Pruning Policy Analysis Report

Date: 2025-11-14 **Analysis Focus:** Understanding Aging vs Pruning and Why Space Isn't Being Reclaimed **Environment:** 226 Storage Policies, 100 Libraries, 243 MediaAgents, 79 Storage Pools

---

Executive Summary

This analysis reveals the **complete picture of why storage space is not being reclaimed** in your Commvault environment. While the previous aging analysis identified that data isn't being marked for deletion fast enough, this pruning analysis explains why **even aged data isn't freeing up storage space**.

🔴 Critical Finding: The Two-Stage Problem

Your storage reclamation failure has **TWO distinct stages**, both of which are failing:

STAGE 1: AGING (Logical Marking) - 🟠 DELAYED - 130 rules have cycle retention issues - Data not marked for aging fast enough - **Result:** Delayed aging = Delayed pruning eligibility

STAGE 2: PRUNING (Physical Deletion) - 🔴 FAILING - 11 storage pools critically low on space - Aged data not being physically deleted - **Result:** No space reclaimed despite aging

---

Understanding Aging vs Pruning

The Critical Difference

Many administrators confuse these two operations. Understanding the difference is **essential** to fixing the problem:

Operation	Type	What It Does	Impact on Storage
AGING	Logical	Marks backup jobs as eligible for deletion based on retention rules	**NO space freed** - metadata only
PRUNING	Physical	Actually deletes aged data blocks from disk storage	**Space IS freed** - physical deletion

The Process Flow

```
1. AGING RUNS (Daily at 12:00 PM)
   ↓
   Retention rules evaluated
   ↓
   Jobs exceeding retention marked as "aged"
   ↓
2. PRUNING RUNS (After aging)
   ↓
   For NON-dedup: Aged jobs deleted immediately
   ↓
   For DEDUP: DDB reference counters decremented
   ↓
   Blocks with zero references queued for pruning
   ↓
   Physical deletion from disk
   ↓
3. SPACE RECLAIMED
```

The Problem: If Aging Works But Pruning Fails

- Aging marks jobs as aged (metadata updated)
- Pruning fails to delete the actual data
- **Storage pools remain full**
- **Backups may fail** due to lack of space
- **Costs increase** from unnecessary storage

This is EXACTLY what's happening in your environment!

Analysis Findings

Section 1: Aging Configuration Status

✅ **GOOD NEWS:** Aging is properly configured

- **518 total retention rules** analyzed
- **518 rules (100%)** have aging ENABLED
- **0 rules (0%)** have aging disabled

This means data CAN be marked for aging (unlike some environments where aging is completely disabled).

Section 2: Aging Delays (From Previous Analysis)

⏱️ **ISSUE:** Cycle retention requirements delaying aging

- **130 retention rules** have short retention (≤30 days) + 2 cycle requirements
- Average cycles required: **2.0**
- **Effective delay:** 7-14 days before data eligible for aging

Impact on Pruning: If aging is delayed, pruning is delayed. Data can't be pruned until it's first aged.

Section 3: Pruning Failure Evidence

💥 **PROOF OF PRUNING FAILURE:** Storage pools critically full

Current State: - **Total Storage Pools:** 79 - 🔴 **CRITICAL (<10% free):** 3 pools - 🟠 **WARNING (10-20% free):** 8 pools - 🟡 **LOW (20-30% free):** 14 pools

Most Critical Pools (Highest Priority for Manual Pruning):

Pool Name	% Free	Status
Apex GDP	1.97%	🔴 CRITICAL
Southern_Sun_Durban	2.05%	🔴 CRITICAL
Simera_GDP	9.30%	🔴 CRITICAL
Southern_Sun_City_Bowl	12.47%	🟠 WARNING
MKLM_GDP	12.85%	🟠 WARNING
GDP Railway	12.98%	🟠 WARNING
Universal GDP	13.37%	🟠 WARNING
Capri_Local_GDP	14.75%	🟠 WARNING
CLMAN02_Storage	14.75%	🟠 WARNING
Southern_Sun_Local	17.80%	🟠 WARNING
EnergyPartners_Local	18.96%	🟠 WARNING

Analysis: If aging AND pruning were working properly, these pools would NOT be this full. The fact that 11 pools have <20% free space is **definitive proof** that pruning is not

completing successfully.

