

História do Transistor

Versão 1.0 (Nov. 2000)
Mauricio Massazumi Oka

O Computador

Durante a 2ª Grande Guerra foi desenvolvido o conceito de **Sistemas**, que permite a um grande número de pessoas trabalharem juntas para resolverem problemas complexos. Com isso surgiu a necessidade de métodos mais eficientes para armazenar, recuperar e controlar dados.

O problema de automatizar processamentos de dados e cálculos matemáticos recebeu grande atenção desde 1800 até meados dos anos 1900. Computadores mecânicos e eletromecânicos de complexidade cada vez maior eram construídos. Nos anos 30, dispositivos com cartões perfurados eram largamente usados para processamento digital de informações financeiras (business). Analisadores diferenciais haviam sido desenvolvidos para computação científica analógica. Princípios de controle de sistemas por realimentação estavam em uso em máquinas auto-reguladas.

Em 1943, iniciam-se os trabalhos com o ENIAC, o primeiro computador de grande escala, construído com válvulas: 18.000 válvulas, consumindo 140 kW. Foi o primeiro de uma série de máquinas similares construídas depois da 2ª Grande Guerra.

Outra revolução ocorreu nos anos 80 com o advento dos PCs.

Em 1964 o custo de um transistor integrado tornou-se menor que o de um transistor discreto.

No começo dos anos 70 a memória em semicondutor tornou-se mais barata que as memórias com núcleo de ferrite.

O Transistor

Alguns marcos da aplicação da "eletrônica" pré descoberta do elétron:

- 1889: é fabricado o primeiro comutador telefônico e, no mesmo ano, é criada uma conexão telefônica entre Nova York e Chicago, IL.
- 1896: é criado o primeiro discador (dialer) telefônico;
- 1893: primeiro gravador magnético de som;
- 1874: F. Brown observou o efeito retificador de um contato entre metal e galena (sulfeto de chumbo (??) -- "lead sulfide" -- PbS);
- 1896: Marconi patenteia o sistema de rádio.

A descoberta do elétron em 1897, por J. J. Thomson marca o nascimento da eletrônica tal qual a conhecemos.

À época da invenção do transistor existia:

- uma série de válvulas -- diodos, pentodos, CRTs, klystron e "travelling-wave tubes" -- guias de onda(?);
- equipamentos de rádio -- AM, FM e microondas;
- conexão telefônica de alcance mundial com emprego da transmissão de rádio;
- TV P&B era comercial e a TV colorida estava a caminho;
- máquina de FAX (?), calculadora, gravadores de áudio e de vídeo;
- tecnologia digital estava começando a aparecer.

Logo após a 2ª Guerra, M. Kelley da Bell Telephone Laboratories (BTL) vislumbrava limitações da tecnologia usando válvulas e relés eletromecânicos para o futuro da telefonia:

- baixa velocidade dos relés eletromecânicos;
- vida curta e alto consumo de potência das válvulas.

No verão de 1945 Kelly monta um grupo de pesquisa com o intuito inicial de entender os semicondutores. O grupo era liderado por W. Shockley e S. Morgan e era contava ainda com J. Bardeen, W. Brattain, G. Pearson, B. Moore e R. Gibney.

Conhecimento sobre semicondutores na época:

- conhecimento empírico de diodos (de "cat's whiskers" usados em receptores de rádio a diodos de microondas usados durante a 2ª Grande Guerra em detetores de rádio e radares);
- retificadores de potência com diodos de óxido de cobre;
- semicondutores usados: selênio, sulfeto de chumbo (galena), óxido de cobre, germânio e silício;
- conceito de bandas de energia;
- conceito de dois tipos de condução (já chamados de tipo n e tipo p), atribuídos à presença de pequenas concentrações de impurezas. Theurer havia detectado a presença de fósforo (pelo cheiro de fosfina) em Ge tipo n.

Os semicondutores eram impuros (o Si mais puro era de 99,8 %) e nenhum era monocristalino (eram no máximo policristalinos, mas em geral eram usados na forma de pó).

Existiam várias propostas para amplificadores em estado sólido. Uma delas, baseada no mecanismo do efeito de campo, foi proposta por Lilienfeld em 1925.

Decisões importantes feitas em 1945:

- focar no Ge e Si como materiais a serem estudados (por serem mais simples);
- pesquisar o efeito de campo.

1947: Bardeen propõe que as tentativas anteriores de fabricar FET não funcionaram porque existem cargas na superfície que blindam o corpo do semicondutor. Isto veio a ser confirmado experimentalmente. Forma-se uma camada de inversão tipo p sobre a superfície do semicondutor tipo n. Não sei qual é o mecanismo exato para este comportamento, nem sei se é algo peculiar do Ge.

Alguns fatos sobre o contato metal-semicondutor que corroboravam a hipótese da existência de cargas na superfície do semicondutor:

1. a propriedade retificadora do contato metal semicondutor não variava em função da função de trabalho do metal;
2. não havia diferença do potencial de contato ("contact potential" ??) entre amostras de Ge tipo n e tipo p, nem entre amostras de Si tipo n e tipo p;
3. características IxV observadas por S. Benzer (Purdue University) sugeriam que para temperaturas suficientemente baixas os elétrons deveriam ser "congelados" nos estados de superfície e o efeito de campo poderia ser observado experimentalmente. Experiências realizadas por Pearson e Bardeen confirmaram esta teoria.

Tentam modificar os estados da superfície do semicondutor aplicando inicialmente dielétrico líquido e depois eletrólitos na superfície, ao redor do contato de ponto (que é, naturalmente, isolado do eletrólito). Observam que aplicando um potencial reverso no contato de ponto surge uma corrente (que correspondia à corrente de lacunas). Depois, seguindo sugestões de Gibney, eles aplicam tensões no eletrólito e observam mudanças consideráveis no estado da superfície, alterando a corrente. Depois foi estudada uma estrutura com ouro evaporado, circundando o contato de ponto, até que se chegou à estrutura com dois contatos de ponto muito próximos (duas tiras de metal separadas por poucos mils pressionados contra o Ge) com que se observou o efeito transistor, às vésperas do natal de 1947. Com um dos contatos reversamente polarizado e o outro diretamente polarizado o dispositivo foi usado para criar um oscilador, uma clara prova de que havia ganho de potência.

