



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
UNIDADE ACADÊMICA DE EDUCAÇÃO A DISTÂNCIA E TECNOLOGIA
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

Aluizio Felipe Miranda da Silva

**SISTEMAS DE BACKUP: UM COMPARATIVO ENTRE BACULA
COMMUNITY E GOOGLE DRIVE PARA EMPRESAS**

Surubim - PE
2022

Aluizio Felipe Miranda da Silva

**SISTEMAS DE BACKUP: UM COMPARATIVO ENTRE BACULA
COMMUNITY E GOOGLE DRIVE PARA EMPRESAS**

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Sistemas de Informação da Unidade Acadêmica de Educação a Distância e Tecnologia da Universidade Federal Rural de Pernambuco como requisito parcial à obtenção do grau de Bacharel.

Orientadora: Prof^ª Ma. Adalmeres Cavalcanti da Mota

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

S586s

Silva, Aluizio Felipe Miranda da
Sistemas de backup: um comparativo entre bacula community e google drive para empresas / Aluizio Felipe
Miranda da Silva. - 2022.
74 f. : il.

Orientadora: Adalmeres Cavalcanti da Mota.
Inclui referências, apêndice(s) e anexo(s).

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal Rural de Pernambuco, Bacharelado em
Sistemas da Informação, Recife, 2022.

1. Segurança da informação. 2. Backup. 3. Bacula. 4. Google drive. I. Mota, Adalmeres Cavalcanti da, orient. II. Título

CDD 004

ALUÍZIO FELIPE MIRANDA DA SILVA

**SISTEMAS DE BACKUP: UM COMPARATIVO ENTRE BACULA
COMMUNITY E GOOGLE DRIVE PARA EMPRESAS**

Trabalho de Conclusão de Curso apresentado em cumprimento às exigências do curso de Bacharelado em Sistemas de Informação da Unidade Acadêmica de Educação a Distância e Tecnologia/UFRPE para obtenção do título de Bacharel em Sistemas de Informação, sob a orientação do (a) Profª Ma. Adalmeres Cavalcanti da Mota.

Aprovado em 28 de junho de 2022.

Profª Ma. Adalmeres Cavalcanti da Mota – UAEADTec/UFRPE

Profª Dra. Juliana Regueira Basto Diniz – UAEADTec/UFRPE

Profª Ma. Jeneffer Cristine Ferreira – UFRPE

Surubim – PE, 2022

AGRADECIMENTOS

Agradeço primeiramente à Deus por ter me dado capacidade de chegar até aqui.

À minha orientadora, Adalmeres Cavalcanti, e demais professores, que fizeram com que eu conseguisse compilar todos os conhecimentos do curso neste trabalho.

À minha esposa, Fernanda Carolina, pelo apoio e paciência ao longo de todo o curso, e minha mãe, Katia Maria, por não ter medido esforços para me proporcionar uma educação de qualidade.

RESUMO

No mundo informatizado que vivemos hoje, em plena revolução da sociedade 5.0, os cuidados básicos com os dados ainda são uma forte preocupação dos profissionais de TI. Backups de dados são a forma mais eficaz de se tratar as recuperações de desastres. É um conceito antigo, mas que para acompanhar os novos formatos de dados, mídias e ataques mal-intencionados sofisticados, evolui a cada dia. Com os dados se tornando, nos últimos anos, um dos ativos mais valiosos das organizações, é necessário sempre estar atento às melhores soluções de backups. Este trabalho é um estudo de caso que faz um comparativo entre duas soluções de backup conceituadas e acessíveis do mercado, estabelecendo variáveis desde a instalação, passando pela configuração e operação em produção destas soluções. Ao final apresentando um questionário que serve de roteiro para apoiar a decisão de novos profissionais de TI, que possam estar em dúvida sobre qual a solução mais adequada para sua instituição.

Palavras-chave: Segurança da Informação. Backup. Bacula. Google Drive.

ABSTRACT

In the computerized world we live in today, in the midst of a 5.0 society revolution, basic data care is still a strong concern for IT professionals. Data backups are the most effective way to treat disaster recoveries. It's an old concept, but to keep up with new data formats, medias and sophisticated malicious attacks, be evolving every day. With data becoming, in recent years, one of the most valuable assets of organizations, it's necessary to always be on the lookout for better backup solutions. This work is a case study that makes a comparison between two reputable and accessible backup solutions in the market, establishing variables from the installation, through the configuration and operation in production of these solutions. At the end, presenting a questionnaire that serves as a guide to support the decision of new IT professionals, who may have doubts about the most suitable solution for their institution.

Keywords: Information Security. Backup. Bacula. Google Drive.

LISTA DE FIGURAS

Figura 1 - Esquema de módulos do <i>Bacula</i> por arquivos de configuração	28
Figura 2 - Rack principal do IFPE Palmares	30
Figura 3 - <i>Workstation</i> onde são alocados os volumes do <i>Bacula</i>	31
Figura 4 - Passos para download do Drive <i>File Stream</i>	33
Figura 5 - Drivers virtuais criados pelo aplicativo após instalação	34
Figura 6 - Estatísticas de uso por usuário no Google Drive	49
Figura 7 - Estatísticas de tipos de arquivos compartilhados.....	50
Figura 8 - Painel de controle do <i>Bacula</i>	51
Figura 9 - Estatísticas de Volumes do <i>Bacula</i>	52

LISTA DE QUADROS

Quadro 1 - Ciclo da Torre de Hanói	22
Quadro 2 - Comparativo entre as Soluções	42
Quadro 3 - Tabela de Pontuação	46
Quadro 4 - Resultados do Questionário	48

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
BSD	Berkeley Software Distribution
CD	Compact disc
DAT	Digital Audio Tape
DVD	Digital Versatile Disc
GFS	Grandfather-Father-Son
IEC	International Electrotechnical Commission
IFPE	Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco
ISO	International Organization of Standardization
LAMP	Linux, Apache, MySQL, PHP/Perl/Python
LZO	Lempel-Ziv-Oberhumer
NAS	Network Attached Storage
NBR	Norma Brasileira
PC	Personal Computer
PDF	Portable Document Format
PSI	Política de Segurança da Informação
RAID	Redundant Array of Inexpensive disks
RH	Recursos Humanos
RPO	Recovery Point Objective
RTO	Recovery Time Objective
SGBD	Sistema Gerenciador de Banco de Dados
SO	Sistema Operacional
SQL	Structured Query Language
STM	Superior Tribunal Militar
TCC	Trabalho de Conclusão de Curso
TI	Tecnologia da Informação
TTy	TeleTYpewriter
VPN	Virtual Private Network

SUMÁRIO

1	INTRODUÇÃO	10
1.1	PROBLEMATIZAÇÃO	11
1.2	OBJETIVOS	12
1.2.1	Objetivo Geral	12
1.2.2	Objetivos Específicos	12
1.3	ESTRUTURA DO TRABALHO	12
2	FUNDAMENTAÇÃO TEÓRICA	14
2.1	SEGURANÇA DA INFORMAÇÃO	14
2.2	POLÍTICA DE SEGURANÇA DA INFORMAÇÃO	15
2.3	LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS (LGPD)	16
2.4	BACKUP	18
2.5	POLÍTICA DE BACKUP	19
2.6	ESTRATÉGIAS DE BACKUP	20
2.6.1	FIFO	20
2.6.2	GFS	21
2.6.3	A Torre de Hanói	21
2.7	SOLUÇÕES DE BACKUP	22
2.7.1	Amanda Network Backup	23
2.7.2	ARCserve Backup	23
2.7.3	Storages NAS	24
2.7.4	Azure Backup	24
2.7.5	Acronis Cyber Protect Cloud	24
2.7.6	Amazon S3	25
2.7.7	Veeam Backup & Replication	25
2.7.8	Google Drive File Stream	26
2.7.9	Bacula Community	26
3	PROJETO	29
3.1	METODOLOGIA APLICADA NA ANÁLISE	29
3.2	ANÁLISE PARA IMPLANTAÇÃO E USO DAS SOLUÇÕES	30
3.2.1	Estrutura existente do IFPE Campus Palmares	30
3.2.2	Custos e requisitos iniciais	31

3.3	INSTALAÇÃO E IMPLANTAÇÃO	33
3.3.1	Instalação do Google Drive <i>File Stream</i>	33
3.3.2	Funcionalidades e vantagens do Google Drive <i>File Stream</i>	34
3.3.3	Instalação do <i>Bacula Community</i>	36
3.3.4	Funcionalidades e vantagens do <i>Bacula Community</i>	37
3.4	IDENTIFICAÇÃO DE VARIÁVEIS	37
3.4.1	Infraestrutura	38
3.4.1.1	Complexidade no cálculo de armazenamento – <i>Bacula</i>	38
3.4.1.2	Custo de infraestrutura – <i>Bacula</i>	39
3.4.1.3	Complexidade de instalação e operação de hardware e software - <i>Bacula</i>	39
3.4.2	Sensibilidade dos dados	39
3.4.3	Disponibilidade dos arquivos	40
3.4.4	Sincronização	40
3.4.5	Colaboração	41
3.5	COMPARATIVO ENTRE AS DUAS SOLUÇÕES EM RELAÇÃO AS VARIÁVEIS	41
3.6	PROCESSO DE ELABORAÇÃO DO QUESTIONÁRIO	43
3.7	RESULTADOS DA IMPLANTAÇÃO DAS SOLUÇÕES NO IFPE	48
4	CONCLUSÃO	53
4.1	CONSIDERAÇÕES FINAIS	53
4.2	CONTRIBUIÇÕES DESTE TRABALHO	54
4.3	PROPOSTA PARA TRABALHOS FUTUROS	54
	REFERÊNCIAS BIBLIOGRÁFICAS	55
	ANEXO A – PESQUISA SOBRE EXISTÊNCIA DE SOLUÇÕES DE BACKUP NOS CAMPI DO IFPE	58
	ANEXO B – SCRIPT DE INSTALAÇÃO DO <i>BACULA</i>	63
	ANEXO C – ARQUIVO DE CONFIGURAÇÃO DO <i>DIRECTOR BACULA</i>	65
	ANEXO D – ARQUIVO DE CONFIGURAÇÃO DO <i>FILE DAEMON BACULA</i>	66
	ANEXO E – ARQUIVO DE CONFIGURAÇÃO DOS AGENDAMENTOS (MODULAR)	67
	ANEXO F – ARQUIVO DE CONFIGURAÇÃO DOS <i>POOLS</i> (MODULAR)	68
	ANEXO G – ARQUIVO DE CONFIGURAÇÃO DOS <i>FILESETS</i> (MODULAR)	69
	ANEXO H – ARQUIVO DE CONFIGURAÇÃO DOS <i>CLIENTS</i> E <i>JOBS</i> (MODULAR)	70
	ANEXO I – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO	72

1 INTRODUÇÃO

Presentemente, a informação se tornou um ativo valioso nas instituições, este ativo precisa ser armazenado com segurança, de forma que somente as pessoas autorizadas tenham acesso, ter um nível de disponibilidade satisfatória, de forma que seja acessível sempre que os colaboradores precisarem, e precisa ser consistente ao longo de todo o seu ciclo de vida, sem sofrer alterações não planejadas ao longo dos processos. Pode-se dizer que isso resume os três pilares da segurança da informação, que são respectivamente: Confidencialidade, Disponibilidade e Integridade. Este último pilar está diretamente relacionado com: controle de acesso, armazenamento e recuperação de dados. É com relação a esses últimos dois pontos que o valor do backup está associado. A segurança da informação é necessária para assegurar o bom andamento das atividades da organização, e para isso, soluções de backups e recuperação de dados não podem ser dispensadas.

Não importa a eficiência do seu computador ou do disco rígido externo, um dia eles se tornarão desgastados. Essa é a natureza de todo hardware. O técnico de manutenção de computadores talvez consiga recuperar os dados, mas não há nenhuma garantia disso. Se o backup não for feito, a segurança dos dados não pode ser garantida. E o que é pior, a internet está cheia de ameaças à integridade dos dados. Os vírus e cavalos de troia não só roubam os dados, mas também os excluem em alguns casos. Outro fenômeno crescente e relevante é o Ransomware, no qual um hacker insere um vírus no computador que criptografa os dados. Para descriptografar os dados, será necessário pagar um resgate a esses hackers. Se foi feito o backup dos dados, essa é uma preocupação a menos. Basta limpar o disco rígido e restaurar o backup mais recente. Seja devido a uma falha mecânica, um desastre natural ou uma atividade criminosa, por qualquer desses motivos, a perda de dados será irreparável caso não ocorra o backup. (NORTONLIFELOCK, 2020).

Felizmente, hoje em dia, existem tanto soluções gratuitas quanto pagas para resolver o dilema dos backups. Várias questões como custo do produto e volume de dados pesam bastante na hora de decidir qual solução implantar, muitas vezes essas soluções podem resultar em uma readaptação dos processos para os colaboradores que participam do ciclo de vida da informação, e isso também deve ser levado em conta na hora da escolha. Entre a variedades de ferramentas e serviços que existem, este trabalho seleciona duas muito bem-conceituadas e recomendadas no mercado, uma gratuita, o Bacula Community, e outra paga, o Drive File Stream,

estabelecendo um comparativo entre elas, fazendo um estudo de caso para apoiar a decisão na escolha da melhor para variados cenários.

Para a avaliação da melhor solução o estudo fez uso heurísticas buscando encontrar pontos de influência no processo de implantação e utilização das soluções. Os pontos de influência encontrados, também chamados ao longo do trabalho de variáveis, foram definidos segundo a viabilidade de utilização das ferramentas comparadas levando em conta as especificidades que outras organizações podem ter. As variáveis catalogadas ao longo do estudo foram nomeadas como: Infraestrutura de hardware e software preexistente, complexidade no cálculo do armazenamento necessário, custo de infraestrutura de implantação, complexidade de instalação e operação, sensibilidade dos dados, disponibilidade dos dados, sincronização e manipulação de documentos por multiusuários simultâneos (colaboração).

Através das definições dos pontos de influência o autor elabora um questionário que serve como roteiro para apoio na decisão entre qual solução implantar, tanto para a organização onde o estudo de caso foi realizado quanto para outras organizações, já que através das respostas do questionário são coletadas as especificidades da organização a qual o respondente pertence.

1.1 PROBLEMATIZAÇÃO

No Instituto Federal de Pernambuco (IFPE) - *Campus* Palmares, a equipe de TI conta com duas soluções para manter seus dados administrativos. A primeira é um sistema *Open Source*, o *Bacula Community*, implantado recentemente por este pesquisador. A segunda, uma parceria com a Google, que provém espaço de armazenamento em nuvem ilimitado, através do Google Drive. Em um dos debates entre a equipe de TI, surgiu o questionamento do que seria melhor utilizar, levando em conta os diferentes tipos de arquivos, frequência de alteração e facilidade de adaptação do usuário, pairou sobre a equipe a dúvida de qual solução escolher. Nos demais *campi* do IFPE existe o mesmo dilema, contam tanto com a solução *Bacula*, ou derivados, quanto com o Google Drive. No entanto, não existe orientação ou diretiva sistêmica para definir qual utilizar. Além disso, existe carência de informações na internet que possam servir de comparativo direto entre as duas opções. Diante disso, viu-se que esta não seria uma decisão rápida ou fácil, e que diversos aspectos teriam que ser analisados antes de oficializar qual a melhor solução para o *campus*. Além disso, para verificar o nível de utilização de backup nos demais *campi*, foi realizada uma pesquisa, apresentada no anexo i, onde foram evidenciados variados problemas como: falta de conhecimento para especificar uma solução de backup,

ausência de métodos de backup para os serviços de TI e até mesmo eventos de perda de dados de forma irreversível.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Mapear variáveis para definir qual a melhor solução entre as apresentadas, levando em conta a estrutura, os dados e os usuários do IFPE-Palmares, através da criação de um roteiro para que seja possível identificar essas variáveis em outras Instituições, sejam nos demais *campus* do IFPE, ou em outras organizações, levando em conta as particularidades de cada uma, apoiando na decisão.

1.2.2 Objetivos Específicos

- Compreender detalhadamente as funcionalidades básicas do *Bacula Community 9.6* e *Drive File Stream*;
- Estabelecer variáveis utilizando heurística para análise da adequação das soluções em diferentes cenários institucionais;
- Identificar através do comparativo das soluções qual a mais indicada para o *Campus* IFPE Palmares;
- Estabelecer o roteiro de forma que fique claro nos resultados quando uma das soluções se torna inviável.

1.3 ESTRUTURA DO TRABALHO

Este trabalho foi estruturado em cinco capítulos. O capítulo 1 trata da importância dos dados como um ativo valioso nas organizações, enfatiza a seriedade da segurança da informação e sua relação com o backup de dados, trata da problematização e por fim do objetivo geral e objetivos específicos do trabalho.

O capítulo 2 abordará a fundamentação da segurança da informação, normativas aplicadas ao armazenamento de dados e recuperação de desastres, os conceitos que compõem backup de dados locais e em nuvem, as estratégias utilizadas na realização de backups, as

soluções de backup mais conhecidas no mercado atualmente, além da definição e funcionalidades do Google Drive para empresas e do *Bacula Community*, dando contexto para o que será abordado no capítulo seguinte.

O capítulo 3 é iniciado com uma explicação sobre a metodologia do trabalho, em seguida é relatado o contexto da instituição onde ocorreu o estudo de caso, após isso o capítulo transcorre sobre a implantação das soluções, configurações e operação das mesmas. Ao passo que também são aplicadas as heurísticas para identificação das variáveis que influenciam a escolha entre as soluções, por fim é realizado um comparativo entre as soluções e explicado como foi elaborado e como funciona o questionário que servirá de roteiro.

No capítulo 4 são apresentadas as considerações finais, as principais contribuições do estudo, suas limitações e indicações para trabalhos futuros.

Por fim, são listadas as Referências Bibliográficas, os Anexos e os Apêndices.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo irá abordar conceitos estudados para elaboração do TCC. Serão descritos alguns elementos que compõem o ciclo de vida dos dados, principalmente da fase de armazenamento e segurança.

2.1 SEGURANÇA DA INFORMAÇÃO

A segurança da informação pode ser definida como todas as medidas necessárias para garantir confidencialidade, integridade e disponibilidade das informações. Essas características, conhecidas como os três pilares da segurança da informação podem ser resumidas, de acordo com a NBR ISO/IEC 27002 (ABNT, 2013), como abaixo.