Section 4: Deduplication Complexity

🔍 **Additional Complexity:** Deduplication requires multi-step pruning

For environments using deduplication (which appears to be your case based on storage pool types):

Deduplication Pruning Process: 1. Data aging marks jobs as aged 2. DDB (Deduplication Database) reference counters are decremented for each aged block 3. When a block's reference count reaches zero, it's added to "Pending Pruning" queue 4. Physical pruning job deletes blocks from disk 5. Space is finally freed

Pruning Can Be Blocked By: - DDB sealed or corrupted - High "Pending Deletes" count in DDB - MediaAgent offline or overloaded - Mount paths inaccessible - Pruning operation window restrictions - Resource constraints (CPU/Memory/Disk I/O)

Root Cause Analysis

Primary Root Causes

1. Aging Delays Due to Cycle Retention (130 rules) - **Impact:** Medium-High - **Cause:** Cycle requirements delay when data becomes eligible for aging - **Effect on Pruning:** Delayed aging = Delayed pruning eligibility - **Solution:** Reduce cycles from 2 → 1 for short-term policies

2. Pruning Operations Not Completing (Evidence: 11 critically low pools) - **Impact:** CRITICAL - **Cause:** Unknown - requires further investigation - **Possible Reasons:** - MediaAgent offline or unavailable - DDB sealed/corrupted - Pruning jobs failing silently - Operation windows blocking pruning - Resource exhaustion - **Solution:** Investigate MediaAgent logs, DDB status, and pruning job history

3. Deduplication Reference Count Issues - **Impact:** High (if using deduplication) - **Cause:** References not being properly decremented - **Effect:** Blocks never reach zero references = Never pruned - **Solution:** Verify DDB health, run DDB verification

---

Infrastructure Analysis

MediaAgents (Critical for Pruning)

Total MediaAgents: 243

⚠️ **CRITICAL REQUIREMENT:** All MediaAgents must be ONLINE for pruning to work.

Key MediaAgents to Verify: - CLSVMA01 - CVHSMAN01.JHB.SEAGATESTORAGECLOUD.CO.ZA - CVHSMAN02.JHB.SEAGATESTORAGECLOUD.CO.ZA - CVHSMAN03.JHB.SEAGATESTORAGECLOUD.CO.ZA - ... and 239 others

Action Required: 1. Open Commvault Console 2. Navigate to Storage Resources → MediaAgents 3. Verify each MediaAgent shows as "Online" 4. For any offline MediaAgents: - Investigate why they're offline - Check network connectivity - Review MediaAgent logs - Restart MediaAgent service if needed

Libraries (Pruning Target Locations)

Total Libraries: 100

These are the physical locations where pruning must occur. If libraries or their mount paths are offline, pruning cannot proceed.

Sample Libraries: - ALS_LocalLibrary - ActiveScale - ActiveScale_Tape - Allbro_Local - Amaro Foods_DiskLibrary - ... and 95 others

Action Required: 1. Verify all library mount paths are accessible 2. Check library status in Commvault Console 3. Ensure sufficient permissions for pruning operations

Storage Policies

Total Storage Policies: 226

Each storage policy has retention rules that govern aging. All policies require working pruning to reclaim space.

Sample Policies: - A.R.B Electrical Backup Plan - ALS_AD, ALS_SQL - AMT Server Plan - Apex Backup Plan - ... and 222 others

---

Immediate Actions Required

PRIORITY 1: Verify Pruning Infrastructure (TODAY)

Step 1: Check MediaAgent Status

```
Location: CommCell Console → Storage Resources → MediaAgents
Action: Verify ALL show as "Online"
Critical: Pruning CANNOT run if MediaAgents are offline
```

