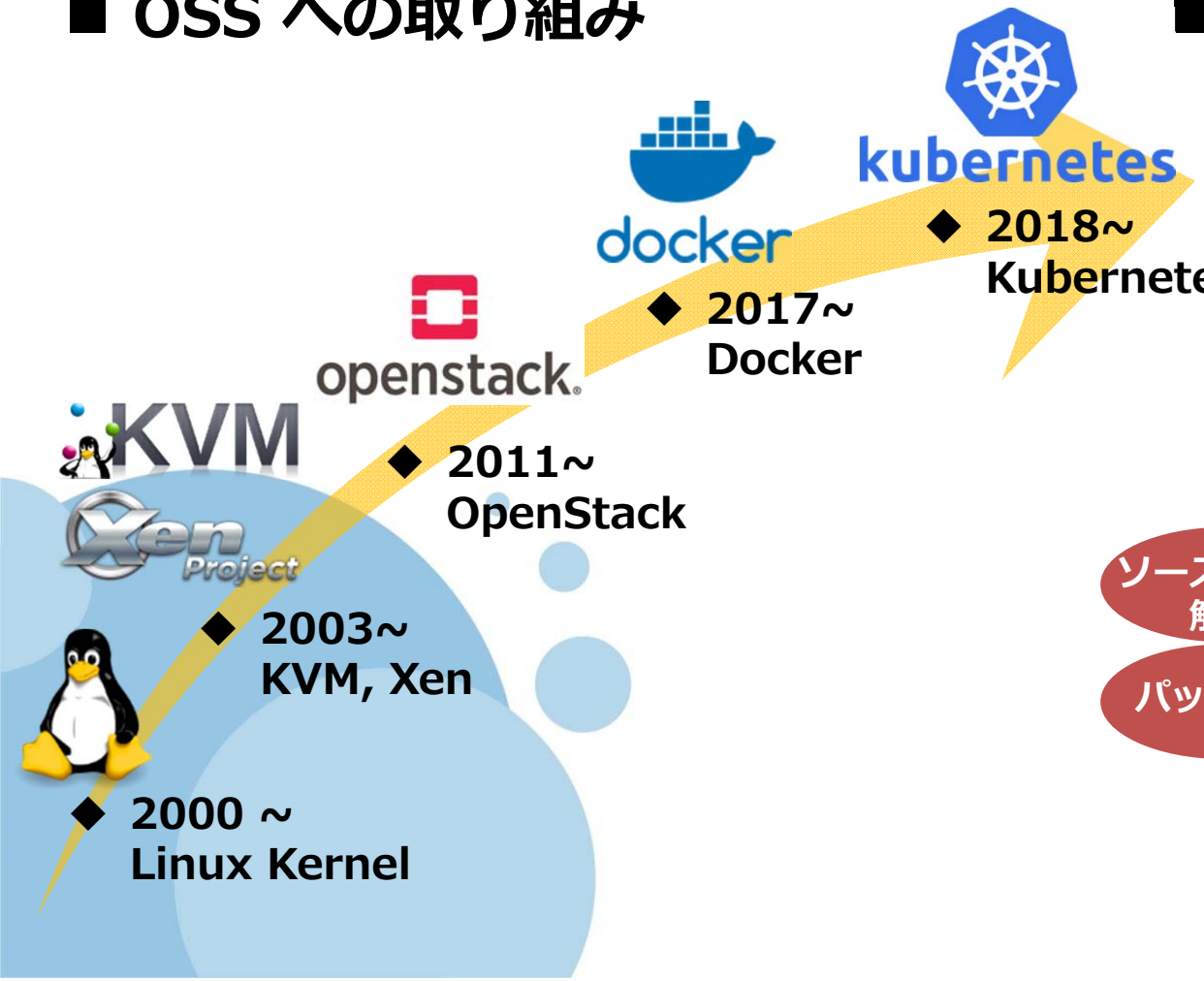
OpenStack Summit Vancouver や KubeCon EU 2018 で おもしろかったもの