Todas as suposições empregadas para se chegar ao transistor baseavam-se no modelo de condução de corrente pela superfície. Acreditava-se que o Ge tipo n possuía uma camada superficial invertida tipo p por onde a corrente poderia ser conduzida. A idéia era mudar a condutividade deste caminho de corrente. Mas em janeiro de 1948 Shockley apresenta o modelo correto, da corrente sendo transportada pelo corpo do semicondutor, esclarecendo o papel dos portadores minoritários. Shockley propõe o transistor de junção com estrutura pnp. Shockley propõe ainda uma estrutura pnpn que viria ser conhecido depois como o tiristor, com característica biestável, e sugere seu uso como chave para aplicação em comutadores telefônicos.

Em 1948 J. Shive prova experimentalmente que a corrente no transistor é conduzida pelo corpo do semicondutor. Para isso Shive moldou o Ge na forma de cunha, aterrou a base, formou dois contatos de "phosphor-bronze" em lados opostos da cunha numa região onde a espessura do Ge era de 0,01 cm e observou o efeito transistor.

Batismo do transistor: John R. Pierce, na Bell Labs, algum tempo antes de 28 de maio de 1948 sugere que numa válvula existe o conceito de transcondutância, então o transistor deveria ter uma transresistência, por analogia. Somou-se a terminação OR comum a vários dispositivos (resistor, capacitor, indutor, termistor, etc).

Seguiram alguns avanços em materiais:

- 1948: G. K. Teal e J. B. Little crescem Ge monocristalino pelo método Cz. O tempo de vida do cristal crescido por Teal era de 10 a 100 vezes maior que no Ge policristalino.
- 1951: B. Pfan desenvolve a técnica de refinamento por zona, onde um lingote de Ge colocado num barquinho de grafite e uma região fundida "passeia" sobre o lingote várias vezes, obtendo pureza de até uma parte para 10^{10} .

Em abril de 1950 Schockley, M. Sparks e Teal crescem um cristal contendo uma camada tipo p entre duas tipo n. O cristal foi cortado em bastões npn, onde foram feitos os contatos elétricos para fabricar o transistor (Transistores de junções crescidas). No início a resposta em frequência era inferior ao do transistor de contato por pontos, mas tinha menor ruído e suas características eram mais uniformes.

Em 1951 a Western Electric, uma divisão da AT&T, começa a fabricar transistores de contato de ponto (o que continuou por 10 anos) que apresentavam todos os problemas dos diodos "cat's whisker" e a produção em massa era difícil. Era o Transistor Tipo A, que tinha mais ou menos o tamanho de um clip de papel e era encapsulado numa latinha de metal.

Depois surgiu o "bead transistor" que era o mesmo transistor encapsulado em plástico. Era muito menor que o transistor tipo A e mais resistente às forças externas, agüentando até 20.000 G, sendo melhor para aplicações militares.

Em 1952 começou a ser produzido comercialmente o transistor de junções crescidas que tinha características bem mais previsíveis e desejáveis, mas usava muito Ge. Era possível formar apenas uma camada de base por puxamento, a dopagem necessitava um controle muito minucioso e era complicado fazer os contatos.

No mesmo ano de 1952 J. E. Saby, da GE anuncia o transistor de junção de liga. A versão original foi feito pela aplicação de pontos de índio, um dopante tipo aceitador, nas faces opostas de uma fatia de Ge tipo n. As dificuldades em produzir este transistor eram o controle das dimensões e o controle da temperatura para formar a liga. Um subproduto deste transistor foi o JFET.

Do Ge para o Si

Vantagens do Si:

1. bandgap do Si sendo de 1,1 eV contra 0,67 eV do Ge, apresenta corrente de fuga menor e menor dependência da características elétricas com a temperatura.

Desvantagens:

1. menor mobilidade dos portadores minoritários no Si. Podia ser facilmente compensada reduzindo a dimensão do dispositivo;
2. A temperatura de fusão do Si (1415°C) era maior que a do Ge (937°C) e seria mais difícil obter um Si monocristalino puro.
3. O Si é quimicamente mais reativo que o Ge.

Em 1952 Teal e E. Buehler produz Si monocristalino pelo método Cz mas era altamente contaminado com oxigênio que vinha do "crucible" de sílica.

Theurer usou a técnica desenvolvida por B. Pfan em 1951 (refinamento por zona, onde um lingote de Ge colocado num barquinho de grafite e a região fundida "passeava" sobre o lingote várias vezes, obtendo pureza de até uma parte para 10^{10}) para purificar o Si. Mas não encontrava um material apropriado para colocar o lingote de Si.

Em 1953 foi inventado o método de Floating Zone.

Em 1954, Teal que havia se mudado para a TI fabrica o primeiro transistor de Si, um ano antes de qualquer outro concorrente.

Substituir a válvula?

No início pensava-se em substituir a válvula pelo transistor, mas era difícil, pois a válvula era um produto bastante maduro. R. Wallace da Bell Labs, propõe aproveitar as propriedades que são inerentemente melhores no transistor: menor dimensão e menor consumo de potência. Isto cria novos horizontes para o transistor na Bell Labs.

Aumentar a velocidade dos transistores

Para aumentar a velocidade dos transistores bipolares seria necessário diminuir a espessura da base. Era conhecido que com a difusão seria possível criar camadas dopadas muito finas.

Difusão em estado sólido para melhorar as características de um diodo havia sido proposto por Bethe em 1942 e publicado por Serin e Stephens em 1946 para melhorar a uniformidade do retificador de silício para detectores de radar e microondas.

Uma junção pn muito fina em Ge foi demonstrado em 1950 por Hall e Dunlap usando impurezas oriundas de uma liga sólida.

