

Evolveum

MidPoint 4.4 – Tasks



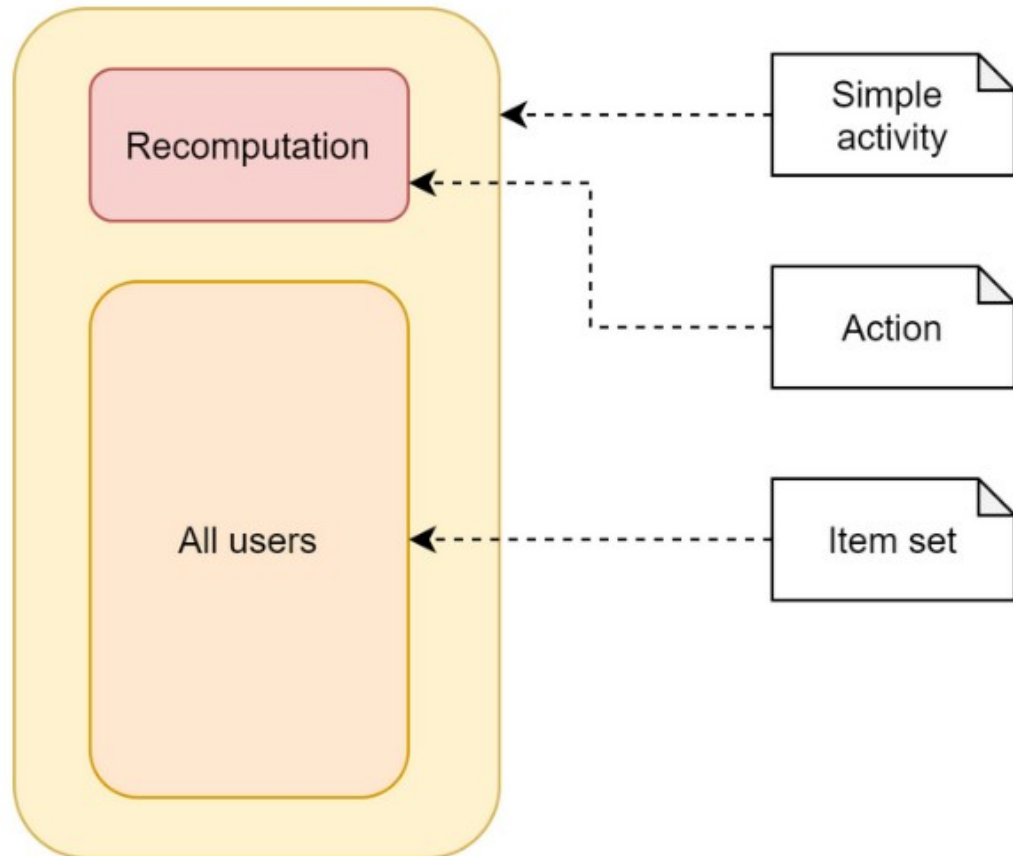
- Activities (big redesign)
- Simplified task definition
- Improvements
- New features
- Migration

- Tasks are here since first versions of midPoint (2011)
 - 3.1 (2015): multiple threads
 - **3.8 (2018): multiple nodes (clustering)**
 - 4.0 (2019): thresholds on a single node (lightweight partitioning)
- Not all the features were easy to use, though.
- After gathering enough experience, we were able to redesign the task execution from scratch.

