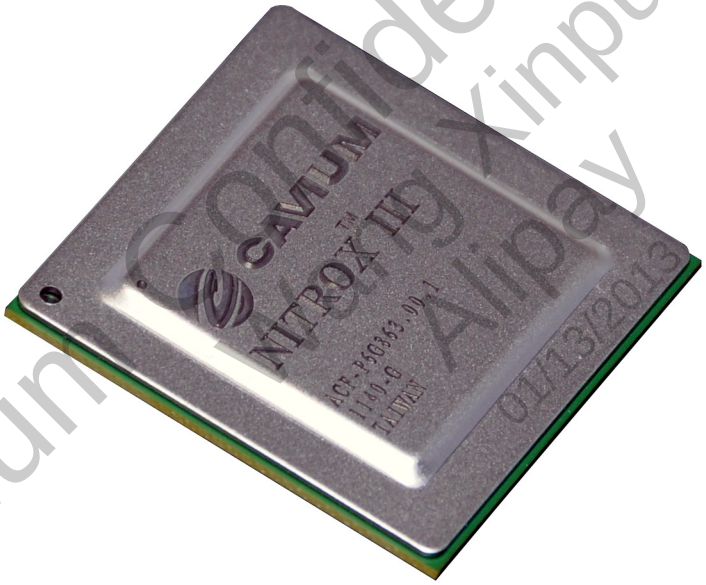
# CNN35XX Design-In Guidelines



PUBLISHED BY
Cavium Inc.
2315 N. First Street
San Jose, CA. 95131
Email : sales@cavium.com
Web : http : //www.cavium.com



History of Changes

| Date | Author | Rev | Comments |
|---|---|---|---|
| 12/14/2011 | PFW | 1.0 | Initial Release |
| | | | |

## Table of Contents

## Table of Figures

# OVERVIEW

The intent of this document is to provide a design-in guideline for the CNN35XX. This document is intended to be used by design and PCB layout engineers of printed circuit boards.

## References and Conventions

Please refer to the CNN35XX Hardware Reference Manual document:

- CNN35XX-HM-V1.0E (HRM)
- CNN35XX Reference Schematics

In the HRM document you will find the following:

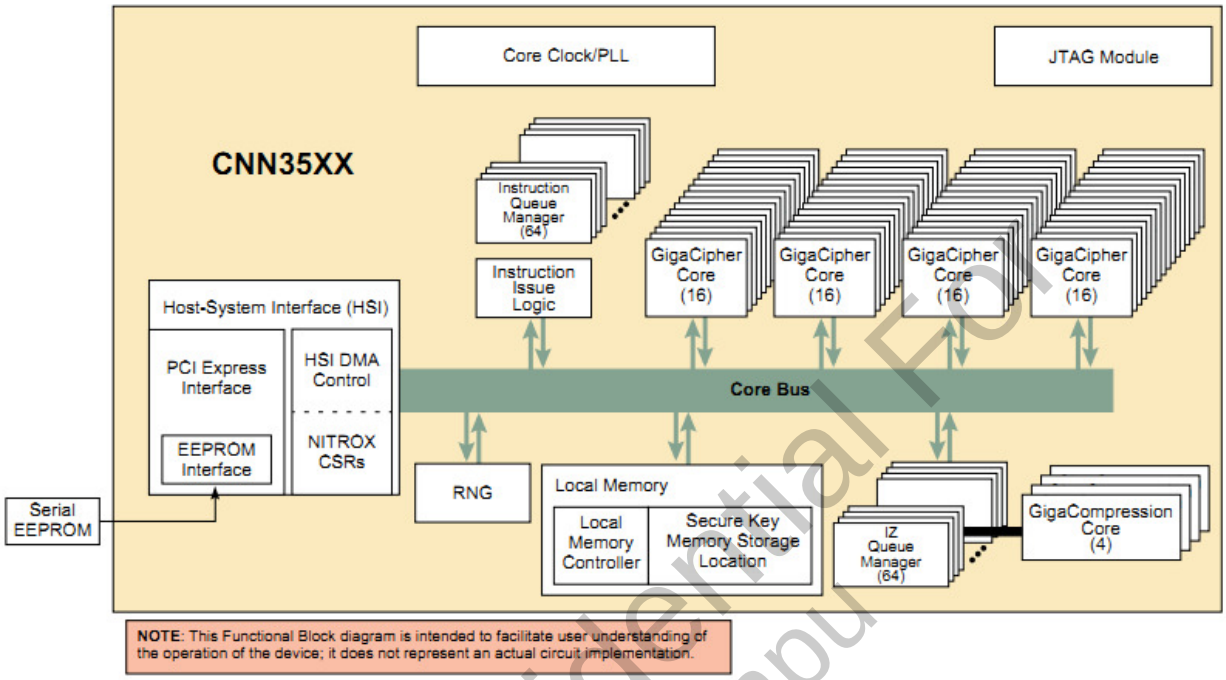
- Signal Descriptions
- AC Characteristics
- DC Characteristics
- BGA Pin-out

Please refer to the following PCI Express documents:

- PCI Express Base Specification Revision 2.0
- PCI Express® Card Electromechanical Specification Revision 2.0

- Signals that are low-active are indicated by the post fix "**_L**"; as in PERST**_L.**
- The term ***PCIe*** will be referred to as ***PCI Express.***
- Unless otherwise stated, the nomenclature ***_n_*** refers to a number as in the following example: EXP_RX_***n***_P; where _n_ could represent EXP_RX_***0***_P, EXP_RX_***1***_P, etc.
- HRM refers to the CNN35XX Hardware Reference Manual

# CNN35XX Block Diagram



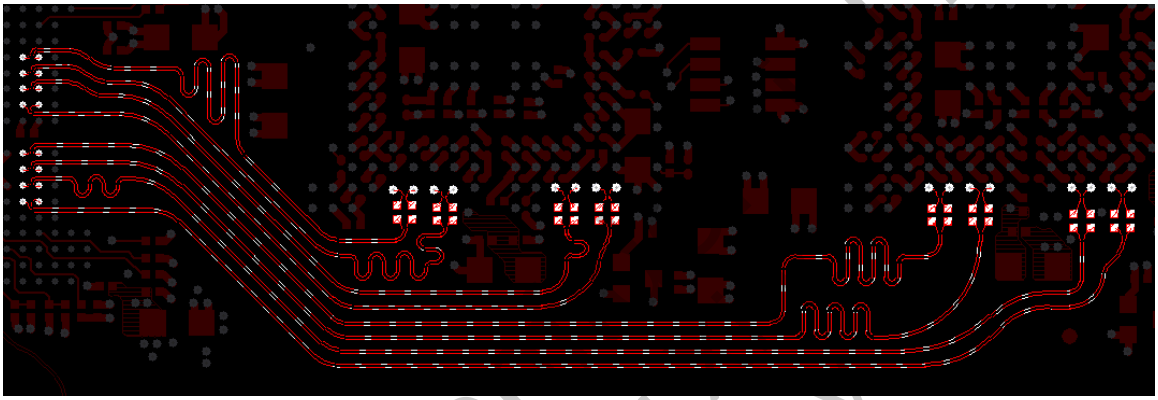

# CNN35XX INTERFACES

## *PCI Express*

Please follow the PCI Express design guidelines. As always, perform board-level signal simulation for the high speed signals.

