JEITA

Technical Report of Japan Electronics and Information Technology Industries Association

JEITA EDR－4704A

Application guide of the accelerated life test for semiconductor devices

Established in May, 2000
Revised in March, 2007

Prepared by
Technical Standardization Committee on Semiconductor Devices

Published by
Japan Electronics and Information Technology Industries Association

11, Kanda-Surugadai 3-chome, Chiyoda-ku, Tokyo 101-0062, Japan
Printed in Japan

Translation without guarantee in the event of any doubt arising, the original standard in Japanese is to be evidence.

JEITA standards are established independently to any existing patents on the products, materials or processes they cover.
JEITA assumes absolutely no responsibility toward parties applying these standards or toward patent owners.

Contents

JEITA EDR-4704A

Technical Report of Japan Electronics and Information Technology Industries Association

Application guide of the accelerated life test for semiconductor devices

1. Trend of the reliability test methodology and theory of the accelerated life test

Recent progresses in science and technology have brought electronic devices closer to our daily life. With respect to the lifecycles of the electronic products, some are expected to have longer life in the important applications, whereas others are used for shorter periods of life. It is remarkable how semiconductor-application marketplaces have expanded and diversified these days. For example, there are handheld terminals which are prone to be replaced frequently, such as cellphones. On the other hand, there are electronic devices which are required long-life durability in the harsh conditions, such as electronic components for automotives.

Most of these electronic products comprise a large number of semiconductor devices without exception. Consisting of many materials (silicon, oxide layers, metal trace materials, dielectrics, etc.) as well as more than one billion transistors and even larger number of contacts, the semiconductor devices including packages that accommodate dice have grown into enormously large-scale systems. The evolution of the semiconductor technology includes the improvement of traces and dielectric layers such as very thin gate dielectric, high-k materials, and Cu/low-k materials, as well as high density packaging technologies including SiP and build-up substrates. Significant progress has been made toward finer patterns, higher integration, higher speed, and less power consumption. The reliability requirements for such complex larger-scale systems are the same as those for the conventional devices or even stringent. In comparison with the mechanical components, the semiconductor devices generally indicate higher reliability. Still the appropriate reliability assurance plans are required in accordance with the applications and purposes of the devices.

On the other hand, the bases of the reliability criteria are not always technically tangible, even though its stress duration or number of stress cycles has already been specified in the qualification program. One of the examples is the stress duration of 1000 hours at the high temperature operation life test or the temperature humidity bias test.

The acceleration test is intended to predict and verify the reliability level. For that purpose, the appropriate acceleration is estimated from the target failure mechanism. Then the relevant test time or stress cycles are determined based on the acceleration factor between the test conditions and the use conditions. As a result of this consideration, the test time or test cycles are in some cases reduced significantly, quite unlike common sense.

The failure mechanisms in the accelerated stress test can be classified into either intrinsic mode or extrinsic mode.

The intrinsic failure is the inherent wear-out failure of semiconductor device depending on the device design, while the extrinsic failure is externally induced failure that harms the inherent reliability, such as particle dust.

The intrinsic failures appear to be in relatively narrow distribution over time, and it is possible to predict the lifetime from the experiment by using a small amount of Test Element Group (TEG) samples. The extrinsic failures appear to be in broad distribution over time depending on how much the parent population of the test samples contains the external detrimental factors against reliability.

Therefore, with regard to the intrinsic failures, the end of life can be evaluated in the extremely accelerated stress conditions with a small number of the TEG samples. The extrinsic failures that cause early failures can be evaluated with a large number of samples but in a short time.

Understanding of both extreme ends of failure modes will benefit those who wish to plan reliability tests in efficient manners such as shortening the reliability test duration or minimizing the sample size.

2. Reliability theory for the semiconductor devices

Theories and concepts are described here for better understanding of the statistical approach in the reliability theory for the semiconductor devices.

2.1 Statistical consideration to the semiconductor reliability

2.1.1 Reliability

Glossary of terms used in dependability, **JIS Z 8115**, defines reliability as “the capability that enables an item to perform its required function under stipulated conditions for the stipulated time period.” Its probability is also defined as the reliability and expressed as “the probability that item provides the required functions under the stipulated conditions for the stipulated time period (t_1, t_2).”

As is apparent from the above definition, it is inevitable to define the use conditions, use period, and failure criteria before discussing reliability.

Note: Failure is defined as “the item looses the capability to provide the required functions.”

(1) Use conditions

In use conditions the semiconductor devices suffer from the operating stresses, such as electric current and voltage that are inevitable for semiconductor devices to function, and the environmental stresses, such as temperature, temperature change, humidity, and vibration.

If the semiconductor devices are used in the extremely harsh conditions, it is obvious that the devices fail to function shortly.

(2) Time period

Reliability is distinguished from the narrow definition of quality by its time dependence. Therefore, the reliability can be also expressed as “retaining the initial quality over time.”

“Over time” here literally means the length of time or the proportional amount such as the number of cycles.

(3) Failures

The failures are categorized into “catastrophic failure” that is a permanent loss of the functions, “degradation failure” that is a gradual degradation of the functions, and “intermittent failure” that is a temporary distortion of the functions or loose contact. The degradation failures and the intermittent failures are difficult to distinguish. Their definitions may differ by application and by person. Thus, the failure criteria of the semiconductor devices shall be clearly specified.

2.1.2 Failure rate

Quantitative indexes are necessary to identify reliability of the items. For semiconductor devices, the failure rate is one of the important measures to indicate the reliability level. The failure rate is the time-dependent function defined by the failure density function and the reliability explained below. As a simple method, the failure rate is often substituted by the average failure rate that is defined by the total operating duration and the total failures in a certain period.

The failure rate function and the average failure rate are described as:

(1) Failure rate function

The failure rate $\lambda(t)$ is the time-dependent probability where a device had performed its function until time t but failed in the subsequent unit of time period. $\lambda(t)$ is expressed by the failure density function $f(t)$ and the reliability function $R(t)$.

Hereupon, the failure density function $f(t)$ is expressed as

$$f(t) = \lim_{\substack{n\to\infty \\ \Delta t\to 0}} \frac{r(t+\Delta t) - r(t)}{n \cdot \Delta t} \quad \cdots\cdots (1)$$

where, n is the sample size and $r(t)$ is the number of failures prior to time t.

In other words, $f(t)$ is the continuous time-dependent function obtained by infinitely large number of sample size n with infinitesimal interval Δt at the histogram of failure occurrences as a function of time, as shown in **Fig. 1**.

The reliability function $R(t)$, the other component that defines the failure rate $\lambda(t)$, is the probability of no failure found prior to time t, and is expressed by the failure density function $f(t)$ as

$$R(t) = \int_t^{\infty} f(t)dt \quad \cdots\cdots (2)$$

Finally, the failure rate $\lambda(t)$ is derived from the above two functions and expressed in Eq. (3), as a time-dependent equation. Generally the failure rates of semiconductor devices are so low that the unit of $\lambda(t)$ is expressed in *FIT* (=10^{-9}/h).

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{-dR(t)/dt}{R(t)} = \frac{-d\ln R(t)}{dt} \quad \cdots\cdots (3)$$

The device failure rate in the field $\lambda_f(t)$ is commonly extrapolated from the results of the accelerated life test.

$$\lambda_f(t) = \frac{\lambda_a(t)}{A_\lambda} \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad (4)$$

where, A_λ is the acceleration factor of failures and $\lambda_a(t)$ is failure rate in this stress test.

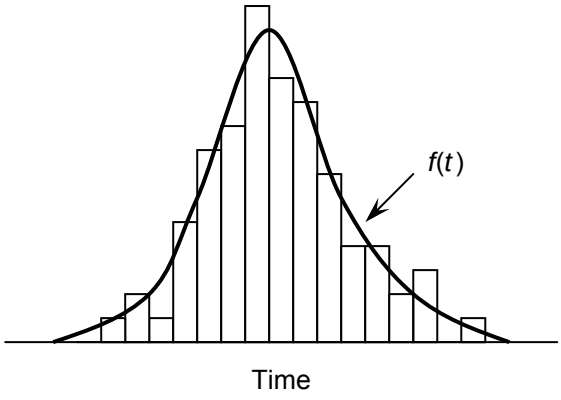

Fig. 1 Typical histogram of failure occurrences as a function of time

(2) Average failure rate

The failure rate originally stands for the aforesaid failure rate function $\lambda(t)$, but Eq. (3) indicates that unknown reliability function $R(t)$ fails to determine $\lambda(t)$. Plotting the wear-out data on some kinds of probability papers including the Weibull probability paper sometimes determine the reliability function $R(t)$ (it is actually the failure distribution function $F(t)$; please refer to **Note**). But it does not always happen. Therefore, the average failure rate in a predetermined period often substitutes for the failure rate (See Eq. (5)). The average failure rate is commonly expressed in *FIT* in the same manner as the failure rate.

$$Average\_failure\_rate = \frac{Total\_number\_of\_failures}{Total\_hours\_of\_operation} = \frac{Total\_number\_of\_failures}{Number\_of\_samples\_\times\_operation\_duration}$$

$\cdots\cdots\cdots\cdots\cdots\cdots$ (5)

The following formula is used to extrapolate the average failure rate in the field from the accelerated test data.

$$Average\_failure\_rate\_in\_the\_field = \frac{Total\_number\_of\_the\_failures}{Number\_of\_samples \times Test\_duration \times Acceleration\_factor}$$

$\cdots\cdots\cdots\cdots\cdots\cdots$ (6)

Note: The failure rate $\lambda(t)$, failure density function $f(t)$, and reliability function $R(t)$ are the key indexes of the reliability level of the semiconductor devices. In addition to these functions, the failure distribution function $F(t)$ is commonly used. It is also referred to as the cumulative failure probability defined as the probability of failures found prior to time *t*.

$F(t)$ and the reliability function $R(t)$ have the following relation:

$$R(t) + F(t) = 1 \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad (7)$$

It is also expressed with the failure density function $f(t)$ as

$$F(t) = \int_0^t f(t)dt \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad (8)$$

Fig. 2 and **Fig. 3** show the relationship of the functions mentioned above.

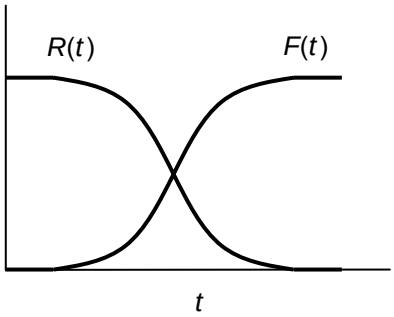

Fig. 2 $R(t)$ vs. $F(t)$

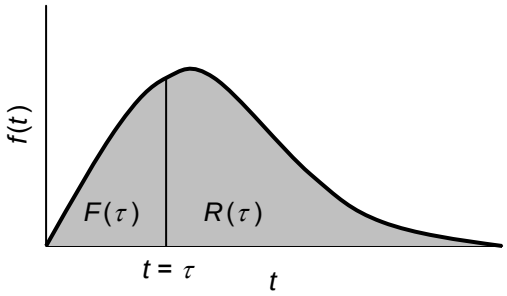

Fig. 3 $f(t)$, $R(t)$, and $F(t)$

2.2 Weibull distribution and bathtub curve

2.2.1 Failure distribution

Generally the failure rates of non-repairable components show the time-dependent change as shown in **Fig. 4**. This failure rate curve can be briefly divided into three periods over time, and from its shape it is referred to as a bathtub curve.

(1) Early failure period

In the early period of the operation the potential defects associated with the process-related troubles and some come to the surface and cause failures (extrinsic failures). It is characterized as a high initial failure rate followed by monotonic decrease as time passes. This trend appears remarkable at the early ramping up stage of manufacturing process.

When the products do not satisfy the predetermined initial failure rate at the early ramping up stage of wafer process, they are screened by burn-in and so on, and failures are removed prior to the shipment. Besides relying on the screening, semiconductor manufacturers make effort to improve the process capability so that they reduce the potential defects that cause the early failures. As a result, the failure rate curves of the commercially available semiconductor devices are more likely to be **Fig. 5**.

(2) Random failure period

In the latter part of the early failure period the downward slope of the bathtub curve turns gently to flat and approaches a certain value; it is because the elimination of the defects results in fewer defects in a population. The period from such a turning point to the onset of the wear-out failures inherent in the material, structure, and process (intrinsic failures) is generally referred to as a random failure period. The failure rate λ in this period is derived from the incidental environmental stresses and the operational overstress to the devices caused by the human errors and other reasons.

(3) Wear-out failure period

After the random failure period, intrinsic failures come to the surface and indicate the end of the inherent lifetime of the device. It is known as a wear-out period and failure rates increases as time passes.

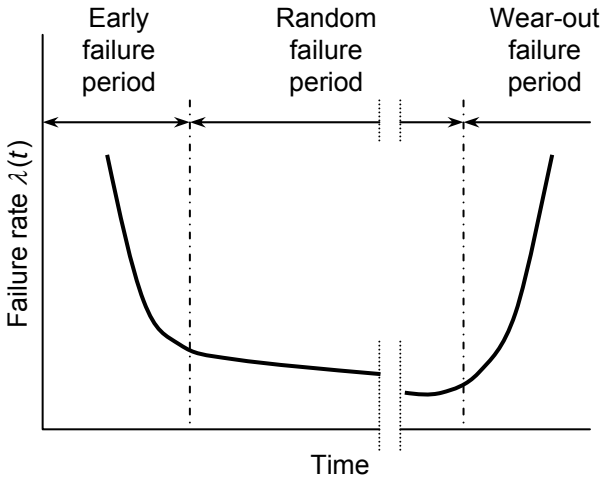

Fig. 4 Bathtub curve (General concept)

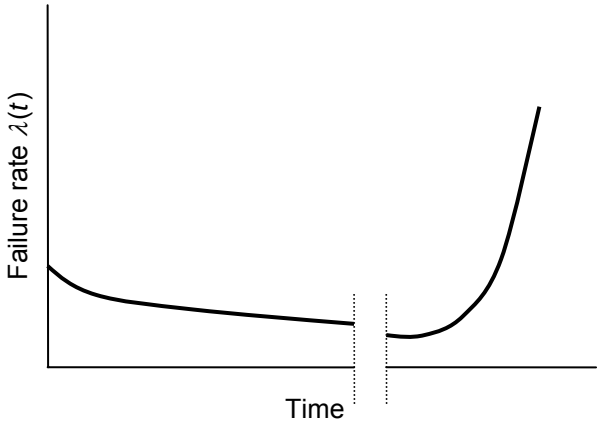

Fig. 5 Actual failure rate curve of the semiconductor devices

2.2.2 Reliability distribution function

Product lifetimes inherently follow some distributions. It is because even an electronic component like semiconductor device is a complex system from the aspects of constituent materials, structures, and processes; therefore, there is a fluctuation among these reliability factors. In addition to that, there are various environmental stresses in the use conditions, such as the self-heat generation, system heat-dissipation design, external EMI noise, and the fluctuation in power and ground voltages during the operations of the semiconductor devices.

For that reason, when the reliability of the products comes up in discussion, it is important to have a clear perception of the lifetime distribution of the parent population, i.e. the failure distribution function. The statistical models of the failure distribution functions for the semiconductor devices are empirically known; exponential distribution, lognormal distribution, and Weibull distribution. Among of them, the Weibull distribution described in the next section is applicable to many failure modes and used most widely.

2.2.3 Weibull distribution and its applications

The Weibull distribution was proposed for the investigation of the metal fatigue failures by W. Weibull of Sweden in 1939. J. H. K. Kao applied this distribution function to the lifetime of the electron tube in 1955, which popularized its use for reliability analysis.

The Weibull distribution is a generalized model of the exponential distribution, and the reliability function $R(t)$, failure density function $f(t)$, and failure rate $\lambda(t)$ are expressed as

$$R(t) = \exp\left\{-\left(\frac{t-\gamma}{\eta}\right)^{m}\right\} \quad \cdots\cdots (9)$$

$$f(t) = \frac{m}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{m-1} \exp\left\{-\left(\frac{t-\gamma}{\eta}\right)^{m}\right\} \quad \cdots\cdots (10)$$

$$\lambda(t) = \frac{m}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{m-1} \quad \cdots\cdots (11)$$

where, m is a shape parameter determining the shape of the distribution curve;

η is a scale parameter (the characteristic life);

γ is a location parameter.

The failure rate $\lambda(t)$ is a decreasing function when m<1, a constant function when m=1, and an increasing function when m>1. In other words, when m is once obtained, it is possible to determine whether $\lambda(t)$ belongs to the early failure period, random failure period, or wear-out failure period.

Provided that it belongs to the random failure period, the reliability function $R(t)$, the failure density function $f(t)$, and the failure rate $\lambda(t)$ are expressed as the following equations after substituting one for m (m=1) in the Eq. (9), (10), and (11). The following equations indicate that these functions follow the exponential distribution (here, γ = 0 for simplicity).

When $\frac{1}{\eta} = \lambda$, each function is described as

$$R(t) = \exp(-\lambda t) \quad \cdots\cdots (12)$$

$$f(t) = \lambda \exp(-\lambda t) \quad \cdots\cdots (13)$$

$$\lambda(t) = \frac{1}{\eta} = \lambda \quad \cdots\cdots (14)$$

When m is roughly two, the failure density function $f(t)$ approximates to the pattern of lognormal distribution. When m is roughly four, $f(t)$ approximates to the pattern of normal distribution.

The patterns of $R(t)$, $f(t)$, and $\lambda(t)$ are shown in **Fig. 6** when m is 0.5, 1, 2, or 4.

As stated above, the Weibull distribution is capable of describing any part of the bathtub curve and even approximating to the lognormal and normal distributions by assigning an appropriate m parameter.

That is the reason why the Weibull distribution can be applied to the most of the reliability data.

Last part of this section touches on the characteristic life η and the location parameter γ:

When substituting η for t in Eq. (12):

$$R(t) = \exp(-1) \quad 0.368 \quad \cdots\cdots\cdots\cdots \quad (15)$$

It means that the characteristic life η is the lifetime whose reliability function is approximately 36.8%.

The location parameter η is an adjustment factor in time and is used when there is the significant difference between the starting time of the observation and that of the failure mechanism for some reason.

Generally the contribution of γ is so small that the Weibull distribution is described with γ = 0 in **section 3** and subsequent sections.

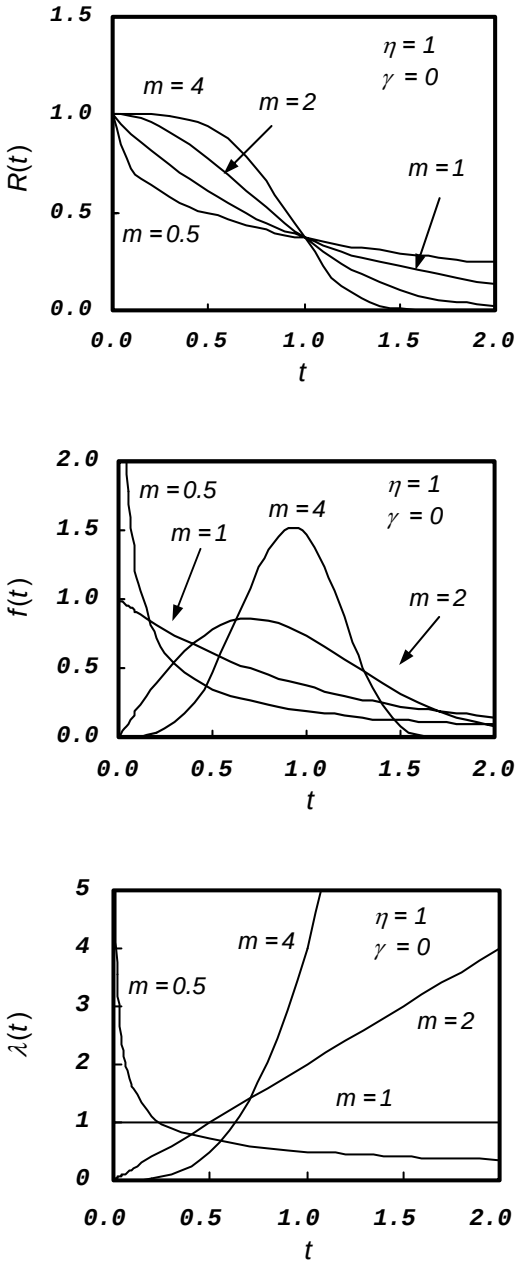

Fig. 6 Weibull distribution

2.2.4 Example of the Weibull distribution derived from the gate-oxide breakdown [1]

This section describes the process how the Weibull distribution is derived for the Time Dependent Dielectric Breakdown (TDDB) of the gate oxide.

