KEITHLEY A GREATER MEASURE OF CONFIDENCE

Getting Started in Parallel Test—Modification of Existing Scribe Line TEGs

Randall G. Lee
Keithley Instruments, Inc.

Initial Strategy

Wafer-level parallel parametric testing involves concurrent execution of multiple tests on multiple scribe line test structures. This has the potential for huge improvements in throughput with existing test hardware.

For many fabs or test cells with mature processes, the most attractive approach to parallel test is to change only the test sequencing on existing TEGs. This approach is usually the best way to achieve significant throughput improvements with a relatively limited investment in analysis, new software, and test sequence modifications. Typically, the process starts with analysis of the TEG and test sequence to find a way to minimize switching time between test pads. Generally, this involves the reordering or regrouping of existing tests on heterogeneous structures.

Analysis of Existing Structures

Typical scribe line test structures are components and groups of components that correspond to the manufacturing process being supported by electrical measurement-based Statistical Process Control (SPC). In order to minimize pad usage, test structure designers frequently connect device terminals together or use other techniques to minimize that amount of space these structures require. From the perspective of parallel testing, this lack of device isolation can create problems. Despite such limitations, experienced test sequence developers have been able to produce parametric test throughput improvements (including prober overhead) ranging from 5% to 40% with existing scribe line test structures. Improvements from 40% to 50% are possible with structure layouts designed to increase the potential for parallel test.

In traditional (i.e., sequential) parametric test programs, each DUT is connected to the measurement instruments one after the other. During the period in which the DUT is connected, forcing signals are applied to it, and then electrical measurements record its response. Once a single test or group of tests for a DUT is complete, the connections are cleared to allow connection to the next DUT. These connect and disconnect times represent some proportion of the overall throughput budget because the relay switching and settling times for high isolation mechanical devices are fixed.

In addition to the relay connect (Conpin) and disconnect (Reinitialize) overheads just described, there is a delay (Delay) overhead, the length of which can vary widely, depending on the DUT type and the measurement conditions. When addressed sequentially, these connect, disconnect, and delay overheads can reduce the overall throughput gains that faster instruments could otherwise deliver. Fortunately, by connecting multiple DUTs to different measurement resources simultaneously (for different types of tests), it's possible to reduce significantly the impact of relay connect and disconnect times on overall throughput (***Figure 1***).

Connecting various measurement instruments to multiple DUTs simultaneously and obtaining reliable data from them is vastly simplified if instruments of the same type (e.g., source-measure units) have identical capabilities. In the case of the DC instrument example shown in ***Figure 2***, each path has uniform signal amplification (via preamp) at the pin. There also are uniform switching

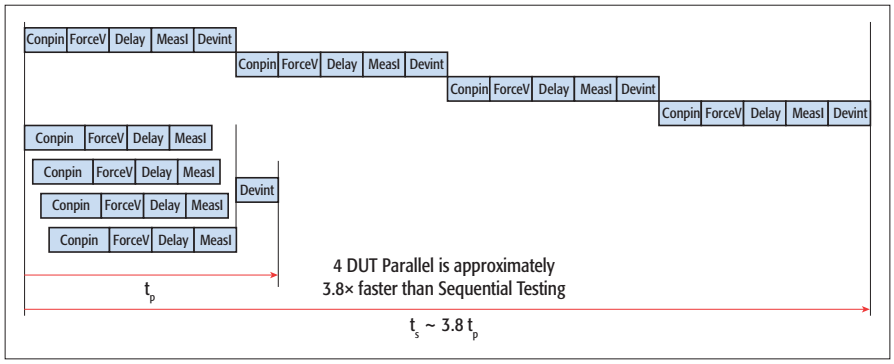

Figure 1. Comparison of elapsed times between sequential and parallel testing of four DUTs. The sequential test time (t_s) is approximately 3.8 times longer than the parallel test time (t_p).

