

# Table of Contents

[The Linux Shell Scripting Handbook -](#The_Linux_Shell_Scripting_Handbo)

[Chapter Summary..........................................................16](#Chapter_Summary)

[Chapter Summary..........................................................21](#Chapter_Summary_1)

[Chapter Summary..........................................................27](#Chapter_Summary_2)

[Chapter Summary..........................................................32](#Chapter_Summary_3)

[Chapter Summary..........................................................37](#Chapter_Summary_4)

[Chapter Summary..........................................................42](#Chapter_Summary_5)

[Chapter Summary..........................................................46](#Chapter_Summary_6)

[Chapter Summary..........................................................51](#Chapter_Summary_7)

[Chapter Summary..........................................................55](#Chapter_Summary_8)

[Chapter Summary..........................................................59](#Chapter_Summary_9)

[Chapter Summary..........................................................63](#Chapter_Summary_10)

[Chapter Summary..........................................................68](#Chapter_Summary_11)

[Chapter Summary..........................................................72](#Chapter_Summary_12)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_1)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_2)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_3)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_4)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_5)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_6)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_7)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_8)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_9)

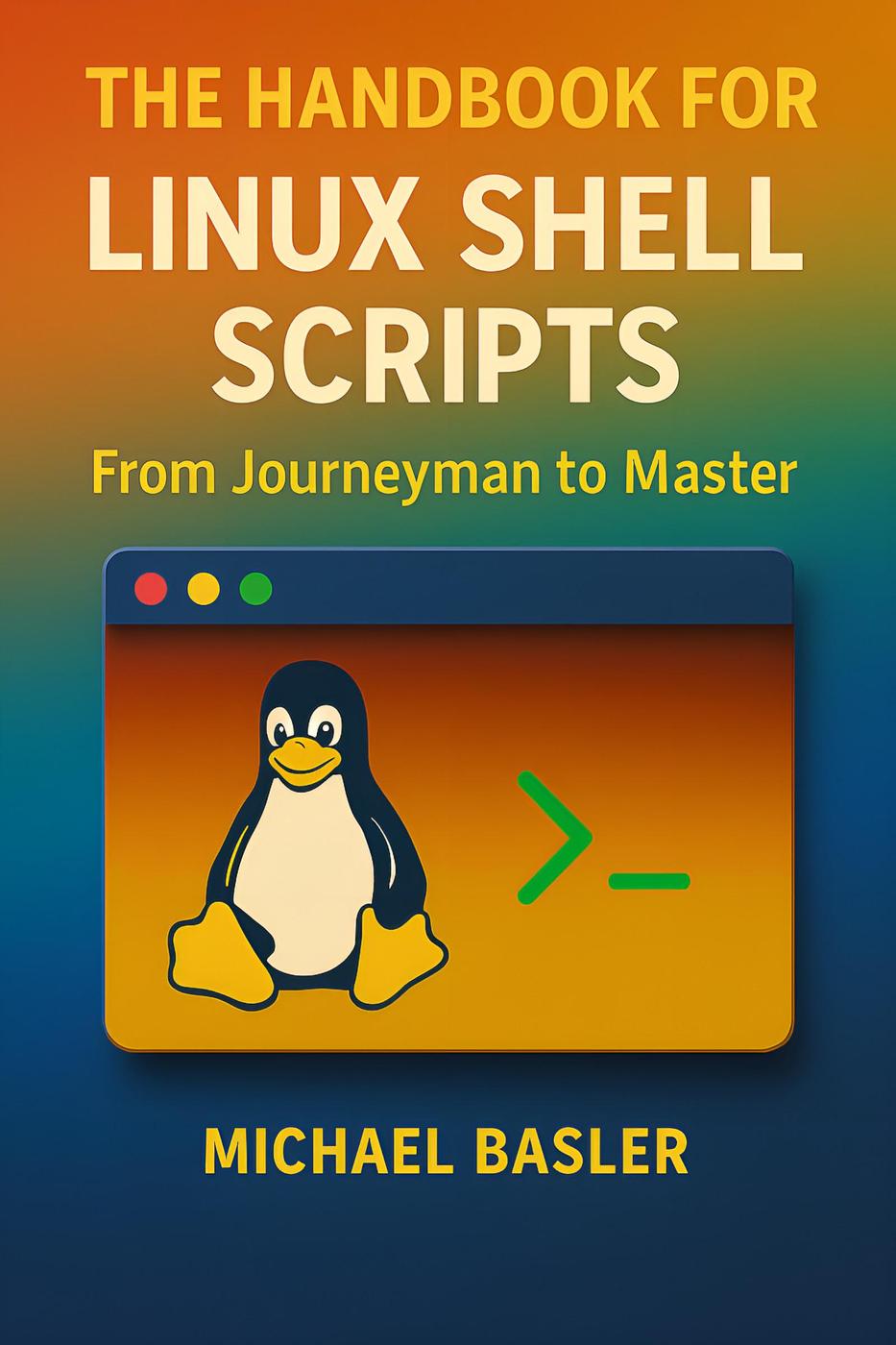
[CHAPTER SUMMARY](#CHAPTER_SUMMARY_10)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_11)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_12)

[CHAPTER SUMMARY](#CHAPTER_SUMMARY_13)

[This project forces you to use almost every concept from this book and](#This_project_forces_you_to_use_a)





# The Linux Shell Scripting Handbook -

From Journeyman to Master

Michael Basler

Published by Michael Basler, 2025.

While every precaution has been taken in the preparation of this book, the publisher assumes no responsibility for errors or omissions, or for damages resulting from the use of the information contained herein. THE LINUX SHELL SCRIPTING HANDBOOK - FROM JOURNEYMAN TO MASTER

First edition. July 1, 2025.

Copyright © 2025 Michael Basler.

Written by Michael Basler.

# Also by Michael Basler

Erste Auflage

[Das Linux Terminal für Fortgeschrittene - Die Kommandozeile leicht](https://www.draft2digital.com/catalog/2859007?distributor=kobo)

[gemacht](https://www.draft2digital.com/catalog/2859007?distributor=kobo)

[Das Linux-Terminal für Einsteiger - Die Kommandozeile leicht gemacht](https://www.draft2digital.com/catalog/2811420?distributor=kobo)

[Linux Shell Scripting – Ein Leitfaden für Anfänger](https://www.draft2digital.com/catalog/2815868?distributor=kobo)

[Switch to Linux – 15 wichtige Tipps](https://www.draft2digital.com/catalog/2875256?distributor=kobo)

First Edition

[The Linux Terminal for Beginners - The Command Line Made Easy](https://www.draft2digital.com/catalog/2812701?distributor=kobo)

[Linux Shell Scripting - A Beginner's Guide](https://www.draft2digital.com/catalog/2814611?distributor=kobo)

[The Linux Terminal for Advanced Users - The Command Line Made Easy](https://www.draft2digital.com/catalog/2858969?distributor=kobo)

[Switch to Linux – 15 Important Tips](https://www.draft2digital.com/catalog/2879783?distributor=kobo)

Standalone

[Die Asiatische Küche Entdecken - Die 10 Besten Rezepte aus den](https://www.draft2digital.com/catalog/2913272?distributor=kobo)

[Philippinen](https://www.draft2digital.com/catalog/2913272?distributor=kobo)

[Discover Asian Cuisine - The 10 Best Recipes from the Philippines](https://www.draft2digital.com/catalog/2913289?distributor=kobo)

[Die Asiatische Küche Entdecken - Die 10 Besten Rezepte aus Afghanistan](https://www.draft2digital.com/catalog/2988144?distributor=kobo)

[Discover Asian Cuisine - The 10 Best Recipes from Afghanistan](https://www.draft2digital.com/catalog/2988168?distributor=kobo)

[Das Linux Shell Skripter Handbuch - Vom Gesellen zum Meister](https://www.draft2digital.com/catalog/3001352?distributor=kobo)

The Linux Shell Scripting Handbook - From Journeyman to Master

# TABLE OF CONTENTS

[Introduction: From Craftsman to Architect.............................................5](#INTRODUCTION__FROM_CRAFTSMAN)

[What to Expect in This Book...................................................5](#WHAT_TO_EXPECT_IN_THIS_BOOK)

[How to Use This Book........................................................6](#HOW_TO_USE_THIS_BOOK)

[1. Mastering the Shell Environment..................................................7](#1__MASTERING_THE_SHELL)

[Introduction: From a Configured System to a Controlled System.......................7](#INTRODUCTION__FROM_A)

[Learning Objectives for This Chapter:............................................7](#LEARNING_OBJECTIVES_FOR_THIS)

[1.1 The Big Picture: Interactive, Non-Interactive, Login, and Non-](#1_1_THE_BIG_PICTURE__INTERACTIVE)

[Login..............8](#1_1_THE_BIG_PICTURE__INTERACTIVE)

[1.2 The Hierarchy of Startup Files............................................8](#1_2_THE_HIERARCHY_OF_STARTUP)

[1.3 The Art of Control: set and shopt..........................................9](#1_3_THE_ART_OF_CONTROL__SET_AND)

[1.4 Using Environment Variables Correctly: The Difference Between](#1_4_USING_ENVIRONMENT_VARIABLES)

[VAR=value and export VAR=value......9](#1_4_USING_ENVIRONMENT_VARIABLES)

[1.5 Hands-on Workshop: Building a Professional .bashrc.........................10](#1_5_HANDS_ON_WORKSHOP__BUILDING)

[Chapter Summary..........................................................11](#CHAPTER_SUMMARY)

[Exercises.................................................................11](#EXERCISES)

[2. POSIX Compliance and Portability – Writing Universal](#2__POSIX_COMPLIANCE_AND)

[Scripts.........................12](#2__POSIX_COMPLIANCE_AND)

[Introduction: The Language Every System Understands............................12](#INTRODUCTION__THE_LANGUAGE)

[Learning Objectives for This Chapter:...........................................12](#LEARNING_OBJECTIVES_FOR_THIS_1)

[2.1 The Shebang Showdown: #!/bin/sh vs. #!/bin/bash...........................13](#2_1_THE_SHEBANG_SHOWDOWN)

[2.2 The POSIX Toolbox: Common Bash-isms and Their](#2_2_THE_POSIX_TOOLBOX__COMMON)

[Alternatives...............13](#2_2_THE_POSIX_TOOLBOX__COMMON)

[2.3 The Unsung Hero: shellcheck............................................14](#2_3_THE_UNSUNG_HERO__SHELLCHECK)

[2.4 The Pragmatic Scripter: When is Portability Overkill?.........................15](#2_4_THE_PRAGMATIC_SCRIPTER)

[Chapter Summary..........................................................16](#CHAPTER_SUMMARY_1)

[Exercises.................................................................16](#EXERCISES_1)

[3. Advanced Parameter Expansion and Quoting – The Shell’s Swiss](#3__ADVANCED_PARAMETER)

[Army Knife.............17](#3__ADVANCED_PARAMETER)

[Introduction: Operating in the Heart of the Shell..................................17](#INTRODUCTION__OPERATING_IN_THE)

[Learning Objectives for This Chapter:...........................................17](#LEARNING_OBJECTIVES_FOR_THIS_2)

[3.1 Substring Removal: The Scalpel for Strings.................................18](#3_1_SUBSTRING_REMOVAL__THE)

[3.2 Search and Replace: sed for In-house Use..................................18](#3_2_SEARCH_AND_REPLACE__SED_FOR)

[3.4 The Master Class: Indirect Expansion and More.............................19](#3_4_THE_MASTER_CLASS__INDIRECT)

[3.5 Quoting – The Safety Net That Must Always Be Up..........................20](#3_5_QUOTING___THE_SAFETY_NET)

[Chapter Summary..........................................................21](#CHAPTER_SUMMARY_2)

[Exercises.................................................................21](#EXERCISES_2)

[4. Functions, Libraries, and Scope – The Path to Modular](#4__FUNCTIONS__LIBRARIES__AND)

[Code...........................22](#4__FUNCTIONS__LIBRARIES__AND)

[Introduction: From Monolithic Scripts to Reusable Tools...........................22](#INTRODUCTION__FROM_MONOLITHIC)

[Learning Objectives for This Chapter:...........................................22](#LEARNING_OBJECTIVES_FOR_THIS_3)

[4.1 The Anatomy of a Professional Function...................................23](#4_1_THE_ANATOMY_OF_A)

[4.2 Scope: The Most Important Rule for Stable Functions.........................24](#4_2_SCOPE__THE_MOST_IMPORTANT)

[4.3 Script Libraries: Organizing and Reusing Code..............................25](#4_3_SCRIPT_LIBRARIES__ORGANIZING)

[4.4 For Advanced Users: Passing Arrays to Functions............................26](#4_4_FOR_ADVANCED_USERS__PASSING)

[Chapter Summary..........................................................27](#CHAPTER_SUMMARY_3)

[Exercises.................................................................27](#EXERCISES_3)

[5. Robust Error Handling and Debugging – When Things Go](#5__ROBUST_ERROR_HANDLING_AND)

[Wrong.......................28](#5__ROBUST_ERROR_HANDLING_AND)

[Introduction: From a Crash to a Controlled Emergency Landing......................28](#INTRODUCTION__FROM_A_CRASH_TO)

[Learning Objectives for This Chapter:...........................................28](#LEARNING_OBJECTIVES_FOR_THIS_4)

[5.1 The Language of Failure: Mastering Exit Codes.............................29](#5_1_THE_LANGUAGE_OF_FAILURE)

[5.2 The Safety Net: Cleaning Up with trap EXIT................................29](#5_2_THE_SAFETY_NET__CLEANING_UP)

[5.3 Proactive Error Detection: trap ERR.......................................30](#5_3_PROACTIVE_ERROR_DETECTION)

[5.4 Advanced Debugging: Beyond set -x......................................30](#5_4_ADVANCED_DEBUGGING__BEYOND)

[5.5 Professional Logging Strategies..........................................31](#5_5_PROFESSIONAL_LOGGING)

[Chapter Summary..........................................................32](#CHAPTER_SUMMARY_4)

[Exercises.................................................................32](#EXERCISES_4)

[6. Mastering Arrays and Data Structures – More Than Just a](#6__MASTERING_ARRAYS_AND_DATA)

[List..........................33](#6__MASTERING_ARRAYS_AND_DATA)

[Introduction: From a Single Value to a Structured Collection........................33](#INTRODUCTION__FROM_A_SINGLE)

[Learning Objectives for This Chapter:...........................................33](#LEARNING_OBJECTIVES_FOR_THIS_5)

[6.1 Indexed Arrays: The Professional Approach................................34](#6_1_INDEXED_ARRAYS__THE)

[6.2 Associative Arrays: The Heart of Structured Data............................34](#6_2_ASSOCIATIVE_ARRAYS__THE)

[6.3 Hands-on Workshop: Building a Configuration Parser........................35](#A_CONFIGURATION_PARSER)

[6.4 Simulating Complex Data Structures......................................35](#6_4_SIMULATING_COMPLEX_DATA)

[Chapter Summary..........................................................37](#CHAPTER_SUMMARY_5)

[Exercises.................................................................37](#EXERCISES_5)

[7. Process Management and Parallelization – Become the](#7__PROCESS_MANAGEMENT_AND)

[Conductor.......................38](#7__PROCESS_MANAGEMENT_AND)

[Introduction: From Serial Execution to Parallel Orchestration........................38](#INTRODUCTION__FROM_SERIAL)

[Learning Objectives for This Chapter:...........................................38](#LEARNING_OBJECTIVES_FOR_THIS_6)

[7.1 More Than Just &: The Art of Job Control in Scripts..........................39](#7_1_MORE_THAN_JUST____THE_ART_OF)

[7.2 Process Substitution: When Pipes Aren’t Enough............................39](#7_2_PROCESS_SUBSTITUTION__WHEN)

[7.3 Coprocesses (coproc): Two-Way Communication............................40](#7_3_COPROCESSES__COPROC___TWO)

[7.4 Hands-on Workshop: Speeding Up Scripts with Parallelization..................40](#7_4_HANDS_ON_WORKSHOP__SPEEDING)

[Chapter Summary..........................................................42](#CHAPTER_SUMMARY_6)

[Exercises.................................................................42](#EXERCISES_6)

[8. The Triumvirate: grep, sed, and awk Revisited.......................................43](#8__THE_TRIUMVIRATE__GREP__SED)

[Introduction: From Pocket Knife to Power Saw...................................43](#INTRODUCTION__FROM_POCKET)

[Learning Objectives for This Chapter:...........................................43](#LEARNING_OBJECTIVES_FOR_THIS_7)

[8.1 grep: The Master of Finding.............................................44](#8_1_GREP__THE_MASTER_OF_FINDING)

[8.2 sed: The Stream Editor, Unleashed........................................44](#8_2_SED__THE_STREAM_EDITOR)

[8.3 awk: The Data-Processing Powerhouse....................................45](#8_3_AWK__THE_DATA_PROCESSING)

[8.4 The Art of the Pipeline: Combining Strengths...............................45](#8_4_THE_ART_OF_THE_PIPELINE)

[Chapter Summary..........................................................46](#CHAPTER_SUMMARY_7)

[Exercises.................................................................46](#EXERCISES_7)

[9. Interacting with Modern Web APIs and Data Formats.................................47](#9__INTERACTING_WITH_MODERN_WEB)

[Introduction: The Shell as a Web Client.........................................47](#INTRODUCTION__THE_SHELL_AS_A)

[Learning Objectives for This Chapter:...........................................47](#LEARNING_OBJECTIVES_FOR_THIS_8)

[9.1 curl: Your Swiss Army Knife for the Web..................................48](#9_1_CURL__YOUR_SWISS_ARMY_KNIFE)

[9.2 Introduction to jq: The sed/awk for JSON..................................48](#9_2_INTRODUCTION_TO_JQ__THE)

[9.3 Advanced jq Expressions...............................................49](#9_3_ADVANCED_JQ_EXPRESSIONS)

[9.4 Hands-on Workshop: Building a GitHub API Client..........................49](#9_4_HANDS_ON_WORKSHOP__BUILDING)

[9.5 A Brief Look at XML and xmlstarlet......................................50](#9_5_A_BRIEF_LOOK_AT_XML_AND)

[Chapter Summary..........................................................51](#CHAPTER_SUMMARY_8)

[Exercises.................................................................51](#EXERCISES_8)

[10. Secure Scripting Practices – Build a Fortress, Not a](#10__SECURE_SCRIPTING_PRACTICES)

[Shed.............................52](#10__SECURE_SCRIPTING_PRACTICES)

[Introduction: From “Does It Work?” to “Is It Secure?”.............................52](#INTRODUCTION__FROM__DOES_IT)

[Learning Objectives for This Chapter:...........................................52](#LEARNING_OBJECTIVES_FOR_THIS_9)

[10.1 The First Commandment: Never Store Secrets in a Script.....................53](#10_1_THE_FIRST_COMMANDMENT)

[10.2 The Plague of Command Injection.......................................53](#10_2_THE_PLAGUE_OF_COMMAND)

[10.3 The Minefield of Temporary Files: mktemp is Your Friend....................54](#10_3_THE_MINEFIELD_OF_TEMPORARY)

[10.4 Permissions and the Principle of Least Privilege............................54](#10_4_PERMISSIONS_AND_THE)

[Chapter Summary..........................................................55](#CHAPTER_SUMMARY_9)

[Exercises.................................................................55](#EXERCISES_9)

[11. Scripting the Network – The Shell as a Network](#11__SCRIPTING_THE_NETWORK___THE)

[Engineer.............................56](#11__SCRIPTING_THE_NETWORK___THE)

[Introduction: From Localhost to the Global Network...............................56](#INTRODUCTION__FROM_LOCALHOST)

[Learning Objectives for This Chapter:...........................................56](#LEARNING_OBJECTIVES_FOR_THIS_10)

[11.1 The Network’s Pulse: Querying Services with netcat and nmap................57](#11_1_THE_NETWORK_S_PULSE)

[11.2 Automation with SSH: The Keys to the Kingdom...........................57](#11_2_AUTOMATION_WITH_SSH__THE)

[11.3 Secure Data Transfer: scp and rsync......................................58](#11_3_SECURE_DATA_TRANSFER__SCP)

[11.4 SSH Tunneling: A Secure Corridor Through the Insecure Net..................58](#11_4_SSH_TUNNELING__A_SECURE)

[Chapter Summary..........................................................59](#CHAPTER_SUMMARY_10)

[Exercises.................................................................59](#EXERCISES_10)

[12. Performance Tuning Your Scripts – From a Tractor to a Race](#12__PERFORMANCE_TUNING_YOUR)

[Car......................60](#12__PERFORMANCE_TUNING_YOUR)

[Introduction: When “It Works” Is No Longer Fast Enough..........................60](#INTRODUCTION__WHEN__IT_WORKS)

[Learning Objectives for This Chapter:...........................................60](#LEARNING_OBJECTIVES_FOR_THIS_11)

[12.1 Measure, Don’t Guess: Finding Bottlenecks with time and](#12_1_MEASURE__DON_T_GUESS)

[strace...............61](#12_1_MEASURE__DON_T_GUESS)

[12.2 The #1 Performance Killer: The Cost of Forking............................61](#12_2_THE__1_PERFORMANCE_KILLER)

