Intel® Xeon™ Processor Multi-Processor Platform Design Guide

November 2002

Document Number: 250397-002

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. Intel products are not intended for use in medical, life saving, or life sustaining applications.

Intel may make changes to specifications and product descriptions at any time, without notice.

Designers must not rely on the absence or characteristics of any features or instructions marked “reserved” or “undefined.” Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them.

This document contains information on products in the design phase of development. Do not finalize a design with this information. Revised information will be published when the product is available. Verify with your local sales office that you have the latest datasheet before finalizing a design.

The Intel® Xeon™ processor MP may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.

*Other names and brands may be claimed as the property of others.

Intel, Pentium, Pentium III Xeon, Intel Xeon, Intel NetBurst, MMX and the Intel logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Copyright 2002 Intel Corporation.

Contents

1 **Introduction** 11

- 1.1 Related Documentation 11
- 1.2 Conventions and Terminology 13

2 **System Overview** 17

- 2.1 The Intel® Xeon™ Processor MP and the Intel® Xeon™ Processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process 17
- 2.2 Bandwidth Summary 18

3 **Processor Quadrant Layout** 19

4 **Platform Stack-Up and Placement Overview** 21

- 4.1 Platform Component Placement 21
- 4.2 4-Way System Stack-Up 22
  - 4.2.1 Design Recommendations 22
  - 4.2.2 Design Considerations 23

5 **Clock Routing Guidelines** 25

- 5.1 System Bus Clocking Guidelines 25
  - 5.1.1 Routing Guidelines for BCLK[1:0] 25

6 **System Bus Routing** 31

- 6.1 Return Path 32
- 6.2 Serpentine Routing 33
- 6.3 System Bus Decoupling Requirements 33
  - 6.3.1 Processor I/O Decoupling Requirements 34
  - 6.3.2 Chipset System Bus I/O Decoupling Recommendations 35
- 6.4 Routing Guidelines for a 4-Way System 35
  - 6.4.1 Topology and Routing 37
    - 6.4.1.1 Design Recommendations 37
    - 6.4.1.2 4X Group (DSTBP [3:0]#, DSTBN [3:0]#, D [63:0]#, DBI [3:0]#) 38
    - 6.4.1.3 2X Address Group (ADSTB [2:0]#, A [35:3]#, REQ [4:0]#) 41
    - 6.4.1.4 Common Clock 41
    - 6.4.1.5 Wired-OR 41
    - 6.4.1.6 Design Considerations 42
  - 6.4.2 Routing Guidelines for Asynchronous GTL+ and Other Signals 43
    - 6.4.2.1 Topology 1: Asynchronous GTL+ Signals Driven by the Processors; FERR#, IERR#, PROCHOT# and THERMTRIP# 44
    - 6.4.2.2 Topology 2: Asynchronous GTL+ Signals Driven by the Chipset; A20M#, IGNNE#, INIT#, LINT[1:0], PWRGOOD, SLP#, SMI#, and STPCLK# 46
    - 6.4.2.3 Topology 3: VID[4:0] 46
    - 6.4.2.4 Topology 4: SMBus Signals 46
    - 6.4.2.5 Topology 5: BR[3:0]# Signals 47
    - 6.4.2.6 Topology 6: COMP[1:0] Signals 48
    - 6.4.2.7 Topology 7: ODTEN Signal 48
    - 6.4.2.8 Topology 8: TESTHI[6:0] Signals 48

6.4.2.9 Topology 9: SKTOCC# Signal 49
6.4.3 Debug Port Signals Routing Guidelines for 4-way Configurations 49

7 Mechanical and EMI Design Considerations 51

7.1 Retention Mechanism Placement and Keep-Outs 51
7.2 Electromagnetic Interference Considerations 54
7.2.1 Introduction 54
7.2.2 Terminology 55
7.2.3 Brief EMI Theory 55
7.2.4 EMI Regulations and Certifications 55
7.2.5 EMI Design Considerations 56
7.2.6 Spread Spectrum Clocking (SSC) 56
7.2.7 Differential Clocking 57
7.2.8 Heatsink Effects 58
7.2.9 EMI Ground Frames and Faraday Cages 58
7.2.10 EMI Test Capabilities 61
7.2.11 Summary 61

8 Processor Power Distribution Guidelines 63

8.1 Introduction 63
8.2 Terms 63
8.3 Power Delivery Overview 64
8.4 Processor Power Delivery Ingredients 64
8.5 System Design 65
8.5.1 Multiple Voltages 65
8.5.2 Voltage Sequencing 65
8.5.3 Block Diagrams with Voltage Regulator Modules 65
8.6 Processor Load 66
8.6.1 Processor Voltage Tolerance 66
8.7 Voltage Regulator 67
8.7.1 Voltage Regulator Design 67
8.7.2 Voltage Regulator System Matching 67
8.7.2.1 Voltage Regulator Output 67
8.7.2.2 Voltage Regulator Input 67
8.7.2.3 Voltage Regulator Cooling 67
8.7.2.4 Voltage Regulator Remote Sense Connection 68
8.7.2.5 Voltage Regulator Module ISHARE Connection 68
8.7.2.6 Voltage Regulator Module OUTEN Connection 68
8.8 Power Planes 68
8.8.1 Layer Stack-Up 68
8.8.2 Sheet Inductance/Resistance and Emission Effects of Power Plane 69
8.9 Decoupling Capacitors 71
8.9.1 Decoupling Technology and Transient Response 71
8.9.2 Location of High-Frequency Decoupling 71
8.9.3 Location of Bulk Decoupling 72
8.9.4 Decoupling Recommendation 72
8.10 Component Placement and Modeling 73
8.10.1 Component Models 73
8.10.2 Processor Socket-Package Lump Model 73
8.10.3 Multi-Processor Component Placement and Models 77
8.11 Validation Testing 83

8.12 Generating and Distributing GTLREF[3:0] 83
8.12.1 GTLREF [3:0] 84
8.13 Filter Specifications for V_{CCA}, $V_{CCIOPLL}$, and V_{SSA} 85

9 Methodology for Determining Topology and Routing Guidelines 91

9.1 Timing Methodology 92
9.1.1 Source Synchronous 92
9.1.1.1 Setup Time 93
9.1.1.2 Hold Time 94
9.1.2 Common Clock 95
9.1.2.1 Setup Margin 96
9.1.2.2 Hold Margin 98
9.1.3 Data and Address Setup Time to BCLK 98
9.1.4 Timing Spreadsheet 98
9.2 Simulation Methodology 99
9.2.1 Design Optimization 99
9.2.2 Signal Categories and Topology Options 99
9.2.3 Sensitivity Analysis 99
9.2.4 Signal Quality Metrics 100
9.2.4.1 Noise Margin 100
9.2.4.2 Ringback 101
9.2.5 Timing Metrics 101
9.2.5.1 Setup Flight Time 101
9.2.5.2 Calculating Flight Time for Signals with Corrupt Signal Quality 101
9.2.5.3 Incorporating Package Effects into the Flight Time 103
9.2.6 Parameter Sweeps and Monte Carlo Analysis 103
9.2.6.1 Parameter Sweeps 104
9.2.6.2 Final Solution Space 106

10 System Theory 107

10.1 AGTL+ Logic 107
10.2 Inter-Symbol Interference 107
10.3 Crosstalk 109
10.3.1 Single Line Equivalent Model (SLEM) 112
10.3.2 Serpentine Traces 113

11 Design Checklist 115

11.1 Processor Family Connection Checklist 115

Figures

3-1 Top View - Intel® Xeon™ Processor MP Socket Quadrant Layout 19
3-2 Top View - Intel® Xeon™ Processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process Socket Quadrant Layout 20
4-1 4-Way Processor Server Component Placement Example in a Midrange SSI Form Factor 21
4-2 Twelve Layer Stack-Up for a 4-Way System 22
4-3 4-Way Stack-Up Example 23
5-1 4-Way Processor BCLK Topology 26
5-2 Source Shunt Termination 27
5-3 Agent-to-Agent BCLK Skew 29
5-4 Dielectric Height to Trace Width Diagram 29
6-1 Serpentine Spacing - Diagram of Spacing to Reference Plane Height Ratio 33
6-2 System Bus I/O Decoupling Guidelines for the Processor 34
6-3 System Bus I/O Decoupling Guidelines for the Chipset 35
6-4 4-Way Processor System Bus Topology 37
6-5 Cross-Sectional View of 3:1 Ratio for Symmetric Stripline (Edge-to-Edge Trace Spacing vs. Trace to Reference Plane Height) 38
6-6 L1 vs. L2 Length Dependencies 39
6-7 Wired-OR Topology 42
6-8 0.025" Via Pad with 50% of Trace over Reference Plane 43
6-9 Topology 1 for 4-Way Configuration 45
6-10 Example Voltage Translator Circuit 45
6-11 Topology 2 for 4-Way Configuration 46
6-12 BR[3:0]# Connection for 4-Way Configuration 47
7-1 Retention Mechanism Outline 52
7-2 Retention Mechanism Placement and Keep-Out Overview 53
7-3 Retention Mechanism Ground Ring 54
7-4 Spread Spectrum Modulation Profile 56
7-5 Impact of Spread Spectrum Clocking on Radiated Emissions 57
7-6 Cancellation of H-fields through Inverse Currents 57
7-7 Conceptual Processor Ground Frame 59
7-8 EMI Ground Pad Size and Locations 60
8-1 Power Distribution Block Diagram for 4- Way System Motherboard with Voltage Regulator Modules 66
8-2 Suggested Twelve Layer Stack-Up for Four Processor Systems 69
8-3 1206 Capacitor Pad and Via Layouts 72
8-4 Connections to Via Patterns 72
8-5 Processor Lump Model Schematic 74
8-6 Processor Lump Model Drawing 75
8-7 “Row” Pattern with Voltage Regulator Module 76
8-8 “Row” Pattern with Voltage Regulator Module Schematic 77
8-9 “Square” Pattern with Voltage Regulator Module 78
8-10 “Square” Pattern with Voltage Regulator Module Schematic 79
8-11 “Row” Pattern with Voltage Regulator Module 81
8-12 “Row” Pattern with Voltage Regulator Module Schematic 82
8-13 GTLREF 83
8-14 Suggested GTLREF Generation 85
8-15 Filter Topology 86

8-16 Filter Specifications 87
8-17 Implementation 1 Using Discrete R 88
8-18 Implementation 2 No Discrete R 89
8-19 Example of Decoupling for a Microstrip Baseboard Design 89
9-1 Simulation Methodology Flowchart 91
9-2 Source Synchronous Timing Diagram for Setup Time 94
9-3 Source Synchronous Timing Diagram for Hold Time 95
9-4 Circuit Used to Develop the Common Clock Timing Equations 96
9-5 Timing Diagram Used to Determine the Common Clock Setup Timing Equations 97
9-6 Timing Diagram Used to Determine the Common Clock Hold Timing Equations 98
9-7 Traditional Method of Calculating Rising Edge to Rising Edge Flight Time, Assuming a Linear Edge from V_{IL} Through V_{IH} at the Receiver 102
9-8 Method of Calculating Setup Flight Time When the Edge Rate Seen at the Receiver is Slower than the Minimum Edge Rate 102
9-9 Traditional Method of Calculating Flight Time Assuming a Nonlinear Edge from V_{IL} Through V_{IH} at the Receiver 103
9-10 Traditional Method of Calculating Flight Time Assuming a Ringback Violation from V_{IL} Through V_{IH} at the Receiver 103
9-11 Example of Sweeps Used to Evaluate the Length Limits of Trace L2 and L3 105
9-12 Monte Carlo Analysis Should Be Performed on These Areas of the Phase 1 Solution Space 105
9-13 Results of Targeted Monte Carlo (TMC) Analysis and the Resultant Phase 2 Solution Space for Variables L2 and L3 106
10-1 Example of ISI Impact on Timing and Signal Integrity 108
10-2 Propagation on Aggressor Network 109
10-3 Aggressor and Victim Networks 109
10-4 Transmission Line Geometry of Microstrip and Stripline 110
10-5 Cross-Section of a 3-Conductor System Used to Create a SLEM Model 113

Tables

1-1 References 11
1-2 Platform Conventions and Terminology 13
2-1 Intel® Xeon™ Processor MP Feature Set Overview 18
2-2 Platform Bandwidth Summary 18
4-1 Placement Assumptions for Server Configurations 22
5-1 BCLK[1:0]# Routing Guidelines 27
6-1 System Bus Routing Summary for 4-Way Processor Configurations 31
6-2 System Bus Signals 36
6-3 Source Synchronous Signal Groups 38
6-4 Asynchronous GTL+ and Miscellaneous Signal List 43
8-1 Various Component Models Used at Intel (Not Vendor Specifications) 73
8-2 Processor Lump Model Component Values 75
8-3 “Row” Pattern with Voltage Regulator Module Schematic Values 77
8-4 “Square” Pattern with Voltage Regulator Module Schematic Values 80
8-5 “Row” Pattern with Voltage Regulator Module Schematic Values 82
8-6 Component Recommendation - Inductor 87
8-7 Component Recommendations - Capacitor 88
9-1 System Variables to Consider for Sensitivity Analysis 100
10-1 Example Backward Crosstalk Coupling Factors 111
11-1 Processor Connection Checklist 115

Revision History

Revision	Draft/Changes	Date
-001	Initial version	March 2002
-002	Added Intel® Xeon™ Processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process information.	November 2002

This page is intentionally left blank.

Introduction 1

This design guide documents Intel's design recommendations for systems based on the *Intel® Xeon™ processor MP* and *Intel® Xeon™ processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process* using non-Intel designed chipsets. In addition to providing motherboard design recommendations such as layout and routing guidelines, this document will also address possible system design issues such as processor power delivery, layout considerations for mechanical pieces, EMI design impacts and system bus decoupling. Design issues such as thermal considerations should be addressed using specific thermal documentation for the Intel Xeon processor MP.

Carefully follow the design information, board schematics, debug recommendations, and system checklist presented in this document. These design guidelines have been developed to ensure maximum flexibility for board designers while reducing the risk of board-related issues. The design information provided in this document falls into one of the two categories below.

Design Recommendations are items based on Intel's simulations and lab experience to date and are strongly recommended, if not necessary, to meet the timing and signal quality specifications.

Design Considerations are suggestions for platform design that provide one way to meet the design recommendations. They are based on the reference platforms designed by Intel. They should be used as an example, but may not be applicable to your particular design.

Note: The guidelines recommended in this document are based on Intel's experience developing Intel Xeon processor MP based systems. The recommendations are subject to change.

1.1 Related Documentation

Reference the following documents for more information. Contact your Intel representative to receive the latest revisions of collateral where the document number is not listed.

Table 1-1. References (Sheet 1 of 2)

Document	Document Number[5]
AP-485, Intel® Processor Identification and the CPUID Instruction	241618
IA-32 Intel Architecture Software Developer's Manual 1. *Volume 1: Basic Architecture* 2. *Volume 2: Instruction Set Reference* 3. *Volume 3: System Programming Guide (Preliminary)*	 245470 245471 245472
CK00 Clock Synthesizer/Driver Design Guidelines	249206
VRM 9.1 DC-DC Converter Design Guidelines	298646
ITP700 Debug Port Design Guide	249679
Intel® Xeon™ Processor MP Datasheet	290740
Intel® Xeon™ Processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process Datasheet	251931
Intel® Xeon™ Processor Thermal Design Guidelines	298348

Table 1-1. References (Sheet 2 of 2)

Intel® Xeon™ Processor Multiprocessor (MP) Thermal Design Guidelines	298650
603 Pin Socket Design Guidelines	249672
Intel® Xeon™ Processor Thermal Solution Functional Specification	249673
Intel® Xeon™ Processor Signal Integrity Model (IBIS format)	developer.intel.com
Intel® Xeon™ Processor Overshoot Checker Tool	developer.intel.com
Mechanical Drawings in Pro-E* Format for the Intel® Xeon™ Processor Enabled Solutions	developer.intel.com
Mechanical Drawings in IGES* Format for the Intel® Xeon™ Processor Enabled Solutions	developer.intel.com
Intel® Xeon™ Processor FloTherm* Thermal Models	developer.intel.com
Intel® Xeon™ Processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process Core Boundary Scan Descriptive Language (BSDL) Model	developer.intel.com
System Management Bus Specification, rev 1.1	www.sbs-forum.org/ specs
Wired for Management 2.0 Design Guide	developer.intel.com

1.2 Conventions and Terminology

This section defines conventions and terminology that will be used throughout this document.

Table 1-2. Platform Conventions and Terminology (Sheet 1 of 3)

Convention/ Terminology	Definition
4-way	Used to specify a system configuration using four processors.
Aggressor	A network that transmits a coupled signal to another network is called the aggressor network.
AGTL+	The processor system bus uses a bus technology called AGTL+, or Assisted Gunning Transceiver Logic. AGTL+ buffers are open-drain and require pull-up resistors that provide the high logic level and termination. AGTL+ output buffers differ from GTL+ buffers with the addition of an active pMOS pull-up transistor to “assist” the pull-up resistors during the first clock of a low-to-high voltage transition.
Bus Agent	A component or group of components that, when combined, represent a single load on the AGTL+ bus.
Corner	Describes how a component performs when all parameters that could impact performance are adjusted simultaneously to have the best or worst impact on performance. Examples of these parameters include variations in manufacturing process, operating temperature, and operating voltage. Performance of an electronic component may change as a result of (including, but not limited to): clock to output time, output driver edge rate, output drive current, and input drive current. The “slow” corner means a component is operating at its slowest, weakest drive strength performance. The “fast” corner means a component is operating at its fastest, strongest drive strength performance. Operation or simulation of a component at its slow corner and fast corner is expected to bound the extremes between slowest, weakest performance and fastest, strongest performance.
Crosstalk	The reception on a victim network of a signal imposed by aggressor network(s) through inductive and capacitive coupling between the networks. Backward Crosstalk – coupling that creates a signal in a victim network that travels in the opposite direction as the aggressor’s signal. Forward Crosstalk – coupling that creates a signal in a victim network that travels in the same direction as the aggressor’s signal. Even Mode Crosstalk – coupling from single or multiple aggressors when all the aggressors switch in the same direction that the victim is switching. Odd Mode Crosstalk – coupling from single or multiple aggressors when all the aggressors switch in the opposite direction that the victim is switching.

Table 1-2. Platform Conventions and Terminology (Sheet 2 of 3)

Convention/ Terminology	Definition
Flight Time	Flight time is a term in the timing equation that includes the signal propagation delay, any effects the system has on the T_{CO} of the driver, plus any adjustments to the signal at the receiver needed to guarantee the setup time of the receiver. More precisely, *flight time* is defined to be: Time difference between a signal at the input pin of a receiving agent crossing the switching voltage (adjusted to meet the receiver manufacturer's conditions required for AC timing specifications; e.g., ringback, etc.) and the output pin of the driving agent crossing the switching voltage when the driver is driving a test load used to specify the driver's AC timings. Maximum and Minimum Flight Time – Flight time variations can be caused by many different parameters. The more obvious causes include variation of the board dielectric constant, changes in load condition, crosstalk, power noise, variation in termination resistance and differences in I/O buffer performance as a function of temperature, voltage and manufacturing process. Some less obvious causes include effects of Simultaneous Switching Output (SSO) and packaging effects. Maximum flight time is the largest acceptable flight time a network will experience under all conditions. Minimum flight time is the smallest acceptable flight time a network will experience under all conditions.
GTL+	GTL+ is the bus technology used by the Intel® Pentium® Pro processor. This is an incident wave switching, open-drain bus with pull-up resistors that provide both the high logic level and termination. It is an enhancement to the GTL (Gunning Transceiver Logic) technology.
ISI	Inter-symbol interference is the effect of a previous signal (or transition) on the interconnect delay. For example, when a signal is transmitted down a line and the reflections due to the transition have not completely dissipated, the following data transition launched onto the bus is affected. ISI is dependent upon frequency, time delay of the line, and the reflection coefficient at the driver and receiver. ISI can impact both timing and signal integrity.
Manageability Features	Circuits incorporated into the processor that allow system administrators to monitor processor status and information including temperature, stepping, cache size, and more. They are accessed through the System Management Bus.
Network	The network is the trace of a Printed Circuit Board (PCB) that completes an electrical connection between two or more components.
Network Length	The distance between agent 0 pin and the agent pin at the far end of the bus.
Overshoot	Maximum voltage observed for a signal at the device pad. Measured with respect to V_{CC_CPU}.
Pad	The electrical contact point of a semiconductor die to the package substrate. A pad is only observable in simulation.
Pin	The contact point of a component package to the traces on a substrate, like the motherboard. Signal quality and timings can be measured at the pin.
Power-Good	*"Power-Good"* or "*PWRGOOD*" (an active high signal) indicates that all of the supplies and clocks within the system are stable. PWRGOOD should go active a predetermined time after system voltages are stable and should go inactive as soon as any of these voltages fail their specifications.
Ringback	The voltage that a signal rings back to after achieving its maximum absolute value. Ringback may be due to reflections, driver oscillations, or other transmission line phenomena.
Setup Window	The time between the beginning of Setup to Clock (T_{SU_MIN}) and the arrival of a valid clock edge. This window may be different for each type of bus agent in the system.

Table 1-2. Platform Conventions and Terminology (Sheet 3 of 3)

Convention/ Terminology	Definition
SSO	Simultaneous Switching Output (SSO) Effects refers to the difference in electrical timing parameters and degradation in signal quality caused by multiple signal outputs simultaneously switching voltage levels (e.g., high-to-low) in the opposite direction from a single signal (e.g., low-to-high) or in the same direction (e.g., high-to-low). These are respectively called odd-mode switching and even-mode switching. This simultaneous switching of multiple outputs creates higher current swings that may cause additional propagation delay (or “push-out”), or a decrease in propagation delay (or “pull-in”). These SSO effects may impact the setup and/or hold times and are not always taken into account by simulations. System timing budgets should include margin for SSO effects.
Stub	The branch from the bus trunk terminating at the pad of an agent.
System Bus	Used in this document to refer to the scalable system bus.
Test Load	Intel uses a 50 Ω test load for specifying its components.
Trunk	The main connection, excluding interconnect branches, from one end agent pad to the other end agent pad.
Undershoot	Minimum voltage observed for a signal to extend below V_{SS} at the device pad.
V_{CC_CPU}	V_{CC_CPU} is the processor power for both the core and the system bus I/O.
Victim	A network that receives a coupled crosstalk signal from another network is called the victim network.
GTLREF Guardband	A guardband defined above and below GTLREF to provide a more realistic model accounting for noise such as VTT and GTLREF noise.
VRM 9.1	*“VRM 9.1”* refers to the Voltage Regulator Module for the processor. It is a DC-DC converter module that supplies the required voltage and current to a single processor.

This page is intentionally left blank.

System Overview 2

2.1 The Intel® Xeon™ Processor MP and the Intel® Xeon™ Processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process

The Intel Xeon processor MP is the next generation IA-32 microprocessor for servers. The processor is based on Intel® NetBurst™ microarchitecture but maintains the tradition of complete compatibility with IA-32 software. The Intel Xeon processor MP, like its P6 predecessors, support Intel® MMX™ technology instructions for enhanced media and communication performance. In addition, the processor also supports Streaming SIMD Extensions, introduced in the Intel® Pentium® III Xeon™ processor.

The Intel Xeon processor MP is intended for high performance server systems with up to four processors on one bus. The processor is available for 1-4 way and greater than 4-way designs. Intel Xeon processors MP employ a 256-KB L2 cache, and are available with 512-KB or 1-MB of L3 cache. The processor includes manageability features and new IA-32 instructions. Components of the manageability features include an Scratch EEPROM and Processor Information ROM, which are accessed through a SMBus, interface and contain information relevant to the particular processor and system in which it is installed. In addition, enhancements have been made to the Machine Check Architecture.

The Intel Xeon processor MP with up to 2-MB L3 cache is based on 0.13-micron process technology, and is a follow-on to the Intel Xeon processor MP. The Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process processor, like the Intel Xeon processor MP, maintains IA-32 software compatibility and is based on the Intel NetBurst microarchitecture. The Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process supports 1-4 way and greater than 4-way designs. These processors are intended for high performance workstations and server systems utilizing up to four processors on one bus. All of these larger L3 cache processors employ a 512 KB L2 cache.The Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process processors will be available with either 1-MB or 2-MB integrated L3 cache. The processor will include the manageability features found on the Intel Xeon processor such as the Scratch EEPROM and Processor Information ROM (PIROM).

For server applications, the Streaming SIMD Extensions improves TCP/IP and memory/cache performance, benefits connected applications such as web hosting, security, and e-commerce, as well as traditional server applications such as database, enterprise resource planning, and media applications. Many popular operating systems, databases, and other applications are being tuned to take advantage of these new Streaming SIMD Extensions to deliver significant performance gains.

Intel Xeon processors MP and Intel Xeon processors MP with up to 2-MB L3 cache on the 0.13 micron process include Streaming SIMD Extensions 2 instructions. For server applications, these new instructions are expected to improve memory/cache performance, boost the number of authenticated transactions in e-commerce environments, as well as the traditional server applications like database, enterprise resource planning, and computer/telephony integration.

The Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process use a new system bus protocol referred to as the “system bus” in this document. This system bus utilizes a split-transaction, deferred reply protocol similar to that of the P6

processor family system bus, but is not compatible with the P6 processor family system bus. The system bus uses Source-Synchronous Transfer (SST) of address and data to improve performance. Whereas the P6 processor family data transfer is once per bus clock, the Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process transfer data four times per bus clock (4X data transfer rate, as in AGP 4X). Along with the 4X data bus, the address bus can deliver addresses two times per bus clock and is referred to as a "double-clocked" or 2X address bus. In addition, the Request Phase completes in one clock cycle. Working together, the 4X data bus and 2X address bus provide a data bus bandwidth of up to 3.2 GB per second. Finally, the system bus also introduces transactions that are used to deliver interrupts.

Table 2-1. Intel® Xeon™ Processor MP Feature Set Overview

Feature	Intel® Xeon™ Processor MP	Intel® Xeon™ Processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process
L1 Cache	On-die	On-die
L2 Cache	256 KB on-die	512-KB on-die
L3 Cache	512 KB or 1 MB on-die	1-MB or 2-MB on-die
Data Bus Frequency	400 MHz	400 MHz
Multi-processor Support	1-4 way	1-4 way
Manageability Features	PIROM, Scratch EEPROM, and thermal sensor on package	PIROM, Scratch EEPROM, and thermal sensor on package
Processor Core Voltage	1.7 V	1.5 V
Processor Socket	603-pin	603-pin
Processor Package	603-pin micro-PGA	603-pin micro-PGA

2.2 Bandwidth Summary

Table 2-2 documents the system bus bandwidth.

Table 2-2. Platform Bandwidth Summary

Interface	Clock Speed (MHz)	Samples per Clock	Data Width (Bytes)	Bandwidth (MB/s)
System bus (4-way)	100	4	8	3200

Processor Quadrant Layout 3

Figure 3-1 illustrates the quadrant layout of the Intel Xeon processor MP. Figure 3-2 illustrates the quadrant layout of the Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process. In the event that this layout conflicts with the respective *Intel Xeon Processor MP Datasheets*, the datasheets supersede.

The quadrant layout figures below do not show the exact component ball count, only the general quadrant information. Only the exact ball assignment should be used to conduct routing analysis. Refer to the processor datasheets for specific pin assignment information.

Figure 3-1. Top View - Intel® Xeon™ Processor MP Socket Quadrant Layout

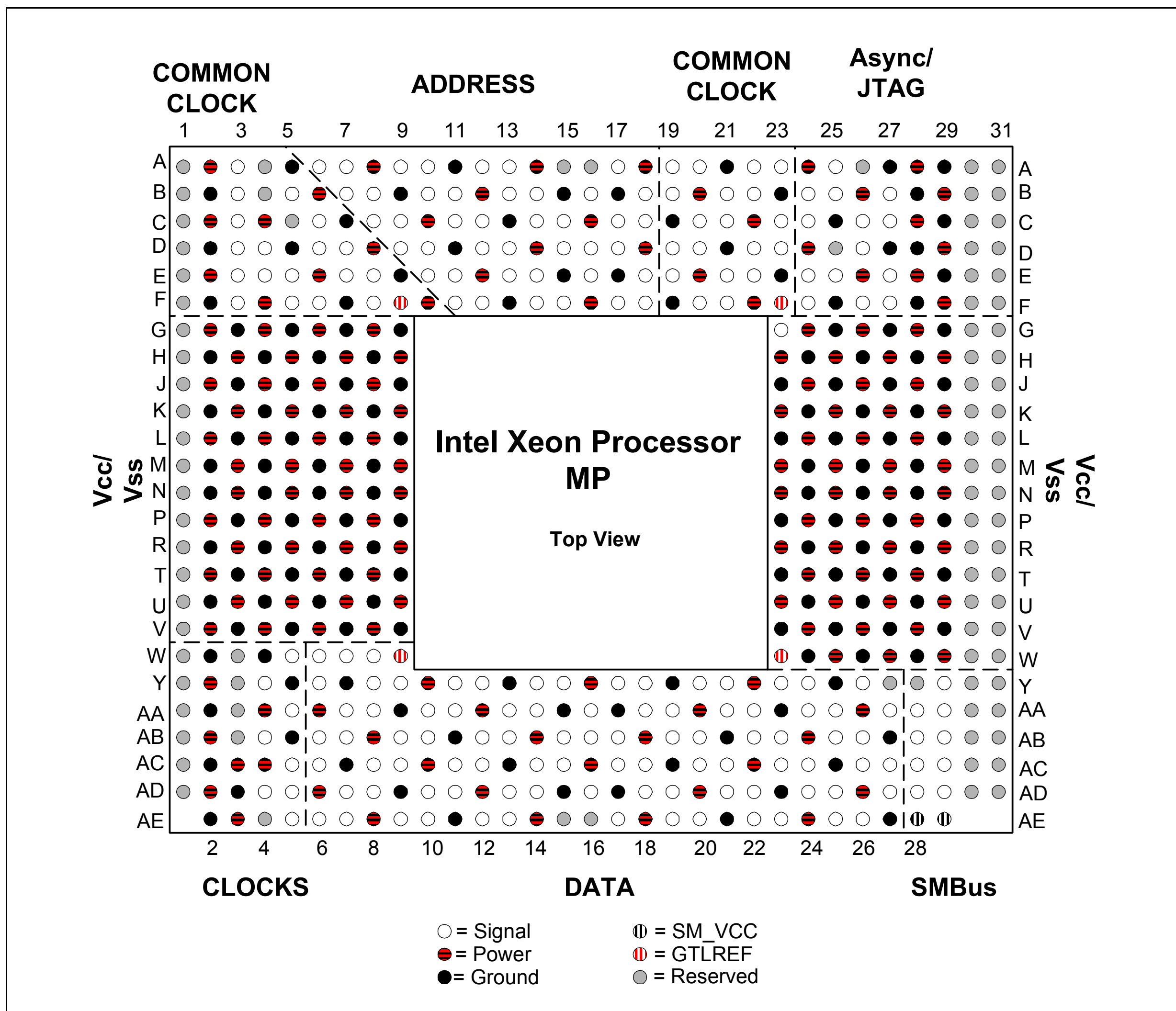
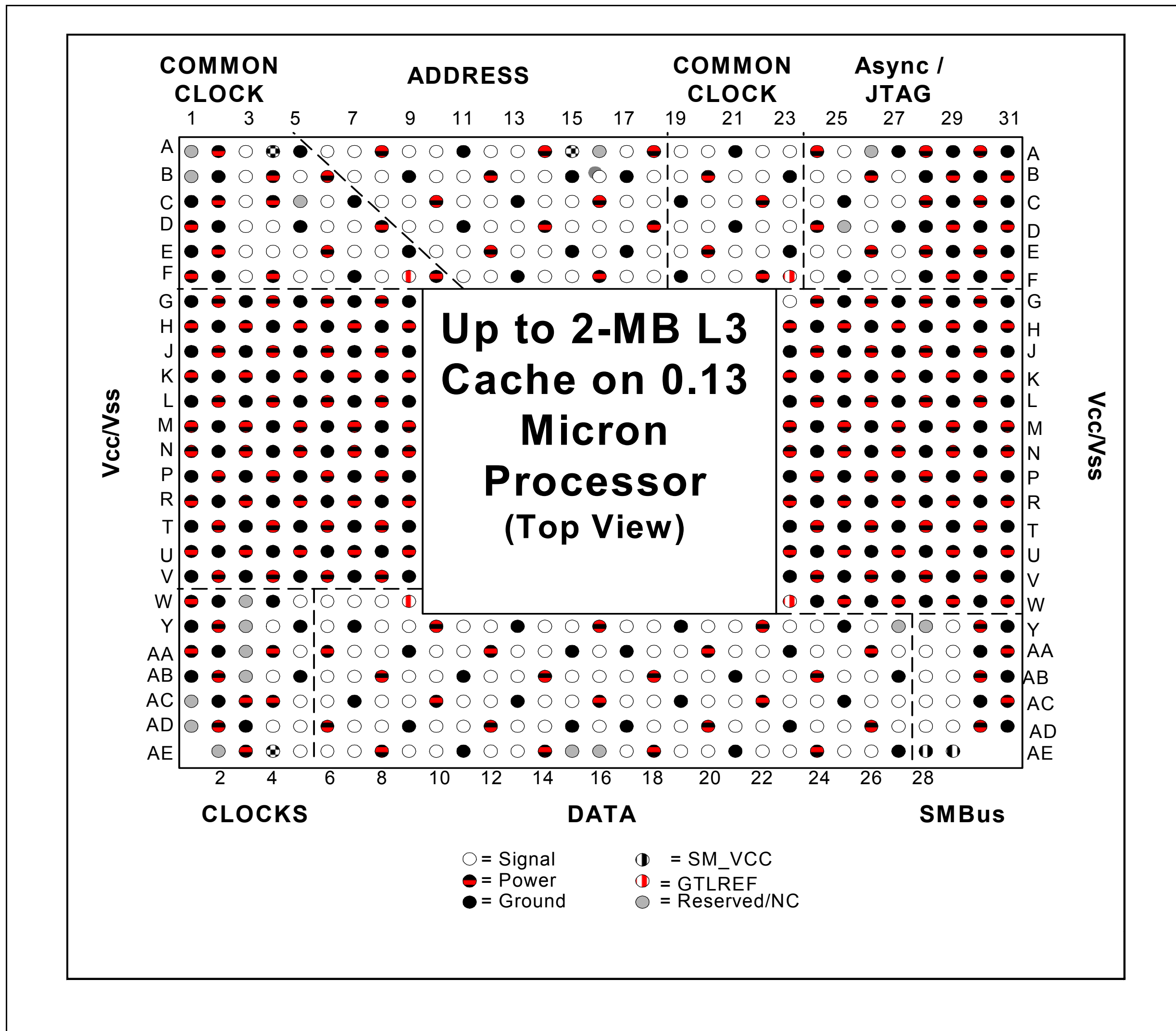


Figure 3-2. Top View - Intel® Xeon™ Processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process Socket Quadrant Layout



Platform Stack-Up and Placement Overview 4

4.1 Platform Component Placement

The figures below illustrate general component placement for server systems. The assumptions used for the component placement are documented in Table 4-1.

