Chapter 1

Introduction

1-1 High dielectric constant gate oxides for SONOS flash memory

1-1.1 Scaling and gate capacitance

The complementary metal oxide semiconductor (CMOS) field effect transistors (FET) made from silicon are the most important and popular electronic devices. [1]. These devices possess the advantage of low power consumption. In addition, the device's performance has been improved over forty years according to the prediction of Moore's Scaling Law. This law notes that the number of devices on an integrated circuit increases exponentially, doubling over a 2–3 year period. The minimum feature size in a transistor decreases exponentially each year. The semiconductor Roadmap defines how each design parameter will scale in future years to continue this trend, as shown in Table 1 and Figure 1-1.

Year	2001	2003	2005	2007	2009	2012	2016	2018
Node	130	100	80	65	45	32	22	18
ASIC 1/2 pitch	150	107	80	65	45	32	25	18
Physical gate length	65	45	32	25	20	13	9	7
T_{ox} high power	1.5	1.3	1.1	0.9	0.8	0.6	0.5	0.5
T_{ox} low power		2.2	2.1	1.6	1.4	1.1	1.0	0.9
Gate oxide		Oxynitride			HfO_x; Si, N		$LaAlO_3$	
Gate metal		Poly Si			Metal gate, e.g. TaC_x, $TaSiN_x$			

Table 1. Summary of 2003 Roadmap. Node, gate length, EOT of high power (CPU) and low standby power devices (mobile), gate oxide material and gate electrode material. Earliest introduction of high K would be late in 45 nm node.

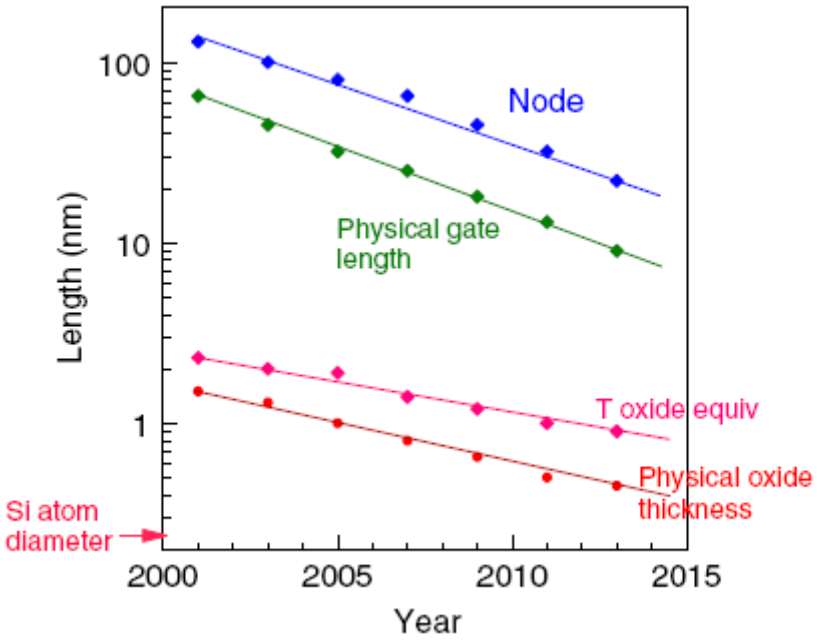

Fig. 1-1: The scaling of feature size, gate length and oxide thickness according to the 2003 semiconductor Roadmap.

The scaling cannot go on forever, and the limits to Moore's law are often ascribed to lithography the need for very short wavelengths of light to pattern the minimum feature size. It turns out that materials are now a key constraint. [2-3] First, the maximum current density in interconnects between transistors recently led to copper replacing aluminium as the conductor. Then, RC time delays led to the SiO_2 inter-metal dielectric being replaced by materials of lower dielectric constant such as SiO_2 : F or porous SiOCH alloys. But the most serious problem in logic circuits is now the FET 'gate stack', which is the gate electrode and the dielectric layer between the gate and the silicon channel.

