DZone Refcardz

CONTENTS INCLUDE:

Deployment Automation Patterns

By: James Betteley

ABOUT THIS REFCARD

Deploying software to a production environment is usually the final step in the product delivery lifecycle. In an ideal world, the deployment is simple, the experience is enjoyable, it works the first time, and we all go to the pub afterwards to celebrate yet another successful production deployment (yay!).
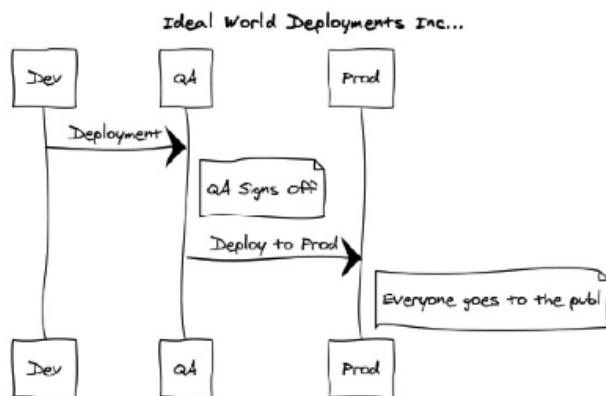
And now back to reality. Quite often, when we do production deployments, it's to fix something that's already broken, or we're releasing a project that's already overdue, or there is simply a great deal of pressure from the business to see the next great piece of functionality go live. The pressure is on, and all eyes are on you. To add to your list of problems, the deployment process is long winded, manually intensive, unreliable, and you've never done it before. You're staring down the barrel of an all-nighter, and you're already on your fifth cup of coffee.

Reality sucks.

But it doesn't have to be that way! With the application of some fairly simple good practices, production deployment can be just a formality. The only pressure you'll have is deciding who's buying the first round.

DEPLOYMENT PATTERNS

I will outline 7 patterns for software deployment to be regarded as generally reusable solutions to common issues within software deployment. Below is a table of patterns, and the common issues they mitigate.

Pattern	Mitigates
Automate deployments	Error prone manual deployments, unclear requirements, lack of auditability
The 5 Rs of application deployment	Time consuming deployments, high risk changes to production, human error, messy production systems, complicated roll-backs.
Standardize where you can	Repetition of similar tasks

Pattern	Mitigates
Make your deployments granular	Large scale deployments for small-scale changes
Treat configuration files as code	Configuration files being different on different environments ("but it works fine on my machine!")
Sanity test your deployments	Inconsistencies/bugs built into the deployment process
KISS!	Overly complicated production environments, troubleshooting nightmares!

PATTERN #1: AUTOMATION IS KEY

PATTERN: Automated deployments using tools and scripts

ANTI-PATTERN: Manual deployments by hand

It's too much to expect a person to manually deploy a complex software solution to an equally complex production environment time after time and never make a mistake. The odd mishap is what makes us human. So let's leave the machines to do the stuff they're good at: the repetitive, labor intensive tasks — tasks like deploying software!

So the first good practice is to automate software deployments. In its most simple form, this could mean simply writing a script to perform the deployments, or using a specific tool to do the leg-work for you.

But what about the environment we're deploying our applications on? How do they get deployed? If we're deploying our infrastructure changes by hand, then we're not fully leveraging the power of deployment automation.

Automating deployments brings with it a heap of other benefits as well, such as increased speed, greater reliability, and built-in audit trails. These benefits can be built into your automated software deployments with very little effort for maximum reward.
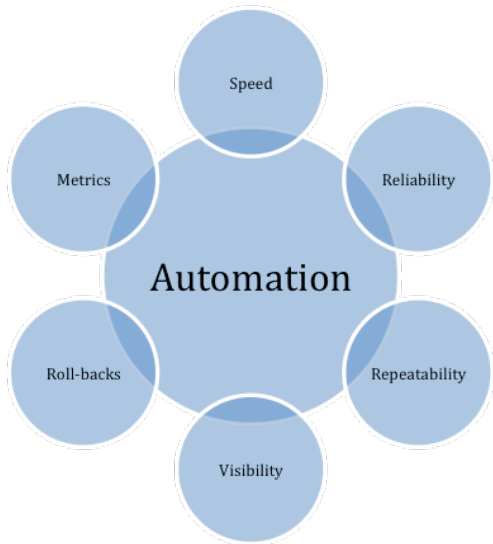
Figure 1 Automation lies at the center of good deployment practice

Why Should I Script It?

