Availability, Disaster Recovery and Data Protection Technology Strategy

CIA rating and Availability

Summary PPTX Deck

Strategy document in word format

Reference Documents/Standards

Cybersecurity Application and Supporting Infrastructure Classification Standard	https://spo.bhpbilliton.com/sites/GRPBISTECHSTDS/Collection/Technology%20Cybersecurity%20Application%20and%20Supporting%20Infrastructure%20Classification%20Standard%20TSS00005.pdf
Cyber Security Backup Standard	https://spo.bhpbilliton.com/sites/GRPBISTECHSTDS/Collection/Forms/AllItems.aspx?id=%2Fsites%2FGRPBISTECHSTDS%2FCollection%2FTechnology%20Cybersecurity%20Backup%20and%20Restoration%20Standard%20TSS00008%2Epdf&parent=%2Fsites%2FGRPBISTECHSTDS%2FCollection
Cyber Security Immutability Standard	https://spo.bhpbilliton.com/sites/GRPTECHTechInfra/technicalwiki/SitePages/Immutable-Backups.aspx
Service Continuity Management and Disaster Recovery Standard	https://spo.bhpbilliton.com/sites/GRPBISTECHSTDS/Collection/Technology%20Service%20Continuity%20Management%20and%20Disaster%20Recovery%20Standard%20TOM00035.pdf?CT=1649725965324&OR=OWA-NT&CID=d0757580-3465-1277-79d5-9db9271c7c28
Cloud Workload Placement Standard (draft)	https://spo.bhpbilliton.com/:p:/r/sites/cloud/_layouts/15/Doc.aspx?sourcedoc=%7BB3063481-6C6F-4DEA-B95B-78190CD5A55F%7D&file=Workload%20Placement%20Cloud%20v2.7.pptx&action=edit&mobileredirect=true&DefaultItemOpen=1&cid=9f1ce28a-f312-4699-9284-1d4ea57c136b
Cloud Platform Security Framework	https://spo.bhpbilliton.com/sites/CyberSecure/Shared%20Documents/Cloud%20Security%20Framework/BHP_Cloud%20Platform%20Security%20Framework.pdf
Information Governance and Controlled Documents	https://spo.bhpbilliton.com/sites/DW/GLD/Information%20Governance%20and%20Controlled%20Documents.pdf
Information Protection Framework Guidance Note	https://spo.bhpbilliton.com/sites/DW/GLD/Information%20Governance%20and%20Controlled%20Documents%20-%20Information%20protection%20framework%20guidance%20note.pdf
Database Backup and Recovery Standard	https://spo.bhpbilliton.com/sites/GRPBISTECHSTDS/Collection/Forms/AllItems.aspx?id=%2Fsites%2FGRPBISTECHSTDS%2FCollection%2FTechnology Database Backup and Recovery Standard TECH125803%2Epdf&parent=%2Fsites%2FGRPBISTECHSTDS%2FCollection
Incident Management Standard	https://spo.bhpbilliton.com/sites/GRPBISTECHSTDS/Collection/Technology%20Incident%20Management%20Standard%20TOM00016.pdf
Cloud Application Monitoring Standard	https://spo.bhpbilliton.com/sites/GRPBISTECHSTDS/Collection/Technology%20Cloud%20Application%20Monitoring%20Standard%20APP00008.pdf
Technology Cybersecurity Standard & Control Verification Procedure	https://spo.bhpbilliton.com/sites/GRPBISTECHSTDS/Collection/Technology%20Cybersecurity%20Standard%20%26%20Control%20Verification%20Procedure%20TSP00003.pdf

