Thin Solid Films, *220* (1992) 1–7

Integration of chemical–mechanical polishing into CMOS integrated circuit manufacturing

Howard Landis, Peter Burke, William Cote, William Hill, Cheryl Hoffman, Carter Kaanta, Charles Koburger, Walter Lange, Micheal Leach and Stephen Luce

IBM Technology Products, Essex Junction, VT 05452 (USA)

Abstract

Planarization by chemical–mechanical polishing (CMP) has been exploited by IBM in the development and manufacture of CMOS products since 1985. Among the products that use this technology are the 4-Mbit DRAM (which uses polysilicon, oxide, tungsten-line and tungsten-stud planarization) and its logic family (which uses four oxide and four tungsten-stud planarization steps). CMP is also used in the planarization of oxide shallow isolation trenches, as in the 16-Mbit DRAM.

Reduced sensitivity to many types of defects is possible with CMP. A wafer that is truly flat is easier to clean, eliminates step coverage concerns, provides for better photolithographic and dry etch yields, and generally minimizes complications from prior level structures. Oxide CMP reduces sensitivity to certain pre-existing defects, such as crystalline inclusions or foreign material in an interlevel dielectric. Metal CMP can reduce the incidence of intralevel shorts relative to conventional RIE processing. Random defects associated with CMP, such as slurry residues and mechanical damage, are controlled by careful optimization of the post-polish clean and of the polish process itself. Systematic defects, such as incomplete planarization over very large structures, are controlled by process optimization and prudent design limitations. These include such things as constraints on the image size, the distance between images, and/or the local pattern density.

Since its introduction in the 4-Mbit DRAM, there has been a steady increase in the use of chemical–mechanical polishing in IBM CMOS products. The number of steps, processes and materials polished continue to rise, both in current and planned future products. Individual applications range from the simple removal of back-side films to complex insulator or metal planarization requiring high removal uniformity. The process tolerances delivered by CMP have decreased faster than image size, even in the face of dramatic increases in circuit and layout complexity. CMP tools are installed in IBM semiconductor manufacturing and development sites worldwide. Chemical–mechanical polish processes and applications provide unique leverage to IBM products, and are a crucial part of both current and planned IBM CMOS technologies.

1. Introduction

In semiconductor manufacturing, the pace of technological change requires decreasing the cost yet increasing the performance of each new product. This is accomplished by both decreasing the size and increasing the complexity of the structures built on silicon. Future DRAM and CMOS logic products will have smaller devices and denser packing, as well as more levels of photolithography and more levels of wiring. As more layers are built on the silicon surface, it is found that problems with non-planarity become severe [1–6]. Three that impact yield and performance are poor step coverage of deposited films, depth-of-focus of high performance photo tools and residual material after dry etching (Fig. 1).

Precision grinding and polishing have long been used for fine glass- and metal-working [7–11]. Although wafer polishing has had an ongoing role in the manufacture of semiconductor substrates [12], and despite numerous research publications [13–23] and patents [24–28], polishing has only recently begun to be recognized as a viable tool for large scale manufacturing of integrated circuits [29–36]. Today, many of the skills of the lensmaker and silversmith are put to use in the chemical–mechanical polishing (CMP) areas of IBM's semiconductor manufacturing lines.

A CMP tool, reduced to its essentials, consists of a polishing pad affixed to a circular polishing table, and a carrier to hold a wafer against the pad (Fig. 2). Both the table and the carrier are made to rotate (usually in the same direction) as the front of the wafer is pressed down against the pad and polishing slurry is dispensed onto the pad. A typical slurry is water-based and contains both abrasive and chemical additives. The combination of mechanical abrasion and chemical etch causes material to be removed from the surface of the wafer. Protrusions erode more efficiently than recessed areas. This leads to planarization of the wafer surface.