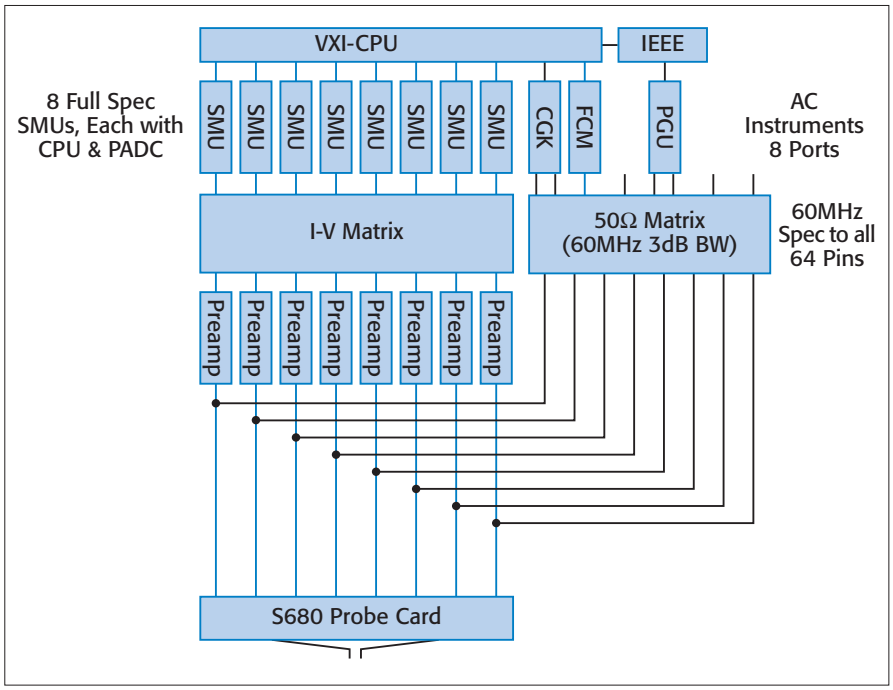

Figure 2. All instruments of the same type must have identical capabilities. In this configuration, all 64 I-V/C-V paths are identical, and all paths provide lab-grade resolution. (This configuration is based on the use of a Keithley S680 Parametric Test System running KTE software.)

characteristics in a relay matrix and uniform SMUs having a full dynamic measurement range, along with dedicated precision analog-to-digital converters (PADCs).

The examples in ***Table 1*** offer one way of evaluating the general benefit of parallel parametric test. In these examples, a fixed set of instrument resources is applied to discrete devices to perform a set of measurements (in this case, threshold voltage, terminal leakages, drive current, and breakdown voltage), which characterize the dynamic measurement range required for the application. In this example, tests of the same type, performed on electrically isolated, identical DUTs, produced maximum (near-theoretical) throughput improvements.

In many cases, however, existing test structure designs lack sufficient electrical isolation (due to shared DUT pins) to produce optimal results. However, empirical data taken on wafer-level structures that are not electrically isolated (***Table 2***) still show throughput benefits, although somewhat reduced from isolated discrete devices. This is significant because it indicates no instrument interference occurs in these parallel test cases.

Table 1. Examples of Parallel Test Throughput Improvements on Discrete Device Measurements

Parameter	Total Sequential Test Time	No. of DUTs Tested in Parallel	Total Parallel Test Time	Throughput Improvement Due to Parallel Testing
V_T	553 ms	2	296 ms	1.87×
I_{DSS}	120 ms	2	67 ms	1.80×
I_{DLEAK}	1154 ms	4	294 ms	3.92×
I_{GLEAK}	1125 ms	4	295 ms	3.81×
B_{VDSS}	1101 ms	4	299 ms	3.68×
TOTAL	**4053 ms**		**1251 ms**	**3.24×**

Table 2. Actual Parallel Test Program Examples (Keithley Series S600 Customers)

Sequential vs. Parallel Time Comparison	Example 1: Lots of Shared Pins	Example 2: Lots of Shared Pins	Example 3: Lots of Shared Pins	Example 4: Isolated Devices	Example 5: Isolated Devices
Sequential Time*	320	8.14	100	210	433
Parallel Time*	230	5.76	62	100	110
Time Savings	28% (1.4×)	29% (1.4×)	38% (1.6×)	52% (2.1×)	77% (3.9×)

*Times are cited in arbitrary units.

Limitations in Assigning Parametric Tester Resources

There are a number of general limitations on how instruments are assigned to various pins, as well as some limitations specific to the parametric test system being used. For example, with the Keithley S680 test system, it's unrealistic to assign a single instrument to multiple pins and expect to collect the same valid measurement results from all pins simultaneously as can be achieved with individual pins. A single SMU instrument assigned to multiple DUTs can only apply one force signal at a time, and cannot measure current for individual DUTs through shared pads and lines. When parallel test is implemented on an S680 system, testing is limited to eight VXI communication threads and one GPIB communication thread[1]. Multiple instruments on the single GPIB communication channel can't be used simultaneously in different threads.

