Testing summary

MAMAS 0.6.0 performance tests.

MAMAS 0.6.0 is the first MAMAS version to receive proper testing as it is the first to go to production.
There are some limited testing results from beta (0.2.0 - 0.5.0) versions however the methodology was changed.

There are three different test types in MAMAS testing:

- stability tests ... runs for a number of hours.
- benchmark tests ... to measure different HW / VMs for comparability and clue for sizing deployments.
- peek tests ... to discover MAMAS SW and deployments limits and breaking points.

Next to the test results, HW specs for environments, test description and methodology is given in the document.

The MAMAS 0.6.0 early bird version is designed for small to mid-sized deployments (up to 3000 or up to 10000 devices).
As such, not all defined tests were performed. Also, not all environments were ready at the time of MAMAS 0.6.0 testing.
Find out more in test overview.

General MAMAS 0.6.0 performance testing verdict:

PASSED	2022-06-24	Initial test set results: Stab-2-2:C1; Stab-4-2:C2; Stab-8-2:C3; Bench-4-noSSL:C1,C2,C3,C4,H1; Bench-16-noSSL:C1,C2,C3,C4,H1; Bench-32-noSSL:C1,C2,C3,C4,H1; Bench-4-SSL:C1,C2,C3,C4,H1; Bench-16-SSL:C1,C2,C3,C4,H1; Bench-32-SSL:C1,C2,C3,C4,H1;

Document updates:

PASSED	2022-07-04	Added test results for: Stab-4-48:H1

Testing overview

As mentioned earlier (Testing summary), there are three types of performance tests. All are here because we need to measure:
- number of accepted and dropped requests (and resulting percentage) FOR ALL TESTS
- time for peak and benchmark tests
- memory usage for stability tests
- CPU utilization (%) for stability tests
- mamasd log review to make detail search for warnings and errors (for stability tests)
- number of pushing threads causing requests drops+resends (peak test)

Next to measurements there are a few important calculated metrics:
- ideal maximal throughput (benchmark test)
- typical maximal throughput (20% of ideal) (benchmark test)
- number of pushes per second (benchmark and peak test)
- maximal recommended # of devices per deployment type (HW / VM specs)

Environment specs		CPU cores	CPU threads	RAM MiB	stability test	benchmark	peak test
Linux container (Proxmox)	C1	1	1	512			
	C2	2	2	1024			
	C3	4	4	2048			
	C4	8	8	4096			
AWS instance	A1 (small)						
	A2 (medium)						
	A3 (big)						
Baremetal server Core i7-8C-32G	H1	8	16	32768			

Tests specs	ID	SSL?	# of devices	# of threads	# of loops	# of minutes	Human-time	Notes
Stability	Stab-2-2	Yes	1000	2	N/A	120	2 hours	
	Stab-4-2	Yes	3000	4	N/A	120	2 hours	
	Stab-8-2	Yes	10000	8	N/A	120	2 hours	
	Stab-16-2	Yes	30000	16	N/A	120	2 hours	

Testing overview

	Stab-4-48	Yes	3000	4	N/A	2880	48 hours	
	Stab-4-672	Yes	3000	4	N/A	40320	4 weeks	
Benchmark	Bench-4-noSSL	No	3000	4	10	N/A		
	Bench-16-noSSL	No	10000	16	10	N/A		
	Bench-32-noSSL	No	30000	32	10	N/A		
	Bench-4-SSL	Yes	3000	4	10	N/A		
	Bench-16-SSL	Yes	10000	16	10	N/A		
	Bench-32-SSL	Yes	30000	32	10	N/A		
Peak	PeakR-64-max	No	60000	64	N/A	10		
	PeakR-256-max	No	60000	256	N/A	10		
	PeakR-1024-max	No	60000	1024	N/A	10		
	PeakD-64-max	No	60000	64	N/A	10		
	PeakD-256-max	No	60000	256	N/A	10		
	PeakD-1024-max	No	60000	1024	N/A	10		
	PeakF-64-max	No	60000	64	N/A	10		
	PeakF-256-max	No	60000	256	N/A	10		
	PeakF-1024-max	No	60000	1024	N/A	10		

Test run 0.6.0	C1	C2	C3	C4	A1	A2	A3	H1
Stab-2-2	Yes							
Stab-4-2		Yes						
Stab-8-2			Yes					
Stab-16-2								
Stab-4-48								Yes
Stab-4-672								
Bench-4-noSSL	Yes	Yes	Yes	Yes				Yes
Bench-16-noSSL	Yes	Yes	Yes	Yes				Yes
Bench-32-noSSL	Yes	Yes	Yes	Yes				Yes
Bench-4-SSL	Yes	Yes	Yes	Yes				Yes
Bench-16-SSL	Yes	Yes	Yes	Yes				Yes
Bench-32-SSL	Yes	Yes	Yes	Yes				Yes
PeakR-64-max								
PeakR-256-max								

PeakR-1024-max								
PeakD-64-max								
PeakD-256-max								
PeakD-1024-max								
PeakF-64-max								
PeakF-256-max								
PeakF-1024-max								

