Flash Memory Summit

What Exactly are SSD Low Power States and How Can They Be Tested and Measured?

Presenter: Anthony Andrews, Principal Software Engineer
SerialTek, LLC

August 10, 2023

Overview

- Understand the L1 state of the LTSSM
- Understand the L1.1 and L1.2 PM states
- Case Study 1- Measuring Power with Quarch and SerialTek
- Case Study 2- Testing Low Power with SANBlaze and SerialTek

Understanding L1 LTSSM

- From Section 4.2.6.8 of the PCIe 6.0 Spec it says:

  "L1 is intended as a power savings state.

  The L1 state allows an additional power savings over L0s at the cost of additional resume latency.

  The entry to L1 occurs after being directed by the Data Link Layer and receiving an EIOS."

- L1 only has 2 substates, L1.Entry and L1.Idle, it does **not** have L1.0/L1.1/L1.2 substates

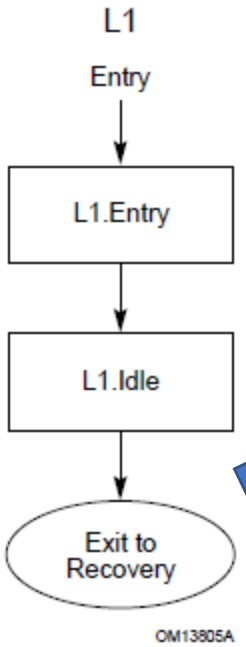


Figure 4-54 L1 Substate Machine

L1 Substates

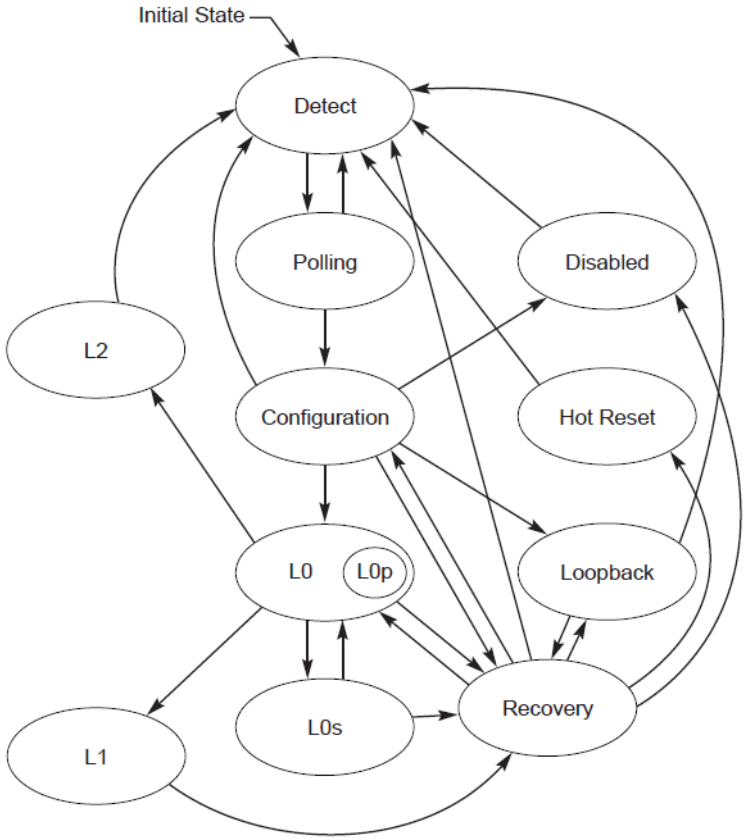
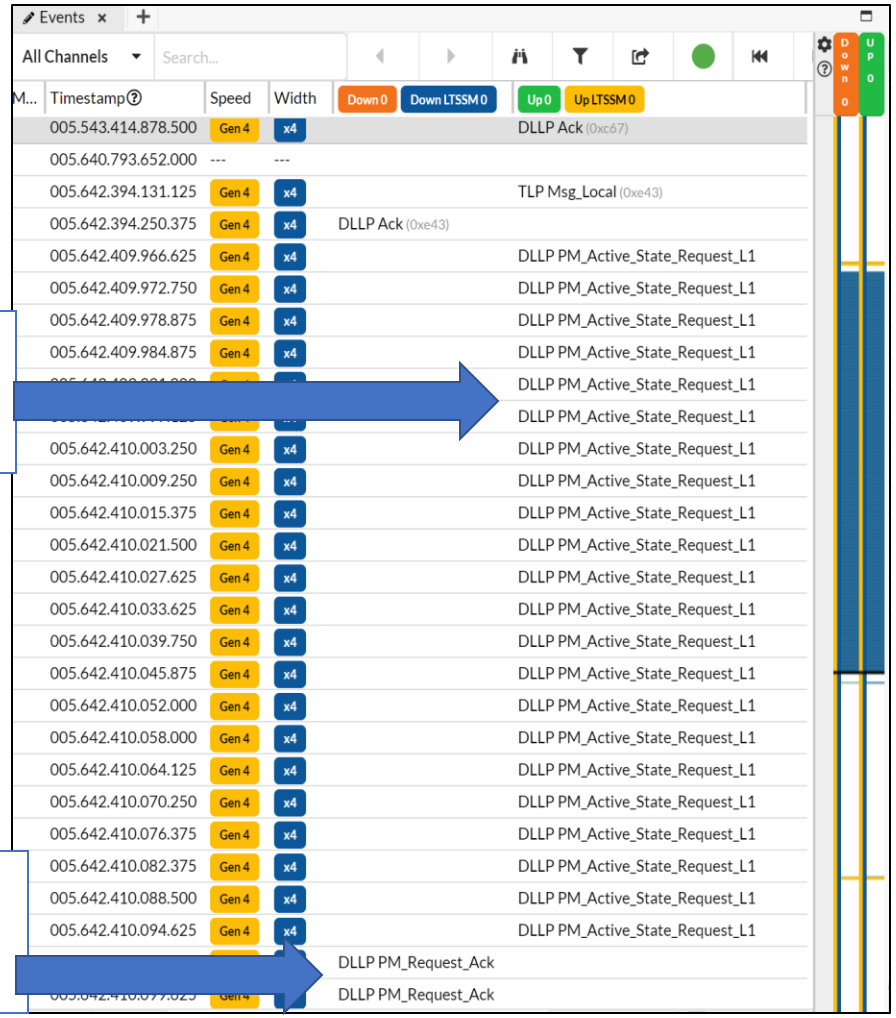


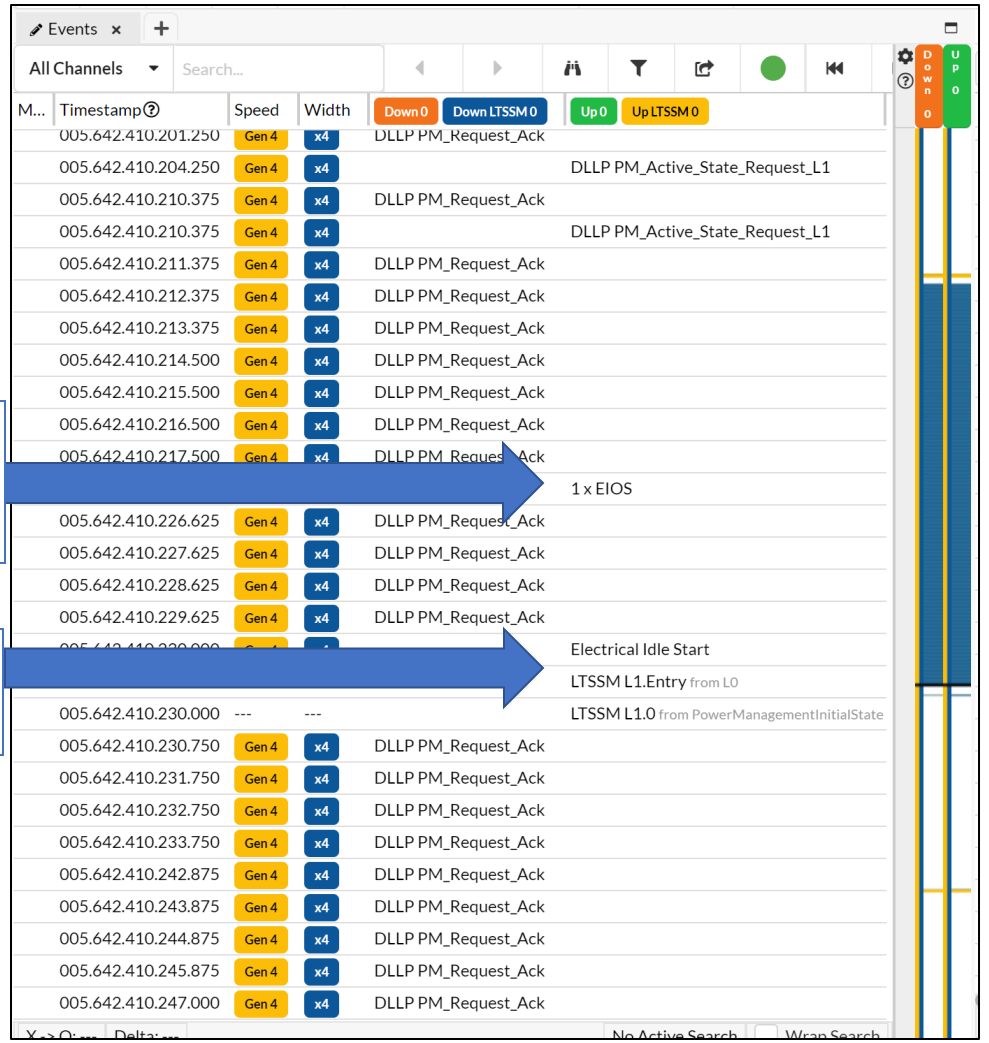
Figure 4-48 Main State Diagram for Link Training and Status State Machine

Understanding L1 LTSSM- Entry



1. Device sends PM DLLP Requesting L1

2. Host responds with PM DLLP Acknowledging the request



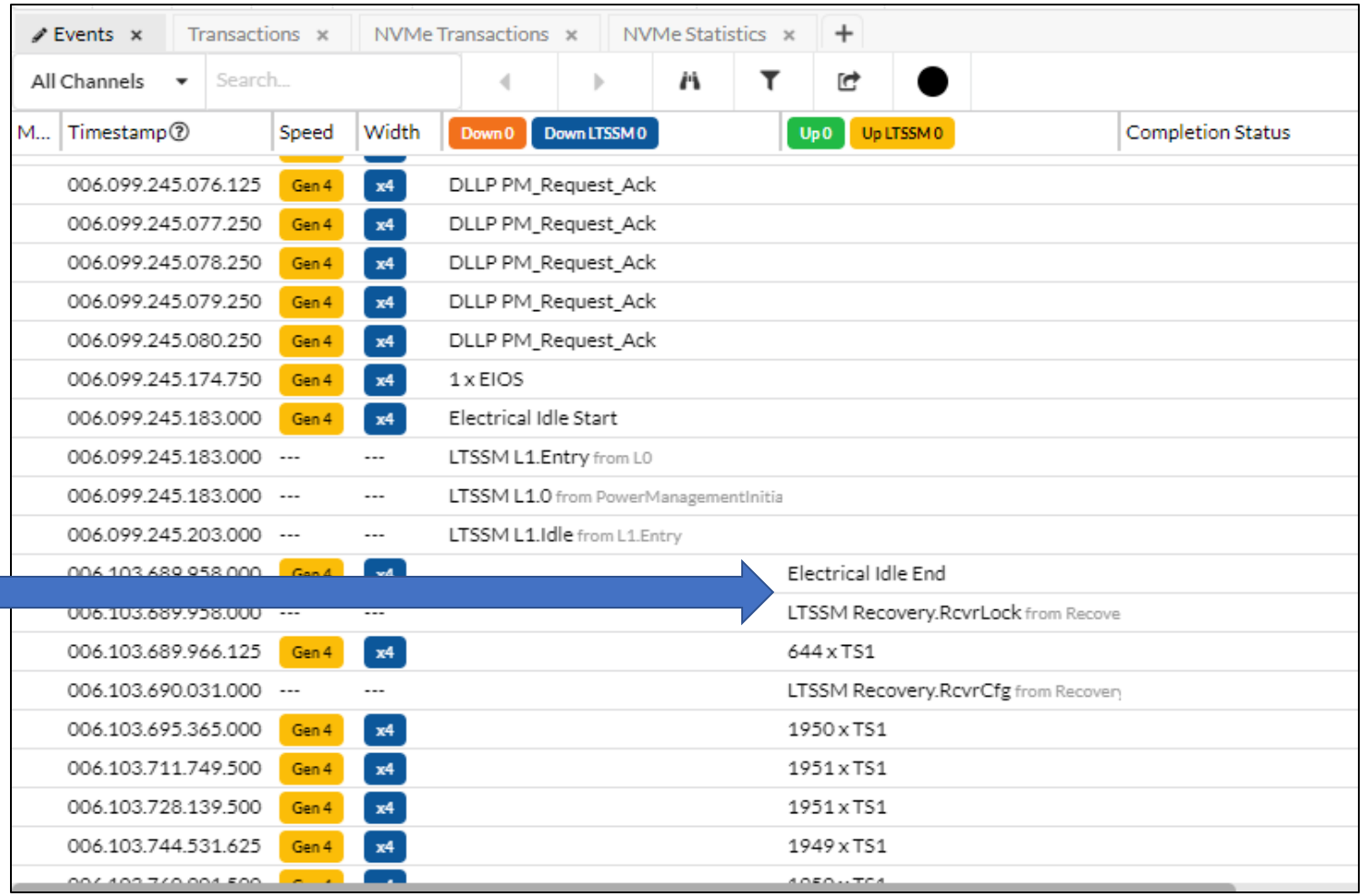
3. EIOS Ordered Set to transition into L1 sent by the device