In parallel tests, instruments and pins are essentially "owned" by the first thread that uses them. The tasks running in parallel can't share instruments that vary force conditions or measurements. Similarly, they can't share pins unless they are fixed bias or ground pins as set within the master test sequence. The GPIB (IEEE-488) bus is considered an instrument, so multiple threads can't share its use simultaneously. For example, an IEEE-based capacitance meter or pulse generator can't be operated on two or more different threads at once because the bus is under the control of only a single thread. Once single tests complete in a thread, however, the instruments

1 Thread: The context and code path in which program execution takes place, from start to finish, through a series of tasks.

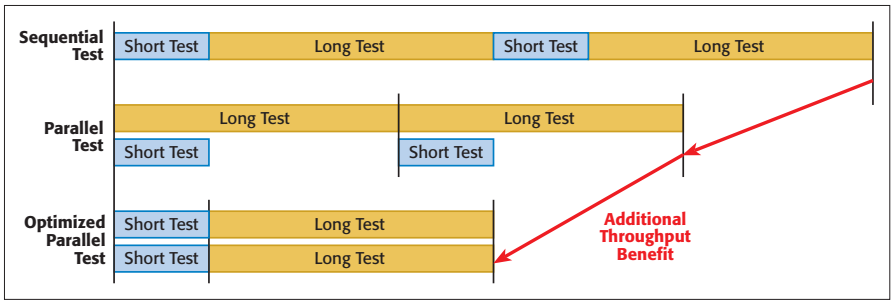

Figure 3. Group tests of similar lengths in order to achieve the highest throughput benefit from parallel test.

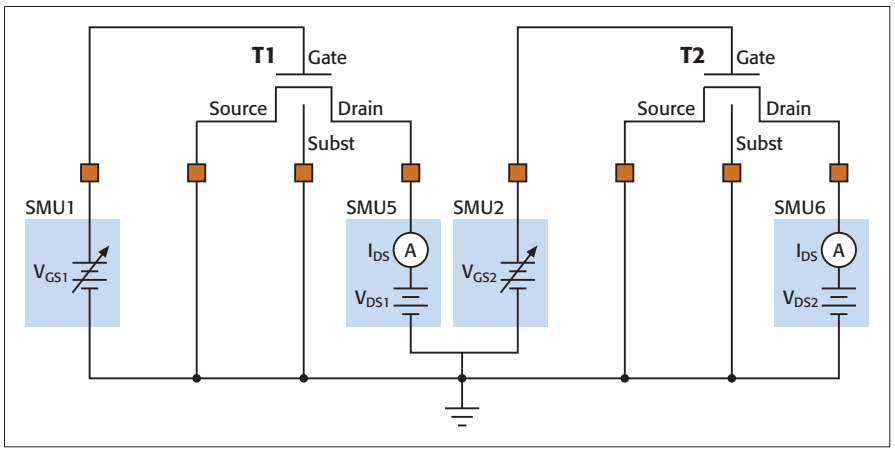

Figure 4. Two totally separate (discrete) transistors are easy to test in parallel.

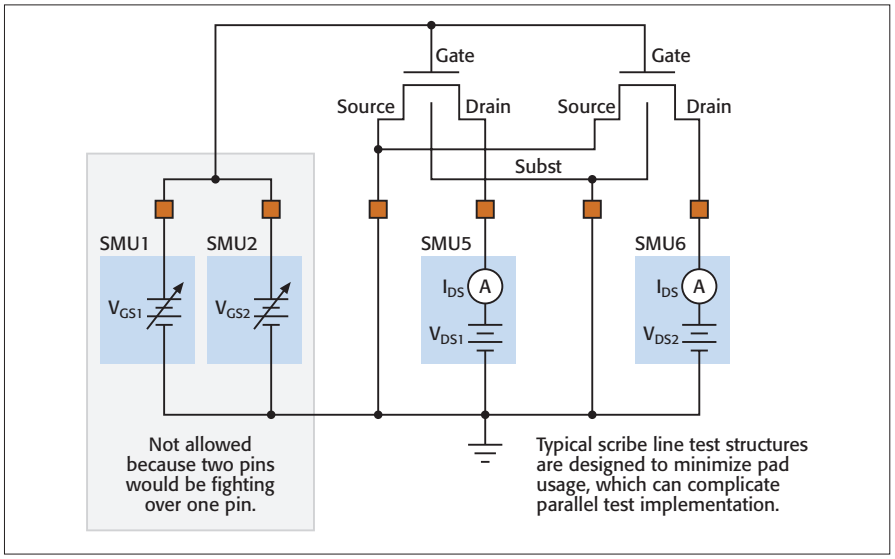

Figure 5. Disallowed test pairings often involve direct connections between structures, such as a shared gate, source, or substrate.