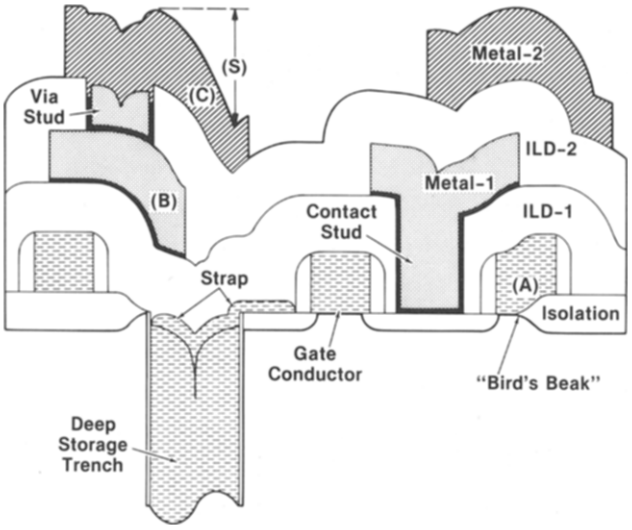

Fig. 1. Schematic of a trench-style DRAM highlighting some problems associated with non-planarity. Poor step coverage and etch residuals are concerns of locations (A), (B) and (C). The step size (S) should be much less than the depth of focus of the photo tool.

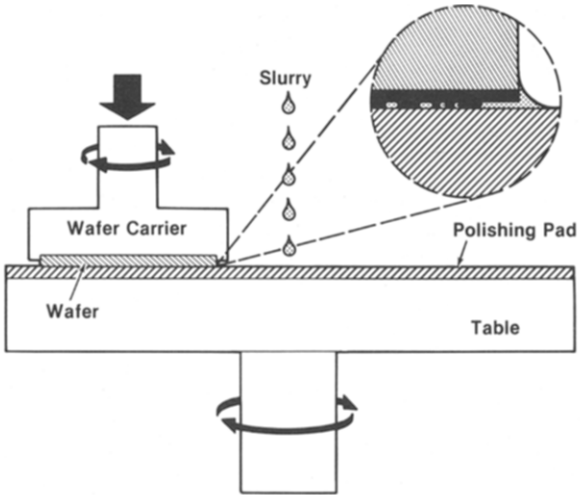

Fig. 2. Schematic diagram of a chemical–mechanical polish tool.

Among the significant tool parameters are the wafer pressure, the table and carrier speeds, and the mechanics of the wafer carrier. The important process variables include the polish pad and its condition, the slurry, the slurry flow, and the process temperature. Besides these obvious controls, there are many subtle tool- and process-specific variables that profoundly influence the character of the CMP process. Thoroughly understanding and properly controlling these subtle influences are crucial for successful implementation of CMP in manufacturing.

2. Applications

Figure 1 depicts a cross-section of an unplanarized deep-trench-style DRAM, as might be found in the 4-Mbit generation. From the deep trench to the metal levels, each structure that introduces topography makes subsequent processing more difficult. CMP is used to improve the planarity of these structures.

In deep-trench processing, when a vertical reactive ion etch (RIE) process is used to clear polysilicon fill from a wafer surface, a center seam is propagated down into the trench, and a step forms at the trench edge during overetch. These defects make it difficult to reliably cover the trench with another film, such as a strap. Chemical–mechanical polishing, however, is a horizontal process: with polysilicon CMP, a seam does not propagate, and a step does not form at the trench edge. Because polysilicon slurries can be quite selective to oxide and nitride masking films [16], the polish process stops short once the silicon nitride layer is reached (Fig. 3). Polysilicon CMP was judged to be a more cost-effective and robust method than RIE for removal of the polysilicon overburden. This CMP process was adopted early in the development cycle of the IBM 4-Mbit DRAM and subsequent programs.

