

Google のインフラに見る基盤標準化と DevOps の真実

Etsuji Nakai
Cloud Solutions Architect at Google
2017/04/18 ver1.0



\$ who am i

- Etsuji Nakai

Cloud Solutions Architect at Google

Twitter @enakai00



Google における世界規模でのインフラ標準化

- 全世界のすべてのデータセンターで共通化されたインフラの提供
- OS レイヤーを隠蔽して、アプリケーションレベルでの管理に集中
- リソーススケジューラーによるアプリケーションデプロイの最適化
- アプリケーションのオートスケールとサービスの抽象化
- データストアとアプリケーションの分離

分散ソフトウェア技術で実現した
Datacenter as a Computer

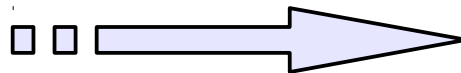




ユーザー企業における IT の変遷

ユーザー企業における IT の変遷

「専門家」に最適な技術を提供してもらう世界



ユーザー企業が必要な技術を自分で作り出す世界

メインフレーム

Unix 分散システム

Linux/OSS

オープンイノベーション

シングルベンダーによる
ワンストップサービス

マルチベンダー環境の
インテグレーション
(SI によるワンストップ
サービス)

標準技術の採用で
ベンダーロックインを排除

コモディティ技術を
ベースとした技術革新



The Apache
Software Foundation



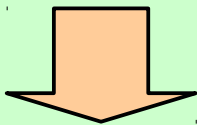
1990

2000

2010

オープンイノベーションとは？

「オープンイノベーション」



コモディティ技術＋組織を超えて生み出された知識・知見
を基礎とした技術革新

- コモディティ技術を基礎として、新しい「アーキテクチャー」「開発手法」「ハードウェア/ソフトウェア技術」を自ら作り出す。
 - ⇒ 「自分たちが必要な技術は、自分たちが一番良く知っている」
- 成果をオープンコミュニティで共有して、技術の進化を加速する。
 - ⇒ 「大切なのは技術そのものではなくて、それを活用する独自のノウハウ」

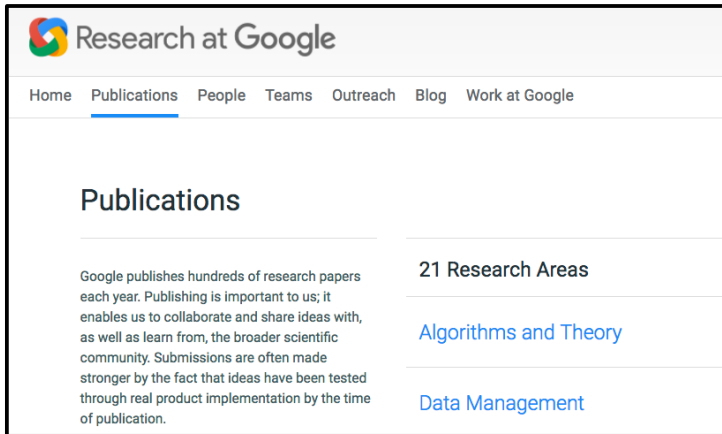
Google の技術が基礎となったオープンソースの例

- Hadoop (MapReduce を用いた分散データ処理基盤)
 - 2004 年に Google が公開した MapReduce の論文を元に、外部のエンジニアが独自にオープンソースとして再実装
- HBase (分散 Key-Value ストア)
 - 2006 年に Google が公開した Bigtable に関する論文を元に、外部のエンジニアが独自にオープンソースとして再実装
- Apache Beam (ストリーミング型分散データ処理ライブラリ)
 - Google が社内で独自開発していた MillWheel, FlumeJava の知見を元に、Google のエンジニアがオープンソースとして新規実装。Cloud Dataflow の処理基盤上で利用可能。
- Kubernetes (コンテナオーケストレーション・システム)
 - Google が社内で独自開発していた Borg, Omega の知見を元に、Google のエンジニアがオープンソースとして新規実装。
- TensorFlow (機械学習のソフトウェアライブラリ)

公開論文から読み解くインフラ技術の「思想」

- 「謎技術」の実体は、徹底的な合理主義
- 「技術的制約」に対する恐ろしいほどの洞察力

- この制約を受け入れることが何が可能になるのか？
- この制約を打破することで何が可能になるのか？



The screenshot shows the 'Research at Google' website. At the top is the Google logo and the text 'Research at Google'. Below this is a navigation bar with links: Home, Publications (highlighted), People, Teams, Outreach, Blog, and Work at Google. The main content area is titled 'Publications' and contains a paragraph about Google's research publishing efforts. To the right, there is a section for '21 Research Areas' with links for 'Algorithms and Theory' and 'Data Management'.

