

GT ZFS

Comment monitorer/superviser mon zfs, et les divers caches ?
Une présentation orientée Linux

Sylvain MAURIN

29 Novembre 2023



Connaitre l'espace

Connaitre l'espace

- ▶ Taux de foisonnement
- ▶ Compression
- ▶ Dé-duplication
- ▶ Distribution des blocs



Connaitre l'espace

Connaitre l'espace

- ▶ Taux de foisonnement
- ▶ Compression
- ▶ Dé-duplication
- ▶ Distribution des blocs



Connaitre l'espace

Connaitre l'espace

- ▶ Taux de foisonnement
- ▶ Compression
- ▶ Dé-duplication
- ▶ Distribution des blocs



Connaitre l'espace

Connaitre l'espace

- ▶ Taux de foisonnement
- ▶ Compression
- ▶ Dé-duplication
- ▶ Distribution des blocs



Connaître la latence

Connaître la latence

- ▶ Taille d'une transaction
 - ▶ Lecture
 - ▶ Écriture



Connaître la latence

Connaître la latence

- ▶ Taille d'une transaction
- ▶ Lecture
- ▶ Écriture

Connaître la latence

Connaître la latence

- ▶ Taille d'une transaction
- ▶ Lecture
- ▶ Écriture

- ▶ Organisation du RAID
- ▶ Bus d'attache
- ▶ Médias

- ▶ Organisation du RAID
- ▶ Bus d'attchement
- ▶ Médias

- ▶ Organisation du RAID
- ▶ Bus d'attachement
- ▶ Médias



Retrouver ses sources

Les sources informations

- ▶ RTFM
- ▶ Solaris Porting Layer
- ▶ Les sources

```
man 4 spl ; man 4 zfs  
ls /proc/spl/kstat/zfs
```



Boite à outils ZFS

En CLI

- ▶ zfs get all, compressratio,... etc.
- ▶ zpool history, iostat,... etc.
- ▶ arcstat
- ▶ arc summary
- ▶ zdb



Boite à outils ZFS

En CLI

- ▶ `zfs get all, compressratio,...` etc.
- ▶ `zpool history, iostat,...` etc.
- ▶ `arcstat`
- ▶ `arc summary`
- ▶ `zdb`



Boite à outils ZFS

En CLI

- ▶ zfs get all, compressratio,... etc.
- ▶ zpool history, iostat,... etc.
- ▶ arcstat
- ▶ arc summary
- ▶ zdb



Boite à outils ZFS

En CLI

- ▶ `zfs get all, compressratio,... etc.`
- ▶ `zpool history, iostat,... etc.`
- ▶ `arcstat`
- ▶ `arc summary`
- ▶ `zdb`



Boite à outils ZFS

En CLI

- ▶ `zfs get all, compressratio,...` etc.
- ▶ `zpool history, iostat,...` etc.
- ▶ `arcstat`
- ▶ `arc summary`
- ▶ `zdb`



Outils de supervision

Les exporteurs de données classiques

- ▶ Prometheus (exemple:
<https://github.com/matusnovak/prometheus-zfs>)
- ▶ snmpd



Outils de supervision

Les exporteurs de données classiques

- ▶ Prometheus (exemple:
<https://github.com/matusnovak/prometheus-zfs>)
- ▶ snmpd



Effet démo

Essayons la connection SSH...



```
Histogramme de la distribution des blocs  
# for ZPOOL in `zpool list -Ho name` ; do zdb -bbb \${ZPOOL} ; done
```



```
exemple snmpd :
#Zpool Usage
extend .1.3.6.1.4.1.2021.88 zpool_name /bin/bash -c "/sbin/zpool list -H -o name"
#ARC Stats
extend .1.3.6.1.4.1.2021.89 arc_hits /bin/bash -c "grep -m 1 hits /proc/spl/kstat/zfs/arcstats |awk -F' ' '{print $(NF)}'"
extend .1.3.6.1.4.1.2021.89 arc_misses /bin/bash -c "grep misses /proc/spl/kstat/zfs/arcstats |awk -F' ' '{print $(NF)}'"
#L2ARC Stats
extend .1.3.6.1.4.1.2021.89 arc_l2_hits /bin/bash -c "grep l2_hits /proc/spl/kstat/zfs/arcstats |awk -F' ' '{print $(NF)}'"
extend .1.3.6.1.4.1.2021.89 arc_l2_misses /bin/bash -c "grep l2_misses /proc/spl/kstat/zfs/arcstats |awk -F' ' '{print $(NF)}'"
```



