

Chapter 10

The SOA Supervisor

In This Chapter

- ▶ Somebody must be in charge
 - ▶ All about plumbing
 - ▶ The SOA supervisor and plumbing
 - ▶ What the SOA supervisor does
-

SOA is about end-to-end processing — all the time. Here's what goes on:

Someone (or possibly some software) requests a business service. The service broker goes into action, consulting the registry and then orchestrating the connection of all the necessary components of the service. It binds them all together, probably with the help of an enterprise service bus to manage all the messages and the service runs. Hurrah.

Now that this miracle of modern software architecture is up and running, though, what's keeping track of the whole set of computers and networking resources and software so that it continues to run without any problems?

Why, the SOA supervisor of course! (Not surprising, given the title of this chapter; were you expecting maybe a Ouija board?) It's a big role and, to be honest, the SOA supervisor can't do it on its own. It needs help from a whole set of infrastructure services to keep the situation under control. It needs help from "the plumbing."

The Plumbing

Throughout this book, we refer to the complex technical software that keeps a whole data center and network running as *the plumbing*. It's our view that you can find out a lot about service oriented architectures without having to

120 Part II: Nitty-Gritty SOA

know lots of technical details, so why not just think of it as a collection of pipes that run under the floorboards? It might be horrifically complicated, and you might need to employ a current-day Albert Einstein just to keep it running, but as long as you can pay his salary, it does stay running and everybody's happy.

But frankly, business leaders have never been particularly happy with the plumbing. Here's why:

Every now and then, the CIO delivers the bad news to the folks on the business side of things. "Business folks," he (or she) says, "you're going to need to spend millions of dollars more just to keep everything working the way that you want it working with acceptable levels of service." It isn't that the business folks don't believe him; they've been through the exercise in which they examine every item in the IT budget and get the CIO to tell them exactly what it's for. And it isn't that they don't end up doing more and more with IT every year. They may remember a time when they didn't even have e-mail and now an avalanche of it swamps them every morning; next quarter, they'll be implementing a new Voice-over IP system that's bound to save money on the phones.

The problem is that the company makes *widgets*. It manufactures widgets, it markets widgets, and it sells widgets. It's admittedly a big business, but the things that really matter to this business are its key business services: the widget manufacturing process, the sales process, the customer ordering process, the delivery process, and the after-care program. When a company invests millions of dollars in IT, it wants to know how it is improving these processes — and if, in fact, it is.

This has been a problem with IT since Pontius was a Pilate. You can spend millions of dollars on big servers or storage area networks (SANs) or high-speed networking or ERP systems or databases, but you can't easily tie the whole investment back to "the widget manufacturing process."

This genuine sore point became sore enough for many of the major IT vendors that they changed the nature of their technology products so that they focused on "aligning IT with the business." In fact, so many vendors did this that "aligning IT with the business" has become both an IT marketing mantra and an IT marketing cliché.

The truth is that IT was always aligned with the business, in the sense that technology was always purchased for sensible business reasons even when poor purchasing decisions were made. But the links between the business process and the technology were missing. The vendors of software management products (plumbing products) led the market in delivering the necessary change.

The change of direction did not happen because of a desire to accommodate service oriented architectures. In fact, it wasn't driven by SOA at all. It was driven by the need for service level management — the need to properly manage IT service levels in line with the business service they support. (We talk a lot about service levels in Chapter 4.) Nevertheless, this change of direction was very convenient for service oriented architecture, and, as this became clear to the IT industry, SOA and service level management started to play nicely with each other.

Layers upon layers upon layers

Figure 10-1 shows the three major layers of IT, which we like to think of as the IT layer cake. The plumbing layer is in the middle, and the first thing to note about it is that it is the least visible part of the cake.