and pins are freed up to be claimed by the next thread and test that needs them.

Optimizing Throughput

Once the general and hardware-specific limitations are understood, the first step in the process of throughput optimization is establishing a sequential performance baseline—in other words, characterizing how long it takes to complete each portion of a specific set of tests in sequential mode. The sequential measurement results should also be compared with data obtained from running the same tests in parallel to make sure there are no anomalies. While performing any portions of the tests in parallel will result in some throughput improvement, it typically doesn't achieve all the potential time savings (***Figure 3***). It's critical to group tests with similar test times to get the greatest benefit from parallel testing. Engineers must take these factors into consideration when generating test programs.

Parallel Testing of Legacy Test Structures

When evaluating legacy test structures (i.e., the existing scribe line structures for wafers already in production) for their suitability for a parallel test strategy, electrically isolated devices are obviously the best choices. For example, the two discrete transistors illustrated in ***Figure 4*** are well suited for testing in parallel. Much more common, however, is the space-saving, shared-terminal type of legacy structure (***Figure 5***), which often presents problems in parallel testing. Testing this device in parallel would require applying different voltages to the two gates, which is clearly impossible.

Measuring resistor chains in parallel requires special attention. ***Figure 6*** illustrates a test structure where resistor chains can present measurement problems.

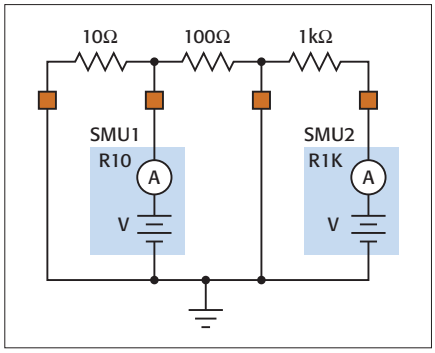

Figure 6. Resistor chains present special parallel test challenges.

Chain structures must also be recognized as shared pin structures that are subject to the same interference issues as structures with shared pins and best treated as a single DUT. Even if the chains are measured without the proper consideration of current flows, tests on some structures within them may produce valid results, such as the 1kΩ resistor shown in ***Figure 7***.

However, tests on other chain resistors, such as the 10Ω resistor in ***Figures 7*** and ***8***, will produce invalid results. This is shown by the calculation in ***Figure 8***.

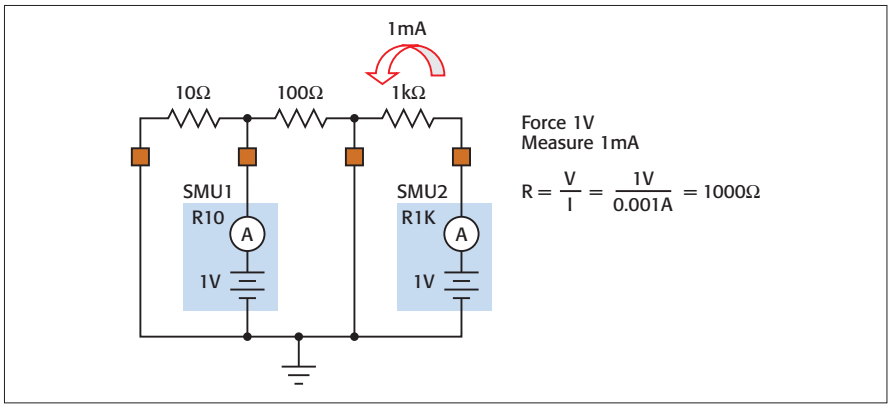

Figure 7. Despite the use of a measurement technique that's inappropriate for parallel test, the measurement of the 1kΩ resistor produces a valid result.

