

veeAMON 2023

Architecting Veeam Backup for Microsoft 365 at Scale



Falko Banaszak



Solution Architect
@Falko_Banaszak
<https://virtualhome.blog>



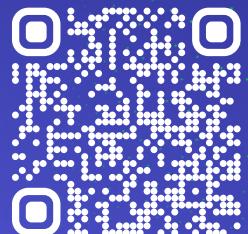
Jim Jones



Senior Product Infrastructure Architect
11:11 Systems
@k00laidIT
<https://koolaid.info>

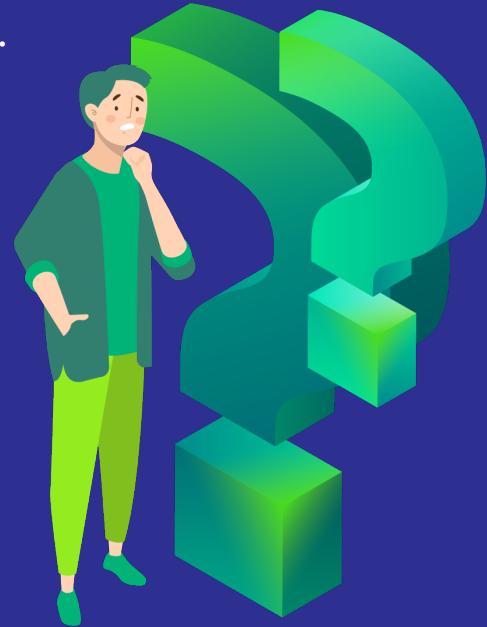
Agenda

- 10 • Planning and preparation
- 15 • Proxy, repository and job design
- 15 • Monitoring & troubleshooting
- 5 • Questions



Why all this ?

- Four different Microsoft 365 workloads with their own “I/O pattern”.
Those are almost impossible to size properly at the beginning and can change.
- Microsoft tends to silently make changes to the Microsoft 365 platform.
- In most cases VB365 is being onboarded in parallel to the M365 migration.
- A proper architecture helps maintain your RTO and flexibility.
- SP style multi-tenancy has its own challenges.





Planning and preparation

<https://bp.veeam.com/vb365>



Organization size discovery

Multi- geo organization – consider licensing, regulation.

Customer/admin managed groups are good.

- At this point only using for sizing/ to discover size of a particular group.
- Delimit users.
- Delimit geographical boundaries.

Use a good calculator as early as possible:

- <https://calculator.veeam.com/vb365>
Overestimates proxy design IMHO.
- <https://111systems.com/catalyst/>

Veeam Backup for Microsoft 365 Capacity Calculator

Import Settings Export Settings

Primary mailbox	Archive mailbox	OneDrive	SharePoint
Total database size, TB: 31	20	20	50
Weekly change rate, %: 5	5	0	0
Mailboxes: Quantity: 2500	Mailboxes: Quantity: 2500	Accounts: Quantity: 2500	Sites: Quantity: 100
Retention policy			
5 years	5 years	5 years	5 years
Teams			
100			
Local repository size: First month 152.43 TB	Object storage size: First month 69.29 TB 1 418.97 USD	Total API calls: First month 51.19 M 255.97 USD	Infrastructure sizing: Deployment: All in one: <input checked="" type="checkbox"/> Off VB365 servers: 1 Proxies: 1
Last month 1308.92 TB	Last month 594.96 TB 12 184.87 USD	Last month 5 274.33 K 26.37 USD	

Environment placement

- Cloud vs. on-premises:
 - Compute: keep close to storage.
 - Azure: F or D type instances.
 - AWS: M type.
 - Secondary copies and where to keep them.
 - Egress, API, retention... OH MY!
- Environmental observations:
 - Networking.
 - Compute.
 - Scale out.
- Best practices: [BP/design/placement](#)





