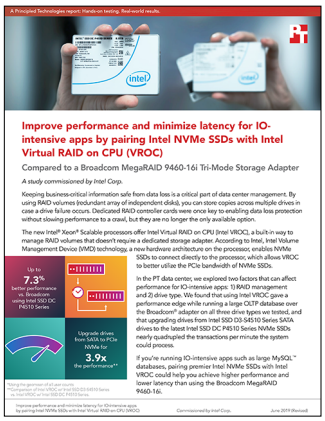

The science behind the report:

Improve performance and minimize latency for IO-intensive apps by pairing Intel NVMe SSDs with Intel Virtual RAID on CPU (VROC)

This document describes what we tested, how we tested, and what we found. To learn how these facts translate into real-world benefits, read the report Improve performance and minimize latency for IO-intensive apps by pairing Intel NVMe SSDs with Intel Virtual RAID on CPU (VROC).

We concluded our hands-on testing on February 21, 2019. During testing, we determined the appropriate hardware and software configurations and applied updates as they became available. The results in this report reflect configurations that we finalized on January 10, 2019 or earlier. Unavoidably, these configurations may not represent the latest versions available when this report appears.

Our results

The tables below presents our findings in detail.

Latency results

The tables below present the median run for each user count for each configuration.

Intel VROC with Intel SSD DC P4510 Series

Users: 64				TPM: 408,882	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	167,442,255	27.78%	21,934	7,633	168,849,943
payment	157,010,115	26.05%	21,478	7,310	158,428,474
delivery	138,603,530	23.00%	2,256	61,437	138,748,265
slev	120,900,231	20.06%	2,201	54,929	120,901,659
ostat	18,179,471	3.02%	2,307	7,880	18,180,879
Geometric mean				17,150	

Users: 96				TPM: 456,140	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	162,831,518	27.14%	16,799	9,692	163,981,704
payment	160,026,157	26.67%	16,657	9,607	161,130,441
delivery	139,103,129	23.19%	1,699	81,873	139,215,583
slev	120,588,961	20.10%	1,648	73,172	120,590,018
ostat	16,980,662	2.83%	1,677	10,125	16,981,732
Geometric mean				22,407	

Users: 128				TPM: 459,580	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	161,823,645	26.92%	13,580	11,916	162,723,990
payment	157,105,535	26.13%	13,260	11,848	158,025,917
delivery	140,264,324	23.33%	1,332	105,303	140,352,851
slev	123,261,909	20.50%	1,277	96,524	123,262,699
ostat	18,455,609	3.07%	1,375	13,422	18,456,462
Geometric mean				28,637	

Broadcom with Intel SSD DC P4510 Series

Users: 64				TPM: 367,709	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	167,110,717	27.88%	19,673	8,494	168,398,718
payment	156,690,856	26.15%	19,470	8,047	157,933,288
delivery	140,709,517	23.48%	2,031	69,280	140,838,338
slev	116,265,951	19.40%	1,926	60,366	116,267,113
ostat	18,060,135	3.01%	2,116	8,535	18,061,421
Geometric mean				18,944	

Users: 96				TPM: 428,547	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	160,497,295	26.72%	15,411	10,414	161,533,837
payment	161,497,201	26.88%	15,689	10,293	162,519,988
delivery	135,925,671	22.63%	1,610	84,425	136,030,138
slev	124,808,297	20.78%	1,519	82,164	124,809,220
ostat	17,611,177	2.93%	1,547	11,384	17,612,111
Geometric mean				24,295	

Users: 128				TPM: 433,597	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	152,308,928	25.33%	12,397	12,285	153,130,114
payment	160,857,017	26.75%	12,452	12,918	161,683,384
delivery	144,476,820	24.03%	1,286	112,345	144,599,618
slev	126,987,195	21.12%	1,245	101,997	126,987,968
ostat	16,418,069	2.73%	1,236	13,283	16,418,867
Geometric mean				29,964	

Intel VROC with Intel SSD D5-P4320 Series

Users: 64				TPM: 285,799	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	151,935,516	25.27%	14,600	10,406	152,876,902
payment	159,118,612	26.47%	14,728	10,803	160,108,564
delivery	139,856,781	23.27%	1,480	94,497	139,951,951
slev	131,750,930	21.92%	1,498	87,951	131,751,865
ostat	18,124,092	3.01%	1,486	12,196	18,125,001
Geometric mean				25,783	

Users: 96				TPM: 348,670	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	152,779,660	25.40%	12,635	12,091	153,610,873
payment	161,270,612	26.81%	12,554	12,846	162,099,321
delivery	140,127,339	23.29%	1,240	113,005	140,246,627
slev	129,590,808	21.54%	1,248	103,838	129,591,557
ostat	17,480,351	2.91%	1,259	13,884	17,481,101
Geometric mean				30,244	

