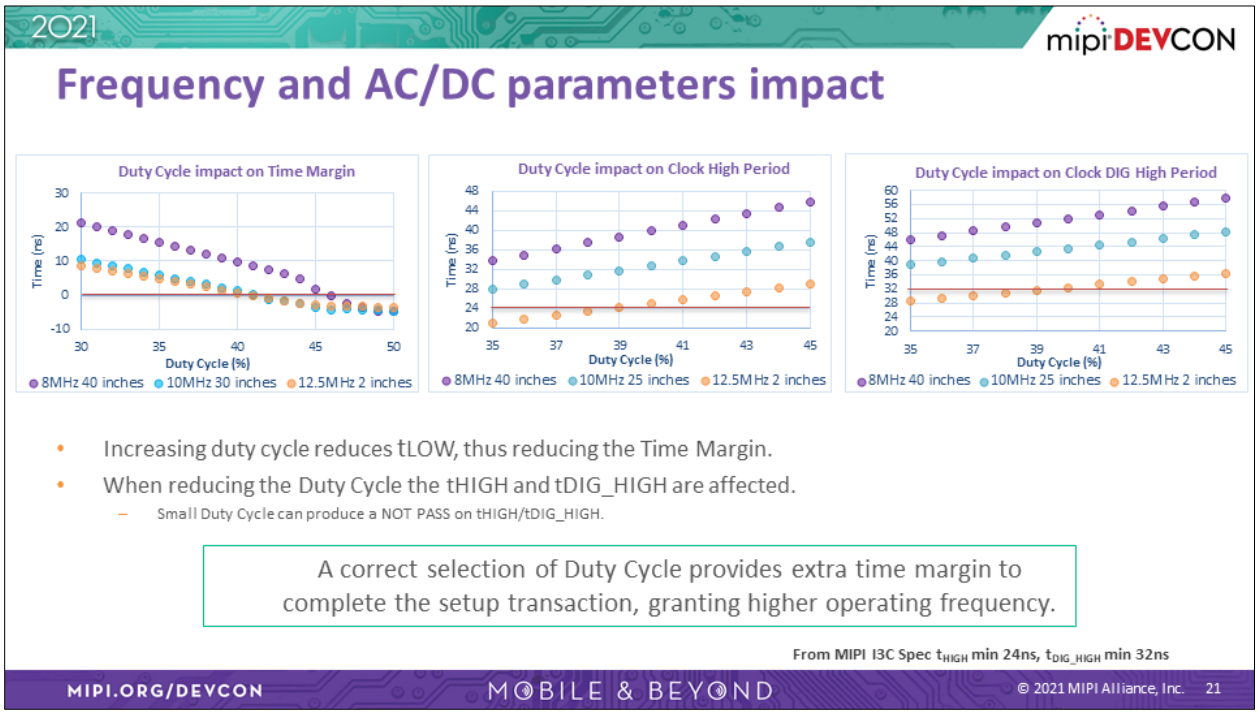
Frequency	10	MHz
Duty cycle	35	%
tLOW	65	ns
tflt_CLK_PRIMARY-HUB_fall	19.1	ns
tpd_CLK_HUB	6	ns
tflt_CLK_HUB-TARGET_fall	4.7	ns
tco_DATA_TARGET	12	ns
tflt_DATA_TARGET-HUB_rise	1.972	ns
tpd_DATA_HUB	6	ns
tflt_DATA_HUB-PRIMARY_rise	9.8	ns
tsu_PRIMARY_max	3	ns
Setup Margin	2.253	ns

- 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.