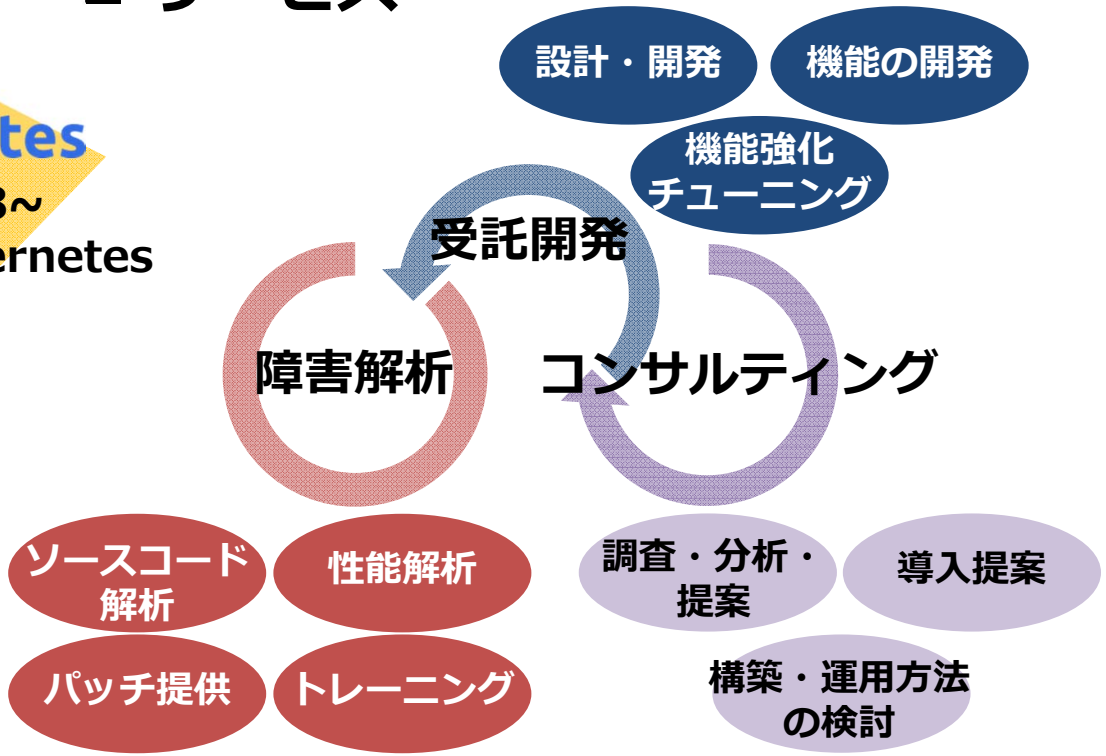
2018年8月3日
技術部
クラウド基盤エキスパート
工学博士　岩本 俊弘

VA Linux Systems Japan について

■ OSS への取り組み



■ サービス

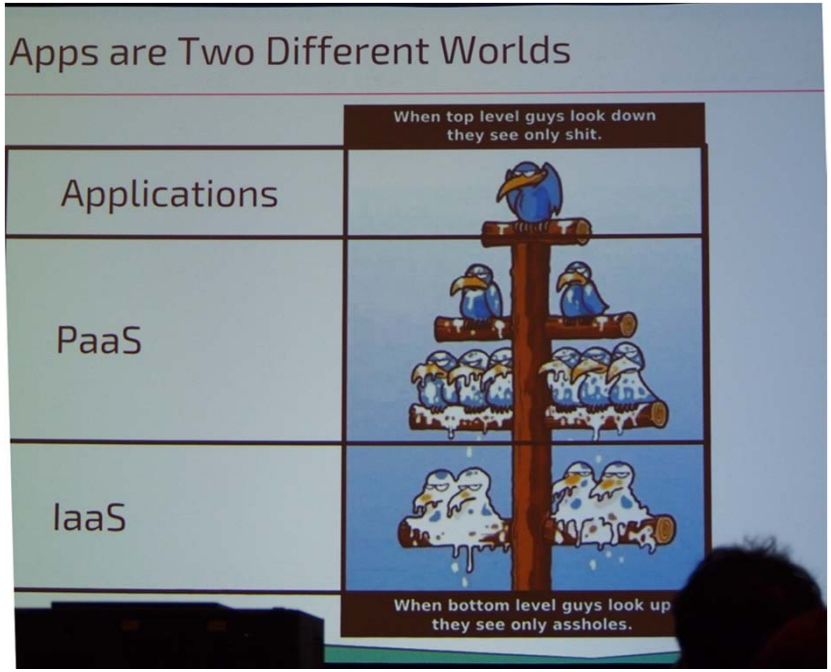


OpenStack Summit @Vancouver

- 3年前と同会場
  - 規模は半分くらい。おそらく3,000人程度の参加者
- Open Infra がキーワード
- Edge 流行らせようとしていた

OpenDev CI/CD

- OpenStack 用にがんばって作った Zuul を他でも流行らせたい
- 他に Spinnaker など
- OpenStack Summit の参加費を払わないでも参加できる



Standalone Cinder

- Nova や Keystone なしで Cinder だけ動かす
  - RabbitMQ とか DB も(?)
- Docker や Kubernetes 環境で Cinder Driver を使いたいという動機
  - (参考) CSI プラグイン
- Attachment をどうするかといった話をしていた
- その後の IRC meeting では cinderlib があるよといった話になっていた
  - cinderlib-csi というのもある
  - 先は長そう

Pre-emptible Instances

- Spot インスタンス (AWS), Preemptible VM (GCP) 相当
- CERN のプライベートクラウドの利用率を上げたい
  - Quota として予約されてるけど使われてない資源など
- 論点
  - VM にどうシャットダウンを通知するか
  - Quota の扱い preemptible かどうかで変えるかどうか
  - シャットダウンする VM をどう選ぶか (Reaper)
  - Reaper をいつ動かすか
- プロトタイプを作って議論中
  https://blueprints.launchpad.net/nova/+spec/preemptible-instances
  https://review.openstack.org/#/c/438640/