Exemples de lectures sur 2 stockages

root@srv-sto-01:~# zpool iostat -v1 5 2

pool	capacity		operations		bandwidth		total_wait		disk_wait		syncq_wait		asyncq_wait		scrub wait	trim wait
	alloc	free	read	write	read	write	read	write	read	write	read	write	read	write		
cncnfs4	585T	230T	768	1.16K	283M	36.2M	47ms	6ms	12ms	1ms	35us	663ns	1ms	6ms	52ms	-
raidz2-0	41.8T	16.4T	53	65	20.2M	1.81M	48ms	3ms	12ms	1ms	33us	459ns	1ms	1ms	52ms	-
slot_01	-	-	13	16	5.05M	464K	48ms	3ms	12ms	1ms	32us	565ns	1ms	1ms	53ms	-
slot_02	-	-	13	16	5.05M	464K	48ms	3ms	12ms	1ms	33us	435ns	1ms	1ms	52ms	-
slot_31	-	-	13	16	5.05M	464K	48ms	3ms	12ms	1ms	34us	420ns	1ms	1ms	52ms	-
slot_32	-	-	13	16	5.05M	464K	47ms	3ms	12ms	1ms	31us	418ns	1ms	1ms	51ms	-
raidz2-1	41.7T	16.5T	53	69	20.2M	1.87M	48ms	3ms	12ms	1ms	34us	439ns	1ms	1ms	53ms	-
slot_03	-	-	13	17	5.06M	479K	49ms	3ms	12ms	1ms	31us	444ns	1ms	1ms	54ms	-
slot_04	-	-	13	17	5.05M	479K	48ms	3ms	12ms	1ms	34us	422ns	1ms	1ms	53ms	-
slot_33	-	-	13	17	5.05M	479K	48ms	3ms	12ms	1ms	36us	444ns	1ms	1ms	53ms	-
slot_34	-	-	13	17	5.06M	479K	48ms	3ms	12ms	1ms	35us	447ns	1ms	1ms	52ms	-
raidz2-2	41.8T	16.5T	55	69	20.3M	1.87M	48ms	3ms	12ms	1ms	39us	452ns	1ms	1ms	53ms	-
slot_05	-	-	13	17	5.07M	479K	48ms	3ms	12ms	1ms	39us	478ns	1ms	1ms	54ms	-
slot_06	-	-	13	17	5.07M	479K	49ms	3ms	12ms	1ms	42us	451ns	1ms	1ms	54ms	-
slot_35	-	-	13	17	5.07M	479K	47ms	3ms	12ms	1ms	38us	441ns	1ms	1ms	53ms	-
slot_36	-	-	14	17	5.07M	479K	46ms	3ms	11ms	1ms	37us	438ns	1ms	1ms	51ms	-
raidz2-3	41.9T	16.3T	53	57	20.3M	1.69M	44ms	2ms	12ms	1ms	30us	452ns	1ms	1ms	46ms	-
slot_07	-	-	13	14	5.07M	433K	46ms	3ms	12ms	1ms	28us	496ns	1ms	1ms	48ms	-
slot_08	-	-	13	14	5.07M	433K	46ms	3ms	12ms	1ms	34us	436ns	1ms	1ms	49ms	-
slot_37	-	-	13	14	5.07M	433K	45ms	3ms	12ms	1ms	34us	438ns	1ms	1ms	47ms	-
slot_38	-	-	14	15	5.07M	433K	40ms	1ms	11ms	1ms	26us	439ns	1ms	791us	41ms	-
raidz2-4	41.8T	16.4T	54	59	20.2M	1.73M	50ms	3ms	12ms	1ms	34us	462ns	1ms	1ms	55ms	-
slot_09	-	-	13	14	5.05M	444K	50ms	3ms	12ms	1ms	34us	502ns	1ms	1ms	55ms	-
slot_10	-	-	13	14	5.06M	444K	50ms	3ms	12ms	1ms	37us	444ns	1ms	1ms	55ms	-
slot_39	-	-	13	14	5.06M	444K	51ms	3ms	12ms	1ms	35us	452ns	1ms	1ms	56ms	-
slot_40	-	-	13	14	5.05M	444K	50ms	3ms	12ms	1ms	30us	449ns	1ms	1ms	55ms	-
raidz2-5	41.8T	16.4T	54	63	20.2M	1.80M	49ms	3ms	12ms	1ms	34us	451ns	1ms	1ms	54ms	-
slot_11	-	-	13	15	5.06M	461K	48ms	3ms	12ms	1ms	32us	502ns	1ms	1ms	53ms	-
slot_12	-	-	13	15	5.06M	461K	49ms	3ms	12ms	1ms	37us	434ns	1ms	1ms	55ms	-
slot_41	-	-	13	15	5.06M	461K	49ms	3ms	12ms	1ms	37us	433ns	1ms	1ms	55ms	-
slot_42	-	-	13	15	5.06M	461K	49ms	3ms	12ms	1ms	31us	437ns	1ms	1ms	54ms	-
raidz2-6	41.8T	16.4T	53	59	20.2M	1.74M	47ms	3ms	12ms	1ms	34us	453ns	1ms	1ms	51ms	-
slot_13	-	-	13	14	5.05M	446K	48ms	3ms	12ms	1ms	37us	499ns	1ms	1ms	52ms	-
slot_14	-	-	13	14	5.05M	446K	47ms	3ms	12ms	1ms	31us	436ns	1ms	1ms	51ms	-



