

Software Architectures

When we design a software architecture, there are lots of strategies can enhance performance, optimize cost, and simplify its effort to implement/operate the system.

Architectural Patterns

When we design a software architecture, there are lots of strategies can enhance performance, optimize cost, and simplify its effort to implement/operate the system.

An architectural pattern is a general, reusable solution to a commonly occurring problem in software architecture within a given context. Architectural patterns are similar to software design patterns but have a broader scope.

- [Client-Server architectural pattern](#) — Client-server architecture (client/server) is a network architecture in which each computer or process on the network is either a client or a server.
- [Layered architectural pattern](#) — Layered Architecture is commonly using when design J2EE application - it is something like defacto standard and widely known by most architects, designers, and developers. The layered architecture pattern closely matches the traditional IT communication and organizational structures found in most companies, making it a natural choice for most business application development efforts.
- [Master-Slave architectural pattern](#) — Master/slave is a model of communication where one device or process has unidirectional control over one or more other devices. In some systems a master is selected from a group of eligible devices, with the other devices acting in the role of slaves.
- [Pipe-filter architectural patterns - useful approach for analytics based on daisy chain](#) — A very simple, yet powerful architecture, that is also very robust. It consists of any number of components (filters) that transform or filter data, before passing it on via connectors (pipes) to other components. The filters are all working at the same time. The architecture is often used as a simple sequence, but it may also be used for very complex structures.
- [Broker architectural pattern - usually used in Message Middlewares like Apache ActiveMQ, Apache Kafka, RabbitMQ](#) — The Broker architectural pattern can be used to structure distributed software systems with decoupled components that interact by remote service invocations.
- [Peer-to-peer architectural pattern - usually used in file sharing like BitTorrent or multimedia protocols like P2PTV or PDTP](#) — Peer-to-peer architectural pattern (P2P architecture) is a commonly used computer networking architecture in which each workstation, or node, has the same capabilities and responsibilities. It is often compared and contrasted to the classic client/server architecture, in which some computers are dedicated to serving others.
- [Event-bus architectural pattern - used in Android app development and Push Notification Service](#) — Event-bus architecture pattern is a distributed asynchronous architecture pattern to create highly scalable reactive applications. The pattern suits for every level application stack from small to complex ones. The main idea is delivering and processing events asynchronously.
- [MVC\(Model-view-controller\) architectural pattern - used in web application development like Django and Rails](#) — MVC(Model-view-controller) is an architectural pattern commonly used for developing user interfaces that divides an application into three interconnected parts. This is done to separate internal representations of information from the ways information is presented to and accepted from the user. The MVC design pattern decouples these major components allowing for efficient code reuse and parallel development.
- [Blackboard architectural pattern - used in NLU\(Natural Language Understanding\) and car recognition](#) — Blackboard architectural pattern usually used in Artificial Intelligence Software to implement Non-Deterministic Algorithm.
- [Interpreter architectural pattern - used in database query language such as SQL](#) — This pattern is used for designing a component that interprets programs written in a dedicated language. It mainly specifies how to evaluate lines of programs, known as sentences or expressions written in a particular language. The basic idea is to have a class for each symbol of the language.
- [Microservice architectural pattern aka MSA - one of the trends of service architecture in 2018-](#) — MSA (Microservice architecture), or simply microservices, is a kind of methodology to develop service by divide and conquer - a distinctive method of developing software systems that tries to focus on building single-function modules with well-defined interfaces and operations. The trend has grown popular in recent years as Enterprises look to become more Agile and move towards a DevOps and continuous testing. Microservices can help create scalable, testable software that can be delivered weekly
- [Active-Active / Active-Stand-by Architectural Pattern](#) — Active-Active / Active-Stand-by architectural patterns are commonly used when we consider fail-over in server deployment design.

AWS Architecture References

Amazon Web Services is a subsidiary of Amazon that provides on-demand cloud computing platforms to individuals, companies and governments, on a paid subscription basis. The technology allows subscribers to have at their disposal a virtual cluster of computers, available all the time, through the Internet.



