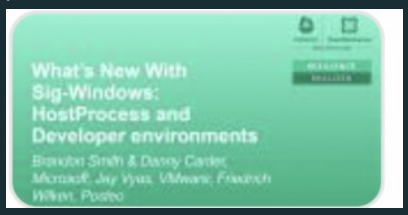
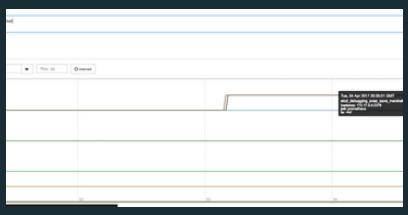
PORTFOLIO

A visual history of a few things I've been lucky enough to get paid to build...

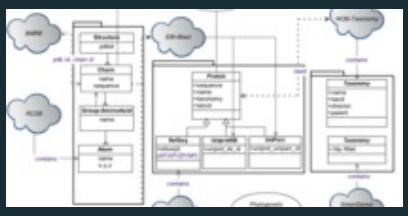


live

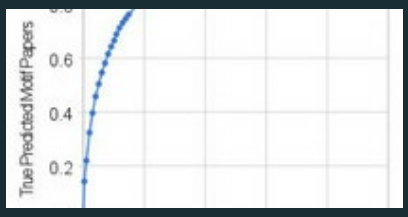
bioinformatics



etcd



datamarts



ROC

Jay Vyas

Leading a variety of customer facing Kubernetes initiatives at VMware around Edge compute, CNI, Upstream Networking, and Windows. Bioinformatician, Author, Engineering leader focusing on radically collaborative ownership.

978-405-4461

Concord, MA

jay@apache.org

jayunit100.blogspot.com

MAJOR ACCOMPLISHMENTS

- Moved Blackduck to a fully containerized, cloud native SaaS platform on GKE. Adoption = 80% of all customers.

- Delivered VMWare's CNI solution (antrea) to the core Tanzu solution and aligned it with open source networking API evolution standards.

- Published 50 patents, 12 publications (fields of bioinformatics, microservices, large data sets)

- Virtualized NMR data processing, seeding the NMRBox project and grant https://grantome.com/grant/NIH/P41-GM111135-06-8673

- Published Books: Container Network Security for Dummies (vmware) and Core Kubernetes (manning).

- Built extremely collaborative teams across a in different environments with different types of people.

- Turned my camera on in most of my Zoom calls this year :)

WORK EXPERIENCE

July 2019 - Current

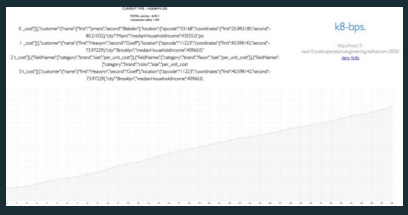
Member of Technical Staff, Staff Engineer

Leading a variety of productized, open technologies that are essential to the VMWare Tanzu Portfolio, specifically around:
- Antrea (our CNI)
- Kubernetes on Windows
- Release Engineering and Conformance Testing
- Edge compute offering

As an engineering leader at VMWare I serve in the following capacities:

- Engineering Lead for "Thick" Edge and Antrea Kubernetes related initiatives at VMWare.
- Engineering Lead for Kubernetes on Windows (VMWare Tanzu) and our broader Extended Infrastructure initiative. Brought Windows from a POC to a fully supported TKG feature, defined upstream and downstream SLAs around windows. Defining windows conformance for the world https://github.com/kubernetes/enhancements/issues/2578 .

Simulators



- Managing customer escalations around networking, windows , and CNI providers, some of which were extremely time and business sensitive.
- Building teams, pairing w/ new engineers, keeping morale high in the face of continuous change, creating a culture of radical transparency, continuous collaboration, and openness, every day.

Meanwhile, I spend alot of time in the K8s community, driving several upstream initiatives:

- TL for sig-windows (CNCF) - working on the windows conformance initiative
- Lead for several sig-network (CNCF) sub projects, started both the network policy and kube proxy refactoring initiatives.
- Windows Conformance https://github.com/kubernetes/enhancements/tree/master/keps/sig-windows/2578-windows-conformance
- Network Policy Validation https://github.com/kubernetes/enhancements/tree/master/keps/sig-network/1611-network-policy-validation
- Default Namespace labels: https://github.com/kubernetes/enhancements/tree/master/keps/sig-api-machinery/2161-apiserver-default-labels

- Note I Joined in 2021 as an MTS4, was promoted to a Staff Enginneer in 2022... as an MTS4 my main focus was on Upstream K8s networking, Antrea as a CNI, and initial investments into TKG Window

January 2019 - July 2019

Kubernetes Engineer

Platform9

Just joined my newest startup !

