Avaliação de arquiteturas de soluções de backup para a nuvem: cloud connector e cloud gateway

15/05/2018 (v1.1)

Kleber Silva

CMG imPACt 2018 – 13º CMG Nacional

(Apresentação atualizada do CMG Nacional ano 2014)

Apresentação

- ✓ Membro palestrante CMG desde 2011
- ✓ 19 anos na área de Infraestrutura de TI: Storage, Backup, Virtualização, Hiperconvergência e Nuvem
- ✓ Mestre em Engenharia da Computação pelo IPT (Instituto de Pesquisas Tecnológicas da USP)
- ✓ Professor de Graduação e Pós-graduação: Estácio, Sumaré, FECAF, SENAC, FEI
- ✓ Membro SNIA Brasil desde 2014
- ✓ Gerente de pré-vendas e sócio: Pise4 Tecnologia
- ✓ https://www.linkedin.com/in/kleberjs/

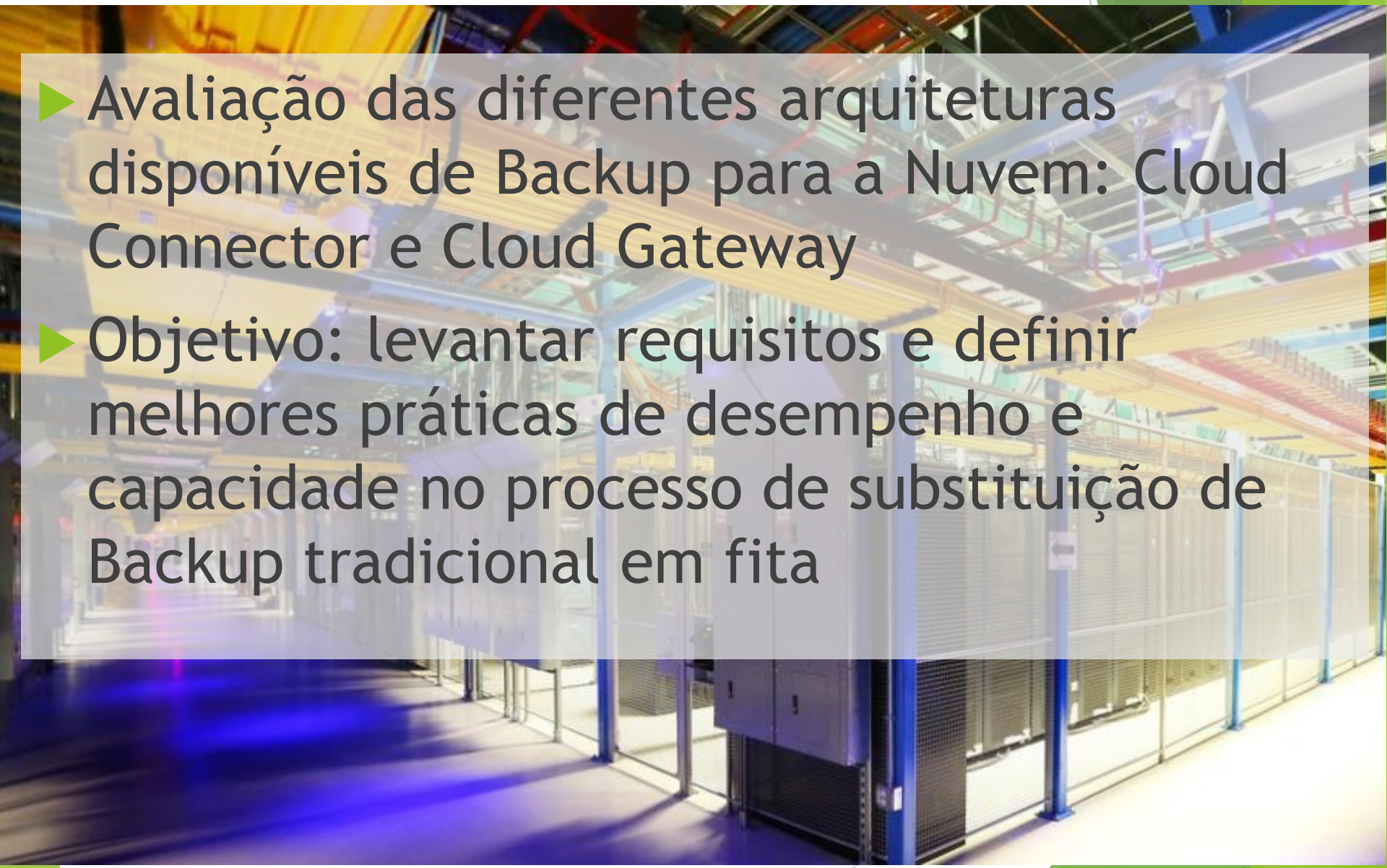
Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

Agenda

1. INTRODUÇÃO
   - 1.1 ESCOPO
2. TRABALHOS RELACIONADOS
3. MOTIVAÇÃO
4. BACKUP PARA A NUVEM
   - 4.1 CLOUD CONNECTOR
   - 4.2 CLOUD GATEWAY
5. CONCLUSÃO
   - 5.1 TRABALHOS FUTUROS

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

1 - Introdução



- Avaliação das diferentes arquiteturas disponíveis de Backup para a Nuvem: Cloud Connector e Cloud Gateway
- Objetivo: levantar requisitos e definir melhores práticas de desempenho e capacidade no processo de substituição de Backup tradicional em fita

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

1.1 - Escopo

- Backup de dados de Servidores, físicos ou virtuais on-premises para a nuvem pública

- Fora do escopo do estudo:
  - Backup de dados pessoais/individuais
  - Backup de servidores (IaaS) e sistemas (SaaS) na nuvem pública
  - Archiving ou servidor de arquivos na nuvem

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

2 – Trabalhos relacionados

- ZRNEC A.; LAVBIC, D. **Comparison of Cloud vs. Tape Backup Performance and Costs with Oracle Database**. University of Ljubljana, Eslovênia. JIOS, VOL. 35, NO. 1 2011. PP. 135-142

Abstract

Current practice of backing up data is based on using backup tapes and remote locations for storing data. Nowadays, with the advent of cloud computing a new concept of database backup emerges. The paper presents the possibility of making backup copies of data in the cloud. We are mainly focused on performance and economic issues of making backups in the cloud in comparison to traditional backups. We tested the performance and overall costs of making backup copies of data in Oracle database using Amazon S3 and EC2 cloud services. The costs estimation was performed on the basis of the prices published on Amazon S3 and Amazon EC2 sites.