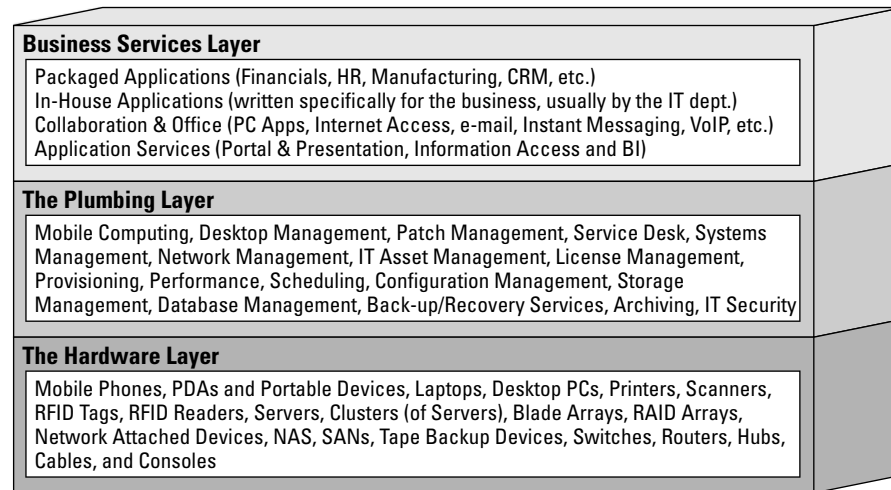


Figure 10-1:
The IT
layer cake.

The bottom layer, the *hardware layer*, consists of things that you can see. You can touch them, turn them on and off with a switch, and kick them. The diagram provides a list of some of these things. None of these items are useful unless you can use them in conjunction with software.

The top layer, which we call the *business services layer*, consists of all the software that is directly useful to the business. This layer contains specific business applications or provides useful ad hoc services, such as presentation services (presenting an application's interface on different devices, such

122 Part II: Nitty-Gritty SOA

as PDAs and PCs, or through a portal) and inquiry services (services that allow you to search through [and find] information). This is the software that all the business users see, know, and sometimes love.

The middle layer, the plumbing layer, is what makes it possible for the top layer to run on the bottom layer.

We could give you a long list of every possible kind of product in this layer, with a description of what it does, and we would doubtless bore you to death in no time at all. So we don't do that. We just tell you what we think you need to know about the plumbing layer for SOA.

The plumbing service

The IT trend of “aligning IT with the business” is about service levels. You can easily link the costs of some of the plumbing layer activities to the applications that use them. For instance, the cost of activities like backing up and archiving files, including the cost of the hardware used, can be attributed to specific applications according to the actual usage of each service by each application. In fact, you can do this with a fair amount of the plumbing layer because what it does for an application is easy to understand.

But you can't tie the actual usage of *every* piece in the plumbing layer to specific applications.

You also need to understand the fact that attributing costs is not really about business applications anyway; it's about business services. To know how much the full IT costs are for a business service, you need to have an accurate map of all the applications that are used to deliver the service and all the hardware involved. Until the IT industry began to focus on service level management, such business process maps were rarely created.



The key point is that some business processes are a matter of life or death for the company, and some are not. Manufacturing processes, for example, usually fall into the “life or death” category, whereas HR systems rarely do.

Life or death availability

What do you think counts as “life or death” availability?

Some commentators think in terms of “five nines” as being the “life or death” level. Five nines means 99.999 percent, which translates to being out of action for no more than 5.25 minutes in a year. Others think in terms of “seven nines,” 99.99999 percent availability, which means being out of action for no more

than about 3 seconds in a year. The need for availability at such levels is rare, and that's a good thing because providing a "five nines" level of availability is very expensive, and providing "seven nines" is very, very, very expensive.

Providing such levels of availability can mean having devoted "hot standby" systems that are ready to run at any moment in the event of a failure somewhere and having software that will transparently fire up these systems and connect users to them when a failure occurs. It requires a lot of plumbing. It even requires a bank of diesel generators in the back of the building to guard against black-outs, brown-outs, and an outright collapse of the power grid.

But providing any level of availability for an application beyond about 98 percent requires making well-thought-out technology choices and organizing a little bit of plumbing. (As a rough guide, 98 percent is what you can get without trying — but that means an application is likely to be out of action for over 7 days in a year.)

Realistically, there will be a pecking order of importance among the various business processes of an organization, and the IT availability demands will be greater for the most important ones — within the range of 98 percent to 99.999 percent.

Response times and customer satisfaction

A long time ago, before the age of the PC, IBM researched the productivity impact of having fast computer-to-human response times. Officially, a *response time* is measured as the time between giving a command to the computer (by a mouse click or by pressing Enter) and the information appearing on your screen in response.

Any response time that is less than one tenth of a second may as well be instant because we humans cannot distinguish the difference. Anything that is "sub-second" is good for productivity, and anything worse than that has a negative impact.

But actually, it's a little more complicated than that because experience also depends on expectation. Nevertheless, most applications nowadays shoot for a one-second response time or better. Whether an application can deliver such a response time consistently depends on the plumbing.