The thickness of the SiO_2 layer presently used as the gate dielectric is now so thin (under 1.4 nm) that the gate leakage current due to direct tunnelling of electrons through the SiO_2 becomes too high, exceeding

$1Acm^{-2}$ at 1V (Figure 1-2), so that power dissipation increases to unacceptable values [4-10]. In addition it becomes increasingly difficult to fabricate and measure accurately such thin films. Finally, the reliability of SiO_2 films against electrical breakdown declines in thin films. These reasons lead to a desire to replace SiO_2 as a gate oxide.

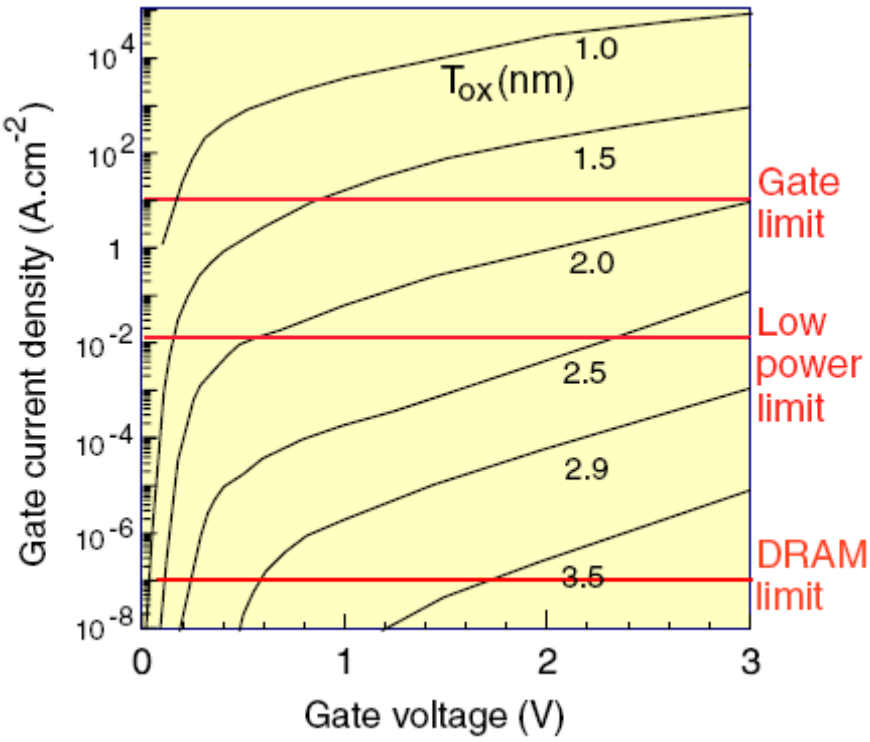

Fig. 1-2: Leakage current versus voltage for various thicknesses of SiO_2 layers.

Tunnelling currents decrease exponentially with increasing distance. An FET is a capacitance-operated device, where the source–drain current of the FET depends on the gate capacitance

$$C = \frac{\varepsilon_0 K A}{t}, \qquad (1)$$

where ε_o is the permittivity of free space, K is the relative permittivity, A is the area and t is the oxide thickness. Hence, the solution for the tunnelling problem is to replace SiO_2 with a physically thicker layer of new material of higher dielectric constant (permittivity) K (Figure 1-3). This high K material will keep the same capacitance but decrease the tunnelling current. These new gate oxides are called 'high K oxides'.

