Sevgin Oktay

Parametric Study of Temperature Profiles in Chips Joined by Controlled Collapse Techniques*

Abstract: Parameters governing the temperature profiles of typical semiconductor chips joined to circuit module substrates by controlled chip collapse (flip-chip bonding) techniques are discussed. These include the physical and geometric properties of various layers of metals and non-metals that form the chip-to-substrate interconnection. The importance of the bond between the interconnection and the substrate from the point of view of interfacial thermal resistance is indicated. Also, the "thermal pinch" effects of voids in controlled chip collapse interconnections are discussed. The various thermal impedances as obtained from computer simulated temperature profiles are given graphically as functions of the parameters. The derivation of a semi-empirical expression for predicting the transient response of junctions on joined chips is shown.

Introduction

With the advent of integrated circuit technology, the problem of dissipating heat from miniaturized electronic components becomes increasingly complex. Consequently, the design engineer must adopt more sophisticated tools for assessing more accurately the cooling requirements of computer hardware. The present state of the art in the electronics industry has reached a point where one cannot always freely invoke air cooling for heat dissipation with an adequate margin of safety. Accordingly, cooling design approaches must be scrutinized, and decisions to adopt, say, liquid cooling must be based on sound analyses.

In this study, a computer analysis of the parameters governing the temperature profiles in "flip-chips" joined by the controlled chip collapse bonding techniques is presented. A description of the computerized numerical methods is given in the Appendix.

The need for a numerical approach is shown first by considering the complicated structure of flip-chips,[1] the generation of heat in the body of the chip and, more importantly, the manner in which heat is removed via its boundaries. In this analysis the thermal path from the chip to the ultimate sink, and the thermal environment of the chip, are identified. The various "thermal resistances" and "thermal impedances" are also defined.

Next, both the steady-state and transient temperature profiles on a flip-chip are shown. These profiles are constructed from computer simulations of random heat sources of varying strengths in the chip. From the temperature profiles, thermal spreading impedance across the chip and the impedance from the chip level to the bottom of the pads are deduced. The influence on thermal impedance of parameters describing the physical and geometrical properties of the various metallurgical layers that form the chip-to-substrate interconnection are discussed. The variations in thermal impedance as a function of tolerances on pad dimensions are calculated. Projections as to the severity of temperature gradients on high-power large chips are made. These temperature gradients are discussed from the point of view of thermal "pinch" effects over the pads. Also, an analysis of transient temperature profiles on flip-chips is shown. In this connection, the derivation of a closed-form expression for predicting the transient response of the chip is indicated.

The immediate environment with which the flip-chip is in contact is the module.[2] The module, comprising chip sites, a ceramic substrate, pins and the can, is analyzed as the final step. First, a temperature profile of a module is obtained with the assumption that the preferred heat path is an exit through the pins. With the same boundary condition, another temperature profile is obtained, but this time with interfacial contact resistance between the chip pads and the substrate included. The importance of the interfacial bond between the pads and the substrate is indicated. Next, with the same boundary condition and an assumed contact resistance, steady-state temperature profiles are constructed for modules containing several chips at varying distances from each other. In this manner, the noninteraction of chips as a function of chip power is established. Comparisons are made of temperature profiles

The author is located at the IBM Components Division laboratory, Poughkeepsie, New York 12602.

* A briefer version of this paper was presented at the 1969 Electronic Components Conference, IEEE (G-PMP) and Electronic Industries Association, Washington, D.C., April 30-May 2, 1969.

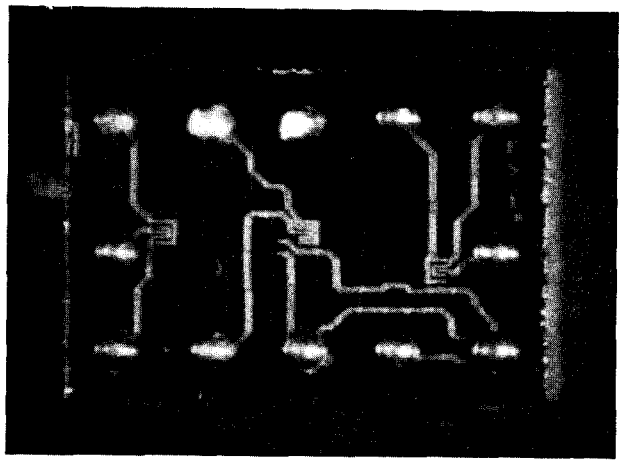


Figure 1 Photograph of a chip.

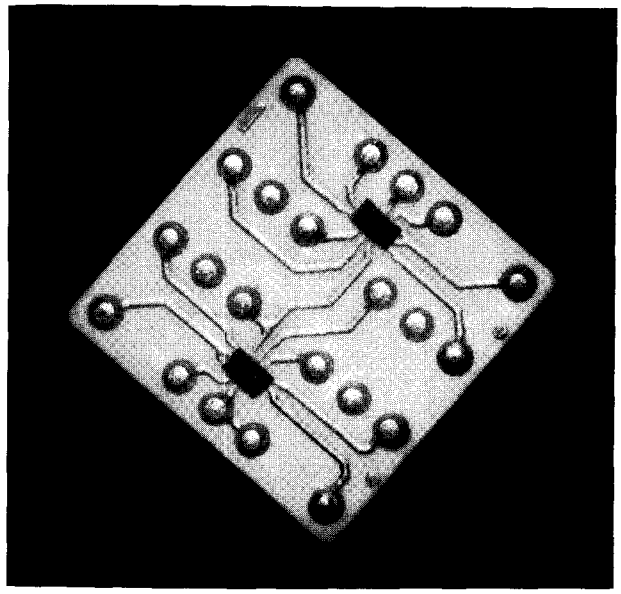


Figure 2 Chip joined to a ceramic substrate by the controlled collapse technique.

and thermal impedances obtained on modules with different heat paths and boundary conditions.

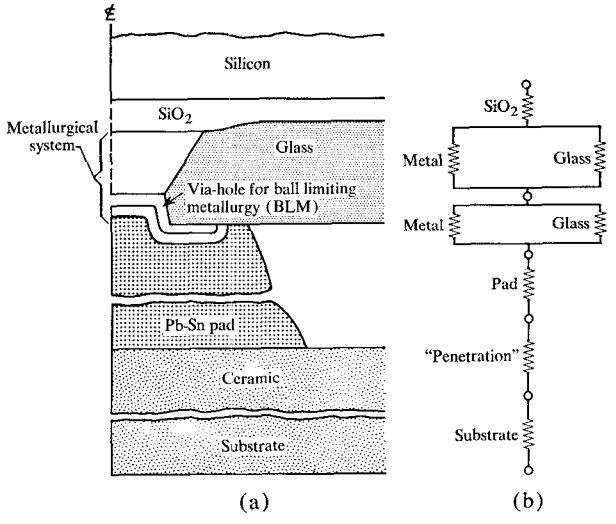
The thermal impedance offered by the substrate is also related to the pad dimensions, the number of pads and the chip size. The results are generalized to suggest an analytical approach for computing the thermal impedance for finite heat sources placed on a finite medium.

Thermal path in controlled chip collapse interconnections (pads)

Figure 1 shows a flip-chip with twelve pads. In Fig. 2, two chips are shown with pads joined to the ceramic substrate by the controlled collapse technique. With the enclosing can in place, the chips communicate thermally with the ultimate sink, i.e., the earth, via the pins and the walls of the module. The heat generated in the chips is transferred to the boundaries of the module by conduction and then transported away or exchanged by any one or more of the well-known modes of heat transfer; i.e., convection, change of phase, conduction or radiation. The actual transport of heat may be accomplished with gaseous or liquid media, depending upon, among other things, the heat flux crossing the boundaries of the module package.

It is clear that the temperature level of the heat-generating chips is determined by the temperature field established on the boundaries of the module in accordance with the nature of the thermal boundary layer between the module and its environment. The actual temperature rise between the heat-generating components in the chip and the boundaries of the module is, however, directly related to the "thermal impedance" offered by the conductive path connecting the chip to the boundaries. Thus, with a given ambient temperature, a lower thermal impedance will lead directly to a lower chip temperature—which is the objective of any cooling design.

Figure 3 (a) Cross-section of a flip-chip; (b) resistive thermal path.



In as much as as the controlled collapse bonding technique forms an important conductive path between the chip and the substrate, this study will primarily address the "thermal impedance" internal to the module as differentiated from the external thermal impedance between the module and its environment.