- [Image Moderation Chatbot - a serverless chatbot on AWS that monitors chat channels and removes images suggestive or explicit image](#) — The below architecture shows you how to build a serverless chatbot on AWS that monitors your chat channels and removes images containing suggestive or explicit content.
- [Web Application Hosting provides the reliable, scalable, secure, and high-performance infrastructure](#) — High available and scalable web hosting can be complex and expensive. Dense peak periods and wild swing in traffic patterns result in low utilization of expensive hardware. Amazon Web Services provides the reliable, scalable, secure, and high-performance infrastructure required for web applications while enabling an elastic, scale-out and scale-down infrastructure to match IT costs in real time as customer traffic fluctuates.
- [Live streaming / Static Content & Media Service based on AWS](#) — AWS provides a suite of services specifically tailored to deliver high-performance media serving. Each service features pay as you go pricing on an elastic infrastructure - meaning that you can scale up and down according to your demand curve while paying for only the resources you use. BEcause this infrastructure is programmable, it can react quickly.
- [Batch processing composed of job manager, job store and worker based on AWS](#) — Batch processing on AWS allows for the on-demand provisioning of a multi-part job processing architecture that can be used for instantaneous or delayed deployment of a heterogeneous, scalable "grid" of worker nodes that can quickly crunch through large batch oriented applications in place today that can leverage this style of on-demand processing, including claims processing, large scale transformation, media transcoding and multi-part data processing work.
- [Fault Tolerance & High Availability based on multi Availability zones in AWS](#) — Most of the higher-level services, such as Amazon Simple Storage Service(S3), Amazon SimpleDB, Amazon Simple Queue Service (SQS), and Amazon Elastic Load Balancing (ELB), have been built with fault tolerance and high availability in mind. Service that provide basic infrastructure, such as Amazon Elastic Compute Cloud (EC2) and Amazon Elastic Block Store (EBS), provide specific features, such as availability zones, elastic IP address, and snapshots, that a fault-tolerant and high available system
- [Large scale computing based on huge data sets in AWS](#) — These scenarios involve huge data sets collected from scientific equipment, measurement device, or other compute jobs. After collection, these data sets need to be analyzed by large-scale compute jobs to generate result data sets. Ideally, results will be available as soon as the data is collected. Often, these results are then made available to a larger audience.
- [Advertisement serving in AWS](#) — Internet advertising services need to serve targeted advertising and must do so under limited time. These are just two of multiple technical challenges they face.
- [Disaster recovery solution in AWS for local applications](#) — Disaster recovery is about preparing for an recovering from any event that has a negative impact on your IT systems. A typical approach involves duplicating infrastructure to ensure the availability of spare capacity in the event of a disaster.
- [File synchronization service architecture in AWS](#) — Given the straightforward, stateless client-server architecture in which web services are viewed as resources and can be identified by their URLs, development teams are free to create file sharing and syncing applications for their departments, for enterprises, or for consumers directly.
- [Media sharing service architecture in AWS](#) — Media sharing is one of the hottest markets on the Internet. Customers have a staggering appetite for placing photos and videos on social networking sites, and for sharing their media in custom online photo albums.
- [Synchronous online gaming architecutre in AWS - Highly Available, Scalable & Elastic to support Millions of Players](#) — The below architecture is intended for online mobile & games. These workloads are a natural fit for running on Amazon Web Services, due to unexpected traffic patterns & high demanding request rates.
- [Financial services grid computing in AWS](#) — Financial services grid computing on the cloud provides dynamic scalability and elasticity for operation when compute jobs are required, and utilizing services for aggregation that simplify the development of grid software.
- [Web front-end architecture in AWS for e-commerce website](#) — Maintaining an e-commerce website with a large product catalog and global customer base can be challenging. The catalog should be searchable, and individual product pages should contain a rich information set that includes, for example, images, a PDF manual, and customer reviews.
- [Checkout service architecture in AWS for e-commerce website](#) — Managing the checkout process involves many steps, which have to be coordinated. Some steps such as credit card transactions, are subject to specific regulatory requirements. Other parts of the process involve manual labor, such as picking, packing, and shipping items from a warehouse.
- [Product Marketing & Recommendation Architecture in AWS for e-commerce website](#) — Product Marketing & Recommendation is one of features that online service websites has to create more user engagement, so online company can try something new together with users.
- [Time series processing architecture in AWS](#) — When data arrives as a succession of regular measurements, it is known as time series information. Processing of time series information poses systems scaling challenges that the elasticity of AWS services is uniquely positioned to address.
- [AWS Cost Calculator for estimating your required monthly budget](#) — Whenever you design your software architecture based on AWS, you may need to know how much cost it will consume while serving your service - following service will help you to get your estimated cost.
- [Secure and Scalable LAMP Service Architecture in AWS](#) — LAMP means Linux + Apache + MySQL + PHP that is commonly used to develop web service for Start-ups that does not require paying license fee, because all of technologies are from open source.
- [Content Distribution for VOD streaming services in AWS](#) — Below architectural diagram shows how to distribute video content to CDN (Amazon CloudFront) in AWS.
- [Large scale messenger service architecture in AWS - Slack case study](#) — Below slack architecture shows how to compose large scale messenger service based in AWS.
- [Enterprise Data Lake Architecture in AWS](#)
- [AWS VPC Setup Steps](#) — Not like architectural diagram, setting up a VPC environment is not that simple, because it requires to click lots of menus in AWS Console after creating that VPC. Based on my experience, that is not easy to remember, because it's rarely happening process and its complex clicking labors. So I think easy video guide is good enough.
- [Importing a wild card domain SSL certificate into AWS Certificate Manager](#) — AWS Certificate Manager requests three items to register your SSL certificate - you will need to copy & paste your STAR.cert to "Certificate body", your_private.key to "Certificate private key", and STAR.ca-bundle to "Certificate chain"
- [Check incomplete multipart uploads if AWS S3 storage size is abnormally big huge compared to the origin storage](#) — You should check incomplete multipart uploads if your Metrics in the S3 Management Console shows larger size than you actually use.
- [What cause a 502 bad gateway error in AWS environment?](#) — A 502 bad gateway message indicates that one server got an invalid response from another. In essence, you've connected with some kind of interim device (like an edge server) that should fetch all of the bits you need to load the page. Something about that process went wrong, and the message indicates the problem.
- ["502 Bad Gateway" Error easy to happen at the combination of nginx and node.js in AWS Elastic Beanstalk](#) — Elastic Beanstalk is a Platform as a Service (PaaS) offered by AWS that simplifies the deployment and management of applications. It supports various programming languages, frameworks, and services. Nginx, on the other hand, is a popular web server and reverse proxy server known for its performance, scalability, and configuration flexibility.
- [Amazon ECS services in a VPC](#) — Using Amazon ECS containers in a VPC, you can break up large applications into smaller pieces that can be launched and scaled separately in a safe environment. But making sure that all of these parts, inside and outside of a VPC, can talk to each other can be hard. There are different ways to make conversation easier, and each has its own pros and cons.