A second source of topography is the isolation level. Shallow oxide-filled isolation trenches are an attractive alternative to standard recessed oxide (ROX). Besides improving planarity, shallow trench isolation (STI) also eliminates the ROX "bird's beak" (Fig. 1), allowing greater circuit density. The STI planarization process for the 16-Mbit DRAM [30] employs a combination of

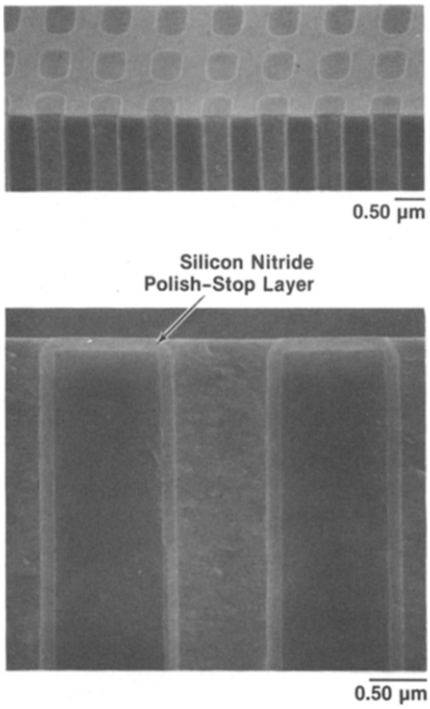

Fig. 3. SEM cross-sections of polysilicon deep trenches planarized by CMP.

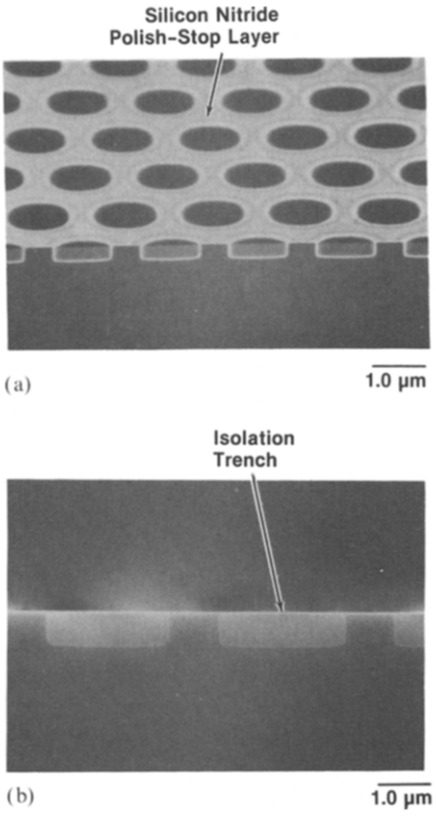

Fig. 4. SEM cross-sections of submicrometer shallow trench isolation test structures. A very thin nitride polish-stop (a) was used to improve flatness after the silicon nitride polish-stop layer was stripped (b).

photolithography, spin-apply and blanket etch-back to remove most of the oxide overburden. Oxide polishing is used for final planarization down to a silicon nitride polish stop (Fig. 4). Because CMOS devices are very sensitive to damage in the silicon substrate, the STI polish process must have very high uniformity and low defect densities.

After isolation, the next level that can obviously benefit from planarization is ILD-1, the interlevel dielectric underneath Metal-1. Reflow of a doped glass [37–41] is often used to minimize the surface angle of this film at gate conductor edges, reducing defects from poor step coverage at Metal-1. But viscous flow can only produce local smoothing [42, 43]; planarization over many micrometers is needed to eliminate metal-etch residuals, and planarization over many millimeters is required to alleviate photolithographic depth-of-focus limitations. Oxide polishing proved capable of providing global planarization over these ranges [44–47]. Since there is no built-in stop-layer for a "blind polish" of this type, high repeatability and uniformity (across the chip, wafer and lot) are critical to the success of ILD polishing. Substantial increases in final chip yields caused the 4-Mbit and subsequent programs to incorporate ILD-1 polishing in their process flows.

Because the high temperatures needed for reflow can be impractical after the first metal level is defined, there is an even greater need for using CMP to planarize ILD-2, the dielectric underneath Metal-2. The adoption of this process by the 4-Mbit program marked the first use of oxide polishing in an IBM DRAM [29].

Vertical vias are critical to wiring metal levels on a tight pitch. Just as with the deep storage trenches, a RIE process to remove the overburden of metal produces two characteristic via defects: the center seam and the edge step. Metallic liners such as titanium can recess quickly during tungsten RIE, causing a possible third via defect, the edge seam (Fig. 1). In addition, CVD tungsten films are very rough, demanding increased overetch and thereby worsening the three via defects described above. On the other hand, tungsten chemical–mechanical polish has proven to be a reliable high-yield process for via formation. Reliability results from the improved flatness afforded by CMP, while yield improvements arise from the ability of tungsten chemical–mechanical polishing to remove random defects (foreign material) from the surface of the wafer. Tungsten-stud polishing is used in the 4-Mbit DRAM and subsequent programs.