The thermal path between the heat-generating components in the chip and the substrate is shown in the cross-sectional view of Fig. 3. The origin of the various thermal

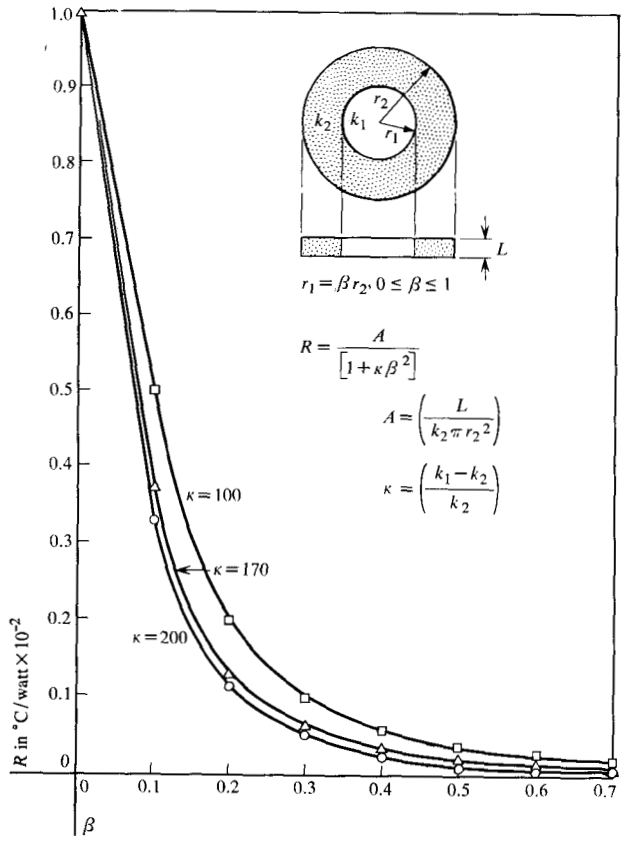


Figure 4 Thermal resistance around glass via hole.

resistances that together make up the total impedance of a module may be inferred from Figs. 3a and 3b. Figure 3a shows a schematic of the various strata of deposition[3] on the chip from the SiO_2 layer down to the bottom of the pad. Figure 3b is a simplified circuit diagram showing how the various resistances combine to form the resistive path of heat flow down to the bottom of the substrate.

Assuming that the chip is an isotherm and that the heat passes from the chip through the various layers shown in Fig. 3a into the substrate, one may use the formula $R = L/kA$ for calculating the resistance of each segment of the thermal path. Here A, the cross-sectional area of the pad, may be assumed to be the effective heat flow area as "seen" by the chip. Thus, with L denoting the thickness, and k the thermal conductivity, one may assess the relative magnitude of the resistances offered by each layer. It is found, for example, that the important parameters governing the thermal resistance between the device layer in a monolithic chip and the substrate are the diameter of the glass via hole, the pad dimensions and the thermal oxide thickness, in that order. Some of the results are shown in Fig. 4, where the thermal resistance in the area of the glass via hole is plotted against the ratio of the hole diameter to the diameter of the area normal to the direction of heat flow. The dimensionless parameter κ contains the thermal conductivities of both the material in the hole and the surrounding material.

The assumption that the chip is an isotherm and that the heat flow is one-dimensional may sometimes lead to erroneous results. This is particularly true when calculating the various thermal impedances. "Thermal impedance" is an arbitrarily-defined quantity that gives the difference in temperature between two points in a bounded space per unit rate of heat flow through that space, and not necessarily between the two points. Since the rate of heat flow in a given direction may be altered by the introduction of sources or sinks, the "thermal impedance," as defined, is not an intrinsic property of the material space but is a function of its boundary conditions. Consequently, the often used formula, $R = L/kA$ which gives the intrinsic resistance of a material with certain material and geometrical properties, cannot be expected to give meaningful results for computing specific thermal impedances on a complicated structure such as the monolithic chip or the module where the heat flow is three-dimensional. Instead, its counterpart $\Delta T/q$, where ΔT denotes the temperature difference and q the total rate of heat flow, should be used. We shall call the quantity $\Delta T/q$ thermal impedance and denote it by Z. Z will vary depending upon the boundary conditions and the location of sources and sinks. The notation R will be used to denote the intrinsic thermal resistance of materials.

Of course, there are sophisticated analytical approaches to obtain temperature distributions in solids. For example, one may use conformal mapping or Green's function[4] for obtaining temperature profiles in solids with point heat sources on its boundaries, or solve the interior problem of Dirichlet in Laplace's equation when the temperature distribution is specified on the boundaries. For the transient case, the three-dimensional diffusion equation may be used. However, most analytical solutions are for idealized bodies and are not of closed form. If numerical methods are used, the arithmetic is lengthy and involved. From an engineering point of view, it is desirable that the solutions be applicable to real situations and relatively amenable to quick quantitative results. For these requirements to be met, it is evident that one must resort to the computer as a tool.

In the following section, computer analyses of both the steady-state and transient temperature profiles on a flip-chip are shown. Various thermal impedances are computed from the temperature distributions.

Temperature profiles in joined chips

A description of the computer program[5] is given in the Appendix. The program is applicable to the three modes of heat transfer (conduction, radiation and convection) in a

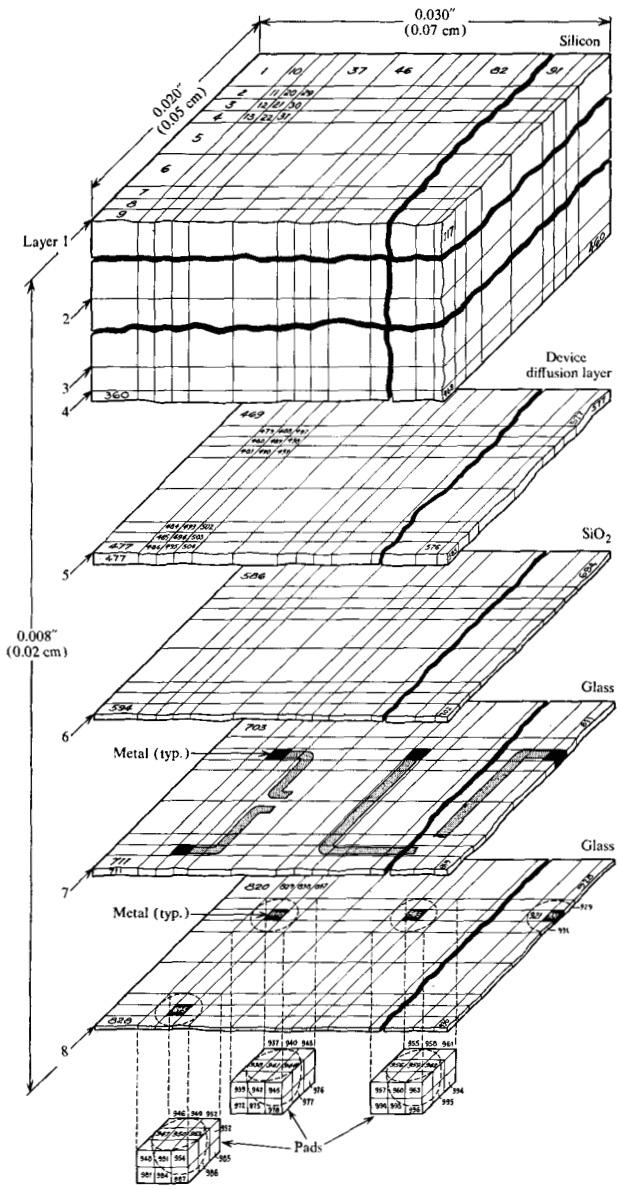


Figure 5 Chip node matrix.

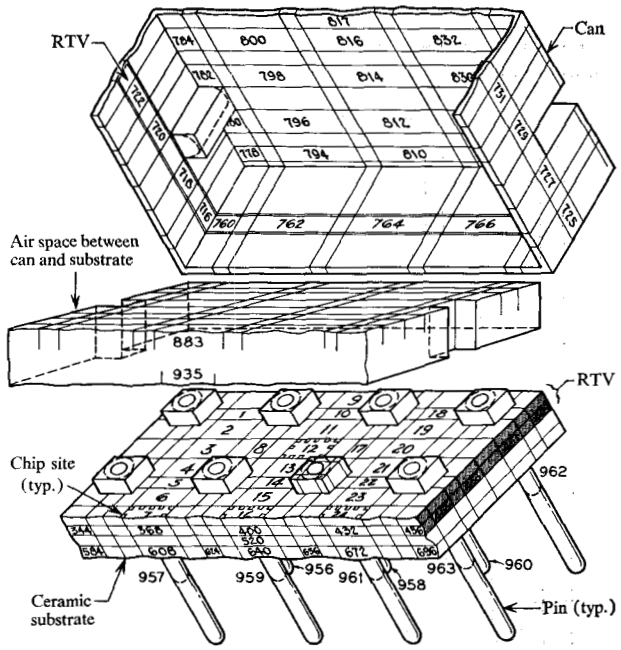


Figure 6 Module node matrix.

three-dimensional rectangular coordinate system for both steady-state and transient problems under various boundary conditions.

The "building blocks" of a given problem are the parallelepipeds or nodes. For example, Fig. 5 shows one quarter of a flip-chip divided into several nodes. Each of the six faces of each node can be subjected to different boundary conditions independently. Also, any node may be assigned heat generation and surface contact resistances.

Because of the nature of the problem of obtaining temperature profiles on a complicated structure such as the chip-mounted module, it is necessary to divide the analysis into two sections or programs. The first program is formulated to yield temperature distributions in flip-chips, while the second program is formulated to yield temperature distributions in modules.

Figures 5 and 6 show the mathematical models used for describing the chip and the module, respectively. Although for this study the chip has the dimensions of 0.040 in. × 0.060 in. × 0.008 in. (0.10 cm × 0.15 cm × 0.02 cm), while the module consists of a ceramic substrate of dimensions 0.50 in. × 0.50 in. × 0.062 in. (1.3 cm × 1.3 cm × 0.16 cm) and an aluminum can which is sealed with RTV ru bber, the results will be generalized to different sizes of chips and substrates. The pins in Fig. 6 protrude approximately 0.100 in. (0.25 cm) from the substrate and are made out of copper.

In Fig. 5, one quarter of the chip is divided into eight different slabs of appropriate materials and thicknesses, each slab being further divided into 117 nodes. Each of the three pads on the quarter chip is also divided into nine nodes. It will be noted in the same figure that the nodes in layer 5 (the device diffusion layer) represent device junctions on a real chip. The depth of layer 5 corresponds to the specified diffusion depths on a real chip.