| Activity-based Conceptual Model

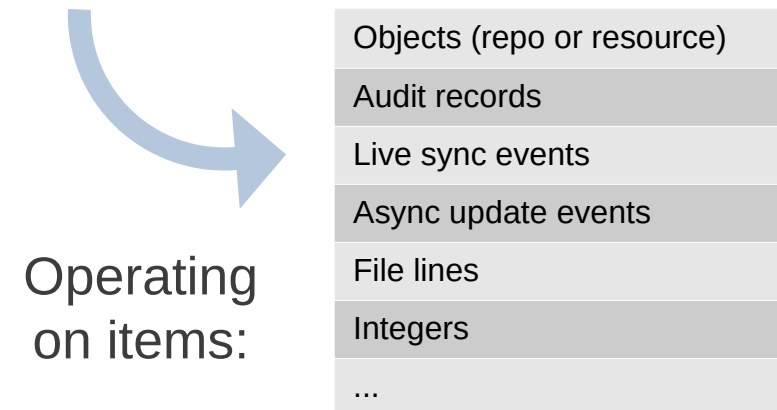
Activity = work that needs to be done

Simple activity = **action** + **item set**



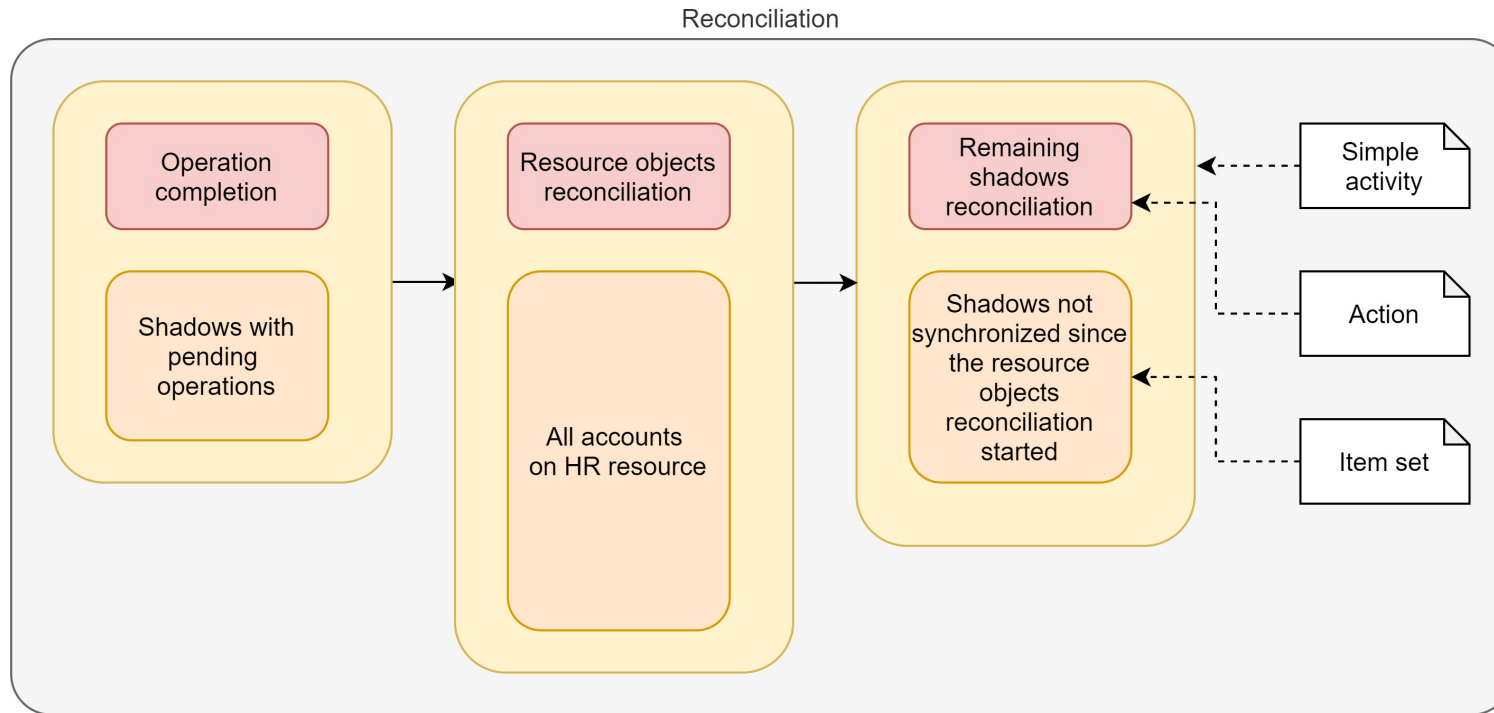
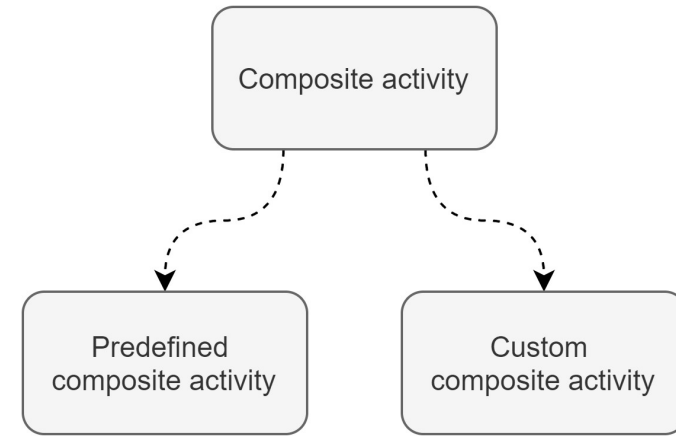
Actions for simple activities

recomputation	triggerScan
import	shadowRefresh
asynchronousUpdate	shadowCleanup
liveSynchronization	reindexing
reportExport	objectIntegrityCheck
reportImport	shadowIntegrityCheck
iterativeScripting	activityAutoScaling
nonIterativeScripting	propagation
changeExecution	multiPropagation
noOp	