I know, it's simple. You could do it in your sleep, there's no need to write a script to do it, is there? Yes, there is.

Let's say you're deploying a jar file to a directory and changing a line in a config file. That's simple enough to do! But if you're doing it manually, then it's also simple enough to get it completely wrong. It's called human error and the best thing about it is you don't even know you're doing it.

Scripting your deployments gives you a nice cookie trail of what you've just done, so if things do go awry, you can look at the script and step through it. You can't replay and step through random human errors!

You can use just about any scripting language to script your software deployments, but again there are some good practices which should be brought into consideration.

- Verbosity – You probably don't want to have to read the world's most verbose scripts when you're troubleshooting your deployments.
- Clarity/readability – Pick a language that's readable and not ambiguous.
- Support – First thing you do when your script gives you an error? Google it, of course! The bigger the support community, the better (sometimes).
- Personal taste – Never overlook people's personal taste when it comes to choosing a scripting language. If the whole team want's to use Perl, then maybe they will just feel a lot happier using Perl!

Here's an example of a simple task which is quite commonly done at the beginning of a deployment – working out the free disk space. Three different scripting languages provide three fairly different scripts:

```
Shell :
        #!/bin/sh
        df –H | awk '{ print $5 " " $1 " " $6 }'

Ruby (First you need to install Ruby and ruby gems. Then sys–
filesystem):
        require 'rubygems'
        require 'sys/filesystem'
        include Sys
        stat = Sys::Filesystem.stat("/")
        mb_available = stat.block_size * stat.blocks_available
        / 1024 / 1024
        print mb_available
        print  "MB available!\n"

Perl:
        use Filesys::DiskFree;
        $value = new Filesys::DiskFree;
        $value->df();
        print "Available space ".$value->avail("/")." bytes\n";
        print "Total space ".$value->total("/")." bytes\n";
        print "Used space ".$value->used("/")." bytes\n";
```

Each script varies in its relative complexity and verbosity, and the output is subtly different from each one. In this very basic example, the Ruby script requires more effort to setup and write, while the shell script is very straightforward. However, deployment scripts are a lot more complex than this, and a slightly more elegant language like Ruby might come into its own depending on the requirements. Ultimately you need to choose a scripting language which is fit for purpose and which the users feel most comfortable with.

Automating Infrastructure Deployments

Deploying servers can be an onerous and highly manual task. Or you could automate it and make it a simple manual one! Thankfully, there are a number of tools available to help us do this. VMware is a popular choice of virtualization software, and can be used to deploy and configure anything from individual vms, to large vm farms. PowerCLI is a command line tool which allows us to automate these tasks. There's a wealth of information, code snippets and examples in the communities to help get you up and running. Here's an example of how to deploy a number of VMs from a single template, and apply some guest customizations:

```
$vmcsv = import–csv resources/vms.csv

ForEach ($line in $vmcsv){

New–VM –VMHost $line.vmhost –Name $line.vmName
–GuestCustomisation $line.guestCustomisation –Template $line.
template

}
```

The script reads from a CSV file containing information such as the host to deploy to, the new vm name to use, the guest customization to apply and the template to use. The CSV file will look similar to this:

```
host, name, customisation, template

esx01, mynewxpvm01, IE6, xpTemplate

esx01, mynewxpvm02, IE7, xpTemplate

esx02, mynewxpvm03, IE8, xpTemplate
```

The guest customization script can do numerous basic tasks such as setting time zones and registering the VM on a domain. However, we can automate even further to perform tasks such as installing software by using the PowerCLI script to invoke another script that resides on the template, and passing in relevant vm-specific parameters using PowerCLIs Invoke-VMscript.

If VMware and PowerCLI are good tools for vm deployment, then tools such as Chef, Puppet and CfEngine are great for configuring them. The question should not be whether or not to use them, rather, which one should I use?