Figure 6 depicts one half of a module which is divided into numerous nodes. The chip sites on the ceramic substrate are readily discernible.

A sample computer printout of steady-state temperature distribution on the chip is shown in Fig. 7. Note that the temperatures are indicated next to the node numbers.

CHIP- UNIFORM HEAT FLUX. 1 MW AT EACH POINT —IN THE DEVICE LAYER

TEMPERATURES IN DEGREES C

STEADY STATE INDICATED

S S	NODE	TEMPERATURES									
1	0.1240E 02	2	0.1254E 02	3	0.1253E 02	4	0.1255E 02	5	0.1259E 02		
6	0.1254E 02	7	0.1239E 02	8	0.1233E 02	9	0.1229E 02	10	0.1266E 02		
11	0.1256E 02	12	0.1254E 02	13	0.1258E 02	14	0.1270E 02	15	0.1267E 02		
16	0.1247E 02	17	0.1238E 02	18	0.1234E 02	19	0.1274E 02	20	0.1263E 02		
21	0.1242E 02	22	0.1266E 02	23	0.1281E 02	24	0.1281E 02	25	0.1260E 02		
26	0.1252E 02	27	0.1247E 02	28	0.1287E 02	29	0.1274E 02	30	0.1277E 02		
31	0.1282E 02	32	0.1298E 02	33	0.1301E 02	34	0.1285E 02	35	0.1278E 02		
36	0.1274E 02	37	0.1318E 02	38	0.1316E 02	39	0.1318E 02	40	0.1323E 02		
41	0.1337E 02	42	0.1350E 02	43	0.1347E 02	44	0.1345E 02	45	0.1344E 02		
46	0.1368E 02	47	0.1369E 02	48	0.1373E 02	49	0.1381E 02	50	0.1404E 02		
51	0.1436E 02	52	0.1450E 02	53	0.1455E 02	54	0.1457E 02	55	0.1404E 02		
56	0.1399E 02	57	0.1403E 02	58	0.1415E 02	59	0.1448E 02	60	0.1494E 02		
61	0.1514E 02	62	0.1522E 02	63	0.1526E 02	64	0.1427E 02	65	0.1422E 02		
66	0.1426E 02	67	0.1439E 02	68	0.1475E 02	69	0.1524E 02	70	0.1546E 02		
71	0.1555E 02	72	0.1559E 02	73	0.1455E 02	74	0.1452E 02	75	0.1458E 02		
76	0.1468E 02	77	0.1509E 02	78	0.1552E 02	79	0.1575E 02	80	0.1584E 02		
81	0.1588E 02	82	0.1511E 02	83	0.1516E 02	84	0.1522E 02	85	0.1532E 02		
86	0.1558E 02	87	0.1600E 02	88	0.1827E 02	89	0.1638E 02	90	0.1635E 02		
91	0.1583E 02	92	0.1593E 02	93	0.1599E 02	94	0.1607E 02	95	0.1624E 02		
96	0.1656E 02	97	0.1675E 02	98	0.1683E 02	99	0.1687E 02	100	0.1606E 02		
451	0.1688E 02	452	0.1763E 02	453	0.1775E 02	454	0.1774E 02	455	0.1720E 02		
456	0.1761E 02	457	0.1847E 02	458	0.1869E 02	459	0.1878E 02	460	0.1709E 02		
461	0.1799E 02	462	0.1809E 02	463	0.1810E 02	464	0.1748E 02	465	0.1781E 02		
466	0.1968E 02	467	0.1912E 02	468	0.1922E 02	469	0.1270E 02	470	0.1255E 02		
471	0.1250E 02	472	0.1259E 02	473	0.1270E 02	474	0.1262E 02	475	0.1232E 02		
476	0.1218E 02	477	0.1211E 02	478	0.1266E 02	479	0.9772E 01	480	0.9455E 01		
481	0.9763E 01	482	0.1270E 02	483	0.1265E 02	484	0.9617E 01	485	0.9143E 01		
486	0.9073E 01	487	0.1270E 02	488	0.9521E 01	489	0.8444E 01	490	0.9540E 01		
491	0.1277E 02	492	0.1275E 02	493	0.9414E 01	494	0.8191E 01	495	0.8813E 01		
496	0.1287E 02	497	0.9935E 01	498	0.9824E 01	499	0.9967E 01	500	0.1297E 02		
501	0.1306E 02	502	0.9925E 01	503	0.9460E 01	504	0.9397E 01	505	0.1328E 02		
506	0.1315E 02	507	0.1313E 02	508	0.1322E 02	509	0.1349E 02	510	0.1364E 02		
511	0.1351E 02	512	0.1343E 02	513	0.1340E 02	514	0.1375E 02	515	0.1362E 02		
516	0.1368E 02	517	0.1377E 02	518	0.1417E 02	519	0.1470E 02	520	0.1511E 02		
521	0.1522E 02	522	0.1526E 02	523	0.1397E 02	524	0.1076E 02	525	0.1049E 02		
526	0.1090E 02	527	0.1452E 02	528	0.1554E 02	529	0.1639E 02	530	0.1658E 02		
531	0.1666E 02	532	0.1416E 02	533	0.1059E 02	534	0.9415E 01	535	0.1073E 02		
536	0.1475E 02	537	0.1588E 02	538	0.1680E 02	539	0.1701E 02	540	0.1710E 02		
541	0.1455E 02	542	0.1118E 02	543	0.1086E 02	544	0.1133E 02	545	0.1508E 02		
546	0.1613E 02	547	0.1704E 02	548	0.1725E 02	549	0.1734E 02	550	0.1527E 02		
551	0.1520E 02	552	0.1523E 02	553	0.1538E 02	554	0.1580E 02	555	0.1643E 02		
556	0.1699E 02	557	0.1714E 02	558	0.1722E 02	559	0.1626E 02	560	0.1660E 02		
561	0.1670E 02	562	0.1675E 02	563	0.1669E 02	564	0.1706E 02	565	0.1758E 02		
566	0.1773E 02	567	0.1781E 02	568	0.1691E 02	569	0.1772E 02	570	0.1783E 02		
571	0.1783E 02	572	0.1731E 02	573	0.1764E 02	574	0.1855E 02	575	0.1877E 02		
576	0.1886E 02	577	0.1719E 02	578	0.1813E 02	579	0.1822E 02	580	0.1824E 02		
581	0.1754E 02	582	0.1786E 02	583	0.1903E 02	584	0.1927E 02	585	0.1937E 02		
586	0.1827E 02	587	0.1255E 02	588	0.1250E 02	589	0.1254E 02	590	0.1270E 02		
591	0.1262E 02	592	0.1232E 02	593	0.1218E 02	594	0.1211E 02	595	0.1266E 02		
596	0.9005E 01	597	0.8733E 01	598	0.9014E 01	599	0.1270E 02	600	0.1265E 02		
801	0.1781E 02	802	0.1691E 02	803	0.1780E 02	804	0.1787E 02	805	0.1787E 02		
806	0.1731E 02	807	0.1764E 02	808	0.1855E 02	809	0.1877E 02	810	0.1884E 02		
811	0.1714E 02	812	0.1806E 02	813	0.1813E 02	814	0.1813E 02	815	0.1754E 02		
816	0.1786E 02	817	0.1903E 02	818	0.1927E 02	819	0.1937E 02	820	0.1270E 02		
821	0.1254E 02	822	0.1250E 02	823	0.1254E 02	824	0.1270E 02	825	0.1262E 02		
826	0.1232E 02	827	0.1218E 02	828	0.1211E 02	829	0.1266E 02	830	0.5593E 01		
831	0.5522E 01	832	0.5598E 01	833	0.1270E 02	834	0.1265E 02	835	0.5495E 01		
836	0.5347E 01	837	0.5267E 01	838	0.1270E 02	839	0.5550E 01	840	0.6315E 01		
841	0.5598E 01	842	0.1277E 02	843	0.1275E 02	844	0.5474E 01	845	0.6130E 01		
846	0.5221E 01	847	0.1286E 02	848	0.5863E 01	849	0.5593E 01	850	0.5479E 01		
851	0.1287E 02	852	0.1300E 02	853	0.5626E 01	854	0.5481E 01	855	0.5199E 01		
856	0.1328E 02	857	0.1315E 02	858	0.1313E 02	859	0.1322E 02	860	0.1349E 02		
861	0.1364E 02	862	0.1351E 02	863	0.1343E 02	864	0.1340E 02	865	0.1375E 02		
866	0.1362E 02	867	0.1368E 02	868	0.1376E 02	869	0.1417E 02	870	0.1470E 02		
871	0.1511E 02	872	0.1522E 02	873	0.1526E 02	874	0.1397E 02	875	0.1127E 01		
876	0.6067E 01	877	0.6186E 01	878	0.1452E 02	879	0.1554E 02	880	0.1639E 02		
881	0.1658E 02	882	0.1666E 02	883	0.1416E 02	884	0.6126E 01	885	0.7010E 01		
886	0.6187E 01	887	0.1479E 02	888	0.1588E 02	889	0.1680E 02	890	0.1701E 02		
891	0.1710E 02	892	0.1450E 02	893	0.6304E 01	894	0.6243E 01	895	0.6231E 01		
896	0.1308E 02	897	0.1615E 02	898	0.1704E 02	899	0.1725E 02	900	0.1734E 02		
901	0.1527E 02	902	0.1520E 02	903	0.1523E 02	904	0.1538E 02	905	0.1580E 02		
906	0.1643E 02	907	0.1698E 02	908	0.1714E 02	909	0.1722E 02	910	0.1626E 02		
911	0.1660E 02	912	0.1670E 02	913	0.1675E 02	914	0.1669E 02	915	0.1706E 02		
916	0.1758E 02	917	0.1773E 02	918	0.1781E 02	919	0.1691E 02	920	0.1784E 02		
921	0.1790E 02	922	0.1790E 02	923	0.1731E 02	924	0.1764E 02	925	0.1855E 02		
926	0.1877E 02	927	0.1886E 02	928	0.1714E 02	929	0.1802E 02	930	0.1813E 02		
931	0.1808E 02	932	0.1754E 02	933	0.1786E 02	934	0.1903E 02	935	0.1927E 02		
936	0.1937E 02	937	0.4130L 01	938	0.4145E 01	939	0.4132E 01	940	0.4161E 01		
941	0.4442E 01	942	0.4165E 01	943	0.4166E 01	944	0.4183E 01	945	0.4173E 01		
946	0.4052E 01	947	0.4019E 01	948	0.3928E 01	949	0.4094E 01	950	0.4316E 01		
951	0.3965E 01	952	0.4121E 01	953	0.4089E 01	954	0.4000E 01	955	0.4501E 01		
956	0.4934E 01	957	0.4536E 01	958	0.4565E 01	959	0.4894E 01	960	0.4597E 01		
961	0.4597E 01	962	0.4627E 01	963	0.4630E 01	964	0.1792E 02	965	0.1794E 02		
966	0.1794E 02	967	0.1795E 02	968	0.1797E 02	969	0.1797E 02	970	0.1501E 01		
971	0.1505E 01	972	0.1501F 01	973	0.1507E 01	974	0.1536E 01	975	0.1508E 01		
976	0.1506E 01	977	0.1513E 01	978	0.1507E 01	979	0.1488E 01	980	0.1484E 01		
981	0.1470E 01	982	0.1496E 01	983	0.1516E 01	984	0.1477E 01	985	0.1497E 01		
986	0.1494E 01	987	0.1480E 01	988	0.1561E 01	989	0.1568E 01	990	0.1566E 01		
991	0.1572E 01	992	0.1606E 01	993	0.1577E 01	994	0.1574E 01	995	0.1591E 01		
996	0.1579E 01	997	0.0000F-38	998	0.0000E-38	999	0.0000E-38	***	0.0000E-38		