4. Device is now sending Electrical Idle as part of L1

Understanding L1 LTSSM- Exit

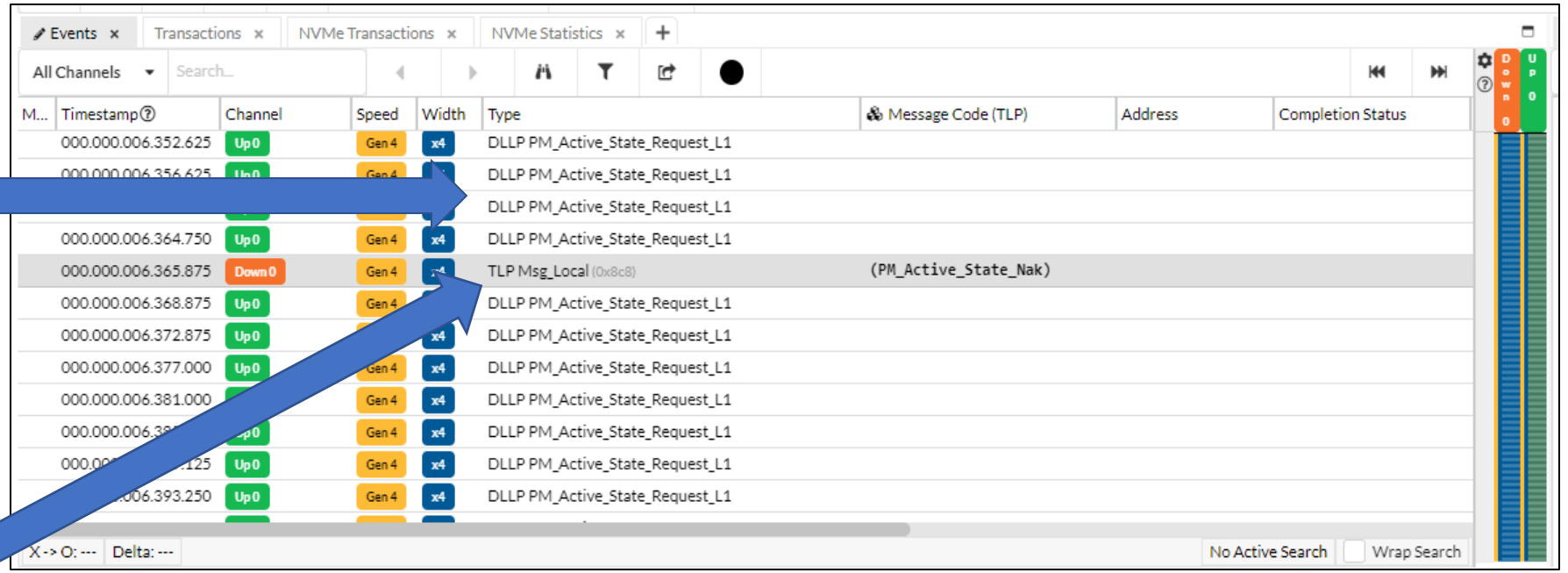
Flash Memory Summit

Device leaves Electrical Idle and starts Sending TS1 Ordered Sets.



SerialTek
an ellisys company

Understanding L1 LTSSM-Nak

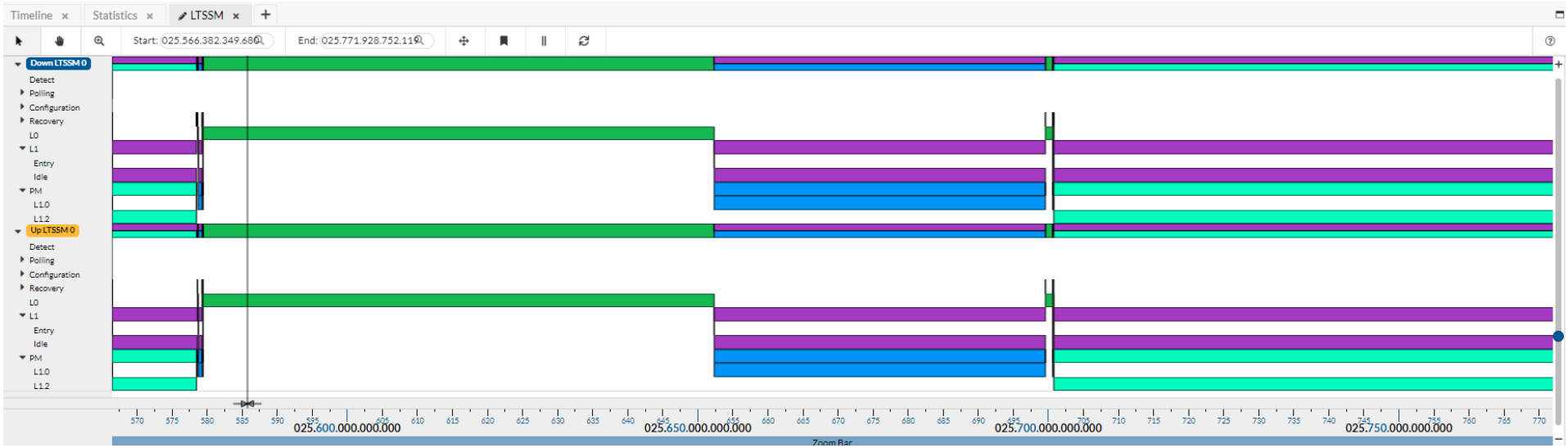


PM DLLP Requesting L1

Message TLP Nak to reject the request

Understanding L1.1/L1.2 PM States

- Section 5.1 explains about Power Management (PM) states.
  - D states are associated with a particular Function
  - L states are associated with a particular Link
- L states contain the following states
  - L0 is the operational state
  - L0p is a reduced power sub-state of L0
  - L0s, L1, L1.0, L1.1 and L1.2 are various low power states



Understanding L1.1/L1.2 PM States

- ***L1.0*** substate
  - The L1.0 substate corresponds to the conventional L1 Link state. This substate is entered whenever the Link enters L1. The L1 PM Substate mechanism defines transitions from this substate to and from the L1.1 and L1.2 substates.
  - The Upstream and Downstream Ports must be enabled to detect Electrical Idle exit as required in § Section 4.2.7.7.2 .
- ***L1.1*** substate
  - Link common mode voltages are maintained.
  - Uses a bidirectional open-drain clock request (CLKREQ#) signal for entry to and exit from this state.
  - The Upstream and Downstream Ports are not required to be enabled to detect Electrical Idle exit.
- ***L1.2*** substate
  - Link common mode voltages are not required to be maintained.
  - Uses a bidirectional open-drain clock request (CLKREQ#) signal for entry to and exit from this state.
  - The Upstream and Downstream Ports are not required to be enabled to detect Electrical Idle exit.

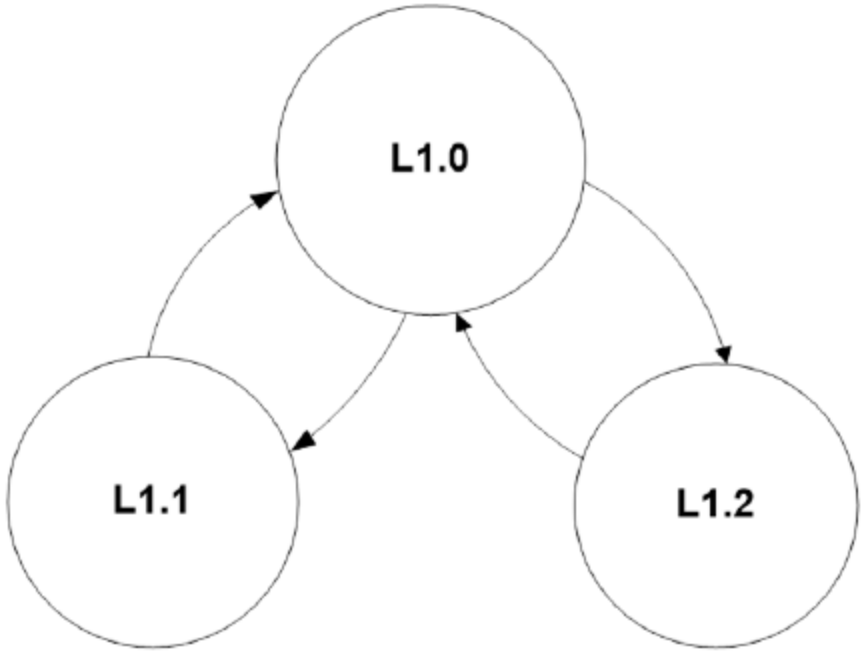
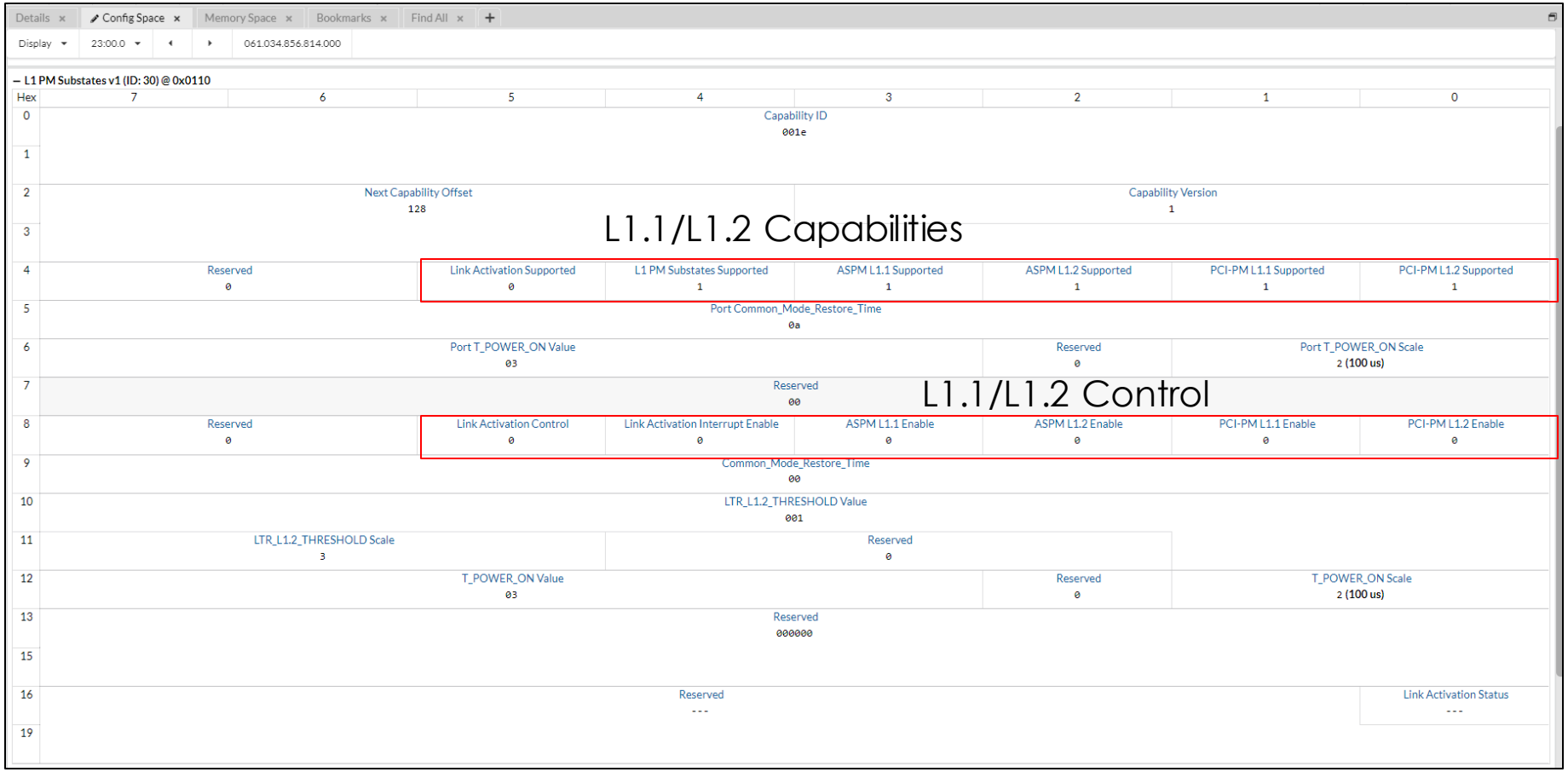