Proxy, repository and job design

Proxy design

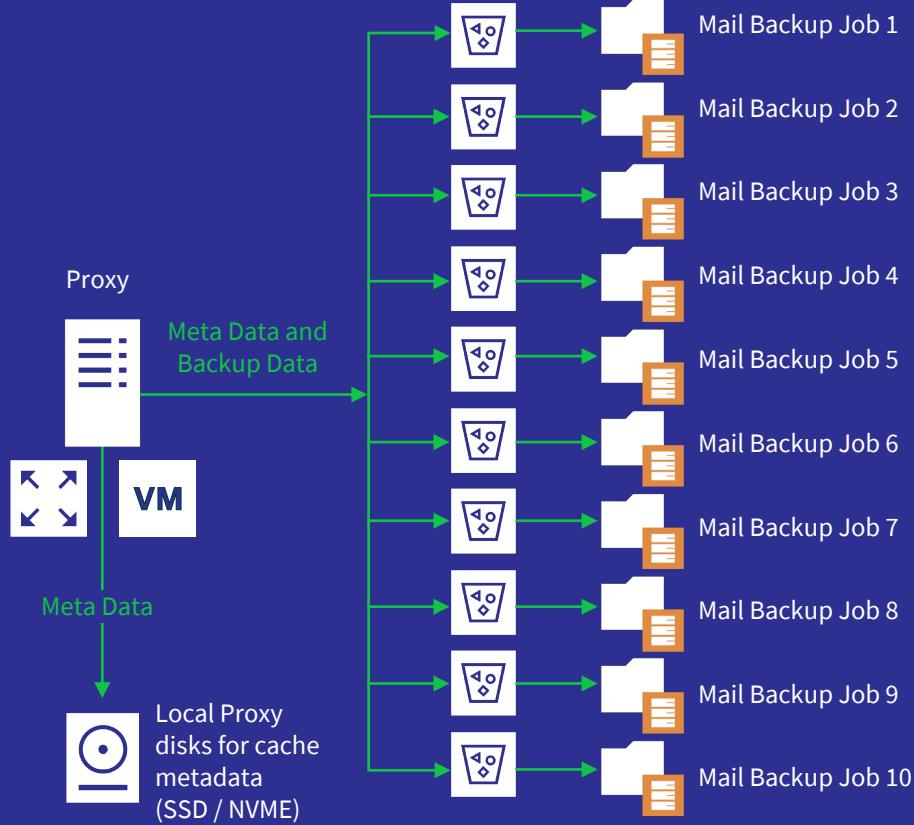
- Compute: 8 vCPU, 32GB RAM.
- Local secondary SSD for cache.
- Be conservative against BP guidance, needs will grow.
 - [BP/design/sizing proxy server](#).
 - [BP/design/configuration maximums](#).
- Start at 2,000-2,500 objects per proxy.
- Proxy distribution – group by workload type.
 - Mailboxes can be denser than SP data types.
- Least taken advice ever.

“For optimum performance, we recommend storing **no more than 300,000 files** in a single OneDrive or team site library.”



Proxy to repository relationship

- Recommended way on scaling a backup proxy with its backup jobs and repositories.
- Repeat this for OneDrive and Archive Mail (user-based objects).
- Workload type proxies are getting the same “I/O pattern”.
- Highest flexibility.



