

Установка Victoria Metrics

Цели практики

Целью данной практической работы является демонстрация процесса настройки и работы Victoria Metrics. В рамках данного урока в Kubernetes кластере будет установлен и настроен Victoria Metrics в режиме Cluster, а также будет проведена настройка уже установленных в кластере Grafana и Prometheus.

Установка Victoria Metrics будет производиться с помощью Helm Chart.

С чем будем работать

- Helm chart victoria metrics - необходим для установки Prometheus server в кластер

Где будем выполнять практику

ВНИМАНИЕ: Перед тем, как работать со стендом, его необходимо запустить из Личного кабинета

- Подключаемся по SSH к sbox.slurm.io:

```
ssh s<номер вашего логина>@sbox.slurm.io
```

- Подключаемся к своему кластеру Kubernetes:

```
ssh master-1.s<ваш номер логина>
```

- Становимся root:

```
sudo -s
```

Практика

Подготовительная работа

1. Клонирование манифестов

Для выполнения практической части, потребуются манифесты. Их необходимо склادировать с gitlab, сделать это можно выполнив команды:

```
cd ~  
git clone git@gitlab.slurm.io:edu/lmk8s.git
```

2. Подключаем репо Victoria Metrics

```
helm repo add vm https://victoriametrics.github.io/helm-charts/
```

```
helm repo update
```

3. Создаем namespace

Victoria Metrics будет устанавливаться в отдельный namespace: victoria. Для его создания, необходимо выполнить команду:

```
kubectl create ns victoria
```

Установка Victoria Metrics в кластер

Теперь все готово для установки и достаточно выполнить команду:

```
helm upgrade -i vm-cluster -n victoria -f  
~/lmk8s/4.prometheus_advanced/4.2.victoria_metrics/values.yml  
vm/victoria-metrics-cluster
```

В результате будут установлены следующие компоненты:

- vmSelect - будет установлен как Deployment с значением replicas: 2
- vmInsert - будет установлен как Deployment с значением replicas: 2
- vmStorage - будет установлен как StatefulSet с значением replicas: 2

Ключевые изменения в values.yml

Обратите внимание, что в `values.yml` для чарта были сделаны изменения, вот наиболее ключевые из них:

- **Заданы ресурсы для все объектов** - это общее хорошее правило для всего, что запускается в Kubernetes кластере. Во первых, заданные `requests` и `limits` позволяют scheduler корректно распределять Pod по node. Во вторых, это важно для избегания ситуации утечки ресурсов, а в случае с Prometheus это особенно актуально, так как он является достаточно "прожорливым".
- Задан **podDisruptionBudget** - данная настройка влияет на то, сколько Pod может быть одновременно выключено (распространяется только на eviction API). Она позволяет гарантировать, что при обслуживании кластера Kubernetes не будут выключены все Pod'ы с приложением.
- **podAntiAffinity** - данная настройка позволяет гарантировать, что Pod из StatefulSet и Pod из Deployment не будут запущены на одной node. Без данной настройки возможна ситуация, когда все Pod будут запущены на одной node, и в случае выхода из строя этой node, кластер останется без мониторинга (Pod для statefulset не перезапускаются автоматически)
- **vmselect.extraArgs** - добавлены следующие параметры:

`search.maxQueueDuration: "10s"` - ограничивает время выполнения запросов. Когда VictoriaMetrics пользуется большое число команд, это позволяет отсекать слишком длительные запросы. Ведь они, в свою очередь, могут сказаться на производительности всей БД.

`search.maxQueryLen: "16384"` - ограничивает максимальную длину запроса. Опять же, когда VictoriaMetrics пользуется сразу несколько команд, это позволяет отсекать слишком большие запросы, которые могут сказаться на производительности БД.

`replicationFactor: "2"` - задает сколько копий time series имеют доступ в vmStorage. Должен совпадать со значением для vmInsert.

`dedup.minScrapeInterval: "1ms"` - data samples, которое отличаются на значение, указанное в этом поле и менее, будут "дедуплицированы". Этот

параметр обязательно использовать при значении `replicationFactor` 2 и больше.

- **vmInsert.extraArgs** - добавлены следующие параметры:

`replicationFactor: "2"` - задает сколько копий `time series` должны быть сохранены в `vmStorage`.

