

SIGILOSO

Modelo Tecnológico

Solução de Backup Remoto

Cliente: SERPRO

Versão 1.6

HISTÓRICO DE VERSÕES

Versão	Data	Autor da Versão	Alterações da Versão
1.6	22/03/18	Euriam Barros de Araújo	<ul style="list-style-type: none"> – Inclusão do nome DNS do TSM Server (item 3); – Revisão da política de backup (item 4); – Alteração da Tabela 1 e inclusão de texto sobre porta utilizada (item 3).
1.5	03/02/17	Euriam Barros de Araújo	<ul style="list-style-type: none"> – Alteração do item 2.3.
1.4	20/12/16	Diogo Jacobs Flávio Xavier Mota Euriam Barros de Araújo	<ul style="list-style-type: none"> – Complementação da introdução (item 1).
1.3	07/11/16	Euriam Barros de Araújo	<ul style="list-style-type: none"> – Alteração da topologia, incluindo replicação e armazenamento em fita (Brasília e São Paulo); – Inclusão da política de backup, retenção e restauração de dados.
1.2	10/08/16	Euriam Barros de Araújo	<ul style="list-style-type: none"> – Atualização dos padrões de nomenclatura (item 3).
1.1	15/06/16	Euriam Barros de Araújo	<ul style="list-style-type: none"> – Alteração da topologia (item 2.2) e nos padrões de nomenclatura (item 3).
1.0	10/06/16	Euriam Barros de Araújo	<ul style="list-style-type: none"> – Versão original.

RESUMO DO DOCUMENTO

Descrição:	Este documento descreve a arquitetura técnica do serviço de Backup Remoto a partir dos Escritórios e Regionais, com replicações de dados e armazenamento em fitoteca nas Regionais Brasília e São Paulo.
Local de Publicação:	GTIC SUPOP → MGADOP → SUPOP → Administração de Rede Local → Modelos Tecnológicos/Manuais de Configuração e Operação
Validade da Versão:	22/03/2019
Baseado no Modelo de Publicação Versão 2.1	

SUMÁRIO

Histórico de Versões.....	2
Resumo do Documento.....	2
1. Introdução.....	4
1.1. Contexto.....	4
1.2. Premissas.....	5
1.3. Escopo.....	5
2. Definição da Solução.....	5
2.1. Especificações Técnicas.....	6
2.2. Topologia.....	7
2.3. Armazenamento em Storage.....	9
3. Padrões de Nomenclatura.....	10
4. Política de Backup, Retenção e Restauração de Dados.....	11
4.1. Tipos de Arquivos Excluídos da Rotina de Backup.....	11
4.2. Tipos de Retenção.....	11
4.3. Replicação e Armazenamento em Fita:.....	11
4.4. Restauração de Dados.....	11
5. Segurança da Solução.....	12
6. Documentos Correlacionados.....	12
Glossário.....	13
Referências Bibliográficas.....	14
Ficha Técnica.....	15

1. INTRODUÇÃO

A Superintendência de Operações (SUPOP), na busca de manter contínua modernização, otimização e adequação do seu ambiente tecnológico no que diz respeito ao armazenamento de dados, construiu uma solução de *backup* e *restore* focando nos aspectos de gerenciamento, redundância de dados armazenados, mitigação do ponto único de falha¹, RTO (tempo máximo de recuperação) e RPO (tolerância para perda de dados – medida em tempo), recursos de deduplicação², *archiving*³, replicação e administração/gerenciamento centralizado.

Essa Solução, nomeada como Backup Remoto, atenderá as Regionais e Escritórios do SERPRO, mas também poderá atender os Clientes, sob contrato, de acordo com as necessidades dos mesmos.