Composite Activities

Composite activity: contains sub-activities



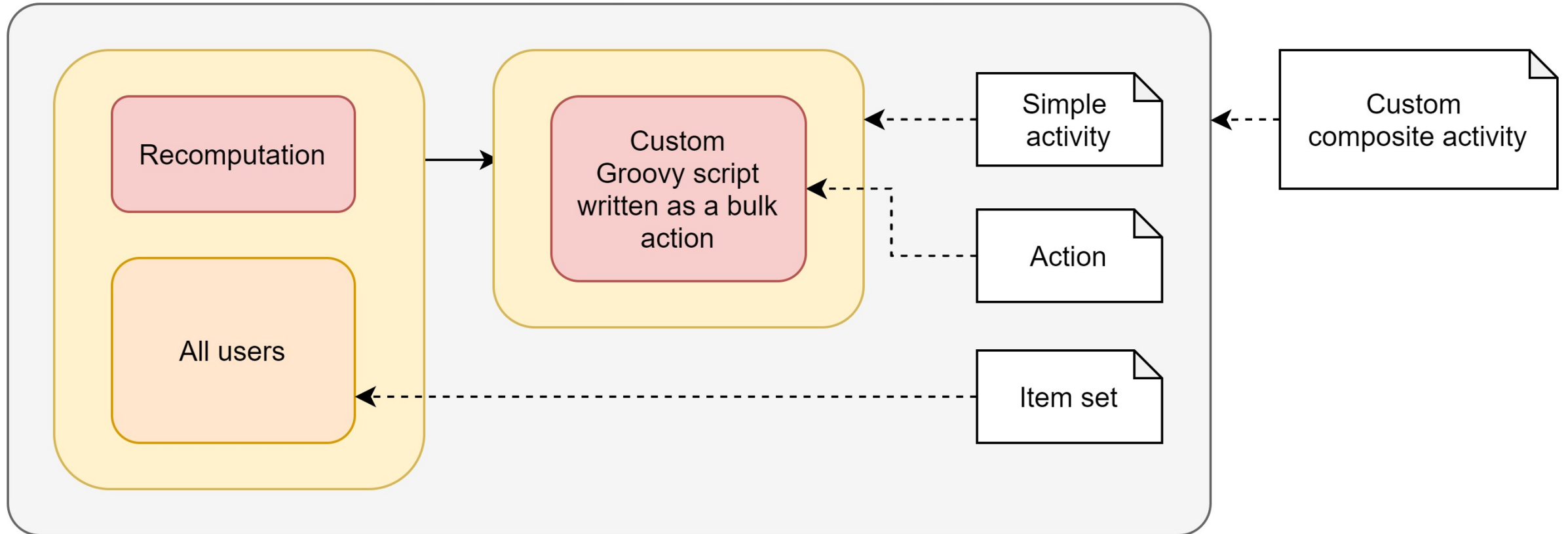
Types of predefined composite activities:

- reconciliation
- cleanup
- distributedReportExport
- focusValidityScan

Custom Composite Activities



Custom composition (recomputation + script execution)



Simplified Task Definition

```
<task
  xmlns="http://midpoint.evolveum.com/xml/ns/public/common/common-3"
  xmlns:ext="http://midpoint.evolveum.com/xml/ns/public/model/extension-3">

  <name>Reconciliation</name>
  <extension>
    <ext:kind>account</ext:kind>
    <ext:intent>default</ext:intent>
  </extension>
  <ownerRef oid="a59e4b45-12e7-44e8-bf8c-b51d3a1bbc3e" type="UserType"/>
  <executionState>runnable</executionState>
  <handlerUri>http://midpoint.evolveum.com/xml/ns/public/model/synchronization/task/reconciliation/handler-3</handlerUri>
  <objectRef oid="9125fb57-9f62-4c82-9122-839fa12cf74b" type="ResourceType"/>
</task>
```

- Generic structures: extension, handler URI, object reference
- Dynamic & flexible, but **hard to write**
(e.g. what extension items are supported?)

| The Clustered Reconciliation Example

```
<task oid="..." xmlns="...">
  ...
  <extension>
    <mext:objectclass>ri:AccountObjectClass</mext:objectclass>
    <mext:workerThreads>4</mext:workerThreads>
  </extension>

  <handlerUri>http://.../partitioned-reconciliation/handler-3</handlerUri>
  <objectRef oid="c845f703-b3c9-..." type="ResourceType" />

  ...
  <workManagement>
    <partitions>
      <partition>
        <index>2</index> <!-- 2 = resource reconciliation -->
        <handlerUri> ... </handlerUri> <!-- optional -->
        <workManagement>
          <taskKind>coordinator</taskKind>
          <buckets>
            <stringSegmentation> ... </stringSegmentation>
          </buckets>
          <workers>
            <workersPerNode> ... </workersPerNode>
          </workers>
        </workManagement>
      </partition>
    </partitions>
  </workManagement>
</task>
```

