HPCネットワークの多様化に挑む
マルチベンダー×マルチOSで支えるHPCネットワーク運用の実際

（事前資料）

さくらインターネット株式会社

黒澤　潔裕

当日のAgenda

はじめに

アーキテクチャの試行

ホワイトボックスへの挑戦

自動化への取り組み

まとめ

今日の話は、
こんな人に向けています

SAKURA internet

Closを「理想的な構成」だと思っている人

新たなベンダーへの挑戦はコスト削減策だと思っている人

自動化は“楽をするため”だと思っている人

さくらインターネットが挑戦した結果、皆さんがどう思うかを議論したい

構築済みHPCクラスタをマネージドサービスとして提供

- **DC基盤**
  - 電力（MW級）、冷却（空冷/液冷）、PUE最適化
- **計算リソース**
  - 計算用のGPU（B200 / H200 / H100）
- **ストレージ**
  - Lustre
- **ソフトウェア基盤**
  - Slurm / MPI / CUDA・ROCm / NCCL
- **ネットワーク**
  - Multi-Tenancy/Traffic/Lossless Ethernet/Cabling

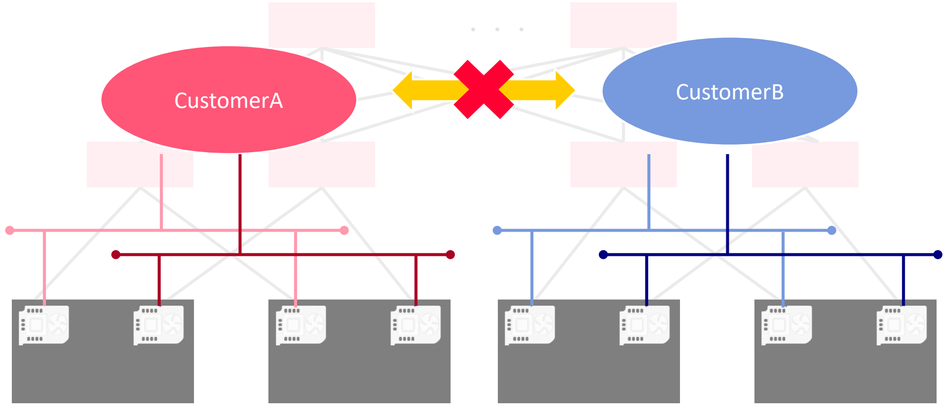
透明性の高い技術による実装を重視

SAKURA internet

GPUクラスタにおける、高帯域かつLosslessの実現

Multi-Tenancy

VRF
VLAN
EVPN/VXLAN

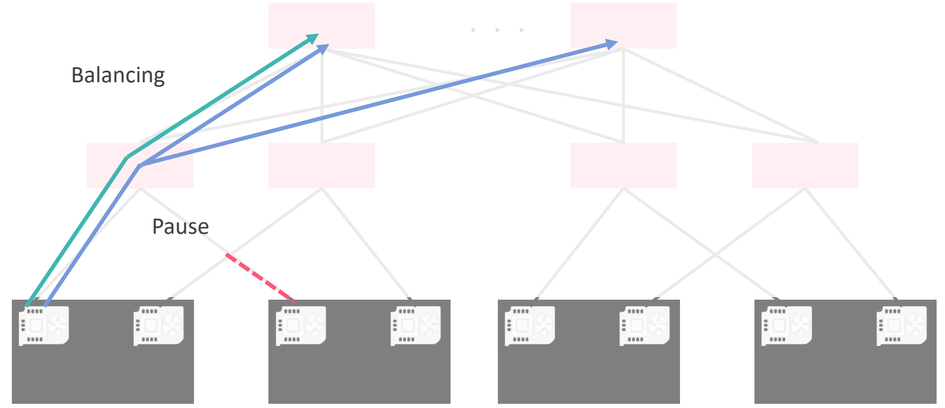


High Traffic

Switch間接続 800Gbps
GPU接続　　400Gbps

Lossless Ethernet

RoCEv2(ECN/PFC/CNP)
Dynamic Load Balancing



Cabling

高密度収容を実現するケーブル実装
Switch: 800GBASE-SR8
GPU:　400GBASE-DR4

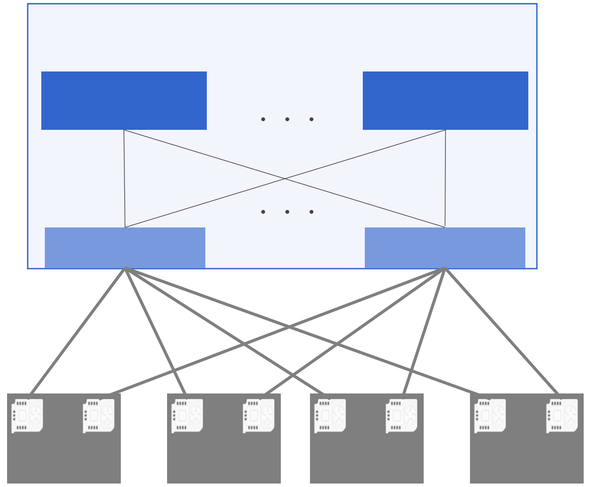