A tecnologia de difusão a partir de fonte gasosa foi inventada por Scaff e Theurer em 1951, que foi posteriormente desenvolvido por Fuller e Ditzenberger em 1954.

Em 1952, C. S. Fuller publica um artigo sobre difusão de impurezas doadoras e aceitadoras no Ge.

Em 1954, C. A. Lee fabricou o primeiro transistor pnp difundido de Ge. Ele difundiu As até $1,5\mu\text{m}$ de profundidade, para formar a base, e o emissor foi criado formando uma liga de alumínio de $25\mu\text{m} \times 50\mu\text{m}$ e $0,5\mu\text{m}$ de profundidade. Foi criada uma estrutura mesa para definir a área da junção base-coletor.

Em 1956 Tannenbaum e Thomas demonstraram que um transistor npn com características elétricas boas poderiam ser fabricados pelo processo de dupla difusão. No seu processo uma estrutura npn foi criada sobre toda a lâmina. Tiras de Al eram definidas por evaporação e um processo térmico criava o contato ôhmico com a base enterrada através do emissor na superfície. Círculos eram definidos por graxa pegando o Al e parte do Si tipo n da superfície que seria o emissor, funcionando como máscara para a corrosão das camadas do emissor e da base.

A máscara de óxido

Em 1955, C. J. Frosch e L. Derick estavam tentando resolver o problema de "pitting" (formação de pequenos furos no Si) durante a difusão e descobrem que seria necessário favorecer o crescimento de SiO_2 em detrimento de SiO , o que foi feito com a adição de vapor de água no processo. Logo observaram que um óxido de alguns milhares de ângstrons podia inibir a dopagem e que poderiam definir regiões difundidas abrindo janelas neste óxido. Os resultados foram publicados em 1957. Logo depois, J. Andrus e W. L. Bond mostraram que alguns fotoresistes eram capazes de proteger regiões do óxido contra corrosão. Estava inventada a fotolitografia e com isso o processamento em "batch".

Não foi encontrado nenhum material equivalente ao SiO_2 para o Ge. O óxido de germânio, por exemplo, dissolve em água.

O transistor era sensível ao ambiente, particularmente à umidade, e era necessário selar o dispositivos hermeticamente. Isto era feito usando a técnica de "metal-to-glass seal", que era também usado em válvulas.

Em 1959 M. M. Atalla confirma que uma camada de óxido de silício é capaz de reduzir a densidade de estados na superfície e poderia ser usado para proteger a superfície. Mas tinham ainda dificuldade em controlar o processo.

Um subproduto do uso de óxido de silício foi o FET. O primeiro transistor MOS foi fabricado por Atalla e D. Kahng em 1960. O problema era o alto valor da tensão de limiar em decorrência do uso de Al como material de porta, que seria resolvido com o uso de silício poli como material de porta.

Por volta do final de 1957 e início de 1958 J. A. Hoerni, da Fairchild, inventou o processo planar, publicado em 1960. Era usado um único metal (alumínio) para criar contatos tanto com o silício tipo n quanto com o silício tipo p. Era proposto manter o óxido usado como máscara contra difusão como um material para proteger a superfície do semicondutor.

Em 1959 foi demonstrado o primeiro transistor epitaxial por Theurerer, Kleimack, Loar e Christenson.

Um dos contaminantes que não podia ser contido pelo óxido era o sódio. Em 1966 J. V. Dalton descobriu que cobrindo a superfície com nitreto de silício era possível proteger o dispositivo contra o sódio mas tinha a desvantagem de reter hidrogênio. Posteriormente foi descoberto que o oxinitreto era capaz de barrar o nitrogênio e permitir a liberação do hidrogênio.

1958: J. S. Kilby da TI e R. Noyce da Fairchild inventam independentemente o CI. Até então existia um problema para implementar sistemas complexos, pois as interconexões eram sempre causas de defeitos. A chamada "tirania dos números" dizia que quando o número de interconexões aumenta a probabilidade do sistema apresentar algum defeito aumenta excessivamente.

CI de Kilby: mostrou ser possível produzir transistores, diodos, capacitores e resistores sobre um único substrato de silício. O primeiro protótipo tinha cerca de 10 componentes e as interconexões eram feitas por "wire bonding".

CI de Noyce: usava tecnologia planar e substituiu o "wire bond" pelo processo em "batch" de deposição de alumínio.

Outras propostas para montar sistemas complexos existentes na época:

1. micromódulos: pilha de pequenas placas de circuito impresso contendo componentes discretos.
2. eletrônica molecular: procurar fenômenos que executassem funções no lugar dos circuitos (ex.: oscilador a cristal (piezoresistivo).), financiado pela Aeronáutica dos EUA. Os desenvolvimentos não ocorreram com a rapidez suficiente para desbancar o CI.

Transistor MOS: possui um desempenho teórico inferior ao do transistor bipolar. Mas é mais fácil de ser fabricado e por consequência é menor. Um dos poucos casos que o pior desempenho prevalece, na história da engenharia.

1963: Kerwin, Klein e Sarace da Bell Labs propõem o processo MOS com porta de silício.

Em 1963 F. M. Wanlass e C. T. Sah apresentam a idéia do circuito MOS complementar.

Em 1963 a IBM decide desenvolver o FET canal n para criar a memória da próxima geração de mainframes. Culmina com o lançamento comercial do IBM-370/158, em 1973, usando memórias MOS.

1964: Snow et al. identificam o sódio como o íon que se move no interior do óxido de porta causando a flutuação das características elétricas do transistor MOS.

1964: Kerr e Young da IBM East Fishkill Laboratory descobrem que o fósforo do PSG funciona como centro de gettering do sódio estabilizando a característica elétrica.

1964: a Fairchild e RCA fabricam os primeiros transistores MOS. O produto da Fairchild (FI-100) era um MOSFET de modo enriquecimento de canal p para aplicações em chaveamento e circuitos lógicos. Um mês depois a RCA anunciava o MOSFET de modo depleção de canal n para aplicações em pequenos sinais.