Os dados provenientes da realização dos backup dos Escritórios do SERPRO trafegarão pela rede WAN⁴ e serão armazenados e administrados/gerenciados pela solução instanciada na Regional com menor custo de rede. Por exemplo, os backups provenientes dos Escritórios de João Pessoa e Maceió serão armazenados na Regional Recife. Por fim, os backup armazenados em *storage* das Regionais serão replicados (via rede WAN) para *storage* das Regionais São Paulo e Brasília (de acordo com item 2.2), e em sequência realizado armazenamento em fita em equipamentos (fitoteca) compartilhados com a SUPCD.

O Backup Remoto é composto pelo software IBM *Spectrum Protect (Tivoli Storage Manager)*⁵ e por servidores e subsistema de discos *midrange – storage*, sob administração da SUPOP, e fitoteca, sob administração da Superintendência de Centro de Dados (SUPCD).

1.1. CONTEXTO

Para o correto funcionamento do Backup Remoto, as seguintes premissas devem ser seguidas:

- O servidor TSM de cada Regional deve ser alocado na respectiva ZDM Serviços de Clientes;
- Cada instância Regional da SUPOP é responsável pela administração do seu respectivo TSM Server (componente do TSM que administra a solução);
- Cada Regional deve ter um hardware dedicado e exclusivo (Dell PowerEdge R730) para suportar o TSM Server;
- O armazenamento dos backups das Regionais Brasília e São Paulo são migrados posteriormente e gradualmente para o respectivo *storage*, pois as mesmas já executam backup utilizando o software TSM com o armazenamento dos dados em ambiente da SUPCD;
- Os processos de backup são agendados e executados automaticamente, após o término do horário de expediente diurno de cada Regional, salvo excepcionalidades;

Os processos de *restore* serão executados mediante solicitação formal e durante o expediente diurno.

1 Uma única cópia dos dados esteja disponível e ocorra falha na mesma.

2 Evita a cópia de dados redundantes.

3 Armazenamento de dados de longa duração.

4 Respeitando a largura de banda e a taxa de utilização da rede.

5 O contrato firmado com a IBM prevê o uso de licenças ilimitadas.

1.2. PREMISSAS

Em virtude de limitações de determinados ativos, seguem restrições do serviço:

- Não há garantia da manutenção do desempenho de processamento dos servidores de redes, durante a execução dos processos de backup, principalmente quando habilitado e utilizado o recurso de deduplicação;
- Como essa solução foi homologada para backups remotos utilizando links WAN a partir de 34 Mbps (entre SERPRO e Escritórios), para canais de comunicação com velocidades inferiores, não é possível garantir a conclusão do processo de backup em menos de 24 horas.

1.3. ESCOPO

Este Modelo Tecnológico abrange os servidores administrados pela SUPOP nas Regionais e Escritórios do SERPRO.

2. DEFINIÇÃO DA SOLUÇÃO

A Solução do Backup Remoto é composta pelo software TSM, hardwares dedicados (servidores de rede, para suportar o TSM Server, e *storages*) e equipamentos de fita automatizados disponibilizados pela SUPCD (nas Regionais Brasília e São Paulo), além da infraestrutura de rede WAN da Rede SERPRO para servir como meio de tráfego dos backups.

Os dados provenientes dos servidores de rede localizados no ambiente interno em cada Regional, e contemplados na rotina de backup, trafegarão pela rede LAN com destino ao TSM Server Regional, cujo armazenamento ocorrerá no *storage* da respectiva Regional. Posteriormente, os dados armazenados em *storage* serão replicados para as Regionais Brasília ou São Paulo. Já os dados provenientes dos backups das localidades remotas (Escritórios do SERPRO), trafegarão pela rede WAN e serão armazenados em *storage* da Regional a qual está fisicamente vinculada via circuito de dados. A parte final do ciclo de armazenamento de dados se dará com a replicação dos backups de cada *storage*, via WAN, para área alocada também em *storage* nas Regionais Brasília e ou São Paulo, cujas localidades disponibilizarão equipamentos de fita automatizados para etapa final do armazenamento dos backups.