Figure 5-9 State Diagram for L1 PM Substates

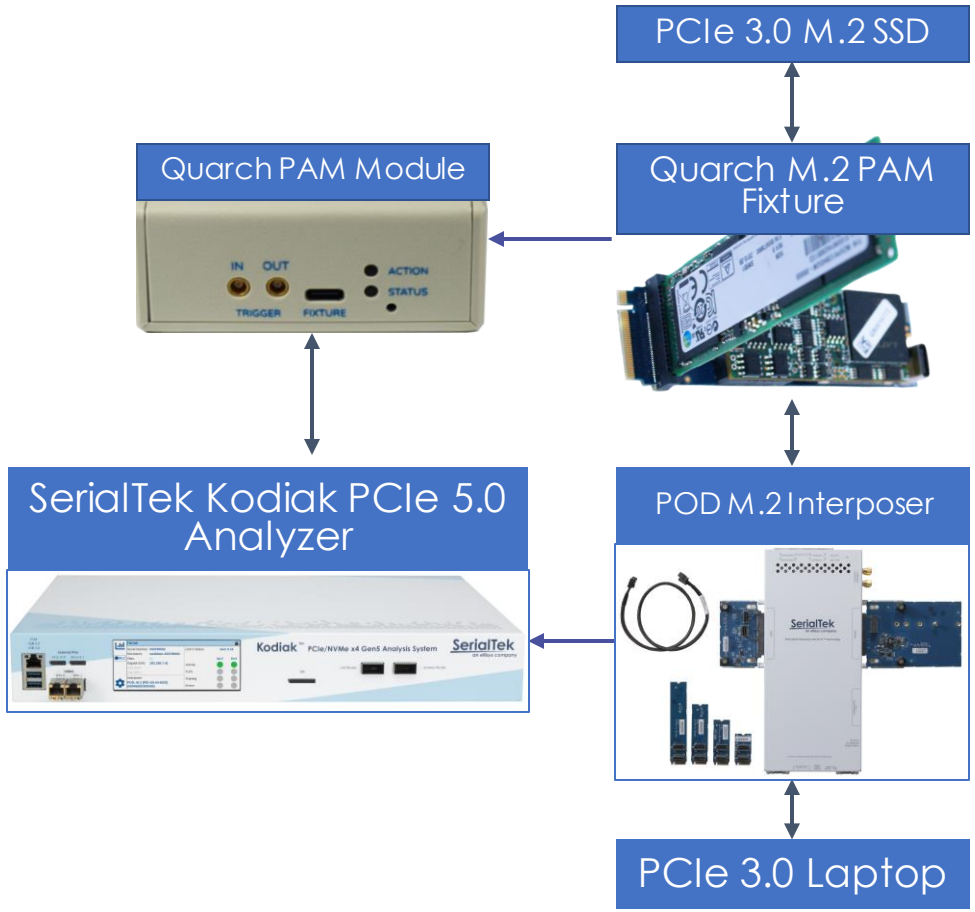
Understanding L1 PM Register Example



L1 PM Substate Extended Capability register is used when enabling the PM substate.

Case Study 1

- Goal
  - **<u>Understand the Power savings in a laptop environment</u>**
- Tools Used
  - PCIe 3.0 Laptop
  - PCIe 3.0 M.2 SSD
  - Quarch Power Analysis Module with PCIe4.0 M.2 PAM Fixture
  - SerialTek Kodiak PCIe Analyzer
- Software
  - lspci & setpci
  - Quarch Power Studio
  - SerialTek BusXpert

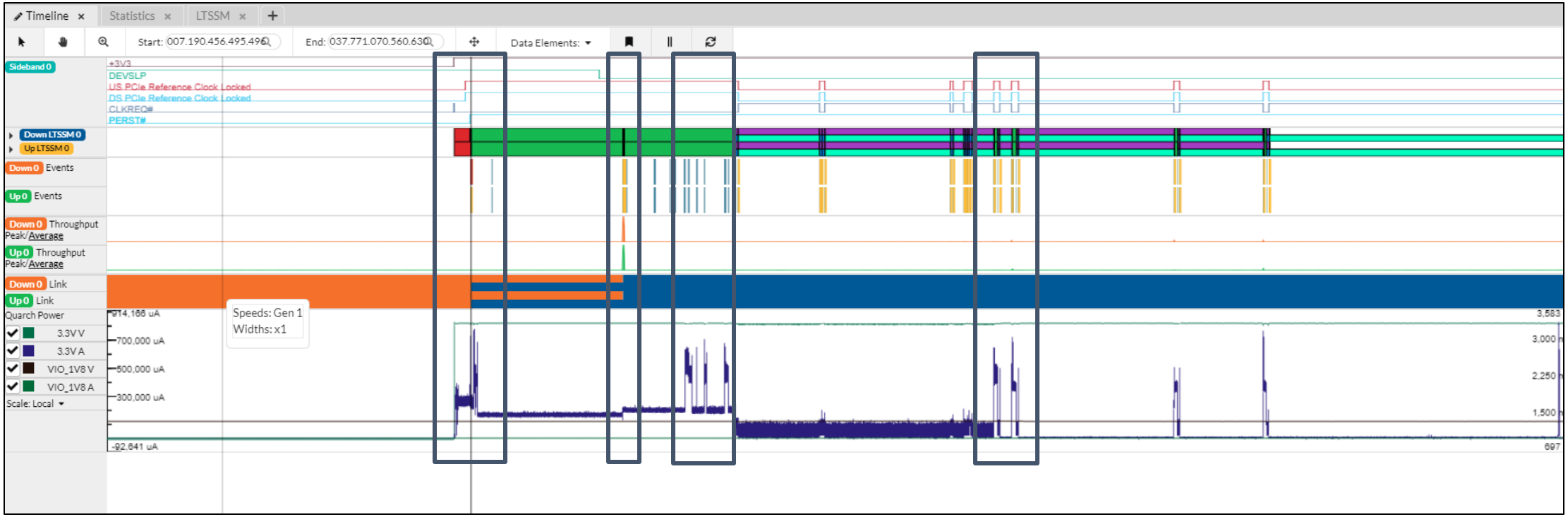


Case Study 1

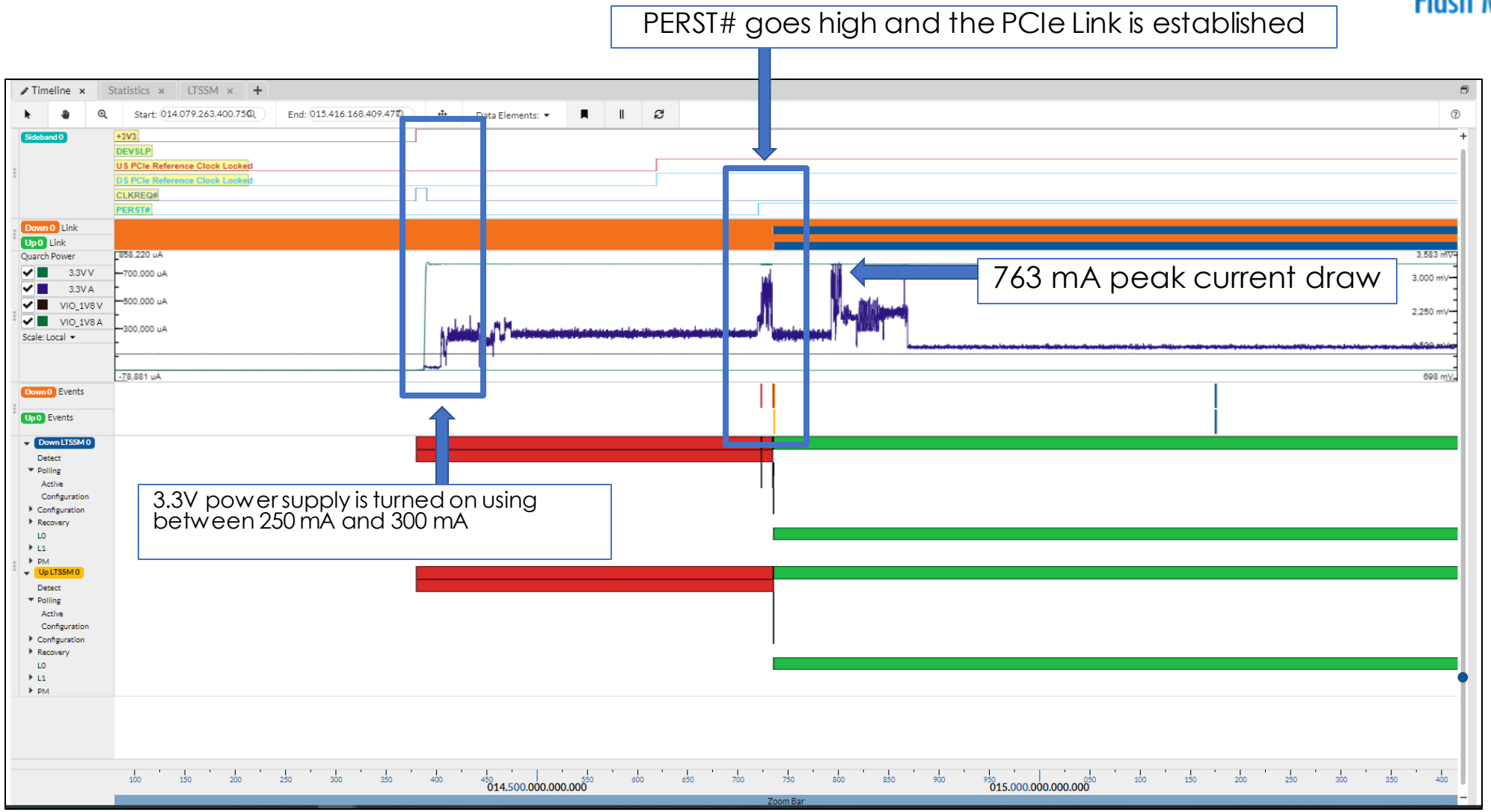
Capture a bootup of the laptop until the OS fully loads.