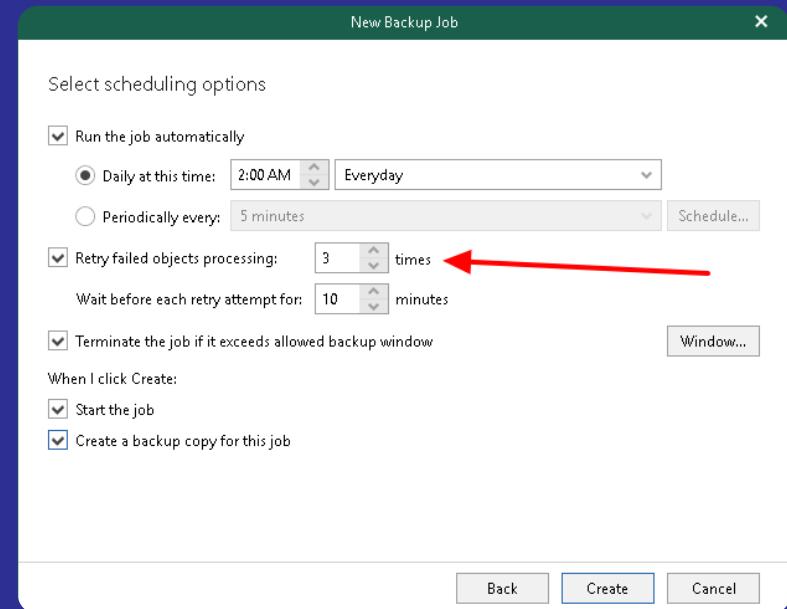
Failure domains within proxy to repository

Keep “failure domains” as small as possible.

- A job with 10K users is a big failure domain.
- A job with 1K users is a smaller failure domain.
- RTO & retrying.

Reassigning objects from one job to another.

- Move to another proxy == another repository!
- Another full sync!
- Psst: check out Mike Ressler’s session tomorrow.



Repository design

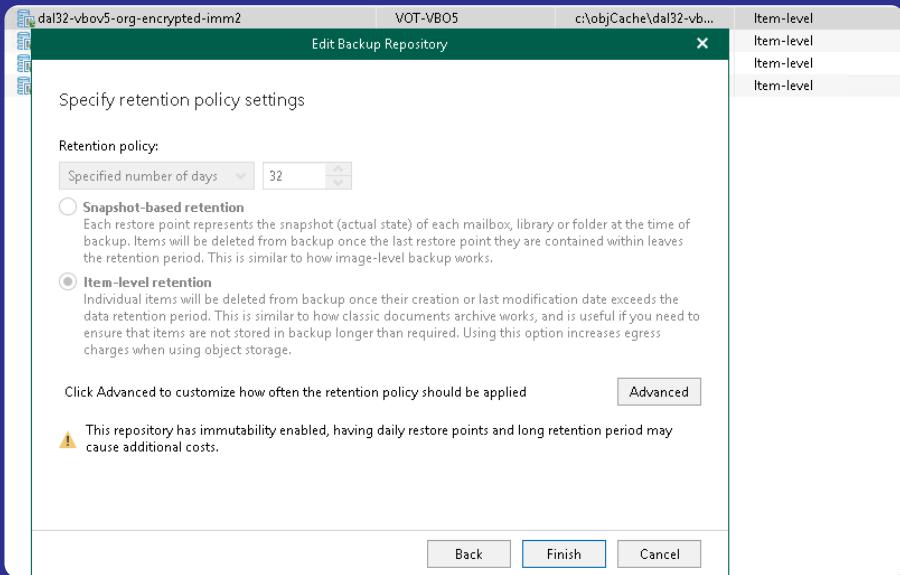
Item level vs. snapshot choices

Retention

- Consider tiered retention policies based on governance.
- ONLY KEEP DATA AS LONG AS MANDATED.
- Maximum in named years: 25 years.
- Maximum years in days: 273 years.

Immutable secondary copies are good! (and bad!)

- Good to have but keep retention low.
- Uses true object-lock, choices have consequences.
- Retention policy MUST match primary copy.