A fim de otimizar o volume de dados provenientes dos backups, uma funcionalidade imprescindível do TSM que será utilizada é a deduplicação de dados por blocos, que só transmitirá e armazenará blocos de informações que ainda não se encontram armazenadas, evitando a redundância e desperdício de espaço ocupado em mídia de backup.

Os processos de backup e *restore* serão administrados e gerenciados pela equipe técnica da SUPOP de cada Regional.

2.1. ESPECIFICAÇÕES TÉCNICAS

A infraestrutura técnica da Solução de Backup Remoto é composta por vários ativos de hardware e software, distribuídos entre as Regionais do SERPRO.

Abaixo segue o detalhamento dos componentes da solução:

- Hardware (por Regional)
 - Todas as Regionais (exceto Brasília e São Paulo): Dell PowerEdge R730 com 02 processadores de 08 núcleos cada, 128 GB RAM, disco rígido de 3 TB em RAID-5. Host que hospedará o TSM Server;
 - Regionais Brasília e São Paulo: Dell PowerEdge R730 com 02 processadores de 08 núcleos cada, 128 GB RAM, disco rígido de 300 GB em RAID-1. Host que hospedará o TSM Server;
 - Dispositivos de armazenamento
 - Serão utilizados como dispositivos de armazenamento em fita os equipamentos automatizados alocados na SUPCD de Brasília e São Paulo, os quais se comunicarão com os respectivos servidores TSM Server com conectividade por switch SAN;
 - Infraestrutura de rede
 - Rede WAN (34 Mbps): rede de longa distância que liga escritórios e respectivas Regionais, permitindo a realização de backup remoto;
 - Rede LAN (1 Gbps): rede local a ser utilizada para o tráfego de dados de backup a partir dos servidores de rede localizados no ambiente das Regionais.
- Software
 - Software de backup
 - TSM Client¹: a ser instalado no servidor de rede contemplado na rotina de backup;
 - TSM Server²: responsável por administrar e gerenciar a solução;
 - Sistema operacional
 - Linux Red Hat³: sistema que hospedará o TSM Server em cada Regional.
- Serviços
 - Desduplicação *inline*: processo executado imediatamente anterior ao de backup que otimiza e reduz os dados redundantes a serem armazenados;
 - Processo de backup: processo de cópia de dados dos servidores de rede contemplados na rotina de backup;
 - Processo de restore: processo de restauração de dados de backup, de acordo com solicitação formal, com destino a área em disco definida a esse fim.

1 A versão utilizada será a 5.3.6 ou superior

2 A versão utilizada será a 7.1.4 ou superior

3 Em virtude da disponibilidade de licenças, pode ser utilizado sistema operacional compatível, como o CentOS

2.2. TOPOLOGIA

A topologia do Backup Remoto encontra-se detalhada abaixo na Ilustração 1.

BACKUP TSM	De	Belo Horizonte	Belém	Boa Vista	Macapá	São Luís	Brasília	Cuiabá	Goiânia	Porto Velho	Curitiba	João Pessoa	Maceió	Recife	Fortaleza	Natal	Teresina	Florianópolis	Porto Alegre	Rio de Janeiro	Vitória	Aracaju	Salvador	Campo Grande	Manaus	Palmas	Rio Branco	São Paulo
	Para	BHE		BLM			BSA			CTA		RCE			FLA		FLS	PAE	RJO		SDR					SPO		
	Replica e B. Fita	BSA												SPO														