Let's look closer at the following:

- When 3.3V is first applied
- Speed change
- At some of the current spikes
- L1 Transitions

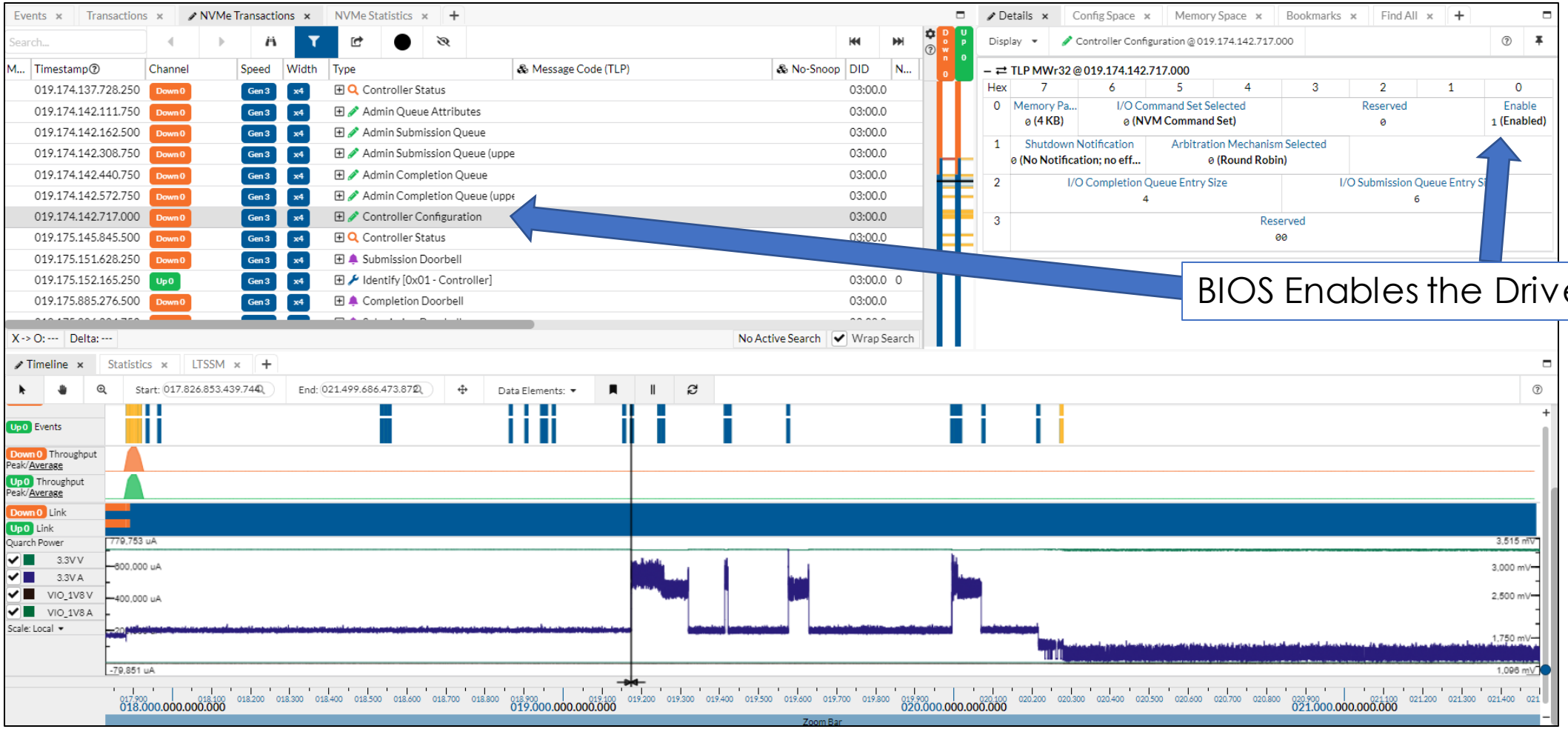


Case Study 1- Power On

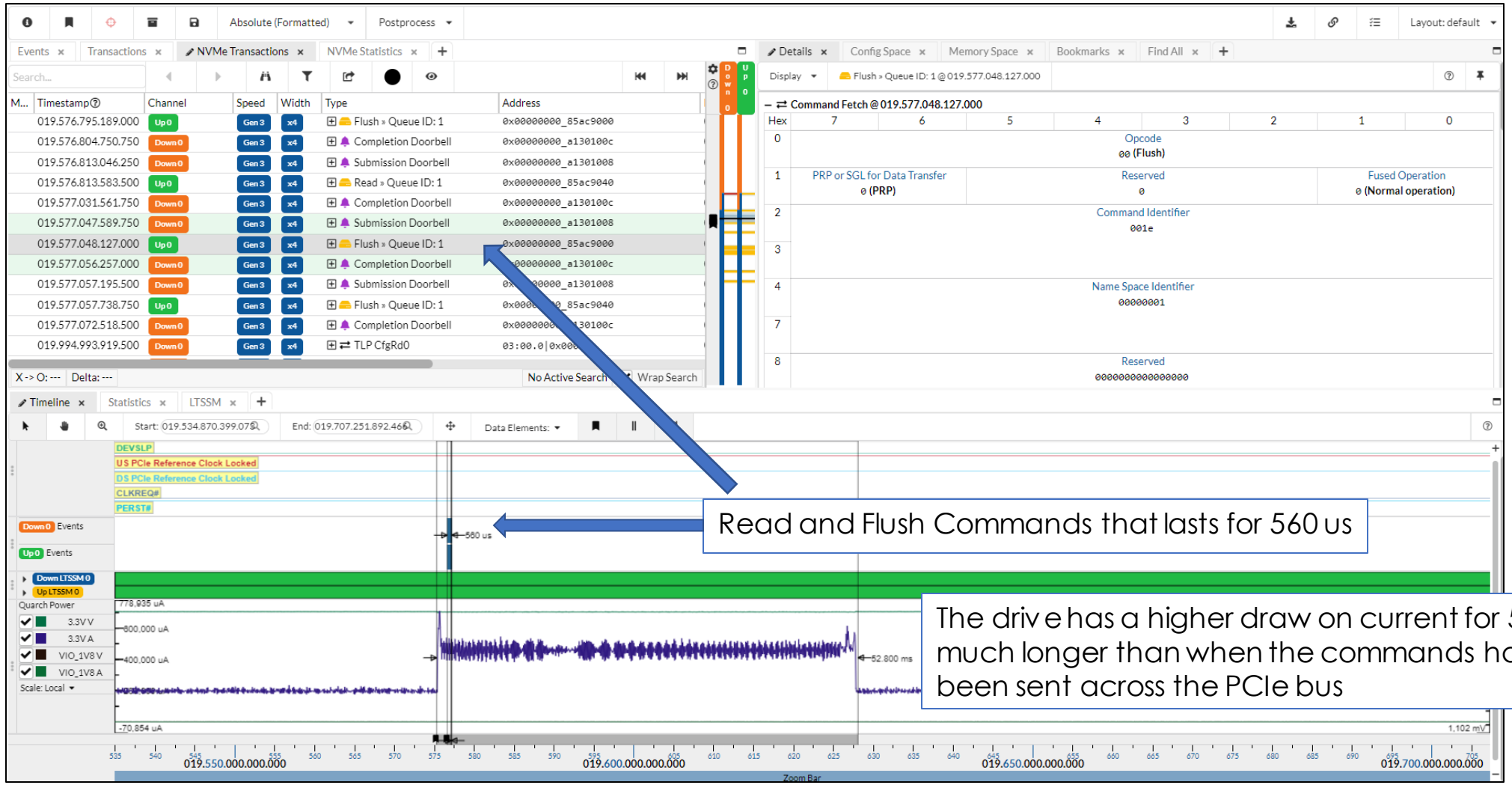


Case Study 1- Speed Change to Gen3

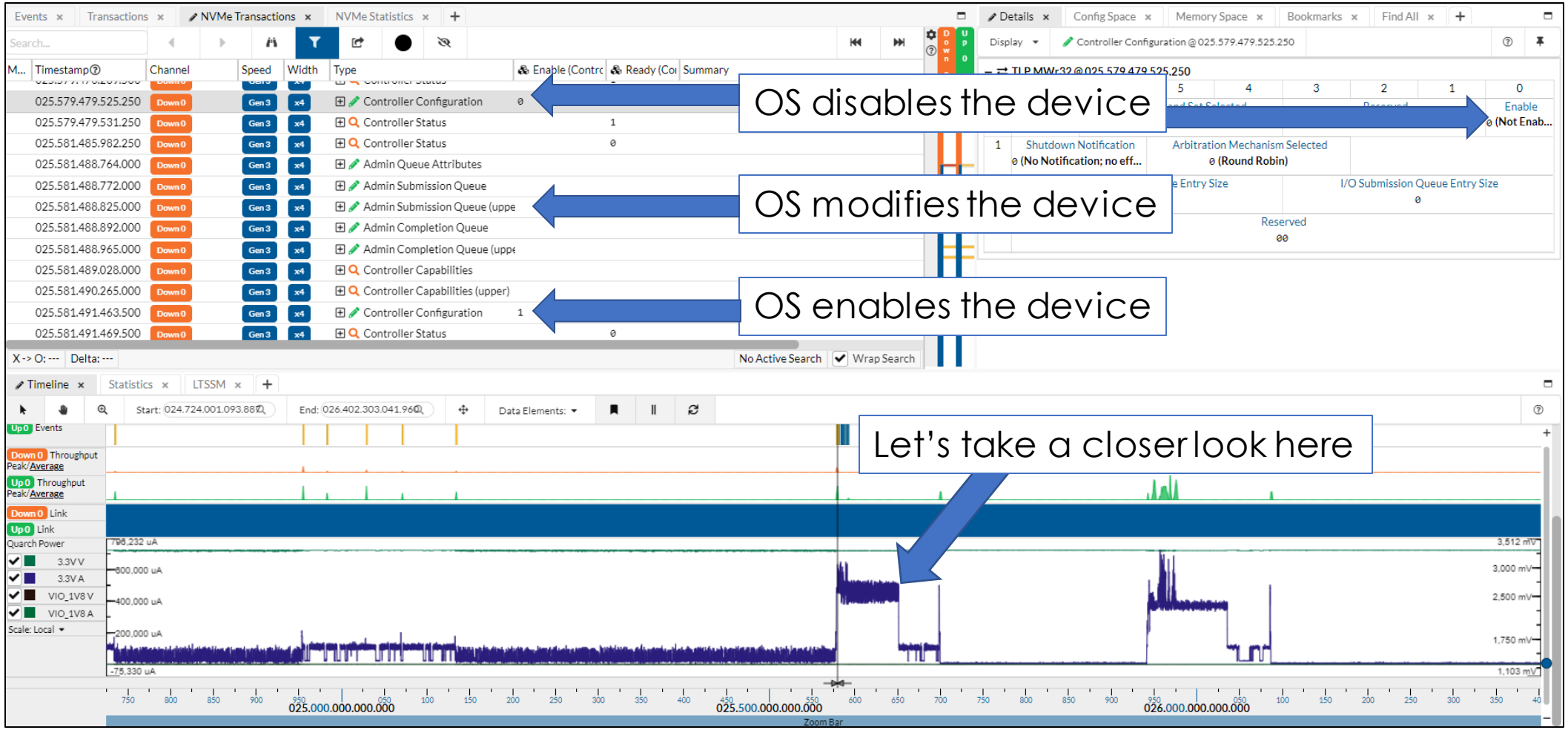
Case Study 1- Analyzing Spike



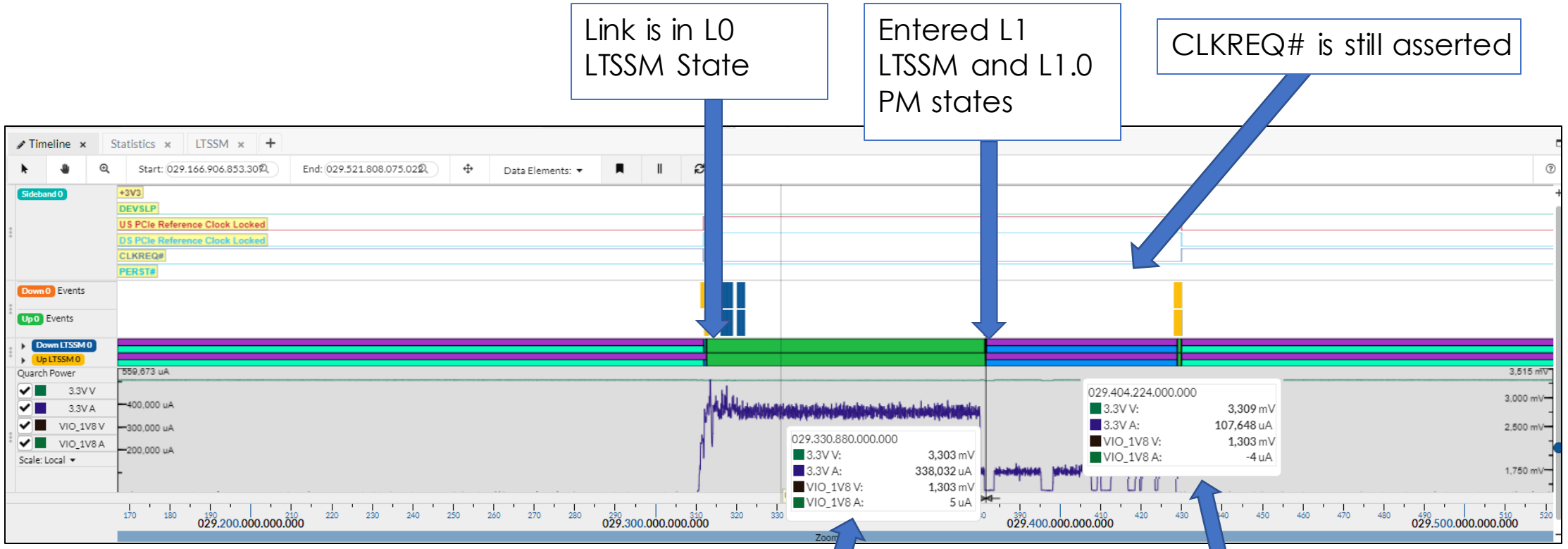
Case Study 1


Read and Flush Commands that lasts for 560 us
The drive has a higher draw on current for 58 ms, much longer than when the commands have been sent across the PCIe bus

