

DDR4: Bereiten Sie sich vor, von Ihrem Lieferanten höhere Preise zu bekommen

Abstrakt

Da die Halbleiterfertigung zunehmend in wirtschaftliche und technische Herausforderungen hineinläuft, werden die Märkte und Anwendungen kleiner, welche es erlauben, die Kosten zu amortisieren und die Risiken der neuer Fertigungsprozesse zu übernehmen. Irgendwann wird die gesamte Branche in einen Legacy- und Spezialmarkt verzweigen - eine, die in der Lage ist, die Kluft zwischen neuen Speichergeräte und Architekturen zu überwinden. DDR4 kann jetzt wie Commodity aussehen, aber sobald die restlichen drei DRAM-Technologie-Entwickler sich auf die nächste große Sache zu bewegen, wird keine Lizenznehmer oder Foundry mehr in der Lage sein, moderne DDR4 in einem wettbewerbsorientierten Markt zu unterstützen. Sehe Sie sich Legacy-DRAM von Tier-2 und fab-less DRAM-Hersteller an. Die Komponenten werden verkauft zu einem Preis 2x bis 5x ihres historischen Tiefs, und dann betrachten ein spezialisierte DRAM's wie RDRAM; diese werden verkauft für einen 5x höheren Preis. DRAM-Hersteller haben lange auf eine ausreichende Marktkonsolidierung gewartet, um endlich ihre Gewinnspannen auf die der CPU-Hersteller-Ebene steigern zu können. Diese Zeit wird früher kommen, als Sie denken.

Einführung

Industrielle Anwendungen sind so ausgelegt, dass sie eine spezifische Funktion in zuverlässiger Art und Weise und in rauen Umgebungen durchführen; die Leistungsanforderung für die Lebensdauer des Produkts ist festgelegt. Es ist wahr, dass ein neueres Ersatzsystem für eine RT-Anwendung eine höhere Leistung haben kann, aber für eine, die im Feld ist, ist dies nicht der Fall. Die meisten embedded Anwendungen profitieren nicht von der schnellen Weiterentwicklung der Systemleistung, indem sie sich auf die neuesten Prozessoren und Speichertechnologien hin bewegen. Sie liefern weiterhin Legacy-CPU's und Controller von etablierten Lieferanten und erfordern langfristige Unterstützung für die Peripherie-Komponenten wie z.B. Speicher.

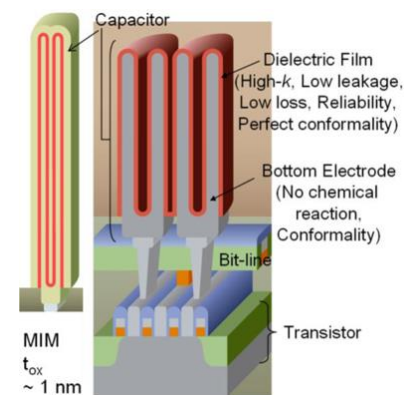
Das Benutzen für diese Anwendungen von fortgeschrittene Halbleiterherstellungsprozessen und schnelleren Gerätekonstruktionen könnte sich als problematisch erweisen. Die neueren Geräte können anfälliger für geringfügige Störungen aus Rauschen und ähnlichen Störungen von Hochenergiepartikeln oder kosmischen Strahlung aufgrund verbesserter Spannungsregelung, niedrigerer Core Spannung und schnelleren Schaltflanken sein. Einige der beobachteten Ausfallmechanismen können in neuen Speichercontrollern oder Implementierungen entworfen oder bearbeitet werden, aber Systeme, die bereits im Feld sind, haben diese Option nicht. Speicher, die in Anwendungen verwendet wurden, die vor Jahren mit jetzt ausentwickelten Steuerungen entworfen wurden, benötigen einen Speicher, der die Zuverlässigkeit und die Qualitätserwartungen des ursprünglichen Designs erfüllt.

