Deposition of ALD-Molybdenum for Flash Memory Wordline Metallization

Joshua Collins
Lam Research Corporation
ALD-CVD Metals
Fremont, CA USA
joshua.collins@lamresearch.com

Sanjay Gopinath
Lam Research Corporation
ALD-CVD Metals
Fremont, CA USA
sanjay.gopinath@lamresearch.com

Kaihan Ashtiani
Lam Research Corporation
ALD-CVD Metals
Fremont, CA USA
kaihan.ashtiani@lamresearch.com

Shruti Thombare
Lam Research Corporation
ALD-CVD Metals
Fremont, CA USA
shruti.thombare@lamresearch.com

Xiaolan Ba
Lam Research Corporation
ALD-CVD Metals
Fremont, CA USA
xiaolan.ba@lamresearch.com

Griffin Kennedy
Lam Research Corporation
ALD-CVD Metals
Fremont, CA USA
griffin.kennedy@lamresearch.com

Juwen Gao
Lam Research Corporation
ALD-CVD Metals
Fremont, CA USA
juwen.gao@lamresearch.com

2025 IEEE International Memory Workshop (IMW) | 979-8-3503-6298-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/IMW61990.2025.11026947

***Abstract*— Multiple metallization applications in both memory and logic are transitioning to fluorine-free, barrierless ALD molybdenum. For the NAND Flash wordline application we have developed a proprietary ALD-Molybdenum (ALD-Mo) process that nucleates well on dielectrics (Al2O3), enabling low wordline resistance in geometrically constrained NAND Flash memory arrays.**

In this paper we demonstrate how chemistry and transport modeling can be used to guide NAND Flash wordline process optimization and to anticipate future requirements as NAND Flash devices transition from 300 to 1000 wordlines.

Keywords—ALD-Mo, chemistry modeling, transport modeling, NAND Flash wordline metallization

I. Low Resistance Metallization Choices and Ranking

Table (1) ranks various metallization choices against the key suitability metrics for NAND Flash wordline metallization. The industry standard for wordline metallization is tungsten (W). When integrated with a conformal titanium nitride (TiN) barrier, ALD-W has outstanding fill capability, near-zero electromigration (high melting point), well developed wet and dry etch processes, and relatively low cost. But the resistivity of the TiN-W metallization stack is rapidly becoming a bottleneck.

The product of (electron mean free path λ) and (metal resistivity ρ) scales well with line resistance, and by this metric all the elements listed in table (1) have the potential to produce lines with lower resistance than TiN-W [1], but we need to consider more than resistance when selecting a replacement metal for NAND Flash wordlines. Elements without suitable wet or dry etch processes for node separation must be eliminated from consideration (Os, Ir). Metals with low melting points or outrageously high cost must likewise be excluded (Cu, Al, Rh). Nickel (Ni), Cobalt (Co), Ruthenium (Ru), and Molybdenum (Mo) are all potentially attractive relative to W. But what really sets Mo above the others is our ability to deposit it directly on dielectrics using ALD. This eliminates the need for a high-resistivity Liner or Barrier layer, such as TiN, freeing up valuable volume for more low resistance metallic conductor. Coupled with its high melting point, low cost, and well developed wet and dry etch processes, molybdenum is the clear choice for wordline resistance scaling.

TABLE I. Ran king of Possible Metallization Choices

Metal	λ x ρ [10^{-10} Ω cm^{-2}]	Melting point [C]	Relative cost/g [vs W]	Wet or Dry etch	Need for barrier
Mo	5.99	2623	0.8	●	●
W	8.20	3422	1.0	●	●
Ru	5.14 / 3.81	2334	26.0	●	●
Ni	4.07	1455	1.6	●	●
Co	4.82	1495	4.0	●	●
Rh	3.23	1963	*520.0*	●	●
Cu	6.70	*1085*	0.12	●	●
Al	5.01	*660*	0.04	●	●
Os	6.41 / 4.33	3045	*260.0*	●	●
Ir	3.69	2447	*440.0*	●	●