Research at Google

Home Publications People Teams Outreach Blog Work at Google

Publications

Google publishes hundreds of research papers each year. Publishing is important to us; it enables us to collaborate and share ideas with, as well as learn from, the broader scientific community. Submissions are often made stronger by the fact that ideas have been tested through real product implementation by the time of publication.

21 Research Areas

Algorithms and Theory

Data Management

<https://research.google.com/pubs/papers.html>



The screenshot shows a blog post from the website 'http://www.school.ctc-g.co.jp/columns/nakai2/'. The post is dated '2016年12月' and features a link to '第8回 スケーラビリティと一貫性を両立した分散データストアMegastore(パート2) New!'. The text below the link mentions a previous article from 2011. A second entry for '2016年12月' links to '第7回 スケーラビリティと一貫性を両立した分散データストアMegastore(パート1) New!', also mentioning the 2011 article. A third entry for '2016年11月' links to '第6回 大規模構造化データに最適化された分散キーバリューストアBigtable(パート2)', mentioning an article from 2006.

2016年12月

第8回 [スケーラビリティと一貫性を両立した分散データストアMegastore\(パート2\) New!](#)

はじめに 前回に続いて、2011年に公開された論文「Megastore: Providing Scalable, Highly Available Stora...

2016年12月

第7回 [スケーラビリティと一貫性を両立した分散データストアMegastore\(パート1\) New!](#)

はじめに 今回からは、2011年に公開された論文「Megastore: Providing Scalable, Highly Available Stora...

2016年11月

第6回 [大規模構造化データに最適化された分散キーバリューストアBigtable\(パート2\)](#)

はじめに 前回に続いて、2006年に公開された論文「Bigtable: A Distributed Storage System for Structured...

<http://www.school.ctc-g.co.jp/columns/nakai2/>



DevOps の視点で考える標準化

そもそも DevOps って何でしたっけ？

- 開発チームと運用チームと一緒に会議すること？
- 開発チームが運用までやっちゃうこと？
- 運用チームがコードを書いて開発すること？

DevOpsとは

2009年にオライリーが開催した「Velocity 2009」というイベントにおいて、Flickrのエンジニアが、「開発と運用が協力することで、1日に10回以上のペースでリリースが可能になること」を紹介しました。いまさまざまなシーンで見かける「DevOps」という言葉は、このプレゼンの中で登場したものです。

DevOpsとは、開発（Development）と運用（Operations）が協力し、ビジネス要求に対して、より柔軟に、スピーディに対応できるシステムを作り上げるためのプラクティスです。多くの人々により議論は続けられていますが、ITILとは異なり、現時点においては、DevOpsに厳密な定義はありません。

<http://www.atmarkit.co.jp/ait/articles/1307/02/news002.html>



DevOps

DevOps（デブオプス^[1]）は、ソフトウェア開発手法の一つ。開発（Development）と運用（Operations）を組み合わせたかばん語であり、開発担当者と運用担当者が連携して協力する開発手法をさす。^[1] また、DevOpsにビジネス部門を追加したBizDevOpsというワードも広がつつある。ただし2013年現時点では厳密な定義は存在しておらず、抽象的な概念に留まっている。^[2]

DevOpsの実現には構成管理ツールであるChef（英語版）やPuppetが用いられることもあり^[3]、DevOpsという語句からこうしたツールがイメージされることもある。^[2]

