Em Busca do Estado da Arte do CGNAT

10/12/2024

Fernando Frediani

GTER 53

GTS 39

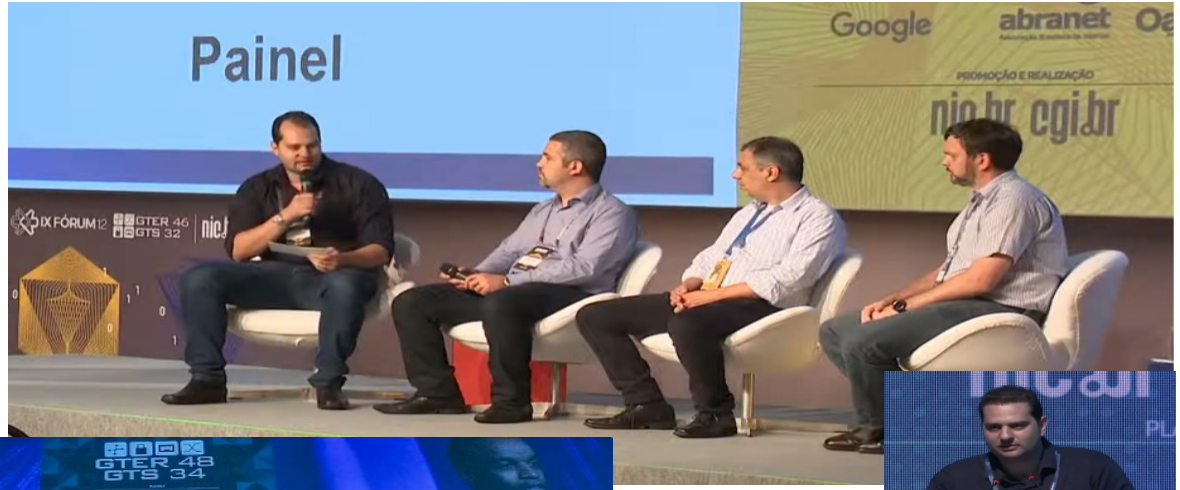
nic.br
Núcleo de Informação e Coordenação do Ponto BR

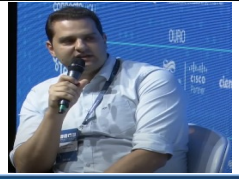
cgi.br
Comitê Gestor da Internet no Brasil

Introdução - Tópicos Abordados

- **Porque se importar com um CGNAT bem feito e organizado**
- **Cases**
- **Características importantes dos Equipamentos**
  - Arquitetura de hardware
  - Capacidade de Throughput
  - Bulk Port Allocation (BPA)
- **Quantidade de Portas por Cliente**
- **Armazenamento de Logs**
- **Alta disponibilidade**
- **Análise de Gráficos e Métricas**

10 anos de GTER e Semana de Infraestrutura





Porque se importar com o CGNAT

- **Já que é algo inevitável vamos fazer da melhor maneira**
  - Reduzir custos operacionais
  - Poder focar na implantação do IPv6