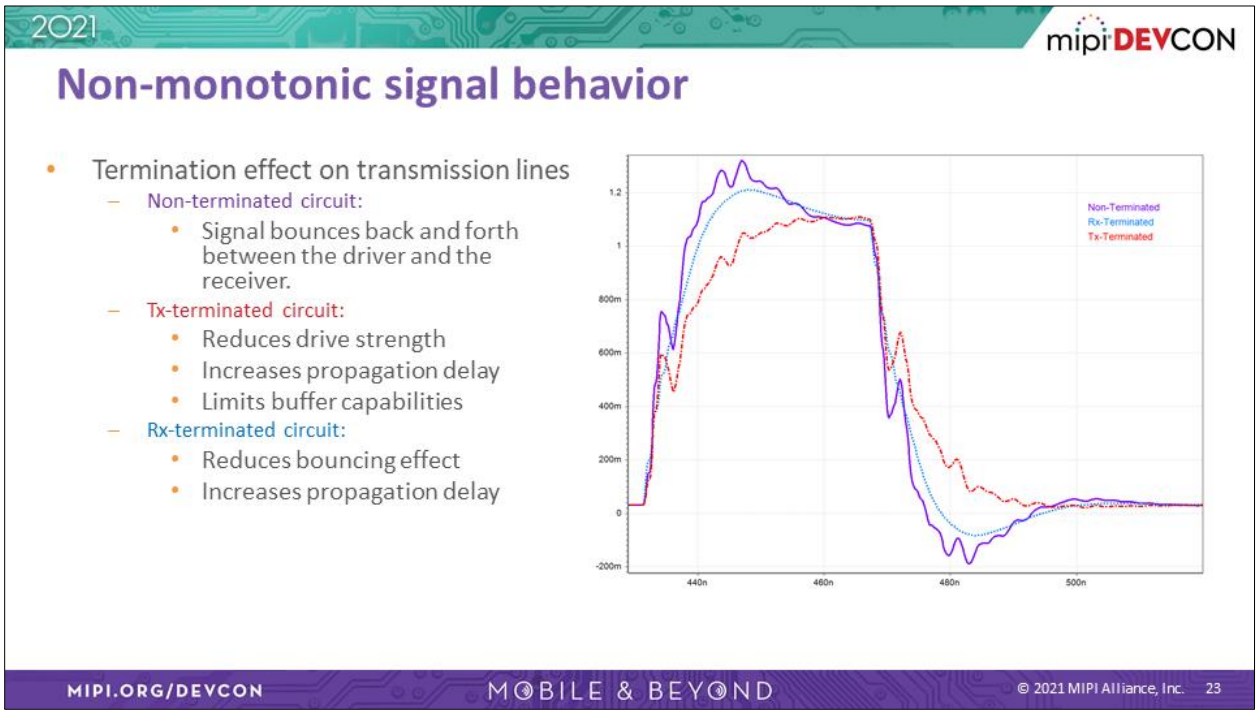


Azusena Lupercio: 各位心中有了這些算式之後，就能釐清該檢查哪些重要參數，以找出各項操作頻率，例如取決於走線長度的飛行時間、各項裝置的內部傳播延遲、tLOW 等，只要走線長度愈長，飛行時間就會增加，因此走線長度若很長，飛行時間就會很長；傳播延遲愈高，時間餘裕就會縮短，但增加 tLOW 可提供額外時間餘裕，下方圖表內共有三種操作頻率：12.5MHz、10 MHz、8 MHz，紅色是 12.5 MHz，即使走線長度很短，時間餘裕也很短，代表訊號傳播的時間不足；10 MHz 的解決方案空間約有 35 吋；8 MHz 的高電容匯流排解決方案空間足夠，可大於 40 吋。



