





# why things fail

From tires to hinges to helicopter  
blades, everything breaks eventually.  
The only question is when.

**BY  
ROBERT  
CAPPS**

**PHOTOS  
BY  
CHRISTOPHER  
GRIFFITH**

## In the corner of Building 4, a massive complex at Ford headquarters in Dearborn, Michigan, the ghostly skeleton of a pickup truck endures a constant torment.

The truck has no wheels, no bed, no seats, and no steering column—it's just a vacant shell and a set of pedals. Inside, a pneumatic piston is positioned to press on the gas pedal over and over again, night and day. It's a test of the whole accelerator assembly, but engineers are focused on one simple part—the hinge that connects the gas pedal to the frame.

Building 4 is Ford's Tough Testing Center, where the company evaluates nearly all of its nonengine parts, from seat belts to axle assemblies. The facility is a monument to a dark truth of manufacturing: Even the best-engineered products fail. Some percentage of all mechanical devices will break before they're expected to. "Companies come to me and say they want to be 100 percent failure-free after three years," says Fred Schenkelberg, whose firm, FMS Reliability, estimates the lifespan of products. "But that's impossible. You can't do it."

Consider a few recent examples. In 2009, Mohawk Industries—one of the largest makers of carpeting in the country—was forced to discontinue an entire line of carpet tiles when the tiles failed unexpectedly, costing the company millions. In 2010, Johnson & Johnson had to recall 93,000 artificial hips after their metal joints started failing—inside patients. In 2011, Southwest Airlines grounded 79 planes after one of its Boeing 737s tore open in midflight. And just this past summer, GE issued a recall of 1.3 million dishwashers due to a defective heating element that could cause fires. Unexpected failure happens to everything, and so every manufacturer lives with some amount of risk: the risk of recalls, the risk of outsize warranty claims, the risk that a misbehaving product could hurt or kill a customer.

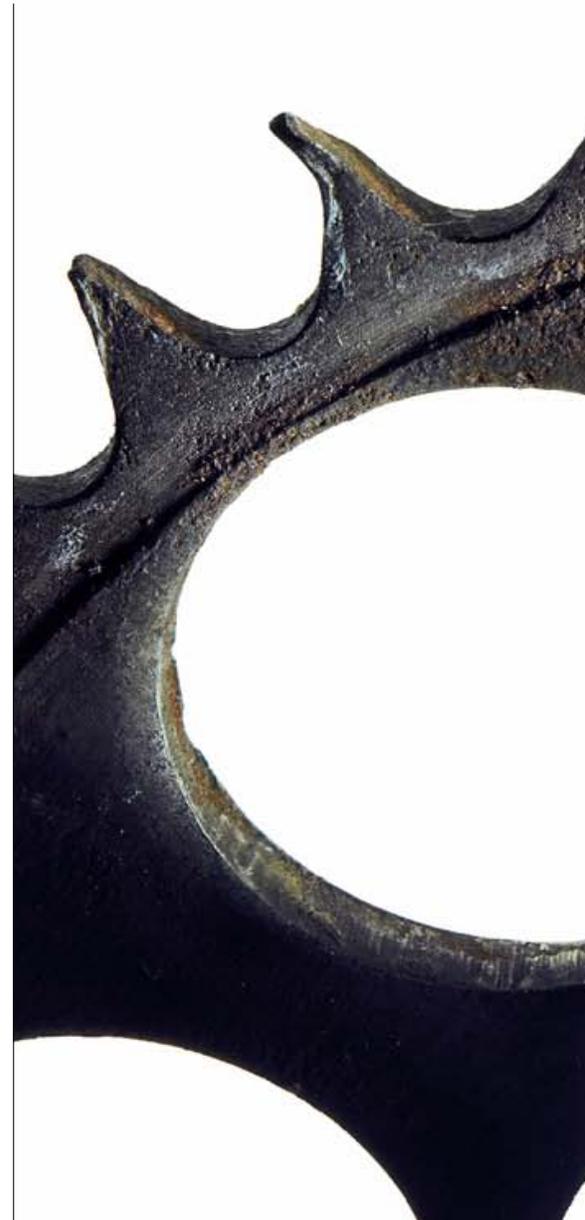
This is why the sprawling hangar-size rooms of Ford's Building 4 are full of machines. Machines that open and close doors, robots that rub padded appendages on seats, treadmills that spin tires until they erupt in a cloud of white smoke. There's even a giant bay where an entire Ford pickup is held up in the air by pistons that violently shake the vehicle by its

suspension. Officially, Building 4 is about reliability, but it's actually more about inevitability. Ford isn't trying to ensure the gas-pedal hinge will never break. The company knows it *will* break; its engineers are trying to understand when—and how and why—this will happen.