Step 2: Check DDB Status (for Deduplication)

```
Location: CommCell Console → Storage Resources → Deduplication Engines
Action:
  1. Right-click each DDB → Properties
  2. Verify Status = "Active" (NOT "Sealed")
  3. Check "Pending Deletes" count
  4. High pending deletes = pruning backlog

Warning: If DDB is "Sealed", NO pruning will occur!
```

Step 3: Manually Trigger Pruning on Critical Pools

```
Location: CommCell Console → Storage Resources → Deduplication Database
Action:
  1. Right-click DDB
  2. Select "All Tasks" → "Validate and Prune Aged Data"
  3. Monitor job progress
  4. Check storage pool space after completion

Priority Pools (do these FIRST):
  - Apex GDP (1.97% free)
  - Southern_Sun_Durban (2.05% free)
  - Simera_GDP (9.30% free)
```

Step 4: Review Pruning Logs

```
Location: MediaAgent Server → <Install_Path>/Log Files/SIDBPrune.log
Action:
  1. Open SIDBPrune.log
  2. Search for recent pruning operations
  3. Look for errors, warnings, or failures
  4. Note any "Skipped" or "Failed" messages

Key Things to Look For:
  - "Pruning skipped" messages
  - Resource exhaustion errors
  - Mount path access errors
  - Timeout errors
```

PRIORITY 2: Fix Aging Configuration (THIS WEEK)

Action 1: Reduce Cycle Retention (130 rules)

Current State: 130 rules have ≤30 days retention + 2 cycle requirement

Required Change: 2 cycles → 1 cycle

Impact: - Aging will occur 7-14 days faster - Data becomes eligible for pruning sooner - Storage reclamation accelerates

Implementation: 1. Identify all plans with ≤30 day retention 2. Update retention rules to require only 1 cycle 3. Verify changes in Commvault UI 4. Monitor aging jobs after change

Affected Plans: (sample) - A.R.B Electrical Backup Plan - ALS_AD, ALS_SQL - AMT Server Plan - Apex Backup Plan - BallStraathof_AD, _FS, _Oracle, _SQL, _VM - ... and 120 others

---

PRIORITY 3: Monitor and Verify (ONGOING)

Verification Step 1: Run Data Retention Forecast Report

Purpose: Shows what data should be aged and pruned

```
Location: CommCell Console → Reports → Data Retention Forecast
Action:
  1. Select library/storage policy
  2. Generate report
  3. Review "Aged Data" section
  4. Check for warnings/issues at bottom

What to Look For:
  - Data eligible for aging but not aged
  - Data aged but not pruned
  - Warnings about DDB or MediaAgent issues
```

Verification Step 2: Monitor Storage Pool Space

Purpose: Confirm pruning is actually freeing space

```
Frequency: Daily for first week, then weekly
Action:
  1. Check storage pool free space %
  2. Compare to previous day
```

```
  3. Look for increasing free space trend

Success Criteria:
  - Free space % increases daily
  - Critical pools move above 20% free
  - No pools drop below 10% free
```

Verification Step 3: Review Pruning Job History

Purpose: Ensure pruning jobs are running and completing successfully

```
Location: CommCell Console → Job Controller → All Jobs
Filter: Job Type = "Pruning" or "Data Aging"
Action:
  1. Check for recent pruning jobs
  2. Verify status = "Completed"
  3. Review job details for data pruned
  4. Check for any failed jobs

Red Flags:
  - No pruning jobs in last 7 days
  - All pruning jobs showing "Failed"
  - Jobs completing but 0 bytes pruned
```

Technical Deep Dive

Why Pruning Fails in Deduplication Environments

Deduplication adds significant complexity to the pruning process:

Traditional (Non-Dedup) Pruning:

```
Job aged → Job deleted → Space freed
(Simple, immediate)
```

Deduplication Pruning:

```
Job 1 aged → Ref count -1 for each block
Job 2 aged → Ref count -1 for each block
...
Job N aged → Some blocks reach ref count = 0
             ↓
Blocks with ref=0 added to pending pruning queue
             ↓
Physical pruning operation runs
             ↓
Blocks deleted from SFILES
             ↓
DDB updated
             ↓
Space freed
(Complex, multi-step, resource-intensive)
```