Keywords: computer center, bandwidth, service, cloud, performance, cost, remote location

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

2 – Trabalhos relacionados

- MKANDAWIRE S. **Improving Backup and Restore Performance for Deduplication-based Cloud Backup Services**. University of Nebraska – Lincoln, USA. 2012

Abstract

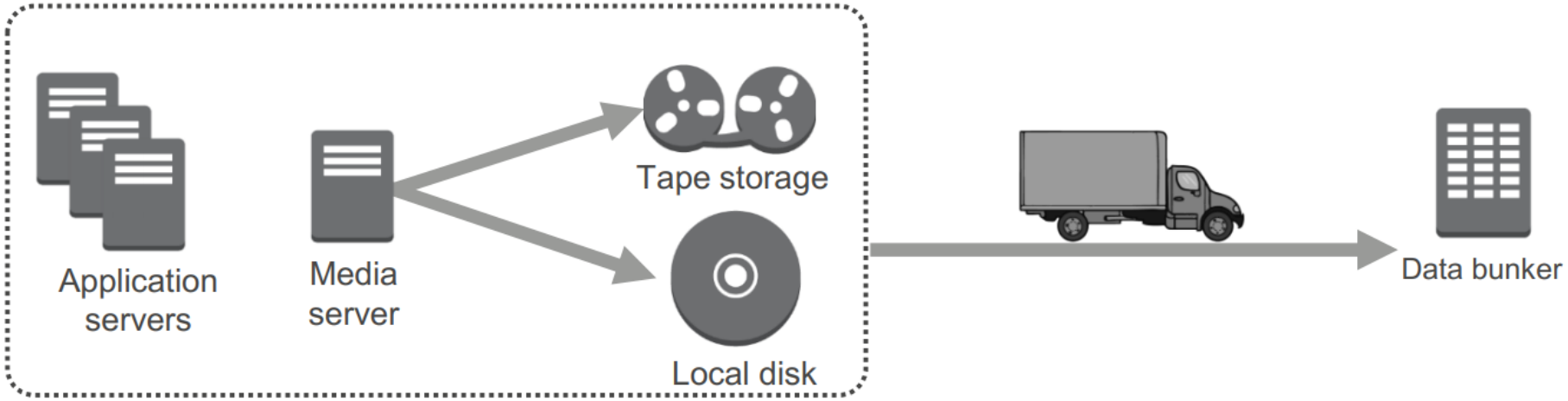
The benefits provided by cloud computing and the space savings offered by data deduplication make it attractive to host data storage services like backup in the cloud. Data deduplication relies on comparing fingerprints of data chunks, and store them in the chunk index, to identify and remove redundant data, with an ultimate goal of saving storage space and network bandwidth.

The restore performance over the relatively low bandwidth wide area network (WAN) links is another drawback in the use of cloud backup services. High speed network connectivity is either too expensive for most organizations or reserved for special applications. Removing redundant data before transmitting over the WAN offers a viable option to improve network throughput during the restore operation.

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

3 – Motivação

- Backup Tradicional



- Software de Backup + Biblioteca de fita (com armazenamento externo) e/ou Appliance de Backup local (Storage/Disco)

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

3 – Motivação

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

3 – Motivação

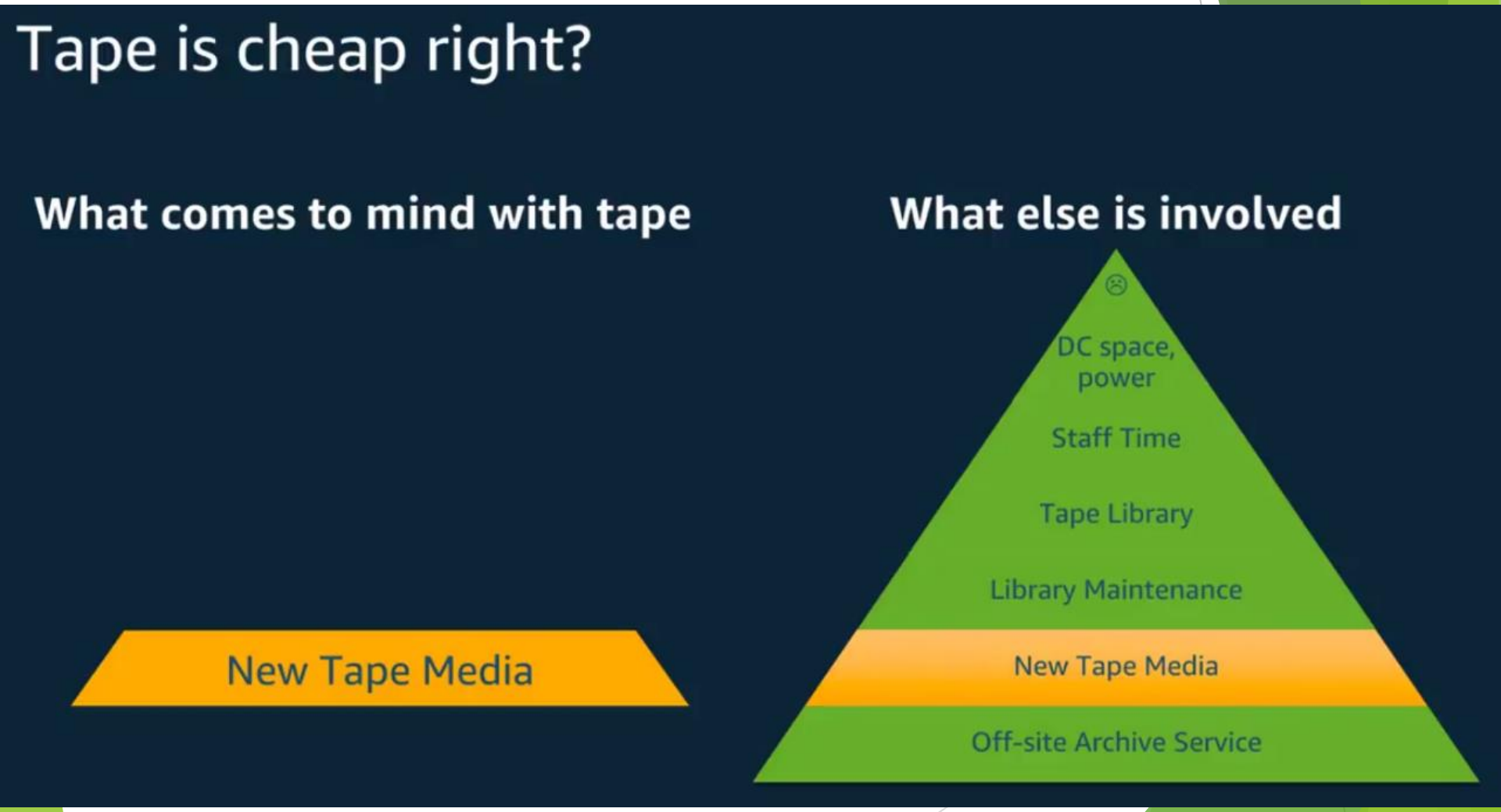
- A fita é uma tecnologia **antiga** que tende a se aposentar