Users: 128				TPM: 374,439	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	146,030,678	24.25%	10,540	13,854	146,767,585
payment	166,223,188	27.60%	10,632	15,634	166,929,058
delivery	143,849,293	23.89%	1,083	132,824	143,920,964
slev	128,027,241	21.26%	1,009	126,885	128,027,910
ostat	17,783,458	2.95%	1,040	17,099	17,784,085
Geometric mean				36,229	

Broadcom with Intel SSD D5-P4320 Series

Users: 64				TPM: 285,060	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	159,575,922	26.56%	14,867	10,733	160,579,242
payment	157,441,545	26.20%	15,042	10,466	158,421,808
delivery	136,571,357	22.73%	1,516	90,086	136,669,880
slev	130,757,279	21.76%	1,482	88,230	130,758,181
ostat	16,150,248	2.69%	1,445	11,176	16,151,124
Geometric mean				25,108	

Users: 96				TPM: 337,045	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	152,571,851	25.36%	12,152	12,555	153,352,216
payment	158,909,844	26.41%	12,379	12,837	159,744,355
delivery	137,890,178	22.92%	1,254	109,960	137,970,827
slev	135,361,061	22.50%	1,247	108,549	135,361,787
ostat	16,617,543	2.76%	1,192	13,940	16,618,223
Geometric mean				30,597	

Users: 128				TPM: 352,647	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	145,847,305	24.22%	9,997	14,589	146,543,523
payment	155,039,974	25.75%	10,078	15,384	155,705,572
delivery	145,611,993	24.18%	1,063	136,982	145,682,031
slev	138,768,492	23.05%	1,008	137,667	138,769,078
ostat	16,600,428	2.76%	1,014	16,371	16,601,031
Geometric mean				36,994	

Intel VROC with Intel SSD D3-S4510 Series

Users: 64				TPM: 116,589	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	146,091,056	24.22%	6,494	22,496	146,541,852
payment	146,091,056	24.22%	6,494	22,496	146,541,852
delivery	146,091,056	24.22%	6,494	22,496	146,541,852
slev	146,091,056	24.22%	6,494	22,496	146,541,852
ostat	146,091,056	24.22%	6,494	22,496	146,541,852
Geometric mean				57,887	

Users: 96				TPM: 96,659	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	135,736,602	22.46%	3,925	34,582	136,028,295
payment	154,843,691	25.63%	4,027	38,451	155,105,211
delivery	145,034,554	24.00%	382	379,671	145,059,270
slev	152,205,144	25.19%	389	391,272	152,205,392
ostat	16,314,240	2.70%	421	38,751	16,314,514
Geometric mean				94,795	

Users: 128				TPM: 100,890	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	130,152,121	21.48%	3,091	42,106	130,394,872
payment	138,433,501	22.85%	3,167	43,711	138,644,595
delivery	167,758,467	27.69%	341	491,960	167,781,039
slev	150,988,676	24.92%	331	456,159	150,988,877
ostat	18,432,186	3.04%	334	55,186	18,432,402
Geometric mean				117,913	

Broadcom with Intel SSD D3-S4510 Series

Users: 64				TPM: 103,950	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	137,275,714	22.74%	5,490	25,004	137,662,972
payment	168,485,088	27.91%	5,606	30,054	168,843,696
delivery	142,728,033	23.64%	566	252,169	142,764,054
slev	138,138,336	22.88%	591	233,736	138,138,714
ostat	17,004,649	2.82%	531	32,023	17,004,955
Geometric mean				67,664	

Users: 96				TPM: 93,945	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	131,898,815	21.82%	3,489	37,804	132,161,431
payment	172,031,750	28.45%	3,546	48,514	172,261,765
delivery	141,856,171	23.46%	345	411,177	141,878,406
slev	137,405,346	22.73%	359	382,744	137,405,576
ostat	21,297,631	3.52%	397	53,646	21,297,874
Geometric mean				109,138	

Users: 128				TPM: 98,565	
PROCNAME	EXCLUSIVETOT	%	CALLNUM	AVGPERCALL	CUMULTOT
neword	129,783,370	21.46%	2,900	44,752	129,972,022
payment	175,737,956	29.06%	2,864	61,361	175,962,163
delivery	145,380,050	24.04%	294	494,489	145,399,245
slev	134,230,837	22.19%	283	474,313	134,231,028
ostat	19,627,609	3.25%	283	69,355	19,627,789
Geometric mean				134,897	

Performance results in transactions per minute

Average values

The tables below present the mean value from the three runs we conducted for each user count for each configuration. We used these numbers for the graphs in the body of the report.