Figure 4-1. 4-Way Processor Server Component Placement Example in a Midrange SSI Form Factor

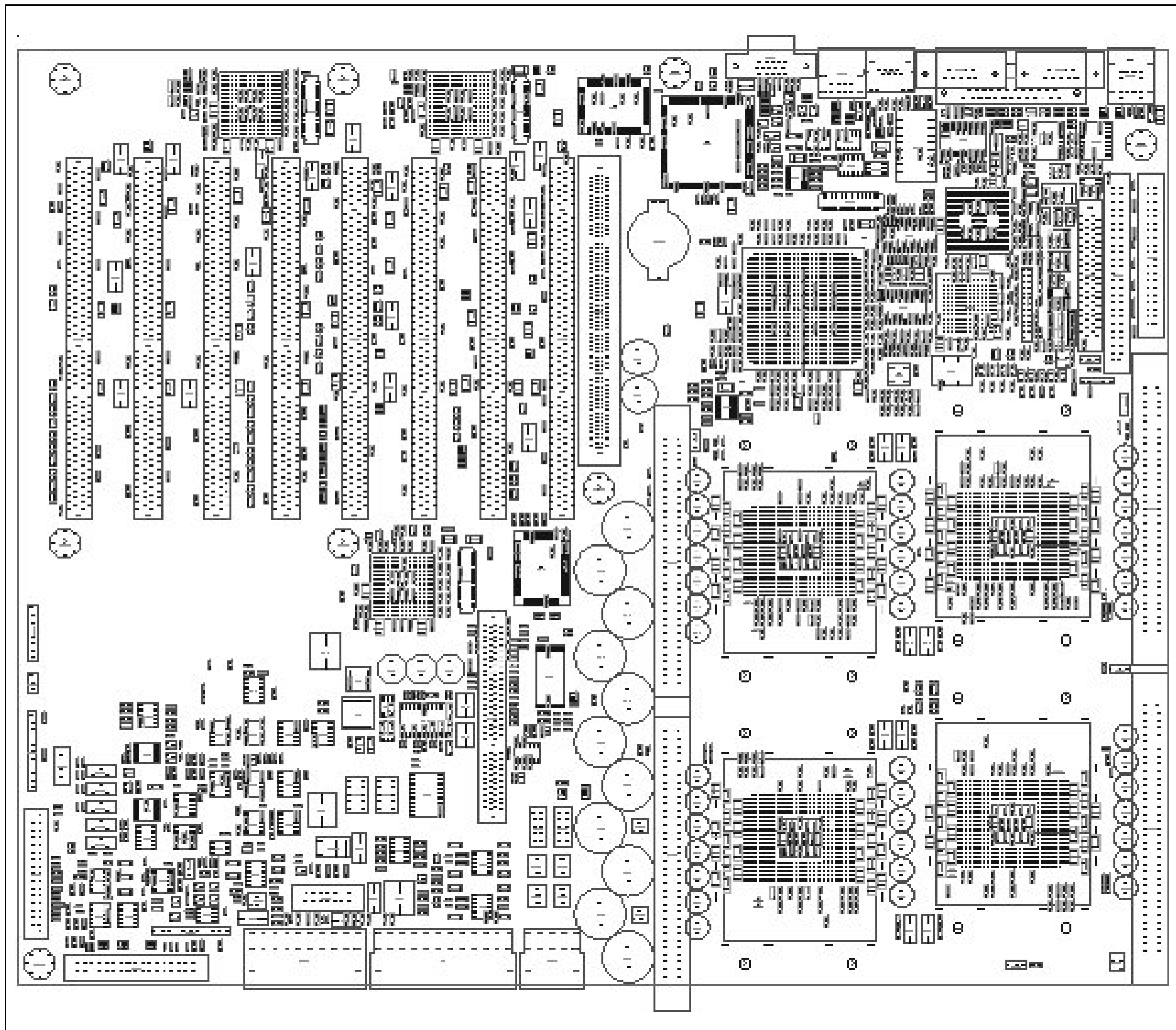


Table 4-1. Placement Assumptions for Server Configurations

System Configuration	Assumptions		
	Form Factor	Number of Total PCB Layers	Assembly
Server (4-way)	Midrange SSI	12 Layers	Double Sided

4.2 4-Way System Stack-Up

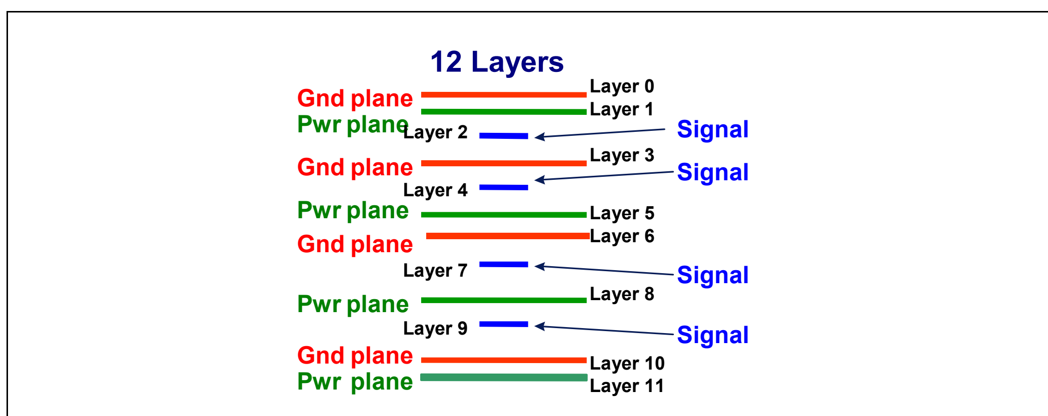
Design recommendations will be presented first followed by design considerations.

4.2.1 Design Recommendations

Figure 4-2 shows the recommended Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process 4-way system stack-up. The processor system bus must be routed in a symmetrical stripline stack-up. This provides clean and equal return paths through V_{SS} and V_{CC_CPU} from the I/O cell of one agent to the next.

Because the surface mount pads of the 603-pin socket create holes in the primary side layer (V_{CC_CPU} plane), these areas do not provide an adequate reference plane. In order to have solid reference planes the top and bottom layers (layers 0 and 11) of the stack-up were added. These layers help to ensure a good return path and minimize crosstalk for layers 2 and 9. Additionally, it may be desirable to design the top and bottom layers to have different trace impedance to allow the routing of other system signals on these layers. The top and bottom layers can also be used to deliver power to the processor, as it is critical to keep a very low inductance for the power path. Please refer to Section 8.8.2 for more information.

Figure 4-2. Twelve Layer Stack-Up for a 4-Way System



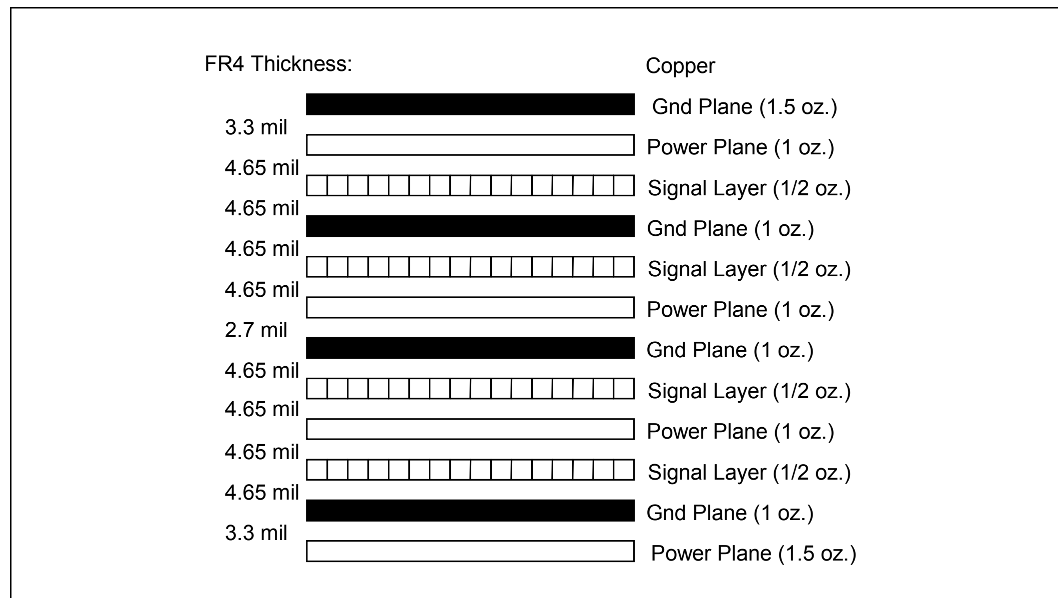
4.2.2 Design Considerations

The following design considerations are based on Intel's intentions for validation systems. These validation systems are targeted to provide a high quality platform with optimized signal integrity, timing margins, and power distribution. They therefore represent Intel's recommended platform design for the Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process. However, excursions from these guidelines can be made to optimize for cost or system-specific designs without violating the specifications of either the processor or chipset. In any design it is up to the designers to ensure that the platform meets all the component specifications. Intel strongly recommends that a comprehensive simulation analysis be performed to ensure all such specifications will be met. This would be particularly important if a design deviated from the following design considerations.

Use the following items and the stack-up in Figure 4-3 as the design considerations for the 4-way system stack-up.

- ½ oz. copper in middle routing layers
- 1 oz. copper for power and ground planes.
- There must be the equivalent total of 2 oz. of copper on both the power and ground planes for power delivery to the processor.
- Vias are 10-mil hole with a 25 mil pad and a 35 mil anti-pad
- Total board thickness is 0.062 inches.
- No less than 4-mil wide traces are recommended to reduce resistive loss in the signal propagation.

Figure 4-3. 4-Way Stack-Up Example



This page is intentionally left blank.

Clock Routing Guidelines 5

5.1 System Bus Clocking Guidelines

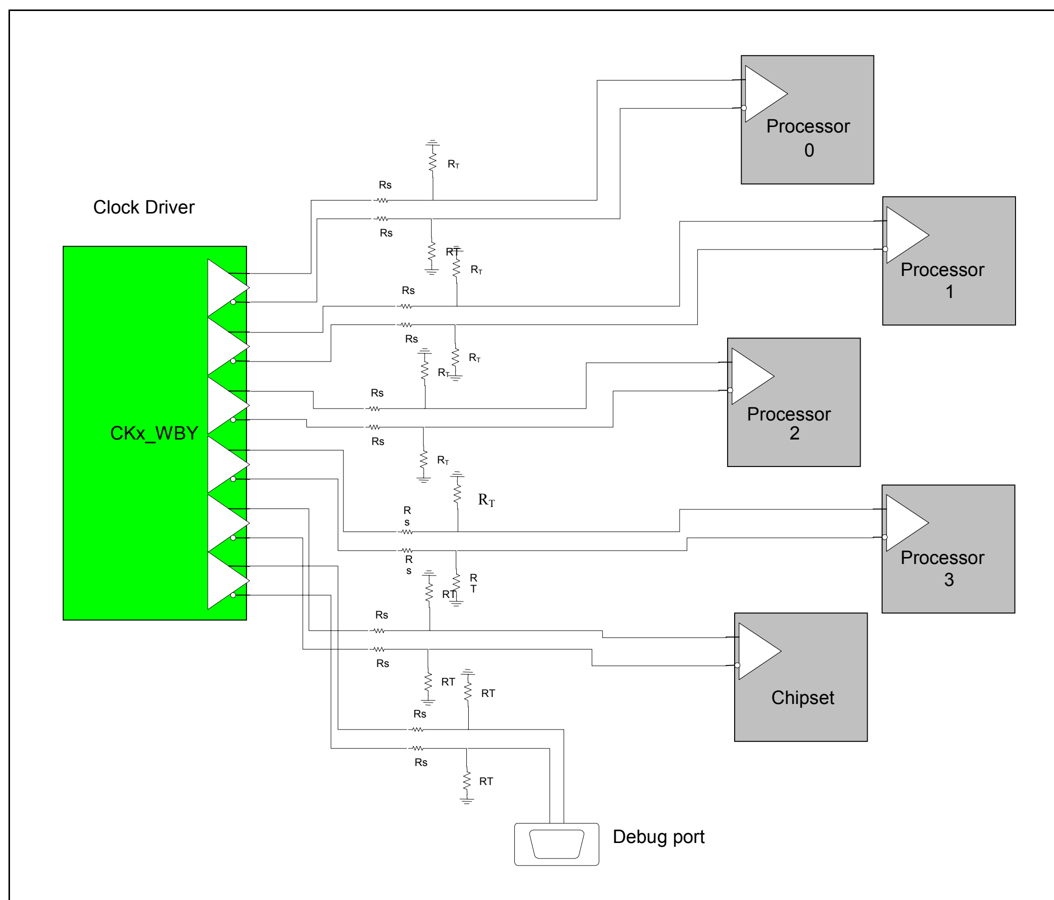
5.1.1 Routing Guidelines for BCLK[1:0]

Note: For designs using non-Intel chipsets please contact the corresponding chipset vendor for specific information regarding clock driver and baseboard design requirements.

To minimize jitter and improve routing, the 4-way Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process system may use a dual-chip clock solution. The first component (referred to as CKx_WBY) is shown in Figure 5-1. In this configuration the main clock generator provides six, 100 MHz differential outputs for all of the system bus agents, one 14.31 MHz output, and one, 66 MHz speed clock that drives a second system clock chip referred to as CKFF, which provides additional system clocks such as PCI. Figure 5-1 shows the implementation of the system bus clocks for this configuration. Refer to the CK00 Clock Synthesizer/Driver Design Guidelines for complete information about CKx_WBY and CKFF.

When connecting CK00 HOST clock and HOST_BAR clock to host agent BCLK0 and BCLK1 be consistent across all host agents. In other words, connect HOST pin to BCLK0 and HOST_BAR pin to BCLK1 for each system bus agent.

Figure 5-1. 4-Way Processor BCLK Topology

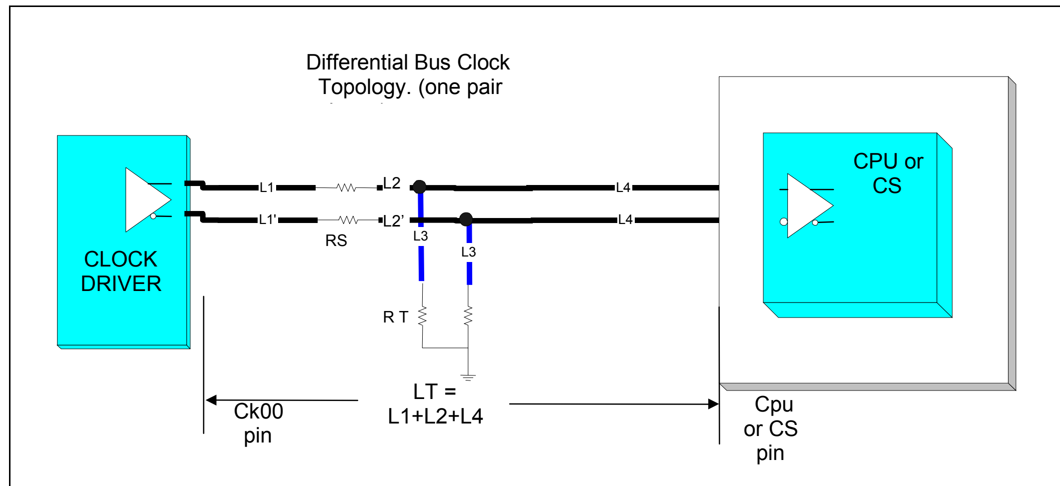


The CK00 clock driver differential bus output structure is a “Current Mode Current Steering” output which develops a clock signal by alternately steering a programmable constant current to the external termination resistors R_T. The resulting amplitude is determined by multiplying IOUT by R_T. The current I_{OUT} is programmable by a resistor and an internal multiplication factor so the amplitude of the clock signal can be adjusted for different values of R_T for impedance matching or to accommodate future load requirements. Refer to the *CK00 Clock Synthesizer/Driver Design Guidelines* for more detailed information.

The recommended termination for the CK00 differential bus clock is a “Shunt Source termination”. Refer to Figure 5-2. Parallel resistors R_T perform a dual function, one converting the current output of the CK00 to a voltage and two matching the driver output impedance to the transmission line. The series resistors R_S provide Isolation from the clock driver's output parasitics, which would otherwise appear in parallel with the termination resistor R_T.

The value of R_T should be selected to match the characteristic impedance of the motherboard and R_S should be between 20 Ω and 33 Ω. Simulations have shown that R_S values above 33 Ω provide no benefit to signal integrity, but do degrade the edge rate.

Figure 5-2. Source Shunt Termination



The goal of constraining all bus clocks to one physical routing layer is to minimize the impact on skew due to variations in Er and the impedance variations due to physical tolerances of circuit board material. Routing on internal layers provides the least amount of Er and impedance variation. If a layer change is necessary, increase the amounts of interconnect skew used in the timing equations for the component receiving that clock.

- Requirement: Do not split up the two halves of a differential clock pair between layers.
- Goal: Route clocks to all agents on same physical routing layer.

If layer transitions are required:

- Make sure that skew induced by the vias used to transition between routing layers is compensated in the traces to other agents.
- Layer transitions should only be made between routing layers of the same configuration, i.e., stripline layer to stripline layer.
- Keep routes to all agents as short as possible to minimize the cumulative effects of Er variations on clock skew.
- Maintain return path referenced to V_{SS} for the pair.
- Match delays (flight time or length) from pad of CK00 to pad of processor input.
- If a layer change must occur, use vias connecting the V_{CC_CPU} planes and/or V_{SS} planes to provide a low impedance path for the return current. Vias should be as close as possible to the signal via.

Table 5-1. BCLK[1:0]# Routing Guidelines (Sheet 1 of 2)

Layout Guideline	Value	Figures	Notes
BCLK Skew assumptions (between agents) 4-way configuration CK_WBY	200 pS total 100 pS for Clock driver 100 pS for interconnect	Figure 5-3	1, 2, 3
Differential pair spacing	4 x W min. to 5 x W max.	Figure 5-4	4, 5
Spacing to other traces	5 x W min.	Figure 5-4	

Table 5-1. BCLK[1:0]# Routing Guidelines (Sheet 2 of 2)

Serpentine spacing	Maintain a minimum S/h ratio of > 5/1 Keep parallel serpentine sections as short as possible Minimize 90 degree bends. Make 45 degree bends if possible.	Figure 5-4	
Motherboard Impedance – Differential	100 Ω typical		6
Motherboard Impedance – single ended	50 Ω ± 10%		7
Processor routing length – L1: CK_WBY/CK_SKS to R_S	0.5” max	Figure 5-2	12
Processor routing length – L4: R_S/R_T to Processor	0 – 12”	Figure 5-2	
Chipset routing length – L1: CK_WBY/CK_SKS to R_S	0.5” max	Figure 5-2	12
Chipset routing length – L4: R_S/R_T to Processor	0 – 12”	Figure 5-2	
L3: Stub length to R_T	Preferably without stubs 200 mils max from fork to R_T pad.	Figure 5-2	12
L2: R_S to fork to R_T	200 mils max from R_S to fork to R_T	Figure 5-2	12
Processor to CS length matching (LT)	0.13” ± 0.010” Chipset LT must be 0.13” longer than Proc. LT.	Figure 5-2	8
Processor to Processor length matching (LT)	± 10 mils	Figure 5-2	9
BCLK0 – BCLK1 length matching	±10 mils		
R_S Series termination value	33 Ω ± 5%	Figure 5-2	10
R_T Shunt termination value	49.9 Ω ±1%	Figure 5-2	11

NOTES:
1. The skew budget includes clock driver output pair to output pair jitter (differential jitter), and skew, clock skew due to interconnect process variation, and static skew due to layout differences between clocks to all bus agents. This number does not include clock driver common mode (cycle to cycle) jitter or spread spectrum clocking.
2. This number assumes all BCLK pairs are routed on the same signal layer.
3. Skew measured at the load between any two bus agents from pad to pad. Measured at the crossing point.
4. Edge to edge spacing between the two traces of any differential pair. Uniform spacing should be maintained along the entire length of the trace.
5. Clock traces are routed in a differential configuration. Maintain the minimum recommended spacing between the two traces of the pair. Do not exceed the maximum trace spacing, as this will degrade the noise rejection of the network.
6. The differential impedance of each clock pair is approximately 2*Z single-ended*(1-2*Kb) where Kb is the backwards crosstalk coefficient. For the recommended trace spacing, the Kb is very small and the effective differential impedance is approximately equal to 2 times the single-ended impedance of each half of the pair.
7. The single ended impedance of both halves of a differential pair should be targeted to be of equal value. They should have the same physical construction. If the BCLK traces vary within the tolerances specified, both traces of a differential pair must vary equally.
8. Length compensation for the processor socket and package delay is added to chipset routing to match electrical lengths between the chipset and the processor at die pad. Therefore, the motherboard trace length for the chipset will be longer than that for the processor. Refer to the processor signal integrity models for exact package compensation lengths.
9. Length of LT for one processor must match the LT of all other BCLK traces to other processors with specified tolerance.

10. R_S values between 20 Ω – 33 Ω have shown to be effective. The value specified is the recommended value.
11. R_T values should match the motherboard trace impedance for BCLK.
12. Minimize the trace lengths from the clock driver pin to R_S, from R_S to R_T stub and the length of the R_T stub.

Figure 5-3. Agent-to-Agent BCLK Skew

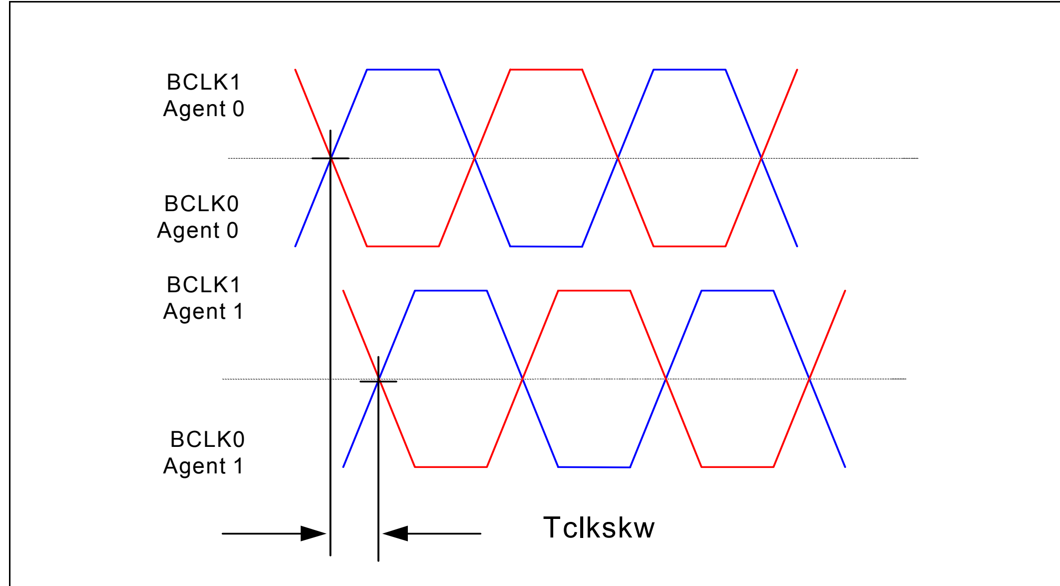
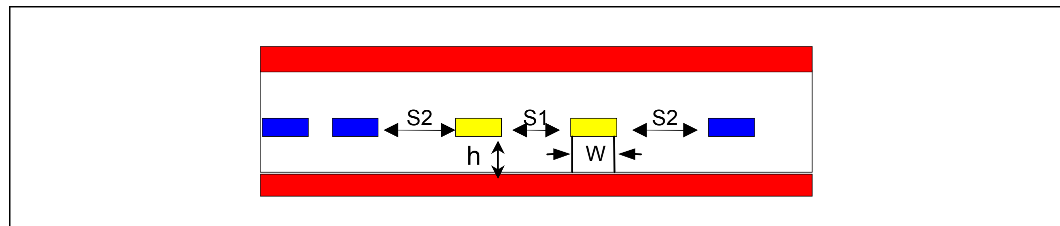


Figure 5-4. Dielectric Height to Trace Width Diagram



This page is intentionally left blank.

System Bus Routing 6

Table 6-1 summarizes the layout recommendations for 4-way processor-based configurations. It should be used for quick reference only. The following sections provide more detailed information about the different system configurations.

Table 6-1. System Bus Routing Summary for 4-Way Processor Configurations (Sheet 1 of 2)

Parameter	4-Way: Intel® Xeon™ Processor MP and Intel® Xeon™ Processor MP with up to 2-MB L3 Cache on the 0.13 Micron Process
Line to line spacing	Greater than 3:1 edge-to-edge spacing vs. trace to reference plane height ratio
4X Signal Group line lengths (agent-to-agent spacing)	3.0” – 6.1” pin-to-pin. **The maximum distance between one and only one set of agents may be up to 6.9 inches, e.g., from processor 2 and processor 3 at the “U-turn” route.** Total bus length must not exceed 20.8” Length must be added to the motherboard trace between agents to compensate for the stub created by the processor package. See Section 6.4.1.1 for details.
DSTBn/p[3:0]# line lengths	DSTB# signals should follow the same routing rules as the Data signals. In addition: A 25-mil spacing should be maintained around each strobe signal (between DSTBp# and DSTBn#, and any other signal.)
2X Signal Group line lengths	Address signals should follow the same routing rules as the Data signals.
ADSTB[1:0]# line lengths	ADSTB# signals should follow the same routing rules at the DSTB# signals.
Common Clock signal line lengths	Common Clock signals should follow the same routing rules at the Data signals, however no length compensation is necessary.
Topology	Daisy chain with chipset at one end of the system bus and Processor 0 at the other (U-turn may exist between Processor 2 and Processor 3.) End processor (Processor 0) must have on-die termination enabled.
Routing Requirements	No motherboard contribution to stub length of middle processors (35 mil max trace via to pad.) Strobes and associated signals must be routed on same layer for entire length of bus.

Table 6-1. System Bus Routing Summary for 4-Way Processor Configurations (Sheet 2 of 2)

Parameter	4-Way: Intel® Xeon™ Processor MP
Reference plane requirements	Signals should be routed in a symmetric stripline configuration. Avoid changing layers when routing system bus signals. If a layer change must occur, use vias connecting the V_{CC} planes and/or V_{SS} planes to provide a low impedance path for the return current. Vias should be as close as possible to the signal via.
Serpentine spacing	S/H ratio greater than or equal to 5 (Figure 6-1) Keep parallel sections as short as possible Minimize 90-degree bends, use 45-degree bends whenever possible
Motherboard Impedance	47 Ω–50 Ω ± 10%

6.1 Return Path

The return path is the route current takes to return to its source. It may take a path through ground planes, power planes, other signals, or integrated circuits. Determination of the return path is based on electro-magnetic field effects. It is useful to think of the return path following a path of least resistance nearest to the signal conductor. Discontinuities in the return path often have signal integrity and timing effects that are similar to the discontinuities in the signal conductor. Therefore, the return paths need to be given similar considerations. A simple way to evaluate return path parasitic inductance is to draw a loop that traces the current from the driver through the signal conductor to the receiver, and then back through the ground/power plane to the driver again. The smaller the area of the loop, the lower the parasitic inductance will be.

If via densities are large and most of the signals switch at the same time (as would be the case when a whole data group switches layers), the layer to layer bypass fails to provide an acceptably short signal return path to maintain timing and noise margins. Experience at Intel indicates that the magnitude of the uncertainty that occurs with shifting return paths is on the same order as the data bus cycle time. Since the signals are routed using symmetric stripline, return current is present on both the V_{CC} and V_{SS} planes. If a layer change must occur, then V_{CC} and V_{SS} vias must be placed as close to the signal via as possible to provide the shortest possible path for the return current.

The following sets of return path rules apply to all designs:

- Always trace out the return current path and provide as much care to the return path as the path of the signal conductor.
- Do not allow splits in the reference planes in the path of the return current.
- Do not allow routing of signals on the reference planes near system bus signals.
- Do not make signal layer changes that force the return path to make a reference plane change even if it is from one V_{SS} layer to another V_{SS} layer.
- Decoupling capacitors do not adequately compensate for a plane split.
- Do not route over via anti-pads or socket anti-pads

If reference plane changes must be made:

- Change from a V_{SS} reference plane to a V_{SS} reference plane and place a via connecting the two planes as close as possible to the signal via. This also applies when making a reference plane change from on V_{CC} plane to another V_{CC} plane.

- For symmetric stripline, return path vias for both V_{SS} and V_{CC} must be provided.
- Do not switch reference from V_{CC} to V_{SS} or vice versa.

6.2 Serpentine Routing

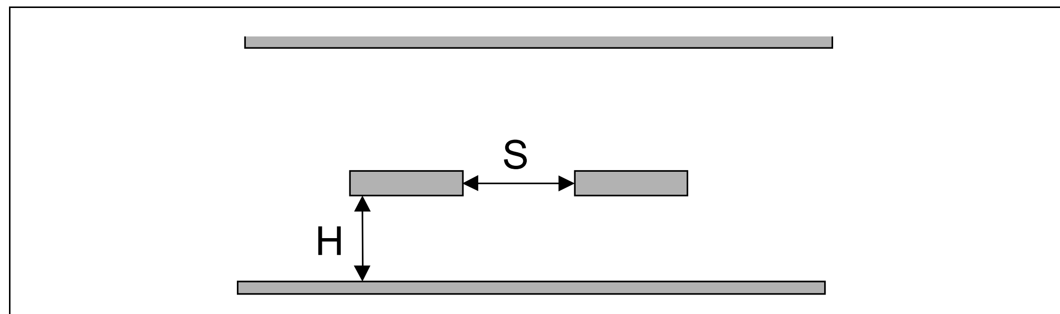
A serpentine net is a transmission line that is routed in such a manner so that sections of the net double back and couple to another segment of the same net.

A serpentining transmission line is sometimes necessary to properly match lengths between nets. It is important to properly control the serpentine in order to avoid signal integrity and timing problems. The primary impact of a serpentine trace is an observed decrease in the flight time when compared to a straight trace of equal length. This decrease in the flight time is a result of the crosstalk between parallel sections of the serpentine net. As the signal travels down the transmission line, a component of the signal will follow the transmission line and behave as though it were a straight line with no serpentine. However, another portion of the energy will propagate perpendicular to the parallel routed portions of the serpentine net via the mutual capacitance and mutual inductance. This creates an extra mode that will arrive at the receiver significantly earlier than the other component of the signal. If the coupling between parallel sections is high, this will cause significant timing skew when attempting to match trace lengths on a bus. Furthermore, if the coupling is very high, significant signal integrity problems can result.

The serpentine guidelines included in this document were based on HSPICE simulations with different spacing between parallel sections. The guidelines were chosen to significantly limit the effect of serpentine nets.

Serpentine spacing S/H ratio should be greater than or equal to 5. The S/H ratio is shown in Figure 6-1.

Figure 6-1. Serpentine Spacing - Diagram of Spacing to Reference Plane Height Ratio



6.3 System Bus Decoupling Requirements

This section contains the motherboard decoupling recommendations to minimize return path discontinuities and provide necessary power delivery for the bus I/O buffers. These are decoupling requirements for the system bus I/O only. This decoupling is not adequate for power delivery. For decoupling requirements for the processor core power, please refer to Section 8.9.

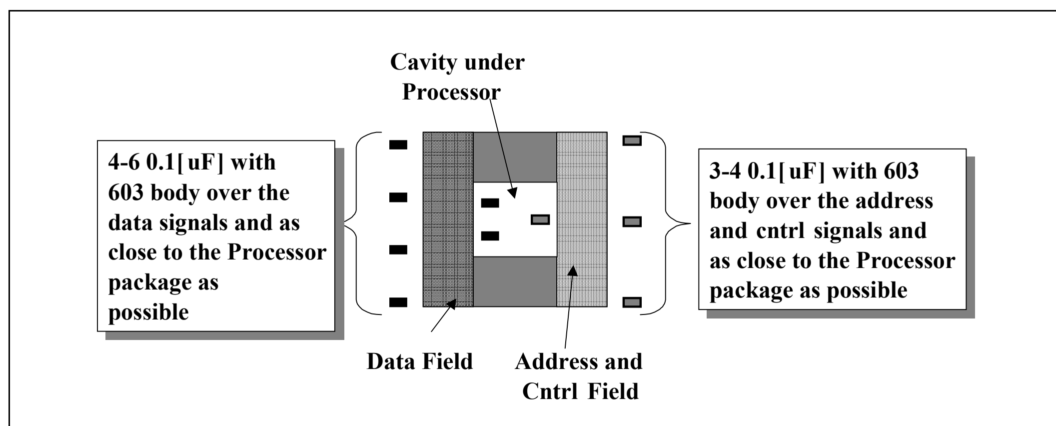
6.3.1 Processor I/O Decoupling Requirements

The primary objective of the processor I/O decoupling guidelines is to minimize the impact of return path discontinuities. The processor power delivery guidelines help insure the processor I/O has adequate power decoupling. The worst-case return path discontinuity anticipated is for systems that use microstrip structures on the motherboard. The processors, from die to package pin, follow a symmetric stripline configuration with V_{CC_CPU} as one reference plane and V_{SS} as the other reference plane. If the motherboard uses symmetric stripline with V_{CC_CPU} and V_{SS} references, then a discontinuity does not exist and additional decoupling is not necessary. If the motherboard routing references only one reference plane (V_{CC_CPU} or V_{SS}), then a return path discontinuity exists between the processor and the motherboard and decoupling capacitors are required.

The decoupling recommendations for each processor is (shown in Figure 6-2):

- Four minimum, six preferred 0.1 µF capacitors with each processor, distributed evenly over the system bus data signals
- Three minimum, four preferred 0.1 µF capacitors with each processor distributed evenly over the system bus address and Common Clock signals
- All capacitors placed as close to the processor as keep-out zones allow

Figure 6-2. System Bus I/O Decoupling Guidelines for the Processor



6.3.2 Chipset System Bus I/O Decoupling Recommendations

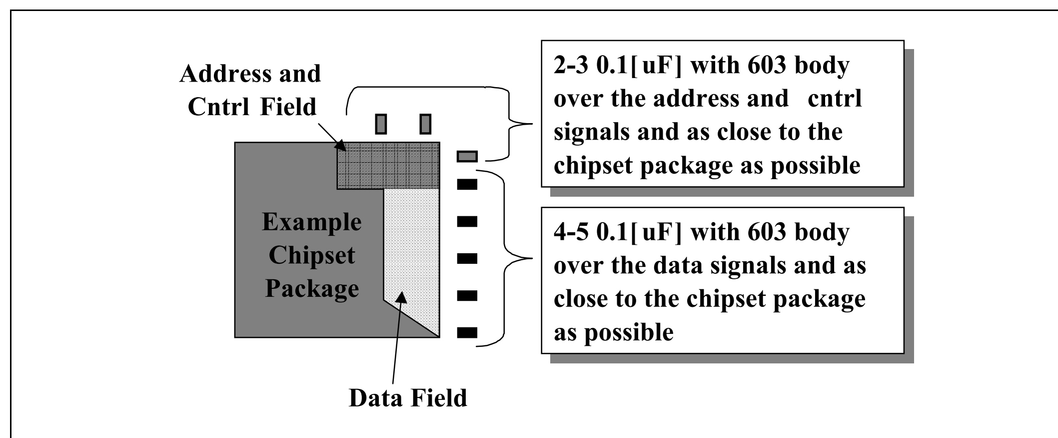
The primary objective of the I/O decoupling recommendations for the chipset is to provide clean power delivery to the system bus I/O buffers. The split power-plane nature of chipsets creates this power delivery concern. A noisy or starved power supply will negatively impact the signal quality and/or drive strength seen on the system bus signals.

The secondary objective of decoupling at the chipset is to minimize the impact of return path discontinuities that may occur between the chipset package and the motherboard. A return path discontinuity occurs in systems whose signals reference either power or ground, but not both. The chipset uses symmetric stripline interconnects that reference the signal to both V_{CC_CPU} and V_{SS}. Systems that have this type of discontinuity should use the larger number of decoupling capacitors listed in the below guidelines for the chipset.

The decoupling recommendations for the chipset are (shown in Figure 6-3):

- Four minimum, five preferred 0.1 µF capacitors with 603 packages distributed evenly over the system bus data lines
- Two minimum, three preferred 0.1 µF capacitors with 603 packages distributed evenly over the system bus address and common clock lines
- All capacitors placed as close to the chipset package as the chipset keep-out zone allows. Double-sided platforms should place the caps on the backside close to the V_{CC_CPU} and V_{SS} pins of the system bus.

Figure 6-3. System Bus I/O Decoupling Guidelines for the Chipset



6.4 Routing Guidelines for a 4-Way System

This section provides more details for 4-way processor systems. This information is subject to change. Both recommendations and considerations are presented.

For proper operation of the Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process system bus it is necessary to meet the timing and voltage specifications of each component on the bus. The most accurate way to understand the signal integrity and timing of the system bus on the platform is by performing a comprehensive simulation analysis. It is conceivable that adjustments to trace impedance, line length, termination impedance, board stack-up and other parameters can be made that improve system performance.

The following recommendations are Intel's best guidelines based on extensive simulation and experimentation based on our reference platform. It is therefore strongly recommended to perform a simulation analysis based on your platform.

Table 6-2 presents all signals interfacing with the processors. This table is included for reference purposes only. Refer to the processor datasheet for current signal interfacing details.

Table 6-2. System Bus Signals

Signal Group	Type	Signals
AGTL+ Common Clock Input	Synchronous to BCLK	BPRI#, BR [3:1]#[1], DEFER#, RESET#[1], RS [2:0]#, RSP#, TRDY#
AGTL+ Common Clock I/O	Synchronous to BCLK	ADS#, AP[1:0]#, BINIT#[2], BNR#[2], BPM[5:0]# [1], BR0#[1], DBSY#, DP[3:0]#, DRDY#, HIT#[2], HITM#[2], LOCK#, MCERR#[2]
AGTL+ Source Synchronous I/O: 4X Data Group	Synchronous to assoc. strobe	D[63:0]#, DBI[3:0]#
AGTL+ Source Synchronous I/O: 2X Address Group	Synchronous to assoc. strobe	A[35:3]#[3], REQ[4:0]#
AGTL+ Strobes	Synchronous to BCLK[1:0]	ADSTB [1:0]#, DSTBN [3:0]#, DSTBP [3:0]#
Asynchronous GTL+ Input[1]	Asynchronous	A20M#, IGNNE#, INIT#[3], lint0/intr, lint1/ nmi, PWRGOOD, SMI#[3], SLP#, STPCLK#
Asynchronous GTL+ Output[1]	Asynchronous	FERR#, IERR#, THERMTRIP#, PROCHOT#
System Bus Clock	Clock	BCLK0, BCLK1
TAP Input[1]	Synchronous to TCK	tck, tdi, tms, trst#
TAP Output[1]	Synchronous to TCK	TDO
SMBus Interface[1]	Synchronous to SM_CLK	SM_EP_A[2:0], SM_TS_A[1:0], SM_DAT, SM_CLK, SM_ALERT#, SM_WP
Power/Other	Power/Other	GTLREF[3:0], COMP[1:0], OTDEN, RESERVED, SKTOCC#, TESTHI[6:0], VID[4:0], V_{CC}, SM_VCC[4], V_{CCA}, V_{SSA}, $V_{CCIOPLL}$, V_{SS}, $V_{CCSENSE}$, $V_{SSSENSE}$

NOTES:
1. These signals do not have on-die termination on the processor. They need to be terminated properly on the motherboard. If they are not connected they will need to be pulled to the appropriate voltage level through a 1 kΩ resistor.
2. These signals are "wired-OR" signals and may be driven simultaneously by multiple agents. For further details on how to implement wired-OR signals refer to the routing guidelines in Section 6.4.1.5.
3. The value of these pins during the active to inactive edge of RESET# determine processor configuration options.
4. SM_VCC is required for correct operation of the Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process processor VID logic.