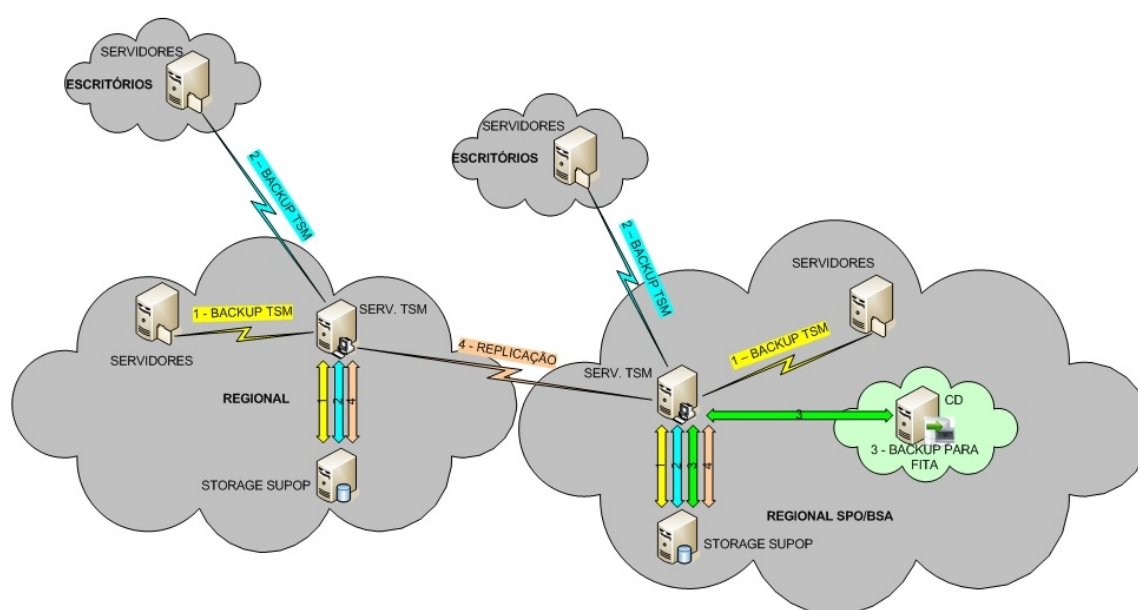


Ilustração 1 - Topologia do Backup Remoto

As Regionais (nuvens inferiores), realizarão backup dos servidores da rede local (LAN) contemplados na rotina de backup, como também dos servidores RLSL (Rede Local de Software Livre) dos respectivos Escritórios (tabela superior). Esses backups serão armazenados no *storage* da respectiva Regional, e de acordo com a política de backup (item 4), e com exceção de Brasília e São Paulo, os dados serão replicados periodicamente para Brasília (Belo Horizonte, Belém, Curitiba e Recife) ou São Paulo (Fortaleza, Florianópolis, Porto Alegre, Rio de Janeiro e Salvador).

Em se tratando das Regionais Brasília e São Paulo, que receberão as replicações dos backups das demais Regionais, os dados serão mantidos no respectivo *storage* por tempo determinado até posteriormente serem realocados em fita (*tape library* ou fitoteca local) da SUPCD, conforme política de retenção (item 4).

Para todas as Regionais (exceto Brasília e São Paulo):

A arquitetura do TSM Server, hospedado em um servidor físico dedicado, contém no disco rígido interno, as seguintes partições específicas da solução: **tsm_db** (bancos de dados DB2) e **tsm_activelog** e **tsm_archive**.

O *storage* o qual o TSM Server está conectado e que armazenará os backups, possui a partição **tsm_backup_db** (armazena cópia redundante do **tsm_db**) e os seguintes pools:

POOL_DEDUP: armazenará os backups tipo “diário” e “mensal” deduplicados, inclusive das localidades (escritórios) remotas associadas; os backups tipo “mensal” serão periodicamente replicados (*replicate node*), via rede WAN, com destino ao POOL_DISCO alocado em *storage* remoto (Brasília ou São Paulo);

POOL_COPIA: armazenará cópia do POOL_DEDUP com os dados originais – não deduplicados;

POOL_REMOTO: redirecionará (*virtual volumes*) os backups tipo “anual” (inclusive das localidades remotas associadas) para o POOL_DISCO armazenado remotamente em *storage* de Brasília ou São Paulo.