II. ALD-Mo Precursor Selection

To avoid substrate damage (Al2O3, SiN), we focused on the chlorides and oxi-chlorides of molybdenum, which are listed in Table (2). MoCl2, MoCl3, and MoCl4 were eliminated from consideration due to their negligible vapor pressures. For the more volatile species another important factor is their ability to switch from Mo deposition to Mo etch depending on precursor concentration. The self-etching behavior of MoOxCly precursors has been observed to roughly correlated with their Cl/Mo ratios. Self-etching in NAND Flash wordlines can lead to upper wordlines being etched away completely (high concentration), normal deposition in middle wordlines (ideal concentration), and precursor depletion resulting in no deposition or etch at the bottom (low concentration), which is clearly disqualifying. This makes MoO2Cl2 the ALD-Mo precursor choice for the NAND Flash wordline application.

TABLE II. ALD-MO PRECURSOR RANKING

Name	Mo Ox State	Pvap	Dep / Etch
MoO2Cl2	*+6*	*High*	*Cl/Mo = 2 → no etch at high conc*
MoOCl4	*+6*	*Low*	*Cl/Mo = 4 → weak etch at high conc*
MoCl5	*+5*	*Low*	*Cl/Mo = 5 → strong etch at low conc*
MoCl4	+4	negligible	• Unsuitable as ALD precursors due to negligible vapor pressure • Possible intermediates in ALD-Mo reactions
MoCl3	+3	negligible	
MoCl2	+2	negligible	

III. OVERCOMING NUCLEATION DELAY AND METAL AGGLOMERATION

A. *Creating a strong dielectric – metal interface*

It is difficult to grow metal films directly on dielectrics because Metal-Metal attraction is typically much stronger than Metal-Dielectric attraction (Fig 1). This can result in long nucleation delays, poor adhesion, and metal agglomeration. Traditionally these problems have been overcome by introducing a metal nitride between the dielectric and the metal (e.g. TiN-W). Metal nitrides create a bridging layer that nucleate well on dielectrics (dielectric-nitride attraction is comparable to nitride-metal attraction) and create strong Nitride-Metal interfaces (Nitride-Metal attraction is comparable to Metal-Metal attraction).

The problem with this approach is that thin Metal-Nitride layers have high resistivities (10^2 *to* 10^3 μΩ-cm) and consume volume that should be filled with low resistivity metals (≈ 10^1 μΩ-cm). 300-pair Flash wordline pillar-pillar opening widths are on the order of 10-20 nm today and shrinking fast. In a typical 300-pair Flash memory array, a 3 nm interfacial liner would leave just 30% of the available wordline volume for low resistivity metal, and increase wordline resistance by over 5-times (Fig. 3, Fig. 4, geometric parameters a, d defined in Fig 2). Even a relatively modest 1nm interfacial liner would have only 75% of its available wordline volume filled with metal and would increase wordline resistance by 50% relative to all-metal fill. Clearly, it is essential to eliminate interfacial liner volume to enable Flash wordline scaling.

Fig. 1. Dielectric-Metal Interface

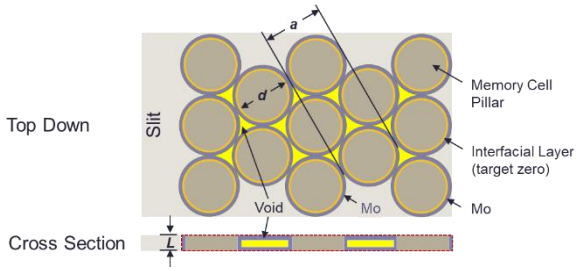

Fig. 2. Geometry of a Flash memory array

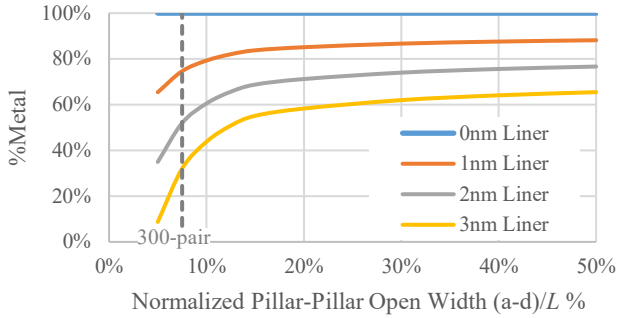

Fig. 3. Fraction of NAND Flash wordline volume occupied by metal (excludes interfacial liner volume)