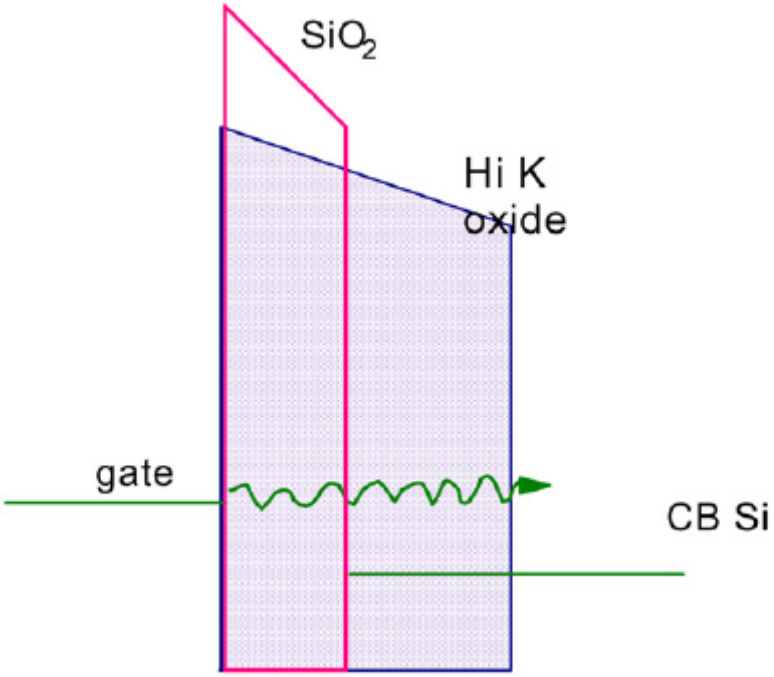

Fig. 1-3: Schematic of direct tunnelling through a SiO_2 layer and the more difficult tunneling through a thicker layer of high K oxide.

It is convenient to define an 'electrical thickness' of the new gate oxide in terms of its equivalent silicon dioxide thickness or 'equivalent oxide thickness' (EOT) as

$$t_{ox} = \mathrm{EOT} = \left(\frac{3.9}{K}\right) t_{hiK}. \qquad (2)$$

Here 3.9 is the static dielectric constant of SiO_2. The aim is to develop high K oxides which allow scaling law to continue to ever lower values of EOT.

The gate leakage problem has been addressed since the late 1990s [11]. However, the criteria for the choice of oxide were not clear. In 2001, the choice of oxide narrowed to HfO_2, but the integration of HfO_2 material into a successful electronic device is still challenge. It was debated whether high K oxides would be used, but instead the device engineers might use novel device designs to circumvent the problem. However, the increasing importance of the low-power sector of electronics, in mobile phones, lap-tops, etc meant that the problem must be confronted [4]. Low standby power CMOS requires a leakage current of below 1.5×10^{-2} A cm^{-2} rather than just 1 A cm^{-2}. There have been many difficulties in manufacturing high K oxide layers of sufficient quality but these have gradually been overcome. Announcements [12] indicated that the problems are now solved such that high K oxides could be implemented in the 45 nm node.

Four main problems have been identified for the successful introduction of high K oxide [13]:

(1) the ability to continue scaling to lower EOTs,

(2) the instabilities caused by the high defect densities,

(3) the loss of carrier mobility in the Si channel when using high K oxides,

(4) the shifts of the gate voltage threshold and the need for metal gates.

1-1.2 Evolution of Flash Memory

The complementary metal oxide semiconductor (CMOS) field effect transistor (FET) fabricated from silicon substrate is the most important electronic device. The CMOS technology can be divided into two categories as depicted in Fig. 1-4:

(i)The volatile memory: this type memory will lose the storage data if the power supply is off, like static random access memory (SRAM) and dynamic random access memory (DRAM).

(ii)The non-volatile memory: this type memory will keep the storage data even if the power supply is off, like electrically programmable read only memory (EPROM), electrically erasable programmable read only memory (EEPROM), and the flash memory.