[12.3 Practical Optimization Techniques.......................................62](#12_3_PRACTICAL_OPTIMIZATION)

[12.4 I/O Optimization: Reading Data Smartly..................................62](#12_4_I_O_OPTIMIZATION__READING)

[12.5 Knowing When to Stop: The Limits of the Shell............................63](#12_5_KNOWING_WHEN_TO_STOP__THE)

[Chapter Summary..........................................................63](#CHAPTER_SUMMARY_11)

[Exercises.................................................................63](#EXERCISES_11)

[13. Creating Professional Command-Line Tools – More Than Just a](#13__CREATING_PROFESSIONAL)

[Script..................64](#13__CREATING_PROFESSIONAL)

[Introduction: From a Private Helper to a Public Utility..............................64](#INTRODUCTION__FROM_A_PRIVATE)

[Learning Objectives for This Chapter:...........................................64](#LEARNING_OBJECTIVES_FOR_THIS_12)

[13.1 Argument Parsing: getopts Instead of $1 Chaos.............................65](#13_1_ARGUMENT_PARSING__GETOPTS)

[13.2 Standard Options:—help,—version, and Friends............................66](#13_2_STANDARD_OPTIONS____HELP)

[13.3 Professional Project Structure...........................................66](#13_3_PROFESSIONAL_PROJECT)

[13.4 The Art of Documentation: Writing man Pages.............................67](#13_4_THE_ART_OF_DOCUMENTATION)

[Chapter Summary..........................................................68](#CHAPTER_SUMMARY_12)

[Exercises.................................................................68](#EXERCISES_12)

[14. Integrating with a Wider Ecosystem – The Shell as the Engine of](#14__INTEGRATING_WITH_A_WIDER)

[Automation.............69](#14__INTEGRATING_WITH_A_WIDER)

[Introduction: From a Tool to an Automated Factory................................69](#INTRODUCTION__FROM_A_TOOL_TO)

[Learning Objectives for This Chapter:...........................................69](#LEARNING_OBJECTIVES_FOR_THIS_13)

[14.1 The Local Guardians: Quality Assurance with Git Hooks.....................70](#14_1_THE_LOCAL_GUARDIANS)

[14.2 The Assembly Line: Shell Scripts in CI/CD Pipelines........................70](#14_2_THE_ASSEMBLY_LINE__SHELL)

[14.3 The Packaging: Shell Scripts in Docker Containers..........................71](#14_3_THE_PACKAGING__SHELL)

[Chapter Summary..........................................................72](#CHAPTER_SUMMARY_13)

[Exercises.................................................................72](#EXERCISES_13)

[15. Extending the Shell and Looking Ahead – The Master’s](#15__EXTENDING_THE_SHELL_AND)

[Horizon.......................73](#15__EXTENDING_THE_SHELL_AND)

[Introduction: The Journey is Complete, The Learning Continues......................73](#INTRODUCTION__THE_JOURNEY_IS)

[Learning Objectives for This Chapter:...........................................73](#LEARNING_OBJECTIVES_FOR_THIS_14)

[15.1 A Look Beyond the Horizon: zsh and fish.................................74](#15_1_A_LOOK_BEYOND_THE_HORIZON)

[15.2 The Art of Knowing When to Stop: When the Shell is the Wrong](#15_2_THE_ART_OF_KNOWING_WHEN_TO)

[Tool..........74](#15_2_THE_ART_OF_KNOWING_WHEN_TO)

[15.3 The Right Tool for the Job: Python, Go, and Beyond.........................75](#15_3_THE_RIGHT_TOOL_FOR_THE_JOB)

[15.4 Final Project: A Modular Deployment Tool................................75](#15_4_FINAL_PROJECT__A_MODULAR)

[Summary and Final Words...................................................76](#SUMMARY_AND_FINAL_WORDS)

[A Final Word from the Author.....................................................77](#A_FINAL_WORD_FROM_THE_AUTHOR)

[A Collection of Useful Shell One-Liners.............................................78](#A_COLLECTION_OF_USEFUL_SHELL)

[File and Directory Management...............................................78](#FILE_AND_DIRECTORY_MANAGEMENT)

[Text processing and data wrangling.............................................78](#TEXT_PROCESSING_AND_DATA)

[Network and Process Management.............................................79](#NETWORK_AND_PROCESS)

[Index.........................................................................80](#INDEX)

[Symbols and Operators......................................................85](#SYMBOLS_AND_OPERATORS)

[Appendix A: Shell Startup File Flowchart (Bash)....................................86](#APPENDIX_A__SHELL_STARTUP_FILE)

[Appendix B: POSIX Portability Quick Reference (English)............................87](#APPENDIX_B__POSIX_PORTABILITY)

[Appendix C: Parameter Expansion Quick Reference..................................88](#APPENDIX_C__PARAMETER)

[Appendix D: Professional Bash Function Template...................................89](#APPENDIX_D__PROFESSIONAL_BASH)

[Appendix E: Useful Variables and Signals for Error Handling..........................90](#APPENDIX_E__USEFUL_VARIABLES)

[Appendix F: Array Operations Cheat Sheet.........................................91](#APPENDIX_F__ARRAY_OPERATIONS)

[Appendix G: Process Management Quick Reference..................................92](#APPENDIX_G__PROCESS)

[Appendix H: grep vs. sed vs. awk – Which Tool to Use?..............................93](#APPENDIX_H__GREP_VS__SED_VS__AWK)

[Appendix I: jq Filter Quick Reference.............................................94](#APPENDIX_I__JQ_FILTER_QUICK)

[Appendix J: Shell Script Security Checklist.........................................95](#APPENDIX_J__SHELL_SCRIPT)

[Appendix K: Network Scripting Command Quick Reference...........................96](#APPENDIX_K__NETWORK_SCRIPTING)

[Appendix L: Performance Anti-Patterns and Their Solutions...........................97](#APPENDIX_L__PERFORMANCE_ANTI)

[Appendix M: Boilerplate for a Professional Command-Line](#APPENDIX_M__BOILERPLATE_FOR_A)

[Tool.......................98](#APPENDIX_M__BOILERPLATE_FOR_A)

[Appendix N: Best Practices for Scripts in Automation](#APPENDIX_N__BEST_PRACTICES_FOR)

[Ecosystems.....................100](#APPENDIX_N__BEST_PRACTICES_FOR)

[Appendix O: Decision Matrix: Shell Script vs. Higher-Level](#APPENDIX_O__DECISION_MATRIX)

[Language..................102](#APPENDIX_O__DECISION_MATRIX)

INTRODUCTION: FROM CRAFTSMAN

TO ARCHITECT

Congratulations! You are holding the next stage of your journey into the world of the Linux shell. If you have worked through the first book, you already command the fundamentals: you can chain commands, automate simple tasks, and harness the power of pipes and redirection to shape data. You have learned to handle the tools of a craftsman—the hammer (rm), the saw (sed), and the screwdriver (grep)—to create functional and useful things.

But true mastery in any craft lies not just in that a piece of work

functions, but in how it functions. Is it robust enough to withstand unexpected stress? Is it built with such precision that it fits into other environments? Is it safe to use and easy to maintain?

This is precisely where this book begins. We are taking the decisive step

from craftsman to architect. An architect doesn’t just think about a single wall; they think about the entire building. They plan the foundation, select the right materials, create a detailed blueprint, and consider all the safety regulations.

Translated to the world of shell scripting, this means:

• Robustness: Your script won’t collapse at the first sign of unexpected input. It will have clean error handling and will clean up after itself. • Portability: Your script won’t just run on your machine, but also in minimalist Docker containers, on macOS, or on legacy enterprise servers. • Maintainability: Your code will be so clearly structured that you (or your colleagues) can still understand and extend it six months from now.

• Security: You will know how to protect secrets, prevent command injection, and operate with the principle of least privilege. • Performance: You will understand which operations are expensive and how to identify and resolve bottlenecks in your scripts.

# WHAT TO EXPECT IN THIS BOOK

We will sharpen the tools you already know and add a range of professional techniques and new utilities to your toolkit. This book is organized into four logical parts, designed to guide you step-by-step from an experienced user to a shell expert:

1. A Foundation of Reinforced Concrete: First, we will solidify the

fundamentals, replacing beginner habits with professional practices. We’ll dive deep into the shell environment, master POSIX compliance for maximum portability, and learn how to make our scripts bulletproof with set -euo pipefail.

2. Mastering Complex Data and Flow Control: Here, you will learn the

building blocks for complex logic. We will design functions and libraries, leverage associative arrays (hashes) as a powerful data structure, and implement watertight error handling with trap.

3. The Shell as a Universal Glue: A huge part of the shell’s power lies in

its ability to control other programs. We will elevate our awk skills to a new level, tap into modern JSON APIs with jq, and allow our scripts to operate securely across networks.

4. From Script to Professional Tool: In the final part, we complete the

transformation into an architect. You will learn how to parse command-line arguments like a pro, create man pages for your own tools, and analyze and optimize the performance of critical scripts.

# HOW TO USE THIS BOOK

This is not a reference manual to be read from cover to cover and then placed on a shelf. It is a training manual. The greatest learning happens when you participate actively.

Type out the examples instead of just copying them. Change them. Try to

break them and figure out why they broke. Apply the concepts to your own, real-world problems. Each chapter concludes with exercises designed to challenge you to apply what you’ve learned—take advantage of these opportunities!

By the end of this book, you will no longer be just writing scripts. You

will be designing robust, thoughtful, and portable tools that will become a reliable and integral part of your automation pipelines and system administration tasks.

So, open your terminal, turn to the first chapter, and let’s begin. The path

from journeyman to master lies before you.

1. MASTERING THE SHELL

ENVIRONMENT

INTRODUCTION: FROM A

CONFIGURED SYSTEM TO A

CONTROLLED SYSTEM

In the first book, you learned how to personalize your shell with aliases and a custom PS1 prompt. You know that .bashrc is your friend. However, in the world of professional system administration and automation, it’s not enough to have an environment that “just works.” We need an environment that is predictable, robust, and secure, whether you’re logging in interactively via SSH or running an unattended script from a cron job.

In this chapter, we will pull back the curtain and illuminate the precise

mechanics behind a shell’s startup process. We will transform your knowledge from “where do I configure what?” to “why is a specific file loaded, when, and in what order?” Mastering these concepts is the cornerstone of writing scripts that function reliably not just on your laptop, but on any Linux system. Stop thinking of your shell environment as a garage with tools lying around; start treating it as a professional workshop where every tool has its place and all safety regulations are followed.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

After reading this chapter, you will be able to: • Confidently distinguish between login, non-login, interactive, and non-interactive shells.

• Deliberately choose the correct startup file (.bash\_profile, .bashrc, etc.) for any given purpose.

• Fortify your scripts with set and shopt options to eliminate common sources of error.

• Deeply understand and strategically use the difference between shell variables and environment variables (export). • Build your own modular and professional shell environment.

1.1 THE BIG PICTURE: INTERACTIVE,

NON-INTERACTIVE, LOGIN, AND

NON-LOGIN

Every time bash starts, it falls into one of four categories. The combination of these categories fundamentally determines the shell’s behavior. • Interactive Login Shell: The classic case. You log in via a text console or SSH to a system. You are presented with a prompt, and the system awaits your input.

◦ Example: ssh user@server

◦ Files Loaded: /etc/profile, then the first one found of ~/.bash\_profile, ~/.bash\_login, ~/.profile.

• Interactive Non-Login Shell: You are already logged in and start a new shell.

◦ Example: Opening a new terminal window in your graphical environment. ◦ Files Loaded: /etc/bash.bashrc (on some systems), then ~/.bashrc. • Non-interactive Non-Login Shell: This is the most common case for script execution.

◦ Example: bash my\_script.sh or a script launched by a cron job. ◦ Files Loaded: Typically none, unless the BASH\_ENV variable is set and points to a file.

• Non-interactive Login Shell: A rarer but important case. ◦ Example: ssh user@server "command"

◦ Behavior: Acts like a login shell (loading profile files) but then executes the command and exits.

Pro Tip: The most common beginner mistake is to place important PATH exports only in .bashrc. These will then be unavailable in cron jobs or during remote SSH command execution! Proper separation is key.

1.2 THE HIERARCHY OF STARTUP

FILES

The loading order is not magic; it’s a clear hierarchy.

1. System-Wide (/etc): Always first. Affects all users.

◦ /etc/profile: For login shells. Sets global paths and environment variables. ◦ /etc/bash.bashrc: For interactive non-login shells. Sets global aliases and functions.

1. User-Specific (~): Overrides or supplements the system-wide settings.

◦ For Login Shells: Bash looks for ~/.bash\_profile, then ~/.bash\_login, then ~/.profile. Only the first one found is executed.

◦ For Non-Login Shells: Only ~/.bashrc is loaded. ◦

The Proven Pattern to Avoid Redundancy: Add the following lines to

the end of your ~/.bash\_profile. This ensures that your .bashrc (with aliases and functions) is also available in login shells.

# ~/.bash\_profile

# Loaded for login shells

# Source global definitions

if [ -f /etc/profile ]; then

. /etc/profile

fi

# Source aliases and functions from .bashrc if it exists

if [ -f ~/.bashrc ]; then

. ~/.bashrc

fi

1.3 THE ART OF CONTROL: SET AND

SHOPT

By default, Bash is very forgiving. This is good for interactive work but a disaster for scripting. Activate “strict mode” at the beginning of every serious script:

#!/bin/bash

# Unofficial Bash Strict Mode

set -euo pipefail

IFS=$'\n\t'

• set -e (errexit): Exits the script immediately if a command exits with a non-zero status.

• set -u (nounset): Exits the script when trying to use an undefined variable. Prevents countless bugs from typos.

• set -o pipefail: The return value of a pipeline is the status of the last command to exit with a non-zero status (or zero if all commands succeed). Without this, erroneous\_command | true would be considered a success! • shopt: Toggles Bash-specific options on or off. One of the most useful is globstar.

◦ shopt -s globstar: Enables recursive file searching with \*\*. Example: ls -l \*\*/\*.log finds all .log files in all subdirectories.

1.4 USING ENVIRONMENT VARIABLES

CORRECTLY: THE DIFFERENCE

BETWEEN VAR=VALUE AND EXPORT

VAR=VALUE

This is a fundamental, and often misunderstood, point. • A shell variable (VAR="value") exists only within the current shell instance. • An environment variable (export VAR="value") is inherited by all child processes launched from that shell.

Practical Example:

1. Create a file test\_var.sh with the content echo "The variable is: $MY\_VAR".

2. Make it executable: chmod +x test\_var.sh.

3. In your terminal, run:

$ MY\_VAR="Hello World"

$ ./test\_var.sh

The variable is: # The variable is empty!

1. Now, with export:

$ export MY\_VAR="Hello World"

$ ./test\_var.sh

The variable is: Hello World # Now it works!

The export command made the variable visible to the child process script test\_var.sh.

1.5 HANDS-ON WORKSHOP: BUILDING

A PROFESSIONAL .BASHRC

Let’s put theory into practice. Here is a well-structured template for your .bashrc.

# ~/.bashrc

# If not running interactively, don't do anything

[[ $- != \*i\* ]] && return

#————————-

# 1. Strict Mode

#————————-

# (Optional for interactive shells, but good practice)

# set -uo pipefail

#————————-

# 2. Environment Variables

#————————-

export EDITOR='vim'

export PAGER='less'

# Add personal scripts to the PATH

export PATH="$HOME/bin:$HOME/.local/bin:$PATH"

#————————-

# 3. Shell Options with shopt

#————————-

# Enable recursive globbing (\*\*)

shopt -s globstar

# Correct spelling errors in directory changes

shopt -s cdspell

#————————-

# 4. Aliases

#————————-

# Source aliases from a separate file for better organization

if [ -f ~/.bash\_aliases ]; then

. ~/.bash\_aliases

fi

#————————-

# 5. Functions

#————————-

# Example: Quickly make a directory and change into it

mcd() {

mkdir -p "$1" && cd "$1"

}

#————————-

# 6. Prompt (PS1)

#————————-

# An informative prompt with Git status (simplified)

if [ -f ~/.git-prompt.sh ]; then

source ~/.git-prompt.sh # (Prerequisite: git-prompt.sh is available) export GIT\_PS1\_SHOWDIRTYSTATE=1

PROMPT\_COMMAND='\_\_git\_ps1 "\[\033[01;32m\]\u@\h\[\033[00m\]:\[\033[01;34m\]\w\

[\033[00m\]" "\\\$ "'

fi

# CHAPTER SUMMARY

In this chapter, you have taken the decisive step from being a user to becoming a conscious architect of your shell environment. You now understand the precise logic of how and when configuration files are loaded and can use this knowledge to write robust and portable scripts. You’ve learned to secure your scripts with “strict mode” and to master the critical difference between local and exported variables. With your new, well-structured .bashrc, you have laid the foundation for all the advanced topics in this book.

# EXERCISES

1. Debugging: Figure out why a script that works perfectly in your

interactive shell fails when run as a cron job. (Hint: check the PATH variable and aliases).

2. Experimentation: Write a short script that checks if it’s running in a

login shell or a non-login shell. (Hint: shopt -q login\_shell is your friend).

3. Refactoring: Restructure your own .bashrc and .bash\_profile according to

the pattern presented in this chapter. Move your aliases into a separate ~/.bash\_aliases file.

2. POSIX COMPLIANCE AND

PORTABILITY – WRITING UNIVERSAL

SCRIPTS

INTRODUCTION: THE LANGUAGE

EVERY SYSTEM UNDERSTANDS

With Bash, you’ve learned a powerful and convenient dialect of the shell language. Features like associative arrays and the [[...]] test operator make life easier. But what happens when your script needs to run not on your familiar Ubuntu desktop, but inside a minimalist Alpine Linux container, on an old enterprise server, or on a colleague’s macOS machine? Suddenly, your perfectly functional script might break with cryptic error messages.

The reason: you were speaking a dialect (Bash) where the lowest

common denominator (POSIX) was required. The POSIX standards define a base specification for shells and command-line tools that is adhered to by virtually every Unix-like system. A script that follows these rules is portable. It’s like a universal travel adapter: it works everywhere.

In this chapter, you will learn to consciously recognize the difference

between Bash-specific extensions (“Bash-isms”) and the portable POSIX standard. You will learn how to write robust, universal scripts that will work today and ten years from now, regardless of the system. This is not an academic exercise—it is an essential skill for anyone developing tools for others or working in heterogeneous environments.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Understand the fundamental difference between #!/bin/sh and #!/bin/bash. • Identify common, non-portable Bash extensions (“Bash-isms”). • Rewrite Bash-specific code into equivalent, portable POSIX code. • Effectively use the shellcheck tool to uncover portability issues. • Make a conscious decision about when maximum portability is necessary and when Bash features are the better choice.

2.1 THE SHEBANG SHOWDOWN:

#!/BIN/SH VS. #!/BIN/BASH

The first line of your script, the shebang, is an explicit instruction. • #!/bin/bash: You are telling the system, “This script requires Bash and all its extensions. Please execute it with /bin/bash.” • #!/bin/sh: You are telling the system, “This script is portable. Please execute it with the default system shell found at /bin/sh.”

The crucial point is that /bin/sh is not always bash.

• On many Debian/Ubuntu systems, /bin/sh is a symbolic link to dash, a very fast, minimalist, and strictly POSIX-compliant shell. • On macOS, /bin/sh is an outdated version of Bash that runs in POSIX mode. • In Alpine Linux (very popular in Docker), /bin/sh is a link to ash, another minimalist shell.Example of a Break: Save this as test.sh and make it executable.

#!/bin/sh

if [[ "a" == "a" ]]; then

echo "Equal"

fi

Run it: sh ./test.sh. On a Debian/Ubuntu system, you’ll get an error: ./test.sh: 2:

[[: not found. The [[...]] operator is a Bash-ism that dash does not understand.

2.2 THE POSIX TOOLBOX: COMMON

BASH-ISMS AND THEIR

ALTERNATIVES

To write portable scripts, you must replace the “luxury features” of Bash with their universal counterparts.

1. Conditional Expressions: [[...]] vs. [...]

• Bash-ism: [[ $var1 == $var2 && $var2 != "foo" ]] ◦ Advantages: Safer against word-splitting, extended operators like == for pattern matching, and &&/|| for logical grouping. • POSIX Standard: [ "$var1" = "$var2" ] && [ "$var2" != "foo" ] ◦ Rules:

▪ Always quote! [ $var ] can fail if $var is empty or contains spaces. [ "$var" ] is safe.

▪ Use a single = for string comparison. == is not standard. ▪ Avoid -a (AND) and -o (OR). They are unreliable. Instead, chain multiple [ tests with && and ||.

2. Arithmetic: ((...)) vs. $((...)) and expr

• Bash-ism: (( i++ ))

• POSIX Standard: i=$((i + 1))

◦ The $((...)) expansion for arithmetic evaluation is part of the POSIX standard and is the preferred, modern method.

• The “Old School” Way: i=$(expr "$i" + 1)

◦ expr is also portable but is more cumbersome and slower, as it’s an external command. You should be able to recognize it, but prefer $((...)).

3. Function Declaration: function vs. ()

• Bash-ism: function my\_func { ... }

• POSIX Standard: my\_func() { ... }

◦ Simply omit the function keyword. The () syntax is universal.