Product failure is deceptively difficult to understand. It depends not just on how customers use a product but on the intrinsic properties of each part—what it's made of and how those materials respond to wildly varying conditions. Estimating a product's lifespan is an art that even the most sophisticated manufacturers still struggle with. And it's getting harder. In our Moore's law-driven age, we expect devices to continuously be getting smaller, lighter, more powerful, and more efficient. This thinking has seeped into our expectations about lots of product categories: Cars must get better gas mileage. Bicycles must get lighter. Washing machines need to get clothes cleaner with less water. Almost every industry is expected to make major advances every year. To do this they are constantly reaching for new materials and design techniques. All this is great for innovation, but it's terrible for reliability.

At Ford, learning exactly when and how things will fail—over many years and across a spectrum of millions of vehicles around the world—can save untold amounts of money and maybe even human lives. So in the stripped-down cab in Building 4, the piston continues to push on the gas pedal, then let up, then push again, over and over. This simple exercise is worth billions of dollars. Look closely enough and you can see all the complexity, perils, and opportunities of managing failure. And, as it happens, you might also catch a glimpse of the future of manufacturing.

**F**ord knows product failure. A little more than 10 years ago, it got as harsh a lesson in the subject as any company in history. The ordeal began in 1999, when a TV news reporter in Houston named Anna Werner started looking into an intriguing story. A local attorney had told her about a fatal car accident that had been caused



by an apparent tire defect. A steel-belted Firestone had inexplicably ripped apart in what's known as a tread separation. This caused the vehicle—a Ford Explorer—to flip over, killing the driver. The woman's family had filed a suit. Curious, Werner started calling other law firms. Eventually she found more than 20 accidents, which killed nearly 30 people, all involving Ford Explorers riding on Firestone tires.

The KHOU story aired in February 2000. Spurred by the media attention, nearly a hundred reports of tread separations flooded into the station and the offices of the National Highway Traffic Safety Administration. Ford and Firestone blamed

---

**ROBERT CAPPS** (@robccaps) is WIRED's articles editor. He wrote about the "Good Enough Revolution" in issue 17.09.



*Manufacturers are constantly adopting new materials and new design techniques. This is great for innovation, but it's terrible for reliability.*

each other. Firestone insisted that the carmaker, in an effort to solve stability problems with the Explorer, had set the vehicle's tire pressure recommendations too low. Ford maintained that the tread separation was caused by a flaw in Firestone's manufacturing process. Lawsuits were filed, congressional hearings held. Eventually more than 14 million tires were recalled. It's estimated that some 192 people died and 500 were injured in tread-separation accidents—most of them involving Ford vehicles.

Ford still doesn't like to talk about the disaster, but it's clear that in its wake, the company overhauled its testing process. The company's warranty costs have

plummeted, and in the *Consumer Reports* annual survey, Ford cars and trucks went from having some of the worst reliability scores in the early 2000s to having some of the best by 2010. Now it regularly competes with the likes of Honda and Toyota. From the embers of the Firestone disaster, Ford rose to become one of the best companies in the world at managing failure.

This achievement can be partly attributed to what's going on in Building 4. But an initial impression of the place can be deceiving. If you watch all those vehicles and parts get pounded, pressed, and shaken, you might come away thinking that Ford is simply trying to make sure its

cars and trucks can withstand enormous levels of abuse. You'd be wrong.

Consider the gas-pedal hinge. All you really want to know is this: How many times does the piston have to press the gas pedal before Ford engineers will be sure the hinge is sound?

"I'm not going to tell you that," says Todd Brooks, one of Ford's engineering supervisors, half laughing, half recoiling from the thought. "Are you kidding me? GM would *love* to get that information." The number of piston presses, it turns out, is a closely held trade secret—and the reason why speaks to the complexity of failure testing.