Chef, Puppet and CFengine all provide automated scripted solutions for deploying applications, policies, accounts, services, etc. to your servers. Their underlying similarity is that they provide users with a centralized system for managing and deploying server configurations on top of your VM.

While the likes of CFengine, Chef and Puppet are focused on configuration, tools such as JumpBox, Capistrano and Fabric are geared more specifically to application deployment. The one thing they share in common is that they all provide automated solutions.

Electric Cloud's ElectricCommander is another automated solution that provisions physical, virtual and cloud infrastructure, automatically spinning up environments and decommissioning them when your tasks are completed.

Continuous Delivery and DevOps

With scripted, automated deployments we can expand our traditional continuous integration (CI) system to include software delivery. If our build passes all the unit tests then we can deploy it to a test environment, and with automated infrastructure deployments, we can even provision those environments automatically. Continuous delivery is the logical extension of continuous integration – if a build passes all the tests on the QA environment, then it's automatically deployed to a UAT environment. If it passes all the tests there, it could be automatically deployed to production.

This system is only made possible with automated deployments. The workflow of builds moving from development all the way through to production can be imagined as a pipeline – indeed, continuous delivery and release pipelines are becoming increasingly frequent bed-partners. Here's an example of a build pipeline in a continuous delivery system:

One of the key attributes in this system is the visible progression of a release from one stage (or environment) to the next. This is akin to a release workflow management process. As the build progresses along the pipeline, from left to right in the picture above, the release moves from development, through QA and UAT and into the hands of the Operations team. It's a seamless progression with no manual handover, and so the development and operations groups must be tightly coupled. This is the foundation of the DevOps movement – breaking down the traditional barriers between development and operations, and once again, automation is at the heart of it!

Naturally a number of tools are available to support this workflow. Their key attributes are:

- Workflow management

- Build tracking/pipelines

- Environment management/procurement

- Reporting

- Auditing

- Artifact management

Tools like Thoughtworks' **Go** provide much of the functionality mentioned above, but Go is generally focused on Continuous Integration. The pipeline visualization provides a certain degree of workflow management and the environment management functionality provides a one-stop-shop for tracking which agents are assigned to particular environments. Go actively encourages collaboration between development and operations; it's designed to be a central tool for developers and operations staff alike – for the developer it provides state of the art CI (with all the usual trimmings of test reporting, build metrics and so on) and for the Operations team it provides a good environment management interface, release tracking, and a simple UI for doing deployments (which can literally be just a click of a button).

Electric Cloud's **ElectricDeploy** is a purpose built software delivery tool built upon their ElectricCommander platform, which firmly embraces the DevOps culture. Like Thoughtworks, Electric Cloud encourages collaboration between development and operations, but they also ensure consistency and visibility across different environments and the whole application delivery lifecycle. ElectricDeploy enables you to take snapshots of your application versions and uses application and environment models to ensure consistent deployments across all environments and any type of infrastructure (physical, virtual or cloud). It leverages the pipeline concept and is unique in its ability to address failure management (allowing users to configure success/failure thresholds etc.).

This space is likely to become rich with enterprise tools as the DevOps movement gathers pace. Their key features are likely to be built on:

- collaboration between teams

- breaking down traditional barriers

- a focus on automation

- continuous delivery

- high visibility

PATTERN #2: THE 5 RS

PATTERN: Build the 5 Rs into your deployment process!

ANTI-PATTERN: Unreliable, slow, manual deployments that cause deployment engineers to lose their hair prematurely!

The 5 Rs of software deployment represent the principles we should follow when we design our software deployment processes, and the criteria we should consider when evaluating existing deployment tools.

Software deployments should be:

- Rapid
- Reliable
- Repeatable

And it should:

- Reduce Risk

And if all else fails:

- Roll-Back!

Make software deployments as rapid as possible. Don't deploy anything you don't have to deploy. Often, software is shipped with various different supporting applications, such as application/web servers and databases. If you are hosting your own application and you don't need to re-deploy your web server each time, then don't! Keep the size of your deployment artifact as small as possible to make the deployment as fast as it can be.