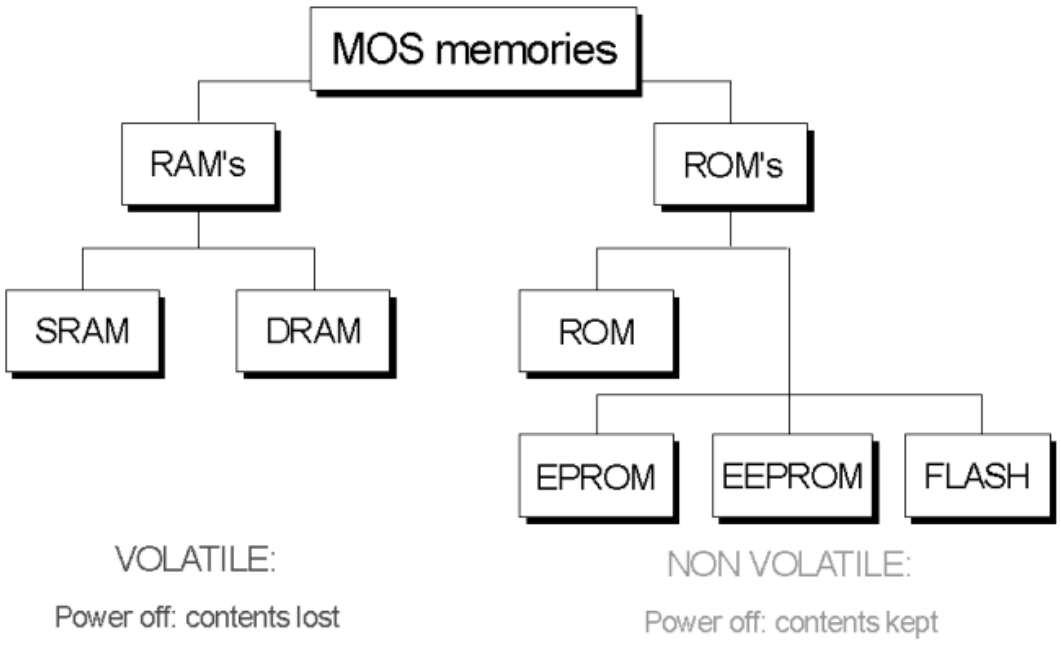

Fig. 1-4: The semiconductor memory tree.

The most explosive growth field of the semiconductor memory is the Flash memory. The advantages of Flash memory are that it can be electrically written more than 100K times with byte programming and sector erasing and with the smallest cell size (one transistor cell) [14-15]. The Flash memory cell is used floating gate (FG) structure, as illustrated in Fig. 1-5(a)

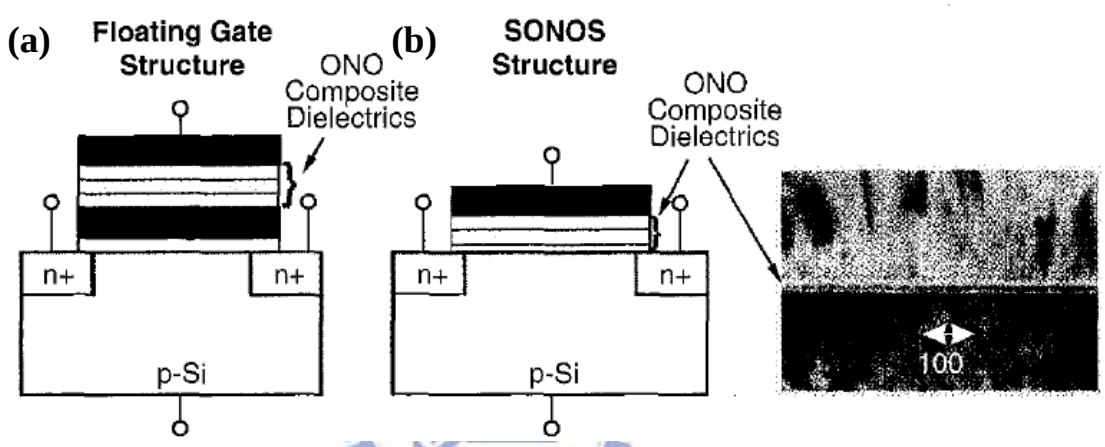

Fig .1-5 : A comparison of (a) floating-gate and (b) SONOS non-volatile semiconductor memory (NVSM) devices.