I'm working on a combination of up and downstream work on our K8s product line.

Recent K8s projects include, https://github.com/carbonrelay/kapture, a Kafka+BigPetStore based load generation library for upstream performance testing of BigData workloads on kubernetes.

Also doing work on various storage and cluster bootstrapping algorithms for kubernetes distributions, and making sure our customers are happy :).

June 2016 - January 2019

Cloud Native Architect (All products)

Blackduck software / Synopsys Inc

Cloud Native architect and lead for the entire synopsys suite of software products. Note: The word architect is not something I generally subscribe to - I beleive in flat, distributed teams that take

individual ownership of microservices.

Brought to blackduck to transition their technology stack to a cloud native offering, worked on 3 core products: our SAAS offering, OpsSight for container scanning at data center scale, and our cloud native product offering for on prem.

- Built, grew, and trained a team of 6 engineers that had virtually no kubernetes experience to being core cloud native team.

- Built and redesigned all cloud native offerings, unified as a single solution: The blackduck operator, driving our SaaS (100+ customers, 100,000,000$ in business) as well as other products.

- Built the team and architecture for the OpsSight Connector (Perceptor) Product: Co-creator of the Perceptor open source platform for cloud native event response, scanning at data center scales, built around its original downstream product, OpsSight scanner.

- Relentless customer support for all products for my team: creating a direct line of sight from customer driven insights to engineering R&D and implementation. Ported 10s of major customers to on-premise OpenShift and Kubernetes offerings.

- Ported a complex, purely docker based application to a cloud native, unprivileged, kubernetes/openshift offering, including putting specific patches into our products microservices so that they behaved better (i.e. least-privileged, backoff capabilities) in a highly secured and commoditized CPU/memory environments.

- Drove our cloud native offerings adoption with several fortune 500 customers, helping them both adopt and embrace their own changing internal infrastructure alongside the idiosyncracies of our own cloud native security scanning offering (the blackduck hub). Whitegloved customers with end to end kubernetes support: from debugging IPTables and DNS platform issues to distributed storage.

- Design, engineering, and internal training on a continous delivery and ephemeral infrastructure pipeline from scratch for verifying cloud native application offerings on all major platforms (openshift, kubernetes, swarm).

- Open sourced several aspects of the business that needed to have broader transparency and ease of adoption in order to be recognized as valid products by customer segments.

- Ported all infrastructure to terraform deployments, which deployed kubernetes using kubeadm on EC2, VMWare, and GCE. Ultimately this saved the company upwards of 20,000$ a month - and led to the entire team of 50+ engineers using an internal version of our customer SAAS offering for performance and feature testing. This platform supported kubernetes, openshift, and docker swarm.

Nov 2012 - June 2016

Principal Software Engineer

Red Hat

Deep technical and distributed systems, community driven engineering which evolved the kubernetes, glusterfs, and hadoop ecosystems in large part towards what they are today.

Kuberenetes/Openshift: Engineer

Worked on core portions of Google's Kubernetes container platform and Red Hat OpenShift.

- Distributed systems engineer. contributor and maintainer on Kubernetes platform (mostly golang). Contributed core features to the HTTP client performance, high availability tooling, developer automation, and E2E testing suites, lead and/or participated in several SIG's with the community.

- Implemented several performance tweaks, optimizations , algorithm improvements to the kubernetes scheduler, including scale, performance, and cache optimizations which are only discoverable in large deployments / clusters of 1000s of nodes/pods.

- Mentored and trained several new developers on the internals of idiomatic kubernetes development and community practices, golang tooling, and on kubernetes framework engineering internals.

- One of the major contributors to kubernetes itself for the first 2 years of its existence.

- Contributed to and architected a broad range of product integration technologies in the bigdata ecosystem, under the Apache Software Foundation: Hadoop, Spark, Cassandra engineering on alternative storage systems, middleware, mostly powered by GlusterFS/HDFS.