One of the examples is shown in **Fig. 7** describing the dielectric breakdown model under the biased gate oxide, referred to as “modified trap to breakdown model”.

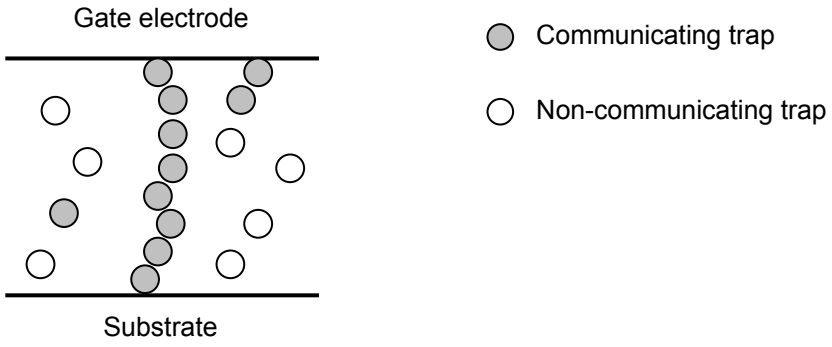

Fig. 7 Breakdown model of the oxide layer

It is assumed that “communicating traps” are formed inside the oxide layer and grow from the gate oxide toward the substrate, and that the formation of continuous path ends up with the dielectric break down. This model can take a weakest-link model from which the Weibull distribution can be derived. The oxide layer is modeled in three dimensional grids made of cubic cells with the length of a_0, as shown in **Fig. 8**.

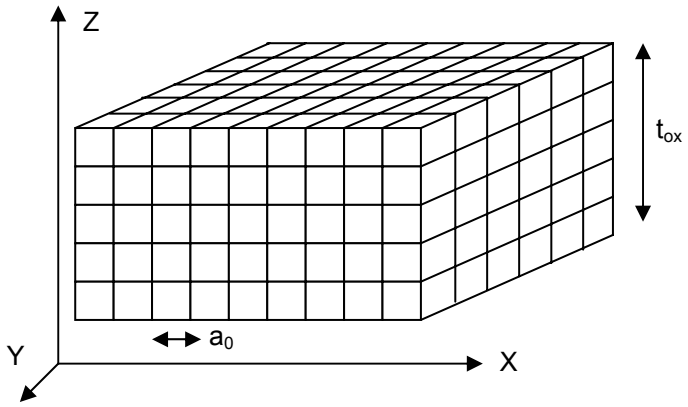

Fig. 8 Three dimensional model of oxide layer made of cubic cells

where, t_{ox} is the thickness of the oxide layer.

When one communicating trap appears in a cubic cell, the occurrence ω of the communicating trap in a cubic cell is expressed as

$$\omega = N_t a_0^{\ 3} \quad \cdots\cdots (16)$$

where, N_t is the defect density of the oxide layer.

Cumulative failure distribution function $F_{cell}(\omega)$ that is associated with the collapse of the cubic cell (generation of a communicating trap) is expressed as

$$F_{cell}(\omega) = \omega \quad \cdots\cdots (17)$$

$F_{coll}(\omega)$ is defined as the cumulative failure distribution function where the oxide layer breaks down due to the collapse of the consecutive n cells in Z axis (the direction toward electrical field) in **Fig. 8**, and is expressed as

$$F_{coll}(\omega) = \left[F_{cell}(\omega)\right]^n = \omega^n \quad \cdots\cdots (18)$$

where, n is obtained by dividing the oxide-layer thickness t_{ox} by cell size a_0.

$$n = \frac{t_{ox}}{a_0} \quad \cdots\cdots (19)$$

The conditions of passing the TDDB test are no defects occurring over the area A_0 of MOS capacitor. The reliability function $R_{BD}(\omega)$ of the MOS capacitor is expressed as

$$R_{BD}(\omega) = 1 - F_{BD}(\omega) = \left[1 - F_{cell}(\omega)\right]^N = \left[1 - \omega^n\right]^N \quad \cdots\cdots (20)$$

where, N is the number of the cells composing the area of the MOS capacitor.

$$N = \frac{A_0}{a_0^{\,2}} \quad \cdots\cdots (21)$$

W_{BD} is defined as

$$\begin{aligned} W_{BD} &\equiv \ln\left[-\ln\left\{1 - F_{BD}(\omega)\right\}\right] \\ &= \ln\left[-N\ln\left(1-\omega^n\right)\right] \end{aligned} \quad \cdots\cdots (22)$$

Then, the relation of $\ln\left(1-\omega^n\right) \approx -\omega^n$ gives

$$W_{BD} = \ln(N) + n\ln(\omega) \quad \cdots\cdots (23)$$

Since ω is a time-dependent function, it is postulated that ω is a function of power of time t:

$$\omega(t) = \omega_0 t^{\alpha} \quad \cdots\cdots (24)$$

where, α is a constant and ω_0 is dependent on the stress condition of TDDB test.

Substituting Eq. (24) for Eq. (23) results in Eq. (25):

$$\begin{aligned} W_{BD} &= \ln(N) + n\ln\left(\omega_0 t^{\alpha}\right) \\ &= \ln(N) + n\ln(\omega_0) + n\alpha\ln(t) \end{aligned} \quad \cdots\cdots (25)$$

This equation is similar to the one of the Weibull distribution shown in the previous section.

The shape parameter *m* in the Weibull distribution is given by

$$m = n\alpha = \alpha\frac{t_{ox}}{a_0} \quad \cdots\cdots (26)$$

It indicates that the shape parameter *m* is proportional to the thickness of the oxide layer; the thinner the oxide layer is, the smaller *m* is.

2.2.5 Identification of the failure-rate behavior in each period of the bathtub curve

The semiconductor devices are qualified based on the data (failure rate) obtained from the reliability test when the devices are developed, and then released to a market. The details of the reliability test are described in the **section 4**, “Practices of the reliability test for the semiconductor devices.”

Before conducting a reliability test, engineers are encouraged to understand which part of the bathtub curve they intend to examine, and then design a test program with rational test durations and sample sizes. The semiconductor devices spend most of their lives in early failure period and then enter the wear-out failure period, followed by the end of life.

The evaluation methods for the failures in the early period are different from the ones in the wear-out period; the differences in the evaluation and the verification methods are described below.

(1) Early failure rate

The semiconductor devices are subjected to the probe test for the electrical inspection after the wafer process. (Some devices are also subjected to the stress test.) Some of the early period failures are rejected as soon as the test begins. After the probe test the devices are conveyed to the packaging process, followed by the final test. Some products are even subjected to the additional burn-in screening. As described in the "**4.2 Concept of Screening**" in details, the burn-in stress accelerates the deterioration of the reliability defects under the high temperature and the high voltage stress. The products are subjected to the stress to remove the early failures until the products achieve the specified reliability level. The early failure rate is generally defined as the failure rate in the period of three months to one year after the shipment. When the failure data obtained from the probe test and the subsequent tests are plotted on the Weibull plot paper as a function of the equivalent duration, these data are expected to align in a straight line. The shape parameter *m* will be typically 0.1 or less when the evaluation includes the probe test, while it will be over 0.1 when the evaluation is burn-in test excluding the probe test data. **Fig. 9** indicates, as an example, how to predict the early failure rate (the cumulative failure rate in one year after the shipment in **Fig. 9**) based on the data from the probe test through the burn-in. "Probe Test-*n*" stands for the *nth* probe test. The field time equivalent to the reliability test duration was projected to the time beyond the product shipment in **Fig. 9**. Typical use period of the semiconductor device is considered to be a part of the early failure period. The results of the high temperature operating life test (HTOL) as a reliability test follow a Weibull distribution, although HTOL itself does not aim at evaluating the early failure rate.

Semiconductor manufacturers predict the early failure rates based on the data from the probe tests, final tests, and occasionally the subsequent burn-in to provide the reliability assurance for their products.

The early failure rate is extrapolated in terms of cumulative failure rate, average failure rate, or instantaneous failure rate that can predict the occurrence of failures in the period of several months to 12 months (for instance) after shipment. Please refer to the section **2.4** for the calculation method in details.

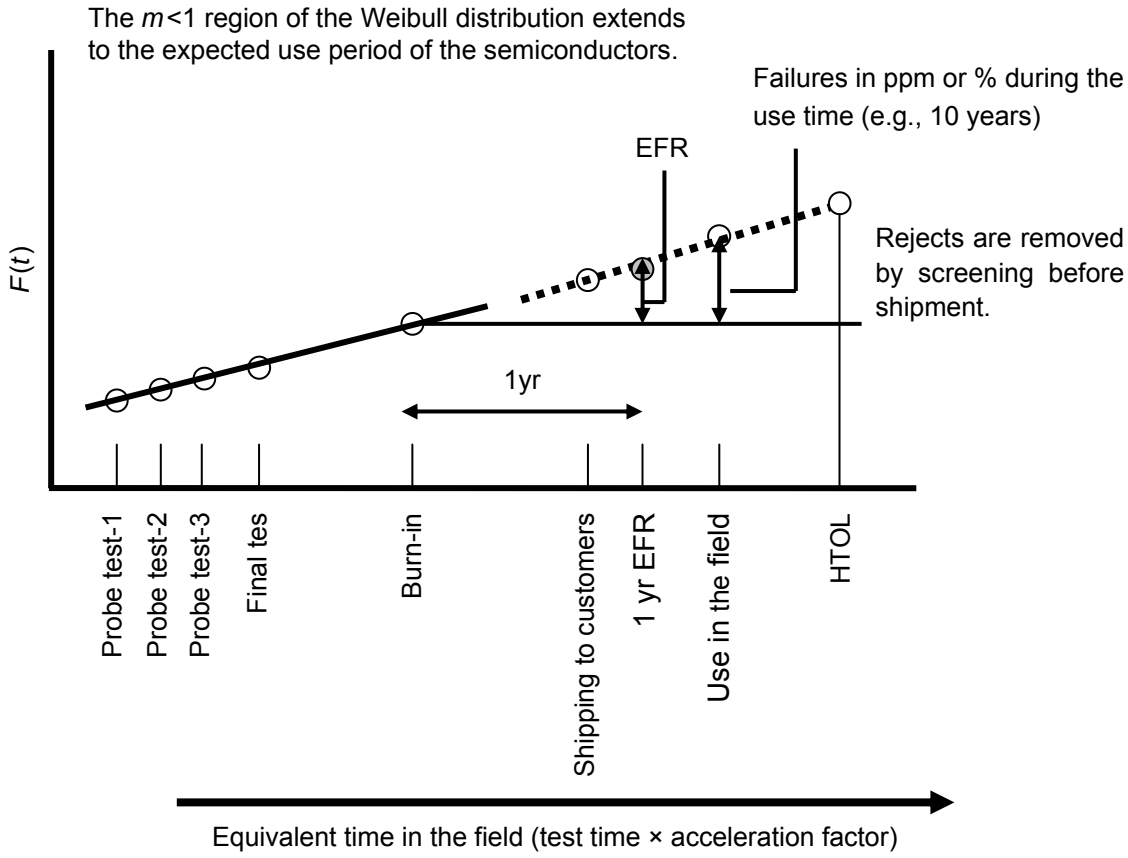

Fig. 9 Prediction of the early failure rate from the probe tests

(2) Wear-out failure rate

The wear-out failure period can be identified by either process reliability test or a part of the product reliability test. Please refer to **section 4**, “Practices of the reliability test of the semiconductor devices” in detail. The process reliability test aims at taking lifetime data in the wear-out failure period to evaluate the maximum capability of the process. Its lifetime is generally defined as the time when the cumulative failure rate reaches 0.1 % (the criteria differs in each case). The lifetimes of the actual products produced in the same process will not exceed the lifetime determined by the process reliability test. On the other hand, the product reliability test aims at verifying that the products are immune from the wear-out failures in the stipulated period, and it is not necessary to take data to find the end of life.

When the cumulative failure rate reaches 0.1%, for example, in the wear-out region, its operating time is defined as useful life. The semiconductor products for general applications are designed to be away from the wear-out failure period in the normal use conditions. The useful life is confirmed to have sufficient time margin before the onset of the wear-out failure period. **Fig. 10** shows the general idea of the bath tub curve from the early failure period through the wear-out failure period.

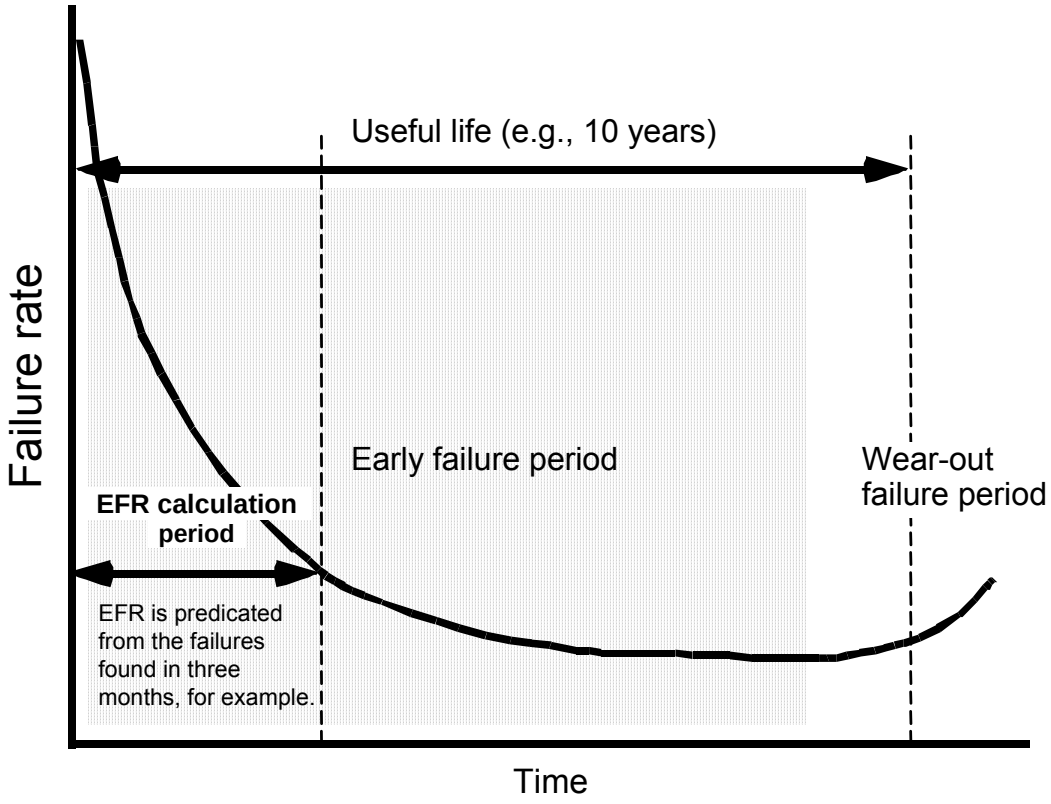

Fig. 10 Bathtub curve and failure rate

2.3 Causes of the failure trend in each period

The failure rates of the electronic components including the semiconductor devices are said to follow the bathtub curve, which is composed of the early failure period, random failure period, and wear-out failure period, as stated in section **2.2**.

It is characterized by the early failure period with initially high but rapidly decreasing failure rate followed by the random failure period with roughly constant failure rate over time, ending up with the wear-out failure period having a rising curve. The causes of being bathtub curve are explained below:

2.3.1 Early failure period

The semiconductor devices are so fine and complex that they are susceptible to the defects generated in the manufacturing process. For that reason, there is a sorting operation at the end of the manufacturing process to weed out defects that do not function properly as requested. The percentage of the passed devices is referred to as a yield. Even after sorting by all possible test items, there still remain some potential defects that were not caught by the programs in passed devices. When the products are in higher yield, they contain less potential defects in general, and vice versa. The use of the semiconductor devices having potential defects may eventually fail in the field due to their defects (Extrinsic mode).

When a production lot contains a small number of devices that are potentially defective but still acceptable, the failure rate will decrease as time passes. It is because the removal of failures having potential defects leaves only the devices that are immune from the defects in the population. In this case the shape parameter *m* of the Weibull distribution is less than one.

In particular, when some potential defects in a production lot failed during the usage of the electronic product, the failed devices were removed by repairs as shown in **Fig. 11**. As a result, only reliable semiconductor devices survive in the lot.

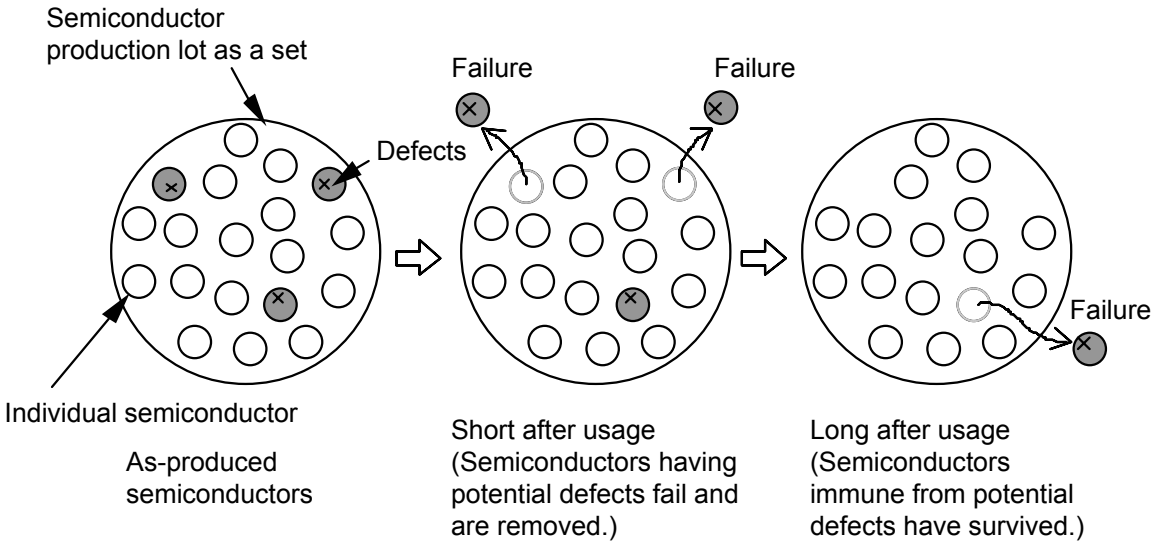

Fig. 11 Failure mechanism in the early failure period

Therefore, the primary countermeasure to such failures is the reduction of defects during the manufacturing process. Another countermeasure is, if it is feasible, the improvement of the product design to be less susceptible to the defects.

On the other hand, there are screening methods such as burn-in, where semiconductor devices are stressed in the harsh conditions so that the potential defects turn into actual failures. Then the failed devices are removed in the sorting process. It means that the burn-in subjects devices to the stress equivalent to the early failure period, and cuts down the early failure rates before shipment. When the defect-reduction activity produces sufficient results, the screening can be gradually relaxed or terminated.

2.3.2 Random failure period

The failure rate shows a downward trend over time when the failures are caused by the potential manufacturing defects, but approximates the exponential distribution with a constant failure rate when the defects are minor. If the devices having defects are removed by the screening measures, it leaves the products whose failure distribution follows the exponential distribution. However, the screening measures may not remove all defects but allow some minor defects remaining in a lot, and they will fail during long use in the field. They are classified as early failures where failure rate decreases over time.