4. Data Structures: Associative Arrays & Co.

• Bash-ism: declare -A my\_hash

• POSIX Standard: There is no direct equivalent. ◦ Solution: This is where you must get creative. You can work with awk, which has built-in associative arrays, or process data in simple text formats (e.g., one line per key-value pair) and access it with grep and cut/sed. This is a point where complexity increases significantly, and you must ask yourself if a shell script is still the right tool for the job.

2.3 THE UNSUNG HERO: SHELLCHECK You don’t have to memorize all the POSIX rules. The shellcheck tool is your personal code reviewer that checks your script for common bugs, stylistic issues, and portability traps.

Installation (Example for Debian/Ubuntu): sudo apt-get install shellcheck Usage: Let’s take a script bad\_script.sh with some Bash-isms:

#!/bin/sh

my\_array=(a b c)

if [[ ${my\_array[0]} == "a" ]]; then

echo "It works"

fi

Run shellcheck on it in POSIX mode (-s sh): shellcheck -s sh bad\_script.sh Potential Output:

In bad\_script.sh line 2:

my\_array=(a b c)

^—SC3010: In POSIX sh, array assignments are not supported.

In bad\_script.sh line 3:

if [[ ${my\_array[0]} == "a" ]]; then

^—SC3014: In POSIX sh, [[ ]] is not supported. Use [ ].

^—SC3058: In POSIX sh, array references are not supported.

shellcheck tells you exactly what isn’t portable. It is an indispensable tool

for any serious shell scripter.

2.4 THE PRAGMATIC SCRIPTER:

WHEN IS PORTABILITY OVERKILL? Maximum portability is not always the ultimate goal. It’s a trade-off between universality and readability/convenience.

• Script for Personal Use: If you’re writing a script only for yourself on your own, known system, feel free to use the Bash features that make your job easier.

• Script for a Team/Company: If the script needs to run on various Linux servers that you control (e.g., all are RHEL 8), you can rely on the Bash version available there.

• Tool for Public Distribution/CI/CD/Docker: This is where portability is critical. If you don’t know where your code will be executed, write for #!/bin/sh and test it with shellcheck.

The art lies in making a conscious decision rather than accidentally

writing non-portable code.

# CHAPTER SUMMARY

You have made the leap from being a pure Bash scripter to a language-aware developer. You now understand that #!/bin/sh is a promise—the promise of portability. You can identify the most common Bash-specific extensions and replace them with their universal POSIX equivalents. With shellcheck, you have a powerful tool at your disposal to ensure the quality and portability of your code. Most importantly, you have adopted a new mindset: you no longer just ask “Does it work?” but rather “Where does it need to work, and how do I ensure that?”

# EXERCISES

1. Refactoring: Take one of your older Bash scripts. Change its shebang to

#!/bin/sh and use shellcheck -s sh to find all portability issues. Then, fix them.

2. Debugging: The following script is supposed to count the number of .log

files in a directory. It works under Bash but fails with sh script.sh. Figure out why and fix it in a portable way.

#!/bin/sh

count=0

for f in ./\*.log; do

((count++))

done

echo "Found $count log files."

1. Tooling: Install shellcheck on your system. Integrate it into your editor

(e.g., via a plugin for VS Code, Vim, etc.) to get instant feedback as you write.

3. ADVANCED PARAMETER

EXPANSION AND QUOTING – THE

SHELL’S SWISS ARMY KNIFE

INTRODUCTION: OPERATING IN THE

HEART OF THE SHELL

Until now, you’ve likely treated variables as simple storage containers: you put a value in (VAR="value") and get it back out (echo "$VAR"). When you needed to manipulate that value—for instance, to separate a filename from its path or replace spaces with underscores—you probably reached for external tools like cut, basename, or sed. This works, but it’s like calling a crane to hammer a nail into a wall. Every call to an external command creates a new process, which is slow and resource-intensive.

However, the shell itself provides an incredibly powerful, built-in set of

tools for string manipulation: parameter expansion. This allows you to operate directly on the contents of variables without ever leaving the shell. You can remove parts of strings, replace patterns, set default values, and even change a string’s case.

In this chapter, you will learn to master this “Swiss Army knife.” In

parallel, we will deepen your understanding of quoting, because only the combination of powerful expansion and correct quoting can protect you from the most insidious bugs in shell programming: unwanted word splitting and pathname expansion.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Efficiently remove parts of strings (prefixes, suffixes) without using cut or sed.

• Replace patterns within strings.

• Write robust scripts that work with default values if variables are not set. • Understand and apply the power of indirect expansion (${!var}). • Internalize the fundamental differences between '' (single quotes), "" (double quotes), and $'...' (ANSI-C quoting). • Understand and apply the “Golden Rule of Quoting” to finally eliminate bugs caused by spaces in filenames.

3.1 SUBSTRING REMOVAL: THE

SCALPEL FOR STRINGS

This is the most common use case. You have a string and want to slice off a portion from the beginning or the end.

Remove Prefix (from the beginning of the string) \* ${variable#pattern}:

Removes the shortest possible matching pattern from the beginning. (non-greedy) \* ${variable##pattern}: Removes the longest possible matching pattern from the beginning. (greedy)

Example: Dissecting a file path

filepath="/var/log/apache2/access.log"

# Shortest pattern: everything up to the first /

echo "${filepath#\*/}"

# -> var/log/apache2/access.log

# Longest pattern: everything up to the last / (extracts the filename) echo "${filepath##\*/}"

# -> access.log (equivalent to `basename "$filepath"`)

Remove Suffix (from the end of the string) \* ${variable%pattern}: Removes

the shortest possible matching pattern from the end. (non-greedy) \* ${variable%%pattern}: Removes the longest possible matching pattern from the end. (greedy)

Example: Removing a file extension

filename="backup\_2023-10-27.tar.gz"

# Shortest pattern: everything from the last .

echo "${filename%.\*}"

# -> backup\_2023-10-27.tar

# Longest pattern: everything from the first .

echo "${filename%%.\*}"

# -> backup\_2023-10-27

3.2 SEARCH AND REPLACE: SED FOR

IN-HOUSE USE

• ${variable/pattern/string}: Replaces the first occurrence of pattern with string. • ${variable//pattern/string}: Replaces all occurrences of pattern with string. • ${variable/#pattern/string}: Replaces pattern only if it’s at the beginning. • ${variable/%pattern/string}: Replaces pattern only if it’s at the end. Example: Sanitizing filenames

ugly\_name=" My Report (final).txt "

# Replace all spaces with underscores

clean\_name="${ugly\_name// /\_}"

echo "$clean\_name"

# -> \_My\_Report\_\_(final).txt\_

# That's not perfect yet. Let's combine!

# Step 1: Replace all spaces with underscores.

# Step 2: Remove parentheses.

temp\_name="${ugly\_name// /\_}"

final\_name="${temp\_name//[()]/}" # Replaces ( or ) with nothing echo "$final\_name"

# -> \_My\_Report\_\_final.txt\_

3.3 Robust Scripts Through Default Values and Error Checking

These expansions are critical for reacting to unset or empty variables.

• ${variable:-default}: If variable is unset or empty, use default. The variable itself is not changed.

• ${variable:=default}: If variable is unset or empty, use default and assign that value back to variable.

• ${variable:?error\_message}: If variable is unset or empty, print error\_message and exit the script (works best with set -u).

Example: Handling optional parameters

#!/bin/bash

set -u # Exit on unset variables

# The first command-line arg is the input; the second is an optional output

input\_file="$1"

# If $2 is not set, derive it from the input, but don't assign back to $2.

output\_file="${2:-${input\_file%.\*}.out}"

# Check if the input was provided. If not, the script aborts.

: "${input\_file:?Error: No input file was provided.}"

echo "Input: $input\_file"

echo "Output: $output\_file"

3.4 THE MASTER CLASS: INDIRECT

EXPANSION AND MORE

• Indirect Expansion: ${!prefix\*} or ${!prefix@} This is one of the most powerful but also most confusing techniques. It allows you to get the name of a variable dynamically from another variable.

var\_name="USER"

echo "The value of the variable whose name is in 'var\_name' is: ${!var\_name}" # -> The value of the variable whose name is in 'var\_name' is: [your username] • Case Modification (Bash 4+):

◦ ${variable^}: First letter to uppercase. ◦ ${variable,}: First letter to lowercase. ◦ ${variable^^}: All to uppercase.

◦ ${variable,,}: All to lowercase.

3.5 QUOTING – THE SAFETY NET

THAT MUST ALWAYS BE UP

You’ve now seen how to manipulate strings. The next step is to use them safely. The biggest problem in the shell is uncontrolled Word Splitting and Pathname Expansion (Globbing).

The Problem:

filename="My Important Report.txt"

touch "$filename" # Creates ONE file

ls -l $filename # ERROR! ls tries to find files "My", "Important", and "Report.txt".

The Solution: Double Quotes! ls -l "$filename" # CORRECT! ls receives the

string as ONE argument.

The Three Types of Quotes:

1. '' (Single Quotes): Absolutely Literal. No expansion of any kind

occurs. $USER remains $USER. Ideal for literal text. echo 'The current user is: $USER' -> The current user is: $USER

2. "" (Double Quotes): The Standard Case. Allows Parameter ($VAR),

Command ($(command)), and Arithmetic ($((...))) expansion, but prevents word splitting and globbing. This is your most important tool! echo "The current user is: $USER" -> The current user is: [your username]

3. $'...' (ANSI-C Quoting): For Special Characters. Allows the

interpretation of backslash escape sequences like \n (newline), \t (tab), or \x41 (hex code). printf $'Line 1\n\tIndented Line 2\n'

The Golden Rule of Quoting: Always quote your variable expansions ("$var") unless you know exactly why you need word splitting or globbing to occur.

# CHAPTER SUMMARY

You have unlocked the shell’s built-in string manipulation capabilities. Instead of launching an external process for every small text-editing task, you can now operate directly and efficiently within your script. You can remove substrings, replace patterns, and make your scripts more robust with default values and checks. More importantly, you have understood the “why” behind quoting. The consistent use of double quotes around your variables is the single most important characteristic that distinguishes a fragile script from a professional, robust tool.

# EXERCISES

1. Backup Script: Write a script that takes a filename as an argument (e.g.,

data.csv). The script should create a “backup” copy of this file named data.csv.YYYY-MM-DD.bak, where YYYY-MM-DD is the current date. Use parameter expansion to manipulate the base name and extension, and the date command for the date.

2. Variable Checker: Create a script that checks if the environment

variables EDITOR and PAGER are set. If EDITOR is not set, it should be set to /usr/bin/vim. If PAGER is not set, it should be set to /usr/bin/less. At the end, print the active values for both. (Hint: ${var:=default} is perfect here).

3. Name Normalizer: Write a script that takes a list of filenames with

spaces and renames them by replacing all spaces with underscores (\_) and converting all letters to lowercase. Your script must handle any filename safely. (Example: "My Report Part 2.DOC" becomes my\_report\_part\_2.doc).

4. FUNCTIONS, LIBRARIES, AND

SCOPE – THE PATH TO MODULAR

CODE

INTRODUCTION: FROM MONOLITHIC

SCRIPTS TO REUSABLE TOOLS

Imagine you’re building a complex piece of furniture. Every time you need a screw, would you melt steel, cast it into shape, and cut the threads? Of course not. You would reach into your toolbox and grab a ready-made screw. The principle is the same in programming: Don’t Repeat Yourself (DRY).

So far, your scripts may have been long, linear sequences of commands.

If you needed to repeat a similar task elsewhere, you probably copied and pasted the code. This leads to bloated, error-prone, and hard-to-maintain scripts. A change in one place has to be hunted down and replicated in many others.

In this chapter, you will learn the art of breaking your code down into

logical, reusable units: functions. We will explore how to design these functions cleanly, organize them into libraries, and—most importantly— control the scope of variables to avoid unintentional and catastrophic side effects. This is the transition from writing scripts to developing software.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Design professional, reusable functions with clear inputs and outputs. • Understand the critical difference between local and global variables and establish local as your default.

• Master the difference between return values (return) for status and stdout for data.

• Create your own script libraries and include them in other scripts using source.

• Safely pass complex data like arrays to functions without destroying them. • Develop a standard for documenting your functions.

4.1 THE ANATOMY OF A

PROFESSIONAL FUNCTION

A function is more than just a grouped block of code. It’s a contract. It has a defined task, defined inputs, and defined outputs.

Input: Processing Arguments Arguments are accessed inside a function

via $1, $2, $@, etc., just like in a script. It is good practice to immediately assign them to descriptive local variables.

# Function to greet a user

greet\_user() {

# Assign the first argument to a local variable

local user\_name="$1"

if [[ -z "$user\_name" ]]; then

echo "Error: No username provided." >&2

return 1 # Status: Failure

fi

echo "Hello, ${user\_name}!"

return 0 # Status: Success

}

Output: The Fundamental Difference Between Status and Data

• Return Value (return): For Status Only! The return command sets the function’s exit code (a number from 0-255). Its sole purpose is to signal success (0) or failure (anything else). It should never be used to return data. greet\_user "Alice" echo $? -> 0

• Standard Output (stdout): For Returning Data! If a function needs to return a result (a string, a number, etc.), it writes that result to standard output (stdout). The caller can then capture this result using command substitution ($()).

get\_current\_date() {

date—iso-8601

}

# Capture the function's data in a variable

today=$(get\_current\_date)

echo "Today's date is: $today"

• Standard Error (stderr): For Error and Debug Messages. As shown in the greet\_user example, error messages should always be redirected to stderr (>&2) so they are not accidentally captured as “data” by the caller.

4.2 SCOPE: THE MOST IMPORTANT

RULE FOR STABLE FUNCTIONS

Scope defines where a variable is visible and valid. In Bash, there are only two scopes: global and local (within a function).

The Danger of Global Variables (The Default Behavior) By default,

every variable you declare is global. This means it is visible and modifiable everywhere in the script—including inside functions.

The Catastrophic Example:

#!/bin/bash

count=10 # Global counter variable

# A function that supposedly counts something else

count\_files\_in\_dir() {

# Programmer forgets 'local' and overwrites the global variable!

count=0

for item in "$1"/\*; do

((count++))

done

echo "Directory '$1' has $count files."

}

echo "Before call: count = $count"

count\_files\_in\_dir "/etc"

echo "After call: count = $count" # Unexpectedly modified!

Output:

Before call: count = 10

Directory '/etc' has 152 files.

After call: count = 152

The global count variable has been destroyed!

The Savior: local The local keyword declares a variable that exists only

within the current function. It does not overwrite global variables and disappears as soon as the function finishes.

Rule Number One for Functions: Declare every variable inside a function with local unless you have a very good, intentional reason to modify a global variable.

The Corrected, Safe Example:

count\_files\_in\_dir() {

local count=0 # This variable is safe and only valid here

# ... rest of the function ...

}

4.3 SCRIPT LIBRARIES: ORGANIZING

AND REUSING CODE

If you have a set of useful helper functions (e.g., for logging, string manipulation, etc.), it’s impractical to copy them into every script. The solution is a library: a separate file containing only function definitions.

Example of a library file logging.sh:

#!/bin/bash

# This is my logging library

# Function to print an informational message with a timestamp

log\_info() {

local message="$1"

echo "[$(date +'%F %T')] [INFO] $message"

}

# Function to print an error message

log\_error() {

local message="$1"

# Write to stderr

echo "[$(date +'%F %T')] [ERROR] $message" >&2

}

A main script main.sh that uses the library: The source command (or its

shorthand, .) reads the specified file and executes its contents in the context of the current shell. This makes all functions from logging.sh available in the main script.

#!/bin/bash

set -euo pipefail

# Source my logging library. The path must be correct.

# `dirname "$0"` refers to the directory the script itself is in.

source "$(dirname "$0")/logging.sh"

log\_info "Script started."

# ... do something ...

if ! do\_something; then

log\_error "Action 'do\_something' failed!"

exit 1

fi

log\_info "Script finished successfully."

4.4 FOR ADVANCED USERS: PASSING

ARRAYS TO FUNCTIONS

This is a common hurdle. Attempting to pass an array like my\_func "${my\_array[@]}" does not work as expected, especially with elements containing spaces. The professional solution (since Bash 4.3) is namerefs (declare -n).

A nameref acts as an alias for another variable.

#!/bin/bash

# Function that accepts a reference to an array

process\_array() {

# 'array\_ref' is now an alias for the variable name

# passed during the call (e.g., 'files' or 'users').

declare -n array\_ref="$1"

log\_info "Processing array '$1' with ${#array\_ref[@]} elements."

for item in "${array\_ref[@]}"; do

echo " -> Element: '$item'"

done

}

# Source the logging library from above

source "$(dirname "$0")/logging.sh"

# Create two different arrays

files=("My Document.txt" "Spreadsheet (final).csv")

users=("Alice" "Bob" "Charlie")

# Call the function with the \*name\* of the array

process\_array files

process\_array users

# CHAPTER SUMMARY

You have made the leap from procedural code to modular programming. You now know how to design clean functions as contracts that have clear tasks and distinguish between status (via return) and data (via stdout). The consistent use of local has become second nature, protecting you from the most common and treacherous bugs. By organizing your functions into libraries, you create reusable and maintainable code. Finally, with namerefs, you have overcome the hurdle of processing complex data structures like arrays safely and efficiently. Your scripts have evolved into small, robust programs.

# EXERCISES

1. Refactoring: Take one of your older, longer scripts. Identify repeated

blocks of code and extract them into one or more functions. Ensure all variables inside the functions are local.

2. Library Creation: Create a string\_utils.sh library with two functions:

to\_lower(), which converts a string to lowercase, and trim(), which removes leading and trailing whitespace from a string. Write a second script that sources this library and tests both functions.

3. Array Processing: Write a function array\_contains\_element that takes two

arguments: the name of an array (as a nameref) and a search string. The function should return with an exit code of 0 (success) if the string is contained in the array, and 1 (failure) if not.

5. ROBUST ERROR HANDLING AND

DEBUGGING – WHEN THINGS GO

WRONG

INTRODUCTION: FROM A CRASH TO

A CONTROLLED EMERGENCY

LANDING

Anyone can write a script that works when everything goes right—the “happy path.” However, the true art of scripting reveals itself when things go wrong: a file doesn’t exist, a disk is full, a network service is unreachable, or the user presses Ctrl+C.

A beginner’s script that relies on set -e will simply crash in such cases. It

might leave behind temporary files, an inconsistent state, or a half-written output file. It’s like an airplane that loses its engines and falls uncontrollably from the sky.

A professional script, in contrast, performs a controlled emergency

landing. It detects the failure, shuts down its systems, sends out a distress signal with precise coordinates, and ensures the “passengers” (your data and your system) remain safe. In this chapter, you will learn how to give your scripts exactly this capability. We will go far beyond set -e and build a robust framework for error detection, cleanup, logging, and deep debugging.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Implement reliable cleanup actions on script exit using trap EXIT. • Catch errors the moment they happen with trap ERR to react immediately. • Generate meaningful error messages with line numbers and function names. • Apply advanced debugging techniques with trap DEBUG that are more powerful than a simple set -x.

• Develop effective logging strategies to direct your script’s output to files or the system log (syslog).

• Understand the importance of exit codes and use them deliberately.

5.1 THE LANGUAGE OF FAILURE:

MASTERING EXIT CODES

Every command returns an exit code upon completion, which is stored in the $? variable. \* 0: Success. Everything is okay. \* 1-255: Failure. Something went wrong.

Using set -e reacts to these codes, but as an advanced scripter, you need to

understand them and set them yourself.

• Setting exit codes explicitly: exit 1 (general error), exit 127 (command not found), etc.

• Importance: Automation tools like cron, systemd, or CI/CD pipelines rely on correct exit codes to decide if a job was successful. Pro Tip: There is a quasi-standard for exit codes documented in /usr/include/sysexits.h. Using these codes (e.g., 64 for bad arguments, 70 for an internal software error) makes your scripts more understandable to other system administrators.

5.2 THE SAFETY NET: CLEANING UP

WITH TRAP EXIT

What happens to your script’s temporary files when it aborts midway? They are left behind as clutter. The trap command intercepts signals and executes a command of your choice when they are received.

The most important trap is trap ... EXIT. It is always executed when the

script ends, regardless of whether it was a successful completion, an error (set-e), or an external signal like Ctrl+C (SIGINT).

The Classic Cleanup Pattern:

#!/bin/bash

set -e

# Create a secure temporary file

# -d creates a directory, which is often better

temp\_dir=$(mktemp -d)

# Define the cleanup function

cleanup() {

echo "=> Performing cleanup..."

# 'rm -rf' is safe here because $temp\_dir is absolute and came from mktemp rm -rf "$temp\_dir"

echo "=> Temporary directory '$temp\_dir' removed."

}

# Register the function for the EXIT signal

trap cleanup EXIT

echo "Temporary directory created: $temp\_dir"

echo "Working with files..."

touch "$temp\_dir/file1.txt"

sleep 5 # Simulate work. Press Ctrl+C here!

echo "Work finished."

# The script ends here, 'trap' is triggered

No matter how you exit this script, the cleanup function will be executed

and the directory will be deleted.

