

Bruno Bonacci

Software Architect · BigData Architect

London, United Kingdom



🏠 blog.brunobonacci.com | 📧 BrunoBonacci | 🌐 brunobonacci |

🐦 @BrunoBonacci | 📧 bbonacci_work |

get an updated CV from: <https://brunobonacci.com/cv>

"Simplicity is the ultimate sophistication." --- Leonardo Da Vinci

Profile

Here a short summary of my experience and personality.

- Previously at: **Apple, Cisco Systems, Tesco, Channel4, Trainline, Betfair.**
- **over 25 years** of commercial experience.
- **Snowflake Certified Architect** with huge experience on many other platforms.
- **at 16 years old**, I designed, developed and sold my first ERP on my own.
- self-taught, self-motivated, self-started, the passion is my drive.
- most of my career I worked on complex distributed systems focusing on **Scalability, High-Availability, Performance, Security and Resilience.**
- I'm particularly intrigued by the challenges of high-volume environments. For example, at **Apple** I worked on a complex event processing system for **over 1Billion of users**. At **Trainline** I've designed and built a real-time ingestion platform which was ingesting **over 3Tb of events data per day**, and sending the data to a **Petabyte-scale Data Lake**. I've worked for **CISCO** where our cloud system was processing **over 3 billion req/day** and recording those requests in a data-warehouse of **over 1 trillion records**. In Italy I worked for **Infostrada**, the second largest Italian telco, processing **over 10 billion req/day** with over 40 million of users. Worked for **Channel4 Television** in UK with a reach of **over 46 million online users**, 9 million registered users and **over 50 TB of analytical data**. At **Tesco** I designed and developed a near real-time analytics platform to analyse metering data from **+500K Hudl2 tablets**; the platform was processing **175Gb/day, 28M events/h**, storing **over 5Billion events** and collecting **over 26TB of data** in less than 3 months.
- I'm the principal architect of **Samsara**, a **OpenSource Real-time analytics solution** (<http://samsara-analytics.io/>). This system has been used in **Tesco** and **Channel4** for **Streaming Real-time analytics**.
- in the 1999 I joined **Wuerth-Phoenix** and designed an application server in **Java** while no other commercial application servers were available. Designed a **Object Persistence System** (such as **Hibernate**) to persist large database design (**over 3000 tables**).
- I started using **NoSQL** at that time (1999) using an **Object Oriented datastore**

called Versant as main database for very large enterprise applications

- Have good experience in developing multi-threading systems with low latency.
- I spend lot of effort in designing effective test automation, continuous integration and continuous delivery.
- I worked with many different languages (listed below) however I'm now focusing on Clojure, Java and JVM languages such as Groovy and I particularly love Clojure and the LISP language design.
- I have good management experience, at one time I was managing over 26 people in geographically distributed teams.
- I worked for many years improving the Quality of many products by redesigning the architecture with more appropriate and scalable designs, rather than patching the broken code only.
- I enjoy designing systems and being hands on technical development.
- I favour Agile methodologies and Lean development principles preferably without the ceremony of the modern consultancy companies.
- I am passionate, I keep myself very up-to-date reading books, blogs and frequenting technical conferences and meetups.
- I love simplicity.

Nowadays my interest goes to Machine Learning, BigData, Cloud Computing, NoSQL, Hadoop/Spark and functional languages such as Clojure.

Technical Knowledge

Programming Languages	Clojure, Java, Groovy, Javascript (basic), Python (basic), C, C++, Visual C++, Visual Basic, Fox Pro, Visual Fox Pro, Clipper, Pascal, Modula/2, ProLog, Pro*C, QBX, Assembly, PHP, Jacl, Tcl, HTML, XML, XSLT/XSL-FO, XML SOAP, SQL, UML, IDL, ASN.1, Bash, sed and AWK.
Operating Systems	Linux, OSX, MS-Dos, Window, HP Unix 11, SCO Unix, OS/2
Databases and Caches	Snowflake, DynamoDB, RedShift, Cassandra (NoSQL), MongoDB, Redis, Oracle 8-9i-10g,11i, Vesant 6-7 (NoSQL), Vertica, ObjectStore, Informix, Interbase, MySQL 3-4-5, FoxPro DB, DBF, Access, PostgreSQL, Memcached, Ehcache, JBoss cache, Coherence, Infinispan, Terracotta.
Web	Ring, Aleph, http-kit, Netty, Vert.x, Nginx, HAProxy, Spring, Tomcat, Apache HTTPD, Grails, Liferay, JBoss, WebSphere, Weblogic, Varnish, Jetty
Data	Snowflake, Spark, Hadoop, Pig, Hive, Spring Batch, Cascalog, Governance
Clustering	ZooKeeper, Exhibitor, etcd, Keepalived (VRRP), JGroup
Events & Streaming	Kafka, Storm, Kinesis, Camel, ActiveMQ, RabbitMQ, ESB, CEP, SEP, SEDA
Build & Test	test.check, Gradle, Jenkins/Hudson, Sonar, Ant, Maven, Spock, BDD, TDD, Gatling, PMD, Checkstyle, Findbugs, JUnit, TestNG, Mockito, EMMA, Cobertura, Midje, Marginalia, Selenium, Code Coverage, Luntbuild JScoverage, Perforce, Git, Mercurial, SVN, CVS, Visual-SourceSafe, Continuous Integration, Continuous Delivery, CruiseControl, SonarQube
Virtualization & Container	Docker, Kubernetes, VirtualBox, VMWare, vSphere, Xen, Proxmox, KVM, Packer, Vagrant
Cloud	Amazon AWS, S3, EC2, SNS, SQS, DynamoDB, RDS, SES, ELB/ALB, EMR, ElasticCache, OpenShift, Azure
Architecture	Microservices, Distributed Systems, SOA, SEDA, Actor, Web-Services, RESTful Service, Governance
Observability	ElastiSearch, NewRelic, Zipkin, Grafana, Splunk, Ganglia, StatsD, Reimann
Methodologies	ROWE, Agile, Scrum, XP, Waterfall
Model Checkers	TLA+