User count	Intel VROC w/ Intel SSD DC P4510 Series	Broadcom w/ Intel SSD DC P4510 Series	Intel VROC w/ Intel SSD D5-P4320 Series	Broadcom w/ Intel SSD D5-P4320 Series	Intel VROC w/ Intel SSD D3-S4510 Series	Broadcom w/ Intel SSD D3-S4510 Series
8	98,772	89,197	59,299	57,697	45,352	37,685
16	176,142	163,478	108,217	110,877	72,956	61,272
24	241,832	220,432	147,929	151,133	90,905	75,923
32	285,896	253,845	184,685	188,833	102,488	84,791
40	333,907	286,079	214,349	216,149	108,557	88,595
48	351,295	323,895	238,147	238,591	110,321	91,624
56	380,479	350,917	258,487	249,438	112,907	92,443
64	401,153	365,355	281,270	282,716	113,079	93,474
72	419,603	367,588	294,185	287,705	109,037	90,087
80	425,584	387,498	310,829	306,699	96,358	92,437
88	429,040	410,680	319,644	318,883	102,389	90,793
96	446,806	425,735	342,583	333,573	96,244	92,105
104	449,622	428,256	348,148	332,348	101,097	92,009
112	453,473	422,215	353,966	342,441	95,463	93,973
120	453,013	430,712	365,945	347,873	99,647	98,302
128	459,036	428,751	365,401	343,184	99,187	93,383
144	467,214	433,270	373,894	325,294	100,682	92,106
160	467,385	440,141	390,343	317,508	100,940	98,583
176	459,196	445,834	391,935	323,627	99,894	96,736
192	459,721	443,604	369,128	325,424	100,007	99,040
208	461,339	428,176	364,152	337,489	97,761	96,938
224	465,401	441,351	362,051	332,325	96,704	96,599
240	468,115	463,366	359,508	324,510	98,217	94,912
256	471,673	468,811	346,946	324,170	93,577	95,463
Median	448,214	423,975	344,765	321,255	99,771	92,440
Average	396,904	371,633	297,960	279,937	97,657	89,136
Geomean	377,083	351,298	275,725	261,922	96,359	87,673
Harmean	344,236	318,361	239,839	231,348	94,500	85,460

Intel VROC with Intel SSD DC P4510 Series

Users	Run 1	Run 2	Run 3	Average
8	98,402	98,650	99,263	98,772
16	180,212	174,577	173,636	176,142
24	239,169	241,987	244,340	241,832
32	292,877	281,249	283,562	285,896
40	330,211	334,920	336,591	333,907
48	335,849	353,767	364,268	351,295
56	368,703	382,531	390,203	380,479
64	393,053	401,525	408,882	401,153
72	411,676	422,869	424,265	419,603
80	423,057	421,355	432,340	425,584
88	431,494	429,007	426,619	429,040
96	436,496	447,783	456,140	446,806
104	447,657	442,322	458,886	449,622
112	454,828	446,781	458,811	453,473
120	457,315	446,251	455,474	453,013
128	458,305	459,580	459,223	459,036
144	466,291	468,416	466,936	467,214
160	468,375	463,020	470,761	467,385
176	461,876	450,094	465,618	459,196
192	467,924	449,733	461,506	459,721
208	469,930	448,051	466,035	461,339
224	478,165	455,202	462,836	465,401
240	478,277	458,389	467,678	468,115
256	475,342	464,907	474,769	471,673
Median	442,077	444,287	455,807	
Average	396,895	393,457	400,360	
Geomean	376,869	374,103	380,169	
Harmean	344,000	341,932	346,587	

Broadcom with Intel SSD DC P4510 Series

Users	Run 1	Run 2	Run 3	Average
8	88,150	90,870	88,572	89,197
16	162,656	165,366	162,412	163,478
24	217,382	222,736	221,177	220,432
32	249,320	260,696	251,519	253,845
40	280,727	290,148	287,361	286,079
48	323,586	326,703	321,397	323,895
56	350,175	351,751	350,826	350,917
64	364,291	367,709	364,064	365,355
72	358,560	373,571	370,633	367,588
80	390,328	382,611	389,555	387,498
88	410,299	407,012	414,728	410,680
96	422,452	426,207	428,547	425,735
104	422,820	436,266	425,683	428,256
112	414,593	431,308	420,745	422,215
120	420,250	436,082	435,805	430,712
128	423,135	429,522	433,597	428,751
144	423,471	447,289	429,050	433,270
160	433,014	447,297	440,112	440,141
176	449,374	455,550	432,578	445,834
192	447,429	430,874	452,509	443,604
208	429,971	438,572	415,985	428,176
224	458,775	441,715	423,562	441,351
240	458,747	467,925	463,425	463,366
256	453,803	481,744	470,885	468,811
Median	417,422	427,865	418,365	
Average	368,888	375,397	370,614	
Geomean	348,513	354,984	350,294	
Harmean	315,574	322,073	317,290	