Case Study 1- OS Configures the Device



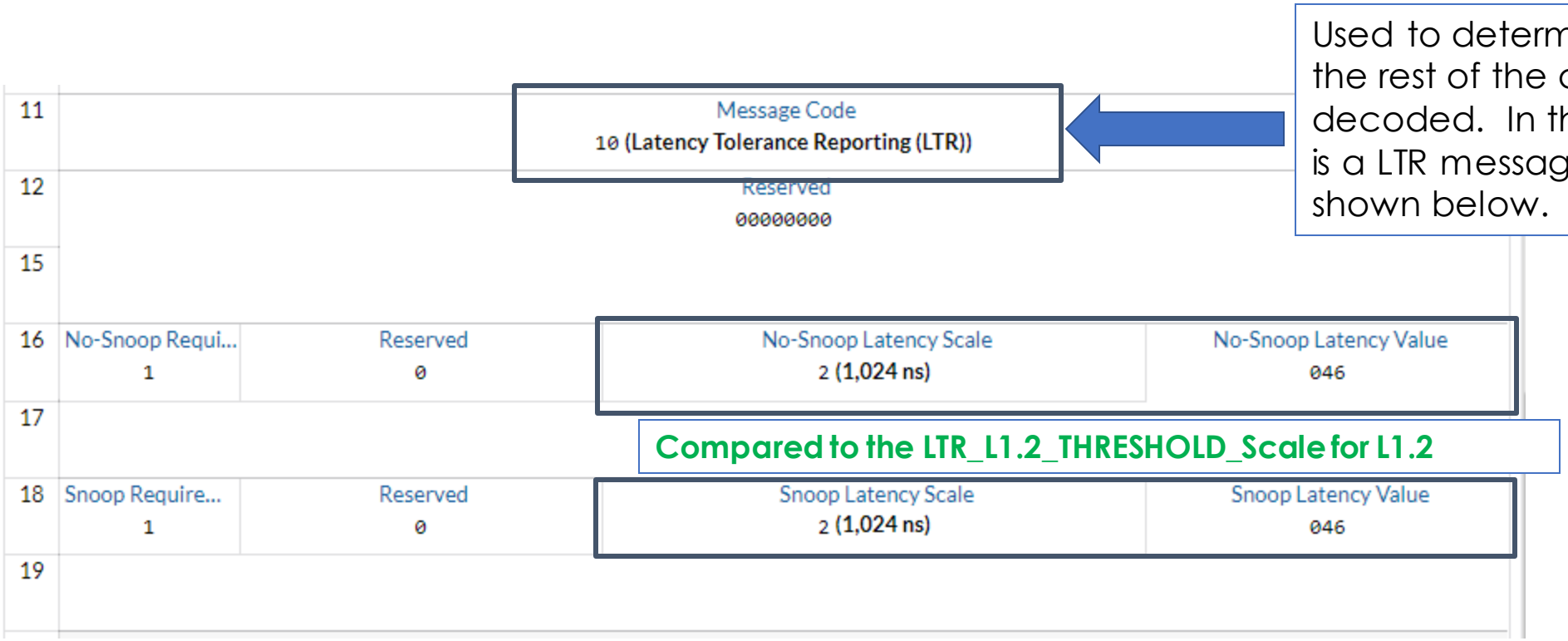
Case Study 1- Closer Look



Case Study 1- Understanding LTR

- Why is CLKREQ# still asserted?
- From section "5.5.1 Entry Conditions for L1 PM Substates and L1.0 Requirements"
  - "When in ASPM L1.0 and the ASPM L1.2 Enable bit is Set, the L1.2 substate must be entered when CLKREQ# is deasserted and all of the following conditions are true:"
    - "The reported snooped LTR value last sent or received by this Port is greater than or equal to the value set by the LTR_L1.2_THRESHOLD Value and Scale fields, or there is no snoop service latency requirement."
    - "The reported non-snooped LTR last sent or received by this Port value is greater than or equal to the value set by the LTR_L1.2_THRESHOLD Value and Scale fields, or there is no non-snoop service latency requirement."

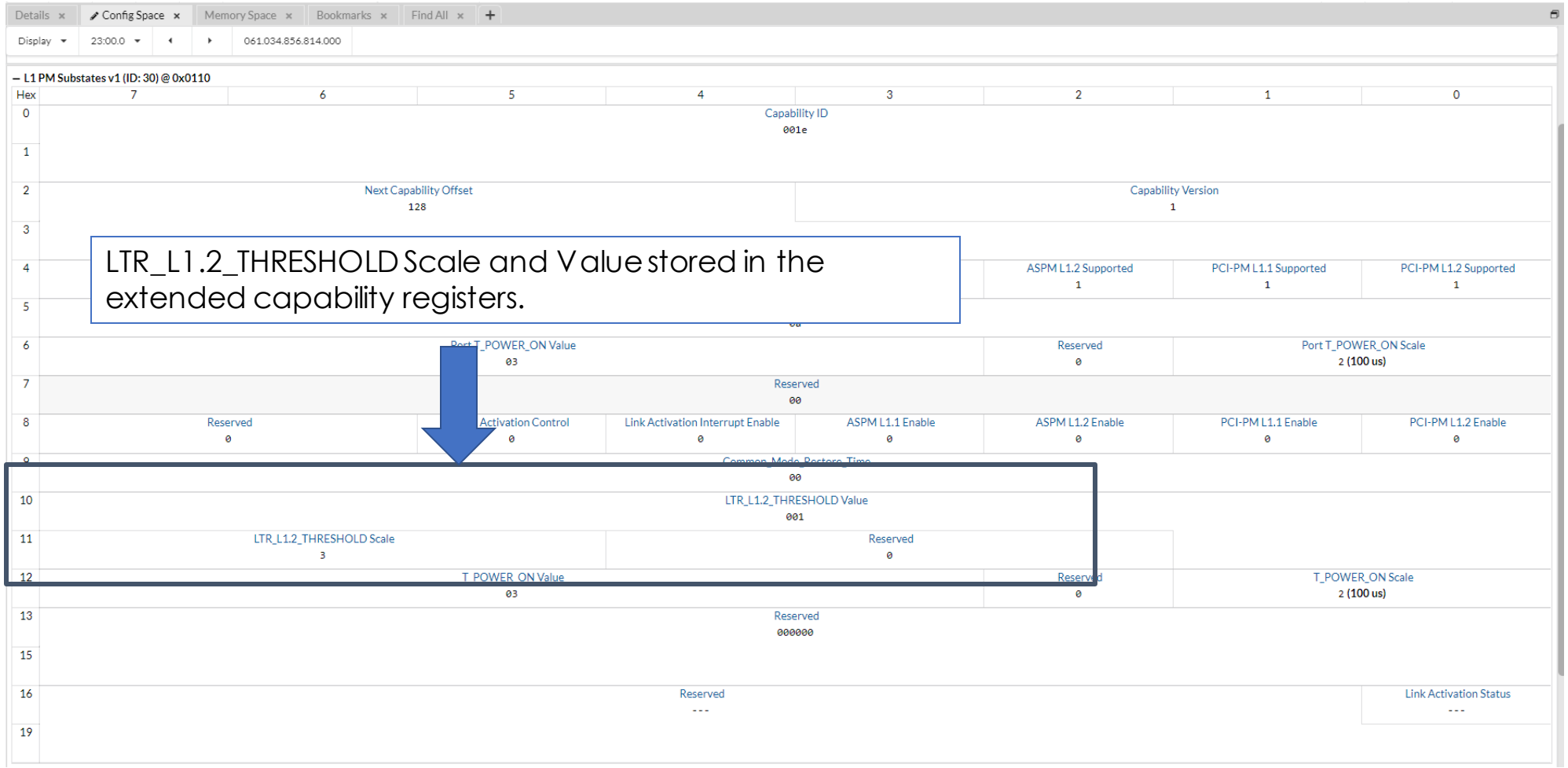
Case Study 1- LTR Message TLPs Decoded



Used to determine how the rest of the data is decoded. In this case it is a LTR message as shown below.

Compared to the LTR_L1.2_THRESHOLD_Scale for L1.2

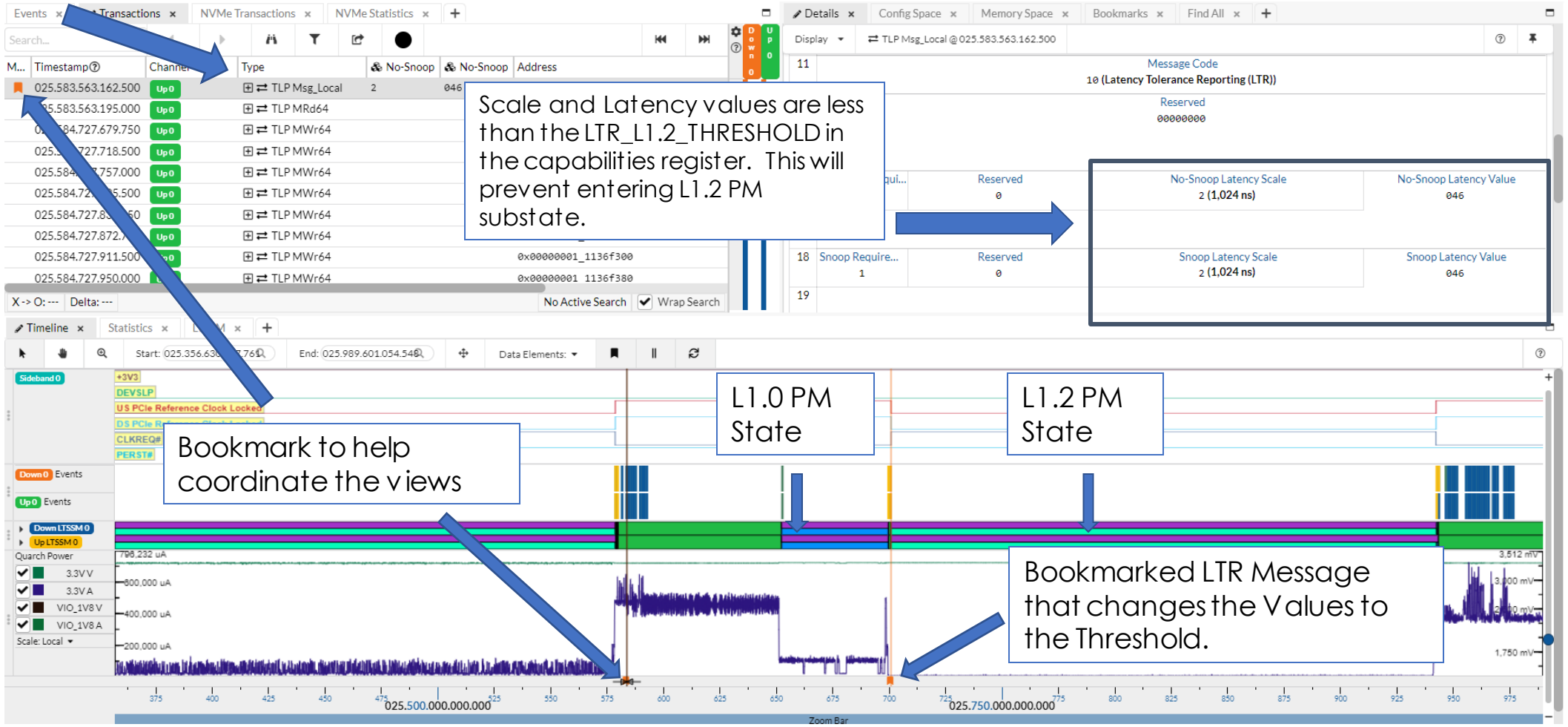
Case Study 1- L1 PM Substates Extended Capabilities