Die großen, hochvolumigen Speicheranbieter werden weiterhin die Legacy-Funktionalität unterstützen, indem sie Komponenten auf eine neuere (kleiner) Prozesstechnologie übertragen, aber nur dann, wenn sie ökonomisch sinnvoll ist. Sie werden versprechen, diese Legacy-Funktionalität auf unbestimmte Zeit zu unterstützen, bis der Tag, kommt, an dem das Ende des Fertigungsprozesses das Ende für ein Produkt bedeutet. Dieses ist dann die letzte Möglichkeit (last time buy), das Produkt noch zu kaufen. Die meisten scharfsinnigen Kunden erkennen die Gefahr, in die sie sich einbringen, und werden bereits im Vorfeld auf alternative langfristige Lieferanten umstellen, so dass Sie auch günstigere Konditionen für eine erweiterte Belieferung vereinbaren können. Zusätzlich zu den offensichtlichen Geschäftsvorteilen gibt es technische Gründe, um Bauteile mit kleineren Prozessgeometrien zu vermeiden.

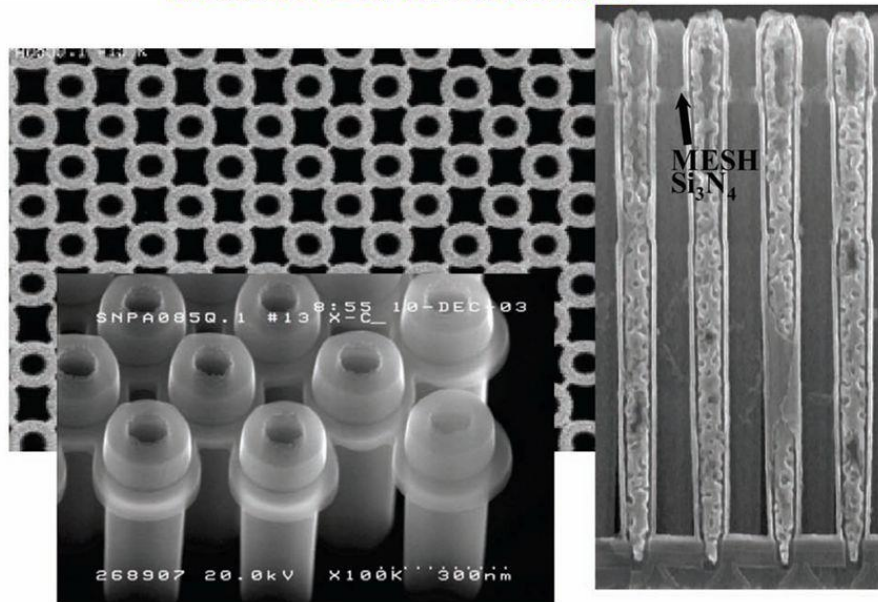
DRAM Design und Fertigung

Ein DRAM speichert Daten auf einem Kondensator, der durch einen 'Pass' Transistor isoliert ist. Diese Kombination wird als die Bitzelle bezeichnet, wobei jede mit x- und y-Dimensionen von einigen zehn Nanometern hat. Der Kondensator wird im Raum oberhalb der Schaltkreise gefertigt, und in modernen DRAM-Bitzellen ähnelt der Kondensator einem großen, dünnwandigen Zylinder, wobei das Dielektrikum sowohl auf der Innen- als auch auf der Außenseite des Rohres liegt. Wenn die Prozessgeometrie schrumpft, wird es einfacher als je zuvor, den Wert auf einer Bitzelle zu stören, entweder durch Störung des Kondensators, des Durchgangstransistors oder der Bitleitung, die mit der Erfassungsschaltung verbunden ist. Der Sense-Betrieb misst Millivolt von Differenzspannungen, die durch einen kleinen Kondensator verursacht werden, der eine große Bitleitung antreibt. Die Bitleitung ist mit Hunderten von Durchgangstransistoren verbunden, und die parasitäre Kapazität der Bitleitung selbst kann mehr als 10x größer sein als der Bitzellenkondensator.