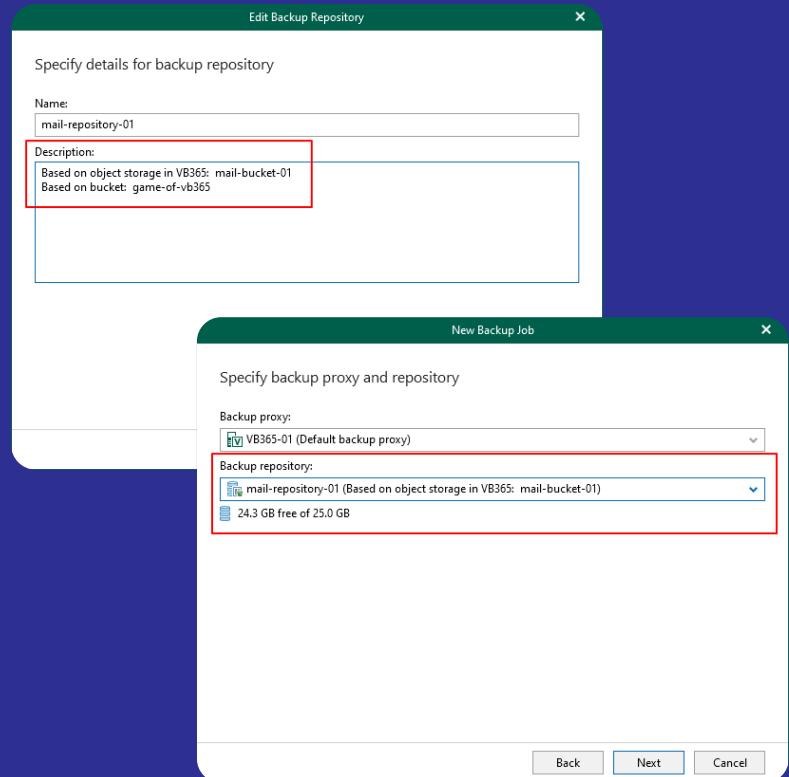


Repository design

- Object storage cache: (5 MB/ObjectTB)*2.
- 1 job = 1 bucket/repository.
- Use description field to show relationships.
- Best practices:

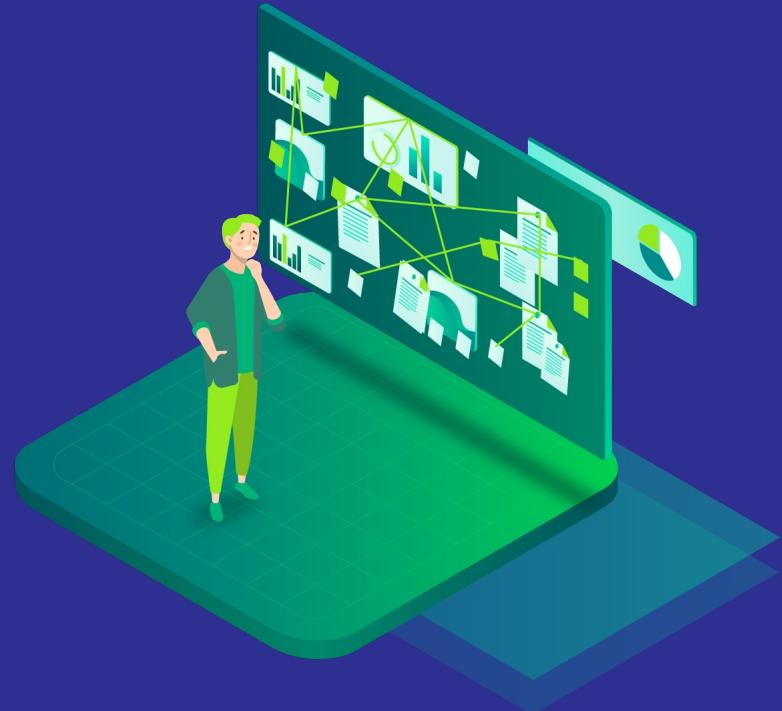
[BP/design/sizing/object storage.](#)

[BP/build & configure/repository/object storage.](#)



Job design

- Design with best practices(ish):
[BP/design/job design.](#)
- Consider staff turnover/growth in design:
 - If low use catch all jobs.
 - If high design for new user growth.
- Always split by workload type as a minimum.
- At scale for user workloads decide how you will split.
- Dynamic groups are great but \$\$\$ for pure M365 orgs.
- Alphabetical jobs not bad but hard to manage.
- Teams: do you really need Teams channels?
- Stagger first runs with no schedule.



veeAMON 2023



Monitoring & troubleshooting

Monitoring

Veeam® ONE™ – dashboards, reports, API!