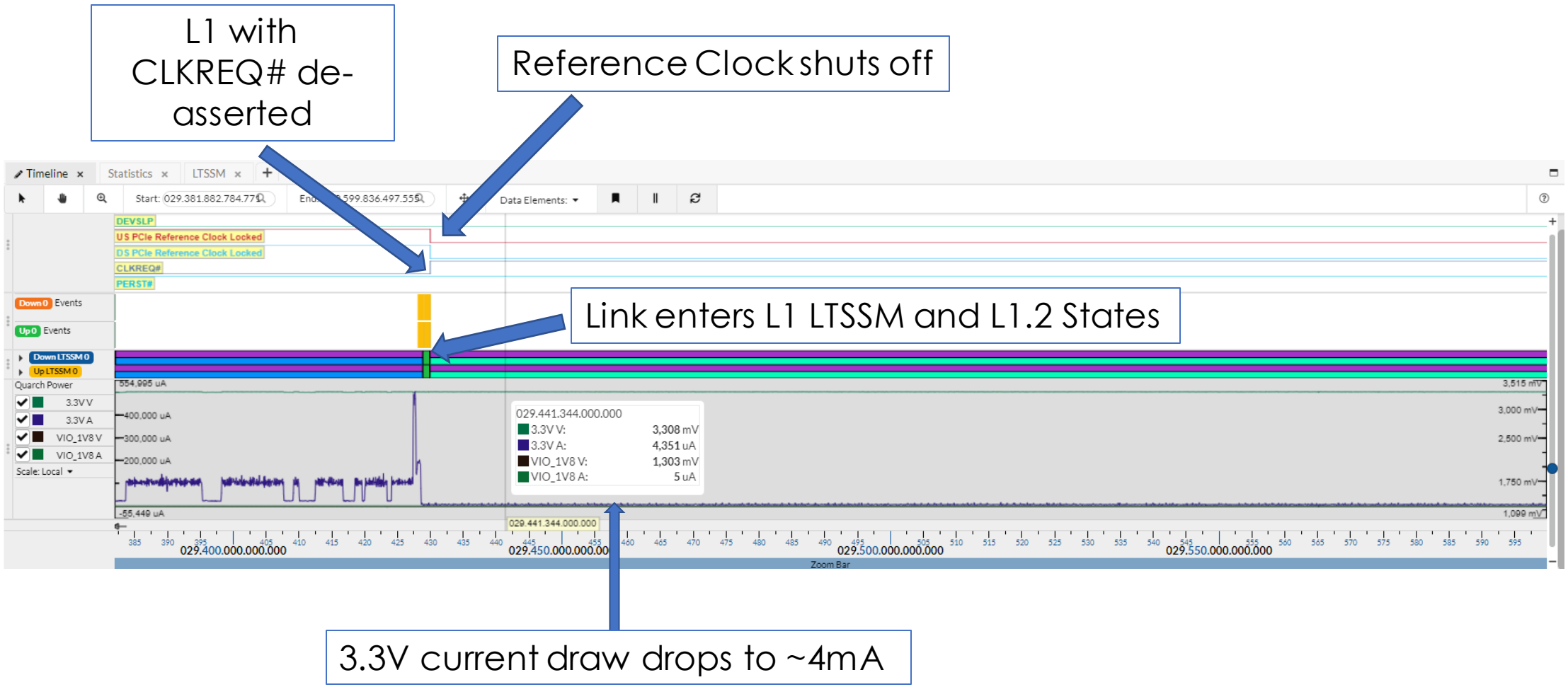
LTR_L1.2_THRESHOLD Scale and Value stored in the extended capability registers.

Case Study 1- LTR and L1.2

LTR Message TLP



Case Study 1- Closer Look at L1.2



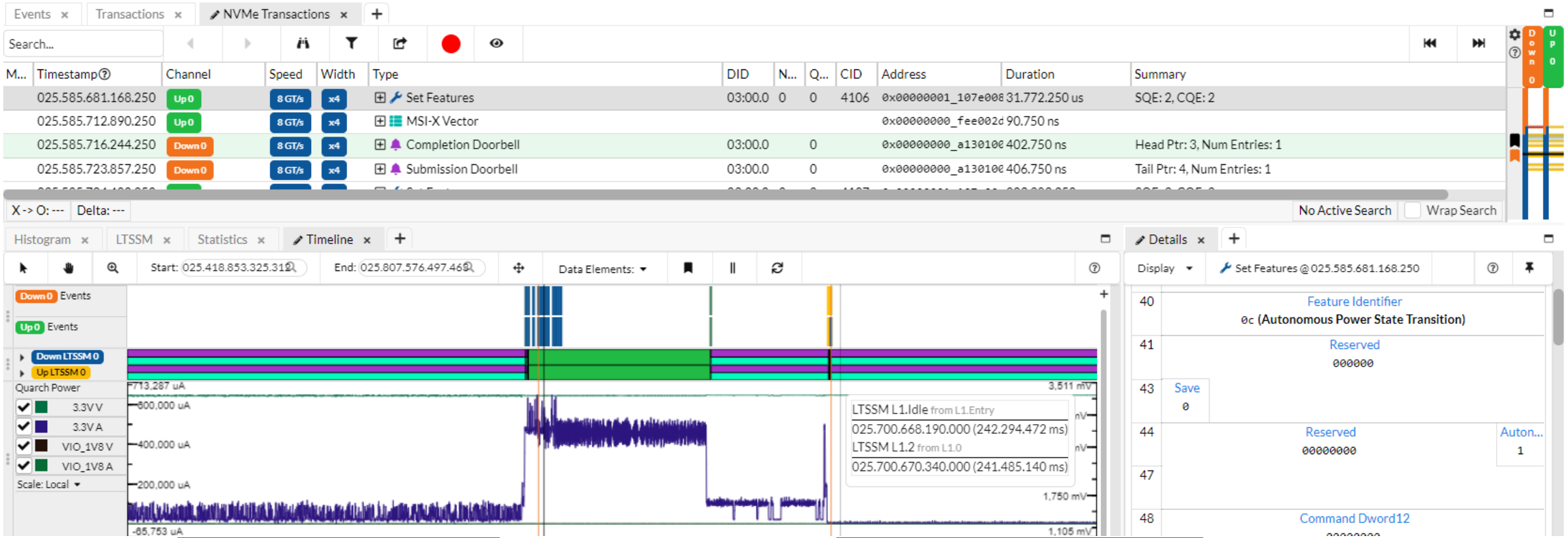
Case Study 1- Power Observations

LTSSM and PM State	Average 3.3V Current draw
L0 LTSSM	380 mA
L1 LTSSM and L1.0 PM State	107 mA
L1 LTSSM and L1.2 PM State	4 mA

- Observation 1: By turning on ASPM L1, the drive can reduce its current draw by 71.8%.
- Observation 2: going into L1.2 PM substate the drive can reduce its current draw by almost 98.9% versus when operating at the L0 state

Case Study 1- Power Observations

- Observation: The host used the Set Features command to tell the device to enable Autonomous Power State Transition. Once the device returned to L1.2 the power draw was lower and more consistent.



L1.2 with inconsistent power draw

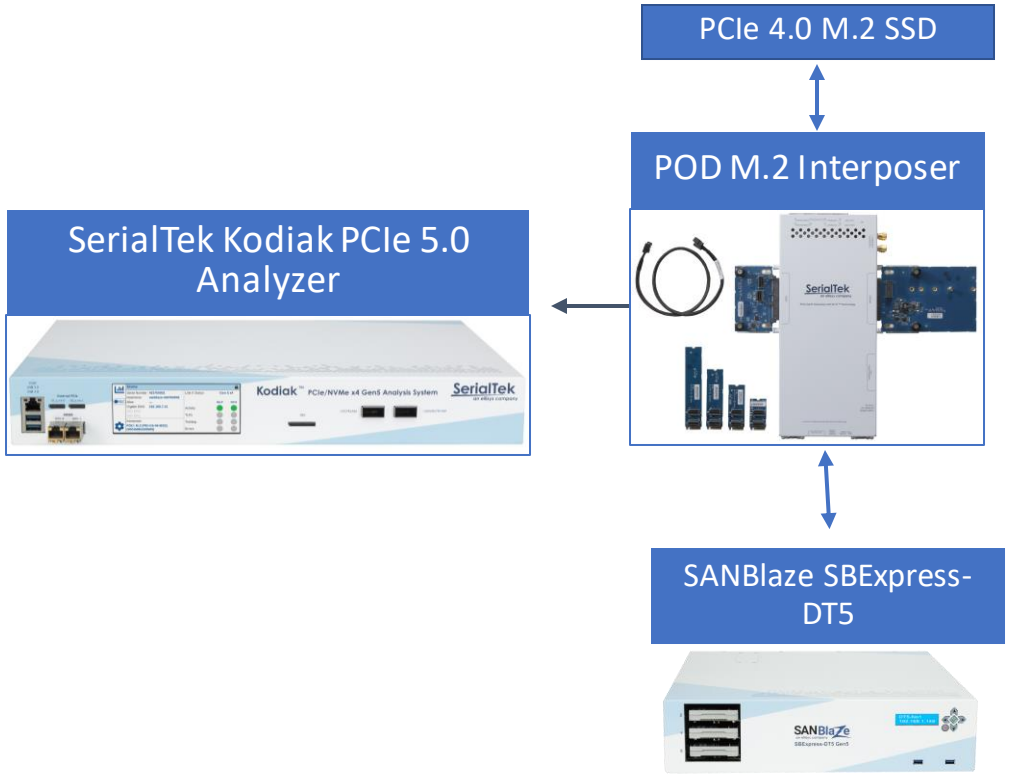
L1.2 with lower consistent power draw

Case Study 1- Other Observations

- When changing speeds that the link operates, the device uses more current.
- By enabling features, this also creates temporary demands on the amount of power the drive consumes.
- Different types of drive operations (i.e. flush command) can also impact the power the drive consumes even when the link activity of the PCIe bus is inactive (sending idle data, flow control DLLPs and skip OS)
- Latency Tolerance Reporting can impact if the device can enter L1.2 PM state which can reduce the power savings of the device.

Case Study 2

- Goal
  - **Test a design to verify that they properly enter lower power states**
- Tools Used
  - SerialTek Kodiak PCIe 5.0 Analyzer + POD interposer configured for M.2
  - SANBlaze SBExpress-DT5
  - PCIe 4.0 M.2 SSD
- Software
  - SerialTek BusXpert
  - SANBlaze



Case Study 2- Test Scripts

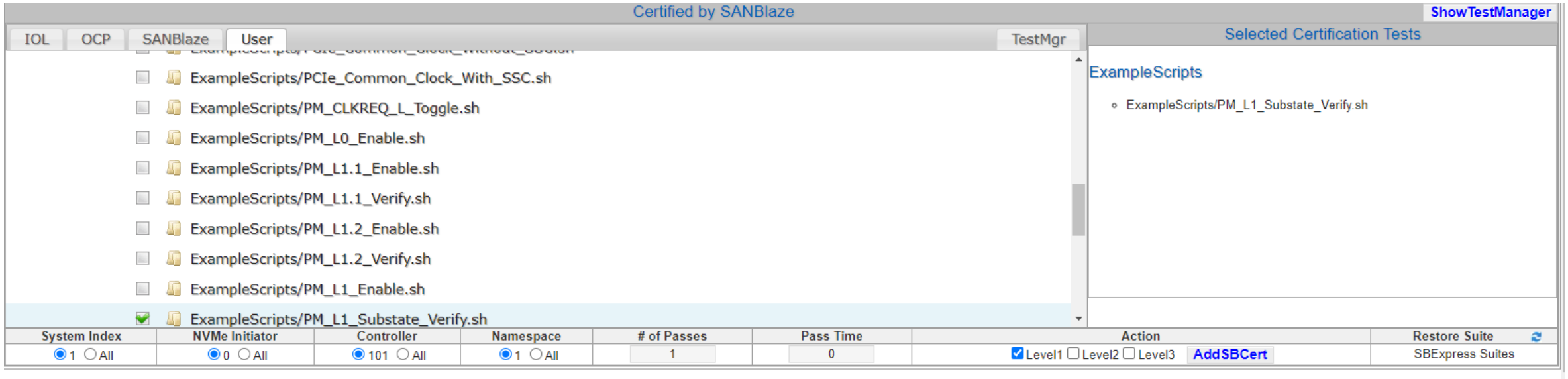
The following scripts are available in the SANBlaze V10.5 software package.

- PM_L1_Substate_Verify.sh
- PM_L0_Enable.sh
- PM_L1.1_Enable.sh
- PM_L1.2_Enable.sh
- PM_L1_Enable.sh

All scripts are variations of the first script PM_L1_Substate_Verify.sh, which walks the DUT from L0.Active to L1.Idle to L1.1 and L1.2. The variations are as follows:

- PM_PM_L0_Enable.sh - Restores the device to L0.Active state
- L1.1_Enable.sh - Leaves the device in L1.1 state
- PM_L1.2_Enable.sh - Leaves the device in L1.2 state
- PM_L1_Enable.sh - Leaves the device in L1 (L1.0) state

Case Study 2- Test Setup

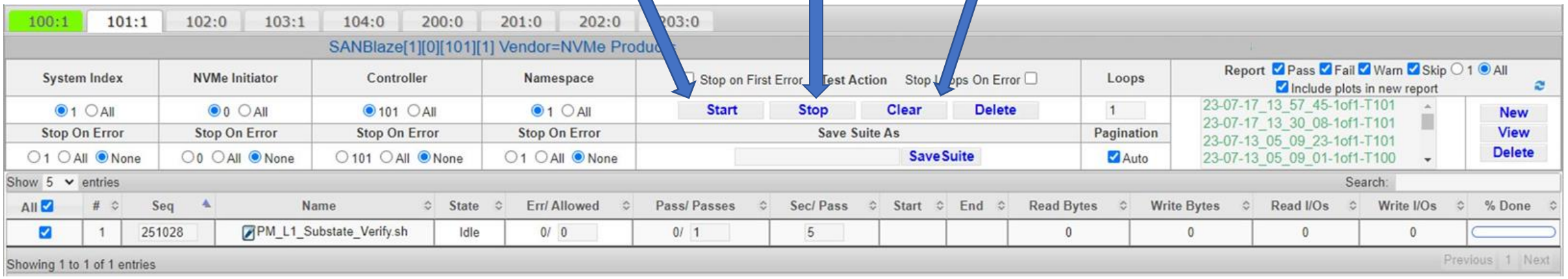


Select the PM_L1_Substate_Verify.sh test.