Studies & Qualifications



- Snowflake SnowPro Advanced: Architect Certification (2024 Snowflake)
- Snowflake SnowPro Core Certification (2024 Snowflake)
- Oracle Database Administration Fundamentals I and II (2006 Oracle)
- Three years of Computer Science Engineering at University of Calabria (Italy)
- High School Science Diploma 58/60 - Science School "E. Fermi" Cosenza (Italy)

Languages

Italian (native)

French (native)

English (fluent)

Work Experience



Redefine.io

Founder

London, UK

June 2020 --- Now

After helping so many companies with their challenges of managing and processing data at scale, I've created a company with the vision of redefining the Data Management in the Cloud. Redefine.io provides services and products to facilitate and enhance the data management capabilities for the teams of tomorrow.

The company offers Data Consultancy and some of the major projects are listed below.

Our mission is to reimagine how data systems are designed today, and build the tools to provide the best experience to the data professionals while working with data. To achieve this ambitious objective, we designed a brand new columnar store database on top of AWS cloud native APIs. The database completely serverless, and it offers high-performances queries on structured and semi-structured data without the need of additional data manipulation. The system is entirely serverless, so the users don't need to provision, maintain, scale and monitor clusters. We designed the database from the ground up, starting from the storage format for the data which needs to be optimised to efficiently make use of the cloud storage APIs, to the serverless query engine and optimiser. We designed a separate control plane which allows us to configure a huge variety of possible solutions for our customers to meet all their regulatory and compliance needs. We developed a SQL dialect to enable users to query the data, but we also developed a richer language to offer an much more intuitive access to the data. We have designed a Serverless User Interface which enables users to perform data exploration tasks and data visualization directly via the console.

Skills Used: Cloud, Data Management, Clojure, Immutability, Function Programming, Database Design, Columnar Store, Key-Value Store, Parallel Computing, AWS, Serverless Distributed System Design



Hyde Housing

Hyde

Data Architect

London, UK

October 2020 --- May 2022

Hyde Housing Association is one of the UK largest housing associations with over 50,000 properties in UK. I was tasked (via Redefine.io) to help the company during their Digital Transformation Programme. My team and I were tasked to design and build a real-time data integration platform and to build a brand new Data Lake on AWS.

We designed a multi-account AWS infrastructure with a Landing Zone, Centralised Logging, and centralised Security. We designed and built a fully automated CI/CD pipeline to Continuously Deploy infrastructure changes and microservices into all accounts and all environments.

The Real-Time Data integration platform had to capture changes in key data records in legacy systems such as Northgate, Keystore and Oracle Finance, and synchronise the data into the new Cloud systems such as Salesforce, Workdays and Total Mobile Connect while keeping the Cloud-based Data Lake up-to-date with the same information. The Data Integration Platform used Confluent Kafka to collect database changes via a Change Data Capture to then distribute the changes downstream.

To avoid coupling with the source data systems (legacy), each records was then mapped to a canonical representation of the key entities from our Unified Data Model. Once the data records were converted into Hyde's Canonical Model, the same record was then distributed to all target systems and to the DataLake.

The design and definition of a Canonical Model helped tremendously with the Data Governance which for a streaming system is much more challenging. The use of Schema Registry and the definition of very strict rules for schema evolution helped to maintain producers and consumers aligned by their data contracts.

We designed a built various microservices, data products, and deployed them into AWS Fargate, designed to capture, or expose data products in real time.

The Cloud-based Data Lake was based on a Medallion Architecture. The raw data was captured as-is and then a series of transformation using AWS EMR, Spark, AWS Glue was producing the Business-Level aggregates ready for analytics and reporting. We designed and build a number of PowerBI dashboards including a dedicated dashboard to analyse and extract insight from arrears.