Red = what to do
Dark red = on which objects
Blue = work distribution

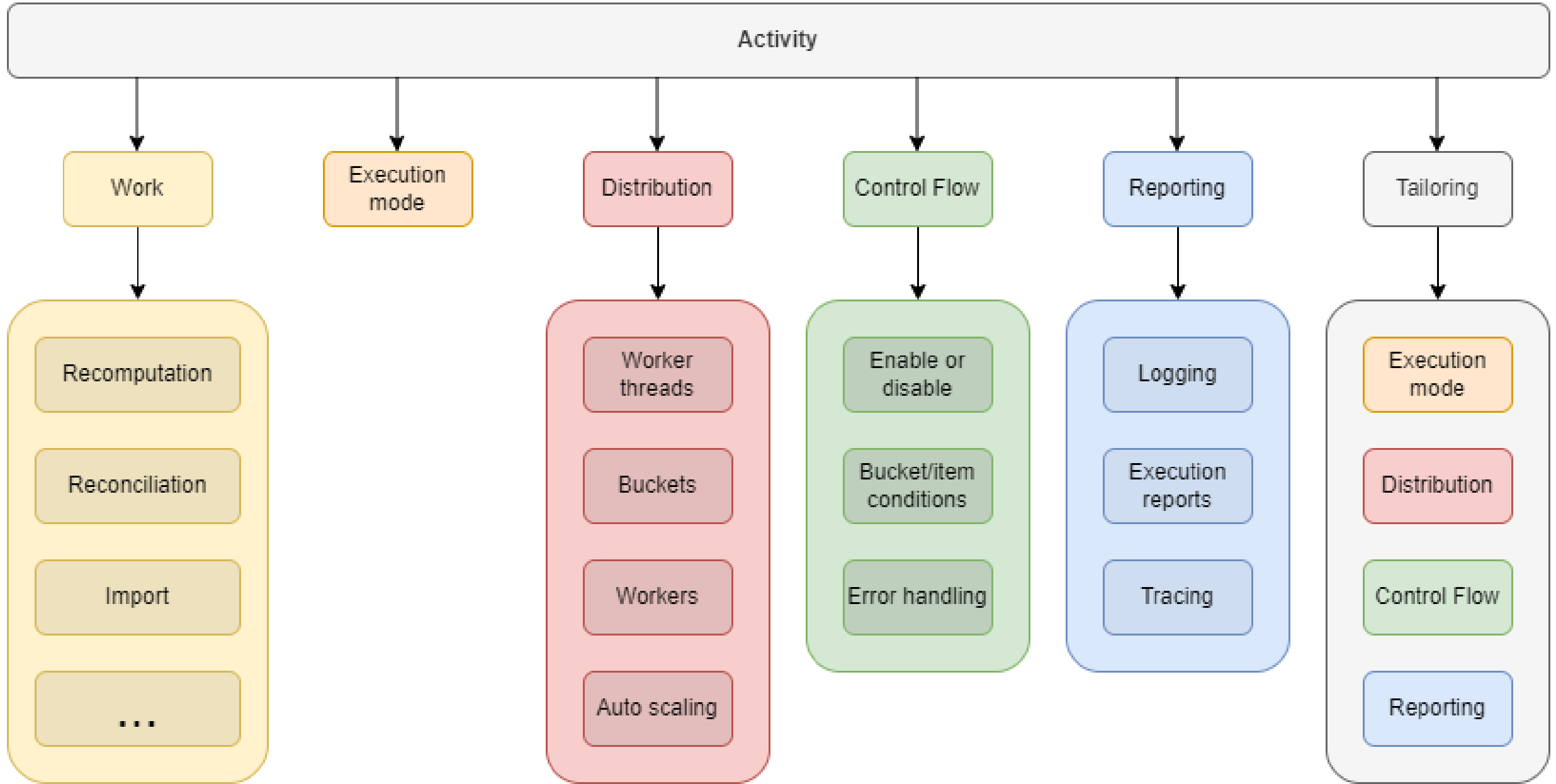
```
<task
  xmlns="http://midpoint.evolveum.com/xml/ns/public/common/common-3">

  <name>Reconciliation</name>
  <ownerRef oid="a59e4b45-12e7-44e8-bf8c-b51d3a1bbc3e" type="UserType"/>
  <executionState>runnable</executionState>
  <activity>
    <work>
      <reconciliation>
        <resourceObjects>
          <resourceRef oid="9125fb57-9f62-4c82-9122-839fa12cf74b"/>
          <kind>account</kind>
          <intent>default</intent>
        </resourceObjects>
      </reconciliation>
    </work>
  </activity>
</task>
```

- Everything in one place
- Statically-typed, yet customer-extensible
- Easy to read, easy to write

Even more important when advanced features (like [clustering](#)) are to be configured!

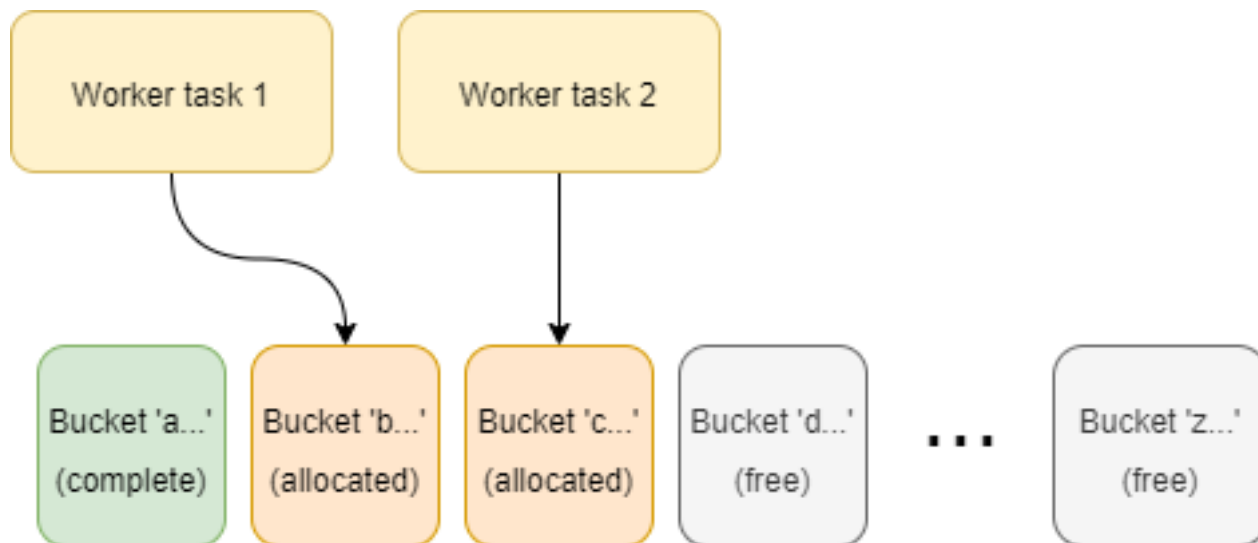
Activity Definition Overview



| Short Demo

| What are Buckets, Anyway?