Add the test to the SBCert

Case Study 2- Test Setup

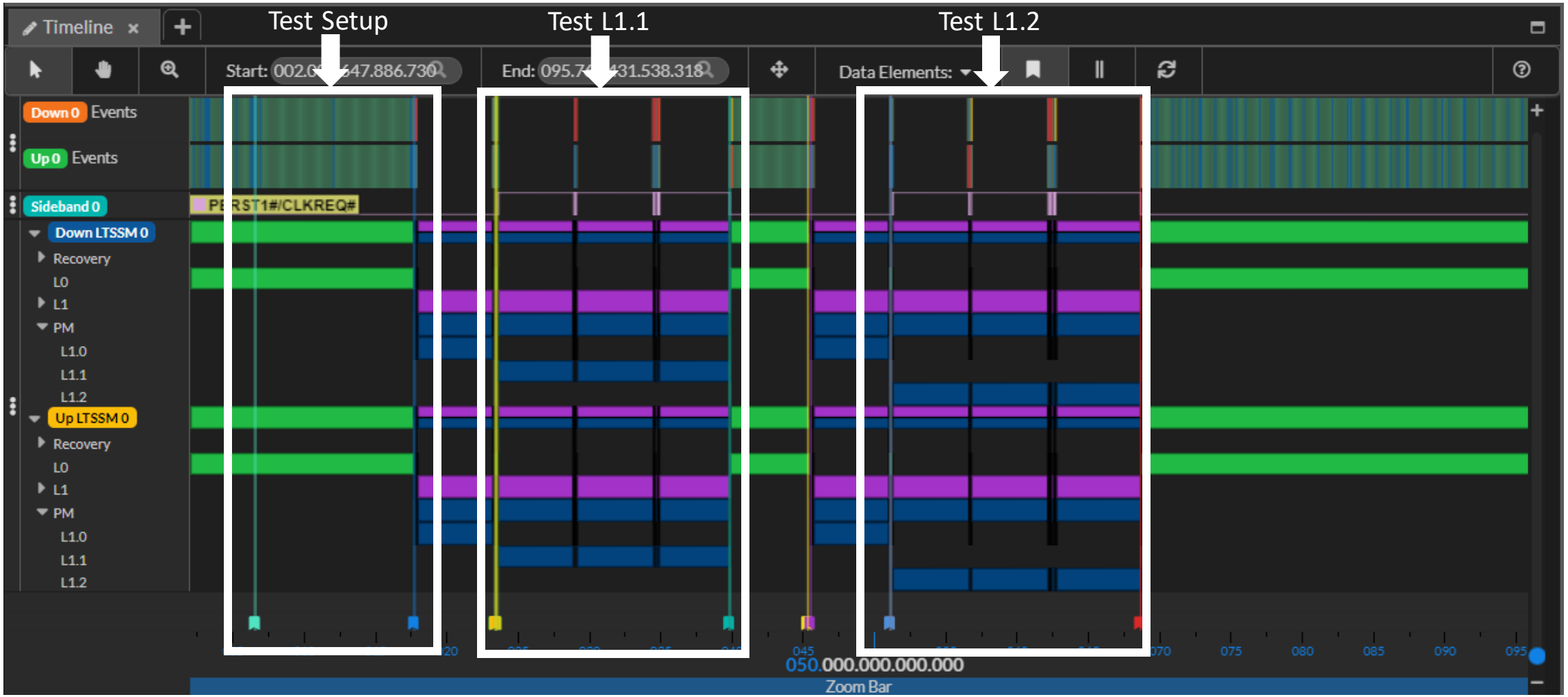
Start the test run | Stop a test run | Reset a test run



List of tests to run

The PM_L1_Substate_Verify.sh test the DUT for the expected power state transition by inspecting the PCIe bus link state and the current state of the CLKREQ# signal.