- Worked in several ASF projects and communities on behalf of Red Hat to increase our interop with open source big data projects, particularly in the Hadoop ecosystem (specifically BigTop, Apache Hadoop). Also wrote open source blueprints for internal evaluation and testing of Flink, Spark, Hadoop.

- Worked at the intersection of our interests in middleware, scalability, and bigdata, building automation and POCs around containerized I/O between big data frameworks (i.e. spark) and middleware data abstractions (TEEID/JDV).

- Built out the GlusterFS integration solution for Apache Hadoop used in Red Hat Storage as an HDFS alternative. Worked with the broader hadoop community to make sure its test coverage was across the entire ecosystem of filesystem semantics (HBase, Hive, Mahout, and so on).

- Built one of the first containerized, generic SparkStreaming blueprint applications (forked and used by 100s of individual developers) and other blueprint applications for bootstrapping bigdata workflows.

- Mentoring and training developers on integration testing for the hadoop ecosystem, microservice architectures, reproducible deployments with vagrant, internal cloud usage. Also spent a large

amount of time with development tooling and engineering higher quality internal application blueprints and deployments.

Apr 2014 - Present

Apache Software Foundation

Member, PMC, and Commiter

ASF BigTop, ASF Hadoop

- Developed, reviewed, maintained code in the ASF as a PMC and Commiter : Apache BigTop (the open source hadoop distribution), and Apache CTakes (the medical text analytics framework).

- Engineered, and maintained large components of ASF BigTop deployment, Integration and Smoke testing frameworks.

Synthetic Data Set Generation and Scale Testing for the ASF

- Built out the BigPetStore application (spark and mapreduce ecosystem apps w/ a synthetic, scale-out data generator).

- Reproducibility engineering around hadoop deployments: Architected complete end-to-end vagrant implementations of Apache BigTop for rapid prototyping and testing of new hadoop ecosystem releases. Also lead the initiative porting ASF BigTop to Docker for container based deployment and testing.

Scaling out CTakes for drug entity extraction at large scales

- Contributed core containerization and scalability initiatives to the Apache CTakes project.

Leading the the HCFS initiative for GlusterFS and Hadoop

- Filesystem and verification improvements across ASF hadoop and ASF bigtop as part of the HCFS initiative (https://wiki.apache.org/hadoop/HCFS/).

- Became a Commiter/PMC in 2016, and later a Member of the ASF in March of 2017

May 2011 - Nov 2012

Lead Data + Ops Engineer @ Peerindex

Founder @ Rudolf Inc

Data (Hadoop) Engineering @ PeerIndex

- Engineered Java and Hadoop pipeline MapReduce PISAE (Peerindex social action engine) platform. Development of our backend MapReduce java systems - largely based on ingestion and feature extraction.

- Enabled Real time streaming for tweets from high profile social entities; EMR and hadoop administratoin on an as needed basis, for 200+ node clusters.2

- In addition to all the deployment and engineering work, founded a

consulting firm, Rudolf Inc. Managed all financial affairs and contracts, presented, delivered technical solutions in the bigdata and web-data spaces to various multi-million dollar companies in Europe and the United States.

Founder @ Rudolf Inc

- Reported directly to the Head of Research and CTO @peerindex on BigData pipeline, status, and implementing architecture changes.

- Coordinated the launch of our new site which was completely driven using a key-value storage (DynamoDB) and HDFS data backend, millisecond latency SLAs.

- Implemented algorithms alongside data scientists, engineering directors, and the front-end team to ensure that data quality standards for 100s of millions of outputs were always met using a rock solid data model contract.

Jan 2004 - Jan 2010

Independent Software consultant, Architect (2006, 2008, 2009)

(later incorporated as Rudolf Inc)

Various one-off solutions built for small / starter companies and academic labs.

- Reverse engineering of UML design for various specifications in accountancy systems for clients, key for scaling engineering efforts across teams at the time.

- Designed custom LISP / Python applications for ensuring transaction integrity in complex, multi-database MySQL federated systems.

- Evaluated of e-commerce solutions (i.e. google checkout, paypal) with home grown product inventory and management systems (seikelceramics.com)

- Built Full stack, rapid software prototying, informatics consulting, and helping individuals get up to speed with data mining and machine learning related initiatives.

- Bioinformatics Consulting with UNLV's Biomedical Sciences department on protein analytics, literature mining software.