slot_43	-	-	13	14	5.05M	445K	47ms	3ms	12ms	1ms	34us	440ns	1ms	1ms	51ms	-
slot_44	-	-	13	14	5.05M	445K	47ms	3ms	12ms	1ms	34us	439ns	1ms	1ms	51ms	-
raidz2-7	41.8T	16.4T	54	63	20.2M	1.80M	48ms	3ms	12ms	1ms	34us	459ns	1ms	1ms	53ms	-
slot_15	-	-	13	15	5.06M	461K	48ms	3ms	12ms	1ms	35us	495ns	1ms	1ms	54ms	-
slot_16	-	-	13	15	5.06M	461K	48ms	3ms	12ms	1ms	35us	433ns	1ms	1ms	54ms	-
slot_45	-	-	13	15	5.06M	461K	47ms	3ms	12ms	1ms	31us	466ns	1ms	1ms	51ms	-
slot_46	-	-	13	15	5.06M	461K	49ms	3ms	12ms	1ms	35us	441ns	1ms	1ms	54ms	-
raidz2-8	41.6T	16.6T	67	73	20.1M	1.95M	39ms	2ms	9ms	1ms	42us	453ns	2ms	1ms	40ms	-
slot_17	-	-	14	18	5.03M	500K	44ms	2ms	11ms	1ms	34us	501ns	1ms	1ms	47ms	-
slot_18	-	-	14	18	5.03M	500K	45ms	2ms	11ms	1ms	42us	436ns	2ms	1ms	49ms	-
slot_47	-	-	14	18	5.04M	500K	44ms	2ms	11ms	1ms	36us	437ns	2ms	1ms	47ms	-
slot_48	-	-	23	18	5.02M	499K	29ms	2ms	6ms	1ms	56us	437ns	2ms	1ms	29ms	-
raidz2-9	41.8T	16.4T	55	68	20.2M	1.89M	49ms	3ms	12ms	1ms	34us	459ns	1ms	1ms	55ms	-
slot_19	-	-	13	17	5.06M	484K	50ms	3ms	12ms	1ms	34us	511ns	1ms	1ms	56ms	-
slot_20	-	-	13	17	5.06M	484K	50ms	3ms	12ms	1ms	35us	446ns	1ms	1ms	56ms	-
slot_49	-	-	13	17	5.06M	484K	49ms	3ms	12ms	1ms	33us	441ns	1ms	1ms	54ms	-
slot_50	-	-	13	17	5.06M	484K	49ms	3ms	12ms	1ms	35us	436ns	2ms	1ms	55ms	-
raidz2-10	41.8T	16.4T	53	60	20.2M	1.76M	47ms	3ms	12ms	1ms	31us	450ns	1ms	1ms	51ms	-
slot_21	-	-	13	15	5.06M	449K	47ms	3ms	12ms	1ms	32us	500ns	1ms	1ms	52ms	-
slot_22	-	-	13	15	5.06M	449K	46ms	3ms	12ms	1ms	29us	433ns	1ms	1ms	50ms	-
slot_51	-	-	13	15	5.06M	449K	47ms	3ms	12ms	1ms	31us	435ns	1ms	1ms	52ms	-
slot_52	-	-	13	15	5.06M	449K	47ms	3ms	12ms	1ms	31us	432ns	1ms	1ms	51ms	-
raidz2-11	41.8T	16.4T	54	64	20.2M	1.82M	49ms	3ms	12ms	1ms	40us	453ns	1ms	1ms	54ms	-
slot_23	-	-	13	16	5.06M	465K	49ms	3ms	12ms	1ms	40us	497ns	1ms	1ms	53ms	-
slot_24	-	-	13	16	5.06M	465K	50ms	3ms	12ms	1ms	37us	436ns	1ms	1ms	56ms	-
slot_53	-	-	13	16	5.06M	465K	49ms	3ms	12ms	1ms	40us	439ns	1ms	1ms	54ms	-
slot_54	-	-	13	16	5.06M	465K	49ms	3ms	12ms	1ms	41us	440ns	1ms	1ms	54ms	-
raidz2-12	41.8T	16.4T	54	69	20.2M	1.91M	49ms	3ms	12ms	1ms	33us	452ns	1ms	1ms	54ms	-
slot_25	-	-	13	17	5.06M	489K	49ms	3ms	12ms	1ms	31us	500ns	1ms	1ms	54ms	-
slot_26	-	-	13	17	5.06M	489K	50ms	3ms	12ms	1ms	34us	435ns	1ms	1ms	56ms	-
slot_55	-	-	13	17	5.06M	489K	49ms	3ms	12ms	1ms	35us	439ns	1ms	1ms	54ms	-
slot_56	-	-	13	17	5.05M	489K	49ms	3ms	12ms	1ms	32us	436ns	1ms	1ms	54ms	-
raidz2-13	41.8T	16.4T	52	62	20.2M	1.77M	49ms	3ms	12ms	1ms	34us	456ns	1ms	1ms	52ms	-
slot_27	-	-	13	15	5.04M	453K	49ms	3ms	12ms	1ms	32us	499ns	1ms	1ms	53ms	-
slot_28	-	-	13	15	5.04M	453K	49ms	3ms	12ms	1ms	36us	443ns	1ms	1ms	52ms	-
slot_57	-	-	13	15	5.04M	453K	49ms	3ms	12ms	1ms	32us	440ns	1ms	1ms	52ms	-
slot_58	-	-	13	15	5.04M	453K	48ms	3ms	12ms	1ms	35us	440ns	1ms	1ms	51ms	-
ogs	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
mirror-16	1.09M	125G	0	141	0	5.39M	2ms	18ms	2ms	173us	789ns	695ns	-	69ms	-	-
pmem0	-	-	0	70	0	2.69M	2ms	18ms	2ms	175us	804ns	748ns	-	69ms	-	-
pmem1	-	-	0	70	0	2.69M	2ms	18ms	2ms	170us	774ns	641ns	-	69ms	-	-



mirror-17	1.70M	125G	0	140	0	5.38M	3ms	16ms	3ms	168us	736ns	696ns	-	59ms	-	-
pmem2	-	-	0	70	0	2.69M	3ms	15ms	3ms	169us	768ns	747ns	-	59ms	-	-
pmem3	-	-	0	70	0	2.69M	3ms	16ms	3ms	166us	704ns	644ns	-	59ms	-	-
cache	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
nvme0n1	2.72T	196G	2	3	192K	312K	434us	3ms	432us	312us	1us	1us	732ns	3ms	731ns	-
nvme1n1	2.72T	197G	2	3	192K	312K	435us	3ms	434us	312us	1us	896ns	727ns	3ms	716ns	-