API options

- Grafana:

<https://github.com/jorgedlcruz/veeam-backup-for-microsoft365-grafana>

- Roll your own reports:

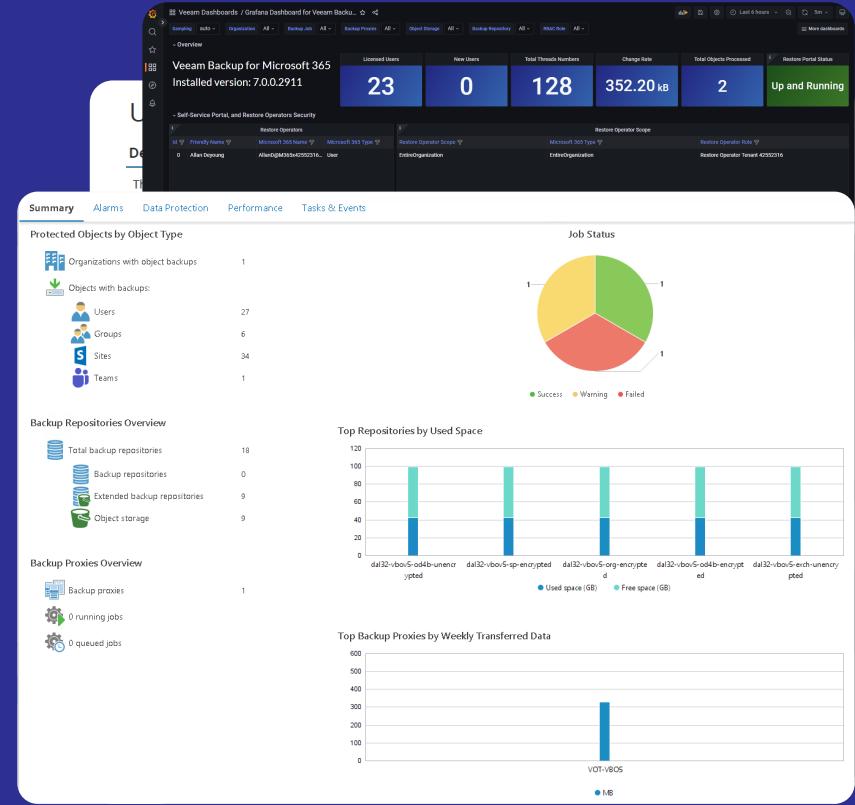
<https://benyoung.blog/tag/vbo/>

New with v7

- Prebuilt VB365 reports:

[Mailbox Protection Report.](#)

[User Protection Report.](#)



Plan well: avoid pains later

Job/data movement

- No supported method for object > object:
Veeam [KB 3067](#).
But... forums / “[Migrate to a another object storage repo](#)”.
- Block > Object = SLOW and jobs must be disabled:
[k00laidIT / Veeam / kb3067.ps1](#) – modified and improved.
[k00laidIT / Veeam / kb3067-validation.ps1](#) – stats and verification.

Graph API for Teams export

- Only way to get Teams Channel data, paid API:
- <https://1111systems.com/tag/microsoft-365-backup/>
- Veeam [KB 4322](#) – request access.
- Veeam [KB 4340](#) – enable on server.

Gather round for a chat about
Graph Teams Export API



Plan well: avoid pains later

Authentication

- Setup separate, well protected account for purpose.
- App secrets = long expiration.
- #1 most common reason for VB365 errors today.

Throttling

- Backup groups support murky today
May actually make processing slower with AppOnly.
- Logs to verify.
- [BP/operate/M365 throttling](#).