SAKURA internet

TopologyとNOSの二軸で新規設計を訴求

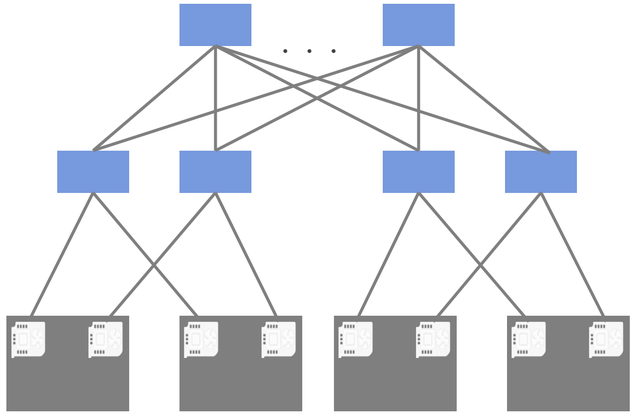
Architecture

収容上限突破を目的としたClos Topology検討

Chassis Switch



Clos Topology



Network OS

デリバリー速度と中立的な技術の採用

EOS ARISTA

SONiC Edge-corE NETWORKS

SAKURA internet

私たちの試行錯誤をベースに、広く下記のような議論を行いたい

アーキテクチャの選定基準

皆さんはどのようなTopologyを選択しますか？

GPU基盤におけるマルチベンダーの検討

皆さんはどのように「マルチベンダー」を実現していますか？

増大する運用コストへのアプローチ

どのような解決を考え、どのようなスキルセットが必要でしたか？

当日は皆さんの考えと共に広く議論ができればと思います

Appendix

Aristaシャーシを用いたGPUクラスタ設計

JANOG54にて発表。シャーシスイッチによる設計当時の登壇資料

JANOG54 meeting

生成AI向けパブリッククラウドサービスをつくってみた話

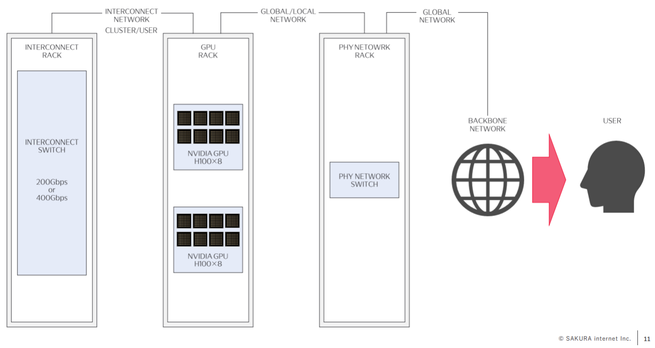
JANOG54 Meeting

生成AI向けパブリッククラウドサービス
をつくってみた話

2024年7月4日
さくらインターネット株式会社
井上 喬視
高峯 誠
平田 大祐

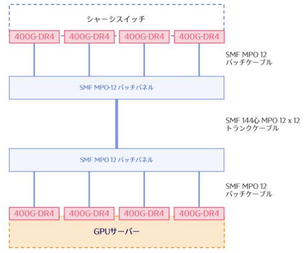
SAKURA internet

全体像



配線の設計

追加整備分では1サーバーあたり 400Gbs x 4 のネットワークに変更



配線のポイント

- SMF MPO-12配線に変更
  - 初回構築時より半年ほど経過した結果、サードパーティ品のOSFP-RHSがいろいろ選べるようになったできるようになった
  - 各種部材の納期が大幅に改善した

