This document details the relationship between CV/I device delay metrics, fan-out-of-4 (FO4) inverter gate delay metrics, and high-performance microprocessor clock frequency trends.

The device CV/I metric is an indicator of overall expected device switching speed since it accounts for the intrinsic device capacitances, the voltage swing of interest, and the drive current supplied by a device. Since device capacitance and drive current are both directly proportional to device size, this metric should be size-independent. For further explanation of this metric, see Y. Taur and T. Ning, Fundamentals of modern VLSI devices, New York, Cambridge University Press, 1998. Note that different definitions of the CV/I delay metric may or may not include parasitic device capacitances (i.e. gate overlap capacitances). The PIDS chapter of the 2003 ITRS includes these parasitic capacitances, including Miller effect, and we use the PIDS definition in this document.

To model microprocessor clock frequency trends in the ITRS, we adopt the concept that clock speeds cannot exceed a value which corresponds to a fixed number of gate delays. For example, in the 2001 ITRS, we used a lower limit of 16 typical gate delays for an on-chip global clock period. A typical gate delay is further defined as the delay of an inverter loaded by 4 identical inverters; this is also called a FO4 inverter delay.

Horowitz and others have quantified a relationship between FO4 delay and the gate length of the process used, whereby the general trend is: FO4 delay = A*L where A is a constant and L is the gate length. Two important points must be made here. First, A has been quoted differently in separate publications so there is no exact constant that universally holds. Second, L refers to the bottom edge poly gate length (physical bottom gate length in the ITRS ORTCs), not to the feature size quoted for the process. For instance, Intel's 90nm process [Proc. IEDM, 2002] has an actual poly gate length of 50nm, or slightly more than 50% of the implied process dimension. This is a key point as use of the feature size will give FO4 delay estimations that are about 100% too large. Typical values for A are 360-400 for typical operating conditions where F is in μm and the delay is given in ps. The main difficulty with this model is that gate delay is not completely determined by gate length. While most technologies with identical poly gate lengths will have comparable device switching speeds, this model cannot comprehend the differences among these processes (e.g. different oxide thicknesses yield very different delay times with identical Lgate values, and threshold voltages can vary widely across processes in the same technology node). We believe that the device CV/I delay metric captures, to a sufficient extent; all the relevant components of delay and can moreover comprehend even minor changes within a process technology node.

The goal of this document is to describe the conversion from device CV/I delay metric to FO4 gate delays. The PIDS ITWG has projected CV/I delays for devices throughout the roadmap based on expected supply voltages, saturation drive currents, and gate capacitances. In order to project clock speeds (based on a steady-state 12*FO4 model), we need only a relationship between CV/I and FO4. To justify this relationship, we note that both the CV/I and FO4 delay metrics consider only gate capacitive loading and no interconnect. Thus, since a FO4 delay is set only by the available drive current and the device input capacitance, there should be excellent correlation between the two metrics.

We examined four Intel Corporation papers from the 1998, 1999, 2000, and 2001 International Electron Device Meetings (IEDM) to obtain ring oscillator (RO) delay values. We then extrapolated these RO delay times to a FO4 estimated delay time. Since a typical RO has a fan-out of just 1, we

need to increase the delay time to compensate for a larger capacitive load. A rule of thumb from [Horowitz, Proc IEEE 01] states that junction (also called diffusion) capacitance for an optimized gate is approximately one-half that of its input capacitance. Normalizing an inverter input capacitance to 1; this gives the output load for a RO of 1.5 units and 4.5 units for a FO4 inverter. Therefore, multiplying the RO delay by 4.5/1.5, or 3, we obtain an estimate of the FO4 delay for these Intel processes. They are given below, along with the CV/I values for each process.

Intel, 1998: FO4 delay = 33ps, CV/I = 2.57ps: ratio = 12.84
Intel, 1999: FO4 delay = 31.5ps, CV/I = 2.00ps: ratio = 15.75
Intel, 2000: FO4 delay = 21.3ps, CV/I = 1.64ps: ratio = 12.99
Intel, 2001: FO4 delay = 18ps, CV/I = 1.34ps; ratio = 13.33
The average ratio between FO4 delay and CV/I device delay found experimentally is then 13.73.

Analytically, we assume that the P/N sizing ratio is 2 (to roughly match PMOS/NMOS drive currents) so that the input capacitance of each inverter is 3 units where 1 unit of capacitance corresponds to that of an NMOS device. Since the junction capacitance is one-half that of each device's input capacitance, a RO will have a total capacitive load of (1+0.5) for junction and (2+1) for the input of the next stage. Thus, we expect the RO delay to be 4.5*CV/I. The FO4 inverter has (1+0.5) units for junction capacitance and 4*(2+1) for gate capacitance for a total of 13.5. This value of 13.5 matches extremely well with the 13.73 value from Intel processes.

Given the Intel 130nm process reported in 2001 IEDM and the plans to top out the Pentium 4 clock speed at 3.4GHz within this process, we get a clock period of 294/18 = 16.3 FO4 inverter delays. Looking forward to the 90nm technology node, according to the PIDS high performance tables, the CV/I for this process in 2003 will be 0.95ps. This translates to an aggressive FO4 delay of 13.73*0.95 = 13ps. The latest Intel processor roadmap projects desktop processors in the 90nm node (the Prescott and Tejas chips) to top out at 6.13GHz. This is equivalent to 163ps/13ps = 12.5 FO4 inverter delays per clock cycle.

Note: PIDS tables have CV/I = 2ps for 180nm, 1.63ps for 130nm, and 0.95ps for 90nm. This is different from the ITRS 2001 PIDS tables.

The recent trend in decreasing numbers of FO4 inverter delays per clock cycle cannot continue. The main reasons are: (1) well-formed clock pulses cannot be generated with period below 6-8 FO4 INV delays; (ii) there is increased overhead (diminishing returns) in pipelining (2-3 FO4 INV delays per standard flip-flop, 1-1.5 FO4 INV delays per pulsed flop). The 2003 ITRS MPU model continues the historical rate of advance for on-chip clock frequencies until 12 FO4 inverter delays (in 2007) and flattens beyond that point.
Current projections based on circuit and architectural advances show that the minimum achievable logic depth is 10-12 FO4 inverters.

U.S. Design Technology Working Group