Length Matching: The signal EXP_RX_n_N should be length-matched to EXP_RX_n_P with a tolerance of less than 5mil. Where "_n_" is the number 0 through 3

Place the AC coupling close to the source.



Differential trace impedance is typically 100 Ohms; however, board level simulation should be the prime criterion to determine the best signal characteristics in term of signal integrity.

## *Serial EEPROM*

The serial EPPROM can be an ATMEL AT93C46W or equivalent. EPR_DI and EPR_DO should be length matched relative to EPR_SK

## *JTAG*

The JTAG interface operates in accordance to the IEEE 1149.1-1990 Standard.
The signal named JTAG_TRST_L should be terminated to GND with a 4.75K Ohm resistor.

# CLOCKS

## *PLL_REF_CLK*

The PLL_REF_CLK is a 50MHz clock. The design engineer should make sure that the clock at the load (CNN35XX) complies with the HRM specification

| Parameter | Description | PLL_REF_CLK = 50MHz | | | Units |
|---|---|---|---|---|---|
| | | Min | Typ | Max | |
| $T_{cyc}$ | Clock Cycle Time | — | 20 | — | ns |
| $T_{high}$ | Clock High Time | 8 | 10 | 12 | ns |
| $T_{low}$ | Clock Low Time | 8 | 10 | 12 | ns |
| $T_{rise}$ | Clock Rise Time[1] | — | — | 2 | ns |
| $T_{fall}$ | Clock Fall Time[2] | — | — | 2 | ns |
| $T_{jitter}$ | Clock Jitter Time[3] | — | — | 100 | ps |

1. 10%-90%
2. 90%-10%
3. cycle to cycle

Shown below is an example a clock generator that produces four individual clocks. Series termination is recommended to adjust for signal reflection due to impedance mismatches. The series termination resistor should be placed at the source. The actual value of the series termination resistor will depend on the output impedance of the clock generator and the trace impedance; the 22.1 Ohm values in the figure below are specific to this particular clock generator and may not necessarily be appropriate for other clock generators.

Keep in mind that even though the component may have favorable specification, poor placement and routing adding to trace parasitics could result in a clock that will be out of HRM specification. It is the CNN35XX needs to see a clock that meets the spec.

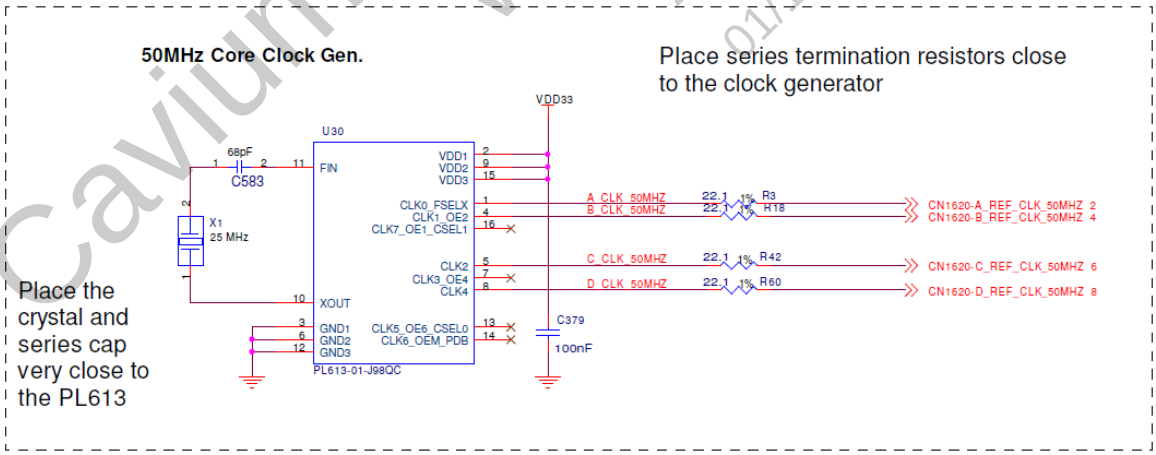



**Figure 1: PLL_REF_CLK Clock Circuit**

## *EXP_REF_CLK*

The EXP_REF_CLK clock is a 100MHz, differential clock. The CNN35XX requires a differential clock with the following specifications:

| Clock Name | Frequency | Frequency Tolerance[1] | Duty Cycle | Rise Time /Fall Time (max)[2] | Phase Jitter (peak-to -peak)[3] | $V_{HIGH}$ Levels | | | $V_{LOW}$ Levels | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Min | Nom | Max | Min | Nom | Max |
| EXP_REF_CLKN/P | 100 MHz | ±300 ppm | 40/60 | 840 ps | 80 ps | 625mV | 700mV | 850mV | –150mV | 0V | 150mV |

The signal EXP_REF_CLK_N should be length-matched relative to EXP_REF_CLK_P with a tolerance of less than 5mil.

When considering a clock generator, follow the manufacturer's placement and routing guidelines. De-coupling caps should be placed close to the clock generator IC.

As mentioned before, even though the component may have favorable specifications, poor placement and routing, which can contribute to trace parasitics, could result in a clock that will be out of specification relative to the CNN35XX-HM.

Some recommendations for PCIe clock generators are as follows. The CNN35XX needs to see a clock within its required specification:

- PhaseLink: PL602-20
- ICS-IDT: ICS9DB202CGLF

An example of a clock scheme is shown below in systems that employ multiple CNN35XX ASICs.

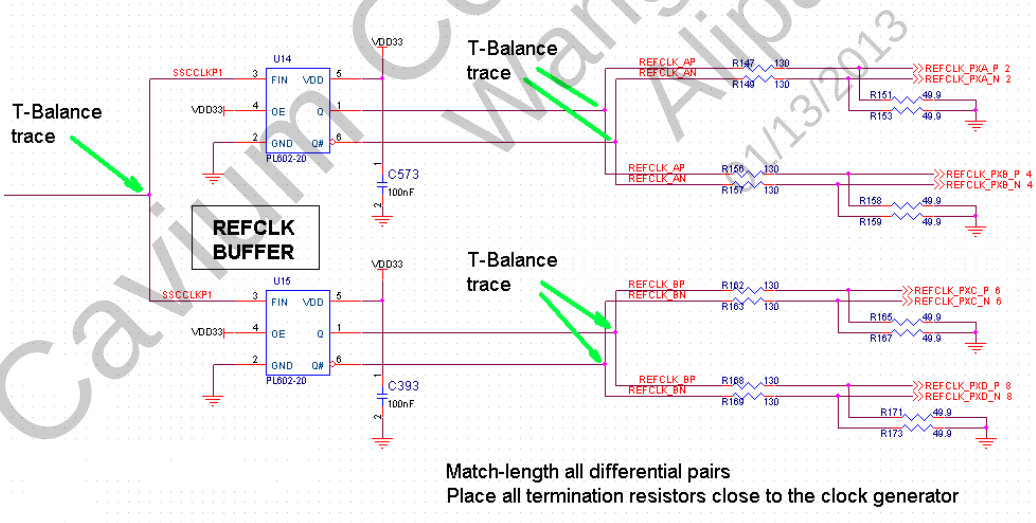



**Figure 2: Differential Pair Clock Circuit**

# POWER

Recommended Operating Supply Voltages

| Parameter | Description | Min | Typ | Max | Units |
|---|---|---|---|---|---|
| CORE_VDD | Core-voltage supply | 0.87 | 0.9 | 0.93 | V |
| EXP_VDD_09V | PCIe core voltage supply | 0.87 | 0.9 | 0.93 | V |
| EXP_VDD_VPH | PCIe 1.8V voltage supply | 1.71 | 1.8 | 1.89 | V |
| EXP_VDD_VTPX | Express analog transmitter termination voltage vptx (1.8V) | 1.71 | 1.8 | 1.89 | V |
| VDD18 | MOS18 supply voltage | 1.71 | 1.8 | 1.89 | V |
| PLL_VDD_18V | PLL supply voltage | 1.71 | 1.8 | 1.89 | V |

The aforementioned table specifies the voltages that the CNN35XX requires.