Intel VROC with Intel SSD D5-P4320 Series

Users	Run 1	Run 2	Run 3	Average
8	58,570	59,997	59,329	59,299
16	107,266	108,509	108,875	108,217
24	145,011	150,578	148,199	147,929
32	183,044	188,121	182,889	184,685
40	214,183	218,104	210,760	214,349
48	240,969	243,185	230,286	238,147
56	259,781	256,414	259,265	258,487
64	285,799	283,137	274,874	281,270
72	298,209	305,229	279,117	294,185
80	322,222	310,809	299,456	310,829
88	313,742	323,908	321,281	319,644
96	348,670	342,743	336,337	342,583
104	354,875	352,998	336,571	348,148
112	352,727	355,057	354,113	353,966
120	369,921	362,347	365,567	365,945
128	374,439	373,100	348,665	365,401
144	381,176	369,480	371,026	373,894
160	388,197	389,271	393,561	390,343
176	397,632	391,081	387,091	391,935
192	370,305	361,805	375,275	369,128
208	364,877	363,029	364,551	364,152
224	365,046	362,159	358,948	362,051
240	363,166	348,337	367,020	359,508
256	340,600	349,622	350,615	346,946
Median	344,635	345,540	336,454	
Average	300,018	298,709	295,153	
Geomean	277,046	276,904	273,118	
Harmean	239,889	241,448	238,014	

Broadcom with Intel SSD D5-P4320 Series

Users	Run 1	Run 2	Run 3	Average
8	57,613	58,027	57,451	57,697
16	111,857	110,368	110,407	110,877
24	151,305	152,291	149,803	151,133
32	190,439	188,227	187,833	188,833
40	218,352	212,801	217,295	216,149
48	244,383	242,775	228,614	238,591
56	248,274	252,170	247,869	249,438
64	283,664	285,060	279,423	282,716
72	298,912	274,035	290,168	287,705
80	310,554	295,794	313,750	306,699
88	327,104	323,700	305,844	318,883
96	337,045	334,054	329,620	333,573
104	338,797	320,157	338,089	332,348
112	352,063	348,213	327,048	342,441
120	364,001	338,540	341,078	347,873
128	352,647	330,619	346,286	343,184
144	344,271	336,804	294,808	325,294
160	335,382	301,140	316,003	317,508
176	335,629	311,556	323,697	323,627
192	332,133	327,735	316,404	325,424
208	339,808	337,786	334,873	337,489
224	341,709	322,133	333,133	332,325
240	329,764	328,093	315,672	324,510
256	315,500	331,643	325,367	324,170
Median	328,434	315,857	314,711	
Average	285,884	277,655	276,272	
Geomean	266,853	260,043	258,663	
Harmean	234,616	230,260	228,876	

Intel VROC with Intel SSD D3-S4510 Series

Users	Run 1	Run 2	Run 3	Average
8	45,459	45,403	45,193	45,352
16	73,186	72,749	72,932	72,956
24	91,137	90,438	91,140	90,905
32	102,505	103,666	101,292	102,488
40	107,872	109,225	108,575	108,557
48	110,308	113,026	107,630	110,321
56	114,737	116,309	107,675	112,907
64	113,179	116,589	109,470	113,079
72	106,812	109,434	110,866	109,037
80	96,399	98,981	93,694	96,358
88	100,121	103,744	103,303	102,389
96	96,659	96,547	95,526	96,244
104	101,799	102,210	99,281	101,097
112	98,587	93,068	94,734	95,463
120	100,380	96,511	102,050	99,647
128	100,890	99,431	97,241	99,187
144	100,321	102,393	99,332	100,682
160	100,277	102,872	99,670	100,940
176	98,370	100,813	100,499	99,894
192	99,336	100,214	100,471	100,007
208	96,452	98,957	97,874	97,761
224	95,947	93,442	100,722	96,704
240	98,431	100,174	96,047	98,217
256	93,397	95,256	92,077	93,577
Median	99,729	100,194	99,501	
Average	97,607	98,394	96,971	
Geomean	96,322	97,006	95,714	
Harmean	94,482	95,046	93,904	

Broadcom with Intel SSD D3-S4510 Series

Users	Run 1	Run 2	Run 3	Average
8	37,435	38,419	37,201	37,685
16	60,721	61,372	61,724	61,272
24	75,158	75,667	76,945	75,923
32	84,700	82,744	86,928	84,791
40	89,591	83,417	92,776	88,595
48	91,829	85,735	97,308	91,624
56	87,526	88,742	101,061	92,443
64	89,404	87,067	103,950	93,474
72	82,609	81,712	105,940	90,087
80	88,508	87,440	101,364	92,437
88	89,234	91,167	91,978	90,793
96	90,883	91,488	93,945	92,105
104	92,710	93,245	90,073	92,009
112	92,994	94,154	94,770	93,973
120	95,523	98,548	100,834	98,302
128	98,565	91,225	90,360	93,383
144	88,731	95,312	92,276	92,106
160	100,849	102,598	92,302	98,583
176	99,232	99,741	91,236	96,736
192	103,957	105,824	87,339	99,040
208	103,015	96,728	91,070	96,938
224	99,674	96,330	93,794	96,599
240	93,213	96,485	95,038	94,912
256	95,988	94,657	95,743	95,463
Median	91,356	91,357	92,539	
Average	88,835	88,326	90,248	
Geomean	87,264	86,826	88,647	
Harmean	84,958	84,680	86,237	

System configuration information

The table below presents detailed information on the systems we tested.