<https://ja.wikipedia.org/wiki/DevOps>

ITproまとめ



DevOps

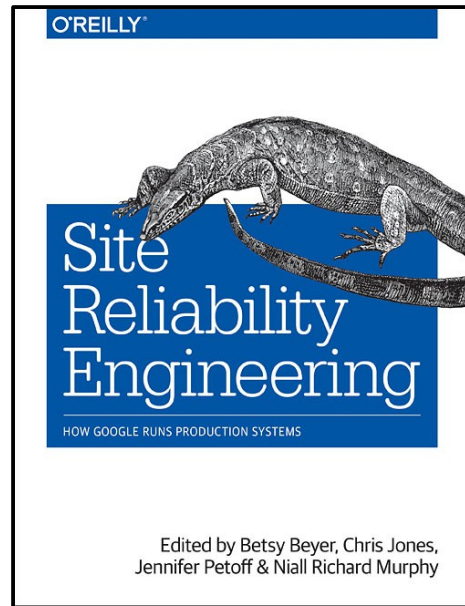
2013/11/15
平田 昌信=ITpro

開発担当者（Dev）と運用担当者（Ops）が連携して協力し合う開発手法のこと。開発担当者と運用担当者の間にあった壁を取り払ってお互いに連携・協力することで、システムの素早いリリースを実現する。「Chef」や「Puppet」など、DevOpsを支援するツールも多数登場している。

<http://itpro.nikkeibp.co.jp/article/COLUMN/20131113/517746/>

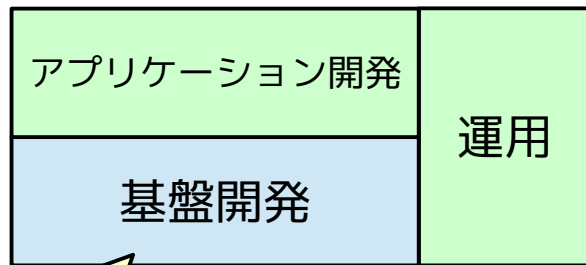
Site Reliability Engineer

- Google の運用チームの名称
 - 開発者と同じスキルセット＋インフラの知識
 - 運用作業 + 運用効率を改善するためのコード開発
 - 運用作業は、業務時間の 50% 以下に制限



理想の DevOps を実現するための隠された視点

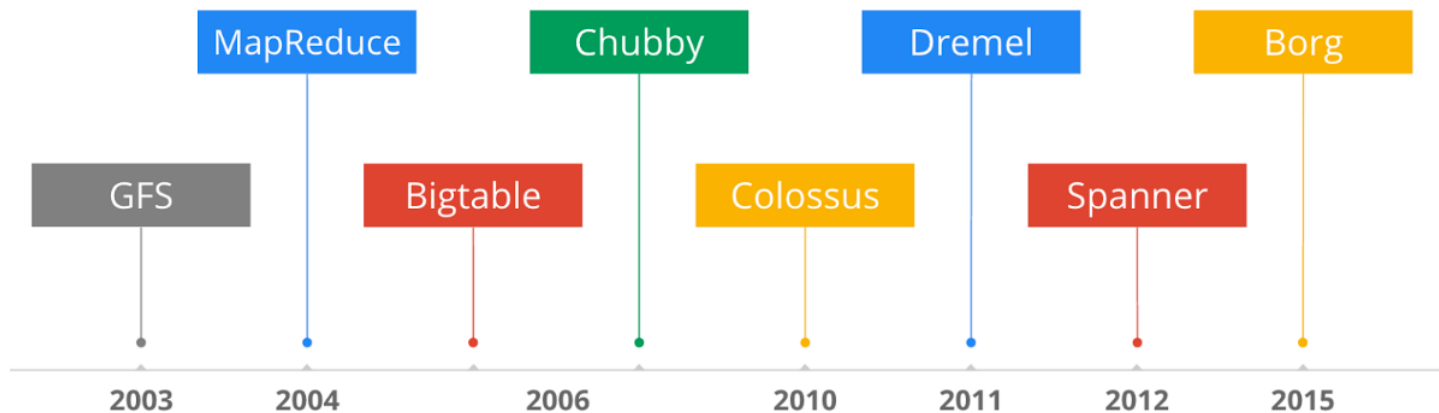
- レイヤーごとの責任分界点を明確にすることで、「本質的でない依存関係」をなくして、全体最適化を実現
 - ・ 無駄な依存関係がないからこそ、インフラ・開発・運用の3チームが健全な協力関係を確立可能に
- その上で「真に重要な依存関係」に叡智を結集
 - ・ スケーラブルで運用効率性の高いアプリケーションに必要なインフラ技術の提供
 - ・ 運用段階での効率性や安定性、スケーラビリティの確保を前提としたインフラ/アプリケーションの設計