Etched tungsten lines are plagued by some of the same defects as traditionally etched vertical tungsten vias. Because of these problems, an inlaid metal, or "damascene", process was developed and implemented in 4-Mbit manufacturing [34]. In this process, trenches for the Metal-1 conductor are etched in ILD-1 just after the stud holes. A single CVD tungsten deposition then fills the stud holes and the Metal-1 trenches, and a single tungsten polish removes the tungsten overburden (Fig. 5). Tungsten damascene not only reduced intra-level shorts, it reduced the overall process cost as well. (With damascene Metal-1, the wafer surface remains flat, so there is no need for ILD-2 planarization.)

A scanning electron microscopy (SEM) cross-section of a 4-Mbit DRAM chip is shown in Fig. 6. Chemical–

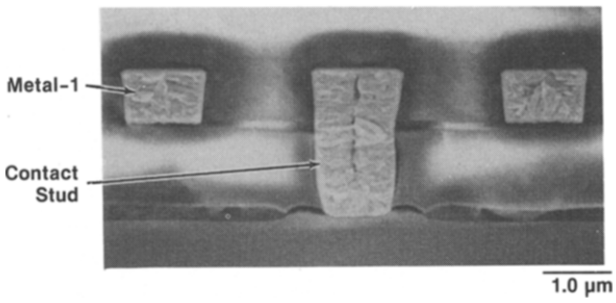

Fig. 5. SEM cross-section of Metal-1 lines and contact studs from a 4-Mbit DRAM chip.

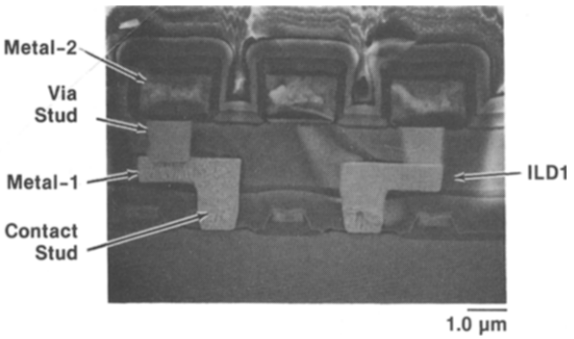

Fig. 6. SEM cross-section of a 4-Mbit DRAM chip.

mechanical polishing was used to planarize the polysilicon deep trench (not shown), oxide ILD-1, tungsten Metal-1 and vias.

3. Manufacturability

To understand how CMP is used in a manufacturing environment, it is necessary to first examine the limitations of the process. The four phenomena that can most constrain the CMP process are non-uniformity, rounding, dishing and erosion [34]. These effects are illustrated in Fig. 7. Non-uniformity refers to polish-rate variations, which can occur across a wafer, across a lot, and lot-to-lot. Rounding refers to the ineffective planarization of wide features, and is most common in oxide polishing. Dishing refers to the thinning of the fill material in the center of a wide inlaid trench (STI or metal damascene). Erosion results from the inability of the polish stop to completely arrest the polish process; it is most severe when the trench area fraction is high (as in some damascene applications) or when the polish selectivity is low (as with STI). Both dishing and erosion are worsened by overpolishing (Fig. 8).