Server configuration information	Intel® Xeon® processor-based server
BIOS name and version	00.01.0015
Non-default BIOS settings	(VROC only) Riser2, Slot1 Volume Management Device(CPU2, IOU2): Enabled VMD Port 2A: Enabled VMD Port 2B: Enabled VMD Port 2C: Enabled VMD Port 2D: Enabled Riser_Slot_2 Bifurcation: x4x4x4x4
Operating system name and version/build number	RHEL 7.5 kernel-3.10.0-862.14.4.el7.x86_64 microcode_ctl-2.1-29.16.el7_5.x86_64
Date of last OS updates/patches applied	1/18/19
Power management policy	CPU Power and Performance Policy: Performance Workload Configuration: I/O Sensitive Package C-State: C6(non Retention) state C1E: Disabled Set Fan Profile: Performance
Processor	
Number of processors	2
Vendor and model	Intel Xeon 8164 Platinum
Core count (per processor)	26
Core frequency (GHz)	2.00
Stepping	H0
Memory module(s)	
Total memory in system (GB)	192
Number of memory modules	12
Vendor and model	Hynix® HMA42GR7AFR4N-TF
Size (GB)	16
Type	DDR4-2133
Speed (MHz)	2,133
Speed running in the server (MHz)	2,133
Storage controller #1 (VROC configuration)	
Vendor and model	Intel 4-Port PCIe Gen3 x16 Retimer AIC (AXXP3RTX16040)
Cache size (GB)	N/A
Firmware version	N/A
Driver version	mdadm-4.0-13.el7.x86_64

Server configuration information	Intel® Xeon® processor-based server
Storage controller #1 (Broadcom configuration)	
Vendor and model	Broadcom MegaRAID 9460-16i
Cache size (GB)	4
Firmware version	50.7.0-1614
Driver version	kmod-megaraid_sas-07.707.03.00_el7.5-1.x86_64
Storage configuration #1 (SATA)	
Number of drives	4
Drive vendor and model	Intel SSD D3-S4510 SSDSC2KB038T8
Drive size (TB)	3.84
Drive information (speed, interface, type)	6Gbps SATA SSD (TLC)
Firmware version	XCV10100
Storage configuration #2 (QLC NVMe)	
Number of drives	4
Drive vendor and model	Intel SSD D5-P4320 SSDPE2NV076T8
Drive size (TB)	7.68
Drive information (speed, interface, type)	NVMe SSD (QLC)
Firmware version	3DV10110
Storage configuration #3 (NVMe)	
Number of drives	4
Drive vendor and model	Intel SSD DC P4510 SSDPE2KX080T8
Drive size (TB)	8.0
Drive information (speed, interface, type)	NVMe SSD (TLC)
Firmware version	VDV10131
Network adapter	
Vendor and model	Intel X722
Number and type of ports	2x 10Gb Ethernet
Driver version	i40e 2.1.14-k

How we tested

Configuring the operating system

We installed Red Hat Enterprise Linux on both servers and the client machine, and executed the following steps.

```
# Disable SELinux and Firewall
sed -i 's/SELINUX=.*/SELINUX=disabled/' /etc/selinux/config
systemctl disable firewalld

# Register with RHN (prompts for credentials), set release, and clear yum cache
subscription-manager register --auto-attach
subscription-manager release --set=7.5
yum clean all
rm -rf /var/cache/yum

# Install updates and reboot
yum update -y
reboot

# Install the EPEL repository and enable the optional and extras repositories
yum install -y https://dl.fedoraproject.org/pub/epel/epel-release-latest-7.noarch.rpm
subscription-manager repos --enable "rhel-*-optional-rpms" --enable "rhel-*-extras-rpms"
yum update -y

# Install extra packages
yum install -y xfsprogs tuned numactl wget vim openssh-clients nfs-utils man unzip ipmitool OpenIPMI
sysstat nmon lz4 device-mapper-multipath hdparm pciutils yum-utils
```