Case Study 2- Examining the Test



Case Study 2- Enabling Lowest Power State

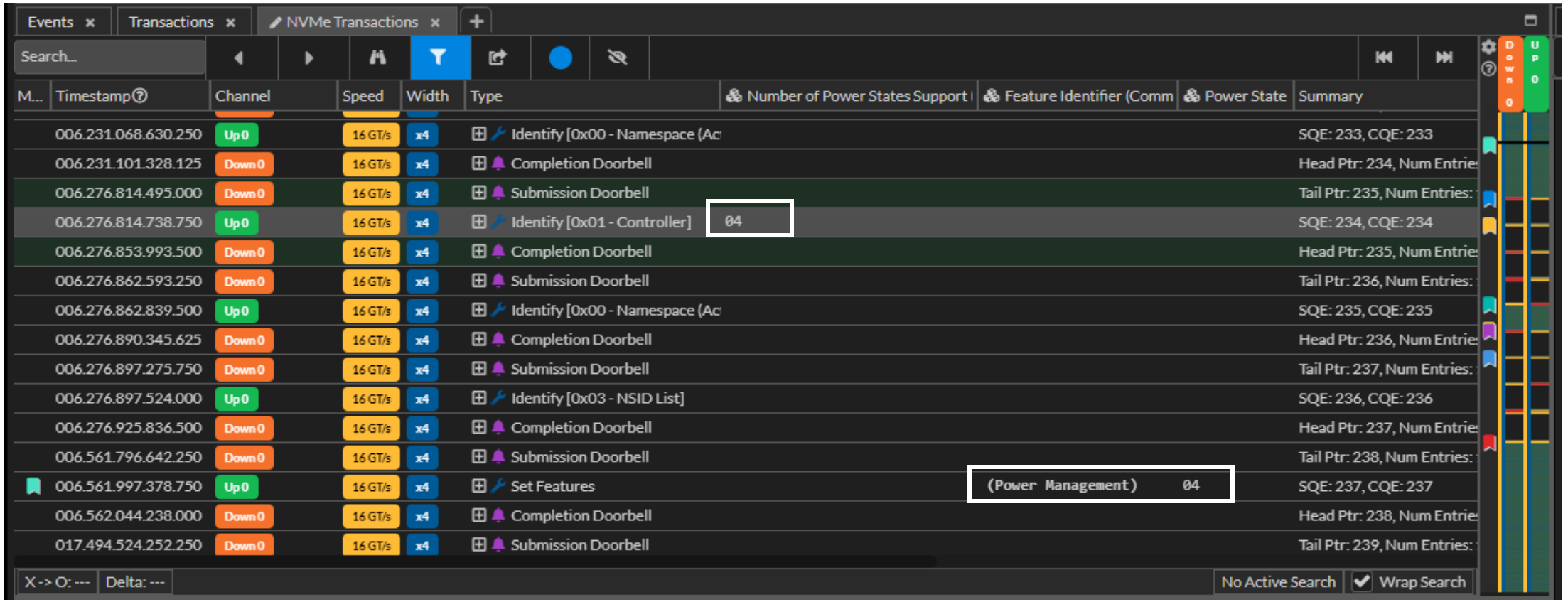
Flash Memory Summit

Get the number of power states supported from the Identify command

- **cat /iport0/target100 |grep ^NPSS**
  NPSS=4

Issue Set Features command

- **io /iport0/target103 SetFeatures -fid 2 -d 4 -w**

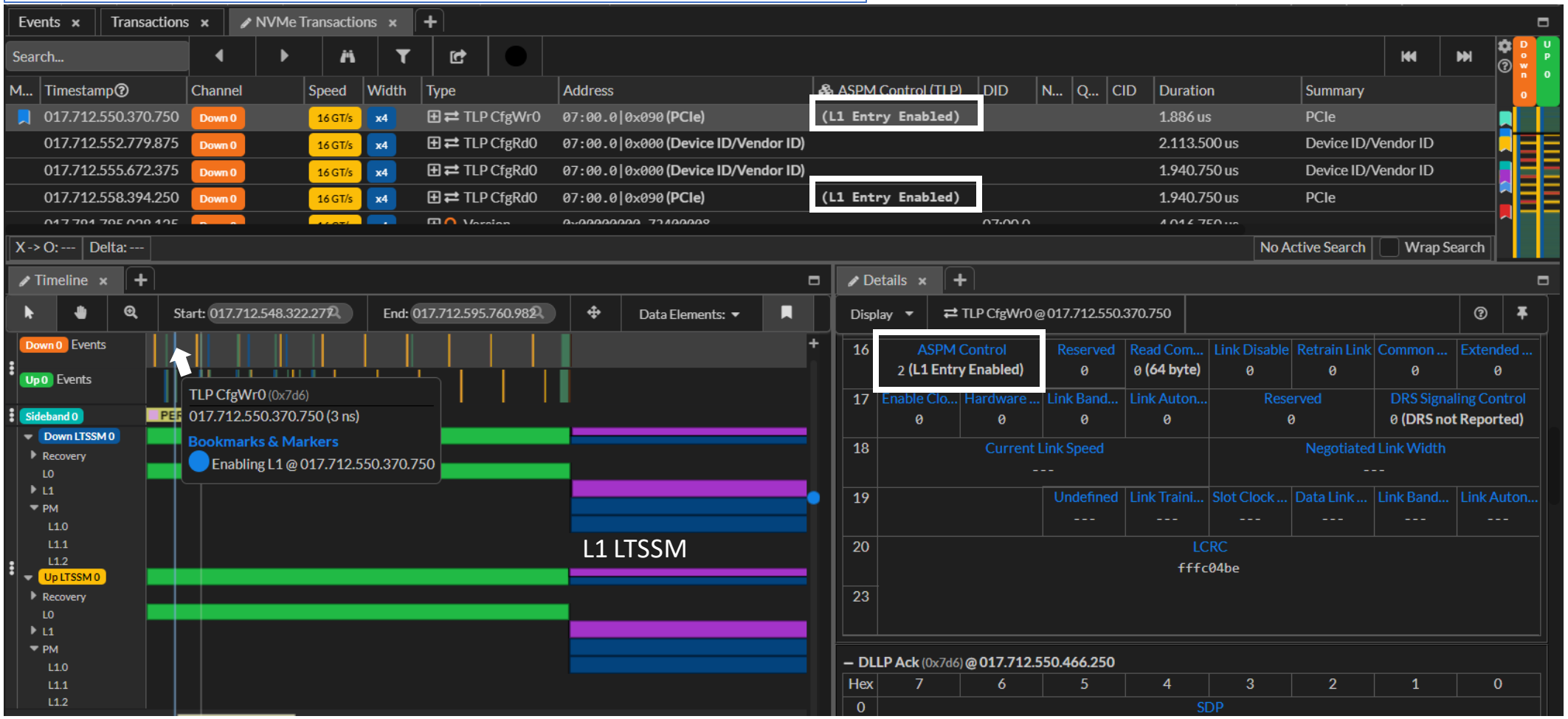


SerialTek
an ellisys company

Case Study 2- Turn on ASPM

Enable ASPM L1 on the DT5 with the following command

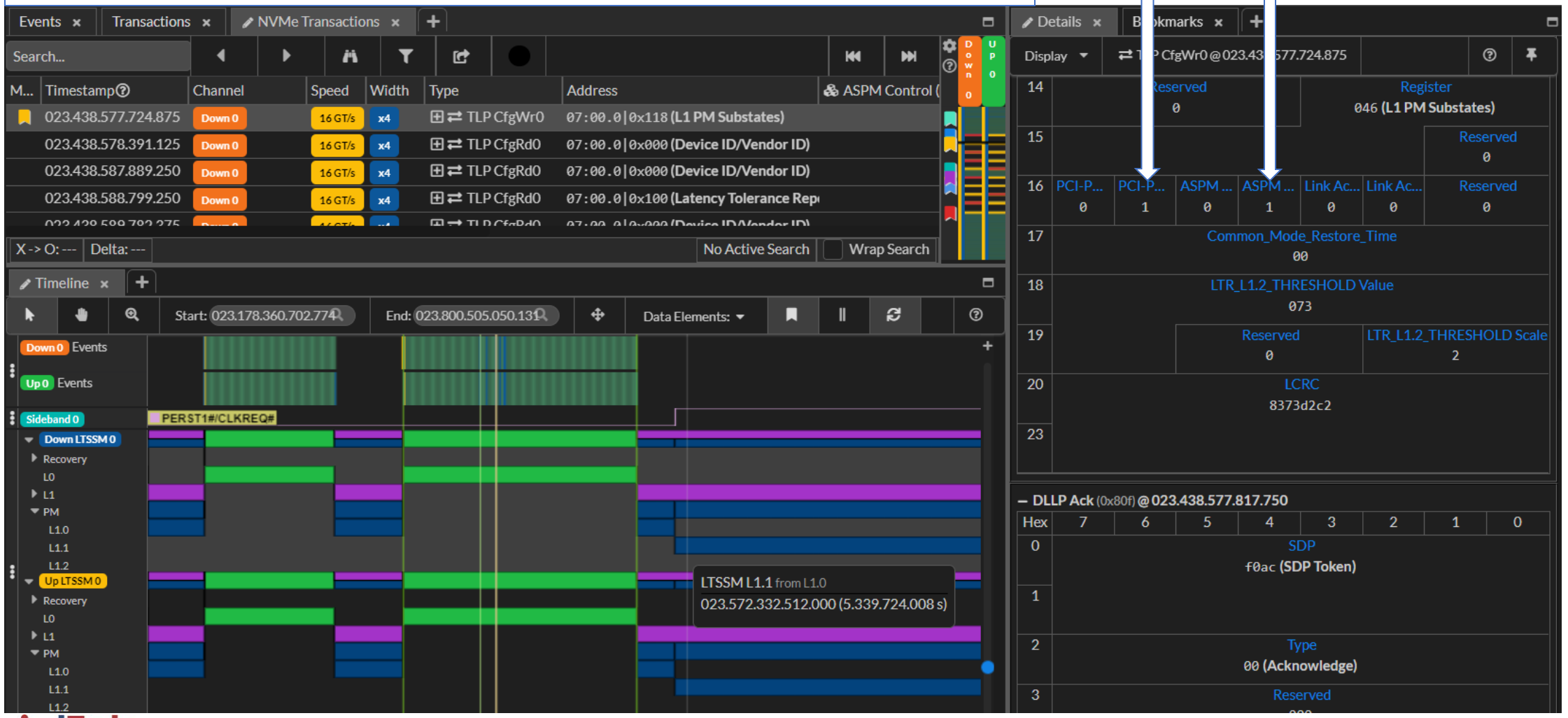
- **sbecho ASPMControl=2 > /iport0/target100**



Case Study 2- Turn on L1.1

Command to enable L1.1 in the LP Control register

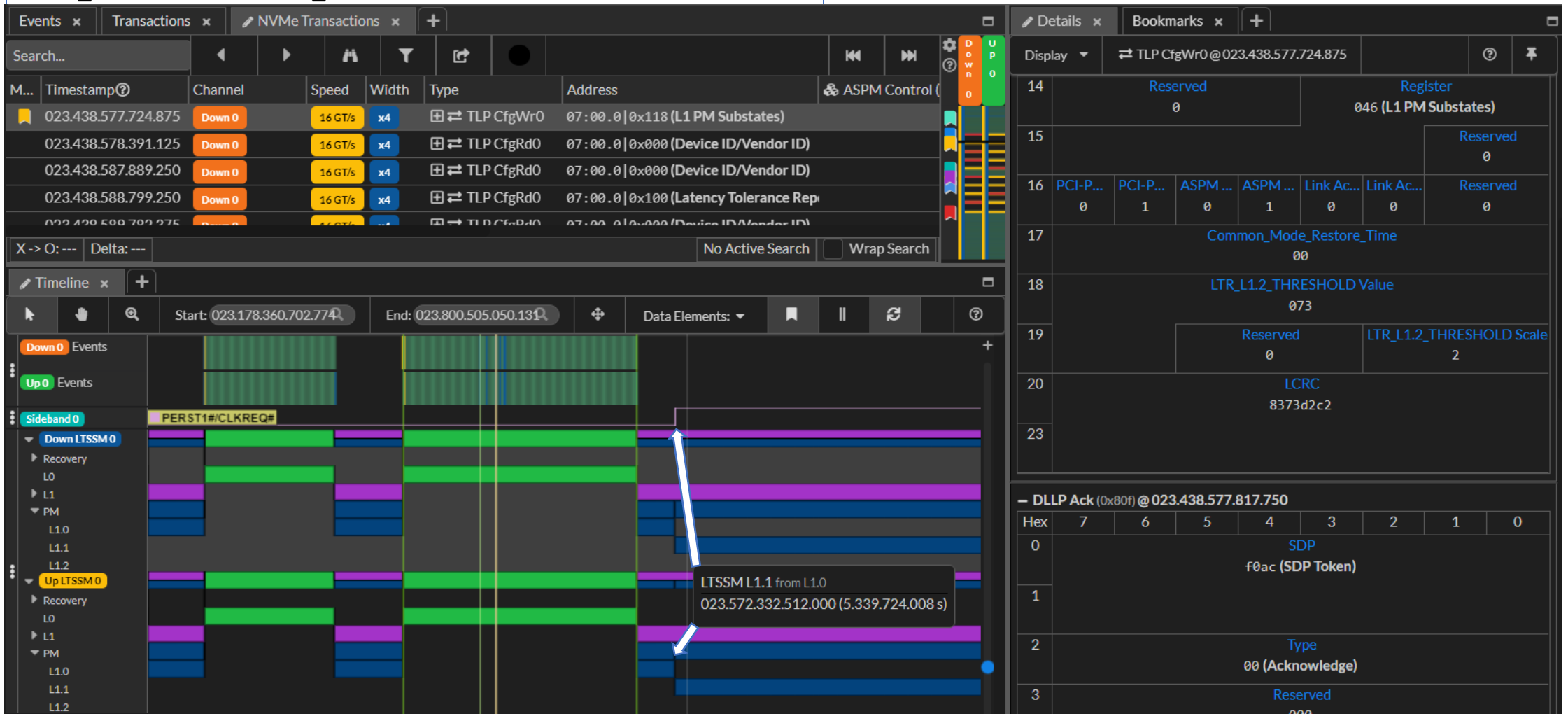
- **sbecho L1PMControl=4 > /iport0/target100**



Case Study 2- De-assert CLKREQ#

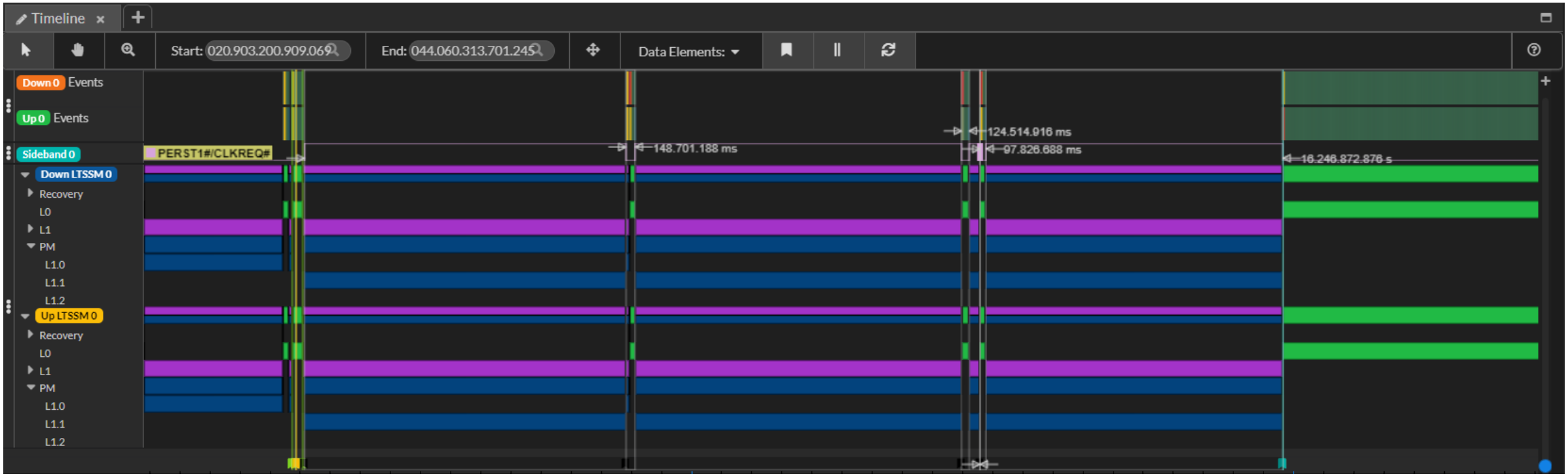
De-assert CLKREQ# by running the following command

- **sb_12x2 –d 0 –f CLKREQ_L –w 1**



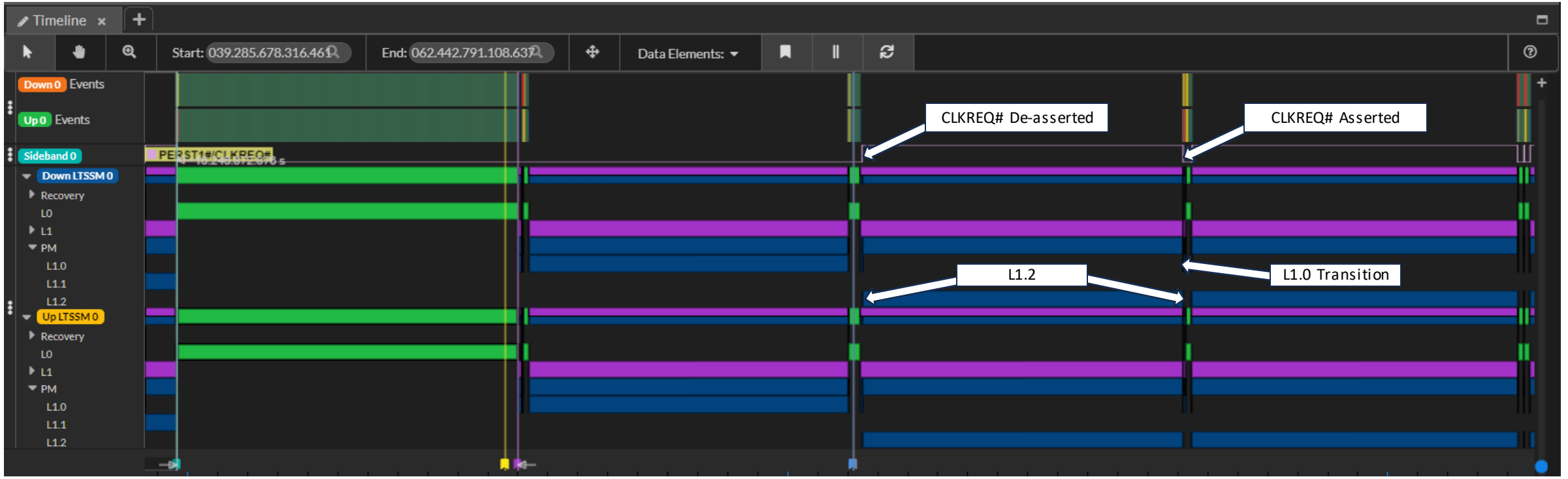
Case Study 2- Verify CLKREQ#

We can now verify that CLKREQ# has been asserted and de-asserted various times as it transitions back and forth from L1 -> L0 -> L1 along with the timings



Case Study 2- Setup Test for L1.2

For L1.2 we change the config space to use L1.2 and then repeat the steps for L1.1 and de-assert CLKREQ# to transition to L1.2.



Case Study 2 Summary

- SANBlaze software allows for easy way to test your devices for support of L1 LTSSM substate and L1 PM substates
- Paired with a SerialTek Kodiak Analyzer, the Sanblaze tool allows you verify that link is properly entering/exiting L1.
- Manually testing L1 LTSSM substates can be difficult, but using a tool such as the SANBlaze SBExpress-DT5 can simplify this testing.

Summary

- There appears a contradiction in L1.Idle/L1.Entry vs L1.0/L1.1/L1.2 substates due to the misunderstanding of the specification
- L1.0/L1.1/L1.2 or Power Management Substates, not LTSSM L1 substates.
- When combined with L1.1 and L1.2 Power management states, devices can achieve significant power savings.
- Using the Quarch PAM and SerialTek Kodiak Analysis system, users can help validate power management changes.
- Using the SANBlaze SBExpress-DT5 and SerialTek Kodiak Analysis system, users can test their devices to verify that the device conforms to the specifications when performing Power Management Substates.