pool	capacity		operations		bandwidth		total_wait		disk_wait		syncq_wait		asyncq_wait		scrub	trim
	alloc	free	read	write	read	write	read	write	read	write	read	write	read	write		
cncnfs4	585T	230T	0	220	0	4.37M	-	1ms	-	1ms	-	886ns	-	531us	-	-
raidz2-0	41.8T	16.4T	0	22	0	373K	-	1ms	-	843us	-	-	-	585us	-	-
slot_01	-	-	0	5	0	93.4K	-	1ms	-	531us	-	-	-	582us	-	-
slot_02	-	-	0	5	0	93.4K	-	1ms	-	1ms	-	-	-	530us	-	-
slot_31	-	-	0	5	0	93.4K	-	1ms	-	983us	-	-	-	611us	-	-
slot_32	-	-	0	5	0	92.6K	-	1ms	-	515us	-	-	-	617us	-	-
raidz2-1	41.7T	16.5T	0	16	0	314K	-	767us	-	407us	-	-	-	336us	-	-
slot_03	-	-	0	3	0	78.2K	-	821us	-	474us	-	-	-	359us	-	-
slot_04	-	-	0	4	0	79.8K	-	786us	-	411us	-	-	-	356us	-	-
slot_33	-	-	0	4	0	80.6K	-	630us	-	352us	-	-	-	295us	-	-
slot_34	-	-	0	4	0	75.8K	-	861us	-	411us	-	-	-	345us	-	-
raidz2-2	41.8T	16.5T	0	6	0	34.3K	-	367us	-	355us	-	-	-	1us	-	-
slot_05	-	-	0	1	0	9.58K	-	327us	-	327us	-	-	-	725ns	-	-
slot_06	-	-	0	1	0	10.4K	-	452us	-	412us	-	-	-	3us	-	-
slot_35	-	-	0	0	0	6.38K	-	353us	-	353us	-	-	-	691ns	-	-
slot_36	-	-	0	1	0	7.98K	-	308us	-	308us	-	-	-	822ns	-	-
raidz2-3	41.9T	16.3T	0	9	0	77.4K	-	720us	-	442us	-	-	-	291us	-	-
slot_07	-	-	0	2	0	18.4K	-	643us	-	446us	-	-	-	268us	-	-
slot_08	-	-	0	2	0	19.2K	-	835us	-	442us	-	-	-	377us	-	-
slot_37	-	-	0	2	0	20.0K	-	851us	-	524us	-	-	-	311us	-	-
slot_38	-	-	0	2	0	20.0K	-	559us	-	362us	-	-	-	212us	-	-
raidz2-4	41.8T	16.4T	0	19	0	462K	-	507us	-	428us	-	-	-	83us	-	-
slot_09	-	-	0	4	0	117K	-	487us	-	432us	-	-	-	72us	-	-
slot_10	-	-	0	5	0	117K	-	529us	-	453us	-	-	-	87us	-	-
slot_39	-	-	0	4	0	117K	-	540us	-	442us	-	-	-	94us	-	-
slot_40	-	-	0	4	0	112K	-	471us	-	385us	-	-	-	77us	-	-
raidz2-5	41.8T	16.4T	0	34	0	673K	-	1ms	-	506us	-	-	-	637us	-	-
slot_11	-	-	0	8	0	165K	-	1ms	-	484us	-	-	-	601us	-	-
slot_12	-	-	0	8	0	168K	-	1ms	-	470us	-	-	-	518us	-	-
slot_41	-	-	0	8	0	170K	-	1ms	-	565us	-	-	-	738us	-	-
slot_42	-	-	0	8	0	169K	-	1ms	-	506us	-	-	-	691us	-	-



raidz2-6	41.8T	16.4T	0	13	0	121K	-	7ms	-	6ms	-	408ns	-	625us	-	-
slot_13	-	-	0	3	0	31.1K	-	3ms	-	3ms	-	624ns	-	141us	-	-
slot_14	-	-	0	2	0	29.5K	-	10ms	-	8ms	-	336ns	-	2ms	-	-
slot_43	-	-	0	2	0	28.7K	-	5ms	-	5ms	-	288ns	-	161us	-	-
slot_44	-	-	0	3	0	31.9K	-	8ms	-	8ms	-	384ns	-	113us	-	-
raidz2-7	41.8T	16.4T	0	17	0	267K	-	3ms	-	3ms	-	348ns	-	404us	-	-
slot_15	-	-	0	4	0	64.6K	-	2ms	-	1ms	-	384ns	-	486us	-	-
slot_16	-	-	0	4	0	67.8K	-	4ms	-	3ms	-	384ns	-	421us	-	-
slot_45	-	-	0	4	0	67.8K	-	4ms	-	3ms	-	336ns	-	405us	-	-
slot_46	-	-	0	4	0	66.2K	-	4ms	-	3ms	-	288ns	-	306us	-	-
raidz2-8	41.6T	16.6T	0	8	0	139K	-	701us	-	567us	-	288ns	-	275us	-	-
slot_17	-	-	0	2	0	35.1K	-	720us	-	557us	-	336ns	-	270us	-	-
slot_18	-	-	0	1	0	34.3K	-	786us	-	668us	-	336ns	-	393us	-	-
slot_47	-	-	0	2	0	33.5K	-	786us	-	643us	-	240ns	-	337us	-	-
slot_48	-	-	0	2	0	35.9K	-	518us	-	411us	-	240ns	-	119us	-	-
raidz2-9	41.8T	16.4T	0	9	0	270K	-	1ms	-	468us	-	-	-	665us	-	-
slot_19	-	-	0	1	0	67.8K	-	1ms	-	471us	-	-	-	634us	-	-
slot_20	-	-	0	2	0	67.0K	-	1ms	-	449us	-	-	-	767us	-	-
slot_49	-	-	0	2	0	67.0K	-	1ms	-	491us	-	-	-	558us	-	-
slot_50	-	-	0	2	0	67.8K	-	1ms	-	464us	-	-	-	679us	-	-
raidz2-10	41.8T	16.4T	0	10	0	77.4K	-	886us	-	524us	-	-	-	336us	-	-
slot_21	-	-	0	2	0	19.9K	-	907us	-	514us	-	-	-	287us	-	-
slot_22	-	-	0	2	0	18.4K	-	786us	-	491us	-	-	-	295us	-	-
slot_51	-	-	0	2	0	19.2K	-	929us	-	643us	-	-	-	286us	-	-
slot_52	-	-	0	2	0	19.9K	-	917us	-	471us	-	-	-	446us	-	-
raidz2-11	41.8T	16.4T	0	4	0	54.3K	-	933us	-	540us	-	-	-	387us	-	-
slot_23	-	-	0	1	0	13.6K	-	983us	-	524us	-	-	-	336us	-	-
slot_24	-	-	0	1	0	13.6K	-	917us	-	589us	-	-	-	336us	-	-
slot_53	-	-	0	1	0	13.6K	-	954us	-	505us	-	-	-	520us	-	-
slot_54	-	-	0	0	0	13.6K	-	865us	-	550us	-	-	-	325us	-	-
raidz2-12	41.8T	16.4T	0	31	0	790K	-	2ms	-	1ms	-	-	-	1ms	-	-
slot_25	-	-	0	8	0	196K	-	1ms	-	1ms	-	-	-	487us	-	-
slot_26	-	-	0	7	0	197K	-	5ms	-	3ms	-	-	-	2ms	-	-
slot_55	-	-	0	7	0	199K	-	1ms	-	796us	-	-	-	597us	-	-
slot_56	-	-	0	7	0	198K	-	1ms	-	915us	-	-	-	851us	-	-
raidz2-13	41.8T	16.4T	0	3	0	102K	-	869us	-	703us	-	-	-	109us	-	-
slot_27	-	-	0	0	0	25.5K	-	1ms	-	983us	-	-	-	74us	-	-
slot_28	-	-	0	0	0	26.3K	-	786us	-	629us	-	-	-	118us	-	-
slot_57	-	-	0	0	0	25.5K	-	786us	-	629us	-	-	-	118us	-	-
slot_58	-	-	0	0	0	24.7K	-	786us	-	629us	-	-	-	118us	-	-