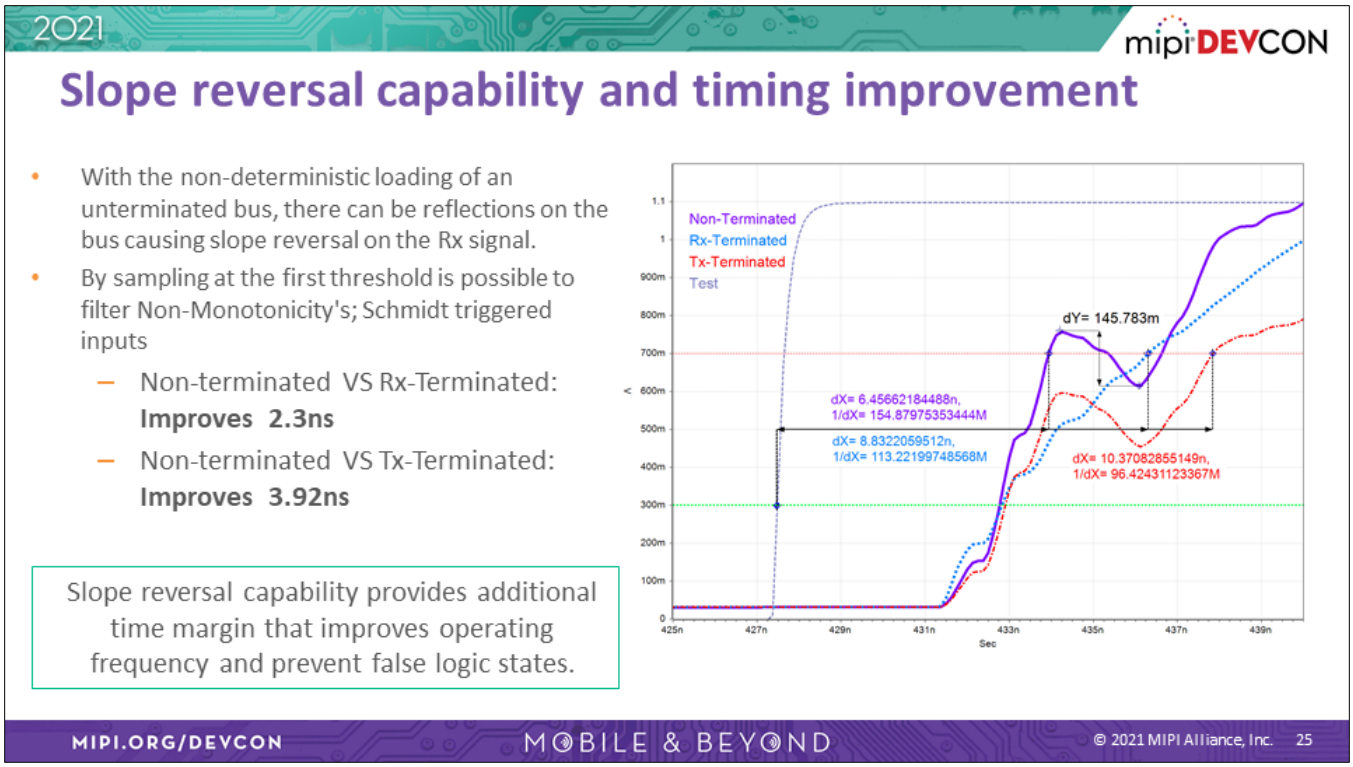
Azusena Lupercio: 了解 tLOW 為重要參數之後，可增加另外三項新參數，例如工作周期、tHIGH、tDIG_HIGH，工作周期變長，會縮短 tLOW，也會減少時間餘裕；若縮短工作周期，tHIGH 和 tDIG_HIGH 會受影響，換言之，工作周期短無法符合 tHIGH 和 tDIG_HIGH 規格，tHIGH 至少需 24 奈秒，tDIG_HIGH 至少需 32 奈秒，若選擇適當工作周期，不僅能提供額外時間餘裕，可完成設定作業，也能確保操作頻率更高。