- **vmStorage.StorageClass** - задает имя `StorageClass`. При этом в `default values` этого поля вообще нет, а вот в `Helm template` он используется.

В поле `extraArgs` для каждого компонента задаются ключи запуска для этого компонента. При этом, большинство значений по умолчанию являются оптимальными. Полный список ключей запуска указан в последнем уроке данного раздела.

Настройка Prometheus

Prometheus уже установлен в кластере, установка производилась с помощью `Helm Chart`. Для внесения изменений, необходимо получить текущие `values`. Сделать это можно выполнив команду:

```
helm get values -n monitoring prometheus > prometheus_values.yml
```

В результате `values` будут сохранены в файл: `prometheus_values.yml`. Этот файл необходимо открыть на редактирование, найти в нем строку: `remoteWrite: []` и заменить ее на:

```
remoteWrite:
```

```
- url:
```

```
http://vm-cluster-victoria-metrics-cluster-vminsert.victoria.svc:8480/inse  
rt/0/prometheus/api/v1/write
```

После этого изменения необходимо применить в кластер, выполнив команду:

```
helm upgrade prometheus prometheus-community/prometheus -n monitoring -f prometheus_values.yml
```

URL, куда отправляет данные Prometheus сервер, имеет следующую структуру:

```
/<component>/<clientID>/prometheus/<Prometheus API query>
```

- `<component>` - префикс для компонента. `vmInsert` - `/insert`, `vmSelect` - `/select`
- `<clientID>` - ID клиента, может быть произвольным `int` и должен совпадать для `insert` и `select` запросов. За счет этого реализуется механизм `tenant`, когда в одну Victoria Metrics могут писать различные Prometheus сервера и их данные будут храниться в отдельных `tenants` (аналог `namespace`)
- `<Prometheus API query>` - исходный запрос Prometheus.

Данная структура используется только для Victoria Metrics в режиме **Cluster**. Если вы используете Victoria Metrics в режиме **Single Node**, то запросы будут иметь структуру: `/<Prometheus API query>`

Настройка Grafana

Grafana уже установлена в кластер, для ее настройки необходимо создать `datasource` и импортировать `dashboard` для `node_exporter`. Для этого достаточно применить два манифеста в кластер. Подробнее как это работает, будет рассмотрено в следующих уроках.

```
kubectl create -f ~/lmk8s/4.prometheus_advanced/4.2.victoria_metrics/grafana/ -n monitoring
```

Для того, чтобы Grafana увидела новый `datasource`, нужно перезапустить её Pod. Сделать это можно командой:

```
kubectl delete po -n monitoring grafana-<Tab>
```

На этом всё! Можно проверять результат, для этого откройте в браузере в анонимном режиме адрес `http://grafana.<номер студента>.edu.slurm.io`. Имя пользователя и пароль совпадают с доступами в Личный кабинет.

Ключи запуска компонентов Victoria Metrics

Ключи запуска отсутствуют в документации и единственный способ их получить - это запуск соответствующего компонента с ключом `--help`. А при условии, что основной способ распространения Victoria Metrics это docker-образы, это не самый удобный способ. Ниже предоставлены ключи для основных компонентов:

vmInsert

-csvTrimTimestamp duration

Trim timestamps when importing csv data to this duration. Minimum practical duration is 1ms. Higher duration (i.e. 1s) may be used for reducing disk space usage for timestamp data (default 1ms)

-enableTCP6

Whether to enable IPv6 for listening and dialing. By default only IPv4 TCP is used

-envflag.enable

Whether to enable reading flags from environment variables additionally to command line. Command line flag values have priority over values from environment vars. Flags are read only from command line if this flag isn't set

-envflag.prefix string

Prefix for environment variables if `-envflag.enable` is set

-fs.disableMmap

Whether to use `pread()` instead of `mmap()` for reading data files. By default `mmap()` is used for 64-bit arches and `pread()` is used for 32-bit arches, since they cannot read data

files bigger than 2^{32} bytes in memory. `mmap()` is usually faster for reading small data chunks than `pread()`

-graphiteListenAddr string

TCP and UDP address to listen for Graphite plaintext data. Usually `:2003` must be set. Doesn't work if empty

-graphiteTrimTimestamp duration

Trim timestamps for Graphite data to this duration. Minimum practical duration is 1s. Higher duration (i.e. 1m) may be used for reducing disk space usage for timestamp data (default 1s)