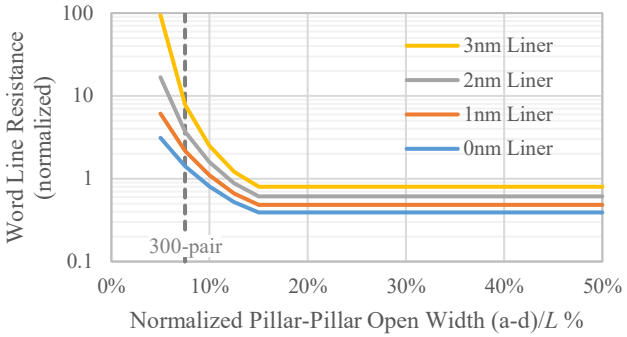

Fig. 4. Normalized NAND Flash Wordline Resistance

B. *Depositing pure ALD-Mo on dielectrics*

Lam ALTUS® Halo ALD-Mo deposition systems utilize deposition processes that create low-resistivity ALD-Mo films without a high-resistance layer at the metal-dielectric interface (Fig 5). This results in ALD-Mo films with zero nucleation delay on dielectrics and low resistivity (Fig. 6, 7). Adhesion and agglomeration resistance are also excellent.

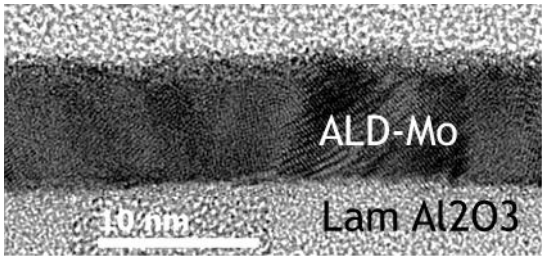

Fig. 5. ALD-Mo on Al2O3 with no interfacial layer

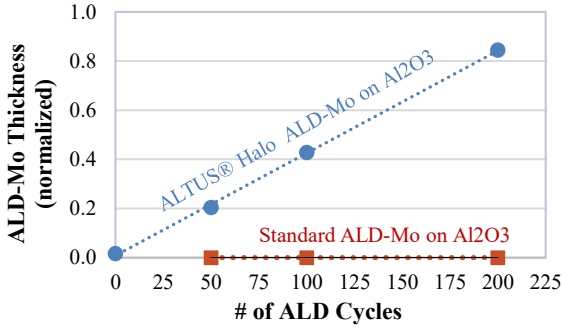

Fig. 6. ALTUS® Halo ALD-Mo achieves zero nucleation delay on Al2O3. Standard ALD-Mo is unable to nucleate directly on dielectrics.

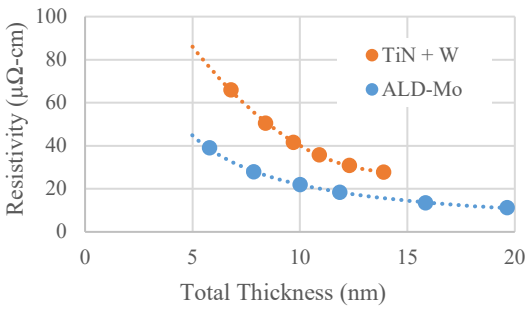

Fig. 7. ALTUS® Halo ALD-Mo resistivity is significantly lower than the resistivity of typical TiN-W film stacks

IV. Optimizing ALD-Mo for Wordline Step Coverage

A. Emperical Chemical Kinetics Modeling

The overall ALD-Mo reaction,

$$MoO_2Cl_2 + 3H_2 \rightarrow Mo + 2H_2O + 2HCl \quad (1)$$

Can be broken down into the following major steps:

1. Dissociative adsorption of MoO2Cl2 on available Mo sites creates adsorbed Cl and surface MoOx – Not Rate Limiting.
2. Dissociative adsorption of H2 on available metallic Mo sites - Not Rate limiting.
3. Reaction of adsorbed H* with surface MoOx – Unlikely Rate Limiting.
4. Desorption of reaction byproducts (H2O, HCl) creates new metal sites for H2, MoO2Cl2 adsorption – Likely rate limiting.