mirror-16	1.77M	125G	0	6	0	375K	-	206us	-	171us	-	1us	-	136us	-	-
pmem0	-	-	0	3	0	188K	-	196us	-	185us	-	2us	-	124us	-	-
pmem1	-	-	0	3	0	188K	-	216us	-	158us	-	945ns	-	148us	-	-
mirror-17	2.27M	125G	0	5	0	346K	-	280us	-	234us	-	1us	-	136us	-	-
pmem2	-	-	0	2	0	173K	-	272us	-	246us	-	1us	-	124us	-	-
pmem3	-	-	0	2	0	173K	-	288us	-	222us	-	907ns	-	148us	-	-
cache	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
nvme0n1	2.72T	196G	0	0	0	0	-	-	-	-	-	-	-	-	-	-
nvme1n1	2.72T	197G	0	0	0	102	-	98us	-	98us	-	-	-	3us	-	-
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



root@disque1-math:~# zpool iostat -v1 5 2

pool	capacity		operations		bandwidth		total_wait		disk_wait		syncq_wait		asyncq_wait		scrub	trim
	alloc	free	read	write	read	write	read	write	read	write	read	write	read	write	wait	wait
CZJ2230DH8_0	135T	18.0T	151	704	47.3M	33.7M	160ms	3ms	8ms	927us	12ms	2us	7ms	3ms	192ms	-
raidz2	19.2T	2.59T	21	62	6.74M	2.97M	165ms	5ms	8ms	1ms	13ms	1us	7ms	4ms	200ms	-
slot_1	-	-	6	20	2.27M	1013K	188ms	6ms	10ms	1ms	9ms	1us	7ms	4ms	250ms	-
slot_2	-	-	6	20	2.22M	1013K	192ms	5ms	10ms	1ms	14ms	1us	11ms	4ms	229ms	-
slot_13	-	-	8	20	2.26M	1013K	127ms	5ms	6ms	1ms	17ms	990ns	4ms	4ms	146ms	-
raidz2	19.3T	2.52T	22	60	6.79M	2.87M	147ms	5ms	8ms	1ms	12ms	980ns	7ms	4ms	174ms	-
slot_3	-	-	5	20	2.25M	981K	192ms	6ms	10ms	1ms	17ms	1us	7ms	4ms	233ms	-
slot_14	-	-	8	20	2.29M	981K	129ms	5ms	7ms	1ms	9ms	959ns	7ms	4ms	153ms	-
slot_15	-	-	8	20	2.25M	981K	133ms	5ms	7ms	1ms	10ms	976ns	7ms	4ms	154ms	-
raidz2	19.2T	2.58T	21	63	6.75M	2.99M	172ms	5ms	8ms	1ms	13ms	1us	7ms	4ms	209ms	-
slot_4	-	-	6	20	2.27M	1022K	190ms	6ms	10ms	1ms	7ms	1us	7ms	4ms	249ms	-
slot_5	-	-	5	20	2.21M	1022K	210ms	6ms	10ms	1ms	13ms	1us	9ms	4ms	252ms	-
slot_16	-	-	8	21	2.27M	1022K	132ms	5ms	6ms	1ms	17ms	1us	7ms	4ms	154ms	-
raidz2	19.3T	2.57T	22	61	6.76M	2.93M	151ms	6ms	8ms	1ms	13ms	1us	9ms	4ms	178ms	-
slot_6	-	-	6	20	2.25M	1001K	185ms	6ms	10ms	1ms	21ms	1us	8ms	4ms	224ms	-
slot_17	-	-	8	20	2.27M	1001K	133ms	6ms	7ms	1ms	6ms	993ns	12ms	4ms	160ms	-
slot_18	-	-	7	20	2.24M	1001K	143ms	6ms	7ms	1ms	11ms	992ns	7ms	4ms	163ms	-
raidz2	19.2T	2.59T	21	64	6.75M	3.02M	168ms	5ms	8ms	1ms	12ms	1us	6ms	4ms	204ms	-
slot_7	-	-	6	21	2.28M	1.01M	177ms	5ms	9ms	1ms	14ms	1us	4ms	4ms	226ms	-
slot_8	-	-	6	21	2.25M	1.01M	194ms	5ms	10ms	1ms	12ms	993ns	6ms	4ms	245ms	-
slot_19	-	-	8	21	2.23M	1.01M	141ms	5ms	7ms	1ms	9ms	999ns	9ms	4ms	160ms	-
raidz2	19.3T	2.55T	22	58	6.77M	2.86M	148ms	6ms	8ms	1ms	12ms	1us	6ms	4ms	175ms	-
slot_9	-	-	6	19	2.27M	978K	167ms	6ms	9ms	1ms	8ms	1us	6ms	4ms	212ms	-
slot_20	-	-	7	19	2.25M	978K	150ms	6ms	7ms	1ms	11ms	1us	8ms	4ms	176ms	-
slot_21	-	-	8	19	2.24M	978K	129ms	6ms	7ms	1ms	16ms	990ns	4ms	4ms	146ms	-
raidz2	19.2T	2.59T	21	64	6.75M	3.04M	170ms	6ms	9ms	1ms	12ms	1us	8ms	4ms	207ms	-
slot_10	-	-	6	21	2.25M	1.01M	187ms	6ms	10ms	1ms	10ms	1us	11ms	4ms	233ms	-
slot_11	-	-	6	21	2.25M	1.01M	184ms	6ms	10ms	1ms	12ms	1us	6ms	4ms	229ms	-
slot_22	-	-	8	21	2.26M	1.01M	146ms	5ms	7ms	1ms	14ms	998ns	6ms	4ms	173ms	-
logs	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
mirror	20.0M	125G	0	134	0	6.52M	68us	308us	65us	102us	3us	2us	-	623us	-	-
pmem0	-	-	0	67	0	3.26M	67us	307us	65us	97us	4us	2us	-	634us	-	-
pmem2	-	-	0	67	0	3.26M	68us	308us	66us	108us	3us	1us	-	611us	-	-
mirror	18.4M	125G	0	135	0	6.53M	64us	303us	63us	102us	4us	2us	-	610us	-	-
pmem1	-	-	0	67	0	3.26M	58us	301us	58us	96us	3us	2us	-	618us	-	-
pmem3	-	-	0	67	0	3.26M	70us	305us	67us	107us	5us	1us	-	602us	-	-
cache	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-