Das Diagramm nach rechts entspricht der Komplexität des DRAM-Kondensator-Arrays nicht. Die Kondensatoren sind so groß, dass sie mit einem Netz mechanisch stabilisiert werden müssen, so dass der Abstand zwischen ihnen nicht variiert, sonst kann die Ladung eines Kondensators eine benachbarte Ladung beeinflussen. Unten ist ein Bild von 2004, wo die Kondensatoren im Vergleich zur aktuellen Gerätetechnik kurz und stumm sind und ein Siliziumnitrid-Mesh enthält, das die Position des Kondensators bei der Geräteherstellung stabilisiert. Ohne das Netz würden sich die Kondensatoren neigen und auf benachbarte Zellen übersprechen.

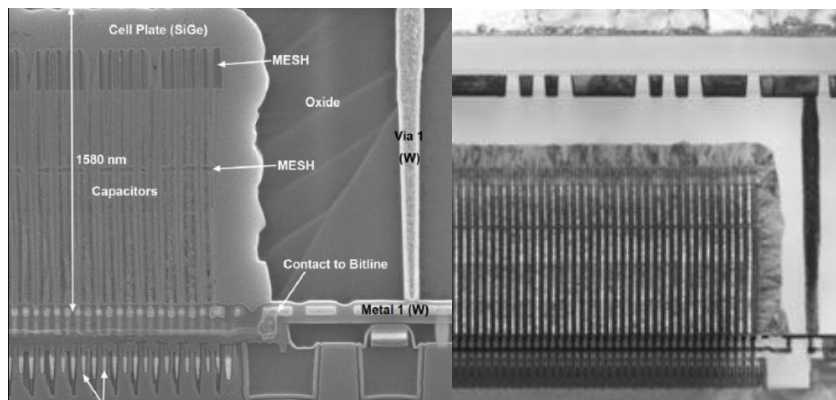


Modern DRAM Structure



[Samsung, sub-70nm DRAM, 2004]

Das Bild unten zeigt ein aktuelleres DRAM. Beachten Sie das viel höhere Seitenverhältnis der Kondensatoren und die beiden Ebenen der Maschen, die für die mechanische Stabilität erforderlich sind. Das Seitenverhältnis einer einzelnen Zelle liegt im Bereich von 30:1 für den Chip auf der linken Seite, und es wird nur schwieriger dadurch.



3D-NAND-Fertigung vs. "3D" DRAM

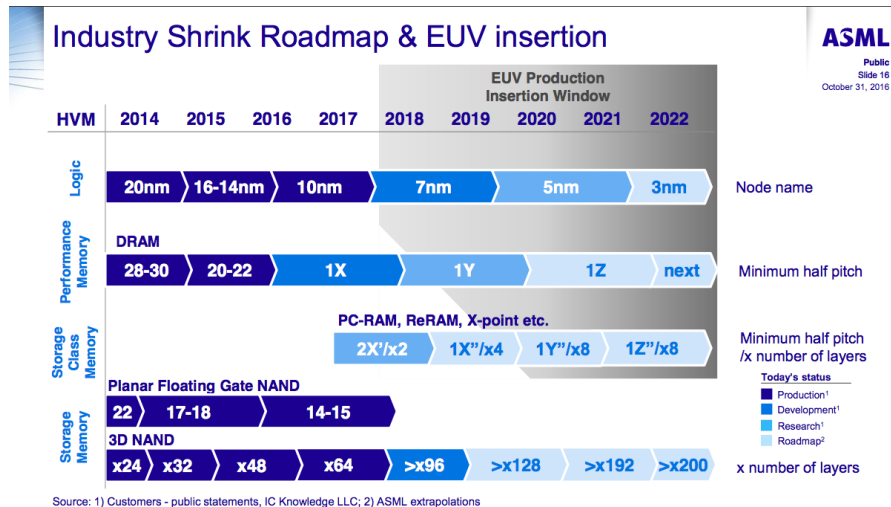
Die NAND-Technologie hat sich von der planaren Konstruktion abgewendet und verwendet stattdessen Stränge von Bitzellen, die vertikal senkrecht zur Oberfläche des Siliziumwafers ausgebildet sind, anstatt die Strings entlang der Oberfläche zu legen. Der Grund für diese drastische Veränderung ist, dass 3D-NAND kostengünstiger ist als sich weiterhin auf fortgeschrittene Lithographie zu verlassen. 3D-NAND addiert Kosten linear basierend auf der Anzahl der Schichten, die im Stapel benötigt werden, während die Fortsetzung des Schrumpfens der Lithographie die Kosten exponentiell erhöht. Die Grundarchitektur eines NAND-Arrays eignet sich für die 3D-Integration, bei der jeder Strang von NAND-Bitzellen in der 3. Dimension in Schleifen gesetzt wird. Dies ist effizient, da jede Bitleitungsschaltung die gleiche Länge hat.