Figure 7 Sample computer print-out of steady-state temperature distribution for a pad-mounted chip.

Figure 8 is a steady-state temperature profile obtained from the computer results. The results shown in Figures 7 and 8 and others are discussed later in this paper.

It should be evident from the description of the program given above that many types of boundary conditions may be imposed upon the chip or the module. In the case of the flip-chip where heat leaves the chip by way of the pads, the lower surfaces of the pads were held at a constant temperature. Different cases were simulated with a different number of heat sources by assigning random heat generation at the proper nodes in the device diffusion layer. Similarly, the module shown in Fig. 6 was simulated under various boundary conditions including placement of heat sinks at the pins, the can, and both. Heat sinks at the pins for example, were simulated by holding at a constant temperature the outer surfaces of the nodes representing the pins, whereas in the case of heat sinks at both the pins and the can, both the pins and the can were held in an infinite heat reservoir of constant temperature.

As indicated earlier, thermal impedances between points are calculated from the relationship $\Delta T/q$, where q is the power dissipation. Now, it is clear that the "q" that must be used in calculating the various thermal resistances at the chip level is four times the total power dissipation assigned to the quarter-chip in the computer analysis. This is because the temperature distribution obtained on a quarter chip dissipating q watts is the image of the temperature distribution on the whole chip, but dissipating $4q$ watts; provided, of course, that the boundary conditions as well as the location and strengths of the corresponding heat sources on all four quarters of the chip are the same. These considerations follow from the fact that $\partial T/\partial n = 0$ at the boundaries of the quadrants dividing the chip into four equal parts. For the same reasons, the q that must be used in calculating the various thermal impedances at the module level is twice the total power dissipation assigned to the half-module in the computer analysis.

Steady-state case

The results shown in Fig. 7 pertain to a pad-mounted chip (a "flip-chip"). That the chip is pad-mounted may be inferred from the pad node temperatures relative to other node temperatures. Thus, nodes 970-996 (see Fig. 7) are kept at an arbitrary temperature of 1.5°C (in a heat reservoir), while all other nodes assume higher steady-state temperatures. In other words, heat leaves the chip through the pads.

As is specified and indicated in Fig. 7, all nodes in the device layer dissipate one milliwatt each. Since there are 117 nodes in the layer, the power dissipation for the quarter chip is 117 mW. For the whole chip then, the total power dissipation q is 468 mW, as explained before. Knowing q, one can now compute the thermal impedance between any two points or nodes on the chip by simply obtaining the temperature difference ΔT between those points from the node temperatures given in Fig. 7 and dividing it by q.

It is clear that in calculating the thermal impedances it is immaterial what level of power is assigned to the chip for a particular configuration of heat sources on the chip. In other words, for a particular chip, the thermal impedances are independent of power. Also, the heat sink temperature at the pads can be of any value since all other node temperatures adjust their levels relative to that reference temperature.

Several generalizations may be made from the results shown in Fig. 7. First, even with uniform distribution of heat sources of equal strength throughout the device layer,

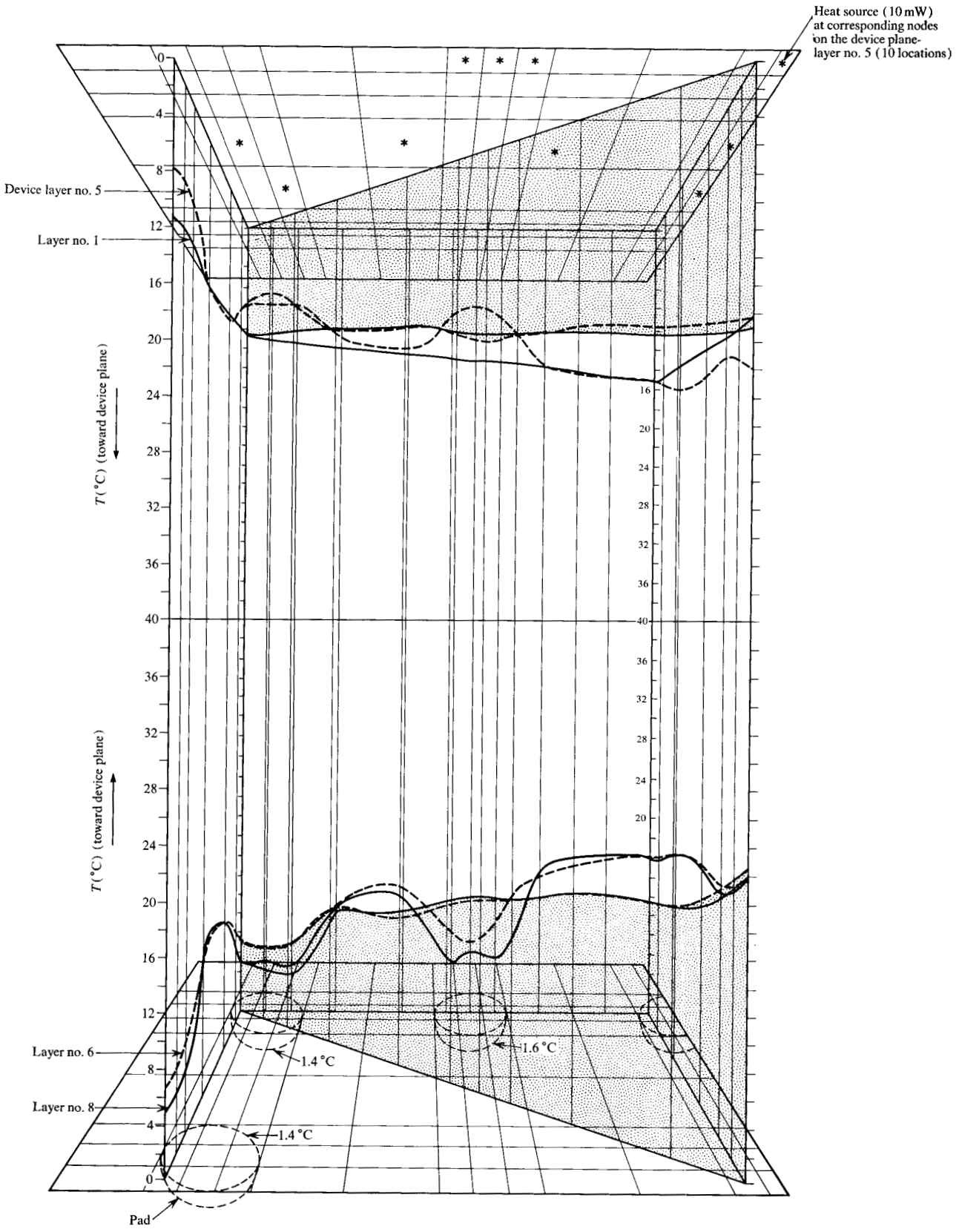


Figure 8 Steady-state chip temperature profile. (The shaded profile curve passes diagonally through the "chip"; the unshaded curves define the profile along the perimeter.)