Kubernetes on Supporting $8 Trillion Card Payments in China

- CaiCloud について (caicloud.io, 杭州才云科技有限公司)
  - 中国でのコミュニティ活動や, TensorFlow, Kubeflow もやっている
- 中国唯一の Interbank Network が助けを求めてきた
  - 年間決済額 14.95兆USD (2017年)
- OpenStack/VM ベースの既存システムを “brown-field” と言っていた
- 素の Kubernetes にはない、以下の項目が必要
  - Multi-cluster/zones
  - Multi-tenancy
  - Multi-networking-plane

Kubernetes on Supporting $8 Trillion Card Payments in China

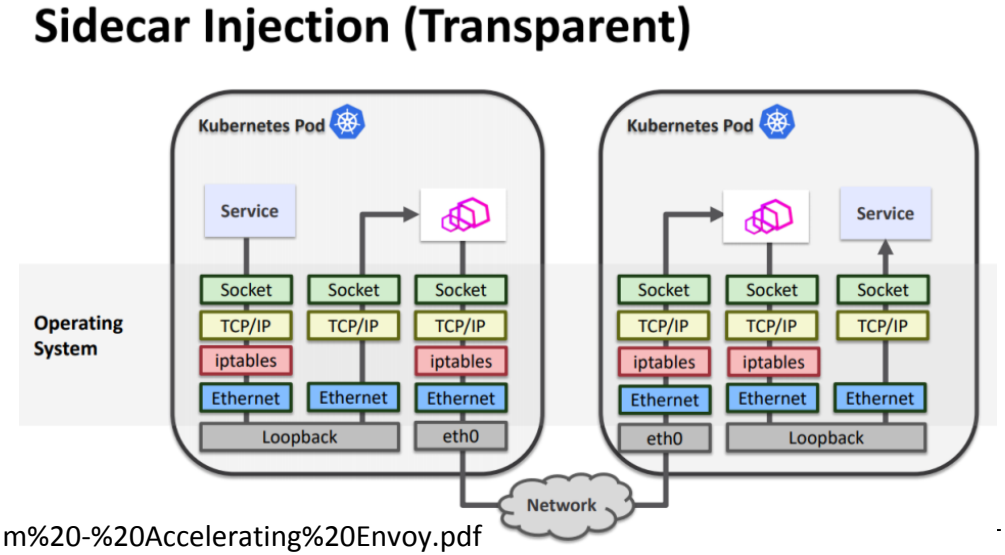
- **経緯**
  - 2015年から Docker を使用。2017年にアドミン負荷急増
  - 7月 kick-off, 3ヶ月で分析と設計、alpha 作りながら調整
  - 2月 staging, 3月から実運用
- **成果物**
  - SSO, マルチテナント, HA, ネットワーク帯域制御
  - NAS, swift 他と統合
- **派生物**
  - CI/CD の導入
  - マイクロサービス化のコンサル
- **教訓**
  - エンタープライズは大変
  - 連日の徹夜による burn-out に注意

Kubernetes on Supporting $8 Trillion Card Payments in China

- **SSO**
  - 一回ログインすれば Kubernetes でも OpenStack でも使える
- **マルチテナント**
  - OpenStack の Project, Network 等と対応
  - 各テナントに 2つの OpenStack Network が対応
- **ネットワーク**
  - Kuryr-kubernetes にマルチテナントと GBP を追加
  - CNI で Pod を Neutron Port と対応づけ
  - Multi-plane: management, control, storage, business
  - planeが違うので pod DNS は使えない
  - HostNetwork の DNS pod
    - ・DNSはテナント分離してない

Accelerating Envoy with the Linux Kernel

- Cilium (https://cilium.io)
- Linux eBPF 速い
  - L3/L4 LB @facebook
  - DDoS mitigation
  - カーネルにバイトコード送りつけて native にコンパイル
- Envoy を使うとこんな風になる



出典: https://schd.ws/hosted_files/kccnceu18/d9/2018%20KubeCon%20EU%20Cilium%20-%20Accelerating%20Envoy.pdf

Accelerating Envoy with the Linux Kernel

- 同一ホスト内の TCP/IP, Ethernet をバイパス
  - 性能 (req/s) が 2-3倍
  - TCP handshake はそのまま
  - データだけ透過的にバイパスする
- CNI plugin
- Cilium 現在 1.0.x, 今の話は 1.1/1.2 から
- kTLS で envoy に CA injection しないでよくなるとか言っている
  - kTLS は TLS の共通鍵暗号をカーネルでやる
- Github で Star を付けた人に抽選で StarWars のレゴをあげるとか言っている (5月4日)

ご清聴ありがとうございました。

VA Linux Systems Japan株式会社
https://www.valinux.co.jp/