Zum Glück für NAND war es möglich, auf 3D zu gehen und die Kosten der Extreme Ultra Violet (EUV) Lithographie zu vermeiden. Im Vergleich dazu hat DRAM seit über 20 Jahren Stapel- oder 3D-Kondensatoren eingesetzt und über diesen Zeitraum wurden die Kondensatoren immer größer mit größeren Oberflächen. Es ist nicht möglich, dass DRAM auf 3D geht, weil es schon dort ist. 3D-NAND ist wie eine einzelne Seite von Papier und verbindet sie in ein 32- oder 64-seitiges Buch mit Drähten, die jeden Briefort auf allen Seiten durch das Papier verbinden. DRAM zu nehmen und es in 3D umzuwandeln, wäre wie ein Buch zu nehmen und sie 32 oder 64 hoch zu stapeln und dann jede Briefposition auf nur einer Seite von jedem Buch zu verkabeln, ohne irgendwelche der anderen Seiten zu stören.

Kann DRAM in eine 3D-Architektur wie NAND umgewandelt werden? **Keine Chance.**

Wie kann DRAM vorankommen?

Genau wie bei einem Logikgerät ist es möglich, die Schaltkreise zu skalieren, aber die große technische Herausforderung mit DRAM beinhaltet die Notwendigkeit, die Kondensatoren zu verkleinern, ohne ihre elektrische Leistung weiter zu verschlechtern. Eine weitere Herausforderung, die die Logik nicht zu leisten vermag, ist die Notwendigkeit, dass der DRAM die Schaltkreise unterhalb der Kondensatoren mit den Metallschichten verbindet und über die Kondensatoren läuft. Es ist nicht die schwierigste Herausforderung in der DRAM-Fertigung, sondern ein weiterer signifikanter Unterschied zur Logik oder NAND-Fertigung.



Es besteht eine erhebliche wirtschaftliche Herausforderung, auch die DRAM-Fertigung voranzutreiben. DRAM (und NAND) sind Multi-Sourcing-Geräte mit Rohstoff-ähnlichen Preisen in Primär- und Sekundärmärkten. Dies hat dazu geführt, dass sich DRAM in eine hochverfügbare Komponente verwandelt hat, das in der einstelligen Preisspanne (\$0 bis \$9) verkauft wird, während Logik- oder CPU-Geräte ähnlicher Chipfläche im mehrstelligen Preisbereich (\$10 bis \$100) verkauft werden. Wirklich große Dice lassen sich verkaufen für \$ 1000, und in einigen Fällen höher noch im Falle von High-End-FPGA-Geräte. In Anbetracht der Bruttomargen der fraglichen Produkte können Rohstoffspeicher nicht die hohen Kosten der fortgeschrittenen Verarbeitung, ob EUV oder fortgeschrittene Multi-Expositions-Musterung, oder den Schrumpfungspfad fortsetzen, denn letztlich ist das Endergebnis keine Kostensenkung. Speicher wird auf einer Dollar per Bit-Metrik verkauft und wenn das nächste Dichte-DRAM mehr kostet pro Bit werden die Kunden weiterhin die alten Komponenten verwenden. Dieser Mythos von Moore's Law ist es, dass Kunden für Leistung bezahlen, wenn in Wirklichkeit Hersteller das gleiche Produkt produzieren für geringer Kosten. Gleichzeitig versuchen die Hersteller, die Kunden zu überzeugen, Komponenten mit der verbesserten Leistung / \$ Metrik zu übernehmen. Die verbesserte Leistung kommt eigentlich kostenlos.