Para as Regionais Brasília e São Paulo:

A arquitetura do TSM Server conterá no seu disco rígido interno apenas **tsm_backup_db** (armazena cópia redundante do **tsm_db**) como partição específica da solução.

O *storage* o qual o TSM Server está conectado e que armazenará os backups, possui as partições: **tsm_db** (bancos de dados DB2), **tsm_activelog**, **tsm_archive** e os seguintes pools:

POOL_DEDUP: armazenará os backups tipo “diário” e “mensal” deduplicados; os backups do tipo “mensal” serão movidos (*move nodedata*) para o POOL_LTO6;

POOL_DISCO: armazenará os backups tipo “mensal” replicados e “anual” das Regionais remotas associadas, além do backup tipo “anual” da sua própria Regional e seus escritórios associados; os dados serão migrados (*migrate stg*) para o POOL_LTO6;

POOL_LTO6: armazenará em fita (fitoteca da SUPCD de Brasília e São Paulo) todos os dados provenientes do POOL_DISCO e os nodes mensais provenientes do POOL_DEDUP.

A Figura 2 ilustra as arquiteturas supracitadas.

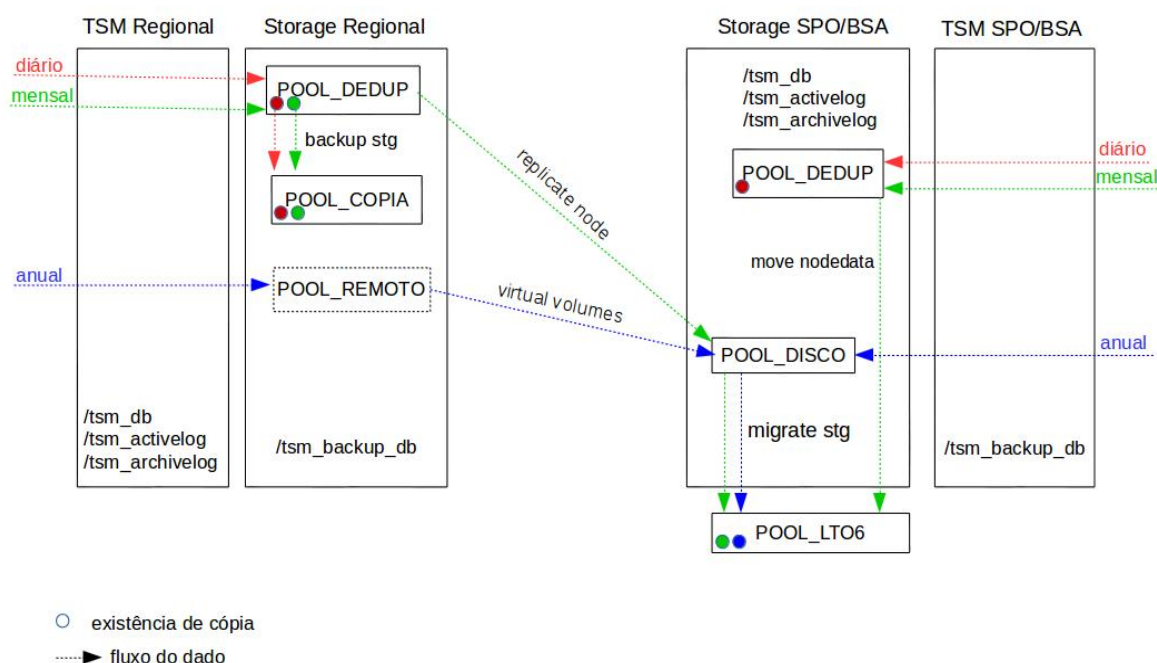


Figura 2: Arquitetura do IBM Spectrum Protect (TSM)

2.3. ARMAZENAMENTO EM STORAGE

De acordo com a topologia do Backup Remoto (Figura 1), os backups são armazenados no *storage* da respectiva Regional.