5.3 PROACTIVE ERROR DETECTION:

TRAP ERR

While trap EXIT cleans up at the end, trap ERR reacts the moment a command returns a non-zero exit code. This allows for much more detailed logging.

The Error Logging Pattern:

#!/bin/bash

set -o errexit # Same as 'set -e'

set -o nounset # Same as 'set -u'

# Error handler function

handle\_error() {

local exit\_code=$?

local line\_number=$1

local command=$2

echo "ERROR in line $line\_number: command '$command' failed with exit code $exit\_code." >&2 }

# Register the handler for the ERR signal

# Pass the line number and the command to the function

trap 'handle\_error $LINENO "$BASH\_COMMAND"' ERR

echo "All good..."

ls /non\_existent\_directory # This command will fail

echo "This line will never be reached."

Output:

All good...

ls: cannot access '/non\_existent\_directory': No such file or directory ERROR in line 16: command 'ls /non\_existent\_directory' failed with exit code 2.

This gives you precise, immediate information about what went wrong

and where.

5.4 ADVANCED DEBUGGING: BEYOND

SET -X

set -x is useful, but its output can be cluttered. With trap DEBUG, you can build your own intelligent tracer. The DEBUG trap is executed before every simple command.

A Smarter Trace Mode:

#!/bin/bash

# Function to be executed before each command

debug\_trace() {

# Print function name, line number, and command

echo "[DEBUG] ${BASH\_SOURCE[0]}:${BASH\_LINENO[0]} -> ${BASH\_COMMAND}" >&2

}

# Check if an environment variable is set to enable debug mode

if [[ "${SCRIPT\_DEBUG:-}" == "true" ]]; then

trap debug\_trace DEBUG

fi

my\_func() {

local name="World"

echo "Hello, $name"

}

echo "Script starting."

my\_func

echo "Script ending."

Invocation: ./script.sh -> Normal output. SCRIPT\_DEBUG=true ./script.sh -> Detailed trace output for every command.

5.5 PROFESSIONAL LOGGING

STRATEGIES

Where should all the echo and error messages go?

1. Logging to a File with exec: To redirect the entire output of a script to a

file without modifying every echo command.

#!/bin/bash

LOG\_FILE="/var/log/my\_script.log"

# Redirect stdout and stderr to the log file

# The 'tee' command allows output to also go to the console

exec > >(tee -a "${LOG\_FILE}") 2> >(tee -a "${LOG\_FILE}" >&2) echo "This info message will go to the log."

ls /error # This error message will also go to the log.

2. Logging to the System Log with logger: The most professional way is

to use the system logger (syslog). The logger command makes this easy.

#!/bin/bash

# The -t tag makes messages easy to filter

logger -t "MyBackupScript" "Backup process started."

# ...

if ! tar -czf /backup.tar.gz /home; then

# -p sets the priority/facility (local0.error)

logger -t "MyBackupScript" -p local0.error "Backup failed!"

exit 1

fi

logger -t "MyBackupScript" "Backup completed successfully."

These messages can then be viewed centrally with journalctl -t MyBackupScript

or in /var/log/syslog.

# CHAPTER SUMMARY

You have learned to equip your scripts with a robust error-handling framework. You no longer rely blindly on set -e, but instead use trap to actively respond to your script’s termination (EXIT) or to errors as they occur (ERR). This allows you to safely release resources and create detailed error logs. With trap DEBUG and the logger command, you now have tools that go far beyond the basics, enabling you to build professional, reliable, and maintainable automation solutions.

# EXERCISES

1. Safe Download Script: Write a script that uses wget to download a file.

It should create a temporary directory with mktemp. Use trap EXIT to ensure this directory is always deleted, even if the download fails or the user aborts the script.

2. Error Reporter: Extend the script from exercise 1. Implement a trap ERR

handler that, in case of an error, prints a message to stderr containing the line number and the error code.

3. Syslog Integration: Modify the script further so that it writes start,

success, and failure messages to the syslog using the logger command. Use a unique tag (e.g., DownloadScript) to make the messages filterable.

6. MASTERING ARRAYS AND DATA

STRUCTURES – MORE THAN JUST A

LIST

INTRODUCTION: FROM A SINGLE

VALUE TO A STRUCTURED

COLLECTION

In simple scripts, a handful of variables is often sufficient to get a job done. But reality is rarely so simple. What if you need to process a list of ten servers, collect metadata from a hundred files, or store the configuration settings of an application? You won’t get far with individual variables. You need a way to group related data and access it efficiently.

You already know the basic form: the indexed array, an ordered list of

items. It’s the backbone for processing lists. But the real power to model reality lies in its more advanced counterpart: the associative array, often called a hash or dictionary. It stores not just values, but the relationship between them by assigning a unique, named key to each value.

In this chapter, we will master both array types. We will briefly review

the professional handling of indexed arrays and then dive deep into the world of associative arrays. You will learn how to use them as configuration stores, as counters, or even to simulate more complex data structures. Mastering these structures is the difference between a script that processes data and a script that understands data.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Handle indexed arrays professionally and safely, especially when iterating. • Securely declare, populate, and read from associative arrays (declare -A) as key-value stores.

• Efficiently iterate over the keys and values of associative arrays. • Implement typical use cases like parsing configuration files or counting elements.

• Understand the limits of shell arrays and learn how to simulate more complex structures (e.g., a list of objects).

6.1 INDEXED ARRAYS: THE

PROFESSIONAL APPROACH

A quick review of the most important concepts with a focus on robust practices.

Declaration and Population:

# Declare and populate in one step

servers=("web-01" "db-01" "api-server-alpha")

# Add individual elements

servers+=( "monitoring" )

Access and Iteration – The Importance of Quoting: Correctly quoting

during expansion is the critical point that separates beginners from professionals.

# Wrong: Will fail for elements with spaces

for server in ${servers[@]}; do ...

# CORRECT: Treats each element as a distinct, single string

for server in "${servers[@]}"; do

echo "Processing server: '$server'"

done

# Access a single element (index starts at 0)

echo "The first server is: ${servers[0]}"

# Number of elements

echo "Number of servers: ${#servers[@]}"

6.2 ASSOCIATIVE ARRAYS: THE

HEART OF STRUCTURED DATA

This is where the real magic begins. Associative arrays must be explicitly declared with declare -A.

Declaration and Population:

# Declaration is mandatory!

declare -A user\_data

# Assign values via key-value pairs

user\_data["name"]="Alice"

user\_data["id"]="1024"

user\_data["email"]="alice@example.com"

user\_data["login\_count"]=42

user\_data["department"]="Quality Assurance" # Key with spaces

Access and Iteration: Accessing keys and values has a special syntax.

# Access a value by its key

echo "Username: ${user\_data[name]}"

# Iterate over all keys (the most important operation!)

echo "All user properties:"

for key in "${!user\_data[@]}"; do

value="${user\_data[$key]}"

printf " %-15s: %s\n" "$key" "$value"

done

# Check if a key exists

if [[ -v user\_data["email"] ]]; then

echo "The email address exists."

fi

# Delete an entry

unset user\_data["login\_count"]

Important: The order of keys in an associative array is not guaranteed! 6.3 HANDS-ON WORKSHOP: BUILDING

A CONFIGURATION PARSER

A classic use case for associative arrays is safely parsing .ini or .conf files.

Example file app.conf:

hostname = db.internal.net

port = 5432

user = admin\_user

# Password shouldn't be here, but for example's sake

password = s3cr\_et!

The Parser Script:

#!/bin/bash

set -euo pipefail

parse\_config() {

local config\_file="$1"

# Nameref to return the result

declare -n result\_hash="$2"

if [[ ! -f "$config\_file" ]]; then

echo "Error: Configuration file '$config\_file' not found." >&2 return 1

fi

while IFS='=' read -r key value; do

# Skip comments and empty lines

[[ "$key" =~ ^\s\*# || -z "$key" ]] && continue

# Trim whitespace around key and value

key=$(echo "$key" | xargs)

value=$(echo "$value" | xargs)

result\_hash["$key"]="$value"

done < "$config\_file"

}

# Declare the destination hash

declare -A db\_config

# Call the function and pass the name of the hash

parse\_config "app.conf" db\_config

# Work with the loaded data

echo "Connecting to host ${db\_config[hostname]} on port ${db\_config[port]}..."

6.4 SIMULATING COMPLEX DATA

STRUCTURES

Shell scripts don’t have native “objects” or “structs.” But what if we have a list of users, each with multiple properties? We can simulate this with clever naming conventions in associative arrays.

Scenario: A list of servers, each with an IP and a role.

The Simulation Technique: We use an index and concatenate it with the

property names.

#!/bin/bash

# Data declaration

declare -A servers

server\_count=0

add\_server() {

local ip="$1"

local role="$2"

servers["${server\_count}\_ip"]="$ip"

servers["${server\_count}\_role"]="$role"

((server\_count++))

}

# Populate the "list of objects"

add\_server "10.0.1.10" "webserver"

add\_server "10.0.1.20" "database"

add\_server "10.0.1.30" "cache"

# Process the data

echo "Server Inventory:"

for ((i=0; i

ip\_key="${i}\_ip"

role\_key="${i}\_role"

echo " - Server $i: IP=${servers[$ip\_key]}, Role=${servers[$role\_key]}"

done

This technique, while verbose, is extremely powerful. However, it also

reveals the limits of shell scripting. For highly complex data structures, languages like Python or Perl are often a better fit.

# CHAPTER SUMMARY

You have mastered the handling of data collections in the shell. You can not only process simple lists safely with indexed arrays but also model complex key-value data with associative arrays. You’ve learned that they are the perfect solution for tasks like parsing configuration files or counting elements. With the technique for simulating object lists, you have even ventured beyond the standard use cases and understood where the strengths— and weaknesses—of the shell lie. With this knowledge, you can now write scripts that not only move data but also structure and interpret it intelligently.

# EXERCISES

1. Frequency Count: Write a script that reads a text file and uses an

associative array to count the frequency of each word. At the end, it should print a sorted list of the words and their counts.

2. Improved Config Parser: Extend the configuration parser from the

chapter. It should accept a second argument: an indexed array of allowed key names. When parsing, only keys that appear in the list of allowed keys should be added to the hash.

3. User Inventory: Write a script that reads the /etc/passwd file. Use the

technique shown in section 6.4 to create a “list of objects,” where each entry has the properties username (field 1), uid (field 3), and shell (field 7). At the end, print a formatted list of all users with a UID greater than 1000.

7. PROCESS MANAGEMENT AND

PARALLELIZATION – BECOME THE

CONDUCTOR

INTRODUCTION: FROM SERIAL

EXECUTION TO PARALLEL

ORCHESTRATION

Until now, your scripts have likely functioned like a single-lane road: one command executes, and only when it’s finished can the next one start. This is simple and predictable. But what if you need to convert 1,000 images, download 50 log files from different servers, or compare the complex output of two commands? Serial execution quickly becomes a crippling bottleneck.

In this chapter, you will learn to evolve from a simple traffic cop into the

conductor of an entire orchestra. You will learn to launch processes in the background and precisely control their execution (wait). You will connect data streams in ways unimaginable with simple pipes (process substitution). And you will harness the dormant power of modern multi-core CPUs to make your scripts many times faster through parallelization.

Mastering these techniques is a quantum leap. It allows you to design

complex workflows that are not only correct but also extremely performant. You will stop just executing commands—you will start orchestrating processes.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Safely and reliably manage background processes in scripts using wait and jobs.

• Understand the power of process substitution (<(...), >(...)) to avoid temporary files.

• Establish two-way communication with background processes using coprocesses (coproc).

• Learn two powerful methods for parallelizing tasks: xargs and the wait loop. • Keep an eye on system load and sensibly throttle the number of parallel jobs.

7.1 MORE THAN JUST &: THE ART OF

JOB CONTROL IN SCRIPTS

You know the & symbol for sending a process to the background from your interactive shell. In scripting, it unleashes its true power only in combination with wait.

• jobs: Displays the jobs running in the background. • wait [job\_id | process\_id]: Pauses the script until the specified job or process has finished. Without an argument, wait waits for all running background processes of the script.

The Basic Pattern for Parallel, Independent Tasks:

#!/bin/bash

echo "Starting all jobs..."

# Launch several long-running processes in the background

sleep 5 && echo "Job 1 finished." &

sleep 3 && echo "Job 2 finished." &

sleep 4 && echo "Job 3 finished." &

echo "All jobs are running. Now waiting for them to complete..."

# 'wait' blocks here until the last background process has finished

wait

echo "All jobs have completed. Script finished."

Without wait, the script would exit immediately, and the echo commands

would execute later, which is rarely the desired behavior.

7.2 PROCESS SUBSTITUTION: WHEN

PIPES AREN’T ENOUGH

A pipe (|) is fantastic for sending the output of one command to the standard input of another. But what if a command expects its input from a file? Example: diff file1 file2. This is where process substitution comes in.

<(command) – A Command That Pretends to Be a File: The shell executes

the command and makes its stdout available via a special file in the /dev/fd directory. The outer command thinks it’s reading from a normal file.

Example: Compare the contents of two directories without temporary

files.

# The "old" way with temporary files:

ls -1 /etc > /tmp/etc.txt

ls -1 /usr/bin > /tmp/usr\_bin.txt

diff /tmp/etc.txt /tmp/usr\_bin.txt

rm /tmp/etc.txt /tmp/usr\_bin.txt

# The elegant, modern way:

diff <(ls -1 /etc) <(ls -1 /usr/bin)

# No temporary files, no cleanup needed!

>(command) – A Command Disguised as a Writable File: This is the

counterpart. It allows the output of a command to be sent directly to the standard input of another, as if it were a file.

Example: Write output to the console and a grep process simultaneously.

# 'tee' writes its input to both stdout and the specified file(s)

echo -e "INFO: Start\nERROR: Critical Failure\nINFO: End" | tee >(grep ERROR)

Output:

ERROR: Critical Failure # Output from grep

INFO: Start

ERROR: Critical Failure

INFO: End # Output from tee to the console

7.3 COPROCESSES (COPROC): TWO-

WAY COMMUNICATION

This is one of the most advanced techniques. A coprocess is a background process with which your main script can communicate via two dedicated pipes: you can send it commands and read replies from it.

Syntax: coproc NAME { commands; } \* NAME[0]: A file descriptor number to

read from the coprocess. \* NAME[1]: A file descriptor number to write to the coprocess.

Example: Interacting with the bc calculator

#!/bin/bash

# Start 'bc' as a coprocess named 'CALC'

coproc CALC { bc -l; }

# Send commands to the coprocess

echo '10 / 3' >&${CALC[1]}

echo 'sqrt(100)' >&${CALC[1]}

# Read the replies from the coprocess

read -r result1 <&${CALC[0]}

read -r result2 <&${CALC[0]}

echo "Result 1: $result1"

echo "Result 2: $result2"

# Close the coprocess

exec {CALC[1]}>&-

Coprocesses are ideal for scenarios where a persistent connection or state

needs to be maintained in the background process.

7.4 HANDS-ON WORKSHOP: SPEEDING UP SCRIPTS WITH PARALLELIZATION Scenario: We have a list of URLs in a file and want to get the HTTP status code for each one.

Method 1: The for Loop with wait and Throttling This method gives us

maximum control. We start a fixed number of jobs and wait for one to finish before starting the next, to avoid overwhelming the system.

#!/bin/bash

URL\_FILE="urls.txt"

MAX\_JOBS=4

# Function that does the actual work

check\_url() {

local url="$1"

local code

code=$(curl -s -o /dev/null -w "%{http\_code}" "$url")

echo "$url -> $code"

}

# Start processing

while read -r url; do

# Start the function in the background

check\_url "$url" &

# Throttling: if the max number of jobs is reached, wait

# for any job to finish before continuing the loop.

if [[ $(jobs -p | wc -l) -ge $MAX\_JOBS ]]; then

wait -n # Wait for the next job to terminate

fi

done < "$URL\_FILE"

# Wait for all remaining jobs

wait

echo "All URLs checked."

Method 2: The Elegant xargs Solution xargs is an extremely powerful tool

built for exactly these kinds of tasks.

• -P max\_procs: Run up to max\_procs processes at a time. • -n num: Pass num arguments to each command invocation.

#!/bin/bash

URL\_FILE="urls.txt"

MAX\_JOBS=4

# The function must be exported so xargs can call it in a subshell export -f check\_url

# Read the file and pipe it to xargs

cat "$URL\_FILE" | xargs -P $MAX\_JOBS -n 1 bash -c 'check\_url "$@"' \_

This xargs variant is often shorter and more performant, while the for loop

offers more flexibility for complex logic.

# CHAPTER SUMMARY

You are now able to unleash the full power of your system. You can not only launch background processes but also control them with wait. With process substitution, you’ve learned a clean and efficient alternative to temporary files for complex data pipelines. You’ve even had a glimpse into two-way communication with coprocesses. Most importantly, you now know how to dramatically speed up repetitive tasks through parallelization, whether through the controlled wait loop or the powerful xargs command. Your scripts are no longer bound by serial execution.

# EXERCISES

1. Parallel Download: Write a script that reads a list of image URLs from

a file and downloads them in parallel using wget or curl. Throttle the number of simultaneous downloads to 5.

2. Comparison of Filtered Data: Use process substitution (<(...)) to

directly compare the output of two different grep commands on a large log file using diff, without creating a single temporary file.

3. Performance Test: Take a script that processes a large number of small

files (e.g., converting images). Measure the execution time of the serial version. Then, implement a parallel version using xargs -P and measure the time again.

8. THE TRIUMVIRATE: GREP, SED,

AND AWK REVISITED

INTRODUCTION: FROM POCKET

KNIFE TO POWER SAW

In your journey as a shell scripter, grep, sed, and awk have been your constant companions. You’ve used grep to find lines, sed to perform simple substitutions, and perhaps awk to extract a specific column. You have mastered them as the digital equivalent of a pocket knife: versatile and useful for countless small tasks.

It is now time to upgrade your tools. This chapter will transform your

pocket knife into a set of specialized power tools. We will revisit this “triumvirate” of text processing, but this time, we will unlock their advanced features. You will learn to use grep for complex, multi-line pattern matching with Perl-Compatible Regular Expressions (PCRE). You will dive into the esoteric but powerful “hold space” of sed to perform non-linear text transformations. And you will discover that awk is not just a column-extractor, but a complete, Turing-complete programming language with arrays, functions, and complex logic.

Most importantly, you will learn the art of combining these three tools

into elegant and powerful pipelines, where each tool does what it does best. Mastering this synergy is the key to solving virtually any text-processing challenge the command line throws at you.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Leverage advanced grep features like PCRE (-P), context control (-A, -B, -C), and performance optimizations.

• Master sed’s hold space to save and recall patterns, enabling complex, multi-line operations.

• Write complete programs in awk, using BEGIN/END blocks, built-in variables, and associative arrays.

• Create powerful and efficient text-processing pipelines by combining the strengths of each tool.

• Decide which tool is the right one for a specific job.

8.1 GREP: THE MASTER OF FINDING

You know grep for finding lines. Now, let’s make it find anything.

PCRE (-P): The Ultimate Pattern Matching Many Linux distributions

ship with a grep that supports Perl-Compatible Regular Expressions. This unlocks incredibly powerful features like lookaheads, lookbehinds, and non-capturing groups.

Example: Find lines that contain an IP address, but not if it’s 127.0.0.1.

# Negative Lookbehind: `(?

grep -P '(?

Context Control: Seeing the Bigger Picture Often, the matching line

alone is not enough. \* -A num (After): Shows num lines of trailing context after a match. \* -B num (Before): Shows num lines of leading context before a match. \* -C num (Context): Shows num lines of context both before and after a match.

Example: Find an error and the 2 lines that led up to it. grep -B 2 "ERROR"

application.log

Performance and Efficiency \* -F (Fixed strings): If you are searching for

a literal string, not a pattern, -F is significantly faster. grep -F "user.login" ... \* -q (Quiet): If you only need to know if a match exists (e.g., in an if statement), -q suppresses all output and exits immediately on the first match. if grep -q "FAIL" /var/log/secure; then ...

8.2 SED: THE STREAM EDITOR,

UNLEASHED

You know s/old/new/. Now, let’s explore sed’s hidden brain: the hold space. The hold space is a secondary, persistent buffer. You can copy or append the current line (from the “pattern space”) into it, and later retrieve it. This allows sed to “remember” things across multiple lines.

Hold Space Commands: \* h, H: Copy/Append pattern space to hold

space. \* g, G: Copy/Append hold space to pattern space. \* x: Exchange pattern space and hold space.

Example: Reverse the order of every two lines in a file.

# For a file with:

# Line 1

# Line 2

# Line 3

# Line 4

# On line 1: Hold it.

# On line 2: Get line 1 from hold space, append line 2, and print. # And so on...

sed -n 'h;n;G;p' input.txt

Output:

Line 2

Line 1

Line 4

Line 3

Example: Add a header only above a specific block. sed '/START\_BLOCK/ { x;

s/.\*/MY\_HEADER/; p; x; }' input.txt

8.3 AWK: THE DATA-PROCESSING

POWERHOUSE

awk is not just for '{print $1}'. It’s a full programming language designed for record-based data processing.