Google Cloud Platform Architecture References

Google Cloud Platform, offered by Google, is a suite of cloud computing services that runs on the same infrastructure that Google uses internally for its end-user products, such as Google Search and YouTube. Alongside a set of management tools, it provides a series of modular cloud services including computing, data storage, data analytics and machine learning. Registration requires a credit card or bank account details. Google Cloud Platform, offered by Google, is a suite of cloud computing services that runs on the same infrastructure that Google uses internally for its end-user products, such as Google Search and YouTube. Alongside a set of management tools, it provides a series of modular cloud services including computing, data storage, data analytics and machine learning. Registration requires a credit card or bank account details.

Google Cloud Platform provides

- Infrastructure as a service
- Platform as a service
- Serverless computing environments.

In April 2008, Google announced App Engine, a platform for developing and hosting web applications in Google-managed data centers, which was the first cloud computing service from the company. The service became generally available in November 2011. Since the announcement of App Engine, Google added multiple cloud services to the platform.

Google Cloud Platform is a part of Google Cloud, which includes the Google Cloud Platform public cloud infrastructure, as well as G Suite, enterprise versions of Android and Chrome OS, and application programming interfaces (APIs) for machine learning and enterprise mapping services.

GCP Products

Category	Product	Functions supported
Compute	App Engine	Platform as a Service to deploy Java, PHP, Node.js, Python, C#, .Net, Ruby and Go applications.
	Compute Engine	Infrastructure as a Service to run Microsoft Windows and Linux virtual machines.
	Kubernetes Engine	Containers as a Service based on Kubernetes.
	Cloud Functions	Functions as a Service to run event-driven code written in Node.js or Python.
Storage & Databases	Cloud Storage	Object storage with integrated edge caching to store unstructured data.
	Cloud SQL	Database as a Service based on MySQL and PostgreSQL.
	Cloud BigTable	Managed NoSQL database service.
	Cloud Spanner	Horizontally scalable, strongly consistent, relational database service.
	Cloud Datastore	NoSQL database for web and mobile applications.
	Persistent Disk	Block storage for Compute Engine virtual machines.
	Cloud MemoryStore	Managed in-memory data store based on Redis.
Networking	VPC	Virtual Private Cloud for managing the software defined network of cloud resources.
	Cloud Load Balancing	Software-defined, managed service for load balancing the traffic.
	Cloud Armor	Web application firewall to protect workloads from DDoS attacks.
	Cloud CDN	Content Delivery Network based on Google's globally distributed edge points of presence. As of June 2018, the service is in Beta.
	Cloud Interconnect	Service to connect a data center with Google Cloud Platform
	Cloud DNS	Managed, authoritative DNS service running on the same infrastructure as Google.
	Network Service Tiers	Option to choose Premium vs Standard network tier for higher performing network.
Big Data	BigQuery	Scalable, managed enterprise data warehouse for analytics.
	Cloud Dataflow	Managed service based on Apache Beam for stream and batch data processing.
	Cloud Dataproc	Big data platform for running Apache Hadoop and Apache Spark jobs.
	Cloud Composer	Managed workflow orchestration service built on Apache Airflow.
	Cloud Datalab	Tool for data exploration, analysis, visualization and machine learning. This is a fully managed Jupyter Notebook service.
	Cloud Dataprep	Data service based on Trifacta to visually explore, clean, and prepare data for analysis.
	Cloud Pub/Sub	Scalable event ingestion service based on message queues.