-http.connTimeout duration

Incoming http connections are closed after the configured timeout. This may help spreading incoming load among a cluster of services behind load balancer. Note that the real timeout may be bigger by up to 10% as a protection from Thundering herd problem (default 2m0s)

-http.disableResponseCompression

Disable compression of HTTP responses for saving CPU resources. By default compression is enabled to save network bandwidth

-http.idleConnTimeout duration

Timeout for incoming idle http connections (default 1m0s)

-http.maxGracefulShutdownDuration duration

The maximum duration for graceful shutdown of HTTP server. Highly loaded server may require increased value for graceful shutdown (default 7s)

-http.pathPrefix string

An optional prefix to add to all the paths handled by http server. For example, if '-http.pathPrefix=/foo/bar' is set, then all the http requests will be handled on '/foo/bar/*' paths. This may be useful for proxied requests. See <https://www.robustperception.io/using-external-urls-and-proxies-with-prometheus>

-http.shutdownDelay duration

Optional delay before http server shutdown. During this delay the server returns non-OK responses from /health page, so load balancers can route new requests to other servers

-httpListenAddr string

Address to listen for http connections (default ":8480")

-import.maxLineLen max_rows_per_line

The maximum length in bytes of a single line accepted by /api/v1/import; the line length can be limited with max_rows_per_line query arg passed to /api/v1/export

Supports the following optional suffixes for values: KB, MB, GB, KiB, MiB, GiB (default 104857600)

-influx.maxLineSize value

The maximum size in bytes for a single Influx line during parsing

Supports the following optional suffixes for values: KB, MB, GB, KiB, MiB, GiB (default 262144)

-influxListenAddr http://<vminsert>:8480/insert/<accountID>/influx/write

TCP and UDP address to listen for Influx line protocol data. Usually :8189 must be set. Doesn't work if empty. This flag isn't needed when ingesting data over HTTP - just send it to http://<vminsert>:8480/insert/<accountID>/influx/write

-influxMeasurementFieldSeparator string

Separator for '{measurement}{separator}{field_name}' metric name when inserted via Influx line protocol (default "_")

-influxSkipMeasurement

Uses '{field_name}' as a metric name while ignoring '{measurement}' and '-influxMeasurementFieldSeparator'

-influxSkipSingleField

Uses '{measurement}' instead of '{measurement}{separator}{field_name}' for metric name if Influx line contains only a single field

-influxTrimTimestamp duration

Trim timestamps for Influx line protocol data to this duration. Minimum practical duration is 1ms. Higher duration (i.e. 1s) may be used for reducing disk space usage for timestamp data (default 1ms)

-insert.maxQueueDuration duration

The maximum duration for waiting in the queue for insert requests due to -maxConcurrentInserts (default 1m0s)

-loggerDisableTimestamps

Whether to disable writing timestamps in logs

-loggerErrorsPerSecondLimit int

Per-second limit on the number of ERROR messages. If more than the given number of errors are emitted per second, then the remaining errors are suppressed. Zero value disables the rate limit

-loggerFormat string

Format for logs. Possible values: default, json (default "default")

-loggerLevel string

Minimum level of errors to log. Possible values: INFO, WARN, ERROR, FATAL, PANIC (default "INFO")

-loggerOutput string

Output for the logs. Supported values: stderr, stdout (default "stderr")

-loggerWarnsPerSecondLimit int

Per-second limit on the number of WARN messages. If more than the given number of warns are emitted per second, then the remaining warns are suppressed. Zero value disables the rate limit

-maxConcurrentInserts int

The maximum number of concurrent inserts. Default value should work for most cases, since it minimizes the overhead for concurrent inserts. This option is tightly coupled with `-insert.maxQueueDuration` (default 24)

-maxInsertRequestSize value

The maximum size in bytes of a single Prometheus `remote_write` API request

Supports the following optional suffixes for values: KB, MB, GB, KiB, MiB, GiB (default 33554432)

-maxLabelsPerTimeseries int

The maximum number of labels accepted per time series. Superfluous labels are dropped (default 30)

-memory.allowedBytes value

Allowed size of system memory VictoriaMetrics caches may occupy. This option overrides `-memory.allowedPercent` if set to non-zero value. Too low value may increase

cache miss rate, which usually results in higher CPU and disk IO usage. Too high value may evict too much data from OS page cache, which will result in higher disk IO usage