- **Confidencialidade:** Preceito da segurança da informação que requer que dados devem somente ser acessados por pessoas autorizadas.
- **Integridade:** Preceito da segurança da informação que diz que dados só devem ser modificados por pessoas e serviços autorizados. Considerando todos os tipos de modificações como falhas de hardware e software, interferências ambientais e intervenção humana.
- **Disponibilidade:** Preceito da segurança da informação que requer que os dados sejam acessados somente por pessoas autorizadas, sempre no momento que são requisitados.

Ainda em se tratando de segurança da informação, temos a ISO 27003:2013 que fornece diretrizes para práticas de gestão de segurança da informação e normas de segurança da informação para as organizações, incluindo a seleção, a implementação e o gerenciamento de controles, levando em consideração os ambientes de risco da segurança da informação da organização (ABNT, 2013).

A ISO além de definir os três pilares da segurança da informação também explora outros como: autenticidade, não repúdio, responsabilidade e confiabilidade.

Segundo a TOTVS (2021), uma das maiores empresas de tecnologia do Brasil, a grande importância da segurança da informação vem do fato que a mesma trabalha para a proteção dos dados contra roubos e danos. Sem uma estratégia de segurança cibernética, qualquer instituição fica vulnerável contra ações maliciosas de violação de dados. As ameaças atuais não afetam

somente dados confidenciais das empresas, mas também podem destruir a sua imagem com os clientes e até mesmo colocá-los em risco legal.

De acordo com a IBM (2021), seu relatório anual de custo da violação de dados revelou que os custos em 2021 foram os maiores registrados nos últimos 17 anos. Além disso, o gasto médio foi de USD 1,07 milhão mais alto em violações em que o trabalho remoto foi um fator que causou a violação, quando comparado com as outras violações. É nítido que os usuários não estavam preparados para o trabalho remoto que foi forçado pela pandemia da COVID-19, ainda segundo o relatório, o vetor de ataque inicial mais comum registrado foi credenciais comprometidas, utilizado em 20% dos casos.

Segundo a Kaspersky Team (2021), o panorama de Ameaças Kaspersky de 2021, levantamento anual feito pela equipe de Pesquisa e Análise, demonstrou que de janeiro até agosto do ano em questão, houve um aumento de 23% dos ciberataques no Brasil em comparação ao ano anterior. A conclusão dos especialistas da empresa é que a segurança do trabalho remoto precisa melhorar, também é necessário expurgar a pirataria das casas e ambientes empresariais. Uma curiosidade é que os ataques de *phishing* parecem ter diminuído em relação ao ano anterior, mas infelizmente o Brasil ainda está no topo do ranking de ataques, com 15,4% dos internautas registrando tentativas.

Ainda segundo a SonicWall (2022), seu relatório de ameaças cibernéticas também reforça a situação, revelando que o Brasil obteve mais de 33 milhões de tentativas de invasões, se tornando o quarto país a mais sofrer ataques de *ransomware*, algo alarmante, principalmente se considerar que no ano passado estava na nona posição com apenas 2,8 milhões de ataques.

Todos os dados denotam a grande e antiga deficiência do Brasil quando se trata na segurança de dados, com tantas mudanças advindas do “novo normal” que a pandemia trouxe, estes problemas foram evidenciados. Para mudar essa situação é necessário que os gestores entendam a importância de investir em soluções de segurança, incentivos para incorporar boas práticas na cultura da empresa, além de capacitação dos usuários.

2.2 POLÍTICA DE SEGURANÇA DA INFORMAÇÃO

A Política de Segurança da Informação (PSI) pode ser definida como um documento que estipula requisitos, princípios, valores, orientações, compromissos e responsabilidades com relação ao padrão a ser seguido para a proteção de informações. Essa política é como um manual de procedimentos que busca trazer eficácia na segurança da informação e que aponta como os

recursos de TI de uma empresa devem ser utilizados e protegidos, visto que, o não estabelecimento de regras pode torná-la vulnerável e inconsistente. Esse estabelecimento de regras e normas de conduta tem o objetivo de reduzir o acontecimento de incidentes que possam causar furto, serviço indisponível ou perda de informações. Sendo assim, a PSI geralmente é construída com base nas necessidades da empresa e são aperfeiçoadas através da experiência do gestor (DANTAS, 2011).

Segundo a NBR ISO/IEC27002 (2013), o recomendado é que a Política de Segurança da Informação (PSI) seja revisada periodicamente, geralmente seis meses, de forma planejada, ou quando acontecerem mudanças significativas, para assegurar que a mesma continua adequada e eficaz.

2.3 LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS (LGPD)

Para a SERPRO (2020?), a Lei nº 13.709, de 14 de agosto de 2018, chamada de Lei Geral de Proteção de Dados Pessoais, entrou em vigência em agosto de 2020. A lei estabelece o que são dados pessoais recomendando e ressaltando o cuidado com dados categorizados como mais sensíveis. Nela também são estabelecidos o órgão fiscalizador para cumprimento da proteção dos dados, a Autoridade Nacional de Proteção de Dados Pessoais - ANPD, e as sanções administrativas que podem ser aplicadas pela mesma. A LGPD também determina que não existe diferença se a sede de uma empresa é do país ou do exterior, se existir manipulação dos dados de pessoas dentro do Brasil, as normas da lei devem ser cumpridas. É claro que o compartilhamento de dados com organizações internacionais também é previsto, porém dentro dos protocolos e exigências da LGPD.

De acordo com a Peduti Advogados (2021), seguir a LGPD traz benefícios, não somente evitando as multas das sanções, mas também garantindo mais segurança aos processos das instituições, protegendo os consumidores finais, ou seja, as pessoas. Ao buscar entrar em conformidade com a lei, as organizações fortalecem suas políticas de privacidade e se tornam mais preparadas para lidar com riscos de desastres e intrusões, e prevenindo ao final as multas e possíveis prejuízos por suspensão de sistemas e bases de dados irregulares. No tocante ao backup, é importante que as empresas saibam gerir seus dados, investindo em soluções e pessoas capacitadas para definir adequadamente o tempo de retenção das informações, para que as informações devidas sejam entregues ao cliente na hora que o mesmo forneça o consentimento para uso de seus dados. Além disso, é necessário avaliar o local onde os dados

serão armazenados, soluções em nuvem podem ser bem eficientes, mas devido à sua natureza de sistema distribuído, tornam difícil a gerência e identificação de onde as informações estão efetivamente, o que também deve ser de ciência do cliente.

Segundo Araujo (2020), vazamentos de informação podem abalar a integridade de qualquer negócio, afetando a imagem e credibilidade no mercado. Levando em conta também as multas e severas restrições que a LGPD prevê em casos de perda ou vazamento de dados. Isso denota mais ainda a importância do backup das informações no cotidiano das empresas, que agora vai além da disponibilidade e recuperação de desastres. Ao entrar em vigor, a LGPD exige que as empresas se fortaleçam em diversos pontos relacionados à segurança, transparência, privacidade, proteção e confidencialidade de dados. Um ponto central nesse tema é o registro das atividades de processamento e tratamento dos dados pessoais, ter um *Access Control List* bem definido, com os *logs* armazenados e protegidos, é o caminho para se adequar à lei. Dessa forma o backup se faz um item muito importante, pois os arquivos de *logs* deverão ter acessos restritos e estarem protegidos com cópias de segurança, através disso a TI de qualquer organização conseguirá gerenciar a classificação dos arquivos pessoais e o tempo de retenção.

Assim sendo, com a lei geral de proteção de dados em vigor, além de uma adequação em relação aos sistemas e dados “circulantes”, as empresas também devem estar de olho na gestão dos dados salvos. Atualmente, as empresas devem ser bem claras sobre onde os dados dos usuários são salvos e por quanto tempo eles serão mantidos, isso impacta em uma adaptação da política de backup nas instituições, fazendo com que o tempo de retenção dos dados, que muitas vezes pode ser algo difícil de calcular, seja uma informação prevista, exata e repassada ao dono dos dados. Além disso, a integridade dos dados, um dos pilares da segurança da informação, foi estritamente reforçada pela nova lei, fazendo com que as instituições sejam obrigadas a ter um controle refinado do acesso e registro de alterações dos dados. Isto torna o investimento em backups seguros, criptografia e boas práticas de senhas algo essencial. Arquivos de *logs* dos sistemas, antes mais utilizados na resolução de problemas, agora se tornam obrigatórios no grande leque de informações que o backup deve reter. A LGPD também garante que a exclusão e atualização dos dados por parte dos usuários seja facilitado, então os sistemas de backup que devem se tornar os mais indicados nos próximos anos são aqueles que permite a fácil gerência, no sentido de que os dados sejam facilmente encontrados e manuseados de forma segura.

2.4 BACKUP

Segundo o Dicio (2022), o termo Backup é da língua inglesa e significa cópia de segurança. Na informática essa palavra é usada para indicar que existem uma ou mais cópias de uma informação salvas em algum tipo de mídia. Quando um indivíduo entende que possui uma informação com um importante valor agregado, o mesmo deve optar pela realização de um backup de forma a proteger o determinado valor. Cópias de segurança não protegem somente da perda de informação. Os dados também podem ser corrompidos ao longo de seu ciclo de vida, devido ao transporte ou transferência da informação entre mídias. Também existe a possibilidade da corrupção dos dados devido à erros humanos, de software ou hardware. Esses erros podem acontecer em um ambiente doméstico, porém são muito mais frequentes dentro das organizações, justamente o lugar onde a informação é um ativo precioso.

Segundo Alencar e Schimiguel (2018), o propósito do backup é criar uma cópia do dado em uma mídia secundária. Esta cópia é armazenada e guardada para uso futuro caso o dado original tenha sido perdido, destruído, corrompido, atacado por vírus ou até mesmo sequestrado. Podemos citar como objetivos do backup: recuperação após desastres, arquivamento, backup operacional, recuperação após dados corrompidos, estar em conformidade (compliance).

O backup ou cópia de segurança serve para restaurar dados perdidos no caso de perda dos arquivos originais. Desse modo, ele pode ser compreendido como um processo de salvaguarda do Sistema de Tecnologia da Informação e Comunicação. Sendo assim, o backup surge para prevenir os equipamentos (hardwares) e banco de dados e arquivos (softwares) (CABRAL, 2015).

Para que exista uma estrutura gerencial de sucesso é necessário se criar uma gestão estratégica da informação. Em se tratando de negócios, as empresas precisam que suas mídias sejam armazenadas com maior proteção e segurança dentro da estrutura de TI. E é nesse processo de armazenamento e segurança dos dados que o backup é fundamental para assegurar a restauração dos dados caso sejam perdidos em operações mal realizadas ou por apagamentos acidentais até desastres com perda de 100% do arquivo original (JUNIOR, 2010).

Entretanto, o backup não é meramente uma cópia dos dados guardados em outro local para serem resgatados quando necessário. Segundo Faria (2020), os backups podem ser divididos em três principais níveis: completo, diferencial e incremental.

- **Completo:** Sempre fará a cópia de todos os diretórios e, independentemente de qualquer outra condição. Será sempre o primeiro backup de qualquer servidor, ainda

que um incremental ou diferencial for submetido (pois não há backups passados para comparar).

- Diferencial: faz o backup apenas dos arquivos modificados a partir do último backup full terminado, por isso normalmente tem um tamanho consideravelmente menor. Exemplo: um backup full ocorreu no sábado; um backup diferencial realizado na segunda só conterá os dados alterados ou criados na segunda; se na terça-feira for gravado outro backup diferencial, ele novamente vai gravar os arquivos alterados ou criados desde que se gravou o backup full, isto é, os arquivos de segunda e terça.
- Incremental: este faz backup de ainda menos arquivos do diferencial ao longo do dia, uma vez que apenas copia os arquivos que foram alterados desde o último backup, independentemente do seu nível. Em outras palavras, os backups incrementais copiam estritamente os arquivos que não estão presentes em qualquer backup anterior.

2.5 POLÍTICA DE BACKUP

Ao planejar o backup de dados é necessário que algumas especificidades sejam identificadas, para tanto é uma boa prática definir uma política de backup, que norteará os limites de quantidade de dados que podem ser retidos e periodicidade em que os backups são feitos.

A política de backup é um documento que retrata as condições básicas de backup, bem como os riscos e procedimentos que podem ser utilizados, sendo construída de forma adaptável para que o grupo técnico da empresa não fique preso a uma tecnologia específica. Esse tipo de política é uma prática recomendada para as corporações, principalmente quando se trata de backups, visto que não é possível estruturar um sistema de backup sem ter uma política. Pois, por exemplo, sem uma informação anuncia sobre tempos de retenção de backup não seria possível o administrador do sistema conhecer o tamanho necessário dos *storages*, robôs de fitas ou a quantidade de fitas necessárias para o armazenamento de backups para um determinado intervalo de tempo (FARIA, 2020).

Conforme Rodrigues (2017), é importante que exista uma política de backup adequada para a realidade da empresa ou instituição. Outro fator a ser considerado antes da estruturação da política de backup é definir o *Recovery Point Objective* (RPO) e *Recovery Time Objective* (RTO). Uma vez que o *Recovery Point Objective* (RPO) está relacionado a quantidade máxima

de dados que a empresa ou instituição pode ter uma perda de dados aceitável, no caso de um desastre, em um intervalo de tempo. Enquanto que o *Recovery Time Objective* determina o tempo concebido de espera pra que o sistema ou arquivos sejam recuperados e fiquem prontos para operar novamente, isso em caso de um desastre.

2.6 ESTRATÉGIAS DE BACKUP

Tentar manter organizadas as mídias de backup, sejam elas fitas, DVDs ou discos magnéticos, é um desafio. O backup não vale muito quando recuperar dados é uma tarefa árdua por conta de inconsistências e perdas em função da desorganização. Para resolver problemas como esses com o tempo foram disseminados métodos que podem ser chamados de estratégias de backups ou esquemas de rotação de backups. Para Mindsight (2017), existem três esquemas de rotação de backup de dados que surgiram ao longo dos anos são eles: FIFO; GFS; e o esquema Torre de Hanói. Eles podem ser descritos como abaixo.

2.6.1 FIFO

Esse método, também chamado de Round Robin, tem seu nome derivado da frase *First In, First Out*, ou Primeiro a Entrar, Último a Sair. Sobre esse método, é possível afirmar que:

Nosso primeiro esquema é o mais simples, e esse é seu melhor trunfo e sua maior desvantagem. No esquema Primeiro a Entrar, Primeiro a Sair, o diretor de TI deve primeiro escolher quantos dias de backups a empresa deseja armazenar - sejam 14, 50 ou 365 backups. O número real de backups no inventário depende inteiramente da preferência da empresa. No entanto, quando uma nova fita é adicionada, a fita mais antiga no inventário é removida, de forma que o número total de fitas no inventário permanece consistente. Como o nome explica, a primeira (ou mais antiga) fita a entrar é a primeira fita a sair. (Mindsight, 2017)

Essa é uma maneira de gerar um padrão de backup para reduzir prejuízos e controlar a entrada e saída das informações. Esse esquema de backup era mais utilizado nos sistemas legado de backup em fitas LTO, as fitas podem ser regravadas e por isso o esquema funcionava da seguinte forma, o operador escolhia a quantidade de dias que desejava reter os arquivos e geralmente essa quantidade de dias estava associada à quantidade de fitas, cada fita representando um dia de backup. Tomando por exemplo uma quantidade de 14 fitas, uma para cada dia de backup, era possível ter uma retenção de arquivos por 14 dias, sendo cada fita adicionada para cada backup diário realizado, ao término de 14 dias, todas as fitas seriam

gravadas. Para que fosse possível manter o tempo de retenção de 14 dias, a primeira fita inserida, referente ao primeiro dia de backup do ciclo, era retirada da fila e seria regravada no backup do 15º dia, mantendo sempre o mesmo número de fitas e tempo de retenção desejado.

2.6.2 GFS

Segundo Faria (2020), *Grandfather-father-son* (avô-pai-filho) é a estratégia de rotação de backup mais comum em todo o mundo. Baseia-se em backups escalonados, o que significa a configuração de diferentes tempos de retenção aplicada aos diferentes níveis de backup. O significado da sigla GFS é apenas uma alusão à execução de backups escalonados. Trata-se de um dos esquemas de rotação mais amigáveis porque baseia-se em fenômenos humanos corriqueiros: dias úteis das semanas e meses, que foram traduzidos para backups com diferentes tempos de retenção, e conseqüentemente *pools* (grupos de volume) distintas: diário, semanal e mensal (por exemplo). A regra básica para GFS é: para começar a reciclagem de volumes para um determinado nível (backups diários), é necessário ter finalizado com êxito um backup do nível mais alto (backups semanais). Se não houve incidentes de perda de dados, neste caso, provavelmente é seguro começar a sobrescrever volumes do backup diário da semana anterior (por exemplo: uma fita escrita segunda-feira pode ser sobrescrita na próxima segunda-feira).

Para Mindsight (2017), o GFS trabalha de maneira um pouco diferente do FIFO. Pois, ao invés de ter um número determinado de backups que serão substituídos um por um, terá “gerações” de backups. Desse modo, segue-se a analogia de que os backups mais recentes são os filhos ou a geração mais jovem. E conforme os filhos ficam mais velhos vão se tornando pais e criando seus próprios filhos, e assim sucessivamente. Em resumo, o “filho” é o backup de ontem, o “pai” é o backup da semana passada, e o “avô” é o backup do mês passado. Com o passar do tempo, esses backups ganham títulos mais altos ao passo que se afastam do ponto atual no tempo.

2.6.3 A Torre de Hanói

Segundo Abwoga (2020), a Torre de Hanói parece ser a maneira mais complexas de todas com relação aos métodos de backup comuns. Esse método tem o nome de um jogo de quebra-cabeça no qual o jogador move uma torre de discos por vez. Os discos geralmente são retirados da torre onde se encontram inicialmente para outra torre, normalmente a terceira,

respeitando a ordem de seus tamanhos, formando o formato de pirâmide. Assim, temos que cada mídia de backup equivale a um disco na torre e cada retirada do disco para uma torre diferente corresponde a um backup dessa mídia. O disco que fica no topo da pilha será mais movimentado, ao contrário do disco na parte inferior da torre. Desse modo, ainda conforme o autor, “a estratégia da Torre de Hanoi é um método bastante eficaz de backup a ser considerado em um caso de uso em que se precisa criar a situação de backup e recuperação mais longa possível com um número limitado de mídia de backup”.