In the strict sense of the word, the failure modes with constant failure rate over time are such as soft errors of memory devices due to alpha particles and electrical overstress (EOS). These overstresses occur at random so that the failure rates stay constant, as shown in **Fig. 12**. The failure rate depends on the tolerance of the device to the overstress. (It depends on the basic specification of the semiconductor.)

When the failure rate is constant, the shape parameter m of the Weibull distribution is equal to one, and the failure rate function follows the exponential distribution.

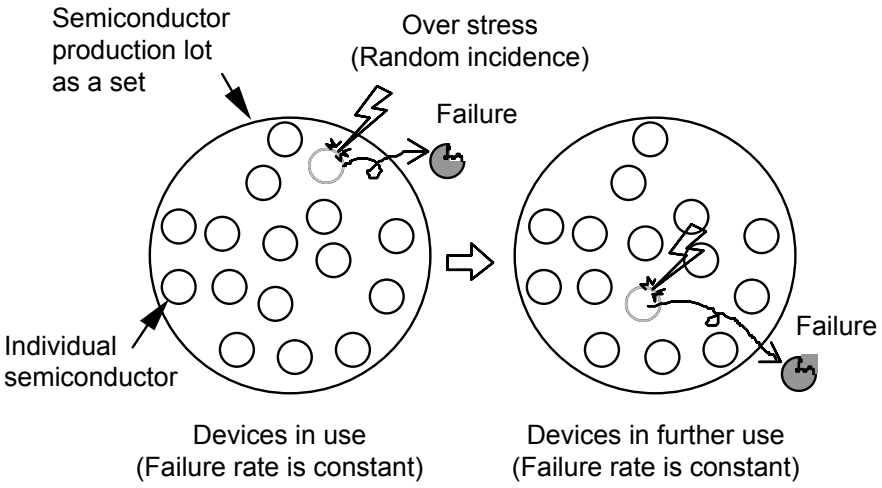

Fig. 12 Failure mechanism in the random failure period

2.3.3 Wear-out failure period

In the long run, any semiconductors eventually wear out to failure, even if they are immune from the potential defects. This is because the semiconductor device reaches inherent durability limit (end of life) and fails (Intrinsic mode). Even though the times to fail vary between devices due to the fluctuations in production and the different stress conditions in use, overall failure rates gradually increase when the operating time comes close to the end of life, as shown in **Fig. 13**.

The useful life is generally far longer (several ten years to several hundred years depending on the basic specification of the semiconductor) than the expected use period of the semiconductors. The endurance for the stresses is evaluated by the various reliability tests. When the failure rate increases over time, the shape parameter m of the Weibull distribution exceeds one. Furthermore, in the case of wear-out failures ($1<m$), it is possible to verify if the failure rate in the use period is sufficiently low from the failure rate function, which is obtained by the test with small sample size (e.g., several pieces to several ten pieces) for the period equivalent to the several times to several hundred times longer than the expected use period.

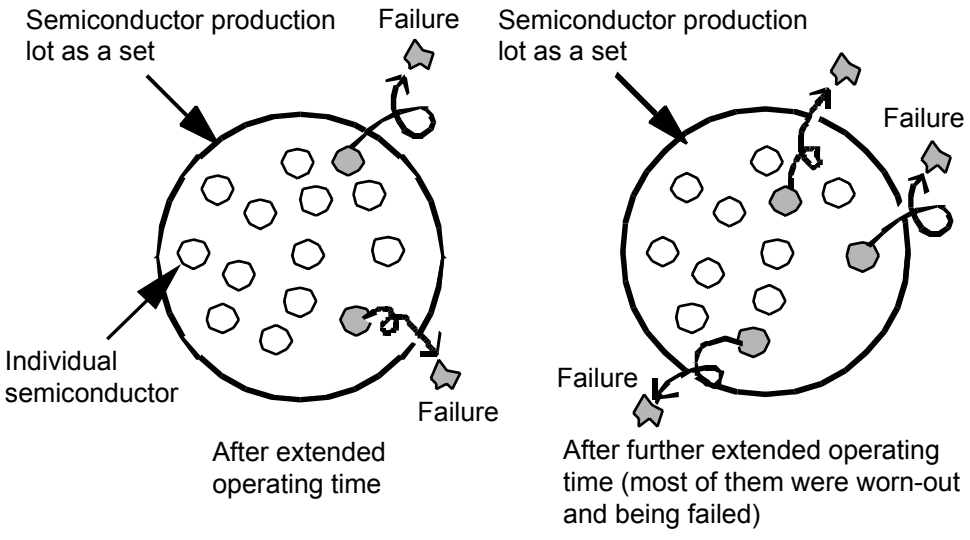

Fig. 13 Failure mechanism in the wear-out failure period

2.4 Prediction of the useful life based on the acceleration test

2.4.1 Reliability test and acceleration factor

The acceleration factor between the test condition and the field condition is necessary to extrapolate the field failure rate from the test data at stress conditions. The operating hours in the field is defined as the product of the test time multiplied by the acceleration factor. It differs between failure modes because of separate physics and separate chemical reaction. It should be noted that the failures found in the same test program may be caused by the different failure mechanisms with the different acceleration factor. The process reliability test mainly aims at predicting the lifetimes for the particular failure mode with the particular acceleration factor. On the other hand, the product reliability test aims at the assurance of the useful time for multiple failure modes. It is evaluated by the reliability test for the necessary duration determined by the representative acceleration factor. Details are described in **section 3**.

2.4.2 Lifetime prediction based on the exponential distribution

Assuming that most of the early failures are removed by some methods such as screening before shipment, the failure rate is expressed in Eq. (27) based on the exponential distribution. The failure rate obtained here is the average failure rate from the beginning to the field time obtained by multiplying the test time and the acceleration factor, as shown in **Fig. 14**.

As far as the wear-out failure does not appear at the test results, devices are considered to reside in the early failure period where failure rate lowers as time passes. With regards to the average failure rate obtained from Eq. (27) and the instantaneous failure rates obtained from the stress test, the latter rate is higher in the first part of the early period, but it becomes lower when the time exceeds the corresponding test time (test time x acceleration factor).

Average failure rate = Total failures/(sample size × test time × acceleration factor) ····· (27)

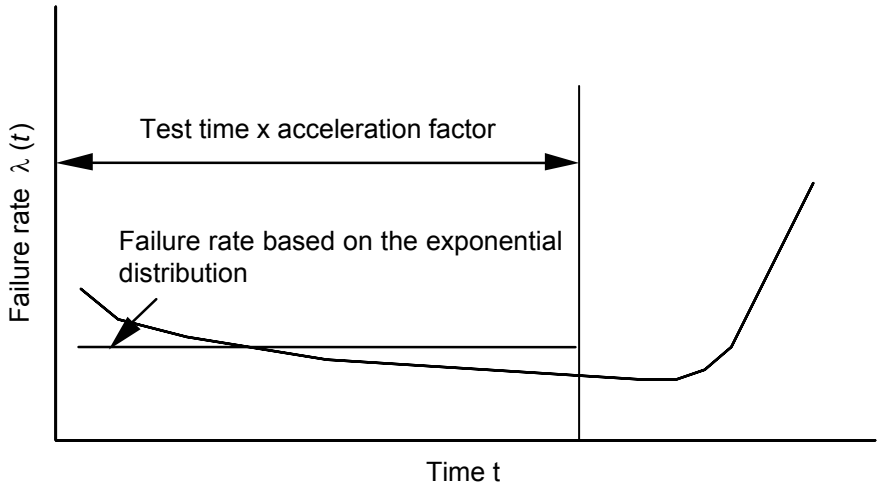

Fig. 14 Failure rate base on the exponential distribution vs. actual failure rate

Sample size for the reliability test is generally several ten pieces, which is sufficient (more than sufficient) to find the wear-out failures but insufficient to figure out the early failures. Therefore, some ingenious plans are necessary to raise the accuracy of the failure rate obtained from Eq. (27), such as the integration of the reliability test data of other products that are manufactured in the same process and specification.

When the life distribution follows the exponential distribution, the failure rate stays constant over time. In other words, the failure rate λ is in the inverse proportion to the Mean Time To Failure (MTTF); λ = 1/MTTF. However, the concept of MTTF does not make sense to the semiconductor devices in general. It is because the semiconductor devices have scarcely been used until they reach MTTF. For example, when the failure rate is 100 FIT, MTTF is $1/(1\times10^{-7})=10^7$ hours as an inverse value of the failure rate, which is an enormous period. Few products have ever been used for such an extended time. Furthermore, since the process reliability test is designed for the durable lifetime of 10 years or more, the lifetimes do not follow the exponential distribution in the range of 10^7 hours. When an electronic system comprises large number of semiconductor devices, the whole system statistically results in lower reliability level, even if each device is highly reliable. On such an occasion, care has to be taken how to deal with the results of the MTTF calculation of the system.

2.4.3 Early failure rate (EFR) prediction based on the Weibull distribution

In the Weibull distribution the failure rates are expressed in Eq. (11). Once m and η are obtained from the experimental data plotted on the Weibull probability paper, it is possible to calculate the failure rate at a certain operating time t. In this case, the value η in the field is the product of the acceleration factor multiplied by η obtained from the stress test.

It is efficient to extrapolate the early failure rate by applying the Weibull distribution to the stress test data including probe test. It provides smaller value of m than that obtained from the data of burn-in only. Since burn-in operation accompanies the expenses for sockets, test boards, etc., simple inexpensive methods are preferred to predict the reliability level.

Unless probe test data are available, burn-in data are utilized. Once m and η are obtained from the Weibull analysis, it becomes possible to predict the reliability level of the products which were screened (shipped).

Shown below is the Weibull analysis applied to the numerical example in which burn-in was performed after wafer-probe test at high temperature. Test conditions and the parameters to determine the acceleration factors used in this example are listed in **Table 1**.

Table 1 Test and burn-in conditions

Test process	Conditions	Notes
Wafer sort	—	Failures in the wafer sort are not regarded as reliability failures.
Wafer-probe test at the accelerated stress conditions	Condition A	Samples are stressed 9 times
Burn-in	125 °C	
Additional burn-in	125 °C	

The following acceleration factor is based on the premises of E_a=0.6 eV and the field conditions of 60 °C and 2.5 V.

Table 2 Test process and acceleration factor

Test process	Acceleration factor	Stress duration
Wafer-probe test at the accelerated stress conditions	40953	80.19 hours (9 times)
Burn-in	23736	2933.29 hours
Additional burn-in	23736	1977.923 hours

Table 3 Field use time and cumulative failure rate

Field use time (h)	cumulative failure rate
0.239	0.031429
0.592	0.036000
1.399	0.038286
2.776	0.039143
5.290	0.041429
11.217	0.043714
22.832	0.046571
45.823	0.048000
80.190	0.049429
3013.480	0.062240
4991.403	0.063308

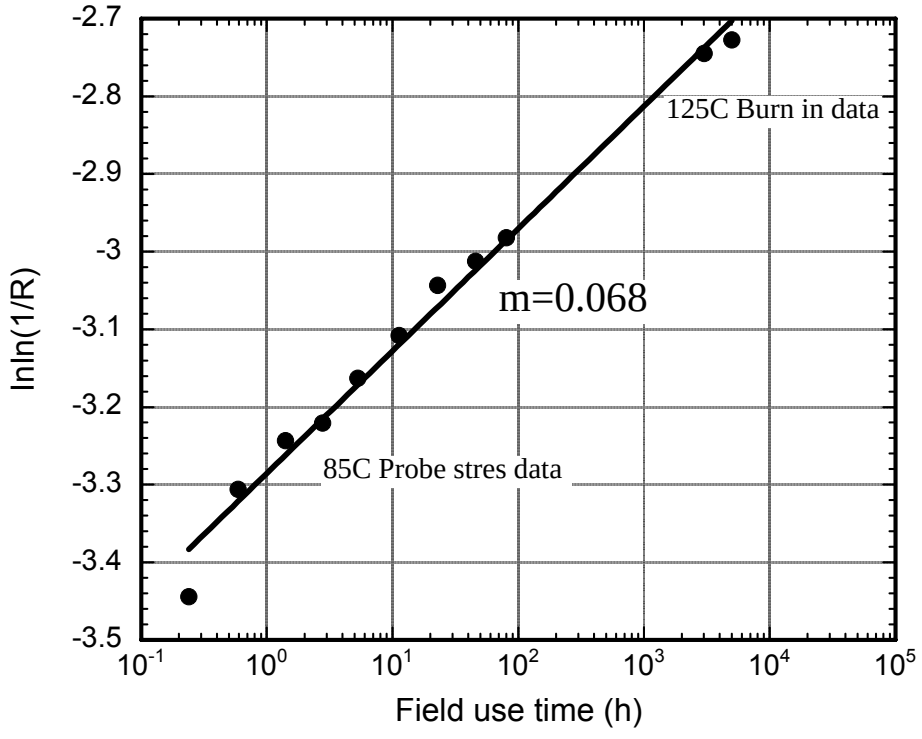

Fig. 15 Weibull plot obtained from the wafer-probe test and burn-in

The repetition of the wafer-probe tests at the stressed conditions enables the reliability prediction as shown in **Fig. 15**. These data imply that the wafer-probe tests alone are also effective for screening the potential defects. This methodology in some cases eliminates the costly burn-in for the reliability prediction. When the wafer-probe test was found insufficient for screening, the products would be subjected to the burn-in for the additional screening. For the purpose of the reliability prediction, the mechanisms of the failures in burn-in shall be the same as the ones in the field.

2.4.4 EFR prediction after screening

When the products were subjected to the screening stress for t_1 period and the failed devices were removed, the reliability level of the population can be predicted by the following procedures.

The probability density function of the products screened for t_1 period, $f(t: t_1)$, is expressed as

$$f(t:t_1) = \frac{f(t+t_1)}{R(t_1)} \quad \cdots\cdots (28)$$

where, $R(t_1)$ is the reliability until time t_1.

When these functions follow the Weibull distribution, the probability density function $f(t)$ is expressed as

$$f(t) = \frac{mt^{m-1}}{\eta^m} e^{-\frac{t^m}{\eta^m}} \quad \cdots\cdots (29)$$

The probability density function of the products screened for t_1 period is obtained by putting Eq. (29) into Eq. (28), as

$$f(t:t_1) = \frac{m(t+t_1)^{m-1}}{\eta^m} \cdot \exp\left(-\frac{(t+t_1)^m - t_1^{\,m}}{\eta^m} \right) \quad \cdots\cdots (30)$$

The cumulative failure probability of the products screened for t_1 period, $F(t: t_1)$, is

$$F(t:t_1) = 1 - \exp\left(-\frac{(t+t_1)^m - t_1^{\,m}}{\eta^m} \right) \quad \cdots\cdots (31)$$

The failure rate $\lambda(t : t_1)$ is expressed as

$$\begin{aligned} \lambda(t:t_1) &= \frac{f(t:t_1)}{R(t:t_1)} \\ &= \frac{m(t+t_1)^{m-1}}{\eta^m} \end{aligned} \quad \cdots\cdots (32)$$

Numerical examples are shown below based on the empirical data:

Table 4 Conditions of test and burn-in

Test process	Condition	Cumulative failure rate
Wafer sort	Condition A: 1.0 sec	0.02
Wafer-probe test at the accelerated stress conditions	Condition B: 2.0 sec	0.025
Burn-in	125 °C, 8 h	0.034
Additional burn-in	125 °C, 4 h	0.0345

The following acceleration factor is based on the premise of E_a=0.6 eV and the field conditions of 55 °C and 5.0 V.

Table 5 Test process and acceleration factor

Test process	Acceleration factor	Stress duration
Wafer sort	4.4	4.4 sec
Wafer-probe test at the accelerated stress conditions	4044	2.25 h
Burn-in	8787	70296 h
Additional burn-in	8787	35148 h

When the data are plotted on the Weibull probability paper, it is known that all data, including the data obtained from the burn-in, align in a straight line, as seen in **Fig. 16**. EFR is calculated from the parameters of the Weibull distribution (m=0.03 and η=4.86E+53).

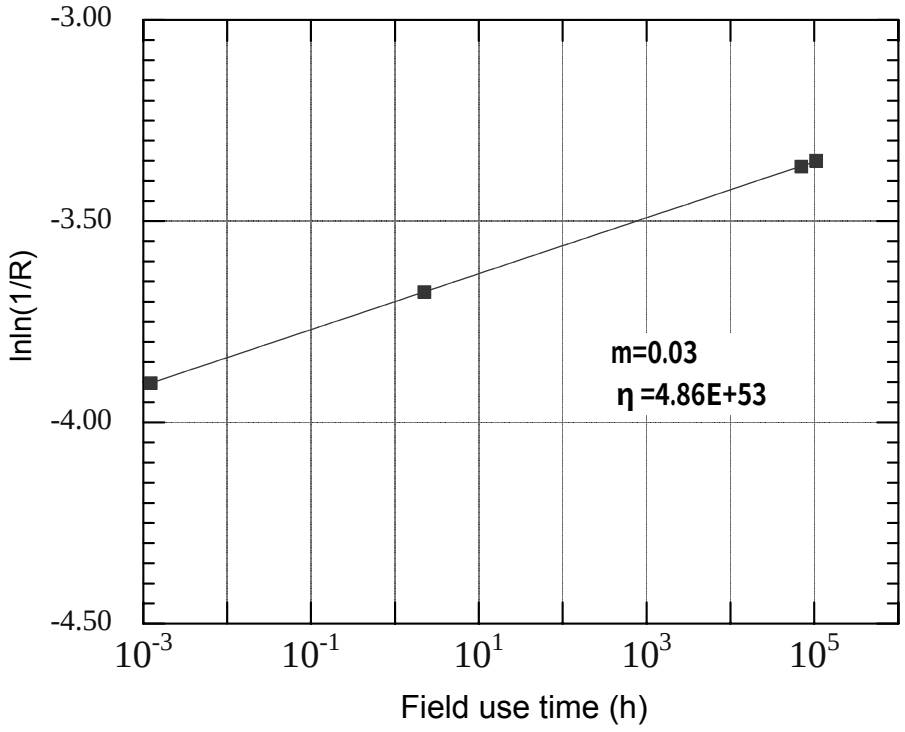

Fig. 16 Example of wafer-probe test at the accelerated stress conditions and burn-in

The failure rate of the products which were sorted out by the probe tests alone is

$$\lambda(t) = \frac{0.03\left(t + \left(2.25 + 1.22 \times 10^{-3}\right)\right)^{0.03-1}}{\left(4.86 \times 10^{53}\right)^{0.03}} = 7.354 \times 10^{-4}\left(t + 2.251\right)^{-0.97} \quad \cdots\cdots (33)$$

The instantaneous failure rate λ is plotted as a function of time, as shown in **Fig. 17**.

Also, the instantaneous failure rate can be calculated for the products which were sorted out by an 8-hour burn-in in the same manner as the above case, where the time t_1 is defined as the product of the burn-in time multiplied by the acceleration factor. The calculation results are shown in **Fig. 18**. The burn-in screening eliminates some early failure rate, and the rate of the failures appears to be flat during the corresponding field period to the burn-in time, followed by the downward curve as time passes.

