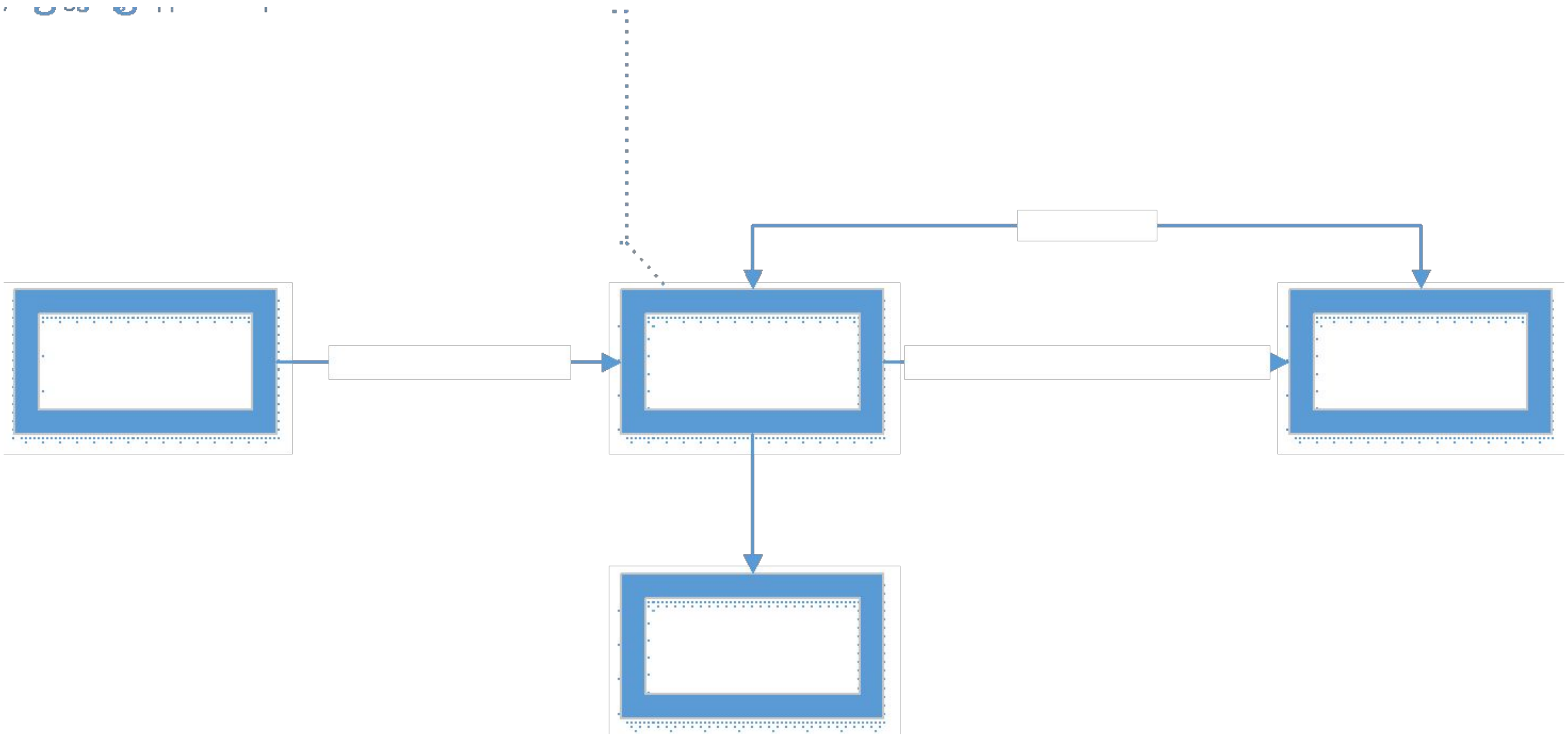
Acronis

ASN and deduplication training

Integration to platform



Vault without deduplication

- Storage
  - TIB files format is the same as on unmanaged vault.
  - TIBs are stored by special path, which contains machine ID, user ID, archive ID, etc:

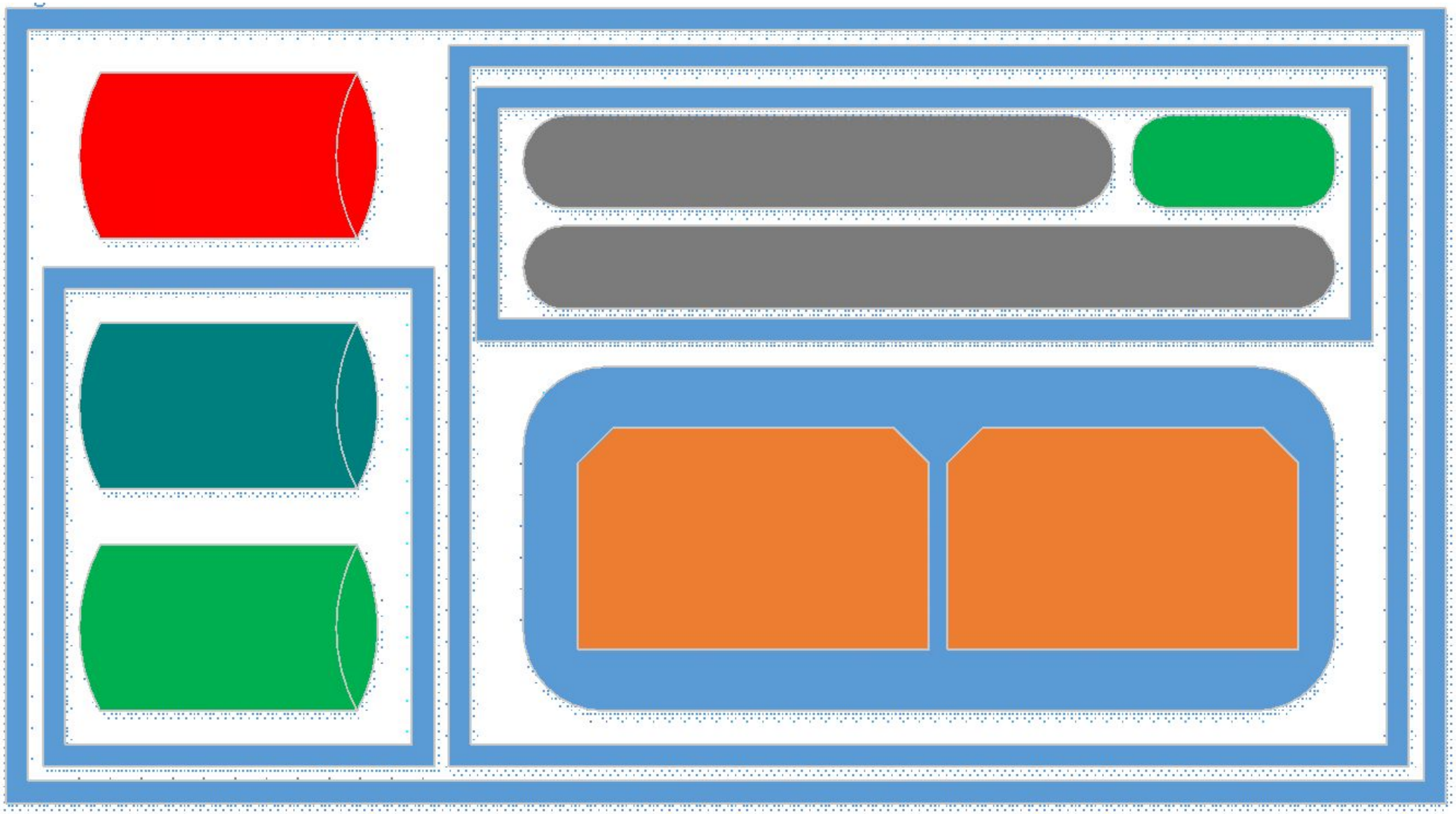
  /computers/MMSCurrentMachineID.InstanceID/users/SID_OF_USER(from windows)/archives/ArchiveID/
  - Each chain, started by FULL stored in separate folder, called “stream” in folders named “1”,”2”. In “1_data” dedup data is stored.
- Vault could be located on
  - ASN local folder
  - Network share
  - NAS
  - SAN

Permissions

User or group	Access to backups		
	Own	Same machine	All
On AMS	YES	YES	YES
Vault administrators	YES	YES	YES
Local administrators	YES	YES	NO
Vault users	YES	NO	NO

For now this works the following way: Vault User (not admin) can list all backups, however he can’t do anything with them. In U6 Vault User will be able to see only his backups (where he is Owner) (that comes from ABA for vCloud)