No quadro 1 abaixo, vemos um exemplo do funcionamento da Torre de Hanói.

Quadro 1 - Ciclo da Torre de Hanói

CICLO DA TORRE DE HANÓI																
DIA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
MÍDIA DE BACKUP	A		A		A		A		A		A		A		A	
		B				B				B				B		
				C								C				
								D								D

Fonte: Compilação do autor, adaptado de Abwoga (2020).

No quadro 1, cada coluna do quadro representa um único dia. O aparecimento de letras (A, B, C e D) em cada um dos dias do ciclo representa um conjunto de mídia de backup. Tem-se que o backup feito na mídia de backup A no primeiro dia é a mesma mídia de backup que será substituída a cada dois dias; ou seja, no dia 1, dia 3, dia 5 e assim sucessivamente. O mesmo A mídia de backup B, por exemplo, será usada para backup após três dias a partir do segundo dia. De acordo com Abwoga (2020), “essa é uma boa maneira de manter baixo o número geral de mídia de backup e, ao mesmo tempo, manter a mídia de backup diária, semanal, mensal, semestral e anual, e garante que não haja sobreposições”.

2.7 SOLUÇÕES DE BACKUP

Para que os backups sejam efetivamente realizados, existem diversos programas que auxiliarão ao operador responsável por esta tarefa a executá-la. Portanto, assim como a política de backup será definida de acordo com as necessidades e objetivos da organização, a definição da aplicação a ser utilizada também será determinada de acordo com o padrão estabelecido pela empresa (SANTOS, 2021).

Existem várias maneiras de se fazer backup corporativo, porém elas podem se resumir em dois tipos, backups locais e backups remotos. Os locais comumente utilizam mídias como discos magnéticos, fitas ou armazenamento flash. É no backup local que se enquadram métodos com pouco ou quase nenhum gerenciamento como puramente cópias realizadas por um profissional de tecnologia da informação, mas também se enquadram as antigas soluções baseadas em fitas, servidores de backups, e *storages* NAS. Já os backups remotos são soluções mais modernas, baseadas em nuvem ou replicação de dados em serviços de terceiros. Algumas das soluções de backup corporativo, as mais bem-conceituadas atualmente, serão exploradas adiante.

2.7.1 Amanda Network Backup

Segundo Preston (2007), Amanda (*Advanced Maryland Automated Network Disk Archiver*), é o software de backup de código aberto mais conhecido. Amanda foi inicialmente desenvolvida na Universidade de Maryland em 1991 com o objetivo de proteger arquivos em um grande número de estações de trabalho de clientes com um único servidor de backup. James da Silva foi um de seus desenvolvedores originais. Ao longo dos anos, mais de 250 desenvolvedores contribuíram para o código Amanda e milhares de usuários forneceram testes e feedback resultando em um pacote estável e robusto. Em abril de 2006, estima-se que Amanda foi implantada em mais de 20.000 locais em todo o mundo. Originalmente Amanda foi implantada em produção principalmente por universidades, laboratórios técnicos e departamentos de pesquisa. Hoje com a ampla adoção do Linux na TI em geral, Amanda é encontrado em muitos outros lugares, especialmente onde o foco está em aplicativos implantados em LAMP (Linux, Apache, MySQL, PHP/Perl/Python).

2.7.2 ARCserve Backup

O ARCserve Backup é um sistema gerenciador de armazenamento que funciona como sistema distribuído, além de ser multiplataforma, ele é capaz de executar backups e recuperar dados em qualquer cliente da rede. Seu painel de gerenciamento dá ao gerenciador uma boa visibilidade dos procedimentos de forma simples, fazendo com que essa solução sirva tanto a pequenas quanto grandes organizações. Depois de implantado o sistema pode executar tarefas a partir do servidor para serem realizadas tanto nos clientes quanto nele mesmo. Também é possível executar tarefas de mesclagem, varreduras, migrações, cópia e comparações. No

ARCserve é apresentada uma lista geral de tarefas onde é possível monitorar o status das tarefas. Todas as informações de relações sobre os dados salvos são armazenadas em um banco de dados de fácil manuseio, onde também são guardadas informações sobre tarefas e sessões. A recuperação de dados é ágil, pois o sistema mantém um catálogo onde são armazenados metadados que auxiliam na busca dos arquivos que se deseja recuperar, não sendo necessária uma consulta demorada no banco de dados (RUSSELL, 2018).

2.7.3 *Storages NAS*

Segundo a Red Hat (2018), o *Network-Attached Storage*, ou armazenamento NAS, tem uma arquitetura baseada em arquivos e por isso facilita o acesso de dispositivos na mesma rede aos dados armazenados. Este tipo de armazenamento estabelece um ponto de acesso único ao armazenamento para as redes, com recursos incorporados de segurança, gerenciamento e tolerância a falhas. O NAS difere dos servidores de uso geral no software, neste tipo de armazenamento o software é implantado em um sistema operacional simples, que normalmente está incorporado ao hardware. Já os servidores de uso geral têm sistemas operacionais complexos que enviam e recebem milhares de solicitações por segundo, sendo que apenas uma fração dessas solicitações são relacionadas ao armazenamento.

2.7.4 *Azure Backup*

Desenvolvido pela Microsoft, essa é uma das soluções de backup de servidores na nuvem mais utilizadas na atualidade. Segundo Passos (2019), o Azure pode ser utilizado tanto para backup de máquinas virtuais hospedadas no próprio Azure, quanto de servidores locais, sejam eles virtuais ou físicos. Ele consiste em uma ferramenta de uso simples, com backup em nuvem para servidores, tendo o menor custo no mercado por GB. Contudo, a sua configuração é muito engessada sem permitir armazenamento externo do backup, como em um HD. Mesmo com o baixo custo, erra por não fornecer uma ferramenta de gerenciamento e monitoramento simples. O Azure Backup é uma boa opção para empresas que possuem um orçamento limitado ou precisam fazer backup de VMs no Azure ou de poucos servidores locais Windows. Porém se o ambiente que precisa de backup não é somente Windows essa solução passa a ser inviável.

2.7.5 *Acronis Cyber Protect Cloud*

Para Passos (2019), a ferramenta mais conhecida pelos profissionais técnicos de informática é a *Acronis*. A sua solução de backup é comercializada no modelo de licenciamento tradicional, tendo como ferramenta mais utilizada o *Acronis Cloud*. Visto que o seu diferencial está em oferecer uma plataforma WEB para gerenciamento dos seus backups, permitindo ao gestor o acesso a esta ferramenta de qualquer lugar que possua internet. Ainda segundo o autor, o *Acronis Cloud* concede a opção backup de máquinas físicas e virtuais com Windows e Linux, além de aplicações como *SQL Server*, *Active Directory*, *Exchange*, Office 365, entre outras. O que a torna uma excelente opção de ferramenta para backup na nuvem de servidores e estações. E ainda permite o armazenamento do backup em *storage* ou disco externo, podendo depois enviar para a nuvem. O seu grande diferencial é a ferramenta para gestão dos backups, que é muito simples de usar.

2.7.6 Amazon S3

O *Amazon Simple Storage Service*, também chamado de Amazon S3, é uma solução muito utilizada por pequenas e grandes organizações na atualidade, o serviço é um armazenamento de objetos em nuvem de alta escalabilidade, disponibilidade e segurança, além de favorecer bastante a performance de diversas aplicações que utilizam esta solução. Com o Amazon S3 é possível armazenar pequenas ou grandes quantidades de dados para diversas aplicabilidades como *data lakes*, aplicações nativas da nuvem e aplicações móveis. E como em outros serviços da Amazon, os dados estão sob várias camadas de segurança, o que dá aos clientes menos preocupação quando se trata da integridade dos dados. A Amazon oferece vários planos com classes de armazenamento econômicas e recursos de gerenciamento fáceis de usar, tornando o serviço mais acessível. No painel de gerenciamento é possível otimizar custos, manipular a organização dos dados e gerenciar o controle de acesso, fazendo com que o Amazon S3 seja bastante adaptável às variadas aplicações (AMAZON, 2006).

2.7.7 Veeam Backup & Replication

Segundo Passos (2019), assim como nas ferramentas gratuitas, a *Veeam* surge com foco na solução para backup e recuperação de ambientes virtualizados com *VMware* e *Hyper-V*. Essa ferramenta passou a ser adotada pelos profissionais de TI por oferecer um custo menor de aquisição em relação aos seus concorrentes, como a *ARCserve*, além de ser mais simples de

implantar e manter. A grande problemática sempre foi o backup de máquinas físicas, bem como sua compatibilidade com sistemas de backup legados em fita. Contudo, em suas versões mais recentes, essa compatibilidade foi adicionada e também oportunidade de backup de servidores físicos. Ainda para o autor, a Veeam “é uma excelente opção para empresas médias, que precisam de uma solução de backup com perfil mais corporativo, pois une um custo baixo de aquisição, com um processo de implantação mais simples”.

2.7.8 Google Drive *File Stream*

Segundo Doms (2020), o Google Drive *File Stream* é um serviço em nuvem do Google para computador, a solução está disponível para clientes que assinaram o GSuite. A partir da instalação todos os arquivos armazenados no Google Drive ficam acessíveis no computador, edições de documentos de texto que não são feitos inteiramente nos editores do Google ficam disponíveis para edição sem precisar abrir o Google Drive no navegador. O seu diferencial é poder utilizar os arquivos sob demanda, funcionando como *stream* de arquivos, o que poupa o armazenamento local do usuário. Caso o usuário deseje, também pode selecionar quais arquivos ficam disponíveis mesmo offline, para casos onde se está em um local sem internet e precisa-se de acesso aos dados. Isso pode ser feito ao selecionar o arquivo com o botão direito, encontrar a opção “Google Drive *File Stream*” e depois selecionar “Disponível off-line”. O aplicativo é basicamente uma extensão virtual de memória no PC, que abstrai o armazenamento online do Drive e torna os arquivos mais próximos do usuário, embora os dados não estando realmente ali, mantendo o mesmo nível de segurança em caso de falha do computador. Mesmo utilizando os arquivos localmente com seu editor de preferência, as alterações realizadas em arquivos como documentos, apresentações e planilhas, são salvas em tempo real. Caso estes arquivos sejam compartilhados, as alterações também são exibidas aos outros usuários que possuem acesso.

2.7.9 Bacula Community

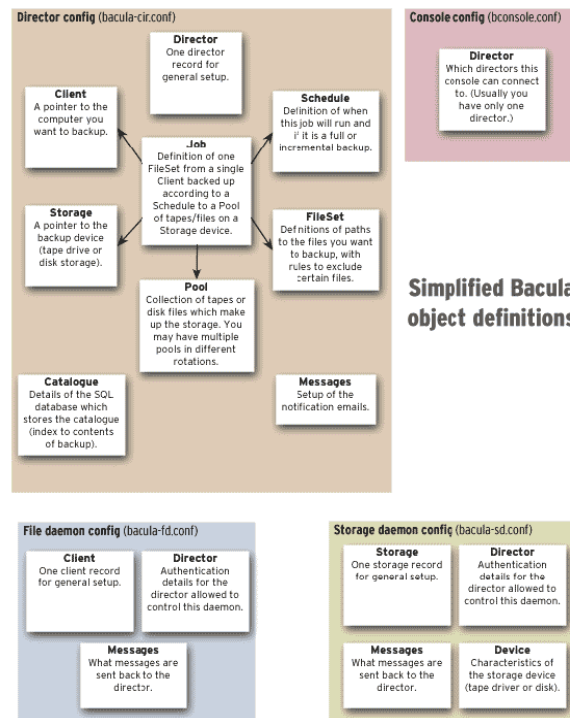
Segundo a bacula.lat (2008), grupo formado pelas empresas representantes do Bacula no Brasil e no restante da América Latina, o *Bacula* se trata de um “software que permite administrar backup, restauração e verificação dos dados de computadores em uma rede de sistemas mistos”. Devido a sua flexibilidade, é considerado um framework de backup, o que

permite a sua implantação em diferentes estruturas corporativas. O *Bacula* nada mais é do que um programa de backup em rede. O seu funcionamento é realizado através de seus vários módulos de serviços, podendo classificar em cinco principais, que são: *Director Daemon*, *Console Manager*, *File Daemon*, *Storage Daemon* e *Catalogue*. Eles podem ser descritos como abaixo.

- *Director Daemon*: É responsável pela administração de todos os processos de backup, restauração, verificação e arquivamento. O Administrador de Sistema usa o *Director Daemon* para efetuar agendamentos de backup e para recuperar arquivos.
- *Console Manager*: É um programa que ajuda o administrador ou o usuário a se comunicar com o *Director Daemon*, pode ser executado em qualquer computador da rede e em sistemas operacionais diferentes.
- *File Daemon*: É o software que é instalado na máquina que vai ser protegida pelo backup, ou seja, ele vai ser responsável por enviar os arquivos solicitados pelo *Director Daemon* pela rede. Ele também é responsável em administrar a gravação dos arquivos de restauração comandados pelo *Director Daemon*.
- *Storage Daemon*: É responsável por administrar a gravação e restauração dos dados e atributos dos backups fisicamente em mídias apropriadas, essas podem ser volume de dados gravados diretamente no disco rígido ou alguma mídia removível (Fita DAT, DVD, CD, etc.)
- *Catalogue*: Também chamado de Serviço de Catálogo é o programa responsável por manter uma indexação de todos os arquivos que são armazenados no backup e gerar uma base de dados dos volumes gerenciados pelo *Director Daemon*. O *Catalogue* agiliza a busca de um arquivo no backup na hora que o administrador de sistema necessita efetuar uma restauração, como ele mantém uma base de indexação dos arquivos gravados, a busca por um arquivo no meio dos volumes é mais rápida.

Na figura 1 podemos visualizar o esquema desses módulos.

Figura 1 - Esquema de módulos do *Bacula* por arquivos de configuração



Fonte: Bacula.lat (2008).

Na figura 1 é possível perceber como a estrutura de arquivos de configuração do *Bacula* está disposta. Os módulos *Director* e *Catalogue* são definidos em um mesmo arquivo, neste arquivo também constam por padrão as definições dos *Jobs*, *Clients*, *Schedules*, *FileSets*, *Pools* e um apontamento para a configuração do módulo do *Storage*. Também é aqui onde a entidade que trata das mensagens do serviço é configurada, podendo configurar tanto as mensagens que são enviadas para o *bconsole* quanto o envio de e-mails com as mensagens que o usuário preferir. Note que a entidade *Job* está associada às outras por meio de setas, isso nada mais é que uma denotação das entidades que precisam estar definidas para serem apontadas na criação de um *Job*. O *Console Manager* da aplicação, também chamado de *bconsole*, possui um arquivo de configuração próprio, que também deve ficar na mesma máquina onde o diretor está instalado. O arquivo de configuração do *File Daemon* se localiza nas máquinas clientes, neste arquivo conta as definições que apontam para o *director* e para o módulo de mensagens. Já o arquivo onde o *Storage Daemon* está definido fica na máquina onde serão guardados os volumes de backup, neste arquivo são definidos apontamentos para os módulos do *director* e de mensagens, como no *File Daemon*, com a diferença que também constam definições sobre o tipo de mídia que será utilizada nos backups.

3 PROJETO

3.1 METODOLOGIA APLICADA NA ANÁLISE

Após levantamento bibliográfico sobre a temática de backup e sua correlação com demais aspectos, o projeto segue dando início a instalação e configuração das soluções. A definição das variáveis que influenciam a implantação e operação da solução é o pilar principal do estudo. Portanto a medida que as soluções foram implantadas, foi feita em paralelo uma pesquisa básica estratégica, descritiva e aplicada, com uma abordagem quali-quantitativa. Para definir as variáveis foi utilizado o método hipotético-dedutivo, estabelecendo heurísticas para definir os pontos de influência que serão chamados de variáveis. Esses pontos chave avaliam e ajudam a denotar o sucesso de implantação e operação das duas soluções, bem como sua viabilidade, levando em conta o ambiente da organização. As variáveis observadas giram em torno dos seguintes temas:

- Custo financeiro para aquisição ou implantação da solução;
- Complexidade de implantação e operação;
- Recursos de infraestrutura;
- Segurança - Sensibilidade dos dados;
- Disponibilidade dos arquivos;
- Sincronização dos Dados;
- Colaboração no manuseio dos dados.

O passo seguinte vem com a classificação das variáveis, a equipe de TI deve levar em conta a estrutura atual para o manutenção dos dados e os custos que a organização pode arcar para uma possível mudança ou implantação de uma nova solução de backup. Possivelmente, em alguns casos, será necessário ter informações sobre orçamento para aquisições de TI. A equipe de TI também deve traçar um perfil de seus usuários, levando em conta a frequência com que os dados são alterados pelos colaboradores, a colaboração na produção conjunta das informações e o volume de dados gerado. Com as variáveis identificadas e classificadas e levando em conta as funcionalidades das duas soluções o resultado será um plano geral onde é possível identificar de forma fácil qual proposta se adequa melhor ao ciclo de vida dos dados.

3.2 ANÁLISE PARA IMPLANTAÇÃO E USO DAS SOLUÇÕES

Essa análise foi dividida em duas etapas como descrito abaixo.

3.2.1 Estrutura existente do IFPE *Campus* Palmares

O campus conta com um servidor físico alocado em um rack, como pode ser visto na figura 2, nele está instalado um software *hipervisor* que controla as diversas máquinas virtuais que abrigam os vários serviços de rede do instituto. Para implantação do *Bacula Community*, foi criada uma máquina virtual afim de ser implantado um SO linux, requisito para instalação do *Bacula*. O SO escolhido foi um *Debian 9*, também se fez necessária a instalação de um SGBD, entre os recomendados na documentação do serviço, foi escolhido o *MySQL*. A versão escolhida do *Bacula* foi a 9.6, sendo esta estável e com mais documentação no momento.

Figura 2 - Rack principal do IFPE Palmares



Fonte: Compilação do autor.

Para persistência dos volumes de backup gerados pelo *Bacula* contamos ainda com outra máquina que pode ser vista na figura 3, esta *workstation* é utilizada somente para este fim e fica em uma sala separada, seguindo as diretrizes de segurança contra possíveis desastres, essa medida faz parte da PSI sistêmica do IFPE. A instituição também conta com uma parceria com a Google, o que lhe possibilita uma licença do *G Suite for Education*, grade de serviços em nuvem da Google onde podemos encontrar o *Google Drive File Stream*.