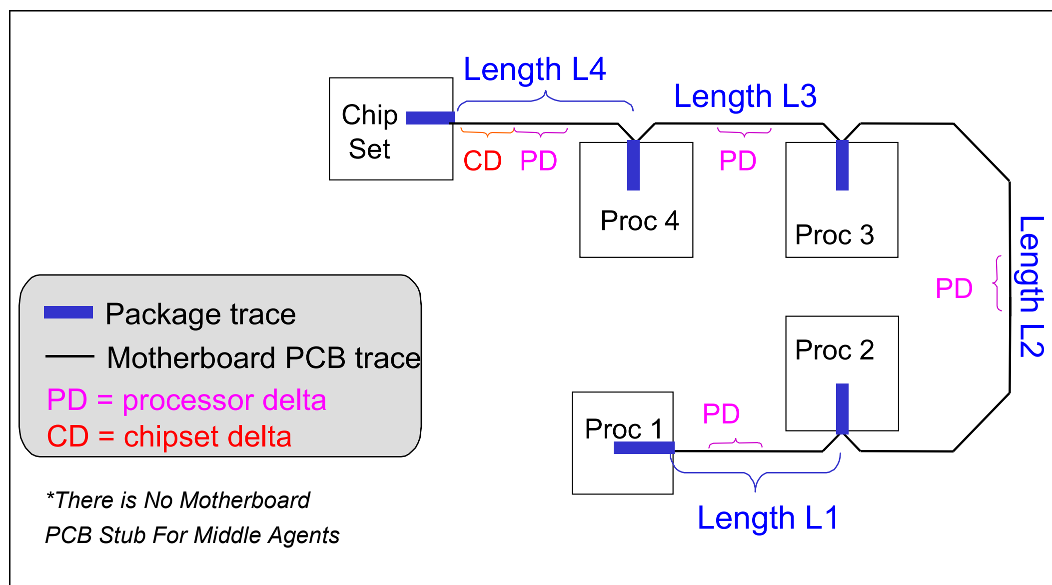
6.4.1 Topology and Routing

Design recommendations will be presented first followed by design considerations. The layout guidelines given in this section are based on specific chipset (I/O buffer, package, and loading) and motherboard properties. Complete simulation and hardware validation is necessary to ensure a robust design.

6.4.1.1 Design Recommendations

Below are the design recommendations for the data, address, strobes, and common clock signals.

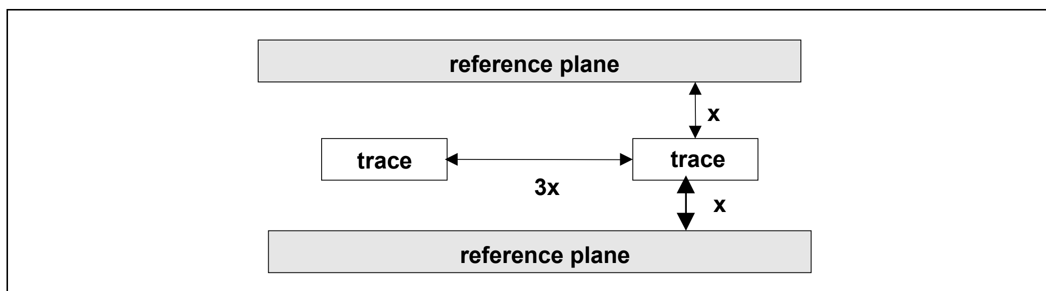
Figure 6-4. 4-Way Processor System Bus Topology



The 4-way processor topology requires that the chipset be at one end of the bus and that no motherboard contribution to the stub length of the processors in the middle of the bus exists. Figure 6-4 shows a schematic of an Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process quad processor daisy chain topology with the chipset at the end. A U-turn may exist between processor 2 and processor 3.

The motherboard trace impedance should be between 47 Ω–50 Ω ± 10%. The traces should maintain a greater than three to one "edge-to-edge spacing" versus "trace to reference plane height" ratio (see Figure 6-5). As the traces pass through the pin fields, the 3:1 requirement may not be achievable. In these areas where the 3:1 ratio is not possible, the separation should be maximized and the distance of the violation should be minimized. Specifically, when routing through the 603-pin socket expand to a 3:1 ratio whenever possible. Do not keep a tighter spacing ratio the entire length of the socket. However, do not route through the V_{CC_CPU} and V_{SS_CPU} pin field as this also has a great potential for noise coupling. Trace spacing to height ratio of 3 to 1 above the reference plane ensures a low crosstalk coefficient. All the effects of crosstalk are difficult and tedious to simulate. Intel has performed extensive simulation and experimentation on the effects of crosstalk to more accurately predict these effects. The timing and layout guidelines for processor have been created with the assumption of 3:1 trace spacing to height above reference plane ratio. A smaller ratio would have a negative impact on both timing and noise margins due to crosstalk.

Figure 6-5. Cross-Sectional View of 3:1 Ratio for Symmetric Stripline (Edge-to-Edge Trace Spacing vs. Trace to Reference Plane Height)



For partially populated systems, the end processor must be populated first. The end processor is that furthest from the chipset. This effectively leaves only the socket as a stub on the bus for the unpopulated agents. Also, the on-die termination must be enabled on the end processor. To enable the on-die termination, connect the ODTEN pin to V_{CC_CPU} through a resistor. On middle agents, disable the ODT by connecting ODTEN to V_{SS} through a resistor.

It is critical that additional stub length is not added on the motherboard to the middle processors.

Each source synchronous group should be routed on the same layer for the entire length of the bus. This results in a significant reduction of the flight time skew since the dielectric thickness, trace width, and velocity of the signals will be uniform across a single layer of the stack-up. There is no guarantee of a relationship of dielectric thickness, trace width, and velocity between two layers.

Additionally, changing layers may create a return path discontinuity, which often leads to unpredictable delay push-out or pull-in and signal quality problems. See Section 6.1, "Return Path" on page 6-32 for more information on the effects of return path discontinuities.

To avoid return path discontinuities, traces must be routed with at least 50% of the trace width directly over a reference plane. This is particularly applicable when routing next to vias in the socket pin field.

Table 6-3. Source Synchronous Signal Groups

Signals	Associated Strobe
REQ [4:0]#, A [16:3]#	ADSTB0#
A [35:17]#	ADSTB1#
D [15:0]#, DBI0#	DDSTBP0##, DSTBN0#
D [31:16]#, DBI1#	DSTBP1#, DSTBN1#
D [47:32]#, DBI2#	DSTBP2#, DSTBN2#
D [63:48]#, DBI3#	DSTBP3#, DSTBN3#

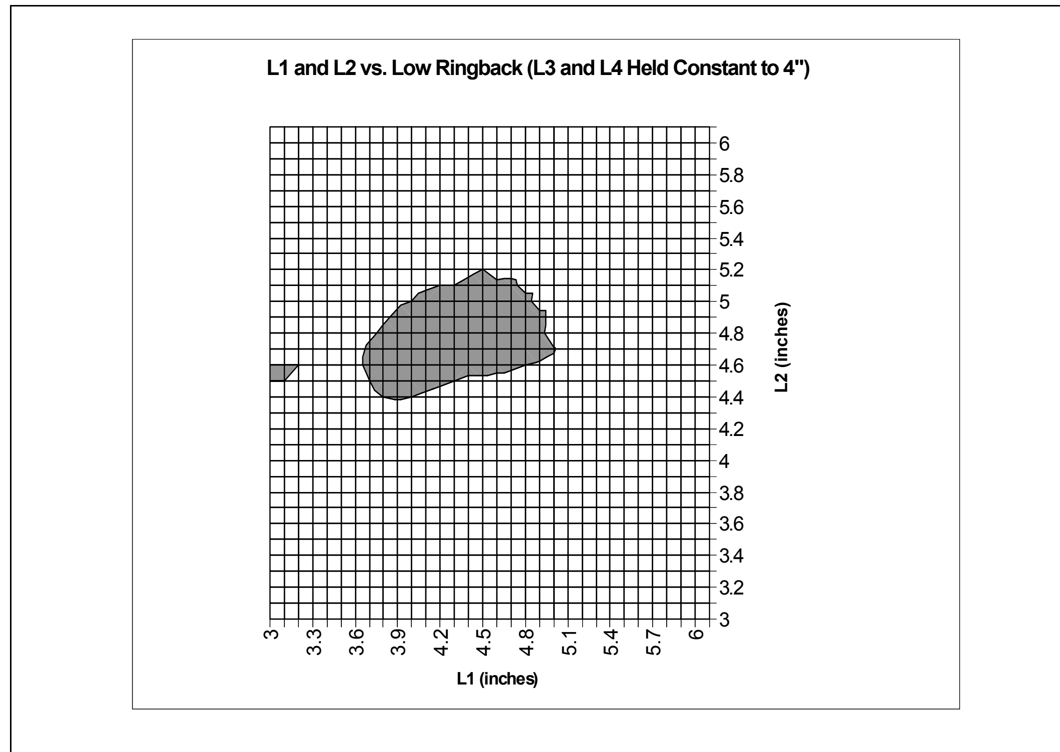
6.4.1.2 4X Group (DSTBP [3:0]#, DSTBN [3:0]#, D [63:0]#, DBI [3:0]#)

Figure 6-4 illustrates the 4-way daisy chain topology with the chipset at the end (i.e., L1, L2, L3, and L4 each must be between 3.0 inches and 6.1 inches). However, the maximum distance between one and only one set of agents may be up to 6.9 inches. A U-turn may exist between processor 2 and processor 3. Total bus length, from end agent to end agent, must not exceed 20.8 inches. The

distance from the package pin of one agent to the package pin of the next should be between 3.0 inches and 6.1 inches. There is an “island” of failing solution space based on L1 and L2 lengths. Refer to Figure 6-6 for a diagram of acceptable routing lengths.

The island in Figure 6-6 represents a configuration where the ringback on the system bus violates specification. This ringback is also a factor of ISI buildup over multiple cycles. This ringback is highly dependent on the length of L1 versus the length of L2. Use the diagram to find acceptable routing lengths for L1 and L2.

Figure 6-6. L1 vs. L2 Length Dependencies



Trace Length Balancing

Length must be added to the motherboard trace between each agent to compensate for package length differences that exist within a source synchronous data group. This length compensation will result in minimizing the source synchronous skew that exists on the system bus. Without the length compensation the flight times between a data signal and its strobe will be different, which results in an inequity between the setup and hold times. Since the strobe has a shorter package length there will be favoritism toward hold time and thus, length compensation is necessary to meet the setup time requirement. Note this will not make the pad-to-pad lengths between all agents equal in length, but it will balance the strobe-to-signal skew in the middle of the setup and hold window between all driver-receiver combinations. The delta between a signal's processor package length (cpu_pkglen) and the longest signal's processor package length (max_cpu_pkglen) in that source synchronous group must be added to the motherboard trace between each agent. The following equations provide a guideline for trace length balancing. Simulation results should provide a more accurate assessment of what these compensation lengths should be for a specific design. See Equation 6-1. This compensation length is added to each of lengths L1, L2, L3, and L4 as shown in Figure 6-4.

Equation 6-1. Processor Package Length Compensation to Be Added to Motherboard Trace

$$PD = delta_{net,group} = \max_cpu_pkglen_{group} - cpu_pkglen_{net}$$

Compensating for the chipset package lengths on the motherboard is also necessary. The amount that should be added can be calculated using Equation 6-2. This length is compensated for on L4.

Equation 6-2. Chipset Package Length Compensation to Be Added to Motherboard Trace

$$CD = cs_delta_{net,group} = \max_cs_pkglen_{group} - cs_pkglen_{net}$$

The routed motherboard lengths within a source synchronous group should match the results of the above equations to ± 25 mils between agent-to-agent and ±100 mils over the entire length of the bus. It is recommended to simulate this skew in order to determine the length that best centers the strobe.

Example: (Note, this example uses hypothetical numbers.)

Consider 2 signals, DSTBP0## and D1#, from the same group. Assume that L1 (motherboard trace from Processor 1 to Processor 2) for both DSTBP0# and D1# is 4 inches. Similarly, assume that the package trace for DSTBP0# is 0.15 inches (cpu_pkglen) and D1# is 0.35 inches (max_cpu_pkglen). Using Equation 6-1, the delta will be 0.20 inches (0.35 - 0.15). The length matching spreadsheet thus requires that an additional length of 0.20 inches be added to signal DSTBP0#. Hence, the new length for DSTBP0# will be 4.20 inches (instead of the current 4 inches). The length matching spreadsheet requires the new length of DSTBP0# to be within ± 0.025 inches (25 mils) of 4.20 inches. A similar calculation is done for L2 and L3. L4 requires the delta from Equation 6-1 and Equation 6-2 in order to determine the new length. Again, the spreadsheet requires that the new length for L4 be met within ± 0.025 inches (25 mils).

Refer to the appropriate Processor Signal Integrity Models and the length matching spreadsheet tool, for the package line lengths and assistance in matching the motherboard trace lengths.

This compensation not only makes up for the flight time difference caused by the difference in package lengths, but it also counteracts the capacitive loading effects caused by stubs on the bus. The stub lengths from the processor package are dispersed at intervals along the bus. These stubs act as capacitive loads, and thus degrade the edge-rate as a signal travels from one end of the bus to the other. Because all stub lengths are not the same, different signals will see varying degrees of degradation. The signals with shorter stubs will see almost no degradation while the signals with longer stubs will see significantly more degradation. For source synchronous signals, the goal is to reduce skew between a data and its strobe. Since the strobe signals typically have short package lengths, they will not see much edge-rate degradation. The other signals can have stub lengths of up to 600 mils and the edge-rate degradation can be dramatic in comparison to that of the strobe. These large differences in the edge-rates at the receiver can result in a very large skew between the signal and the strobe. This could result in a failure to meet the setup time requirements.

Because the length compensation is being added to aid in meeting the setup time when all processors are populated, unpopulated sockets will have greater setup timing margin.

Signals of the same source synchronous group should match the compensated lengths within 25 mils agent-to-agent and 100 mils over the entire length of the bus. It is not necessary to match lengths of one 4X signal group to other 4X signal groups. All signals must meet their setup and hold timing requirements.

In addition, strobes should maintain a 25-mil spacing from all other signals, including other strobes (DSTBn# and DSTBp#). It is also advisable to keep 4X signals away from non-4X signals, particular the asynchronous signals. Strobe signals should be routed following the data routing guidelines above. It is recommended to simulate this skew in order to determine the length that best centers the strobe for a given system.

6.4.1.3 2X Address Group (ADSTB [2:0]#, A [35:3]#, REQ [4:0]#)

The requirements for the 2X address group signals are the same as those for the 4X data group signals.

6.4.1.4 Common Clock

Common Clock signals should follow the same routing rules at the Data signals, however no length compensation is necessary.

The distance from the package pin of one agent to the package pin of the next agent should be between 3.0 inches and 6.9 inches. Figure 6-4 illustrates the 4-way system bus daisy chain topology with the chipset at the end. (I.e., L1, L2, L3, and L4 each must be between 3.0 inches and 6.9 inches) Total bus length, from end agent to end agent, must not exceed 20.8 inches. Simulation of these signals is strongly recommended to ensure they meet the setup and hold times to BCLK [1:0].

6.4.1.5 Wired-OR

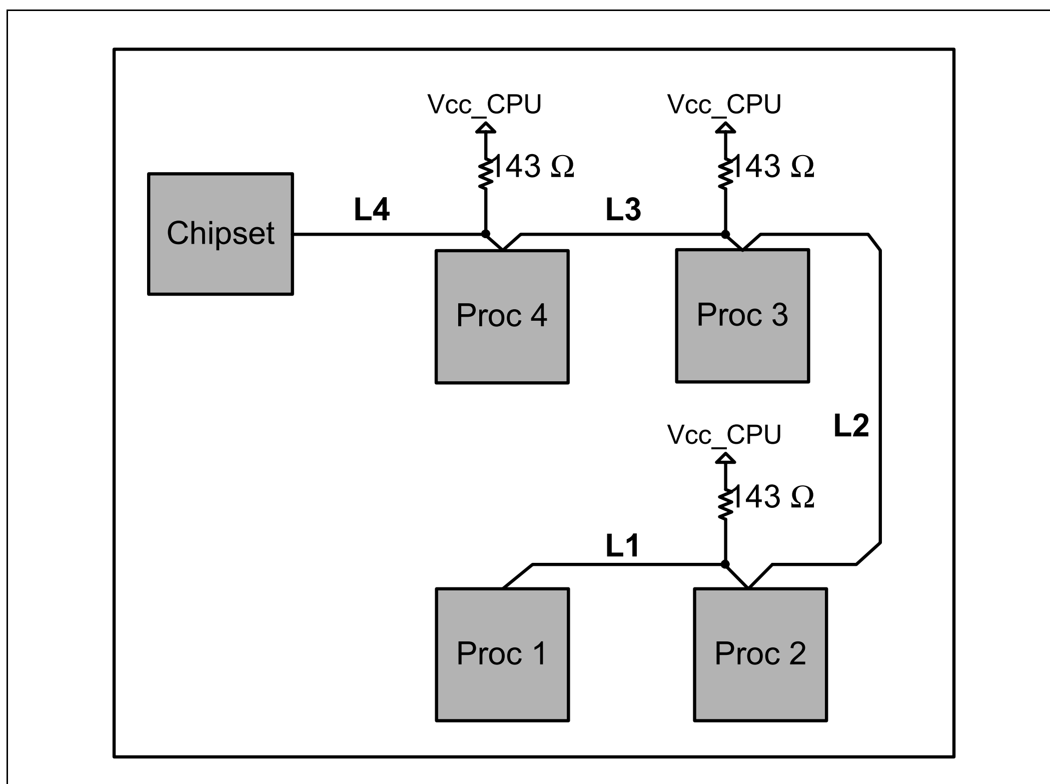
There are five "wired-OR" signals on the system bus. These signals are HIT#, HITM#, MCERR#, BINIT#, and BNR#. These signals differ from the other system bus signals in that more than one agent can be driving the signal at the same time. So, timing and signal integrity must be met for the case where one agent is driving, all agents are driving, or any combination of agents are driving. Therefore, specialized routing guidelines are required to meet signal integrity and timing requirements.

The wired-OR signals should follow the same routing rules as the common clock signals except for the items specified below. It is highly recommended that simulations for these signals be performed for a given system.

All wired-OR signals should have termination to V_{CC_CPU} at the middle agents (see Figure 6-7). The termination should be located as close as possible to the processor pins (< 1 inch) with no stubs. A 143 Ω resistor should be used for the termination. The nominal impedance of the wired-OR signal traces should be 25 Ω ± 10%. The lengths between agents should be:

- L1 = 4.2 inches – 6.2 inches when L4 = 3.0 inches – 6.2 inches OR L1 = 3.0 inches – 6.2 inches when L4 = 4.4 inches – 6.2 inches
- L2 = 3.0 inches – 6.2 inches
- L3 = 3.0 inches – 6.2 inches

Figure 6-7. Wired-OR Topology

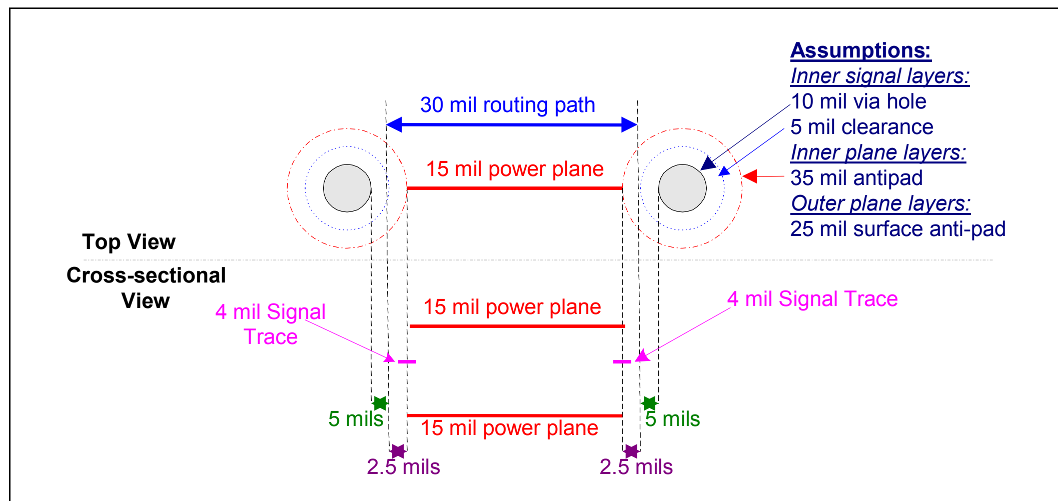


6.4.1.6 Design Considerations

Intel has found that the following recommendations provide one method of designing a 4-way processor platform. This is a baseline configuration only. Modify this baseline as needed while adhering to the above *Design Recommendations*.

- Maintain a trace width of 4.0 mil to provide a 47 Ω trace impedance.
- Trace to trace spacing of 16-mil edge-to-edge (except in component breakout where spacing is constrained)
- Rotate the processor socket such that the data pins are closest to the chipset making data the shortest signals.

Figure 6-8. 0.025" Via Pad with 50% of Trace over Reference Plane



Using the recommendations for via size shown in Figure 6-8 will ensure that two traces can be routed between vias and that the traces are overlapping the reference planes by at least 50%. The overlap should occur only for a short distance. The use of 4.0 mil wide traces allows good trace spacing to be maintained through the socket pin field. Three-dimensional simulation is recommended for determining the impact of signals without a solid reference plane.

6.4.2 Routing Guidelines for Asynchronous GTL+ and Other Signals

This section describes layout recommendations for signals not covered in the previous section. Table 6-4 shows the signals covered in this section. Each topology will be described in detail in the following sections. All signals must meet the AC and DC specifications listed in the processor datasheet.

Table 6-4. Asynchronous GTL+ and Miscellaneous Signal List (Sheet 1 of 2)

Signal Name	Type	Processor I/O Type	Topology Number	Driven by	Received by
A20M#	Asynchronous GTL+	I	2	Chipset	Processor
BR[3:1]#	AGTL+	I	5	Processor	Processor
BR0#	AGTL+	I/O	5	Processor	Processor/ Chipset
BINIT#	AGTL+	I/O	13	Processor/ Chipset	Processor
COMP[1:0]	Other	I	6	Pull-Down	Processor
FERR#	Asynchronous GTL+	O	1	Processor	Chipset
IERR#	Asynchronous GTL+	O	1	Processor	
IGNNE#	Asynchronous GTL+	I	2	Chipset	Processor
INIT#	Asynchronous GTL+	I	2	Chipset	Processor
LINT[1:0]	Asynchronous GTL+	I	2	Chipset	Processor

Table 6-4. Asynchronous GTL+ and Miscellaneous Signal List (Sheet 2 of 2)

Signal Name	Type	Processor I/O Type	Topology Number	Driven by	Received by
ODTEN	Other	I	7	Pull-Up / Pull-Down	Processor
PROCHOT#	Asynchronous GTL+	O	1	Processor	External Logic
PWRGOOD	Asynchronous GTL+	I	2	External Logic	Processor
RESET#					
SKTOCC#	Other	O	9	Processor	External Logic
SLP#	Asynchronous GTL+	I	2	Chipset	Processor
SM_ALERT#[1]	SMBus Interface	O	4	SMBus Agent	SMBus agent/ Processor
SM_CLK[1]	SMBus Interface	I/O	4	SMBus Agent	Processor
SM_DAT[1]	SMBus Interface	I/O	4	SMBus Agent	Processor
SM_EP_A[2:0][1]	SMBus Interface	I	4	Pull-Up / Pull-Down	Processor
SM_TS_A[1:0][1]	SMBus Interface	I	4	Pull-Up / Pull-Down	Processor
SM_WP[1]	SMBus Interface	I	4	Pull-Up / Pull-Down	Processor
SMI#	Asynchronous GTL+	I	2	Chipset	Processor
STPCLK#	Asynchronous GTL+	I	2	Chipset	Processor
TAP signals	TAP		See Section 6.4.3		
TESTHI[6:0]	Other	I	8	Pull-Up	Processor
THERMTRIP#	Asynchronous GTL+	O	1	Processor	External Logic
V_{CCA}	Power	I	See Section 8.13		
$V_{CCIOPLL}$	Power	I	See Section 8.13		
$V_{CCSENSE}$	Other	O	See Section 8.13		
VID[4:0]	Other	O	3	Processor	VRM
V_{SSA}	Power	I	See Section 8.13		
$V_{SSSENSE}$	Other	O	See Section 8.13		
GTLREF	Power	I	See Section 8.12.1		

6.4.2.1 Topology 1: Asynchronous GTL+ Signals Driven by the Processors; FERR#, IERR#, PROCHOT# and THERMTRIP#

These signals should adhere to the following routing and layout recommendations. Figure 6-9 illustrates the recommended topology. When routing to middle agents connect in true daisy chain topology. Do not create a stub to connect to the socket pins. Note that FERR# is the only signal in this group that connects between the processor(s) and the chipset. PROCHOT#, THERMTRIP# and IERR# are connected to other motherboard logic and may need voltage translation logic depending on voltage input thresholds of the motherboard receiver logic devices used. Figure 6-10 shows an example voltage translator circuit.

To protect the processors from damage in over-temperature situations, motherboard and/or chipset logic must ensure that power to the processor core is removed within 0.5 seconds after the assertion of THERMTRIP#. If power is applied to a processor when no thermal solution is attached, normal leakage currents will cause the die temperature to rapidly rise to levels at which permanent silicon damage is possible. This high temperature will cause THERMTRIP# to go active. For details regarding the THERMTRIP# specification, refer to the processor datasheet.

To avoid excessive undershoot seen at the processors, use dual termination on these four signals for a 4-way configuration. Each processor's signals can be routed to its own receiver or they can be wire-ORd together. If routed separately each signal must be terminated, but the signal can be terminated at the receiver end only. Figure 6-9 illustrates the recommended topology.

If the functionality of any of these signals is not required, it is acceptable to not connect the pin (let float).

Trace Zo	Trace Spacing	L1	L2	L3	Rpu
50 Ω	10 mil	4–6”	1–12”	3” max	56 Ω ± 5%

Figure 6-9. Topology 1 for 4-Way Configuration

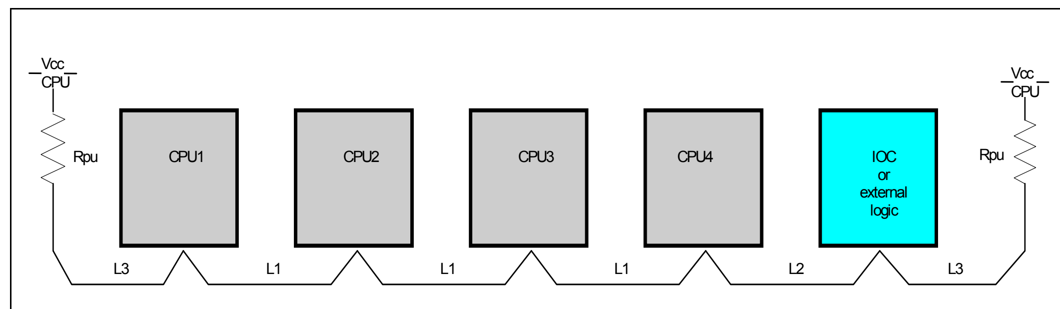
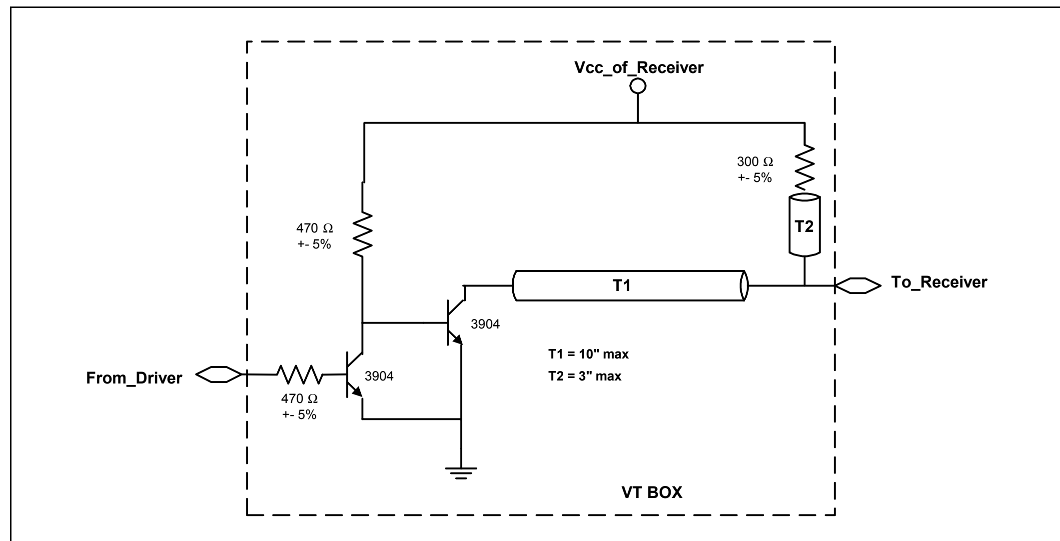


Figure 6-10. Example Voltage Translator Circuit



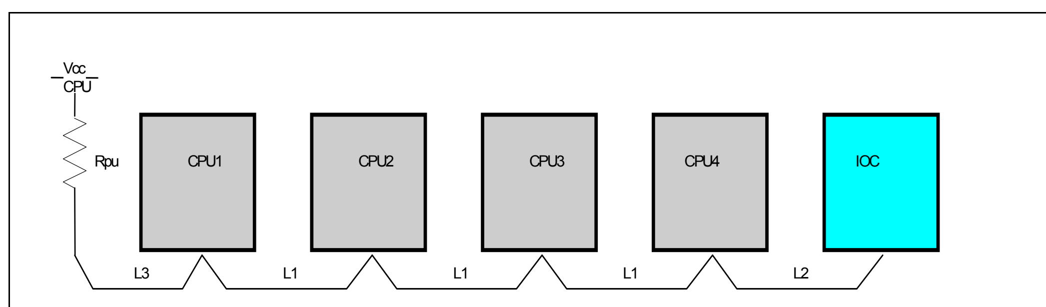
6.4.2.2 Topology 2: Asynchronous GTL+ Signals Driven by the Chipset; A20M#, IGNNE#, INIT#, LINT[1:0], PWRGOOD, SLP#, SMI#, and STPCLK#

These signals should adhere to the following routing and layout recommendations. Figure 6-11 illustrates the recommended topology. When routing to middle agents connect in true daisy chain topology. Do not create a stub to connect to the socket pins.

It may be desirable to isolate PWRGOOD for each VRM and processor pair in order to recognize individual VRM failures.

Trace Zo	Trace Spacing	L1	L2	L3	Rpu
50 Ω	10 mil	4–6”	1-12”	3” max	300 Ω ± 5%

Figure 6-11. Topology 2 for 4-Way Configuration



6.4.2.3 Topology 3: VID[4:0]

The VID[4:0] signals for the Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process should be routed to the VID[4:0] inputs on the voltage regulator controller. The voltage regulator controller should provide internal pull-up resistors for these signals. Refer to the *VRM 9.1 DC-DC Converter Design Guidelines* and the specification of the voltage controller specific to your design for further details.

Since all processors must operate at the same V_{CC} voltage, it is imperative to provide a way to check the VID[4:0] signals to ensure a processor does not operate out of specification.

6.4.2.4 Topology 4: SMBus Signals

The SMBus signals provide access to the manageability features on the processor. The signaling protocol adheres to the specification of the System Management Bus. Refer to the processor datasheet for details on the processor SMBus implementation and addressing scheme.

The SM_ALERT#, SM_CLK, and SM_DAT signals should be connected to SMBus controller in adherence to the *System Management Bus Specification, rev 2.0*. These signals can be connected to other processors on the same SMBus.

The SM_EP_A[2:0] signals set the SMBus address for the memory device on the processor. These signals need to be set at power up with a unique address per bus. They have an internal 10 kΩ pull down. To pull a signal to a logic high level, connect to a 100 Ω resistor tied to SM_VCC.

The SM_TS_A[1:0] signals set the SMBus address for the thermal device on the processor. These signals need to be set at power up with a unique address per bus. The SM_TS_A[1:0] can be set to a logic high, a logic low, or a high impedance state giving nine possible combinations of addresses. Refer to the section on SMBus Device Addressing in the datasheet for addressing details. The SM_TS_A[1:0] signals do not have an internal pull down and thus need to be pulled to V_{SS} or SM_VCC with a 1 kΩ or smaller resistor. Leaving the pins floating achieves a high-Z state.

The SM_WP signal is a write protect signal for the memory device. Pulling this signal to SM_VCC will enable write protection. SM_WP has an internal 10 kΩ pull-down.

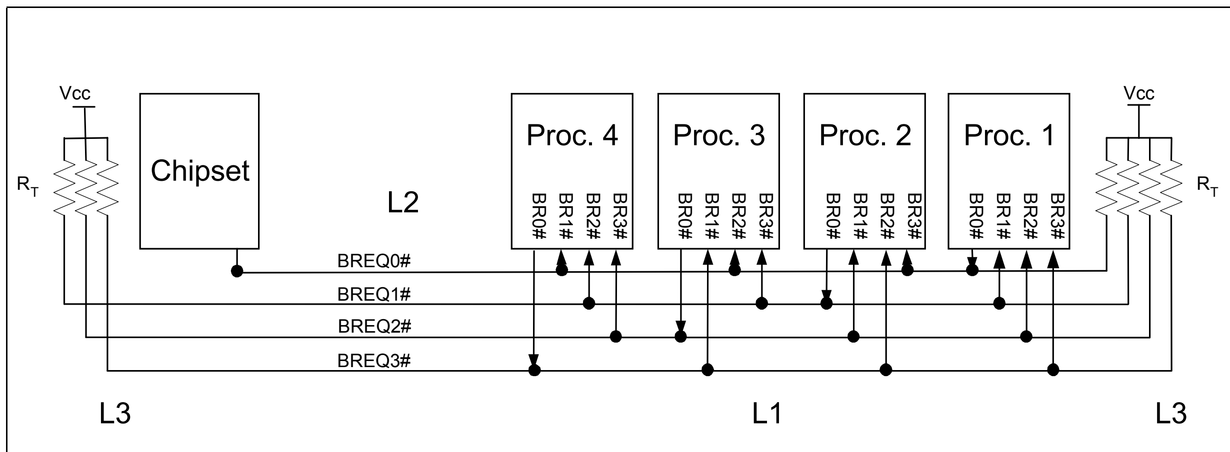
6.4.2.5 Topology 5: BR[3:0]# Signals

Since the Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process does not have on-die termination on the BR[3:0]# signals, it is necessary to terminate using discrete components on the motherboard. Connect the BR[3:0]# signals as in the past by "swizzling" the lines between the processors as shown in Figure 6-12, below. The chipset has on-die termination for BR0# and thus it is necessary to terminate only at the processor end. See the routing guidelines below. For other routing guidelines such as trace spacing and layer referencing follow the guidelines for common clock signals in Section 6.4.1.

Supports	Trace Zo	L1 Agent-to-agent	L2 (BR0#) Agent-to-Chipset	L3 Agent-to-Rpu stub	R_T
Existing Designs	50 Ω	3.0 to 6.1 inches[1]	Up to max total bus length of 20.8 inches	1 inch max	41 Ω ± 5%
New Designs[2]	50 Ω	3.0 to 6.1 inches[1]	Up to max total bus length of 20.8 inches	1 inch max	50 Ω ± 5%

NOTES:
1. .The maximum distance between one and only one set of agents may be up to 6.9 inches.
2. New designs are those platforms that only intend to support Intel Xeon processors MP with up to 2-MB L3 cache on the 0.13 micron process.

Figure 6-12. BR[3:0]# Connection for 4-Way Configuration



6.4.2.6 Topology 6: COMP[1:0] Signals

For details regarding termination of COMP[1:0] pins, please refer to the processor datasheet. Do not wire COMP pins together; connect each pin to its own termination resistor.