If using a discrete switching power supply (i.e. not a POLA or other integrated power supply), make sure that switcher supply components are not placed over high frequency traces; in other words, ***do not route high speed traces underneath the switcher power supply components.***

Power rail measurements should be done with a scope probe that has a very short GND lead. The best place for such measurements is to choose a bulk cap on the bottom side of the board where the CNN35XX is located.

## *CORE_VDD, VDD18, EXP_VDD*

Make sure that the copper pours do not present a significant IR drop that would marginalize the rail voltage. For example, if the margin for a voltage rail is 50mV, then the copper pour should not present a 50mV drop between the output of the regulator and the load. Always measure the voltages at the load.

## *EXP_VDD_09, EXP_VDD_VPH, EXP_VDD_VTPX*

EXP_VDD_09, EXP_VDD_VPH, EXP_VDD_VTPX nets should be filtered. The low-pass filter components should be placed close to the CNN35XX package. Provide at least a 20mil trace width from the ferrite bead to the CNN35XX; in the example below, EXP_VDD_09_FLT connects to the balls of the CNN35XX.

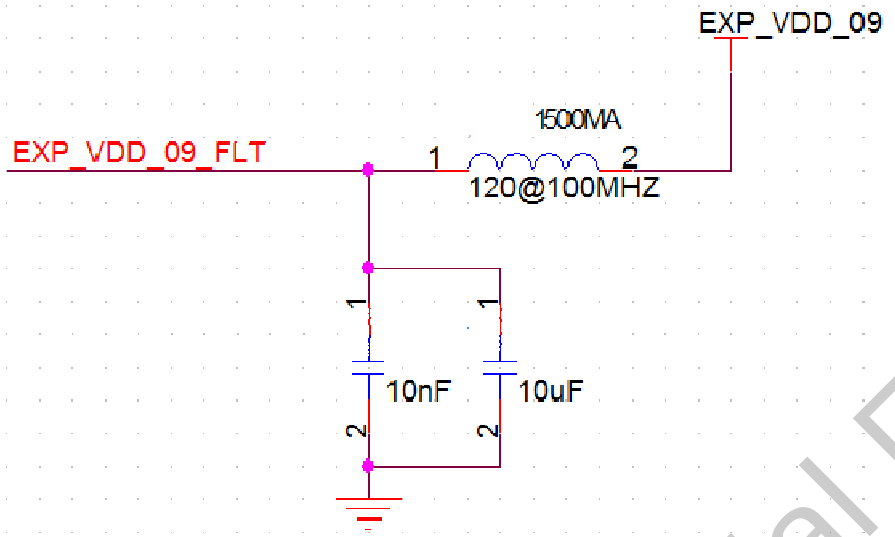



**Figure 3: Rail Low Pass Filter**

## *PLL_E_VDD_18V, PLL_S_VDD_18V, PLL_Z_VDD_18V*

PLL_E_VDD_18V, PLL_S_VDD_18V, PLL_Z_VDD_18V (1.8V) rails are sensitive to noise. Provide at least a 20mil trace width from regulator to the CNN35XX.

PLL_E_VSS_18V, PLL_S_VSS_18V, PLL_Z_VSS_18V should be routed with 20 mil trace widths. The return paths are not to be connected to the GND of the board. The ASIC's substrate has made provisions for the "SS" traces to be internally connected to GND.

See figures below.

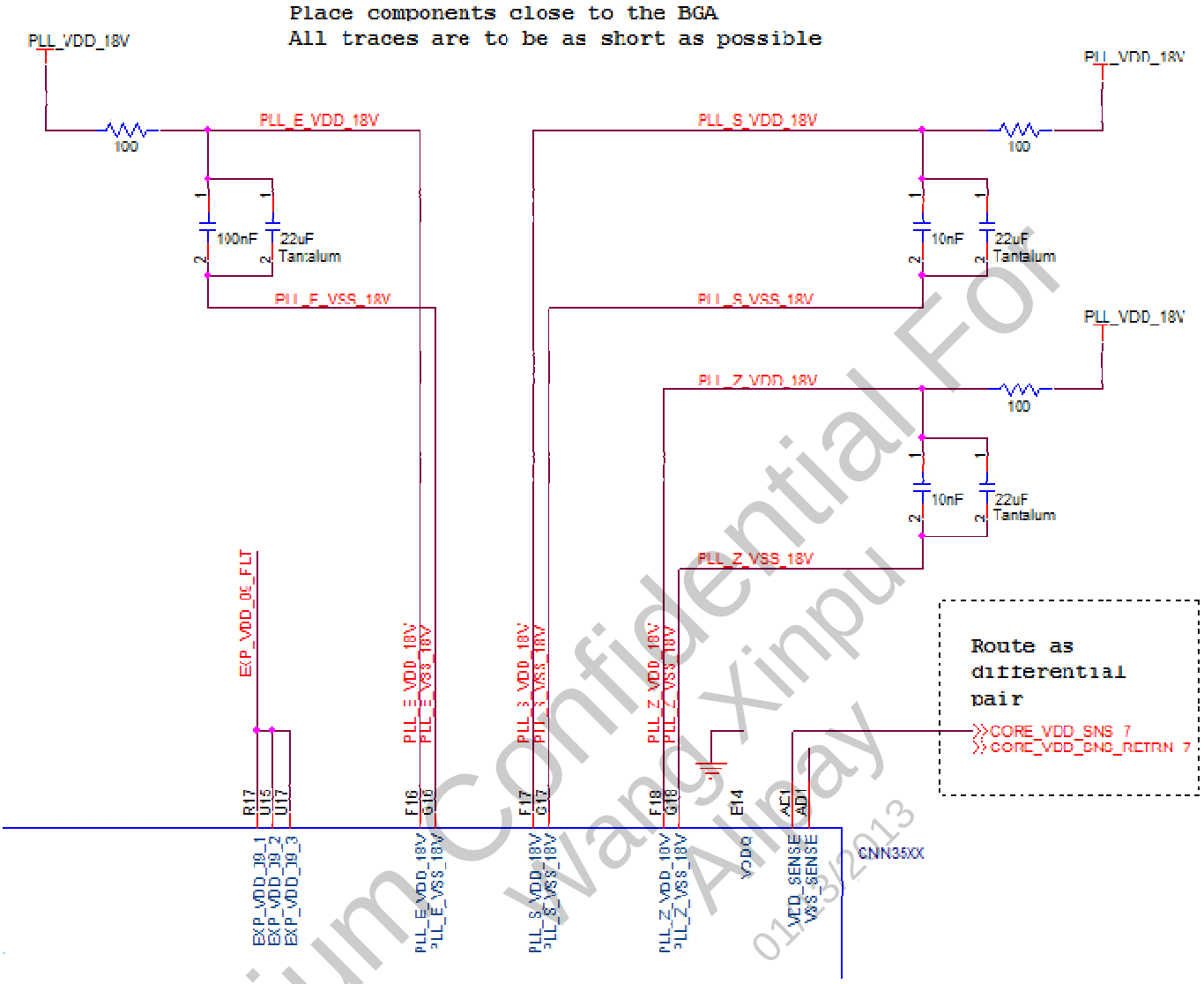



**Figure 4: PLL_E/S/Z_VDD_18V Connections**

## VDD_SENSE

The CNN35XX has two dedicated balls to allow the Core_VDD switching voltage regulator to monitor the voltage at the CNN35XX. All switching power supplies have a Feed Back (FB) input. The FB input requires a voltage that is based on a voltage divider developed using two resistors. Please refer to the specification of the voltage regulator to determine the values of the resistors.