Supports the following optional suffixes for values: KB, MB, GB, KiB, MiB, GiB (default 0)

-memory.allowedPercent float

Allowed percent of system memory VictoriaMetrics caches may occupy. See also -memory.allowedBytes. Too low value may increase cache miss rate, which usually results in higher CPU and disk IO usage. Too high value may evict too much data from OS page cache, which will result in higher disk IO usage (default 60)

-opentsdbHTTPListenAddr string

TCP address to listen for OpenTSDB HTTP put requests. Usually :4242 must be set. Doesn't work if empty

-opentsdbListenAddr string

TCP and UDP address to listen for OpenTSDB metrics. Telnet put messages and HTTP /api/put messages are simultaneously served on TCP port. Usually :4242 must be set. Doesn't work if empty

-opentsdbTrimTimestamp duration

Trim timestamps for OpenTSDB 'telnet put' data to this duration. Minimum practical duration is 1s. Higher duration (i.e. 1m) may be used for reducing disk space usage for timestamp data (default 1s)

-opentsdbhttp.maxInsertRequestSize value

The maximum size of OpenTSDB HTTP put request

Supports the following optional suffixes for values: KB, MB, GB, KiB, MiB, GiB (default 33554432)

-opentsdbhttpTrimTimestamp duration

Trim timestamps for OpenTSDB HTTP data to this duration. Minimum practical duration is 1ms. Higher duration (i.e. 1s) may be used for reducing disk space usage for timestamp data (default 1ms)

-relabelConfig string

Optional path to a file with relabeling rules, which are applied to all the ingested metrics. See <https://victoriametrics.github.io/#relabeling> for details

-replicationFactor int

Replication factor for the ingested data, i.e. how many copies to make among distinct `-storageNode` instances. Note that `vmselect` must run with `-dedup.minScrapeInterval=1ms` for data de-duplication when `replicationFactor` is greater than 1. Higher values for `-dedup.minScrapeInterval` at `vmselect` is OK (default 1)

-rpc.disableCompression

Disable compression of RPC traffic. This reduces CPU usage at the cost of higher network bandwidth usage

-storageNode array

Address of `vmstorage` nodes; usage: `-storageNode=vmstorage-host1:8400`
`-storageNode=vmstorage-host2:8400`

Supports array of values separated by comma or specified via multiple flags.

-tls

Whether to enable TLS (aka HTTPS) for incoming requests. `-tlsCertFile` and `-tlsKeyFile` must be set if `-tls` is set

-tlsCertFile string

Path to file with TLS certificate. Used only if `-tls` is set. Prefer ECDSA certs instead of RSA certs, since RSA certs are slow

`-tlsKeyFile` string

Path to file with TLS key. Used only if `-tls` is set

`vmSelect`

`-cacheDataPath` string

Path to directory for cache files. Cache isn't saved if empty

`-dedup.minScrapeInterval` duration

Remove superfluous samples from time series if they are located closer to each other than this duration. This may be useful for reducing overhead when multiple identically configured Prometheus instances write data to the same VictoriaMetrics. Deduplication is disabled if the `-dedup.minScrapeInterval` is 0

`-enableTCP6`

Whether to enable IPv6 for listening and dialing. By default only IPv4 TCP is used

`-envflag.enable`

Whether to enable reading flags from environment variables additionally to command line. Command line flag values have priority over values from environment vars. Flags are read only from command line if this flag isn't set

`-envflag.prefix` string

Prefix for environment variables if `-envflag.enable` is set

`-fs.disableMmap`

Whether to use `pread()` instead of `mmap()` for reading data files. By default `mmap()` is used for 64-bit arches and `pread()` is used for 32-bit arches, since they cannot read data

files bigger than 2^{32} bytes in memory. `mmap()` is usually faster for reading small data chunks than `pread()`

-graphiteTrimTimestamp duration

Trim timestamps for Graphite data to this duration. Minimum practical duration is 1s. Higher duration (i.e. 1m) may be used for reducing disk space usage for timestamp data (default 1s)

-http.connTimeout duration

Incoming http connections are closed after the configured timeout. This may help spreading incoming load among a cluster of services behind load balancer. Note that the real timeout may be bigger by up to 10% as a protection from Thundering herd problem (default 2m0s)