- A task obtains objects typically through a query.
 - (They are then processed in single or multiple worker threads.)
- If we want to have multiple worker tasks, we need to split the objects set to smaller chunks, called **buckets**.
- **Worker tasks** then allocate and process individual buckets.



```
<distribution>
  <buckets>
    <stringSegmentation>
      <discriminator>attributes/icfs:name</discriminator>
      <boundary>
        <characters>a-z</characters>
      </boundary>
    </stringSegmentation>
  </buckets>
  <workers>
    <workersPerNode>
      <count>1</count>
    </workersPerNode>
  </workers>
</distribution>
```

Visibility Improvements

Progress, Items Processed, and Throughput

- Progress = how far we have got up to now
- Items processed = how many items we have processed
- Throughput = items processed per time unit

Progress + throughput



Estimated time to complete (future)

The screenshot shows the 'User recomputation' interface. At the top left, there is a green gear icon and the title 'User recomputation'. Below the title, it displays '34.6% (89 of 257 buckets) (suspended)' and 'Started 2/1/22, 10:50:12 PM (00:11:02.627 ago)'. A yellow box highlights '35,174 users (not shown)'. On the right, there are three buttons: 'Suspended', 'In Progress', and 'Recomputation task'. Below this is a toolbar with 'Operations' (Back, Save, Save & Run, Delete object, Edit raw), 'Task operations' (Resume, and two charts), and 'Refresh' (Refresh icon, Play icon, and '(no refreshing)'). On the left, a sidebar menu includes 'Basic', 'Activity', 'Schedule', 'Operation statistics', 'Errors', 'Performance', 'Subtasks', 'Result', and 'Advanced options'. The main content area shows 'Objects processed: 36,481 (might contain duplicates in case task was suspended/resumed)' and 'Wall clock average: 17.8 ms, throughput: 3,375.2 per minute'. A pie chart shows 'Success (36,481)', 'Failure (0)', and 'Skip (0)'. A green box on the right contains a checkmark, 'LAST SUCCESS ON TUESDAY, 1. FEB 2022 23:01:13', 'u197630', and '420 milliseconds'.

Activity Tree Overview

Key information in the root task – no more need to “click through” to see it

≡ All tasks Server tasks > All tasks administrator

Name ✕ More... Basic

<input type="checkbox"/>	^ Name	Category	Execution	Object reference	Executing at	Current run time	Scheduled to start again	Progress	Errors	Status	
<input type="checkbox"/>	Cleanup	Cleanup task	Runnable				in 23 hours 5 minutes 40 seconds	Complete	0		
<input type="checkbox"/>	Import	Import task	Closed	CSV 100K		closed at 2/1/22, 10:24:40 PM		100001	0		
<input type="checkbox"/>	Trigger Scanner	System task	Runnable				in 40 seconds	0	0		
<input type="checkbox"/>	User recomputation	Recomputation task	Running		DefaultNode (4)			3.5%	0		
<input type="checkbox"/>	Validity Scanner	System task	Runnable				in 5 minutes 40 seconds	0	0		

1 to 5 of 5 << < 1 > >>

New Features

- Starting or suspending worker tasks when adding/removing nodes
- **Manually** or **automatically**

```
<task>
  <name>Auto scaling</name>
  <!-- ownerRef and execution state -->
  <schedule>
    <interval>60</interval>
  </schedule>
  <activity>
    <work>
      <activityAutoScaling/>
    </work>
  </activity>
</task>
```

	Name	Category	Execution	Object reference	Executing at	Current run time	Scheduled to start again	Progress	Errors	Status	
<input type="checkbox"/>	Cleanup	Cleanup task	Runnable				in 22 hours 29 minutes 1 second	Complete	0	✓	
<input type="checkbox"/>	Import	Import task	Closed	CSV 100K		closed at 2/1/22, 10:24:40 PM		100001	0	✓	
<input type="checkbox"/>	Trigger Scanner	System task	Runnable				in 4 minutes 1 second	0	0	✓	
<input type="checkbox"/>	User recomputation	Recomputation task	Running		DefaultNode (4)			34.6%	0	🕒	
<input type="checkbox"/>	Validity Scanner	System task	Runnable				in 14 minutes 1 second	0			

Context menu for 'User recomputation':

- Suspend
- Delete
- Reconcile workers
- Suspend (root only)
- Resume (root only)
- Suspend and delete workers and work state
- Suspend and delete work state

Cluster-wide Thresholds

- Sometimes we want to block unexpected actions before they cause a catastrophe.
- MidPoint allows to declare limits on actions in the form of policy rules.

```
<role>
  <name>Stop on delete 10 users</name>
  <inducement>
    <policyRule>
      <policyConstraints>
        <modification>
          <operation>delete</operation>
        </modification>
      </policyConstraints>
      <policyThreshold>
        <lowWaterMark>
          <count>10</count>
        </lowWaterMark>
      </policyThreshold>
      <policyActions>
        <suspendTask/>
      </policyActions>
    </policyRule>
    <order>2</order>
  </inducement>
</role>
```

- Roles with such rules are then assigned to tasks.
- After the threshold is reached, a task is suspended.
 - Optionally, a simulation activity can be run right before the execution takes place.
- In midPoint 4.0 the counters were held in memory, so this feature was limited to single node only.
- In 4.4 we maintain the counters directly in the tasks. Therefore, cluster-wide operation is possible.