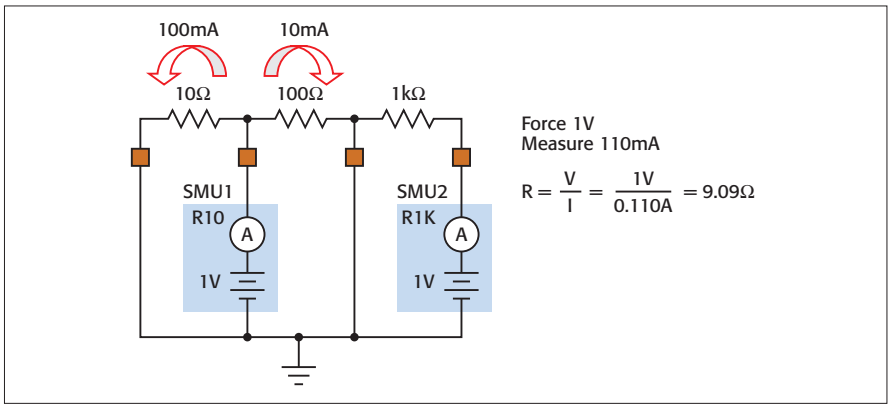

Figure 8. In this portion of the test, the incorrect technique produces an invalid result for the measurement of the 10Ω resistor.

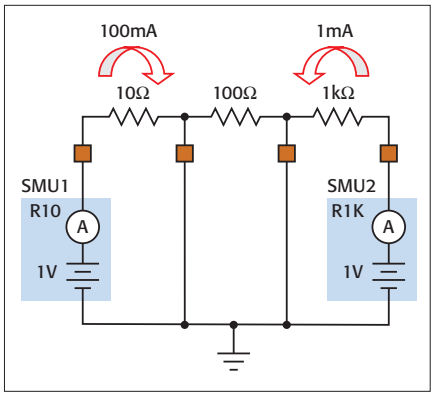

Figure 9. The appropriate measurement technique allows for valid parallel measurements of both the 10Ω and 1kΩ resistors.

When the correct technique is used (***Figure 9***), both the 10Ω and the 1kΩ resistor will be measured correctly. Still, sorting out the unintended measurement paths in legacy test structures typically demands additional engineering effort, which can slow the payback on a parallel test implementation investment.

In some cases, evaluating legacy structures for parallel test requires more than studying their schematic representations; examining a device's cross-section may also be necessary. For example, with a P+ resistor in an N well, a common diffusion means a common measurement path, but some common diffusions can be managed by applying the correct bias (***Figure 10***). Even though the reverse bias leakage paths may produce some surprises, this structure would generally be considered an acceptable candidate for parallel test.

In contrast, examining the cross-section of an N+ resistor in an N well (***Figure 11***) reveals some obvious problems in terms of parallel test. Testing of un-isolated structures creates problems for testing in parallel. For example, a diagnostic program that measures between all pins in a test structure set will quantify the resistance of an unintended measurement path. Clearly, the structure in ***Figure 11*** cannot be used for parallel testing.

Parasitic voltage drop is another important consideration when evaluating legacy structures for parallel test. In a common test structure that contains electrically isolated DUTs distributed along a common ground line, like the set of transistors shown in ***Figure 12***, the source and substrate are frequently connected together. Despite this connection, such a structure is still a good candidate for parallel test if the voltage drops have been characterized.

In ***Figure 13***, the gate and drain voltages must be adjusted to compensate for the voltage drop on the shared source connection. However, in many designs, a long source line effectively acts as a resistor, which means measurements made on one device can affect the results of measurements on others. For example, passing high current through one device can cause a voltage drop in the test structure as a whole, thereby changing the V_{DS} and V_{GS} voltages.

Test structure design changes, although almost never made to wafers already in production, can minimize the effect of cumulative currents on shared ground lines during parallel testing. These changes usually involve increasing the area of the structure. In the case of the structure shown in ***Figure 14***, an alternative to the structure in ***Figure 12***, duplicate ground lines are used to minimize the parasitic resistance drops within the structure. Adding a ground pad at the other end of the structure is another possible option for minimizing the drops. Once these parasitic voltage drop considerations are managed, only the pad contact resistive drops remain to be managed.