It's actually not hard to make a hinge that

will last for a really, really long time. All you have to do is make it a tough, heavy hinge. But that creates several problems. First, a burly hinge will be stiffer and less sensitive than a small, thin hinge, so the pedal won't feel right. Second, and worse, is the excess weight. Slap a big hinge onto the gas pedal and you may add only a couple of ounces and a few cents of overhead to the truck. But multiply that across hundreds of hinges, bolts, handles, door locks, latches, and so on, and suddenly you have a bloated truck that is slow, sluggish, gas-hungry, and expensive. A truck that is, in the parlance of reliability testers, *overengineered*.

can accept is diminishing, and as a result the amount of risk the company must tolerate is increasing. Just as laptops need to get faster, thinner, and more powerful every year, cars need to continually get both more powerful and more fuel-efficient. And one of the best ways to achieve both goals is to focus on weight. Make the car lighter and you've improved both gas mileage and performance in a single stroke. So almost every component of every Ford vehicle gets put on the scale. It's not just that Ford wants a hinge that won't break. It needs a hinge that's as durable as possible while also staying as light and inexpensive as possible. Get

out stopping. Most of his days are spent in his small office, working on a vast array of spreadsheets and PowerPoint slides—files that contain detailed warranty information for 1,107 companies. Collectively, these sheets hold perhaps the most comprehensive accounting of product failures on the planet.

Warranty information is one of the most closely guarded secrets in corporate America. Companies are loath to share how much they spend on warranties and why. It's understandable, as talking about warranties is the same as talking about the fact that your products break when they're not supposed to. Because of this, nobody just gives data to Arnum. He has to dig it out, one company at a time.

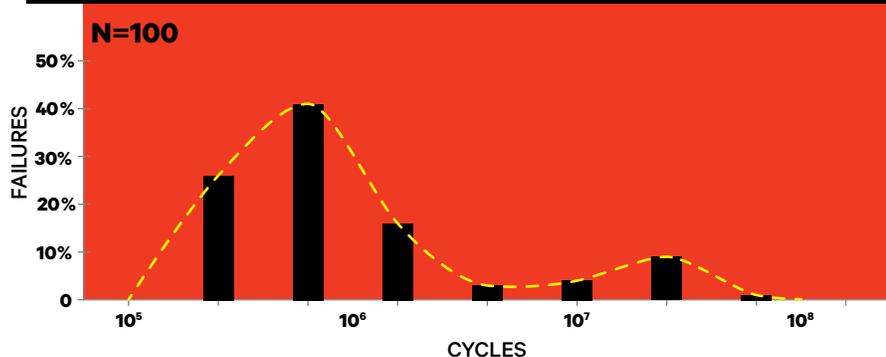
Arnum owes his livelihood to Enron. In the wake of the scandal that took down the energy juggernaut, the Financial Accounting Standards Board made changes to the Generally Accepted Accounting Principles—the rules that, among other things, govern how companies write financial statements. As of November 2002, companies were required to provide a detailed reckoning of their guarantees, including their warranty reserves and payments, in quarterly and yearly filings. The result was that, for the first time in history, someone could look at, and compare, how US public companies handle claims—how much they pay out, how much they hold aside for future payments.

And this is just what Arnum did. He began gathering warranty information, 10-Q filing by 10-Q filing. His job is even harder than it sounds. Because companies are so reluctant to share this information, they often stash their warranty numbers in footnotes. Arnum frequently must pick through an entire hundred-page filing before he finds what he's looking for. Then he enters this information by hand into his spreadsheets.

This meticulous work has produced revelations. Before, even information like the size of the market—how much gets paid out each year in warranty claims—was a mystery. Nobody, not analysts, not the government, not the companies themselves, knew what it was. Now Arnum can tell you. In 2011, for example, basic warranties cost US manufacturers \$24.7 billion. Because of the slow economy, this is actually down, Arnum says; in 2007 it was around \$28 billion. Extended warranties—warranties that customers purchase from a manufacturer or a retailer like Best Buy—account for an

## The Failure Curve

Product failure happens in what's called a Weibull distribution and often looks roughly like a bell curve. Ensuring reliability requires knowing where this curve begins and where it peaks. The chart below shows the logarithmic failure curve of steel bars placed in a fatigue machine. Most fail after 1 million cycles, but if you were to test only a few bars, those failures might occur after 10 million cycles. This might cause you to think the steel is much stronger than it actually is.