- Allows reporting on the execution at the levels of:
 - buckets,
 - items (~ objects),
 - individual ConnId operations,
 - individual internal midPoint operations.
- Easy default configuration, but with great level of flexibility

Buckets Report

#	content	content-from	content-to	size	itemsSuccessfullyProcessed	itemsFailed	itemsSkipped	startTimestamp	startTimestamp-millis	endTimestamp	endTimestamp-millis	duration
1	[5-230004)	5	230004	2935	2927	8	0	2021-10-08T11:00:19.015Z	1633690819015	2021-10-08T11:02:38.528Z	1633690958528	139513
6	[1150000-1379999)	1150000	1379999	2297	2291	6	0	2021-10-08T11:02:38.734Z	1633690958734	2021-10-08T11:04:35.088Z	1633691075088	116354
10	[2069996-2299995)	2069996	2299995	2061	2055	6	0	2021-10-08T11:04:35.239Z	1633691075239	2021-10-08T11:06:17.209Z	1633691177209	101970
14	[2989992-3219991)	2989992	3219991	3747	3734	13	0	2021-10-08T11:06:17.307Z	1633691177307	2021-10-08T11:09:25.689Z	1633691365689	188382
18	[3909988-4139987)	3909988	4139987	1462	1460	2	0	2021-10-08T11:09:25.791Z	1633691365791	2021-10-08T11:10:40.895Z	1633691440895	75104
20	[4369986-4599985)	4369986	4599985	1920	1919	1	0	2021-10-08T11:10:41.018Z	1633691441018	2021-10-08T11:12:15.892Z	1633691535892	94874
23	[5059983-5289982)	5059983	5289982	2426	2419	7	0	2021-10-08T11:12:15.987Z	1633691535987	2021-10-08T11:14:13.690Z	1633691653690	117703
35	[7819971-8049970)	7819971	8049970	505	499	6	0	2021-10-08T11:14:13.805Z	1633691653805	2021-10-08T11:14:38.003Z	1633691678003	24198
38	[8509968-8739967)	8509968	8739967	1468	1461	7	0	2021-10-08T11:14:38.093Z	1633691678093	2021-10-08T11:15:51.241Z	1633691751241	73148
47	[10579959-10809958)	10579959	10809958	17	15	2	0	2021-10-08T11:15:51.406Z	1633691751406	2021-10-08T11:15:52.899Z	1633691752899	1493

```
<reporting>
  <executionReports>
    | <buckets/>
  </executionReports>
</reporting>
```

Configurability:

- Row selection (using filter and/or expression)

Items (Objects) Report



#	name	displayName	type	oid	bucket#	outcome	outcome-qualifier	startTimestamp	startTimestamp-millis	endTimestamp	endTimestamp-millis	duration	errorMessage
0	ht	ht (ACCOUNT - default - AccountObjectClass)	{http://midpoint.evolveum.com/xml/ns/public/common/common-3}ShadowType	e010c326-071a-4666-b791-4be5e8480527	1	SUCCESS		2021-10-19T11:50:18.744+02:00	1634637018744	2021-10-19T11:50:18.832+02:00	1634637018832	88.09	
1	guybrush	guybrush (ACCOUNT - default - AccountObjectClass)	{http://midpoint.evolveum.com/xml/ns/public/common/common-3}ShadowType	22226666-2200-6666-6666-444400004444	1	SUCCESS		2021-10-19T11:50:18.917+02:00	1634637018917	2021-10-19T11:50:19.028+02:00	1634637019028	110.6	
2	daviejones	daviejones (ACCOUNT - default - AccountObjectClass)	{http://midpoint.evolveum.com/xml/ns/public/common/common-3}ShadowType	bb6a968d-3a0c-4753-9b58-6c932b1a5245	1	SKIP		2021-10-19T11:50:19.050+02:00	1634637019050	2021-10-19T11:50:19.051+02:00	1634637019051	0.29	
3	calypso	calypso (ACCOUNT - default - AccountObjectClass)	{http://midpoint.evolveum.com/xml/ns/public/common/common-3}ShadowType	a4793302-0624-42bd-9527-98a39ede2621	1	SKIP		2021-10-19T11:50:19.055+02:00	1634637019055	2021-10-19T11:50:19.056+02:00	1634637019056	0.28	
4	elaine	elaine (ACCOUNT - default - AccountObjectClass)	{http://midpoint.evolveum.com/xml/ns/public/common/common-3}ShadowType	c0c010c0-d34d-b33f-f00d-22220004000e	1	SUCCESS		2021-10-19T11:50:19.061+02:00	1634637019061	2021-10-19T11:50:19.191+02:00	1634637019191	129.55	
5	rapp	rapp (ACCOUNT - default - AccountObjectClass)	{http://midpoint.evolveum.com/xml/ns/public/common/common-3}ShadowType	c70312bb-a6fd-48d7-b2ed-303007a2e190	1	SUCCESS		2021-10-19T11:50:19.206+02:00	1634637019206	2021-10-19T11:50:19.331+02:00	1634637019331	125.03	
6	stan	stan (ACCOUNT - default - AccountObjectClass)	{http://midpoint.evolveum.com/xml/ns/public/common/common-3}ShadowType	22220000-2200-0000-0000-444400004455	1	SUCCESS		2021-10-19T11:50:19.348+02:00	1634637019348	2021-10-19T11:50:19.404+02:00	1634637019404	55.81	