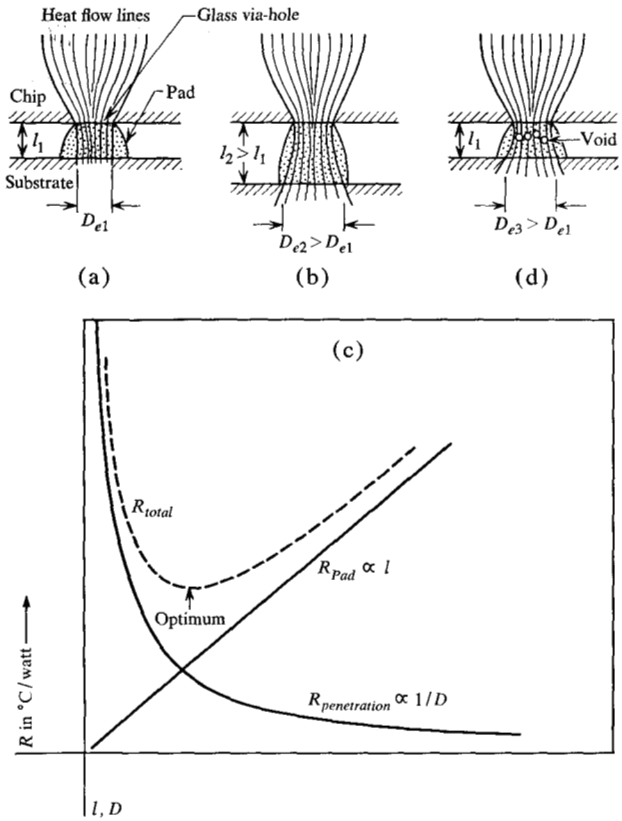


Figure 9 Heat flow lines in pads and penetration resistance.

the chip temperature is not uniform. In other words, it does not necessarily follow that by definition a monolithic chip is an isotherm. This can be explained in terms of interaction between sources and sinks, where in this case the sinks are the pads. In fact, it is shown later in this study that with fewer heat sources, of equivalent strengths and randomly scattered in the device layer, the non-uniformity of temperature field becomes more pronounced. For example, the gradients around the pads become relatively large (see Fig. 8). However, for low powered chips, the temperature profile becomes flatter, resembling an isotherm.

Second, except for the device layer itself, the back side of the chip (nodes 1-117) is, in general, at a higher temperature than the other layers. For example, the average temperature of nodes 11, 12, 13, 20, 21, 22, 29, 30, and 31 in layer 1 (see Fig. 7) is 12.7°C, whereas the average temperature of nodes 830, 831, 832, 839, 840, 841, 848, 849, and 850 in layer 8 is 5.7°C. Coupling this finding with the result that the average temperature of nodes located far from the pads (for example, nodes 920, 921, 922, 929, 930, and 931 in layer 8) is approximately the same as the average temperatures of the nodes at the corresponding location on the back side of the chip, it becomes evident that the heat traverses a path from a heat dissipating junction to the silicon and then back down through the various layers to the pads. This is because the impedance through the silicon towards the back of the chip is less than that laterally through the very thin layer of thermal oxide (SiO_2) and glass (which are two orders of magnitude less conductive than silicon) toward the pads.

Third, a closer examination of temperature distributions in the neighborhood of the pads reveals that large temperature gradients exist at all layers above the glass via hole. For example, node 489, which is in layer 5 directly above the metal in layer 8 (see Fig. 5), is at 8.4°C while its neighboring nodes 488 and 490 are at 9.5°C. The same pattern is observed at all layers above pad locations. (The gradients are even higher in the temperature profile shown in Fig. 8). In comparison, at other nodes the gradients are much smaller.

If one were to draw the corresponding heat flow lines in the neighborhood of the glass holes, one would obtain a potential field similar to that observed near a small orifice in a liquid filled tank. This phenomenon, known as "thermal pinch," contributes an additional impedance, which is sometimes referred to as "volume impedance." It is evident then that the size of the glass hole is important in determining the thermal impedance of pad-mounted chips (see Figs. 4 and 9).

With the temperature distribution given in Fig. 7, the various thermal impedances may now be readily calculated. For example, if the impedance from the device diffusion layer to the bottom of the pads is desired, one may divide the difference between the average temperature in the device layer and at the bottom of the pads by the total power dissipation of the chip, which yields 18°C/watt for the example cited above.

In addition to the thermal impedance from the chip level to the surrounding environment, which we shall denote by $Z_{T-\text{chip}}$ other contributions of interest are the spreading impedance $Z_{T-\text{sp}}$ across a chip, and the "junction" impedance $Z_{T-\text{j}}$ from a heat source to its immediate environment in the chip. The latter is readily simulated on the computer. It may be as high as 250°C per watt of junction dissipation depending upon the size of the diffused junction area.

The $Z_{T-\text{sp}}$ is readily calculated from the results shown in Fig. 7. For example, the temperature at the center of the chip (node 585 in layer 5) is 19.37°C. The average temperature above the pad on the short edge of the chip (nodes 484-486, 493-495, 502-504) is 9.23°C. Now, calculating the impedance between the center of the chip and the pad which are, say, 25 mils (0.06 cm) apart we obtain 0.87°C-watt^{-1} mil^{-1} (348°C-watt^{-1}-cm^{-1}). This means that for a chip dissipating, say, 100 milliwatts total, the chip would be approximately 2°C hotter relative to the locations above the pads in the device layer. If there is a

transistor located at the center of the chip which is dissipating 5 milliwatts, its junction temperature would be hotter by another $0.005 \times Z_{T-j}$, where the junction impedance has the units of °C/watt.

It should be remembered that the Z_{T-sp} is calculated from the temperature distribution in device layer 5; i.e., the layer in which heat is generated. If, on the other hand, one were to contend that the temperature distribution on the back of the chip is indicative of the temperature distribution in the device junction layer, the calculated Z_{T-sp} would be smaller by a factor 2.5 (see node temperatures in layer 1).

If we express these results in terms of the familiar thermal impedance values, what is suggested here is that in addition to the Z_{T-chip} (18°C/watt) calculated previously, there is also the Z_{T-sp} (21.7°C/watt) to the center of the chip. Of course, the impedances to all other locations on the chip are smaller.

Although the temperature differences from one junction to another may become negligibly small for low power dissipating chips, for higher power dissipating chips the situation is different. To illustrate the point, consider the temperature profile shown in Fig. 8. The case shown there corresponds to a hypothetical flip-chip of the same dimensions as in the previous example, but with 40 devices dissipating 400 milliwatts in total. On the quarter chip shown there are 10 junctions, each dissipating 10 milliwatts. Three of the junctions are located near the center of the chip (three of them are behind the shaded profiles). The node mesh plan is the same as that shown in Fig. 5, so that the relative locations of the heat sources may be inferred. The top and bottom planes depict the back and pad sides of the chip, respectively. For clarity, only the temperature profiles on the periphery of the chip, and along the diagonal are shown. Also, only the temperature profiles on four layers are shown for purposes of illustration. The schematic may be considered to represent a chip pulled apart at the device layer such that the upper portion is viewed from the bottom and the lower portion from the top. Since the maximum temperatures are found in the device layer the directions of increasing temperature for the two portions are opposing as indicated in Fig. 8.

It is clear from Fig. 8 that the temperature profile in the device layer (5) is far from being smooth. In contrast, the profile on the back of the chip (layer 1) is relatively smooth. But what is more significant is that the Z_{T-sp}, calculated in the same manner as before, is now 2.17°C-watt^{-1}-mil^{-1} (868°C-watt^{-1}-cm^{-1}). In other words, the impedance to the center of the chip, which turns to be 54.4°C/watt in this case,[6] is more than three times the Z_{T-chip} (18°C/watt), which of course remains the same. Thus, a pad-mounted chip with heat sources randomly located as shown in Fig. 8 cannot be considered to be an isotherm. Therefore, Z_{T-chip} alone would be inadequate in estimating the cooling requirements. An average Z_{T-sp} must also be specified.

The wide variation possible in boundary conditions, and the location of sources and sinks, precludes the use of general rules to predict Z_{T-sp}; fortunately it is not difficult to investigate each case individually with the aid of the computer program.

Figure 8 summarizes the earlier discussion on temperature gradients near the pads. Some other observations of interest are the temperature distributions in the neighborhood of the metals and in the pads.

It will be noticed in the temperature profile of layer 8 in Fig. 8 that there is a "bump" over nodes 840, 845 and 885 right above the pads. That is, temperatures at those nodes are higher than in the neighboring nodes, in contrast to the general trend in all layers above the pads. This same occurrence is observed in all cases, though to a small degree for low power dissipating chips. This observation may be explained by considering that a large portion of the heat is channeled out from the chip through the metal due to the very high conductivity of the same, in comparison with the very low conductivity of the surrounding glass. As a result, the temperature of the metals becomes high.

It can be calculated from the temperature distribution in Fig. 7 that the contribution to thermal impedance of the pads is ≈9°C/watt, which is fifty precent of the total impedance from the device layer to the bottom of the pads. It is of interest to analyze the parameters affecting the nature of the heat path in pads since it has been found that the temperature distributions in the pads are not uniform. The nodes located centrally in the pads, i.e., in the line of sight of the via hole, assume higher temperatures than the nodes located at the periphery. This suggests that most of the heat leaves the pads centrally, rather than spreading out immediately. Thus the effective cross section of the thermal path at a pad is less than the physical contact area of the pad. With heat flow assumed to be as shown in Fig. 9a, the "penetration" thermal impedance into the substrate has a strong inverse dependence on the diameter D of this *effective* cross section. (In the case of a round heat source on an infinite medium, $R \sim 1/D$).[7] In order to decrease the penetration impedance, the D may be increased by making the pads higher as shown in Fig. 9b. The optimum pad height may then be found by determining the decrease in penetration impedance against the increase in the pad impedance due to the increase in the height l (Fig. 9c).