Figura 3 - *Workstation* onde são alocados os volumes do *Bacula*



Fonte: Compilação do autor.

3.2.2 Custos e requisitos iniciais

O fato do IFPE já ter indefinidamente o *G Suite for Education*, conta pontos positivos para a implementação da ferramenta já que isso torna o custo significativamente mais baixo que o normal, onde normalmente são pagas assinaturas por usuário/ano. Para outras situações o cálculo do custo de aquisição do *Google Drive File Stream* pode ser feito utilizando os valores disponibilizados no site oficial da ferramenta.

Ao analisar superficialmente a estrutura da organização já é possível identificar alguns parâmetros que servem para aferir a melhor implantação. Enquanto a análise é focada no *Bacula*, deve-se considerar se a estrutura necessária para implantação dessa solução (servidores físicos, equipe especializada, quantidade disponível de armazenamento) é pré-existente, caso não seja, o primeiro passo é calcular o valor gasto na compra de equipamentos necessários (*Bacula*).

O IFPE Palmares já contava com o maquinário de rede necessário, possuía uma equipe técnica que buscou, através de cursos gratuitos, a especialização na ferramenta e, baseado no cálculo de uso de disco médio de cada usuário, desconsiderando o espaço ocupado pelo SO nos desktops, foi verificado que o armazenamento disponível no servidor seria o suficiente para realizar backup dos arquivos.

Esse cálculo não é tão simples. As duas soluções trabalham com versionamento e restauração de arquivos. Porém, enquanto na ferramenta do Google os backups são feitos geralmente de forma automática e a restauração de arquivos é bastante intuitiva, no *Bacula* eles devem ser bem definidos nas estruturas de *pool*, *filesset* e *schedule*. A retenção de arquivos, que é o tempo em que é viável manter várias versões dos arquivos em relação ao armazenamento disponível, deve ser levada em conta. Esse tempo pode variar bastante em função da frequência com que esses arquivos são modificados, já que manter várias versões por um mês, custa mais armazenamento que manter uma única versão do mesmo arquivo por vários meses. Calcular quanto de armazenamento seria necessário adquirir caso o já existente não fosse o suficiente é inevitável para finalizar o cálculo inicial de custo da implantação do *Bacula*.

É praticamente impossível chegar em um cálculo ideal, e por isso, todos esses valores devem ser estimados pela equipe técnica. Muitas vezes isso se torna algo complexo. Por exemplo, o servidor de arquivos do IFPE é relativamente novo e ainda não está sendo utilizado plenamente por todos os setores, é nítido que com o tempo, o espaço ocupado no *File Server* iria aumentar bastante, por isso foi feita a abordagem de considerar o espaço ocupado médio nas máquinas desktop dos funcionários. Para considerarmos a taxa de modificação dos arquivos foi aferida a média do armazenamento em cada desktop em intervalos regulares de dois dias. Quando a variação calculada mostrava um aumento no tamanho ocupado em disco, a mesma era considerada para cálculo médio de espaço/dia a ser acrescentado ao disco, essa métrica foi aferida por duas semanas. Outro ponto importante é a previsão de contratação de novos funcionários, no caso do IFPE já é algo planejado, isso pode influenciar bastante tanto a longo quanto a curto prazo, tendo em vista o armazenamento imediato e as várias versões com o tempo, portanto isso foi considerado de forma empírica levando em conta a duplicação do quadro de funcionários.

Ao levar esses parâmetros em consideração, a equipe técnica calculou que, com o fluxo de dados atual, seria possível manter um tempo de retenção de dois anos, até que o disco utilizado para backups fosse preenchido próximo aos 90% da capacidade, diante disso, considerando os possíveis aumentos da quantidade e frequência de modificação dos dados no

futuro, foi decidido elaborar um agendamento de backups para três meses. Tempo bastante razoável para recuperação de arquivos que possam ter se extraviado. Não sendo necessário adquirir mais armazenamento.

Até o momento, para a solução do Google, já foi possível identificar o parâmetro de custo de assinatura. Já para o *Bacula* identificou-se os parâmetros: custo de infraestrutura, custo de pessoal especializado para implantação e manutenção, complexidade (a ser considerado ou não pela equipe técnica). Ao tomar nota que a complexidade aqui entra como fator em uma situação como a do IFPE Palmares, onde a equipe técnica dispõe das duas soluções, sendo possível implantá-las sem custos financeiros adicionais.

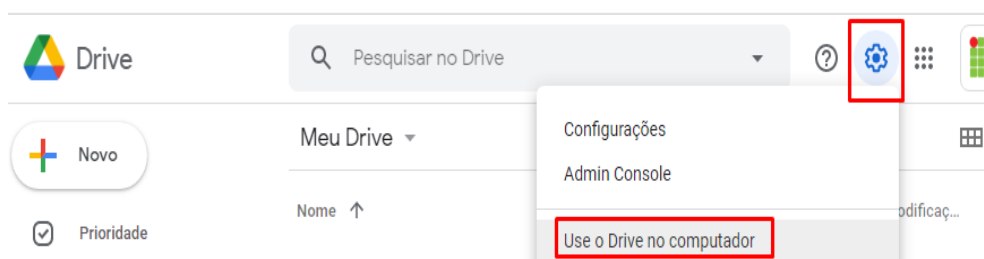
3.3 INSTALAÇÃO E IMPLANTAÇÃO

A seguir veremos quatro tópicos referentes a instalação e implantação.

3.3.1 Instalação do Google Drive *File Stream*

O *G Suite* do IFPE foi implantado e é administrado pela reitoria, aos campi foi dada a administração de um subdomínio, no caso de palmares registrado como “palmares.ifpe.edu.br”. Ao efetuar a criação de um e-mail neste domínio, o usuário passa a ter acesso a várias ferramentas do *G Suite*, entre elas o *Drive File Stream*. A instalação do mesmo pode ser efetuada pelo próprio usuário em qualquer máquina que ele tenha permissão administrativa, basta acessar o Google Drive em um navegador, clicar na engrenagem de opções e depois em “Use o Drive no Computador”, como destacado na figura 4.

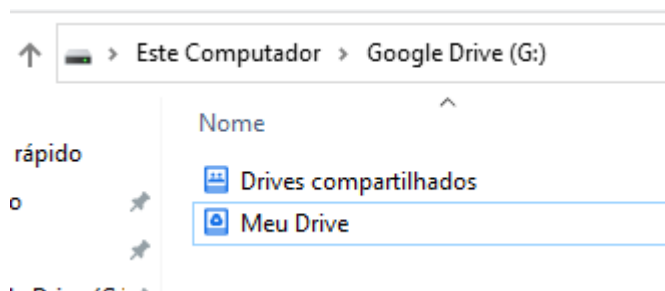
Figura 4 - Passos para download do Drive *File Stream*



Fonte: Compilação do autor.

Feito o download, ao executar o aplicativo de instalação uma janela para autenticação é aberta, inseridas as credenciais, o programa criará um disco virtual, que pode ser visto ao acessar o “Meu Computador” em distribuições Windows, contendo uma pasta para os arquivos do usuário e outra para exibir os drivers compartilhados, como pode ser visto na figura 5.

Figura 5 - Drivers virtuais criados pelo aplicativo após instalação



Fonte: Compilação do autor.

3.3.2 Funcionalidades e vantagens do Google Drive *File Stream*

A praticidade que a ferramenta traz para usuários domésticos é muito grande, porém aqui só serão destacadas funcionalidades e vantagens em um ambiente corporativo.

No Google Drive, o custo de armazenamento não é duplicado. Fica um tanto implícito, mas é bom destacar que na outra solução além de precisarmos de um armazenamento para o servidor de arquivos, o armazenamento dos backups é feito de forma separada e, se tratando de tamanho, precisa ser próximo ao do servidor de arquivos. No caso do IFPE Palmares isso não tem muito peso, pois as duas soluções já estão custeadas, mas para outras situações essa variável deve ser levada em consideração.

A complexidade de implantação é ínfima, somente precisando da ativação de uma conta de gerência do domínio, demandando mais trabalho de configuração e liberação das demais ferramentas do conjunto do *G Suite*. Nesse aspecto essa solução é bem mais vantajosa. A facilidade não é aparente somente para a equipe técnica, os backups automáticos facilitam a vida do usuário que sempre terá ao seu alcance a restauração de qualquer minimamente diferente versão. A acessibilidade dos arquivos também é um ponto importante aqui. Os arquivos estão disponíveis em qualquer lugar com internet. Como a edição é em nuvem, os documentos também são colaborativos, fazendo com que vários usuários possam trabalhar no mesmo arquivo. As alterações aqui são sincronizadas com os demais usuários e salvas em tempo real. Além disso, cada edição é associada a um usuário, sendo possível restaurar uma

versão de um trecho editado por um usuário enquanto as demais atualizações feitas por outros ficam intactas. Resumindo, em termos de versionamento e colaboração a ferramenta do Google se destaca bastante.

O problema começa quando tratamos de documentos sigilosos, de fato mesmo o IFPE sendo uma organização pública e transparente, existem arquivos que devem ser mantidos em sigilo, como em qualquer outra empresa. Como a guarda dos arquivos é feita por terceiros, e muitas vezes vazamentos de dados não são divulgados pela empresa que cometeu a falha, alguns arquivos simplesmente não poderiam ser guardados na nuvem.

Diante disso, em um debate no Superior Tribunal Militar (STM), o secretário de fiscalização de TI do TCU, Márcio Rodrigo Braz, fala sobre a análise que deve ser feita para utilização do serviço de nuvem no âmbito público federal, e diz:

Por isso é fundamental que seja feita a classificação da informação, a criticidade dos arquivos que estão sendo colocados, seja na nuvem, seja transportados por uma rede pública ou privada. É preciso conhecer a sensibilidade dessas informações, para aí sim tomar as medidas e controles adequados (STM, 2019).

De acordo com Brasil (2016), o manual de boas práticas de segurança do ministério do planejamento também trata desse assunto, nele existe a orientação de que caso seja contratado serviço de armazenamento em nuvem, este trata do armazenamento dos arquivos da contratante somente em solo brasileiro. A narrativa é realmente séria, para se ter ideia, as primeiras edições do projeto de Lei 2.126/2011, conforme Brasil (2011), que tratava sobre o marco civil da internet, houve tentativa de fazer com que todas as empresas brasileiras utilizassem somente serviços onde os *data centers* fossem localizados no Brasil, a fim de evitar espionagem internacional, porém com as pressões e polêmicas em torno desse projeto, esse texto foi retirado.

Ao se levar em conta as orientações e normativas do serviço público, ficou claro para a equipe de TI do *campus* Palmares que seria melhor não optar por salvar arquivos sensíveis no Google Drive, já que o mesmo não apresentava a opção de nuvem híbrida ou alocação exclusiva em servidores nacionais, requisitos ideais para seguir as boas práticas estabelecidas. O próprio setor de tecnologia da informação mantém arquivos de configuração de serviços, bancos de senha e arquivos de backups das bases de dados dos variados serviços de rede. Além disso, nenhum desses arquivos se caracteriza por necessitarem de disponibilidade ou serem de cunho colaborativo. Ficou claro o limite do Google Drive para uso no próprio setor, um armazenamento local, onde a segurança física e lógica poderia ser monitorada de perto e

ativamente se fez necessário. Isso levou a identificar a sensibilidade dos dados como sendo uma variável na escolha entre as duas soluções.

3.3.3 Instalação do *Bacula Community*

A instalação do *Bacula* se baseou em um guia de instalação do *bacula.org*. O script adaptado pode ser encontrado no Anexo A. Como dito anteriormente, o SO Linux escolhido para comportar o *Bacula* foi o *Debian*, uma distribuição muito bem-conceituada, estável e com uma comunidade ativa para dar suporte. O problema do *Debian* é que como ele tem versões estáveis muito duráveis, seus pacotes não atualizam com muita frequência. Portanto, para evitar uma instalação do *Bacula* que com o tempo ficará desatualizada, foi utilizada uma instalação com os pacotes do *bacula.org*, o site dá a opção de criar um repositório pessoal, necessitando de um pequeno cadastro. Para poder efetuar o download por pacotes, é preciso registrar-se no site gratuitamente e preencher seu nome, sobrenome e e-mail, no seguinte endereço: <https://www.bacula.org/bacula-binary-package-download/>. Ao final do registro será encaminhada para o e-mail cadastrado uma chave alfanumérica que deverá ser informada durante a execução de alguns comandos do script disponibilizado.

O *Bacula* funciona com três módulos principais, são eles: *director*, *file daemon* e *storage daemon*. Basicamente pode-se resumir o *director* como um controlador que mantém a comunicação e gerenciamento dos *clients* (*file daemons*) com o gestor de armazenamento (*storage daemon*). Como a estrutura de backup do *Bacula* do IFPE Palmares foi pensada levando em conta recuperação de desastres, a instalação do *director* foi feita em uma máquina virtual em um servidor que executa uma versão do *VMware Esxi*, enquanto que o *storage daemon* foi instalado em uma *Workstation* que fica localizada em outra sala. Dessa forma separando o servidor de arquivos do armazenamento de backups. É necessário ressaltar que em condições ideais essa separação deveria ser maior, pois qualquer desastre que comprometa toda a estrutura do prédio onde está alocada a sede provisória do IFPE Palmares, ainda afetaria as duas máquinas em questão. No caso de outra instituição que possui vários prédios ou filiais, o distanciamento ideal seria possível e é fortemente recomendado de acordo com as boas práticas de segurança da informação. Ao longo do script de instalação, a equipe técnica encontrou problemas que levaram a uma série de adaptações e correções do mesmo.

3.3.4 Funcionalidades e vantagens do *Bacula Community*

Os desenvolvedores do *Bacula* costumam dizer em seus artigos e fóruns que o *Bacula* será sempre necessário. Isso porque ele conta com as funcionalidades essenciais para qualquer organização que precisa de uma ferramenta para recuperação de desastres em caso de perda de dados. O sistema ainda por cima não custa nada e é *Open Source* permitindo personalizações pelos usuários para se adequar às especificidades. A comunidade da grande maioria das ferramentas de código aberto é bastante solícita e com o *Bacula* não é diferente, o que dá a equipe de TI um suporte com identificação e correção de erros de forma rápida, além de documentação atualizada de forma constante.

O *Bacula* conta com suporte a algoritmos de compressão livres LZO e GZIP, além de poder ser utilizado com bancos de dados gratuitos como o *MySQL* e *SQLite*. A ferramenta pode ser operada por linha de comando e também conta com uma interface gráfica padronizada, que suporta multiusuários, o que permite um bom gerenciamento através de uma central de funções. A portabilidade aqui não é um problema, isso porque existem módulos construídos para integrar uma vasta quantidade de sistemas operacionais do mercado.

Embora envolva um pouco de complexidade, é possível também a personalização de backups para poder se adequar a qualquer política de backup que o usuário necessite, além de poder construir várias, elas podem funcionar ao mesmo tempo entre os diferentes clientes conectados ao servidor.

Uma função que vale a pena ser mencionada é a possibilidade de executar scripts antes e depois dos *jobs* de *backup/restore*, tanto na parte do cliente quanto no servidor. Isso é bastante útil em diversas situações, pois dá flexibilidade para automatizar ações que possam ser necessárias no ambiente antes de um backup, o que muitas vezes ocorre em máquinas clientes que são também servidores com algum serviço rodando. Por último, é importante dizer que o *Bacula* é 100% compatível com a política de backup GFS, amplamente utilizada no IFPE.

3.4 IDENTIFICAÇÃO DE VARIÁVEIS

A identificação das variáveis foi um processo em paralelo com a implantação e o uso das duas soluções. Elas foram selecionadas levando em consideração as várias situações de infraestrutura, recursos financeiros, tipos de dados, complexidade para a equipe de implantação e *workflow* dos usuários nos *endpoints*. Algumas variáveis podem estar mais associadas para

uma solução que para outra, mas servem para verificar os requisitos básicos e viabilidade de implantação e uso das mesmas. Abaixo as variáveis serão categorizadas e subdivididas para depois serem compiladas em um questionário para apoio a decisão.

3.4.1 Infraestrutura

A infraestrutura está mais relacionada à implantação do *Bacula Community*, isso porque ele precisará ser hospedado na instituição, diferente do Drive que fica nos servidores da Google. Aqui deverão ser observados os requisitos mínimos de Hardware para a operação do *Bacula*. Isto porque os requisitos de Software são acessíveis a qualquer um, já que são gratuitos.

Nesta categoria vamos analisar três variáveis descritas abaixo.

3.4.1.1 Complexidade no cálculo de armazenamento – *Bacula*

Importante observar que o custo crescerá proporcionalmente com o tamanho do armazenamento necessário, por isso o cálculo de quanto de armazenamento será necessário vem antes. Embora o *Bacula* conte com poderosos algoritmos de compactação, o tempo de retenção dos backups faz com que seja necessário armazenar várias versões de um dado. As políticas de backup também são flexíveis, então caso os recursos para aquisição de armazenamento sejam insuficientes, o que pode ser feito é diminuir o tempo de retenção dos arquivos e a periodicidade dos backups. Isto, claro, respeitando os limites para suprir as recomendações da PSI instituída.

O cálculo para se chegar no armazenamento ideal é complicado e muitos ajustes podem ser feitos ao longo dos primeiros meses de implantação. Caso os arquivos que são salvos sejam modificados diretamente por usuários finais, até mesmo o *workflow* dos mesmos pode influenciar na periodicidade dos backups e armazenamento necessário. Portanto, uma “folga” no dimensionamento do armazenamento é extremamente recomendada. No início do capítulo 4 foi apresentado o cálculo feito pela equipe de TI do IFPE *Campus* Palmares, por ser algo variável e a critério da especificidade de cada instituição, não será apresentado um cálculo padrão para todos. Após a definição da quantidade de armazenamento já se pode calcular os demais custos da infraestrutura.

3.4.1.2 Custo de infraestrutura – *Bacula*

No IFPE *Campus* Palmares já existia uma infra com um *rack* portando dois servidores com virtualização hospedando vários serviços. Isto é o equipamento padrão nos demais *Campi*, porém algumas instituições podem não ter a mesma estrutura. Por isso, este é um importante fator a ser considerado.