The VDD_SENSE traces should be routed as differential pairs. Please refer to the CNN35XX Reference Schematics for more detail.
Cavium recommends the use of the VDD_SENSE signals; better voltage regulation is achieved by using the VDD_SENSE signals.

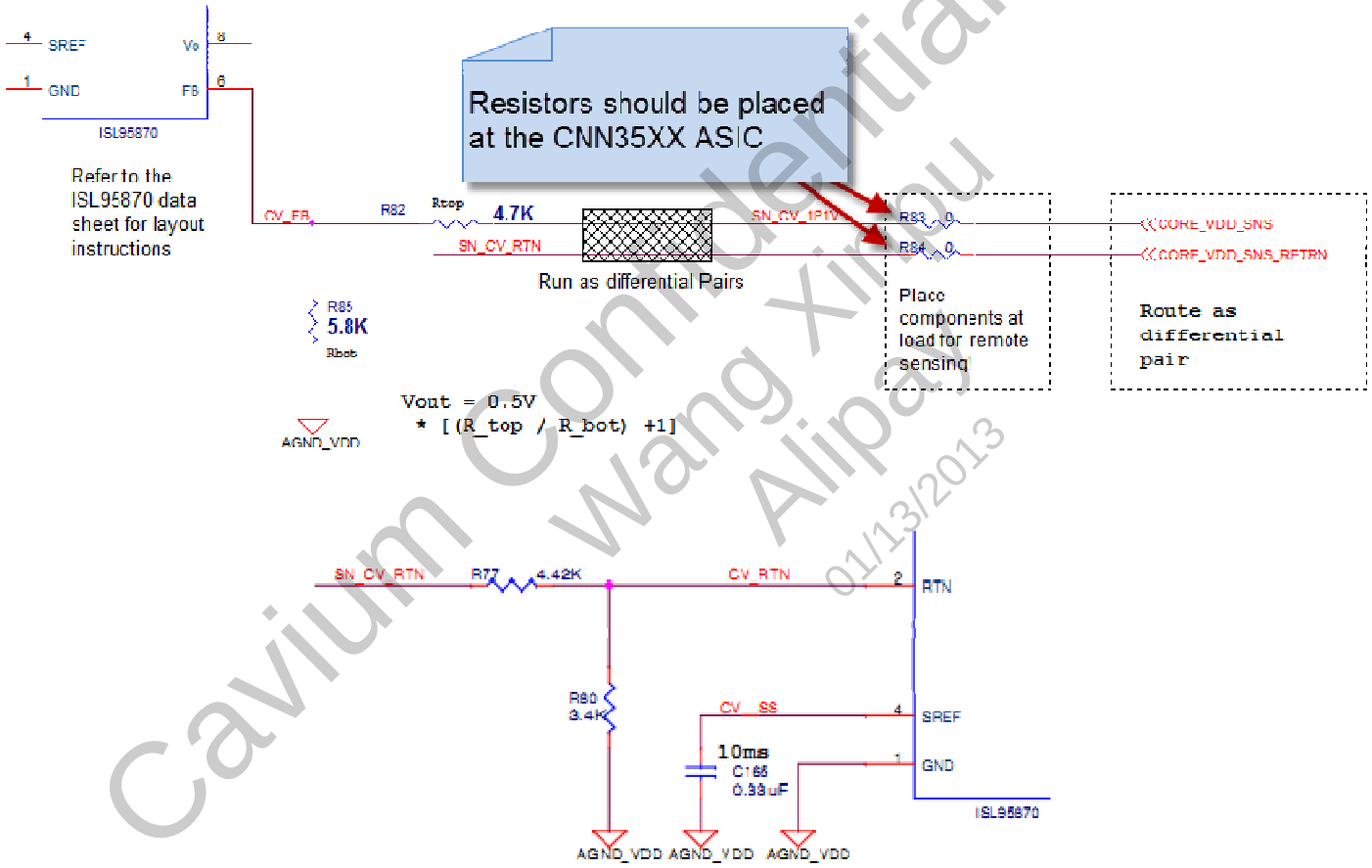
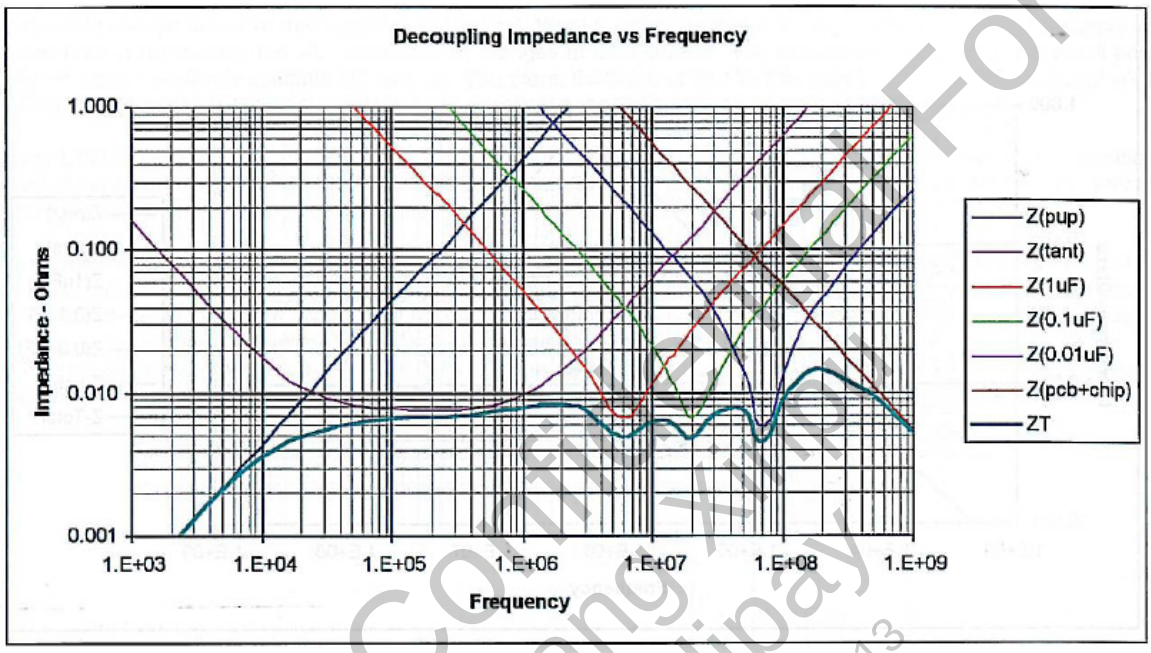



**Figure 5: CORE_VDD_SENSE**

### *Decoupling Caps*

Decoupling depends on the kind of loading on the power supplies. The best method for determining the proper decoupling capacitance is to use a modeling tool. As shown in the diagram below, the PCB impedance should be below 10 m-Ohms. Note that 10nF caps start to become ineffective at frequencies above 100MHz. Plane capacitance must be employed. See the section on board stack-up for more details






## Miscellaneous Circuits

### *PLL_DCOK*

The PLL_DCOK to the CNN35XX means that ***all*** voltage rails are powered up. This design criterion stipulates that the following voltages must be at nominal levels before PLL_DCOK can be raised to a high level. PLL_DCOK must have a rise time faster than 200ns. If any one of the rails drops below HRM specification, PLL_DCOK must drop as well. Refer to the power-up timing diagram below.