L2ARC	4.64T	832G	33	17	2.31M	1.61M	193us	2ms	190us	205us	3us	2us	1us	2ms	2us	-	
capacity		operations		bandwidth		total_wait		disk_wait		syncq_wait		asyncq_wait		scrub	trim		
pool	alloc	free	read	write	read	write	read	write	read	write	read	write	read	write	wait	wait	
CZJ2230DH8_0	135T	18.0T	2	374	384K	8.98M	629us	876us	629us	379us	-	2us	2us	720us	-	-	
raidz2	19.2T	2.59T	0	41	0	1.14M	-	1ms	-	490us	-	832ns	-	864us	-	-	
slot_1	-	-	0	13	0	388K	-	1ms	-	538us	-	864ns	-	920us	-	-	
slot_2	-	-	0	13	0	388K	-	1ms	-	514us	-	768ns	-	849us	-	-	
slot_13	-	-	0	14	0	388K	-	1ms	-	422us	-	864ns	-	827us	-	-	
raidz2	19.3T	2.52T	0	39	0	973K	-	1ms	-	506us	-	736ns	-	802us	-	-	
slot_3	-	-	0	13	0	324K	-	1ms	-	526us	-	672ns	-	799us	-	-	
slot_14	-	-	0	13	0	324K	-	1ms	-	492us	-	768ns	-	828us	-	-	
slot_15	-	-	0	12	0	324K	-	1ms	-	499us	-	768ns	-	776us	-	-	
raidz2	19.2T	2.58T	0	44	128K	1.30M	629us	1ms	629us	526us	-	960ns	2us	741us	-	-	
slot_4	-	-	0	14	128K	444K	629us	1ms	629us	555us	-	768ns	2us	808us	-	-	
slot_5	-	-	0	14	0	444K	-	1ms	-	526us	-	1us	-	734us	-	-	
slot_16	-	-	0	15	0	444K	-	1ms	-	496us	-	864ns	-	683us	-	-	
raidz2	19.3T	2.57T	0	43	51.1K	1.13M	786us	1ms	786us	484us	-	-	3us	635us	-	-	
slot_6	-	-	0	14	0	384K	-	1ms	-	481us	-	-	-	631us	-	-	
slot_17	-	-	0	14	0	384K	-	1ms	-	486us	-	-	-	623us	-	-	
slot_18	-	-	0	14	51.1K	384K	786us	1ms	786us	484us	-	-	3us	650us	-	-	
raidz2	19.2T	2.59T	0	38	76.7K	1.34M	524us	1ms	524us	548us	-	-	2us	860us	-	-	
slot_7	-	-	0	13	0	459K	-	1ms	-	586us	-	-	-	955us	-	-	
slot_8	-	-	0	13	0	459K	-	1ms	-	530us	-	-	-	804us	-	-	
slot_19	-	-	0	12	76.7K	459K	524us	1ms	524us	526us	-	-	2us	818us	-	-	
raidz2	19.3T	2.55T	0	22	128K	1.13M	629us	929us	629us	613us	-	-	3us	295us	-	-	
slot_9	-	-	0	7	128K	385K	629us	1ms	629us	677us	-	-	3us	326us	-	-	
slot_20	-	-	0	7	0	385K	-	900us	-	605us	-	-	-	294us	-	-	
slot_21	-	-	0	7	0	385K	-	786us	-	558us	-	-	-	265us	-	-	
raidz2	19.2T	2.59T	0	29	0	1.25M	-	1ms	-	622us	-	-	-	659us	-	-	
slot_10	-	-	0	10	0	426K	-	1ms	-	609us	-	-	-	735us	-	-	
slot_11	-	-	0	9	0	426K	-	1ms	-	629us	-	-	-	720us	-	-	
slot_22	-	-	0	9	0	426K	-	1ms	-	629us	-	-	-	516us	-	-	
logs	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
mirror	17.4M	125G	0	53	0	380K	-	29us	-	25us	-	2us	-	-	-	-	
pmem0	-	-	0	26	0	190K	-	31us	-	25us	-	3us	-	-	-	-	
pmem2	-	-	0	26	0	190K	-	27us	-	24us	-	1us	-	-	-	-	
mirror	15.8M	125G	0	58	0	387K	-	28us	-	25us	-	2us	-	-	-	-	
pmem1	-	-	0	29	0	193K	-	29us	-	24us	-	3us	-	-	-	-	
pmem3	-	-	0	29	0	193K	-	27us	-	25us	-	1us	-	-	-	-	