Mit Blick auf die ASML-Roadmap ist es durchaus möglich, dass DRAM niemals in EUV migrieren kann. Es ist wirklich erstaunlich, dass es das DRAM so weit gebracht hat, aber jeder gute Lauf muss schließlich zu einem Ende kommen, und wenn es eine einfache und kostengünstige Architektur zur Migration gegeben hätte, wäre es schon passiert.

Was kommt nach DRAM?

DRAM ist am Ende der Straße angelangt, nachdem er sich in eine Technologie-Ecke hin entwickelt hat. Die DRAM-Lithographie ist aufgrund der Komplexität der Kondensatorkonstruktion auf NAND und Logik zurückgeblieben. Die Roadmap von ASML, dem weltweit führenden Unternehmen in der Lithographie, zeigt die Technologie, die in der Produktion für verschiedene Speichertypen verwendet wird. Die erste 3D-NAND von Samsung erschien in der Produktion Ende 2013. Es ist interessant zu sehen, dass NAND in der Produktion seit den frühen 1990er Jahren von Planar zu 3D in nur 3 kurzen Jahren umgewandelt wurde. Es ist wahrscheinlich, dass DRAM einer ähnlichen Umwandlung folgen wird, trotz der Roadmap, die zeigt, dass es für immer weitergehen wird. Es wird strikt eine ökonomische Entscheidung sein. Sobald die Ersatztechnologie eine bessere finanzielle Lösung für die Hersteller ergibt, werden die alten Produkte stark eingeschränkt.

Diese Ersatztechnologie wurde seit Jahren in der allgemeinen Klasse namens Storage Class Memory (SCM) als Zwischentechnologie zwischen DRAM und Speicher diskutiert. SCM ist endlich da, eingeführt von Intel in Form von 3D-Xpoint und Optane SSDs.