1. VDD18, EXP_VDD_VPH, EXP_VDD_VPTX
2. CORE_VDD, EXP_VDD_09

In regards to PLL_DCOK, the timing diagram shows that all rails must be nominal before the signal is raised; be aware that the minimum delay is 3ms (see timing diagram below). Any time delay less than 3ms may result in unpredictable performance.

Note that EXP_REF_CLK (PCIe 100MHz clock) and PLL_REF_CLK (50 MHz) clocks must be toggling before PLL_DCOK is asserted. A clock-detect circuit can be employed to make sure that EXP_REF_CLK is operating at nominal specifications before PLL_DCOK is raised, or the PLL_DCOK delay can be extended. Results from DVT have shown that a 250ms delay for PLL_DCOK is reasonable. There is no maximum delay specification; however, any delay longer than 250ms could raise issues with the Root Complex (RC).

## Power-UP Sequence Timing Diagram

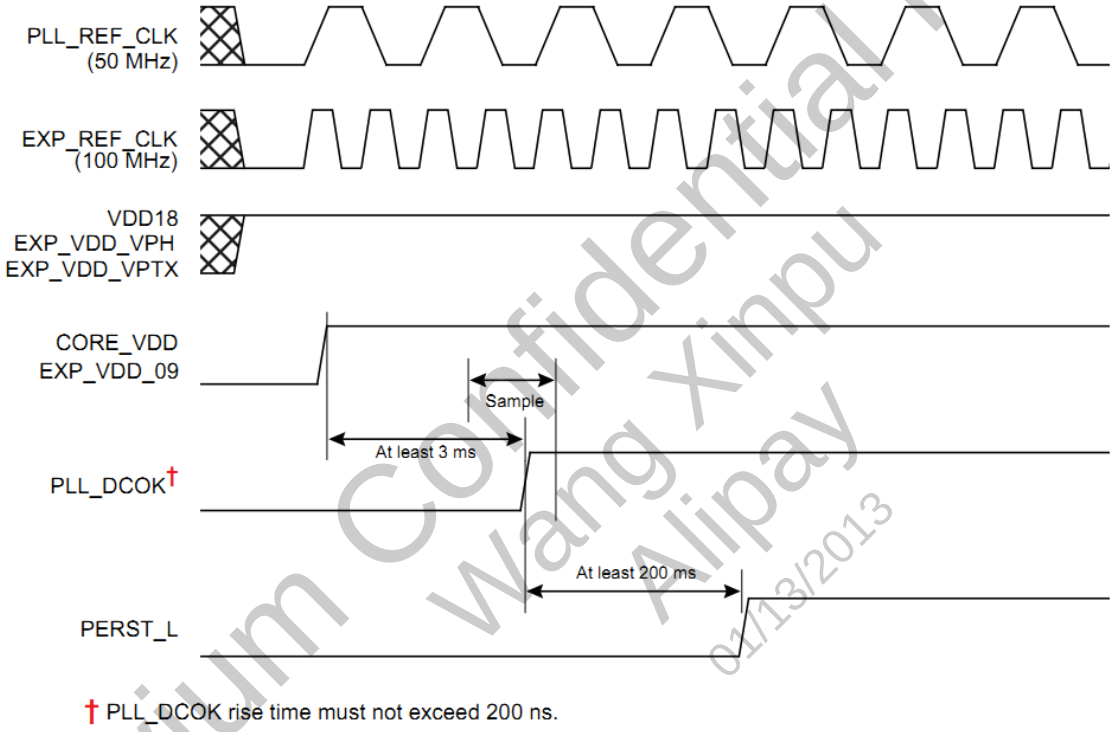



**Figure 6: Power Sequence Timing Diagram**

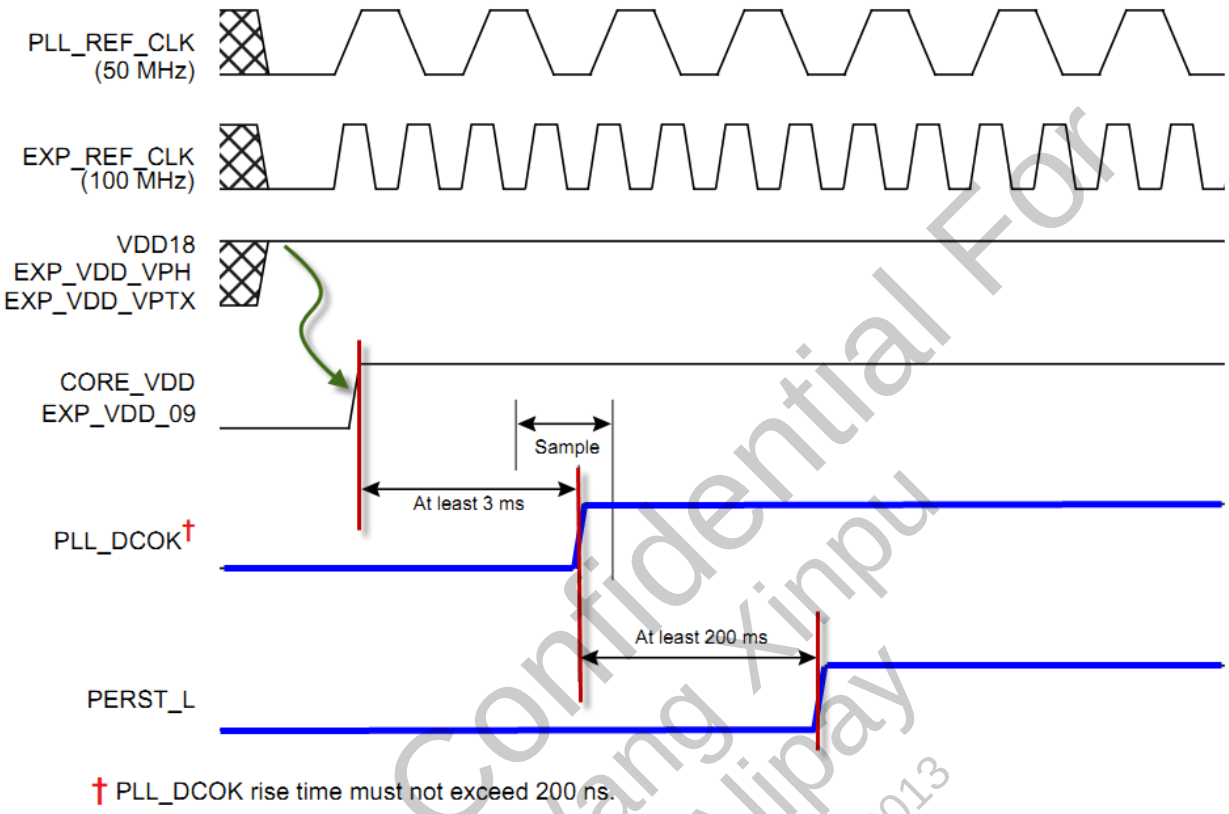
## *PERST_L*

PERST_L should asserted low by any one of the following signals going low:

- PLL_DCOK: ……. Any voltage rail below specification
- PCIE_RESET_L: ... PCIe Reset by the Root Complex or mother board
- PB_RESET_L: …... Manual Push button reset asserted
  Note that push button reset is not a requirement; merely a convenience that the designer may choose to implement.