Deployments to production must be reliable. When we deploy to production, we should simply be repeating the exact same steps as were undertaken when the application was deployed to the dev-test environment, the QA environment, the UAT environment and the Pre-Prod environment. It should use the same deploy script and deploy the same artifact. The only difference is the environment to which we are deploying. The point is that by the time we deploy our artifact to production, we have tested the deployment process several times over, on several other environments, and we should now be confident in the reliability of our deployment scripts. We should never deploy artifacts by copying them from one environment to another - the artifacts could have undergone changes during testing, or someone could have edited the artifacts while they resided on a test environment.

Do this:
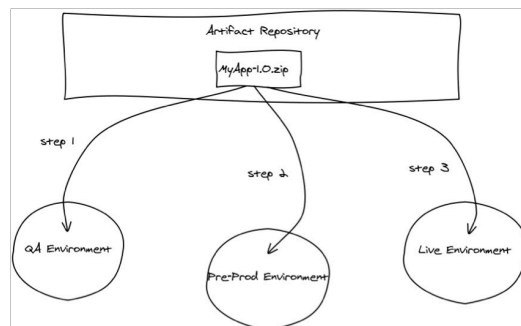
NOT this:
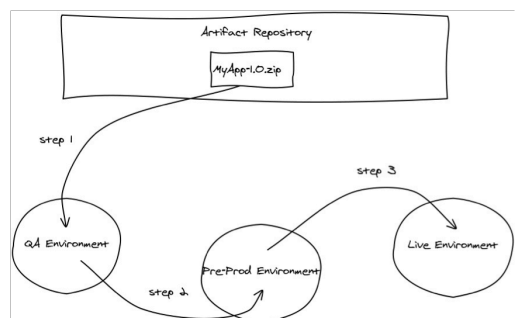
Deployments should be repeatable, that is to say that if we did a deployment of the same artifact 100 times, we should confidently expect to see the same result 100 times. If we deploy version 1.0.0 of MyApp to production, then do some changes to our deploy script and re-deploy 1.0.0 of MyApp, we could very easily see a different result. This is not repeatable. To prevent this situation from occurring, deployment scripts should be treated as code that is shipped with the application. If that script needs changing, then MyApp should be re-packaged, re-versioned and re-deployed to all the environments until it reaches production. The key here is that any new artifact is re-labeled with a new version number – you should never be able to build 2 different artifacts with the same name and version number. Another way of ensuring repeatability is to version the deploy scripts and use a configuration management process to bind a particular version of your application to the deploy script.

Another way of ensuring repeatability is to version the deploy scripts and use a configuration management process to bind a particular version of your application to the deploy script.

An effective versioning system is important for ensuring that no two different artifacts can ever have the same version number. Most Continuous Integration (CI) systems are capable of pulling in change list numbers from source control, as well as generating their own build numbers. Consider using a system that increments a build number every single time a build is initiated. For example:
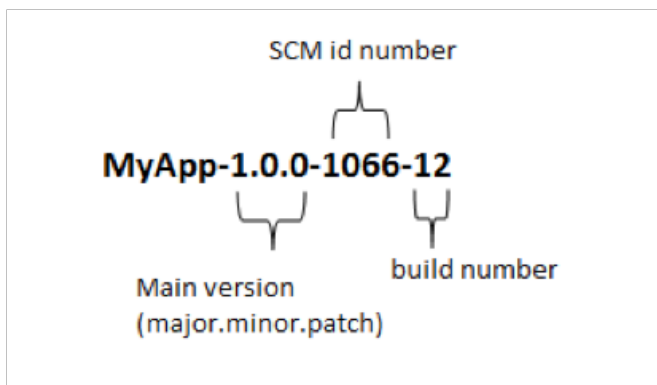
Every single time a build is created, the build number is guaranteed to increment, even if it's a rebuild. This ensures that provided this build is "signed off" on QA/UAT/Pre-Live then it will be identical when we deploy it to production. The SCM id number (usually a commit id) is useful for traceability – you can trace back to the actual code changes which caused the build, and see the areas of your application that have been changed.

Risk is reduced by eliminating human error, and this is done by automating tasks that are better suited to our trusty computers, as mentioned in pattern #1.