The first (FG) nonvolatile semiconductor memory was invented by S. M. Sze and D. Kahng in 1967 [16]. The conventional FG memory (in Fig. 1-5(a)) used polysilicon as a charge storage layer surrounded by the dielectric [14]. The FG structure can achieve high densities, good program/erase speed and good reliability for Flash memory application. However, the FG memory concerns the scaling issue [17]. When the tunneling oxide thickness is below 10nm, the storage charge in the FG is easy to leak due to a defect in the tunneling oxide formed by repeated write/erase cycles or direct tunneling current.

In order to solve the scaling issue of FG memory, the

polySi-Oxide-Nitride-Oxdie-Silicon (SONOS) memory has been proposed recently [17]. SONOS memory has better charge retention than floating gate memory when floating gate bitcell's tunneling oxide is below 10nm. Also, the SONOS memory possesses the spatially isolated deep-level traps. Hence, a single defect in the tunneling oxide will not cause the discharge of the memory cell [17]. The structure of SONOS memory is depicted in Fig. 1-5(b). The SONOS memory uses silicon nitride as a charge trapping layer, and the band diagram is depicted in Fig.1-6. The conduction band offset between tunneling oxide and nitride is 1.05eV. When we apply a positive voltage on the gate, the band will bend downward as illustrated in Fig. 1-6 [18]. The electrons in the Si-sub conduction band will tunnel through the tunneling oxide and a portion of nitride to be trapped in the charge trapping layer. Before electrons are trapped in the nitride, they must tunnel a portion of nitride and this will degrade the program speed. Besides this, the conduction band offset of nitride is only 1.05eV and the trapped electron back tunneling may also occur. To solve these problems, the high-k materials are the possible candidates to replace the traditional silicon nitride as the charge trapping layer.

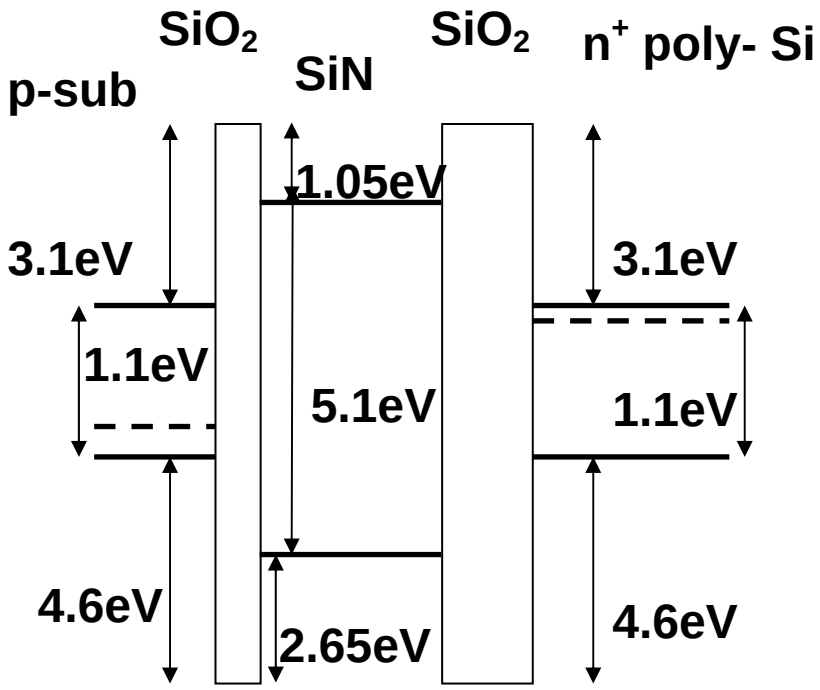

Fig. 1-6: The band diagram of nitride- based SONOS memory.