When PERST_L de-asserts (rises), the rise time must be faster than 200ns

A minimum delay of 200 ms must exist between the rising edge of PLL_DCOK and PERST_L. See the timing diagram below.






Regarding the delay times of PLL_DCOK and PERST_L, the HRM does not specify a maximum delay; however, the Root Complex may have such a delay criteria. Results from DVT have shown that 250ms delay is reasonable.

## *RSVD and RSVD_0*

- All ball names with RSVD_0 must be connected to GND.
- All ball names with RSVD must remain floating; thus, not connected.

### *VDDQ*

- Ball E14 must be connected to GND

### *NC*

- All ball names with NC must remain floating; thus, not connected.

### *PSE Interface*

- Please refer to the HRM for its implementation; otherwise, the signals should be pulled low.

### *ECLK, SCLK and ZCLK_PLL_BYPASS*

- These signals could be pulled low with a 10K, if PLL_BYPASS is not used. Please refer to the CNN35XX-HM

# Board Design Guidelines Basics

| | |
|---|---|
| Stack up | Use "Power Plane Capacitance" to compensate for the ineffectiveness of decoupling caps at high frequency switching currents. Shown below is an example of an eight layer stack-up. Note that the core thickness between layers 2 and 3, also between layers 6 and 7, is 3 mils (0.08mm) |
| IR Drop | Make sure that the copper, for the rail that supplies the current to the load, is wide enough to carry the load current.<br><br>Avoid the "Swiss Cheese" effect where closely grouped multiple VIAs punch through the copper pour effectively reducing the copper conductivity to the BGA.<br><br>Power pours that transitions between layers should have enough vias to provide the required load current. |
| Orthogonal Routing | Traces on layers that have only a Prepreg between them (See stack-up layers 4 and 5) should be routed in an orthogonal fashion; meaning, one layer routed vertically (in the y-direction) and the other layer routed horizontally (in the x-direction). |
| PLL Filter | Low pass PLL filters should use ferrite beads.<br>Filter components should be placed about 100mil from the CPU.<br>The trace width between the CPU and the filter components should be at least 20mils |

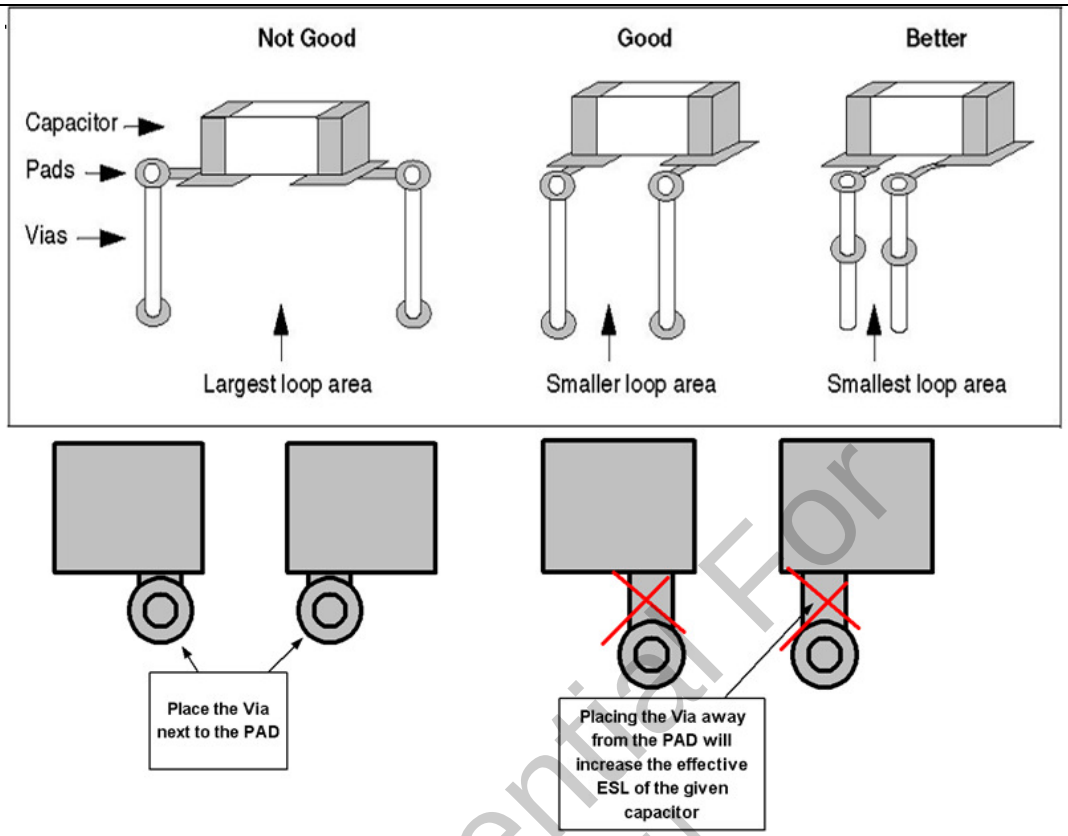
Suggested 8 layer PCB board Stack up

| Conductor name | Layer | Thickness mm | Thickness (mil) |
|---|---|---|---|
| Soldermask | | | 0.7 |
| TOP | 1 Cu 0.5 oz | 0.02 | 0.6 |
| Prepreg | | 0.10 | 4.0 |
| GND - 1 | 2 Cu 1.0 oz | 0.03 | 1.2 |
| Core | | 0.08 | 3.0 |
| Power -1 | 3 Cu 1.0 oz | 0.03 | 1.2 |
| Prepreg | | 0.10 | 4.0 |
| Inner -1 | 4 Cu 1.0 oz | 0.03 | 1.2 |
| Core | | 0.84 | 33.0 |
| Inner -2 | 5 Cu 1.0 oz | 0.03 | 1.2 |
| Prepreg | | 0.10 | 4.0 |
| Power -2 | 6 Cu 1.0 oz | 0.03 | 1.2 |
| Core | | 0.08 | 3.0 |
| GND - 2 | 7 Cu 1.0 oz | 0.03 | 1.2 |
| PrePreg | | 0.10 | 4.0 |
| BOTTOM | 8 Cu 0.5 oz | 0.02 | 0.6 |
| soldermask | | | 0.7 |
| Total | | 1.61036 | 64.8 |

Do not route traces over Voids.