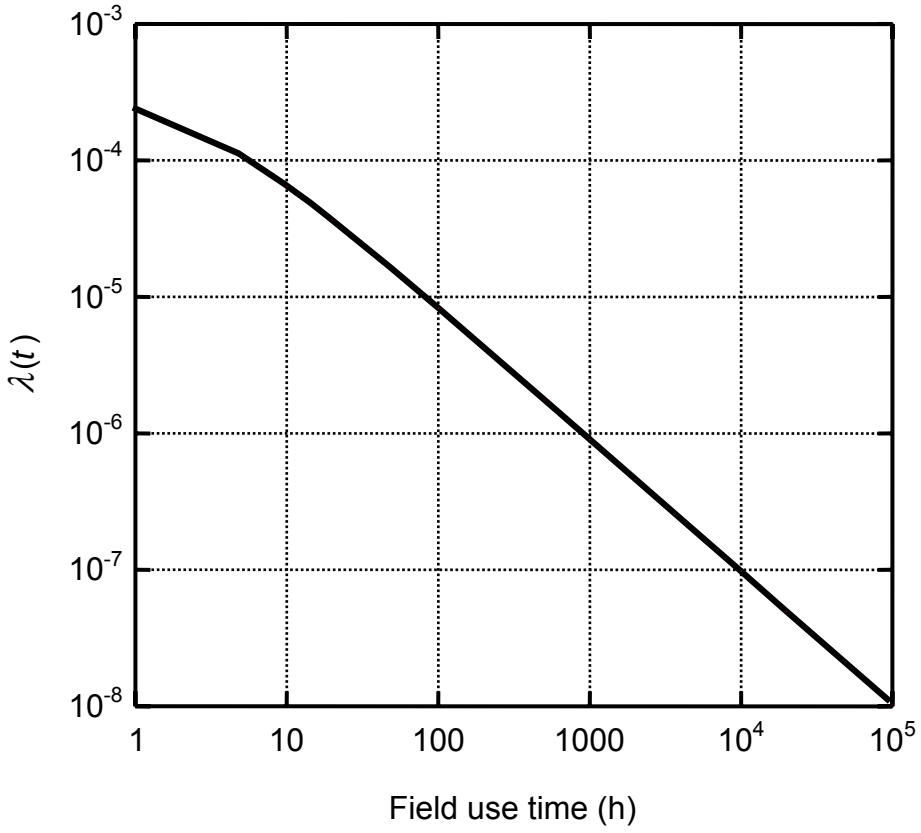

Fig. 17 Instantaneous failure rate of the products screened by the probe tests alone

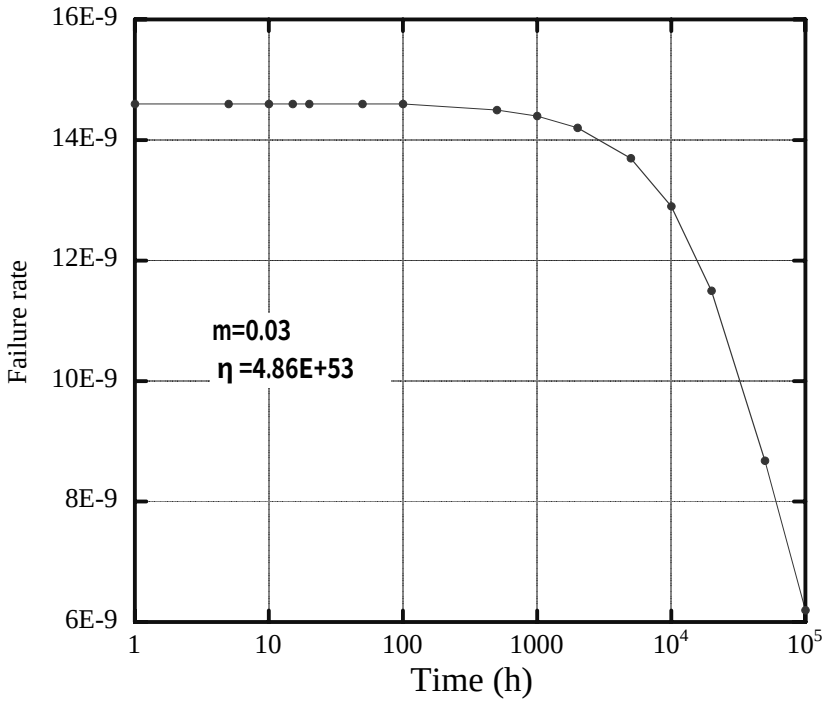

Fig. 18 Instantaneous failure rates of the products screened with an 8-h burn-in stress

2.4.5 Summary: prediction method of EFR

This section compares the prediction methods of EFR described in the previous section.

EFR can be expressed by the failure rate in FIT or the cumulative failure probability in ppm.

a) Predicting EFR form the average failure rate

For example, EFR in a certain period (e.g. 1000 h) can be predicted from the number of the failures found in the burn-in screening.

The average EFR in the period of 1000 h is calculated by

The average failure rate = (Total failures/sample size)/(1000 h x acceleration factor) ····· (34)

b) Predicting EFR from the Weibull distribution

Since the failure rate and the cumulative failure rate are expressed as a function of time, the EFR is the cumulative failure probability in a certain period (e.g. 1000 h). Also the failure rate declines as time passes, the cumulative failure probability is a time-dependent function.

One example indicates the differences between the instantaneous failure rate and the average failure rate for the products screened by an 8-h burn-in with the parameters of m=0.03, η=4.86E+53. It is estimated that the instantaneous failure rate $\lambda(t)$ stays almost constant (14.4 FIT) for a certain period after burn-in, as shown in **Fig. 18**.

Whereas, the cumulative failure rate increases over time and reaches 14.0 ppm at 1000 h in the field operation time. The average failure rate at this point is calculated as

Average failure rate = 14.0E-6/1000 = 14.0 FIT

Therefore, there are few differences between the instantaneous failure rate and the average failure rate of the product whose potential defects were removed by screening.

Take the other case in comparison; the products having the same parameters of m=0.03 and η= 4.86E+53 were sorted out by probe test, not by burn-in.

The comparison between the instantaneous failure rate and the average failure rate is shown in **Fig. 19**.

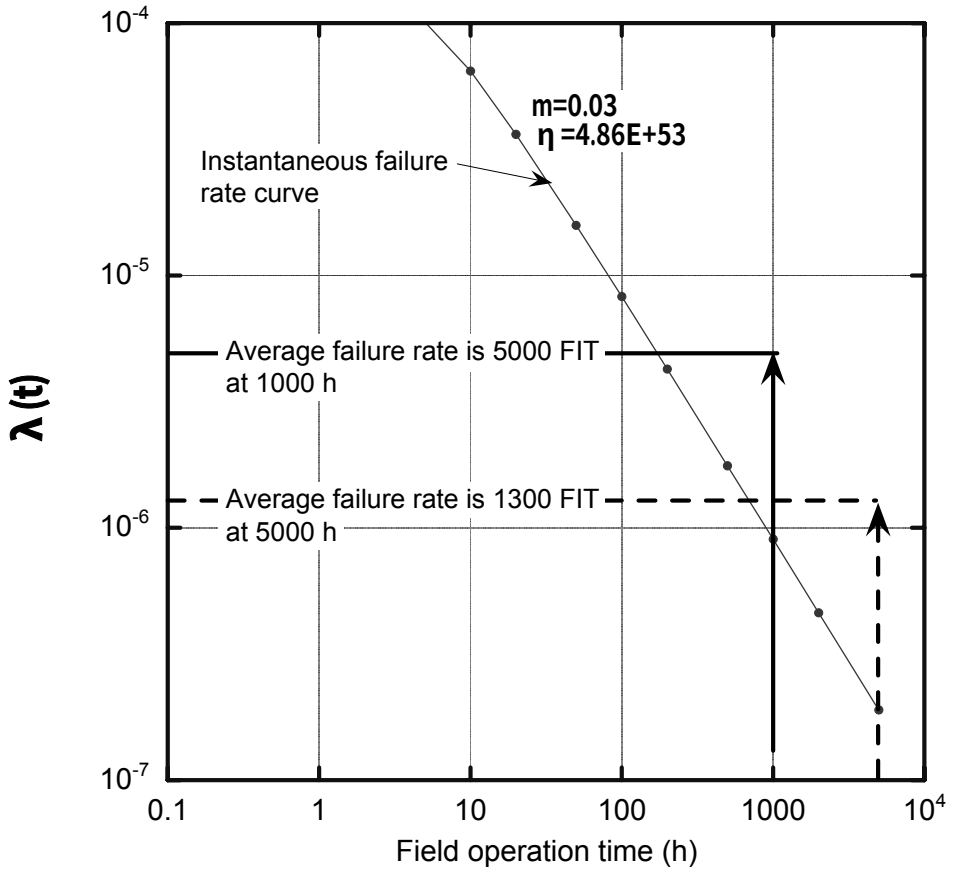

Fig. 19 Instantaneous failure rate vs. average failure rate

It is postulated that the instantaneous failure rate lowers continuously. The average failure rate is, for example, calculated from the failure rate 0.5% at 1000 h.

Average failure rate at 1000 h = 0.005/1000 = 5000 FIT

The instantaneous failure rate at 1000 h is 900 FIT; therefore, the average failure rate in this case appears to be larger than the instantaneous failure rate. The average failure rate at 5000 h, as is obtained in the same way, is 1300 FIT. It is still higher than the instantaneous failure rate, because the failure rate keeps lowering as time passes. The average failure rate projects higher failure rate, which safer side, but the actual status is shown by the instantaneous failure rate.

c) Predicting the cumulative failure probability (ppm)

Here is an example of EFR defined in ppm not in FIT. **Fig. 20** shows the cumulative failure probability $F(t)$ after 8 h of burn-in of the products having the same parameter of m=0.03 and η =4.86E+53. For example, EFR can be estimated to be 14 ppm in the cumulative failure probability at 1000 h, as shown in **Fig. 20**.

Note: EFR in ppm is not necessarily the cumulative failure probability at 1000 h, but at any field time.

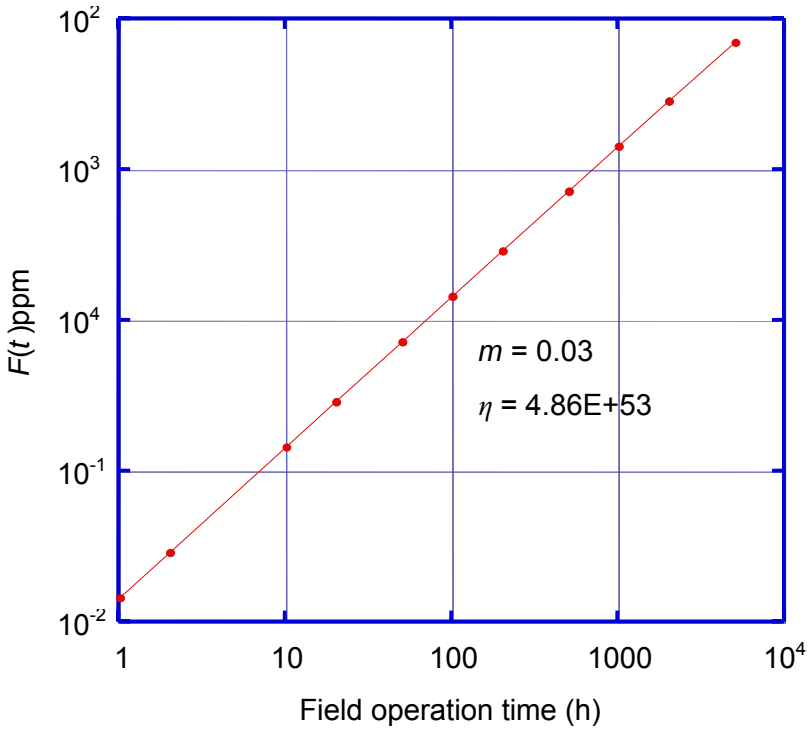

Fig. 20 Prediction of the EFR in the cumulative failure probability (ppm)

2.5 Characteristics of the long-term reliability test

The stress test generally covers the period equivalent to several times to several hundred times of the use time in the field based on its acceleration factor. It is a sufficient period to confirm that none of the wear-out failure occurs during the use time; the extension of the test duration twice or three times would be meaningless to evaluate the failure rate during the use time.

Unless the fundamental structures or the basic process specifications of the semiconductors are modified, significant change will not be seen in the trends of the wear-out failures (inherent failures of the semiconductors). Therefore, it is sufficient to perform only once the long term reliability tests whose duration is equivalent to several times of the use period. If the test duration is equivalent to several times to several hundred times of the use period, and even if the lifetime of a certain lot is found to be half the length of other lots, its lifetime has still an enough margin to the actual use time. From this aspect, it makes little sense to repeat the endurance test for multiple production lots (to find the dispersion) in terms of the failure rate prediction in the use period.

2.6 Significance of the stress test that is not correlative to the field condition

If none of the acceleration characteristics can be found in the stress test, in other words, if there is not any correlation of the stress conditions with the field environment, the stress condition is something unreal and generates failure mechanisms that would not happen in the real operating conditions. Such stresses are meaningless to predict the failure rate in the use environment.

Spending cost to take measures to these failures is also meaningless to improve the product's reliability.

Some examples of the tests which are not correlative to the field conditions are the pressure cooker test with saturated water vapor pressurization, mechanical shock test for the hermetic package, and centrifugal acceleration test for the plastic package. **EIAJ ED-4701** describes the termination of the pressure cooker test with saturated water vapor pressurization and the clarification of the application scope.

3. Accelerated life test

3.1 Principle of the accelerated stress test

The reliability test simulates the stresses in the use conditions (See **Table 6**), but it takes too long to find a failure, or it does not find a failure in a limited test time. Therefore, the stress conditions are employed to accelerate the degradation processes of the devices. It also shortens the time of the evaluation, prediction of the lifetime, and prediction of the failure rates in the use conditions.

Since the acceleration factors differ among failure mechanisms, the properly accelerated stress conditions shall be applied, provided that the failure modes and mechanisms are the same as those in the use conditions. Excessively accelerated stress conditions would induce the different failure mechanisms; care has to be taken to determine the test conditions. In general, the accelerated stress tests aim at finding the failures in the early failure period and in the wear-out failure period.

There are two kinds of accelerated stress tests:

- Constant stress test is to measure a time to failure distribution in the multiple stress conditions.
- Ramp stress test is to raise the stress in equal stress increments at equal time intervals to monitor which level of the stress caused failures.

These test methods are used separately or in combination, and the device lifetime or failure rate is predicted based on the acceleration factors.

The results of the ramp stress tests do not always meet those of the constant stress tests; therefore, enough care shall be taken accordingly.

Some accelerated stress tests are able to determine the end of life, referred to as a limit test. The limit test is conducted for the evaluations, such as the mechanical strength, the durability of the electrical surge and overload, and length of the useful life.

One of the typical accelerated tests is the temperature-stress test, and the relation between the constant stress test and the ramp stress test is shown in **Fig. 21** as an example.

The broken lines that connect the corresponding times at the stepped temperatures are in parallel with the border line that connects the corresponding lifetimes at the test temperatures, indicating that the amount of degradation on each temperature level accumulates step by step. Normally the increments are designed to be large enough to be able to neglect the degradation amount caused in the previous steps so that the test analysis is easier. The dielectric-breakdown test of the oxide layer is a typical ramp stress test, whose correlation with the constant stress test, known as Time Dependent Dielectric Breakdown (TDDB), is shown in **Fig. 21**.

Table 6 Typical environmental and operating stress that semiconductors suffer from

Environmental stress		Operating stress		Types of stress
Natural	Temperature and its change	Electromagnetic	Electric current	Steady state or
	Humidity and its change		Voltage	intermittent
	Rain, water		Power surge, noise	Duration
	Wind		Electric field	Stress frequency
	Thunder		Magnetic field	Cycle time
	Atmospheric pressure (low, high)	Heat	Temperature increase	Complex stress
	Dusts, particles		Hot spot	Complex stress
	Molds, bacteria			Stress fluctuation
	Radial ray			Velocity
	Electromagnetic waves			Stress behavior
	Corrosive atmosphere			
Artificial	Soldering			
	Solvent			
	Shock, drop			
	Vibration, resonance			
	Pull, bending			

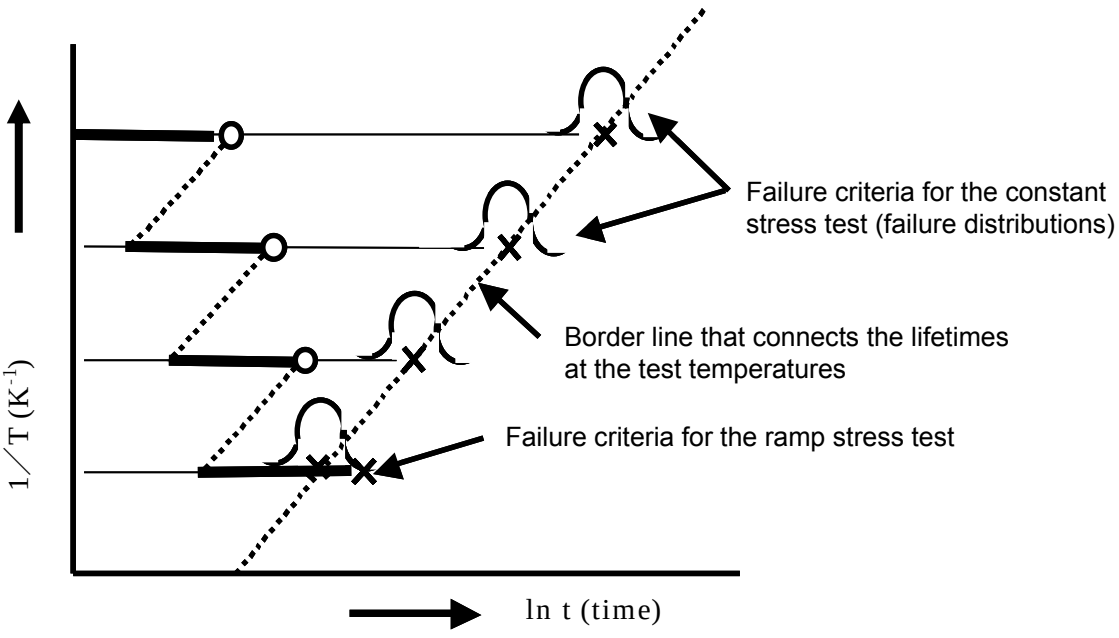

Fig. 21 Accelerated life test (constant stress test vs. ramp stress test)

3.2 Acceleration model

Various acceleration models have been proposed to analyze the data taken from the accelerated stress test. Typical models are described in this section.

3.2.1 Eyring model [2]

Eyring model is an acceleration model dealing with the effects of temperature, voltage and mechanical stresses. The (chemical) reaction rate *K* of the Eyring model is expressed as

$$K = A\left(\frac{kT}{h}\right)\cdot \exp^{-\frac{Ea}{kT}} \cdot \exp^{\left\{ f(s)\cdot\left(C+\frac{D}{kT}\right)\right\}} \quad \cdots\cdots (35)$$

where, *A*, *C*, and *D* are constants;

E_a is activation energy (eV);

k is Boltzmann constant, 8.617×10^{-5} (eV/K);

T is absolute temperature (K);

h is Plank constant;

f(*s*) is non-thermal stress term.

When each term is replaced by

$$A\left(\frac{k}{h}\right)=\Lambda,\quad f(s)=\ln s,\quad F=C+\frac{D}{kT}$$

Then, in the narrow range of T, Eq (35) is approximated as

$$K = \Lambda\ Te^{-\frac{Ea}{kT}} s^{F} \quad \cdots\cdots (36)$$

Acceleration models and acceleration factors for the characteristic parameters (temperature, humidity, voltage, and temperature difference) in the Eyring model are described in **3.2.1.1** through **3.2.1.4**.

3.2.1.1 Temperature acceleration model (Arrhenius Model)/temperature acceleration factor

The functions of the semiconductor devices depend on the chemical and physical reactions on the surface of the solid state substances. Therefore, commonly used is the reaction model where the chemically and physically detrimental reactions make progress over time until devices fails at a certain point. Semiconductor devices are most sensitive to the temperature stress among the electrical, thermal, and mechanical stresses. The dependency of the reaction rate on the temperature stress was found by S. A. Arrhenius, and has been widely used for semiconductor devices as Arrhenius model.

According to the empirical Arrhenius Reaction Rate Model, the reaction rate *K* is expressed as

$$K = A\cdot\exp\left(-\frac{Ea}{kT}\right) \quad \cdots\cdots (37)$$

where, *A* is a constant;

E_a is activation energy (eV);

k is Boltzmann constant, 8.617×10^{-5} (eV/K);

T is absolute temperature (K).

