Kubatbek Alimbekov

TC 362

The Hardware that Powers Etsy

Etsy relies heavily on a large number of MySQL machines to store the data. The servers used are HP DL380 servers, with configurations at 2U of rack space, 2x 8 core Intel E5630 CPUs (@ 2.53ghz), 96GB of RAM (for MySQL buffer cache) and 16x 15,000 RPM 146GB hard discs. Etsy uses only one 4x 1gbit Ethernet port, but all machines have one. All this allows Etsy to keep a balance between the disc space to store user data and RAM to retrieve it efficiently. Etsy also has a number of Solid-State Drives (SSD) in use throughout the infrastructure and they are to be tested/utilized for databases as well in near future.

To promote re-use of hardware and to be able to change machine roles quickly and easily, Etsy generally utilizes one type of machine. For these purposes Etsy uses 2U Supermicro chassis that allows for 4x nodes to share to power supplies with 12 3.5" discs on the front. The configurations are as follows: 2x 8 core Intel E5620 CPUs (@ 2.40ghz), 12GB-96GB of RAM, and either a 600GB 7200pm hard disk or an Intel 160GB SSD. Etsy has been building its Hadoop cluster as well and uses a similar chassis as previously mentioned with a small change of using 24 2.5" discs instead of 12 3.5" on the front. Each node (with 4 in a 2U chassis) has 2x 12 core Intel E5646 CPUs (@ 2.40ghz), 96GB of RAM, and 6x 1Tb 2.5″ 7200rpm disks. That's 96 cores, 384GB of RAM and 24TB per 2U of rack space. Etsy web servers don't have important data and all the logs are constantly sent to Etsy's centralized logging host. This way the web code can always easily be deployed back on the machine.

For search Etsy utilizes Intel's Sandy Bridge CPUs. They use 2x 16 core Intel E5-2690 CPUs clocked at 2.90ghz. This configurations works really well for search purposes since search is CPU dependent and the machines can handle 4 times more workload the nodes listed earlier, while not consuming that much more power. SSDs are used in these machines and nodes have 96GB RAM and 8000GB SSD for the indexes. Also there is a BitTorrent index distribution used to make the index transfer to machines very fast.

For backups Etsy uses 36 disc Supermicro chassis 6047R-E1R36N. It is a 4U chassis with 36X 3.5" discs with 2TB 7200rpm drives and an LSI RAID controller with 1gn of battery backed write back cache. This gives a sequential write throughput of 1.2 gb/sec and 60TB of usable disc space across two RAID6 volumes.

For other purposes where machines don't need much power or disc space but still need RAID for data security (e.g. DNS servers, LDAP servers, Hadoop Namenodes etc) Etsy uses HP DL360 servers with configurations of 1U, 4x 2.5" 15,000rpm 146GB SAS discs.

For security Etsy uses two-factor authentication and full site SSL support. Login history is also available.