Another parameter which can alter the heat path in pads is the void or gas pocket content. Computer simulation of pads with various degrees of void content shows that it is the number of voids and their locations in the pads rather than the total volume of voids which influences

the heat flow path. It is found that small voids scattered across the pad have the effect of "spreading" the heat flow lines and thus increasing the effective thermal contact area on the substrate (see Fig. 9d). Of course, the presence of voids themselves contributes to an increase in the thermal impedance of pads. However, one big void having the same volume as the combined volume of smaller voids offers greater thermal impedance with a degrading effect on the thermal contact areas as well. Voids, big or small, found at the contact surface area between the pad and the substrate can create high interfacial impedance and hence high overall impedance from the interface to the pins and the can. The interfacial impedance is discussed again later in connection with temperature profiles in modules.

Finally, we note that variations in the thermal resistance of pads as a function of process tolerances may be estimated. Assuming, for a first approximation, that the height of the pad, l, and volume of solder, V, are independent,[8] we may recast the relationship $R = L/kA$ into the form

$$R = \frac{l}{kA}\cdot\frac{l}{l} = \frac{l^2}{kV},$$

where $V = Al$ and A is the cross-sectional area.

Variations in the thermal resistance of pads may now be computed as a function of the variations in l and V.

Rewriting the relationship above in the functional form, we have

$$R = f(l,V);$$

then

$$\Delta R = f_{,l}\Delta l + f_{,V}\Delta V,$$

where

$$f_{,l} \equiv \frac{\partial f}{\partial l}, \quad \text{etc.}$$

Or, using the notation $\sigma_R = \Delta R$, $\sigma_l = \Delta l, \sigma_V = \Delta V$, and forming the square of the deviation, we obtain

$$\sigma_R^2 = f_{,l}^2\sigma_l^2 + f_{,V}^2\sigma_V^2.$$

Dividing by R, and rearranging the terms, we have

$$\left(\frac{\sigma_R}{R}\right)^2 = \left(\frac{lf_{,l}}{R}\right)^2\left(\frac{\sigma_l}{l}\right)^2 + \left(\frac{Vf_{,V}}{R}\right)^2\left(\frac{\sigma_V}{V}\right)^2.$$

But,

$$\frac{lf_{,l}}{R} = \frac{l}{R}\cdot\frac{2l}{V} = 2,$$

$$\frac{Vf_{,V}}{R} = -\frac{V}{R}\cdot\frac{l^2}{V^2} = -1;$$

$$\therefore \quad \left(\frac{\sigma_R}{R}\right)^2 = 4\left(\frac{\sigma_l}{l}\right)^2 + \left(\frac{\sigma_V}{V}\right)^2.$$

This result is valid for one pad. For N pads, we obtain

$$\left(\frac{\sigma_R}{R}\right)^2 = \frac{1}{N}\left[\left(4\,\frac{\sigma_l}{l}\right)^2 + \left(\frac{\sigma_V}{V}\right)^2\right],$$

so that the deviation σ_R from the mean R, as a function of the deviations σ_l, σ_V from the mean values of l and V, is

$$\sigma_R = \frac{1}{N^{\frac{1}{2}}}\left[4\left(\frac{\sigma_l}{l}\right)^2 + \left(\frac{\sigma_V}{V}\right)^2\right]^{\frac{1}{2}}R,$$

where R may be the measured or calculated thermal resistance of the pads.

• *Transient case*

The chip temperature profiles discussed in the last section will not apply if the junction power switching period[9] is much shorter than the time constant of the device. This and considerations such as the heating of the junction during testing make it necessary to investigate the transient response of chips to temperature.

It is well known that the temperature history of a point in a semi-infinite slab subjected to an effective temperature potential is governed by the error function. On the other hand, the thermal response of a finite slab subjected to a temperature potential on its well-defined boundaries is governed by an exponential function.[10] For a three-dimensional finite chip where the boundary conditions are of the "mixed" type, the situation is much different. The classical heat conduction solutions become inapplicable. For example, the well known relationship, $\tau = L^2/\alpha$, used in computing the time constant τ for a homogeneous semi-infinite body of characteristic length L and thermal diffusivity α, can lead to erroneous results when applied to a composite finite body such as a chip.

In fact, using $\tau = L^2/\alpha$, where $\alpha = k/\rho\, c_p$ we find for a flip-chip subjected to sudden heating ($k = 0.1$ cal-sec^{-1}-cm-°C^{-1}, $\rho = 2.33$ g-cm^{-3}, $c_p = 0.16$ cal-g^{-1}-°C^{-1}) that $\tau = L^2/0.268$. Taking the representative length L to be the thickness of the chip, 8 mils (0.02 cm), we see that τ turns out to be 1.5 milliseconds. This suggests that in approximately 5 msec, the chip will have reached approximately 95% of its steady-state temperature level. This will be compared with the computer results discussed below.

Figure 10 shows a plot of temperature versus time for a flip-chip with 12 pads. Initially, the chip is held at 1°C while the boundary conditions are such that the pads are kept at 1°C at all times. The remaining boundaries of the chip are insulated. Under these conditions, node 576 at the center of the device layer of the chip is assigned a heat strength of 100mW. Some of the transient temperatures of node 576 at other times, as well as the steady-state temperature, are reproduced in the inset in Fig.10.

Note that after approximately 60 μsec, the curve in Fig. 10 flattens out asymptotically. To a first approximation,

the temperature becomes *almost* a linear function of time. By extrapolating on the basis of linearity to the steady-state temperature obtained separately with the steady-state computer program, it is found that the total time required to reach steady-state conditions is 0.5 msec.

It will be observed that the functional relationship shown in Fig. 10 is complicated. This is expected, since the function represents the transient response of a complicated chip. It will be recalled that the chip is made up of several materials of different physical properties. Furthermore, the heat source is located inside the various layers of materials (see layer 5 in Fig. 5).

Although it is difficult to fit one equation through all the calculated points, it is of interest to analyze portions of the curve and then arrive at an approximate relationship. The nonlinear portion of the curve may be represented by the functional relationship

$$T = f(e^{-t/\tau'}),$$

where τ' is approximately 20 μsec. Interpreting τ' to be the "time constant" for this portion of the curve, we may write

$$(T_f - T) = (T_f - T_a)\, e^{t-/\tau'} \text{ for } t < 3\tau',$$

where T_f denotes the "final" temperature at which the temperature-time relationship becomes approximately linear. T_a is the ambient temperature, and T denotes the temperature at a given time t.

Now, at $t \approx 3\tau'$, $T \approx T_f$, and for $t > 3\tau'$, we may represent the approximate linear portion of the curve by

$$T = T_f + 0.083\ (t - 3\ \tau'), \text{ for } t > 3\ \tau' \text{ and } T \leq T_s$$

where the slope is obtained from the constructed curve.

It is clear from the formulation above that once the T_f is known, the temperature at a junction at any given time may be predicted. Here, $T_f \approx 56$°C, and the steady-state temperature $T_s = 100$°C, so that $T_f/T_s = 0.56$, which is a constant of the transient profile of the chip.

Substituting $T_f = 0.56\ T_s$ and rewriting the equations above, we have

$$T = T_a e^{-t/\tau'} + 0.56\ T_s(1 - e^{-t/\tau'}),\ t < 3\tau';$$

$$T = 0.56\ T_s + 0.083\ (t - 3\tau'),\ t > 3\tau',\ T_{\max} = T_s\ .$$

Employing the step function,* $u(t - 3\tau')$, we may combine these relations to obtain

$$T = T_a e^{-t/\tau'} + 0.56\, T_s(1 - e^{-t/\tau'}) + 0.083\,(t - 3\tau')\,u(t - 3\tau'),$$

where

$\tau' = 20$ μsec and $T_{\max} = T_s$.

* $u \equiv \begin{cases} 0, & \text{when } t < 3\tau' \\ 1, & \text{when } t > 3\tau'. \end{cases}$

t (μsec)	0	5.4	10.8	21.6	32.4	43.2	54.0	64.8	∞
T(°C)	1.0	20.5	28.4	38.3	44.6	48.9	52.0	54.4	99.9

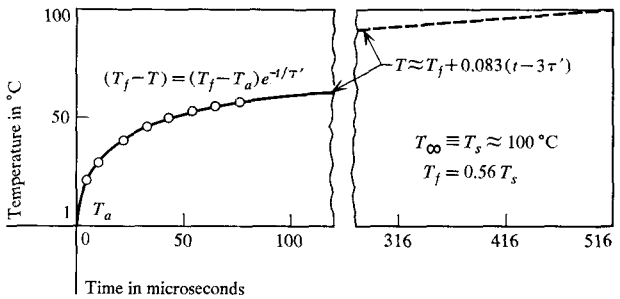


Figure 10 Transient junction temperature profile.

Thus, once the steady-state temperature T_s is known (from measurements or otherwise), an approximate transient profile may be constructed with the aid of the equation derived above.

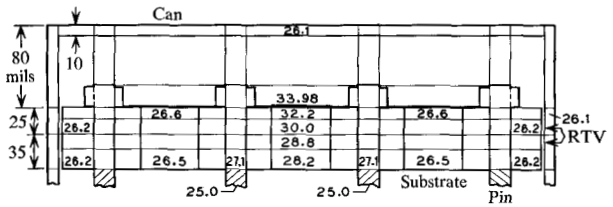
It can be shown that the relation above yields $t_s = 0.516$ msec when $T = T_s$. Thus, we see that the junction attains half its steady-state temperature in $\approx 1/9$ the total time required for it to reach the steady state.