Para a operação da solução, será necessário um servidor para hospedar o serviço, podendo já contar com armazenamento, ou um *Storage* separado atuando com o servidor. Sistemas operacionais e SGBD's que atuam em conjunto com o *Bacula* possuem alternativas de licença gratuita e por isso seu custo não precisa ser calculado aqui. Neste caso, talvez seja melhor optar pelo Google Drive já que existe pouca ou quase nenhuma complexidade na implantação e operação.

3.4.1.3 Complexidade de instalação e operação de hardware e software - *Bacula*

A implantação e operação do *Bacula* por parte dos administradores do sistema requer conhecimento em gerenciamento de servidores Linux, redes de computadores, banco de dados SQL e políticas de Backup. Além de bastante tempo gasto para entender o processo de instalação, configuração e operação do *Bacula*. Caso a equipe não se sinta apta ou não tenha tempo/vontade para se capacitar nesses assuntos isso vai impactar na escolha dessa solução.

3.4.2 Sensibilidade dos dados

Existem dados dentro da organização que podem ser restritos entre setores ou para o público. Dados financeiros, informações pessoais dos empregados, planejamentos da gestão, arquivos que podem ser usados para enfraquecer a segurança dos sistemas, todos esses são dados sensíveis. Para casos assim, soluções como o Google Drive não são bem vistas. Isto porque as boas práticas da segurança da informação recomendam que a guarda das informações não seja feita por terceiros.

No contexto atual existe bastante preocupação sobre quem guarda nossos dados, já que recentemente vieram à tona situações inacreditáveis do quão prejudicial os dados podem ser nas mãos de pessoas mal-intencionadas, como o caso do escândalo de dados do Facebook – *Cambridge Analytica*. Nesse contexto, alguns dados do IFPE, mesmo sendo somente uma

instituição de ensino pública, foram identificados como sensíveis. Arquivos com dados pessoais dos funcionários que constam no RH, arquivos de configurações de nossos servidores que são salvos para no caso de uma recuperação de desastre, entre outros, foram classificados dessa forma. O que fez com que a sensibilidade dos dados se tornasse uma variável importante.

3.4.3 Disponibilidade dos arquivos

Aqui uma pergunta importante é feita: Os arquivos precisam estar acessíveis aos usuários o tempo todo e de qualquer lugar? Alguns funcionários do IFPE às vezes precisam trabalhar em suas casas, como por exemplo os gestores e professores. Isso se tornou o padrão para todos na pandemia. O *Bacula* atua com um servidor de arquivos, os dados mantidos por este servidor ficam disponíveis somente na intranet do *campus* por questões de segurança. Também não é permitido aos funcionários acesso remoto de qualquer tipo para suas estações de trabalho. Exceto pela equipe de TI que mantém um gerenciamento ativo de vários serviços de rede e precisam de uma VPN. Isso faz com que o *Bacula* não seja uma solução que entrega uma grande disponibilidade de arquivos. O que fez com que alguns dados fossem migrados para o drive e outros, mais sensíveis, tivessem que ser acessados com uma visita ao *campus*. Isso gerou o entendimento que entre as duas soluções, em ambientes de empresas que precisam de alta disponibilidade dos arquivos, a segurança deveria ser um pouco diminuída. Seja pelo fato de liberar aos usuários o acesso remoto para ter uma forma de acessar os dados no servidor de arquivos, ou ter que utilizar o Google Drive mesmo com dados mais sensíveis, confiando na guarda de terceiros.

Vale lembrar que dados como arquivos de configuração de servidores e sites não são editados diretamente por usuários e sim pelos próprios serviços, fazendo com que seu backup no Google Drive muitas vezes seja manual, já que o mesmo possui suporte limitado em multiplataformas, tornando para esses casos o *Bacula* mais viável.

3.4.4 Sincronização

Alguns arquivos como documentos de texto e planilha são editados constantemente ao longo de um dia de trabalho, muitas vezes até por mais de um usuário. Alguns deles são bastante complexos e caso sejam corrompidos causaria grande impacto nos setores que trabalham com esses dados. Esses arquivos precisam ser sincronizados com seus backups em um tempo muito

curto. Servidores de backups gratuitos, como o *Bacula*, geralmente não possuem uma função de backup automatizado a cada alteração como acontece com o Google Drive.

O *Bacula* atua com backups agendados, sendo esse agendamento definido pelos administradores do sistema e dependendo do volume de dados e armazenamento da instituição, eles não podem executar muitos backups ao longo do dia. Em casos onde esse tipo de arquivo é de extremo valor, perder uma versão de um curto período atrás impactaria em muito retrabalho, portanto soluções como o Google Drive seria mais indicado, levando em conta claro a sensibilidade dos dados. Essa é uma variável importante e que tem um peso considerável já que se deve pensar no balanceamento entre praticidade e segurança.

3.4.5 Colaboração

A colaboração também está associada ao mesmo tipo de arquivo que o anterior, no entanto, levando em conta o paradigma de que muitas vezes para se ter uma produtividade aceitável arquivos de texto precisam ser editados por mais de uma pessoa ao mesmo tempo, com retornos de suas edições para todos, em tempo real. Quando esse paradigma é indispensável, infelizmente o *Bacula* sai de cena, não tendo outra opção além de *suites* como Google Drive e OneDrive. Muitas reuniões de planejamento dentro IFPE são acompanhadas por edições de arquivos por vários usuários simultaneamente. Isso é algo natural e fez com que a equipe de TI percebesse que o Google Drive seria indispensável à medida que os usuários foram se acostumando com essa funcionalidade.

3.5 COMPARATIVO ENTRE AS DUAS SOLUÇÕES EM RELAÇÃO AS VARIÁVEIS

Após definir os pontos de influência ou variáveis que afetam a decisão entre as soluções, foi elaborado um comparativo entre os dois sistemas, a comparação serviu para melhor visualização das funcionalidades que as duas soluções dispõem, além de ajudar na elaboração do questionário do roteiro, fazendo com que a construção das perguntas focassem na simplicidade, objetivando obter respostas com dados úteis, que pudessem denotar as especificidades do ambiente em que as soluções seriam implantadas. Um esboço desse comparativo pode ser observado abaixo no quadro 2.

Quadro 2 – Comparativo entre as Soluções

	DRIVE FILE STREAM	BACULA
Custo de Licença	Consideravelmente caro , seu custo varia por quantidade de usuários, tamanho do armazenamento e tempo.	Sem custo . O sistema é open source
Necessidade de infraestrutura de hardware	Nenhuma . Os dados são armazenados nos servidores da google. Não existe necessidade de adquirir equipamentos nem de manutenção.	Total . O sistema requer que a organização tenha ou adquira o maquinário necessário para a operação do serviço.
Complexidade para calcular armazenamento necessário	Nenhuma . A solução se encarrega de calcular tanto o armazenamento dos arquivos quanto o necessário para recuperação de dados. Esse tamanho de armazenamento pode ser alterado facilmente, porém tem grande impacto no custo.	Alta . Os responsáveis pela implantação do serviço devem calcular o tamanho do armazenamento e dispor ou adquirir o maquinário necessário. O cálculo é complexo e o tamanho ideal é ajustado após um bom tempo de operação do sistema.
Complexidade de instalação e operação	Baixa . Por ser um sistema em nuvem sua instalação só requer pouca configuração através da central de gerenciamento do GSuite. Caso o operador tenha problemas também pode recorrer a documentação atualizada e suporte remoto.	Alta . O responsável pela instalação precisa ter conhecimentos em redes de computadores e sistemas linux, a configuração do software é todo feito por linha de comando e edição de módulos de texto. Em caso de problemas o operador precisa recorrer a fóruns e documentações geralmente desatualizadas.
Segurança de dados sensíveis	Baixa . Por ser um sistema em nuvem os dados ficam sob a guarda de terceiros, dependendo da sensibilidade dos arquivos o vazamento, perda ou uso indevido podem trazer sérios prejuízos a uma organização. Brechas de segurança e intrusões são mais difíceis de detectar já que o serviço não é hospedado na infraestrutura da organização.	Alta . A instituição faz a guarda de seus dados. A segurança da informação será garantida contanto que os responsáveis sigam as boas práticas de segurança mais atualizadas, fortalecendo as barreiras tanto no meio digital quanto no físico.

Disponibilidade dos dados	Alta. Dados hospedados em nuvem podem ser acessíveis de qualquer lugar que tenha internet.	Baixa. Como a guarda dos dados fica a cargo da organização, disponibilizar os dados a qualquer tempo significa aumentar a possibilidade de brechas de segurança. Por este motivo geralmente os dados podem ser acessados somente da intranet ou via VPN para casos específicos.
Sincronização em tempo real	Possui. No drive file stream os arquivos são salvos a medida que são alterados, a velocidade de sincronização depende somente da velocidade da internet.	Não possui. Os dados nesta solução são salvos em tempos predefinidos, mesmo que seja diminuído o intervalo entre os backups dos dados, isso pode aumentar consideravelmente o custo da infraestrutura.
Manipulação de documentos por multiusuários simultâneos	Possui. No drive file stream os documentos podem ser editados por vários usuários ao mesmo tempo mantendo a possibilidade de reverter o documento inteiro ou partes dele para qualquer versão antes da alteração de qualquer usuário.	Não possui. O sistema só gerencia backups, não possui uma toolbox para edição de arquivos nem sincronização em tempo real, aspectos que seriam necessários para ter a dita funcionalidade.

Fonte: Compilação do autor.

3.6 PROCESSO DE ELABORAÇÃO DO QUESTIONÁRIO

O questionário foi pensado seguindo as diretrizes que foram definidas por processo heurístico à medida que foram concluídas as etapas de implantação, monitoramento e coleta de feedback dos usuários. Alguns cuidados para elaboração do questionário foram definidos, são eles: não usar perguntas abertas, não elaborar perguntas longas, seguir uma ordem nas perguntas que seja de acordo com o que será necessário para a implantação das soluções, categorizar as perguntas para evidenciar os temas ou etapas diferentes da ordem de implantação da ferramenta, ser composto com perguntas e alternativas para que seja capaz de calcular um resultado de uma recomendação entre as soluções propostas, não utilizar alternativas muito abertas para evitar um resultado não conclusivo, chegar em ao menos três resultados possíveis (Mais apto para solução A, Mais apto para Solução B, Inapto para ambas as soluções), informar os aspectos que impossibilitam a utilização quando inapto para uma das soluções.

A seguir está o questionário elaborado pelo autor, junto as alternativas e marcado em negrito estão os pontos de decisão que levam a um resultado final.

Questionário

Tema: Infraestrutura

1) A equipe de TI tem os conhecimentos necessários para implantação e operação de um servidor *Bacula*? Caso não conheça o processo saiba mais em: <https://www.bacula.lat/>

- a) Sim.
- b) Não, mas estou disposto a me capacitar.
- c) Não, e não irei me capacitar para isso.

2) Sua organização possui a infraestrutura necessária para hospedar o serviço do *Bacula* ou tem recursos financeiros para compra dos equipamentos necessários?

- a) Sim.
- b) Não.

3) Sua organização já possui uma licença para o *G Suite* ou tem recursos financeiros para adquirir uma? Caso não saiba os custos, verifique o site: <https://workspace.google.com/intl/pt-BR/pricing.html>

- a) Sim.
- b) Não, e não tenho como adquirir num futuro próximo.

Tema: Sensibilidade dos Dados

4) Sua organização possui dados sensíveis, que precisam de backups, mas que trariam risco e grande impacto a organização em casos de vazamentos?

- a) Sim, mas estou disposto em ter backups deles hospedados pela Google pois confio na segurança do serviço.
- b) Sim, e não estou disposto a ter cópias desses arquivos na nuvem.
- c) Não possuo dados sensíveis que precisam de backups.

Tema: Disponibilidade dos dados

5) Sua organização precisa que os arquivos estejam sempre disponíveis para os funcionários, mesmo que eles estejam fora do local de trabalho?

- a) Sim.

b) Não.

*Responda a próxima questão caso tenha dito “Sim” para a anterior

6) Os funcionários da sua organização têm meios de acessar remotamente seus dados a qualquer momento?

a) Sim.

b) Não, mas estou disposto a implementar/liberar essa função.

c) Não, e não estou disposto a dar esse tipo de acesso aos funcionários.

Tema: Sincronização dos Dados

7) Selecione a opção que se adequa a periodicidade com que seus backups devem estar sincronizados com os dados originais

a) É extremamente necessário, em TODOS OS ARQUIVOS, que a cada alteração seja efetuado um backup em tempo real.

b) Possuo ALGUNS ARQUIVOS que preciso que seus backups sejam atualizados em tempo real.

c) NENHUM dos meus arquivos precisa que seus backups sejam atualizados em tempo real a cada modificação.

Tema: Colaboração

8) Selecione a opção que se adequa ao fluxo de trabalho dos colaboradores da sua organização.

a) Existem momentos em que seria extremamente necessário para o bom andamento das atividades da minha organização, o uso de ferramentas colaborativas de edição de arquivos para alterações em grupo de um mesmo documento, em tempo real.

b) Embora não seja essencial, acredito que a opção de edição colaborativa de documentos agilizaria os procedimentos na minha organização.

c) Nas atividades da minha organização, edições colaborativas de documentos não têm um impacto positivo significativo.

9) Os usuários da sua organização estão familiarizados com as funcionalidades de edições de arquivos de forma colaborativa?

a) Sim.

b) Não, mas estou disposto a capacitá-los para melhorar o fluxo de trabalho deles.

c) Não, e não estou disposto a fazer alguma mudança no fluxo de trabalho deles.

Após responder o questionário o processo para obter a pontuação para cálculo do resultado do roteiro é definido através do quadro 3. O respondente deve comparar as alternativas escolhidas em questão para obter uma pontuação referente a cada solução. No questionário também tem perguntas que buscam saber se a solução é viável ou não para a organização, por conta disso também são verificadas a quantidade de pontos de inviabilidade, para que a organização seja apta para a solução o resultado do teste não deve acumular nenhum ponto de inaptidão para a referida solução.

Quadro 3 - Tabela de Pontuação

QUESTÕES	OPÇÃO	PONTUAÇÃO			INAPTO PARA IMPLANTAR:	
		Bacula	Google Drive	Híbrido	Bacula	Google Drive
Questão 1	a	•				
	b	•				
	c				•	
Questão 2	a	•				
	b				•	
Questão 3	a		•			
	b					•
Questão 4	a		•			
	b	•				
	c		•			
Questão 5	a		•			
	b	•				
Questão 6	a	•				
	b	•				
	c				•	
Questão 7	a				•	

	b			•		
	c	•				
Questão 8	a				•	
	b			•		
	c	•				
Questão 9	a		•			
	b		•			
	c	•				
TOTAL DE PONTOS:						
De acordo com as alternativas assinaladas no questionário, preencha o total de pontos acumulados para cada coluna						

Fonte: Compilação do autor.

Como foi visto acima, embora a intenção com o questionário era ter inicialmente 3 resultados possíveis, fica mais coerente ter 4, sendo eles: **Uso do Bacula recomendado, Uso do Google Drive recomendado, Inapto para ambas as soluções e Solução Híbrida recomendada**. A última opção foi identificada na medida que as duas soluções foram implantadas no IFPE Palmares. Verificou-se que com o novo normal da pandemia do COVID-19, os usuários passaram a usar mais as funções de colaboração em arquivos, além de ter ficado difícil para eles a ida até o *campus* para buscar algum documento no servidor de arquivos, já que o acesso remoto não era liberado em suas estações de trabalho. Outro fator importante foi que, por segurança, a equipe de TI achou melhor não instaurar uma VPN para todos, o que fez aumentar o número de arquivos de texto migrados do servidor de arquivos para a Nuvem, sempre com a recomendação de não fazer isso com dados sensíveis.

Em contrapartida, os arquivos de configuração dos equipamentos de rede e serviços ainda precisavam de backups automatizados e locais, por ser mais prático e seguro, o que também fez do *Bacula* extremamente necessário. Diante disso verificou-se o cenário onde uma solução híbrida era possível e necessária, modificando o questionário.

No quadro 4 é possível ver as relações entre a pontuação obtida ao responder o questionário, e o respectivo resultado. Na coluna de resultado é feita uma recomendação de qual solução é mais adequada ao contexto da instituição, informando alguns pontos-chave que devem ser observados caso o respondente deseje seguir a recomendação.

Quadro 4 - Resultados do Questionário

PONTUAÇÃO	RESULTADO
Maior parte dos pontos para o <i>Bacula</i> e nenhum que o torne inapto.	O uso do <i>Bacula</i> é recomendado. Sua organização apresenta requisitos em que o uso de nuvem não é necessário ou afeta a segurança em níveis que são considerados inaceitáveis, por isso o uso de um servidor de Backups com gerenciamento centralizado pode ser a solução ideal.
Maior parte dos pontos para o <i>Google Drive</i> e nenhum que o torne inapto.	O uso do <i>Google Drive</i> é recomendado. Sua organização apresenta necessidade de funcionalidades que tornam a elaboração e edição de arquivos algo síncrono e colaborativo, tudo isso para manter a produtividade em níveis aceitáveis. Além disso, conseguir a licença para essa solução não parece ser um problema.
Inapto para as duas soluções.	Inapto para ambas as soluções. Sua organização parece não ter recursos no momento para implantar alguma das soluções ou simplesmente não se sente segura com alguma delas. Recomenda-se que procure outras soluções que existem no mercado. O importante é não ficar sem backup.
Ao menos um ponto para solução Híbrida e nenhum ponto de inaptação para qualquer solução.	Solução Híbrida recomendada. Sua organização apresenta condições favoráveis para ambas as soluções e parece precisar de funcionalidades específicas de cada uma delas. Nesse caso, é recomendado o uso de ambas, fazendo uma classificação de quais dados seriam melhor trabalhados em cada solução.

Fonte: Compilação do autor.