Non-uniformity, rounding, dishing and erosion can be mitigated by process and design changes. The tool, the polish pad and the process all strongly affect non-uniformity. The choice of polish pad has the greatest influence on rounding and dishing: hard, incompressible pads planarize best. Erosion depends on the selectivity, so the materials set and the slurry chemistry strongly influence erosion. Design rules can also help insure product quality. Automated layout and checking

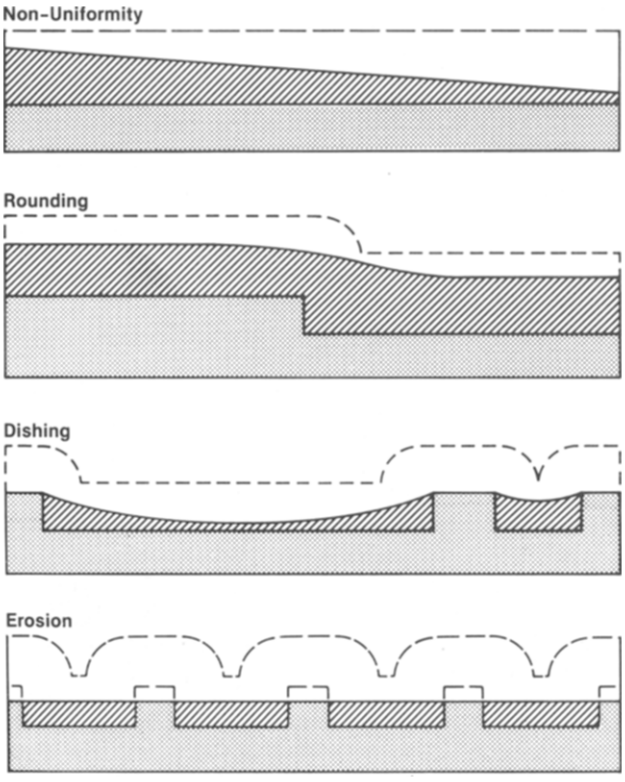

Fig. 7. Non-uniformity, rounding, dishing and erosion can limit the performance of CMP processes.

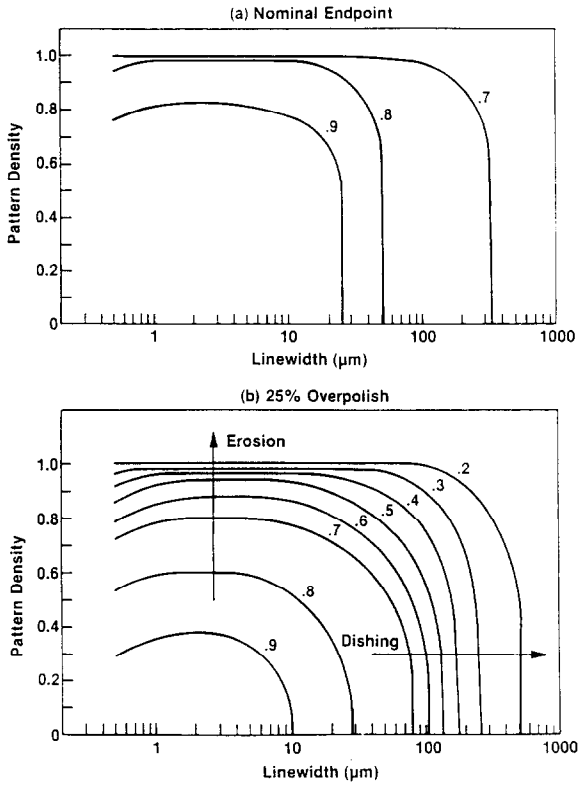

Fig. 8. Sample data from a damascene metal polish process using a hard pad and exhibiting approximately 40:1 polish selectivity. The contours represent the normalized line thickness after CMP at (a) nominal endpoint, and (b) after 25% overpolish. For example, the 0.9 contour indicates that 90% of the original line thickness is maintained after CMP. Wide lines are measured in the center, where they are thinnest.

tools guarantee that feature widths and area fractions remain within appropriate bounds, while technology decisions, such as fully landed contacts, can ease some uniformity requirements.

Uniformity requirements vary from level to level. Certain CMP processes, like tungsten studs and polysilicon deep trench planarization, have a wide process window. This is because a large overpolish can be used to compensate for any non-uniformity. (Small feature sizes limit rounding and dishing, while high polish selectivities minimize erosion.) Other processes, such as STI, ILD and Metal-1, have very tight uniformity requirements. If the 6σ tolerance exceeds the product requirement, closed-loop control can reduce non-uniformity and increase yield. An operator can adjust the polish time if measurements indicate a drift in the polish rate, or an integrated measurement system can provide this feedback for each wafer, but true closed-loop control requires *in situ* endpoint detection. By minimizing within-wafer variations, and by employing closed-loop control where required, the uniformities for all of the CMP processes discussed in this paper have been sufficient to produce parts with consistently high yield.