The advantages of high-k material are larger band offset with tunneling oxide and more trapping site than silicon nitride. For HfO_2 high-k material in Fig. 1-7, the conduction band offset between tunneling oxide and HfO_2 is 1.6eV. When programming, the electron will tunnel shorter distance in HfO_2 than in nitride to be trapped. This can achieve high program/erase speed. Thus, it is beneficial to use a high-k material as the charge trapping layer in a SONOS-type memory device, provided that there are many deep level trapping sites in the high-k material. The electron trap level of ZrO_2 is 1.0eV [19] and 1.5eV of JVD HfO_2 [20], and this is deeper than 0.8eV of nitride. It is desirable to choose a high-k material with large band offset with tunneling oxide and deep trapping level as charge trapping layer to achieve high program/erase speed and good reliability due to deep trapping level. High-k material has large

dielectric constant, a wide band gap, and high trap site density. Hence, it is also suitable for SONOS-type memory application.

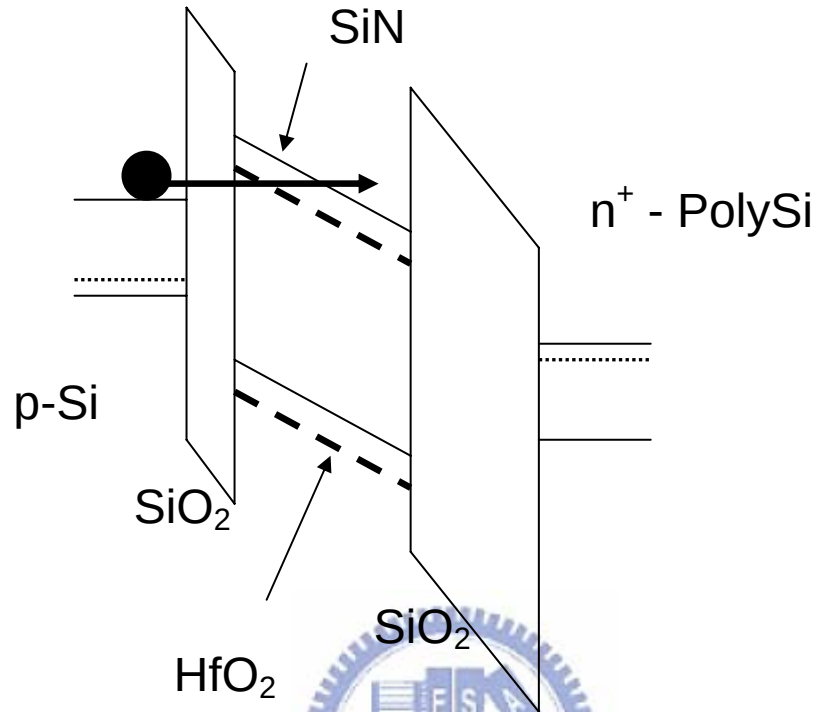

Fig. 1-7: The band diagram comparison of SONOS-type memory of nitride and HfO_2 charge trapping layer when programming (SiN: solid line, HfO_2: dash line).

Although SONOS-type memory is a promising device, the high-K memory faces the electron migration problem in the charge trapping layer [21]. This phenomenon will cause the charge loss and degrade the charge retention performance. The nanocrystal memory can store the trapped charge tightly to avoid the charge loss problem of SONOS memory and also achieve the advantages like fast program/erase speed, low programming voltage and good endurance as SONOS memory [22-23]

1-2 Motivation

1-2.1 The Deposition Method of High-k Material

The deposition method of high-k material must satisfy two requirements: first, to achieve good quality of deposited film for the applications, particularly with respect to the interface controllability. Secondly, the fabrication method is the compatible with the conventional CMOS processes. To date, the technologies applied on high-k film deposition includes physical vapor deposition (PVD), metal organic chemical vapor deposition (MOCVD), and atomic layer deposition (ALD).