運用してみて…

- いまのところ問題なし
  配線もシンプルになり、400G-DR4も安定している

スイッチの運搬と設置

Arista 7816R3のラック設置について

ハードウェアのインストール方法は動画などもあり詳細にまとめられていた

Arista 7800R3 Series Modular Data Center Switches

https://www.arista.com/assets/data/pdf/Datasheets/7800R3-Quick-Look.pdf

ラックマウントについて

- ラック設置する際は推奨のツールがある
  ベンダー推奨のツールを探してみたが、国内では取り扱っている業者が無い様子（スペックダウンモデルであれば取扱い有り
- 自力でのラック搭載は不可と判断
  メーカー推奨の方法が取れないことから自社でのラックマウント作業は困難であると判断、専門の業者に依頼することに

配線の設計

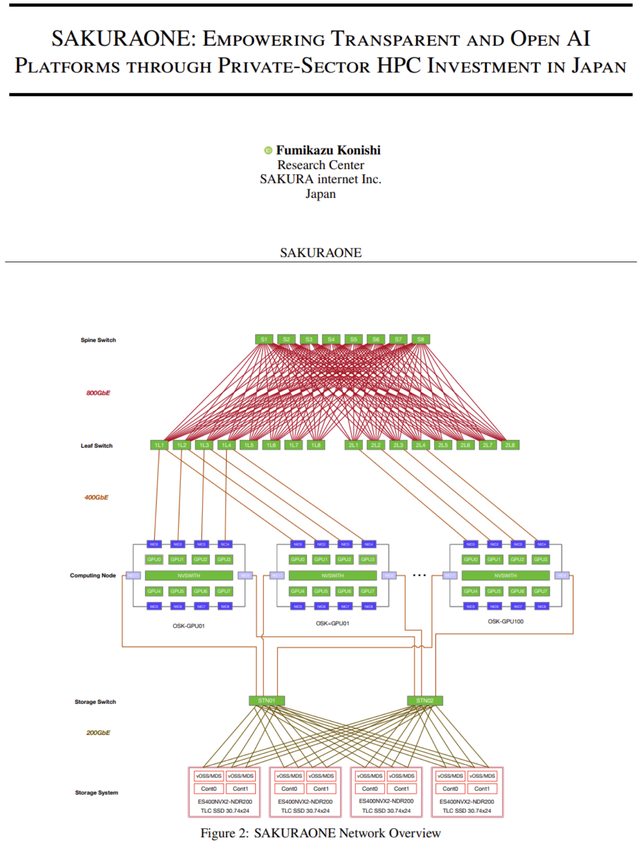
両隣のラックをパッチパネル専用ラックとして利用して配線する設計に

SONiCを用いたGPUクラスタの構成

SONiCにてClos Topologyを実装した際の記事（2025/6 ISC TOP500 49位にランクイン）

Arxiv

SAKURAONE: Empowering Transparent and Open AI Platforms through Private-Sector HPC Investment in Japan
Fumikazu Konishi



SAKURAONE: Empowering Transparent and Open AI Platforms through Private-Sector HPC Investment in Japan

Fumikazu Konishi
Research Center
SAKURA internet Inc.
Japan

SAKURAONE

Figure 2: SAKURAONE Network Overview

The Linux Foudation User stories

Open Networking at Scale: How SAKURA internet Deployed a TOP500 GPU Supercomputer with SONiC

SONiC SAKURA internet

USER STORY

Open Networking at Scale: How SAKURA internet Deployed a TOP500 GPU Supercomputer with SONiC

Organization

SAKURA internet Inc. is an internet company founded in 1996. Under the corporate philosophy of "Turning 'what you want to do' into 'what you can do,'" we develop a variety of services to meet customer needs and propose DX solutions that cater to various industries.