Aug 2005 - Dec 2007

Open-Source Bioinformatics Developer Java Developer & Researcher

UConn Health Center

Designing Bioinformatics federation platforms

- Developer/designer of a highly interactive client side application for fully integrated molecular visualization platform (VENN) which allows for Jmol based 3D analysis of protein evolutionary conservation (See

Publications). Created a cross platform VM for NMR data processing (vagrant, virtualbox, and ubuntu).

- Deployed plug-in based modules for implementation of molecular analysis algorithms in a single computational proteomics framework. Optimized and changed features of the application to suit emergent needs, such as protein domain oriented analyses.

- Data Warehousing/Modelling of Functional Minimotifs Designed a Hibernate API and MySQL data repository along with an java based which integrated proteomic information spanning protein motifs, functional annotations, taxonomy data, sequences, and protein domains.

- Architected an Expert System (MIMOSA) which automatically implemented various text mining and correlation scoring algorithms using ontologies for 5,000,000 publically available medical abstracts. Developed a plug in oriented Graphical User Interface using the Java Swing framework which enabled database driven, high throughput annotation of "minimotifs".

End to End Workflow Engineering Solutions for Protein NMR Analysis

- Collaborated on several database aspects of the NIH funded, publically available, web based Minimotif miner application (http://mnm.engr.uconn.edu). Engineered Java API's for reading/writing of large, binary FID representations present in vendor-specific NMR spectral data types in support of an open-source translation and conversion API.

- Built a Swing-based workflow building environment for time domain NMR data processing, as a custom, visual, 2D graphical application which allowed for on the fly creation of data-processing "actors", with reloadable and persistent state and associated, workflows which triggered offline data processing tasks. All based on a finite state machine.

- Designed, architected of an integrated NMR visual data mining platform, as part of the Rudolf project using and Clojure to wrap existing Java API's.

- Presented work at premier scientific conferences at conferences (ICBMRS, Protein Folding Symposium, New England Structural Biology, Exp Nuclear Magnetic processing conference).

EDUCATION

2007 - 2011

Doctor of Philosophy (PhD)

UConn

Proteomics data federation apps; published several (10+) articles in medium/high tier journals on algorithmic and visualization advances in data-integration of medical, biological, genomic data.

Computational protein analysis structure, sequence, and functional analysis; workflow builders for NMR data processing and annotation of peptide sequences with respect to phenotypes, medical conditions.

2004 - 2006

Master of Science (MS)

RPI

Created an NMR software integration environment for protein structure calculation and data processing. This led to a PhD and several other adventures at the interface of data visualization, integration, and bioinformatics.

2001 - 2004

Bachelor's degree

University of Arizona

Mathematics (major) & Computer Science (minor).

PUBLICATIONS AND PATENTS

Book: **Core Kubernetes**, Jay Vyas and Chris Love May 2022, ISBN 9781617297557 (https://www.manning.com/books/core-kubernetes).

Article: **A Domain-Driven, Generative Data Model for Big Pet Store** http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=7034765. RJ Nowling and Jay Vyas, by RJ Nowling · 2014

Article: **A Pipeline Software Architecture for NMR Spectrum Data Translation.** Heidi J C Ellis, Gerard Weatherby, Ronald J Nowling, Jay Vyas, Matthew Fenwick, Michael R Gryk Computing in Science and Engineering 05/2012; 15(1):76-83. · 1.25 Impact Factor

Conference Paper: **An Open-Source Sandbox for Increasing the Accessibility of Functional Programming to the Bioinformatics and Scientific Communities** M. Fenwick, C. Sesanker, M.R. Schiller, H.J.C. Ellis, M.L. Hinman, J. Vyas, M.R. Gryk Information Technology: New Generations (ITNG), 2012 Ninth International Conference on; 01/2012

Article: **HIVToolbox, an integrated web application for investigating HIV**. David Sargeant, Sandeep Deverasetty, Yang Luo, Angel Villahoz Baleta, Stephanie Zobrist, Viraj Rathnayake, Jacqueline C Russo, Jay Vyas, Mark A Muesing, Martin R Schiller PLoS ONE 05/2011; 6(5):e20122. · 3.53 Impact Factor

Conference Paper: **CONNJUR Workflow Builder: Open Source software for spectral reconstruction of NMR data** Weatherby G, Vyas J, Nowling RJ, Heidi J.C. Ellis, Gryk MR 52th Experimental Nuclear Magnetic Resonance Conference,; 04/2011

Article: **Iterative Development of an Application to Support Nuclear Magnetic Resonance Data Analysis of Proteins**. Heidi J C Ellis, Ronald J Nowling, Jay Vyas, Timothy O Martyn, Michael R Gryk Proceedings of the ...