Key Concepts: \* BEGIN { ... } block: Executed once before any lines are

read. Perfect for initializing variables or printing headers. \* END { ... } block: Executed once after all lines have been read. Ideal for calculations and printing summaries. \* Built-in Variables: NF (Number of Fields), NR (Number of Records/Line Number), FS (Field Separator). \* Associative Arrays: awk’s greatest strength.

Example: Calculate the average size of files listed by ls -l.

ls -l | awk '

# Skip the first line ("total...") and directories

NR > 1 && !/^d/ {

# Add the size (field 5) to the total

total\_size += $5

# Increment the file count

count++

}

END {

if (count > 0) {

# Calculate and print the average

printf "Average file size: %.2f bytes\n", total\_size / count

}

}

'

Example: Count the number of connections from each IP address in a log

file. awk '{ ip\_counts[$1]++ } END { for (ip in ip\_counts) { print ip, ip\_counts[ip] } }' access.log

8.4 THE ART OF THE PIPELINE:

COMBINING STRENGTHS

The true mastery lies in knowing which tool to use for which part of the job and chaining them together.

Scenario: From an Apache access log, find all “404 Not Found” errors,

extract the requested URL (field 7), count how many times each unique URL was requested, and display the top 5 most frequent 404 errors.

The Pipeline:

# 1. grep: Quickly find only the relevant lines. It's faster than awk for simple filtering. grep ' 404 ' /var/log/apache/access.log | \

# 2. awk: Extract the 7th field (the URL). awk is perfect for column-based data. awk '{print $7}' | \

# 3. sort: Sort the URLs so that identical ones are adjacent. This is a prerequisite for `uniq`. sort | \

# 4. uniq -c: Count the occurrences of adjacent identical lines. uniq -c | \

# 5. sort -rn: Sort the results numerically and in reverse (highest first).

sort -rn | \

# 6. head -n 5: Take only the first 5 lines of the sorted output.

head -n 5

This pipeline is a classic example of the Unix philosophy: each program

does one thing well, and they are combined to achieve a complex result. Trying to do this all in a single awk or sed script would be far more complex and less readable.

# CHAPTER SUMMARY

You have taken three familiar tools and rediscovered them as the powerful and specialized instruments they are. You can now use grep for intricate pattern matching, sed for complex, stateful transformations, and awk as a robust data-processing language. More importantly, you have learned the art of synergy—how to build elegant and efficient pipelines by assigning each tool the task it was designed for. You are no longer just processing text; you are engineering data flows.

# EXERCISES

1. Log File Summarizer: Write a pipeline that reads a syslog file. It should

find all lines containing the word “CRON”, extract the username in parentheses (e.g., (root)), count the occurrences of each username, and display a sorted list.

2. sed Hold Space Challenge: Given a file with paragraphs separated by

blank lines, write a sed script that deletes any paragraph that contains the word “legacy”. (Hint: Accumulate lines in the hold space until a blank line is found).

3. Advanced awk Report: Write an awk script that processes the output of

netstat -an. It should generate a report that counts how many connections are in each state (e.g., ESTABLISHED, TIME\_WAIT, LISTEN). The output should be a formatted table.

9. INTERACTING WITH MODERN WEB

APIS AND DATA FORMATS

INTRODUCTION: THE SHELL AS A

WEB CLIENT

In the past, the command line was primarily a tool for managing a single machine. Text files, logs, and system processes were its domain. Today, however, the digital world is a vast, interconnected network of services. Your cloud servers, your CI/CD system, your weather provider, and even your social media feed are all accessible through Application Programming Interfaces (APIs). The universal language of these modern APIs is not plain text, but structured data formats—most commonly, JSON (JavaScript Object Notation).

To remain relevant and powerful, the shell scripter must learn to speak

this language. This chapter will transform your shell from a local system manager into a fully-fledged web client. You will master curl, the undisputed champion of command-line HTTP requests, learning how to handle everything from simple GET requests to complex authenticated POST operations.

Then, you will meet jq, an indispensable tool that is to JSON what sed and

awk are to plain text. You will learn to slice, dice, transform, and extract any piece of data from complex JSON documents with elegant and powerful expressions. By combining curl and jq, you will be able to write scripts that automate cloud infrastructure, query databases, or interact with virtually any modern web service, all from the comfort of your terminal.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Master curl to make sophisticated HTTP requests, including setting headers, handling cookies, and using different HTTP methods (POST, PUT, DELETE). • Understand and implement common API authentication patterns. • Learn the fundamentals of jq to navigate, filter, and extract data from JSON objects and arrays.

• Build complex jq expressions to transform and restructure JSON data. • Create a complete command-line client for a real-world REST API by combining curl and jq.

• Get a brief introduction to handling XML, the other common data format, with xmlstarlet.

9.1 CURL: YOUR SWISS ARMY KNIFE

FOR THE WEB

curl can do much more than just download a file. It’s a powerful tool for any kind of HTTP interaction.

Essential curl Options: \* -X METHOD: Specify the HTTP method, e.g., -X

POST, -X DELETE. \* -H "Header: Value": Add a custom HTTP header. Essential for setting Content-Type or Authorization. \* -d 'data': Send data in the request body, typically for POST or PUT requests. \* -i: Include the HTTP response headers in the output. \* -s: Silent mode. Don’t show progress meters or error messages. \*-L: Follow redirects (3xx status codes).

Example: Making a simple POST request to a JSON API

# We tell the server we are sending JSON and provide the JSON data.

curl -X POST \

-H "Content-Type: application/json" \

-d '{"username":"alice","score":100}' \

https://api.example.com/v1/scores

Handling Authentication: Most APIs require authentication. A common

method is using a bearer token in an Authorization header.

API\_TOKEN="your\_super\_secret\_token\_here"

curl -s -H "Authorization: Bearer ${API\_TOKEN}" \

https://api.example.com/v1/user/profile

(Remember the lessons from Chapter 10: Never hard-code secrets!)

9.2 INTRODUCTION TO JQ: THE

SED/AWK FOR JSON

Imagine receiving this JSON response from the API call above:

{

"user": {

"id": 101,

"username": "alice",

"real\_name": "Alice Smith",

"is\_active": true

},

"roles": ["editor", "contributor"],

"last\_login": "2023-10-28T10:00:00Z"

}

Trying to parse this with grep and sed is fragile and will break. jq is

designed for this.

Basic Filters: \* .: The identity filter; outputs the entire input, pretty-

printed. \* .key: Access a value by its key. \* .[index]: Access an array element by its index.

Piping curl to jq: This is the fundamental workflow.

# Get the username

curl -s ... | jq '.user.username'

# -> "alice" (Note the quotes, it's a JSON string)

# To get the raw string without quotes, use the -r flag

curl -s ... | jq -r '.user.username'

# -> alice

# Get the first role

curl -s ... | jq -r '.roles[0]'

# -> editor

9.3 ADVANCED JQ EXPRESSIONS

jq’s power comes from its ability to chain filters together with a pipe |, just like the shell.

Chaining and Object Construction: You can build new JSON objects

on the fly.

Example: Create a new object with only the user ID and the last login

time.

curl -s ... | jq '{ user\_id: .user.id, login: .last\_login }'

Output:

{

"user\_id": 101,

"login": "2023-10-28T10:00:00Z"

}

Working with Arrays: map and select \* map(expression): Applies an

expression to each element of an array. \* select(condition): Filters elements based on a boolean condition.

Example: From a list of users, get the names of all active\* users. Input

JSON:\*

[

{"name": "alice", "active": true},

{"name": "bob", "active": false},

{"name": "carol", "active": true}

]

jq command:

jq '.[] | select(.active == true) | .name'

# Or, more idiomatically:

jq -r '.[] | select(.active) | .name'

Output:

alice

carol

9.4 HANDS-ON WORKSHOP: BUILDING

A GITHUB API CLIENT

Let’s build a simple script that fetches a list of public repositories for a given GitHub user and displays their names and star counts.

#!/bin/bash

set -euo pipefail

# Check for username argument

if [[ $# -ne 1 ]]; then

echo "Usage: $0 " >&2

exit 1

fi

USERNAME="$1"

# The GitHub API endpoint

API\_URL="https://api.github.com/users/${USERNAME}/repos"

echo "Fetching repositories for user: ${USERNAME}..."

# Call the API with curl, pipe to jq for processing

# We use jq to create a custom, tabular output.

curl -s "$API\_URL" | jq -r '

# For each element in the input array...

.[] |

# ...create a string by combining the name and the star count, separated by a tab. "\(.name)\t\(.stargazers\_count) stars"

'

This single, powerful pipeline replaces complex parsing logic and

produces clean, human-readable output.

9.5 A BRIEF LOOK AT XML AND

XMLSTARLET

While JSON is dominant, you will still encounter XML, especially in older enterprise systems or configuration files (e.g., Maven, Ant). xmlstarlet is the jq equivalent for XML. Its syntax is different and based on XPath, but the principle is the same.

Example: Extracting a value from an XML document. XML file pom.xml:

4.0.0

com.example

my-app

1.0-SNAPSHOT

Command: xmlstarlet sel -t -v "/project/version" pom.xml Output: 1.0-SNAPSHOT

# CHAPTER SUMMARY

You have successfully connected your shell to the modern, API-driven web. You are no longer limited to local files and processes. With curl, you can communicate with any HTTP endpoint, and with the powerful query language of jq, you can parse and transform the resulting JSON data with ease. You have learned to build pipelines that fetch, filter, and format data from web services, turning complex API responses into simple, usable information. This skill set is indispensable for modern automation, cloud management, and DevOps tasks.

# EXERCISES

1. Weather Report: Find a free weather API (like OpenWeatherMap or

WeatherAPI.com). Write a script that takes a city name as an argument, calls the API with curl, and uses jq to print a human-readable weather report, e.g., “Current weather in London: 15°C, clear sky.”

2. jq Transformation: Given a JSON array of objects, each with id, name,

and email, write a jq command that transforms it into a new JSON object where the keys are the user IDs and the values are the user emails.

3. API Error Handling: Modify the GitHub client from the workshop.

Check the HTTP status code from curl. If it’s not 200, print an informative error message (e.g., “Error: User not found (404)”) and exit. (Hint: Use curl -w '%{http\_code}').

10. SECURE SCRIPTING PRACTICES –

BUILD A FORTRESS, NOT A SHED

INTRODUCTION: FROM “DOES IT

WORK?” TO “IS IT SECURE?”

A script that accomplishes its task is good. A script that accomplishes its task without accidentally compromising your entire system is professional. In today’s interconnected world, every line of code that runs on a server is a potential point of entry. An carelessly written script can leak passwords, execute unauthorized code, or delete sensitive data.

In this chapter, we will fundamentally change our mindset. We will stop

building scripts like open sheds that anyone can wander into and start building digital fortresses. Security is not a feature you add at the end; it is the foundation upon which everything else is built.

We will illuminate the most common and dangerous security pitfalls in

shell scripts and give you the tools and techniques to avoid them. You will learn to guard secrets like a spy, treat user input like a suspicious border agent, and operate with the least privilege necessary. Mastering these practices is not just a technical skill—it is a professional responsibility.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Securely manage secrets (passwords, API keys) without ever storing them in your code.

• Understand the deadly threat of command injection and prevent it through rigorous quoting and validation.

• Create temporary files and directories safely and without race conditions using mktemp.

• Understand and use umask to control the default permissions of newly created files.

• Apply the Principle of Least Privilege to minimize the potential damage of your scripts.

10.1 THE FIRST COMMANDMENT:

NEVER STORE SECRETS IN A SCRIPT

This is the most common and dangerous mistake. An API token or database password hard-coded in a script will inevitably end up in a Git repository, a backup, or a public pastebin.

Method 1: Environment Variables (The Standard Way) Secrets should

be passed to the script at runtime via the environment.

# BAD:

DB\_PASSWORD="MySuperSecretPassword123"

mysql—user=admin—password="$DB\_PASSWORD" -e "..."

# BETTER: The script expects the variable to already exist.

# Check if the variable is set, otherwise exit safely.

: "${DB\_PASSWORD:?Error: The DB\_PASSWORD environment variable is not set.}"

mysql—user=admin—password="$DB\_PASSWORD" -e "..."

# Calling the script:

$ export DB\_PASSWORD="MySuperSecretPassword123"

$ ./my\_script.sh

# Or for a single invocation:

$ DB\_PASSWORD="..." ./my\_script.sh

Method 2: Configuration or Credential Files Many tools use

configuration files in the home directory (e.g., ~/.aws/credentials, ~/.my.cnf). you can emulate this pattern.

• Create a file ~/.my\_app.conf with the content API\_TOKEN=... • Important: Set extremely restrictive permissions: chmod 600 ~/.my\_app.conf. • In the script: source ~/.my\_app.conf

Method 3 (Professional): Secret Management Systems In production

environments, tools like HashiCorp Vault or AWS Secrets Manager are used. A script would authenticate with these services and fetch secrets at runtime.

10.2 THE PLAGUE OF COMMAND

INJECTION

This occurs when a script accepts user input (e.g., a filename) and carelessly uses it in a command. This allows an attacker to inject their own code.

The Disastrous Example: A script meant to delete a user-specified report.

report\_name="$1" rm /var/www/reports/$report\_name

An attacker calls the script like this: ./delete\_report.sh "dummy.txt; rm -rf /" The

shell executes two commands: rm /var/www/reports/dummy.txt and rm -rf /!

The Cures:

1. Always, always, always quote! rm "/var/www/reports/$report\_name" Now, the

entire string dummy.txt; rm -rf / is passed as a single filename to rm, resulting in a “file not found” error—the system is safe.

2. Validate your input. Only allow characters that make sense in a

filename. bash if ! [[ "$report\_name" =~ ^[a-zA-Z0-9.\_-]+$ ]]; then echo "Error: Invalid filename." >&2 exit 1 fi

3. Never use eval with user input. eval is the most powerful and dangerous

tool in the shell. It interprets a string as code to be executed. If that string comes from an external source, it’s an open invitation for compromise.

10.3 THE MINEFIELD OF TEMPORARY

FILES: MKTEMP IS YOUR FRIEND

Never, under any circumstances, should you create a temporary file with a fixed or predictable name.

BAD: tmp\_file="/tmp/my\_script\_$$" Between the creation and use of this file, an

attacker can replace it with a symbolic link to a critical system file (like /etc/passwd). This is known as a race condition attack.

The Only Correct Solution: mktemp mktemp creates a file (or directory)

with a guaranteed unique name and secure permissions, and returns the name.

# Create a secure temporary file

tmp\_file=$(mktemp)

# Create a secure temporary directory (often even better)

tmp\_dir=$(mktemp -d)

# IMPORTANT: Combine with `trap EXIT` for cleanup!

trap 'rm -rf "$tmp\_dir"' EXIT

# Work safely inside this directory

...

10.4 PERMISSIONS AND THE

PRINCIPLE OF LEAST PRIVILEGE

1. umask: The Default Birth Certificate for Files The umask (user file-creation mode mask) determines which permissions a newly created file does not get. A security-critical script should set a restrictive umask at the beginning.

# Set a very restrictive umask for this script

# 077 means: Only the owner has read/write/execute permissions.

# Group and Others have NO rights.

umask 077

touch new\_file # This file will be created with 600 (-rw———-) permissions

mkdir new\_directory # This directory will be created with 700 (drwx———)

2. The Principle of Least Privilege A script should always run with the

minimum privileges necessary to do its job. \* A backup script that only needs to back up the files of a specific user (webuser) should not run as root. It should be run as webuser. \* Use sudo -u my\_script.sh to run a script as another user. \* Avoid giving scripts unnecessary setuid permissions.

# CHAPTER SUMMARY

You have learned that security is not a feature, but a fundamental design decision. You now know how to avoid the cardinal sins of script security: you never store secrets directly in code, instead loading them safely from the environment. You treat all external input with extreme suspicion and prevent command injection through rigorous quoting and validation. You now create temporary files only with mktemp and reliably clean them up with trap. By consciously using umask and adhering to the principle of least privilege, you ensure your scripts can do as little potential damage as possible. You are no longer just building scripts—you are building fortresses.

# EXERCISES

1. Refactoring an Insecure Script: Take a script that has a hard-coded

API key. Change it to read the key from an environment variable. The script should exit with an error if the variable is not set.

2. Hardening a Script: You are given the following script for a code

review. Identify and fix at least three distinct security vulnerabilities. bash #!/bin/bash # This script compresses a folder specified by the user TARGET\_DIR=$1 OUTPUT\_FILE="/tmp/backup.tar.gz" tar -czf $OUTPUT\_FILE $TARGET\_DIR echo "Backup created in $OUTPUT\_FILE"

3. Secure Processing Pipeline: Write a script that downloads a large file,

unzips it in a secure temporary directory, and then processes it. The script must guarantee that the temporary directory is deleted under all circumstances (success, error, or interruption).

11. SCRIPTING THE NETWORK – THE

SHELL AS A NETWORK ENGINEER

INTRODUCTION: FROM LOCALHOST

TO THE GLOBAL NETWORK

By now, your scripts have mastered the local machine. They can manipulate files, control processes, and transform data. But the true power of modern systems lies in their interconnectedness. Servers communicate with each other, services offer their functionality over network ports, and administration often requires access to dozens or hundreds of remote machines.

In this chapter, we will break the boundaries of localhost and turn your shell

into a fully-fledged networking tool. You will learn that you don’t need complex programming languages to check network services, transfer data securely, or automate repetitive tasks on remote servers. The tools are already there, waiting for you to conduct them.

We will go beyond a simple ping and speak directly to TCP and UDP ports

with netcat. We will use nmap not as a manual hacking tool, but as a scripted inventory and monitoring assistant. And most importantly, we will master the art of SSH automation to enable passwordless, secure, and scalable remote administration.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Actively and programmatically check network ports and services with netcat (nc) and nmap.

• Use the basics of nmap for automated network discovery and port scanning. • Set up and securely use passwordless SSH connections with key-based authentication.

• Automate repetitive tasks on many servers in parallel using SSH loops. • Transfer files securely and automatically with scp and rsync. • Understand and use SSH tunneling (port forwarding) to secure unsafe services.

11.1 THE NETWORK’S PULSE:

QUERYING SERVICES WITH NETCAT

AND NMAP

netcat (nc): The TCP/IP Swiss Army Knife netcat is a simple yet extremely powerful tool for creating raw TCP or UDP connections. It’s perfect for testing if a service is listening on a specific port. • -z: Zero-I/O mode. Only checks if the port is open, without sending any data.

• -v: Verbose. Prints more information. • -w timeout: Sets a timeout in seconds.

Example: Check if a web server and a mail server are reachable.

#!/bin/bash

HOST="example.com"

WEB\_PORT=80

MAIL\_PORT=25

echo "Checking host: $HOST"

# '-w 3' sets a 3-second timeout

if nc -z -v -w 3 "$HOST" "$WEB\_PORT"; then

echo "Web server on port $WEB\_PORT is reachable."

else

echo "Web server on port $WEB\_PORT is NOT reachable."

fi

nmap: The Network Scanner for Scripts While nc tests one port, nmap can

scan entire networks and thousands of ports. For scripting, the “grep-able” output mode (-oG) is particularly useful.

Example: Find all hosts in a subnet that have the SSH port (22) open.

SUBNET="192.168.1.0/24"

# -p 22: Scan only port 22

# -oG -: Output in grep-able format to stdout

# The '-' as a filename means stdout

nmap -p 22 -oG - "$SUBNET" | awk '/Host: .\* Status: Up/{print $2}'

This output can be directly processed in a loop to, for example, execute a

command on all of these hosts.

11.2 AUTOMATION WITH SSH: THE

KEYS TO THE KINGDOM

Repeatedly typing passwords is the death of automation. The solution is key-based authentication.

Step 1: Create a Key Pair (if you don’t have one) ssh-keygen -t ed25519 (or

rsa). Accept the defaults.

Step 2: Copy Your Public Key to the Target Server ssh-copy-id user@target-

server This command securely appends your public key (~/.ssh/id\_ed25519.pub) to the ~/.ssh/authorized\_keys file on the target server. You will have to enter your password one last time.

Step 3: Connect Without a Password From now on, you can connect

with ssh user@target-server without a password prompt.

Automation with a for loop: Now you can execute commands across a

whole fleet of servers.

#!/bin/bash

SERVER\_LIST=("web-01" "web-02" "api-01")

for server in "${SERVER\_LIST[@]}"; do

echo "-—Updating server: $server—-"

# Execute 'apt update' and 'apt upgrade' on the remote server

# -t forces TTY allocation, which can be useful for interactive commands like 'apt' ssh -t "user@${server}" "sudo apt update && sudo apt upgrade -y" echo "——————————————————"

done

11.3 SECURE DATA TRANSFER: SCP

AND RSYNC

• scp (Secure Copy): Simple and effective for single files. # Copy a local file to a server

scp /local/path/file.txt user@server:/remote/path/

# Download a file from a server

scp user@server:/remote/path/file.txt /local/path/

• rsync: More powerful and intelligent than scp. rsync only transfers the changed parts of files, can recursively synchronize directories, and much more. It is the preferred method for backups and deployments. # Synchronize a local directory with a remote one

# -a: archive mode (recursive, preserves permissions, etc.)

# -v: verbose

#—delete: deletes files in the destination that no longer exist in the source rsync -av—delete /path/to/project/ user@server:/path/to/deployment/

11.4 SSH TUNNELING: A SECURE

CORRIDOR THROUGH THE INSECURE

NET

Imagine you have a database server whose port (e.g., 5432) must not be accessible over the internet for security reasons. But you need to access it from your laptop. An SSH tunnel (also called port forwarding) is the solution.