	Cloud Data Studio	Business intelligence tool to visualize data through dashboards and reports.
Cloud AI	Cloud AutoML	Service to train and deploy custom machine learning models. As of September 2018, the service is in Beta.
	Cloud TPU	Accelerators used by Google to train machine learning models.
	Cloud Machine Learning Engine	Managed service for training and building machine learning models based on mainstream frameworks.
	Cloud Job Discovery	Service based on Google's search and machine learning capabilities for recruiting ecosystem.
	Dialogflow Enterprise	Development environment based on Google's machine learning for building conversational interfaces.
	Cloud Natural Language	Text analysis service based on Google Deep Learning models.
	Cloud Speech-to-Text	Speech to text conversion service based on machine learning.
	Cloud Text-to-Speech	Text to speech conversion service based on machine learning.
	Cloud Translation API	Service to dynamically translate between thousands of available language pairs
	Cloud Vision API	Image analysis service based on machine learning
	Cloud Video Intelligence	Video analysis service based on machine learning
Management Tools	Stackdriver	Monitoring, logging, and diagnostics for applications on Google Cloud Platform and AWS.
	Cloud Deployment Manager	Tool to deploy Google Cloud Platform resources defined in templates created in YAML, Python or Jinja2.
	Cloud Console	Web interface to manage Google Cloud Platform resources.
	Cloud Shell	Browser-based shell command line access to manage Google Cloud Platform resources.
	Cloud Console Mobile App	Android and iOS application to manage Google Cloud Platform resources.
	Cloud APIs	APIs to programmatically access Google Cloud Platform resources
Identity & Security	Cloud Identity	Single sign-on (SSO) service based on SAML 2.0 and OpenID.
	Cloud IAM	Identity & Access Management (IAM) service for defining policies based on role-based access control.
	Cloud Identity-Aware Proxy	Service to control access to cloud applications running on Google Cloud Platform without using a VPN.
	Cloud Data Loss Prevention API	Service to automatically discover, classify, and redact sensitive data.
	Security Key Enforcement	Two-step verification service based on a security key.
	Cloud Key Management Service	Cloud-hosted key management service integrated with IAM and audit logging.
	Cloud Resource Manager	Service to manage resources by project, folder, and organization based on the hierarchy.
	Cloud Security Command Center	Security and data risk platform for data and services running in Google Cloud Platform.
	Cloud Security Scanner	Automated vulnerability scanning service for applications deployed in App Engine.
	Access Transparency	Real-time audit logs providing visibility to Google Cloud Platform administrators.
IoT	Cloud IoT Core	Secure device connection and management service for Internet of Things.
	Edge TPU	Purpose-built ASIC designed to run inference at the edge. As of September 2018, this product is in private beta.
	Cloud IoT Edge	Brings AI to the edge computing layer.
API Platform	Maps Platform	APIs for maps, routes, and places based on Google Maps.
	Apigee API Platform	Lifecycle management platform to design, secure, deploy, monitor, and scale APIs.
	API Monetization	Solution for API providers to create revenue models, reports, payment gateways, and developer portal integrations.
	Developer Portal	Self-service platform for developers to publish and manage APIs.

API Analytics	Service to analyze API-driven programs through monitoring, measuring, and managing APIs.
Apigee Sense	Enables API security by identifying and alerting administrators to suspicious API behaviors.
Cloud Endpoints	An NGINX-based proxy to deploy and manage APIs.