ASN Vault structure



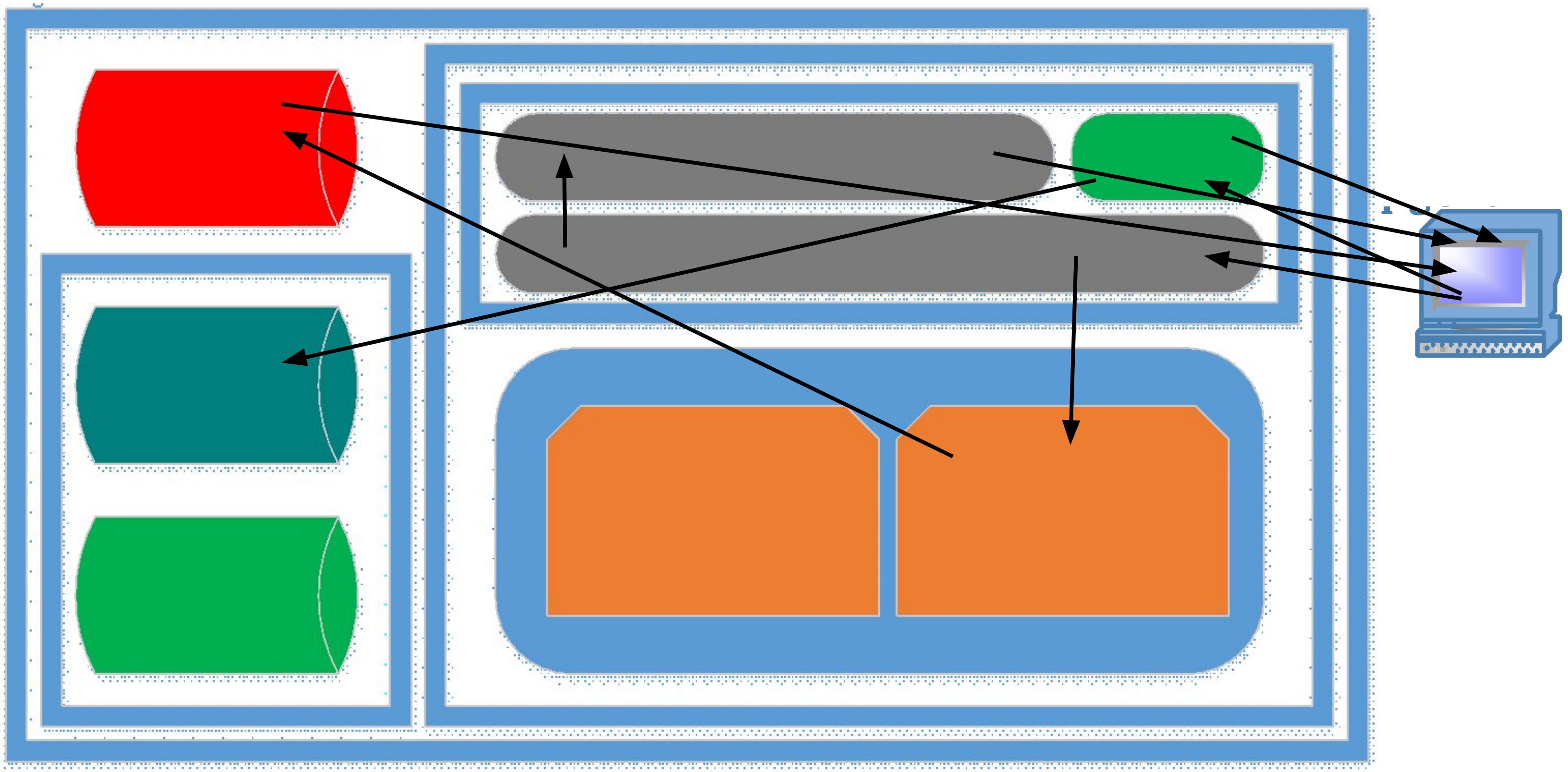
Deduplicated vault

- Recommendations:
  - Put Index (dedup) database on separate storage.
  - Exclude all paths for vault from antivirus scan.
  - Have only one dedup vault on single ASN (It’d share RAM).
- Parts of deduplicated vault:
  - Backup storage (local folder, net share, NAS, SAN)
  - Datastore, LDS, (L)IND, LOC files (stored inside along with vault - Backup storage)
  - Dedup DB (storing on network share is not supported)
  - Catalog (local folder recommended)
  - Vault meta-info DB (Firebird DB stored inside ProgramData)

Backup

- 2 streams (connections):
  - Header/metadata/links are stored in TIB file
  - Actual data Blocks are stored in LDS file,then it is indexed into unified_data
- Deduplication at source – only new blocks are sent
- Connectivity limits (may be changed)
  - Simultaneous backup (Client Connection Limit) – 10
  - Connections to wait in queue (Backup Queue Limit) – 50
- Encrypted backups (by agent) are skipped for deduplication