International Conference on Information Technology: New Generations. International Conference on Information Technology: New Generations. 04/2011;

Article: **CONNJUR spectrum translator: an open source application for reformatting NMR spectral data**. Ronald J Nowling, Jay Vyas, Gerard Weatherby, Matthew W Fenwick, Heidi J C Ellis, Michael R Gryk Journal of Biomolecular NMR 03/2011; 50(1):83-9. · 3.31 Impact Factor

Article: **Extremely variable conservation of γ-type small, acid-soluble proteins from spores of some species in the bacterial order Bacillales.** Jay Vyas, Jesse Cox, Barbara Setlow, William H Coleman, Peter Setlow Journal of bacteriology 02/2011; 193(8):1884-92. · 2.69 Impact Factor

Article: **SciReader enables reading of medical content with instantaneous definitions.** Patrick R Gradie, Megan Litster, Rinu Thomas, Jay Vyas, Martin R Schiller BMC Medical Informatics and Decision Making 01/2011; 11:4. · 1.50 Impact Factor

Article: **A computational tool for identifying minimotifs in protein-protein interactions and improving the accuracy of minimotif predictions**. Sanguthevar Rajasekaran, Jerlin Camilus Merlin, Vamsi Kundeti, Tian Mi, Aaron Oommen, Jay Vyas, Izua Alaniz, Keith Chung, Farah Chowdhury, Sandeep Deverasatty, Tenisha M Irvey, David Lacambacal, Darlene Lara, Subhasree Panchangam, Viraj Rathnayake, Paula Watts, Martin R Schiller Proteins Structure Function and Bioinformatics 09/2010; 79(1):153-64. · 3.34 Impact Factor

Conference Paper: **The CONNJUR Spectrum Translator: Open Source software for converting the format of time-domain NMR data** Nowling RJ, Vyas J, Weatherby G, Ellis HJC, Gryk MR XXIVth International Conference on Magnetic Resonance in Biological Systems; 08/2010

Article: **Biomolecular NMR data analysis**. Michael R Gryk, Jay Vyas, Mark W Maciejewski Progress in Nuclear Magnetic Resonance Spectroscopy 05/2010; 56(4):329-45. · 8.71 Impact Factor

Article: **MimoSA: a system for minimotif annotation.** Jay Vyas, Ronald J Nowling, Thomas Meusburger, David Sargeant, Krishna Kadaveru, Michael R Gryk, Vamsi Kundeti, Sanguthevar Rajasekaran, Martin R Schiller BMC Bioinformatics 01/2010; 11:328. · 2.67 Impact Factor

Article: **A proposed syntax for Minimotif Semantics,** version 1. Jay Vyas, Ronald J Nowling, Mark W Maciejewski, Sanguthevar Rajasekaran, Michael R Gryk, Martin R Schiller BMC Genomics 09/2009; 10:360. · 4.04 Impact Factor

Article: **VENN, a tool for titrating sequence conservation onto protein structures**. Jay Vyas, Michael R Gryk, Martin R Schiller Nucleic Acids Research 09/2009; 37(18):e124. · 8.81 Impact Factor

Article: **Minimotif miner 2nd release: a database and web system for motif search.** Sanguthevar Rajasekaran, Sudha Balla, Patrick Gradie, Michael R Gryk, Krishna Kadaveru, Vamsi Kundeti, Mark W Maciejewski,

Tian Mi, Nicholas Rubino, Jay Vyas, Martin R Schiller Nucleic Acids Research 11/2008; 37(Database issue):D185-90. · 8.81 Impact Factor Article: Viral infection and human disease--insights from minimotifs. Krishna Kadaveru, Jay Vyas, Martin R Schiller

PATENTS

Over 40 patents (most of them in the microservices / distributed systems areas, on behalf of Red Hat Inc.) are browsable at https://patents.justia.com/inventor/jay-vyas.

Note: **I believe innovation is more important than IP, both for businesses as well as the broader technology community.**

In that regard, I'm proud to note that all of my software patents (at least up to 9/01/2018) have been filed under red hat's patent policy, which are never used offensively, and liberally encourage innovation, collaboration, experimentation across business boundaries.