- A fita é barata (baixo OPEX), mas seu custo de **gerenciamento, manutenção** e principalmente seu **risco** é alto

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

3 – Motivação

- Custos do Backup em Fita



Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

3 – Motivação

- Esforço (horas) em migração para manter a compatibilidade (apenas 2 versões anteriores de LTO são compatíveis para leitura e 1 versão para escrita) https://www.lto.org/about-the-lto-program/faq/

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

3 – Motivação

- Os Appliances de Backup (em Disco/ Storage) são mais apropriados para **curta/média retenção**. Alto CAPEX
- Não atendem requisitos (compliance) de armazenamento externo.
- Não podem ser considerados como alternativa única de substituição às fitas
  - A não ser que possuam replicação para outro appliance em um Data Center diferente

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

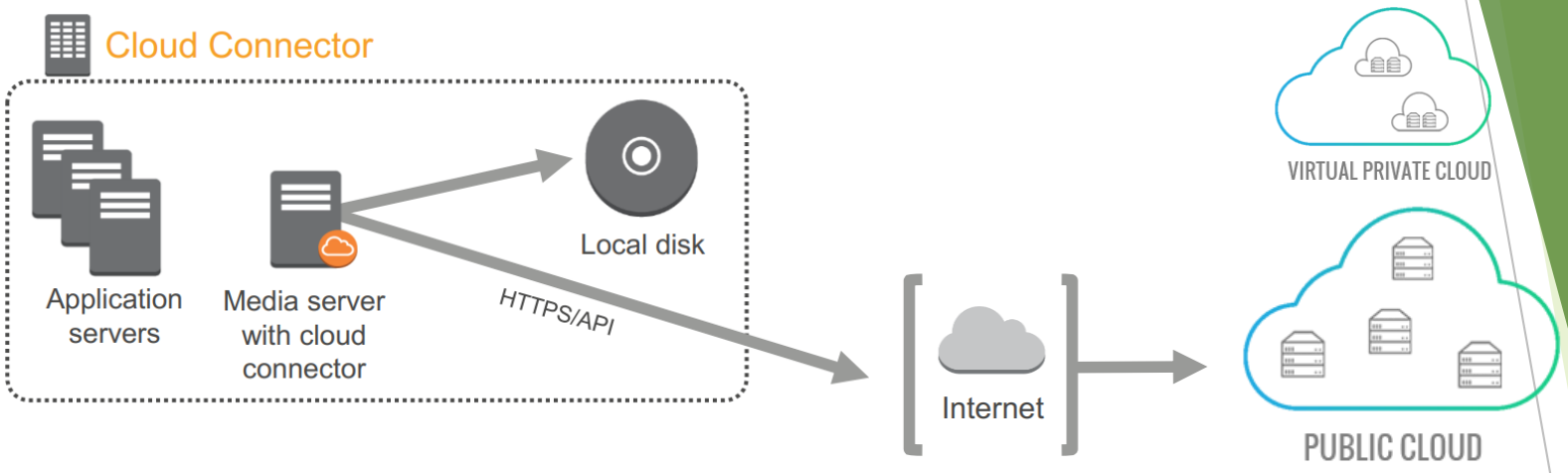
4 – Backup para a nuvem

Requisitos desejáveis

1. Processo eficiente de **restore** dos dados
2. Otimizar o consumo do **link** de tráfego para a nuvem durante o processo do backup e restore
3. Implementar técnicas eficiências de **desduplicação** e **compressão** de dados a fim de reduzir o espaço consumido na nuvem
4. Trafegar e armazenar os dados com **segurança** (criptografia) para a nuvem
5. Possuir área de **staging** local on-premises, com desduplicação e compressão