From kinetic theory, the overall precursor mass flux rate to the substrate ($j_i^{"}$) can be expressed as [2]

$$j_i^{"} = \frac{1}{4} \cdot \omega_i \cdot \rho \cdot \bar{c}_i \cdot \boldsymbol{\varepsilon_i} \quad (2)$$

where $\boldsymbol{\varepsilon_i}$, the precursor sticking coefficient, is also closely related to the step coverage of thin film deposition (ω_i= mass fraction, ρ = density, $\bar{c}_i$ = mean thermal speed). The precursor mass flux ($j_i^{"}$) can also be expressed in terms of a chemical rate expression

$$j_i^{"} \approx k_1 \cdot (\rho \cdot \omega_{Mo} \cdot \Delta t_{Mo\ dose})^a \cdot (\rho \cdot \omega_{H_2} \cdot \Delta t_{H2\ dose})^b \cdot exp\left(\frac{-E_a}{RT}\right) \quad (3)$$

where k_1 is a rate constant, Δt_i are dose times, and E_a is an overall reaction activation energy [adapted from 3]. By combining equations (2) and (3) we can solve for the sticking coefficient governing precursor deposition $\boldsymbol{\varepsilon_i}$, which is a necessary input to ALD-Mo step coverage models.

B. Transport Modeling

The geometry of NAND Flash wordlines is extremely complicated (Fig. 2, Fig 8), but with appropriate simplifications we can model transient wordline vapor transport and extract process trends.

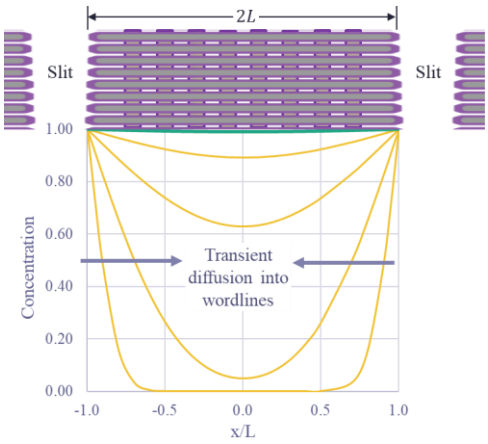

Fig. 8. NAND Flash wordlines in the cell array. Precursor molecules diffuse laterally into the wordline array from vertical slits on either side of the array at the start of a precursor dose.

We use the diffusion equation (4) to model transient precursor mass fractions inside a wordline (ω_{Mo}) during ALD [4]

$$\frac{\partial \omega_{Mo}}{\partial t} = D_{eff} \cdot \frac{\partial^2 \omega_{Mo}}{\partial x^2} \quad (4)$$

The effective diffusivity coefficient D_{eff} in these equations is the free-molecular diffusivity of a precursor molecule in a porous media comprised of a regular array of cylinders in crossflow and ***2L*** is the slit-to slit spacing of the array [5]. With appropriate scaling and boundary conditions, the solution of equation (4) yields useful estimates of transient precursor concentration profiles inside NAND Flash memory arrays, as illustrated in Fig. 8.

Wordline saturation time increases as ALD-Mo deposition progresses towards wordline seam closure as shown in Fig. 9. However, by adjusting deposition conditions from the beginning of the process, when the wordlines are wide open, to seam closure at the end, we can maintain complete saturation with reasonable dose times.