cache	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L2ARC	4.64T	832G	0	2	6.39K	43.6K	196us	144us	196us	124us	13us	-	-	20us	-	-
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



```
root@srv-sto-01:~# arcstat 5 2
  time read miss miss% dmis dm% pmis pm% mmis mm% size  c avail
22:30:00   4   0   0   0   0   0   0   0   0   0 180G 188G 165G
22:30:05 250K   1   0   0   0   0   0 100   1   0 180G 188G 165G

root@disquel-math:~# arcstat 5 2
  time read miss miss% dmis dm% pmis pm% mmis mm% size  c avail
22:30:48   3   0   0   0   0   0   0   0   0 125G 125G  85G
22:30:53 3.7K   0   0   0   0   0   0   0   0 125G 125G  85G
```



```
root@srv-sto-01:~# for SECTION in arc archits l2arc zil ; do arc_summary -s $SECTION ; done
```

```
-----  
ZFS Subsystem Report                               Tue Nov 28 22:35:41 2023  
Linux 5.10.0-20-amd64                             2.1.7-1~bpo11+1  
Machine: srv-sto-01 (x86_64)                     2.1.7-1~bpo11+1
```

```
ARC status:                                       HEALTHY  
Memory throttle count:                          0
```

```
ARC size (current):                               95.6 % 180.5 GiB  
Target size (adaptive):                         100.0 % 188.8 GiB  
Min size (hard limit):                          6.2 % 11.8 GiB  
Max size (high water):                          16:1 188.8 GiB  
Most Frequently Used (MFU) cache size:          64.3 % 108.7 GiB  
Most Recently Used (MRU) cache size:            35.7 % 60.5 GiB  
Metadata cache size (hard limit):              75.0 % 141.6 GiB  
Metadata cache size (current):                  11.5 % 16.2 GiB  
Dnode cache size (hard limit):                  10.0 % 14.2 GiB  
Dnode cache size (current):                     20.7 % 2.9 GiB
```

```
ARC hash breakdown:  
Elements max:                                    451.5M  
Elements current:                               14.5 % 65.4M  
Collisions:                                      6.3G  
Chain max:                                       26  
Chains:                                          17.1M
```

```
ARC misc:  
Deleted:                                         3.3G  
Mutex misses:                                    9.3M  
Eviction skips:                                  4.6G  
Eviction skips due to L2 writes:                 305  
L2 cached evictions:                             45.1 TiB  
L2 eligible evictions:                           11.2 TiB  
L2 eligible MFU evictions:                       20.0 % 2.2 TiB  
L2 eligible MRU evictions:                       80.0 % 9.0 TiB  
L2 ineligible evictions:                         310.7 TiB
```




```
-----  
ZFS Subsystem Report                               Tue Nov 28 22:35:41 2023  
Linux 5.10.0-20-amd64                             2.1.7-1~bpo11+1  
Machine: srv-sto-01 (x86_64)                     2.1.7-1~bpo11+1  
  
ARC total accesses (hits + misses):              804.3G  
  Cache hit ratio:                               99.7 %    802.3G  
  Cache miss ratio:                              0.3 %     2.1G  
  Actual hit ratio (MFU + MRU hits):             99.7 %    802.2G  
  Data demand efficiency:                       98.5 %    17.9G  
  Data prefetch efficiency:                      7.4 %    181.0M  
  
Cache hits by cache type:  
  Most frequently used (MFU):                   97.6 %    783.0G  
  Most recently used (MRU):                     2.4 %     19.2G  
  Most frequently used (MFU) ghost:             < 0.1 %    27.8M  
  Most recently used (MRU) ghost:               < 0.1 %    77.0M  
  
Cache hits by data type:  
  Demand data:                                  2.2 %     17.6G  
  Prefetch data:                               < 0.1 %    13.4M  
  Demand metadata:                             97.7 %    784.0G  
  Prefetch metadata:                           0.1 %    633.7M  
  
Cache misses by data type:  
  Demand data:                                  12.7 %    260.5M  
  Prefetch data:                                8.2 %    167.6M  
  Demand metadata:                             2.3 %     46.6M  
  Prefetch metadata:                           76.9 %     1.6G
```



```
-----  
ZFS Subsystem Report                               Tue Nov 28 22:35:41 2023  
Linux 5.10.0-20-amd64                             2.1.7-1~bpo11+1  
Machine: srv-sto-01 (x86_64)                     2.1.7-1~bpo11+1  
  
L2ARC status:                                     HEALTHY  
  Low memory aborts:                               0  
  Free on write:                                   37.9k  
  R/W clashes:                                     0  
  Bad checksums:                                   0  
  I/O errors:                                       0  
  
L2ARC size (adaptive):                            7.2 TiB  
  Compressed:                                       75.3 %   5.4 TiB  
  Header size:                                       0.1 %   5.5 GiB  
  MFU allocated size:                               63.2 %   3.4 TiB  
  MRU allocated size:                               36.7 %   2.0 TiB  
  Prefetch allocated size:                          < 0.1 %  857.8 MiB  
  Data (buffer content) allocated size:            99.7 %   5.4 TiB  
  Metadata (buffer content) allocated size:        0.3 %   16.4 GiB  
  
L2ARC breakdown:                                  2.1G  
  Hit ratio:                                        6.8 %   140.5M  
  Miss ratio:                                       93.2 %   1.9G  
  Feeds:                                           29.3M  
  
L2ARC writes:  
  Writes sent:                                     100 %   10.4M  
  
L2ARC evicts:  