The amount of overengineering a product can tolerate depends on what the product is. Airplanes, for example, are a classic example of overengineering because the cost of even minor failure is so high. But with this overengineering comes excess weight—and the resulting loss in fuel efficiency makes flights more expensive than they otherwise could be while also causing them to generate greater carbon emissions. On the other hand, some products—like carbon-fiber racing bicycles of the sort you'd see in the Tour de France—are almost entirely about performance, and so they're consciously underengineered. Obviously, the makers of such bikes don't want them to shatter going up l'Alpe d'Huez. But having a few frames that crack earlier than expected is better than adding even a few ounces to a bike.

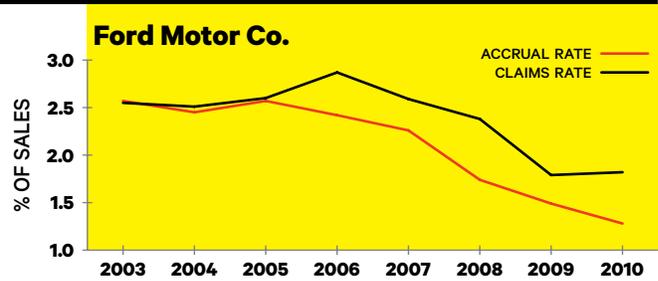
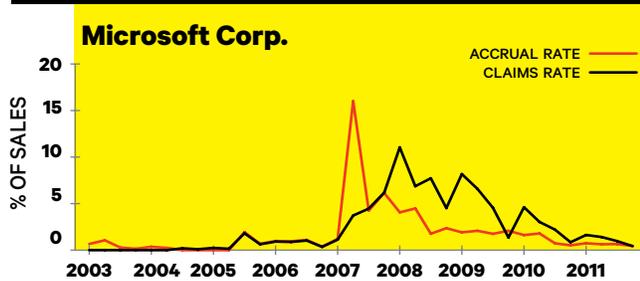
The amount of overengineering Ford

it right and the truck meets the demand for constant improvement: The gas-mileage ratings on next year's window stickers will be higher while the 0-to-60 times might tick down. The problem, of course, is that occasionally the Fords of the world get it wrong. And when they do, they pay a hefty price.

One of the world's foremost experts on the cost of product failure lives and works in a fifth-floor apartment on a modest block in Forest Hills, Queens. His name is Eric Arnum, and he runs a one-man newsletter titled *Warranty Week*. Tall and soft-spoken, he can (and often does) talk about warranty accruals, payment rates, and reimbursement policies for hours with-

## The Billion-Dollar Question

Successfully managing failure can have a huge impact on a company's bottom line, as this data from *Warranty Week's* Eric Arnum shows. (Accruals are how much money a company puts away in anticipation of warranty payments; claims are what they actually pay.) Microsoft missed problems with the Xbox 360 and lost more than a billion dollars. Since 2004, Ford has upped reliability—and has saved a billion.



estimated \$30.2 billion in additional claims payments. Before Arnum, this \$60 billion-a-year industry was virtually invisible.

Then there are the warranty “events.” When a company gets something seriously wrong, it shows up in an Arnum spreadsheet. Asked for a dramatic example, he thinks for a second, then says, “the Xbox 360.”

Microsoft released the Xbox 360 during Thanksgiving week in 2005. Within a day of the machine going on sale, the game consoles were overheating and dying. As time went on, these failures earned a name: the Red Ring of Death. The moniker came from the fact that when an Xbox 360 failed, three lights around its outside power button glowed red rather than the normal green.

The Xbox 360 issues first hit Arnum’s radar in the summer of 2006, when he got a news alert that console owners had petitioned Microsoft to extend the Xbox’s 90-day warranty. Microsoft did extend the warranty to one year but still denied that

there was a problem, insisting the 360’s failure rate wasn’t exceptionally high—3 to 5 percent at most, well within the normal range for a new game console. But there was clearly a problem, and angry gamers were becoming more and more vocal.

Microsoft stonewalled on the issue until Fourth of July weekend in 2007—a full year and a half after the launch. Then Peter Moore, VP of Microsoft’s Interactive Entertainment division, wrote an open letter officially acknowledging the Red Ring of Death. He announced that Microsoft was extending the Xbox warranty to three years for Red Ring issues and said the extension would apply retroactively. Anyone who had previously suffered a Red Ring would be reimbursed for repairs. In a stunning admission of how badly it had messed up, Microsoft also revealed the amount of money it was setting aside for the program: between \$1.05 billion and \$1.15 billion. It was a monumental disaster.