Other Reference Materials and Contributing Content

Cloud Backup Guidance

Guide to Managing Availability on Azure

Azure Backup assessment by MSFT

AWS backup assessment by AWS

Cloud Factory Offerings

Scope and Design Considerations for Enterprise Backup Solution

these lists of AWS and Azure services need to have a backup and restore methodology defined as an outcome of the project. **Before we can select a tool, these services need to be refined to map out what services would be backed up /restore using a tool and what is configuration/IAC as not everything here is relevant.**	**https://spo.bhpbilliton.com/sites/cloud/Shared Documents/Forms/AllItems.aspx?id=%2Fsites%2Fcloud%2FShared Documents%2FStrategy%2FCloud Components Strategy&viewid=11680601-d6ed-43ad-b40c-5395768f7743**	**Azure Cloud Services - Backup and Restore Matrix**	**AWS Cloud Services - Backup and Restore Matrix**	**any other services that are supported by cloud ops or a centralised team that will need to provide and backup and restore story for their service as it is a core dependency for other apps or other recovery plans.** **gitlab** **github** **Active directory** **splunk** **aviatrix** **certificate services**
is the scope continuing to be cloud backup and restore, or will it include on-premises and edge?				
is there a need to consider future state requirements for edge compute?				
how will the tool selection be weighted? preferred IAAS backup and recovery against future state PAAS backup and recovery?				
how important is the operational simplicity in the approach, does it have more weight that cost?				
the azure US regions are not paired and so replication of backup data needs to be done at the backup solution layer, but other azure regions use the region pair (AU EAST/AUSE). do we replicate at the solution/tool layer everywhere (including AWS) or do we use the platform replication where we can?				
a detailed RACI for backup and restore needs to be developed and agreed as a deliverable of the backup/restore program of work.				
Who owns the data, what are the boundaries of duty between ops team and app teams?				
is BHP interested in splitting out the ransomware protection and recovery solution to a separate product, for example look at rubrik as this product which is aligned and focused on cyber protection more than traditional backup and restore and leave the traditional backup and restore to something else.				

There needs to be clear distinction between BCP (Business Continuity Planning) and DR (Disaster Recovery). Note that a declared disaster is usually greater than (or a major conflation of) the incidents covered in the Technology Major Incident Management Procedure and recommended practice is to have a prescriptive testable disaster recovery process that is separate from normal BCP activities. I would remove all references to “disaster” and “disaster recovery”. SLAs and CIA ratings do not normally apply in disaster scenarios. Backup and restore should not be the default recommended availability strategy for BCP of cloud-based assets with the exception of data that has immutability requirements and is contained within systems of record. With the recommendation of using Infrastructure as Code and shift away from Click-Ops, redeployment and reingestion of data from highly available datastores is the quickest, most secure, and least error prone process of recovering system availability.				
As mentioned in the listed questions - responsibility of availability needs to be established between cloud provider (for cloud platform and infrastructure), telecommunications providers (for network access), ops (for virtual infrastructure and managed platforms), and application owners (for applications). Availability of cloud platform and infrastructure is publicly available for all major cloud vendors.				

Backup & Recovery capabilities

Technical features, capabilities and other requirements in no particular order rated, however will need further refinement.

Recovery options	**tools and workload to enable the:** 1. Recovery of entire system, instance, or application 2. Granular single file recovery 3. In-place to same location or out-of-place to different location 4. Latest data in point-in-time		**Must Have**
multiple agents to support cloud services, applications and file systems	Agents are software modules that are installed on computers to access and protect data. The backup and recovery system uses agents to interface with file systems, applications, and databases to facilitate the protection of data on production systems.	file system agents application agents DB agents Archive Agents	Must Have
Hight availability designs that complement regions and AZs to provide availability for the data protection infrastructure.	clustering or distributed deployment model that works with AZ’s to provide HA for the solution for each region		Must Have
Deduplicated data for more efficient data transmission and to lower costs of moving data around on cloud platforms	Source-Side (Client-Side) Deduplication MediaAgent-Side (Storage-Side) Deduplication Global Deduplication Accelerated (deduped) Synthetic Full Backups		Must Have
Policy-driven automation, monitoring and reporting.	out of the box reports that can be attached to the protect workload via policy and automation, see bottom of table for reports.		Must Have
policy based but customized schedules to meet any SLA			Must Have
Customizable short-term retention Customizable long-term retention			Must Have
Support for all major BHP cloud vendors AWS and Azure multiple storage format support, on-premises, cloud storage automatic cloud tiering across hot, cool and archival cloud storage based on policies.	Cloud/Object-based Storage: Amazon Simple Storage Service (S3 Standard, S3-IA, S3-RRS, S3 OZ, S3 Intelligent Tiering) Amazon Elastic File System (EFS) Microsoft Azure Blob Storage Microsoft Azure File Storage		Must Have