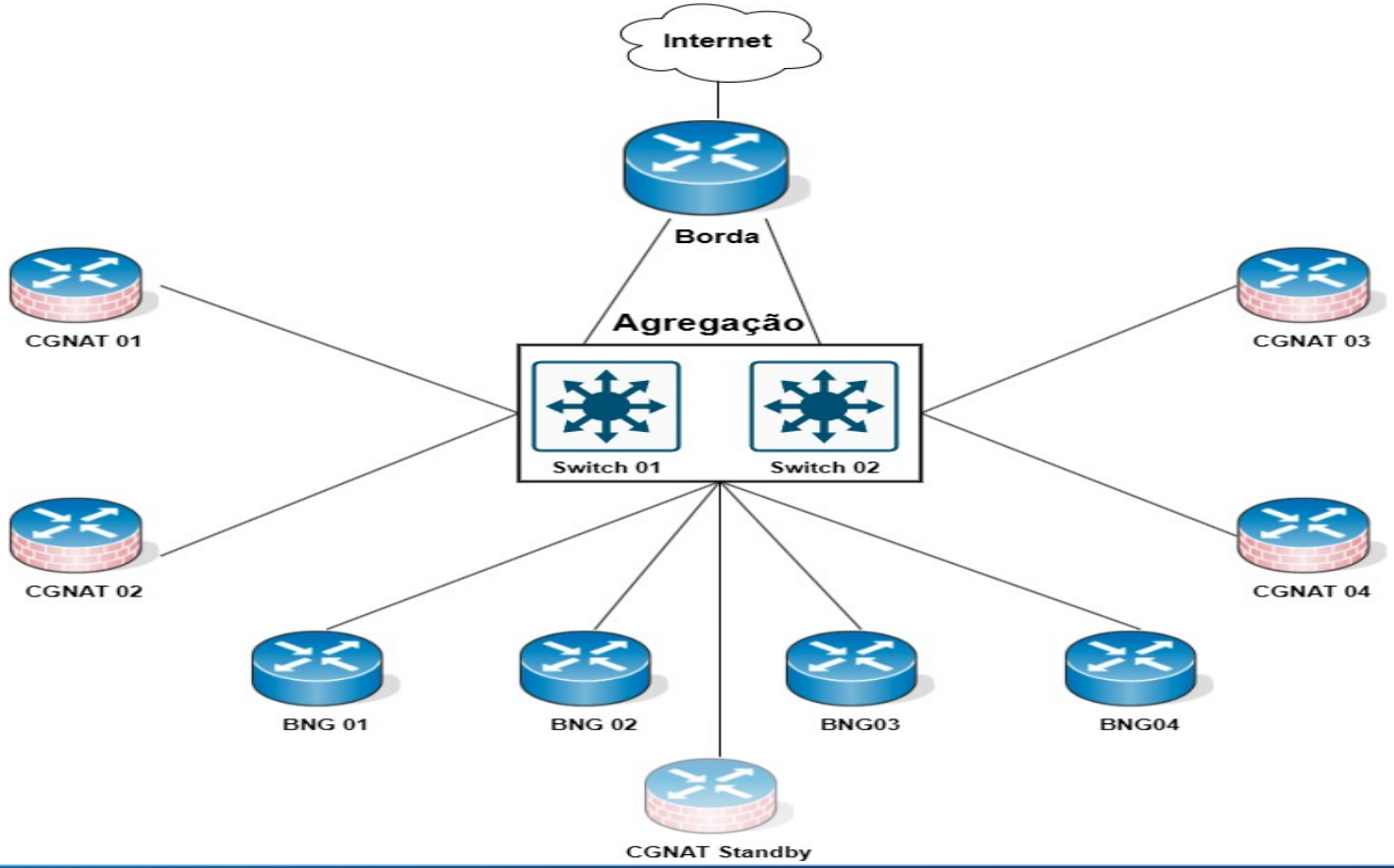
- **Fonte de problemas difíceis de diagnosticar**
  - Blacklists
  - Facilitar o diagnóstico

- **Reduzir a quantidade de IPv4 necessários**
- **Escolha das funcionalidades e equipamentos corretos que facilitem a operação**
- **Implantar de maneira escalável**

Cenário Inicial

- 5 equipamentos menores para CGNAT
- Utilização de uma quantidade razoável de IPv4 Públicos
- Maior dificuldade para gerir tráfegos passando por cada equipamento
- Maior complexidade com failover
- Alta geração de Logs e Archive problemático