The PVD process needs a high-k metal target for sputtering under oxygen ambient to form high-k oxide film. MOCVD is a material synthesis process using a variety of solid or gaseous precursors in which the precursors will thermally decompose into reactive species on the substrate surface and combine to form a thin film. In MOCVD process, a high substrate temperature is necessary to get better film quality and reduce impurity concentration. However, MOCVD process faces a challenge criteria: how to control the impurity like carbon in the film as low as possible [24]. ALD process can control film growth in a layer-by-layer formation at atomic scale. The growth rate of ALD is much less than one monolayer per cycle while in the case of reactant condensation or decomposition the growth rate is greater than one monolayer per cycle. In practice, the growth rate of ALD is less than one monolayer per cycle. But the main issue of ALD process is that: the very limited selections of available precursor sources induce chlorine contamination in the films. Besides this, in the special case of metal

electrode deposition, precursor sources for ALD process are still not available [25-26].

1-2.2 The Sol-Gel Spin Coating Method

A sol is a colloidal suspension of solid particles about 0.1~1 μm in a liquid phase [27]. A gel is a solid material network containing a liquid component [28]. The sol-gel spin coating process includes four steps in the following. Firstly, the desired colloidal particles dissolved in a solvent to form a sol. Secondly, the deposition of sol solution produces the coatings on the substrates by spraying, dipping or spinning. Thirdly, the particles in sol are polymerized through the removal of the stabilizing components and produce a gel in a state of a continuous network. Finally, the final heat treatments pyrolyze the remaining organic or inorganic components and form an amorphous or crystalline coating [29-32].

Sol-gel method has been applied to the fabrication of the organic and inorganic hybrid materials for specific applications. Liquid phase processing enables the molecular scale mixing of precursors, leading to homogeneous, multi-component materials. The most interesting feature of sol-gel processing is its capability to synthesize a new type of materials called inorganic-organic hybrids. In addition, metal oxides with various shapes, such as thin films, porous structures, and particles, can also be formed by sol-gel method, thus increasing the applicability to many specific usages [33]-[35].

Sol-gel spin coating method is used more and more widely in the creation of ceramic fibers, thin films, and aerogel, because it allows the

fabrication of very homogeneous and very thin fibers and films. Figure 1-8 illustrates the process and products of the sol-gel method [36]. These sol-gel ceramic fibers are mostly used in the optical industry as fiber optic cores. For the sol-gel method applied on the thin film deposition, dense film can be made by coating a substrate material with the sol and letting it gel. This leaves a very dense film on the substrate which can have a number of the uses such as catalysts, molecular sieves, chemical sensing, optical devices, and nanoelectronic devices [37]. Aerogels are a class of ceramic materials fabricated from a sol-gel by carefully evacuating the solvent to leave a fragile polymer or oxide network which is 90~99% air by volume. Silica aerogels have interesting applications, among them as thermal insulation materials [38].