1965: Pieter Balk da IBM Yorown Heights Research Laboratory sugere que o hidrogênio eliminar estados de superfície na interface óxido de porta/silício numa estrutura MOS. No mesmo ano Kooi da Philips Research Lab confirma experimentalmente esta idéia.

1965: Balk, Burkhardt e Gregor da IBM, De Lord, Hoffman, e Stringer de Reed College em Oregon e Miura da NEC-Japão descobrem independentemente que a densidade de estados na interface óxido/Si é menor no Si(100) que nos Si(110) e Si(111).

Mais recentemente cloro oriundo de HCl, TCE (trichloroethylene) e TCA (trichloroethane) é usado para limpar sódio dos tubos de quartzo dos fornos a alta temperatura (1000 - 1100 °C). Uma pequena quantidade de cloro de uma dessas fontes também é adicionado durante a oxidação para que o cloro incorporado no óxido pois há indicativos de que o cloro imobiliza o íon de sódio.

1965: Heiman na RCA e um grupo da IBM mostra que uma tensão DC aplicada no substrato pode controlar a tensão de limiar do MOSFET (body effect). Esta propriedade era usada para compensar o efeito da inversão da superfície do substrato tipo p no NMOS. Hoje a implantação iônica permite compensar precisamente a concentração de dopantes e não é necessário aplicar essa tensão.

1965: Schmidt da Fairchild apresenta o primeiro chip de RAM com tecnologia MOS. Era uma SRAM de 64 bits (16 x 4) implementado com uma célula de memória estática pMOS de 6 transistores (ou flip flop).

1965: A Lei de Moore é publicada na "Electronics Magazine". O número de transistores num CI dobra a cada ano. Em 1975 Moore faz uma atualização a esta lei, prevendo que este número dobra a cada 18 meses). O impacto foi no sentido de destinar recursos intelectuais e financeiros para que a lei se mantivesse. Do ponto de vista de projeto, o engenheiro, ao invés de usar uma tecnologia madura e ainda prever uma margem de segurança, precisasse projetar sistemas supondo o aumento da capacidade de processamento do CI no transcorrer do projeto.

Em 1968 Dennard da IBM publica sua invenção, o DRAM de um único transistor.

1968: Sarace, Kerwin, Klein e Edwards da Bell Labs apresentam o nitreto de silício como máscara contra oxidação e máscara contra difusão. Embora não sejam usados com esses propósitos hoje, foram importantes na segunda e terceira geração de chips entre meados das décadas de 70 e 80. O nitreto é usado também como dielétrico na fabricação de capacitores para células de memórias dinâmicas usando MOSFET e em estruturas LOCOS.

1969: a Intel lança RAM de 256 bits implementada usando tecnologia pMOS de porta de silício. A IBM possuía desde um ano antes um protótipo de memória de 512 bits implementado usando tecnologia nMOS com porta de alumínio.

1970: A Intel lança o primeiro produto de produção em massa. DRAM de 1 kbit usando tecnologia pMOS de porta de silício com células de três transistores. Logo seria substituído pela versão de 4 kbit (1972). No DRAM de 4 kbit a Intel usou a célula de Dennard com regiões de fonte e dreno formados por difusão.

Por volta de 1970 foi lançada uma das primeiras aplicações da tecnologia CMOS (lógica inversora) no chip de relógio (como divisor de frequência). Valia-se do baixo consumo de potência. Nas primeiras versões o produto era comprometido pelo visor de LED mas passou a durar mais de um ano com o uso de LCD.

1971: Dov Frohman-Bentchowsky da Intel consegue implementar o EPROM, criando uma estrutura de "floating gate", onde uma camada de óxido enterrado na porta é carregada por tunelamento. Ao contrário da célula do DRAM, a informação é guardada de uma forma não-volátil. O dispositivo foi inicialmente batizado de FAMOS (Floating gate Avalanche-injection MOS) mas hoje é mais conhecido por UV-EPROM (Ultra Violet light Erasable Programmable Read Only Memory). Posteriormente, engenheiros da Intel desenvolveram a idéia inicialmente proposta por Kahng em 1967, desenvolvendo o EEPROM ou FLOTOX (Floating-gate Tunnel Oxide). O tempo de escrita foi melhorado para 1ms em comparação aos 10 ms que era típico, nos chamados flash EPROM.

1973: Dennard da IBM demonstram que é possível diminuir o tamanho do transistor sem comprometer as características $I \times V$. É a lei de escalamento.

1976: aumento da densidade da memória de 4 para 16 kbits. A dimensão crítica passa de 8 μm para 5 μm .

1979: passagem de memórias de 16 k para 64 kbit.

1. Uso de placas planas (ao invés da inversão de superfície);
2. Dielétrico dual: uso de uma camada de nitreto sobre a de óxido;
3. Introdução do plasma etching;
4. Uso de stepper: a dimensão crítica passa de 3 μm para menos de 2 μm .

A Conferência de Difusão ("Diffusion Conference")

Bell Labs decidiu licenciar as empresas que estivessem dispostas a pagar US\$ 25.000 com antecedência a título de direitos de propriedade intelectual. Para transferir a tecnologia foram realizados seminários no laboratório de Murray Hill, NJ, e na facilidade de fabricação de Allentown, PA. O primeiro dos seminários ocorreu em abril de 1952, com a presença de mais de 30 companhias. As primeiras empresas são listadas no Apêndice 1. Existiam dois grupos de empresas: as que produziam válvulas e as que não produziam.

Em 1955 por imposição do governo, a Western Electric precisaram licenciar as suas patentes de semicondutores às empresas licenciadas, incluindo as patentes da Bell Labs. Isto foi feito num congresso que revia toda a tecnologia até então desenvolvida, na "Conferência de Difusão -- Diffusion Conference".

As Empresas de Microeletrônica

Nos anos 50 os fabricantes tinham mais a postura de vender o que produziam que produzir o que poderia ser vendido. Um batch típico de transistor de liga tinha um rendimento ("yield") de 20% para características boas, sendo tais transistores vendidos a US\$6 cada. 20% adicionais com desempenho inferior era vendido a US\$ 1,50. Os 40 % inferiores eram vendidos como dispositivos de "qualidade-entretenimento" a US\$ 0,50 e o resto era inoperante.