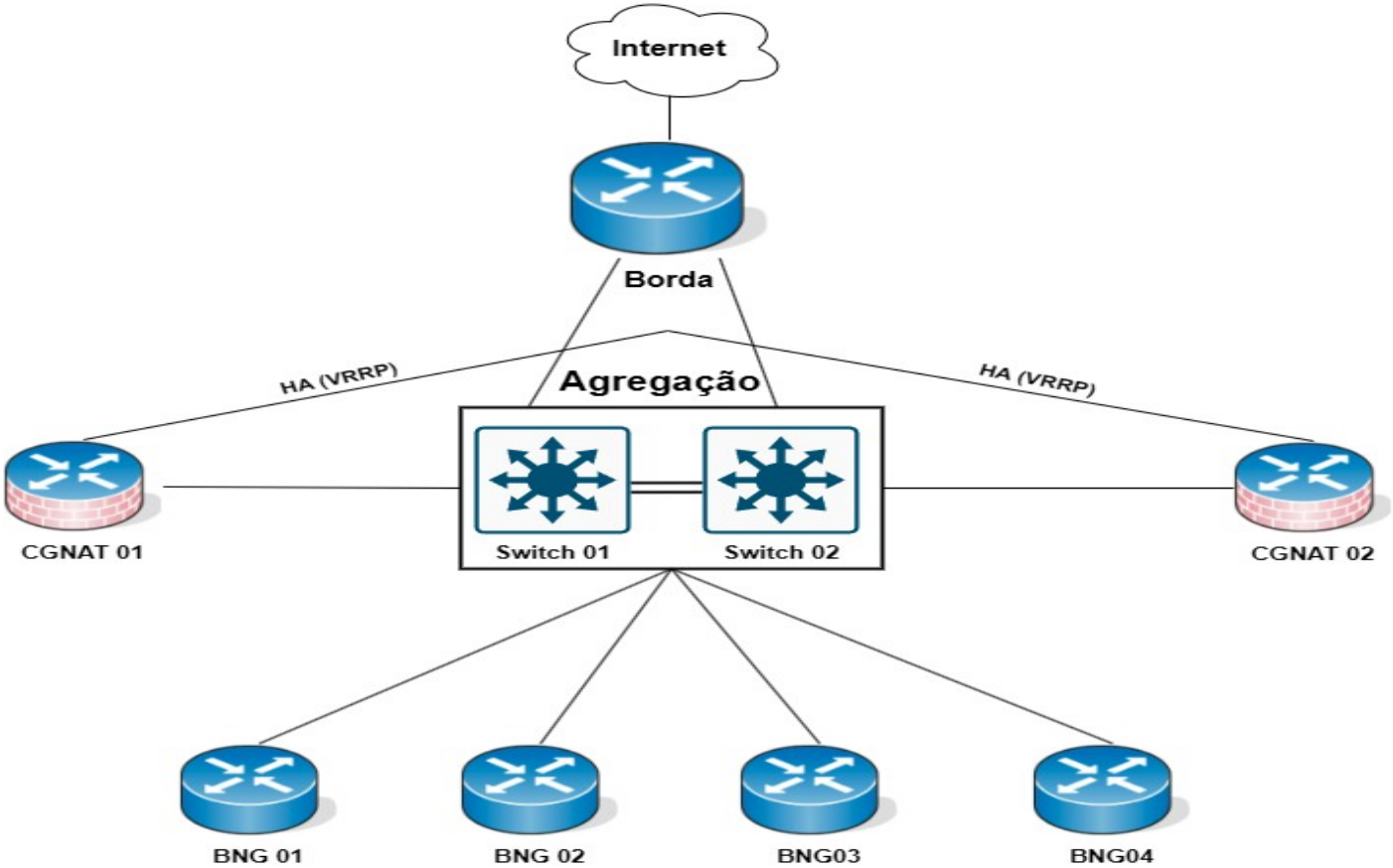
Cenário Inicial



Cenário após o Upgrade

- 2 equipamentos maiores
- Redução de +50% dos IPv4 em uso para atender a mesma base de clientes
- Equipamentos com capacidade suficiente para operarem sozinhos em cenário de Failover
- Geração de Logs com Bulk Port Allocation (pouco Log)

Cenário após o Upgrade



Características dos Equipamentos

- Possuem ASIC para diversas funcionalidades além do encaminhamento de tráfego
- Interfaces de Uplink não são limitadas por licença (suporte nominal de capacidade)
- Portas de 100Gb x 40Gb
- Suporte à Bulk Port Allocation, EIM/EIF, etc