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.1 – Cloud Connector

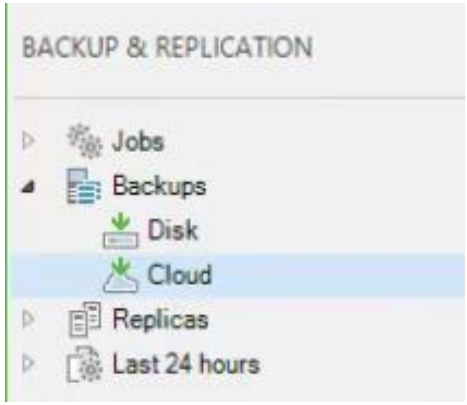


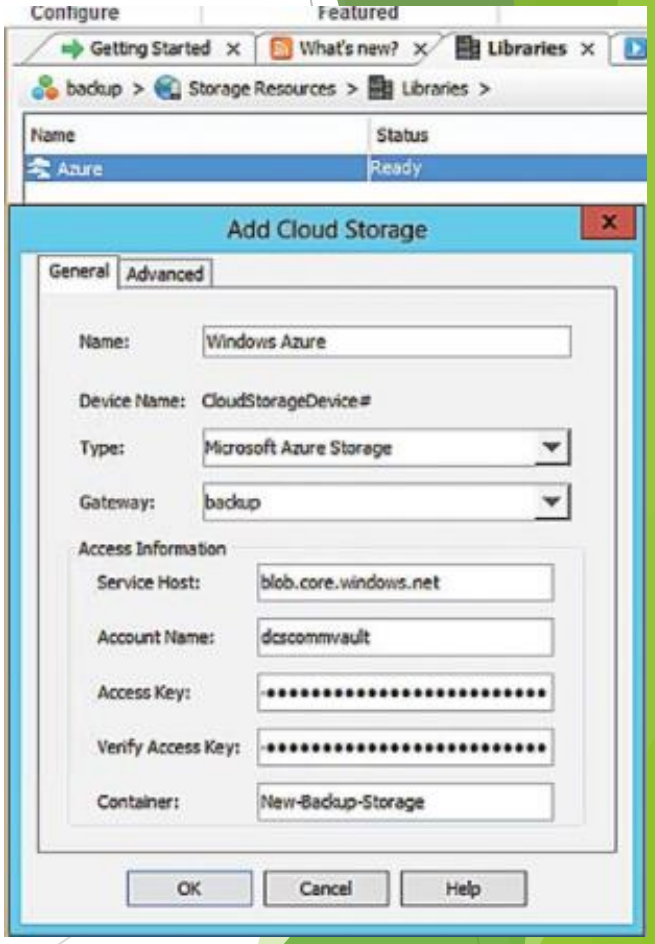
- Connector fornecido pelo fabricante do software de backup (Veeam, Commvault, Veritas, etc) ou pelo provedor de nuvem (AWS, Azure, Box, etc)
- Plugin/Módulo/Agente adicional (licenciado ou não) para nuvens públicas e privadas (S3, etc)

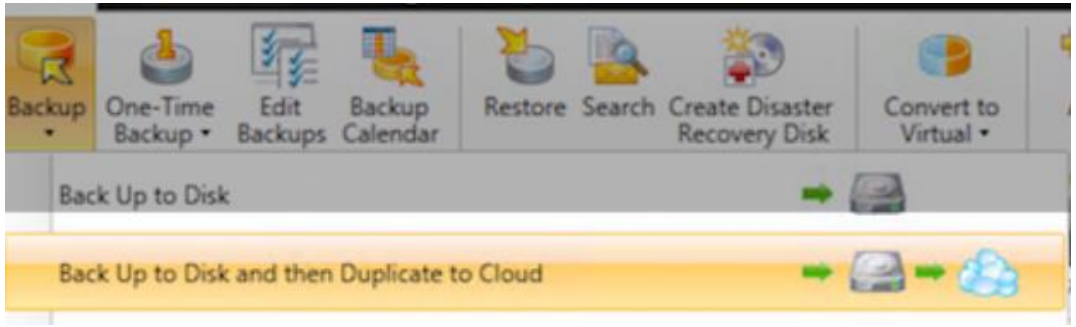
Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.1 – Cloud Connector