-http.disableResponseCompression

Disable compression of HTTP responses for saving CPU resources. By default compression is enabled to save network bandwidth

-http.idleConnTimeout duration

Timeout for incoming idle http connections (default 1m0s)

-http.maxGracefulShutdownDuration duration

The maximum duration for graceful shutdown of HTTP server. Highly loaded server may require increased value for graceful shutdown (default 7s)

-http.pathPrefix string

An optional prefix to add to all the paths handled by http server. For example, if '-http.pathPrefix=/foo/bar' is set, then all the http requests will be handled on '/foo/bar/*' paths. This may be useful for proxied requests. See

<https://www.robustperception.io/using-external-urls-and-proxies-with-prometheus>

-http.shutdownDelay duration

Optional delay before http server shutdown. During this delay the server returns non-OK responses from /health page, so load balancers can route new requests to other servers

-httpListenAddr string

Address to listen for http connections (default ":8481")

-loggerDisableTimestamps

Whether to disable writing timestamps in logs

-loggerErrorsPerSecondLimit int

Per-second limit on the number of ERROR messages. If more than the given number of errors are emitted per second, then the remaining errors are suppressed. Zero value disables the rate limit

-loggerFormat string

Format for logs. Possible values: default, json (default "default")

-loggerLevel string

Minimum level of errors to log. Possible values: INFO, WARN, ERROR, FATAL, PANIC (default "INFO")

-loggerOutput string

Output for the logs. Supported values: stderr, stdout (default "stderr")

-loggerWarnsPerSecondLimit int

Per-second limit on the number of WARN messages. If more than the given number of warns are emitted per second, then the remaining warns are suppressed. Zero value disables the rate limit

-memory.allowedBytes value

Allowed size of system memory VictoriaMetrics caches may occupy. This option overrides -memory.allowedPercent if set to non-zero value. Too low value may increase cache miss rate, which usually results in higher CPU and disk IO usage. Too high value may evict too much data from OS page cache, which will result in higher disk IO usage

Supports the following optional suffixes for values: KB, MB, GB, KiB, MiB, GiB (default 0)

-memory.allowedPercent float

Allowed percent of system memory VictoriaMetrics caches may occupy. See also -memory.allowedBytes. Too low value may increase cache miss rate, which usually results in higher CPU and disk IO usage. Too high value may evict too much data from OS page cache, which will result in higher disk IO usage (default 60)

-replicationFactor int

How many copies of every time series is available on vmstorage nodes. See -replicationFactor command-line flag for vminsert nodes (default 1)

-search.cacheTimestampOffset duration

The maximum duration since the current time for response data, which is always queried from the original raw data, without using the response cache. Increase this value if you see gaps in responses due to time synchronization issues between VictoriaMetrics and data sources (default 5m0s)

-search.denyPartialResponse

Whether to deny partial responses if a part of `-storageNode` instances fail to perform queries; this trades availability over consistency; see also `-search.maxQueryDuration` and `-search.storageTimeout`

-search.disableCache

Whether to disable response caching. This may be useful during data backfilling

-search.latencyOffset duration

The time when data points become visible in query results after the collection. Too small value can result in incomplete last points for query results (default 30s)

-search.logSlowQueryDuration duration

Log queries with execution time exceeding this value. Zero disables slow query logging (default 5s)

-search.maxConcurrentRequests int

The maximum number of concurrent search requests. It shouldn't be high, since a single request can saturate all the CPU cores. See also `-search.maxQueueDuration` (default 6)

-search.maxExportDuration duration

The maximum duration for `/api/v1/export` call (default 720h0m0s)

-search.maxLookback duration

Synonym to `-search.lookback-delta` from Prometheus. The value is dynamically detected from interval between time series datapoints if not set. It can be overridden on per-query basis via `max_lookback` arg. See also `'-search.maxStalenessInterval'` flag, which has the same meaning due to historical reasons

-search.maxPointsPerTimeseries int

The maximum points per a single timeseries returned from the search (default 30000)

-search.maxQueryDuration duration

The maximum duration for query execution; see also -search.storageTimeout (default 30s)

-search.maxQueryLen value

The maximum search query length in bytes

Supports the following optional suffixes for values: KB, MB, GB, KiB, MiB, GiB (default 16384)

-search.maxQueueDuration duration

The maximum time the request waits for execution when -search.maxConcurrentRequests limit is reached; see also -search.maxQueryDuration (default 10s)