Azusena Lupercio: 非單調性訊號行為



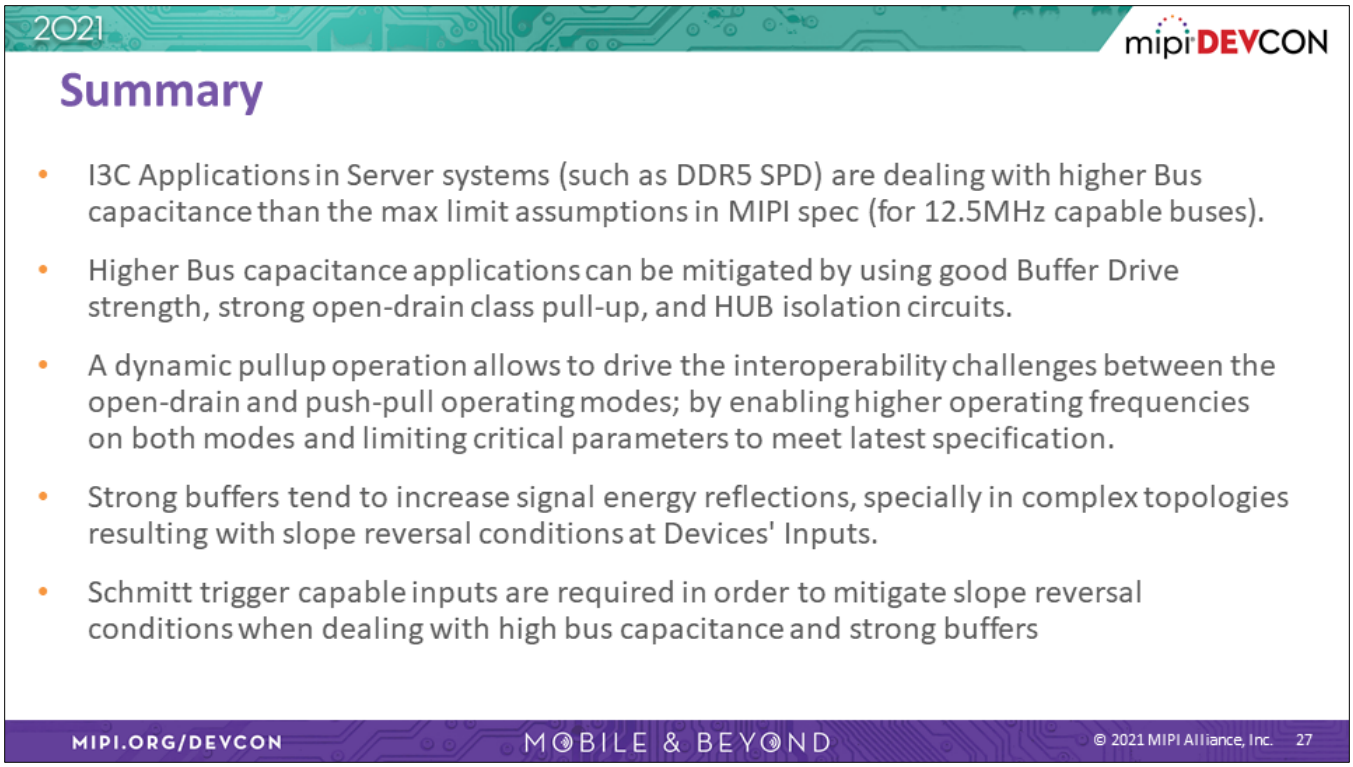
Azusena Lupercio: 非單調性訊號可透過決定性效果排除，若無終接線路，訊號會不斷在驅動器和接收器之間來回，由於訊號存在非單調性，傳輸器的終接線路降低驅動器強度、增加傳輸延、限制緩衝能力；這項訊號無單調性存在，傳輸器的終接線路減少反彈效果，但時脈傳輸延遲很高，藍線代表無單調性存在。

Azusena Lupercio: 斜率反轉能力與時序改善



Azusena Lupercio: 由於無終接線路匯流排的非決定性負載，可能反射在匯流排上，造成接收器訊號斜率反轉，若在第一門檻取樣，有可能過濾非單調性或施密特觸發器輸入，若比較終接與無終接接收器，飛行時間可改善約 2.3 奈秒；若比較無終接線路和終接傳輸器，飛行時間可改善約 3.9 奈秒，斜率反轉能力提供額外時間餘裕，可改變操作頻率，也避免錯誤邏輯狀態。

Azusena Lupercio: 總結



Azusena Lupercio: 在 DDR5 SPD 等伺服器系統內，I3C 需處理的匯流排電容，大於 I3C MIPI 規格針對 12.5MHz 匯流排的高低限制，應用匯流排電容較高時，緩和途徑包括緩衝驅動器強度、強開漏、上拉電阻、HUB 隔離電路，動態推挽操作可改善開漏與推挽操作模式之間的互用性，提高兩種模式的操作頻率，並限制重要參數，以符合最新規格。強大緩衝會增加訊號能量反射，尤其在複雜組態中，導致裝置輸入出現斜率反轉情況。需要施密特觸發器輸入，以緩和斜率反轉情況，也同時處理高匯流排電容及緩衝。

Azusena Lupercio: 感謝大家聆聽，接下來歡迎各位提問。