- [Common Web Service Architecture with High Availability in GCP](#) — Below architectural diagram is for common web service composed of web server and database with high availability.
- [Continuous Delivery with Spinnaker in GCP](#) — CI(Continuous Integration) and CD(Continuous Delivery) are one of the important technologies in terms of Agile Development's point of view.
- [Continuous Integration by Jenkins, automated solution, in GCP](#) — Jenkins is an open source automation server written in Java. Jenkins helps to automate the non-human part of the software development process, with continuous integration and facilitating technical aspects of continuous delivery. It is a server-based system that runs in servlet containers such as Apache Tomcat. Below architecture shows continuous integration based in GCP.
- [Scale testing with Kubernetes + Locust in GCP](#) — Locust is an open source tool for quickly scaling up and down load tests of web frontend as REST APIs. It allows testers to quickly create test profiles that can be orchestrated by multiple nodes in order to stress an HTTP endpoint.

GCP Reference Architectures: <https://gcp.solutions>

Programmatic Advertising Architecture

Traditionally programmatic advertising is also called as online advertising. Programmatic advertising includes real-time bidding as a part of its solution.

Programmatic architecture allows advertisers to deliver a personalized ad with the help of DMPs(Data Management Platforms) that store a user's information such as age, demographics, and location.

- Programmatic advertising is the automated buying and selling of online advertising
 - This automation makes transactions efficient and more effective, streamlining the process and consolidating your digital advertising efforts in one technology platform.
 - Targeting tactics are used to segment audiences using data so that advertisers only pay for ads delivered to the right people at the right time, and depend less on the "spray and pray" method of digital advertising

Gliffy Macro Error

You do not have permission to view this diagram.

Note that **RTB = Real-Time Bidding**

You may often in confusion to distinguish SSP and DSP. Basically those two things are designed for advertising. We can simply think **DSP is for advertiser** and **SSP is for application/web site owner**.

Supply Side Platform (SSP)

- SSP is for web/app owners
- A technology that enables web owners to manage their ad impressions inventory and maximize their income from digital media.
- SSP is used by web owners to maximize their revenue and offer ad impressions for sale to many ad exchanges and DSPs directly

Demand Side Platform (DSP)

- DSP is for advertisers
- A platform that allows advertisers to buy multiple ad spaces from different publishers.
- DSPs helps online marketers and advertising agencies to access the best ad spaces. This is through audience targeting and sorting the web content through their parameters.
- DSP makes the searching and analyzing of the audience segments (location, gender, age, income level, etc.) fast and effective cheap thanks to automation, which boost ROI.

External References

<https://www.iab.com/> - IAB(The Interactive Advertising Bureau) is one of the popular websites can get good information. Some people ignore / make a fool if you have no information about iab and its events as long as you work on advertising industry.

Technical Elements for Programmatic Advertising

- [Android AAdIDs, Apple IDFA, Safari & Firefox IDs](#) — There are various type of Identifiers widely used in AD Tech or Audience Measurement Industry - IDFA, AAdID, Hashed Device ID, Safari ID, Firefox ID and etc. Unfortunately these formats are not compatible each other except IDFA and AAdID. The difference between IDFA and AAdID is lowercase/uppercase in Latin characters.
- [ATD\(Agency Trading Desk\) Architecture](#) — ATD(Agency Trading Desk) is a kind of Advertiser solutions that uses DSP(Demand-Side Platform) as a tool. Trading desks are media buyers and re-sellers within an ad agency that help advertisers execute programmatic media buys.
- [DMP\(Data Management Platform\) Architecture](#) — DMP(Data Management Platform) is used by digital advertising buyers and publishers to store and manage audience data, often from multiple sources.
- [DSP\(Demand-Side Platform\) Architecture](#) — A Demand-Side Platform (DSP) is a real-time bidding system that connects media buyers with data exchanges and Supply-Side Platforms(SSPs) through a single interface.
- [False Positive Error](#) — A false positive error, or in short a false positive, commonly called a "false alarm", is a result that indicates a given condition exists, when it does not.
- [ITD\(Independent Trading Desk\) Architecture](#) — ITD(IndependentTrading Desk) is a kind of Advertiser solutions that uses DSP(Demand-Side Platform) as a tool.
- [SSP\(Supply-Side Platform\) Architecture](#) — SSP(Supply-Side Platform) is a kind of publisher solutions as a centralized platform to sell/manage inventory, access many advertisers, DSPs, etc.
- [VMAp for Video Ad Insertion](#) — Video Multiple Ad Playlists (VMAp) are used to return multiple video ads to be played during a video. It offers content owners control over the placement and timing of ads, especially when they lack control or ownership of the player.