6.4.2.7 Topology 7: ODTEN Signal

The end processor in a 4-way processor system must have its on-die termination enabled. The middle agent should disable the on-die termination. To enable, pull the ODTEN pin to a high state by terminating it to V_{CC_CPU} through a resistor. To disable, pull the ODTEN pin to a low state by terminating it to ground through a resistor. There are two options for choosing the pull-up and pull-down resistor values. While both options are suitable for this platform, Option 1 is preferred over Option 2. The two available options are:

- Option 1 (preferred): Enable ODT (on-die termination) on Processor 0 (end processor) by pulling up to V_{CC_CPU} with a resistor that matches the motherboard trace impedance within ± 20%. Disable ODT on Processor 1, Processor 2, and Processor 3 by pulling down to V_{SS} with a resistor that matches the motherboard trace impedance within ± 20%. For example, since the recommended nominal trace impedance is 50 Ω, resistor values within the range of 50 Ω ± 20% should be used for the pull-up and pull-down.
- Option 2: Enable ODT on Processor 0 (end processor) by pulling up to V_{CC_CPU} with a 1 kΩ resistor. Disable ODT on Processor 1, Processor 2, and Processor 3 by pulling down to V_{SS} with a 1 kΩ resistor.

6.4.2.8 Topology 8: TESTHI[6:0] Signals

For each processor, all TESTHI[6:0] pins must be connected to V_{CC_CPU} via pull-up resistors. TESTHI[3:0] and TESTHI[6:5] may all be tied together at each processor and pulled up to V_{CC_CPU} with a single resistor, if desired. However, boundary scan testing will not be functional if any TESTHI pins are connected together. TESTHI4 must always be pulled up independently from the other TESTHI pins regardless of the usage of boundary scan. The TESTHI[6:0] signal group must not be connected between processors. There are four options for choosing the pull-up and pull-down resistor values. While four options are suitable for this platform, Intel recommends new designs or designs undergoing design updates follow the trace impedance matching termination guidelines given in Option 1a or Option 2a. The four available options are:

- Option 1a (preferred): All TESTHI[6:0] pins may be individually pulled-up to V_{CC_CPU} with resistors. For optimum noise margin, the pull-up resistor value should have a resistance value within ± 20% of the impedance of the board transmission line traces. Since the recommended nominal trace impedance is 50 Ω, use resistors that fall within the range of 50 Ω ± 20%.
- Option 1b: All TESTHI[6:0] pins may be individually pulled-up to V_{CC_CPU} with 1 kΩ ± 5% resistors.
- Option 2a (preferred): TESTHI[3:0] and TESTHI[6:5] may all be tied together and pulled up to V_{CC_CPU} with a single resistor. For optimum noise margin, the pull-up resistor value should have a resistance value within ± 20% of the impedance of the board transmission line traces. Since the recommended nominal trace impedance is 50 Ω, use resistors that fall within the range of 50 Ω ± 20%. However, utilization of boundary scan test will not be functional if these pins are connected together. TESTHI4 must always be pulled up independently from the other TESTHI pins.
- Options 2b: TESTHI[3:0] and TESTHI[6:5] may all be tied together and pulled up to V_{CC_CPU} with a single 1 kΩ – 4.7 kΩ resistor if desired. However, utilization of boundary

scan test will not be functional if these pins are connected together. TESTHI4 must always be pulled up independently from the other TESTHI pins.

6.4.2.9 Topology 9: SKTOCC# Signal

The SKTOCC# signal is an output from the processor used as an indication of whether a processor is installed or not. It will be asserted low when a processor is installed in the socket and will float when there is no processor present. SKTOCC# can be used to disable the VRM output for unpopulated processor sockets, the system power supply output when no processors are installed and other features.

6.4.3 Debug Port Signals Routing Guidelines for 4-way Configurations

The debug port design information can be found in a separate document. The routing of the signals, the signal levels, and all other information required to develop a debug port on this platform can be found in the *ITP700 Debug Port Design Guide*. See Section 1.1, “Related Documentation” on page 1-11 for more details.

This page is intentionally left blank.

Mechanical and EMI Design Considerations 7

7.1 Retention Mechanism Placement and Keep-Outs

The retention mechanism (RM) requires two keep-out zones, one for the EMI ground pads and another for a limited component height area under the RM as shown in Figure 7-1. Figure 7-2 shows the relationship between the RM mounting holes and pin one of the socket. In addition it also documents the ground pads and keep-out zones. Figure 7-3 details the ground pad locations and the associated limited height areas due to the ground frame.

The EMI ground pads under the retention mechanism should have a minimum of 8 vias connecting the pad to the baseboard ground plane. The retention holes should be a non-plated hole.

The ground pads for the EMI ground frame should have a minimum of 6 vias each connecting the pads to the ground plane. The suggested via size is 0.012. This should allow sufficient clearance to route traces between the vias on the secondary side of the PB or on internal layers.

Figure 7-1. Retention Mechanism Outline

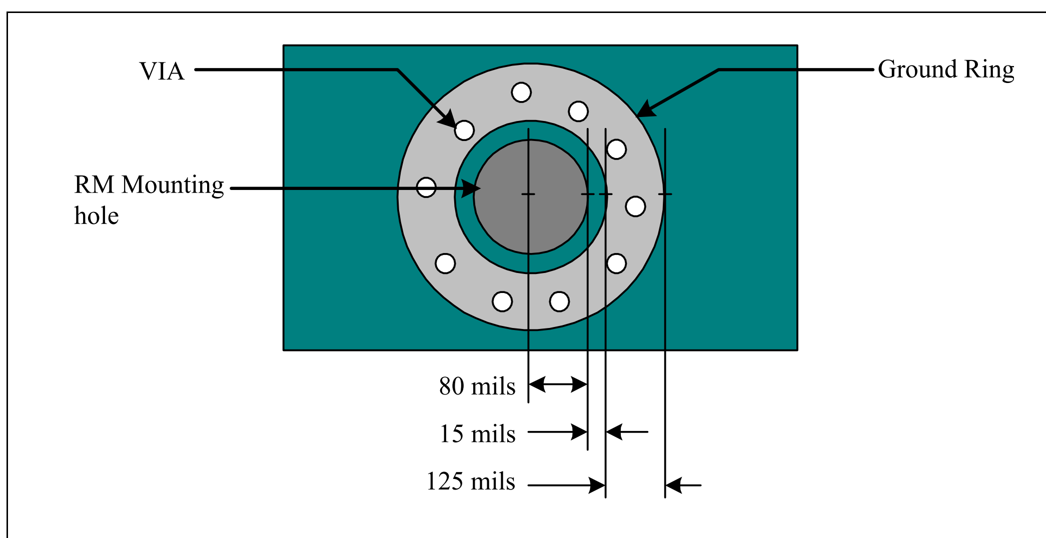
Figure 7-2. Retention Mechanism Placement and Keep-Out Overview

Intel recommends the following for constructing the mounting holes for the enabled retention mechanism.

- All four RM mounting holes must have ground pad rings.
- Ground pad annular ring should be no less than 125 mils wide. Try to cover the entire keep-out zone, if possible. See the following illustration for better dimensions.
- Place 8-12 vias in the annular ring, which connects the pad to internal ground planes.
- Anodizing or any form of insulated coating of the heat sink is strongly discouraged.

Refer to Figure 7-3 for specific details regarding the ground pads that are required to utilize this reduction technique.

Figure 7-3. Retention Mechanism Ground Ring



7.2 Electromagnetic Interference Considerations

7.2.1 Introduction

As microprocessor amperage and speeds increase, the ability to contain the corresponding electromagnetic radiation becomes more difficult. Frequencies generated by these processors will be in the low gigahertz (GHz) range, which will impact both the system design and the electromagnetic interference (EMI) test methodology.

This section is intended to provide electrical and mechanical design engineers with information that will aid in developing a platform that will meet government EMI regulations. Heatsink grounding, processor shielding, differential and spread spectrum clocking and the test methodology impact to FCC Class B requirements are specifically discussed.

Designers should be aware that implementing all the recommendations in this guideline will not guarantee compliance to EMI regulations. Rather, these guidelines may help to reduce the emissions from processors and motherboards and make chassis design easier.

7.2.2 Terminology

Electromagnetic Interference (EMI) - electromagnetic radiation from an electrical source that interrupts the normal function of an electronic device.

Electromagnetic Compatibility (EMC) - the successful operation of electronic equipment in its intended electromagnetic environment.

7.2.3 Brief EMI Theory

Electromagnetic energy transfer can be viewed in four ways: radiated emissions, radiated susceptibility, conducted emissions and conducted susceptibility. For system designers, reduction of radiated and conducted emissions is the way to achieve EMC compliance. Susceptibility is typically not a major concern in the desktop PC environment although it may be more important in an industrial environment.

The main component of EMI is a radiated electromagnetic wave, which consists of both electric (E-fields) and magnetic (H-fields) waves traveling together and oriented perpendicular to one another. Although E- and H-fields are intimately tied together, they are generated by different sources. E-fields are created by voltage potentials while H-fields are created by current flow. In a steady state environment (where voltage or current is unchanging), E- and H-fields are also static and of no concern to EMI. Changing voltages and currents are of concern since they contribute to EMI. If a dynamic E-field is present then there must be a corresponding dynamic H-field, and vice versa. Motherboards with fast processors will generate high-frequency E- and H-fields from currents and voltages present in the component silicon and signal traces.

Two methods exist for minimizing E- and H-field system emissions: prevention and containment. Prevention is achieved by implementing design techniques that minimize the ability of the motherboard to generate EMI fields. Containment is used in a chassis environment to contain radiated energy within the chassis. Careful consideration of board layout, trace routing and grounding may significantly reduce the motherboards radiated emissions and make the chassis design easier.

7.2.4 EMI Regulations and Certifications

Original Equipment Manufacturers (OEMs) ensure EMC compliance by meeting EMI regulatory requirements. System designers must ensure that their computer systems do not exceed the emission limit standards set by applicable regulatory agencies. Regulatory requirements referenced in this document include:

United States Federal Communication Commission (FCC) Part 15 Class B

International Electrotechnical Commission's International Special Committee on Radio Interference (CISPR) Publication 22 class B limits

The FCC rules are viewed to require any OEM who sells an “off-the-shelf” motherboard in the United States to pass an open chassis test. Open chassis testing is defined as removing the chassis cover (or top and 2 sides) and testing for EMI compliance (although permitted emission levels are allowed to be higher). Removing the cover greatly reduces the shielding provided by the chassis and increases the amount of EMI radiation. The purpose of this regulation is to ensure that system boards have reasonable emission levels since they are one of the main contributors to EMI.

7.2.5 EMI Design Considerations

The following sections discuss design techniques that may be applied to minimize EMI emissions. Some ideas have been incorporated into Intel-enabled designs (differential clock drivers, selective clock gating, etc.) and some must be implemented by motherboard designers (trace routing, clocking schemes, etc.).

7.2.6 Spread Spectrum Clocking (SSC)

Spread Spectrum Clocking is defined as continuously ramping (or modulating) the processor clock frequency over a predefined range (see Figure 7-4). SSC reduces radiated emissions by spreading the radiated energy over a wider frequency band (see Figure 7-5). Thus, instead of maintaining a constant system frequency, SSC modulates the clock frequency along a predetermined path (or modulating profile, Figure 7-4) with a predetermined modulation frequency. The modulation frequency is usually selected to be larger than 30 kHz (above the audio band) while small enough not to upset system timings (less than 0.8% of the clock frequency). SSC has been demonstrated to effectively reduce peak radiation levels, making EMC compliance easier to achieve.

To conserve the minimum period requirement for bus timing, the SSC clock is modulated between fnom and (1-d) fnom where fnom is the nominal frequency for a constant frequency clock. d specifies the total amount of spreading as a relative percentage of fnom. The modulation percentage is always a function of 1-d and not 1+d, as increasing the clock frequency above the rated speed of the processor may cause unpredictable operation.

Figure 7-4. Spread Spectrum Modulation Profile

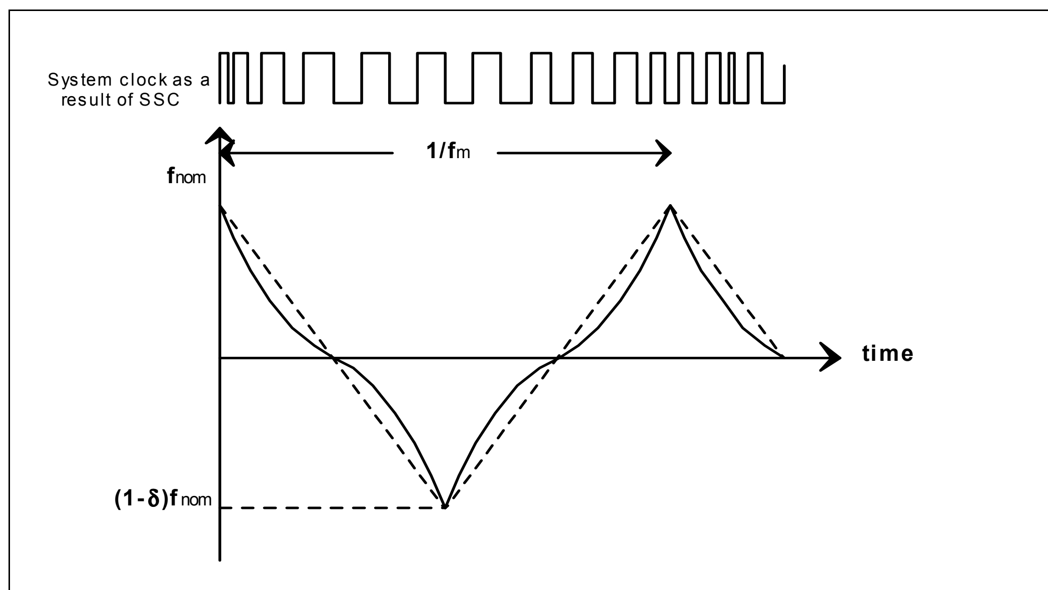
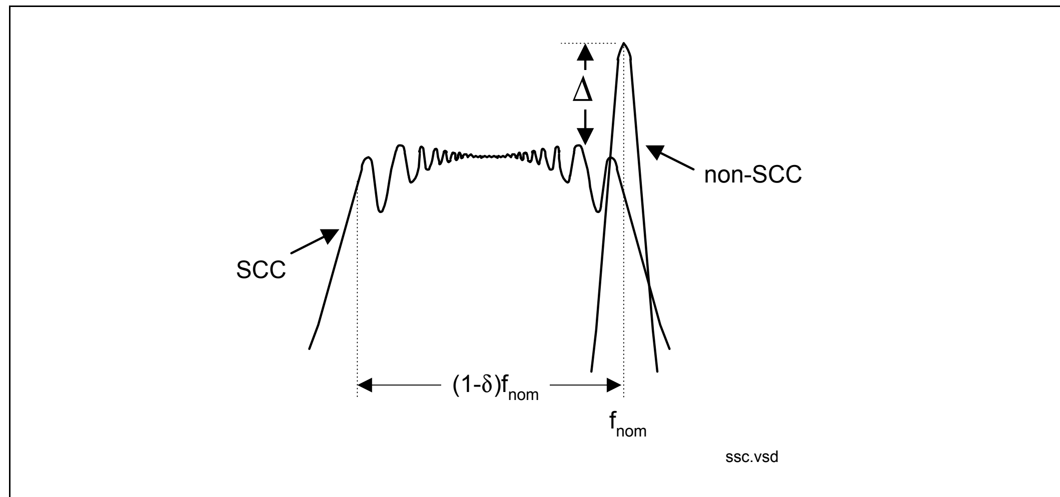


Figure 7-5. Impact of Spread Spectrum Clocking on Radiated Emissions

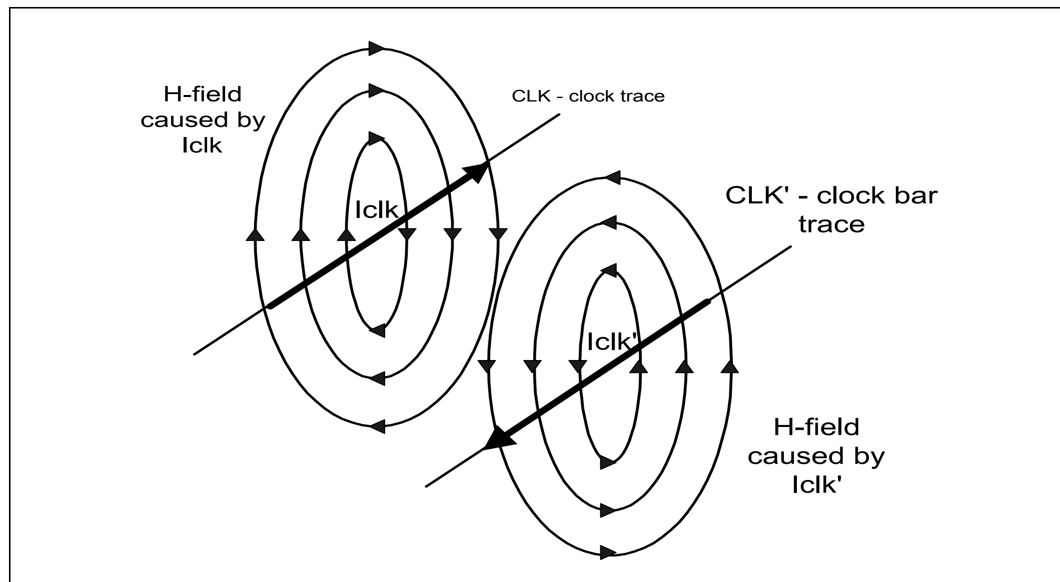


7.2.7 Differential Clocking

Differential clocking requires that the clock generator supply both clock and clock-bar traces. Clock-bar has equal and opposite current as the primary clock and is also 180 degrees out of phase. To maximize the benefit of differential clocking, both clock lines must be routed parallel to each other for their entire length. Devices connected to the clock must also be designed to accept both the clock and clock-bar signals.

EMI reduction due to differential clocking is caused by H-field cancellation. Since H-field orientation is generated by and is dependent upon current flow, two equal currents flowing in opposite directions and 180 degrees out of phase will have their H-fields cancelled (see Figure 7-6). Lower H-fields will result in reduced EMI radiation.

Figure 7-6. Cancellation of H-fields through Inverse Currents



Differential clocking can also reduce the amount of noise coupled to other traces, which improves signal quality and reduces EMI. I/O signals are particularly important because they often leave the system chassis (serial and parallel ports, keyboards, mouse, etc.) and will radiate noise that has been induced onto them. A single-ended clock's return path is usually a reference plane, which is shared by other signals/traces. When noise is created on a single-ended clock, the noise will appear on the reference plane and may be coupled to I/O traces. A differential clock's return path is the clock-bar signal/trace, which is more isolated than the reference plane and minimizes potential I/O trace coupling.

For best results, the trace lengths and routing of the clock lines must be closely matched and spacing between the two traces should be kept as small as possible. This will minimize loop area and maximize H-field cancellation. In addition, the real and parasitic terminations of each signal of a differential pair should be the same. Also, the skew between the signal level transitions on the two lines must be small compared to the rise time of the level transitions.

Placing ground traces on the outside of the differential pair may further reduce emissions. Intermediate vias to ground may be needed to reduce the opportunity for re-radiation from the ground traces themselves. Distance between vias should be less than ¼ of a wavelength of the fifth harmonic of the processor core frequency.

7.2.8 Heatsink Effects

Heatsink grounding may be an effective way to reduce system EMI emissions. Noise coupled from the processor package to the heatsink may cause it to act as an antenna and re-radiate the noise. Heatsink size, shape, fin pattern, orientation and material may all impact its ability to reradiate the high-frequency signals. Designers will have to experimentally investigate the behavior of a particular heatsink to determine its EMC performance.

Grounding of the heatsink through the Intel processor package is not possible with the current package implementation but may be an option at some time in the future. As such, system designers must design their own heatsink grounding solution.

When designing a grounding mechanism for the heatsink, care must be taken to minimize the impedance and distance between the ground paths. Typical guidelines suggest ground points should be separated by less than ¼ wavelength of the third harmonic of the processor core frequency.

Grounding materials should be selected to eliminate galvanic action between the various metals in contact. Oxidation of the various materials should also be considered as some oxides are non-conductive (for example, aluminum oxide) and will degrade EMC performance over time. Manufacturing process residue or coatings to prevent oxidation should also be checked for conductivity, especially at high frequencies.

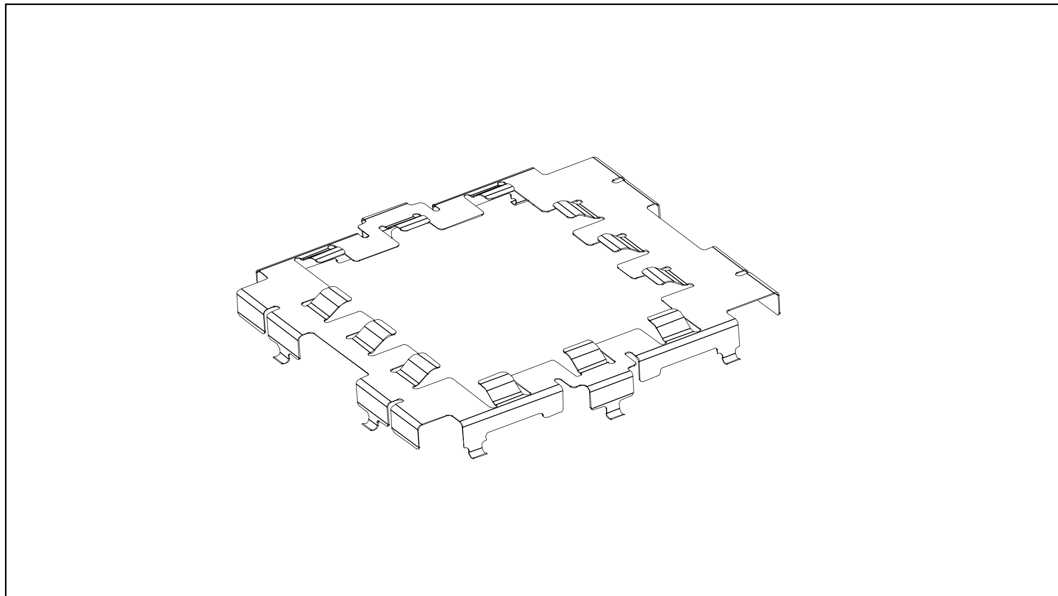
7.2.9 EMI Ground Frames and Faraday Cages

Grounding of heatsinks may reduce EMI, but that alone may not be sufficient to pass the required tests. Additional shielding of the processor itself may be necessary. A Faraday cage placed around the processor may provide a reduction in radiated noise and make chassis design easier.

A true Faraday cage would completely surround the source of radiation and contain all radiated energy. Within the limitations of processor packaging and motherboard assembly it is not possible to create a true Faraday cage around the processor. By using the heatsink and motherboard ground plane as two sides of the cage and a metal frame to enclose the remaining four sides, a reasonable approximation of a Faraday cage can be achieved.

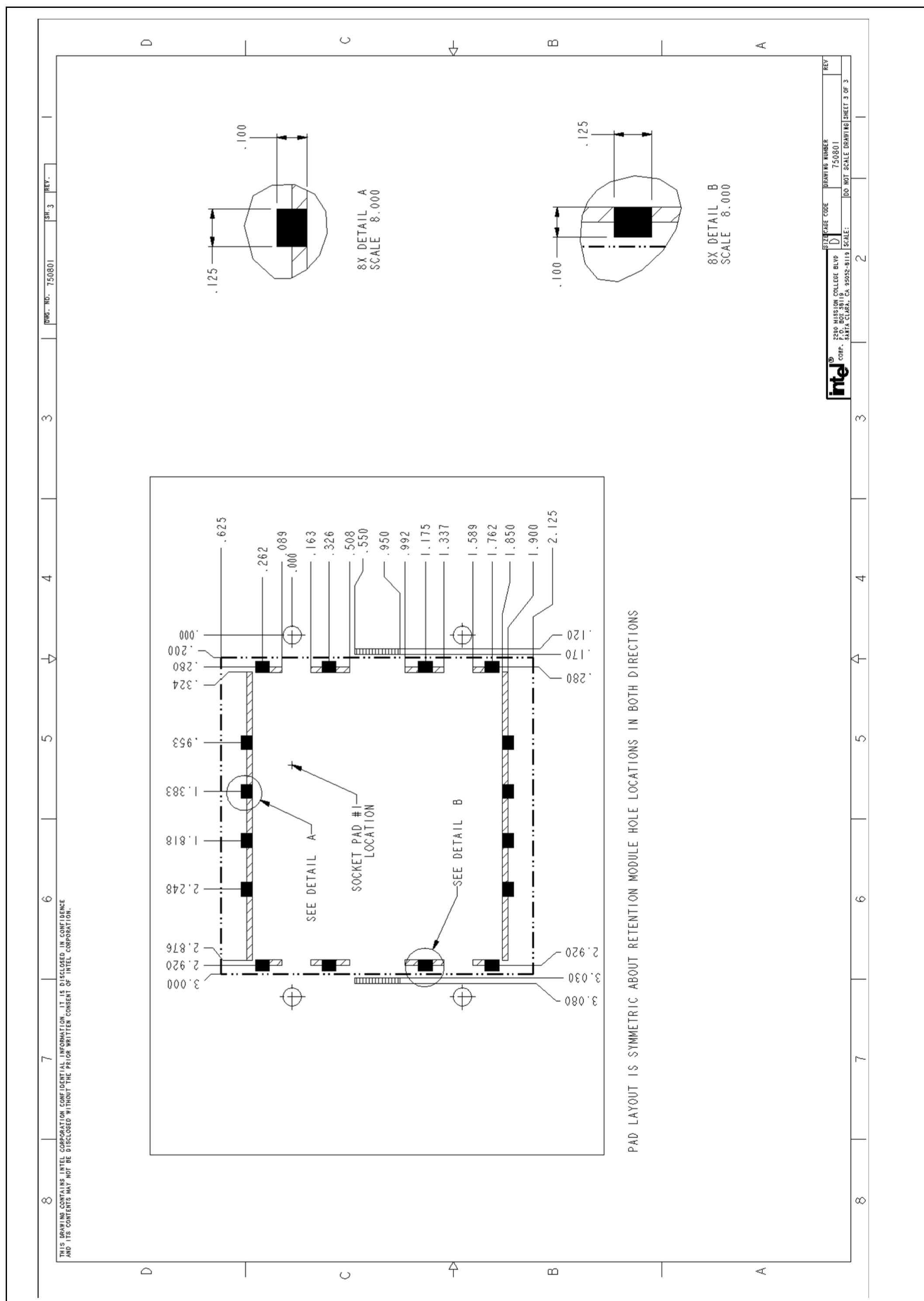
Intel has designed a “picture frame” type of grounding device, called an EMI ground frame, that fits between the processor and heatsink (see Figure 7-7). With this implementation, it is unnecessary to design a separate heatsink grounding mechanism, as the frame will provide this capability. OEMs who choose to use the Intel designed grounding frame will need to provide ground pads on the top layer of the motherboard around the processor socket (see Figure 7-8). These pads will provide the necessary ground continuity to complete the Faraday cage. Exact physical dimensions of the frame and the material used are provided in the processor IGES and Pro-E models for enabled components. The required motherboard ground pad descriptions are provided in the Mechanical Design Considerations chapter of this Platform Design Guide.

Figure 7-7. Conceptual Processor Ground Frame



The EMI ground frame is meant to provide grounding of AC currents seen on the heatsink and has been shown to be the most effective design in EMI reduction for the processor. The metal frame will be installed after the processor and retention mechanisms have been inserted. It will fit around the processor and inside of the retention mechanisms. Fingers on the top of the metal frame will provide contact to the heatsink, and fingers along the bottom will contact the ground pads on the motherboard.

Figure 7-8. EMI Ground Pad Size and Locations



7.2.10 EMI Test Capabilities

FCC regulations in the United States specify the maximum test frequency for products with clocks in excess of 1 GHz is 5 times the highest clock frequency or 40 GHz, which ever is lower. OEMs are advised to inquire into the capabilities of their preferred EMC test lab to ensure they are able to scan up to the required frequency range.

History indicates that processor performance and frequency double approximately every two years. With this in mind, it would be advisable to be prepared for the frequencies that will need to be scanned in the next few years.

Since the FCC Rules ultimately require testing to 40 GHz, commercial test equipment has been developed which is capable of making measurements to that frequency. Although it will be some time before processors require testing at this frequency, it may be cheaper to upgrade to 40 GHz now rather than making several intermediate steps.

It is also possible to upgrade various parts at different times. The spectrum analyzer may be upgraded to 40 GHz today while only obtaining the necessary antennas to support the initial processor frequencies. As processor speed increases, the necessary antennas and cables could be purchased which would support testing to the higher levels. Cost flexibility in antenna selection is probably the greatest, as different antenna designs are necessary for different frequency ranges.

7.2.11 Summary

High-speed clock frequencies within the platform will make EMI compliance more challenging. In order to facilitate successful chassis and motherboard designs, Intel has developed a number of components and techniques to reduce or contain EMI emissions. OEMs are advised to verify that their preferred EMC test facility is capable of measuring the required frequencies.

This page is intentionally left blank.

Processor Power Distribution Guidelines 8

Note: Intel recommends systems utilize modules based on *VRM 9.1 DC-DC Converter Design Guidelines* for Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process-based designs. These recommendations are required to meet the current requirements of the processor.

8.1 Introduction

As computer performance demands increase, new, higher speed logic with increased density is developed to fulfill these needs. To reduce their overall power dissipation, modern microprocessors are being designed with lower voltage implementations. This in turn requires power supplies to provide lower voltages with higher current capability. Because of this, processor power is now becoming a significant portion of the system design and demands special attention. Now more than ever, power distribution requires careful design practices. Intel Xeon processors MP and Intel Xeon processors MP with up to 2-MB L3 cache on the 0.13 micron process have unique requirements for voltages supplied to them. Their system bus implementation based on AGTL+ technology, the processor core, and the cache are being powered from the same voltage supply. The demand on the supply current and transient specification has increased drastically by the processor core. As the differences in processor current between the low power state and the high power state increase, the cost of the power distribution system becomes significant enough to merit careful calculation. Centralized distribution of power, for example, is no longer the effective solution to power distribution.

The intent of this chapter is to familiarize the reader with the power requirements of the Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process. In addition, the chapter discusses simulation and power implementation techniques. It is assumed the reader is familiar with power distribution issues of the Pentium III, Pentium II Xeon and Pentium III Xeon processors.

8.2 Terms

"Flexible Mother Board" or "FMB" is an estimation of the worst-case value the processor family will have over the lifetime of the product. The value is only an estimate and actual specifications for future processors may differ. System designers should meet the FMB values to ensure their systems will be compatible with all future processor family offerings.

"Power-Good" or "PWRGOOD" (an active high signal) indicates that all of the supplies and clocks within the system are stable. PWRGOOD should go active a predetermined time after system voltages are stable and should go inactive as soon as any of these voltages fail their specifications.

"V_{CC}" in this 000, refers to the appropriate processor core V_{CC}, cache supply voltage and Assisted Gunning Transceiver Logic + (AGTL+) supply voltage. With Intel Xeon processors MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process, the core and cache are on the same silicon and are powered from the same power plane unlike Pentium II Xeon and Pentium III Xeon processors, which required different power planes.

"VRM 9.1" refers to the Voltage Regulator Module for the processor. It is a DC-DC converter module that supplies the required voltage and current to a single processor.

8.3 Power Delivery Overview

Power distribution is generally thought of as supplying power to the components that require it. Most digital designers typically assume that an ideal supply will be provided. The printed circuit board (PCB) designers attempt to create this ideal supply with two power planes in the PCB or by using large width traces to distribute power. High-frequency noise created when logic gates switch is typically controlled with high-frequency ceramic capacitors, which are recharged from lower frequency bulk capacitors. Various rule of thumb methods exist for determining the amount of each type of capacitance that is required. For Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process-based designs, the system designer needs to design beyond the rule of thumb and architect a power distribution system that meets appropriate processor specifications.

Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process-based systems have the capability to use four or more processors. The processor core and all of the caches are operating at the same voltage level, i.e., V_{CC}. On-die termination is used to pull the AGTL+ bus up to V_{CC} to control reflections on the transmission line. The data bus must route over a uniform power plane because of signal quality constraints. Consequently in a multiprocessor system design a single power plane should be used for power delivery to all processors. Hence, the mixing of processors operating at different voltages is not supported and will not be validated by Intel.

8.4 Processor Power Delivery Ingredients

Discussion of processor power delivery may be broken down into seven ingredients:

1. System Design.
2. Processor Load
3. Voltage Regulator
4. Power Planes
5. Decoupling Capacitors
6. Component Placement and Modeling
7. Validation Testing

8.5 System Design

8.5.1 Multiple Voltages

The voltage regulator modules that provide V_{CC} supply to processor and have the capability of supplying voltages from +1.1 V to +1.85 V. The V_{CCA} supplies power to the processor core and on-die termination used for AGTL+ bus.

Multiple voltages required for Intel Xeon processor MP based systems are V_{CC_MAX} = 1.7 V and SM_V_{CC} = 3.3 V. Similarly, for and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process processors V_{CC_MAX} = 1.475 V and SM_V_{CC} = 3.3 V. $V_{CCIOPLL}$, V_{CCA}, and V_{SSA} are the power supplies to the internal PLL. $V_{CCIOPLL}$, V_{CCA} and V_{SSA} must be connected to V_{CC} through a discrete RLC filter as described in Section 8.13. Refer to the processor datasheet for the pin location of these voltage supplies and specifications for all processor voltage supplies.

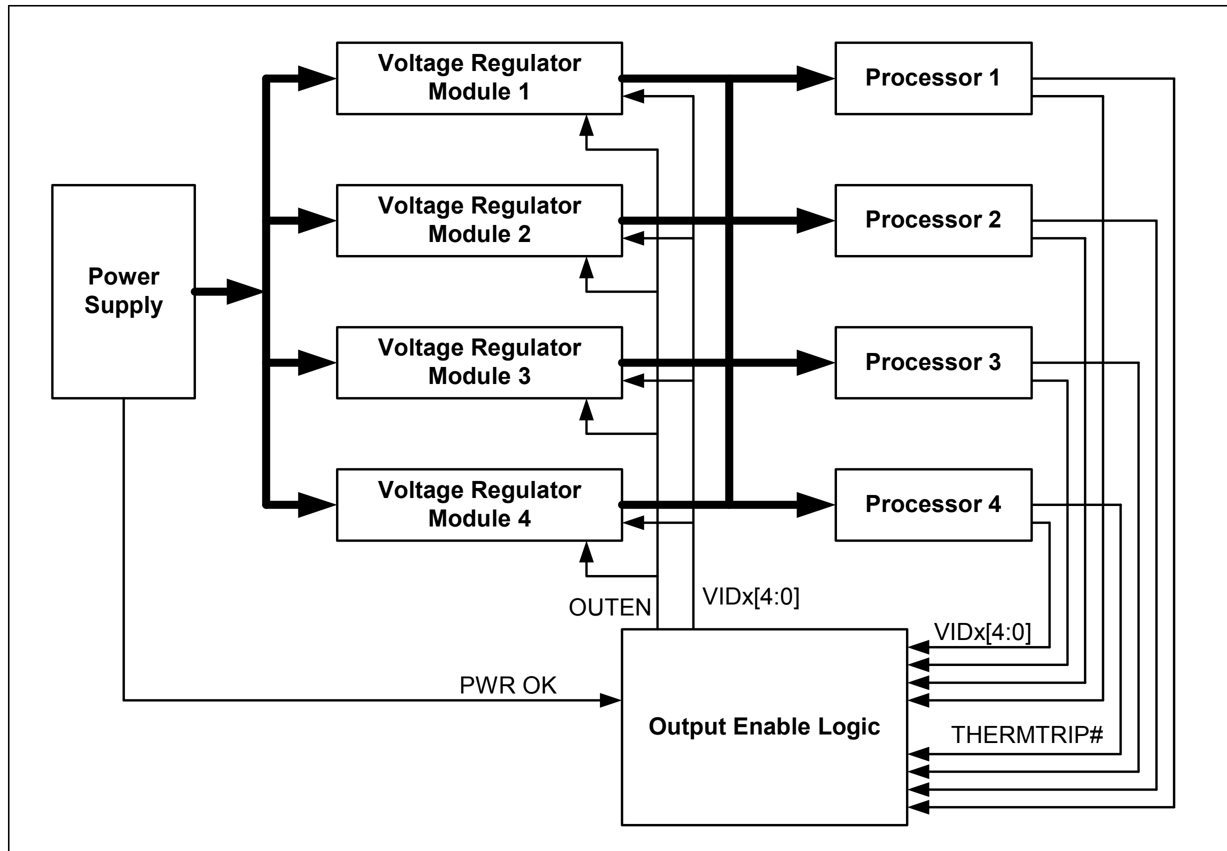
8.5.2 Voltage Sequencing

When designing a system with multiple voltages, there is always the issue of ensuring that no damage occurs to the system during voltage sequencing. Voltage sequencing is the timing relationship between two or more voltages, such as V_{CC} and SM_V_{CC}. Sequencing applies to the power voltage levels and the levels of certain other crucial signals when the user turns on or off the power supply, or the system enters a failure condition. V_{CC} from the voltage regulators should be enabled after assertion of the Power Supply Power Good signal and disabled upon de-assertion of the Power Supply Power Good signal. In addition, in the event of any processor asserting its THERMTRIP# signal, V_{CC} must be disabled within 0.5 s. Please contact the chipset vendor for the recommended circuit to disable power to the processor.