Configuring the VROC storage controller solution

On the server running VROC, install mdadm, and reboot the server.

```
yum install -y mdadm

# Reboot
reboot

# Enable blk_mq for SCSI devices
sed -i 's/rhgb quiet/scsi_mod.use_blk_mq=y/' /etc/default/grub
grub2-mkconfig -o /boot/efi/EFI/redhat/grub.cfg
reboot

# Create NVMe VROC RAID volumes (assumes drives nvme{0,1,2,3}n1)
mdadm --zero-superblock /dev/nvme{0,1,2,3}n1
mdadm -C /dev/md/imsm0 -e imsm -n4 /dev/nvme{0,1,2,3}n1
mdadm -C /dev/md0 /dev/md/imsm0 -n4 -l5 -N db_data -z 6979321856 -c 64
mdadm -C /dev/md1 /dev/md/imsm0 -n4 -l10 -N db_log -c 128


# Create SATA VROC RAID volumes (assumes drives sd{b,c,d,e})
mdadm --zero-superblock /dev/sd{b,c,d,e}
mdadm -C /dev/md/imsm0 -e imsm -n4 /dev/sd{b,c,d,e}
mdadm -C /dev/md0 /dev/md/imsm0 -n4 -l5 -N db_data -z 3489660928 -c 64
mdadm -C /dev/md1 /dev/md/imsm0 -n4 -l10 -N db_log -c 128


# Format volumes and increase sync speed
mkfs.xfs -L db_data /dev/md0
mkfs.xfs -L db_log  /dev/md1
echo 4 > /sys/block/md0/md/group_thread_cnt
echo 5000000 > /sys/block/md0/md/sync_speed_max
echo 5000000 > /sys/block/md1/md/sync_speed_max
```

```
# Tune VROC (NVMe)
echo 4 > /sys/block/md125/md/group_thread_cnt

for DRIVE in nvme{0,1,2,3}n1 ; do
 echo $DRIVE
 echo "0" > /sys/block/${DRIVE}/queue/rotational
 echo "none" > /sys/block/${DRIVE}/queue/scheduler
 echo "2048" > /sys/block/${DRIVE}/queue/nr_requests
 echo "1" > /sys/block/${DRIVE}/queue/nomerges
 echo "0" > /sys/block/${DRIVE}/queue/add_random
 echo "2" > /sys/block/${DRIVE}/queue/rq_affinity
 blockdev --setra 0 /dev/${DRIVE}
done

# Tune VROC (SATA)
echo 4 > /sys/block/md125/md/group_thread_cnt

for DRIVE in sd{b,c,d,e} ; do
 echo $DRIVE
 echo "0" > /sys/block/${DRIVE}/queue/rotational
 echo "none" > /sys/block/${DRIVE}/queue/scheduler
 echo "2048" > /sys/block/${DRIVE}/queue/nr_requests
 echo "1" > /sys/block/${DRIVE}/queue/nomerges
 echo "0" > /sys/block/${DRIVE}/queue/add_random
 echo "2" > /sys/block/${DRIVE}/queue/rq_affinity
 blockdev --setra 0 /dev/${DRIVE}
done
```

Configuring the Broadcom controller solution

On the server running Broadcom, install drivers and the MegaRAID tool, and reboot the server.

```
# Install BCOM Drivers
mkdir ~/megaraid
cd ~/megaraid
wget https://docs.broadcom.com/docs/MR_LINUX_DRIVER_7.7-07.707.03.00-1.tgz
tar -xf MR_LINUX_DRIVER_7.7-07.707.03.00-1.tgz
cd rhel7_oel7_centos7/rpms-1/
yum install -y kmod-megaraid_sas-07.707.03.00_el7.5-1.x86_64.rpm

# Install BCOM MegaRAID tool
cd ~/megaraid
wget https://docs.broadcom.com/docs/007.0709.0000.0000_Unified_StorCLI.zip
unzip 007.0709.0000.0000_Unified_StorCLI.zip
unzip Unified_storcli_all_os.zip
cd Unified_storcli_all_os/Linux/
yum install -y storcli-007.0709.0000.0000-1.noarch.rpm
ln -s /opt/MegaRAID/storcli/storcli64 /usr/local/sbin/storcli

# Reboot
reboot

# Enable blk_mq for SCSI devices
sed -i 's/rhgb quiet/scsi_mod.use_blk_mq=y/' /etc/default/grub
grub2-mkconfig -o /boot/efi/EFI/redhat/grub.cfg
reboot

# Create BCOM RAID volumes
#NVMe
storcli /c0 add vd type=raid5Size=19.5TB name=db_data drives=74:4,6,8,10 pdcache=off wt nora direct
Strip=64 force
storcli /c0 add vd type=raid5 name=db_log drives=74:4,6,8,10 pdcache=off awb nora direct Strip=128
```

```
force
#SATA
storcli /c0 add vd type=raid5Size=10TB name=db_data drives=134:0,1,2,3 pdcache=off wt nora direct
Strip=64 force
storcli /c0 add vd type=raid5 name=db_log drives=134:0,1,2,3 pdcache=off awb nora direct Strip=128
force

# Force full initialization to start
storcli /c0/v0 start init full force
storcli /c0/v1 start init full force

# Check initialization status
storcli /c0/v0 show all | grep Operations
storcli /c0/v1 show all | grep Operations

# Wait for initialization to finish before running these commands (requires storcli to identify
volume)
# Format volumes
mkfs.xfs -L db_data $(storcli /c0/v0 show all | awk '/OS Drive Name/{print $5}')
mkfs.xfs -L db_log $(storcli /c0/v1 show all | awk '/OS Drive Name/{print $5}')


# Tune Broadcom (requires storcli)
for VD in 0 1 ; do
 DRIVE=$(storcli /c0/v${VD} show all | awk '/OS Drive Name/{print substr($5,6)}')
 echo "0" > /sys/block/${DRIVE}/queue/rotational
 echo "none" > /sys/block/${DRIVE}/queue/scheduler
 echo "2048" > /sys/block/${DRIVE}/queue/nr_requests
 echo "256" > /sys/block/${DRIVE}/device/queue_depth
 echo "0" > /sys/block/${DRIVE}/queue/nomerges
 echo "0" > /sys/block/${DRIVE}/queue/add_random
 echo "0" > /sys/block/${DRIVE}/queue/rq_affinity
 tail /sys/block/${DRIVE}/queue/*
 tail /sys/block/${DRIVE}/device/queue_*
done
blockdev --report
```