While implementing parallel test demands rethinking a variety of preconceptions about parametric test structures, many of these issues only arise when new structures are being designed. By examining the parallel vs. sequential mode program correlation on existing device layouts, most of these issues can be resolved without the need to examine the structure itself. In fact, Keithley engineers developed *pt_execute*, a test program characterization and optimization tool that, among many other things, allows identifying correlation problems quickly.

The *pt_execute* tool automates many of the decisions a test program developer would make. It is based on the experience of Keithley application engineers in implementing high throughput parametric test systems at customer sites. This tool is now part of the Keithley's standard parallel parametric test product offering.

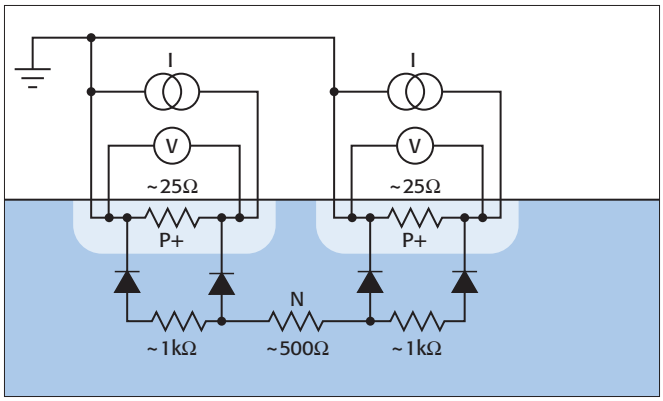

Figure 10. P+ resistor in an N well with a common diffusion.

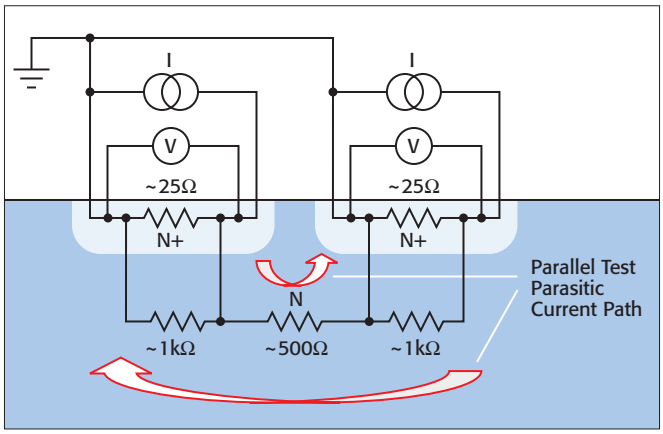

Figure 11. N+ resistor in an N well represents a problem for parallel test.

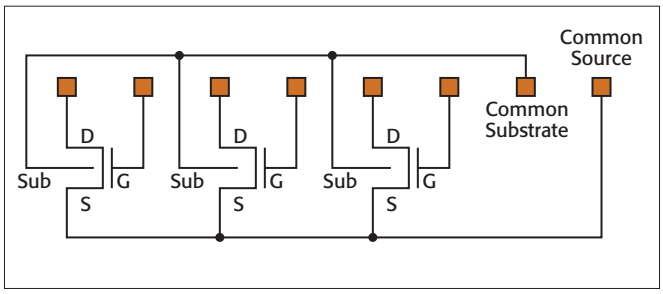

Figure 12. A set of transistors is a common test structure. This one is a good candidate for parallel test if the voltage drops have been characterized.

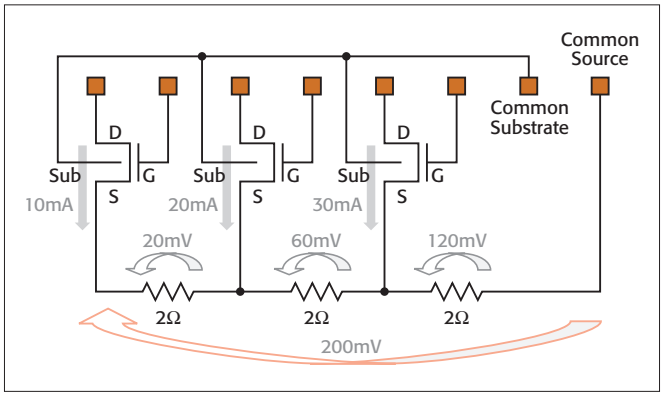

Figure 13. Measurements made on one device often affect the results of measurements on others.