To this day, Microsoft has never acknowledged the cause of the problem, but it’s generally assumed to be overheating. The processing unit would heat up the inside of the 360 to the point that the circuit board it was placed on started to warp. This caused the solder joints—made with lead-free solder to meet new European environmental standards—to break.

The Xbox 360 was one of the most public warranty debacles in the past decade, but it was hardly the only one. “There’s an Xbox in every industry,” Arnum says. “They try their best to keep it quiet, to minimize it, whatever they have to do.”

But, of course, there is also good news in Arnum’s data. I ask him to show me his slide on Ford. It clearly confirms that the company’s warranty payments have declined. At first it looks slightly unremarkable. But then Arnum puts it in context: “This,” he says, pointing out how much Ford is saving on warranties today compared with where it was a few years ago, “is a billion dollars.”

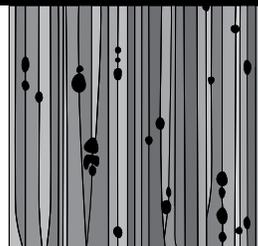
Whenever a new part—like that gas-pedal hinge—is designed, the first question an engineer must ask is, how long does it need to last? Ford’s standard warranty guarantees all parts for three years and engines and transmissions for six. But Ford wants to be sure its products last longer than this. To ensure that parts easily surpass warranty claims (and hopefully ensure that buyers feel they own a reliable product), Ford aims to have everything last 10 years. Upholstery, transmissions, paint—all of it is built to last at least a decade. Ford has not only constructed nearly all of its elaborate lab testing around the 10-year mark, it has also built tracks that are designed to, **I CONTINUED ON PAGE 204**

## Failure Under the Microscope

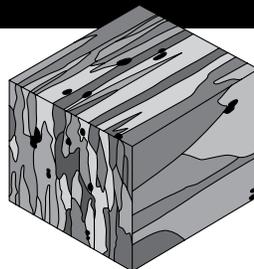
One of the biggest challenges in predicting when a product will fail is understanding the material it’s made from. Every material, from metals to composites to ceramics, will have microscopic variations from unit to unit that affect a product’s lifespan. One company, Vextec, hopes to solve this problem—by creating statistically accurate computer models, down to the grains, voids, and crystals that make up a material’s microstructure.



AN ACTUAL PHOTO OF A METAL’S MICROSTRUCTURE.



VEXTEC’S SIMULATION, COMPLETE WITH VOIDS, GRAINS, AND IMPURITIES.



THE 3-D VERSION WILL SHOW HOW AND WHEN CRACKS FORM.



## Why Things Fail

CONTINUED FROM PAGE 167

over a number of runs, roughly simulate a decade of regular driving. The problem, of course, is that it's impossible to make a product that lasts exactly 10 years. But setting this goal provides a concrete minimum to work with. And establishing that minimum—the point where it's OK to start seeing the first product failures—is one of the most vital parts of reliability engineering.

If you chart failures over time, you will almost always see some form of bell-shaped curve: A few units will fail early, most will fail in a cluster in the middle of the chart, and a few will last much longer than expected. Knowing when the first failures will happen is vital to guaranteeing reliability. On Ford parts, the very first fails aren't supposed to happen until just after the 10-year mark (with most of them occurring much later).

The problem is figuring how to be sure something will last for 10 years. Obviously you can't test for 10 years. Instead, you have to simulate 10 years of use.

The standard solution to this problem is to start building hinges, pressing on them, and seeing how long they last. This is the test-to-failure method. But it's hardly a perfect solution. If you break one hinge, you get one data point—you only really know when that one hinge, with its particular material composition, broke. (And because you broke it, you're never going to actually use that specific hinge.) You have no idea where the failure falls on the curve. Was it a first fail? A long-laster? Somewhere in the middle? So you break more hinges to get more data points. But it turns out that you have to break a lot of hinges to get a satisfying graph. In fact, to even start to get statistically significant results, you have to break thousands of hinges. That might sound somewhat doable with hinges, but it gets

horrendously expensive when you move up to things like engines.