Do not route traces over Voids.

| | |
|---|---|
| Controlled Impedance | Maintain controlled impedance throughout the PCB.<br>Generally speaking, use 50 Ohms as the controlled impedance for single ended signal traces<br>Generally speaking, use 100 Ohms controlled impedance for differential signal traces. |
| Clocks | Place series termination at the source.<br>Where required, follow clock vendor's specific regarding termination requirements.<br>Do **not** route clocks close or underneath components of the switcher power supplies.<br>Provide at least a 3X the trace width for the air gap between the clock and any trace adjacent to the clock trace |
| SERDES | Length-match the trace of the negative signal relative to its corresponding trace of the positive signal to a tolerance within +/-5 mil.<br>Route the traces on as few layers as possible. VIAs that are used to transition the trace amongst the layer will present adverse effects on the signal integrity of the trace. Refer to the PCI Express Specifications R2.0 regarding Lane-to-Lane Skew, Differential Data Trace Impedance. |
| Differential Clocks | Length-match the trace of the negative signal relative to its corresponding trace of the positive signal to within a tolerance of +/-5 mil. Maintain a controlled impedance of 100 Ohms. In reference to the Stack-up in this document, a trace width of 4mil and an air gap of 5mil has shown favorable results; however, board simulation results should be the primary guidelines. |
| Voids (air gap) | **Do not** route traces over copper void or split power planes. Doing so will create an impedance discontinuity that will result in unwanted EMC issues. |
| Series Termination | Series resistor should be placed at the source. |

| | |
|---|---|
| Decoupling Caps.<br><br>Place the decoupling caps close power pin of to the BGA or IC.<br><br>Cap's via placement shown to the right. |  |

# Mechanical Specifications

## *BGA Package*

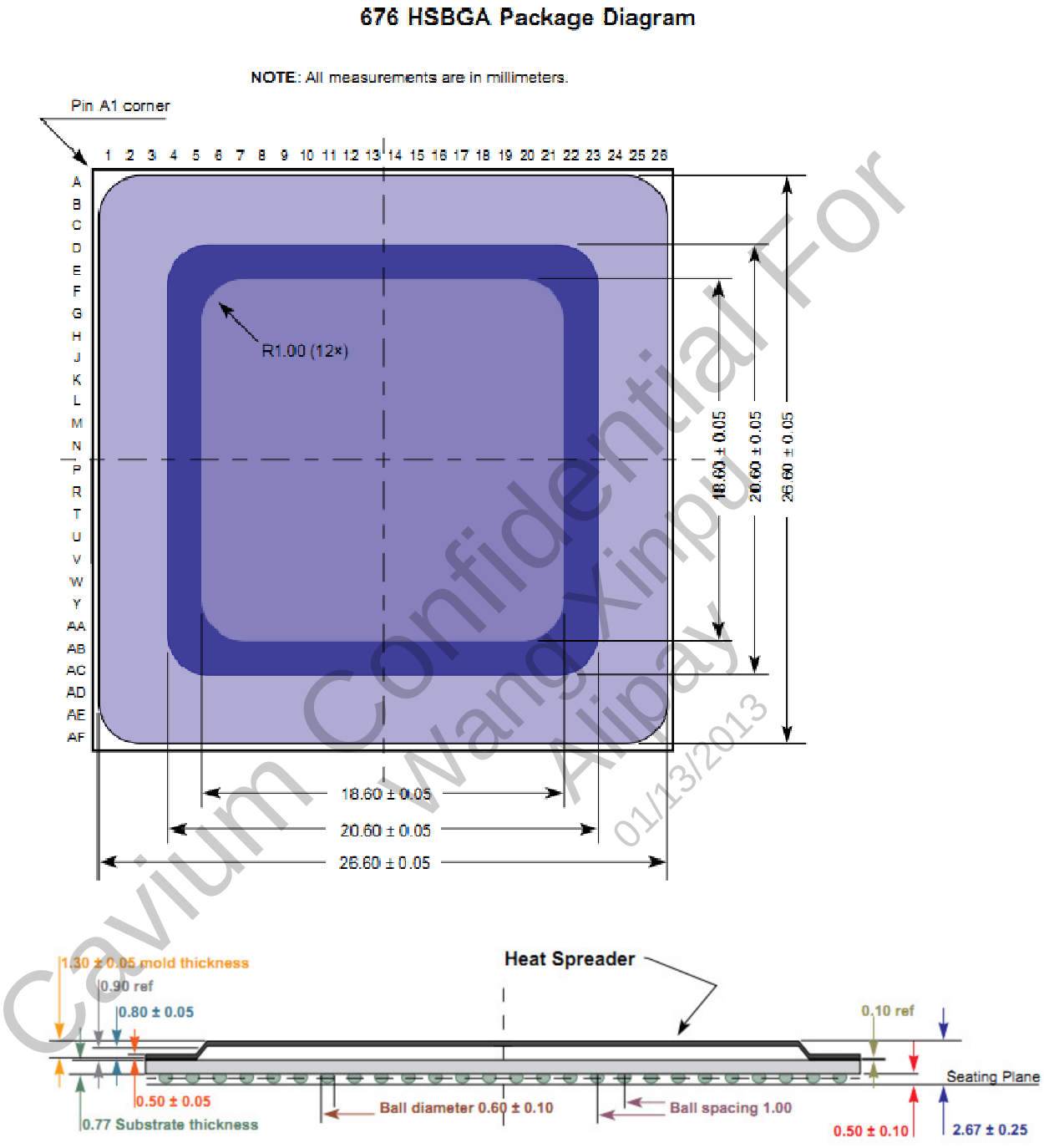



**Figure 7: BGA Mechanical Specifications**

# Thermal Solution Considerations.

Cavium Networks does not provide an off-the-shelf heatsink solution for the CNN35XX. It is recommended that the lead engineer determine the environmental condition into which the CNN35XX will be placed by establishing the following parameters:

- LFM (Linear Feet per Minute): air flow over the CNN35XX
- Ta (Ambient temperature): the maximum temperature (Celsius) in which the CNN35XX will operate.
- P (power dissipation): refer to the HRM.
- Mechanical Constraints: PCI Add-In cards allows a max height of 14.67mm on the top side of the board; thus the heatsink height plus the BGA's height must take that height specification into consideration.

Be advised that the thermal solution should be based on Tc (case temperature) of the CNN35XX. The HRM provides package thermal specifications. It is the case temperature (Tc) that can be measured during DVT. Observed values of Tc can be used to compare with the specification. Case temperature of the CNN35XX must never exceed the temperature rating listed in the HRM specification.

Recommended Operating Temperatures

| Frequency | Junction Temperature ($T_J$) | | Case Temperature ($T_C$) | | Units |
|---|---|---|---|---|---|
| | Min | Max | Min | Max | |
| 500 MHz | 0 | 105 | 0 | 95 | °C |
| 700 MHz | 0 | 105 | 0 | 90 | °C |

Thermal Package Specification for CNN35XX

| Parameter | Description | Max | Units |
|---|---|---|---|
| $\theta_{JA0}$ | Thermal resistance – junction to ambient at 0 m/s or 0 LFM | 9.8 | °C/W |
| $\theta_{JA1}$ | Thermal resistance – junction to ambient at 1 m/s or 200 LFM | 8.6 | °C/W |
| $\theta_{JA2}$ | Thermal resistance – junction to ambient at 2 m/s or 400 LFM | 8.0 | °C/W |
| $\theta_{JC}$ | Thermal resistance – junction to case | 0.6 | °C/W |
| $\theta_{JB}$ | Thermal resistance – junction to board | 4.0 | °C/W |



With the aforementioned system parameters and using the thermal equations, calculate the $\Theta_{SA}$ that is needed for your thermal solution. $\Theta_{SA}$ is the specification of the heatsink

that is needed to keep the case temperature of the CNN35XX below its specification. Choose a heatsink that has a $\Theta_{SA}$ equal to or better than the calculated $\Theta_{SA}$.