We also designed a Machine Learning and AI pipeline to extract essential information from Gas Safety Certificates, EICR and other important documents and enrich the data in the Data Lake and other systems of records.

More information about this project available in our blog

<https://redefine.io/blog/transforming-a-housing-association/>

Skills Used: Cloud, AWS, Data Management, Real-time streaming, Kafka, Confluent, Clojure, Immutability, Function Programming, Database Design, Governance, Data Lake, Data Modelling, Data Quality, PowerBI, Data Analytics, Realtime Data Integration, Machine Learning, IaC - Infrastructure as Code, CI/CD Pipeline, Microservices, AWS Fargate, Container, Docker, REST API.

VIOOH VIOOH Ltd

London, UK

Head of Architecture | Head of Engineering (interim)

January 2019 --- June 2020

VIOOH is a spin-off of JCDcaux born to optimize and automate the scheduling and delivery of Advertisement campaigns in the Out-Of-Home space. I designed and built ground up a real-time DataLake which ingests and process events from Digital displays deployed in 15 countries, events from CMS systems, events from Campaign scheduling and Optimisations Systems. I designed and built an in-house, fault-tolerant, self-repair Kafka cluster which ingests over 1,800 msg/s and 1 billion events per day.

The company was formed with the acquisition and merging of separate monolithic systems. I've redesigned and promoted a Microservice Architecture, promoted the transition from monolith to micro-services in gradual steps. I've migrated the inter-service communication to a Kafka message exchange. Schemas between services served as contracts for reliable communication and safe evolution. All schemas were described in a Common Model.

I've designed and built ground up a Continuous Deployment system to replace the legacy Continuous Integration in Jenkins. The in-house solution I've built was completely serverless (AWS Lambda + CodeBuild), it required no project configuration and completely driven via Slack commands. It offered the easiest and fastest solution for developers to control deployments without requiring direct access to clusters.

I've been a strong promoter of Chaos-Engineering and continuous Chaos Testing practises. I've promoted a change in the organisation to match the micro-service structure, small teams 3-5 people with the ownership of one or more micro-services.

I've been driving the recruitment of the technical staff and designed various take-home technical challenges. I've been the promoter of Open-Source contribution and open-sourced a few internal tools we build. I've started and promoted #UpSkillFridays initiative to provide engineering team time and organised groups for collaborative learning.

Skills Used: Kafka, AWS, EC2, S3, ECS, Docker, DynamoDB, Kinesis, AWS Lambda, Serverless architecture, EMR, Hadoop, Spark, ElasticSearch, Kibana, Terraform, Packer, Clojure, Functional Programming, REST services, Microservices, Swagger/OpenAPI, Presto, AWS Glue, Parquet, AWS Athena, AWS API Gateway, Jenkins, Mircoservices, Chaos-Engineering.



Apple

London, UK --- California, USA

Applied Machine Learning Engineer

December 2017 --- December 2018

As part of the Applied Machine Learning team, I worked on a platform that enables Data

Scientist to define, develop, deploy Machine Learning algorithms to analyse behavioural patterns of user and machine interactions. The system was initially built as a fraud detection system, but it later expanded to map and analyse behaviours from many lines of business within Apple such as ApplePay, iTunes, Online store, Retail, Manufacturing and many other Apple systems. The platform is an **extremely highly concurrent complex event processing system** processing many tens of thousands of events per second from over 1 billion users in **real-time**. It implements a rule-engine which can be used to analyze, alert or block actions within various systems. I designed and extended a Clojure DSL used by the Data Scientist to interact with the platform. I designed a new approach that enabled compile-time optimizations on the rule engine DAG, re-implemented ground up a dynamic isolation classloader for the runtime engine evaluator.

Skills Used: Clojure, Java, Netty, Cassandra



Trainline.com

BigData Consultant

London, UK

May 2016 --- November 2017

I've joined the Trainline.com to help them to build the next generation data systems, leverage the enormous amount of behavioural data at their disposal and build real-time recommendation, classification, and personalisation systems. I was tasked to design and build a platform to allow fast prototyping and reliable deployment of data products such as predictive models, classification models and recommendation systems. I was also tasked to collect, organise data into a Data Lake, to design and build real-time collection systems, data validation and enrichment pipelines. I've used serverless architecture and AWS Lambdas to create restful services. I've designed and built a real-time ingestion platform to ingest events from over 300 microservices and feed the data-lake as well as design a messaging system to decouple service-to-service communication. The system ingests **over 3Tb of events data per day** and dispatches the events to multiple destinations including the data lake. I've designed and built a **Petabyte-scale data lake**. I've designed and built a datastore for predictive models, a K/V Store with transactional guarantees on bulk load (soon to open-source). I was lucky to join the company at the beginning of their data programme and set the foundations for a platform which allows data scientists to quickly prototype a data model and deploy it to production with a simple "one-click" deployment solution.