Since statistically accurate test-to-failure simulations are prohibitively costly, what Ford ends up doing is essentially taking an educated guess at how long a part should last. It then runs a few tests simulating real-world conditions to help reassure the company that the parts last *long enough* (no breaking required). But this approach creates a new problem: What is 10 years of use? How many times will that gas pedal be pressed, on average, in 10 years? How many times will extremely active drivers press it? How do you know you're not pressing too many times—simulating, say, 20 years of use and thereby ending up with an overly heavy and expensive hinge?

Mike Herr, an engine durability expert at Ford, has a chart he uses to illustrate the problems with physical testing—in this case as it applies to engines. It's a pyramid, and the top triangle is labeled Vehicle Testing. This is what happens at the company's proving grounds—Ford builds an entire car or truck and drives it over harsh terrain to see how it performs in its native environment. Below that is Engine Testing. This takes place in Ford's Dynamometer Lab, another sprawling, labyrinthine complex, in this case filled with rooms where engines run continuously, undergoing their own versions of the hinge test. The next layer in the chart is Subsystem Testing, which focuses on, say, just the airflow system of an engine, which can be done on a lab table. Below this is Analytical Validation (computer models), and at the bottom is Design Rules. This last category is simply the rules that Ford uses when it starts designing engines.

The higher you are on the pyramid, Herr explains, the more expensive—and laborious—the testing. Building and testing a full vehicle is a pricey and time-consuming affair, which is why Herr and his team constantly struggle to push their testing lower and lower down the pyramid. They're always asking themselves if they can, say, get more out of engine testing, so the company can do less full-vehicle testing. If Ford just built the same engine over and over, ensuring reliability would be easy—the company would simply know how to build its engine. But under pressure to constantly improve performance and efficiency, Ford must always be designing and trying new iterations. So the real target

is the second layer from the bottom of the pyramid: Analytical Validation. Engineers want to be able to test as much as possible in the computer.

It helps that everything starts in silicon. Almost all Ford parts begin their lives as CAD files. So the geometry of the components is already in digital form. The next step is predicting stress, and computers are actually very good at this too. You can import CAD models directly into statistical stress-modeling software called Finite Element Analysis, programs that use complex equations to simulate things like pressure and temperature being applied to the CAD models. When the piston presses the gas pedal and engages the hinge, engineers already know—thanks to Finite Element Analysis—the exact amount of stress each part of the hinge will experience and how the energy will travel through the hinge.

But once you know the stress, the next thing you need to determine is the strength of the hinge—and here is where computers falter. "Actual material behavior is simply more complex than people can model," says Drew Nelson, a professor of mechanical engineering at Stanford University who works on material fatigue. "At a microstructural level, the mechanisms that cause cracks to form are not fully understood." Because of variations in the raw material and manufacturing process (how much heat was applied, how much dust it was exposed to, and so on), every hinge is unique in subtle ways. Even very small changes, like tiny shifts in the size and orientation of metal grains, can alter how the material performs.

Models tend to assume components that are identical in their material composition. The result is that virtual components tend to fail at the same time in every simulation. But actual failures will occur in that bell-shaped distribution. If you could simulate that curve in software, you could finally get the upper hand on risk.

Five hundred miles south of Detroit, in Nashville, Robert Tryon understands the problem as well as anyone. For years Tryon was charged with predicting the life expectancy of aircraft engines for General Motors. He was constantly frustrated with the methods available for assessing materials. After deciding what kind of metal it wanted to

use in an engine, GM would have a smooth, round bar of that metal made for testing. The engineers would then repeatedly pull the ends of the bar until it broke. This, in theory, provided a failure point for that material.

The problem, again, was getting enough of these data points. "You need to test 3,000 parts to get a reliable 1-in-1,000 number," Tryon says. In other words, to statistically predict that one bar in a thousand that is going to be a first fail—snapping at the beginning of the failure curve—you would need to test 3,000. But this was utterly impractical. "We were elated if we got 25 bars to test," Tryon says. The solution was to test what bars they could, then build in a margin of error by dividing the load under which the bar broke by three or four. This made their estimates extremely rough—especially given that no components are actually shaped like smooth, round bars.