アプリケーション開発に真に必要な
コンポーネントを厳選して開発・提供

Google が開発した分散ソフトウェア技術の例

- 全世界のデータセンターで共通化されたインフラの提供
- スケーラブルで運用効率性の高いアプリケーションを実現する機能を提供
- インフラを隠蔽して、アプリケーションレベルでの開発／管理に集中



Google のインフラを一般開放した Google Cloud Platform



STACKDRIVER



IDENTITY AND
ACCESS
MANAGEMENT



COMPUTE ENGINE



PREEMPTIBLE VMS



CUSTOM MACHINE TYPES



APP ENGINE



CONTAINER ENGINE



CLOUD STORAGE



NEARLINE



CLOUD SQL



DATASTORE



BIGTABLE



VIRTUAL NETWORK



LOAD BALANCING



CDN



DNS



INTERCONNECT



BIGQUERY



DATAFLOW



DATAPROC



DATALAB



PUB/SUB



CLOUD ML



SPEECH API



VISION API



TRANSLATE API



NATURAL
LANGUAGE API

Management

Compute

Storage

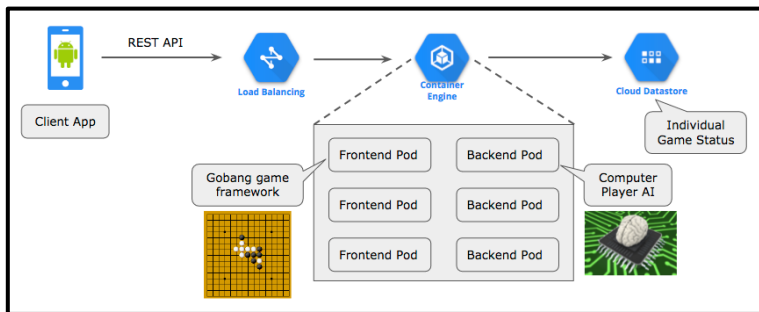
Networking

Data

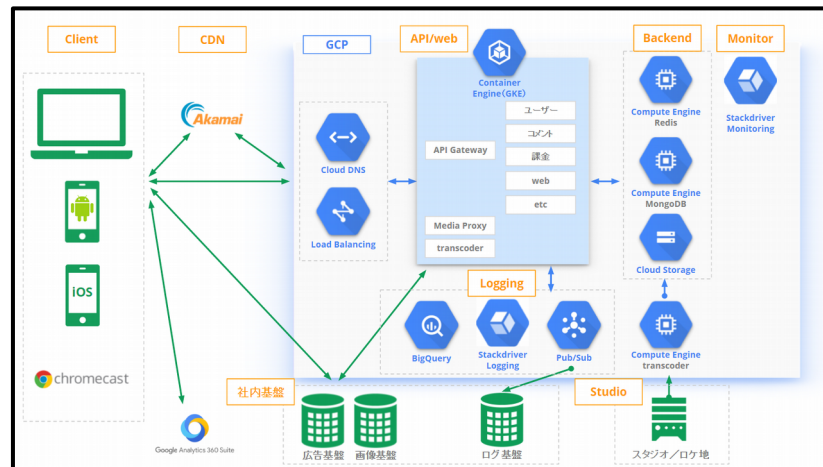
Machine
Learning

技術者に向けたメッセージ

- インフラを構成するソフトウェアの特性を深く理解して、最適なアプリケーション・アーキテクチャを見極めることが重要
- それぞれの技術要素を根本から理解して、システム全体のアーキテクチャを俯瞰できる能力が必要
 - 重要なのは、役割ではなく、知識範囲としてのフルスタック