High sync wait times

- [- BP/operate/common issues](#).



Log diving - throttling

HTTP Error 500 internal server error.

HTTP Error 429 too many requests.

HMM
LOOKS FAMILIAR

```
PS C:\ProgramData\Veeam\Backup365\Logs> Get-ChildItem -Recurse -Filter *.log | select-string "throttled [^0]"  
  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:483:[25.04.2023 01:27:43.519] 108 (6844) 5fc1069d-6d24-4c54-9b39-ddac46fec11b: requests 5970, throttled 332  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:484:[25.04.2023 01:27:43.519] 108 (6844) 064ef8b2-9e7d-4c5c-8390-20a3c027f59f: requests 5681, throttled 125  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:485:[25.04.2023 01:27:43.519] 108 (6844) 9f3d6d34-d939-42ea-b84f-75027ea7b6c9: requests 5938, throttled 257  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:486:[25.04.2023 01:27:43.519] 108 (6844) cdf15d00-b322-448b-84f1-7cfab63a51e6f: requests 5013, throttled 72  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:487:[25.04.2023 01:27:43.519] 108 (6844) ec79f765-6bf8-45e3-8c93-6c7e1b7a2713: requests 4716, throttled 53  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:515:[25.04.2023 01:35:47.280] 124 (3928) 5fc1069d-6d24-4c54-9b39-ddac46fec11b: requests 2327, throttled 53  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:516:[25.04.2023 01:35:47.280] 124 (3928) 064ef8b2-9e7d-4c5c-8390-20a3c027f59f: requests 2197, throttled 39  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:518:[25.04.2023 01:35:47.280] 124 (3928) cdf15d00-b322-448b-84f1-7cfab63a51e6f: requests 2314, throttled 72  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:642:[25.04.2023 02:06:53.851] 111 (3664) 5fc1069d-6d24-4c54-9b39-ddac46fec11b: requests 6344, throttled 101  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:643:[25.04.2023 02:06:53.851] 111 (3664) 064ef8b2-9e7d-4c5c-8390-20a3c027f59f: requests 5687, throttled 126  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:644:[25.04.2023 02:06:53.851] 111 (3664) 9f3d6d34-d939-42ea-b84f-75027ea7b6c9: requests 6193, throttled 223  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:645:[25.04.2023 02:06:53.851] 111 (3664) cdf15d00-b322-448b-84f1-7cfab63a51e6f: requests 6116, throttled 146  
Veeam.Archiver.Proxy_2023_04_25_00_42_55.log:646:[25.04.2023 02:06:53.851] 111 (3664) ec79f765-6bf8-45e3-8c93-6c7e1b7a2713: requests 6562, throttled 273  
  
PS C:\ProgramData\Veeam\Backup365\Logs> Get-ChildItem -Recurse -Filter *.log | select-string "too many requests" | Measure-Object -Line  
  
Lines Words Characters Property  
----- ----- ----- -----  
1872  
  
PS C:\ProgramData\Veeam\Backup365\Logs>
```



Log diving - high sync times

- Use a direct internet connection whenever possible.
- Try to avoid traffic shaping, next-generation firewalling and all sorts of "logic" on the internet connection uplink .
- "Normal sync times" for a direct connection are expected to be in the range of 50 to 150 **milliseconds**.

```
select-string "Sync time: [^0]" *
```

```
<JOB NAME>_2022_04_14_08_59_59.log:2732:[14.04.2022 09:00:08] 59 (8132) Sync time: 143.7938527
```

```
<JOB NAME>_2022_04_14_08_59_59.log:2737:[14.04.2022 09:00:08] 49 (4600) Sync time: 324.8092703
```

Yes, those are seconds not milliseconds - and guess what the next log line say:

14.04.2022 09:00:08 75 (4600) No changes

veeAMON2023

Questions



<https://github.com/k00laidIT/VeeamON2023/MIA03>



<https://github.com/k00laidiT/VeeamON2023/MIA03>