Roll-Backs are our safety net. If anything goes wrong, it is comforting to know that we can reliably restore to a previous working version with minimum fuss. This is where symlinks (symbolic links – analogous to shortcuts on Windows) are very popular. Deploy your application to a versioned directory and use a symlink to point to it. If the deployment fails, simply re-point the symlink to the older version. A note of warning though – don't leave too many old versions lying around on the system. They can get in the way, take up valuable space, and hamper troubleshooting.

PATTERN #3: STANDARDIZE WHERE YOU CAN

PATTERN: Emphasize convention over configuration. Look for common attributes in your deployments, and standardize their behavior using scripts. Encourage common deployment behavior within the application development teams.

ANTI-PATTERN: Writing a new deploy script for every application.

Do not continually re-invent the wheel when it comes to deployments. Many deployment tasks are fairly similar, they often involve copying one file or directory to an application server and starting it up. If this behavior is common to a number of applications, then extract it into a common file and use it across your different projects. Alternatively, use one of the numerous deployment tools available on the market - one might well fit your deployment needs. For example, Capistrano and Fabric provide ready-made deployment wrappers for deploying applications over ssh (Capistrano being Ruby while Fabric uses Python).

Sample Fabric script for deploying to a collection of servers:

```
from fabric.api import *
from __future__ import *

env.hosts = ['server1', 'server2', 'server3' ]

def deploy():
src_dir = '/my/src/dir'
with cd(src_dir):
        run("git pull")
```

Rather than use different deploy scripts for each application or for different environments, Fabric (as with Capistrano) allows you to reuse the same common logic to deploy applications to numerous different servers or environments simply by executing a script. In the example above, we're pulling down changes from git onto our 3 servers, but only executing the script once.

PATTERN #4: MAKE YOUR DEPLOYMENTS GRANULAR

PATTERN: Deploy the smallest module of your stack if there is a valid business need to do so.

ANTI-PATTERN: Deploying your whole IT stack just to make a single file change

If your production suite consists of numerous web applications, a database and an application server or two, then it's hard to justify deploying the whole stack just to be able to correct a spelling mistake on your homepage. Doing that would be time consuming, and potentially more risky.

Strange as it may sound, doing full stack deployments isn't as rare as you might think. Nevertheless, it does lack a lot of flexibility. It would be favorable to be able to deploy the smallest possible component, a single file for example – but there is a trade-off of course. The more granular your deployments are, the more effort it requires to deploy your whole stack, should the need arise.

What we must do is determine exactly how granular our deployments need to be. Are we more likely to deploy parts of our stack, or our whole stack? Usually it's the former. And we can break that down even further. We can go as far as deploying individual libraries if we need to as long as there's a strong enough business case for it, and the libraries are versioned.

Generally speaking, the smaller the deployment the easier it is to deploy, and the easier it is, the less risk there is involved. It is also quicker to deploy smaller components than a whole stack. One thing that we need to consider is the business requirement for our releases: do the customers need us to be able to deploy rapidly and frequently? If so, then the fact that small-sized deployments are generally less risky and more rapid should be taken into account. If we have no business requirement to deploy frequently and rapidly – for instance if the customers don't want any changes but want high service availability, then a quarterly or half-yearly release schedule might be more appropriate. Given these timescales, by the time a release comes around, it's likely that you're going to want to release a large proportion of your stack – in which case a full-stack deployment might be well suited. These are generalizations and not a firm rule – there are some very efficient companies who do full stack deployments on a frequent basis.

Requirement	Granularity
Daily Releases	Small - Individual jars, dlls
Weekly Releases	Medium - Small self-contained sites, small apps
Monthly Releases	Large - whole applications, large sites, whole stack

PATTERN #5: TREAT CONFIGURATION FILES AS CODE

PATTERN: Use tokens or placeholders in your config files (the dev "values" could actually act as the tokens themselves).

ANTI-PATTERN: Editing config files by hand in-situ.

Along with databases, config files are what make deployments really interesting. They're the one difference between a deployment to production and a deployment to any other environment. As such, we need to have a great deal of visibility and control over config file changes.

Application configuration files should be stored in a central repository, ideally alongside the application source code (although for various reasons this isn't always possible). They should never be edited in-situ. As with any source file, changes should be committed to source control and tested on each environment before going live. Editing config files on the production environment quickly leads to a maintenance nightmare, and you will find it hard to make any changes to the environment for fear of breaking something.