Future NAND Flash scaling to structures with more and more wordline pairs will also necessitate process scaling, as shown in Fig. 10. We do not, of course, know exactly what future 1000-pair NAND Flash will look like, but even rough scaling enables us to anticipate and prepare for future device requirements.

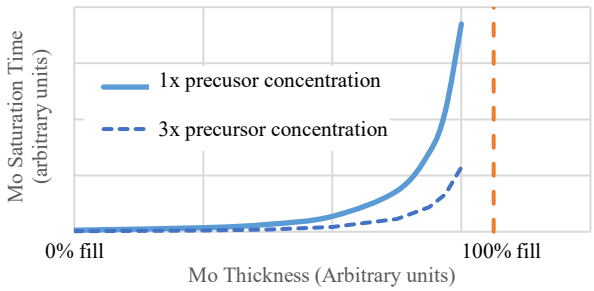

Fig. 9. ALD-Mo precursor saturation time scaling vs deposited metal thickness in the wordline (vertical dashed line indicates thickness at which wordlines close).

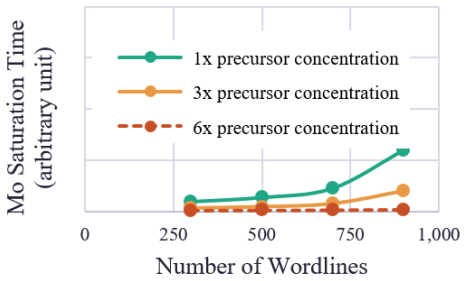

Fig. 10. ALD-Mo precursor saturation time at start of deposition relative to the # of wordlines (fixed geometry,)

V. Optimizing Metal Deposition Equipment for ALD-Mo

As one can see from in-feature modeling, vapor transport changes dramatically during the ALD-Mo fill process. Therefore ALD systems must enable recipe optimization at the beginning, middle, and end of the process. Lam ALTUS® Halo ALD-Mo systems utilize a 4-station chamber architecture, which allows users to optimize wafer temperature, gas timing, and gas flows from station to station. This process adaptability is likely to become more important as Flash memory scales beyond 300 wordlines.

VI. Conclusions

ALD-TiN + ALD-W wordline metallization has been one of the key process modules enabling NAND Flash memory. But as Flash memory arrays shrink, the wordline volume consumed by TiN is rapidly becoming an obstacle to further scaling. We can now produce conformal, low resistivity ALD-Mo directly on dielectric substrates (no TiN). This change is one of the key drivers for the current transition from TiN-W to ALD-Mo.

Achieving good metal fill inside NAND Flash wordlines is another ongoing challenge. At Lam we guide our process development with chemistry and transport modeling to help speed our solutions to market. Empirical chemical kinetics modeling provides a framework to explain observed process. Wordline transport modeling helps us estimate ALD dose and purge requirements and points us towards promising process windows.

Acknowledgment

The authors wish to thank our ALTUS® Halo customers for their tremendous efforts as we work together to validate and improve ALD-Mo performance.

We also wish to thank the Lam Semiverse™ and Lam Light modeling teams for their outstanding system-level and device-level modeling.

References

[1] Diane Gall, “Electron mean free path in elemental metals,” J. Appl. Phys. 28 February 2016; 119 (8): 085101, February 2016.

[2] J.W. Moore and R.G. Pearson, “Kinetics and Mechanism”, 3rd ed, John Wiley and Sons, 1981

[3] E.J. McInerney, D.C. Smith, and G. Ramanath, “Kinetic Rate Expression for Tungsten CVD in Different WF6 flow Regimes from Step Coverage Measurements”, Z. Metallkd. 91-7, pp 573-580, 2000

[4] J. Crank. (1975). “The Mathematics of Diffusion” (2nd Edition). Oxford University Press, Oxford, 1975

[5] Bernhard Tjaden, Samuel J Cooper, Daniel JL Brett, Denis Kramer and Paul R Shearing, “On the origin and application of the Bruggeman correlation for analysing transport phenomena in electrochemical systems”, Current Opinion in Chemical Engineering, Edited by W S Winston Ho and Kang Li, 12:44–51, March 2016,