Local Port Forwarding (-L) This command opens a port on your local

machine and securely forwards all traffic sent there through the SSH connection to the destination server, and from there on to the target service.

Syntax: ssh -L LOCAL\_PORT:TARGET\_HOST:TARGET\_PORT BASTION\_HOST Example: ssh -L 8000:db-server.internal:5432 user@bastion-host.com

• What happens?

1. You establish a normal SSH connection to bastion-host.com.

2. SSH opens port 8000 on your laptop (localhost).

3. When you now connect to localhost:8000 with a database client, SSH

intercepts this connection.

4. The traffic is sent, encrypted, through the tunnel to bastion-host.com.

5. The bastion-host.com server decrypts the traffic and forwards it to db-

server.internal:5432.

You can now work with your local database tool as if the database were

on your own machine, even though it’s safely behind a firewall.

# CHAPTER SUMMARY

Your shell is no longer confined to its own machine. You have learned to feel the pulse of the network with netcat and nmap and to actively monitor services. By mastering key-based SSH authentication, you can now manage dozens of servers as easily as one, and with rsync, you can synchronize data intelligently and efficiently. With the knowledge of SSH tunneling, you can even secure unsafe services and create secure access to internal resources. You are now capable of writing scripts that control and automate not just single systems, but entire network infrastructures.

# EXERCISES

1. Port Scanner: Write a script that takes a list of hostnames and a list of

ports. The script should check each port on each host using nc and print a clear table showing which port is “open” or “closed” on which host.

2. Parallel uptime: Take a list of servers to which you have passwordless

SSH access. Write a script that logs into all servers in parallel, executes the uptime command, and displays the collected output.

3. Automated Database Backup: Write a script that establishes an SSH

tunnel to a bastion host to access an internal database. It should then use the pg\_dump (or mysqldump) command to back up the database through the tunnel and save the backup file locally. The tunnel should be torn down after the backup is complete. (Hint: Start SSH in the background with -fN and save the PID to kill it later).

12. PERFORMANCE TUNING YOUR

SCRIPTS – FROM A TRACTOR TO A

RACE CAR

INTRODUCTION: WHEN “IT WORKS”

IS NO LONGER FAST ENOUGH

Your script does what it’s supposed to do. It’s robust, secure, and well-structured. But when it has to work with 10,000 files instead of 10, it suddenly takes hours instead of seconds to run. In the world of automation and data processing, speed is often not a luxury, but a necessity. A backup script that runs too long could impact a production system. A data processing job that takes hours could delay critical business decisions.

This chapter is your pit stop. Here, you will learn to transform your

scripts, which have been reliable tractors until now, into sleek race cars. We will uncover a fundamental but crucial truth about the shell: its greatest strength—the ability to call other commands—is also its greatest performance weakness. Every external command is an expensive pit stop.

We will learn how to pinpoint bottlenecks with professional tools like time

and strace. You will master the art of replacing external commands with lightning-fast, built-in shell operations. And you will discover how to offload computationally intensive tasks to specialized tools like awk to drastically reduce the number of “pit stops.”

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Precisely identify performance bottlenecks in scripts with time and strace. • Understand the high cost of “forking”—why every external command call is expensive.

• Speed up scripts by consistently using shell builtins and parameter expansion.

• Recognize and avoid common performance traps like the “Useless Use of cat.”

• Rewrite computationally intensive loops into more performant awk scripts. • Recognize the limits of shell performance and know when switching to another language is the right move.

12.1 MEASURE, DON’T GUESS:

FINDING BOTTLENECKS WITH TIME

AND STRACE

Optimization without measurement is a waste of time. You need to know where your script is slow.

The time Command: The First Look Under the Hood The time

command measures the execution time of a command and breaks it down: \* real: The actual elapsed wall-clock time. \* user: The CPU time spent in user-space (in the process itself). \* sys: The CPU time spent in kernel-space (for system calls like reading/writing files).

$ time sleep 3

real 0m3.004s

user 0m0.000s

sys 0m0.003s

A high sys value often indicates many I/O operations or process startups. strace: The Microscope for System Calls strace logs every system call a

process makes. The -c option is invaluable: it counts the calls and provides a summary.

Example: The difference between a builtin and an external command.

# External command in a loop

$ strace -c -e 'trace=process' bash -c 'for i in {1..1000}; do /bin/true; done' % time seconds usecs/call calls errors syscall

———————-—————-————-————-—————————

100.00 0.001673 2 1000 clone

———————-—————-————-————-—————————

Total 0.001673 1000 1 total

# Shell builtin in a loop

$ strace -c -e 'trace=process' bash -c 'for i in {1..1000}; do :; done'

% time seconds usecs/call calls errors syscall

———————-—————-————-————-————————— ———————-—————-————-————-—————————

Total 0.000000 0 0 total

The result is dramatic: 1000 clone calls (the start of the forking process) for

the external command, but zero for the builtin.

12.2 THE #1 PERFORMANCE KILLER:

THE COST OF FORKING

Every time your script calls an external command (grep, sed, cut, ls...), the kernel must: 1. fork(): Create an exact copy of the shell process in memory. This is expensive. 2. exec(): Load the code of the new command into this copied process and execute it.

This process is orders of magnitude slower than an operation that happens

directly within the shell process.

The Golden Rule of Shell Performance: Minimize the number of

processes launched. Every fork avoided is a win.

12.3 PRACTICAL OPTIMIZATION

TECHNIQUES

1. Use Shell Builtins Instead of External Commands This is the most important technique. Many tasks for which beginners use external commands can be done with parameter expansion (see Chapter 3). • Bad (slow): filename=$(basename "$filepath") • Good (fast): filename="${filepath##\*/}" • Bad (slow): ext=$(echo "$filename" | cut -d . -f 2) • Good (fast): ext="${filename#\*.}"

2. Avoid the “Useless Use of cat” (and similar patterns) A classic

beginner mistake that creates an unnecessary process. • Bad (1 unnecessary process): cat file.txt | grep "pattern" • Good (efficient): grep "pattern" file.txt • Bad (2 unnecessary processes): cat file.txt | sort | uniq -c • Good (efficient): sort file.txt | uniq -c

3. Offload Heavy Work to awk Sometimes, a loop in the shell itself is the

bottleneck, especially for arithmetic. awk is an external command, but it’s only launched once and then runs its own, much faster, internal loop.

Scenario: Sum the numbers in a file.

The slow shell loop (never forks, but shell arithmetic is slow):

total=0

while read -r num; do

((total += num))

done < numbers.txt

echo "$total"

The fast awk solution (forks once, but awk is optimized for numbers):

awk '{ total += $1 } END { print total }' numbers.txt

For large files, the awk version can be 10 to 100 times faster.

12.4 I/O OPTIMIZATION: READING

DATA SMARTLY

Avoid reading the same file over and over. If you need to perform multiple operations on the data in a file, read it once into an array or a variable, then work with the data in memory.

• Bad (reads the file 3 times): bash count=$(wc -l < file.txt) first\_line=$(head -n 1 file.txt) last\_line=$(tail -n 1 file.txt)

• Good (reads the file 1 time): bash readarray -t lines < file.txt count=${#lines[@]} first\_line="${lines[0]}" last\_line="${lines[-1]}"

12.5 KNOWING WHEN TO STOP: THE

LIMITS OF THE SHELL

The shell is a fantastic “glue” for connecting commands and for I/O-heavy tasks. However, for CPU-intensive tasks (complex algorithms, heavy math, recursive calculations), it is the wrong tool for the job.

Rules of thumb for switching to another language (Python, Go, Perl,

Rust):

• When you need complex data structures that go beyond associative arrays. • When your script consists mainly of computationally intensive for or while loops.

• When performance is absolutely critical and even an optimized awk solution is too slow.

• When the script requires extensive error handling and testability that becomes cumbersome in the shell.

A good developer knows the strengths and weaknesses of their tools.

# CHAPTER SUMMARY

You have learned to see your scripts through the lens of a performance engineer. You now know that the key to speed is minimizing process startups (forks). You can pinpoint the slow parts of your code with time and strace. You actively replace external commands with fast shell builtins, avoid unnecessary pipes, and offload computationally intensive tasks to the optimized tool awk. You also understand that the shell has its limits and when it’s time to switch to a faster, compiled language. Your scripts are now not only robust and secure, but also as fast as they can be in the world of shell scripting.

# EXERCISES

1. Bottleneck Analysis: Find an older, slow script of yours. Use time and

strace -c` to identify which commands are called most frequently and consume the most time.

2. Refactoring for Performance: Rewrite a script that uses basename and

dirname in a loop over 1000 files. Replace the external calls with parameter expansion and measure the speed difference with time.

3. Shell vs. awk: Write a script that calculates the average value of all

numbers in a large file, once with a pure shell loop and once with awk. Compare the execution times.

13. CREATING PROFESSIONAL

COMMAND-LINE TOOLS – MORE

THAN JUST A SCRIPT

INTRODUCTION: FROM A PRIVATE

HELPER TO A PUBLIC UTILITY

So far, you have written scripts that accomplish a task. You know how to call them and what arguments they expect. But what if you are creating a tool that others will use? Or what if your script becomes so complex that it needs different modes of operation, configuration options, and flags? At this point, a simple my\_script.sh arg1 arg2 is no longer sufficient.

Users expect behavior they are accustomed to from standard tools like ls,

grep, or git: clearly defined options (-v, —verbose), a help output (—help), and clean error handling for incorrect input.

In this chapter, we complete the transformation from scripter to tool-

smith. You will learn how to parse command-line arguments robustly and flexibly using the POSIX standard getopts. We will design a professional directory structure for your tools that cleanly separates code, libraries, and documentation. And finally, you will learn how to write man pages (manual pages)—the official form of documentation in the Unix world—so that your own creations integrate seamlessly into the system.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Robustly parse complex command-line options (flags, arguments) with the getopts builtin.

• Professionally implement standard options like —help and —version. • Design a sensible and scalable project structure for your tools (bin/, lib/, man/). • Write and install your own man pages for your scripts. • Create a clear and helpful user experience (UX) for your command-line tools.

13.1 ARGUMENT PARSING: GETOPTS

INSTEAD OF $1 CHAOS

Processing arguments with $1, $2, etc., quickly becomes messy and error-prone. What if options come in a different order? What about combined flags like -vf? getopts is the built-in, POSIX-compliant solution.

The getopts Loop: getopts is used in a while loop that runs as long as options

are found.

Syntax: while getopts ":optstring" opt; do ... \* optstring: Defines the allowed

options. A colon after a letter means that option expects an argument (e.g., i:o:vh). The leading colon : enables silent error mode, which we want to handle ourselves. \* opt: A variable that holds the found option letter in each iteration. \* $OPTARG: Contains the argument of an option if one is required. \* $OPTIND: The index of the next argument to be processed.

Example: A script that accepts an input file, an output file, and a verbose

mode.

#!/bin/bash

# Default values

VERBOSE=false

OUTPUT\_FILE=""

INPUT\_FILE=""

usage() {

echo "Usage: $0 -i [-o ] [-v] [-h]"

}

while getopts ":i:o:vh" opt; do

case $opt in

i) INPUT\_FILE="$OPTARG" ;;

o) OUTPUT\_FILE="$OPTARG" ;;

v) VERBOSE=true ;;

h) usage; exit 0 ;;

\?) echo "Invalid option: -$OPTARG" >&2; usage; exit 1 ;;

:) echo "Option -$OPTARG requires an argument." >&2; usage; exit 1 ;;

esac

done

# Remove the processed options from the arguments

shift $((OPTIND - 1))

# The rest of the arguments are now in $@

# e.g., for filenames at the end

REMAINING\_ARGS=("$@")

# Main logic of the script...

echo "Input: $INPUT\_FILE"

echo "Output: $OUTPUT\_FILE"

echo "Verbose: $VERBOSE"

echo "Remaining args: ${REMAINING\_ARGS[\*]}"

13.2 STANDARD OPTIONS: —HELP, —

VERSION, AND FRIENDS

A professional tool should always respond to —help and —version. Since getopts does not support long options (—...), we use a preceding case statement or a for loop construct to handle them before getopts starts.

Extending the example above:

#!/bin/bash

VERSION="1.0.0"

usage() { ... }

version() { echo "$0 Version $VERSION"; }

# Handle long options manually

for arg in "$@"; do

shift

case "$arg" in

"—help") usage; exit 0 ;;

"—version") version; exit 0 ;;

\*) set—"$@" "$arg"

esac

done

# Now, the getopts loop for short options...

while getopts ...

Note: The external command getopt (without ‘s’) supports long options, but it is not POSIX-standardized and behaves differently on various systems (Linux vs. BSD/macOS). For maximum portability, manual handling or forgoing long options is the safer path.

13.3 PROFESSIONAL PROJECT

STRUCTURE

Stop cramming everything into one file. A good structure promotes maintainability and reusability.

A Typical Structure:

my-tool/

├── bin/

│ └── my-tool # The executable main script

├── lib/

│ └── utils.sh # A library with helper functions

├── man/

│ └── my-tool.1 # The man page

└── README.md

The main script in bin/my-tool would then source the library:

#!/bin/bash

# bin/my-tool

# Source the library relative to the script's path

source "$(dirname "$0")/../lib/utils.sh"

# ... rest of the script ...

This structure makes installation easier (e.g., copying bin/ to /usr/local/bin)

and keeps the code cleanly separated.

13.4 THE ART OF DOCUMENTATION:

WRITING MAN PAGES

The man page is your tool’s business card within the system. It is formatted with groff macros. It looks cryptic at first but follows a simple pattern.

Basic groff Macros: \* .TH TITLE SECTION: The title header (e.g., .TH MY-TOOL

1). \* .SH NAME: A new section header with the name NAME (e.g., .SH SYNOPSIS). \* .B text: Print text in bold. \* .I text: Print text in italics. \* .P: A new paragraph. \* .IP tag indent: An indented paragraph, useful for option descriptions.

Example for man/my-tool.1:

.TH MY-TOOL 1 "October 2023" "1.0.0" "User Commands"

.SH NAME

my-tool \- an example tool for demonstration

.SH SYNOPSIS

.B my-tool

[\-h] [\-v] [\-o file] \-i file [argument...]

.SH DESCRIPTION

.B my-tool

is a script that demonstrates the professional creation of command-line tools. It reads from an input

file and writes...

.SH OPTIONS

.IP "\-i,—input file"

Specifies the input file to read (required).

.IP "\-o,—output file"

Specifies the output file to write (optional).

.SH AUTHOR

Your Name

Displaying the local man page: man -l ./man/my-tool.1

# CHAPTER SUMMARY

You have taken the final step from a simple script to a full-fledged, professional command-line tool. You can now use getopts to create robust and flexible interfaces for your users that match the behavior of standard Unix commands. You structure your projects cleanly into separate directories for executables, libraries, and documentation. With the ability to write man pages, you ensure that your tools are first-class citizens, well-documented and seamlessly integrated into the system environment. Your creations no longer feel like foreign objects, but like a native part of the operating system.

# EXERCISES

1. Refactoring with getopts: Take an older script of yours that accepts

multiple arguments. Refactor it to use getopts to process options like -i , -o , and -f (force). Add a —help function.

2. Writing a man Page: Write a complete man page for the script you

refactored in exercise 1. Include the sections NAME, SYNOPSIS, DESCRIPTION, OPTIONS, and EXAMPLES.

3. Building a Structured Tool: Create a new tool with the bin/, lib/, man/

structure. The tool should perform a small task (e.g., count words in a text file). The counting logic should reside in a function in the lib/ library. The main script in bin/ should only handle argument parsing and calling the library function.

14. INTEGRATING WITH A WIDER

ECOSYSTEM – THE SHELL AS THE

ENGINE OF AUTOMATION

INTRODUCTION: FROM A TOOL TO

AN AUTOMATED FACTORY

Until now, you have viewed scripts as tools that you invoke manually to accomplish a specific task. You are the craftsman reaching for a hammer. In modern software development and systems administration, however, the most important processes are no longer initiated by hand. They run automatically, triggered by events: a git push, a new pull request, or the need to deploy a new version of an application.

In this chapter, we will transform your tools into the engines of an

automated factory. We will see how your shell scripts come to life and act as an integral part of the three pillars of modern DevOps workflows: 1. Local Development (Git Hooks): Automation directly on the developer’s machine to ensure quality before code even leaves the system. 2. Continuous Integration/Continuous Deployment (CI/CD): The heart of team automation, where code is automatically built, tested, and delivered. 3. Containerization (Docker): The packaging and delivery of applications in standardized, portable environments.

You will discover that the shell is not just “nice to have” here, but the

universal glue and the executive power that drives all these systems. The ability to write robust scripts for these ecosystems is one of the most sought-after skills for a modern administrator and developer.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Create Git Hooks to enforce code quality locally and automatically. • Understand and implement the role of shell scripts in CI/CD pipelines (e.g., GitHub Actions).

• Learn best practices for writing ENTRYPOINT and CMD scripts in Dockerfiles. • Understand the difference between the “shell form” and the “exec form” in Dockerfiles and its impact on signal processing.

14.1 THE LOCAL GUARDIANS:

QUALITY ASSURANCE WITH GIT

HOOKS

Git Hooks are scripts that Git automatically executes when certain events occur (e.g., before a commit or before a push). They reside in the .git/hooks/ directory of every repository.

The pre-commit Hook: A Final Check Before Checking In This is the

most popular hook. It runs before a commit is created. If the script exits with a non-zero exit code, the commit is aborted. It’s the perfect place to run linters like shellcheck.

Example: A pre-commit hook that checks all modified shell scripts with

shellcheck.

Create the file .git/hooks/pre-commit and make it executable:

#!/bin/bash

set -eo pipefail

echo "Running pre-commit hook..."

# Find all .sh files that are "staged" for the commit

STAGED\_SH\_FILES=$(git diff—cached—name-only—filter=ACMR | grep '\.sh$' || true) if [[ -z "$STAGED\_SH\_FILES" ]]; then

echo "No shell scripts to check."

exit 0

fi

echo "Checking the following shell scripts with shellcheck:"

echo "$STAGED\_SH\_FILES"

# Run shellcheck on each file. If one fails, the script exits due to 'set -e'. echo "$STAGED\_SH\_FILES" | xargs -n 1 shellcheck

echo "All scripts are clean. Proceeding with commit."

exit 0

With this hook in the repository, no team member can accidentally

commit a faulty shell script again.

14.2 THE ASSEMBLY LINE: SHELL

SCRIPTS IN CI/CD PIPELINES

CI/CD platforms like GitHub Actions, GitLab CI, or Jenkins are essentially highly sophisticated execution environments for shell scripts. Their YAML configuration files describe which commands should be run, when, and in which environment.

GitHub Actions as an Example: A workflow is defined in a .yml file

under .github/workflows/.

Example: A simple CI pipeline that runs linters and tests on every push.

File: .github/workflows/ci.yml

name: Linux Shell CI

on: [push, pull\_request]

jobs:

build-and-test:

runs-on: ubuntu-latest

steps:

- name: Check out code

uses: actions/checkout@v3

- name: Run ShellCheck Linter

run: |

echo "Checking all shell scripts in the repository..."

# Find all .sh files and run shellcheck on them

find . -type f -name "\*.sh" -print0 | xargs -0 shellcheck

- name: Run Tests

run: |

echo "Running the test suite..."

# Here you would call your test framework

# For our example, we just call a test script

./tests/run\_tests.sh

Each run: block is essentially a small, embedded shell script. Complex

logic should be outsourced to separate script files to keep the YAML file clean.

14.3 THE PACKAGING: SHELL

SCRIPTS IN DOCKER CONTAINERS

Docker is the de facto standard for deploying applications. Shell scripts play a crucial role here.

RUN vs. CMD vs. ENTRYPOINT \* RUN: Executes commands during the image

build. Used for installing packages and setting up the environment. \* ENTRYPOINT: Defines the main command that is always executed when the container starts. \* CMD: Provides default arguments for the ENTRYPOINT. Can be easily overridden by the user when starting the container (docker run ...).

The “exec form” vs. the “shell form” – A Critical Difference \* Shell

form: CMD my\_program -arg Docker starts a shell (/bin/sh -c) which then starts your program. Your program does not run as PID 1 in the container, and signals (like docker stop) are not forwarded correctly. This is almost always the wrong choice. \* Exec form: CMD ["my\_program", "-arg"] Docker starts your program directly. It runs as PID 1 and receives signals correctly. This is the preferred method.

The ENTRYPOINT Script Pattern An entrypoint.sh script is a common

practice to configure a container before the main application starts.

Example: An entrypoint.sh script that waits for a database.

#!/bin/bash

set -e

# Wait for the database to be ready

until nc -z -v -w30 "$DB\_HOST" "$DB\_PORT"; do

echo "Waiting for database..."

sleep 1

done

echo "Database is ready. Starting application..."