A) Softwares de Backup



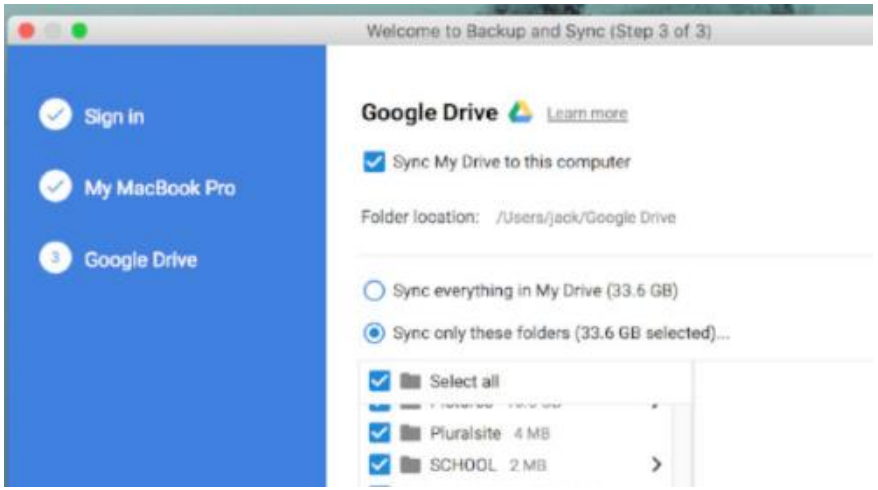


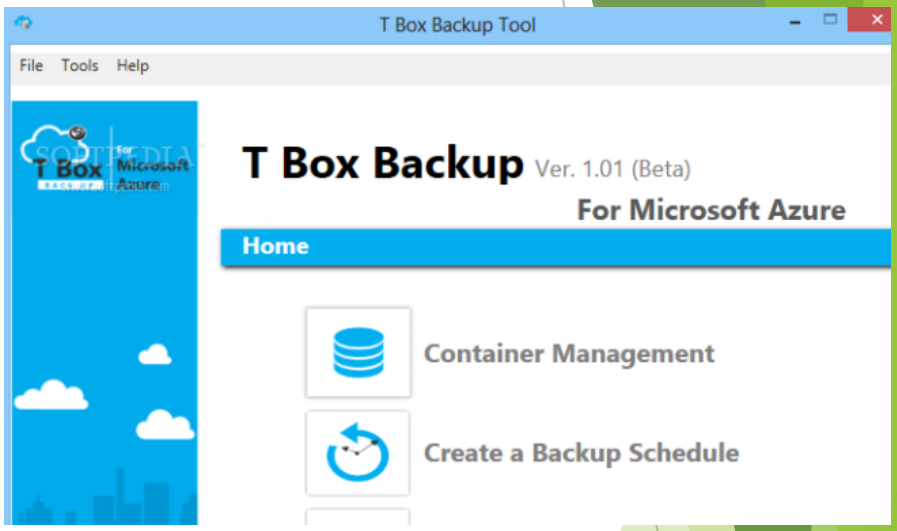


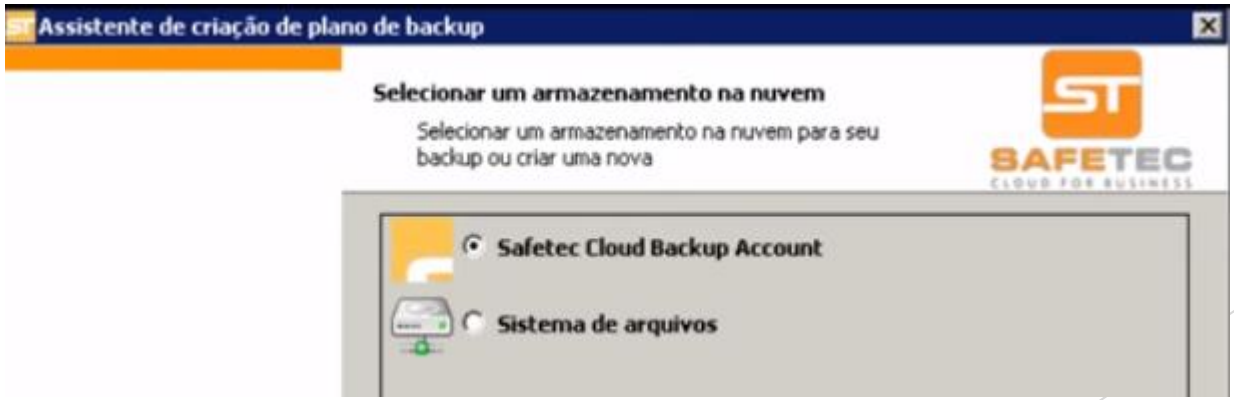
Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.1 – Cloud Connector

B) Provedores de nuvem pública







Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.1.1 – Avaliação

- Normalmente já é licenciado com o software de backup existente: não apresenta custo adicional
- Quando disponibilizado pelos provedores de nuvem e hosting: custo baixo
- Implementação mais simples
- Possui desduplicação/compressão apenas na área de staging on-premises. Na nuvem o dado é “reidratado” e consome mais espaço de armazenamento no provedor: maior custo mensal

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.2 – Cloud Gateway

- “A cloud gateway emulates a disk array, a block-based device or a file server. The device is placed at the customer's premises and it's delivered either as a piece of hardware or as software that's loaded onto hardware. They're particularly interesting because they solve a couple of problems. Things like backup, archive and file sharing are more appropriate for cloud gateways.”