But how can you test config files when they're different for each environment? Well, actually, the files are often very similar on each environment. Only passwords, server names, connection strings and the like tend to differ, and these should be managed in the form of token substitution during the deployment process.

Use a single tokenized configuration file for development, and during your deployment process simply replace the tokens with the relevant value for the environment. An example:

This is a configuration file for a test environment:

```
<add key="DB:Connection" value="Server=TestServer;Initial
Catalog=TestDB;User id=Adminuser;password=pa55w0rd"/ >
```

The master version kept in source control could look like:

```
<add key="DB:Connection" value="Server=%DB_SERVER%;Initial
Catalog=%DB_NAME%;User id=%DB_UID%;password=%DB_PWD%"/ >
```

Then, during the deployment process, the deploy script will replace the tokens with the relevant values for the environment you are deploying to. Here's an example of a sed script which would replace the tokens in the example above:

```
s/%DB_SERVER%/TestServer/i

s/%DB_NAME%/TestDB/i

s/%DB_UID%/Adminuser/i

s/%DB_PWD%/pa55w0rd/i
```

In some cases using tokens is not convenient for development, and so the tokens themselves are replaced with actual development values. The deployment process then searches for and replaces these development values at deploy time.

Token substitution is the underlying mechanism by which many common deployment tools operate. They essentially map values to environments, and substitute the relevant values when you deploy to a particular environment. Octopus Deploy (an automated deployment solution for .Net applications) uses configuration file transforms and variable substitution to manage environmentally sensitive configuration settings such as connection strings and passwords.

PATTERN #6: SANITY TEST YOUR DEPLOYMENTS

PATTERN: Have a detailed expectation of what your deployed system should look like (this is the acceptance criteria) and test for it.

ANTI-PATTERN: Assuming what you have deployed is correct!

Before we do a production deployment, we really ought to know exactly how our system should look and behave once the deployment is complete. This should be treated as our "acceptance criteria" and unless we can prove that it has been met then the job isn't done.

Perhaps we might have a list of files and folders that we expect to see on our Live system when the deployment is done – in that case we can simply write a test to make sure what we get is what we are expecting.

Likewise we ought to know the md5 checksums of the binaries we deploy – this too can be tested once the deployment is complete. These tests should be automated and built into our deployment process. For example, we could write a verifier script which checks that all the directories we deployed have been granted the correct permissions. We could simply execute this verifier script at the end of our deployment.

Shell script to verify directory permissions (let's call it verifier.sh):

```
cat results/temp_dir_info.txt | grep -v drwxr-xr-x | grep rwxr-
-r-- > results/permission_report.txt

if cat results/temp_dir_info.txt | grep -v drwxr-xr-x | grep
rwxr--r-- > $null ; then

        echo "Some permissions don't match expectations"

        echo "Please check the results/permission_report.txt
        for more information"

        echo "----------------------------------------------
------------------------------------"

else

        echo "Congratulations!!! All files and folders have the
        expected permissions"

        echo "----------------------------------------------
------------------------------------"
fi
```

Executing the verifier at the end of our Fabric deployment script:

```
from fabric.api import *

from __future__ import *

env.hosts = ['server1', 'server2', 'server3' ]

def deploy():

src_dir = '/my/src/dir'

with cd(src_dir):

        run("git pull")

        run("verifier.sh")
```

Below is a table containing some common sanity checks that we might want to perform post-deployment.

Verification	Acceptance Criteria
Check correct versions of supporting infrastructure are present	OS version, the version of java/.NET installed on the target, the versions of IIS, Tomcat etc. all as expected. Machine architecture is correct (64 or 32 bit)
Check symlinks/shortcuts	Symlinks/shortcuts are successfully created Symlinks/shortcuts point to the correct location
Check web ports	Ports are listening/responding
Check permissions	Sites have correct permissions directories/files have the right permissions
Check the binaries	Binaries are versioned correctly
Check the files & directories	File and directory listing matches the expected list (or matches the QA/UAT environments)

PATTERN #7: KISS: KEEP IT SIMPLE, SYSADMINS!

PATTERN: Keep your deployment architecture as simple as possible.