Stability tests						
ID / Environment →			Stab-2-2 / C1	Stab-4-2 / C2	Stab-8-2 C3	Stab-4-48 H1
Test specs		Devices	1000	3000	10000	3000
		Emulator threads	2	4	8	4
		Set: Time / minutes	125	125	125	2880
		Set loop time / sec	260	260	260	260
HW specs		CPU cores	1	2	4	8
		RAM / MiB	512	1024	2048	32768
Resources	Pre-test	Used MiB PX	134.62	134.45	164.12	
		Used MiB OS	141	141	167	
		mamasd RSS	34508	34360	34444	37116
	Post-test	Total MiB PX	168.71	216.43	394.00	
		Total MiB OS	169	216	394	
		mamasd RSS / KiB	49411	54006	111897	60604
		AVG mamasd % CPU	0.41	1.05	3.31	1.31
Results	Test numbers	Processesd loops	26	25	25	572
		Loop time / sec	288	300	300	303
		Data-pushes	26000	75000	250000	1716000
	Errors	Data ERR/retries	0	0	0	0
		Crashes	0	0	0	0
		Reboots	0	0	0	0
		Log entries	3	4	0	15
		Log analysis	delayed sync	delayed sync		delayed sync
	Calculations	RSS increase KiB	14903	19646	77453	23488
		RSS increase %	43.19	57.18	224.87	63.28
		KiB per device	14.90	6.55	7.75	7.83
		RAM utilization %	27.54	13.77	8.15	
		Max safe # of devs	3631	21787	122635	
	Summary	RAM utilization note	medium	low	low	low
		CPU utilization note	imperceptible	very low	very low	very low
		Verdict	PASSED	PASSED	PASSED	PASSED

Log analysis explained

delayed sync	Under a load, some DB-backed data synchronization to DB may be (and was) delayed – especially in artificial testing conditions. All DB-backed tables went to fully-synchronized mode during the testing period (the sync delay was intermittent – shorter than 2 minutes)

Benchmark tests		bench-4-SSL	bench-16-SSL	bench-32-SSL	bench-4-noSSL	bench-16-noSSL	bench-32-noSSL
Devices		3000	10000	30000	3000	10000	30000
Loops		10	10	10	10	10	10
Threads		4	16	32	4	16	32
Regular loop time		260	260	260	260	260	260
Container 1	CPUs / MiBs	1 / 512					
	Elapsed time	210.66	467.85		88.76	172.58	
	Pushes / sec	142.41	213.74		337.97	579.44	
	Ideal max # devs	37026	55573		87873	150655	
	Rec. max # devs	7405	11115		17575	30131	
	Notes	RAM size					
Container 2	CPUs / MiBs	2 / 1024					
	Elapsed time	165.34	336.91		82.84	168.61	
	Pushes / sec	181.44	296.81		362.15	593.07	
	Ideal max # devs	47175	77171		94158	154199	
	Rec. max # devs	9435	15434		18832	30840	
	Notes	CPU limit			2T limit	2T limit	
Container 3	CPUs / MiBs	4 / 2048					
	Elapsed time	159.10	334.74		83.74	170.62	
	Pushes / sec	188.56	298.74		358.23	586.10	
	Ideal max # devs	49024	77673		93141	152385	
	Rec. max # devs	9805	15535		18628	30477	
	Notes	2T limit	2T limit		2T limit	2T limit	
Container 4	CPUs / MiBs	8 / 4069					
	Elapsed time	158.45	333.13		82.28	168.01	
	Pushes / sec	189.34	300.18		364.62	595.19	
	Ideal max # devs	49228	78047		94800	154750	
	Rec. max # devs	9846	15609		18960	30950	
	Notes	2T limit	2T limit		2T limit	2T limit	
Baremetal H1	CPUs / MiBs	8 / 32768					
	Elapsed time	155.58	311.74		75.94	155.51	
	Pushes / sec	192.83	320.78		395.06	643.04	
	Ideal max # devs	50135	83403		102717	167191	
	Rec. max # devs	10027	16681		20543	33438	

Benchmark tests

	Notes	2T limit	2T limit		2T limit	2T limit	
Baremetal H1 MAMAS v0.5	CPUs / MiBs	8 / 32768					
	Elapsed time						
	Pushes / sec	153					
	Ideal max # devs	39780					
	Rec. max # devs	7956					
	Notes						
% increase 0.6 to 0.5		26.03					

Notes explained	
2T limit	Bigger container size has no or negligible impact on benchmark because of the MAMAS 0.6.0 limitation to two request handling threads.
RAM size	Due to container limited RAM, the lower resulting number is taken as recommended maximum number of devices.
CPU limit	Container has only 2 CPUs that could be fully loaded by MAMAS 0.6.0 request handling threads, causing lack of time for 6 to 8 additional (non-request handling) MAMAS threads.

Peak tests are for 60000 devices which is outside of scope for MAMAS 0.6.0