Support for all major snapshot hardware vendors Automated snapshot backup and recovery Customized snapshot retention			Must Have
Cloud / Infrastructure as a Service (IaaS)	Amazon Elastic Compute Cloud (EC2) Microsoft Azure Virtual Machines		Must Have
No downtime to production systems or minimized impact to cloud resource during the backup process			Should Have
Data portability between clouds	backups taken in one cloud provider should be transferable to another provider and be able to be restored, azure to aws and aws to azure.		Must Have
Avoid vendor lock-in			Should Have
Self-service Integrations and Automation	PowerShell Module TerraForm Web base portal for self-service access for project or app teams integration into service now to create incident tickets		Must Have
Identify and Access Management integration	CyberArk	credentials should integrate with CyberArk and support password rotation.	Must Have
Identify and Access Management	granular RBAC Azure Active Directory	1. permissions should be granular and defined as roles 2. roles should be assigned to identities 3. identity source should be with AD or AAD	Must Have
3rd Party Key Management	Amazon KMS HashiCorp Vault Microsoft Azure Key Vault		Must Have
Security Information and Event Management (SIEM) & Security Orchestration, Automation, and Response (SOAR)	BHP Splunk Azure event hubs and log analytics integration AWS Cloudtrail and Cloudwatch-logs		Must Have
Hypervisor Platforms:	Amazon Outpost Microsoft Azure Stack Microsoft Hyper-V VMware vSphere		Should Have Could Have

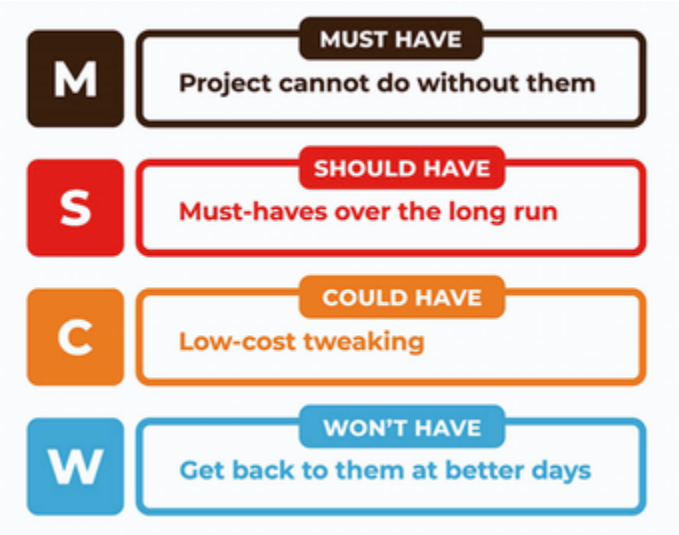
Container Platforms	Docker Kubernetes – Hosted (Amazon EKS, Azure Kubernetes Service (AKS), Kubernetes – PaaS	even though containers should be stateless, there may be use cases where there is data on the container that need to be protected. there will also be storage mounted to containers, this storage can be protected through the storage API and not the container, but there does need to be a level or orchestration between container and storage to ensure app consistency. Also, data protection can be used to migrate containers between blue/green environments and between clusters in different clouds - so BHP may want to explore this option.	Could Have
Filesystems	Windows File System NAS Netapp NFS shares UNIX/Linux File Systems SMB/CIFS shares		Must Have Should Have Should Have Must Have Must Have
Cloud Database (DBaaS):	Azure SQL Database Azure Database for PostgreSQL Azure Database for MySQL Azure Database for MariaDB Azure Cosmos DB Amazon RDS for Microsoft SQL Amazon RDS for Oracle Amazon RDS for PostgreSQL Amazon RDS for MySQL Amazon RDS for MariaDB Amazon Aurora Amazon DynamoDB Amazon Redshift Amazon DocumentDB		Must Have
Enterprise Databases	Microsoft SQL Server - SQL clusters, SQL Availability Group MySQL? Oracle? Oracle RAC? PostgreSQL? SAP HANA?		Must Have
Enterprise Applications:	Active Directory Splunk, Aviatrix		Must Have
Software as a Service (SaaS) Applications:	gitlab github Office 365 - Exchange Online? Office 365 - Sharepoint Online? Office 365 - OneDrive for Business? Office 365 - Microsoft Teams?		Must Have