The solution to the problem seemed obvious: Find a way to model the strength of a component—with all its material variability—in the computer, the way you can model stress. GM wanted such a tool so badly that it sent Tryon off to do research in this area as an engineering PhD student at Vanderbilt University. While there, Tryon met Animesh Dey, who was pursuing a doctorate in civil engineering, and the two began working on developing a material simulator. But by the time Tryon presented his thesis, GM had sold the division he worked for and, in essence, laid him off. So he and Dey started their own company, Vextec, to see if they could use their new simulation techniques to help manufacturers better predict failure. They call their software tool Virtual Life Management. Vextec has attracted a number of large clients—including American Airlines, the US Army, and medical-device maker Boston Scientific—and its predictions have proved eerily accurate.

How? Most reliability studies today use physical testing to create a model of how a material will perform. Again, the problem with that concept—whether testing gas-pedal hinges or airplane fuselages—lies in correctly establishing the complete failure curve and knowing when the first fails will happen. Virtual Life Management, by contrast, bases its predictions on the microstructure of materials; it models the variations that happen at the microscopic level. In the case of metals, the microstructure consists of small crystals, and every-

thing about these crystals—how they're shaped, how they line up, where spaces appear between them—affects the properties of the material. Indeed, the shape of the failure curve essentially derives from the particular way that the microstructure within a material varies from inch to inch, from part to part. So Virtual Life Management tries to model these little crystal grains and simulate a pattern of variation in them that roughly conforms to the variations that happen in the real world.

To start the process, a Vextec client pulls a sample of its product off the production line, after which the component is cut open, polished, dipped in acid, and examined under a scanning electron microscope. The result is essentially an image of the component's microstructure. Vextec's algorithms then assess this microstructure: What are the grain sizes and orientations? How often do voids appear and in what shape? How frequently do particles of dust or other contaminants appear? The algorithms create a set of rules for the material—a statistical model of every aspect of the microstructure. The rules are used to create multiple virtual versions of the material whose microstructures vary within the rough range that the client could expect to see in manufacturing.

This process allows Vextec to make a set of virtual models—hundreds, thousands, or even millions of them—each with a similar but not identical microstructure. Combine these models with information from Finite Element Analysis and suddenly you have the ability to fully simulate a component, and to do it until cracks start forming. Vextec's software even predicts how cracks will move through the material. Now that it's possible to run simulations on a thousand virtual samples, clients can have enough data points to get a statistically valid failure curve. This can be done in minutes at very little cost. And it works with almost any material, from alloys to composites and plastics to ceramics. In one instance, Vextec was asked to examine a transmission box on a helicopter. The simulation predicted that after a set number of cycles—flights—the gearing would start to crack. In the field, the actual helicopters had broken exactly when (and how) Vextec's software predicted they would. The company has had similar success examining medical devices, manufacturing equipment, and turbo-charged engines.

**V**extec modeling software is still new and in the very earliest stages of adoption. It's hard to say if it will always work as advertised. Certainly there are detractors who still don't believe such sophisticated modeling is possible. But the future of both failure prediction and manufacturing clearly lies in a simulation tool like Vextec's. Once you can model everything in the computer, all sorts of new opportunities arise. You can alter the shape and thickness of a product and see its estimated lifespan change, all on the fly. You can even create bespoke materials, subtly tweaking alloys in software until you find one that performs the way you want.

When that future arrives, Ford will no longer need a piston to push on that hinge. It might not even need a Building 4. Everything will happen in software.

And this will be a very big deal. If you want to see how a hinge can matter, just look at Toyota. In 2007 the company started getting reports that vehicles were unexplainably accelerating, even when the driver wasn't pressing the gas. Like Microsoft did with the Xbox, Toyota initially downplayed the problem. It was just floor mats, the company said, sliding out of place and jamming the accelerator. This answer, however, didn't satisfy the National Highway Traffic Safety Administration. For a while, the investigation focused on software—was Toyota's new drive-by-wire system faulty? Was there a bug in the software that caused the cars to accelerate? In the end, it did turn out to be floor mats (or driver error), but they also discovered a good old-fashioned mechanical problem: Pedals could actually get stuck. And the problem was in the hinge.

The culprit was a "shoe" in the hinge assembly. Material in the shoe was wearing over time, creating friction and, eventually, sticking. In time, pedals could stick so much that they wouldn't disengage. If Toyota had been able to simulate this material in software, to see how it was affected by wear over time, the company might have spotted the problem before a single faulty shoe was made—and saved itself a recall of more than 4 million vehicles.

Can a hinge be worth a billion dollars? Absolutely. Which is why, in Building 4 of Ford's Dearborn campus, the piston presses on. 