The transition from normal state to degradation state in chemical reaction requires externally provided energy to surmount the energy barrier between them, as shown in **Fig. 22**.

This energy barrier is referred to as activation energy.

Provided that lifetime L is in inverse proportion to the reaction rate K, the temperature acceleration factor α_T is expressed by using Eq. (37) as

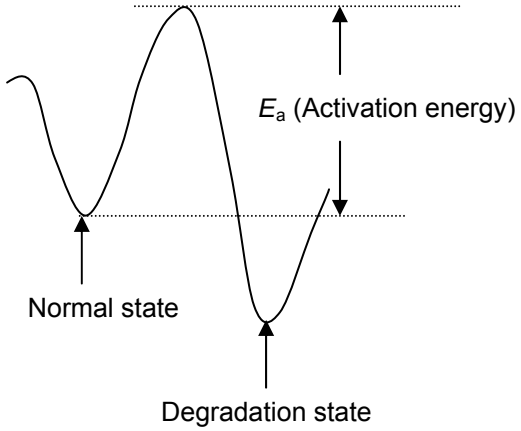

Fig. 22 Activation energy

$$\alpha_T = \frac{L_1}{L_2} = \exp^{\frac{Ea}{k}\left(\frac{1}{T_1}-\frac{1}{T_2}\right)} \quad \text{..... (38)}$$

where, T_1 is temperature in the reference state, e.g., typical field condition;

L_1 is lifetime in the reference state, e.g., typical field condition;

T_2 is temperature in the accelerated condition;

L_2 is lifetime in the accelerated condition.

This equation indicates that the acceleration factor is dependent on the activation energy E_a.

The relationship between the acceleration factor and the activation energy is described in **Fig. 23** as fractions of the acceleration factor at 25 °C. Since the activation energy is inherent to the activation process of individual phenomena, the cause of the failure can be estimated from its activation energy. The failure mechanisms and their activation energy of the semiconductor devices are listed in the **Table 7**.

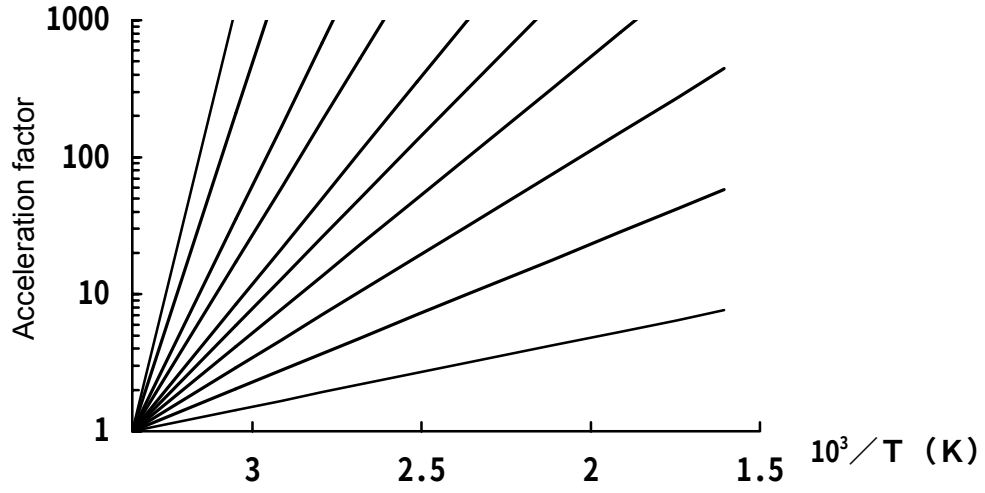

Fig. 23 Acceleration factor vs. activation energy

Table 7 Failure mechanisms and their activation energy of the semiconductor devices

Failure	Failure mechanism	Evaluated device	Activation energy (eV)
Open	Formation of Au-Al intermetallic compound	IC	1.0 to 1.26 [3], [4], [5]
	Electromigration of Al Pure Al Al-Si Al-Si-Cu	IC	0.55 to 0.6 [6]-[11] 0.50 to 0.6 [6]-[11] 0.60 to 0.65 [6]-[11] When these metals are deposited on the high melting temperature metals such as Ti, TiN, Ta, and W, E_a increases to 0.7 to 1.0 V.
	Electromigration of Cu	IC	0.8 to 1.0 [12]-[19]
	Stressmigration of Al	IC	0.9 to 1.1 V [5], [20] for intra-grain diffusion in bamboo-like single grain. 0.5 to 0.7 V [5], [20] for grain boundary diffusion between small grains.
	Stressmigration of Cu	IC	0.74 [21]
	Al corrosion (water penetration)	IC (Plastic)	0.7 to 0.9 [22]-[25]
Short	Time-dependent dielectric breakdown (TDDB)	MOS device	0.3 to 1.1 [5], [26]-[31]
Increased leakage current	Formation of the inversion layer	MOS device	0.8 to 1.0 [32], [33]
Memory endurance	Leakage through silicon oxide layer	IC (EPROM)	0.8 to 1.15 [34], [35]
Threshold voltage shift	Polarized phosphorus glass	MOS device	1.0 [36]
	Drift/diffusion of Na ion in SiO_2	MOS device	0.75 to 1.8 [5]
	NBTI (P^+ GATE PMOSFET)	MOS device	0.4 to 1.1 [37]-[45] E_a differs substantially between reports. See **section 3.2.3**.
	Hot carrier injection (NMOSFET)	MOS device	-0.1 to -0.2 [5]

3.2.1.2 Humidity acceleration model/humidity acceleration factor

Humidity acceleration stress is larger, when,

(a) absorbed moisture amount becomes larger (relative humidity, absolute humidity),

(b) water diffusion rate becomes higher (temperature, humidity), or

(c) resolution or chemical reaction rate becomes higher (temperature, electrical field, current density).

There are many moisture resistance tests that reflect the above conditions. A single test method is not sufficient for humidity acceleration evaluations, because so many humidity-related factors and humidity acceleration models have been reported. Absolute water vapor pressure acceleration model and relative humidity acceleration model #1 and #2 are described here as representative models.

(1) Absolute water vapor pressure acceleration model/water vapor pressure acceleration factor

The time L is the period from the beginning to the certain cumulative failure rate, and a function of the water vapor pressure V_p [46], [47]:

$$L = A \cdot \left(V_p\right)^{-n} \quad \cdots\cdots (39)$$

where, A is a constant and n = 2 (for reference).

The humidity acceleration factor α_{Vp} is expressed as

$$\alpha_{VP} = \frac{L_1}{L_2} = \left(\frac{V_{p2}}{V_{p1}}\right)^n \quad \cdots\cdots (40)$$

where, V_{p1} is vapor pressure and L_1 is lifetime in the reference state, and V_{p2} is vapor pressure and L_2 is lifetime in the accelerated condition.

(2) Relative humidity acceleration model #1/humidity acceleration factor

The time L is the period from the beginning to the certain cumulative failure rate, and a function of the relative humidity RH (%) and temperature T (°C) [48]:

$$L = A \cdot (RH)^{-n} \times \exp\left(\frac{Ea}{kT}\right) \quad \cdots\cdots (41)$$

where, A is a constant; E_a is activation energy (eV); k is Boltzmann constant, 8.617×10^{-5} (eV/K). For reference, E_a = 0.8 eV to 1.2 eV and n is a constant between 4.0 and 6.0.

The humidity acceleration factor α_H is given by

$$\alpha_H = \frac{L_1}{L_2} = \left(\frac{RH_2}{RH_1}\right)^n \quad \cdots\cdots (42)$$

where, RH_1 is relative humidity and L_1 is lifetime in the reference state, and RH_2 is relative humidity and L_2 is lifetime in the accelerated condition.

(3) Relative humidity acceleration model #2 (Lycoudes Model)/humidity acceleration factor

The time L is the period from the beginning to the certain cumulative failure rate, and a function of the relative humidity RH (%) and temperature T (°C) [48].

$$L = A \times \exp\left(\frac{B}{RH}\right) \times \exp\left(\frac{Ea}{kT}\right) \quad \cdots\cdots (43)$$

where, A is a constant; E_a is activation energy (eV); k is Boltzmann constant, 8.617×10^{-5} (eV/K); B differs significantly by failure mode. For reference, E_a=0.8 eV to 1.2 eV, n is a constant between 4.0 and 6.0.

The humidity acceleration factor α_H is expressed as

$$\alpha_H = \frac{L_1}{L_2} = \exp^{B\left(\frac{1}{RH_1}-\frac{1}{RH_2}\right)} \quad \cdots\cdots (44)$$

where, RH_1 is relative humidity and L_1 is lifetime in the reference state, and RH_2 is relative humidity and L_2 is lifetime in the accelerated condition.

3.2.1.3 Voltage acceleration model [49]/voltage acceleration factor

When the effective voltage acceleration model cannot be determined by the experimental results, the following voltage acceleration model has empirically been used often to obtain the reaction rate K.

$$K = A \cdot \exp\left[\left(\frac{B}{t}\right) \cdot V\right] = A \cdot \exp(\beta \cdot V) \quad \cdots\cdots (45)$$

where, A is a constant, B is an electric field constant (cm/V), V is voltage, $\beta=B/t$, and t is a thickness of the gate oxide.

The voltage acceleration factor α_V is expressed as

$$\alpha_V = \frac{L_1}{L_2} = \exp\left[\beta \cdot (V_2 - V_1)\right] \quad \cdots\cdots (46)$$

where, V_1 is voltage and L_1 is lifetime in the typical field condition, and V_2 is voltage and L_2 is lifetime in the stressed condition.

3.2.1.4 Temperature-difference acceleration model/temperature-difference acceleration factor

Eyring model is in general used for the temperature cycle acceleration model, and the relation between the repetition of the temperature difference and the lifetime (cycle number) is expressed as

$$L = A \cdot (\Delta T)^{-n} \quad \cdots\cdots (47)$$

where, L is lifetime, A is constant, ΔT is temperature difference, and n is temperature-difference acceleration factor.

The temperature difference acceleration factor $\alpha_{\Delta T}$ is expressed as

$$\alpha_{\Delta T} = \frac{L_1}{L_2} = \left(\frac{\Delta T_2}{\Delta T_1}\right)^n \quad \cdots\cdots (48)$$

where, ΔT_1 is temperature difference and L_1 is lifetime in the typical field condition, and ΔT_2 is temperature difference and L_2 is lifetime in the stressed condition.

Typical failure modes of the semiconductor devices and the temperature-difference acceleration factors n are shown in **Table 8**.

Table 8 Failure modes and the temperature-difference acceleration factors

Failure mode	n	Literature
Aluminum slide	6	Internally agreed data in the Subcommittee on Semiconductor Device Reliability
Aluminum slide	7.5	Toshiba semiconductor reliability handbook
Passivation crack	11	Internally agreed data in the Subcommittee on Semiconductor Device Reliability
Passivation crack	4.4	SMD assembly and its reliability (Hitachi)
Thin film cracking	8.4	IRPS, 1997, p110
Interlayer dielectric cracking	5.5±0.7	IRPS, 1997, p110
Gold wire breakage	5.2	Hitachi semiconductor reliability handbook
Gold wire breakage	5.3	SMD assembly and its reliability (Hitachi)
Gold wire breakage ϕ25 μm	4	IEICE Technical report R85-16, 1985
Gold wire breakage ϕ30 μm	7	IEICE Technical report R85-16, 1985
Gold wire breakage ϕ50 μm	7	IEICE Technical report R85-16, 1985
Fractured-intermetallic bond	4	IRPS, 1990, p252
Chip-out bond	7	IRPS, 1990, p252
Flip chip solder joint	2.1	Internally agreed data in the Subcommittee on Semiconductor Device Reliability
Degradation of the soldering joint of the die bond	3.4	Reliability and maintainability symposium (Jun. 1, 1991), Union of Japanese Scientists and Engineers (JUSE)
Package crack	5	Toshiba semiconductor reliability handbook

3.2.2 Hot carrier injection

When the strength of the electric field becomes higher in close proximity to the drain of MOS FET, carriers (electrons and holes) gain much higher energy in the event of flowing into this strong electric field. Then some carriers that have higher energy levels than the potential barrier of the Si-SiO_2 interface are injected into the gate oxide. Although the majority of the hot carriers are collected afterward by the gate electrode and generate the gate current, some carriers are trapped in the gate oxide. Hot carrier injection induces the structural instability (interface-state generation) at the Si-SiO_2 boundary. These charges produced in the gate oxide form the spatial distribution and drift the properties (Vth, gm, etc.) of FET. This fluctuation is one of the causes of the reliability degradation.

The following models have been reported with regard to the hot carrier-induced degradation of characteristics.

Note: Although lifetime τ is proportional to $\exp(E_a/kT)$, the temperature acceleration factor is negligibly small because the activation energy E_a is negative in the range of -0.1 eV to -0.2 eV (for reference).

3.2.2.1 Prediction of the DC stress-limited lifetime

(1) Exp(1/Vds) model for Nch and Pch MOSFET [50]

For both Nch MOSFET and Pch MOSFET, when the hot carrier-limited lifetimes τ are plotted on a semi-logarithmic graph as a function of reciprocal numbers of stress voltage V_{ds}, the test data align in a straight line. It means their relation is expressed as

$$\tau = A \cdot \exp\left(\frac{B}{V_{ds}}\right) \quad \cdots\cdots (49)$$

where, A is a constant, and B is a constant in the range of 100 to 200 for Nch MOS FET for reference.

(2) Substrate current model for Nch MOSFET [51]

A double logarithmic plot of the hot carrier-limited lifetime τ and the substrate current I_b in the stress test will yield a straight line for Nch MOSFET. They are expressed as

$$\tau = C \times I_b^{-m} \quad \cdots\cdots (50)$$

where, C is a constant, and m is a constant in the range of 2 to 4 for Nch MOSFET.

For the Nch MOSFET with the submicron technology of 0.5 μm and below, the following equation is said to fit in better.

$$\tau \times I_d = D \times \left(\frac{I_b}{I_d}\right)^{-m} \quad \cdots\cdots (51)$$

where, D is a constant, and Id is the drain current.

Any Nch MOS FET with different gate lengths are said to regress to a common single straight line based on this equation.

(3) Gate current model for Pch MOSFET [52][53]

A double logarithmic plot of the hot carrier-limited lifetimes τ and the gate current I_g in the stress test will yield a straight line for Pch MOSFET. They are expressed as

$$\tau = C \times I_g^{-m} \quad \cdots\cdots (52)$$

where, C and m are constants.

3.2.2.2 Prediction of the AC stress-limited lifetime

(1) Duty method [54]

Duty method aims at extrapolating from the data obtained from the DC acceleration stress test to the actual use conditions. It assumes that the amount of degradation during rising time period t_r and falling time t_f is equivalent to those at the worst DC accelerated stress condition and that the rest of AC-stress time does not cause any degradation.

$$t(AC) = t(DC)\frac{t_{cycle}}{(t_r + t_f)} \quad \cdots\cdots (53)$$

where, $t(AC)$ is the AC stress-limited lifetime to a certain degradation level;

$t(DC)$ is the DC stress-limited lifetime to a certain degradation level;

$t(AC)/t(DC)$ is referred to as duty ratio.

Although this method tends to overrate the degradation rate, this calculation is simple and its accuracy is mostly satisfactory to put into practical use. The duty method includes the method that utilizes the fractions of t_r and t_f or other method that utilizes one of t_r or t_f for calculation, because whole period of t_r and t_f does not contribute to the degradation.

(2) Quasi-static approach [55] [56]

The substrate current changes significantly during AC operations. Switching process is divided into short segments with the substrate current data, and the amount of degradation in each segment of AC operation is regarded as the amount at the same substrate current in DC stress test. The degradation over a cycle time of the waveform is obtained by aggregating the degradation amount of all segments in one cycle t_{cycle}.

The lifetime of AC operation is expressed as

$$t(AC) = \frac{t_{cycle} \times C}{\int_0^{t_{cycle}} [I_b(t)]^m \, dt} \quad \cdots\cdots (54)$$

here, $I_b = a \times I_d \times E_{xd} \times \exp\left(-\frac{b}{E_{xd}}\right)$

$$E_{xd} = \left[\alpha^2 (V_{ds} - V_{dsat})^2 + E_C{}^2\right]^{1/2}$$

where, a, b, α, E_c are fitting parameter;

V_{dsat} is the saturated voltage.

This formula does not properly fit in the higher range of gate-source voltage V_{gs}; therefore, "V_d - $0.6V_{dsat}$" sometimes substitutes for "V_d - V_{dsat}" or the term of exponential function is replaced by the power function of E_{xd} for better conformity.

(3) BERT [57]

Berkeley Reliability Tool, BERT, is a reliability prediction tool developed by C. H. Hu of University of California, Berkeley.

The hot carrier degradation model of BERT is shown as ΔD, degradation amount of the I_d, gm, etc., and its change over time is expressed by the power of time t:

$\Delta D = Kt^n$

The lifetime τ of the Nch MOSFET is given by

$$\tau\left[\frac{I_d}{W}\right] = \frac{H}{K} \Delta D^{1/n} \left[\frac{I_b}{I_d}\right]^{-m} \quad \cdots\cdots (55)$$

The lifetime τ of the Pch MOSFET is likewise given by

$$\tau = \frac{H}{K} \Delta D^{1/n} \left[\frac{I_g}{W}\right]^{-m} \quad \cdots\cdots (56)$$

H and m are both dependent on the drain-gate voltage to be exact, only H is, however, normally taken as the voltage-dependent parameter:

$H = E(V_{ds} - V_{gs}) + F$, or

$\log(H) = E(V_{ds} - V_{gs}) + F$

3.2.3 Negative Bias Temperature Instability (NBTI) in Pch MOSFET

At the interface between Si and SiO_2, there are strains of Si-Si bonds and atomic bonds such as Si-H and Si-OH bonds. These bond powers are so weak that the reactions with injected electrons and holes cause the bond dissociation and form the fixed charges and interface traps, which are the cause of the Vth shift, I_d decline, and G_m degradation.

Particularly, the significant degradation happens under the negative bias temperature (NBT) stress, i.e., when negative bias is provided at the gate electrode of Pch MOSFET at high temperature.

Some acceleration models reported in a past are described below [58].

Voltage acceleration is expressed as

$$L \propto \exp\left(-b \cdot |V_a|\right) \quad \cdots\cdots (57)$$

where, L is the time before reaching a certain amount of the characteristic change (Vth, I_d, Gm);

b is the voltage acceleration factor; b=0.8 for reference;

V_a is the applied voltage.

Temperature acceleration is expressed by Arrhenius model:

$$L \propto \exp\left(\frac{E_a}{kT}\right) \quad \cdots\cdots (58)$$

where, L is the time until reaching a certain amount of the characteristic change (Vth, I_d, Gm);

E_a is the activation energy in the range of 0.4 to 1.0 eV for reference;

T is the absolute temperature (K).

For the buried channel Pch MOSFET with N^+ gate, NBTI was not a major problem in a past. But NBTI has been drawing attention as a reliability problem of gate oxide in association with the development of the finer process technology toward thinner gate oxide and shorter gate length as well as the introduction of the surface channel Pch MOSFET with P^+ gate.

Two acceleration models are proposed for NBTI.

(a) 1/E model [59], [60], [61]

$$\Delta I_d = A \cdot t^n \cdot \exp\left(-\frac{C}{Eox}\right) \cdot \exp\left(-\frac{E_a}{kT}\right) \quad \cdots\cdots (59)$$

where, ΔI_d is the drain current shift from the initial value;

C is the electrical field-dependent constant.

Here, ΔI_d is expressed by the voltage V_G instead of E_{ox} in the reports, [59] and [61].

$$\Delta I_d = A \cdot t^n \cdot \exp\left(-\frac{C}{V_G}\right) \cdot \exp\left(-\frac{E_a}{kT}\right) \quad \cdots\cdots (60)$$

(b) E model [62], [63], [64]

$$\Delta I_d = A \cdot t^n \cdot \exp(D \cdot E_{OX}) \cdot \exp\left(-\frac{E_a}{kT}\right) \quad \cdots\cdots (61)$$

where, ΔI_d is the amount of change from the initial drain current;

D is the electrical field dependent constant.