8.5.3 Block Diagrams with Voltage Regulator Modules

Figure 8-1 depicts the recommended four-way baseboard solutions involving local voltage regulator modules (VRMs). The block diagrams also recommend the implementation of logic for monitoring the VID [4:0] of all processors. This logic should determine that all of the installed processors are requesting the same V_{CC}. If mixed voltage processors are detected the output enable signal (OUTEN) of all voltage regulators must be disabled. Note that if a processor is not installed the VID [4:0] of that processor are high, and this should not cause disabling of the output of other VRMs.

Figure 8-1. Power Distribution Block Diagram for 4- Way System Motherboard with Voltage Regulator Modules



8.6 Processor Load

8.6.1 Processor Voltage Tolerance

Refer to the processor datasheet for die voltage tolerance specifications. Failure to meet these specifications on the low-end tolerance can result in system error or lock up. Not meeting these specifications on the high-end tolerance can cause damage or reduce the life of the processor.

Refer to the *VRM 9.1 DC-DC Converter Design Guidelines* for voltage regulator tolerance specifications (regulation requirements at the voltage regulator remote sense point located at the geometric center of the processors).

8.7 Voltage Regulator

Intel recommends that the processor power provided by a voltage regulator module (one per processor) meets the specifications as described by the *VRM 9.1 DC-DC Converter Design Guidelines*. The voltage regulator definition includes Remote-Sense, Current Share and Output Enable features. Voltage regulator designers must provide these features as well as meeting voltage and current requirements set forth in the regulator design guidelines. The voltage regulator output slew rate is generally less than 50 A/µs. The slew rate at the processor socket pins can be as high as 450 A/µs. The system designer needs to provide adequate bulk and high-frequency decoupling on the baseboard to meet the processor required slew rate. Generally, the main power source for the voltage regulator is 12 V +5%, -8%. This voltage is supplied by a conventional computer power supply through a cable to the baseboard.

8.7.1 Voltage Regulator Design

It is outside the scope of this document to provide all the nuances involved in creating a high performance voltage regulator. Intel provides a list of enabled VRM suppliers. Also, manufacturers of voltage regulator IC controllers provide application notes, demo modules, design reviews and other forms of customer support including on site design assistance.

8.7.2 Voltage Regulator System Matching

8.7.2.1 Voltage Regulator Output

Some voltage regulator modules may be purchased with output capacitance included although many voltage regulator modules available today rely entirely on the baseboard for their output filtering capacitance. Regardless, the transient response or bandwidth of the voltage regulator must complement the combined output decoupling and storage capacitors so that the DC to 10 MHz impedance seen by the processor socket is less than 0.25 mΩ for a four processor system.

8.7.2.2 Voltage Regulator Input

Some voltage regulator modules may be purchased with input power filtering included although many voltage regulator modules available today rely entirely on the baseboard for their input power filtering. Input filtering requirements are dependant upon the power source and transient tolerance of common loads. However, for a typical 12 V power source, the voltage regulator input filter must limit its power source's current rate of change to less than 100 mA/µs. Other peripheral 12 V components should be powered separately from the voltage regulators.

8.7.2.3 Voltage Regulator Cooling

High performance voltage regulators generally operate at less than 85% efficiency. Careful attention must be given to providing adequate cooling air or thermal conduction paths. If using voltage regulator modules, the temperature of the module board must not exceed 90 °C at the connector interface. Specifically, to maintain the connector within its operating temperature range, the VRM board must not exceed 90 °C at any point within 2.54 mm of the top of the connector.

8.7.2.4 Voltage Regulator Remote Sense Connection

The system board is to include a positive and a negative SENSE input for each voltage regulator module. The round trip trace resistance should not be greater than one ohm. These voltage sense lines draw little current and there should be only a minute voltage drop from the remote sense connection and the voltage regulators.

In a multiprocessor system the SENSE lines from the voltage regulator modules should be routed to a point in the middle of and equidistant from all processors. At this remote sense point all positive SENSE lines should be tied together and connected to V_{CC_CPU} and all negative SENSE lines should be tied together and connected to V_{SS_CPU}.

All SENSE lines should be routed directly between the remote sense point and the voltage regulator and should not exceed 5 inches in trace length. If the SENSE lines are routed parallel to signal lines, the SENSE lines should be shielded.

8.7.2.5 Voltage Regulator Module ISHARE Connection

If voltage regulator modules are used it is necessary that the ISHARE lines of the modules be connected. The round trip resistance of this connection should be less than 1 Ω. The I_{SHARE} connection traces should be shielded from or separated from noisy signal traces.

8.7.2.6 Voltage Regulator Module OUTEN Connection

The voltage regulators' OUTEN input is used to disable the regulators' output voltage. The system designer should disable the output of the voltage regulator in a multiprocessor system when processors with different voltage settings are installed. The block diagram, Figure 8-1, shows this implementation based on VID [4:0] signals. Note that the VID lines are pulled up internal to the voltage regulators. No pull-ups are allowed on the baseboard. If the designer adds pull-ups on the baseboards, voltage-sequencing problems can occur with unpredictable results.

8.8 Power Planes

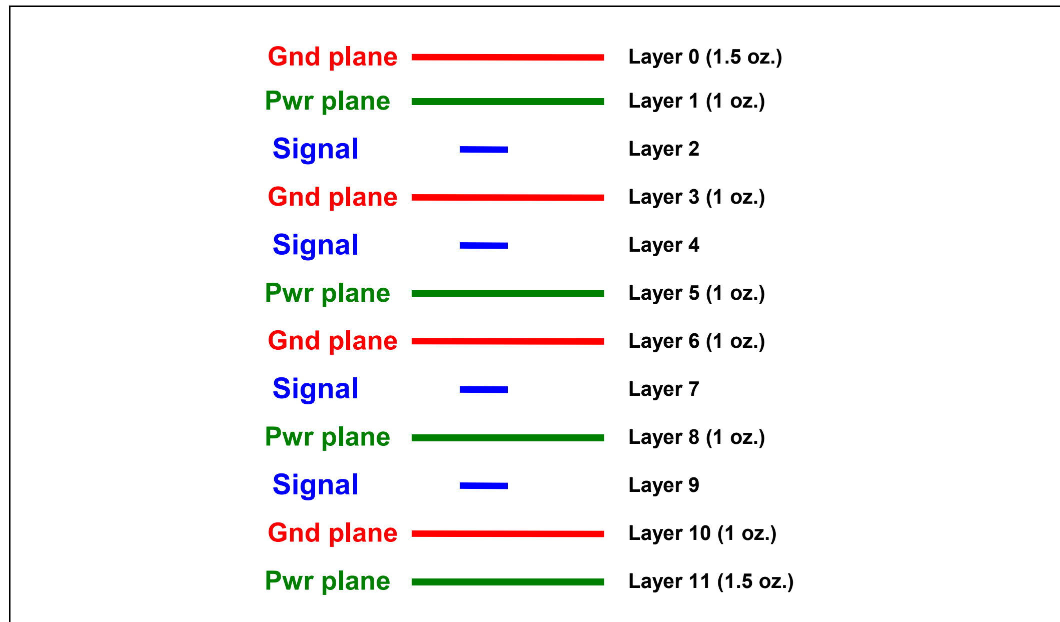
V_{CC} static and transient tolerances of the processor, and the corresponding voltage regulator tolerances assume power distribution paths with round trip resistances no greater than 300 mΩ and inductances any greater than 100 pH. Power must be distributed as a plane. This plane can be constructed as an island on a layer used for other signals, on a supply plane with other power islands, or as a dedicated layer of the PCB. Processor power should never be distributed by traces alone.

Due to the fact that processor voltage is unique to most system designs, a voltage island will probably be the most cost-effective means of distributing power to the processors. This island from the source of power to the load should not have any breaks, so as to minimize inductance in the plane. It should also completely surround all of the pins of the source and all of the pins of the load.

8.8.1 Layer Stack-Up

Intel recommends an absolute minimum of two ounce copper power plane for both V_{CC} and V_{SS}. The goal should be at 3.5 ounces or more for four processor systems. This can be implemented using multiple layers as shown in example Figure 8-2.

Figure 8-2. Suggested Twelve Layer Stack-Up for Four Processor Systems

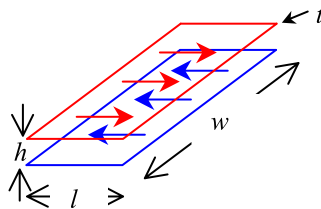


8.8.2 Sheet Inductance/Resistance and Emission Effects of Power Plane

The imperfections of the power planes themselves may introduce unwanted resistance and inductance into the power distribution system. Assuming layer thickness is smaller than skin depth, the metal layer resistance can be calculated as:

$$R = \rho \cdot \frac{l}{w \cdot t}$$

Where ρ is the copper resistivity (ρ = 0.667 mΩ-mil), *l*, *w*, and *t* are the length, width and thickness of the metal layer, respectively.



The loop inductance can be calculated as:

$$L = 31.9 \frac{\text{pH}}{\text{mil}} \cdot \frac{l \cdot h}{w \cdot (N - 1)}$$

Where N is the number of V_{CC}/V_{SS} planes. To minimize parasitic layer inductance, it is important to reduce the distance from decoupling capacitors to the processor socket (reducing *l*) and to use islands for power distribution (increasing *w*). To reduce *h*, it is recommended to select the V_{CC}/V_{SS} planes in the layer stack up that are interleaved and have small spacing in between. As a practical matter it is impossible to get the requisite baseboard inductance without locally dedicating at least 4 planes to carry power from the baseboard capacitors to the power pins of the processor.

There are impedance consequences for signals that cross over or under the edges of the Power Island that exists on another layer. While neither of these may be necessary for most designs, there are two reasonable options to consider which can protect a system from these consequences.

Processor power islands can be isolated from signals by one of the solid power plane layers such as the ground layer. This forces a particular stack-up model.

Another option that helps, but does not completely eliminate radiation effects, is to decouple the edges of the processor power islands to ground on regular intervals of about 1 inch, using good high-frequency decoupling capacitors (1206 packages). This requires more components but does not require any particular board stack-up.

In either case, for controlling emissions, all planes and islands should be well decoupled. The exact board layout, and the chassis design will determine the amount of decoupling required for controlling emissions. One should plan ahead by allowing additional pads for capacitors to be added in case they are found to be necessary during EMI testing.

Signals routed over power islands or islands in the ground plane create a discontinuity in the return path of that signal. This discontinuity can have detrimental effects on the timing and signal quality of that signal and other signals referencing the same planes. Avoid routing signals over splits in power planes or ground planes at all times.

Example:

Given power bussing area from the regulator to the socket approximated as a rectangle, with the following dimensions for the power and ground plane: l = 0.279 inch; w = 2.09 inches; t = 1.24 mils (1 oz. copper):

$$R = 0.677\ \text{m}\Omega \cdot \text{mil} \cdot \frac{0.279''}{2.09 \cdot 1.24\ \text{m}il} = 0.073\ \text{m}\Omega$$

The total resistance of the round trip is:

$$R = 2 \cdot 0.073\ \text{m}\Omega = 0.15\ \text{m}\Omega$$

With the V_{CC}-V_{SS} separated by 4.5 mils, the loop inductance is:

$$L = 31.9\ \frac{\text{pH}}{\text{mil}} \cdot \frac{0.279\ \text{mil} \cdot 4.5\ \text{mil}}{2.09\ \text{mil} \cdot (2 - 1)} = 19.2\ \text{pH}$$

8.9 Decoupling Capacitors

8.9.1 Decoupling Technology and Transient Response

The inductance of the system due to cables and power planes slows the power supply's ability to respond quickly to a current transient. Decoupling a power plane can be broken into several independent parts. The closer to the load the capacitor is placed, the more inductance that is bypassed. By bypassing the inductance of leads, power planes, etc., less capacitance is required. However, closer to the load there is less room for capacitance. Therefore trade-offs must be made.

Intel Xeon processors MP cause very large switching transients. These sharp surges of current occur at the transition between low power mode and high power mode. It is the responsibility of the system designer to provide adequate high-frequency decoupling to manage the highest frequency components of the current transients. To lower total board inductance and resistance, Intel Xeon processors MP are designed with approximately 141 V_{CC} and 141 V_{SS} (ground) pins. Intel Xeon processors MP with up to 2-MB L3 cache on the 0.13 micron process are designed with 188 V_{CC} and 189 V_{SS} (ground) pins. The designer needs to support a current slew rate of 450 A/µs at the socket pins. Larger bulk storage such as OSCON capacitors, supply current during longer lasting changes in current demand by the component, such as coming out of an idle condition. Similarly, they act as a storage well for current when entering an idle condition from a running condition.

All of this power bypassing is required due to the relatively slow speed at which a DC-to-DC converter can react. A typical voltage converter has a reaction time on the order of 1 µs to 10 µs while the processor's current steps are on the order of 100 ns to 200 ns. Bulk capacitance supplies energy from the time the high-frequency decoupling capacitors are drained until the power supply can react to the demand. More correctly, the bulk capacitors in the system slow the transient requirement seen by the power source to a rate that it is able to supply, while the high-frequency capacitors slow the transient requirement seen by the bulk capacitors to a rate that they can supply.

A load-change transient occurs when coming out of or entering a low power mode. This load-change transient can be on the order of 55 Amps. These are not only quick changes in current demand, but also long lasting average current requirements. This occurs when the processor enters or leaves a low power state. Please refer to the processor datasheet for more information on the low power states. Note that even during normal operation, the processor current requirements can change by as much as 70% (± 10%) of the max current very quickly.

Maintaining voltage tolerance, during these changes in current, requires high-density bulk capacitors with low Effective Series Resistance (ESR) and low Effective Series Inductance (ESL). Use thorough analysis when choosing these components.

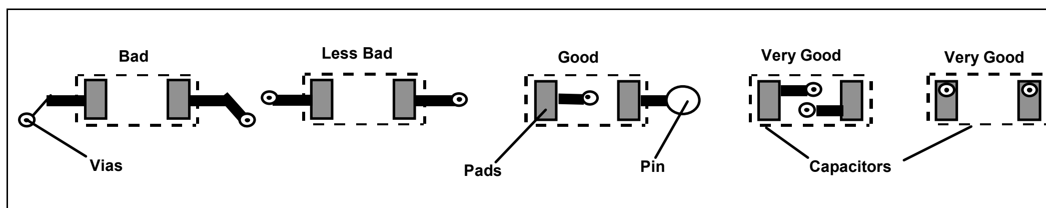
8.9.2 Location of High-Frequency Decoupling

A system designer for Intel Xeon processors MP and Intel Xeon processors MP with up to 2-MB L3 cache on the 0.13 micron process should properly design for the high-frequency decoupling. High-frequency decoupling should be placed as close to the power pins of the processor as physically possible. Use both sides of the board if necessary for placing components in order to achieve the optimum proximity to the power pins. This is vital as the inductance of the board's metal plane layers could cancel the usefulness of these low inductance components.

Another method to lower the inductance that should be considered is to shorten the path from the capacitor pads to the pins that it is decoupling. If possible, place the vias connecting to the planes within the pad of the capacitor. If this is not possible, keep the traces as short and wide as is

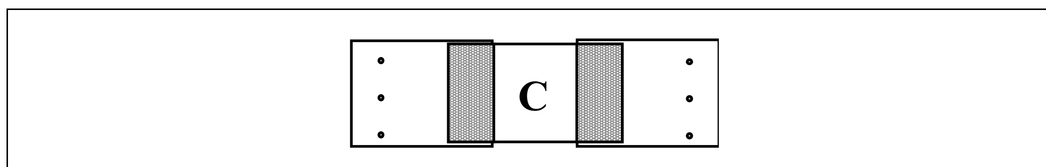
feasible. Possibly one or both ends of the capacitor can be connected directly to the pin of the processor without the use of via. Even if simulation results look good, these practical suggestions can be used to create an even better decoupling situation where they can be applied in layout. Figure 8-3 illustrates these concepts.

Figure 8-3. 1206 Capacitor Pad and Via Layouts



If polymer capacitors or large ceramics are being used, avoid the loss of the low ESL characteristic by connecting via patterns as wide as the capacitor with multiple via holes per connection, as shown in Figure 8-4.

Figure 8-4. Connections to Via Patterns



8.9.3 Location of Bulk Decoupling

The location of bulk capacitance is not as critical as the high-frequency decoupling since more inductance is already expected for these components. However, to achieve better performance, good placement of these components will affect the transient response of the system for the better, as shown in simulation. In addition to the bulk capacitors on the voltage converter module, which are electrically behind the inductance of the converter pins, several bulk capacitors need to be placed close to the processor socket.

8.9.4 Decoupling Recommendation

Intel recommends that the baseboard design incorporates at least nine 560 µF OSCON bulk capacitors and twenty 22 µF ceramic capacitors per processor. The bulk capacitors should be placed, half on one side of the processor and half on the other as close to the processor package as the keep-out zone allows. One quarter of the ceramic capacitors should be placed on one side of the processor, one quarter on the other side, and half in the processor cavity using both sides of the board. See Section 8.10 for placement options. Check with the voltage regulator designer for optimal choice of bulk capacitors. Some very high switching regulators are better served by replacing the OSCON bulk capacitors with additional high-frequency ceramics. Table 8-1 provides the parameters for bulk and high-frequency capacitors.

8.10 Component Placement and Modeling

Intel recommends using simulation to design and verify Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process-based systems. The models in the following sections can be used to piece together a complete base board spice circuit.

The maximum distance between each processor and its voltage regulator module should not be greater than 1.5 inches. To be more specific, the distance between the facing edges of the VRM connector and the socket should be no more than 0.5 inch. The bulk capacitors can be placed close to and the high-frequency capacitors should be placed next to the processors. Distribute the bulk and high-frequency capacitors equally on both sides of the socket where the power/ground pins are located (the east and west side).

Intel Xeon processor MP and Intel Xeon processor MP with up to 2-MB L3 cache on the 0.13 micron process sockets have 603-pins with 50-mil pitch. The routing of the signals, power and ground pins will require creation of lots of vias. These vias cause a “Swiss cheese” effect in the power and ground planes beneath the processor resulting in increased inductance of these planes. It is recommended to place as many high-frequency capacitors as possible inside the cut out of the processor socket. The remaining high-frequency capacitors should be placed next to the processor, specifically the power/ground pins.

Processors should be placed with respect to the voltage regulator modules and bulk decoupling capacitors such that current to one processor does not flow in the same path as that of any other processor.

8.10.1 Component Models

Acquire component models from their respective manufacturers. Intel cannot guarantee the specifications of another manufacturer's components. This section contains some of the models developed by Intel for internal simulations.

Table 8-1. Various Component Models Used at Intel (Not Vendor Specifications)

Component of Simulation	ESR (Ω)	ESL (nH)
0.1 µF Ceramic 0603 package	0.006	0.63
1 µF Ceramic 0805 package	0.080	0.702
10.0 µF Ceramic 1206 package	0.010	0.880
22.0 µF Ceramic 1210 package	0.010	1.1
560 µF OSCONS	0.012	3.1

8.10.2 Processor Socket-Package Lump Model

Figure 8-5 shows the lump electrical model for the high-frequency baseboard capacitors, the processor socket, and the processor package. Figure 8-6 shows a physical pictorial of the model. This model serves as a sub-circuit for the following baseboard models. Table 8-2 lists the model's component values. L2 and L6 refer to the inductance and resistance of the power plane next to the processor socket area, i.e., “before Swiss cheese” area. L3 and L5 refer to the inductance and resistance of the power plane within the processor socket area, i.e., “after Swiss cheese” area. The

inductance and resistance of the power plane between voltage regulator and bulk capacitors and the processor is shown as L1 and L7. A current step from 0 A to 55 A should be applied with a rise time of 308 ns or use the PWL values.

Figure 8-5. Processor Lump Model Schematic

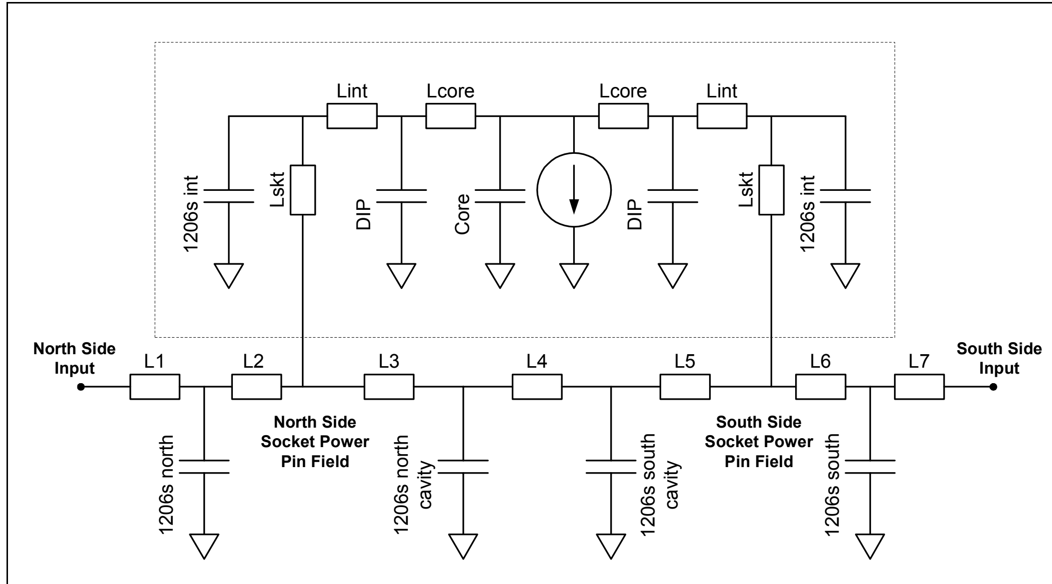


Table 8-2. Processor Lump Model Component Values

Component	Description	Value		
		Resistance	Inductance	Capacitance
1206s North/South	Five 22 µF MLCC	10 mΩ / 5	1.1 nH / 5	5 * 22 µF
1206s North/South Cavity	Five 22 µF MLCC	10 mΩ / 5	1.1 nH / 5	5 * 22 µF
1206s int	Interposer MLCC	833 µΩ	45 pH	120 µF
DIP Capacitors	Package Capacitors	270 µΩ	2.35 pH	36 µF
Core capacitors	Die Capacitance	146 µΩ	0	541 nF
L1	North Side Input	170 µΩ	23 nH	-
L2	North Side Pin Field Input	150 µΩ	23 nH	-
L3	North Side Cavity Input	120 µΩ	18 nH	-
L4	Cavity	130 µΩ	20 nH	-
L5	South Side Cavity Input	120 µΩ	18 nH	-
L6	South Side Pin Field Input	150 µΩ	23 nH	-
L7	South Side Input	170 µΩ	23 nH	-
Lskt	Socket Impedance	326 µΩ	24 pH	-
Lint	Interposer Impedance	125 µΩ	12 pH	-
Lcore	Package Impedance	25 µΩ	1 pH	-

Figure 8-6. Processor Lump Model Drawing

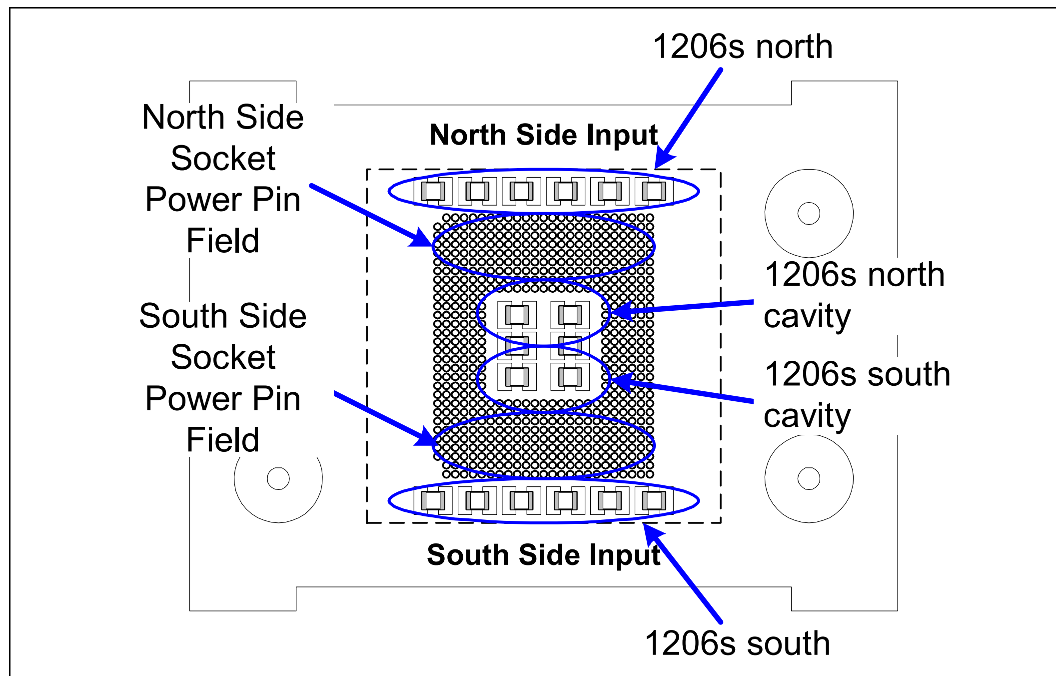


Figure 8-7. “Row” Pattern with Voltage Regulator Module

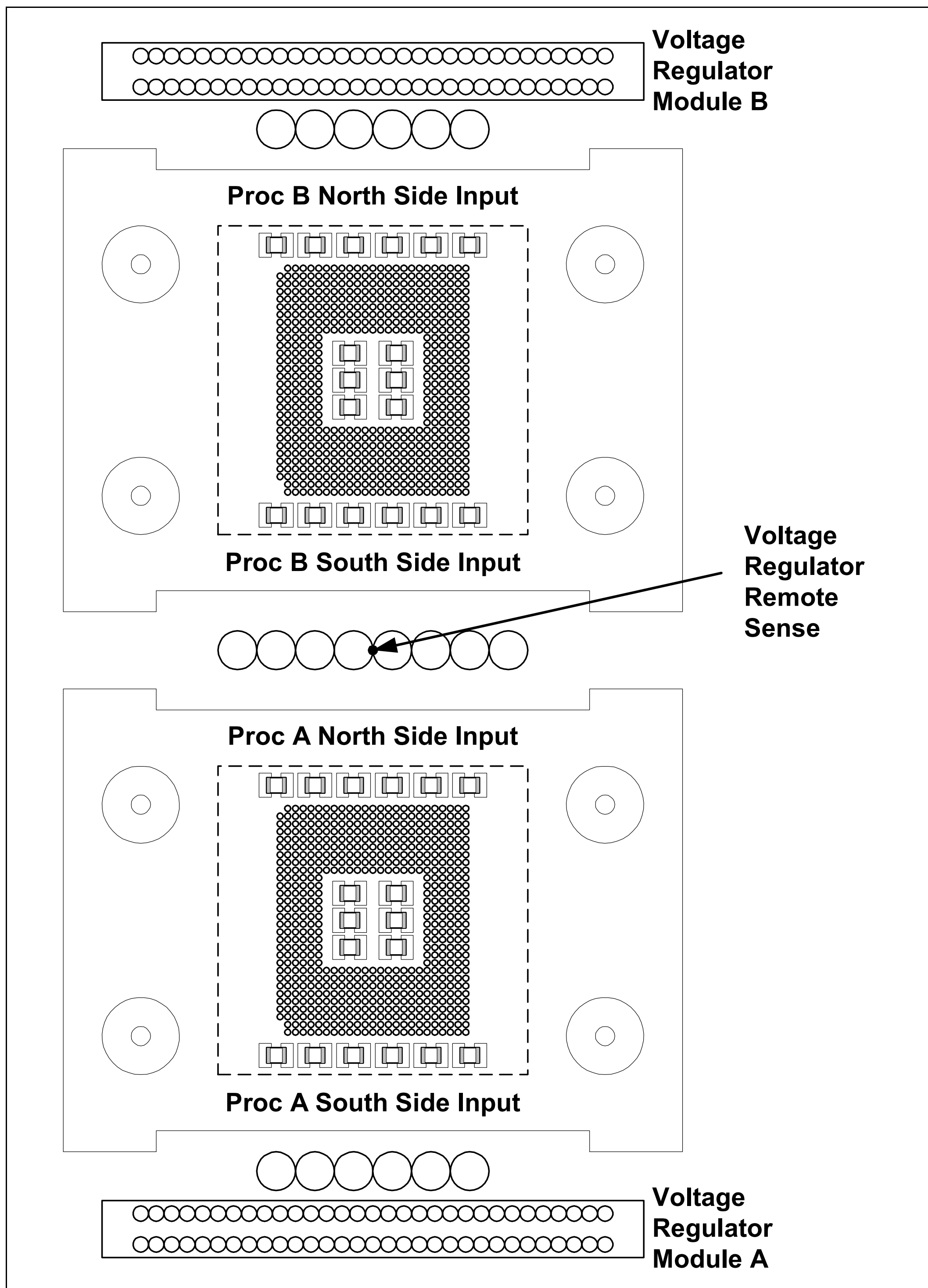


Figure 8-8. “Row” Pattern with Voltage Regulator Module Schematic

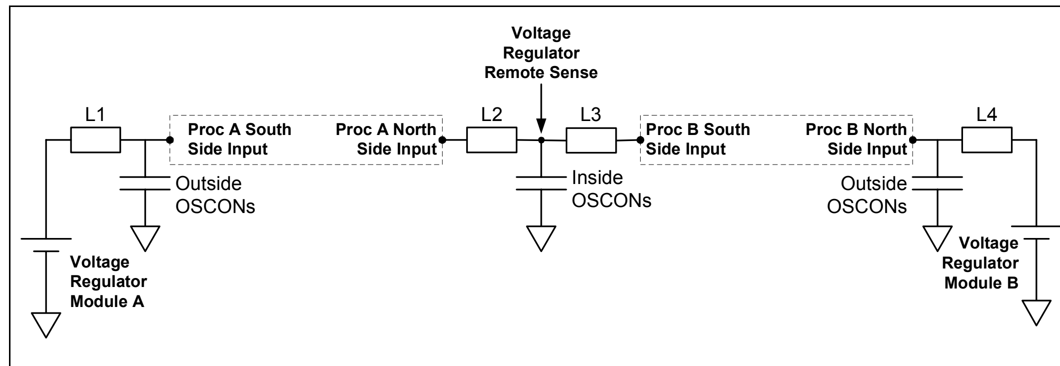


Table 8-3. “Row” Pattern with Voltage Regulator Module Schematic Values

Component	Description	Values		
		Resistance	Inductance	Capacitance
Outside OSCONs	Bulk Capacitors	12 mΩ / 6	3.1 nH / 6	6 × 560 µF
Inside OSCONs	Bulk Capacitors	12 mΩ / 8	3.1 nH / 8	8 × 560 µF
L1	VRM A – Proc A south	75 µΩ	20 pH	-
L2	Proc A north - sense	75 µΩ	20 pH	-
L3	Proc B south - sense	75 µΩ	20 pH	-
L4	VRM B – Proc B north	75 µΩ	20 pH	-

8.10.3 Multi-Processor Component Placement and Models

This section provides recommended placement diagrams and lump electrical model schematics for four processor systems. Baseboard impedances are estimates based on minimum copper weight requirements.