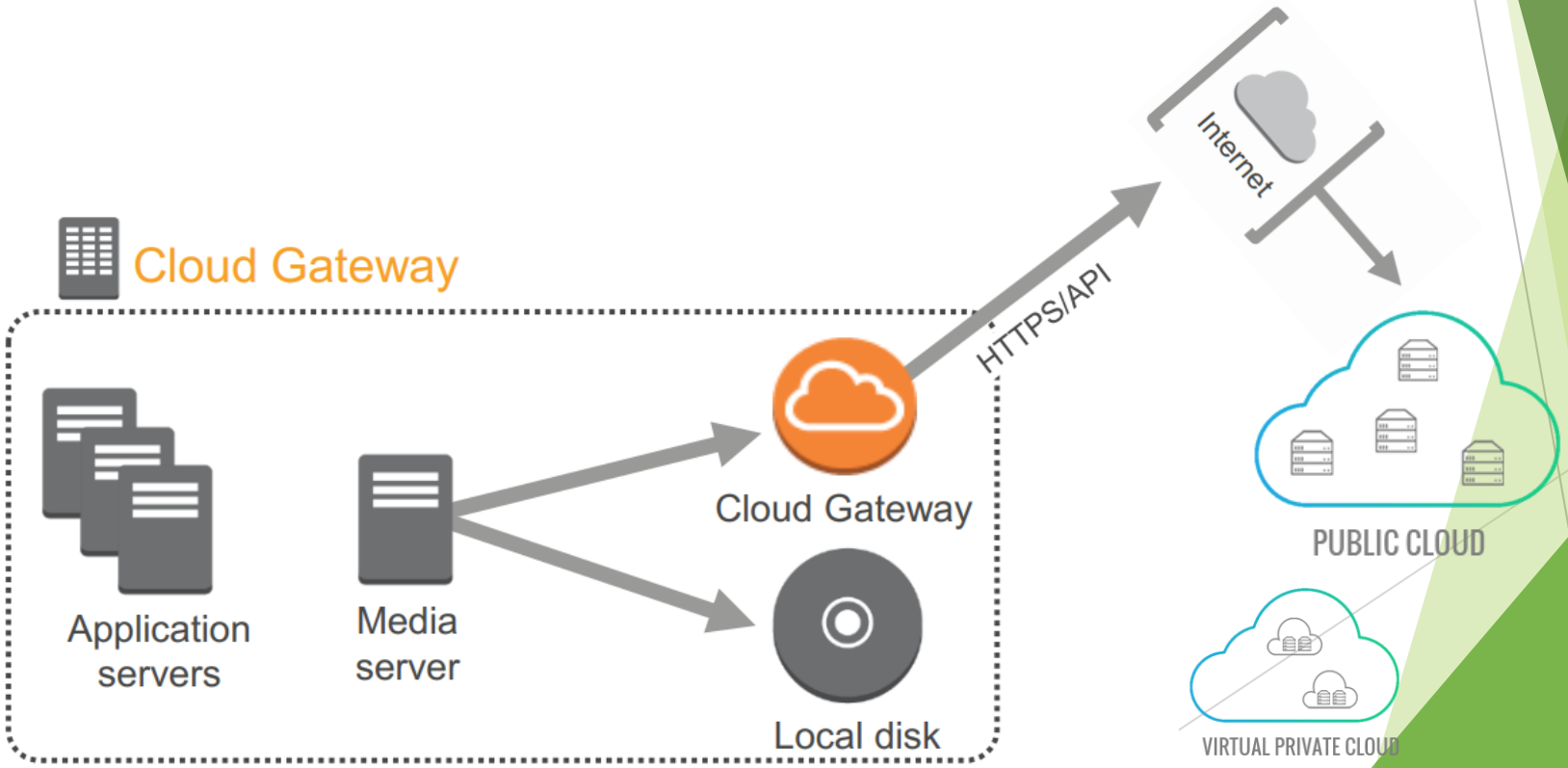
Gene Ruth

Research Director at Gartner Inc.

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.2 – Cloud Gateway

- Gateway físico ou virtual, instalado on-premises, fornecido por fabricantes de Storage ou software de Backup ou mesmo pelos provedores de nuvem pública



Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

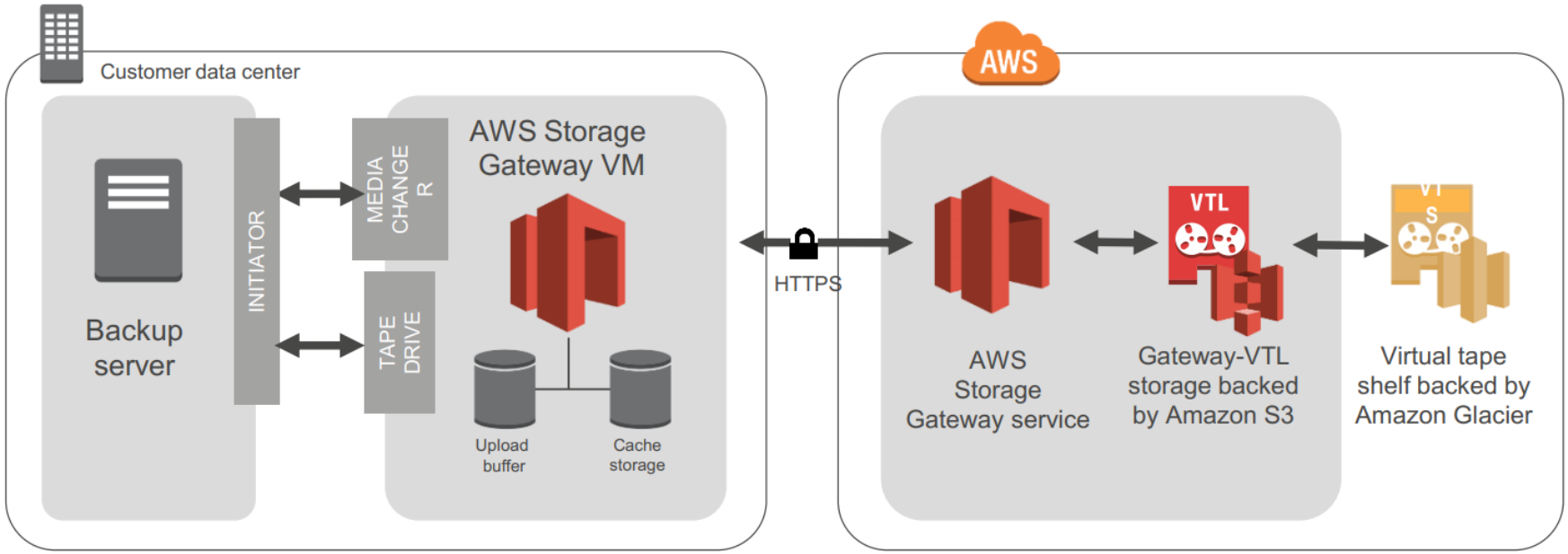
4.2 – Cloud Gateway

A) Provedores de nuvem pública

- amazon web services

Cloud Gateway: AWS Storage Gateway

GATEWAY-VIRTUAL TAPE LIBRARY ARCHITECTURE

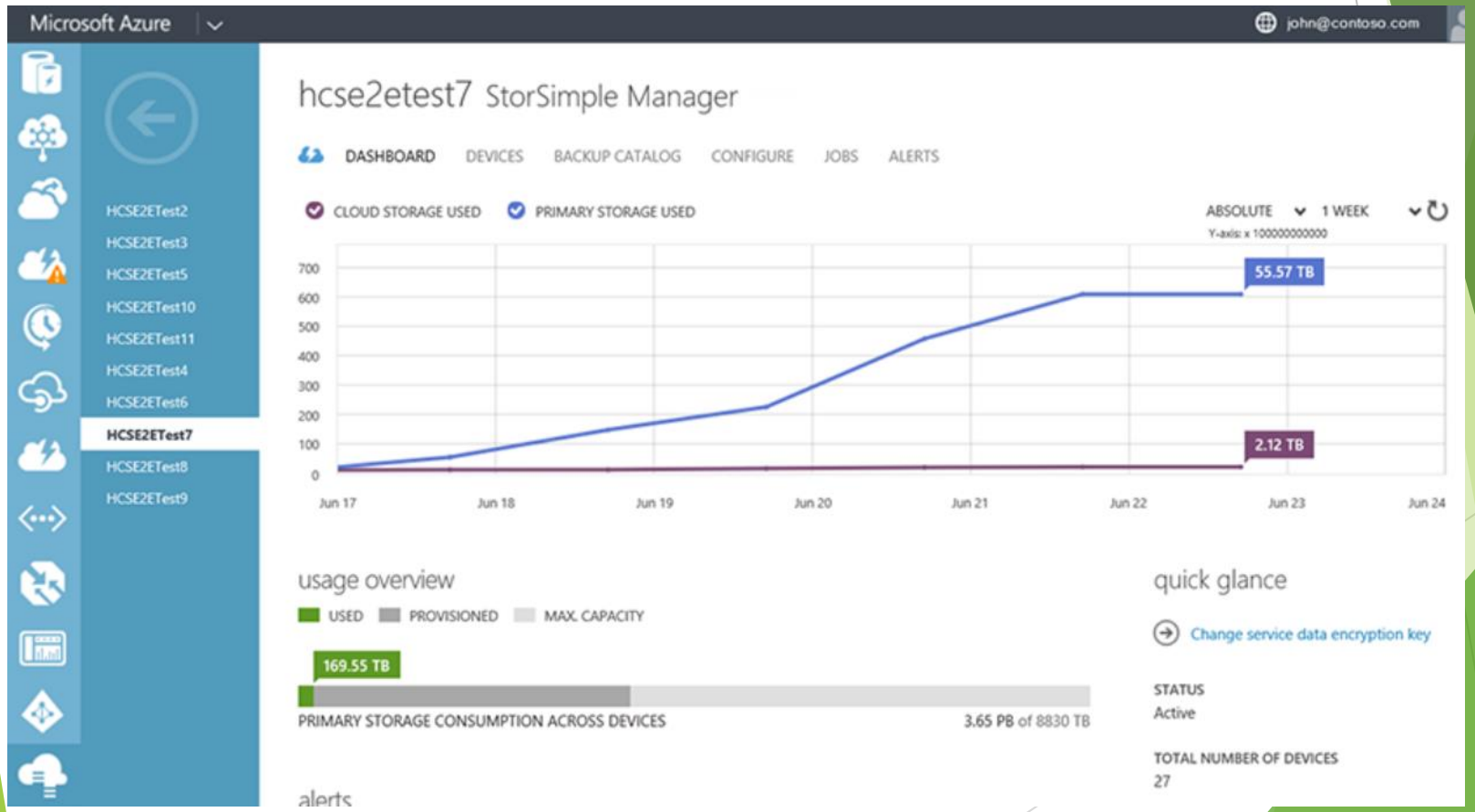


Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.2 – Cloud Gateway

A) Provedores de nuvem pública

- Microsoft Azure

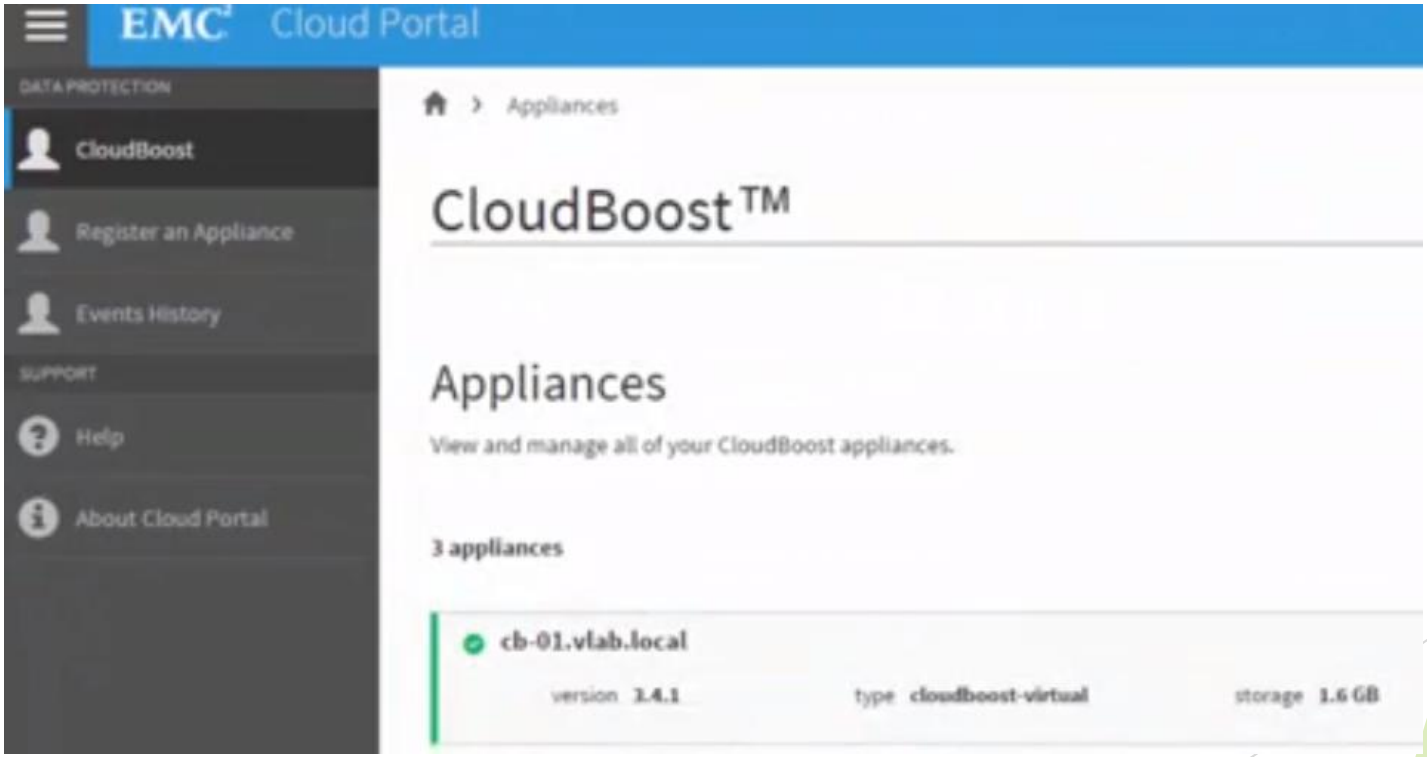


Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.2 - Cloud Gateway

B) Fabricantes de Storage

- DELL EMC

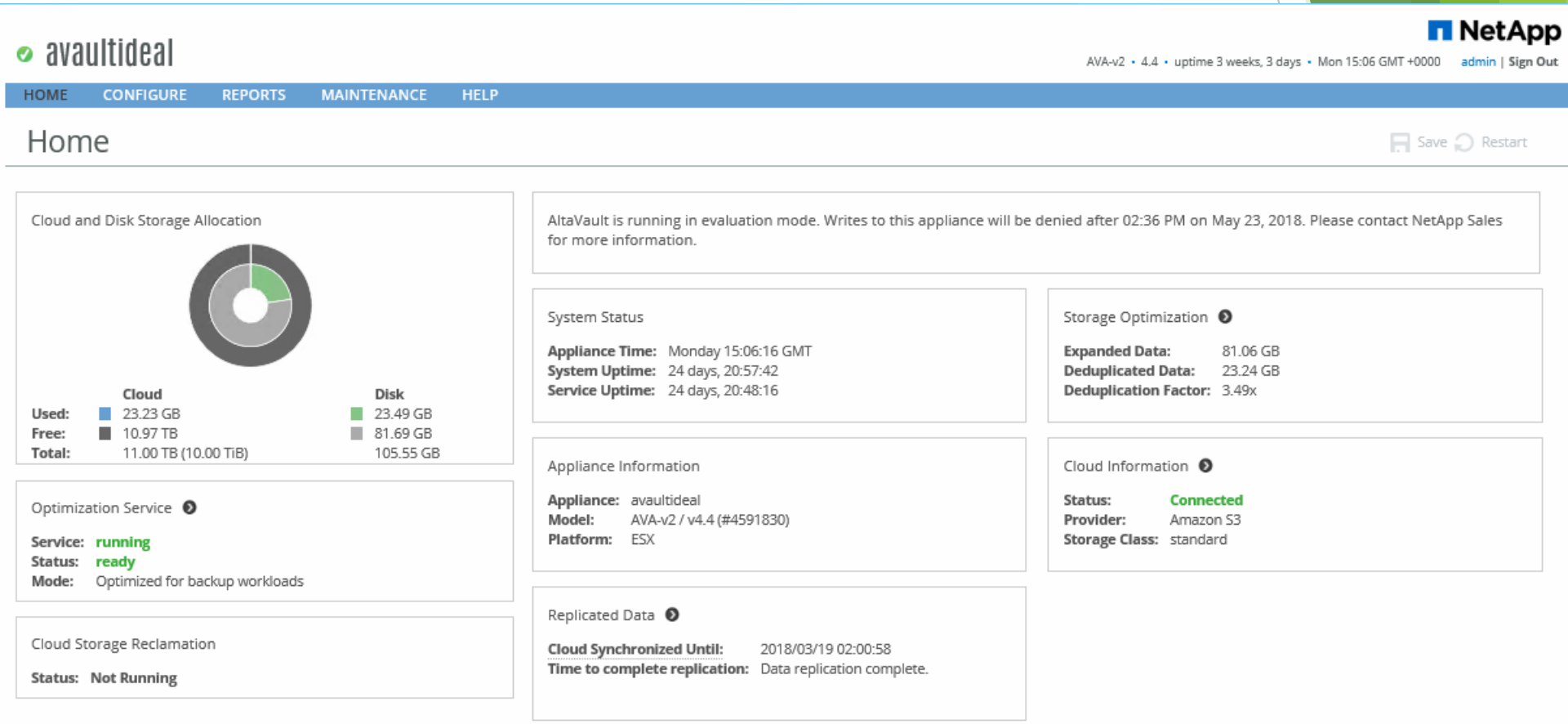


Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.2 - Cloud Gateway

B) Fabricantes de Storage

- NetApp®



Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.2 – Cloud Gateway

C) Outras empresas/startups de soluções de Cloud Storage

ctera
your files. your cloud.

NASUNI

panzura

TwinStrata

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

4.2.1 – Avaliação

- Integra com os principais softwares de backups disponibilizando uma library virtual
  - Maioria apenas via ethernet, sem conexão FC (SAN)
- Dos fabricantes de Storage não amarra o provedor de nuvem pública e/ou privada a ser escolhido
- Dos provedores de nuvem disponibiliza um ambiente de DR no próprio proveedor
- Fornece/requer área em disco on-premises para staging do backup e restore rápido