A Raytheon Corporation, tradicional fabricante de válvulas, tornou-se a primeira grande fabricante de dispositivos semicondutores. Em março de 1953 a Raytheon produzia 10,000 transistores de germânio por mês, vendido a US\$ 9 cada, para o mercado de auxiliar-auditivo ("hearing-aid"). Contudo nunca conseguiram encontrar novos mercados e sua tecnologia tornou-se rapidamente obsoleta.

Dentre as primeiras empresas de semicondutores que entrou no mercado sem ter prévia experiência na fabricação de válvulas destaca-se a Texas Instruments (TI). A TI iniciou suas atividades como provedora de serviços contratados para empresas exploradoras de petróleo nos anos 30. Este mercado era altamente instável e a direção da empresa procurava outros campos de atuação. A TI participou do primeiro seminário de transistores, de abril de 1952, e em 1953, fundou o laboratório de semicondutores, recrutando Teal do Bell Labs para comandá-lo. Os objetivos da empresa eram: produzir transistores de junção crescida, fabricar transistores de germânio em massa para usar em rádios portáteis e tornar-se

fornecedor de silício grau-semicondutor para consumo próprio e para vendas no mercado. Os objetivos foram alcançados em 1956.

Uma série de empresas tradicionalmente fabricantes de válvulas fracassaram em entrar no mercado de semicondutores. Em todos os casos as empresas entravam no mercado, cresciam rapidamente, entravam em decadência e saíam do mercado. É preciso lembrar que logo após a 2ª Grande Guerra havia uma enorme demanda por válvulas, cujo mercado não deixou de crescer até meados da década de 60. Entre 1954 e 1956, 28 milhões de transistores haviam sido fabricados, rendendo US\$ 55 M. No mesmo intervalo havia sido produzido 1,3 bilhões de válvulas, rendendo mais de US\$ 1 B.

Uma outra empresa que teve destaque nos anos 50 foi a Transitron. Foi fundada em 1952 por dois irmãos, David e Leo Bakalar. David vinha da Bell Labs e Leo que era um bem sucedido empresário da área de plásticos. Transitron optou por fabricar diodos e seu produto, com contatos de outro, tinha características de ruptura superiores que os produtos similares da época encontrando um mercado cativo na área militar, tornando-se lucrativa em pouco tempo. Mas sendo empresa de um único produto, voltada para um mercado restrito e com poucos investimentos em P&D acabou desaparecendo.

Shockley Semiconductors

Em 1955 Shockley saiu da Bell Labs (BTL) para juntar-se à Beckman Instruments Inc. para fundar a Schockley Semiconductor Laboratory, em Palo Alto, CA. Marca o início do forte crescimento do Silicon Valley. Em 1958 a empresa mudou de nome para Schockley Transistor Corporation. O principal objetivo de Schockley era, aparentemente, fabricar chaves eletrônicas com aplicações para telefonia (tiristores). Entre 1958 e 1959 Schockley produziu e vendeu diodos, mas era difícil produzir grandes quantidades com características reprodutíveis. Em abril de 1960 a empresa foi comprada pela Clevite Corporation e Schockley tornou-se diretor da Shockley Laboratories, uma unidade da Clevite Transistor da Clevite Corporation, de Ohio. Num dos projetos de pesquisa, Schockley havia dividido seu grupo de pesquisas em dois, dos quais apenas um tinha conhecimento do projeto. G. E. Moore era do grupo sem conhecimento e aparentemente isto veio a motivá-lo a sair da empresa mais tarde.

Fairchild

Em setembro de 1957 um grupo de 8 pessoas, encabeçados por G. E. Moore saíram da Schockley Semiconductor e fundaram a Fairchild Semiconductor Corporation, posteriormente a Divisão de Semicondutores da Fairchild Camera and Instrument Corporation. A Fairchild Camera and Instrument Corporation era uma empresa sediada em Nova York, esta via interessada entrar no mercado de mísseis e sistemas de satélite, e investiu US\$ 1,3 M por 18 meses no grupo formado por S. Roberts (metalurgista), J. Hoerni, J. Last e R. Noyce (físicos), V. Grinich (engenheiro industrial) E. Kleiner (engenheiro mecânico) e J. Blank e Gordon E. Moore (químico-físicos). Dentre os fundadores da Fairchild, apenas R. Noyce tinha algum conhecimento prévio de semicondutores, mas com Ge. O objetivo inicial era desenvolver, produzir e vender transistor de silício com dupla difusão. As principais contribuições da Fairchild foram:

1. primeiro transistor mesa;
2. primeiro transistor planar;
3. primeiro circuito integrado comercial;
4. muitos resultados de pesquisa que levaram à fabricação do transistor MOS.

E. Baldwin da Hughes Aircraft foi contratado como administrador da empresa. Um prédio de 14.000 pés quadrados foi alugado por "leasing" em 844 Charleston Road, em Palo Alto. Alguns cuidados foram tomados para garantir a limpeza dos processos e diminuir o particulado. Por exemplo, nitrogênio, oxigênio e forming gas passavam de uma fonte central de cilindros através do prédio por tubos de cobre limpos soldados sem o uso de fluxo, para minimizar a contaminação. Foi necessário puxar silício para consumo próprio. A fotolitografia era adaptação do processo usado para fabricar placas de circuito impresso. Os fornos para difusão precisavam de boa uniformidade, mas não havia produtos comerciais com tais características e precisavam ser projetados. Era preciso projetar um evaporador com ambiente de vácuo. Precisava criar técnicas para encapsular os transistores para protegê-los do ambiente.

O primeiro sistema de fotolitografia criado na Fairchild permitiu dimensões críticas de aproximadamente 125 μm .

As primeiras junções criadas na Fairchild tinham características IV muito ruins. Foi inventado o processo de "gettering" usando metalização por níquel nas costas da lâmina.