Bulk Port Allocation x Determinístico

- **É um dos principais pontos para se atentar em cenários médios e grandes**

- **BPA**
- **Proporciona uma economia significativa de endereços IPv4 Públicos**
  - Alocação mais eficiente de portas de origem
- **Baixa geração de Log**
  - Simplifica o armazenamento
  - Facilita nas identificações de usuários

- **Determinístico**
- **Alocação fixa de portas independente do uso**
- **Maior necessidade de IPv4 para uma mesma base de usuários**
- **Maior complexidade para cenários de failover**

Quantidade de Portas por Cliente

- **Em um cenário Determinístico no mínimo 2000 portas por cliente/IP**
  - Quantidade fixa que pode ser desperdício ou problema se exauridas
  - Maior necessidade de IPv4 Públicos conforme a base de clientes

- **Com Bulk Port Allocation a alocação é dinâmica**
  - Alocação inicial de 512 portas por bloco para cada cliente/IP
  - Máximo de 8 blocos para cada cliente/IP (4000 portas)

- **Maioria dos clientes usam uma quantidade menor de portas**

Geração e Armazenamento de Logs

- **Só gera logs nos seguintes eventos:**
  - Alocação de um novo bloco (primeiro ou adicionais)
  - Atualizações periódicas de blocos alocados
  - Remoção de alocação devido à inatividade

- **Cliente permanece com alocação por várias horas ou dias**
- **Quantidade de logs muito baixa – menos que 1GB / dia**
- **Simplifica o armazenamento (archive)**
  - Menor espaço necessário e tempo de processamento

- **Facilita a busca de informações para identificações de usuários**

Geração e Armazenamento de Logs

- **3 eventos: ALLOC, INTERIM e RELEASE**

- Dec 5 08:30:20 CGNAT01 RT_NAT: **RT_SRC_NAT_PBA_ALLOC:** Subscriber **100.64.10.85** used/maximum **[1/8]** blocks, allocates port block **[47104-47615]** from **203.0.113.22** in source pool POOL_IPS_PUBLICOS lsys_id: 0 epoch 0x67523e99

- Dec 6 08:30:20 CGNAT01 RT_NAT: **RT_SRC_NAT_PBA_INTERIM:** Subscriber **100.64.10.85** used/maximum **[1/8]** blocks, interim port block **[47104-47615]** from **203.0.113.22** in source pool POOL_IPS_PUBLICOS lsys_id: 0 epoch 0x674e4b60

- Dec 6 21:00:02 CGNAT01 RT_NAT: **RT_SRC_NAT_PBA_RELEASE:** Subscriber **100.64.10.85** used/maximum **[0/8]** blocks, releases port block **[47104-47615]** from **203.0.113.22** in source pool POOL_IPS_PUBLICOS lsys_id: 0 epoch 0x67523d74

Geração e Armazenamento de Logs

- **ALLOC (multiplos blocos) e RELEASE**

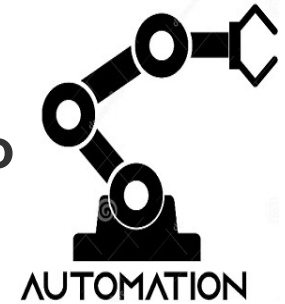
- Dec 5 09:32:00 CGNAT01 RT_NAT: **RT_SRC_NAT_PBA_ALLOC:** Subscriber **100.64.12.20** used/maximum **[1/8]** blocks, allocates port block **[20992-21503]** from **203.0.113.51** in source pool POOL_IPS_PUBLICOS lsys_id: 0 epoch 0x67523e9a
- Dec 5 11:15:02 CGNAT01 RT_NAT: **RT_SRC_NAT_PBA_ALLOC:** Subscriber **100.64.12.20** used/maximum **[2/8]** blocks, allocates port block **[3584-4095]** from **203.0.113.51** in source pool POOL_IPS_PUBLICOS lsys_id: 0 epoch 0x67523e9f