- Desduplicação dos dados locais e na nuvem: menor custo de storage (CAPEX) e do provedor (OPEX)

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

5 – Conclusão

Item	Cloud Connector	Cloud Gateway
Otimização do link	Sim ☺ ±	Sim ☺
Criptografia	Sim ☺	Sim ☺
Custo de licença	Não ☺	Sim ☹
Desduplicação local	Sim ☺ ±	Sim ☺
Desduplicação na nuvem (redução de custo)	Não ☹	Sim ☺
Integra com vários software de Backup	Não ☹ (somente o próprio software)	Sim ☺
Integra com vários provedores de nuvem pública e privada	Não ☹ (somente com os principais)	Sim ☺ (compatível com S3 e outros protocolos objetos)
Restauração de dados	Da nuvem ☹	On-premises e da Nuvem (mais rápido) ☺

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

5.1 – Trabalhos futuros

- Avaliar as soluções de backup dos dados já existentes na nuvem: IaaS e SaaS
- Alguns produtos disponibilizam o backup (versões) dos arquivos, e-mails, etc por alguns dias, porém não atenderiam compliance de retenção por um prazo maior. Ex.: Office365 = 14 dias ("Microsoft takes backups of site collections every 12 hours and keeps these backups for 14 days“)

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil

Perguntas/Comentários

Obrigado pelo seu tempo!

kleber@pise4.com.br

(11) 98970-3801

Proibida cópia ou divulgação sem permissão escrita do CMG Brasil