Skills Used: AWS, EC2, S3, ECS, Docker, DynamoDB, Kinesis, Firehose, AWS Lambda, Serverless architecture, EMR, Hadoop, Spark, Elasticsearch, Kibana, Terraform, Packer, Cascalog, Clojure, Functional Programming, REST services, Microservices, Swagger, Hive, Presto, Parquet, AWS Athena, AWS API Gateway, Jenkins.



Channel 4

BigData Architect in R&D

London, UK

March 2015 --- March 2016

I've been working within the R&D department to explore, find and promote, BigData technologies with the aim of improving Channel4's capabilities to leverage the large amount of data at their disposal.

In this context I've designed and developed a solution for Spot placement optimisation based on programme subtitles. In other words I utilised the TV programme subtitles to understand what the programme is talking about and use this information to place commercials which are more relevant to the topic discussed. This process it is called Topic Modelling and the methodology used is the Latent Dirichlet Allocation (LDA) which extracts topics or classifies documents of a given corpus. In order to improve the matching the system was complemented with the Princeton's WordNet database which provided lexicographic information about the relations between words.

I've also designed and developed a real-time streaming platform which enables Channel4 to make sense of huge amounts of data in real-time. The system is based on my open-source solution called samsara-analytics.io which combines the power of open-source technologies such as Apache Kafka and ElasticSearch to provide a much simpler approach to bespoke analytics solutions. The solution I've designed and built demonstrated how with a single solution you can analyse multiple streams of real-time data, and not only make sense of what's happening "right-now", but also be able to analyse historical data with an interactive interface, and effortlessly extracting insights from billions of records in seconds. The system is entirely developed in Clojure and deployed in the AWS cloud. Using the scalable streaming capabilities of the solution, I've also designed and built the building blocks of a Multi-Variant-Testing (MVT) solution which enables the company to dynamically make decisions based on user's behaviour. Additionally I've designed and built a real-time recommendation system, for the video-on-demand platform, based on co-occurrences Machine-Learning algorithm which allows the system to react immediately (3 millis) to user interactions and dynamically learn user's preferences. Currently new recommendations are computed with overnight batch jobs.

Skills Used: AWS, EC2, S3, Cloud, BigData, ElasticSearch, Kibana, Clojure, Stanford CoreNLP, Topic Modelling (LDA), Mallet, Sentiment analysis, Machine Learning (ML), Spark, Hadoop, EMR, Docker, Fig, Docker-compose, Neo4j, Graph DB, Natural Language Processing, AWS, EC2, S3, ELB, EMR, Cloud, BigData, ElasticSearch, Kibana, Clojure, Samsara, Terraform, Docker, Packer, Consul, Spark, Spark-Streaming, Cascalog, Hadoop, Real-time Analytics, Stream Events Processing (SEP), Unified Log, Machine Learning, Recommendations engines.

 **Tesco**
BigData Architect

London, UK

May 2014 -- February 2015

I've designed and built a near real-time analytics platform to collect, analyse and correlate metering data coming from Tesco's Hudl2 tablet. Hudl2 is an Android tablet awarded with the

“Best Value Tablet” at the 2014 TrustedReviews Awards.

The platform was composed by an async RESTful layer which was ingesting the metering data and device events. The APIs were designed to achieve a 1500 req/s with single digit latency. A data processing pipeline, in real-time, would then enrich the events with additional data streams, correlate independent events deriving higher level/aggregated ones. The data processing pipeline and the ingestion APIs were both written entirely in Clojure. Stateful event stream processing was achieved with the support of an in-memory datastore which was then backed by Cassandra. Once the raw data was processed by the data-pipeline the output was then indexed into an ElasticSearch cluster which offered real-time query and aggregation capabilities. The platform collected over **26TB** in less than **5 months**. At peak time it processed **28 million events per hour**, and receives **175GB per day** from over half million **devices**. We ran our analytics queries on over 5 billion events in ElasticSearch in which we were keeping last 90 days. The platform was entirely designed, developed, tested and maintained in operation by just 2 people (myself and one more person).

Skills Used: BigData, Clojure, Cloud, Azure, Analytics, real-time, Kafka, Cassandra, ElasticSearch, Kibana, Graphite, StatsD, Linux, Cascalog, Hadoop, ZooKeeper, Exhibitor, Midje, Marginalia, HDInsight, Hortonworks



Postreet.io

London, UK

Co-founder and Principal Architect

June 2013 --- March 2015

Postreet.io is a platform for geo-localised mobile advertisement. It allows to send messages to a mobile device based on the device's location. The service provides RESTful webservice interface and a mobile SDK to access the APIs. We had a very successful trial of the system and we are looking for more commercial engagement with companies. My duties were to design the overall architecture, implement the all server-side infrastructure, and maintain the services in operations.