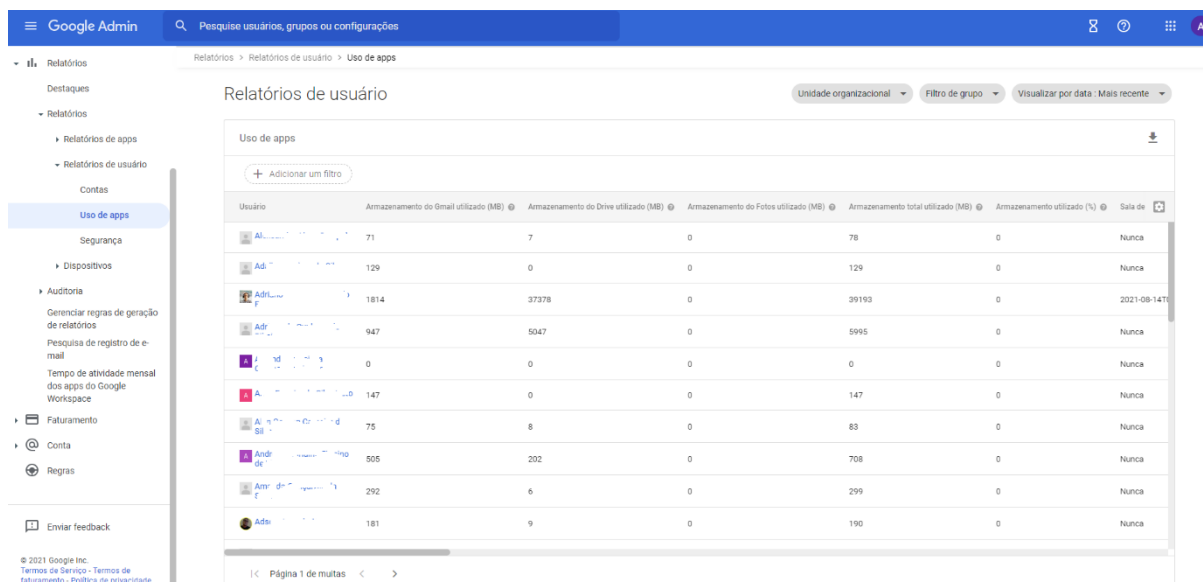
3.7 RESULTADOS DA IMPLANTAÇÃO DAS SOLUÇÕES NO IFPE

Abaixo estão imagens dos sistemas implantados mostrando as estatísticas de uso por parte dos funcionários e os backups efetuados por parte do sistema.

No painel de controle do *GSuite*, o *Google Admin*, é possível verificar os relatórios da ferramenta. Na figura 6 é possível ver o relatório de usuário por uso de aplicativos, um desses aplicativos é o *Drive File Stream*. A solução permite visualizar a quantidade e tipo de documentos armazenados por cada usuário, além de informar o armazenamento utilizado por

cada um. Também é possível ver informações de armazenamento em outros aplicativos como o Gmail. Através da funcionalidade de relatórios pode-se ter de forma transparente noções de utilização do serviço tanto desde a implantação quanto em períodos específicos.

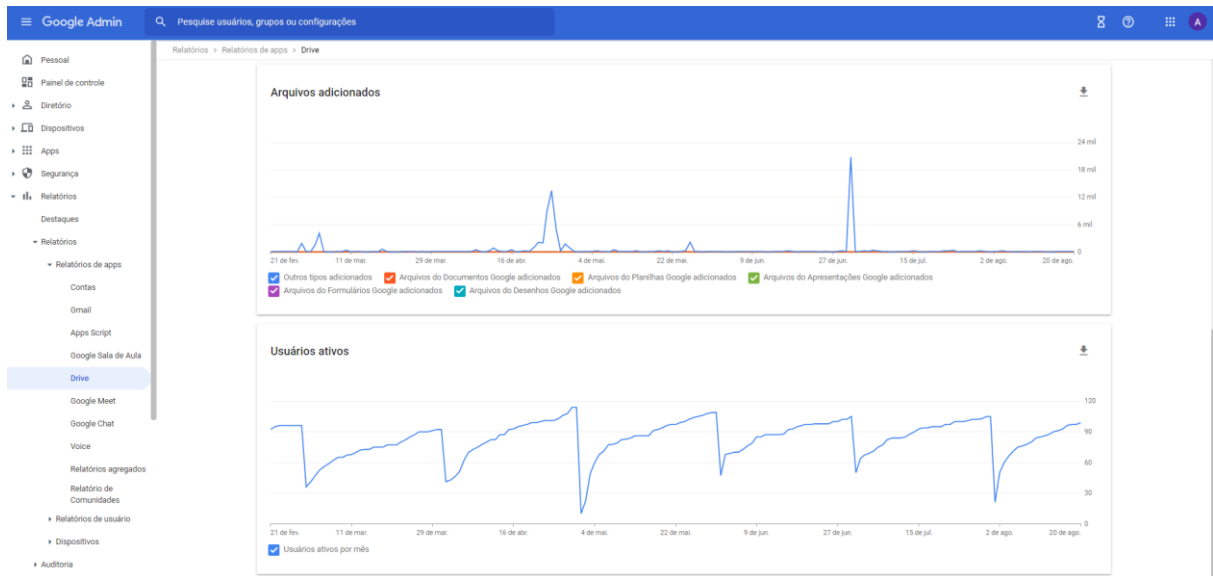
Figura 6 - Estatísticas de uso por usuário no Google Drive



Fonte: Compilação do autor.

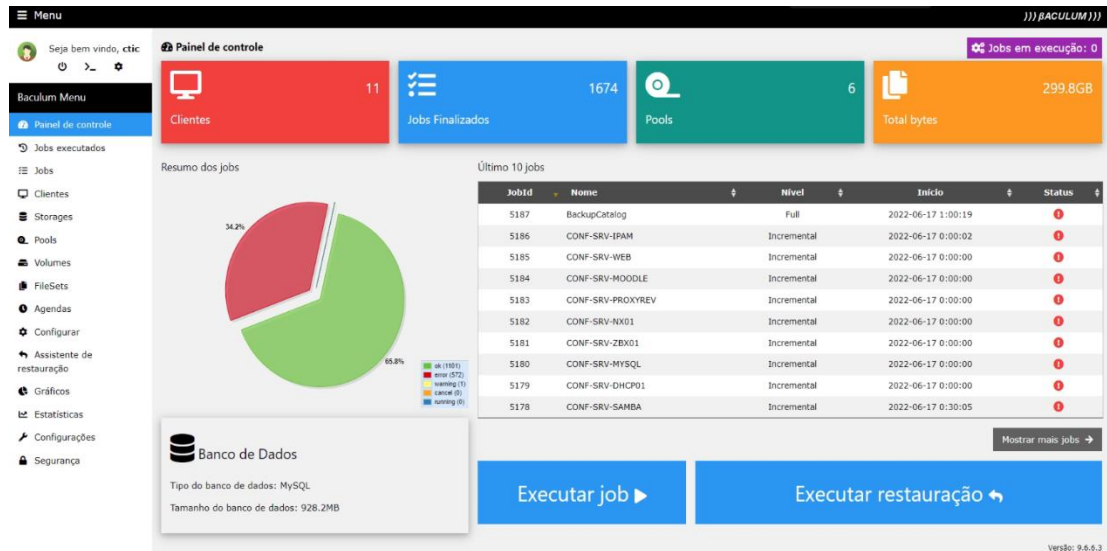
Na figura 7, logo abaixo, ainda é possível detalhar mais ainda o relatório, selecionando o relatório de apps e escolhendo exatamente o aplicativo que se deseja monitorar, neste caso específico foi selecionado o app do Drive *File Stream*. No primeiro gráfico foi possível verificar as datas onde houve pico de upload de informações, também tendo a visão de qual o tipo dos arquivos que foram armazenados. No segundo gráfico é apresentada a quantidade de usuários ativos em função do tempo.

Figura 7 - Estatísticas de tipos de arquivos compartilhados



Fonte: Compilação do autor.

Na figura 8 é exibido o painel de controle do *Bacula Community*, logo na página inicial é possível ver o grande número de *Jobs* de Backup e *Restore* efetuados desde a implantação. No gráfico pizza, pode-se notar em verde os *Jobs* finalizados com sucesso e em vermelho os que finalizaram com erro. A grande quantidade de erros se dá por dois principais motivos, em primeiro está o fato de que quando o *Director* e os *Clients* estão em fase de configuração muitos *Jobs* de teste são executados, e a maioria deles termina com erro até que a configuração correta seja alcançada. Em segundo, na época de implantação, o *campus* passou por problemas na rede elétrica por diversos dias, o que ocasionava desligamento não planejado de alguns servidores.

Figura 8 - Painel de controle do *Bacula*

Fonte: Compilação do autor.

Na figura 9, temos a exibição dos volumes de backup criados pelo *Bacula Community* desde sua implantação, embora não seja regra, pode-se notar que a maioria dos volumes com rótulo de semanal é maior que os com rótulo de diário, isso se deve ao fato da configuração de política de backup GFS ter sido adotada. Como os *Jobs* diários são realizados com backup incremental, os semanais com backup diferencial e os mensais com backup *full*, isto é claramente refletido no tamanho dos arquivos de volumes. Vale ressaltar que na imagem os rótulos dos volumes também têm numeração e, embora a listagem apresentada mostre os últimos volumes utilizados, a numeração dos rótulos algumas vezes não é essencialmente crescente. Isso se dá pelo fato do *Bacula* requerer do operador a configuração de um limite de tamanho para os volumes, e se o mesmo não é atingido, o *Bacula* reutiliza o volume até o mesmo ser totalmente preenchido.

Figura 9 - Estatísticas de Volumes do *Bacula*

The screenshot shows the Bacula web interface. The top bar includes a menu icon, the user name 'Seja bem vindo, ctic', and the Bacula logo. The sidebar menu on the left lists various system components, with 'Volumes' highlighted. The main content area displays a table titled 'Lista de Volumes' with columns for MediaId, Nome, Slot, InChanger, Status, Mediatype, Última gravação, Utilizado, and Ações. The table contains 13 rows of data, showing details for various weekly backup volumes. Below the table, there are pagination controls and a search bar.

MediaId	Nome	Slot	InChanger	Status	Mediatype	Última gravação	Utilizado	Ações
130	semanal-0130	0	Não	Used	File1	2022-05-25 0:00:31	981.3MB	Detalhar
129	semanal-0129	0	Não	Used	File1	2022-04-06 0:00:05	60.6KB	Detalhar
128	semanal-0128	0	Não	Used	File1	2022-03-19 0:01:26	121.5KB	Detalhar
127	semanal-0127	0	Não	Used	File1	2022-03-16 0:00:22	120.7KB	Detalhar
126	semanal-0126	0	Não	Used	File1	2022-03-13 0:41:54	3.3GB	Detalhar
125	semanal-0125	0	Não	Used	File1	2022-03-12 0:00:52	84.1KB	Detalhar
124	semanal-0124	0	Não	Used	File1	2022-03-06 0:53:12	3.3GB	Detalhar
123	semanal-0123	0	Não	Used	File1	2022-05-11 0:00:33	981.3MB	Detalhar
122	semanal-0122	0	Não	Used	File1	2022-05-04 0:00:26	11.6MB	Detalhar
121	semanal-0121	0	Não	Used	File1	2022-05-01 0:08:24	2.4GB	Detalhar

Fonte: Compilação do autor.

Depois da implantação das soluções, um pouco do potencial de cada ferramenta pôde ser observado. O *Bacula*, que é um serviço ubíquo para os colaboradores do IFPE, parece um tanto coadjuvante. Afinal, só a equipe de TI entende totalmente como o serviço funciona e tem acesso ao mesmo. Ainda assim, em duas ocasiões importantes ele pôde brilhar. Certa vez quando uma planilha, de tamanho maior que o comum, foi corrompida e o sistema foi capaz de restaurá-la, e uma segunda vez quando uma queda de energia fez um dos *switches* do *campus* perder suas configurações. É importante ressaltar que sistemas como esse devem ser constantemente testados, e por isso, mesmo sem uma necessidade real de uma restauração de arquivos, a equipe de TI faz testes semanais dessa função.

Quanto ao Drive, foi notório o impacto que teve nos colaboradores, muitos ficaram maravilhados com a possibilidade de compartilhar um arquivo uma única vez com seu colega e não precisar mais trocar e-mails com atualizações do mesmo. Outros, perceberam a vantagem de não precisar esperar alguém terminar a edição em uma planilha no servidor de arquivos, afinal quando compartilhada pelo Drive os resultados eram instantâneos e visíveis para todos que mantinham ele aberto. São funções simples, mas que alguns funcionários com mais idade ainda não tinham explorado. As funções avançadas, como um histórico de edição por cada usuário, também foram muito bem aceitas pelos gestores, sem falar na possibilidade de poder voltar para qualquer versão que qualquer usuário tenha feito durante todo o ciclo de vida do documento.

4 CONCLUSÃO

Este TCC apresentou um estudo de caso para efetuar uma comparação entre duas ferramentas de backups consolidadas no mercado, estabelecendo ao longo do estudo, instalação e operação, pontos primordiais das ferramentas para o bom funcionamento da organização em que foram implantadas, extrapolando os comparativos para os possíveis cenários em outras instituições, a fim de elaborar um questionário para apoiar a decisão da escolha mais adequada entre as duas ferramentas em seu próprio ambiente empresarial.

Algumas dificuldades foram encontradas nesse processo, a maior parte na instalação e configuração do *Bacula*, porém essa dificuldade foi importante para entender a importância do impacto da complexidade na hora de escolher uma das soluções. A análise das variáveis que deveriam ser destacadas para elaboração do roteiro para apoio de decisão, também foi algo desafiador, e foi um trabalho não só do aluno iniciador da pesquisa, mas também de toda a equipe de TI do IFPE *Campus* Palmares.

Mesmo não sendo um dos pontos-chave da pesquisa, é importante ressaltar que embora a solução do Google tenha sido implantada bem antes do *Bacula*, seu uso para backups no Google Drive era mínimo. Acredita-se que o novo panorama que a pandemia trouxe impactou bastante na aceitação das funções e nuvem, isso é um feliz acidente e um avanço para atividades não presenciais. Depois da pandemia, o projeto de teletrabalho que o IFPE tinha criado há anos, porém sem execução, ganhou força. Os gestores do IFPE já afirmam que, assim que o projeto entrar em vigor, mesmo depois da pandemia acabar, será possível que seus colaboradores usem o teletrabalho ao menos um dia na semana. Isso denota que o uso do Google Drive não é algo de momento e é uma solução exitosa. Importante ressaltar também que a solução do *Bacula* também é exitosa, através de conversas entre as equipes de TI do IFPE, o *Campus* Palmares já recebeu pedidos de ajuda para a implantação do *Bacula* em outros *Campi* mais novos.

4.1 CONSIDERAÇÕES FINAIS

O *Bacula* e o Google Drive são realmente duas ótimas ferramentas que também podem trabalhar bem em conjunto, cada uma com um foco complementar de acordo com o tipo dos arquivos e perfis dos usuários. É de fato uma boa surpresa que uma abordagem híbrida possa ser bem sucedida, embora seja algo realmente custoso em boa parte das instituições, isso pode ser replicado sem problemas nos demais *campi* do IFPE, já que não implica em mais custo

financeiro, dando flexibilidade e agilidade aos processos da organização, sem minar a segurança da informação. Recomenda-se fortemente a aplicação do questionário do item 4.4 a qualquer profissional que esteja em dúvida entre as ferramentas, isso pode poupar um bom tempo para entendê-las e verificar se algum requisito torna a aplicação de uma delas algo impossível no contexto. E, embora não seja o foco deste estudo ser um tutorial de instalação de um servidor de backups, o script de instalação apresentado nos anexos é mais atualizado que a maioria dos que se pode encontrar na internet e os arquivos de configuração do servidor *Bacula* foram feitos de forma modular para facilitar o entendimento e manutenção do serviço, por isso seriam bons artefatos para qualquer um que deseja instalar uma versão aproximada do *Bacula*.

4.2 CONTRIBUIÇÕES DESTE TRABALHO

Este TCC aplicou um estudo de caso comparando duas ferramentas de backup desde a instalação até uso em produção, criando um roteiro para agilizar e apoiar a escolha entre as duas ferramentas à medida que informações sobre o ambiente institucional são fornecidas. Acredita-se que não somente o roteiro elaborado aqui seja útil, mas também as experiências descritas no processo sirvam para entender melhor o funcionamento e desafios de cada solução.

4.3 PROPOSTA PARA TRABALHOS FUTUROS

Como indicação para elaborações futuras fica a possibilidade de aprofundamento do roteiro, identificando outras variáveis que não foram expostas ou algumas futuras que se apresentem com os avanços das tecnologias de backup. O desenvolvimento de um site que substitua o roteiro em papel e seja mais dinâmico, apresentando aspectos importantes à medida que cada resposta é dada e exemplificando os cenários de cada temática na qual a pergunta atual esteja envolvida. Também é interessante envolver outras ferramentas líderes do mercado, estabelecendo novas comparações e dando um leque maior de opções nas recomendações.

REFERÊNCIAS BIBLIOGRÁFICAS

ABWOGA, Andrew. **THE TOWER OF HANOI BACKUP STRATEGY**. Backup.ninja. 2020. Disponível em: <https://backup.ninja/news/tower-hanoi-backup-strategy>. Acesso em: 2 set. 2021.

ALENCAR, Jesus Guilherme Bindi; SCHIMIGUEL, Juliano. **IMPLEMENTAÇÃO DE BACKUP COMO PROCESSO DE SEGURANÇA DA INFORMAÇÃO**. 2018. Disponível em: <http://www.eumed.net/2/rev/atlante/2018/02/backup-seguranca-informacao.html>. Acesso em: 10 ago. 2021.

AMAZON. **AMAZON SIMPLE STORAGE SERVICE USER GUIDE**. 2006. Disponível em: <https://docs.aws.amazon.com/AmazonS3/latest/userguide/s3-userguide.pdf#Welcome>. Acesso em: 3 set. 2021.

ARAUJO, Juarez. **QUAIS OS PRINCÍPIOS DA LGPD E POR QUE ADEQUAR SUA POLÍTICA DE BACKUP**. Dbacorp Blog. 2020. Disponível em: <https://blog.dbacorp.com.br/2020/07/15/sua-politica-de-backup-esta-adequada-com-lgpd/>. Acesso em: 26 jan. 2022.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). **NBR ISO/IEC 27002 DE 11/2013**. [2013]. Disponível em: <https://www.normas.com.br/visualizar/abnt-nbr-nm/21529/abnt-nbriso-iec27002-tecnologia-da-informacao-tecnicas-de-seguranca-codigo-de-pratica-para-controles-de-seguranca-da-informacao>. Acesso em: 22 jul. 2021.

BACULA.LAT. **O QUE É O BACULA?**. 2008. Disponível em: <http://www.bacula.lat/o-que-e-o-bacula/>. Acesso em: 22 jul. 2021.

BACULA.ORG. **Pacotes Deb e rpm**. [2022?]. Disponível em: <https://www.bacula.org/bacula-binary-package-download/>. Acesso em: 20 jun. 2022.