O tamanho e a alocação desse espaço lógico está definido no Modelo Tecnológico rede SAN – Storage EMC VNX 5400¹, porém de acordo com o crescimento gradual do volume de dados de backup, pode ser necessária a expansão do espaço inicialmente/anteriormente alocado.

¹ <https://portalgtic.serpro.gov.br/clientes/gtic/supop/mgadop/supop/administracao-de-rede-local/modelos-tecnologicos-manuais-de-configuracao-e-operacao/55.6.5.Modelo%20Tecnologico%20SAN%20v1.5.pdf>

3. PADRÕES DE NOMENCLATURA

Os *nodes* dos TSM Clients instalados nos servidores das Regionais e Escritórios, que possuem uma identificação única para cada tipo de backup, devem adotar o seguinte padrão de nomes:

UF_CLIENTE_CNLSERVICO_TIPO

Onde:

- **UF**: Sigla do Estado ou do Distrito Federal (PE, SP, DF, etc)
- **CLIENTE**: Nome do Cliente (SERPRO)
- **CNL**: Código de três caracteres das cidades (RCE para Recife, NTL para Natal, SPO para São Paulo, etc)
- **SERVICO**: Nome do principal serviço hospedado no *host* contemplado na rotina de backup (RLSL, DIRETORIO, FREERADIUS, etc)
- **TIPO**: referente ao tipo de retenção do backup (DIARIO, MENSAL ou ANUAL)

A porta utilizada para administração, comunicação, transporte de dados de backup (TSM Server e TSM client) e replicação de backup (entre TSM Servers) é a **1950/tcp**.

Quanto aos TSM Servers das Regionais, para o correto funcionamento da Solução de Backup Remoto, esses devem seguir o padrão da Tabela 1.

<i>Regional</i>	<i>Nome do TSM Server</i>
Recife	TSM_OPRCE
Curitiba	TSM_OPCTA
Belo Horizonte	TSM_OPBHE
Salvador	TSM_OPSDR
Fortaleza	TSM_OPFLA
Florianópolis	TSM_OPFNS
Porto Alegre	TSM_OPPAE
Belém	TSM_OPBLM
Rio de Janeiro	TSM_OPRJO
São Paulo	TSM_OPSP0
Brasília	TSM_OPBSA

Tabela 1: Padrões para uso do TSM (Tivoli Storage Manager)

O nome DNS do TSM Server seguirá o padrão conforme item 3.5 do documento **Manual Operacional – Padrões de Nomenclatura**¹.

¹ https://portalgtic.serpro.gov.br/clientes/gtic/supop/mgadop/geral/administracao-de-rede-local/modelos-tecnologicos-manuais-de-configuracao-e-operacao/51.2.ManualOperacional_PadroesNomenclatura_v1.3.pdf

4. POLÍTICA DE BACKUP, RETENÇÃO E RESTAURAÇÃO DE DADOS

4.1. TIPOS DE ARQUIVOS EXCLUÍDOS DA ROTINA DE BACKUP

Como pré-configuração do IBM *Spectrum Protect* (TSM), não será realizado backup dos seguintes formatos de arquivos de mídia: mp3, mp4, avi, mkv, wmv e mpg.

Em caso da necessidade da realização de backup de arquivos com quaisquer desses formatos, deve ser solicitada formalmente (com a justificativa da necessidade) para a devida inclusão da(s) extensão(ões) na rotina de backup, e o tempo de permanência.

4.2. TIPOS DE RETENÇÃO

- ✓ **Backup diário:** o armazenamento ocorrerá em *storage* da respectiva Regional e obedecerá a retenção de 35 dias;
- ✓ **Backup mensal:** o armazenamento ocorrerá em *storage* da respectiva Regional e obedecerá a retenção 12 meses;
- ✓ **Backup anual:** o armazenamento ocorrerá em *storage* das Regionais Brasília e São Paulo, migrado para armazenamento em fita, quando obedecerá a retenção de 05 anos.