The same pecking order that we mention in the previous section is likely to apply between the various business processes that an organization has. More investment of time and money will be made to deliver consistent response times for the mission-critical business processes.

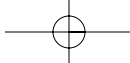
124 Part II: Nitty-Gritty SOA

Business service management

And so it came to pass that the vendors of plumbing products began to focus on delivering specific service levels to actual business processes. The much-hoped-for benefit was that organizations would be able to make their IT investments where they could make the most difference.

We have been at pains to point out that there is a great deal of plumbing underneath the floorboards of any IT network. From a service oriented architecture perspective, the aspects you need to be aware of concern

- ✔ **Monitoring service levels:** With SOA, service levels can and should be defined for all business services. There is no point in defining such service levels without monitoring them, so one service that the plumbing must provide is the monitoring of service levels.
- ✔ **Identifying faults and failures:** All faults and failures anywhere in the network, whether hardware or software related, have the potential to impact service levels for one or more business services. The plumbing products that provide this service are called *system management products*. They work by planting software “agents” on various computers to monitor local events, and these products also listen to and analyze network traffic in order to assemble a comprehensive picture of what is going on.
- ✔ **Fault management:** Depending on the actual circumstances, one of two different courses of action is taken if either a fault is detected or service monitoring information indicates that service levels are likely to be violated. Either the circumstances are reported to someone within the IT department — basically because the plumbing can’t deal with the situation all on its own — or the plumbing figures out what to do and does it. In the real world, far too much is still reliant on human intervention, but the movement toward self-healing, automated remediation is well underway.
- ✔ **Automatic provisioning/remedial action:** Some plumbing products can respond automatically to some situations. You can find provisioning software that can automatically make extra computer resources available to an application (and hence a business service). Such software manages a pool of servers, deploying them when needed and removing them from service when they are no longer required. Cluster management software can share an application across a cluster of servers and automatically compensate if one of them bites the dust. It can also balance application workloads across multiple servers.
- ✔ **Performance modeling and optimization:** Some plumbing products can gather information on the way application workloads are changing and model different ways of using the network’s resources to meet demands — both for specific applications and taking all applications into account. Such



a capability is important for planning future capacity needs of an organization. It's also important when failures occur, and you may have to run a degraded service for a while. Such products can guide you as to which service levels to let slip.

- ✓ **Management reporting:** Consolidation and reporting products can gather relevant information about the performance of all business services so that senior management can know the status of the IT service.

The SOA Supervisor

If you've read the entire chapter up to now, we're pretty confident that you've read everything you need to know about plumbing to be able to make sense of the SOA supervisor, so we'll just go ahead and put the icing on this particular cake. In a service oriented architecture, the SOA supervisor is, in fact, the chief plumber. It *orchestrates* the plumbing.

We can't overemphasize how important this role is. Back in Chapter 2, we define a service oriented architecture as

A software architecture for building applications that implement business processes or services by using a set of loosely coupled black-box components orchestrated to deliver a well-defined level of service.

The truth is that loose coupling is not as effective for delivering good service levels as tight coupling. The price paid in decreased effectiveness is worth paying, but it is still a price that has to be paid. That's why the SOA supervisor is important. It focuses on making sure that service levels are acceptable.

The SOA supervisor also has a great future. Right now, most businesses that are experimenting with SOA are implementing SOA in a limited way. They are not doing wall-to-wall SOA — and we don't recommend that they do. But as their ambitions expand, the job of the SOA supervisor expands because more applications are included, and the SOA supervisor simply has to keep more balls in the air.



Soon, some businesses will connect directly to other businesses through SOAs; it will be SOA to SOA. And at that point, the SOA supervisor is likely to take on legal obligations — because the service levels that it is charged with guaranteeing will be covered by legal contracts.

SOA supervising: The inside view

Take a look at how the SOA supervisor operates.



It is important to understand that what we describe here is the goal toward which IT vendors are moving SOA. In other words, it is an ideal that has yet to be fully achieved. But we'll go ahead and describe it anyway.

The first thing to say is that the SOA supervisor is active as long as any service within the SOA environment is operating. For all practical purposes, that means 24/7 with no holidays.

Our story begins with the service broker sending a message to the SOA supervisor saying that the service broker has in fact threaded together and started up yet another business process. We show the business process the broker is talking about, all happily threaded together, at the bottom of Figure 10-2.