- Dec 5 15:40:10 CGNAT01 RT_NAT: **RT_SRC_NAT_PBA_RELEASE:** Subscriber **100.64.10.20** used/maximum **[1/8]** blocks, releases port block **[3584-4095]** from **203.0.113.51 in** source pool POOL_IPS_PUBLICOS lsys_id: 0 epoch 0x67523d74
- Dec 6 07:25:14 CGNAT01 RT_NAT: **RT_SRC_NAT_PBA_RELEASE:** Subscriber **100.64.10.20** used/maximum **[0/8]** blocks, releases port block **[20992-21503]** from **203.0.113.51** in source pool POOL_IPS_PUBLICOS lsys_id: 0 epoch 0x67523d7c

Script/Automação para Identificação de Usuários

- **Logs gerados em diretórios organizados de maneira hierárquica por Ano, Mês e Dia**

- **Foi elaborado script que automatiza toda a identificação**



  - Realiza parsing no arquivo do dia/mês/ano respectivo
  - Encontra a informação do IP de CGNAT (através do IP Público + Porta de Origem)
  - Busca informação na base de Accounting do Radius (radacct)
  - Consulta o ERP da empresa através de API e gera as informações no formato para responder para o jurídico

ATA de CGNAT

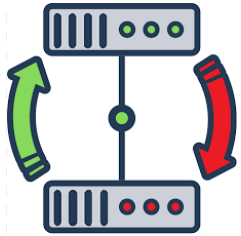
- **Elaborar uma ATA com todos os IPv4 Públicos e Privados utilizados em cada equipamento**
  - Importante para garantir um processo de identificação sem falhas

- **Anotar as mudanças temporárias na relação IPv4 Público x Privado em cenários de manutenção e failover**
  - Failover do tráfego de um equipamento de CGNAT para os outros
  - Alocação de novos blocos para equipamento de CGNAT
  - Migração de blocos de IPs Públicos entre equipamentos

- **Criar a prática de manter a ATA atualizada e consultar quando necessário para identificação de usuários**

Alta Disponibilidade (VRRP)

- **Cada equipamento possui capacidade para assumir toda a demanda do outro sozinho caso necessário**
  - Throughput e Quantidade de Endereços IPv4



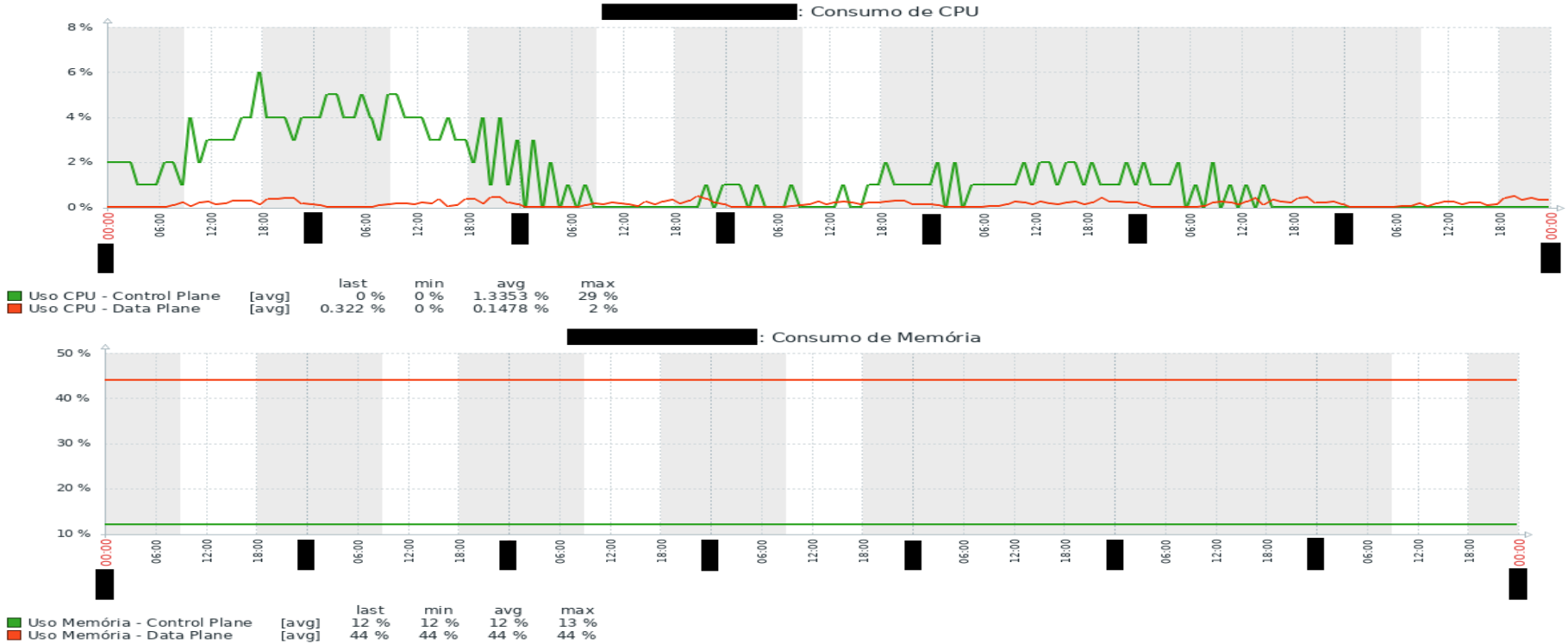
- **Utilizado VRRP entre os equipamentos e balanceamento entre os BNGs através dos VIPs**
  - BNG01 e BNG02 → CGNAT01
  - BNG03 e BNG04 → CGNAT02