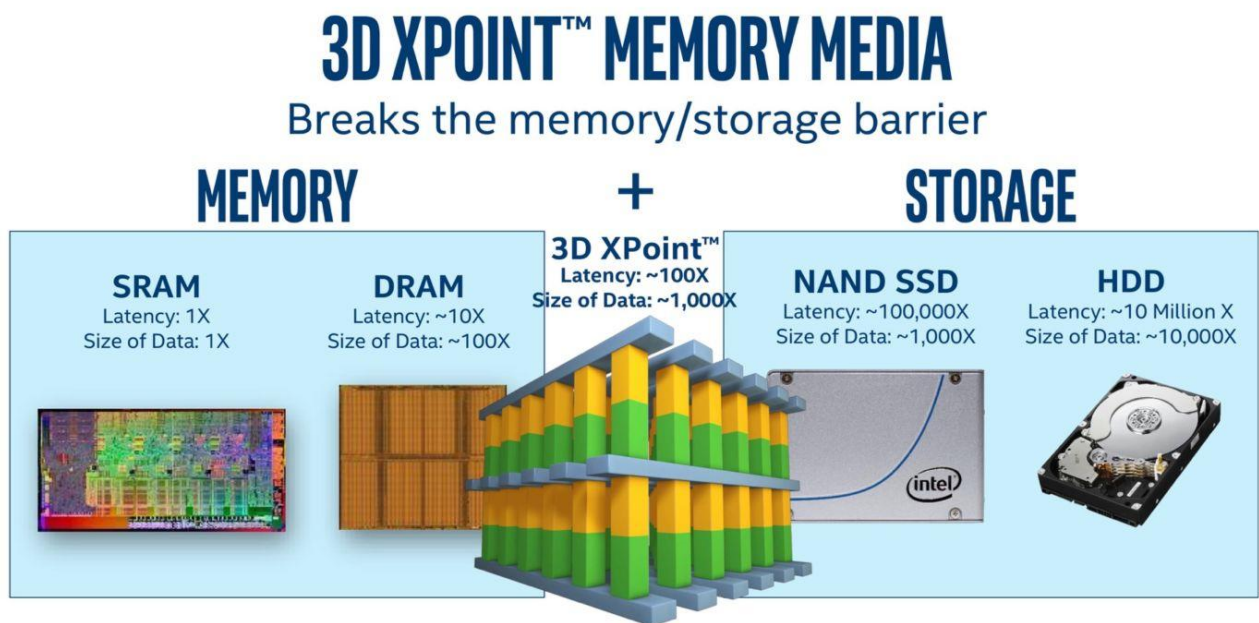
Intels M.2 Optane SSD wurde im März 2017 mit Lieferungen ab April eingeführt. 3D-XP ist teurer als NAND und plaziert 3D-XP hinter einer NVMe-Schnittstelle mit beschränkter Performance ist die Lösung. Aber diese ersten Produkte sind die Hälfte der Kosten und volumetrisch dichter der DRAM. Da der DRAM schwieriger zu skalieren ist, bedeutet das, dass die Kosten nicht so schnell fallen, wie es historisch erwartbar wäre, Lösungen wie 3D-Xpoint finden ihren Weg in die Speicherhierarchie. Wie bei jeder neuen Speichertechnologie oder -schnittstelle stellen die anfänglichen Fähigkeiten nicht das Leistungspotential dar, aber die neue Technologie stellt eine Lösung dar, die in die Zukunft skalierbar ist, so wie 3D-NAND das planare NAND ersetzt hat.

Natürlich hat 3D-XPoint einen weiteren signifikanten Vorteil gegenüber DRAM: Wie 3D-NAND können die 3D-XP-Schaltungen und Bitcells während des Herstellungsprozesses gestapelt werden. Der Name des Produktes impliziert es und die ASML-Roadmap bestätigt. Anstelle von 32, 48 oder 64 Schichten wie 3D-NAND zeigt die 3D-XP Roadmap eine 2, 4, 8 Progression. Wiederum, das Stapeln der Schichten

während der Herstellung, addiert Kosten linear statt exponentiell. Es gibt viel Weg für die 3D-XP-Roadmap und keine für die DRAM-Fertigung.