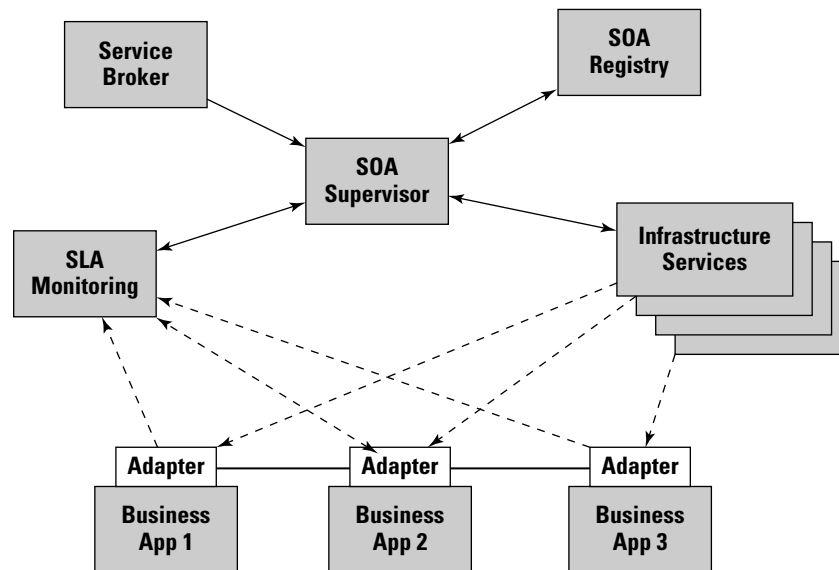


Figure 10-2:
The SOA supervisor.

The SOA supervisor immediately consults the SOA registry to get the details of the full business process so that it can set up monitoring software to monitor all the necessary components. It delegates the job of doing the monitoring to a utility that we have labeled SLA Monitoring, where the “SLA” stands for Service Level Agreement. This component now activates agents that are

local to the various application components shown at the bottom of the diagram. We represent the reporting of the performance of these components by the dotted lines.

The SLA Monitoring utility sends regular performance bulletins to the SOA supervisor, which passes them on so that they can be reported in real time to a console, which is probably manned by a bored operator. If everything is hunky-dory, and the service level agreements are being obeyed, that is all that happens.

However, should the SLA Monitoring utility provide information that indicates that some business process is running into problems and may be about to go belly-up, the SOA supervisor calls in the cavalry in the form of various infrastructure services that we describe earlier (for example, fault management and automatic provisioning). These services do what they can to save the day. This may involve a little load balancing or a little provisioning.

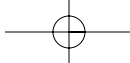
If a problem arises that needs some kind of human action, the potentially bored operator will probably be woken up by a nasty attention-grabbing noise and will be given detailed information about what the situation is and, possibly, where the problem lies.

Getting real

We don't want to mislead you into believing that the technology is more advanced than it actually is. We have described the role of the SOA supervisor as supervising the service level agreements (SLAs) and orchestrating an automatic response if any SLA violation looks likely to occur. Here's the reality in most enterprises today:

- ✓ Most enterprises don't have formally defined or well-defined SLAs.
- ✓ Most enterprises don't have a fully functional suite of infrastructure software.
- ✓ Most enterprises don't have maps of all their business processes and how they work.
- ✓ Most enterprises don't have a full inventory of all the computer equipment and software that they have deployed.

The truth is that most enterprises have quite a lot to do before they will be able to implement the kind of SOA supervision that we're describing here. It may seem like that puts a wrinkle in the works as regards SOA, but actually it doesn't. The truth is that, in most enterprises, the business applications are running reliably.



128 Part II: Nitty-Gritty SOA

As organizations begin to adopt SOA, they will do it gradually. They will slowly move toward the kind of environment that we are discussing here. Right now, applications run reliably because organizations buy a lot more computer hardware than they actually need for the workloads that they run. And to be honest, in many companies, the loyal folk in the IT trenches are putting in long hours, living with beepers, and having sporadic nervous breakdowns to keep everything copacetic. The performance of applications is not meticulously managed in the way that a SOA supervisor manages a business service.

It is also the case that no IT vendors can claim to have a complete SOA solution right now, and this is particularly the case as regards managing the plumbing so that it is both effective and efficient. There are sophisticated SOA registries and brokers and sophisticated enterprise service buses, and there are also some sophisticated SOA supervisors that do provide an SLA monitoring service. But right now, a comprehensive set of well-integrated infrastructure services and monitoring services that link to a SOA supervisor and cater to every kind of computer platform is not yet available anywhere. But it will be, in time.