No transistor npn com estrutura mesa da AT&T era usado Al para fazer contato ôhmico com o Si tipo p e Au-antimônio para fazer contato ôhmico com o Si tipo n, uma vez que o antimônio era um dopante tipo n no Si. Por sugestão de Noyce, foi feito contato ôhmico no Si tipo n altamente dopado do Si. Não se sabia ainda na época o mecanismo físico do tunelamento. No transistor pnp, onde o contato ôhmico com a base de tipo n precisava ser feita numa região menos

dopada, foi necessário criar um processo mais complexo (qual??).

Logo depois da fundação da Fairchild os russos lançam o Sputnik e a reação histórica dos americanos cria da noite para o dia um novo mercado para dispositivos de baixo peso e alta frequência proposta na nova empresa. Menos de um ano depois era lançado no mercado o primeiro lote de transistores de silício com base difundida e ao final de 1958 a empresa já era lucrativa.

O primeiro produto da Fairchild foi o 2N696 (seguido de um transistor equivalente, com ganho maior denominado de 2N697) capaz de chavear corrente de 150 mA, suficiente para fazer funcionar memórias de núcleo magnético, começou a ser vendido para a Divisão de Sistemas Governamentais da IBM, em Owego, NY.

Para aumentar o tempo de resposta dos transistores com estrutura mesa, Hoerni evaporou ouro nas costas da lâmina que funcionava como "life time killer".

Os transistores eram encapsulados em latinhas e selados em ambiente de nitrogênio mas muitos apresentavam problemas depois de encapsulados. B. Robertson, um dos técnicos abriu um dos transistores com defeito e observou um ponto luminoso na parede lateral da estrutura mesa quando o transistor era polarizado para ruptura ("breakdown"). Ele percebeu que o ponto luminoso era uma partícula grudada naquele ponto, atraído pelo campo elétrico intenso. Removendo a partícula o transistor voltou a funcionar corretamente. Com isso surgiu a idéia de eliminar particulados durante a fabricação do transistor. O primeiro teste que foi inventado era dar uma pancadinha com lápis na latinha do encapsulamento. O processo veio a ser automatizado depois. O problema só veio a ter a solução definitiva com a invenção do processo planar, de Hoerni.

O transistor planar exigia quatro máscaras contra três da estrutura mesa. O transistor planar era também mais suscetível a pinholes que a estrutura mesa e necessitou uma limpeza maior dos processos. Um dos problemas da tecnologia planar ocorria em transistores pnp. No coletor levemente dopado o Si tipo p podia ficar invertido na superfície, fazendo com que a base se estendesse até a extremidade do silício. Para evitar este problema, os primeiros transistores da Fairchild tinham a base circular que circundava todo o emissor.

Os dois primeiros transistores planares BJT 2N613 (nnp) e 2N869 (pnp) foram produzidos a partir do final de 1961.

A tecnologia planar foi empregada para criar circuitos integrados. Um dos grandes problemas era isolar os coletores. Para isso foi feita uma grade com difusão de boro de face a face, cuja junção com o coletor era reversamente polarizada, para isolar as regiões de coletor. Numa lâmina com 80 μm de espessura era feita difusão por 24 horas. A primeira família de CIs chamava-se Micrologic e consistia de 5 funções lógicas simples e foi introduzida no mercado em 1961. Usava DCTL (Direct Coupled Transistor Logic), com acoplamento direto transistor-transistor e continha quatro ou mais transistores por chip.

A idéia de crescimento epitaxial havia sido proposta em 1960 por um grupo de pesquisadores da Bell Labs (H. C. Theuerer, J. J. Kleimack, H. H. Loar e H. Christensen). Embora inicialmente proposto para reduzir a capacitância do coletor, um substrato altamente condutor tipo p com uma fina camada de Si tipo n na superfície passou a ser usado para fabricar CIs sem a necessidade de difundir dopante no Si de face a face.

Dois produtos foram responsáveis por grandes lucros da Fairchild, que permitiram aumentar o ritmo de desenvolvimento tecnológico. o transistor de chaveamento npn 2N706 era uma versão mais rápida (16 ns) do transistor npn 2N696, e era fabricado com dopagem de ouro possuía uma área menor ($1,2 \times 10^{-4} \text{ cm}^2$). A Control Data Corporation (CDC) assinou um contrato para o desenvolvimento de um transistor ainda mais rápido (tempo de chaveamento de 3 ns), no valor de US\$ 500.000. O objetivo era criar o computador científico CDC-6600 com capacidade de 10 MHz (3 MIPS). Outro requisito era confiabilidade, uma vez que seriam necessários 600.000 transistores no CPU. Este contato foi seguido de um contrato de produção de US\$ 5M para 10 milhões de unidades dos transistores de alta velocidade e 2,5 milhões de unidades de diodos de alta velocidade, em 1964. Os engenheiros da Fairchild reduziram a área da geometria circular da estrutura mesa do 2N706 para uma geometria de três dedos e usou passivação por óxido para estabilizar o transistor. Eles aumentaram também o rendimento usando uma camada epitaxial para melhor controlar a resistividade. Com isso atingiram a marca do tempo de chaveamento de 3 ns e confiabilidade exigida (2000 horas de operação de CPU antes de falhar um transistor). Com isso tornou-se viável o primeiro computador científico da história. É interessante lembrar que em 1960 era necessário um investimento de US\$ 250.000 para instalar uma fábrica de semicondutores.

O segundo produto de grande impacto na Fairchild foi a versão do transistor 2N696 com tecnologia planar. Sua capacidade de operar em altas temperaturas (comparado aos transistores de Ge) e confiabilidade tornavam-no próprio para os programas dos mísseis "Minuteman". As amostras dos transistores foram vendidos a US\$ 250 a unidade e os transistores produzidos em volume, a US\$ 100 a unidade para um custo de produção de menos de 50 cents por unidade.