The values of activation energy E_a of NBTI in P^+ poly-gate devices differ significantly among the reports [37] through [45].

The degradation model of NBTI is often expressed by the characteristics degradation as an index, such as ΔVth, ΔI_d, and $\Delta I_d/I_d$, which are seen in Eq. (59) through (61). The characteristic degradation shows the time dependency in proportion to t^n (n is normally 1/5 to 1/4). When it is expressed by the degradation time as an index, as seen in Eq. (58), the activation energy in Eq. (58) will be 1/n times as high activation energy obtained by characteristic degradation (1/n=4 to 5 times). Thus, whenever activation energy E_a is quoted, the index shall be clearly mentioned if it is the characteristic degradation or the degradation time. The degradation time is used as the index in **Table 7** to facilitate the comparison with other mechanisms.

3.2.4 Time dependent dielectric breakdown (TDDB) [65]-[69]

Time Dependent Dielectric Breakdown (TDDB) refers to the phenomenon that the dielectric layer ends up with the breakdown as a result of long lasting stress, such as the voltage stress or the current stress. The typical acceleration models for TDDB, **(1)** through **(4)** in general, are proposed.

(1) E_{ox} model

E_{ox} model claims that *TTF* is obtained by the following empirical equations with the functions of temperature and electric field strength across the oxide layer.

(a) $$TTF = A \cdot \exp\left(\frac{Ea}{kT}\right) \cdot \exp\left(-\gamma \cdot E_{ox}\right) \quad \cdots\cdots (62)$$

or

(b) $$TTF = A \cdot \exp\left(\frac{Ea}{kT}\right) \cdot 10^{-\gamma \cdot E_{ox}} \quad \cdots\cdots (63)$$

where, *TTF* is the Time To Failure; *A* is a constant; E_{ox} is the electric field strength at the oxide layer (MV/cm); *k* is Boltzmann constant, 8.617×10^{-5} (eV/K); γ is the field acceleration parameter; and E_a is the activation energy (eV).

The electric field acceleration factors γ have been reported along with these models, but these values differ significantly among reports. It indicates that the failure model is not as simple as the equation above.

For reference, γ is in the range of 2 to 4 cm/MV for the case **(a)**

γ is in the range of 0.87 eV to 1.74 cm/MV for the case **(b)**, and

E_a is in the range of 0.3 eV to 1.1 eV [5].

The references [26] through [31] indicate that

E_a is in the range of 0.6 eV to 1.0 eV for T_{ox} of 6 nm or thicker, and

E_a is in the range of 0.75 eV to 1.0 eV for T_{ox} of 6 nm or thinner.

JEDEC Publication 122C [5] indicates that

E_a is in the range of 0.6 eV to 0.9 eV for intrinsic failures, and

E_a is 0.3 eV for extrinsic defects (obtained by burn-in).

The data plot in **Fig. 24**, activation energy E_a vs. thickness of the gate oxide, indicates how they spread in the range of E_a=0.3 eV to 1.1 eV. Also, **Fig. 25** indicates that the activation energy E_a, i.e. temperature dependency, is affected by the strength of the electric field.

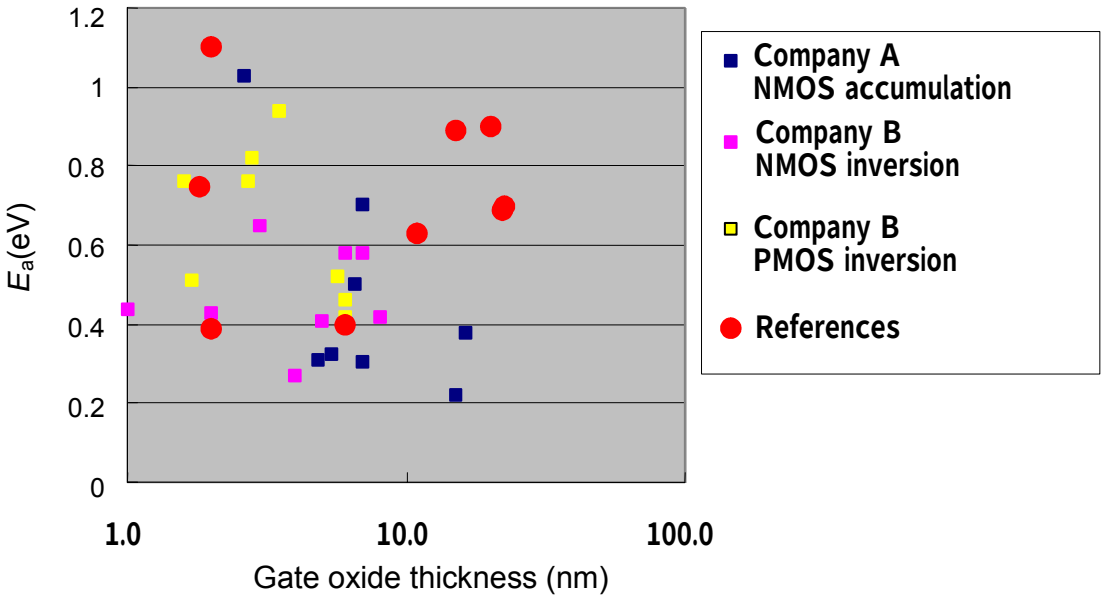

Fig. 24 E_a vs. Gate oxide thickness (References [72]-[73] and the data from **JEITA** affiliates)

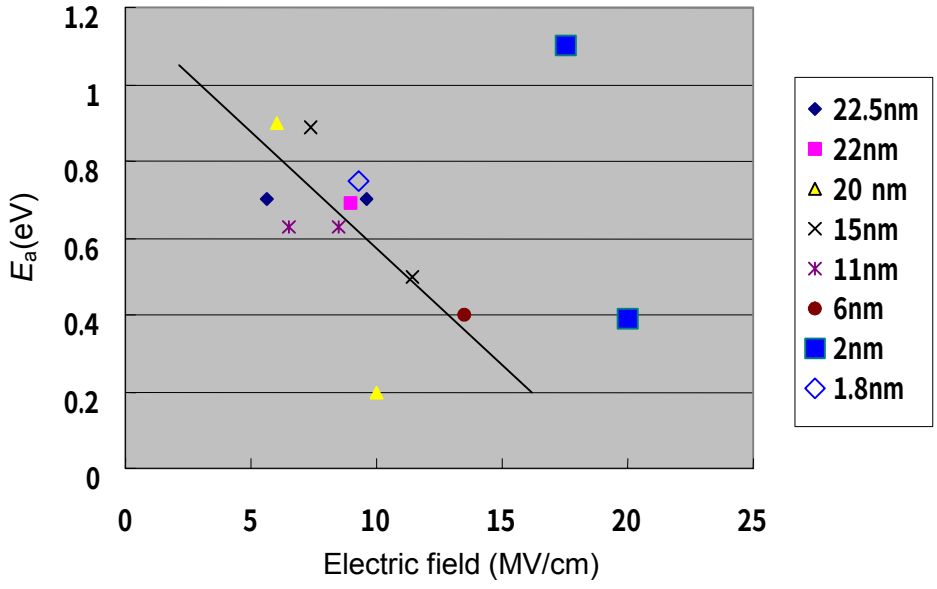

Fig. 25 E_a vs. Electric field (References [26]-[31])

(2) $1/E_{ox}$ model [70]

The $1/E_{ox}$ model claims that *TTF* is expressed by Eq. (64) with the functions of temperature and the electric field strength across the oxide layer. This acceleration model is based on the following postulation: Carriers are injected into the oxide dielectric through Fowler-Nordheim (FN) tunneling at the electric field strength from 5 to 6 MV/cm or higher, which produces an electric current. Then, the injected holes and electrons are accelerated inside oxide layer. The collisions of them induce ionization and the generated holes cause the dielectric breakdown.

$$TTF = A \cdot \exp\left(\frac{Ea}{kT}\right) \cdot \exp\left(\frac{G}{E_{ox}}\right) \quad \cdots\cdots (64)$$

where, *TTF* is the Time To Failure; *A* is a constant; E_{ox} is the electric field strength at the oxide layer (MV/cm); *k* is Boltzmann's constant, 8.617×10^{-5} (eV/K); *G* is the field acceleration parameter (MV/cm); and E_a is the activation energy (eV).

The $1/E_{ox}$ model provides more optimistic lifetime prediction than the E_{ox} model does. Both models, however, have been used because their failure mechanisms have not been clarified which model to use.

(3) V_G model [71]

The following model is proposed for the very thin gate oxide whose thickness is 5 nm or thinner. The TDDB lifetime has been described as the functions of the ambient temperature, total oxide area, and electric field strength across oxide. Typical TDDB models are the E model explained in (1) and 1/E model explained in (2), where E stands for the electric field strength. But when the oxide layer is 5 nm or thinner [72], the *TTF* follows the applied voltage rather than applied electric field strength.

$$TTF = A \cdot \exp\left(\frac{Ea}{kT}\right) \cdot \exp\left(-\gamma \cdot V_G\right) \quad \cdots\cdots (65)$$

where, *A* is a constant; *k* is Boltzmann's constant, 8.617×10^{-5} (eV/K); V_G is the gate voltage (V); γ is voltage acceleration factor (1/V); and E_a is the activation energy (eV).

(4) V_G^{-n} model (Power law model) [72]

The following model has also been proposed for the oxide layer with the thickness of 2 nm or less ($T_{ox} \leq 2$ nm).

$$TTF = A \cdot \exp\left(\frac{Ea}{kT}\right) \cdot \left(V_G\right)^{-n} \quad \cdots\cdots (66)$$

where, *A* is a constant; k is Boltzmann's constant, 8.617×10^{-5} (eV/K); *n* is the voltage acceleration exponent; V_G is the gate voltage (V); and E_a is the activation energy (eV).

3.2.5 Aluminum and copper electromigration

Electromigration is a phenomenon that metal ions migrate by momenta exchange between the current carrying electrons and the host metal lattice when they collide together. In fact, metal-ion migration generates voids or hillocks in the metallization, which result in open- or short-circuit failures.

The acceleration model is commonly expressed by the Black model:

$$TTF = A \cdot J^{-n} \exp\left(\frac{Ea}{kT}\right) \quad \cdots\cdots (67)$$

where, A is a constant; J is the current density; and n is the current-density acceleration exponent.

The development of the finer pattern process adopted copper metallization rather than aluminum metallization. For reference, E_a and n are quoted below from numerous reports [6]-[19].

E_a of pure Al : 0.55 eV - 0.6 eV

E_a of Al-Si : 0.50 eV - 0.6 eV

E_a of Al-Si-Cu: 0.60 eV - 0.65 eV

n : 1 - 3

When these metals are deposited on the under-barrier metals such as Ti, TiN, W, etc., E_a takes larger values in the range of 0.7 to 1.0 eV.

E_a of copper : 0.8 eV - 1.0 eV

n : 1 - 2

3.2.6 Stress migration

Stressmigration typically occurs when a sample is stressed at high temperature with no electric current. The end result is the open-circuit failures due to the formation of voids in the metal traces.

Basically the stress migration occurs in the mitigation process of the tension in the metal traces.

The ratio of the voids caused in a trace, width of which is 2 μm or less, cannot be ignored because its tensile strength becomes larger. Void generation phenomena are categorized into the high-temperature mode and the low-temperature mode with respect to the criteria of 673 K.

The relationship between the volume percentage of the voids and the temperature is shown in **Fig. 26** [20].

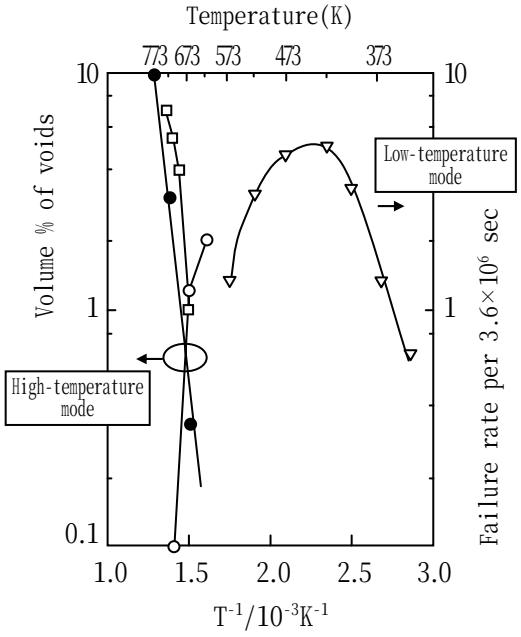

Fig. 26 Volume % of voids vs. temperature

The high-temperature mode happens during wafer fabrication process. It is mainly due to the heat treatment such as dielectric-layer formation, passivation-layer formation, and molding process in packaging. Voids will stop growing between 1 and 10^4 s.

The volume percentage of the voids in the low-temperature mode peaks in the range of 150 to 200 °C (423 to 473 K), which is harmful in the actual use conditions of LSI.

The acceleration model is typically expressed as

$$TTF = A \cdot (T_0 - T)^{-n} \cdot \exp\left(\frac{Ea}{kT}\right) \quad \cdots\cdots (68)$$

The Arrhenius model can be applied as an acceleration model, when the environmental temperature is lower than the temperature at which the void percentage was the highest.

$$TTF = A \cdot \exp\left(\frac{Ea}{kT}\right) \quad \cdots\cdots (69)$$

where, A is a constant and T_0 is the stress-free temperature (K).

The stress-free temperature is close to the metal-deposition temperature, thus, there is little stress difference between the dielectric layer and each metal trace. Some examples of E_a and n are shown below for reference.

For aluminum traces **JEDEC Publication 122C** [5] gives a description that

n is from 2 to 3,

E_a is from 0.5 to 0.6 eV for small grains, and

E_a is approximately 1 eV for bamboo-like single grain.

For copper traces, there are a few reports [21] citing that n is 3.3 and E_a is 0.74 eV for the voids locating under the vias.

4. Practices of the reliability tests for the semiconductor devices

Semiconductor devices are one of the products which have relatively longer lifetimes among industrial products. For that reason, the reliability tests for the semiconductor devices are often performed from a different viewpoint point than common practices. This section describes the concept peculiar to the reliability test for the semiconductors.

4.1 Reliability test for the semiconductor devices and the failure distributions

There are mainly two kinds of the reliability tests to be performed when the new semiconductor devices are being developed.

One of them is “process reliability test” for the newly developed process by using the test element group (TEG) that is specifically designed to evaluate the process.

The other is “product reliability test” for the qualification of the products or the samples whose die sizes or circuit scales are similar to the products. In either case the test time will be minimized by employing the appropriate accelerated life tests. The failure mechanisms are accelerated under the stress conditions, which provide the prediction of the product lifetimes in the use conditions based on the correlations of these with the lifetimes at the stress conditions. Generally failures that belong to either “early failure period” or “wear-out failure period” can be found at the accelerated stress conditions.

"Random failure period" is hardly distinguished from other periods in the reliability test data and often identified as the extended period of the "early failure period." In other words, it is common to perform the reliability test focusing on two periods; "early failure period" and "wear-out failure period."

(1) Process reliability test

The process reliability test aims at identifying the reliability limitations of the new process, establishing the design rules, and verifying those. The reliability limit is the time when a device becomes worn out and no more functioning, being determined by the evaluation of the moment-to-moment changes in the particular parameters that were assigned to each test item in the accelerated life test. The process reliability test examines the potential constraints on the product lifetime, such as electromigration (EM), time dependent dielectric breakdown (TDDB), and hot carrier injection (HCI). This test evaluates the wear-out period where the shape parameter m of the Weibull distribution is greater than one and determines the maximum lifetime of the products produced in the evaluated process.

Note: The TDDB test is used to evaluate the defect density in the gate dielectric in the early failure period (m<1) as well as the reliability limit in the wear-out period.

(2) Product reliability test

The reliability test of the semiconductor device (product reliability test) aims at predicting the failure rate of the product in the use conditions or assuring the product reliability, and it shall be distinguished from the process reliability test described above. When the test failures are plotted on the Weibull sheet, they are expected to be categorized in the early failure mode with m<1 and should be confirmed to be immune from the wear-out failure mode with m>1 on the sheet as well.

a) Evaluations in m>1 region

The wear-out failures shall not appear in the normal use conditions. For confirmation of the failure mode, the test time has to be extended at least to the time equivalent to the product's lifetime plus some allowance time with the consideration of the acceleration factor.

b) Evaluations in m<1 region

The stress duration of the reliability test is normally 1000 h maximum in the case of the high temperature operation test and all failures found in this stress duration usually fall in m<1 region. This test duration corresponds to the period of 10 to 30 years in the typical use conditions depending on the stress conditions.

The assurance period of 10 years is normally sufficient for the electronic products. If longer assurance period is required in the use conditions, the reliability designs for such requirements will be performed separately. When m<1, the failure rate declines as time passes. For example, the prediction of the failure rate based on the data within 1000 h in the high temperature operation test is always larger than that based on the data beyond 1000 h. Therefore, it is meaningless to extend the test duration longer than necessary in the region of m<1.

4.2 Concept of the screening

As stated before, the failure rate of the semiconductor devices appears high just after produced but decreases as time passes. With regard to MOS IC (See **Note 1**), semiconductor manufacturers have the reliability targets (maximum failure rates at the shipment) set for the applications based on the customer's requirements (See **Note 2**). When it were found that the failure rates of the products were higher than the criteria, semiconductor manufacturers would carry out screening at the final inspection to ensure the reliability target.

For the purpose of removing the potential defects which may fail in the early failure period, as stated in section **2.2.5**, the high voltage screening is occasionally performed at the probe test before conveying wafers to the packaging process other than burn-in screening after packaging.

For new process development, it is so important to understand the failure-rate transition from time to time that semiconductor manufacturers repeat the burn-in for several thousands to several ten thousands pieces of products, e.g., test → burn-in for 8 h → test → burn-in for 16 h → test…. Then, plotting failure rates of the burn-in cycles on the Weibull probability paper will provide the transition curve over time. The test time will be several hours to several ten hours in total.

The reliability target of the semiconductor device can be stipulated in terms of the failure-rate transition, the annual average failure rate, or the annual cumulative failure rate (See **Note 2**). In either case, the indexes shall be expressed in the use conditions, i.e. the acceleration factor for the burn-in shall be predetermined to compare the reliability target with the failure rate transition derived from the burn-in cycles.

Since the acceleration factors difer among failure modes, the failures found in burn-in cycles shall be analyzed and categorized to determine the appropriate acceleration factors. Chief failure causes are gate oxide leak and metal foreign materials.

The both failure modes of the gate oxide leak and metal foreign materials are considered to follow the same acceleration model and acceleration factor (See **Fig. 27**) which are predetermined by the TDDB test of gate oxide at the process development. Burn-in aims at removing the early failures (extrinsic defects) derived from the manufacturing faults and so on. Therefore, it is ideal to utilize the acceleration factor obtained from the $m<1$ region of the Weibull distribution obtained from the TDDB test. But obtaining the reliability acceleration factor from TDDB test requires a large number of the TEG samples that are designed to detect the extrinsic defects; it is practically difficult. Therefore, it is common to utilize the acceleration factor obtained from the $m>1$ region as the burn-in acceleration factor.

Since the burn-in cycle evaluation is conducted when a new process is developed, it is not necessary to repeat the burn-in cycle test for every new product. It is because the failure rates of the new products made in the same process can be extrapolated from the process test data and products specifications such as the normalized die area, number of gates, transistor density. The burn-in screening conditions of the product are determined from the reliability target, acceleration factor, and the results of the burn-in cycle evaluation. The failure rates of the particular products are monitored in the specified burn-in screening conditions. It intends to find the abnormal lots that do not meet the expectations and to prevent the unintentional shipment of the failures.