Figure 8-9. “Square” Pattern with Voltage Regulator Module

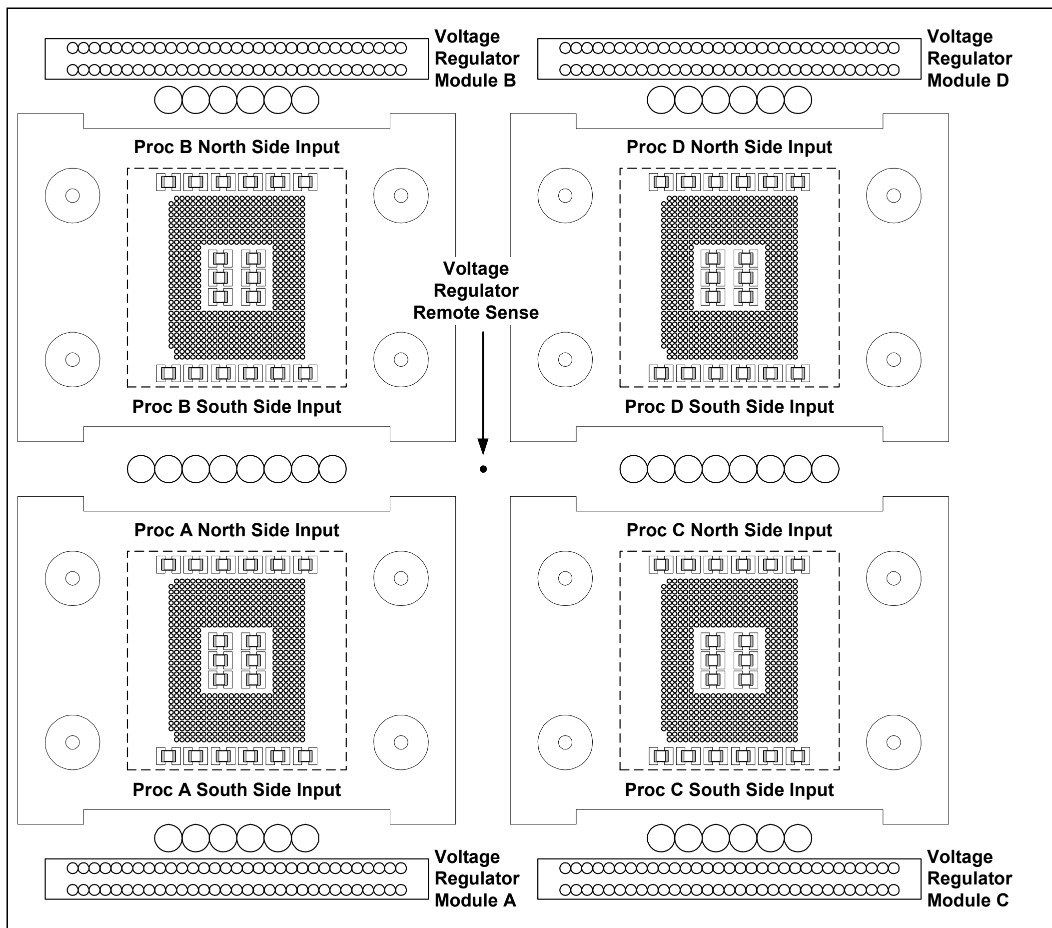


Figure 8-10. "Square" Pattern with Voltage Regulator Module Schematic

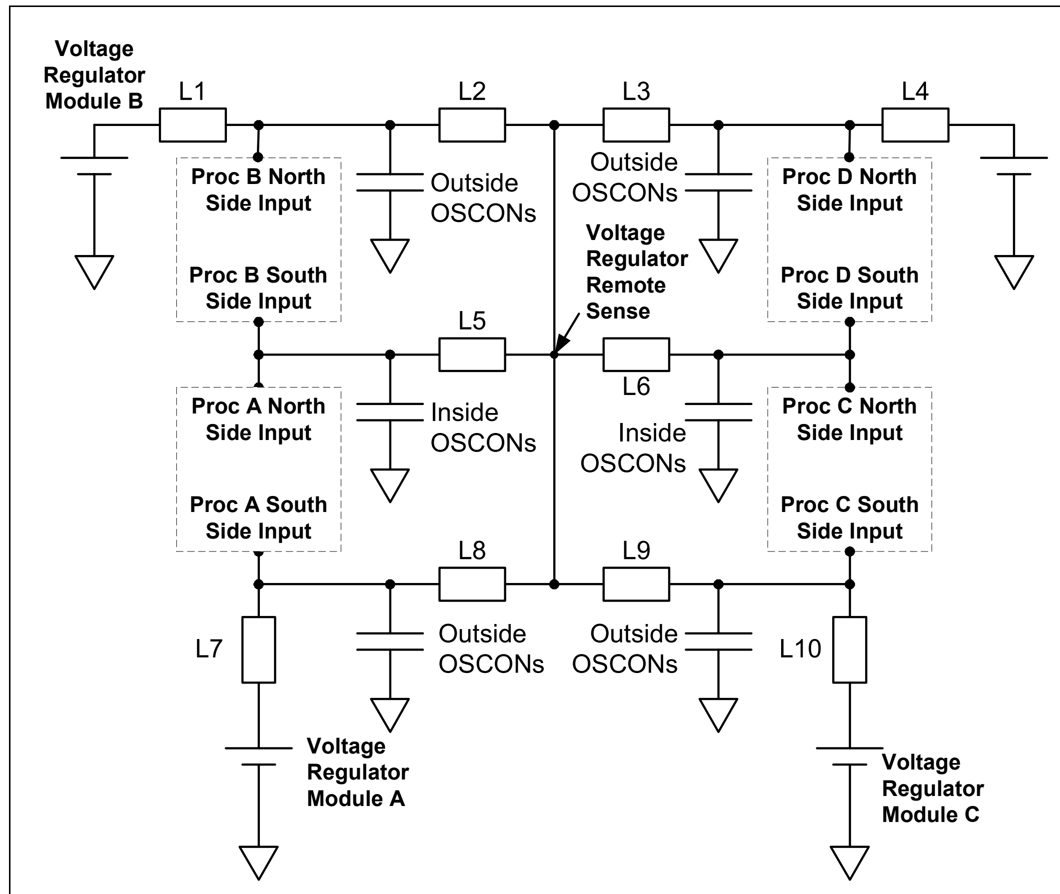


Table 8-4. "Square" Pattern with Voltage Regulator Module Schematic Values

Component	Description	Values		
		Resistance	Inductance	Capacitance
Outside OSCONs	Bulk Capacitors	12 mΩ / 6	3.1 nH / 6	6 × 560 µF
Inside OSCONs	Bulk Capacitors	12 mΩ / 8	3.1 nH / 8	8 × 560 µF
L1	VRM B – Proc B north	75 µΩ	20 pH	-
L2	Proc B north - sense	300 µΩ	80 pH	-
L3	Proc D north - sense	300 µΩ	80 pH	-
L4	VRM D – Proc D north	75 µΩ	20 pH	-
L5	Proc A north - sense	300 µΩ	80 pH	-
L6	Proc C north - sense	300 µΩ	80 pH	-
L7	VRM A – Proc A south	75 µΩ	20 pH	-
L8	Proc A south - sense	300 µΩ	80 pH	-
L9	Proc C south - sense	300 µΩ	80 pH	-
L10	VRM C – Proc C south	75 µΩ	20 pH	-

Figure 8-11. “Row” Pattern with Voltage Regulator Module

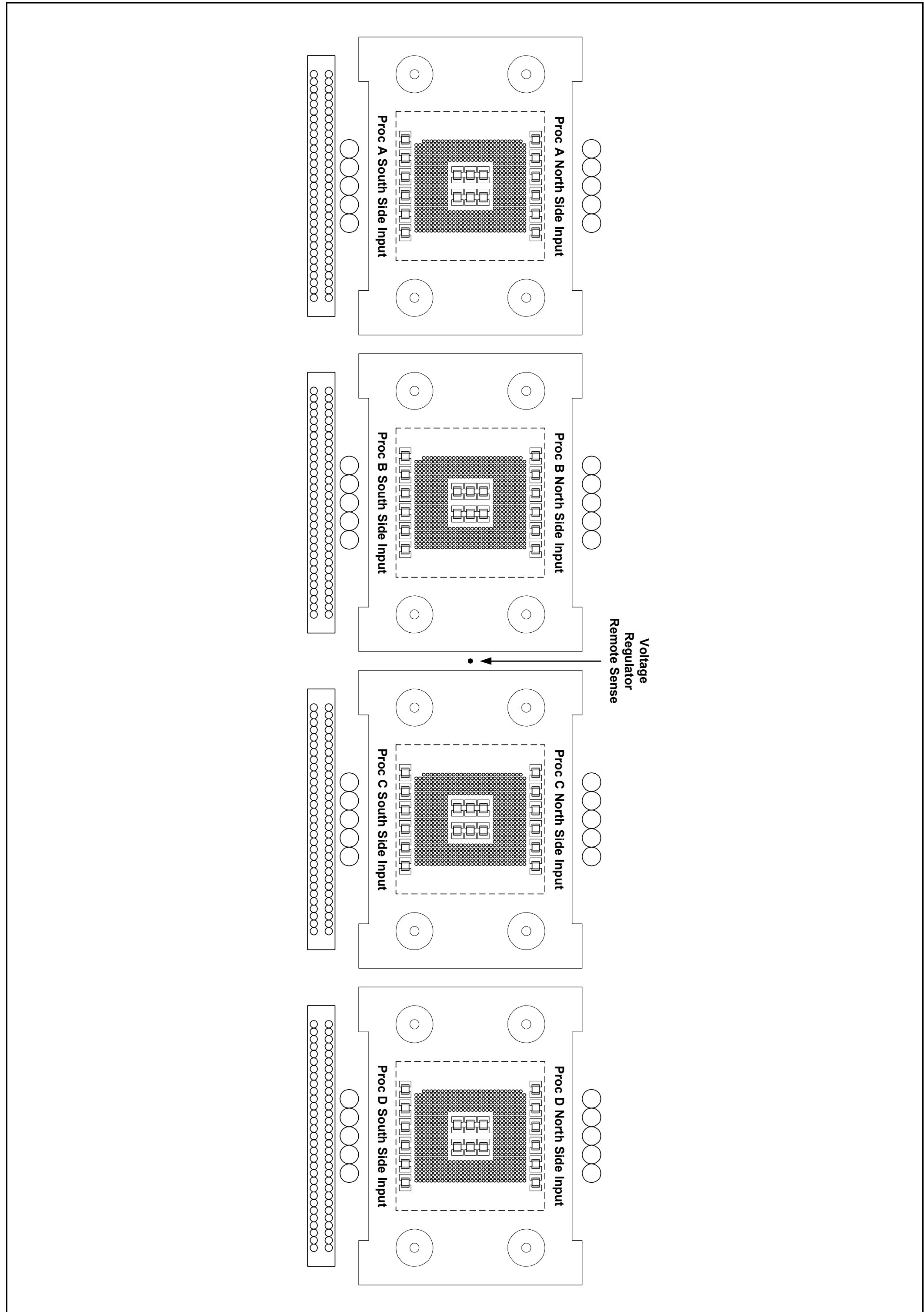


Figure 8-12. “Row” Pattern with Voltage Regulator Module Schematic

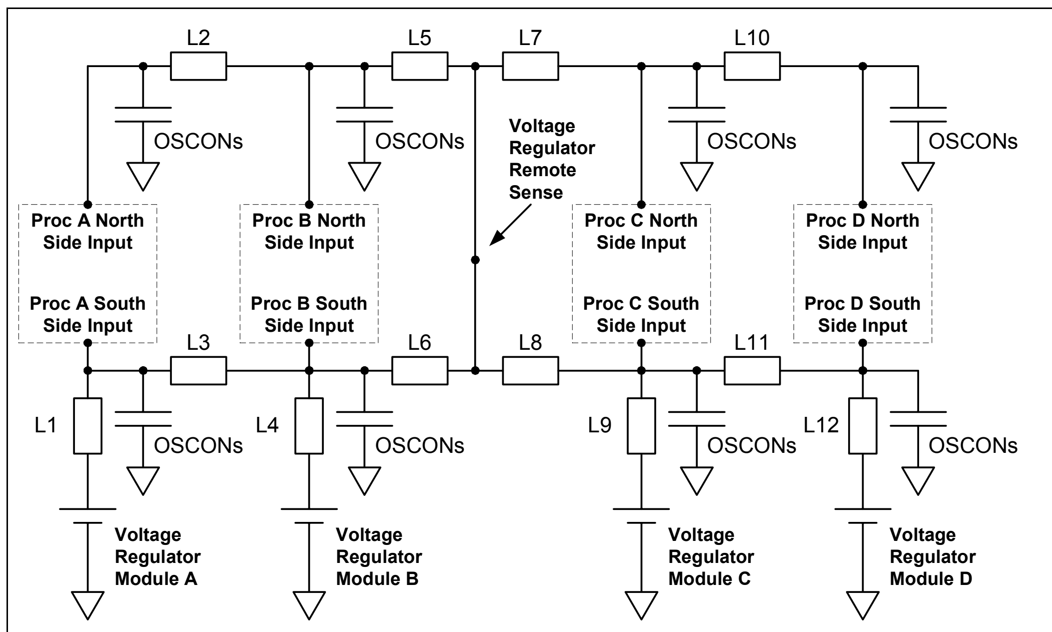


Table 8-5. “Row” Pattern with Voltage Regulator Module Schematic Values

Component	Description	Values		
		Resistance	Inductance	Capacitance
OSCONs	Bulk Capacitors	12 mΩ / 5	3.1 nH / 5	5 × 560 µF
L1	VRM A – Proc A south	75 µΩ	20 pH	-
L2	Proc A north - Proc B north	600 µΩ	160 pH	-
L3	Proc A south - Proc B south	600 µΩ	160 pH	-
L4	VRM B – Proc B south	75 µΩ	20 pH	-
L5	Proc B north - sense	300 µΩ	80 pH	-
L6	Proc B south - sense	300 µΩ	80 pH	-
L7	Proc C north - sense	300 µΩ	80 pH	-
L8	Proc C south - sense	300 µΩ	80 pH	-
L9	VRM C – Proc C south	75 µΩ	20 pH	-
L10	Proc C north - Proc D north	600 µΩ	160 pH	-
L11	Proc C south - Proc D south	600 µΩ	160 pH	-
L12	VRM D – Proc D south	75 µΩ	20 pH	-

8.11 Validation Testing

The processor $V_{CCSENSE}$ and $V_{SSSENSE}$ pins should be routed to vias. The vias should be as close to the socket pins as possible and should be connected with a low impedance trace. As these signals provide measurement points to verify adherence to the processor's V_{CC} specifications, the vias need to be accessible to measurement equipment.

Intel recommends the following guideline when measuring the transients on the processor V_{CC}. The measurement should be done across the V_{CC} and V_{SS} pins on processor socket. Use an oscilloscope with 100 MHz bandwidth, 1.5 pF maximum probe capacitance, and 1 mΩ minimum impedance. The maximum length of ground wire on the probe should be less than 5 mm. Ensure external noise from the system is not coupled in the scope probe. Some probes have a very significant level of inherent noise. Attempt to minimize noise by investigating different probes. Use a differential probe to make the voltage measurements. The bandwidth of the probe should be no less than the oscilloscope. Ensure all connections from oscilloscope to motherboard pin are good and have a very low contact resistance.

8.12 Generating and Distributing GTLREF[3:0]

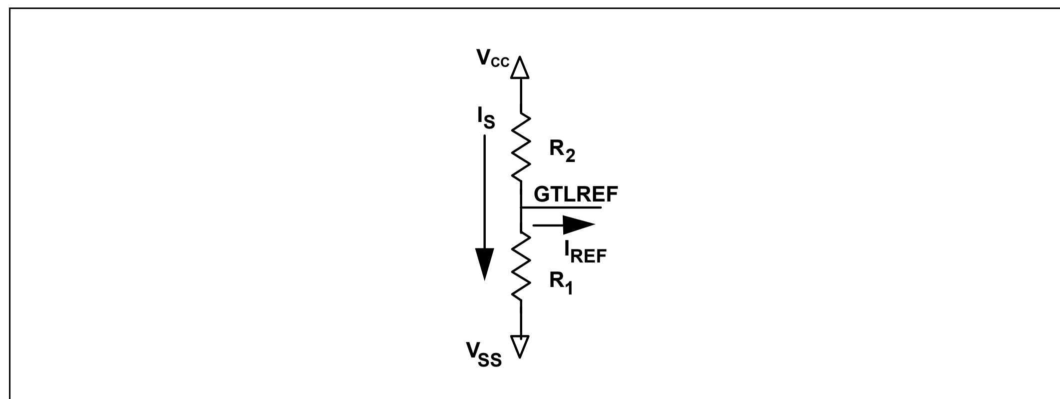
GTLREF[3:0] are low current inputs (less than 15 µA each) to the differential receivers within each of the components on the AGTL+ bus. A simple voltage divider can generate GTLREF[3:0]. The GTLREF[3:0] inputs need to meet the 2% specification.

Equation 8-1 uses $R_1 = 2 \times R_2$ to generate a GTLREF set at a nominal value of 2/3 V_{CC}. Figure 8-13 illustrates using 1% resistors to generate the GTLREF specification of 2/3 V_{CC} ± 2%.

Equation 8-1. Creating GTLREF of 2/3 V_{CC}

$$V_{REF} = V_{CC} \times \frac{R_1}{R_1 + R_2} = V_{CC} \times \frac{2 \times R_2}{2 \times R_2 + R_2} = \tfrac{2}{3} V_{CC}$$

Figure 8-13. GTLREF



R_1 and R_2 should be small enough values that the current drawn by the GTLREF inputs (I_{REF}) is negligible versus the current caused by R_2 and R_1.

A complete analysis of this circuit's currents into and out of the center node, as in Equation 11, will provide the final GTLREF of the circuit. n is the number of I_{REF} inputs supplied by the divider.

Equation 8-2. Node Analysis

$$I(R_2) = I(R_1) + n \times I_{\mathrm{REF}}$$

Plugging in for the currents and rearranging gives:

Equation 8-3. Node Analysis in Terms of Voltage

$$\frac{V_{\mathrm{CC}} - GTLREF}{R_2} - \frac{GTLREF}{R_1} = n \times I_{\mathrm{REF}}$$

Which leads to:

Equation 8-4. Solving for GTLREF

$$GTLREF = \frac{V_{\mathrm{CC}}/R_2 - n \times I_{\mathrm{REF}}}{1/R_2 + 1/R_1}$$

The worst case GTLREF should be analyzed with IREF at the maximum and minimum values determined for the number of loads being supplied. If the number of loads can change from model to model because of upgrades, this should be taken into account as well. Analyze Equation 8-4 with R_1 and R_2 at the extremes of their tolerance specifications.

8.12.1 GTLREF [3:0]

Intel recommends two voltage dividers for each processor and one for the chipset component. Assume a maximum of 15 µA of leakage current per load. Note that these leakage currents can be positive or negative.

The following discussion illustrates using a single voltage divider to support two GTLREF Loads assuming V_{CC} of 1.7 V. Using 1% resistors for the voltage divider in Figure 8-14 make R_1 a 100 Ω resistor, and use 49.9 Ω for R_2. This creates a static usage of 10.7 mA (1.7 V / 149.9 Ω) per voltage divider. After looking at all combinations of R_1 and R_2 (above and below tolerance) and I (± 30 µA), the worst case solution for Equation 8-4 can be found with I_{REF} at 30 µA, R_1 at the low end of its tolerance specification (99 Ω), and R_2 at the high end of its tolerance specification (50.4 Ω). This yields:

Equation 8-5. Resistor Tolerance Analysis

$$V_{REF} = \frac{1.7/50.4 - .000030}{1/50.4 + 1/99} = 1.1255V$$

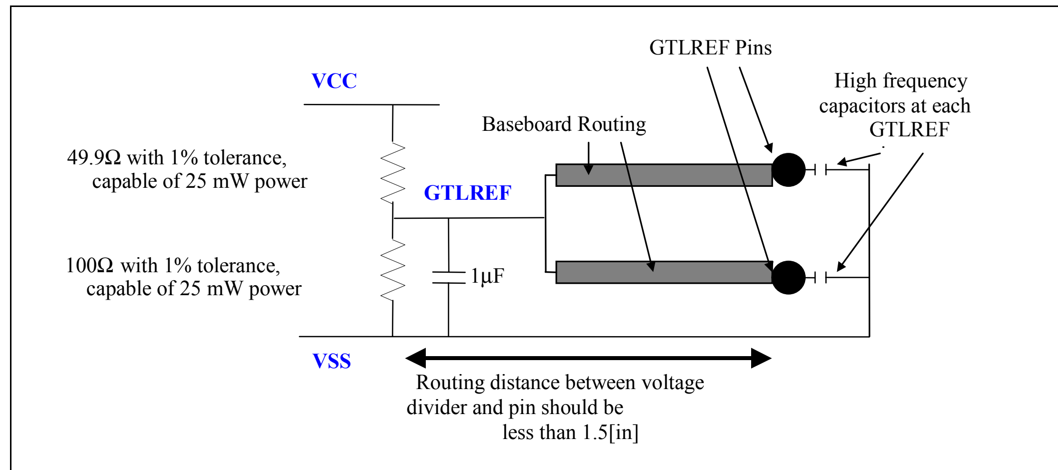
Since the target of 2/3 of V_{CC} is 1.133 V, this setting is within 0.7% of the 2/3 point and satisfies the 2% specification. A spreadsheet program allows the reader to easily verify the other corners. Varying over its tolerance range has minimal effect.

These values chosen for R_1 and R_2 have additional benefits: The parallel combination terminates the GTLREF line to 33 Ω. These generally available resistance values reduce resistor cost.

Decouple GTLREF[3:0] at each pin with a 220 pF capacitor to V_{SS}. Decoupling GTLREF to V_{SS} at the voltage dividers with a 1 µF capacitor may further enhance the ability for GTLREF to track V_{CC}.

When routing GTLREF to the pins, use a 30-50 mil trace (the wider the better) and keep it as short as possible (less than 1.5 inches). Also, keep all other signals at least 20 mils away from the GTLREF trace. This provides a low impedance line without the cost of an additional plane or island. Due to the placement of the GTLREF pins on the processor, it may not be possible or convenient to route all four pins from one voltage divider. It is acceptable to use more than one voltage divider with decoupling at each voltage divider and each pin.

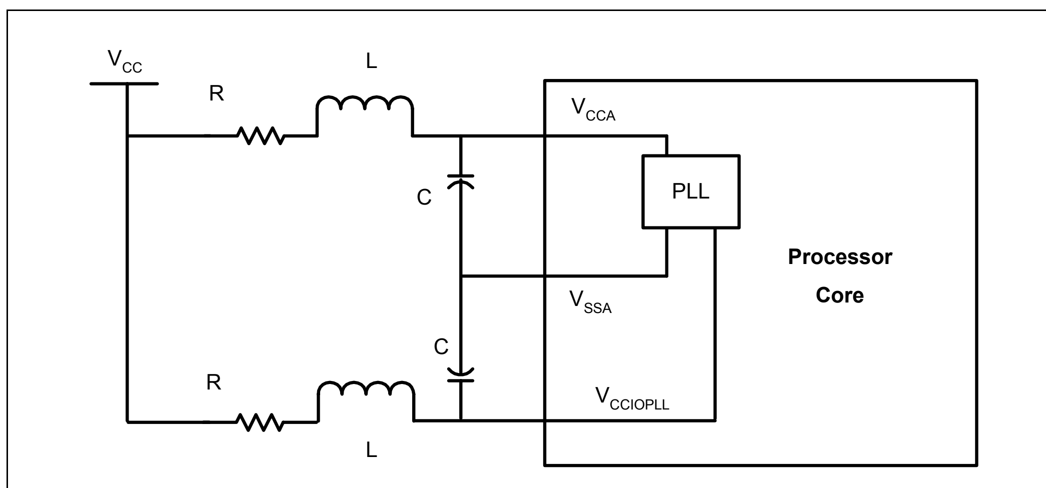
Figure 8-14. Suggested GTLREF Generation



8.13 Filter Specifications for V_{CCA}, $V_{CCIOPLL}$, and V_{SSA}

V_{CCA} and $V_{CCIOPLL}$ are power required by internal PLL. These powers are low passed filter V_{CC}. Intel Xeon processors MP have internal PLL clock generators, which are analog in nature and require quiet power supplies for minimum jitter. Jitter is detrimental to a system; it degrades external I/O timings, as well as internal core timings (i.e., maximum frequency). The general desired topology is shown in Figure 8-15. Not shown in the figure are the parasitics of connecting traces, circuits, and components.

Figure 8-15. Filter Topology



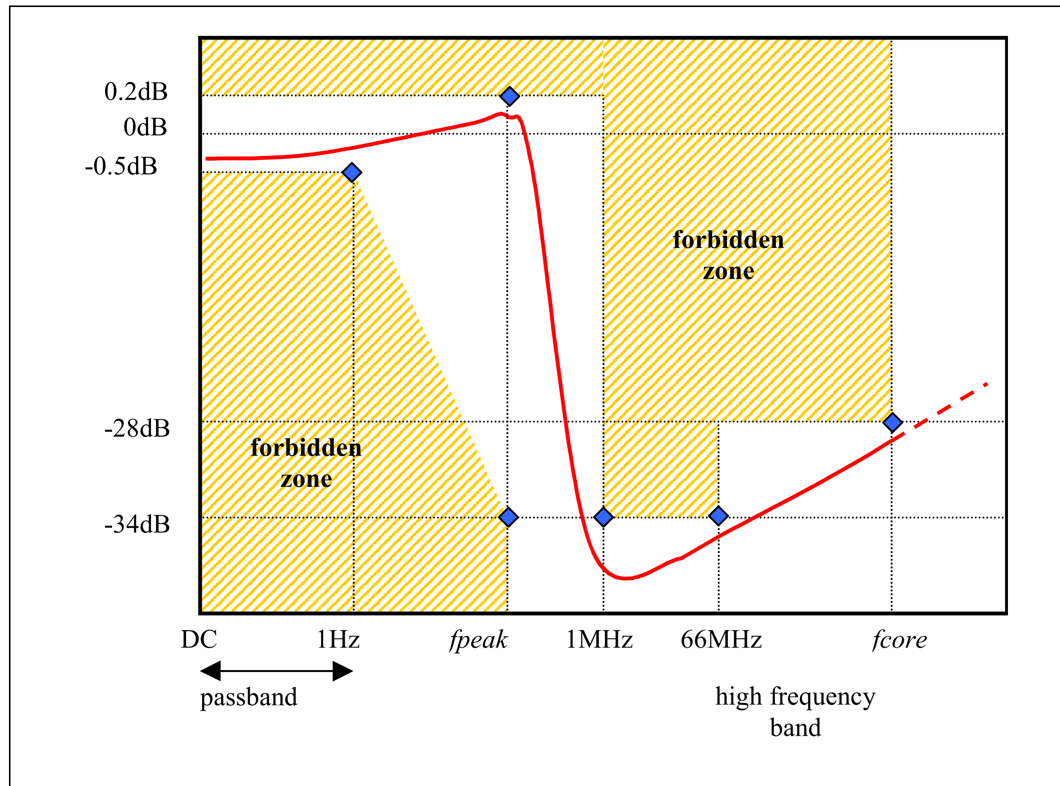
The function of the filter is two-fold. It protects the PLL from external noise through low-pass attenuation. It also protects the PLL from internal noise through high-pass filtering. In general, the low-pass description forms an adequate description for the filter. For simplicity we are addressing the recommendation for V_{CCA} filter design. The same characteristics and design approach is applicable for $V_{CCIOPLL}$ filter design.

The AC low-pass specification, with input at V_{CC} and output measured across the capacitor, is as follows:

- < 0.2 dB gain in pass band
- < 0.5 dB attenuation in pass band < 1 Hz (see DC drop in next set of requirements)
- 34 dB attenuation from 1 MHz to 66 MHz
- 28 dB attenuation from 66 MHz to core frequency

The filter specification (AC) is graphically shown in Figure 8-16.

Figure 8-16. Filter Specifications



NOTES:
1. Diagram not to scale.
2. No specification for frequencies beyond fcore.
3. Fpeak, if existent, should be less than 0.05 MHz.

Other requirements:

- Use shielded type inductor to minimize magnetic pickup.
- Filter should support at least 30 mA DC current.
- DC voltage drop from V_{CC_CPU} to processor interposer pin V_{CCA} should be < 33 mV, which in practice implies series R < 1.1 Ω; also means pass band (from DC to 1 Hz) attenuation < 0.5 dB for V_{CC_CPU} = 1.1 V, and < 0.35 dB for V_{CC_CPU} = 1.7 V.

The following tables list some recommended components for the filter.

Table 8-6. Component Recommendation - Inductor

Part Number (Ref Designator)	Value	Tol	SRF	Rated I	DCR
TDK MLF2012A4R7KT (L1)	4.7 µH	10%	35 MHz	30 mA	0.56 Ω (1 Ω max)
Murata LQG21N4R7K10 (L2)	4.7 µH	10%	47 MHz	30 mA	0.7 Ω (± 50%)
Murata LQG21C4R7N00 (L3)	4.7 µH	30%	35 MHz	30 mA	0.3 Ω max

Table 8-7. Component Recommendations - Capacitor

Part Number	Value	Tolerance	ESL	ESR
Kemet T495D336M016AS	33 µF	20%	2.5 nH	0.225 Ω
AVX TPSD336M020S0200	33 µF	20%	TBD	0.2 Ω

To satisfy damping requirements, total series resistance in the filter (from V_{CC_CPU} to the top plate of the capacitor) must be at least 0.35 Ω. It includes the DCR of the inductor and any resistance (routing or discrete components) between V_{CC_CPU} and capacitor top plate. Keep the routing short and wide. If the total is less than 0.35 Ω, add a discrete resistor to make up the difference. For example, if the selected filter inductor has a minimum of 0.1 Ω DCR, and a negligible routing resistance (less than 10 mΩ), add a discrete resistor R_1 of approximately 0.3.Ω.

The total maximum resistance in each route cannot be more than 1.1 Ω as defined from V_{CC_CPU} (more specifically, the baseboard via that connects the PLL filter to the V_{CC_CPU} plane) to the processor interposer pin. Refer to Figure 8-17 and Figure 8-18 for recommended filter circuits. This path includes the total trace resistance (denoted “R-TRACE” in the following figures), discrete resistor (if needed), inductor DCR, and Socket 603 resistance (0.025 Ω). It is important to note that “R-TRACE” includes the total trace resistance between V_{CC_CPU} and the processor socket pin, but is represented in the figures as a single resistor to simplify the circuit representation.

Other routing requirements:

- C should be as close as possible to V_{CCA} and V_{SSA} pins in the socket (typically < 0.02 Ω per route).
- Route away from clocks and fast switching signals
- V_{CCA} route should be parallel and next to V_{SSA} route (to minimize loop area)
- $V_{CCIOPLL}$ route should be parallel and next to V_{SSA} route (to minimize loop area)
- L should be close to C; any routing resistance should be inserted between V_{CC_CPU} and L
- Any discrete R should be inserted between V_{CC_CPU} and L

Figure 8-17. Implementation 1 Using Discrete R

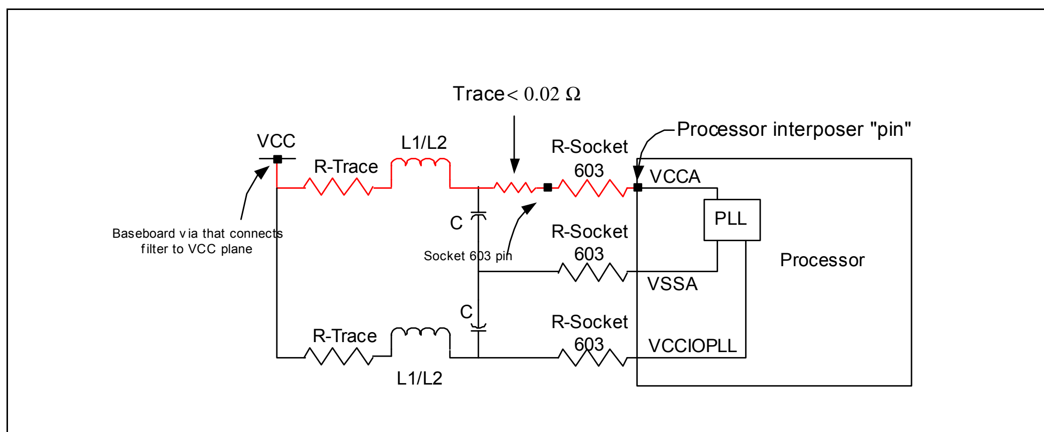
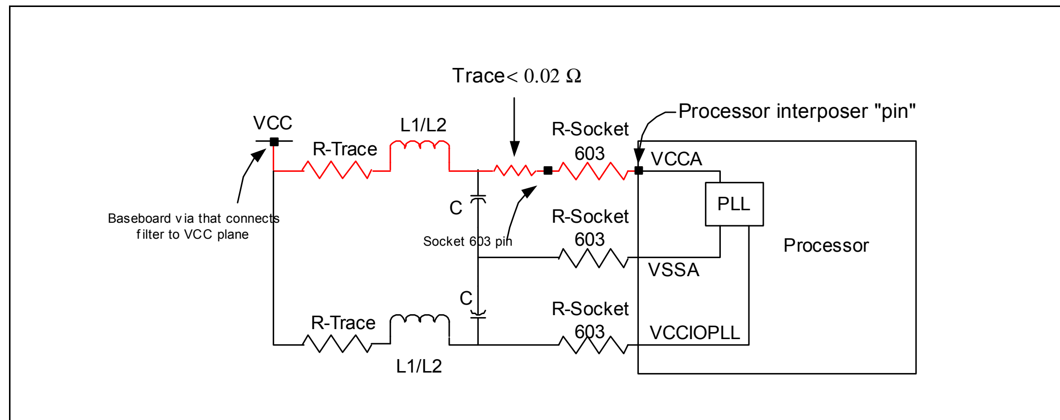


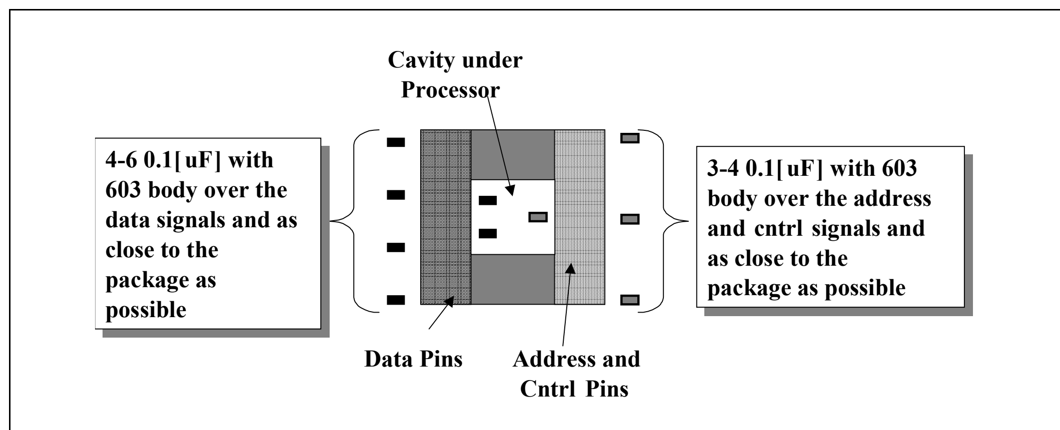
Figure 8-18. Implementation 2 No Discrete R



In addition, high-frequency decoupling may be required for signal integrity. System boards designed using striplines with V_{CC_CPU} and V_{SS} references should not require high-frequency decoupling beyond that recommended above. For systems using microstrip configurations a return path discontinuity will exist between the processor and the baseboard due to the baseboard traces having only one reference plane. These systems should distribute decoupling capacitors as shown in Table 8-1 and described as follows:

- Four minimum, six preferred 1 µF capacitors with 0805 packages distributed evenly over the data lines.
- Three minimum, four preferred 1 µF capacitors with 0805 packages distributed evenly over the address and control lines.
- All capacitors placed as close to the processor package as the keep-out zone allows.

Figure 8-19. Example of Decoupling for a Microstrip Baseboard Design

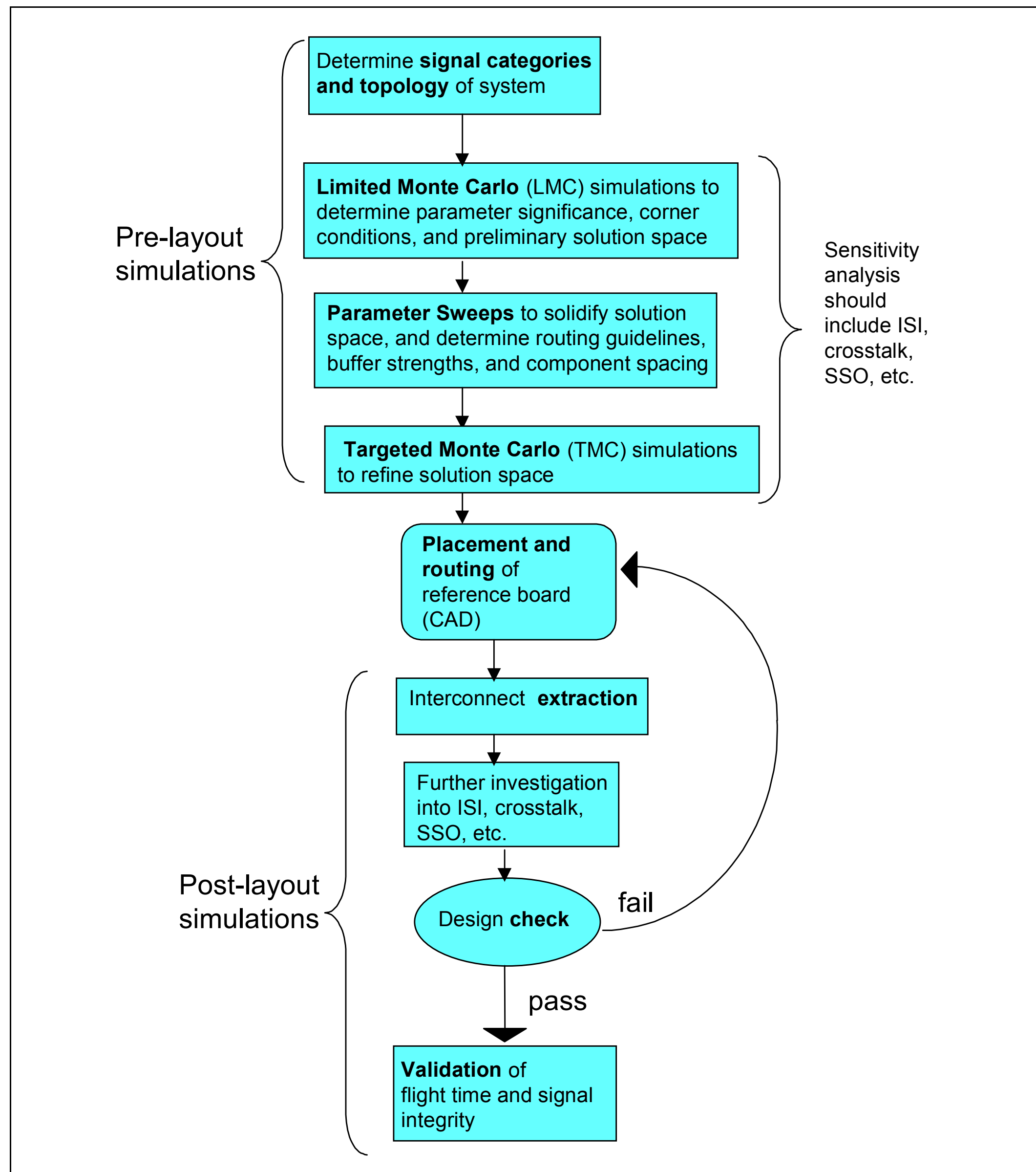


This page is intentionally left blank.

Methodology for Determining Topology and Routing Guidelines 9

This section documents the simulation methodology that was used to derive the topology and routing guidelines presented in this design guide.

Figure 9-1. Simulation Methodology Flowchart



The design process should begin with an initial timing analysis and topology definition. Pre-layout analog simulations should be performed. These pre-layout simulations will help define routing rules prior to placement and routing. After routing, the interconnect database can be extracted and post-layout simulations can be performed to refine the timing and signal integrity analysis. The analog simulations should be validated when actual systems become available. Target measurements of the high-significance parameters to ensure they fall within simulated boundaries.

Pre-layout simulations provide a detailed picture of the working solution space that meets flight time and signal quality requirements. By basing board layout guidelines on the solution space, the iterations between layout and post-layout simulation can be reduced.

Intel recommends running simulations at the device pads for signal quality and for timing analysis. However, simulation results at the device pins may be used later to correlate simulation performance against actual system measurements.

Pre-layout analysis includes a sensitivity analysis using parametric sweeps. Parametric sweep analysis involves varying one or two system parameters while all others such as driver strength, package Z_0, and S_0 are held constant. This way, the sensitivity of the proposed bus topology to varying parameters can be analyzed systematically. Sensitivity of the bus to flight times and signal quality should be covered. Suggested sweep parameters include trace lengths, termination resistor values, and any other factors that may affect flight time, signal quality, and feasibility of layout. Minimum flight time and worst signal quality are typically analyzed using fast I/O buffers and interconnect. Maximum flight time is typically analyzed using slow I/O buffers and slow interconnects. However, fast I/O buffers and various baseboard and package combinations have been found to have violating signal quality. It is advisable to perform some level of Monte Carlo analysis, which includes all possible parameters.

Outputs from each sweep should be analyzed to determine which regions meet timing and signal quality specifications. To establish the working solution space, find the common space across all the sweeps that result in passing timing and signal quality. The solution space should allow enough design flexibility for a feasible, cost-effective layout.

The effects of ISI and return path irregularities are difficult and tedious to simulate with 100% accuracy. Intel has found through experimentation and targeted simulations that these effects can have a significant impact on the primary signal. Due to the findings of this work, Intel has been able to more accurately predict the effects. Given the complexity of a quad-pumped source synchronous bus architecture, it was necessary for Intel to tighten the component timings and bus requirements in order to provide a viable routing solution space. Because of this, Intel strongly recommends adhering to the design guideline requirements and component specifications.