BRASIL. Câmara dos Deputados. **Projeto de Lei Nº 2.126, de 2011**. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Brasília: Câmara dos Deputados, 2011. Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=918144&filename=Avulso+-PL+2126/2011. Acesso em: 20 jun. 2022.

BRASIL. Ministério do Planejamento, orçamento e gestão. Secretaria de Tecnologia da Informação. **Boas práticas, orientação e vedações para contratação de Serviços de Computação em Nuvem**. Brasília, DF, 2016. Disponível em: <https://www.gov.br/economia/pt-br/assuntos/drei/consultas-publicas/consultas-encerradas/2018/arquivos/guia-de-orientacoes-para-contratacao-de-servicos-em-nuvem.pdf>. Acesso em: 20 jun. 2022.

CABRAL, Ismael. **SEGURANÇA DA INFORMAÇÃO EM BIBLIOTECAS UNIVERSITÁRIAS FEDERAIS**: Um levantamento sobre ferramentas e técnicas utilizadas. 2015. 80 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Universidade Federal de Santa Catarina, Florianópolis. Disponível em: <https://repositorio.ufsc.br/xmlui/handle/123456789/133969>. Acesso em: 18 jun. 2022.

DANTAS, Marcus Leal. **SEGURANÇA DA INFORMAÇÃO: UMA ABORDAGEM FOCADA EM GESTÃO DE RISCOS**. 1 ed. Olinda: Livro rápido, 2011.

DICIO. Dicionário Online de Português. **Backup**. [2022?]. Disponível em: <https://www.dicio.com.br/backup/>. Acesso em: 13 mar. 2022.

DOMS, Caroline. **COMO FUNCIONA GOOGLE DRIVE FILE STREAM? VEJA COMO FAZER DOWNLOAD DA NUVEM**. Techtudo. 2020. Disponível em: <https://www.techtudo.com.br/listas/2020/07/como-funciona-google-drive-file-stream-veja-como-fazer-download-da-nuvem.ghtml>. Acesso em: 22 jul. 2021.

FARIA, Heitor M. **BACULA COMMUNITY & ENTERPRISE**. 4 ed. Rio de Janeiro: Brasport, 2020.

IBM. **QUANTO CUSTA UMA VIOLAÇÃO DE DADOS?**. 2021. Disponível em: <https://www.ibm.com/br-pt/security/data-breach>. Acesso em: 22 fev. 2022.

JUNIOR, Braz Pereira. **GERENCIAMENTO CENTRALIZADO DE BACKUPS DISTRIBUÍDOS**. 2010. 70 f. Trabalho de Conclusão de Curso (Ciência da Computação) – Universidade do Vale do Itajaí, Santa Catarina. Disponível em: https://silo.tips/queue/universidade-do-vale-do-itajai-centro-de-ciencias-tecnologicas-da-terra-e-do-mar-97?&queue_id=-1&v=1655599142&u=NDUuMTY2LjQ5LjU2. Acesso em: 18 jun. 2022.

KASPERSKY TEAM. **CIBERATAQUES CRECEM 23% NO BRASIL EM 2021**. Kaspersky Daily. 2021. Disponível em: <https://www.kaspersky.com.br/blog/panorama-ciberameacas-brasil-2021-pesquisa/18020/>. Acesso em: 22 fev. 2022.

MINDSIGHT. **HOW IT WORKS: DATA BACKUP ROTATION SCHEMES**. Mindsight. Chicago, 2017. Disponível em: <https://gomindsight.com/insights/blog/data-backup-schemes/>. Acesso em: 2 set. 2021.

PASSOS, Eduardo. **QUAL O MELHOR SOFTWARE DE BACKUP PARA SERVIDORES?** Infobusiness Informática, 2019. Disponível em: <https://infob.com.br/qual-o-melhor-software-de-backup-para-servidores/>. Acesso em: 2 set. 2021

PEDUTI. **ADEQUAÇÃO DA POLÍTICA DE BACKUP À LGPD**. Peduti Blog. 2021. Disponível em: <https://blog.peduti.com.br/politica-de-backup-a-lgpd/>. Acesso em: 12 fev. 2022.

PRESTON, W. Curtis. **BACKUP & RECOVERY: INEXPENSIVE BACKUP SOLUTIONS FOR OPEN SYSTEMS**. 1 ed. Sebastopol: O'Reilly Media, 2007.

RED HAT. **O QUE É ARMAZENAMENTO NAS?** RED HAT. 2018. Disponível em: <https://www.redhat.com/pt-br/topics/data-storage/network-attached-storage>. Acesso em: 15 mar. 2022.

RODRIGUES, Wilson Flávio. **ANÁLISE DOS PROCEDIMENTOS DE BACKUP DOS INSTITUTOS FEDERAIS**. 2017. 140 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Pernambuco, Recife. Disponível em:

<https://repositorio.ufpe.br/bitstream/123456789/25612/1/DISSERTA%C3%87%C3%83O%20Wilson%20Fl%C3%A1vio%20Rodrigues.pdf>. Acesso em: 29 jul. 2021.

RUSSELL, Erin. **ARCERVE BACKUP**. 2018. 136 f. Dissertação (Pós-graduação *Stricto Sensu* em Gestão do Conhecimento e Tecnologia da Informação) – Universidade Católica de Brasília, DF. Disponível em: <https://www.arcserve.com/br/data-protection-solutions/arcserve-backup/>. Acesso em: 2 set. 2021.

SANTOS, Bruno Belarmínio Aparecido dos. **BACKUP CORPORATIVO COM ALTA RETENÇÃO: SUBSÍDIOS PARA CONSTRUÇÃO DA ARQUITETURA**. 2018. Disponível em: <https://bdtd.ucb.br:8443/jspui/handle/tede/2616>. Acesso em: 30 set. 2021.

SERPRO. **O QUE MUDA COM A LGPD**. [2020?]. Disponível em: <https://www.serpro.gov.br/lgpd/menu/a-lgpd/o-que-muda-com-a-lgpd>. Acesso em: 12 fev. 2022.

SONICWALL. **2022 CYBER THREAT REPORT**. [2022?]. Disponível em: <https://www.sonicwall.com/medialibrary/en/white-paper/2022-sonicwall-cyber-threat-report.pdf>. Acesso em: 22 fev. 2022.

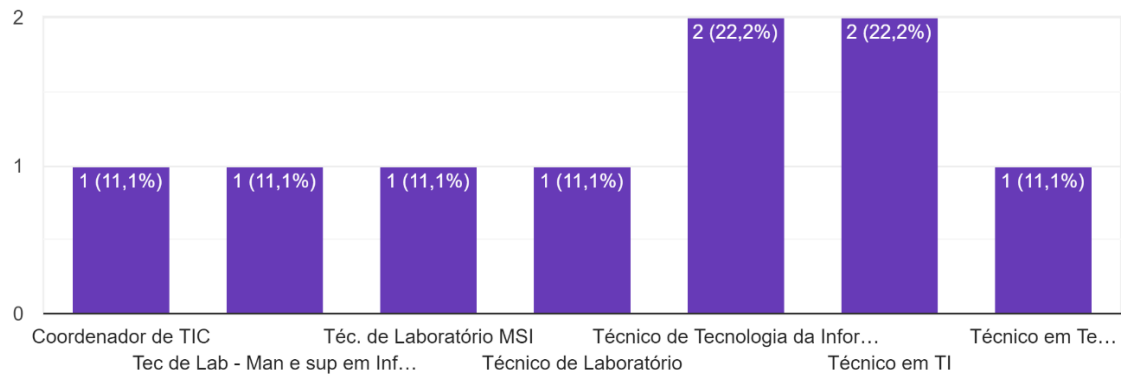
SUPERIOR TRIBUNAL MILITAR (STM). **Especialistas e gestores de TI debatem no STM uso do serviço em nuvem, por órgãos públicos federais**. Disponível em: <https://jusdecisum.com.br/especialistas-e-gestores-de-ti-debatem-no-stm-uso-do-servico-em-nuvem-por-orgaos-publicos-federais/>. Acesso em: 20 jun. 2022.

TOTVS. **SEGURANÇA DA INFORMAÇÃO: O QUE É E BOAS PRÁTICAS**. TOTVS. 2021. Disponível em: <https://www.totvs.com/blog/negocios/seguranca-da-informacao/>. Acesso em: 16 mar. 2022.

ANEXO A – PESQUISA SOBRE EXISTÊNCIA DE SOLUÇÕES DE BACKUP NOS

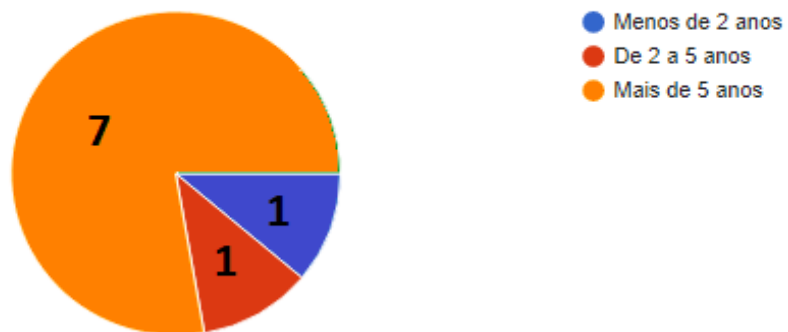
Qual o seu cargo ou função?

9 respostas



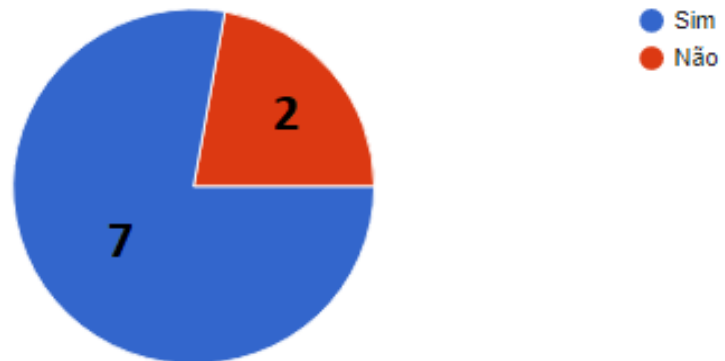
Há quanto tempo atua no IFPE

9 respostas



Na infraestrutura do seu campus existe um Servidor de Arquivos?

9 respostas



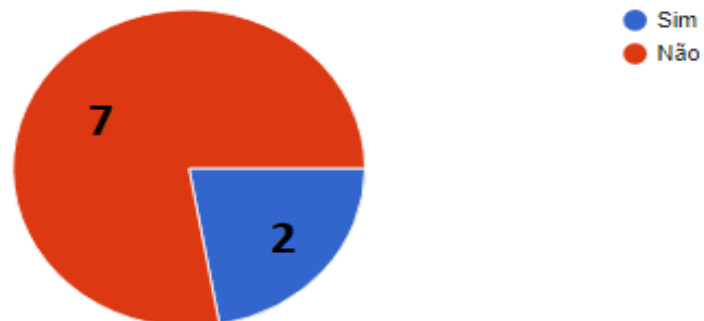
Em seu campus é utilizada alguma versão do pacote do GSuite?

9 respostas



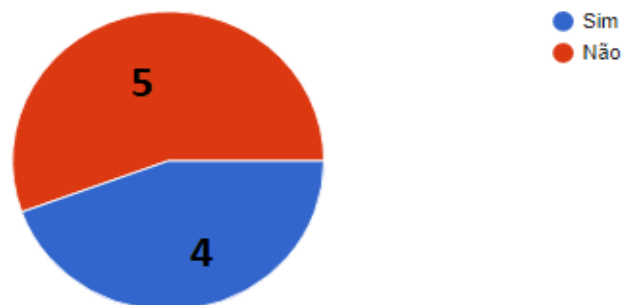
Em seu campus é utilizada alguma solução de backup para end-points?

9 respostas



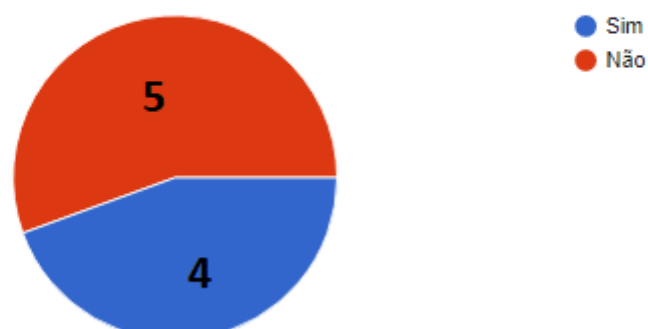
Na infraestrutura do seu campus existe alguma solução de backup para os serviços de TIC?

9 respostas



Em seu campus já houve algum evento que ocasionou perda de dados de forma irreversível?

9 respostas



Meu campus já possui solução de backup para serviços de TIC

Informe qual solução de backup seu campus utiliza.

4 respostas

Bacula

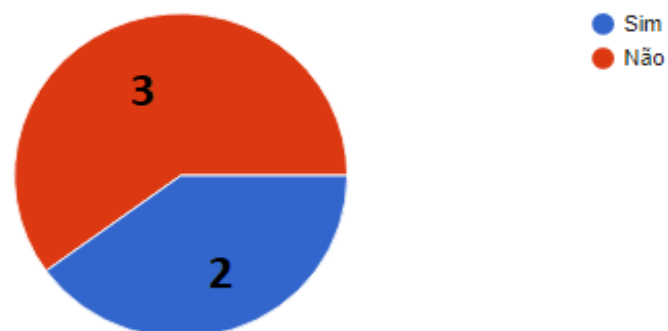
Proxmox Backup Server

Bacula e Google Drive

Meu campus não possui solução de backup para serviços de TIC

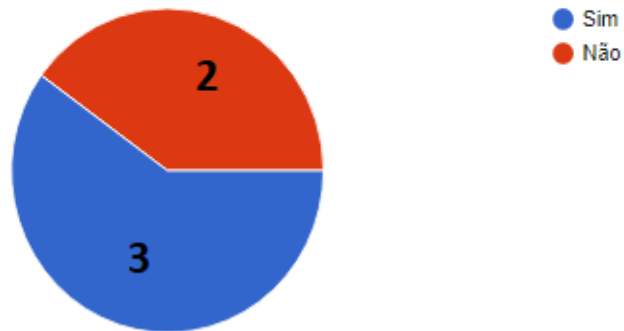
Existe previsão para ser implantada uma solução de backup no seu campus?

5 respostas

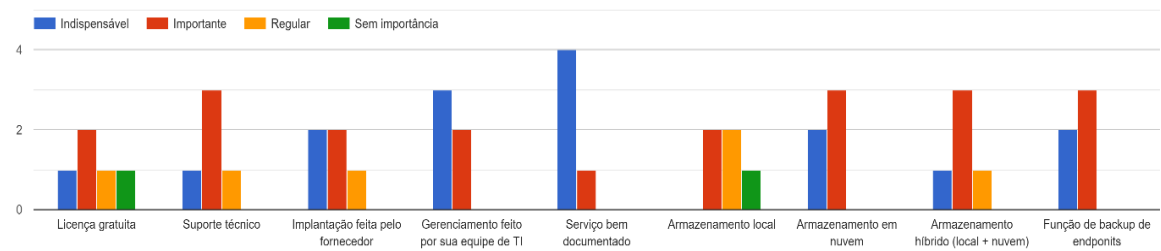


Você possui conhecimento necessário para especificar uma solução de backup para seu campus?

5 respostas



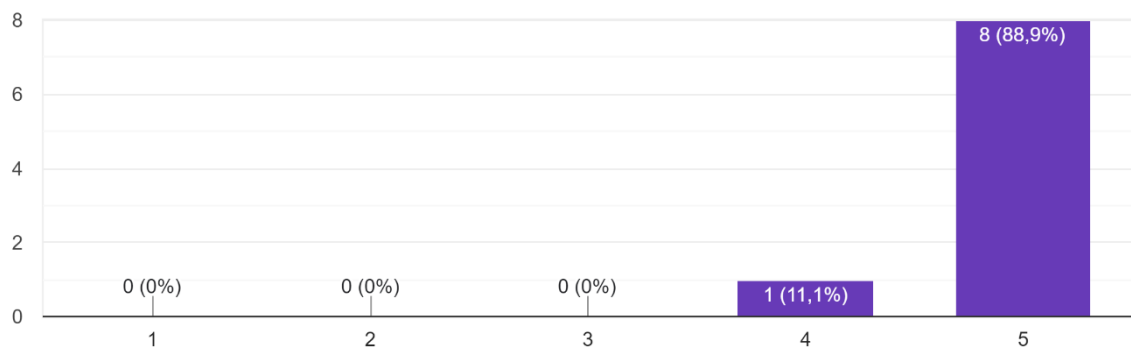
Escolha o grau de importância dos seguintes aspectos para uma solução de backup ser implantada em seu campus



Obrigado por chegar até aqui!

Para você, qual a utilidade de uma ferramenta que, ao responder algumas perguntas sobre sua infraestrutura, recomendasse as melhores soluções de backup para sua organização?

9 respostas



ANEXO B – SCRIPT DE INSTALAÇÃO DO BACULA

Bacula Community

O que é?

Bacula é um *sistema open-source* que permite gerenciar o backup, a restauração e a verificação dos dados em uma rede de computadores de diferentes tipos.

De acordo com as estatísticas do Source Forge (classificação e downloads), o Bacula é de longe o programa de backup de código aberto mais popular. Possui 2.5 milhões de downloads e milhares de colaboradores em todo o mundo.

Neste livro vamos abordar a instalação do servidor Bacula, instalação de clientes, instalação de storage, configuração de backups e restores, agendamento e política de backups. Nessa abordagem será utilizado o Debian 9, portanto tenha ele já instalado e configurado antes de prosseguir. Além disso, serão utilizadas as versões Bacula 9.6.3 e MySQL 5.7.



Instalação dos serviços no Servidor

Iremos utilizar os repositórios do bacula community para que nosso servidor de backups não fique dependente dos pacotes nativos do debian, que podem ficar desatualizados eventualmente.