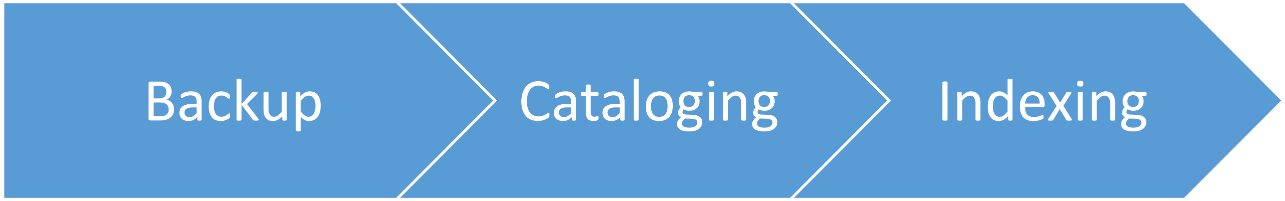
Workflow



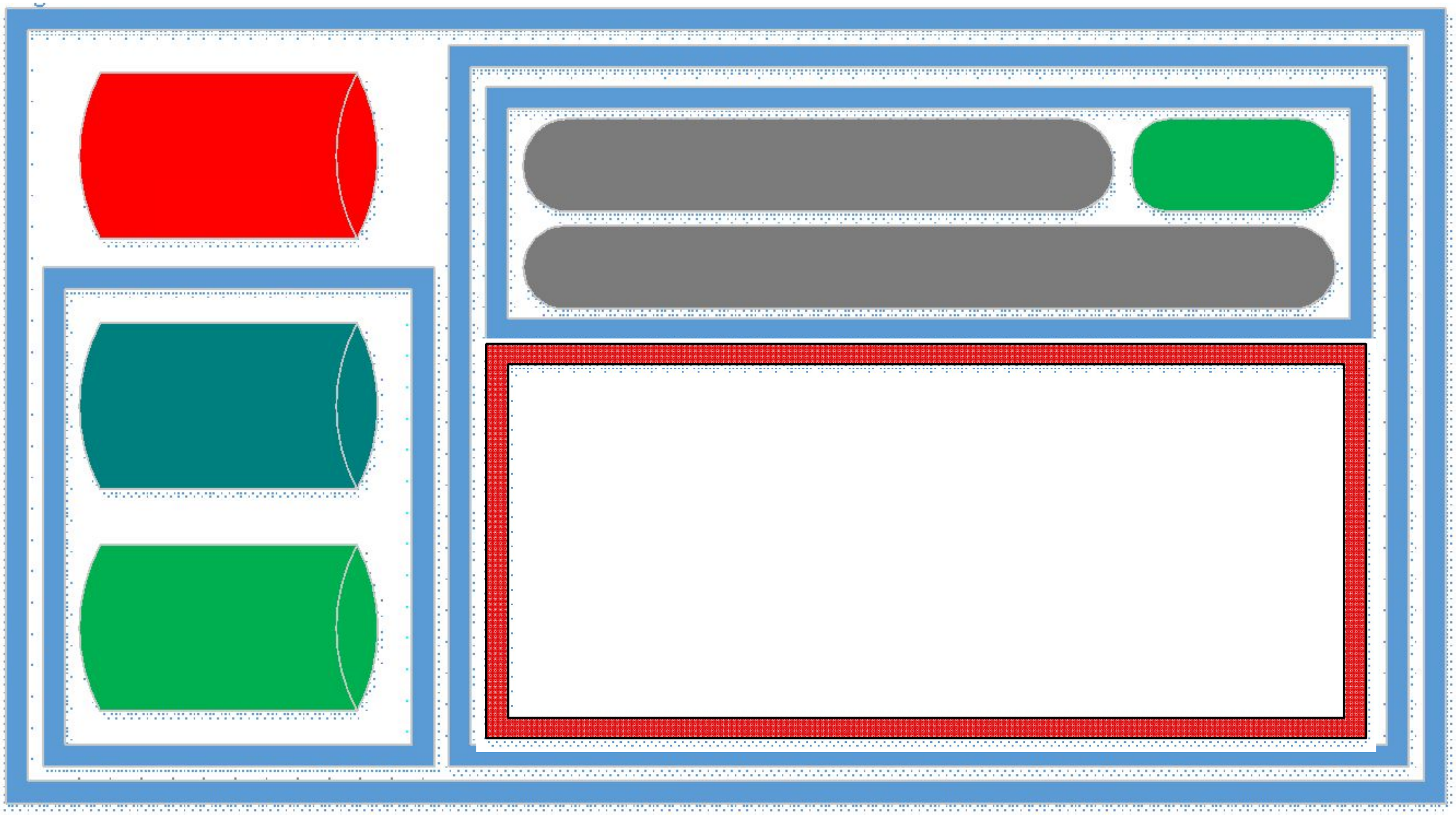
[illegible] (aka Repack)

Indexing

- Indexing moves unique blocks from LDS file (backup contents) to Datastore.
- Indexing is queued for each backup. Queue is rebuilt on service restart.
- Local Index (L)IND is created If only recovery/validation/convert to VM was requested *before* Indexing.