The platform offered the basic building block for a variety of location based services which we sought to implement with different business partners.

Skills Used: BigData, Java, SOA, Microservices, RESTful Web-Services, Hadoop, Pig, Hive, Cascalog, Clojure, Redis, Cloud, AWS, EC2, EMR, S3, ELB, Route53, Glacier, MongoDB, RabbitMQ, HAProxy, Vagrant, Docker, Splunk, Analytics, Graphite, StatsD, Event-sourced, Midje



Channel 4

London, UK

BigData hacker / BigData Architect

May 2013 --- January 2014

I've been working within the R&D department to explore, find and promote, BigData technologies with the aim of improving Channel4's capabilities to leverage the large amount of data at their disposal. During this period I have worked on the following projects:

- Built a system able combine recommendations and personalisations in real-time from several event streams and able to support over 1500 req/s using Storm in DRPC mode.
- I've designed a recommendations system based on Mahout/Pig for C4 video content for ~9 million of users and ~200M of user's actions per month. The system is based on Mahout's Item-based Collaborative Filtering which analyses the user's behaviour to predict/recommend new interesting content to watch. I've also designed a built a system to evaluate comparatively the recommendations against the currently used system.
- During this project I've also explored Prediction.IO and ways to scale it using AWS/EMR.
- C4 data analysts use a home grown platform for BigData called BigData Control Panel (BDCP) which is backed by Hive/Hadoop and AWS EMR to process over 50 TB of analytical data.
- To reduce the query latency I've built a RedShift cluster with ~8 billions records and measured a 100x-700x speedup compared to a cost equivalent solution with Hive/Hadoop.
- I've worked along with the Data Scientist team in a demographic targeting of C4 user base to scale and automate their SVDs computations. Other technologies briefly explored are: Cloudera CDH4/5, Impala, Parquet format, Cascalog, Docker.

Skills Used: BigData, Java, SOA, Groovy, Vert.x, Web-Services, Hadoop, Pig, Hive, Cascalog, Clojure, Redis, Cloud, AWS, EC2, EMR, S3, MongoDB, Storm, HAProxy, Whirr, PredictionIO, Mahout, Route53, Riak.

FIS Global | Debit & Prepaid cards

London, UK

Senior Technical Architect

December 2011 --- December 2012

FIS is a global company providing services for card payment for banks and financial institutions. FIS offers a Card Processor Service to provide realtime authorization of cards payments in connection with Visa and MasterCard. My role is to support development and operational teams with enterprise design of SOA services and outline an architectural roadmap to fulfil scalability and high-availability objectives with technical recommendations. I've defined the solution and technical architecture, ground-up, for the entire web and service infrastructure, replacing proprietary containers and high-end hardware with open-source, container-less solutions and commodity hardware. My focus was to improve scalability, performances, resilience and maintaining a simple SOA/RESTful design. This was a green field project for which I've selected the full stack of tools and technologies to use, favouring small light-weight asynchronous web framework to J2EE containers. I've been leading the analysis and resolution of critical production issues. I was responsible for reviewing and analyzing penetration test reports as well as improving security by correcting design flaws. I was responsible for analyzing architecture and code of critical batch processes and improve performances (from over 32 hours of execution down to 4 hours circa). I was responsible to select/recruit a team in India, as well as selecting team members for London's office,

conducting interviews and preparing technical tests. I was the promoter of technical innovation and evangelisation, by organising the "Tech@Lunch" initiative. A series of technical presentation at lunch time with the intent to talk about new technologies, demoing new languages, products and tools, and stimulate innovation.

I used Splunk (BigData log analyzer) in development and staging during root cause analysis and finding correlated events causing live issues. Built service dashboard extracting useful metrics in real-time, offering a platform for log exploration for non-technical teams and business members. Analyzed logs from production for auditing and data-forensic purposes after outages discovering potential DOS attacks. Extracted useful information and metrics about user traffic on customer websites.

Skills Used: Java, J2EE, SOA, Web-Services, Multi-threading, networking, Tomcat, Apache, Liferay, Informix, Spring, Continuous Integration, WebSphere. Green-Field project: HA-Proxy, Keepalived, Nginx, Varnish, Vertx, Jetty, Redis, Jackson, Gradle, Groovy, Grails, Spock, Gatling, Camel, ActiveMQ/RabbitMQ, EDA, ESB, Jenkins, Continuous Delivery, Polyglot systems, Storm, Hadoop, Pig/Hive, Cassandra, VirtualBox, Vagrant, Splunk.