ANTI-PATTERN: Overdoing the symlinks/shortcuts and leaving too many old files/directories lying around on the production system.

When we get overly familiar with our systems and applications, it becomes easy to see past their complexity. But to an outsider the complexity can be blinding. As production sysadmins we tend to know the Live environment better than we know the backs of our own hands, but wouldn't it be better if we just didn't have to know all this useless complexity?

Troubleshooting complex production environments can be very frustrating. When the live site is down and you're trying to diagnose the problem, finding symlinks more than 2 deep should be a crime punishable by law! It's overly complex and usually 100% avoidable. In deployment scripts, using variables where variables are not needed just makes it harder to follow. Likewise, writing conditionals for every edge case you can think of is not practical in a deploy script.

Here's an example of pointless variables in a deploy script:

```
12    <property name="java_version" value="1.6.0_07" />
13    <property name="java" value="/opt/java/jdk${java_version}/bin/java"/>
14    <property name="psql_location" value="/usr/local/bin/"/>
15    <property name="psql_exec" value="${psql_location}/psql"/>
```

Lines 12 and 14 might have seemed a good idea when writing the script, indeed they may look sensible, but in fact their values are only used once in the remainder of the script (lines 13 and 15), so we might as well simply replace this with:

```
12    <property name="java" value="/opt/java/jdk1.6.0_07/bin/java"/>
13    <property name="psql_exec" value="/usr/local/bin/psql"/>
```

So we've reduced the size of our sample by 50% just by rationalizing our use of variables! This might not look like much in isolation, but it could mean the difference between troubleshooting a 1000 line deploy script and a 500 line script.

Symbolic links can be useful and frustrating in equal measure. They act as shortcuts to other locations on Unix based systems, so for instance we can have a link to Java in our application's home directory, and our application's configuration file need only point to its home directory to find the java installation – the symlink will do the rest. However, using symlinks can quickly get out of hand, and we can end up chasing around the system just to find our java installation (for instance). When you're troubleshooting a production issue, this can be unimaginably frustrating!

```
java > /usr/bin/java_latest > /usr/bin/java/jdk1.6 > /usr/bin/
java/jdk1.6.0_24/bin > GO TO JAIL!
```

Keep your production system as clean as possible. When doing deployments, don't leave any remnants of the last build on the file system. Either deploy to a clean directory and use symlinks to point to the "latest" version, or delete (or move) the existing files in their entirety. Trying to do complex partial upgrades is often more trouble than it is worth.

Not only the production system needs to be kept as simple as possible – the deployment process needs to be simple as well.

Make your deployment processes even easier by putting a front end on your deployment scripts, eliminating room for human error. It's becoming increasingly popular to use CI tools to trigger production deployments; or you could easily write a simple interface to drive your deploy scripts. Alternatively there are numerous tools available (open source as well as enterprise) which can be used to drive your deployments.

ABOUT THE AUTHOR

For the last 10 years James Betteley has been a keen believer in the benefits of automation. When he's not busy in his day job as a Change and Release Manager, you might find James blogging or speaking at various DevOps and Continuous Delivery talks around London.

RECOMMENDED BOOK
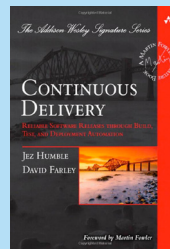
The authors introduce state-of-the-art techniques, including automated infrastructure management and data migration, and the use of virtualization. For each, they review key issues, identify best practices, and demonstrate how to mitigate risks. Whether you're a developer, systems administrator, tester, or manager, this book will help your organization move from idea to release faster than ever—so you can deliver value to your business rapidly and reliably.

BUY HERE
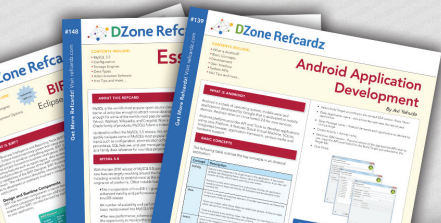
DZone communities deliver over 6 million pages each month to more than 3.3 million software developers, architects and decision makers. DZone offers something for everyone, including news, tutorials, cheat sheets, blogs, feature articles, source code and more. **""DZone is a developer's dream","** says PC Magazine.