data protection tool has a scale out and scale up methodology for azure cloud and AWS cloud	there should be a detailed and documented scale up approach and scale out approach for each cloud topology which address scaling of the solution to cope with backup and restore load		Must Have
periodic replication or continuous replication	Replication Monitor Replication alerting Replication across regions within the solution to work around azure region pair replication.	Periodic - replication that is performed on a scheduled basis. Continuous - block-level replication that is performed continuously. granular replication scope defined through policy	Must Have
sizing of hardware and storage calculator	tool or documented approach for sizing the solution components, network and storage		Must Have
DR/Replication Sizing Calculator	tool or documented approach for sizing a replication use-case		Must Have
Two-Factor Authentication	Two-Factor Authentication integration with BHP identity system		Must Have
Web Console access for remote users, end-users or self-service users.	Google Chrome Microsoft Edge		Must Have
immutable storage	immutable vault that is backed by cloud storage service that is natively integrated and outside the BHP tenancy and cloud environment. Access to the vault is through a service principle that is used 1 time at setup to establish the vault, the backup solution uses this security context to copy data into the vault, RBAC applied to this one service principle to perform this role. all other roles and permissions do not have any ability to change the immutable vault, only read and restore access - a type of break glass account. auditing of immutable vault process, access and admin process logged and ingested to Splunk SIEM encryption in transit and rest automatic and integrated workflow to copy immutable backups in vault based on tags configured via policy on protected workloads, no special policy or action required, the backup into vault is just another storage location. solution identical across AWS and Azure - solution should be the same across the clouds with the only change being the backing storage.		Must Have

Built-in ransomware protection	analytics tools and dashboards that provide early warning alerts of suspicious and malicious activities: • A single interface to easily monitor, manage, protect, and secure your environment. • Actively monitor for abnormal activities for more signicant insights, alerting, and faster response. • Fastest detection of ransomware and other suspicious activities. • Track user accountability by monitoring all resources and activities. dual authorization to implement changes to process or data. multi-factor authentication (MFA) options automatically updated credentials for backup jobs if changed in the management console isolating networks and data management away from the central backup infrastructure through the use of proxies or appliances in each landing zone with well-defined network flows.		Must Have
ransomware recovery to enable restoration of clean data quickly	During recovery events avoid ransomware le reinfections by surgically deleting suspicious or unnecessary les. Automate recoveries with streamlined recovery operations through machine-learning and orchestrated workows. Ensure clean le recoveries by quickly isolating suspected backup copies or restore to a safe location (isolated recovery environments) for deep cleaning or forensic analysis.		Must Have
Configuration Audit reports	Global settings, Basic Properties, Index Cache Properties, Firewall Properties, Network Throttle Properties, Job management configurations, Media management properties Browse, search, and recovery parameters		should have
Storage Policy Report	The number and type of storage policies A comparison of retention rules, data paths, media, data verification, and content indexing settings for all storage policies of the same type. Comparison of storage policies across the environments and between instances in each cloud.		should have
Chargeback Report	Chargeback Report provides information about the size of backup jobs and the size and cost of data saved on storage media, it should be filterable based on the data for a particular application or workload, so that at least all of the backup data sets on all storage media for that workload can be reported and charged back. other options if available would include network traffic estimation where the network bandwidth and transfer is reported for charge back as well as other infrastructure cost such as compute time for shared resources such as proxies or appliances.		should have
Activity Report	Successful, partially successful, and failed backup jobs over the last 7 days, including counts for each day. Summary of client sucess, failure, problematic counts and throughput rate. List of top 10 largest clients over the last 30 days. List of the top 10 most frequent errors in the last 24 hours. Successful, partially successful, and failed backup jobs over the last 12 months, including counts for each month. Ability to click on any of the error codes to view a KnowledgeBase article that explains how to resolve the error.		Must Have