AKQA [AKQA \(akqa.com\)](http://akqa.com)

London, UK

Senior Technical Architect

April 2011 --- August 2011

AKQA is a global digital agency specialised in designing high-performance car configurators, and global sites. My responsibilities were to define the architecture for parts of the global site for one of the world largest car manufacturer **Volkswagen**. This included front-end architecture as well as back-ends web services. This position was very hands on development, where I've been focusing on developing critical parts of the system, as well as review and improve the code with developers. As architect I've participated in the governance programme ensuring that new design and implementations adhered to all defined defined rules and conventions, and join the governance committee to review global architectural decisions. The site was built using portals technology powered on back end by a network of web services. The environment was Agile with Scrum, CI, and automated tests.

Skills Used: Java, J2EE, SOA, Web-Services, Multi-threading, networking, Tomcat, Apache, Liferay, Mysql, Spring, Camel, Cloud Services, Continuous Integration, Governance.

HouseHub.org

London, UK

Founder and Principal Architect

December 2010 --- December 2013

Househub.org bridges the most popular sites for properties renting in the London's area to enable the final users to search for a property across all sites at once. It uses the passive search paradigm (alert system), in other words the users define his search criteria, and passively he gets all new property ads matching those criteria as soon as they are advertised. The system back-end uses the Staged Event Driven Architecture approach (SEDA) and

implemented using a lightweight ESB. The front-end is used to define users' search criteria and it is implemented in Groovy and Grails. The systems checks continuously the most popular property sites, when it finds new ads, it scans them and retrieves relevant information. Then it matches the information with users search profiles and it notifies them via email. The user saves a lot of time on property hunting because he doesn't need to visit and search on all property sites, but he gets notified when something relevant come up in the market.

I've designed and integrated a recommendation/similarity algorithm with the k-nearest-neighbours approach using a mix Pig/Hadoop and SEDA strategies. The Hadoop job cross-matches over **3 million adverts in a 9 trillion dataset** (over 1.5 Petabytes). The "distance" algorithm is a unidirectional, multi-category normalized computation over the following properties: geographical distance, price, property size and advert "freshness". Due to the complexity and scale of the problem I've then changed the approach and calculate similarities in ingestion (lambda architecture). In ingestion adverts are added to a memory-mapped file hash-trie of geohash tags that serves as partition base for computation of the distances. This solution allowed me to pre-compute similarities using a single EC2 small instance rather than a large EMR cluster.

I've integrated Splunk not only for the log analysis but as a platform for application metrics, building service dashboards and using it for root cause analysis. I used the tool for strategic decisions based on trend analysis. Currently it collects hundreds of service/application metrics offering a powerful data-visualization and enabling a fast and accurate operation support.

Skills Used: Java, J2EE, EDA, ESB, Multi-threading, networking, Groovy, Grails, Tomcat, Apache, Mysql, Spring, Camel, HA-Proxy, ActiveMQ, Cloud Services, SEO, Pig, Hadoop, MongoDB, Cassandra, NoSQL, Splunk, Ganglia, OpenVPN, AWS, S3, EC2, SNS, SQS, DynamoDB, SES, ELB, Redis, CEP, SEP, Hadoop, Pig, EMR



Betfair

London, UK

Quality Assurance Manager

September 2010 --- End of

November 2010

Betfair provides a Sport-Exchange Platform on a global scale. I was QA Manager of the Sport Platform API. I manage a team of ten automation engineers (permanent and contractors) throughout a global release of a new service within the Sports Platforms API (SOA). I've introduced a Continuous Integration in my functional area based on Hudson and Sonar.

Skills Used: Java, J2EE, SOA, Webservice, Continuous Integration, Agile, Scrum, Hudson, Sonar, Jira, RallyDev, Perforce.



CISCO is the world leader provider of SaaS security. CISCO provides a real-time traffic scanning to block viruses, malware, spyware and unwanted content. My duties are to establish the QA department defining process, rules, guidelines and review architecture. Introduce a Continuous Integration based on Hudson, Sonar, Groovy, Luntbuild, Maven 1 and 2, Junit/TestNG/Qunit, Selenium and JSCoverage, static analysis of source code (FindBugs, PMD, CheckStyle, Sonar).

The environment is Java 6, Spring 2.5, Oracle, Vertica. The strategy applied is based on Automated Software Quality via JUnit and Code Coverage (Emma/Cobertura). My duties cover also managing a small team of 3 people with functional and technical coordination (1 offshore). The QA department has to do strict code review, verify the implementation and tests written by the development team, review the architecture in order to improve scalability, performances and security. The QA is in charge of creating performances and load/stress test suites, performing fine tuning on JVM Garbage Collector, to create framework and basic infrastructure to simplify the implementation of automatic test cases for the developers. For example I reviewed the architecture and implementation of a multi-threaded software proxy, and improved its design resulting in 5+ times faster.

My duties include review architectures and implementation, refactoring for performances improvements, manage development projects of application to support the development process, integrate Agile methodologies and guarantee high quality for functional delivery. Code analysis, identification of vulnerability (functional and security), DoS protection and design test harness to verify reliability under load. I've designed and implemented a Continuous Integration system in Groovy using Maven and Hudson. The development process itself is fully automated and supported by a set of systems developed in Groovy/Java/Spring designed and implemented by myself and my team.