4.3. REPLICAÇÃO E ARMAZENAMENTO EM FITA:

- ✓ **Backup mensal (exceto Brasília e São Paulo):** após os backups, os dados serão replicados para *storage* das Regionais Brasília ou São Paulo e, na sequência, migrados para fita;
- ✓ **Backup mensal (Brasília e São Paulo):** serão periodicamente migrados de *storage* para fita;
- ✓ **Backup anual:** backups armazenados em *storage* de Brasília ou São Paulo e periodicamente migrados para fita.

4.4. RESTAURAÇÃO DE DADOS

Após a solicitação formal da restauração, o(s) dado(s) será(ão) alocado e disponibilizado em espaço reservado do *storage* da respectiva Regional solicitante.

Caso o(s) dado(s) esteja(m) armazenado(s) remotamente (*storage* ou fitoteca de Brasília ou São Paulo), o(s) mesmo(s) trafegará(ão) via rede WAN até a Regional solicitante.

5. SEGURANÇA DA SOLUÇÃO

Em função dos vários ativos de software e hardware que compõem a solução, alguns riscos de segurança devem ser destacados:

- Caso o TSM Client dos servidores de rede contemplados na rotina de backup não esteja em execução ou desativado, o processo de backup não pode ser realizado;
- Caso haja problema de comunicação com a rede WAN, os processos de backup remoto entre Escritórios e Regional ficam impossibilitados de execução;
- Deve haver disponibilidade de fita nas Regionais Brasília e São Paulo para armazenamento dos backups mensais e anuais.

6. DOCUMENTOS CORRELACIONADOS

[1] SERPRO/SUPOP/OPGTC/OPTSE. Manual de Instalação e Configuração – IBM Spectrum Protect – TSM Server, versão 1.1, 2018.

[2] SERPRO/SUPOP/OPGTC/OPTSE. Manual de Instalação e Configuração – IBM Spectrum Protect – TSM Client, versão 1.1, 2018.

[3] SERPRO/SUPOP/OPGTC/OPTSE. Manual Operacional do IBM Spectrum Protect – TSM, versão 1.0, 2018.

GLOSSÁRIO

Desduplicação: processo de eliminação de dados ou bloco de dados duplicados.

Desduplicação *inline*: processo de eliminação de dados ou bloco de dados duplicados, imediatamente anterior ao processo de backup.

Archiving: armazenamento e retenção de dados por longa duração, com possibilidade de exclusão dos dados da origem.

REFERÊNCIAS BIBLIOGRÁFICAS

IBM Spectrum Protect. Disponível em:

<<https://www.ibm.com/br-pt/marketplace/data-protection-and-recovery>>. Acesso em:
abril de 2017.

FICHA TÉCNICA

SERVIÇO FEDERAL DE PROCESSAMENTO DE DADOS

Diretora Presidente

Maria da Glória Guimarães dos Santos

Diretor de Operações

Iran Martins Porto Júnior

Superintendente de Operações – SUPOP

Gilberto de Oliveira Netto

Departamento de Gestão dos Serviços de Rede de Acesso e Comunicação – SUPOP/OPGTC

Weldson Queiroz de Lima

Divisão de Tecnologias de Acesso e Comunicação – SUPOP/OPGTC/OPTAC

Régison Rodrigo Martins

Rede de Suporte Técnico de Linux e Virtualização

Paulo Roberto dos Santos Arruda

Elaboração

Euriam Barros de Araújo – DIOPE/SUPOP/OPGTC/OPTSE

Colaboração

Diogo Jacobs – DIOPE/SUPCD/CDINF/CDARS

Flávio Xavier Mota – DIOPE/SUPES/ESTAR/ESAIT

Revisão da Versão

Priscilla Lanne Cordeiro de Leao – DIOPE/SUPOP/OPGTC/OPTAC

Versão 1.6

março / 2018