The conventional method to estimate the average failure rate is the high temperature operation test over 1000 h with the sample size of 15 to 77 pieces. The failure rate curve obtained from the high temperature operation test fits the failure rate transition curve over time (Bathtub curve) extrapolated from the burn-in cycles, even though the test time extends to 1000 h. In other words, it is valid to predict the failure rate based on the results of the burn-in cycles in a shorter test time instead of conducting a long-term reliability test over 1000 h.

Note 1: For bipolar IC and discrete semiconductors, typical method is not burn-in cycles but the estimation of the average failure rate. Even for MOS IC the burn-in cycles are not applied in some cases.

2: The reliability targets are such as the failure rate at shipment, the average failure rate in a one-year period after shipment, or cumulative failure rate over a one-year period after shipment. The burn-in conditions can be chosen to be equivalent to the corresponding reliability targets.

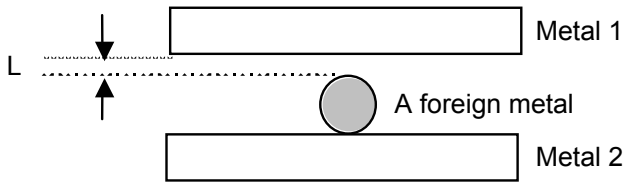

Fig. 27 Functional failures due to a foreign metal

Originally two metal traces were kept insulated, but high temperature and electric field during burn-in induce the dielectric breakdown and cause the short-circuits. Since the gap L in **Fig. 27** is unknown, it is postulated that the thickness of the gate oxide is L and that the acceleration factor is the same as the one for TDDB.

4.3 Recommended operating conditions and reliability assurance

The recommended operating conditions are specified for each semiconductor device, and there are close relationship between the reliability and operating temperature or voltage.

In these years thermal-dissipation designs have become more challenging in association with miniaturization of the electronic products that contain semiconductor devices. It may lead to the problem of the increase in the device junction temperature.

For example, assuming that the recommended operating temperature range is between -40 °C and +85 °C and the activation energy E_a=0.7 eV, the acceleration factor between the stress at the average operating temperature of 35 °C and that of 85 °C is given by

$$\alpha_T = \exp\left\{\frac{E_a}{k}\left(\frac{1}{T_1}-\frac{1}{T_2}\right)\right\}, \qquad E_a = 0.7eV \quad \cdots\cdots (70)$$

This equation indicates 39.5 times larger thermal stress (temperature acceleration factor) at 85 °C than that at 35 °C. Therefore, it should be noticed that the continuous usage at the upper limit temperature, even if it is inside the recommended operating temperature, reduces the device lifetime significantly. It is important to introduce the heat dissipation design with reasonable derating to the system with the consideration of the thermal resistances of the package and PWB.

For example, assuming that the recommended voltage range is between 2.7 V and 3.6 V and the voltage acceleration constant B=3.0 V^{-1}, the acceleration factor between the voltage stress at 3.0 V on average and that at 3.6 V on average is given by

$$\alpha_V = \exp\{\mathrm{B}(V_2 - V_1)\}, \qquad \mathrm{B} = 3.0 \quad V^{-1} \quad \cdots\cdots (71)$$

This equation indicates 6.0 times larger voltage stress (voltage acceleration factor) at 3.6 V than the stress at 3.0 V.

Again it should be noticed that the actual use temperature and voltage, even if it is inside the recommended operating condition limits, affect the device lifetime significantly.

For the special applications that require the reliability assurance for the continuous high temperature operation, the separate considerations and specifications are required.

4.4 Assumption of the application environment

The use conditions of the semiconductor devices depend on the electronic product that contains them and are quite different. The concept of the application environment is described here.

4.4.1 Junction temperature

When the device is used in the typical environment, the ambient temperature Ta is assumed to be 25 °C to 55 °C in average. Extreme application conditions require the separate specifications as stated in the previous section. When the semiconductor devices are used at a certain ambient temperature, the additional increment of the die operating temperature should be taken into consideration. The temperature increase during the device operation depends on the power consumption of the die as well as the thermal resistance of the package and chassis, and is expressed by equation (72).

$$\Delta T = \theta \times P_C \quad \cdots\cdots (72)$$

where, ΔT is temperature increase (°C); θ is thermal resistance of package (°C/W); and P_c is power consumption of the die (W).

4.4.2 Application conditions

The expected lifetimes of the semiconductor devices differ among their applications. The reliability assurance level should be considered separately for each application with the appropriate expected lifetime. Examples of the semiconductor device applications and their expected lifetimes are shown in **Table 9** (Quoted from **JEDEC JESD 94** (January 2004) p.16).

Table 9 Illustrative application conditions (examples only) for a range of applications

Applications	Application Conditions							
	Operating Life (POH)	Field Lifetime (Years)	Environmental & Power Cycles	Environmental Relative Humidity Range (% RH)	Environmental Temperature Range (°C)	Operational Temperature Cycle Range (°C)	Chip Junction Temperature (Tj) Typicall Max. (°C)	Device Nominal Operating Voltage (V)
Desk Top Computer with Energy Saving Features	13,000*	5 years	Main: 1/ day Mini: 17/ day Short: 1/ day	10- 80%	10° – 38°C	Main: 20° – 60°C Mini: 52° – 60 °C Short : 40° – 60 °C	70°C / 105°C	12.0 V
High End Server	94,000*	11 years	4 / year	10 – 80%	10° – 38°C	14° - 55°C	70°C / 105°C	1.2 V
Avionic Electroninca in Cockpit	>150,000*	~ 23 years	Power: 21,500 2.5 / day	5 – 86%	-20° – 60°C	5° – 50°C	70°C / 105°C	3.3V/ 5V
Telecom Hand Held	43,800*	5 years	Talk: 20 / day Standby/ Off: 1/ day	10 – 95%	-40° - 40°C	Talk: 32° - 70°C Standby/ Off: 30° - 32°C	30°C / 70°C	1.8V / 3.3V
Telecom Uncontrolled	131,000*	15 years	Power: 1/ month Environ: 1/ day	85%	-40° - 85°C	Power: Δ 85°C Environ: Δ 25°C	85°C / 110°C	1.2 V
Telecom Controlled	131,000*	15 years	Power: 1/ month Environ: 1/ day	70%	0° - 70°C	Power: Δ 85°C Environ: Δ 6°C	85°C / 110°C	1.2 V
Automotive Underhood (Grade 0)	8200*	15 years	Power: 5 / day	0 – 100%	-40° - 125°C	-40° - 150°C	100°C / 150°C	12.0 V
* POH value assumes worst case 100% power-on over the life of the application. Actual application use POH may be less.								

4.5 Acceleration factors and the efficient accelerated test methods

This section describes some examples of acceleration factors between the stress conditions and the application conditions and discusses the implementation of the efficient accelerated lifetime tests.

4.5.1 Numerical examples of the calculation of the acceleration factors

4.5.1.1 Temperature bias acceleration

Assumptions:

- Mean temperature in the use conditions is 55 °C;
- Electric field across the oxide layer in the use condition is 3.3MV/cm;
- Temperature at accelerated stress condition is 125 °C;
- Electric field strength across the oxide layer at the accelerated condition is 4.2 MV/cm;
- Activation energy E_a is 0.5 eV;
- Voltage acceleration exponent β is 3.5/V;
- Boltzmann constant k is 8.617×10^{-5} eV/K.

The acceleration factor can be derived from the Eq. (38) in section **3.2.1.1** and Eq. (46) in section **3.2.1.3**:

$$
\begin{aligned}
AF &= \alpha_T \times \alpha_V \\
&= \exp\left\{\frac{E_a}{k}\left(\frac{1}{T_1}-\frac{1}{T_2}\right)\right\}\times\exp\{\beta(V_2-V_1)\} \\
&= \exp\left\{\frac{0.5}{8.617\times10^{-5}}\left(\frac{1}{(55+273)}-\frac{1}{(125+273)}\right)\right\}\times\exp\{3.5(4.2-3.3)\} \quad \cdots\cdots\cdots\cdots \quad (73) \\
&\cong 22.45\times 23.34 \\
&\cong 524.0
\end{aligned}
$$

In this case, 1000 h in the high temperature operation test is equivalent to 5.240 x 10^5 h (60 years) in the use conditions.

4.5.1.2 Temperature difference acceleration

Assumptions:

- The mean temperature difference in the use conditions is 50 °C.
- The temperature difference at accelerated stress test is 150 °C - (-65 °C) = 215 °C.

The temperature acceleration factor is derived from the Eq. (48) in section **3.2.1.4**:

$$\alpha_{\Delta T} = \left(\frac{\Delta T_2}{\Delta T_1}\right)^n = \left(\frac{215}{50}\right)^4 \cong 341.9 \quad \text{..... (74)}$$

It indicates that 300 cycles in temperature cycling (-65 °C /+150 °C) is equivalent to 1.0257 x 10^5 cycles in the use conditions.

When the use conditions are supposed to be 5 cycles a day, it is equivalent to 2.0514 x 10^4 days (56 years).

4.5.1.3 Humidity acceleration

There are water vapor pressure-based method and relative humidity-based method to predict the humidity-limited lifetime.

(1) Water vapor pressure-based method

From the section **3.2.1.2**, the water vapor pressure can be obtained from Eq. (40).

Assumptions:

- Temperature and humidity in the use conditions is 30 °C, 60 %RH
  → Water vapor pressure is 0.25 × 10^4 Pa
- Temperature and humidity at test conditions is 85 °C, 85 %RH
  → Water vapor pressure is 4.91 × 10^4 Pa

Water vapor pressure-based humidity acceleration factor is

$$\alpha_{VP} = \left(\frac{V_{p2}}{V_{p1}}\right)^n = \left(\frac{4.91\times10^4}{0.25\times10^4}\right)^2 \cong 385.7 \quad \text{..... (75)}$$

It indicates that 1000 h in temperature humidity bias test is equivalent to 3.857 x 10^5 h (44 years) in the use conditions.

Note: When the water vapor pressure-based method is applied to predict the humidity-limited lifetime, the results obtained from the dry conditions and those obtained from the wet conditions may be inverted in some cases; care have to be taken.

For example,

40 °C, 90 %RH → Water vapor pressure is 0.66 × 10^4 Pa.

85 °C, 20 %RH → Water vapor pressure is 1.16 × 10^4 Pa.

The latter case results in higher water vapor pressure, in other words, it seems to be severer condition than the former case. But the latter case is apparently dry condition, while the former condition is wet condition and has higher humidity acceleration factor.

(2) Relative humidity-based method

The relative humidity-based method is applied with the Arrhenius equation.

Assumptions:

- Average temperature and humidity in the use conditions are 30 °C, 60 %RH;
- Temperature and humidity at the accelerated stress test condition are 85 °C, 85 %RH;
- Boltzmann's constant k is 8.617 × 10^{-5} eV/K.

By substituting the assumptions in Eq. (42) in the section **3.2.1.2** and Eq. (38) in the section of **3.2.1.1**, humidity acceleration factor is

$$a = \alpha_{\mathrm{H}} \times \alpha_{\mathrm{T}} = \left(\frac{RH_2}{RH_1}\right)^n \times \exp\left\{\frac{\mathrm{E}_a}{k}\left(\frac{1}{\mathrm{T}_1} - \frac{1}{\mathrm{T}_2}\right)\right\}$$

$$= \left(\frac{85}{60}\right)^4 \times \exp\left\{\frac{0.8}{8.617\times10^{-5}}\left(\frac{1}{(30+273)} - \frac{1}{(85+273)}\right)\right\} \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad (76)$$

$$\cong 4.03 \times 110.8$$

$$\cong 446.5$$

It indicates that 1000 h in the temperature humidity bias test is equivalent to 4.465 x 10^5 h (51 years) in the use conditions.

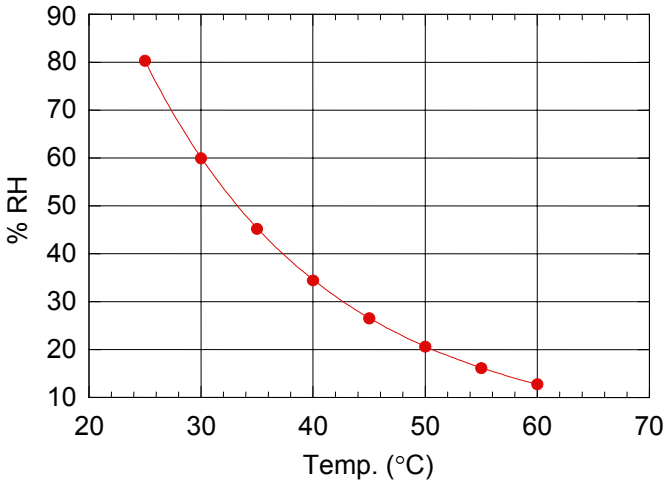

Fig. 28 Temperature vs. humidity

A relative humidity curve which passes through the point of 30 °C and 60 %RH is shown in **Fig. 28** as a function of temperature. It indicates that rising temperature is associated with lowering relative humidity and less influence of humidity.

Therefore, in contrast with the high temperature operation life, it is not practical to premise high temperature in the application environment.

In other words, it is relevant to postulate the field temperature to be 30 °C to predict the humidity-limited lifetime from the results of acceleration test.

4.5.2 Designing efficient acceleration stress tests

The above numerical examples indicate that the traditional criteria, such as 1000 h in the high temperature operation life test or 1000 h in the temperature humidity bias test, are equivalent to the field time that is far beyond either reliability target or expected lifetime.

On the other hand, more efficient test plans are available by taking advantage of acceleration factor with respect to the reliability target.

One example is given here for further explanation. It is postulated that the reliability target is 0.1 % or less in failure rate at the end of product life and that failures are in wear-out mode in this case.

The tangible example is the solder joint reliability of BGA balls. It is presumed that m>5 when the data are simulated to the Weibull distribution. Actually, some references have reported 5.29<m<8.99 [74]. Also it is postulated that the acceleration factor (AF) to the use condition is calculated from the Norris-Landzberg Model (Modified Coffin-Manson model), e.g., AF=4.0.

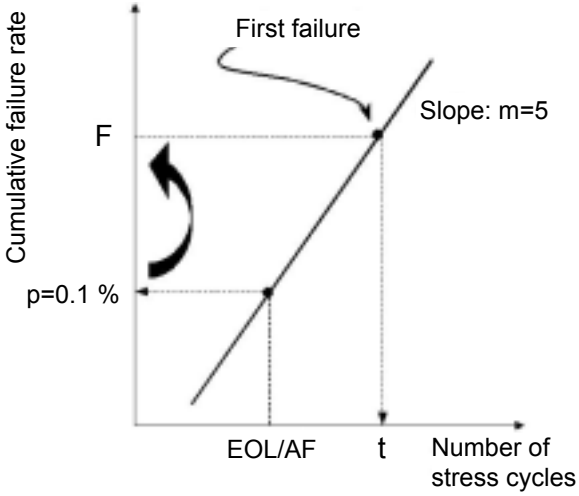

Fig. 29 Weibull plot

When their application is for high-end servers, the typical use conditions are four cycles a year and 11 years in lifetime from **Table 9**; therefore, the devices typically experience 44 cycles until the end-of-life (EOL).

Another assumption is that no failure was found up to t cycles but one failure found just after t cycles. When the number of cycles until the end-of-life is expressed as “EOL cycles” in the use conditions, the number of cycles at the stress test is EOL/AF, where AF is the acceleration factor of the stress test. When the data are plotted on the Weibull probability paper, as shown in **Fig. 29**, the straight line that shows the worst case passes through two points given by

1) number of cycles equivalent to the end of life (EOL/AF) and the cumulative failure rate of 0.1%, and

2) number of cycles when the first failure was found and the cumulative failure rate for one piece of failure.

When the first failure was found out of 30 pieces of samples at t cycles (N=30), there are three methods to estimate the cumulative failure rate of the parent population F.

(1) The simplest method is to regard the cumulative failure rate of the samples as that of the parent population, i.e., F=1/N=1/30=3.33%

(2) The first failure found just after t cycles is regarded as the first failure in the order statistics, and then the cumulative failure rate of the parent population is obtained through median rank method. The approximation formula, F=(1-0.3)/(N+0.4), is used for this calculation.

(3) This is further generalization of (2). When the cumulative failure rate F is strictly estimated at 90% rank, it can be obtained from the inverse beta function.

$$F=\begin{cases}(1)\ \dfrac{1}{N} & (direct\ \ estimate)\\ (2)\ \dfrac{1-0.3}{N+0.4}\ \ or\ \ \begin{pmatrix} Beta\\ inverse\ of\ \ median\ \ rank\\ function\end{pmatrix} & \\ (3)\ \begin{pmatrix} Beta\\ inverse\ of\ \ 90\%\ \ rank\\ function\end{pmatrix} & \end{cases} \quad \cdots\cdots (77)$$

When N=30, F will be

$$F=\begin{cases}3.33\%\ \ (direct\ \ estimate)\\ 2.28\%\ \ (median\ \ ranking)\\ 7.39\%\ \ (90\%\ \ ranking)\end{cases} \quad \cdots\cdots (78)$$

The number of cycle t can be obtained from the straight line that passes through two points; one is the cumulative failure rate of 0.1% at the number of cycle equivalent to the lifetime, and the other is F at t cycles.

$$\begin{cases} p=1-\exp\left\{-\left(\dfrac{EOL/AF}{\eta}\right)^m\right\}\\ F=1-\exp\left\{-\left(\dfrac{t}{\eta}\right)^m\right\}\end{cases} \quad \cdots\cdots (79)$$

Cycle t is given by

$$t=\left(\frac{EOL}{AF}\right)\times\left\{\frac{\ln(1-F)}{\ln(1-p)}\right\}^{\frac{1}{m}}=\left(\frac{44}{4}\right)\times\left\{\frac{\ln(1-F)}{\ln(1-0.001)}\right\}^{\frac{1}{5}}$$

$$=\begin{cases}23\ \ cycles & \left(direct\ \ estimate\right)\\ 21\ \ cycles & \left(median\ \ ranking\right)\\ 27\ \ cycles & \left(90\%\ \ rannking\right)\end{cases} \quad \cdots\cdots (80)$$

Therefore, based on the extremely conservative view of 90% rank, if no failure was found out of 30 pieces of samples at 27 test cycles, the parent population is assumed to meet the reliability target. Similar calculations are available for any sample size of *N*.

4.5.3 Grouping concept for simplifying the reliability test

(1) Grouping

The reliability test for semiconductor devices are mostly carried out at the development phase of the process or the product. The objects of the evaluations are new processes, new products, new packages, etc. The test plan can be simplified by grouping objects having common features, choosing representatives for each group, and regarding the test results as those for the entire objects in the same group. This is referred to as the reliability test on the basis of the grouping concept.

(2) Examples of grouping

This section describes the case of qualifying the objects (product groups) that are fabricated through the same wafer process. The products produced in the different processes may often require the separate qualifications, but what makes it different can be determined individually at each test.

The examples of the reliability tests where grouping concepts are introduced are as follows:

a) Grouping by product as a silicon die

From the view point of the following characteristics, products are classified into groups, and then at least one representative combination of the characteristics is subjected to the reliability test.

Examples:

- Products having the similar functions and characteristics
- Designs based on the same technology (standard cell, gate array, etc.)
- Designs with the common library

When the device features a new functional cell by means of standard cell and so forth, the necessity of the reliability test can be considered individually.

b) Grouping by die size

Bigger die typically comprises larger number of transistors and tends to have more defects that may affect the reliability. Therefore, in some cases the qualification of the larger-size die represents the smaller-size dice as well.

c) The result of the reliability test for the product having larger size memory circuits on a die can represent the qualifications of the products having smaller memory sizes, as long as the memory-circuit configurations are similar.

d) Grouping by package type

Although there are various package types, such as QFP and CSP (e.g., package A and B), it is possible to group the package types that comprise the same materials. Some examples are shown below. Grouping concept of the package-level lifetime test and environmental tests is different from the board-level reliability test; care has to be taken. Even if the packaging materials are the same, the reliability test has to be done separately if the packaging manufacturers are different.

The following perspectives have to be considered for the reliability test.