Datastore



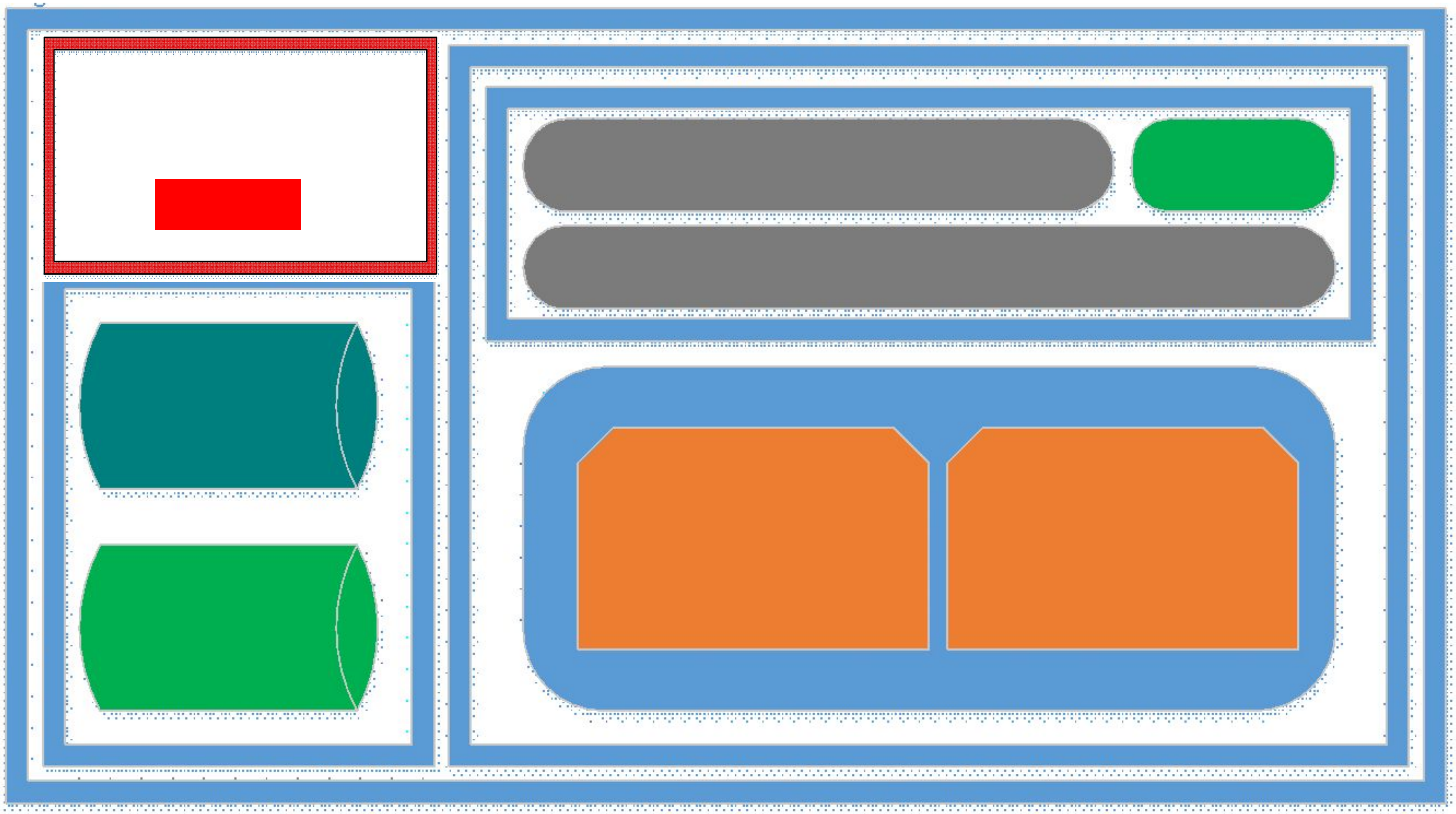
Datastore

- Datastore stores blocks
- Single datastore for all backup kinds
- Blocks are stored in two datastore files (unified_data):
  - Active – during indexing data is written there
  - Passive – during compacting unique blocks moved to active
- Datastore:
  - Is transactional (rollback on failure/crash)
  - Is always compressed
  - Could be encrypted (encrypted vault)
- For 1 TB of unique Disk Backup data we need 3 Gb of RAM
- If data is mixed: File, Disk, Exchange, then dedup DB will be growing much faster and much more RAM for 1 TB will be needed.