  Lock retries:                                    21.7k  
  Upon reading:                                    0
```



```
-----  
ZFS Subsystem Report                               Tue Nov 28 22:35:41 2023  
Linux 5.10.0-20-amd64                             2.1.7-1~bpo11+1  
Machine: srv-sto-01 (x86_64)                      2.1.7-1~bpo11+1  
  
ZIL committed transactions:                       5.3G  
  Commit requests:                               3.0G  
  Flushes to stable storage:                     3.0G  
  Transactions to SLOG storage pool:             136.6 TiB  4.0G  
  Transactions to non-SLOG storage pool:         0 Bytes    0
```



```
root@disque1-math:~# for SECTION in arc archits l2arc zil ; do arc_summary -s $SECTION ; done
```

```
-----
ZFS Subsystem Report                               Tue Nov 28 22:38:19 2023
Linux 5.10.0-23-amd64                             2.0.3-9+deb11u1
Machine: disque1-math (x86_64)                   2.0.3-9+deb11u1

ARC status:                                       HEALTHY
  Memory throttle count:                          0

ARC size (current):                              100.3 % 126.1 GiB
  Target size (adaptive):                         100.0 % 125.8 GiB
  Min size (hard limit):                          6.2 %   7.9 GiB
  Max size (high water):                           16:1 125.8 GiB
  Most Frequently Used (MFU) cache size:           96.8 % 104.4 GiB
  Most Recently Used (MRU) cache size:             3.2 %   3.5 GiB
  Metadata cache size (hard limit):                75.0 %  94.3 GiB
  Metadata cache size (current):                   46.5 %  43.8 GiB
  Dnode cache size (hard limit):                   10.0 %   9.4 GiB
  Dnode cache size (current):                       64.9 %   6.1 GiB

ARC hash breakdown:
  Elements max:                                    101.4M
  Elements current:                                80.8 %   81.9M
  Collisions:                                       3.1G
  Chain max:                                        16
  Chains:                                           23.5M

ARC misc:
  Deleted:                                         1.3G
  Mutex misses:                                    5.5M
  Eviction skips:                                  1.6G
```



```
-----  
ZFS Subsystem Report                               Tue Nov 28 22:38:19 2023  
Linux 5.10.0-23-amd64                             2.0.3-9+deb11u1  
Machine: disque1-math (x86_64)                   2.0.3-9+deb11u1  
  
ARC total accesses (hits + misses):                503.9G  
  Cache hit ratio:                                99.8 %    503.1G  
  Cache miss ratio:                               0.2 %    822.4M  
  Actual hit ratio (MFU + MRU hits):               99.8 %    503.1G  
  Data demand efficiency:                          85.9 %     2.8G  
  Data prefetch efficiency:                        69.4 %    274.6M  
  
Cache hits by cache type:  
  Most frequently used (MFU):                      99.1 %    498.8G  
  Most recently used (MRU):                        0.8 %     4.2G  
  Most frequently used (MFU) ghost:                < 0.1 %    17.4M  
  Most recently used (MRU) ghost:                  < 0.1 %   107.0M  
  
Cache hits by data type:  
  Demand data:                                     0.5 %     2.4G  
  Demand prefetch data:                            < 0.1 %   190.4M  
  Demand metadata:                                 99.1 %    498.8G  
  Demand prefetch metadata:                        0.3 %     1.7G  
  
Cache misses by data type:  
  Demand data:                                     47.7 %    392.6M  
  Demand prefetch data:                            10.2 %    84.2M  
  Demand metadata:                                 8.5 %    70.3M  
  Demand prefetch metadata:                        33.5 %   275.4M
```



```
-----  
ZFS Subsystem Report                               Tue Nov 28 22:38:19 2023  
Linux 5.10.0-23-amd64                             2.0.3-9+deb11u1  
Machine: disque1-math (x86_64)                    2.0.3-9+deb11u1  
  
L2ARC status:                                     HEALTHY  
  Low memory aborts:                             0  
  Free on write:                                 825.3k  
  R/W clashes:                                   1  
  Bad checksums:                                 0  
  I/O errors:                                    0  
  
L2ARC size (adaptive):                            8.4 TiB  
  Compressed:                                    55.0 %   4.6 TiB  
  Header size:                                   0.1 %   6.8 GiB  
  
L2ARC breakdown:                                  822.4M  
  Hit ratio:                                      47.1 %   387.1M  
  Miss ratio:                                    52.9 %   435.3M  
  Feeds:                                          13.0M  
  
L2ARC writes:                                       
  Writes sent:                                   100 %    7.9M  
  
L2ARC evicts:                                       
  Lock retries:                                  1.4k  
  Upon reading:                                  0
```



```
-----  
ZFS Subsystem Report                               Tue Nov 28 22:38:19 2023  
Linux 5.10.0-23-amd64                             2.0.3-9+deb11u1  
Machine: disque1-math (x86_64)                   2.0.3-9+deb11u1  
  
ZIL committed transactions:                       2.2G  
  Commit requests:                               1.1G  
  Flushes to stable storage:                     1.1G  
  Transactions to SLOG storage pool:              68.6 TiB  1.5G  
  Transactions to non-SLOG storage pool:          0 Bytes  0
```