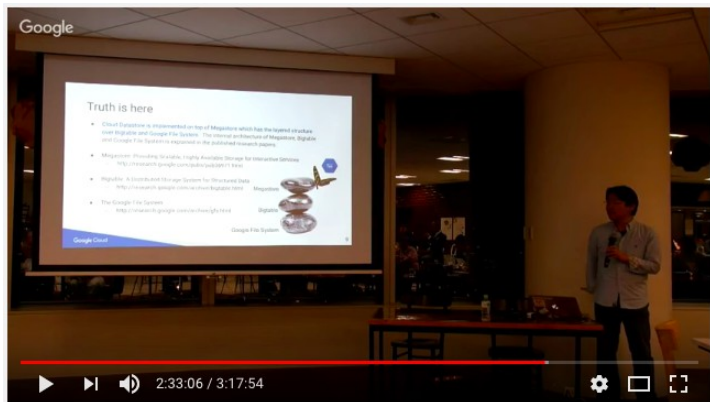


<https://github.com/GoogleCloudPlatform/gke-gobang-app-example>



<http://www.slideshare.net/strsk/google-container-engine-kubernetes>

技術を根本から理解して使いこなす喜び



<https://www.youtube.com/watch?v=H-tZUZGBo60&start=8540>



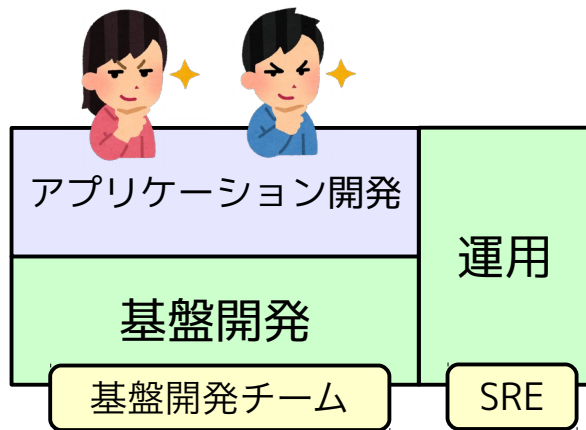
<http://togetter.com/li/1046340?page=6>



まとめ

理想の DevOps を実現するための隠された視点

- レイヤーごとの責任分界点を明確にすることで、「本質的でない依存関係」をなくして、全体最適化を実現
 - ・ 無駄な依存関係がないからこそ、インフラ・開発・運用の3チームが健全な協力関係を確立可能に
- その上で「真に重要な依存関係」に叡智を結集
 - ・ スケーラブルで運用効率性の高いアプリケーションに必要なインフラ技術の提供
 - ・ 運用段階での効率性や安定性、スケーラビリティの確保を前提としたインフラ/アプリケーションの設計



決して「謎技術」ではありません！

Google のインフラを一般開放した Google Cloud Platform



STACKDRIVER



IDENTITY AND
ACCESS
MANAGEMENT



COMPUTE ENGINE



PREEMPTIBLE VMS



CUSTOM MACHINE TYPES



APP ENGINE



CONTAINER ENGINE



CLOUD STORAGE



NEARLINE



CLOUD SQL



DATASTORE



BIGTABLE



VIRTUAL NETWORK



LOAD BALANCING



CDN



DNS



INTERCONNECT



BIGQUERY



DATAFLOW



DATAPROC



DATALAB



PUB/SUB



CLOUD ML



SPEECH API



VISION API



TRANSLATE API



NATURAL
LANGUAGE API

Management

Compute

Storage

Networking

Data

Machine
Learning

Google Cloud Platform ユーザー事例（一部）

- 株式会社テレビ朝日の導入事例：ニュースポータル「favclip」の先進的な開発環境を Google App Engine が支える
 - <https://cloudplatform-jp.googleblog.com/2017/01/tv-asahi-favclip-google-app-engine.html>
- AbemaTV での GKE 運用事例のご紹介
 - <http://ameblo.jp/principia-ca/entry-12164657818.html>
- 株式会社ディー・エヌ・エーの導入事例：エンジニアを開発に集中させるために Google App Engine を起用！
 - <https://cloudplatform-jp.googleblog.com/2016/12/google-app-engine.html>
- 株式会社メルカリの導入事例：先端技術を手軽に活用できる Google Cloud Platform はベストな選択肢
 - https://cloudplatform-jp.googleblog.com/2016/10/google-cloud-platform_28.html
- 株式会社 gloops の導入事例：ログ解析を Google BigQuery に移行してコストを 1/100 に。
 - <https://cloudplatform-jp.googleblog.com/2016/08/gloops-google-bigquery-1100.html>



Google Cloud Platform

Thank you!