Eine andere 3D-Lösung: Package und Die Stacking

DRAM hat noch gewisse Leistungsvorteile gegenüber 3D-Xpoint. Es hat unendliche Schreibausdauer, fast symmetrische und niedrige Lese-Schreib-Latenzen und hohe Bandbreite. DRAM kann niemals vollständig durch 3D-XPoint ersetzt werden. Stattdessen erwarten wir, dass es auf dichte Verpackungslösungen entweder in Form von LPDDR mit Paket auf Paket- (POP) -Lösungen oder High Bandwidth Memory (HBM) mit der Silicon Via (TSV) Verpackungstechnologie, wie sie bei Intel (Altera) in Verbindung steht, zu migrieren Mit EMIB Verpackung und Xilinx FPGAs mit Silicon Interposer Technologie. Diese Direct-Attach-DRAM-Lösungen, bei denen der Host-Prozessor eine direkte Verbindung zum DRAM hat, waren seit jeher die Hauptstütze der Off-Chip-Speicherlösung und haben die beste Leistung für den Chip-Chip geliefert. Speicher hat sich auf kostengünstige Kapazität konzentriert und wurde entworfen, um die Leistung innerhalb dieser Limitierungen zu maximieren. Jetzt, mit der Entstehung von SCMs, ist die Kapazität von DRAM nicht so kritisch wie die Verbesserung der Leistung. Erwarten Sie zukünftigen DRAM, um auf niedrigere Latenz und höhere Bandbreite zu konzentrieren, anstatt maximale Kapazität und niedrigere Leistung. Denken Sie daran, dass DRAM-Performance heute sehr schlecht ist im Vergleich zu On-Chip-SRAM. Die Verbesserung der Latenz und der Bandbreite wird helfen, wird aber durch die Chip-to-Chip-Interconnect begrenzt. Die Lösung, um dies zu überwinden, besteht darin, den DRAM in Form eines Multi-Chip-Pakets oder heterogener Integration auf das Paket zu bringen.



Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.

Implikationen für eingebettete Märkte und langfristige Unterstützung

Jeder Speicherhersteller steigert die Investition in die NAND-Produktion und verringert diese für DRAM. Kombiniert mit der Bedrohung durch 3D-Xpoint ist es leicht möglich, dass DRAM so plötzlich enden wird wie planarer NAND, so dass die Branche mit DDR4 für Server, Graphics DRAM und mobilen DRAM bleibt. Eingebettete und Legacy-Märkte für Speicher werden niemals in dem Volumen laufen, das benötigt wird, das benötigt wird, um einen Die auf die fortschrittlichsten Prozesstechnologien zu schrumpfen. Seien Sie vorsichtig, einen Technologiepfad mit dem Versprechen für langfristigen support vom Primärlieferanten beschreiten, aber auch einen, der zu teuer für sekundäre fabless Lieferanten ist, dieses zu unterstützen.

Auch ohne die drei großen Hersteller, DDR3 ist Multi-Sourced und wird hergestellt von zwei verschiedenen DRAM-Foundries. Ohne die drei großen Anbieter scheint DDR4 auf nur eine Foundry, eine Dichte und ganz zwei mögliche Lieferanten beschränkt zu sein, die gleichen Komponenten zu liefern. Es wurde öffentlich bekannt gegeben, dass diese Foundry Geschäftsbeziehungen mit ihrem Prozesstechnologie-Anbieter beendet hat. Es ist also wahrscheinlich, mit 8Gb Kapazität oder DDR5-Technologie Schritt zu halten, wenn das jemals materialisiert, aber es ist ein riskantes Angebot. DRAM-Prozess-Technologie ist hoch spezialisiert, stark vor Patentverletzung geschützt, und, wie bei jedem Halbleiter-Prozess, ist es unmöglich, schnell und unabhängig zu entwickeln. Die Chance, dass ein neues Unternehmen in das DRAM-Geschäft eintritt, ist null. Dies bedeutet dass DDR4 ein hohes Risiko für eine langfristige Versicherung der Versorgung darstellt und wird von daher nicht für Systeme und Märkte empfohlen, die eine langfristige Unterstützung erfordern.

HAFTUNGSAUSSCHLUSS: Alle Produkt-, Produktspezifikationen und Daten können ohne vorherige Ankündigung geändert werden, um die Zuverlässigkeit, Funktion oder Design zu verbessern oder auf andere Weise. Die hierin enthaltenen Informationen entsprechen dem Besten des Insignis Technology Corporation. Keine Haftung für Fehler, Fakten oder Meinungen wird akzeptiert. Die Kunden müssen sich hinsichtlich der Eignung dieses Produkts für ihre Anwendung befriedigen. Keine Verantwortung für Verluste infolge einer Person, die sich auf das hierin enthaltene Material bezieht, wird akzeptiert.