Comparing these results with the theoretical result stated before ($t_s = 5$ msec), we see that calculations based on the assumption that the chip is a homogeneous slab may differ by as much as an order of magnitude from the results obtained on a model representing the real chip more precisely.

Module temperature profiles

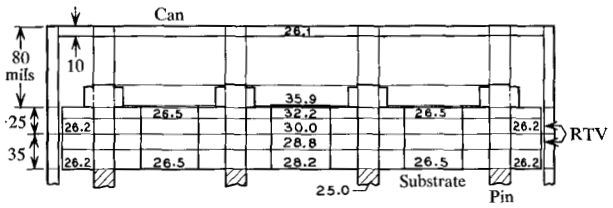
Temperature profiles on a module are simulated by assigning power dissipations to any one of the nodes 1-24 shown in Fig. 6. The chip sites indicated in the figure may be considered to consist of the pads on a chip. Thus, (the equivalent of) these pads act as heat sources on the substrate. Accordingly, thermal impedances to any desired point in the module, such as to the pins or the can, are found with respect to these heat sources by dividing the temperature difference between the respective points by the chip power dissipation. Then the total thermal impedance is found by adding to these impedances the $Z_{T\text{-chip}}$, i.e., the impedance from the chip level to the bottom of the pads.

Figure 11a shows a partial temperature distribution through the cross section of a module which is cooled at the pins.[11] The twelve-pad chip is placed centrally on the module and, for the example shown, the chip power dissipation is 400 milliwatts. The thermal impedance from the pads to the bottom of the substrate below the chip is, for example, (33.98-28.2) /0.4 = 14.4°C/watt. (Note that the substrate is not of uniform temperature.)



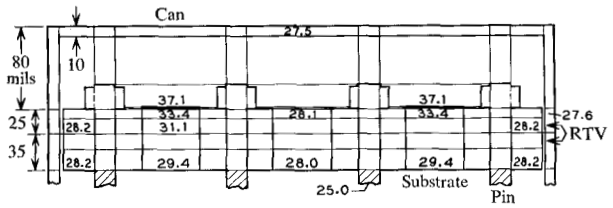
Heat sink: Pins
Chips: (1) at center (no pad contact resistance)
Module power: $1 \times 400 = 400$ mW

(a)



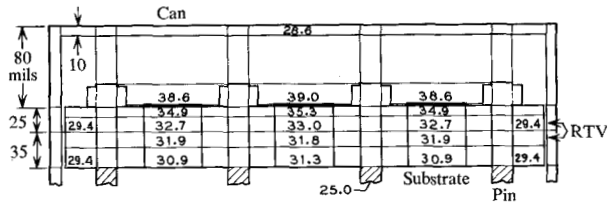
Heat sink: Pins
Chips: (1) at center
Module power: $1 \times 400 = 400$ mW

(b)



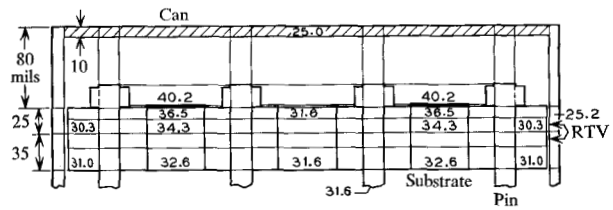
Heat sink: Pins
Chips: (2) at left, right
Module power: $2 \times 400 = 800$ mW

(c)



Heat sink: Pins
Chips: (3) at left, center, right
Module power: $3 \times 400 = 1.2$ watts

(d)



Heat sink: Can
Chips: (2) at left, right
Module power: $2 \times 400 = 800$ mW

(e)

Indicates heat sink
Note: Temperatures are in °C

Figure 11 Temperature distributions in modules.

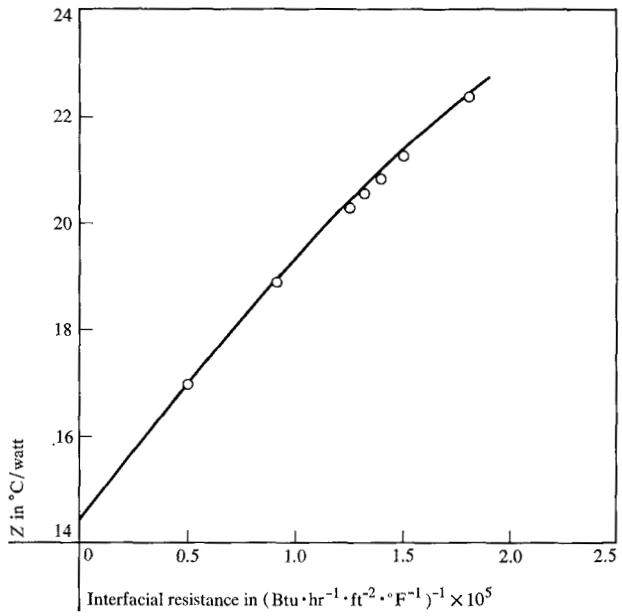


Figure 12 Module thermal impedance vs. interfacial resistance between twelve pads and substrate.

Figure 11b shows the temperature distribution in a module subjected to the same conditions as in Fig. 11a, but with an assumed included interfacial surface conductance of 1.74×10^5 Btu-hr^{-1}-ft^{-2}-°F^{-1} (100 watts-cm^{-2}-°C^{-1}) between the pads and the substrate.[12] All cases depicted in Fig. 11, b through e, include the same interfacial resistance (reciprocal of conductance).

It is observed in Fig. 11b that the assumed thermal contact resistance introduces an additional resistance of approximately 5°C/watt. It is found experimentally as well that the nature of the bond between the pad and the substrate affects the interfacial resistance considerably. Figure 12 shows a computer calculated thermal resistance for the modules in Fig. 11b as a function of the interfacial resistance.

Modules in Fig. 11, b through d, contain one, two, and three chips respectively, each dissipating 400 mW. In all three cases the mode of heat sink placement at the pins, for example, is the same. For the three-chip case where the chips are located at $\frac{1}{8}$ inch (0.32 cm) centers, it is seen that the outer two chips begin to influence the temperature of the center chip. It can be shown that at lower power dissipations, all three chips assume the same temperature within the accuracy indicated in the figure. Hence, within the power range considered in the example, the sphere of non-interaction of chips is $\approx \frac{1}{8}$ inch (0.32 cm).

The module in Fig. 11e differs from the others in that heat is removed from its can surface.[13] In other words, the boundary condition is different. However, we find that the thermal impedance from the chips to the bottom of the substrate directly below the chips is the same as those

found in the previous cases. From what thermal impedance is usually understood to mean, one might conclude that the chips in all these cases (with the same interfacial resistance and power dissipation) should assume the same temperature. This clearly is not the case, as can be seen in Fig. 11.

The reasons for the ambiguity in using the indicated thermal impedance for predicting true chip temperature may be explained by considering the various resistive heat paths from the chip to the ambience. Ultimately, heat leaves the module either by way of the pins or by way of the can. However, whether the preferred heat path is out through the pins or the can, the substrate is always included in the path.[14] Since the reference point for calculating the thermal impedance is on the substrate, the temperature difference between that point and the chip will always remain constant, although the temperature levels at the points may change depending upon the direction and type of the preferred heat path.

Thus, the substrate thermal impedances calculated with respect to the bottom of the substrate become useful in predicting the chip temperature when the substrate temperature is known by measurement or otherwise.

The results discussed above pertain to a flip-chip with twelve pads joined to a substrate. It is of interest to be able to estimate the substrate impedance for any size chip with any number of pads. In general, the thermal impedance of the substrate is a function of the number of pads, the size of the pads, and the area enclosed by the pads. Unfortunately, the classical equations of heat conduction do not lend themselves to predicting the desired impedances for finite heat sources placed on finite composite bodies subject to "mixed" boundary conditions.

Figure 13 shows a plot of substrate impedance Z as a function of a new parameter $\zeta \equiv P/NA$ which contains information on the number of pads N, the size of the contact area of the pads A, and the perimeter, P, of the area enclosed by the pads. It is found from computer results that for $0 < \zeta < 1$, the functional form of $Z = Z(\zeta)$ is linear. This is based on the assumption that the pads are placed along the perimeter and that they are equally spaced. Thus, for the chip shown in Fig. 5, for which $A = \pi(2)^2 = 12.6$ mil^2, $N = 12$, $P = 100$ mils, and $\zeta = 0.66$, the thermal impedance from the pads to the bottom of the substrate shown in Fig. 6 is 18°C/watt. On the other hand, for 6 mil diameter pads, $\zeta = 0.294$ and the impedance is found to be $\approx$ 10°C/watt. We also deduce from the parameter ζ that for the same Z, the numbers of pads of different diameters are related by $N_2/N_1 = A_1/A_2$, which reduces to $N_2/N_1 = (r_1/r_2)^2$ for round pads, provided that $0 < \zeta < 1$.