Para utilização dos repositórios do bacula community você deve efetuar um registro no site <https://www.bacula.org/bacula-binary-package-download/>

Após o registro você receberá uma chave para ter acesso ao repositório. Em seguida, você pode iniciar a instalação do Bacula e MySQL com os comando abaixo.

```
# Requisitos para instalar o Bacula por pacotes  
apt-get install -y zip wget bzip2 apt-transport-https  
# Download da chave do repositório  
wget -c https://www.bacula.org/downloads/Bacula-4096-Distribution-Verification-key.asc -O /tmp/Bacula-4096-Distribution-Verification-key.asc  
# Adicionar chave no repositório local
```

```
apt-key add /tmp/Bacula-4096-Distribution-Verification-key.asc
```

Entre no modo root e execute os comandos abaixo.

```
# Criar variáveis (como root)
# Uma bacula_key pode ser requisitada no link: # https://www.bacula.org/bacula-binary-package-
download/
bacula_version="9.6.3"
bacula_key="xxxxxxxxxxxxx"
# O linux_name deve ser o nome da versão do debian
linux_name="stretch"
export DEBIAN_FRONTEND=noninteractive
# Criar o repositório do Bacula Community (como root)
echo "# Bacula Community
deb http://www.bacula.org/packages/${bacula_key}/debs/${bacula_version}/${linux_name}/amd64/ ${lin
```

Ainda em modo root configure o repositório para instalação do MySQL.

```
##### Fazer como root#####
# Instalar o MySQL
wget -c https://repo.mysql.com/RPM-GPG-KEY-mysql -O /tmp/RPM-GPG-KEY-mysql --no-check-certificate
apt-key add /tmp/RPM-GPG-KEY-mysql
# O linux_name deve ser o nome da versão do debian
linux_name="stretch"
echo "deb http://repo.mysql.com/apt/debian/ ${linux_name} mysql-apt-config
deb http://repo.mysql.com/apt/debian/ ${linux_name} mysql-5.7
deb http://repo.mysql.com/apt/debian/ ${linux_name} mysql-tools
deb http://repo.mysql.com/apt/debian/ ${linux_name} mysql-tools-preview
deb-src http://repo.mysql.com/apt/debian/ ${linux_name} mysql-5.7" >
/etc/apt/sources.list.d/mysql.list#####
```

Agora vamos instalar o MySQL e o Bacula. Também iremos configurar as variáveis de ambiente para facilitar o uso do Bacula. Você já pode sair do modo root e executar os comandos abaixo com o sudo.

ANEXO C – ARQUIVO DE CONFIGURAÇÃO DO *DIRECTOR BACULA*

```

1 Director {
2   Name = "pmr-srv-bacula-dir"
3   Messages = "Daemon"
4   QueryFile = "/opt/bacula/scripts/query.sql"
5   WorkingDirectory = "/opt/bacula/working"
6   PidDirectory = "/opt/bacula/working"
7   MaximumConcurrentJobs = 20
8   Password = "xxx"
9 }
10
11 # Definição de Clients e Jobs
12 @/opt/bacula/etc/bacula-dir-clients-and-jobs.conf
13
14 Catalog {
15   Name = "MyCatalog"
16   Password = "xxx"
17   User = "bacula"
18   DbName = "bacula"
19 }
20
21 # Definição de Agendamentos
22 @/opt/bacula/etc/bacula-dir-schedules.conf
23
24 # Definição de Filesets
25 @/opt/bacula/etc/bacula-dir-filesets.conf
26
27 # Definição de Pools
28 @/opt/bacula/etc/bacula-dir-pools.conf
29
30 Messages {
31   Name = "Daemon"
32   MailCommand = "/opt/bacula/bin/bsmtp -h localhost -f \"(Bacula) <tr>\" -s \"Bacula daemon message\" %r"
33   Mail = root = All, !Debug, !Saved, !Skipped, !Terminate
34   Append = /opt/bacula/log/bacula.log = All, !Debug, !Saved, !Skipped
35   Console = All, !Debug, !Saved, !Skipped
36 }
37 Messages {
38   Name = "Standard"
39   MailCommand = "/opt/bacula/bin/bsmtp -h localhost -f \"(Bacula) <tr>\" -s \"Bacula: %t %e of %c %l\" %r"
40   OperatorCommand = "/opt/bacula/bin/bsmtp -h localhost -f \"(Bacula) <tr>\" -s \"Bacula: Intervention needed for %j\" %r"
41   Mail = ti@palmareis.ifpe.edu.br = All, !Debug, !Saved, !Skipped, !Terminate
42   Append = /opt/bacula/log/bacula.log = All, !Debug, !Saved, !Skipped
43   Console = All, !Debug, !Saved, !Skipped
44   Operator = root = Mount
45   Catalog = All, !Debug, !Saved, !Skipped
46 }
47 Console {
48   Name = "pmr-srv-bacula-mon"
49   Password = "xxx"
50   CommandAcl = "status"
51   CommandAcl = ".status"
52 }
53 Storage {
54   MaximumConcurrentJobs = 10
55   Password = "xxx"
56   Name = "Storage"
57   Address = "IP"
58   Device = "FileChgrl-Dev1"
59   MediaType = "File1"
60 }
61

```

ANEXO D – ARQUIVO DE CONFIGURAÇÃO DO *FILE DAEMON BACULA*

```

1  #
2  # Default Bacula File Daemon Configuration file
3  #
4  # For Bacula release 9.6.3 (09 March 2020) -- debian 9.11
5  #
6  # There is not much to change here except perhaps the
7  # File daemon Name to
8  #
9  #
10 # Copyright (C) 2000-2020 Kern Sibbald
11 # License: BSD 2-Clause; see file LICENSE-FOSS
12 #
13 #
14 #
15 # List Directors who are permitted to contact this File daemon
16 #
17 Director {
18     Name = pmr-srv-bacula-dir
19     Password = "xxx"
20 }
21 #
22 #
23 # Restricted Director, used by tray-monitor to get the
24 # status of the file daemon
25 #
26 Director {
27     Name = pmr-srv-bacula-mon
28     Password = "xxx"
29     Monitor = yes
30 }
31 #
32 #
33 # "Global" File daemon configuration specifications
34 #
35 FileDaemon {                                # this is me
36     Name = pmr-srv-bacula-fd
37     FDport = 9102                            # where we listen for the director
38     WorkingDirectory = /opt/bacula/working
39     Pid Directory = /opt/bacula/working
40     Maximum Concurrent Jobs = 20
41     Plugin Directory = /opt/bacula/plugins
42 }
43 #
44 # Send all messages except skipped files back to Director
45 Messages {
46     Name = Standard
47     director = pmr-srv-bacula-dir = all, !skipped, !restored
48 }
49

```

ANEXO E – ARQUIVO DE CONFIGURAÇÃO DOS AGENDAMENTOS (MODULAR)

```
1 # When to do the backups, full backup on first sunday of the month,
2 # differential (i.e. incremental since full) every other sunday,
3 # and incremental backups other days
4
5 # Agendamento do SERVIDOR DE ARQUIVOS (GFS)
6 Schedule {
7     Name = CicloTrimestral
8     Run = Level=Incremental Pool=Diario mon-tue at 00:00
9     Run = Level=Incremental Pool=Diario thu-fri at 00:00
10    Run = Level=Differential Pool=Semanal wed at 00:00
11    Run = Level=Differential Pool=Semanal 2nd-5th sat at 00:00
12    Run = Level=Full Pool=Mensal 1st sat at 00:00
13 }
14
15 # Agendamento dos Arquivos de Configurações dos Servidores
16 Schedule {
17     Name = CicloDosServidores
18     Run = Level=Incremental Pool=Diario mon-fri at 00:00
19     Run = Level=Full Pool=Semanal sun at 00:00
20 }
21
22 # Agendamento dos Backups do Catalogo. Feito sempre depois dos backups
23 Schedule {
24     Name = CicloAposBackup
25     Run = Level=Full sun-sat at 00:10
26 }
27
```

ANEXO F – ARQUIVO DE CONFIGURAÇÃO DOS *POOLS* (MODULAR)

```

1  # Default pool definition
2  Pool {
3      Name = Default
4      Pool Type = Backup
5      Recycle = yes                    # Bacula can automatically recycle Volumes
6      AutoPrune = yes                  # Prune expired volumes
7      Volume Retention = 365 days      # one year
8      Maximum Volume Bytes = 50G      # Limit Volume size to something reasonable
9      Maximum Volumes = 100           # Limit number of Volumes in Pool
10 }
11
12 # Definição do Pool Diário
13 Pool {
14     Name = Diario
15     Pool Type = Backup
16     Recycle = yes
17     AutoPrune = yes
18     Volume Use Duration = 23 hours
19     Volume Retention = 90 days
20     Maximum Volume Bytes = 20G
21     Maximum Volumes = 120
22     # LabelFormat = "diario-#{NumVols}"
23     LabelFormat = "diario-"
24 }
25
26 # File Pool definition
27 Pool {
28     Name = Semanal
29     Pool Type = Backup
30     Recycle = yes
31     AutoPrune = yes
32     Volume Use Duration = 23 hours
33     Volume Retention = 90 days
34     Maximum Volume Bytes = 50G
35     Maximum Volumes = 60
36     LabelFormat = "semanal-"
37 }
38
39 # File Pool definition
40 Pool {
41     Name = Mensal
42     Pool Type = Backup
43     Recycle = yes
44     AutoPrune = yes
45     Volume Use Duration = 23 hours
46     Volume Retention = 90 days
47     Maximum Volume Bytes = 100G
48     Maximum Volumes = 9
49     LabelFormat = "mensal-"
50 }
51
52 Pool {
53     Name = Catalogo
54     Pool Type = Backup
55     Volume Retention = 90 days
56     Recycle = yes
57     AutoPrune = yes
58     LabelFormat = "catalogo-"
59 }
60
61 # Scratch pool definition
62 Pool {
63     Name = Scratch
64     Pool Type = Backup
65 }

```

ANEXO G – ARQUIVO DE CONFIGURAÇÃO DOS *FILESETS* (MODULAR)

```

1  # This is the backup of the catalog
2  # Definição dos arquivos de backup da base do bacula
3  FileSet {
4      Name = "Catalog"
5      Include {
6          Options {
7              signature = MD5
8          }
9          File = "/opt/bacula/working/bacula.sql"
10     }
11 }
12
13 FileSet {
14     Name = "Full Set Restore"
15     Include {
16         Options {
17             signature = SHA1
18             compression = GZIP
19             verify = pinl
20             onefs = no
21         }
22         File = /mnt/samba
23     }
24
25     Exclude {
26         File = /var/lib/bacula
27         File = /nonexistent/path/to/file/archive/dir
28         File = /proc
29         File = /tmp
30         File = /.journal
31         File = /.fsck
32     }
33 }
34
35 FileSet {
36     Name = "Full Set Ifpe"
37     Include {
38         Options {
39             signature = SHA1
40             compression = GZIP
41             verify = pinl
42             onefs = no
43         }
44         File = /mnt/samba/IFPE
45     }
46
47     Exclude {
48         File = /var/lib/bacula
49         File = /nonexistent/path/to/file/archive/dir
50         File = /proc
51         File = /tmp
52         File = /.journal
53         File = /.fsck
54     }
55 }

```

ANEXO H – ARQUIVO DE CONFIGURAÇÃO DOS *CLIENTS* E *JOBS* (MODULAR)

```

1 JobDefs {
2   Name = "DefaultJob"
3   Type = Backup
4   Level = Incremental
5   Client = PMR-SRV-BACULA-fd
6   FileSet = "Full Set Ifpe"
7   Schedule = "CicloTrimestral"
8   Storage = Storage
9   Messages = Standard
10  Pool = Diario
11  SpoolAttributes = yes
12  Priority = 10
13  Write Bootstrap = "/opt/bacula/working/%c.bsr"
14 }
15
16 # Backup the catalog database (after the nightly save)
17 Job {
18   Name = "BackupCatalog"
19   JobDefs = "DefaultJob"
20   Level = Full
21   FileSet="Catalog"
22   Schedule = "CicloAposBackup"
23   # This creates an ASCII copy of the catalog
24   # Arguments to make_catalog_backup.pl are:
25   # make_catalog_backup.pl <catalog-name>
26   RunBeforeJob = "/opt/bacula/scripts/make_catalog_backup.pl MyCatalog"
27   # This deletes the copy of the catalog
28   RunAfterJob = "/opt/bacula/scripts/delete_catalog_backup"
29   Write Bootstrap = "/opt/bacula/working/%n.bsr"
30   Priority = 11           # run after main backup
31 }
32
33 # Standard Restore template, to be changed by Console program
34 # Only one such job is needed for all Jobs/Clients/Storage ...
35 Job {
36   Name = "RestoreFiles"
37   Type = Restore
38   Client=PMR-SRV-BACULA-fd
39   Storage = Storage
40   # The FileSet and Pool directives are not used by Restore Jobs
41   # but must not be removed
42   FileSet="Full Set Restore"
43   Pool = Diario
44   Messages = Standard
45   Where = /var/lib/bacula/bacula-restores
46 }
47

```

```
47
48 # Client (File Services) to backup
49 Client {
50     Name = PMR-SRV-BACULA-fd
51     Address = "IP"
52     FDPort = 9102
53     Catalog = MyCatalog
54     Password = "xxx"           # password for FileDaemon
55     File Retention = 60 days   # 60 days
56     Job Retention = 6 months   # six months
57     AutoPrune = yes           # Prune expired Jobs/Files
58 }
59
60 # SRV-SAMBA
61 Client {
62     Name = PMR-SRV-SAMBA
63     Address = "IP"
64     Password = "xxx"
65     @/opt/bacula/etc/basic-client.conf
66 }
67 Job {
68     Name = "IFPE"
69     JobDefs = "DefaultJob"
70     FileSet = "Full Set Ifpe"
71     Client = PMR-SRV-SAMBA
72 }
73
```

ANEXO I – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Título da Pesquisa: Sistemas de Backup: Um Comparativo Entre Bacula Community e Google Drive Para Empresas

Nome do Orientador(a): Prof. M.e Adalmeres Cavalcanti

Nome do Pesquisador Principal: Aluizio Felipe Miranda da Silva

1. A Justificativa e os Objetivos: O motivo que nos leva a estudar o problema vem do fato de

muitas organizações, públicas e privadas, terem uma crescente massa de informações geradas pelos seus processos internos, enquanto que há ausência de uma política de backup dessas informações ou até mesmo a aplicação de soluções não adequadas àquele ambiente, a pesquisa se justifica na existência de diversas soluções de Backup no mercado com foco e funcionalidades bastante específicas, ao mesmo tempo que a natureza dos dados das organizações é complexa, tornando o processo de escolha de uma ferramenta desgastante e algumas vezes errôneo. O objetivo desse projeto é estabelecer procedimentos para compor um roteiro de análise, que servirá para identificar as particularidades de uma organização, e a partir disso apoiar na decisão entre as soluções de backup estudadas.

2. Participantes da pesquisa: Os participantes da pesquisa são um integrante da equipe de TI e o Diretor Geral do campus, este último sendo o Gestor Máximo da Organização e responsável pela autorização da pesquisa.

3. Envolvimento na pesquisa: Ao participar deste estudo os srs. permitirão que o pesquisador Aluizio Felipe Miranda da Silva tenha acesso à visualização e edição das configurações dos serviços de rede do campus Palmares, além de permitir acesso de visualização aos dados sensíveis de todos os setores, para fins de avaliação do tipo, tamanho e frequência de manipulação dos arquivos. Os srs. tem liberdade de se recusar a participar e ainda se recusar a continuar participando em qualquer fase da pesquisa, sem qualquer prejuízo para os mesmos. Sempre que quiserem poderão pedir mais informações sobre a pesquisa através do telefone do pesquisador do projeto e, se necessário, através do telefone do Comitê de Ética em Pesquisa.

4. Riscos e desconforto: a participação nesta pesquisa não traz complicações legais. Por se tratar de uma pesquisa que deverá conceder acesso administrativo à servidores de rede e dados sensíveis, existe o mínimo risco de comprometimento da integridade da informação e até mesmo perda da mesma. Porém é importante ressaltar que os acessos pelo pesquisador serão monitorados pelo coordenador de TI da instituição, e que o próprio pesquisador é integrante da equipe de TI, portanto, é um servidor público lotado na organização. Fato que faz com que o pesquisador esteja sob pena em caso de infrações administrativas, como estabelecido no o artigo 127 da Lei no 8.112/90. Os procedimentos adotados nesta pesquisa obedecem aos Critérios da Ética em Pesquisa com Seres Humanos conforme Resolução no. 196/96 do Conselho Nacional de Saúde. Nenhum dos procedimentos usados oferece riscos à sua dignidade.

5. Confidencialidade: todas as informações sensíveis coletadas neste estudo são estritamente confidenciais. Somente o pesquisador e o (a) orientador (a) terão conhecimento dos dados. Podendo serem publicados na pesquisa dados que não comprometam a organização agora ou no futuro.

7. Benefícios: ao participar desta pesquisa os srs. não terão nenhum benefício direto. Entretanto, esperamos que este estudo traga informações importantes sobre aplicabilidade de soluções de backup de forma eficiente e eficaz em outros ambientes, como os demais campi do Instituto Federal de Pernambuco, de forma que o conhecimento que será construído a partir desta pesquisa possa ajudar tanto

os profissionais de TI iniciantes, dando um ponto de partida, quanto profissionais mais experientes, diminuindo a complexidade do processo de escolha de uma solução, onde pesquisador se compromete a divulgar os resultados obtidos.

8. Pagamento: os srs. não terão nenhum tipo de despesa para participar desta pesquisa, bem como nada será pago por sua participação.

Após estes esclarecimentos, solicitamos o seu consentimento de forma livre para participar desta pesquisa. Portanto preencha, por favor, os itens que se seguem.

Obs: Não assine esse termo se ainda tiver dúvida a respeito.

Consentimento Livre e Esclarecido

Tendo em vista os itens acima apresentados, eu, de forma livre e esclarecida, manifesto meu consentimento em participar da pesquisa. Declaro que recebi cópia deste termo de consentimento, e autorizo a realização da pesquisa e a divulgação dos dados obtidos neste estudo.

Assinatura do Participante da Pesquisa

Assinatura do Participante da Pesquisa

Assinatura do Pesquisador

Assinatura do Orientador

Pesquisador Principal: Aluizio Felipe Miranda da Silva.
Comitê de Ética em Pesquisa: Rua Dom Manuel de Medeiros, s/n - Dois Irmãos, Recife/PE, 1o andar do Prédio Central da Reitoria da UFRPE (ao lado da Secretaria Geral dos Conselhos Superiores).
Telefone do Comitê: (81) 3320-6638
E-mail do Comitê: cep@ufrpe.br