9.1 Timing Methodology

The timing equations used for both source synchronous and common clock parameters are derived in the following sections.

9.1.1 Source Synchronous

In a source synchronous bus the clock (or strobe) is driven from the same source as the signal it will sample. The strobe and the signal both propagate to the receiver via the PCB. The receiver then uses the strobe to sample the signal. This eliminates the need to account for worst-case flight times and, in theory, will significantly increase the maximum bus speed.

9.1.1.1 Setup Time

Figure 9-2 shows the setup timing diagram for a source synchronous bus design. Equation 9-1 gives the total loop equation derived from the timing diagram.

Equation 9-1. Source Synchronous Setup Time

$$T_{co}(strobe) + T_{flight}(strobe) - T_{setup} - T_{margin_setup} - T_{co}(data) - T_{flight}(data) = 0$$

- T_{co}(strobe)[(data)] is the driver delay of the strobe [data]
- T_{flight}(strobe)[(data)] is the flight time of the strobe [data] interconnect
- T_{setup} is the receiver's setup requirement
- T_{margin} is the available timing margin for the setup time

The loop equation can be simplified and solved for T_{margin_setup}. The equation can be broken into two parts, valid before and interconnect skew. Then, the setup margin can be determined.

Equation 9-2. Source Synchronous, Valid Before

$$T_{vb} = T_{co}(data)_{\max} - T_{co}(strobe)_{\min}$$

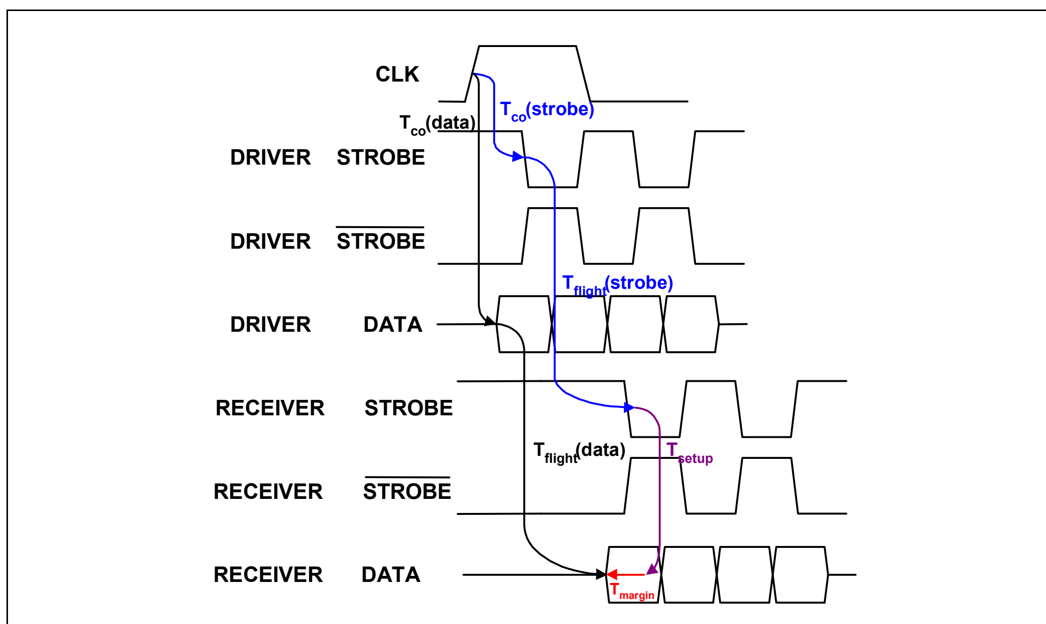
Equation 9-3. Source Synchronous, Interconnect Skew

$$T_{skew,\max} = T_{flight}(data)_{\max} - T_{flight}(strobe)_{\min}$$

Equation 9-4. Source Synchronous Setup Margin

$$T_{margin_setup} = -T_{vb,\min} - T_{skew,\max} - T_{setup}$$

Figure 9-2. Source Synchronous Timing Diagram for Setup Time



9.1.1.2 Hold Time

The hold timing diagram for a source synchronous bus design is shown in Figure 9-3. The total loop equation is derived from the hold timing diagram.

Equation 9-5. Source Synchronous Loop Equation for Hold Timing Diagram

$$T_{co}(data) + T_{flight}(data) - T_{margin_hold} - T_{hold} - T_{flight}(strobe) - T_{co}(strobe) = 0$$

- T_{co}(strobe)[(data)] is the driver delay of the strobe [data]
- T_{flight}(strobe)[(data)] is the flight time of the strobe [data] interconnect
- T_{hold} is the receiver's hold time requirement
- T_{margin} is the available timing margin for the hold time

The loop equation can be simplified and solved for T_{margin_hold}. The equation can be broken into two parts, valid before and interconnect skew. Then, the hold margin can be determined.

Equation 9-6. Source Synchronous, Valid After

$$T_{vb} = T_{co}(data)_{\max} - T_{co}(strobe)_{\min}$$

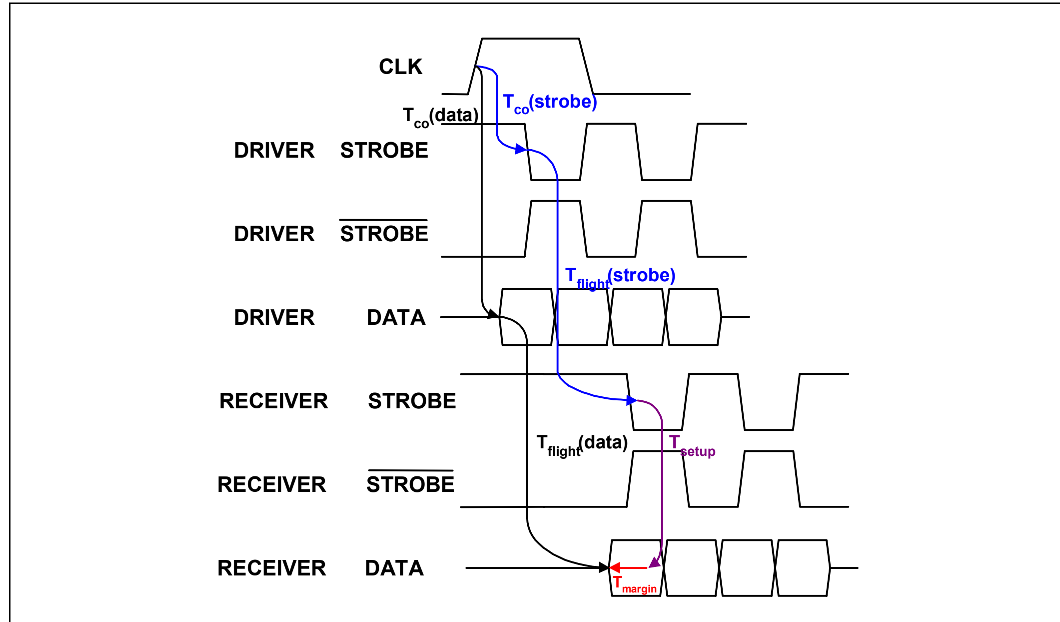
Equation 9-7. Source Synchronous, Interconnect Skew

$$T_{skew,\max} = T_{flight}(data)_{\max} - T_{flight}(strobe)_{\min}$$

Equation 9-8. Source Synchronous, Hold Margin

$$T_{margin_setup} = -T_{vb,\min} - T_{skew,\max} - T_{setup}$$

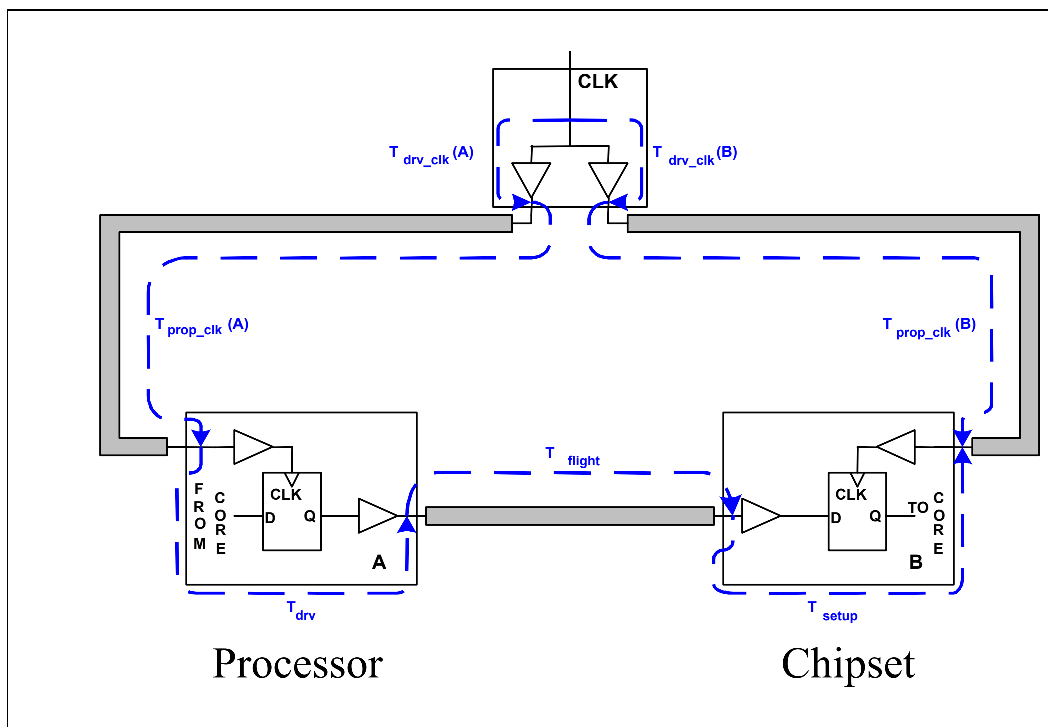
Figure 9-3. Source Synchronous Timing Diagram for Hold Time



9.1.2 Common Clock

A block diagram of a circuit that was used to develop the basic timing equations is shown in Figure 9-4.

Figure 9-4. Circuit Used to Develop the Common Clock Timing Equations



9.1.2.1 Setup Margin

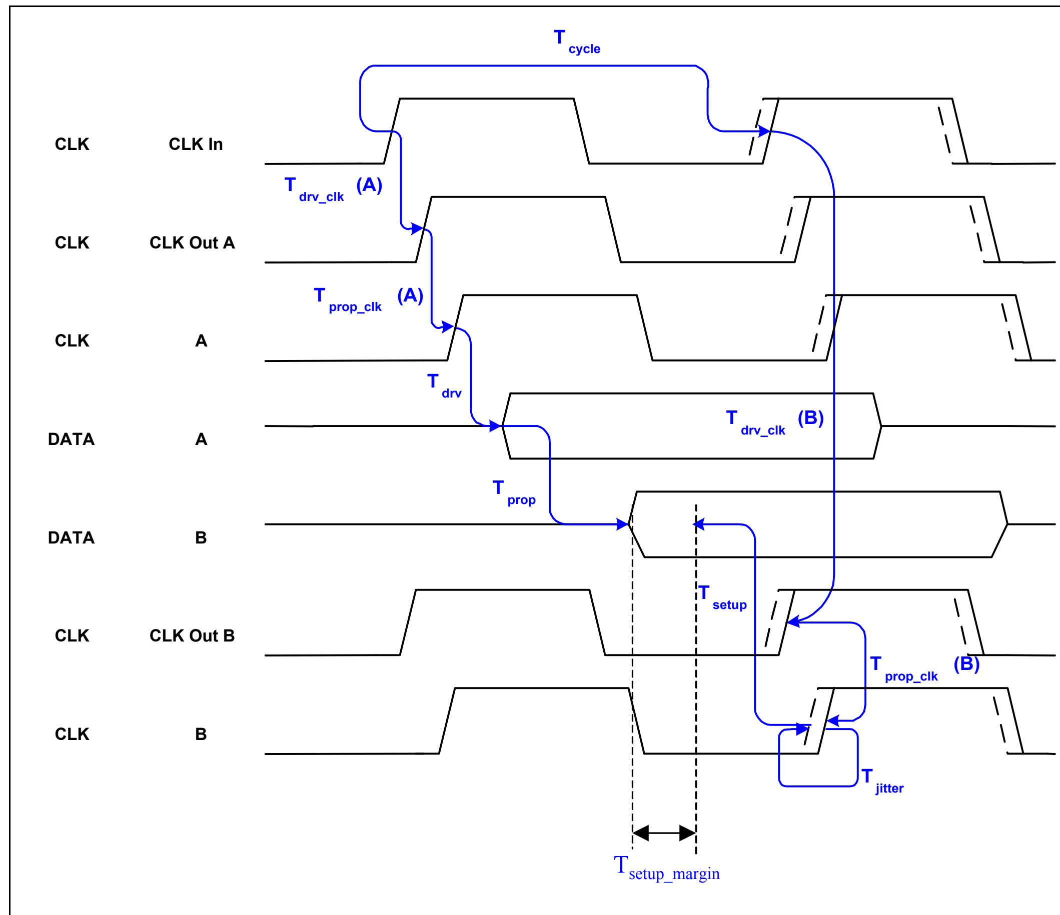
Figure 9-5 shows the setup timing diagram that was used to develop the final timing equations for the setup margin.

Equation 9-9. Common Clock Loop Equation

$$T_{cyvle} + T_{drv_clk}(B) + T_{prop_clk}(B) - T_{jitter} - T_{setup} - T_{margin} - T_{prop} - T_{drv} - T_{prop_clk}(A) - T_{drv_clk}(A) = 0$$

- T_{cycle} is the cycle time
- T_{drv_clk}(A)[(B)] is the delay of the clock buffer circuit connected to device A [B]
- T_{prop_clk}(A)[(B)] is the delay of the interconnect between the clock buffer and device A [B]
- T_{drv} is the delay of the output buffer for the data signal on device A (TCO)
- T_{PROP} is the interconnect delay between device A and B
- T_{setup} is the setup time required by the buffer
- T_{jitter} is the clock cycle-to-cycle jitter

Figure 9-5. Timing Diagram Used to Determine the Common Clock Setup Timing Equations



Equation 9-8 can be simplified by defining the clock delay and the clock skew as shown in Equation 9-9 and Equation 9-10. After simplification, Equation 9-8 is solved for the setup margin as shown in Equation 9-12.

Equation 9-10. Common Clock Delay

$$T_{clk} = T_{drv_clk} + T_{prop_clk}$$

Equation 9-11. Common Clock Skew

$$T_{skew_setup} = T_{clk}(A) - T_{clk}(B)$$

Equation 9-12. Common Clock Setup Margin

$$T_{margin_setup} = T_{cycle} - T_{drv} - T_{setup} - T_{prop} - T_{skew_setup} - T_{jitter}$$

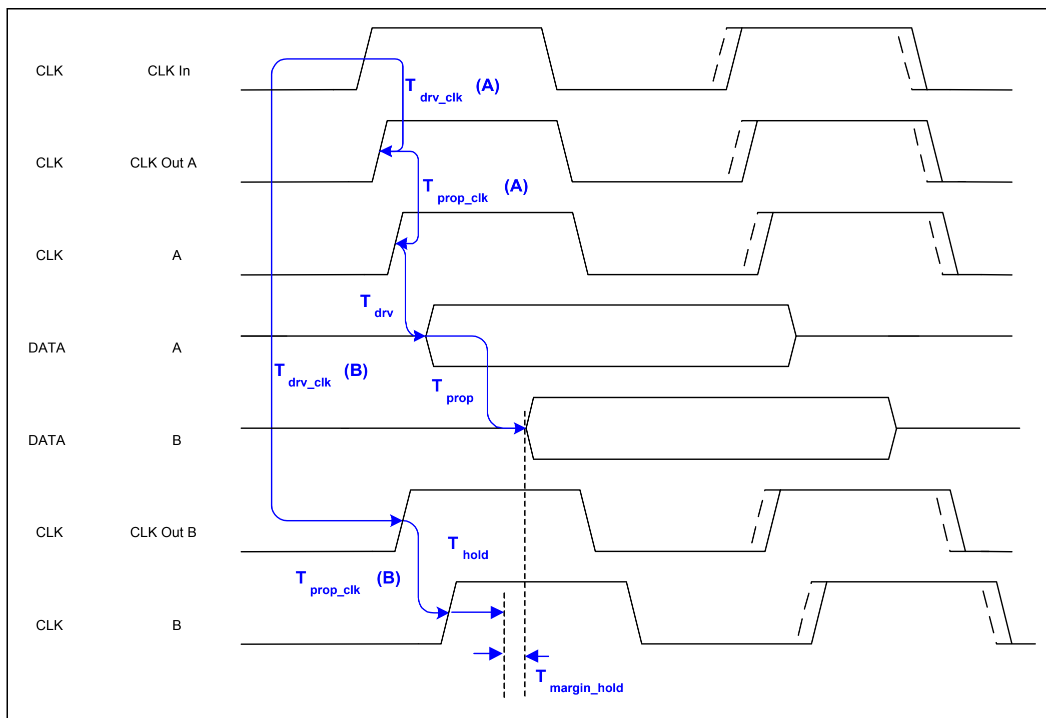
9.1.2.2 Hold Margin

Figure 9-6 illustrates the timing diagram that was used to develop the final timing equations.

Equation 9-13. Common Clock Hold Margin

$$T_{margin_hold} = T_{drv} + T_{prop} - T_{hold} - T_{skew_hold}$$

Figure 9-6. Timing Diagram Used to Determine the Common Clock Hold Timing Equations



9.1.3 Data and Address Setup Time to BCLK

The Intel Xeon processor MP has an additional timing requirement for the data and address signals. The data and address signals must meet the specified setup time to the processor BCLK[1:0]. This should be calculated as a simple common clock setup time as in Equation 9-12.

The specified setup time can be found in the processor datasheet.

9.1.4 Timing Spreadsheet

A timing spreadsheet should be created to keep track of each signal's timing margins.

To effectively manage the timing spreadsheet the following recommendations should be adhered to.

- All assumptions should be stated in the spreadsheet file
- Simulated and measured timings should be tracked separately

- Each timing component has an owner and revision date
- Rising and falling edges should be tracked separately

9.2 Simulation Methodology

This sections outlines the simulation methodology used to determine the topology and routing guidelines.

9.2.1 Design Optimization

The layout for a high-speed bus design can be complex. High-frequency phenomena that previously had second or third order effects on system level performance are becoming first order as bus speeds continue to increase. It should be noted that for a high-speed bus, fixing a problem in one area of the board might create another problem in a different area.

The design recommendations of this design guide have been written to provide enough detail to allow a platform designer to go right into layout designs and only perform post-route simulations. If any of the recommendations are not followed, then it is advisable to follow the complete simulation process described below in order to accurately quantify your solution space.

9.2.2 Signal Categories and Topology Options

The first section of the bus design process is to determine all the signal categories contained within the design. Categories should be defined by signal buffer type, timing requirements, and topology similarities. The bus or component specification should provide help in this categorization.

Once signals have been categorized, all possible interconnect topologies for each signal group must be determined. This requires significant collaboration with the layout engineer and will be the direct result of a layout study. The optimum part placement and all possible interconnect solutions should be determined. The layout study should produce a solution space that lists all possible interconnect topology options including line lengths, widths and spacing. Extensive simulations during the sensitivity analysis will be used to limit the solution space determined from the layout study. This limited solution space becomes a final design solution that meets all timing and signal quality specifications.

9.2.3 Sensitivity Analysis

A sensitivity analysis is used to determine the solution space for all aspects of the design. Every parameter in the system bus should be varied in simulations. The performance metrics, such as flight time, flight skew, and signal integrity are observed while each one of the variables is swept. The performance as a function of each variable is compared to the timing and signal quality specifications. As a result, limits are placed on each of the components. This produces a solution space that places strict limits on the system variables such as trace lengths, spacing, impedance, etc. The solution space will lead to design guidelines for the PCB and routing. Table 9-1 lists the primary system variables that should be considered in the system bus sensitivity analysis. The following table indicates the relative effect of each variable on system performance

Table 9-1. System Variables to Consider for Sensitivity Analysis

System Variable	Impact on Timings and/or Signal Integrity
Trace/stub lengths	High
Trace impedance variations	High
Buffer impedance variations	High
Buffer capacitance variations	Moderate
Er variations	Low (variation is usually small for stripline. Higher for microstrip)
Pattern dependency	Low to High (high for long lines)
Ground return path discontinuities	Potentially high
Trace to trace spacing	High
AC losses	Moderate
Receiver capacitance variations	High
Package parasitic effects	High
Termination resistor variations	Low (if high tolerance resistors are used)
Layer to layer Zo and Er variations	Moderate
Serpentine spacing	Moderate
Simultaneous Switching Outputs (SSO)	Moderate
Inter-Symbol Interference (ISI)	Moderate

9.2.4 Signal Quality Metrics

The tight timing and low voltage characteristics of the system bus require clean reception of all signals. Ringing below receiver thresholds, non-monotonic signal edges, and excessive voltage swing will adversely effect system timings. Excessive ringback, and signal non-monotinicity cannot by tolerated, since these phenomena may inadvertently advance receiver state machines. Excessive signal swings (overshoot and undershoot) are detrimental to silicon gate oxide integrity, and can cause device failure if absolute voltage limits are exceeded. Additionally, over/undershoot can cause timing degradation due to the build up of inter-symbol interference (ISI) effects.

For these reasons, it is important that the designer work to achieve a solution that provides acceptable signal quality across all systematic variations encountered in volume manufacturing.

This section documents signal quality metrics used to derive topology and routing guidelines.

9.2.4.1 Noise Margin

The receiver buffers are designed to switch at the threshold voltage. Due to several variables such as processor variations and system noise, the threshold voltage may change. This variation in the threshold voltage is known as the noise margin. For the processor the noise margin is assumed to be 100 mV above and 100 mV below the reference voltage, GTLREF.

Signal quality is measured by observing the linearity of the signal as it passes through the transition region (GTLREF ± 100 mV) and by observing any signal ringing into the noise margin region. The upper and lower noise margin levels are referred to as V_{IH} and V_{IL} respectively.

An edge is defined to be linear if it exhibits a linear shape between V_{IH} and V_{IL}. Figure 9-7 depicts a linear edge. Figure 9-9 depicts a nonlinear edge. For non-clock and non-strobe signals, the signal seen at the receiver should be linear between V_{IH} and V_{IL}. On a case-by-case basis where it can be shown that the timing can be met with no potential data corruption, it may not be necessary for the signal to be linear between V_{IH} and V_{IL}.

9.2.4.2 Ringback

Ringback is defined as the amount of voltage that “rings” back towards the threshold voltage and is measured at the receiver.

For non-clock and non-strobe signals, the signal at the receiver should not ringback into the noise margin region unless it can be shown that the timing can be met with no potential data corruption.

9.2.5 Timing Metrics

The timing metrics consists of flight time and flight time skew. Flight time is defined as the amount of time between the point where the signal on the unloaded driver (or loaded with a test load) crosses a certain threshold and the time where the signal crosses the threshold at the receiver. Flight time skew is the difference in flight times between two nets. Figure 9-7 illustrates the definition of flight time (assuming a linear edge from V_{IL} through V_{IH}). The flight time should be evaluated at V_{IL}, $V_{threshold}$, and V_{IH}. The methodology is to measure flight time at these three points and record the worst case. This is valid as long as the edge rate seen at the receiver is equal to the edge rate at which it was characterized (the tester edge rate). If the device was characterized at a different edge rate than the system edge rate, then the following procedures for setup and hold time calculations should be used.

9.2.5.1 Setup Flight Time

If the edge rate seen at the receiver is faster than the specified edge rate, then the traditional method should be used. If the edge rate seen at the receiver is slower than the specified edge rate, then the flight time must be extrapolated from V_{IL} or V_{IH} to $V_{threshold}$ at the specified edge rate. See Figure 9-8 for more details.

9.2.5.2 Calculating Flight Time for Signals with Corrupt Signal Quality

If nonlinearity or ringback occurs between V_{IL} and V_{IH}, then the last crossing should be extrapolated back at the specified minimum edge rate to $V_{threshold}$. See Figure 9-9 and Figure 9-10.

Figure 9-7. Traditional Method of Calculating Rising Edge to Rising Edge Flight Time, Assuming a Linear Edge from V_{IL} Through V_{IH} at the Receiver

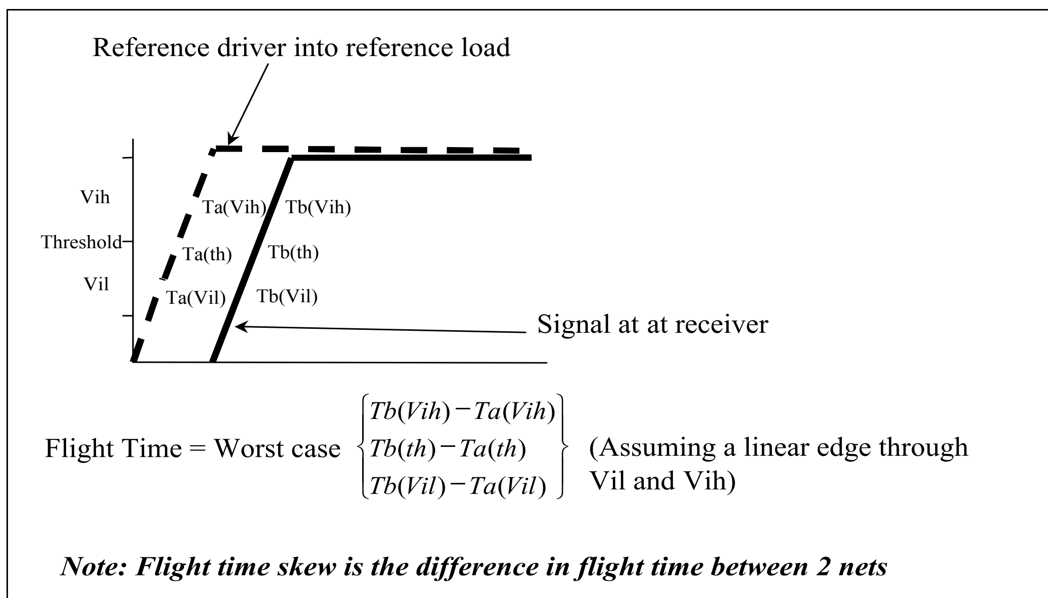


Figure 9-8. Method of Calculating Setup Flight Time When the Edge Rate Seen at the Receiver is Slower than the Minimum Edge Rate

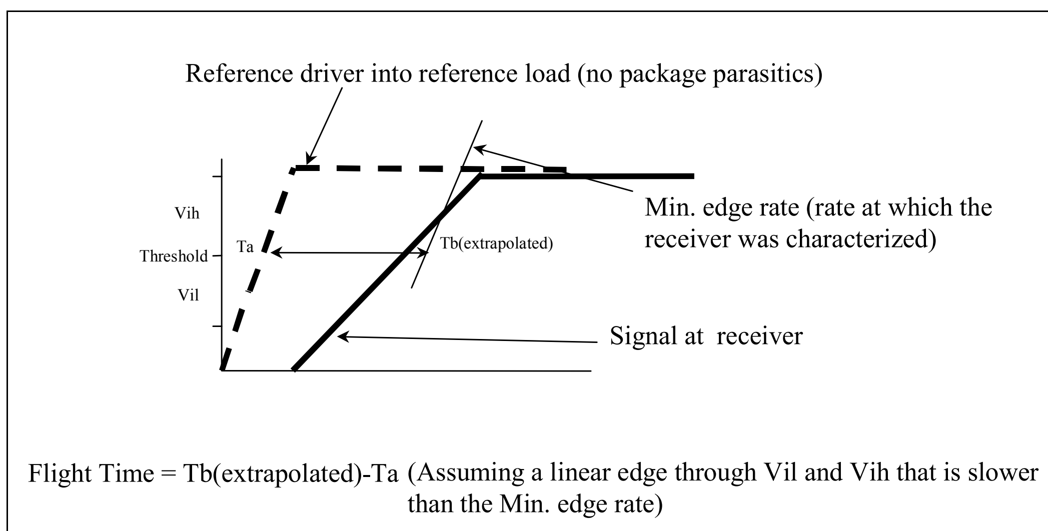


Figure 9-9. Traditional Method of Calculating Flight Time Assuming a Nonlinear Edge from V_{IL} Through V_{IH} at the Receiver

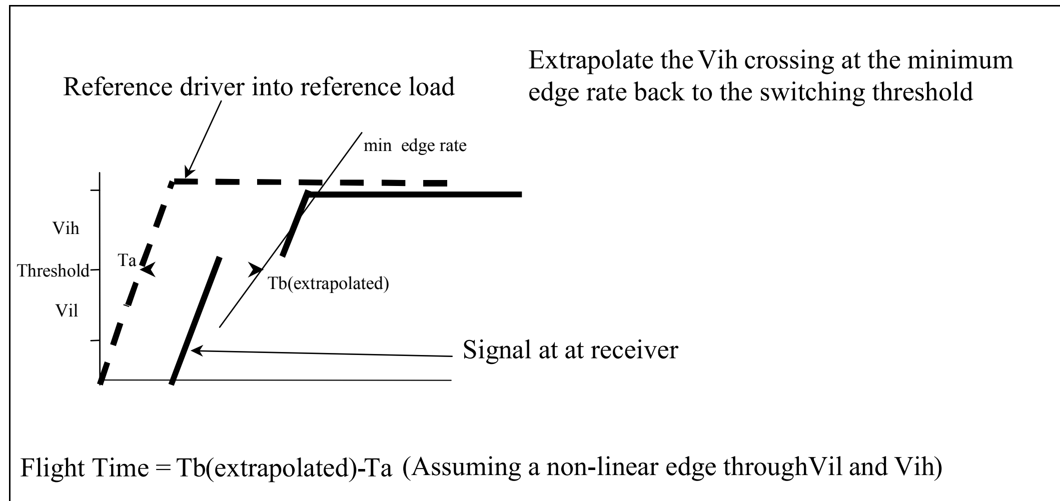
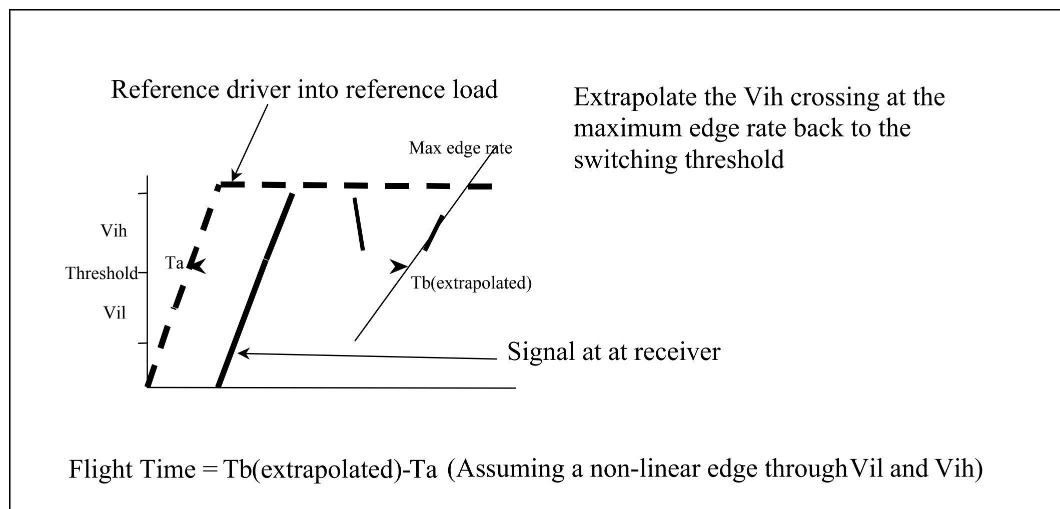


Figure 9-10. Traditional Method of Calculating Flight Time Assuming a Ringback Violation from V_{IL} Through V_{IH} at the Receiver



9.2.5.3 Incorporating Package Effects into the Flight Time

Flight time should be simulated beginning and ending at the pad of the silicon, not at the package pin. This allows the skew due to the package trace length differences to be accounted for. Additionally, the traces on the motherboard should be skewed appropriately to cancel any skews built into the package.

9.2.6 Parameter Sweeps and Monte Carlo Analysis

This part of the sensitivity analysis constitutes the bulk of the pre-route design. In this section of the design phase, all system variables shown in Table 9-1 are varied, and the solution space for the design is determined.

9.2.6.1 Parameter Sweeps

The bulk of the sensitivity analysis consists of parameter sweeps. During a parameter sweep all parameters are held constant except for one or two. The system performance is then observed as the specific variables are being swept. Surface plots can be generated from the results of the parameter sweep. Then the timing and signal quality specifications can be superimposed onto the surface and design limits such as line length, buffer strength, line impedance, etc. can be determined. Figure 9-11 shows an example of surface plots based on simulated flight times and undershoot. The upper and lower left surface plots show flight time as a function of line lengths L2 and L3. Additional planes are incorporated into the plots that represent the upper and lower flight time specifications. The upper right plot depicts a signal quality metric as a function of L2 and L3 line length. The surfaces of all three of these plots are intersected with the specifications and the resultant solution space for these variables under the conditions of the simulation is shown in the lower right hand side of Figure 9-11.

It should be noted that the sweeps should be performed using different switching patterns in order to capture the majority of ISI effects. If N is the fastest switching speed for a given bus, then simulations should be performed at a frequency of N, ½N, and ¼N. The difference between the timings at the different switching frequencies is a good approximation of the ISI noise. Final checks on fully coupled models with long worst-case bit patterns should be performed in order to find any ISI effects not captured in this analysis. This sweeping technique can be used extensively to get initial bounds on all variables in the system. The resultant solution space will be known as the “phase 1 solution space.”

The drawback of this method is that the sweeps are only good for evaluating two variables at a time. While the two parameters of interest are being varied, the other parameters in the system are held constant at a value that may not yield worst-case performance. For example, if the line lengths are being swept all line and buffer impedances, package parasitic effects, receiver capacitance, etc. are held at fixed values. Every effort should be made to set these parameters so that the performance will approach worst case. Use Intel's recommendations as a baseline, and work from there to refine corner conditions specific to your environment. In order to ensure that the worst-case performance is captured, all system variables must be varied simultaneously. This can be accomplished using a targeted Monte Carlo analysis.

9.2.6.1.1Targeted Monte Carlo Analysis

Targeted Monte Carlo (TMC) analysis can be used to further refine the phase 1 solution space and ensure that the worst-case performance has been captured. Performing a full Monte Carlo analysis over all system variables is inefficient because a large number of simulations must be performed to statistically guarantee that all worst case conditions are captured. If a TMC analysis is performed on the boundaries of the phase 1 solution space determined by the parametric sweeps, the number of simulations required will decrease dramatically and the solution space can be refined by changing the variables that were held constant during the sweep.

Figure 9-12 shows the area where the TMC analysis was performed to refine the phase 1 solution space illustrated in Figure 9-11. Figure 9-13 shows the results of the TMC analysis and shows the final solution space. This final solution space will be referred to as the “phase 2 solution space.” It should be noted that the worst-case bit pattern should be included in the TMC analysis.