Configuring the file system mount and OS tunings for MySQL

On both systems, mount the filesystems and tune for MySQL as specified below.

```
# /etc/fstab mount config
LABEL=db_data  /mnt/db_data  xfs  defaults,noatime,nodiratime,x-systemd.device-timeout=10,nofail,_
netdev  0 0
LABEL=db_log  /mnt/db_log  xfs  defaults,noatime,nodiratime,x-systemd.device-timeout=10,nofail,_
netdev  0 0


# Tune OS for MySQL
# Add the following to /etc/sysctl.conf
vm.swappiness = 0
kernel.sem = 250 32000 100 128
fs.file-max = 6815744
net.ipv4.ip_local_port_range = 9000 65500
net.core.rmem_default = 262144
net.core.rmem_max = 4194304
net.core.wmem_default = 262144
net.core.wmem_max = 1048576
fs.aio-max-nr = 1048576
vm.nr_hugepages = 50000


# Run these commands to configure MySQL permission access to huge pages
sed -i '/vm.hugetlb_shm_group/d' /etc/sysctl.conf
echo vm.hugetlb_shm_group=$(id -g mysql) >> /etc/sysctl.conf

reboot
```

Installing and configuring MySQL

On both systems, mount the filesystems and tune for MySQL as specified below.

```
# Install MySQL (Ver 8.0.13 for Linux on x86_64)
yum install -y https://dev.mysql.com/get/mysql80-community-release-el7-1.noarch.rpm
yum install -y mysql-community-server

# Configure MySQL
systemctl stop mysqld
systemctl disable mysqld
cp -p /etc/my.cnf{,.orig}
# Refer to attached my.cnf config file

# Cleanup, create, and configure data and log directory
rm -rf /mnt/db_data/mysql_data
mkdir /mnt/db_data/mysql_data
chown -R mysql:mysql /mnt/db_data/mysql_data

rm -rf /mnt/db_log/mysql_innodb_log
mkdir /mnt/db_log/mysql_innodb_log
chown -R mysql:mysql /mnt/db_log/mysql_innodb_log

# Initialize MySQL (this will create the log files)
systemctl start mysqld

# Set root password
grep password /var/log/mysqld.log
# Example output: A temporary password is generated for root@localhost: <uAdom9lyo/s
# Use the temporary password to log in
mysql -u root -p

ALTER USER 'root'@'localhost' IDENTIFIED BY 'mysql';
use mysql;
update user set host='%' where user='root' and host='localhost';
flush privileges;
grant all on *.* to root@'%';
quit
```

Changing the my.cnf file

After MySQL is installed and initial configuration is completed, modify the my.cnf file with the parameters below.

```
datadir=/mnt/db_data/mysql_data
socket=/var/lib/mysql/mysql.sock

log-error=/var/log/mysqld.log
pid-file=/var/run/mysqld/mysqld.pid

large-pages
skip-log-bin
default_authentication_plugin=mysql_native_password
port=3306
# general
max_connections=4000
table_open_cache=8000
table_open_cache_instances=16
back_log=1500
default_password_lifetime=0
ssl=0
performance_schema=OFF
max_prepared_stmt_count=128000
```

```
skip_log_bin=1
character_set_server=latin1
collation_server=latin1_swedish_ci
transaction_isolation=REPEATABLE-READ
# files
innodb_file_per_table
innodb_log_file_size=1024M
innodb_log_files_in_group=32
innodb_open_files=4000
# buffers
innodb_buffer_pool_size=96000M
innodb_buffer_pool_instances=16
innodb_log_buffer_size=64M
# tune
innodb_doublewrite=0
innodb_thread_concurrency=0
innodb_flush_log_at_trx_commit=2
innodb_max_dirty_pages_pct=90
innodb_max_dirty_pages_pct_lwm=10
join_buffer_size=32K
sort_buffer_size=32K
innodb_use_native_aio=1
innodb_stats_persistent=1
innodb_spin_wait_delay=6
innodb_max_purge_lag_delay=300000
innodb_max_purge_lag=0
innodb_flush_method=O_DIRECT_NO_FSYNC
innodb_checksum_algorithm=none
innodb_io_capacity=40000
innodb_io_capacity_max=200000
innodb_lru_scan_depth=9000
innodb_change_buffering=none
innodb_read_only=0
innodb_page_cleaners=4
innodb_undo_log_truncate=off
# perf special
innodb_adaptive_flushing=1
innodb_flush_neighbors=0
innodb_read_io_threads=16
innodb_write_io_threads=16
innodb_purge_threads=4
innodb_adaptive_hash_index=0
# monitoring
innodb_monitor_enable='%'

innodb_log_group_home_dir=/mnt/db_log/mysql_innodb_log
validate_password.policy=LOW
validate_password.length=5
```