ConnId Operations Report



#	itemName	itemOid	bucket#	identifier	resourceRef	resourceRef-name	objectClass	operation	status	message	size	startTimestamp	startTimestamp-millis	endTimestamp	endTimestamp-millis	duration
1	guybrush	22226666-2200-6666-6666-444400004444	1	1634637018962-0-1	10000000-0000-0000-0000-0000000000004	Dummy Resource	{http://midpoint.evolveum.com/xml/ns/public/resource/instance-3}AccountObjectClass	ICF_GET				2021-10-19T11:50:18.962+02:00	1634637018962	2021-10-19T11:50:18.962+02:00	1634637018962	0.0
1	guybrush	22226666-2200-6666-6666-444400004444	1	1634637018965-0-1	10000000-0000-0000-0000-0000000000004	Dummy Resource	{http://midpoint.evolveum.com/xml/ns/public/resource/instance-3}AccountObjectClass	ICF_UPDATE				2021-10-19T11:50:18.965+02:00	1634637018965	2021-10-19T11:50:18.965+02:00	1634637018965	0.0
1	guybrush	22226666-2200-6666-6666-444400004444	1	1634637019003-0-1	10000000-0000-0000-0000-0000000000004	Dummy Resource	{http://midpoint.evolveum.com/xml/ns/public/resource/instance-3}AccountObjectClass	ICF_GET				2021-10-19T11:50:19.003+02:00	1634637019003	2021-10-19T11:50:19.004+02:00	1634637019004	1.0
4	elaine	c0c010c0-d34d-b33f-f00d-22220004000e	1	1634637019094-0-1	10000000-0000-0000-0000-000000000104	Dummy Resource Red	{http://midpoint.evolveum.com/xml/ns/public/resource/instance-3}AccountObjectClass	ICF_GET				2021-10-19T11:50:19.094+02:00	1634637019094	2021-10-19T11:50:19.095+02:00	1634637019095	1.0
4	elaine	c0c010c0-d34d-b33f-f00d-22220004000e	1	1634637019099-0-1	10000000-0000-0000-0000-000000000204	Dummy Resource Blue	{http://midpoint.evolveum.com/xml/ns/public/resource/instance-3}AccountObjectClass	ICF_GET				2021-10-19T11:50:19.099+02:00	1634637019099	2021-10-19T11:50:19.099+02:00	1634637019099	0.0
			1	1634637018739-0-1	10000000-0000-0000-0000-0000000000004	Dummy Resource	{http://midpoint.evolveum.com/xml/ns/public/resource/instance-3}AccountObjectClass	ICF_SEARCH				2021-10-19T11:50:18.739+02:00	1634637018739	2021-10-19T11:50:19.414+02:00	1634637019414	0.0



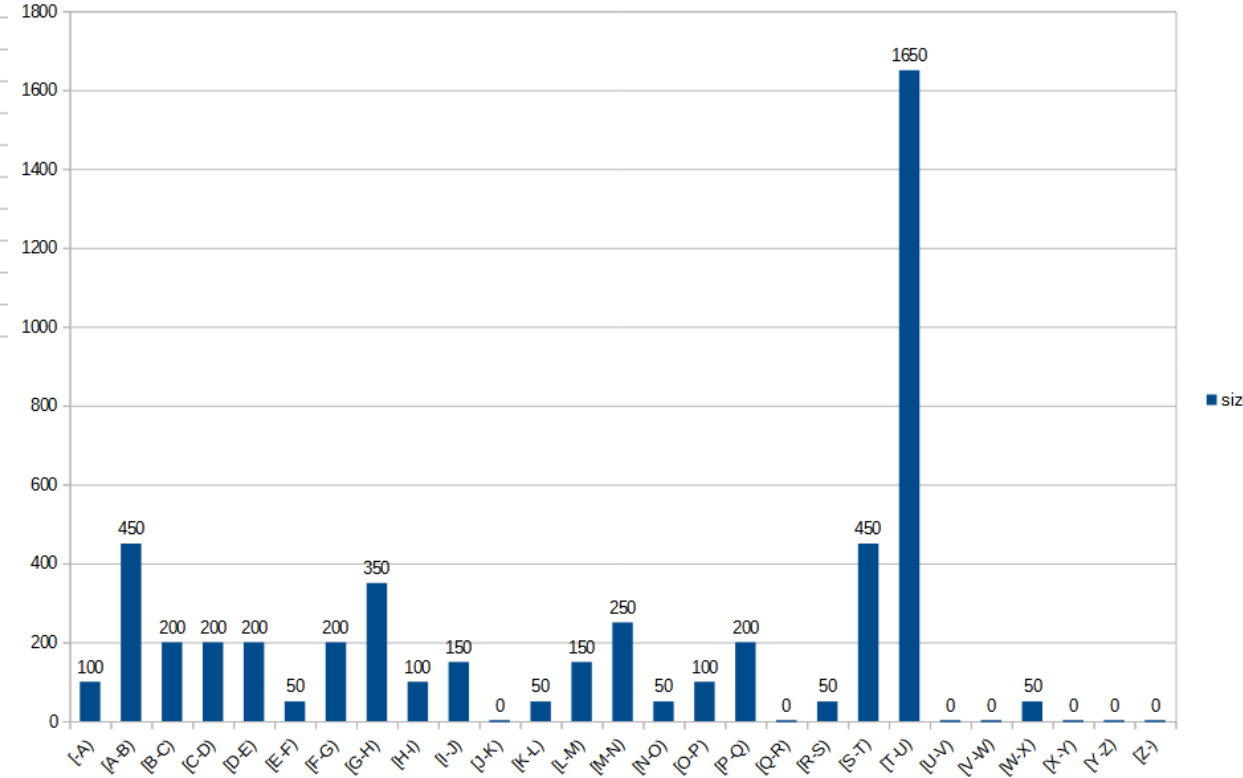
- Important for attaining top performance
 - Large/few buckets
 - inefficient distribution, coarse %, wasteful restarts
 - Small/many buckets
 - large overhead (including the contention)
 - Uneven distribution
 - short & long runners
- Options: ex-post, ex-ante