Muito trabalho foi também realizado sobre a estrutura MOS depois que foi apresentado o artigo de Kahng e Atalla. O grande problema da estrutura MOS era a flutuação da tensão de limiar ao longo do tempo. Foi observado que quando o metal de porta era depositado por e-beam esta flutuação era menor que no caso da deposição por evaporação sugerindo que o problema era devido a alguma impureza. Posteriormente descobriu-se que eram os íons alcalinos, principalmente o sódio. O estudo da interface silício-óxido era obviamente importante. O estudo ficou inicialmente a cargo de Sah e depois contou com a adesão de A. S. Grove, E. Snow e B. Deal que foram contratados para este propósito. Sah foi para a Universidade de Illinois e a liderança do grupo ficou a cargo de Grove.

Simultaneamente eram estudadas as potencialidades dos transistores MOS. Destaca-se F. Wanlass que juntou-se à Fairchild em 1962, que propôs juntamente com C. T. Sah, em 1963, a estrutura CMOS.

Em 1968 G. Moore funda a Intel Corporation.

O Papel dos Militares

O lançamento do Sputnik pelos russos a 4 de outubro de 1957 demonstrava que estavam anos à frente dos americanos na tecnologia de foguetes e havia um sério risco da tecnologia de transistores passar ao controle do Departamento de Defesa (DOD -- Department Of Defense), que buscava meios para miniaturizar sistemas. Uma semana antes do anúncio oficial do transistor pela Bell Labs houve uma reunião da empresa com a marinha e com o exército, chegando à promessa de que a tecnologia não viria a se tornar secreta, o que poderia dificultar o desenvolvimento deste conhecimento em fase embrionária. O DOD usava os transistores como substitutos das válvulas, para miniaturizar equipamentos de comunicação e era um grande mercado consumidor, tanto que dos 90.000 transistores produzidos em 1952, a grande maioria foi comprada pelos militares. Contudo uma série de problemas dos primeiros transistores como confiabilidade, repetibilidade das características elétricas, operação a altas temperaturas e disponibilidade só realçavam a necessidade de maiores avanços antes de se tornar prático.

Na época, os recursos militares eram muito mais voltados ao aumento da capacidade produtiva que a P&D. A U.S. Army Signal Corp que representava todos os ramos do desenvolvimento de produção para uso militar dos transistores despendeu entre 1954 e 1964, entre 0,5 e 1 milhão de US\$ para P&D. No mesmo período os fundos para o "production engineering measures" (PEM's) cujo objetivo era aumentar a disponibilidade de dispositivos semicondutores totalizavam mais de US\$ 50 M. A condição para receber o fundo para produção era comprometer-se a produzir pelo menos 3.000 unidades por mês de cada componente especificado no fundo PEM.

A primeira empresa a receber os fundos da PEM foi a Western Electric, em 1952. No ano seguinte juntaram-se a General Electric, Raytheon, Sylvania e a Radio Corporation of America. Em 1956 o financiamento começou a contemplar empresas menos tradicionais como a Motorola, TI e Transitron. Ao longo dos anos, as empresas fabricantes de válvulas beneficiaram-se muito mais dos fundos governamentais para P&D que as empresas que dedicavam-se exclusivamente a semicondutores. Em 1959 as velhas companhias abocanhavam 78 % dos fundos governamentais apesar de possuírem apenas 37% do mercado.

Em suma, o papel fundamental dos militares foi criar o mercado de semicondutores. Vale destacar que a TI logo cedo percebeu que para semicondutores o volume de produção e custo eram interdependentes e que o segredo para reduzir custos era aumentar o volume de produção.

As pessoas envolvidas

William Shockley

No começo de 1954, depois de passar um período como professor visitante na Caltech, Shockley foi para Washington assumir a posição de diretor de pesquisa da Weapons Systems Evaluation Group -- um grupo de elite composto por civis e cientistas que aconselhavam o "Joint Chiefs of Staff" em sistemas bélicos avançados. Mas logo se entediou com o trabalho burocrático e resolveu voltar à iniciativa privada. Mas não para o Bell Labs. Resolveu colocar em prática um velho sonho: criar uma empresa de tecnologia, focado em semicondutores. As primeiras consultas com a TI e com a Raytheon não resultaram em nada. Shockley negociou por um mês com a "venture capital firm" de Laurence Rockefeller, mas também sem sucesso. A 29 de julho de 1955 recebeu um recado do colega da época da Caltech, Arnold Beckman, que havia sido professor de química por lá nos anos 30. Beckman era também um empresário bem sucedido, tendo inventado um medidor de pH que produzia e comercializava. Em meados dos anos 50 a Beckman Instruments empregava perto de 2.000 pessoas nos EUA, Canadá e Alemanha Ocidental, especializando-se na fabricação de instrumentos analíticos para controle de processos industriais. Pelas mãos desses dois homens nascia o Shockley Semiconductor Laboratory. Frederick E. Terman, professor e reitor da faculdade de engenharia da Universidade de Stanford, fica sabendo dos planos de Shockley, e considera que seria o parceiro ideal para se estabelecer no Stanford Industrial Park, nas terras da universidade, que já abrigava iniciativas de seus alunos, Hewlett e Packard, e os irmãos Varian, cujas empresas já estavam instaladas desde a época da 2ª Grande Guerra. Comisso pretendia mudar a cara da

região da Baía de São Francisco, que até então destacava-se pelo cultivo de pêssegos. A Shockley Semiconductor Laboratory foi estabelecida no endereço 391 South San Antonio Road, em Mountain View, CA, nos limites de uma antiga base do exército americano que estava sendo transformado num shopping center. Shockley alugou o prédio por leasing pagando US\$ 325 ao mês.

Nesse meio tempo Shockley percorria os EUA recrutando pesquisadores. No começo de 1955 Shockley abordou alguns pesquisadores da Bell Labs, incluindo Ross, que preferiu permanecer em New Jersey. Questionado depois por que não tinha acompanhado Shockley na empreitada Ross respondeu: "nós o conhecíamos bem demais!". Shockley encontrou Gordon Moore, um físico químico da Caltech, então pesquisando a estrutura da chama em John Hopkins Applied Physics Laboratory, nas proximidades de Washington DC. Encontrou também R. Noyce, físico treinado no MIT, trabalhando com transistores de alta frequência na Philco.