# exec "$@": Replaces the shell process with the command

# passed as arguments (from the Dockerfile's CMD).

# This is crucial for proper signal handling!

exec "$@"

The Corresponding Dockerfile:

FROM ubuntu:22.04

# ... install netcat, etc.

COPY entrypoint.sh /usr/local/bin/

RUN chmod +x /usr/local/bin/entrypoint.sh

# The entrypoint script is always executed

ENTRYPOINT ["entrypoint.sh"]

# The CMD command is passed as arguments ($@) to the entrypoint script

CMD ["/app/my\_application", "—config", "/etc/app.conf"]

# CHAPTER SUMMARY

You have seen that your shell scripting skills are at the center of modern software development cycles. You can now integrate automatic quality checks directly into the development process with Git Hooks. You understand how CI/CD pipelines orchestrate shell commands to build and test code. And you know how to configure Docker containers with robust ENTRYPOINT scripts that ensure clean initialization and proper process management. Your scripts are no longer just isolated tools; they are the indispensable cogs in the engine of modern IT automation.

# EXERCISES

1. Create a pre-push Hook: Create a Git hook that runs before a git push. This

hook should execute a test script (./run\_tests.sh) and only allow the push if the tests are successful (exit code 0).

2. Design a Build Pipeline: Create a GitHub Actions or GitLab CI

pipeline for one of your script projects. The pipeline should have two stages: build and test. The build stage could, for example, create a man page, and the test stage would then run the tests.

3. Dockerizing a Tool: Take one of the command-line tools you designed

in Chapter 13. Write a Dockerfile that packages this tool into a minimal image (e.g., based on alpine) so that it can be called with docker run my-tool-image—help.

15. EXTENDING THE SHELL AND

LOOKING AHEAD – THE MASTER’S

HORIZON

INTRODUCTION: THE JOURNEY IS

COMPLETE, THE LEARNING

CONTINUES

Congratulations! You have reached the end of this book. You are no longer just a user typing commands; you are an architect of automation solutions. You can design robust, secure, portable, and performant scripts that integrate seamlessly into professional development and operations environments. You have mastered the shell as the powerful tool that it is.

But a true master knows not only the strengths of their tool but also its

limitations. They know when to use it and, more importantly, when to set it aside and reach for a more specialized tool.

This final chapter is your compass for the journey ahead. We will take a

brief look beyond the horizon at other powerful shells that might be better suited for certain tasks. We will answer the crucial question: “When is a shell script no longer the right solution?”. And finally, we will design a final, comprehensive project where you can consolidate and apply all your acquired knowledge. This is the last step on your path from journeyman to master— the development of judgment.

# LEARNING OBJECTIVES FOR THIS

CHAPTER:

• Understand the key advantages of other shells like zsh and fish for scripting. • Recognize the red flags that signal a problem is becoming too complex for a shell script.

• Know when to switch to a higher-level programming language like Python or Go.

• Design and implement a comprehensive final project that combines the concepts of the entire book.

• Place your skills in the context of the broader IT landscape.

15.1 A LOOK BEYOND THE HORIZON:

ZSH AND FISH

Bash is the universal standard, but other shells offer interesting alternatives, especially for scripting.

• zsh (Z Shell): The “Power-User” Bash zsh is largely backward-compatible with Bash but offers a wealth of improvements that are also useful for scripting:

◦ More Powerful Globbing Patterns: ls \*\*/\*.log(.) finds only regular files; ls \*\*/\*(/) finds only directories.

◦ Better Array Handling: Indexes start at 1 by default (can be more intuitive), and slicing is easier.

◦ Floating-Point Arithmetic: zsh natively supports calculations with decimal numbers.

• fish (Friendly Interactive Shell): The User-Friendly Alternative fish intentionally breaks with POSIX compatibility to offer a simpler and more consistent syntax.

◦ Simpler Syntax: Less need for cryptic quoting. for i in (ls); ... just works. ◦ Lists Instead of Strings: Variables can hold true lists of values, making IFS hacks obsolete.

◦ Disadvantage: Scripts are not portable and won’t run on systems without fish.

Conclusion: For scripts that require maximum portability, #!/bin/sh

(POSIX) remains the gold standard. For complex, internal tools on systems where zsh is standard, its features can be a blessing.

15.2 THE ART OF KNOWING WHEN TO

STOP: WHEN THE SHELL IS THE

WRONG TOOL

This is the most important lesson for an advanced scripter. A shell script is the wrong choice when:

1. Data structures become too complex. We learned to simulate objects

with associative arrays. But if you need nested structures, trees, or graphs, the code quickly becomes an unreadable and error-prone nightmare.

2. Performance is CPU-critical. The shell is fantastic for I/O-heavy tasks

(reading files, starting commands). For computationally intensive algorithms (complex math, image processing, large in-memory data transformations), its interpreted nature and process overhead are far too slow.

3. You need extensive external libraries. Do you need to communicate

with a specific database via a native driver? Build a complex web application? Access a vast ecosystem of machine learning libraries? This is what languages with package managers (like Python’s pip or Node’s npm) are for.

4. Robust, automated testing is a must. Although you can test shell

scripts (e.g., with bats), the testing frameworks and the culture of unit testing are far more mature in languages like Python, Go, or Java.

15.3 THE RIGHT TOOL FOR THE JOB:

PYTHON, GO, AND BEYOND

When you hit the limits of the shell, these are the logical next steps: • Python: The de facto standard for system administration, DevOps, and data “glue” code. It’s easy to learn, has a huge standard library (os, subprocess, json), and a gigantic ecosystem. It’s the perfect all-rounder and often the next step after the shell.

• Go (Golang): The best choice for high-performance, concurrent command-line tools and network services. It compiles to a single, static binary that runs on any system without dependencies—a huge advantage for deployment. • Perl: The original “Swiss-army chainsaw” for text processing. Although often superseded by Python for new projects, its ability to process regular expressions and text is still legendary and unparalleled.

15.4 FINAL PROJECT: A MODULAR

DEPLOYMENT TOOL

It’s time to bring it all together. Your task is to create a simplified deployment tool called deploy-mate that deploys an application from a Git repository to one or more servers.

Requirements:

1. Configuration (lib/config.sh): The tool reads its configuration from a file

(/etc/deploy-mate.conf or ~/.deploy-mate.conf) which is stored in an associative array (Chapter 6).

2. Command-Line Interface (bin/deploy-mate): The tool accepts arguments

like deploy-mate -a -e (e.g., staging or production) and a — dry-run flag (Chapter 13).

3. Error Handling and Logging (lib/utils.sh): The tool uses trap EXIT for

cleanup and logs important steps and errors with a timestamp (Chapters 4, 5).

4. Network Automation (lib/ssh\_helpers.sh): It connects via passwordless

SSH to the servers defined in the configuration (Chapter 11).

5. Deployment Logic (lib/deploy.sh): ◦ Creates a secure temporary directory on the target server (mktemp) (Chapter 10).

◦ Clones or updates the application’s Git repository. ◦ Runs build steps (e.g., npm install). ◦ Synchronizes the built application to the target directory with rsync. ◦ Restarts the application service (systemctl restart ...).

This project forces you to use almost every concept from this book and

combine them into a coherent, professional tool.

# SUMMARY AND FINAL WORDS

Your journey through advanced shell scripting has now come to an end. You have acquired the tools and, more importantly, the mindset to solve complex automation problems. You understand the nuances of portability, security, performance, and modularity. You are capable of not just executing commands, but orchestrating systems.

The horizon is now wide open. Whether you continue to refine your shell

skills, dive into the world of Python and Go, or specialize in cloud automation, the foundation you have laid is solid. The shell will always be a part of your toolkit, as it is the lingua franca of all Unix-like systems.

Use your knowledge wisely. Automate the tedious, secure the vulnerable,

and build tools that make your life and the lives of others easier. Good luck on your continued journey!

# A FINAL WORD FROM THE AUTHOR

Congratulations!

You’ve made it. You are not just holding the end of this book in your

hands, but also the reward for countless hours of learning, experimenting, and understanding. If you have arrived here, you have completed a demanding journey that goes far beyond the basics. Take a moment to be proud of this milestone.

When we stood at the beginning of this book, I promised you we would

walk the path from craftsman to architect. You have learned to write not just functional scripts, but to design robust, secure, and thoughtful tools. You no longer think merely in commands, but in concepts like portability, error handling, modularity, and efficiency. This is a fundamental shift in how you solve problems on the command line—and this mindset is far more valuable than the knowledge of any single command.

The true sign of mastery, however, is not the feeling of knowing

everything, but the realization of how much there is still to discover. The world of IT never stands still. New tools emerge, and new challenges arise. Therefore, do not view the knowledge from this book as a destination, but as your springboard. You have built a solid foundation and—more importantly —the ability to independently learn and master new topics. Your most important skill is no longer just knowing how to do something, but possessing the curiosity and competence to tackle any new challenge that comes your way.

So, what’s next?

That answer lies with you. Build the tool you’ve always wished you had

at work. Automate the process that costs you and your colleagues the most time. Take a look at Python or Go to see how the concepts from this book are implemented in other languages. Share your knowledge with others who are at the beginning of their journey—because teaching is the best way to learn. Engage with others in forums and communities.

The command line is your canvas. Be creative, be curious, and never stop

learning.

It has been an honor and a great pleasure to accompany you on this path.

Thank you for your trust, your time, and your perseverance.

I wish you countless #!/bin/bash moments filled with success and the joy of

discovery.

All the best on your continued journey,

Michael Basler

# A COLLECTION OF USEFUL SHELL

ONE-LINERS

The true elegance of the shell is often revealed in its ability to solve complex problems with a single, well-thought-out command chain. This collection serves as a source of inspiration and a cheat sheet for everyday challenges. Many of these one-liners combine the concepts presented in this book into powerful tools.

# FILE AND DIRECTORY MANAGEMENT

1. Find the 10 largest files in the current directory and all

subdirectories:

find . -type f -printf "%s\t%p\n" | sort -rn | head -n 10

• Explanation: find prints the size (%s) and path (%p) of each file. sort -rn sorts this list numerically and in reverse order. head displays the top 10 results.

1. Rename all .jpg files to .png (bulk renaming):

for f in \*.jpg; do mv—"$f" "${f%.jpg}.png"; done

• Explanation: A for loop iterates over all .jpg files. The parameter expansion ${f%.jpg} removes the .jpg extension from the filename before appending the new .png extension. — protects against filenames that begin with a hyphen.

1. Find all duplicate files in a directory:

find . -type f -exec sha256sum {} + | sort | uniq -w 64 -d

• Explanation: sha256sum calculates a unique checksum for each file. sort sorts the output so that identical checksums appear next to each other. uniq -w 64 -d compares only the first 64 characters (the length of the SHA256 hash) and displays only the lines that are duplicates.

1. Create a directory and change to it immediately:

mcd() { mkdir -p "$1" && cd "$1"; }

• Explanation: This is a function for your .bashrc. mkdir -p creates the directory (and all parent directories, if necessary). Only if this succeeds (&&) is cd executed.

1. Pack all files older than 90 days into an archive:

find . -type f -mtime +90 -print0 | xargs -0 tar -czvf old\_files.tar.gz

• Explanation: find -mtime +90 finds the files. The combination -print0 and xargs -0 is the safest way to reliably pass filenames (even those containing spaces or special characters) to xargs.

# TEXT PROCESSING AND DATA

WRANGLING

1. Sum up all numbers in the first column of a file: awk '{s+=$1} END {print s}' file.txt

• Explanation: awk is optimized for column-based data. This one-liner adds the value of the first column ($1) of each row to the variable s and prints the total at the end (END).

1. Display the 10 most frequently used commands from your Bash

history:

history | awk '{print $2}' | sort | uniq -c | sort -rn | head -n 10

• Explanation: The canonical pipeline for frequency analysis: awk extracts the command, sort prepares for uniq, uniq -c counts, sort -rn sorts by frequency, and head displays the top 10.

1. Search and replace a string in all .txt files: find . -type f -name "\*.txt" -exec sed -i 's/old\_pattern/new\_pattern/g' {} + • Explanation: find finds the files. -exec sed -i ... {} + is a high-performance method for invoking sed because + bundles as many filenames as possible into a single sed invocation.

1. Query your own public IP address via a web service:

curl -s ifconfig.me

• Explanation: curl retrieves the contents of the web page. Services like ifconfig.me are specialized to return only your IP address as text. -s suppresses the loading bars.

1. Remove all blank lines from a file:

sed '/^$/d' file.txt > new\_file.txt • Explanation: The sed command applies the delete command (d) to all lines that match the pattern ^$ (an empty line).

# NETWORK AND PROCESS

MANAGEMENT

1. Start a quick, temporary web server in the current directory: python3 -m http.server 8000

• Explanation: This is a fantastic trick that uses Python’s built-in web server module to quickly share the files in the current directory over the network.

1. \*\*Find out which process is using a specific port bash sudo lsof -i :443 • Explanation: lsof -i lists all open internet files (network connections). The colon : indicates a port number.

1. \*\*Find all “zombie” processes and display their parent processes:

bash ps aux | awk ‘$8==“Z” {print “Zombie PID:”, $2, “Parent PID:”,

$3}’

• Explanation: ps lists all processes. awk filters for lines where the eighth column (status) is “Z” and displays a formatted message.

1. Watch the output of a command live, every second:

watch -n 1 'df -h'

• Explanation: watch executes the specified command repeatedly (every 2 seconds by default, -n 1 changes this to 1 second) and displays the output in fullscreen mode. Perfect for monitoring.

1. Kill all processes of a specific user:

pkill -u username

• Explanation: pkill is a more modern and secure alternative to constructs like ps | grep | xargs kill. -u filters by username.

This list is just the beginning. The real power lies in understanding

these patterns and creatively combining them to create new, unique solutions for your specific problems.

# INDEX

This index lists the most important commands, concepts, and variables covered in this book. Special symbols and operators can be found at the end of the index.

A

• awk

◦ BEGIN block

◦ END block

◦ Arithmetic

◦ Associative arrays in awk

◦ NF (Number of Fields)

◦ NR (Number of Records)

◦ as a performance tool

• API (Application Programming Interface) ◦ Authentication

◦ JSON processing

◦ REST

• Arguments, command-line

◦ $@

◦ $\*

◦ $1, $2, ...

◦ manual parsing

◦ see also getopts

• Array

◦ Declaration

◦ Adding elements

◦ Deleting elements (unset)

◦ Iteration

◦ Length (${#array[@]})

◦ see also Indexed Array, Associative Array • Associative Array (Hash)

◦ declare -A

◦ Iterating over keys ("${!array[@]}") ◦ Iterating over values ("${array[@]}") ◦ Key-value pairs

◦ Simulating objects

B

• Background Process (see Process Management) • basename (and portable alternative) • Bash

• Bash-isms (see Portability)

• BASH\_COMMAND

• BASH\_SOURCE

• Built-in (see Shell Built-in)

C

• cat

◦ Useless Use of cat

• cdspell (see shopt)

• CI/CD (Continuous Integration/Deployment) ◦ GitHub Actions

◦ GitLab CI

◦ YAML

• Command Injection

• Configuration files, parsing

• coproc (see Coprocess)

• Coprocess (coproc)

• curl

◦ -d (send data)

◦ -H (header)

◦ -X (HTTP method)

D

• dash (Debian Almquist Shell)

• Data Structures (see Array)

• Debugging

◦ set -x

◦ see also trap DEBUG

• declare

◦ -A (Associative Array)

◦ -n (Nameref)

• dirname (and portable alternative)

• Docker

◦ CMD

◦ Dockerfile

◦ ENTRYPOINT

◦ exec form vs. shell form

◦ RUN

E

• echo (and printf alternative)

• Environment Variable

• errexit (see set -e)

• Error Handling

◦ set -e

◦ set -o pipefail

◦ see also trap

• eval (and its dangers)

• exec

◦ in Docker entrypoints

◦ for redirecting script output

• Exit Code

◦ $?

◦ exit command

◦ return command

• export

F

• fish (Friendly Interactive Shell) • for loop

◦ C-style (for ((...)))

◦ over arrays

• Forking (process creation)

• Functions

◦ Arguments ($1, $@)

◦ local variables

◦ Return value (return)

◦ stdout for data

G

• getopts

◦ OPTARG

◦ OPTIND

• Git Hooks

◦ .git/hooks

◦ pre-commit

◦ pre-push

• Globbing

◦ \*, ?, []

◦ globstar (\*\*)

• grep

◦ -A, -B, -C (context)

◦ -F (fixed strings)

◦ -P (PCRE)

◦ -q (quiet)

H

• Hash (see Associative Array)

• Here String (<<<)

• Hold Space (see sed)

## I

• IFS (Internal Field Separator)

• Indexed Array

• Indirect Expansion (${!var})

J

• jobs

• JSON (JavaScript Object Notation) • jq

◦ -r (raw output)

◦ map

◦ select

L

• Least Privilege, Principle of

• Libraries, shell

◦ source

• local

• Logging ◦ logger

◦ syslog

• Login Shell

M

• man page

◦ groff macros (.TH, .SH, etc.)

• mktemp

N

• Nameref (declare -n)

• nc (see netcat)

• netcat (nc)

• Network Scripting

• nmap

• Non-Login Shell

P

• Parallelization

◦ wait

◦ xargs -P

• Parameter Expansion

◦ ${var#pattern}, ${var##pattern}

◦ ${var%pattern} , ${var%%pattern}

◦ ${var/pattern/string}

◦ ${var:-default}, ${var:=default}

◦ ${var:?error}

• PATH

• Performance

• Permissions

◦ chmod

◦ setuid

• pipefail (see set -o pipefail)

• Pipeline (|)

• Portability

• POSIX

• printf

• Process Management

• Process Substitution

◦ <(...)

◦ >(...)

• Python (as an alternative)

Q

• Quoting (see Quotes)

R

• Race Condition

• read

• readarray

• Redirection (>, >>, <)

• Regular Expression (Regex)

• return (see Exit Code)

• rsync

S

• Scope

• scp

• Secrets

◦ Environment variables

◦ Vaults

• Security

• sed

◦ Hold space (h, H, g, G, x)

◦ s/old/new/g

• set

◦ -e (errexit)

◦ -u (nounset)

◦ -x (debug)

◦ -o pipefail

• Shell Built-in

• Shell Variable

• shellcheck

• shopt

◦ globstar

• Signal (see trap)

• source

• SSH (Secure Shell)

◦ ssh-copy-id

◦ ssh-keygen ◦ Automation

◦ Tunneling (Port Forwarding)

• strace

• syslog (see Logging)

T

• tee

• test ([ ])

• time

• trap

◦ DEBUG

◦ ERR

◦ EXIT

U

• umask

• unset

• until

## V

• Variable (see Environment Variable, Shell Variable)

W

• wait

• while loop

• Word Splitting

## X

• xargs

◦ -P (parallelize)

Z

• zsh (Z Shell)

# SYMBOLS AND OPERATORS

• & (Background process)

• $((...)) (Arithmetic expansion)

• $() (Command substitution)

• $(()) (Arithmetic command, see set) • $'...' (ANSI-C quoting)

• "" (Double quotes, weak quoting) • '' (Single quotes, strong quoting)

• # (Comment, prefix removal pattern) • % (Suffix removal pattern)

• | (Pipe)

• >, >>, < (Redirection)

• >(), <() (see Process Substitution)

• ? (Exit code of last command, wildcard) • \* (Wildcard, globbing)

• [ (Test command, see test)

• [[ ... ]] (Extended test, Bash-specific) • !

◦ as logical NOT

◦ in array expansion ("${!array[@]}")

• :

◦ Null command (:), useful for variable checks

APPENDIX A: SHELL STARTUP FILE

FLOWCHART (BASH)

This diagram visualizes the decision-making process that a Bash shell follows at startup to determine which configuration files to load.

+—————————+

| Bash Startup |

+—————————+

|

v

+———————————-+

+———| Is it a login shell? |———+

| +———————————-+ |

| (Yes) (No) |

v v

+——————————+ +—————————————+

| /etc/profile | | Is the shell interactive?|

| is executed | +—————————————+

+——————————+ | (Yes) | (No)

| v v

v +————————-+ +————————-+

+——————————+ | /etc/bash.bashrc| | Is BASH\_ENV |

| ~/.bash\_profile |—(exists?)->| (if present) | | set? |

| is searched | +————————-+ +————————-+

+——————————+ | | (Yes)

| (No) v v

v +————————-+ +————————-+

+——————————+ | ~/.bashrc | | File in BASH\_ENV|

| ~/.bash\_login |—(exists?)->| is executed | | is executed |

| is searched | +————————-+ +————————-+

+——————————+ | |

| (No) v v

v ... ...

+——————————+

| ~/.profile |—(exists?)->

| is searched |

+——————————+

|

...

Notes on the Flowchart:

• Arrows (->): Indicate the flow of control. • Conditions ((Yes), (No)): Represent decision points in the process. • ~/.bash\_profile Chain: Bash executes only the first file it finds in this chain (.bash\_profile, .bash\_login, .profile). Using ~/.bash\_profile is the most common convention.