Bucket Sampling

```



<activity>
  <work>
    <noOp>
      <steps>10000</steps>
      <delay>10</delay>
      <stepInterruptibility>full</stepInterruptibility>
    </noOp>
  </work>
  <executionMode>bucketAnalysis</executionMode>
  <distribution>
    <buckets>
      <implicitSegmentation>
        <!-- 10000 steps divided into 200 buckets -->
        <numberOfBuckets>200</numberOfBuckets>
      </implicitSegmentation>
      <sampling>
        <regular>
          <sampleSize>10</sampleSize>
        </regular>
      </sampling>
    </buckets>
  </distribution>
  <reporting>
    <executionReports>
      <buckets/>
    </executionReports>
  </reporting>
</activity>
  
```

sequentialNum	content	content-from	content-to	size
1	[A]		A	100
2	[A-B]	A	B	450
3	[B-C]	B	C	200
4	[C-D]	C	D	200
5	[D-E]	D	E	200
6	[E-F]	E	F	50
7	[F-G]	F	G	200
8	[G-H]	G	H	350
9	[H-I]	H	I	100
10	[I-J]	I	J	150
11	[J-K]	J	K	0
12	[K-L]	K	L	50
13	[L-M]	L	M	150
14	[M-N]	M	N	250
15	[N-O]	N	O	50
16	[O-P]	O	P	100
17	[P-Q]	P	Q	200
18	[Q-R]	Q	R	0
19	[R-S]	R	S	0
20	[S-T]	S	T	0
21	[T-U]	T	U	450
22	[U-V]	U	V	0
23	[V-W]	V	W	0
24	[W-X]	W	X	50
25	[X-Y]	X	Y	0
26	[Y-Z]	Y	Z	0
27	[Z-]	Z		0




Other options: counting all buckets, using random sampling (by sample size or probability), or using custom expression.

Beware: Counting may not be supported by all resources!

- Failed objects re-processing (since 4.3)
 - Re-process failed objects during separate task execution
 - Re-synchronize failed accounts automatically after given time
- Buckets and items logging configuration
 - E.g. logging full information for each bucket completion, and brief one for each item
- Flexible tracing (before/after conditions), e.g. 
 - Trace each 100-th item (fully) and each 10-th item (briefly)
 - Trace all failed items
 - Trace all items that took longer than 5,000 milliseconds
- Buckets and items processing conditions 
 - Choosing what buckets and/or items to process, and what to skip.

Migration from 4.0.x/4.3.x

- Legacy-style task **execution** is supported for **simple tasks**
 - but no GUI support for these tasks,
 - and no support for advanced features (buckets, workers, partitions) at all.
- Task **state** data structures has changed
 - so it is strongly advisable to re-import even simple tasks,
 - an exception: live sync token and last scan timestamp are migrated automatically.
- The guide to tasks migration is in docs
 - <https://docs.evolveum.com/midpoint/reference/tasks/activities/migration/>
- But midPoint Studio can help with the migration! 

| Migration: An Example

- Assuming you are running 4.0.4 or 4.3.2
- Before stopping midPoint:
 - suspend the tasks,
 - export the tasks (unless kept externally e.g. in Studio).
- Do the upgrade according to the documentation
- Start midPoint and either:
 - migrate the tasks in Studio and re-import them,
 - or simply resume them – but there will be warnings in logs and some pieces missing in GUI.

| Migration Demo

Conclusion

| Conclusion

- Tasks were substantially changed in midPoint 4.4.
- Besides quality improvements, there are a lot of new features.
- Have a look at them, try them, and let us know.
 - (We are interested especially in the experimental ones.)
- Everything is in the docs!
 - <https://docs.evolveum.com/midpoint/reference/tasks/activities/>



This project has received funding from the European Union's Horizon 2020 research and innovation programme under the NGI_TRUST grant agreement no 825618.

- Customizing GUI, February 10, 2022 by Katarína Bolemant
- Native reports, February 16, 2022 by Lukáš Škublík

Thank you for your time!

See other talks at <https://docs.evolveum.com>

Also **follow us** on our social media for further information!



/Evolveum



/Evolveum



/Evolveum



@Evolveum



/Evolveum

Evolveum