Since our founding, which began with the provision of shared server services, we have expanded to offer services such as "Koukaryoku" to support generative AI, and "SAKURA Cloud," which has been conditionally certified for use in government cloud systems. A key feature of our company is that we handle everything from development to operations in-house.

Overview

SAKURA internet is responding to the growing demand for computational infrastructure driven by the rapid adoption of generative AI by continuously procuring next-generation GPUs and strengthening reliable operational systems in our own data centers. As a digital infrastructure company contributing to the sustainable development of the digital society, our mission is to provide cloud services for generative AI.

To continuously meet the increasing demand, we recognized the necessity of resolving the following challenges:

- Ensuring vendor-neutral supply to mitigate risks
- Adopting technologies with high neutrality
- Accelerating delivery speed

Why SONiC?

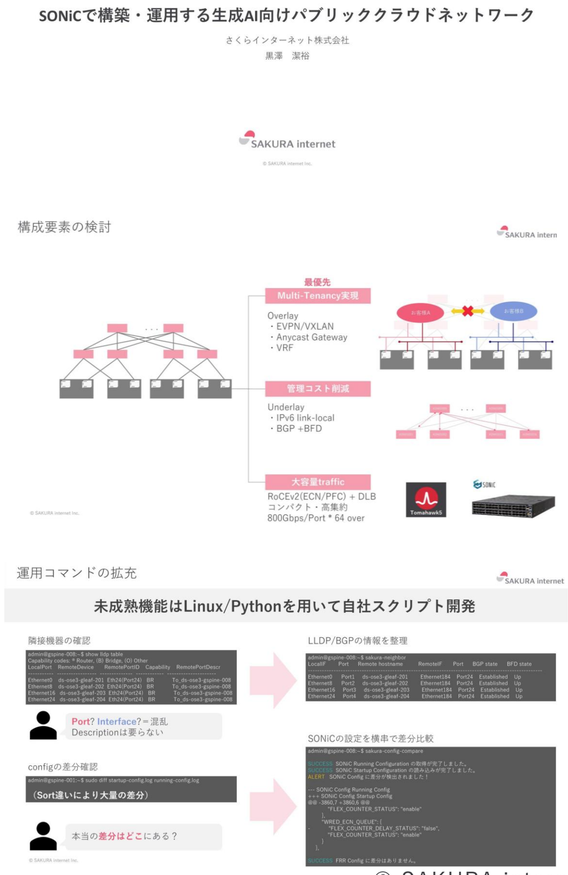
We selected SONiC for our new 800-GPU cluster because it directly addressed these needs. SONiC provided:

- High transparency, as it is implemented on a Debian/ Linux platform
- Active development of new features supported by the global community
- The ability to streamline operations by leveraging the same technologies used for Linux servers

In addition, SAKURA internet has a corporate culture of leveraging OSS and bottom-up initiatives. This culture supported our adoption of SONiC and enabled us to launch a GPU cloud infrastructure in a very short period of time.

SONiC Workshop Japan 2025

SONiCで構築・運用する生成AI向けパブリッククラウドネットワーク



SONiCで構築・運用する生成AI向けパブリッククラウドネットワーク

さくらインターネット株式会社
黒澤　潔裕

SAKURA internet

構成要素の検討

最優先
Multi-Tenancy実現
Overlay
・EVPN/VXLAN
・Anycast Gateway
・VRF
管理コスト削減
Underlay
・IPv6 link-local
・BGP +BFD
大容量traffic
RoCEv2(ECN/PFC) + DLB
コンパクト・高集約
800Gbps/Port * 64 over

運用コマンドの拡充

未成熟機能はLinux/Pythonを用いて自社スクリプト開発

隣接機器の確認
LLDP/BGPの情報を整理
Port? Interface? = 混乱
Descriptionは要らない
configの差分確認
（Sort違いにより大量の差分）
本当の差分はどこにある？
SONiCの設定を構造で差分比較