-search.maxStalenessInterval duration

The maximum interval for staleness calculations. By default it is automatically calculated from the median interval between samples. This flag could be useful for tuning Prometheus data model closer to Influx-style data model. See <https://prometheus.io/docs/prometheus/latest/querying/basics/#staleness> for details. See also '-search.maxLookback' flag, which has the same meaning due to historical reasons

-search.minStalenessInterval duration

The minimum interval for staleness calculations. This flag could be useful for removing gaps on graphs generated from time series with irregular intervals between samples. See also '-search.maxStalenessInterval'

-search.queryStats.lastQueriesCount /api/v1/status/top_queries

Query stats for `/api/v1/status/top_queries` is tracked on this number of last queries.
Zero value disables query stats tracking (default 20000)

-search.queryStats.minQueryDuration /api/v1/status/top_queries

The minimum duration for queries to track in query stats at `/api/v1/status/top_queries`.
Queries with lower duration are ignored in query stats

-search.resetCacheAuthKey string

Optional authKey for resetting rollup cache via `/internal/resetRollupResultCache` call

-search.storageTimeout duration

The timeout for per-storage query processing; this allows returning partial responses
if certain `-storageNode` instances slowly process the query; see also

`-search.maxQueryDuration` and `-search.denyPartialResponse` command-line flags

-search.treatDotsAsIsInRegexps

Whether to treat dots as is in regexp label filters used in queries. For example,
`foo{bar=~"a.b.c"}` will be automatically converted to `foo{bar=~"a\\.b\\.c"}`, i.e. all the dots
in regexp filters will be automatically escaped in order to match only dot char instead of
matching any char. Dots in `".+"`, `".*"` and `".{n}"` regexps aren't escaped. Such escaping
can be useful when querying Graphite data

-selectNode array

Addresses of vmselect nodes; usage: `-selectNode=vmselect-host1:8481`

`-selectNode=vmselect-host2:8481`

Supports array of values separated by comma or specified via multiple flags.

-storageNode array

Addresses of vmstorage nodes; usage: `-storageNode=vmstorage-host1:8401`

`-storageNode=vmstorage-host2:8401`

Supports array of values separated by comma or specified via multiple flags.

-tls

Whether to enable TLS (aka HTTPS) for incoming requests. `-tlsCertFile` and `-tlsKeyFile` must be set if `-tls` is set

-tlsCertFile string

Path to file with TLS certificate. Used only if `-tls` is set. Prefer ECDSA certs instead of RSA certs, since RSA certs are slow

-tlsKeyFile string

Path to file with TLS key. Used only if `-tls` is set

vmStorage

-bigMergeConcurrency int

The maximum number of CPU cores to use for big merges. Default value is used if set to 0

-dedup.minScrapeInterval duration

Remove superfluous samples from time series if they are located closer to each other than this duration. This may be useful for reducing overhead when multiple identically configured Prometheus instances write data to the same VictoriaMetrics. Deduplication is disabled if the `-dedup.minScrapeInterval` is 0

-denyQueriesOutsideRetention

Whether to deny queries outside of the configured `-retentionPeriod`. When set, then `/api/v1/query_range` would return '503 Service Unavailable' error for queries with 'from' value outside `-retentionPeriod`. This may be useful when multiple data sources with distinct retentions are hidden behind query-tee

-enableTCP6

Whether to enable IPv6 for listening and dialing. By default only IPv4 TCP is used

-envflag.enable

Whether to enable reading flags from environment variables additionally to command line. Command line flag values have priority over values from environment vars. Flags are read only from command line if this flag isn't set

-envflag.prefix string

Prefix for environment variables if -envflag.enable is set

-finalMergeDelay duration

The delay before starting final merge for per-month partition after no new data is ingested into it. Final merge may require additional disk IO and CPU resources. Final merge may increase query speed and reduce disk space usage in some cases. Zero value disables final merge

-forceFlushAuthKey string

authKey, which must be passed in query string to /internal/force_flush pages

-forceMergeAuthKey string

authKey, which must be passed in query string to /internal/force_merge pages

-fs.disableMmap

Whether to use pread() instead of mmap() for reading data files. By default mmap() is used for 64-bit arches and pread() is used for 32-bit arches, since they cannot read data files bigger than 2^{32} bytes in memory. mmap() is usually faster for reading small data chunks than pread()