Block size

- Block size
  - Image backups: 4 Kb
  - File backups: 1b – 256Kb
- Blocks are compared by fingerprint (block MD5 hash).
- Blocks content is stored in Datastore.
- Offsets and sizes of blocks are stored in Dedup DB.
- Partitions with block less than 4 Kb or not multiple of 4Kb are skipped for deduplication.

Deduplication Database



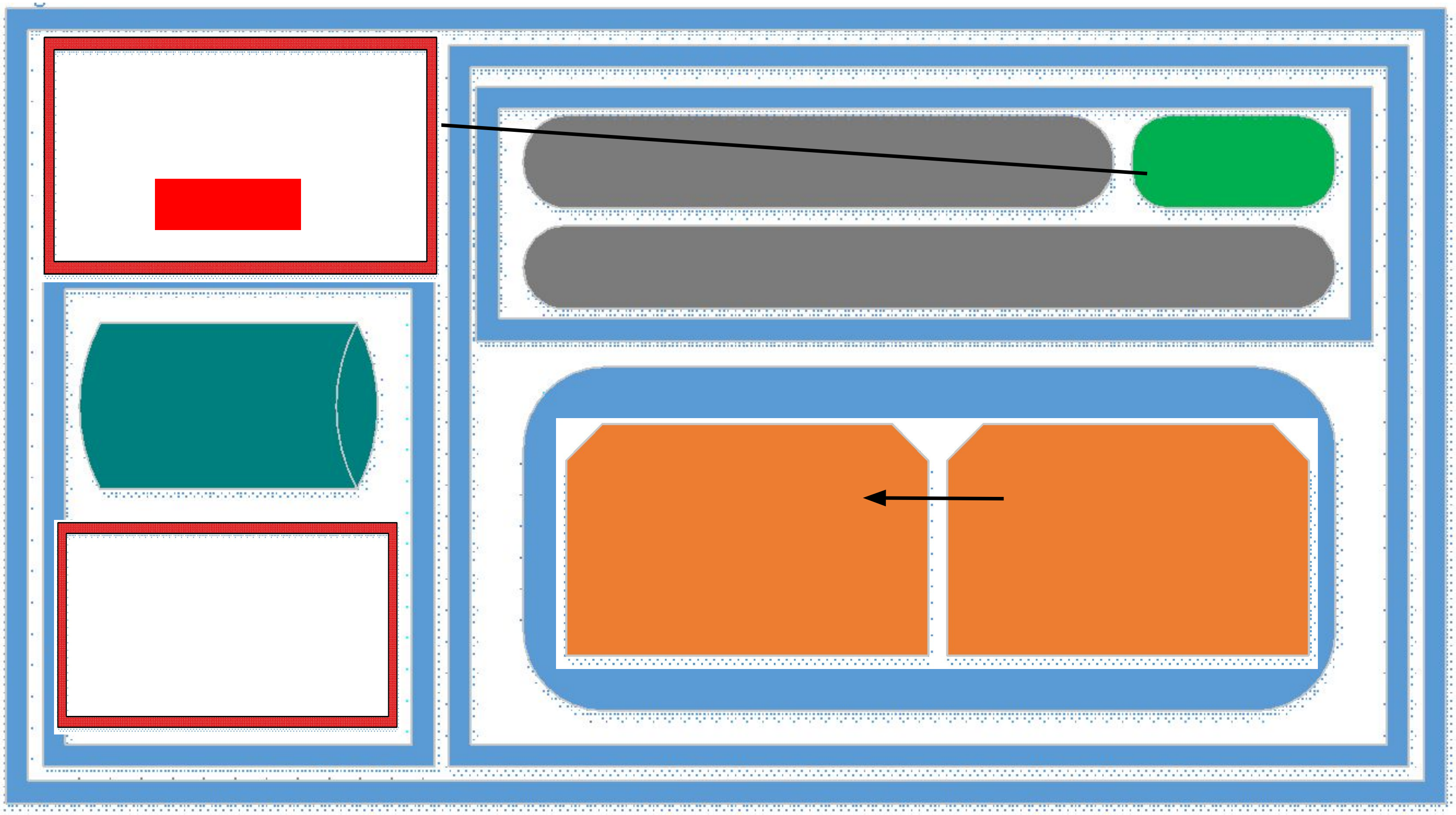
Deduplication Database

- Dedup DB is required for fast blocks access by fingerprints.
- It stores HASH of block and its offset it datastore.
- RAM is used mostly for LOCALITY index
- 80% of free physical memory used by default.
- RAM is locked by ASN even if locality is small
- Adjustable -DatastoreIndexCacheMemoryPercent
- More than 1 dedup vault on the same machine can be a problem.