Shockley exigiu de todos os potenciais empregados uma bateria de testes para avaliar sua criatividade. Realizou também testes psicológicos para avaliar o QI, capacidade de trabalhar em equipe, entre outras capacidades. Baseava-se no seu artigo intitulado "On the statistics of individual variations". Seu objetivo era empregar os mais talentosos cientistas e engenheiros e pagar os melhores salários do mercado; se eles gastarem apenas metade do tempo trabalhando para a empresa e a outra metade em pesquisa e publicação de artigos a empresa ainda seria altamente lucrativa.

Shockley tinha um temperamento autoritário e havia um clima de descontentamento generalizado entre os funcionários. Com o tempo, os objetivos iniciais de pesquisar diodos e transistores difundidos em silício haviam perdido a importância, não se sabe exatamente como e por quê, e o principal objetivo de Shockley passou a ser as chaves eletrônicas com aplicações para telefonia (tiristores). Os custos de pesquisa eram elevados na Beckman Instruments e estavam escapando ao controle, em especial no Shockley Semiconductor Lab. Numa reunião de Beckman com o pessoal da Shockley Labs para conter os gastos Shockley saiu gritando "Se você não gosta do que estamos fazendo, posso pegar o grupo e arranjar apoio em qualquer outro lugar". Quatro meses depois do evento, em setembro de 1957, oito dos seus principais homens deixaria a empresa para criar a Fairchild.

Em 1958 a empresa mudou de nome para Shockley Transistor Corporation. Entre 1958 e 1959 Shockley produziu e vendeu diodos de silício, mas era difícil produzir grandes quantidades com características reprodutíveis. Em abril de 1960 a empresa foi comprada pela Clevite Corporation e Shockley tornou-se diretor da Shockley Laboratories, uma unidade da Clevite Transistor da Clevite Corporation, de Ohio. Depois de ganhar uma posição de "chair professorship" na Universidade de Stanford, CA, Shockley continuou como consultor da Shockley Laboratories da Clevite Transistor até ser vendida para a International Telephone and Telegraph (ITT), em 1965.

Em 1963 Shockley abandonou sua carreira de empresário, quando Terman ofereceu uma cadeira de professor na Universidade de Stanford (Alexander M. Poniatoff Professor of Engineering and Applied Physics). Em 1965 Shockley voltou a trabalhar na BTL como consultor executivo. Em 1975 Shockley aposentou-se da BTL e de sua posição de catedrático da Universidade de Stanford. Tornou-se professor emérito até falecer em 1989 de câncer de próstata.

Shockley dividiu o Prêmio Nobel com Bardeen e Brattain, em 1956.

Em setembro de 1957 oito dos pesquisadores da Shockley Semiconductor Laboratory fundam a Fairchild Semiconductor, encabeçados por G. E. Moore, para produzir transistores de junção.

John Bardeen

Bardeen tinha 39 anos em 1947, quando inventou o transistor e era o mais quieto dos membros do grupo. Era conhecido como "whispering John" por alguns amigos. Bardeen nasceu em 1908 em uma família com tradição acadêmica, filho do reitor da escola de medicina da Universidade de Wisconsin. Ele entrou na Universidade de Wisconsin aos 15 anos e recebeu os graus de bacharelado e mestre. Depois de passar alguns anos na iniciativa privada ele foi para a Universidade de Princeton para o doutorado em física. Bardeen juntou-se à Bell Labs no outono de 1945 e dividiu uma sala com Brattain, de temperamento totalmente opostos. Mas tornaram-se grandes amigos e até jogaram golf juntos nos fins de semana. Bardeen ficou fascinado com os trabalhos sobre semicondutores em Brattain estava envolvido desde o início dos anos 30. Logo o cerebral Bardeen começou a apresentar propostas que eram implementadas prontamente por Brattain. Brattain havia se tornado um pesquisador experimental de muitos recursos.

Julius Edgar Lilienfeld

Lilienfeld nasceu na Polónia em 1882 e era professor na Universidade de Leipzig de 1910 a 1926. Ele imigrou para os EUA em 1926 e foi diretor da Ergon Research Laboratory em Malden MA. Naturalizou-se americano em 1935 e aposentou-se logo depois, indo para Virgin Island com sua esposa americana, Beatrice Ginsberg, falecendo a 28 de agosto de 1963, aos 81 anos.

Apêndice-1: Lista das primeiras empresas licenciadas pela Bell Labs para produzir dispositivos semicondutores.

1. Arnold Engineering
2. Brush Electronics
3. Globe Union
4. Hanovia Chemical and Manufacturing Company
5. T. R. Mallory Company
6. Microwaves Associates
7. Minneapolis Honeywell
8. Raytheon Manufacturing
9. Radio Receptor Company Inc.
10. Sprague Electric Company
11. Texas Instruments
12. Tung-Sol Electric Company
13. Automatic Electric
14. Automatic Telephone & Electric Company Ltd.
15. The Baldwin Company
16. Bower Inc.
17. British Thomson-Houston Company, Ltd.
18. Bulova Watch Company
19. Crane Company
20. L. M. Ericsson
21. Felton Guillaume Carlswerk
22. General Electric Company, Ltd.
23. Hughes Tool Company
24. IBM Corporation
25. IT&T Corporation
26. Lenkurt Electric
27. National Cash Register Company
28. National Fabricated Fabrics
29. N. V. Philips
30. Pye, Ltd.
31. Radio Development and Research Corporation
32. Siemens & Halske
33. Telefunken Gesellschaft
34. Transistor Products Inc.
35. English Electric Company, Ltd.

Referências Bibliográficas

1. Proceeding of the IEEE, Jan. 1998.
2. Michael Riordan e Lillian Hoddeson, "The Moses of Silicon Valley", Physics Today, Dec. 1997, pp. 42 - 47.
3. C-T Sah; "Evolution of the MOS transistor - from conception to VLSI"; Proceeding of the IEEE, V.76, n. 10; Oct., 1988, pp. 1280-1326.