Figure 9-11. Example of Sweeps Used to Evaluate the Length Limits of Trace L2 and L3

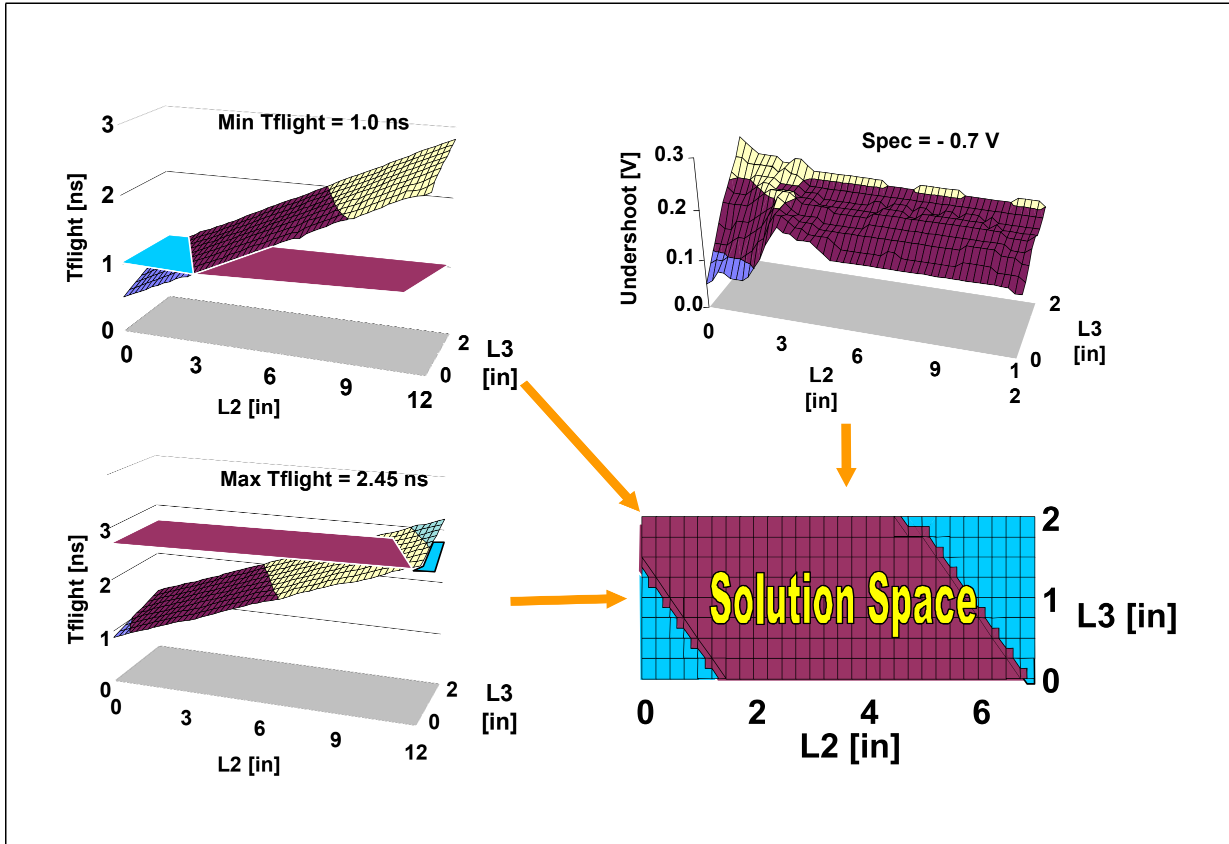


Figure 9-12. Monte Carlo Analysis Should Be Performed on These Areas of the Phase 1 Solution Space

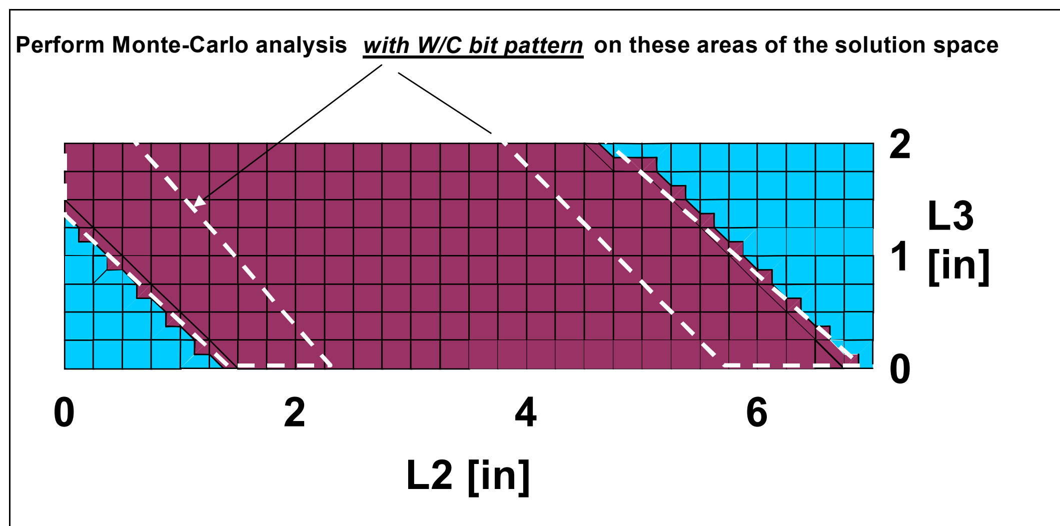
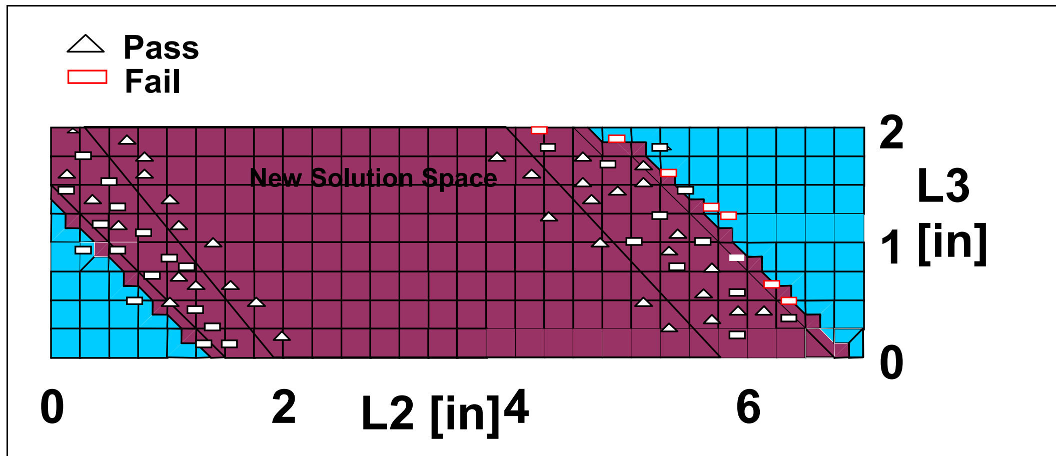


Figure 9-13. Results of Targeted Monte Carlo (TMC) Analysis and the Resultant Phase 2 Solution Space for Variables L2 and L3



9.2.6.2 Final Solution Space

The final solution space will be referred to as the "phase 3 solution space." This phase incorporates effects that are too computationally demanding to easily include in the phase 1 or phase 2 solution spaces. A final check on the worst case nets under fast and slow conditions should be performed. The sweeps and the Monte Carlo analysis already performed should allow the worst-case conditions to easily be chosen. These simulations should be performed using a fully coupled crosstalk model. The results of the final check should be used to further narrow the solution space.

The routing guidelines should incorporate the entire solution space determined during the sensitivity analysis in order to provide the maximum amount of flexibility. When some portions of the routing guidelines cannot be met due to physical real estate limitations or manufacturing concerns, new solutions need to be determined.

System Theory 10

This section provides in-depth information about signal technology and system signal interference.

10.1 AGTL+ Logic

AGTL+ is the electrical system bus technology. It is an incident wave switching, open-drain bus with integrated pull-up resistors (p-channel FETs) that provide both the high logic level and the termination.

The end agents on the system bus will always have their termination on. Middle agents' pull-ups will turn on only when needed to drive a signal to its high state. The ODTEN pin on the processor will be used to determine whether a processor is an end agent and thus needs to enable its termination. It is up to the baseboard designer to pull this pin to the appropriate logic level (high to enable).

10.2 Inter-Symbol Interference

Inter-symbol interference (ISI) is the effect of a previous signal (or transition) on the interconnect delay. When a signal is transmitted down a transmission line and the reflections due to the transition have not completely dissipated, the following data transition launched onto the bus is affected. ISI can impact both the timing and the signal integrity. ISI is dependent upon frequency, time delay of the line, and the reflection coefficient at the driver and receiver. Thus, ISI is a major concern in any high-speed design where the period is smaller than the delay of the transmission line. Figure 10-1 shows an example of how ISI can affect timing. In this example the starting voltage of the driver is different from the idle state starting voltage. This figure illustrates the ISI effect on both timing and signal integrity.

One method of capturing most of the timing impact due to ISI is to perform parameter sweeps at the fastest bus period, and then at 2Xand 3X multiples of the fastest bus period. For example, if the fastest frequency at which the bus will operate is 400 MHz, then the pulse duration of a single bit is 2.5 ns (5 ns period). The data pattern should be repeated with pulse durations of 5 ns and 7.5 ns (10 ns and 15 ns periods). This represents the following data patterns transitioning at the highest bus rate.

01010101010101

00110011001100

00011100011100

The worst-case results of these patterns can be used to produce the phase 1 solution space. The maximum difference in flight time between these patterns produces a first order approximation of the ISI impact.

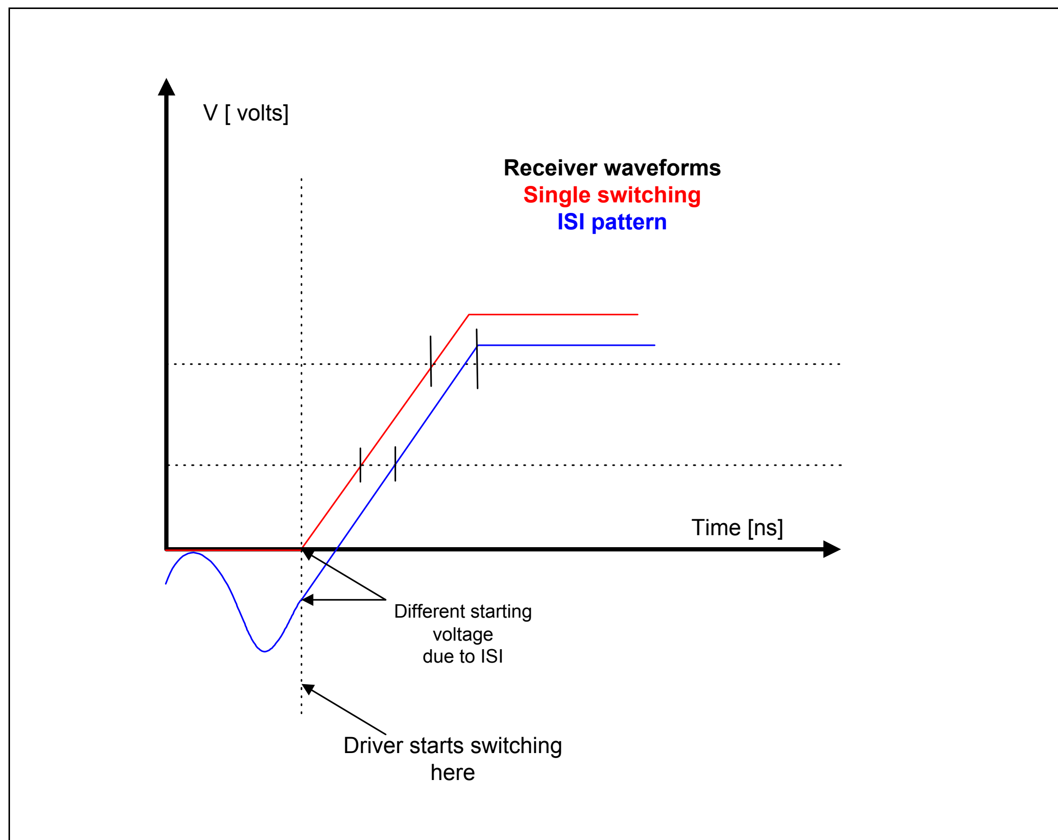
The final solution space must account for the full ISI variations. This can be done by performing targeted simulations at the edges of the phase 2 solution space using a long pseudo-random pulse train. If the timing impact due to ISI does not violate any timing or signal integrity specifications,

then the phase 2 solution space is acceptable. If timing violations do occur then steps should be taken to minimize reflections on the bus, which will reduce the ISI. Typically, the best way to limit reflections is to reduce impedance variations and minimize discontinuities (e.g., by shortening stubs and connectors, matching impedance between packages and motherboard traces, etc.).

The worst-case ISI can be evaluated using the following procedure:

- Simulate the longest net on the bus using a long pseudo-random bit pattern for both the fast and slow cases.
- Take the first transition of the ISI simulation as the baseline.
- Determine the rising and falling delays for each bus transition.
- Subtract the minimum and maximum delays from the baseline delays and find the worst-case difference.
- Take the smallest negative and the greatest positive difference. This should be the worst case ISI impact on timing.

Figure 10-1. Example of ISI Impact on Timing and Signal Integrity



10.3 Crosstalk

Crosstalk is caused through capacitive and inductive coupling between networks. Crosstalk can be backward or forward. Backward crosstalk creates an induced signal on a victim network that travels in a direction opposite that of the aggressor's signal. Forward crosstalk creates a signal that travels in the same direction as the aggressor's signal. On an AGTL+ bus a driver on the aggressor network is not necessarily at the end of the network. Therefore, it sends signals in both directions on the aggressor's network. The signal propagating in each direction causes crosstalk on the victim network. This effect is illustrated in Figure 10-2, which shows a driver on the aggressor network and a receiver on the victim network. Figure 10-3 shows two aggressors on each side of the victim. Additional aggressors are possible in the z-direction if adjacent signal layers are not routed in mutually perpendicular directions. Because coupling coefficients decrease rapidly with increasing separation, it is rarely necessary to consider aggressors that are at least five line widths separated from the victim. Additionally, there is crosstalk internal to the IC packages, which can also affect the signal quality.

Figure 10-2. Propagation on Aggressor Network

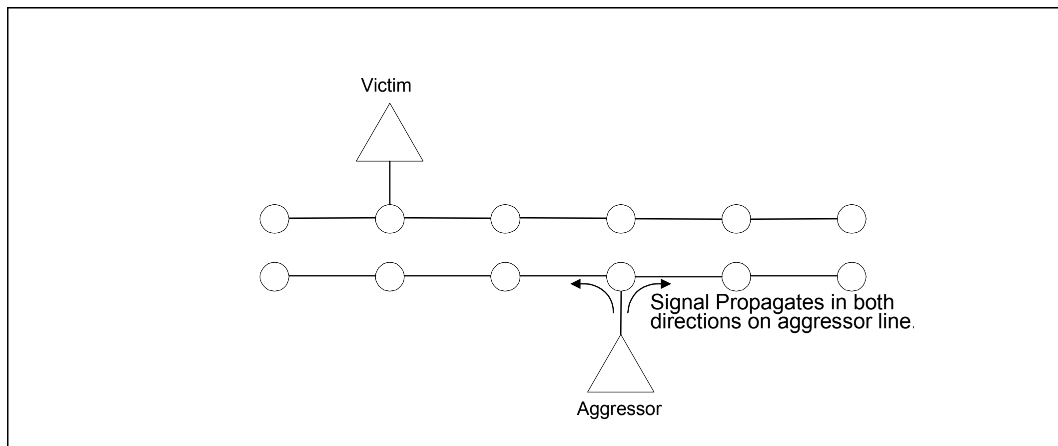


Figure 10-3. Aggressor and Victim Networks

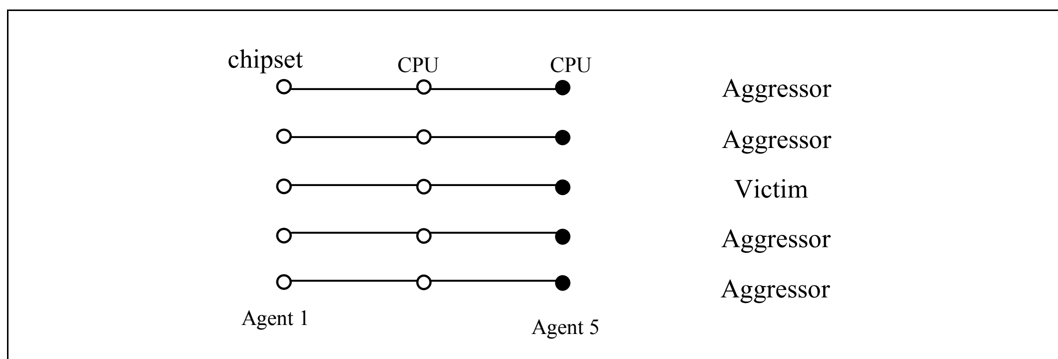
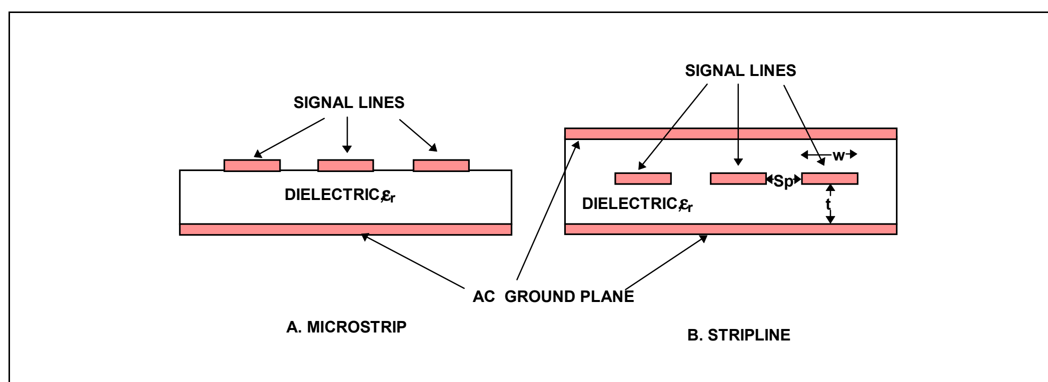


Figure 10-4. Transmission Line Geometry of Microstrip and Stripline



Backward crosstalk is present in both stripline and microstrip geometry. The backward-coupled amplitude is proportional to the backward crosstalk coefficient, the aggressor's signal amplitude, and the coupled length of the network. Backward crosstalk reaches a maximum (and remains constant) when the propagation time on the coupled network length exceeds one half of the rise time of the aggressor's signal. Assuming the ideal ramp on the aggressor from 0% to 100% voltage swing and the rise time on an unloaded coupled network, then the following equation applies.

Equation 10-1. Length for Maximum Backward Crosstalk

$$Length\,for Max Backward Crosstalk = \frac{\frac{1}{2} \times RiseTime}{Board Delay Per Unit Length}$$

An example calculation if fast corner fall time is 1.5 V/ns and board delay is 175 ps/inch (2.1 ns/foot) follows:

Fall time = 1.5 V/1.5 V/ns = 1 ns

Length of maximum backward crosstalk = ½ * 1 ns * 1000 ps/ns /175 ps/in = 2.86 inches

Agents on the AGTL+ bus drive signals in each direction on the network. This will cause backward crosstalk from segments on two sides of a driver. The pulses from the backward crosstalk travel toward each other and will meet and add at certain moments and positions on the bus. This can cause the voltage (noise) from crosstalk to double. Table 10-1 provides example coupling factors for various stripline space to width to dielectric thickness ratios (see Figure 10-4) with dielectric constant $\varepsilon_r = 4.5$, $V_{OH_MAX} = 1.5$ V, and $Z_0 = 65\ \Omega$. Note that the fast edge rates of falling edges place limits on the maximum-coupled length allowable. Also, it should be noted that multiple parallel-coupled lines will increase the impact on the noise budget.

Forward crosstalk is absent in stripline topologies, but present in microstrip. This is for the ideal case with a uniform dielectric constant. In actual boards, forward crosstalk is nearly absent in stripline topologies, but abundant in microstrip. The forward coupled amplitude is proportional to the forward crosstalk coefficient, the aggressor's signal edge rate (dv/dt), and the coupled network's electrical length. The forward crosstalk coefficient is also a function of the geometry. Unlike backward crosstalk, forward crosstalk can grow with coupled section length, and may transition in a direction similar to or opposite to that of the aggressor's edge. Unlike backward crosstalk, forward crosstalk on the victim signal will continue to grow as it passes through more coupled length before the aggressor's wave front is absorbed by the termination.

To minimize crosstalk:

- Route adjacent trace layers in different directions (orthogonal preferred) to minimize the forward and backward crosstalk that can occur from parallel traces on adjacent layers. This reduces the source of crosstalk.
- Maximize the spacing between traces. Where traces have to be close and parallel to one another, the distance that they are close together should be minimized, and the distance between sections that have close spacing should be maximized. Routing close together could occur where multiple signals have to route between a pair of pins. When this happens the signals should be spread apart where possible. Also note that routing multiple layers in the same direction between reference planes can result in parallel traces that are close enough to each other to have significant crosstalk.
- Minimize the variation in board impedance (Z_0). For the example topologies covered in this guideline, either 50 Ω ± 10% was assumed for multi-processor based designs respectively.
- Minimize the nominal board impedance within the AGTL+ specification while maintaining the same trace width/spacing ratio. For a given dielectric constant, this reduces the trace width/trace height ratio, which reduces the backward and forward crosstalk coefficients. Having reduced crosstalk coefficients reduces the magnitude of the crosstalk.
- Minimize the dielectric constant used in the PCB fabrication. As above, all else being equal, this puts the traces closer to their reference planes and reduces the magnitude of the crosstalk.
- Watch out for voltage doubling at a receiving agent, caused by the adding of the backward crosstalk on either side of a driver. Minimize the total network length of signals that have coupled sections. If there has to be closely spaced/coupled lines, place them near the center of the net. This will cause the point in time that voltage doubling occurs to be before the setup window.
- Route synchronous signals that could be driven by different components in separate groups to minimize crosstalk between these groups. The processor uses a split transaction bus with five independent sub buses (arbitration, request, snoop, response, and data). This implies that in a given clock cycle, each sub bus could be driven by a different agent. If these two agents are at the opposite process corner (one fast and one slow), then separating the bus types will reduce the impact of crosstalk.

Table 10-1. Example Backward Crosstalk Coupling Factors

Space: Width: Thickness	Coupling Factor	Maximum Crosstalk
24:4:8	0.65%	9.8 mV
20:4:8	1.3%	19. 5 mV
16:4:8	1.75%	26.2 mV
14:4:8	2.5%	37.5 mV
12:4:8	3.4%	51.0 mV
8:4:8	6.55%	98.2 mV
4:4:8	13.5%	202.5 mV

NOTE: Assumes ε_r = 4.5, V_{OH_MAX} = 1.5 V, and Z_0 = 65 Ω

Crosstalk will be incorporated in the processor system bus analysis methodology in two ways. Initially, for the sensitivity analysis sweeps, the SLEM model will be utilized. This method allows quick computation and is well-suited to the bulk of the sensitivity analysis. After the initial solution space is found using the SLEM method to account for crosstalk, fully coupled simulations will be performed as a final check.

10.3.1 Single Line Equivalent Model (SLEM)

The SLEM method of modeling crosstalk is a technique where a single line can be used to represent the effects seen by a victim line that is routed in the middle if a multiple conductor bus. The impedance and the velocity of the single line are altered depending on the propagation mode. This method allows quick computation and is well-suited to large parameter sweeps. Furthermore, the method has been accurately correlated to measurements.

The method involves calculating the odd and even mode impedance and velocities of a 3- or 5-line system and modeling a single line with these variations. For example, consider the following 3-line system as illustrated in Figure 10-5. The worst-case propagation modes will be odd and even modes. The target (or victim) line is conductor B. The SLEM model is calculated Equation 10-2 through Equation 10-5.

Equation 10-2. Even Mode Impedance Calculation

$$Z_{B,even} = \sqrt{\frac{L_{22} + L_{12} + L_{23}}{C_{22} - C_{12} - C_{23}}}$$

Equation 10-3. Even Mode Time Delay Calculation

$$TD_{B,even} = \sqrt{(L_{22} + L_{12} + L_{23})(C_{22} - C_{12} - C_{23})}$$

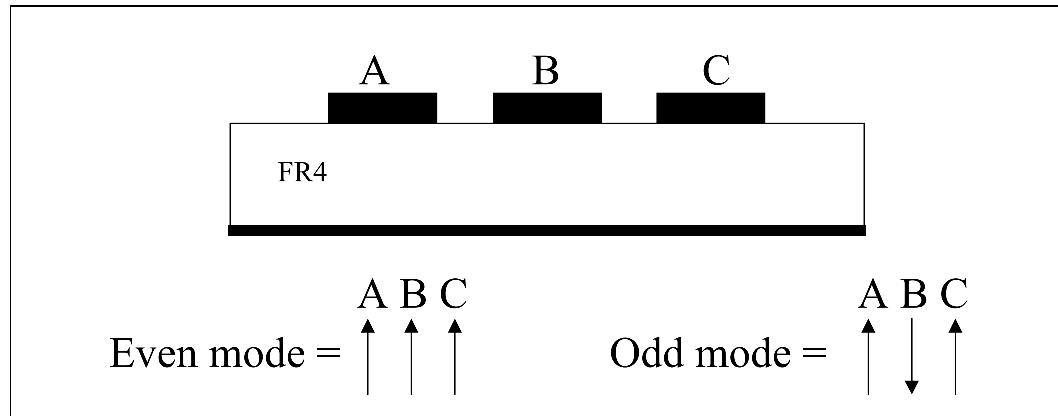
Equation 10-4. Odd Mode Impedance Calculation

$$Z_{B,odd} = \sqrt{\frac{L_{22} - L_{12} - L_{23}}{C_{22} + C_{12} + C_{23}}}$$

Equation 10-5. Odd Mode Impedance Calculation

$$TD_{B,odd} = \sqrt{(L_{22} - L_{12} - L_{23})(C_{22} + C_{12} + C_{23})}$$

Figure 10-5. Cross-Section of a 3-Conductor System Used to Create a SLEM Model



Using equations Equation 10-2 through Equation 10-5 a 3-conductor system can be replaced with a single transmission line model. The odd and even modes will be the worst-case patterns for mode dependent signal integrity and velocity differences.

The SLEM method has some drawbacks. The SLEM method will only model the impedance and velocity differences due to crosstalk. Noise coupling due to excessive coupling is not modeled. However, experience indicates that the impedance and velocity differences are the dominant effect of crosstalk in a system. After the SLEM model is used to determine the initial solution space and the majority of the simulations have been completed, it is necessary to perform a small number of fully coupled simulations to catch any timing impacts due to effects not captured by the SLEM model.

10.3.2 Serpentine Traces

A serpentine net is a transmission line that is routed in such a manner so that sections of the net double back and couple to another segment of the same net.

A serpentine transmission line is sometimes necessary in order to properly match lengths between nets. It is important to properly control the serpentine in order to avoid signal integrity and timing problems. The primary impact of a serpentine trace is an observed decrease in the flight time when compared to a straight trace of equal length. This decrease in the flight time is a result of the crosstalk between parallel sections of the serpentine net. As the signal travels down the transmission line, a component of the signal will follow the transmission line and behave as though it were a straight line with no serpentine. However, another portion of the energy will propagate perpendicular to the parallel routed portions of the serpentine net via the mutual capacitance and mutual inductance. This creates an extra mode that will arrive at the receiver significantly earlier than the other component of the signal. If the coupling between parallel sections is high, this will cause significant timing skew when attempting to match traces length on a bus. Furthermore, if the coupling very high, significant signal integrity problems can result.

The serpentine guidelines included in this document were based on HSPICE simulations with different spacing between parallel sections. The guidelines were chosen to significantly limit the effect of serpentine net.

This page is intentionally left blank.

Design Checklist 11

Use the following checklists as a final check to ensure the motherboard incorporates solid design practices. These lists are only a reference. For correct operation, all of the design guidelines within this document must be followed.

The following tables are quick checklists for platform design. They are created to provide a reminder for key design points or easily forgotten items. These lists are by no means comprehensive, nor do they attempt to explain routing and layout rules. Please refer to the sections provided for detailed instructions.

11.1 Processor Family Connection Checklist

Table 11-1. Processor Connection Checklist (Sheet 1 of 5)

Processor Pin	Signal Type	Pin Connection	Section No.
A[35:3]#	Source synch AGTL+	Connect to all system bus agents. Balance signal lengths within strobe group.	Section 6.4.1
A20M#	Asynch GTL+	Connect to chipset or transition logic. Requires 300 Ω pull-up.	Section 6.4.2
ADS#	Common Clock	Connect to all system bus agents.	Section 6.4.1
ADSTB[1:0]	Source synch AGTL+	Connect to all system bus agents. Balance signal lengths within strobe group. Maintain 25 mil spacing from all other signals.	Section 6.4.1
AP[1:0]#	Common Clock	Connect to all system bus agents.	Section 6.4.1
BCLK[1:0]	Bus Clock	Connect the clock driver. BCLK's to all processors should be length matched, and the BCLK to the chipset should be offset accordingly. Maintain proper spacing.	Section 5.1.1
BINIT#	Common Clock	Connect to all system bus agents, and the chipset if supported. Wired-OR signal. All wired-OR signals should have 143 Ω termination to VCC_CPU at the middle agents (see Figure 6-7). The termination should be located as close as possible to the processor pins with no stubs.	Section 6.4.1

Table 11-1. Processor Connection Checklist (Sheet 2 of 5)

Processor Pin	Signal Type	Pin Connection	Section No.
BNR#	Common Clock	Connect to all system bus agents, and the chipset if supported. Wired-OR signal. All wired-OR signals should have AC termination to VCC_CPU at the middle agents (see Figure 6-7). The termination should be located as close as possible to the processor pins with no stubs.	Section 6.4.1
BPRI#	Common Clock	Connect to all system bus agents.	Section 6.4.1
BR[0]#	Common Clock	Connect to all system bus agents. Swizzle signals between processors. BR0# should be connected to the chipset. Terminate using a 50 Ω pull-up resistor at the processor end.	Section 6.4.1
BR[3:1]#	Common Clock	Connect to all 4 processors. Swizzle signals between processors. Dual terminations using a 50 Ω pull-up resistor.	Section 6.4.1
COMP[1:0]	Power/Other	Refer to processor datasheet.	Section 6.4.2
D[63:0]#	Source synch AGTL+	Connect to all system bus agents. Balance signal lengths within strobe group.	Section 6.4.1
DBI[3:0]#	Source synch AGTL+	Connect to all system bus agents. Balance signal lengths within strobe group.	Section 6.4.1
DBSY#	Common Clock	Connect to all system bus agents.	Section 6.4.1
DEFER#	Common Clock	Connect to all system bus agents.	Section 6.4.1
DP[3:0]	Common Clock	Connect to all system bus agents.	Section 6.4.1
DRDY#	Common Clock	Connect to all system bus agents.	Section 6.4.1

Table 11-1. Processor Connection Checklist (Sheet 3 of 5)

Processor Pin	Signal Type	Pin Connection	Section No.
DSTBN[3:0]#	Source synch AGTL+	Connect to all system bus agents. Balance signal lengths within strobe group. Maintain 25 mil spacing from other signals.	Section 6.4.1
DSTBP[3:0]#	Source synch AGTL+	Connect to all system bus agents. Balance signal lengths within strobe group. Maintain 25 mil spacing from other signals.	Section 6.4.1
FERR#	Asynch GTL+	Connect to chipset or translation logic. Pull-up at both ends of the signal with 56 Ω resistors.	Section 6.4.2
GTLREF	Power/Other	Set to 2/3 $V_{CC.}$ Use multiple GTLREF circuits. Place as close as possible to pin.	Section 8.12.1
HIT#	Common Clock	Connect to all system bus agents, and the chipset if supported. Wired-OR signal: all wired-OR signals should have AC termination to VCC_CPU at the middle agents (seeFigure 6-7). The termination should be located as close as possible to the processor pins with no stubs.	Section 6.4.1
HITM#	Common Clock	Connect to all system bus agents, and the chipset if supported. Wired-OR signal: all wired-OR signals should have AC termination to VCC_CPU at the middle agents (see Figure 6-7). The termination should be located as close as possible to the processor pins with no stubs.	Section 6.4.1
IERR#	Asynch GTL+	If supported, connect to all system bus agents and terminate at both ends with a 56 Ω pull-up. If not supported, leave as no-connect.	Section 6.4.2
IGNNE#	Asynch GTL+	Connect to chipset or translation logic. Pull up with a 300 Ω resistor at processor end of signal.	Section 6.4.2
INIT#	Asynch GTL+	Connect to chipset or translation logic. Pull up with a 300 Ω resistor at processor end of signal.	Section 6.4.2
LINT0/INTR	Asynch GTL+	Connect to chipset or translation logic. Pull up with a 300 Ω resistor at processor end of signal.	Section 6.4.2
LINT1/NMI	Asynch GTL+	Connect to chipset or translation logic. Pull up with a 300 Ω resistor at processor end of signal.	Section 6.4.2
LOCK#	Common Clock	Connect to all system bus agents.	Section 6.4.1
MCERR#	Common Clock	Connect to all system bus agents, and the chipset if supported. Wired-OR signal: all wired-OR signals should have AC termination to VCC_CPU at the middle agents (see Figure 6-7). The termination should be located as close as possible to the processor pins with no stubs.	Section 6.4.1

Table 11-1. Processor Connection Checklist (Sheet 4 of 5)

Processor Pin	Signal Type	Pin Connection	Section No.
ODTEN	Power/Other	Option 1 (preferred): Enable ODT (on-die termination) on Processor 0 (end processor) by pulling up to VCC_CPU with a resistor that falls within the range of 50 Ω ± 20%. Disable ODT for middle agent processors (Processors 1-3) by pulling down to VSS with a resistor that falls in the range of 50 Ω ± 20%. Option 2: Enable ODT on Processor 0 (end processor) by pulling up to VCC_CPU with a 1 kΩ resistor. Disable ODT for middle agent processors (Processors 1-3) by pulling down to VSS with a 1 kΩ resistor.	Section 6.4.2
PROCHOT#	Asynch GTL+	Connect to chipset GPIO or external logic. Pull-up at both ends of the signal with 56 Ω resistors.	Section 6.4.2
PWRGOOD	Asynch GTL+	300 Ω ± 5% pull-up to VCC. Connect to power good circuitry of chipset and/or VRM.	Section 6.4.2
REQ[4:0]#	Source synch AGTL+	Connect to all system bus agents. Balance signal lengths within strobe group.	Section 6.4.1
Reserved		Must remain unconnected.	
RESET#	Common Clock	Use a 50 Ω ± 5% pull-up to VCC. Connect to chipset.	Section 6.4.1
RS[2:0]#	Common Clock	Connect to all system bus agents.	Section 6.4.1
RSP#	Common Clock	Connect to all system bus agents.	Section 6.4.1
SKTOCC#	Power/Other	Connect to external logic as needed. Connect pin of the second processor to BUSPARK on the chipset. This pin maybe used to enable bus paring feature if the chipset has an input pin. If chipset uses software to enable bus parking, Intel recommends the use of the software option.	Section 6.4.2
SLP#	Asynch GTL+	Connect to chipset or translation logic. Pull up with a 300 Ω resistor at processor end of signal.	Section 6.4.2
SM_ALERT#[1]	Power/Other	Pull-up to SM_VCC and connect to SMBus master. The number of devices on the SMBus determines pull-up value.	Section 6.4.2
SM_CLK[1]	Power/Other	Pull-up to SM_VCC. The number of devices on the SMBus determines pull-up value.	Section 6.4.2
SM_DAT[1]	Power/Other	Pull-up to SM_VCC. The number of devices on the SMBus determines pull-up value.	Section 6.4.2
SM_EP_A[2:0] [1]	Power/Other	Leave as no connect to set bit low. Pull up to SM_VCC using a < 1 kΩ resistor to set the bit high. Use these address bits to set a unique SMBus address for the memory devices on the processor. See the processor EMTS for more details.	Section 6.4.2

Table 11-1. Processor Connection Checklist (Sheet 5 of 5)

Processor Pin	Signal Type	Pin Connection	Section No.
SM_TS_A[1:0] [1]	Power/Other	Pull-up to VCC_SMBus with < 1 kΩ resistors to set bit high. Pull-down to VSS through < 1 kΩ resistors to set bit low. Use these address bits to set a unique SMBus address for the processor thermal sensing device. See the processor datasheet for more details.	Section 6.4.2
SM_VCC	Power/Other	Must be connected to 3.3 V power supply and should follow the power sequencing routine.	Section 6.4.2
SM_WP[1]	Power/Other	If used, drive with 3.3 V compatible logic.	Section 6.4.2
SMI#	Power/Other	Connect to chipset or translation logic. Pull-up with a 300 Ω resistor at processor end of signal.	Section 6.4.2
STPCLK#	Asynch GTL+	Connect to chipset. Pull-up with a 300 Ω resistor at processor end of signal.	Section 6.4.2
TESTHI[6:0]	Power/Other	1 kΩ pull-up to VCC_CPU. If boundary scan is not required, TESTHI[0:3] may be tied together and pulled up to VCC_CPU with a single 1 kΩ resistor, and TESTHI[5:6] may be pulled up to VCC_CPU with a single 1 kΩ resistor. TESTHI4 must always be pulled up separately on each processor. Do not connect between processors.	
THERMTRIP#	Power/Other	Use a 56 Ω ± 5% pull-up to V_{CC} and connect to external logic to disable processor VCC_CPU supply within 0.5 seconds after the assertion of THERMTRIP# to protect the processors from damage in over-temperature situations.	Section 6.4.2
TRDY#	Common Clock	Connect to all system bus agents.	Section 6.4.1
V_{CCA}	Power/Other	Connect through appropriate discrete filter.	Section 8.13
$V_{CCIOPLL}$	Power/Other	Connect through appropriate discrete filter.	Section 8.13
$V_{CCSENSE}$	Power/Other	Place via next to processor pad for measurement of V_{CC}. Do not connect to sense logic. Utilize this pin for power delivery validation.	Section 8.13
VID[4:0]	Power/Other	Connect to on-board VRM. X-OR signals together to ensure all processors operate at the same voltage. Optional: Connect to comparison logic to compare VID of all installed processors	Section 6.4.2
V_{SSA}	Power/Other	Connect to V_{CCA} and $V_{CCIOPLL}$ filter.	Section 8.13
$V_{SSSENSE}$	Power/Other	Place via next to processor pad for measurement of V_{CC}/V_{SS}. Do not connect to sense logic. Utilize this pin for power delivery validation.	Section 8.13