Installing and configuring the HammerDB client, dataset, and the test harness

On the client system, Install HammerDB and build the database by using these commands.

```
# Tune OS for MySQL
# Add the following to /etc/sysctl.conf
vm.swappiness = 0
kernel.sem = 250 32000 100 128
fs.file-max = 6815744
net.ipv4.ip_local_port_range = 9000 65500
net.core.rmem_default = 262144
net.core.rmem_max = 4194304
net.core.wmem_default = 262144
net.core.wmem_max = 1048576
fs.aio-max-nr = 1048576

yum install -y https://dev.mysql.com/get/mysql80-community-release-el7-1.noarch.rpm

# Install HammerDB client
yum install -y tcl
cd ~
wget https://downloads.sourceforge.net/project/hammerdb/HammerDB/HammerDB-3.1/HammerDB-3.1-Linux.tar.
gz
tar -xf HammerDB-3.1-Linux.tar.gz
cd HammerDB-3.1/
yumdownloader --enablerepo=mysql57-community mysql-community-libs-5.7.24-1.el7.x86_64
rm -rf ./usr
rpm2cpio mysql-community-libs-5.7.24-1.el7.x86_64.rpm | cpio -idmv ./usr/lib64/mysql/libmysqlclient.
so.20*
mv -f ./usr/lib64/mysql/libmysqlclient.so.20* .
rm -rf ./usr


# Create file innodbbuild.tcl
#!/bin/tclsh
puts "SETTING CONFIGURATION"
dbset db mysql
diset connection mysql_host test01-direct
diset connection mysql_port 3306
diset tpcc mysql_user root
diset tpcc mysql_pass mysql
diset tpcc mysql_count_ware 100000
diset tpcc mysql_partition true
diset tpcc mysql_num_vu 100
diset tpcc mysql_storage_engine innodb
print dict
buildschema
exit

# Run the build command:
./hammerdbcli < innodbbuild.tcl
```

Backing up the dataset

Run the procedures below to backup the dataset and restore it between runs.

```
# Create backup (100K)
systemctl stop mysqld
sync
cd /mnt
umount /mnt/net_backup
mount -v /mnt/net_backup
rm -rf /mnt/net_backup/mysql_tpcc_100k
mkdir -p /mnt/net_backup/mysql_tpcc_100k/
echo 1 > /proc/sys/vm/drop_caches
echo 0 > /proc/sys/vm/nr_hugepages
time cp -av db_log /mnt/net_backup/mysql_tpcc_100k/
time cp -av db_data /mnt/net_backup/mysql_tpcc_100k/
sync


# Restore backup (100K)
systemctl stop mysqld
sync
cd /mnt
rm -rf db_data/mysql_data
rm -rf db_log/mysql_innodb_log/*
umount /mnt/net_backup
mount -v /mnt/net_backup
echo 1 > /proc/sys/vm/drop_caches
echo 0 > /proc/sys/vm/nr_hugepages
time cp -av /mnt/net_backup/mysql_tpcc_100k/db_data/mysql_data db_data/
sync
systemctl start mysqld
systemctl stop mysqld
date
reboot
```

Running the tests

We used the following sequence in our testing:

1. Restore the test database.
2. Reboot the server.
3. Allow the server to sit idle for two hours.
4. Complete a warm-up run consisting of 100 threads executing 100,000 transactions (10,000,000 total iterations).

```
Run ./hammerdbcli < client_step-up.tcl on HammerDB. See below for HammerDB commands to execute the
tests
#!/bin/tclsh
proc runtimer { seconds } {
set x 0
set timerstop 0
while {!$timerstop} {
incr x
after 1000
  if { ![ expr {$x % 60} ] } {
          set y [ expr $x / 60 ]
          puts "Timer: $y minutes elapsed"
  }
update
```

```
if {  [ vucomplete ] || $x eq $seconds } { set timerstop 1 }
    }
return
}
puts "SETTING CONFIGURATION"
dbset db mysql
diset connection mysql_host test01
diset connection mysql_port 3306
diset tpcc mysql_driver timed
diset tpcc mysql_total_iterations 1000000
diset tpcc mysql_rampup 5
diset tpcc mysql_duration 5
diset tpcc mysql_allwarehouse true
diset tpcc mysql_timeprofile true
vuset logtotemp 1
loadscript
puts "SEQUENCE STARTED"
foreach z { 8 16 24 32 40 48 56 64 72 80 88 96 104 112 120 128 144 160 176 192 208 224 240 256 } {
puts "$z VU TEST"
vuset vu $z
vucreate
vurun
runtimer 900
vudestroy
after 5000
        }
puts "TEST SEQUENCE COMPLETE"
exit
```

5. Measure and report TPM for each user count noted above.

Read the report at http://facts.pt/yjr4kn5 ▶

This project was commissioned by Intel.