Skills Used: Java6, J2EE, Spring, JPA, Hibernate, Oracle, Vertica, ActiveMQ, WebServices, Groovy, Maven1/2, Hudson, Sonar, Luntbuild, JUnit, TestNG, Code Coverage (Emma/Cobertura), FindBugs, PMD, CheckStyle, Selenium, JSCoverage, Qunit, Agile, Scrum, Continuous Integration, Multi-threading, OO Design, SOA Design, Java networking, C, C++, Full development lifecycle, Management, Team leading, Project Management, Networking, HTTP, FTP, ICAP protocols.



My duties were to design internal processes and define team collaboration for the whole Wuerth Phoenix International group. The software development process was defined using Agile concepts from Extreme Programming and Continuous Integration. As Senior Core

Architect I've participated to the design of an application server called Zeus (project started in 1998 in my previous employment in this company), for which we received a nomination for the "International OMG Awards 2003" in London for "Best MDA (Model Driven Architecture) approach", for our software architecture and development methodology based on UML model design. As Development Manager I had to design and realize MDA framework support for unit and integration testing, design a framework for automated performance measurement under Continuous Integration environment.

This role included both management and hands on design and development activities. From the management side my duties were to establish the QA department in the company, recruit team members, plan budgets for the department, define department annual targets and coordinate the communication and collaboration between all locations. As team manager my duties included work assignment, annual reviews, coaching, define deliverables, salary definition, objectives and bonus assignment.

From the technical side my duties included to review architectural design of the application server, develop parts of the framework and the application server using Java/J2EE APIs, take care of performances, scalability and security issues, and check the multi-threading operations. Define criteria for security as password management, ACL and access to customer sensible data. The development included database connectivity with JDO/JPA for Oracle and Postgre, and JVI for Versant, CORBA and RMI for communication layers. The applications were deployable in Tomcat, WebSphere or JBoss with rich clients (Canoo/ULC), and the enterprise integration with legacy systems through JMS, EJBs, and Web services.

Projects realized here:

- Design of software development processes and methodologies for Wuerth Phoenix with the target to define clearly responsibilities and information exchange between development teams organised in several locations (2 branch offices in Italy, 2 in Switzerland, 3 in Germany, 1 in Hungary, 1 in Netherlands, 1 in USA e 1 in China), looking after language differences and time zone difference. This process definition gave the possibility to Wuerth Phoenix to support a very large growth (from 135 people in 2003 to 465 in 2004) and to optimise resources improving response time for our customers (average response time improved of 63%).
- Design of internal product Zeus, an application server and development framework used to design and develop all other company products. This software was developed totally in Java/J2EE using the MDA approach.
- Design of Automated Software Quality systems (ASQ) based on concept of continuous integration with static and dynamic analysis of source code, performances analysis, load analysis, functional analysis for multi-platform certification (Windows, Linux, HPUX), multi-database tests (Oracle, Versant, PostgreSQL,...), and semantic check on models and architecture break checks. With this approach the company reduced the number of bugs by the 93% in 14 months.
- Organisation of training and designing certification process for our products into company's academy.
- Design of systems for development process auditing with graphical project trend

analysis, activity analysis for design, development, QA and support phases, analysis of service level agreements (SLA) and response time for support teams.

- Responsible for the Quality Assurance department for all products developed in Wuerth Phoenix with 22 people in 4 branch offices (Italy [Rome and Bolzano], China and Germany).
- Design and implementation of an Enterprise Integration System (EIS-ETL) for heterogeneous systems (integration also in real-time). The system was a part of Zeus and was using a SQL-like language to define the output formats to export data. The focus of this system was to provide ETL capability to the Zeus application server, with an easy and flexible language as well as a very powerful and robust tool to migrate huge database structures.

Skills Used: Java, J2EE, Groovy, C/C++, CORBA, RMI, XML, UML, MDA, XSLT, Oracle, Versant (NoSQL), MySQL, ActiveMQ, PostgreSQL, JUnit, Luntbuild, CruiseControl, ANT, Maven 1&2, Code Coverage (Emma), SVN, CVS, MDA, ASQ, Linux, Windows, HP-UX, Design Patterns, Object Oriented, SOA design, ULC, Swing, Applet, Servlet, HTML, JMS, EJB, Tomcat, WebSphere, JDO, JPA, SQL, Agile, Scrum, XP, Continuous Integration, Multi-threading, Full development lifecycle, Management, Team leading.