Backup Job Summary Report	The number of jobs that completed, failed, were killed, or were delayed. Each backup job that ran during the specified time period, including the size of the job and the percentage of deduplication savings. The number of each type of agent that was backed up over the specified period of time		Must Have
Capacity Usage Growth Report	The amount of space consumed by backed up and archived data over the last year, up to the current time you run the report. A prediction of the amount of data that your organization will consume over the next 6 months, 12 months, and 18 months. The date on which the organization will exceed license capacity, based on the current growth rate.		Must Have
Capacity License and Usage by Client Report	reports on the Front-End data usage. Information includes the front-end size and indicates whether the client is a virtual machine for a workloads or application		should have
Health Report	Summary information about the data protection instance including the status of backup operations (SLA) over the last 30 days and disaster recovery settings. Details about all clients including failure/incomplete cancelled counts, the largest clients, and clients that have jobs with errors. Job statistics, which include backup and restore jobs, long running backup jobs, and the top 10 errors over the last day. Deduplication rates and performance, such as the size of Dedeup DB partitions and replication copies or immutable copies that have fallen behind. Status of the schedule policy, if applicable Usage statistics including the amount of data currently backed up, the number of clients, mailboxes, and virtual machines, and snapshot engine statistics. Capacity information, such as the free space available on mount paths and disk libraries. Value assessment and efficiency of the data protection instance, such as the percentage of backup clients with long running backup jobs, the percentage of backup clients with long running restore jobs, and the percentage of backup clients that are protected over the network and use source side deduplication.		should have
Infrastructure Load Report	The amount of physical memory used the amount of virtual memory used the number of processes running the number of threads used the disk I/O metrics		should have
File Search Report	information about files that are on one or more backup clients, including a file list, file count, file size, file type distribution, file size distribution, and modified time distribution. Filters for the report include client group, file name, file size, modification date, maximum items to show, and options to show deleted items and archived items.		should have
Audit Trail Report	Monitor operations in the data protection instances. Track operations performed by particular users.		should have
SLA Report	reporting on SLA met or SLA missed		should have

Alerting	Alerts provide automatic notification about operations, such as failed jobs. Alerts displayed on the Triggered Alerts dashboard. Users defined in the alert definition receive an email /SMS/something else notification when an alert is triggered. alerts should also trigger a SNOW ticket creation.		Must Have



