Resume


```
Name: Dhruva Krishnamurthy
Email: dhruvakm@gmail.com
Phone: +1 (408) 623-0605
Location: California, United States
```


- Resume: https://mechanicker.github.io/resume.html
- GitHub profile: https://github.com/mechanicker
- LinkedIn profile: https://www.linkedin.com/in/dhruvakm/

Objective

20+ years of software design & development experience involving cross-site collaboration across multiple technologies and domains, having led numerous initiatives improving the performance of products and services, delivering impactful customer value with tangible cost savings. Thrive in situations demanding out-of-the-box innovative problem-solving approaches. I am seeking opportunities to leverage my expertise in a dynamic and innovative environment with high customer impact.

Skills

Proficiency

- Distributed systems and developing cloud-native backend microservices on `AWS`
- Design and implement scalable `CI/CD` systems, data pipelines, and data lifecycle management
- File systems and systems programming using the POSIX interface
- System programming, performance troubleshooting, and related tool development on both Unix and Microsoft Windows (`Win32` API)
- Source code management systems, `git` and `libgit2` internals
- C, C++, Go, Python, and Java
- Troubleshooting and debugging distributed systems, debuggers, and profilers across operating systems

Working knowledge

- Traditional and NoSQL databases
- Memcached, Redis and BerkeleyDB key/value stores
- Boost C++ library
- Parser development using Lex & Yacc and similar tools

Experience

Principal engineer, Atlassian

Jan 2017 - Current

Bitbucket

- Working on an initiative identifying Bitbucket workflows that can significantly benefit from heuristics
    - Exploring ML-based potential enhancements to customer workflows
- Designing storage layout for better performance, data protection, and enabling data residency for future expansion into different regions.
- Exploring opportunities to better integrate Bitbucket into storage vendor-provided extensions to enable the development of enhanced data management solutions for our customers
    - Leading an initiative to improve disaster recovery aspects of Bitbucket with multi-million cost savings by optimizing resources
    - Implemented `gRPC` routing service to increase locality of data access and benefit from file system caching resulting in reduced latency
- Maintainer of Bitbucket fork of git with enhances to increase robustness in a distributed storage environment.
    - Implemented enhancements to recover from crashes with a reliable cleanup, completely eliminating support requests to recover from corrupt git repositories
    - Improved observability into problems git operations run into operating in distributed environments with shared data access
    - Undertook git upgrade from an old version to the current version, designed and implemented progressive rollout support, and streamlined the git upgrade process for Bitbucket
- Led various streams of work as part of migrating Bitbucket customer repositories to the cloud, a critical part of the long-term strategy to make Bitbucket multi-region and increase reliability
    - Architecting Bitbucket git repository disaster recovery, a key enabler for Bitbucket cloud transition plan
    - Implemented near real-time cross-region fault-tolerant data replication for recovering customer data during disasters like region outages
    - Delivered high-performant git object caching to reduce NFS latencies, a critical requirement for moving Bitbucket to the cloud
- Implemented parallel unarchiving tool `puntar` to speed up restoring PostgreSQL database backup archives

- Submitted patches to upstream libgit2 to improve the performance of repositories accessed over the network file system
- Implemented complex `CICD` pipelines to orchestrate build and deployments of Bitbucket microservices
- Prototyped non-intrusive Bitbucket pipelines abuse detection and prevention mechanism to combat Bitcoin mining

Cross product search

As part of implementing GDPR compliance, we decided to consolidate user-identifiable information and enforce tighter access controls. User search is a critical component used across Atlassian services.

- Responsible for the user search data pipeline for ingesting user information and indexing for search
- Implemented core parts of user search backend services to manage the lifecycle of search data with a denormalized permissions model to support permission-aware search

Trust & Identity

Unified role-based permission model for seamless user experience across Atlassian services.

- Implemented and operationalized a highly scalable and resilient data pipeline for ingesting real-time permission data
- Built detection and optimal recovery mechanisms during partial/total data loss in the pipeline due to upstream failures

Senior engineer, NetApp

Feb, 2008 - Jan, 2017

NFSv4 server performance lead

With an open mandate to improve NFSv4 server performance and make NFSv4 a compelling alternative to NFSv3, I worked on streamlining performance testing followed by performance improvements.

- Establish predictable and repeatable baseline performance metrics.
- Identify custom workloads for performance improvements.
- Served as a liaison between NFS and the wider performance engineering team
- Improved NFSv4 IO performance by ~40% by implementing in-kernel buffer caching
- Simplified NFSv4 server-side state lifecycle management using C++ `RAII`

SAP Hana over NFSv4
Designed and implemented a client-side performance improvement library to increase overall IO throughout by ~50%, critical for SAP Hana workloads. SAP qualifying NetApp NFSv4 depended on these performance improvements.

- Developed IOtrap library to transparently use `asyncio` to improve performance via `IO` interception. This was later incorporated into the SAP Hana core engine. SAP Hana now supports configuring `async` IO parameters
- NetApp published best practices for SAP Hana over NFSv4 "Configuration of Performance Test Tool".

Scaleout NAS storage

As a core member of the architecture group designing distributed scaleout NAS storage `Infinite Volume`, I played a significant role in developing object storage, cluster-wide file system metadata search, and file system recovery. Infinite Volume allowed NetApp better compete with other storage vendors (Isilon).