- **Failover imperceptível para os usuários**

- **Não existência de limite de tráfego por licença permite acomodar cenários de failover mais facilmente**

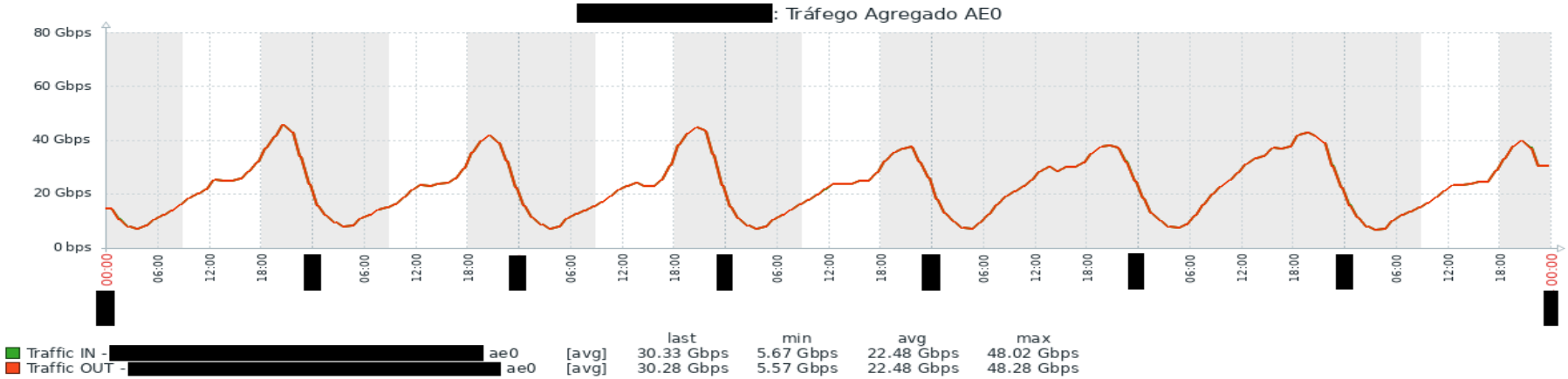
Gráficos e Métricas

- CPU e Memória
  - Diferença no que é processado no ASIC e na CPU



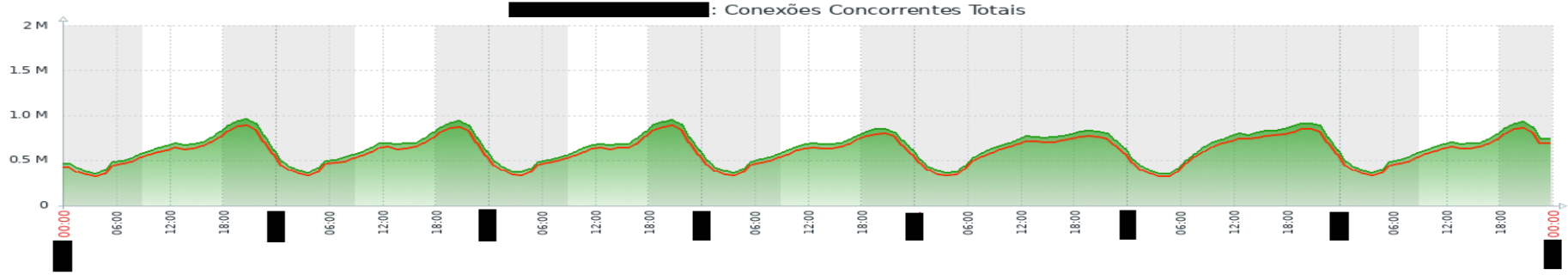
Gráficos e Métricas

- Gráficos de Throughput



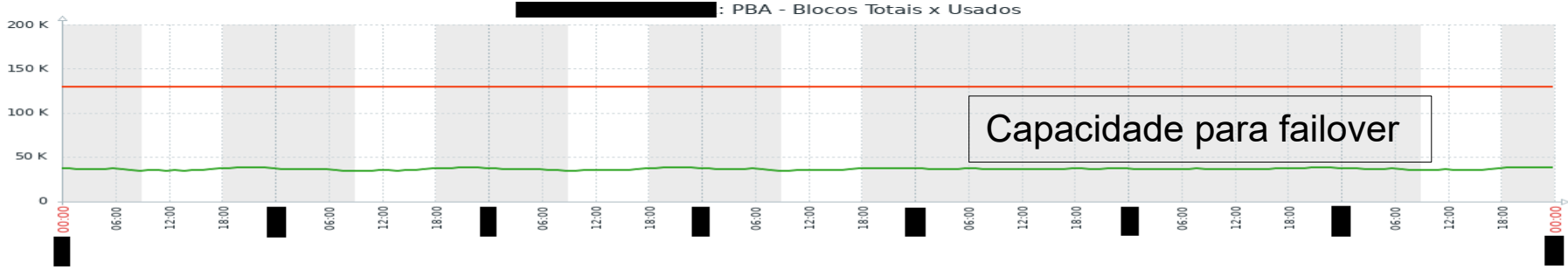
Gráficos e Métricas

Quantidade de Conexões Concorrentes



		last	min	avg	max
Conexões Correntes Totais	[avg]	729.06 K	330.18 K	634.43 K	976.78 K
Conexões Concorrentes ASIC	[avg]	687.76 K	309.94 K	593.98 K	926.72 K
Portas Públicas Alocadas para CGNAT	[avg]	66.06 M	66.06 M	66.06 M	66.06 M