Deduplication Database

- Index is rebuilt after compacting. Rebuilding of index works fast (with disk reading speed).
- On every ASN load the whole LOCALITY file is read. That takes time. Vault will be showing “Not ready for use”.
- About 1/3 of LOCALITY is loaded into RAM.
- If there is not enough RAM, everything will work except Indexing. It will fail asking for RAM. There is no performance degradation with Dedup DB growth.

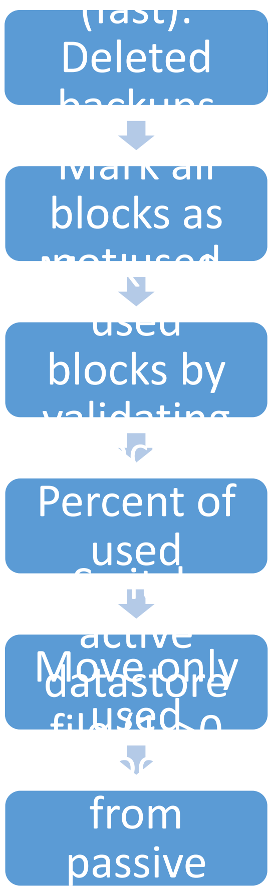
Compacting



Compacting: [illegible]

Compacting

Algorithm



Details

- Fine checks tuning
  - Compacting trigger rough estimation threshold
  - Compacting trigger threshold
- Simultaneous indexing and compacting are not allowed (handled automatically)
- Compacting requires 1 GB of space to start

Acronis

Export / Replication

- Backups are being un-deduplicated
- Possible to Export to local folder without agent installed
- Deduplication at source is enabled during export/replication
- It is slow, we know it. ABR-69401

Validation

- Validation of backups/archives validates only existence of hashes in Dedup DB (on disk and file archives at least)
- Validation of “Vault” validates all archives and then datastore.
- Theoretically there is a chance that info in dedup DB does not match datastore. In this case validaton of vault succeeds but recovery of backups fail. In this case escalate.

Attach / detach

- Detach
  - Vault meta-info db (.fdb) is copied to vault (storage) location.
- Attach
  - During attach it’s recommended to copy Index and Catalog from last location
    - Storage path (it is obligatory)
    - Index (deduplication) db path
    - Catalog path
  - If Index or Catalog paths contain no Index – it will be recreated. Recreation of index is going to be done with disk writing speed.

After attach/detach ASN syncs with AMS. So the vault appears/disappears from AMS with a delay.

Deduplication at source

- Faster backups (up to x6)

Backup pattern	Backup time [min]	RAM (peak), [Mb]	RAM (avg), [Mb]	CPU
w/o client-side	53	362	340	46%
with client-side (no data changes)	**9**	**350**	**332**	60%
with client-side (100% data changed)	53	472	416	45%

- Bandwidth saved (up to x200)

Backup pattern	Data Size [Mb]	Sent [Mb]	Received [Mb]	Sent+Received [Mb]	Stored on server [Mb]
vault w/o dedup	10240	10638	182	10819	10242
dedup w/o client-side		10573	161	10734	10243
dedup with client-side (0% new)		**25**	**24**	**50**	**1**
dedup with client-side (100% new)		10604	162	10766	10243

Compression

- Normal level – best choice for most cases.
- When deduplication is provided by Filesystem or hardware compression should be set to “none”

Compression level	Size (traffic) overhead	CPU time overhead	Backup time overhead	Traffic (size) [byte]	Backup time [sec]	MMS CPU [%]	Service process CPU time [sec]
normal				6284024874	244	44-66	380
none	42%	76%	21%	8953568582	295	32-40	290
high	-6%	168%	90%	5924419135	464	39-58	640
maximum	-7%	203%	135%	5865404133	573	33-60	772

Vault meta structure

- Vault meta files are located in: \BackupAndRecovery\ASN\.meta
- 1 file per vault. There is also 1 file with a list of vaults.

```
<?xml version="1.0" encoding="UTF-8" ?>
<metainfo id="26D7D967-C222-4B8E-927B-F8CF4FE1F410" type="location">
        <catalog_database_uri> -- Path to Catalog. You can use it to change catalog path
C:\ProgramData\Acronis\BackupAndRecovery\ASN\Catalog\. (should be done on stopped ASN)
        </catalog_database_uri>
        <database_uri>-- Path to Firebird (.FDB) database. Do not touch.
                C:\ProgramData\Acronis\BackupAndRecovery\ASN\VaultMetadataDatabases\
        </database_uri>
        <description />
        <index_database_uri> -- Path to dedup DB
                \\\
        </index_database_uri>
        <location_id> -- Location ID. Must match file name and metainfo id in top.
                26D7D967-C222-4B8E-927B-F8CF4FE1F410
        </location_id>
        <name>  -- Vault name
                testvlt
        </name>
        <storage_uri> -- Vault path
                F:\mng
        </storage_uri>
</metainfo>
```

These files are found in sysinfo and help to determine if the current vault has dedup enabled, where it is located, etc.