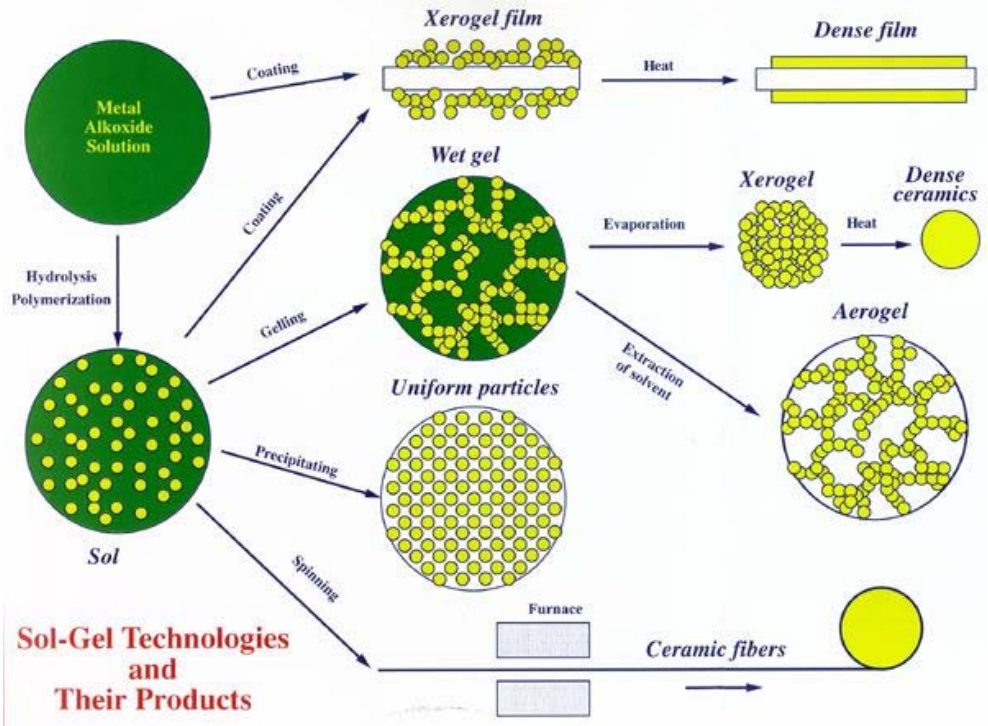

Fig. 1-8: Three main applications of sol-gel method.

1-2.3 Purpose of This Study

The sol-gel spin coating method using metal halides hydrolyzed in organic or colloidal solvents to form precursor compound and undergo hydrolysis, condensation, and polymerization steps to form metal oxide networks. The advantages of using sol-gel method to fabricate high-k film are its cheaper precursor and low cost tool than ALD, PVD, and MOCVD, and its ability to synthesize various types of thin films. However, no reports exist that describe the fabrication of SONOS type memory with sol-gel spin coating method.

In this thesis, the high-k charge trapping layer of SONOS-type memory deposited by sol-gel spin coating method is proposed. We fabricated two SONOS-type memories with different high-k charge trapping layer by using different precursors. We use sol-gel method to combine two or three high-k precursors, i.e. $HfCl_4$, $ZrCl_4$ and $SiCl_4$ together, to form hafnium silicate nanocrystal memory and co-existed hafnium silicate and zirconium silicate in memory devices. As mentioned above, the great advantage of the sol-gel spin coating method is its capability to synthesize new types of nanocrystal materials. We mixed these precursors to fabricate the new types high-k materials. After sol-gel spin coating, we used high-k rapid thermal annealing (RTA) at 900^oC 1min in O_2 ambient to form high-k dielectric film. The device performance like Id-Vg, data retention, endurance, and program/erase speed is obtained to check the quality of the high-k charge trapping layer deposited by sol-gel method. The better experimental results suggest the sol-gel spin coating can be applied to the high-k dielectrics deposition

.

1-3 Thesis Organization

This thesis includes four chapters. We evaluate the device performance of the next generation SONOS-type memory using high-k dielectrics as charge trapping layer deposited by novel sol-gel spin coating method.

In Chapter 1, we introduce the evolution of the Flash memory and explain why SONOS-type nanocrystal memory with high-k charge trapping sites is aimed to replace the traditional floating gate memory. The sol-gel spin coating technique and the purpose of this thesis are also mentioned in this chapter.

In Chapter 2, we develop sol-gel method to combine two high-k precursors of HfO_2 and SiO_2, and deposit thin film consist of two high-k material as a charge trapping layer for SONOS-type memory. X-ray photoelectron spectrometer (XPS) and transmission electron microscopy (TEM) are measured to obtain the physical characteristics of the binary high-k thin film. Furthermore, the electrical characteristics for the memory devices are also evaluated.

In Chapter 3, we apply our sol-gel spin coating to fabricate the SONOS-type memory with co-existed hafnium silicate and zirconium silicate nanocrystals. From the TEM identification, the nanocrystals are formed as the charge trapping layer after 900^{o}C 1 min RTA. We verify the chemical bonding of nanocrystal from the XPS analysis. We also evaluate the electric properties in terms of program/erase speed, charge retention,

and endurance measurement.

At the end of this thesis, the conclusion is given in the final Chapter .