• .bashrc in Login Shells: As described in the chapter, ~/.bashrc is not automatically loaded by a pure login shell. It is a common and recommended practice to explicitly source it from within ~/.bash\_profile (using source ~/.bashrc or . ~/.bashrc).

APPENDIX B: POSIX PORTABILITY

QUICK REFERENCE (ENGLISH)

This table serves as a quick reference for replacing common, non-portable Bash extensions (“Bash-isms”) with their portable POSIX equivalents. Use this as a checklist when writing scripts for maximum compatibility with #!/bin/sh.

POSIX

Bash-ism

Task / Standard

(Convenient, Key Notes

Feature (Portable,

not portable)

safe)

Inside [, you must

Conditional

[[ ... ]] [ ... ] or test always quote your variables

Test

("$var").

String POSIX uses a single =

[[ $a == "foo" ]] [ "$a" = "foo" ]

Comparison for string comparison.

Logical [[ $a -gt 5 && $b [ "$a" -gt 5 ] Avoid using -a inside [

AND = "x" ]] && [ "$b" = "x" ] tests; it’s unreliable.

Logical OR "y" ]] [ "$b" = "y" ] tests. [[ $a -lt 5 \|\| $b = [ "$a" -lt 5 ] \|\| Avoid using -o inside [

Function function my\_func() { ... Simply omit the function

Declaration my\_func { ... } } keyword.

Regular [[ $var =~ $regex grep -qE echo "$var" \| In POSIX, you must

Expressions —"$regex" grep ]] rely on external tools like

or awk.

POSIX

Bash-ism

Task / Standard

(Convenient, Key Notes

Feature (Portable,

not portable)

safe)

The $((...)) expansion is

(( i++ )) or let

Arithmetic i=$((i + 1)) POSIX-compliant, but ((...))

i++

as a command is not.

Emulate with space-

Not

Arrays my\_arr=(a b c) separated strings and loops,

supported.

or use set—.

Associative declare -A Not Use external tools (awk)

Arrays my\_hash supported. or parse text files manually.

The behavior of echo is

echo with

echo -e "\n" printf "\n" not standardized. printf is

options

always preferred.

APPENDIX C: PARAMETER

EXPANSION QUICK REFERENCE

This appendix provides a quick-reference “cheat sheet” for the parameter expansion syntax covered in this chapter.

Expansion Example

Description

Syntax (var="path/to/file.txt.gz")

Substring

Removal

Remove shortest matching

${var#\*/} to/file.txt.gz

prefix \*/

Remove longest matching

${var##\*/} file.txt.gz

prefix \*/

Remove shortest matching

${var%.\*} path/to/file.txt

suffix .\*

Remove longest matching

${var%%.\*} path/to/file

suffix .\*

Search and

Replace

${var/txt/log} Replace first txt with log path/to/file.log.gz

${var//./-} Replace all . with - path/to/file-txt-gz

Default Values

(unset var)

& Errors

Expansion Example

Description

Syntax (var="path/to/file.txt.gz")

Use “default” if var is

${var:-"default"} default

unset/empty

Use and assign “default” if default (and now $var

${var:="default"}

unset/empty is “default”)

Show “error” and exit if

${var:?"error"} bash: var: error

unset/empty

Length &

(var="File Name")

Case

${#var} Length of the string 9

Convert to all uppercase

${var^^} FILE NAME

(Bash 4+)

Convert to all lowercase

${var,,} file name

(Bash 4+)

Indirect (name="USER";

Expansion USER="admin")

Use the value of name as the

${!name} admin

variable to expand

APPENDIX D: PROFESSIONAL BASH

FUNCTION TEMPLATE

Use this template as a starting point for your own functions to ensure a consistent, readable, and well-documented style.

#!/bin/bash

#######################################

# Converts a filename into a "safe" format.

# Replaces spaces and special characters with underscores and

# converts the entire name to lowercase.

# Globals:

# None

# Arguments:

# $1 - The filename string to sanitize.

# Outputs:

# Writes the sanitized filename to stdout.

# Returns:

# 0 on success.

# 1 if no argument was provided.

#######################################

sanitize\_filename() {

# 1. Validate arguments and assign to local variables

local original\_filename="$1"

if [[ -z "$original\_filename" ]]; then

echo "Error: sanitize\_filename requires a filename as an argument." >&2

return 1

fi

# 2. Main logic using local variables

local sanitized\_name

# Convert to lowercase

sanitized\_name="${original\_filename,,}"

# Replace unwanted characters (anything not a-z, 0-9, ., \_, -) with \_ sanitized\_name="${sanitized\_name//[^a-z0-9.\_-]/\_}"

# Reduce multiple underscores to a single one

sanitized\_name="${sanitized\_name//\_\_\*/\_}"

# 3. Output data to stdout

echo "$sanitized\_name"

# 4. Return success status

return 0

}

# Example call:

# new\_name=$(sanitize\_filename "My! Bad Filename (2).TXT")

# echo "$new\_name" # -> my\_bad\_filename\_2\_.txt

APPENDIX E: USEFUL VARIABLES

AND SIGNALS FOR ERROR HANDLING

Element Description Example Usage

Common trap

Signals

Always executed when the

EXIT trap cleanup EXIT

script exits (success, error, Ctrl+C).

Executed for any command that trap 'handle\_error

ERR returns a non-zero exit code. $LINENO' ERR

Executed before every simple trap 'trace\_command'

DEBUG command. DEBUG

Signals for

INT, trap 'echo "Aborted!"' TERM, HUP interruption/termination from

INT

external sources (e.g., Ctrl+C, kill).

Magic

Variables

The exit code of the last

$? if [[ $? -ne 0 ]]; then ...

command.

The current line number in the echo "Error on line

$LINENO script. $LINENO"

An array containing the names

$FUNCNAME of functions in the current call echo "Error in function

${FUNCNAME[0]}"

stack.

Element Description Example Usage

The command currently being trap 'echo "Command:

$BASH\_COMMAND executed (useful in trap DEBUG and $BASH\_COMMAND"'

trap ERR). DEBUG

The filename where the code echo "Script:

$BASH\_SOURCE resides (useful in libraries). ${BASH\_SOURCE[0]}"

APPENDIX F: ARRAY OPERATIONS

CHEAT SHEET

Indexed Array Associative Array (declare

Task

(my\_arr=(a b c)) -A h; h[k]=v )

my\_arr=(a b c) or declare -a declare -A my\_hash

Declaration

my\_arr (mandatory)

Add Element my\_arr+=("d") my\_hash["new\_key"]="new\_value"

${my\_hash["key"]} (reads

Read Element ${my\_arr[0]} (reads ‘a’)

value of ‘key’)

All

"${my\_arr[@]}" "${my\_hash[@]}"

Elements/Values

All "${!my\_arr[@]}" (returns "${!my\_hash[@]}" (returns all

Indices/Keys 0 1 2 ...) keys)

Number of

${#my\_arr[@]} ${#my\_hash[@]}

Elements

Delete

unset my\_arr[1] unset my\_hash["key"]

Element

Key Exists? [[ -v my\_arr[2] ]] [[ -v my\_hash["key"] ]]

APPENDIX G: PROCESS

MANAGEMENT QUICK REFERENCE

Concept / Syntax /

Main Use Case

Command Example

Job Control

Starts a process in the background,

& (in a script) long\_command allowing the script to continue

&

immediately.

Waits for all or a specific background

wait or wait

wait process to finish. Essential for

$PID

parallelization.

Lists the Process IDs of running

jobs -p pids=$(jobs -p)

background jobs. Useful for throttling.

Process

Substitution

diff <(cmd1) Treat the output of a command like a

<(...)

<(cmd2) temporary input file.

... \| tee > Treat the input of a command like a

>(...)

(cmd1) temporary output file.

Start a persistent background process

Coprocess coproc NAME to communicate with via file descriptors

{ cmd; }

(NAME[0], NAME[1]).

Parallelization

Concept / Syntax /

Main Use Case

Command Example

Simple and high-performance

xargs -P parallelization of tasks for a list of ... \| xargs -P 8

-n 1 cmd

arguments.

APPENDIX H: GREP VS. SED VS. AWK

– WHICH TOOL TO USE?

The

When you need

Best Tool Because... Example

to...

is...

It is highly

Find lines

optimized, simple,

containing a grep grep -i 'error' log.txt

and fast for this one

pattern

job.

Perform simple sed syntax is concise The s/old/new/

substitutions on a sed 's/\/home\/user/\/usr\/local/' line-by-line basis and designed for file

this.

It is built around

Extract specific ls -l \| awk '{print $9,

columns of data awk the concept of fields $5}'

and records.

It has built-in

arithmetic,

Perform ... \| awk '{sum+=$1}

calculations on data awk variables, and END {print sum}'

mathematical

functions.

Perform sed ’s hold space

complex, multi-line sed or can handle it, but it sed -n 'h;n;G;p' text Perl/Python gets complex transformations quickly.

The

When you need

Best Tool Because... Example

to...

is...

Generate a BEGIN/END

formatted report blocks and printf awk 'BEGIN{...} {..} from structured END{...}' awk

make it ideal for

data reporting.

Filter and

Its if/else logic is

process data based awk '$3 > 100 && $9

on multiple awk much more powerful ~ /\.log$/'

than grep.

conditions

APPENDIX I: JQ FILTER QUICK

REFERENCE

Example Example

Task jq Filter

Input Output

Basic Selection {"a":1, "b":[2,3]}

Select key value .a (above) 1

Select array element .b[1] (above) 3

1 (newline) 2

Iteration/Unpacking .[], .[]? [1, 2, 3]

(newline) 3

Object Construction {key1: .val1, {"val1": "x", {"key1": "x", "k2":

k2: .val2} "val2": "y"} "y"}

(with map(. \* 2))

Array Filtering map(expr) [1, 2]

[2, 4]

(with select(. >

select(cond) [1, 2, 3]

1)) 2 (newline) 3

Built-in Functions

Get length length [1,2,3] or 3 "abc"

Get object keys keys {"a":1, "b":2} ["a", "b"]

Check for value has("key") {"a":1} true

-r (with jq -r val (without

String Output (command-

'.key') {"key":"val"} quotes)

line flag)

Example Example

Task jq Filter

Input Output

String Interpolation "\(.key1) is \ {"k1":"A", "A is B"

(.key2)" "k2":"B"}

APPENDIX J: SHELL SCRIPT

SECURITY CHECKLIST

Before deploying a script to a production environment, run through this checklist.

Area Check How to Implement

Does the script contain Externalize secrets to

Secrets any passwords, API keys, environment variables or secure

or other secrets? credential files (chmod 600).

Always quote variables

Is all external input

Input ("$var"). Validate input against

(arguments, read) treated as

Validation allowed character sets (regex

potentially malicious?

match).

Is eval used with

Avoid eval. Always use "$var".

Command variable data? Are

Use — to separate options from

Injection unquoted variables used in

arguments.

commands?

Are temp files created Only use mktemp. Always set

Temporary

with predictable names in trap 'rm -rf "$tmpdir"' EXIT for

Files

/tmp? cleanup.

Does the script run Apply the Principle of Least

Permissions with more privileges than Privilege. Run the script as an

it needs? unprivileged user.

Area Check How to Implement

Are files created with

File Set a restrictive umask 077 at

overly permissive

Creation the beginning of the script.

permissions?

Error Does the script fail Use set -euo pipefail at the start

Handling safely on error? of the script.

APPENDIX K: NETWORK SCRIPTING

COMMAND QUICK REFERENCE

Command Example Main Use Case

nc -zvw nc -zvw 3 Quickly and script-friendly check if a

example.com 443 TCP port is open.

nmap -p 80,443 Scanning specific ports across many

nmap -p

192.168.1.0/24 hosts.

ssh-copy-id The easiest and most secure way to

ssh-copy-id

user@server install a public SSH key.

ssh ssh webadmin@host Executing a single, non-interactive

@

"cmd" "tail -f /var/log/app.log" command on a remote server.

scp file.zip Simple copying of files to or from a

scp

user@server:/backups/ remote host.

Intelligent, fast synchronization of

rsync -av directories. Preferred for rsync -av local\_dir/

user@server:/remote\_dir/

backups/deployments.

ssh -L Local Port Forwarding: Forward a

ssh -L 8080:localhost:80

user@host local port to a remote port.

ssh -R Remote Port Forwarding: Forward

ssh -R 8080:localhost:80

user@host a remote port to a local port.

ssh -D 1080 Creates a dynamic SOCKS proxy for

ssh -D

user@host flexible tunneling through an SSH host.

APPENDIX L: PERFORMANCE ANTI-

PATTERNS AND THEIR SOLUTIONS

Slow Anti-Pattern Fast Solution

Why It’s Faster

(many forks) (few/no forks)

filename=$(basename Shell Builtin (Parameter

filename="${path##\*/}"

"$path") Expansion)

Shell Builtin (Parameter

dir=$(dirname "$path") dir="${path%/\*}"

Expansion)

cat file \| grep "pattern" grep "pattern" file Saves a whole cat process.

wc -l < file or awk Saves a cat process. awk is

cat file \| wc -l

'END{print NR}' file often even faster.

head -n1 file read -r line < file read is a shell builtin.

for ((i=1; i<=1000; i++)); C-style for loop is a

for i in $(seq 1 1000); do ...

do ... builtin; seq is external.

Parameter expansion is

var=$(echo "$str" \| sed ...) var="${str/.../...}"

faster than echo + sed.

while read...; do ((c++)); awk is optimized for

awk 'END{print NR}'

done processing entire files.

APPENDIX M: BOILERPLATE FOR A

PROFESSIONAL COMMAND-LINE

TOOL

Copy this code as a starting point for your new tools. It includes a robust getopts loop, long option handling, a usage function, and a clear structure.

#!/bin/bash

set -euo pipefail

# -—Configuration and Default Values—-

VERSION="1.0.0"

VERBOSE=false

# ... other default values ...

# -—Functions—-

usage() {

cat <

Usage: $(basename "${BASH\_SOURCE[0]}") [-h] [-v] [—version] args... Script description goes here.

Available options:

-h,—help Print this help and exit

-v,—verbose Print script debug info

—version Print version and exit

EOF

exit

}

# -—Argument Parsing—-

# Manual long option parsing before getopts

for arg in "$@"; do

shift

case "$arg" in

"—help") usage ;;

"—verbose") set—"$@" "-v" ;;

"—version") echo "$(basename "${BASH\_SOURCE[0]}") $VERSION"; exit ;;

\*) set—"$@" "$arg"

esac

done

# getopts for short options

while getopts ':vh' opt; do

case "$opt" in

v) VERBOSE=true ;;

h) usage ;;

:) printf "Missing argument for -%s\n" "$OPTARG" >&2; usage ;;

\?) printf "Invalid option: -%s\n" "$OPTARG" >&2; usage ;;

esac

done

# Remove the processed options

shift $((OPTIND - 1))

# -—Main Logic—-

main() {

if [[ "$VERBOSE" = true ]]; then

echo "Verbose mode is enabled."

fi

echo "Remaining arguments: $@"

# ... Your main logic here ...

}

main "$@"

APPENDIX N: BEST PRACTICES FOR

SCRIPTS IN AUTOMATION

ECOSYSTEMS

This checklist summarizes the most important patterns and rules for writing shell scripts used in Git Hooks, CI/CD pipelines, and Docker containers.

Best

Ecosystem Practice / Why It’s Important Example

Rule

Always

Be

The execution provide full paths

explicit; do

environment (e.g., a CI or ensure $PATH is

General not rely on

runner) is minimal. It is set correctly.

the

not your personal shell. Explicitly install

environment.

all dependencies.

In automated systems,

there is no user to notice set -euo pipefail at

Use set -euo

errors. A script must fail the start of every

pipefail .

immediately and script. unambiguously on error.

Hooks block the Use fast linters

developer’s workflow (git (shellcheck) and only

Keep

Git Hooks commit, git push). Slow hooks run tests on

hooks fast.

will be ignored or changed files, not

disabled. the entire project.

Best

Ecosystem Practice / Why It’s Important Example

Rule

If a hook fails, the

Provide echo "Error: developer must

clear error shellcheck found issues immediately know why

in $file. Please fix them."

messages. and what to do to fix the >&2; exit 1

problem.

Outsource

complex logic to

Keep YAML is for declaring

CI/CD separate files .sh

YAML files steps, not for complex

Pipelines and call them from

clean. logic.

the pipeline: run: ./scripts/deploy.sh

Use your

CI/CD platform’s

Repeatedly

caching

Use downloading dependencies

mechanisms to

caching. or building artifacts is

store directories

slow and expensive.

like node\_modules or ~/.m2 between jobs.

Best

Ecosystem Practice / Why It’s Important Example

Rule

The exec form

CMD ["/usr/bin/my-

Prefer (["command", "arg"]) starts

the exec form your process as PID 1 and app", "—foreground"]

Docker instead of CMD

over the shell forwards signals correctly.

/usr/bin/my-app—

form. This is crucial for graceful

shutdowns ( foreground docker stop).

Allows for preparing

Use an the environment (e.g., A script that

entrypoint.sh waiting for a database, performs script for adjusting config files) configuration and initialization. before the main ends with exec "$@".

application starts.

This command

Use replaces the entrypoint The last exec

script’s shell process with command in an

"$@" at the

the main application. This entrypoint.sh script

end of

ensures the application that executes a

entrypoint.sh .

becomes PID 1 and CMD instruction. receives signals.

Best

Ecosystem Practice / Why It’s Important Example

Rule

Use multi-

Smaller images are

Build stage builds and

faster to transfer, have a

small start with minimal

smaller attack surface, and

images. base images like

require less disk space.

alpine or scratch.

APPENDIX O: DECISION MATRIX:

SHELL SCRIPT VS. HIGHER-LEVEL

LANGUAGE

Use this table as a guide to decide which tool is best suited for a given task.

Best

Better

Task / Choice:

Choice: Rationale

Requirement Shell

Python/Go/etc.

Script

Chaining

This is the absolute core

Commands / Yes No

competency of the shell.

“Glue Code”

The shell is unbeatable for

Rapid Yes

Yes small prototypes. Python is a

Prototyping (Python)

close second.

Shell pipelines are

Log Yes

Yes extremely powerful. Python

Analysis / Text (with grep,

(Python) offers more structure for

Manipulation sed, awk)

complex logic.

Complex

The shell only has simple

Data Structures

No Yes arrays. Higher-level languages

(Objects,

are essential for this.

Graphs)

Best

Better

Task / Choice:

Choice: Rationale

Requirement Shell

Python/Go/etc.

Script

The interpreted nature of the

CPU-

Yes (Go, shell is a massive disadvantage

Intensive No

Rust, C++) here. Compiled languages are

Calculations

necessary.

Yes Yes Good for simple queries,

Interacting

(with curl (Python + but libraries offer better error

with Web APIs

+ jq) requests) handling, sessions, etc.

Requires a

Completely outside the

GUI or Web No Yes

scope of the shell.

Frontend

Large, Higher-level languages

Team- offer better testability,

No Yes

Maintained modularity, and tools for Project collaboration.

Maximum

Yes A POSIX shell is available

Portability on No

(#!/bin/sh) on virtually every system.

Unix Systems

# Also by Michael Basler

Erste Auflage

[Das Linux Terminal für Fortgeschrittene - Die Kommandozeile leicht](https://www.draft2digital.com/catalog/2859007?distributor=kobo)

[gemacht](https://www.draft2digital.com/catalog/2859007?distributor=kobo)

[Das Linux-Terminal für Einsteiger - Die Kommandozeile leicht gemacht](https://www.draft2digital.com/catalog/2811420?distributor=kobo)

[Linux Shell Scripting – Ein Leitfaden für Anfänger](https://www.draft2digital.com/catalog/2815868?distributor=kobo)

[Switch to Linux – 15 wichtige Tipps](https://www.draft2digital.com/catalog/2875256?distributor=kobo)

First Edition

[The Linux Terminal for Beginners - The Command Line Made Easy](https://www.draft2digital.com/catalog/2812701?distributor=kobo)

[Linux Shell Scripting - A Beginner's Guide](https://www.draft2digital.com/catalog/2814611?distributor=kobo)

[The Linux Terminal for Advanced Users - The Command Line Made Easy](https://www.draft2digital.com/catalog/2858969?distributor=kobo)

[Switch to Linux – 15 Important Tips](https://www.draft2digital.com/catalog/2879783?distributor=kobo)

Standalone

[Die Asiatische Küche Entdecken - Die 10 Besten Rezepte aus den](https://www.draft2digital.com/catalog/2913272?distributor=kobo)

[Philippinen](https://www.draft2digital.com/catalog/2913272?distributor=kobo)

[Discover Asian Cuisine - The 10 Best Recipes from the Philippines](https://www.draft2digital.com/catalog/2913289?distributor=kobo)

[Die Asiatische Küche Entdecken - Die 10 Besten Rezepte aus Afghanistan](https://www.draft2digital.com/catalog/2988144?distributor=kobo)

[Discover Asian Cuisine - The 10 Best Recipes from Afghanistan](https://www.draft2digital.com/catalog/2988168?distributor=kobo)

[Das Linux Shell Skripter Handbuch - Vom Gesellen zum Meister](https://www.draft2digital.com/catalog/3001352?distributor=kobo)

The Linux Shell Scripting Handbook - From Journeyman to Master