Vault meta files

Inside of the vault there is .meta folder. The folder contains meta for each archive and 1 meta for the vault itself.
Archive meta is similar to XML on unmanaged vault.
Vault meta is similar to the .meta on ASN <?xml version="1.0" encoding="UTF-8" ?>

```
<metainfo id="92FCE129-B465-4ECA-B6F5-DF4CC3DE1682" type="location"> -Location ID. Must match location ID from .meta on ASN
        <admin_ids>
                S-1-5-32-544 –SIDs of vault administrators
        </admin_ids>
        <attached_to>
                00000000-0000-0000-0000-000000000000–ASN machine ID. If 000 – then the vault is detached
        </attached_to>
        <compression_level>
                none –compression
        </compression_level>
        <deduplicated>
                1 –1 – deduplicated. 0 – non deduplicated
        </deduplicated>
        <description />
        <location_id>
                92FCE129-B465-4ECA-B6F5-DF4CC3DE1682 -Location ID. Must match location ID from .meta on ASN
        </location_id>
        <location_version>
                abr11
        </location_version>
        <name>
                sharevault
        </name>
        <unified_store_data_size>
                14286848 –dedup unified_data_ds size
        </unified_store_data_size>
        <user_ids>
                S-1-1-0 –SIDs of vault users
        </user_ids>
</metainfo>
```

If vault is encrypted it will also have fingreprint of data to ensure if correct key was entered.

DML Database

- ASN DML Database is located in \BackupAndRecovery\ASN\DmlDatabase\asn_dml_objects.db3

It is used for infrastructure integration of ASN. Due to a known issue it grows: KB 47170

In the worst case it can be removed (on stopped ASN).

ASN logs

ANS logs are located in
\BackupAndRecovery\ASN\Logs
And
\BackupAndRecovery\ASN\events.db3
For events.db3 use Yalp.
It is worth checking both logs sources for each case.

ASN and Tapes

ASN is the service that writes to tape.
ARSM is responsible for:
1. Moving tapes.
2. Inventoring tapes.
3. Operations with ARSM.sqlite
Starting from U4 ASN is using ARSM.sqlite as vault database.
4. Delays after backup, before replication starts. Almost Fixed in u6.

ASN and OB

When backing up to ASN and replicating to cloud here is how it works:

Agent backs up to ASN.

After the backup Agent downloads the data from ASN and sends it to cloud.

ASN is only functioning as storage in this case.

Metadata Issues

- Fixing issues:

  1. Reindex: acrocmd reindex vault –loc=bsp://ASN_IP/vault
  2. Ultimate reindex:

Detach vault

remove FDB from vault

Attach vault.

Vault is corrupted

When ASN says that vault is corrupted check events.db3

- .tmp files in .meta in the vault (fixed in u5)
- Multiple “location” meta files In .meta in vault.
- Vault is on NAS. Access to NAS fails.
- Vault is attached but “ASNID” in .meta in vault is different from ASN or is 00000.
- Vault is corrupted due to known issues on 43916. Rebackup lost blocks or recreate vault.

Storage Node is busy.

Usually a deadlock. Most likely caused not by connection limiter itself.
If it is really a very heavily loaded environment and ASN runs many activities then temporary workaround is to set:

- HKLM\SOFTWARE\Acronis\ASN\Configuration\StorageNode\ClientConnectionLimit to 30
- HKLM\SOFTWARE\Acronis\ASN\Configuration\StorageNode\FastOperationConnectionLimit to 100
- HKLM\SOFTWARE\Acronis\ASN\Configuration\StorageNode\FastOperationQueueLimit to 500
- HKLM\SOFTWARE\Acronis\ASN\Configuration\StorageNode\BackupQueueLimit to 150

SSL on ASN

- Before U3 there was AES256 encryption.
- After U3 it is AES128 with HW optimization
- Still there is a slowdown from SSL so in this case disable it as said in KB
- MAKE SURE YOU READ RED WARNING FROM KB.