- Body size of the package

  If the reliability test is performed for the package with a certain body size, the same package type with the same body size or smaller is often exempt from the duplicated test. The same package type with a different package thickness may require the separate reliability test in some cases. In other words, the definition of the body size shall be determined separately whether it is three-dimensional or two dimensional (disregards to the package thickness).

- Ball count and ball pitch of CSP

  When the reliability test is performed for a package, the same package type with the same body size but different ball count and different ball pitch is sometimes exempt from the duplicated test. However, these packages may be required the separate board-level test; therefore, grouping has to be done accordingly.

An example of grouping CSPs

- A group is defined as 388 pin CSP (lead-free balls) featuring a two-layer substrate, the same body size, the same body thickness, and the same ball pitch. Even if the packages have the same conditions (the same materials as well) but the different ball counts, the additional reliability test is not required for package-level test.

e) Combination of packaging and wafer processes

The reliability tests are sometimes conducted to find whether there are any reliability problems in the combinations of packaging (package types, packaging manufacturer) and wafer processes. It is possible to choose the appropriate combination by which the whole combinations are represented.

The illustrative figure of the overall grouping concept is shown in **Fig. 30**. It indicates that qualification is conducted for every wafer process technology of the die in a package.

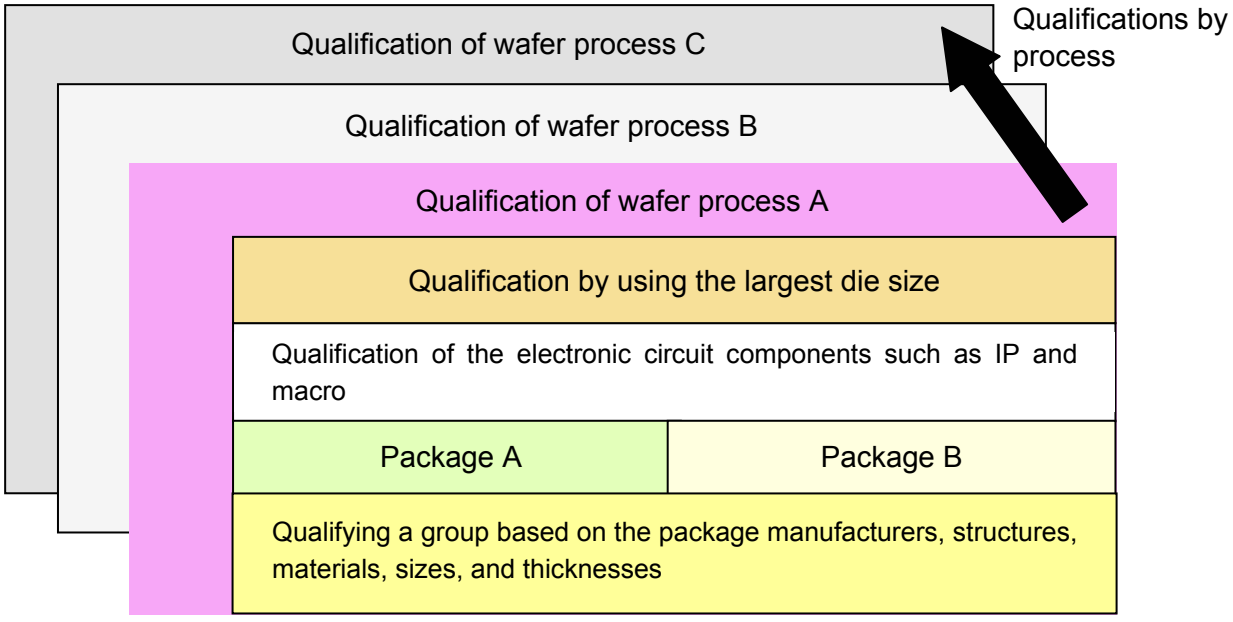

Fig. 30 Qualification scheme by grouping products

References

[1] D. J. Dumin, “OXIDE RELIABILITY A Summary of Silicon Oxide Wearout, Breakdown, and Reliability”, World Scientific, p.179, 2002.

[2] Renesas Technology, “Semiconductor Reliability Handbook (2006)”

[3] J. Wood, et al., “Reliability and Degradation of Silicon Devices and Integrated Circuits, in Reliability and Degradation”, edited by M. H. Howes and D. V. Morgan, Wiley, New York, 1981.

[4] D. S. Peck, “Practical Applications of Accelerated Testing-Introduction”, Proc. Of IRPS, p.253, 1975.

[5] JEDEC PUBLICATION JEP122C, MARCH 2006,“Failure Mechanisms and Model for Semiconductor Devices”

[6] J. M. Towner, “Electromigration-Induced Short Circuit Failure,” Proc. Of IRPS, p.81, 1985

[7] J. C. Ondrusek, et al., “Effective Kinetic Variations With Stress Duration For Multilayered Metallizations”, Proc. Of IRPS, p.179, 1988

[8] J. G. Cottle, et al., “A Comparison Between Noise Measurement And Conventional Electromigration Reliability Testing”, Proc. Of IRPS, p.203, 1988

[9] C. A. Martin, et al., “Electromigration Performance of CVD-W/Al-Alloy Multilayered Metalizations”, Proc. Of IRPS, p.31, 1990

[10] J. A. Maiz, et al., “A Resistance Change Methodology For The Study Of Electromigration In Al-Si Interconnects”, Proc. Of IRPS, p.209, 1988

[11] L. M. Ting, et al., “AC Electromigration Characterization and Modeling Of Multilayered Interconnects”, Proc. Of IRPS, p.311, 1993

[12] C. Ryu, et al., “Effect of Texture on Electromigration of CVD Copper”, Proc. Of IRPS, p.201, 1997

[13] L. Arnaud, et al., “Microstructure and Electromigration in Copper Damascene Lines”, Proc. Of IRPS, p.263, 1999

[14] Reliability Qualification of Copper BEOL, IBM MRS2001

[15] Electromigration Performance of Multi-level Damascene Copper, NEC MicroRel 2001

[16] Electromigration Induced Incubation, Drift and Threshold in Single Damascene Copper Interconnects, NEC IITC2002

[17] K.-D. Lee, et al., “Electromigration Study of Cu/low k Dual- Damascene Interconnects”, Proc. Of IRPS, p.322, 2002

[18] C.-K. Hu, “Bimodal Electromigration Mechanism in Dual-Damascene Cu Line/Via on W”, IBM IITC2002

[19] “Electromigration Failure Mechanism Studies on Copper Interconnects”, Inferion IITC2002

[20] H. Okabayashi, “Stress migration for aluminum”, Material, Vol.36, pp.565-570(1997)

[21] E. Ogawa, et al., “Stress-induced Voiding under Vias connected to Wide Cu Metal Leads”, Proc. Of 2002 IRPS, pp312-321(2002)

[22] C. G. Shirley, et al., “Optimal Acceleration Cycle THB tests for Plastic packaged Devices”, Proc. Of IPRS, p.12, 1991

[23] D. Peck, “Comprehensive Model for Humidity Testing Correlation”, Proc. Of IPRS, p.44, 1986

[24] D. Peck, “Recent Humidity Accelerations, Base for Testing Standards”, Quality and Reliability Engineering 1991

[25] S. M. Tan, “Demonstrated Reliability of Plastic-encapsulated microcircuits for Missile Applications”, IEEE Transactions on Reliability 1995

[26] M. Takayanagi, et al., “Gate Voltage Dependent Model for TDDB lifetime Prediction under Direct Tunneling Regime”, VLSI Tech. Sympo. 2001

[27] J. S. Suehle, et al., “Temperature Dependence of Soft Breakdown and Wear-out in Sub-3nm SiO_2 Films”, Proc. Of IRPS, p.33, 2000

[28] N. Shiono, et al., “A Lifetime Projection Method Using Series Model and Acceleration Factors for TDDB failures of Thin Gate Oxides”, Proc. Of IPRS, p.1, 1993

[29] J. S. Suehle, et al., “Field and Temperature Acceleration of Time-Dependent Dielectric Breakdown in Intrinsic Thin SiO_2”, Proc. Of IRPS, p.120, 1994

[30] J. Prendergast, et al., “TDDB Characterization of Thin SiO_2 Films With Bimodal failure Populations”, Proc. Of IPRS, p.124, 1995

[31] P. Chaparala, et al., “Electric Field Dependent Dielectric Breakdown of Intrinsic SiO_2 Films under Dynamic Stress”, Proc. Of IPRS, p.61, 1996

[32] W. H. Schroen, et al., “Reliability Improvement by Process Control”, Proc. of IRPS, p.42, 1972.

[33] Matsuda, et al., “Recent EPROM having improved data-retention and rewriting characteristics”, Nikkei Electronics, No.216, p.121, 1979. (In Japanese)

[34] Otsuki, et al., “Reliability of the Microprocessor LSI”, Technical group of reliability, The Institute of Electronics, Information and Communication Engineers, SSD76-23, p.35, 1976. (In Japanese)

[35] Miyamoto, et al., “Reliability of CMOS EPROM”, Reliability and maintainability symposium, 9th Union of Japanese Scientists and Engineers (JUSE) p.67, 1979. (In Japanese)

[36] E. H. Snow, et al., “Polarization Phenomena and Other Properties of Phosphosilicate Glass Film on Silicon”, J.Electrochem.Soc., vol.113, p.263, 1966.

[37] Y. F. Chen, et al., “Negative Bias Temperature Instability (NBTI) in Deep Sub-micron P^+-gate PMOSFETs”, IRW FINAL REPORT, p.98, 2000

[38] P. Chaparala, et al., “Threshold Voltage Drift in PMOSFETS due to NBTI and HCI”, IRW FINAL REPORT, p.95, 2000

[39] 11^{th} RCJ Reliability seminar text

[40] C. H. Liu, et al., “Mechanism and Process Dependence of Negative Bias Temperature Instability (NBTI) for PMOSFETs with Ultrathin Gate Dielectric”, IEDM, 39.2, 2001

[41] K. Sasada, et al., “The influence of SiN Films on Negative Bias Temperature Instability and Characteristics in MOSFETs”, IEEE Conference on Microelectronics Test Structure, 1998

[42] N. Kimizuka, et al., “NBTI enhancement by nitrogen incorporation into ultra thin gate oxide for 0.1um gate CMOS generation”, VLSI Technology Digest for Technical Paper, p.92, 2000

[43] P. S. Doyle, et.al., IEDM 1991,p.529

[44] C.-H. Ang, Jpn. J. Appl. Phys, vol.41, pL314, 2002

[45] Kato et al., RCJ Reliability Symposium 2002. (In Japanese)

[46] F. N. Sinnadurai, Microelectronics and Reliability, vol. 13, p.23, 1974.

[47] Toshiba semiconductor reliability handbook (1998)

[48] Internally agreed data in the Subcommittee on Semiconductor Device Reliability

[49] JEDEC Board Ballot JCB-99-86 “EARLY LIFE FAILURE RATE CALCULATION PROCEDURE FOR ELECTRONIC COMPONENTS”

[50] E. Takeda, “Hot-Carrier and Wear-out Phenomena in Submicron VLSI,S”, Proc. VLSI Technology Symp., p.2, 1985.

[51] C. Hu, “Hot-Electron-Induced MOSFET Degradation-Model, Monitor, and Improvement”, IEEE Trans Electron Devices, Vol.32 No.2, p.375, 1985.

[52] T.-C. Ong, et al., “P-MOSFET Gate Current and Device Degradation”, Proc. Of IRPS, p.178, 1989.

[53] D.-H. Huang, et al., “Characterization of Hot-Carrier-Induced Degradation in P-Channel MOSFETs by Total Injected Charge Techniques”, Proc. Of IRPS, p.34, 1994.

[54] F-C. Hsu, “Hot-Carrier Substrate-Current Generation During Switching Transients”, IEEE Trans. Electron Devices, Vo1.32 No.2, p.394, 1985.

[55] S. Aur, “Kinetics of Hot-Carrier Effects for Circuit Simulation”, Proc. Of IRPS, p.88, 1989.

[56] R. Rakkhit, et al., “A Comparison ofInverter-type Circuit Lifetime and Quasi-Static Analysis of NMOSFET lifetime”, Proc. Of IRPS, p.112, 1991.

[57] K. N. Quader, et al., “Simulation of CMOS Circuit Degradation Due to Hot-Carrier Effects”, Proc. Of IRPS, p.16, 1992.

[58] Edited by Matsushita Electronics Corporation, “Reliability engineering of the semiconductor devices” (In Japanese)

[59] G. LaRosa, et al., “NBTI-Channel Hot Carrier Effects in PMOSFETs in Advanced COMS Technologies”, Proc. Of IRPS, p.282, 1997

[60] Kato, et al., “Consideration for the prediction of MBTI degradation in PMOSFET”, 12th Reliability Symposium of the electronics devices”, Nov. 2002. (In Japanese)

[61] JEDEC/FSA JOINT PUBLICATION JP-001 “FOUNDRY PROCESS QUALIFICATION GUIDELINES (Wafer Fabrication Manufacturing Sites)”

[62] J. W. Mcpherson, “Underlying physics of the thermochemical E model in describing low-field time-dependent breakdown in SiO_2 thin films”, J. Appl. Phys., Vol.84, pp.1513-1523 (1998)

[63] D. J. DiMaria, “Impact ionization trap creation, degradation, silicon”, J. Appl. Phys., Vol.73, pp.3367-3384 (1993)

[64] D. J. DiMaria, “Defect generation under substrate-hole-electron injection into ultra-thin silicon dioxide layers”, J. Appl. Phys., Vol.86, pp.2100-2109 (1999)

[65] D. L. Crook, “Method of Determining Reliability Screens for Time Dependent Dielectric Breakdown”, Proc. Of IRPS, p.1, 1979.

[66] E. S. Anolick, “Low Field Time Dependent Dielectric Integrity”, Proc. Of IRPS, p.8, 1979.

[67] C. M. Osburn, “Accelerated Dielectric Breakdown of Silicon Dioxide Films”, J. Electrochemical Soc.,120, 1377, 1973.

[68] Yamabe, “Thin SiO_2 layer- Dielectric breakdown defects,” Semiconductor research, 22, Kogyo Chosakai Publishing, 75, 1985. (In Japanese)

[69] J. W. McPherson, et al., “Acceleration Factor for thin Gate Oxide Stressing”, Proc. Of IRPS, p.1, 1985.

[70] J. C. Lee, “Modeling and Characterization of Gate Oxide Reliability”, IEEE Trans. Electron Devices, 35, 2268, 1988.

[71] H. Mase, et al., 1999 Int. workshop on Dielectric Thin Films for Future ULSI Devices.

[72] P. E. Nicollian, et al., “Experimental Evidence for Voltage Driven Breakdown Models in Ultrathin Gate Oxides”, Proc. Of IRPS, p.7, 2000

[73] E. Wu, et al., “Nonlinear Characteristics of Weibull Breakdown Distributions and Its Impact on Reliability Projection for Ultra-Thin Oxides”, IEDM Tech. Dig., 18-2, 1999

[74] J. Lau, et.al., 2004 Electronic Components and Technology Conference, p.718-725.

Explanation

1. Purpose of this establishment

This document is intended to describe the reliability concept particular to the semiconductor devices, which is not fully mentioned in the general descriptive literature, and provide better understanding of the appropriate reliability designs and evaluations to the semiconductor applications.

Especially, with regard to the recent trend toward the excessively stressed conditions in reliability tests, much efficient tests can be implemented by understanding the acceleration factors of physics and designing appropriate stress conditions and test durations that meet the reliability targets in the field conditions.

This document was written by the reliability-engineering experts who have been working for reliability tasks for many years (as well as members of the Subcommittee on Semiconductor Devices Reliability of **JEITA**) and is worth naming a summary of the common sense in the field of the semiconductor industry.

This document provides the explanation from the aspects of both traditional and unique concepts, and is expected to be useful for the semiconductor users as well as reliability engineers in semiconductor suppliers.

The reliability engineering advances along with the development of the semiconductor process.

However, it is true that the fine processing technology is approaching the dead end of the solid state properties. We hope this document contributes to many engineers to breakthrough this problem.

2. History of the deliberation

“Group a” of the Subcommittee on Semiconductor Devices Reliability started the deliberation on Jun. 1998 for the aforesaid purpose. The deliberations had been held every other month with absorbing information from the academia and industrial trend. This document was finally summarized as a guideline on Jun. 1999 on the basis of the latest reliability technology trend.

In 2004, the Subcommittee started amending this document so that it reflects the latest fine processing trend, increasing complexity of the devices, and further streamlining the reliability test. The amendment is primarily intended to promote the integrated theory that describes the early failure period through the wear-out failure period by means of the Weibull distribution, and to extend the reliability concept to the newest processes. This document was created in cooperation with **JEDEC** members through joint meetings.

3. Contents of the deliberation

The appropriateness of the duration and the stress cycles of the reliability test were raised as the first controversial theme and its technical background was discussed. One of the examples is the test duration of 1000 h in high temperature operating life or in temperature humidity bias test. When the appropriate acceleration factor is obtained from the failure mechanism, test time can be often cut down significantly with respect to the "common sense." The extrapolation procedure of the lifetime is described in "**4.5 Acceleration factors and the efficient accelerated test methods**" in this document.

Another theme discussed here was the failure mechanisms of the accelerated stress test, which are classified into two groups, i.e., intrinsic mode and extrinsic mode.

The intrinsic failures are inherent wear-out due to the design rules of semiconductor devices, while the extrinsic failures are induced by the external causes, such as particles, that degrade the inherent reliability. Once we understand the both ends of the concept, we can design the reliability test more efficient in a short duration and accurate. This document describes the popular conventional "Weibull distribution and bathtub curve" and the procedures how to predict the EFR or how to determine the sampling size at wear-out failure period by making positive use of the Weibull distribution that can describe the reliability behavior of the semiconductor in the integrated manner. Also, the acceleration models and factors in the table were updated by reference to the information from the recent reports and papers that describe the latest fabrication process.

Finally the section "**4.5 Acceleration factors and the efficient accelerated test methods**" described the numerical examples for the convenience of semiconductor reliability engineers and some other people such as semiconductor users.

4. Deliberation members

This technical report was deliberated by the Subcommittee on Semiconductor Device Reliability, Semiconductor Reliability Group, TSC on Semiconductor Devices.

Participated members are:

<Technical Standardization Committee on Semiconductor Devices>

Chairman	Hisao Kasuga	NEC Electronics Corp.

<Semiconductor Devices Reliability Group>

Chairman	Kazutoshi Miyamoto	Renesas Technology Corp.

<Sub-Committee on Semiconductor Devices Reliability>

Chairman	Tetsuaki Wada	Matsushita Electric Industrial Co.,Ltd.
Vice Chairman	Junichi Mitsuhashi	Renesas Technology Corp
Member	Youichi Iga	NEC Electronics Corp.
	Yasuhito Anzai	Oki Electric Industry Co., Ltd.
	Osamu Nakayama	Kawasaki Microelectronics, Inc.
	Kazutoshi Kitazume	Sanyo Semiconductor Co., Ltd.
	Makoto Kanayama	Shindengen Electric Mfg. Co., Ltd.
	Shinichi Ikezoe	New Japan Radio Co., Ltd.
	Hiroyoshi Odaira	Seiko Epson Corp.
	Kenji Sasaki	Sony Corp.
	Tooru Ueda	Toshiba Corp.
	Yasuyuki Igarashi	IBM Japan, Ltd.
	Masayoshi Takani	Texas Instruments Japan Limited
	Toshiki Yamaguchi	Fujitsu Ltd.
	Naohiro Yasuda	Fuji Electric Holdings Co., Ltd.
	Masashi Kusuda	Mitsumi Electric Co., Ltd.
	Kohki Ohara	Ricoh Co., Ltd.
	Takahiro Ito	Rohm Co., Ltd.
Special Members	Takeshi Watanabe	NEC Electronics Corp
	Yasuhiro Fukuda	Oki Electric Industry Co., Ltd.
	Kouji Obinata	Sony Corp.