In general, it is more difficult to planarize a wafer than it is to keep it planar. With ILD oxide polishing there is no stop-layer, so the uniformity of the polished film is tightly bound to the uniformity of the polish process itself. But if the wafer is already flat, a damascene process can decouple the final film thickness from the polish uniformity. Assuming adequate selectivity, a stop-layer will locally arrest the polish process when endpoint is reached, allowing other areas of the wafer to "catch up". This effect can provide a substantial process window for well-designed damascene levels [34]. But the wafer must be truly flat for the damascene process to succeed. The high conformality of the deposited material guarantees that any void or local depression on the wafer surface would be filled, and the high polish selectivity of a good damascene process insures that the unwanted fill material would remain after polishing.

An obvious manufacturability issue is cleanliness. Successful cleaning procedures are designed to take advantage of the flatness of a polished wafer. Since slurry abrasive is the only type of particle on a polished wafer, cleaning procedures are further optimized for the different CMP slurries. The success of the many diverse polishing operations used at IBM attests to the low defect densities achievable with CMP.

Scratches in the polished films or etch stops are another concern. Polishing pressures often exceed 5×10^4 N/m^2 (7 lb in^{-2}), so any foreign particles between the wafer and the polish pad can cause damage. Shallow trench isolation is a particularly sensitive test.

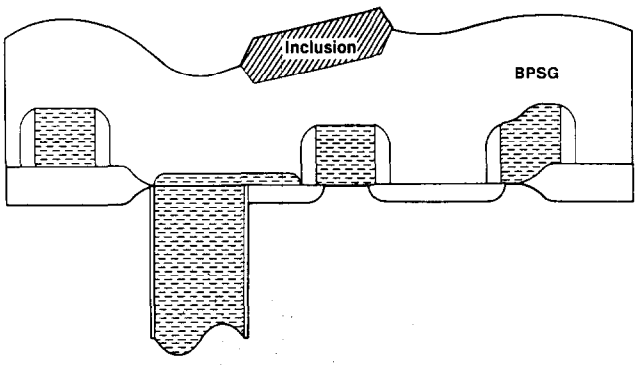

Fig. 9. Crystalline inclusions that occasionally form in BPSG during reflow anneal can be a source for "stringers", or etch residuals, after metal RIE. Polishing the ILD-1 dielectric proved to be very effective at planarizing these random defects.

Damage to the isolation edge or nitride stop-layer can result in an electrical failure if the underlying silicon is affected [48]. STI microscratches are controlled by eliminating process-generated defects and by insuring the cleanliness of the polishing environment.

Instead of creating defects, CMP more often acts to remove existing defects. Metal "stringers", or etch residuals, can form at the edges of large crystalline inclusions that sometimes appear in BPSG during reflow (Fig. 9). Oxide CMP proved effective at planarizing these inclusions. Before damascene Metal-1 was implemented in our 4-Mbit program, this source of tungsten etch residuals was eliminated by the fully planar surface provided by oxide CMP. In many other cases, including tungsten polishing, CMP has demonstrated a unique ability to effectively remove many pre-existing random defects.

4. Extendability

CMP can be extended to more levels, more materials, smaller dimensions and tighter tolerances to satisfy the requirements of future semiconductor products.

Oxide and tungsten-stud CMP are used repeatedly to make 0.7-μm CMOS logic chips with four levels of metal (Fig. 10) [33]. The planarization and vertical studs provided by CMP allow the first three metal levels to be fabricated at nearly the same wired pitch (the last level is built at a relaxed pitch for other reasons). Good early yields [33] are evidence of the ready extendability of CMP to many levels.

As successive generations move to smaller structures and tighter tolerances, two general rules become apparent. (1) Small isolated features, planarize best. Wide features (because of rounding and dishing) and dense areas (because of erosion) are more difficult. (2) The

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.