**The best thermal solution is the result of simulation using a modeling tool such as ICEPAK.**

## The Empirical Side of Measuring Case temperature (Tc)

One method of measuring case temperature (Tc) is by placing a thermal spot sensor between the case of the CNN35XX and the TIM (Thermal Interface Material) as shown below. Be advised that if the spot sensor is not seated properly, the empirical results may not reflect the actual case temperature.

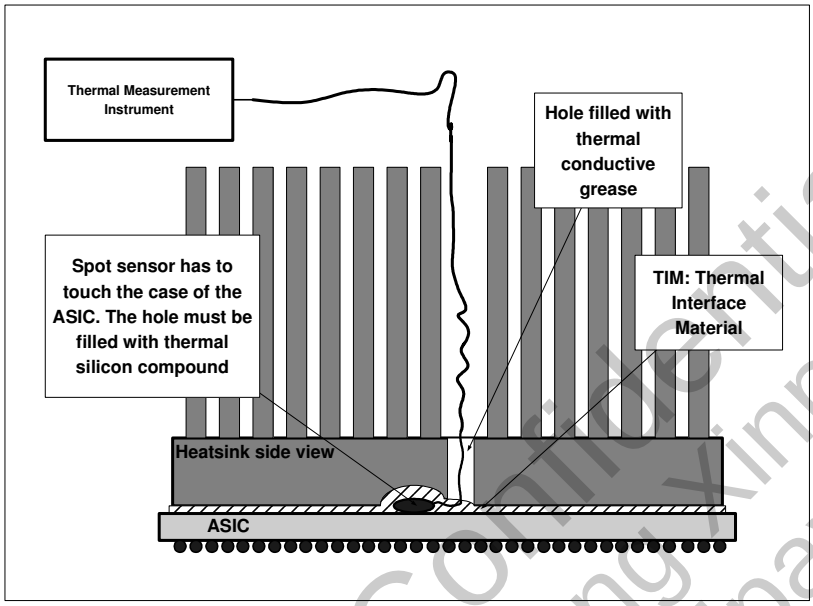



**Figure 8: Thermocouple placement**

## Sample Thermal Calculations

The calculations do not take into consideration the heat developed by surrounding components, convection or forced air flow. To produce a thermal solution, the engineer must perform thermal simulations followed by an empirical DVT.

| CNN35XX | Power Calculation | | | | |
|---|---|---|---|---|---|
| 500 MHz | | Voltage | Current | | Power |
| 56 Cores | Rail | V | mA | | W |
| 16 lanes | CORE_VDD | 0.90 | 18400.00 | | 16.56 |
| | EXP_VDD_09V | 0.90 | 320.00 | | 0.29 |
| | EXP_VDD_VPH | 1.80 | 210.00 | | 0.38 |
| | EXP_VDD_VPTX | 1.80 | 600.00 | | 1.08 |
| | PLL_VDD_18V | 1.80 | 12.00 | | 0.02 |
| | VDD18 | 1.80 | 50.00 | | 0.09 |
| | | | Total Power | | 18.42 |

Power calculations for a CNN35XX; consider the following:

- Core Frequency at 500 MHz
- 56 cores
- 16 PCIe lanes

Power consumption is 18.42 W

Calculated $\Theta_{SA}$ for the heatsink is 1.98 °C/W

| | | |
|---|---|---|
| Ta | Ambient temperature in Celcius | |
| Tj | Junction temperature in Celcius | |
| Tc | Case temperature of the ASIC in Celcius | |
| P | Power dissipated by the ASIC | |
| Theta_sa = [ (Tc-Ta) / P ] - Theta_cs | | |
| | | |
| Theta_sa | 1.98 | C/W - Heat Sink to Air |
| Theta_cs | 0.46 | C/W - Case to Sink (TIM) |
| Theta_jc | 0.60 | C/W - Junction to Case |
| Ta (Ambient) | 50.00 | Celcius |
| Tc (max Case | 95.00 | Celcius |
| Power (max) | 18.42 | W |

Heatsinks are rated by the following:

- Air Flow in unit of LFM (Linear Feet per Minute) or M/S (Meters per Second)
- $\Theta_{SA}$ (Thermal resistance) expressed in units of °C/W

Lowering the LFM increases the thermal resistance.
Ambient temperature is considered to be the temperature of the air within a distance of about 1000mil from the heatsink.

If there are no off-the-shelf heatsinks that have the required $\Theta_{SA}$ at the specified air flow in LFM or M/S, then a custom heatsink that is rated at the desired LFM must be developed.

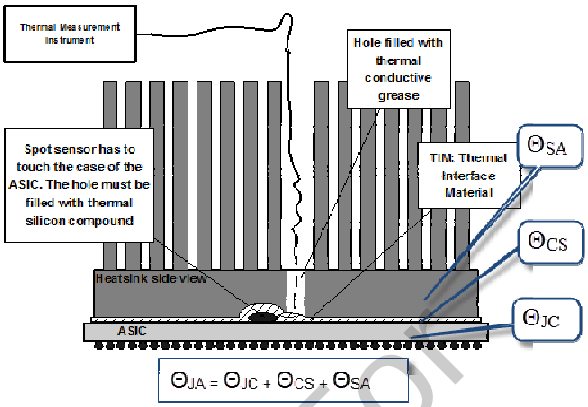
## *Thermal Definitions*

$\Theta_{JA}$ = Thermal resistance from the die's junction to ambient air expressed in °C/W

- $\Theta_{JA} = \Theta_{JC} + \Theta_{CS} + \Theta_{SA}$

$\Theta_{JC}$ = Thermal resistance of the die junction to package case expressed in °C/W

$\Theta_{CS}$ = Thermal resistance of the package case to the heatsink expressed in °C/W; also known as the TIM (Thermal Interface Material).




$\Theta_{SA}$ = Thermal resistance of the heatsink to the ambient air expressed in °C/W

$\Theta_{SA} = [(T_C - T_A) / P] - \Theta_{CS}$

$T_A$: Ambient temperature (°C). Temperature of the ambient air around the device.

- $T_A = T_J - P * \Theta_{JA}$
- $T_A = T_C - P * (\Theta_{JA} - \Theta_{JC})$

$T_C$: Case temperature (°C). Temperature of the case of the device package; the measurement is made using thermocouples placed on the warmest point of the package (in the middle of the upper cover; see figure 8).

- $T_C = T_A + P * (\Theta_{JA} - \Theta_{JC})$

$T_J$: Junction temperature (°C). Average junction temperature of the die within the package.

- $T_J = P * \Theta_{JA} + T_A$
- $T_J = P * \Theta_{JC} + T_C$

P: Total power dissipation of the device (W); the sum of power dissipation from all of the device's power supplies.