-http.connTimeout duration

Incoming http connections are closed after the configured timeout. This may help spreading incoming load among a cluster of services behind load balancer. Note that the real timeout may be bigger by up to 10% as a protection from Thundering herd problem (default 2m0s)

-http.disableResponseCompression

Disable compression of HTTP responses for saving CPU resources. By default compression is enabled to save network bandwidth

-http.idleConnTimeout duration

Timeout for incoming idle http connections (default 1m0s)

-http.maxGracefulShutdownDuration duration

The maximum duration for graceful shutdown of HTTP server. Highly loaded server may require increased value for graceful shutdown (default 7s)

-http.pathPrefix string

An optional prefix to add to all the paths handled by http server. For example, if '-http.pathPrefix=/foo/bar' is set, then all the http requests will be handled on '/foo/bar/*' paths. This may be useful for proxied requests. See <https://www.robustperception.io/using-external-urls-and-proxies-with-prometheus>

-http.shutdownDelay duration

Optional delay before http server shutdown. During this delay the server returns non-OK responses from /health page, so load balancers can route new requests to other servers

-httpListenAddr string

Address to listen for http connections (default ":8482")

-loggerDisableTimestamps

Whether to disable writing timestamps in logs

-loggerErrorsPerSecondLimit int

Per-second limit on the number of ERROR messages. If more than the given number of errors are emitted per second, then the remaining errors are suppressed. Zero value disables the rate limit

-loggerFormat string

Format for logs. Possible values: default, json (default "default")

-loggerLevel string

Minimum level of errors to log. Possible values: INFO, WARN, ERROR, FATAL, PANIC (default "INFO")

-loggerOutput string

Output for the logs. Supported values: stderr, stdout (default "stderr")

-loggerWarnsPerSecondLimit int

Per-second limit on the number of WARN messages. If more than the given number of warns are emitted per second, then the remaining warns are suppressed. Zero value disables the rate limit

-memory.allowedBytes value

Allowed size of system memory VictoriaMetrics caches may occupy. This option overrides -memory.allowedPercent if set to non-zero value. Too low value may increase cache miss rate, which usually results in higher CPU and disk IO usage. Too high value may evict too much data from OS page cache, which will result in higher disk IO usage

Supports the following optional suffixes for values: KB, MB, GB, KiB, MiB, GiB (default 0)

-memory.allowedPercent float

Allowed percent of system memory VictoriaMetrics caches may occupy. See also `-memory.allowedBytes`. Too low value may increase cache miss rate, which usually results in higher CPU and disk IO usage. Too high value may evict too much data from OS page cache, which will result in higher disk IO usage (default 60)

-precisionBits int

The number of precision bits to store per each value. Lower precision bits improves data compression at the cost of precision loss (default 64)

-retentionPeriod value

Data with timestamps outside the retentionPeriod is automatically deleted

The following optional suffixes are supported: h (hour), d (day), w (week), y (year). If suffix isn't set, then the duration is counted in months (default 1)

-rpc.disableCompression

Disable compression of RPC traffic. This reduces CPU usage at the cost of higher network bandwidth usage

-search.maxTagKeys int

The maximum number of tag keys returned per search (default 100000)

-search.maxTagValueSuffixesPerSearch int

The maximum number of tag value suffixes returned from `/metrics/find` (default 100000)

-search.maxTagValues int

The maximum number of tag values returned per search (default 100000)

-search.maxUniqueTimeseries int

The maximum number of unique time series each search can scan (default 300000)

-smallMergeConcurrency int

The maximum number of CPU cores to use for small merges. Default value is used if set to 0

-snapshotAuthKey string

authKey, which must be passed in query string to /snapshot* pages

-storageDataPath string

Path to storage data (default "vmstorage-data")

-tls

Whether to enable TLS (aka HTTPS) for incoming requests. -tlsCertFile and -tlsKeyFile must be set if -tls is set

-tlsCertFile string

Path to file with TLS certificate. Used only if -tls is set. Prefer ECDSA certs instead of RSA certs, since RSA certs are slow

-tlsKeyFile string

Path to file with TLS key. Used only if -tls is set

-vminsertAddr string

TCP address to accept connections from vminsert services (default ":8400")

-vmselectAddr string

TCP address to accept connections from vmselect services (default ":8401")