Infostrada Spa

Milan, Italy

Development Manager, Senior Enterprise Architect, Senior Developer

May 2001 --- August 2002

For Infostrada Spa (the second largest Italian telecommunication company) I've created a MDS Collector Manager to collect CDRs (Call Details Record) from telecommunication hubs located all over national territory. Analysis, software architecture and development of a daemon Unix (HP-UX11 x SuperDome) in C++ entirely done by me. The system, declared as Mission Critical (max downtime certified), was organised into a main daemon that controls several distributed agents via IPC. Communication protocols with telecommunication hubs were specialised for each kind of vendor/producer of hub, the MDS Collector Manager supports a plug-in architecture system to be able to add new agent for new hubs in an easy and secure way. The biggest design efforts was directed to obtain a certified High Reliability proved by over 700 automatic test cases able to check a large number of possible problems. Moreover I had also to specify some procedures of Disaster Recovery and Incident Management, and define methodologies for active live monitoring.

I've done the analysis of organisation solutions for configuration management. Within this project I've coordinated defined strategies of CM via CVS for all development teams.

I've been the Development Manager of a team of 15 people. I've integrated systems for traffic management during the merging between Infostrada and Wind SPA. The existing system of Infostrada MDS Mediation Device System was selected to manage the whole traffic of both companies. From the technical point of view the MDS of Infostrada was a C++ application designed for HP-UX Superdome with Oracle 9 database, Wind SPA network was based on Ericsson's technology, and the MDS was designed to work with Alcatel network. My role

included the team management as well as redesign part of the MDS architecture and extend it in order to include the new platform traffic. Design and implement plug-ins to extend the functionalities of BillingGateway (Ericsson).

Skills Used: C, C++, Oracle, HP-Unix, Daemon, CVS, ASN.1, HP-Superdome, UML

AAA Bank Spa

Milan, Italy

Lead Architect

August 2002 --- April 2003

I've designed and realised an online-trading and online-banking software. The system based on J2EE (WebSphere/Oracle) was designed as a backend service (SOA) to support online trading for multi-channel banks, and the possibility to plug several real-time Info provider (MilanoFinanza and Reuters) for the market stock quotes. The online-trading service was providing the most common capabilities (like: portfolio management, stock alert, planning, etc). The online banking offered two options: front-end and back-end fully implemented in Java implementation, or a live connection with existing back ends like a DeltaDator spa solution (AS/400) or a remote connection via EJB services. The front-end application was based on JSP/Servlet architecture driven by a decision maker engine able to provide a full control of the application workflow as well as taking care of security issues. The identification and authentication mechanism was driven by a Single Sign-On (SSO) solution. Due to the strict security policies we had to implement a customized middle-ware with the aim to protect sensitive data, and build test harness to verify the correct operatively as well as making intrusion tests.

Skills Used: Java, WebSphere, SOA, J2EE, EJB, JSP/Servlet, SSO, UML

Data Mark Brokerage Services srl

Milan, Italy

Senior Architect

May 2002 --- August 2002

I've made the analysis and design and implementation of a skill inventorying system realised in PHP, Mysql and Java/Lucene.

Skills Used: PHP, MySQL, Java, Lucene



Würth s.r.l.

Rome, Italy

Senior Technical Architect, Senior Developer

February 1998 --- May 2001

As a software architect and developer in an international project. Here I've designed an ERP for middle sized and big sized (more than 3000 employees and more than 500 system users) trading companies. For this project we designed an application server called Zeus in Java at a time when there wasn't a open-source application server for Java. The analysis and design phase was entirely done in UML by using an MDA approach. I was responsible for designing part of application server and of analysis, design and development of an procurement

engine able to analyse warehouse good stocks and make fully automatic purchase orders for more than 300.000 products, looking at worldwide suppliers for best conditions (price, delivery time, packaging, etc etc).

Skills Used: Java, Swing, CORBA, Versant, NoSQL, Object-Store, Multi-Threading, UML, Unix, HP-UX.

AZ Computers Srl

Cosenza, Italy

Head of Development | Lead Developer

February 1997 --- February 1999

Responsible for the Software Development division, analyst, software architect and programmer. This small software company developed 6 products like: ERP system for fashion companies (retail and distributors), ERP for salesman networks, ERP of hospitals, financial and accounting system. I've been a trainer in several training session for our customers.

Skills Used: Fox, Fox-Pro, Clipper, Sco-Unix, Windows, Visual Fox-Pro

Teorema Ufficio

Cosenza, Italy

Head of Technology

January 1996 --- February 1997

Responsible for the computer and technology division. My duties covered from contact with supplier, purchase orders, assembling computers, network installations, and customer support.

Skills Used: MS-DOS, Windows, Networking, Novell Netware, Hardware assembling

Freelance --- Self

Cosenza, Italy

Senior Developer

1992 --- 1996

I've realized several projects for small companies and banks. Collaboration with Olivetti branch office managing the networking division.

Skills Used: QBX, Basic, Assembly, C, C++, MS-DOS, DR-DOS, Windows

Interests

- Passion for photography. I made three exhibitions in Bolzano (Italy).
- Mountaineering and Rock Climbing which I do practice regularly.
- Travelling, books, music.

References available upon request.