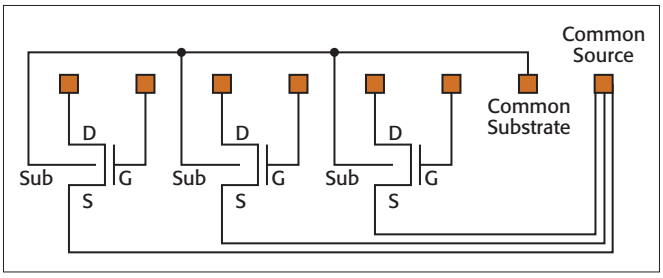

Figure 14. Duplicate ground lines can be used to minimize the parasitic resistance drops within a structure.

The *pt_execute* software automatically detects the test hardware configuration of the parameter test system on which it is installed. It groups tests based on the instrument resources available, so there's no need to keep a running tally of how many SMUs are available to apply to a specific test. It also allows for easy switching in and out of parallel test mode, which simplifies analysis of throughput improvements and detection of sequential vs. parallel correlation issues. Keithley S600 users can easily group tests into sequential and parallel modes when correlation issues arise. This allows specific tests to run in sequential mode if necessary, while the rest run in parallel to increase throughput to the maximum extent possible.

More information on wafer level parallel parametric test and *pt_execute* can be found in Keithley's Parallel Test Technology handbook, available at http://www.keithley.com/at/508. KEITHLEY

About the Author

Randall Lee is a Senior Industry Market Manager in Keithley Instruments' Semiconductor Test Group where he is responsible for parametric tester product marketing and management. He received a BS in Electrical Engineering from Rensselaer Polytechnic Institute, and has 25 years of experience in semiconductor lithography and wafer characterization and test.

KEITHLEY

A GREATER MEASURE OF CONFIDENCE

KEITHLEY INSTRUMENTS, INC. ■ 28775 AURORA ROAD ■ CLEVELAND, OHIO 44139-1891 ■ 440-248-0400 ■ Fax: 440-248-6168 ■ 1-888-KEITHLEY ■ www.keithley.com

BELGIUM
Sint-Pieters-Leeuw
Ph: 02-3630040
Fax: 02-3630064
info@keithley.nl
www.keithley.nl

CHINA
Beijing
Ph: 8610-82255010
Fax: 8610-82255018
china@keithley.com
www.keithley.com.cn

FINLAND
Espoo
Ph: 09-88171661
Fax: 09-88171662
finland@keithley.com
www.keithley.com

FRANCE
Saint-Aubin
Ph: 01-64532020
Fax: 01-60117726
info@keithley.fr
www.keithley.fr

GERMANY
Germering
Ph: 089-84930740
Fax: 089-84930734
info@keithley.de
www.keithley.de

INDIA
Bangalore
Ph: 080-26771071, -72, -73
Fax: 080-26771076
support_india@keithley.com
www.keithley.com

ITALY
Peschiera Borromeo (Mi)
Ph: 02-5538421
Fax: 02-55384228
info@keithley.it
www.keithley.it

JAPAN
Tokyo
Ph: 81-3-5733-7555
Fax: 81-3-5733-7556
info.jp@keithley.com
www.keithley.jp

KOREA
Seoul
Ph: 82-2-574-7778
Fax: 82-2-574-7838
keithley@keithley.co.kr
www.keithley.co.kr

MALAYSIA
Penang
Ph: 60-4-656-2592
Fax: 60-4-656-3794
chan_patrick@keithley.com
www.keithley.com

NETHERLANDS
Gorinchem
Ph: 0183-635333
Fax: 0183-630821
info@keithley.nl
www.keithley.nl

SINGAPORE
Singapore
Ph: 65-6747-9077
Fax: 65-6747-2991
koh_william@keithley.com
www.keithley.com.sg

SWEDEN
Solna
Ph: 08-50904600
Fax: 08-6552610
sweden@keithley.com
www.keithley.com

SWITZERLAND
Zürich
Ph: 044-8219444
Fax: 044-8203081
info@keithley.ch
www.keithley.ch

TAIWAN
Hsinchu
Ph: 886-3-572-9077
Fax: 886-3-572-9031
info.kei@keithley.com.tw
www.keithley.com.tw

UNITED KINGDOM
Theale
Ph: 0118-9297500
Fax: 0118-9297519
info@keithley.co.uk
www.keithley.co.uk

Printed in the U.S.A. No. 2950 04.28.08