- Core contributor to distributed search service for file system metadata using NoSQL (BerkeleyDB)
    - Designed and implemented optimized tree-based complex query expression execution engine using Intel TBB (Threading Building Blocks) library based on recursive parallelism
    - Tuned the query engine and delivered ~5x improvements for complex queries involving multiple fields and logical operators
    - Protyped porting BerkeleyDB to kernel space and benchmarked against running search service in user space
- Designed and implemented various components of CDMI object versioning using NetApp file system technologies
- `HTTPS` support for CDMI by tunneling from the kernel network layer to external SSL terminating service
- Led the initiative to implement CDMI conformance and performance testing suite
- Design and implement core aspects of file system metadata search engine

Storage management

- Implemented statistical methods using C++ Boost libraries for NetApp performance advisor. This enabled the frontend team to implement advanced performance data analytics using the NetApp storage management suite.
- Designed and implemented zero-configuration storage controller discovery in a data center based on ICMP/ICMPv6 packet broadcast/response mechanism.
- Developed a prototype of scaleout distributed storage management to overcome limitations in the storage management suite allowing managing 500+ controllers from 50 with a single clustered installation of the storage management suite.
- Designed and implemented a realtime SQL parser to translate Sybase specific syntax to Oracle

Technical Specialist, McAfee

Feb, 2006 - Feb, 2008

- As a lead performance engineer, built a team of 4 engineers to focus on performance engineering.
- Developed methodologies to measure performance

- Implemented custom memory allocators to reduce lock contentions in multi-threaded service resulting in ~25% increase in scan rate
    - Evaluated and benchmarked custom allocators from MicroQuill and Hoard

Technical Specialist, HP

Jan, 2006 - Feb, 2007

- Led initiative to port Samba, an open-source CIFS file server on the VMS operating system
- Implemented missing core POSIX APIs emulation on VMS required for porting Samba
- Ported `cvs` source code management tool to VMS and transitioned tracking Samaba on VMS to `cvs`
    - Refactored large parts of scattered VMS-specific changes to reduce merge conflict
    - This led to the reduction in time to merge upstream changes and made it easier to stay in sync with upstream

Engineering Manager, Bosch

Dec 2004 - Dec 2005

- Led a cross-site team developing a navigation point-of-interest data compiler chain for Blaupunkt car navigation systems. Improved resiliency of multi-stage compiler and delivered cost optimizations by reducing processing time from multiple weeks to days.
- Team building, requirements gathering, project tracking, and coordination with the team in Germany
- Transition data compiler pipeline from Sun SPARC servers to SUSE Linux, eliminating expensive annual maintenance contract with Sun Microsystems

Technical Lead, Delmia (Dassault Systemes)

Feb 1998 - Dec 2004

Developed core features in a 3D simulation-based robotics and factory floor simulation software in C++ and cross-platform CATIA CAA V5 architecture.

- Team lead for integrating PLM solution `Process Engineer` and CATIA
- Core engineer developing CATIA V5 Composites
- Designed and implemented bi-directional import/export of DMIS programs for CMM machines with vendor-specific language variants
    - Designed and implemented a module to import DMIS (CMM/CNC programs) and create a 3D simulation model
    - Implemented DMIS program generation from a 3D simulation model
- Ported LXR to run on Microsoct Windows by using SWISH-E for free text search. Maintained code search based on the port with custom search result sorting/ranking for internal use

Production Engineer, Wipro Fluid Power

Oct 1996 - Feb 1998

As part of production engineering, I worked in a hybrid environment involving production on the shop floor and automating processes for production. This exposed me to real-world manufacturing problems and an opportunity to solve them using software-based technologies.

- Implemented a CNC machine path simulator in `C` using Borland graphics primitives
- Implemented an inventory tracking system in `C` and dBase

Awards and recognition

Patent 10812313 (granted)

A federated namespace of heterogeneous storage system namespaces

Abstract

The patent covers a system and computer-based method for performing a data transaction in a network storage system by offering a federated file system namespace with a POSIX-compliant file system interface to access data stored across distributed decoupled heterogeneous storage entities along with policy-based data lifecycle management leveraging different storage tiers.

Recognition at work

- Appreciation and peer bonus for implementing significant performance improvements using ASYNC IO, leading to around $30k savings per month
- Improved cache utilization leading to cost savings by reducing the fleet of servers required to handle the load
- Implemented transparent cloud-scale `git` object caching layer to significantly lower NFS IO
- Secured 2nd place in NetApp-wide `hackathon` (leading to patent filing)
- Recognition and award for implementing search service for file system metadata at NetApp
- Received employee excellence award for `1999 - 2000` at Delmia

Education & certifications

Bachelor of Engineering, Mechanical with CAD/CAM & Robotics

08/1992 - 08/1996

National Institute of Engineering,
University of Mysore, Karnataka, India

- Designed and developed finite element solver using Skyline matrix in `C` for 2D structures and thermal distribution
- Developed robot path simulator

Intel VTune

Underwent hands-on training on using Intel VTune for troubleshooting performance bottlenecks in performance-critical multi-threaded applications

- Effectively implemented the learnings to solve performance problems in McAfee Antivirus
- Applied learnings during the implementation of the search query execution engine at NetApp

Unix & C

- 6-month practical training program to learn Unix system programming using `C`

Other projects and links to referred articles

- Patent 10812313, "Federated namespace of heterogeneous storage system namespaces": https://patents.justia.com/patent/10812313
- Parallel unarchive command line tool `puntar`: https://github.com/mechanicker/puntar
- `IOtrap` library for `IO` througput improvement: https://github.com/mechanicker/iotrap
- Win32 function call profiler: https://github.com/mechanicker/cramp
- NetApp best practices for SAP Hana using `asyncio`: https://www.netapp.com/media/8991-tr4290.pdf