The parameter ζ was arrived at by a heuristic generalization of the relation $R = L/kA$ where k is the thermal conductivity as previously discussed. Thus, the reciprocal

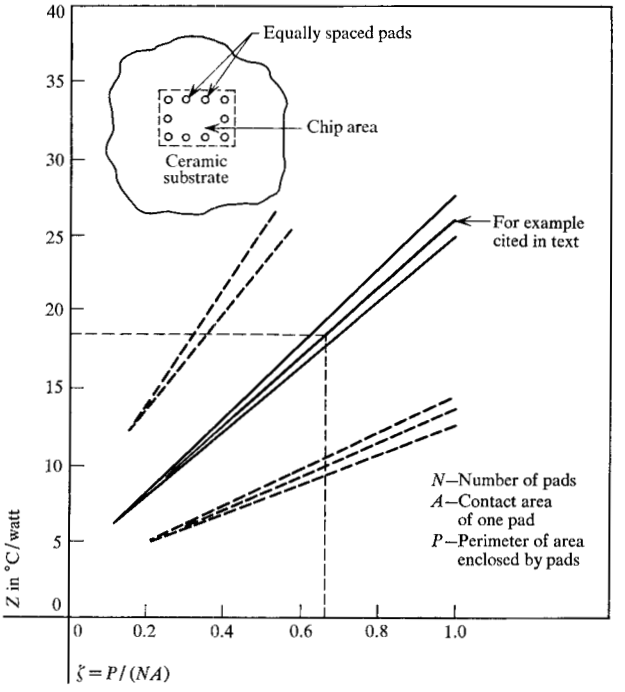


Figure 13 Thermal impedance for pads joined to a substrate.

of the slope of the curves in Fig. 13 may be interpreted as the "equivalent thermal conductivity" of the substrate subjected to a particular boundary condition. For substrates of different geometrical and physical properties, and of different boundary conditions, different slopes are obtained (see dotted lines in Fig. 13). The slopes are influenced by the size of the chip as well. This is because the area enclosed by the pads, which is determined by the size of the chip, also forms a part of the boundary of the substrate.

In order to construct a curve for a particular module and a chip of certain size, only two values of the thermal impedance are needed. First, thermal impedance is found (by measurement or calculation) for a single heat source of the size of the chip. Then another impedance is found for a heat source consisting of a number of pads on the same chip. When these points are plotted against ζ, other thermal impedances as a function of different pad numbers and dimensions may be estimated from a straight line drawn between these two points. In practice, this is tantamount to making two measurements on a single flip-chip. One is made by back-bonding the chip, and the other by bonding the pads to the substrate. The error involved reaches $\pm 20\%$ at $\zeta = 1$, and becomes smaller as $\zeta \rightarrow 0$ (see dashed envelope in Fig. 13).

The findings above suggest that the parameter $\zeta = P/NA$ can be generalized to characterize thermal impedances for finite heat sources placed on general bodies, which, heretofore, has not been possible.

Conclusions

Temperature profiles in flip-chips are largely governed by parameters related to the geometrical and physical properties of the interconnections formed by the controlled chip collapse bonding technique. The dimensions of the metallurgical system comprising the heat path from the heat generating device in chips to the pads are important. The dimensions and number of pads, the size of the chip, and the quality of the bond between the pads and the substrate determine the nature of thermal impedance to heat flow into the substrate.

In particular, it is found that a monolithic chip is not necessarily an isotherm. The nature of the chip temperature profile is governed by the thermal impedance of the heat path from the device layer in the chip to the chip boundaries. Evidently, the heat traverses the path from a heat dissipating junction to the bulk silicon and then back down through the various layers out to the pads.

Before the heat reaches the pads, it encounters a "thermal pinch" impedance that is determined by the size of the glass hole. At the pads, it is found that most of the heat leaves the pads centrally, rather than spreading out immediately. The degree of thermal spreading is influenced, among other things, by the number of voids and their locations in the pads.

In general, analysis of thermal characteristics of flip-chips on modules is complicated by the fact that the thermal paths traverse composite bodies. Furthermore, even the idealized mathematical models of the heat-flow structure are such that classical analytical approaches are not applicable. A case in point is the model of a number of finite heat sources placed on the boundaries of a finite body. Although computer programming is useful in solving individual problems, generalized results such as parametric solutions are needed for efficient design procedures.

Appendix: Description of the numerical computer program

Numerical methods are extremely useful in solving problems of steady-state and transient diffusion of heat in bodies of complex shapes and unusual boundary conditions. Mathematically, the numerical finite-difference method consists of replacing the pertinent differential equation by a finite-difference equation. Physically, it is equivalent to replacing the continuous medium by a network of finite elements or nodes. In most cases, it is simpler to derive the difference equations by direct application of pertinent conservation laws to each node. In the numerical computer program* used for this analysis, the basic equations

$$Q_c = -kA\Delta T/L \qquad \text{(conduction)};$$

$$Q_r = \sigma F(T_0^4 - T_\infty^4) \qquad \text{(radiation)};$$

$$Q_v = hA\Delta T \qquad \text{(convection)};$$

and

$$Q_s = c_p(\Delta T/\Delta t) \qquad \text{(stored energy)*}$$

are applied to each node, and an energy balance is made.

Application of the pertinent equation to each node results in a set of algebraic equations that must be solved simultaneously for the unknown temperature. In the program cited above, two numerical procedures are used for this purpose. In the steady-state case a relaxation technique is used, and in the transient case an iterative technique. In both methods the temperatures in the interior are guessed, based on the known boundary conditions, and then these temperatures are adjusted so that the appropriate conditions are satisfied.

Details of the programming of the numerical methods mentioned above may be found in the excellent description given with the *NATA* program. We point out that steady-state problems are characterized by boundary conditions being specified over the entire domain of the solutions, which are called "jury" solutions† because they must satisfy a jury composed of the boundary conditions. In transient problems, solutions are not bounded in time since initial conditions are prescribed, but conditions at a later time are determined by the progress of the solutions themselves; they are sometimes called "marching" solutions.

The *NATA* program is written for an IBM 7090 system. As mentioned earlier, the program is applicable to the three modes of heat transfer (conduction, radiation, and convection) in a three-dimensional rectangular coordinate system for both steady-state and transient problems under various boundary conditions.

The "building blocks" of a given problem are the nodes. For example, Fig. 5 shows one-quarter of a flip-chip divided into 996 nodes. The maximum number of allowable nodes in the program is 1,000. Each of the six faces of each node can be subjected to different boundary conditions independently. Also, any node may be assigned heat generation and surface contact resistance.

The computer program calls for the specification of the relative location of each node with respect to its neighboring nodes. This is done by simply punching on cards, in a certain format, the numbers of the six nodes adjacent to the six faces of the given node. Thus, for a matrix containing 1,000 nodes, it is necessary to enter 6,000 numbers. Furthermore, the dimensions as well as the thermal conductivity of each node must be entered. If a transient

* *NATA, T4IBM0031*, IBM SHARE Library, written for the IBM 7090 System.

* In the equations stated, σ is the Stephan-Boltzmann constant; F the net radiation factor; h the convective heat transfer coefficient; c_p the heat capacity; and t the time.

† L. F. Richardson, *Trans Royal. Soc. (London)* **A226**, 229 (1927).

case is to be run, the density and specific heat of each node must also be entered.

Once the geometrical as well as the physical properties of the nodes are specified, the nature of a particular problem is established by defining the heat generation and the boundary conditions.

Again following a certain format, any node can be assigned heat generation. Thus, in Fig. 5, the nodes in the device diffusion layer are assigned desired power dissipations for simulating temperature profiles on flip-chips. As for the boundary conditions, the six faces of any node may be separately assigned any desired boundary condition by entering the appropriate heat transfer coefficient corresponding to natural convection, forced convection, and so on. Furthermore, radiation may also be included between nodes if the proper radiation factors are entered into the data.

It will be observed that the number of nodes needed for the description of a particular problem depends on, among other things, the accuracy needed.

A sample computer print-out of a steady-state temperature distribution is shown in Fig. 7. Note that the temperatures are indicated next to the node numbers. Fig. 8 is a steady-state temperature profile obtained from computer results.

References and footnotes

1. L. F. Miller, *IBM J. Res. Develop.* **13,** 239 (1969, this issue).
2. E. M. Davis et al, *IBM J. Res. Develop.* **8,** 102 (1964).
3. J. A. Perri, "Glass Encapsulation," *SCP and Solid-State Technology* **8** (5), 19 (1965).
4. See, for example, H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids,* Oxford University Press, London 1947.
5. An earlier version of this program is available at the IBM SHARE Library: K. W. Lallier, *NATA T4IBM-0031.*
6. Note that the "intrinsic" thermal resistance from the center of a flip-chip to its edges is about 26°C/watt, as calculated from the formula $R = L/kA$. However, this value does not have significant meaning since the thermal impedance, defined by $\Delta T/q$, may change depending upon the boundary conditions and the relative location of sources and sinks.
7. S. S. Kutateladze, *Fundamentals of Heat Transfer,* Academic Press, New York 1963.
8. L. S. Goldmann, *IBM J. Res. Develop.* **13,** 251 (1968, this issue).
9. Of course, some devices may be "on" all the time while others may switch from "on" to "off" depending upon operational requirements. Note also that heat may be generated under both conditions, but at different levels.
10. R. Byron Bird et al, *Transport Phenomena,* John Wiley and Sons, Inc., New York 1960.
11. The boundaries on the module where heat sinks are allowed are indicated by cross-hatched markings on the Figures.
12. Literature is meager on interfacial conductance values for joints. For some values, however, see H. Fenech and W. M. Rohsenow, "Prediction of Thermal Conductance of Metallic Surfaces in Contact," *J. Heat Transfer* **85,** 15 (1963).
13. Other boundary conditions, such as the attachment of sinks to both the can and the pin sides at the same time, can be simulated with the aid of the computer program.
14. Whether the preferred path is through the pins, through the can, or both, depends upon the relative magnitudes of the heat transfer coefficients on the respective sides of the can.

Received December 13, 1968