Backup and restore use-cases for BHP

Scenario	Context and Priority	data protection workflow	restore workflow
worst case: Backup and restore persistent data, non-immutable app components, COTS applications, legacy configuration primarily to meet requirements for recovery time objectives (RTO) and recovery point objectives (RPO) because the app architecture is traditional/legacy in nature, cannot be automated through IAC and does not support application layer data replication and consistency.	BCP and DR protect and recovery process - priority 1 data protection for applications on virtual machine that are legacy, they are constrained to lower availability because they are limited to fault and update domains only, they cannot take advantage of ASR/DRS and CloudEndure or deployment stamps across AZ's. This is a legacy BCP workflow.	use backup agents to backup these legacy applications as best you can, RTO and RPO must be based on reality and is a product of the app architecture, the ability to copy data across the network, the size of the data set being protected. SLA, RTO and RPO should be set based on reality by what cloud ops can manage, not by the business, if the business doesn't accept this, raise a risk and manage the risk.	BCP recovery is based on the CIA RTO and RPO, leverage the platform availability such as fault domains/update domains or AZ's (if you can), use backup and restore to recover the failed workload or from any accidental user deletion/data corruption. If DR event is enacted by BHP, use backup and restore to recover the entire landscape into a new region.

Common use-case: Backup and restore persistent data and non-immutable app components, COTS applications, and legacy configuration primarily to protect and recover to application to another region. the app has some or a lot of automation through IAC and supports platform or application layer data replication and consistency.	DR protect and recovery process - priority 1 takes advantage of ASR/DRS and CloudEndure or deployment stamps across AZ's, has IAC automation - BCP process. Backup and restore is for accidental user deletion/data corruption and for DR recovery.	CIA RTO and RPO is handled by the deployment stamp architecture or replication using ASR/DRS and CloudEndure. Backups are taken based on CIA rating; short term retention is stored within the region and everything including long term retention is replicated to the secondary region.	BCP recovery is based on the CIA RTO and RPO, leverage IAC or replication to redeploy app components or failover to replicated components, adjust traffic flows so new components are live. Use backup and restore to recover from any accidental user deletion /data corruption. If DR event is enacted by BHP, use backup and restore to recover the entire landscape into a new region, along with IAC redeployment.
Common use-case: Backup and restore persistent data for modern immutable apps primarily to protect and recover to application to another region, the BCP recovery approach is focused on release or redeployment of immutable deployment stamps and auto-healing. Data replication and consistency is managed by the application architecture.	DR protect and recovery process - priority 1 takes advantage of cloud resilience patterns, self-healing with deployment stamps and app layer consistency and replication, full IAC automation - BCP process. Backup and restore is for accidental user deletion/data corruption and for DR recovery.	CIA RTO and RPO is handled by the deployment stamp architecture or replication Backups are taken based on CIA rating; short term retention is stored within the region and everything including long term retention is replicated to the secondary region.	BCP recovery is based on the CIA RTO and RPO, leverage IAC or replication to redeploy app components or failover and auto heal to warm or live components. Use backup and restore to recover from any accidental user deletion /data corruption. If DR event is enacted by BHP, use backup and restore to recover the entire landscape into a new region, along with IAC redeployment.
Compliance (long term retention)	archival and compliance process - priority 2 short term compliance data should be managed by the application team. long term compliance data should be stored on archival tier cloud storage by the data protection tool based on policies applied to that application or workload.	archival storage pools attached to the data protection solution. policy to move long term archival data onto the storage pools based on policies reporting and management dashboards to report on archival data, charge back and audit information	
Backup and restore is used as a last layer of defense for cyber-attacks, particularly ransomware recovery.	cyber protect and recovery process - priority 1 BHP need to determine if this is a BCP procedure or a DR procedure, what is the trigger, what is the procedure, the team, RACI and define RTO and RPO's around this particular use-case.	immutable storage pools attached to the data protection solution. policy to move data copies onto the immutable storage pools based on policies reporting and management dashboards to report on archival data, charge back and audit information	BHP need to determine a process around restoration for ransomware recovery, this should be very detailed and include RACI.
Backup and restore can be used as a release management tool for making copies of a production system in test environments or by restoring production data/database into a test environment.	migration process - priority 3		
Backup and restore can be used as a migration tool, for example if workloads are running in US region and will be moved to a new region, the backup tooling could be leveraged to restore into a new region as a migration technique.	migration process - priority 4		