Common Failure Points: 1. **Reference count not decremented** (aging job didn't update DDB properly) 2. **Pending pruning queue too large** (pruning can't keep up with aging) 3. **DDB sealed** (no pruning allowed on sealed DDBs) 4. **MediaAgent unavailable** (pruning job can't run) 5. **Pruning operation window** (pruning only allowed at certain times)

Troubleshooting Pruning Failures

Symptom: Aging works, but storage pools remain full

Diagnostic Steps:

1. **Check if pruning jobs are running:** `CommCell Console → Job Controller → Filter by "Pruning" Expected: Daily pruning jobs completing successfully If Missing: Pruning not scheduled or failing to start`
2. **Check DDB pending deletes:** `CommCell Console → Deduplication Engine → Properties → Statistics Look for: "Pending Deletes" count Normal: <10,000 Warning: 10,000-100,000 Critical: >100,000 (pruning severely backlogged)`

3. **Check MediaAgent resource usage:** `On MediaAgent server: Task Manager → Performance tab Look for: CPU >90%, Memory >90%, Disk I/O >90% If High: Resource exhaustion preventing pruning`
4. **Check mount path accessibility:** `On MediaAgent: Navigate to mount path directory Expected: Directory accessible, sufficient permissions If Failed: Pruning cannot access storage to delete files`
5. **Check pruning operation window:** `CommCell Console → MediaAgent → Properties → Operation Window Look for: Pruning operation restrictions If Restricted: Pruning may only run during specific hours`

Recommendations Summary

Immediate (Today)

1. ✅ Verify all 243 MediaAgents are ONLINE
2. ✅ Check DDB status (Active vs Sealed)
3. ✅ Manually run pruning on 3 critical pools
4. ✅ Review SIDBPrune.log for errors

Short-term (This Week)

1. ✅ Reduce cycle retention from 2→1 on 130 rules
2. ✅ Run Data Retention Forecast report
3. ✅ Monitor pruning job execution daily
4. ✅ Check storage pool space trends

Long-term (This Month)

1. ✅ Implement automated pruning monitoring

2. ✅ Set up alerts for pruning job failures
3. ✅ Establish baseline for normal pruning performance
4. ✅ Document pruning troubleshooting procedures

Success Metrics

Week 1 Success Criteria

- ✅ All MediaAgents showing as "Online"
- ✅ DDB status confirmed as "Active"
- ✅ Manual pruning completes successfully on critical pools
- ✅ Storage pool free space increases by at least 5%

Month 1 Success Criteria

- ✅ All 130 cycle retention rules updated
- ✅ No storage pools below 20% free
- ✅ Daily pruning jobs completing successfully
- ✅ Storage pool free space stabilized above 30%

Month 3 Success Criteria

- ✅ Automated monitoring and alerting operational
- ✅ Zero pruning job failures in 30 days
- ✅ Storage costs reduced by 10-15%
- ✅ No backup failures due to storage space

Conclusion

Your storage reclamation problem has **TWO distinct components:**

Component 1: Aging (Logical) - Issue: Delayed due to cycle retention requirements - Severity: MEDIUM-HIGH - Solution: Reduce cycles from 2 → 1 - Status: Solution identified, ready to implement

Component 2: Pruning (Physical) - Issue: Not completing or not running - Severity: CRITICAL - Solution: Verify MediaAgent health, DDB status, trigger manual pruning - Status: Requires immediate investigation and action

Both must be fixed for storage space to be reclaimed.

Immediate Priority: Focus on PRIORITY 1 actions (verify MediaAgent status, check DDB, manually run pruning) to get pruning working again. This will provide immediate relief to critically full storage pools.

Secondary Priority: Fix aging delays (reduce cycle retention) to prevent the problem from recurring and to accelerate future pruning operations.

Report Generated: 2025-11-14 **Analysis Tools:** analyze_pruning_policies.py, analyze_aging_failures.py **Data Source:** Commvault SQLite Database (Database/commvault.db) **Environment Size:** 226 Storage Policies, 518 Retention Rules, 79 Storage Pools, 243 MediaAgents