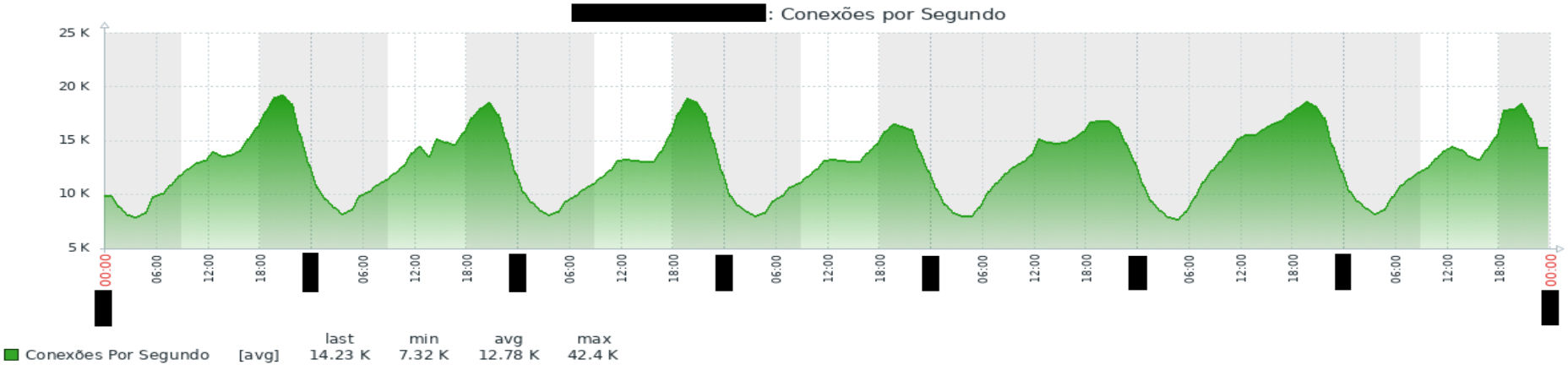
Quantidade de Blocos Alocados



		last	min	avg	max
PBA - Blocos Usados	[avg]	37.74 K	31.98 K	35.94 K	39.27 K
PBA – Blocos Totais	[avg]	129.02 K	129.02 K	129.02 K	129.02 K

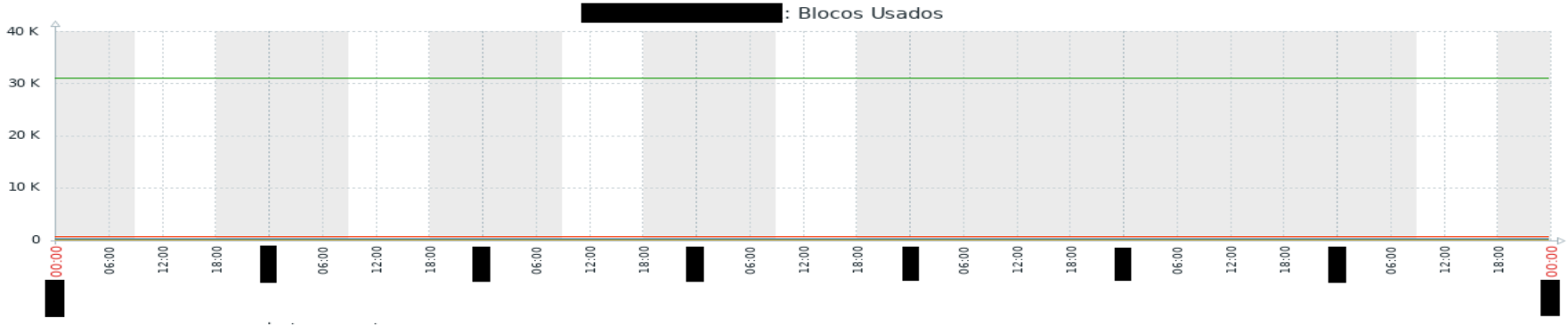
Gráficos e Métricas

- ## Conexões por Segundo



Gráficos e Métricas

- **Quantidade de Blocos Usados**



		last	min	avg	max
Qtd Bloco Usados - 1	[avg]	31.14 K	31.14 K	31.14 K	31.14 K
Qtd Bloco Usados - 2	[avg]	684	684	684	684
Qtd Bloco Usados - 3	[avg]	122	122	122	122
Qtd Bloco Usados - 4	[avg]	48	48	48	48
Qtd Bloco Usados - 5	[avg]	23	23	23	23
Qtd Bloco Usados - 6	[avg]	23	23	23	23
Qtd Bloco Usados - 7	[avg]	24	24	24	24
Qtd Bloco Usados - 8	[avg]	64	64	64	64

Conclusões e Recomendações

- É possível ter um CGNAT bem feito sem precisar de um número grande de endereços IPv4
- É possível reduzir a quantidade de IPv4 necessários para atender bem e com qualidade uma mesma base de clientes
- Com a tecnologia do BPA é possível ter uma maior concentração de usuários por IPv4 de maneira segura
- Construir de maneira que facilite escalar e para cenários de failover
- Não ignorar a importância do IPv6 para o sucesso do CGNAT
- Monitorar / Possuir métricas sobre a saúde do CGNAT

Perguntas ?

Contato: fhfrediani@gmail.com

Obrigado

Contato: fhfrediani@gmail.com