120 TByte and 120 CPUs in 10 rack height units

Pfaffhausen, 12th May 2011: *Until two years ago, we did not deliver any computers capable of handling more than 320 GB of data. In the year before last, we achieved the 8 TB mark for the first time with the Summit and Matterhorn ArchivistaBoxes. We are delighted to introduce you to the new ArchivistaVM SwissRocket cluster generation.*

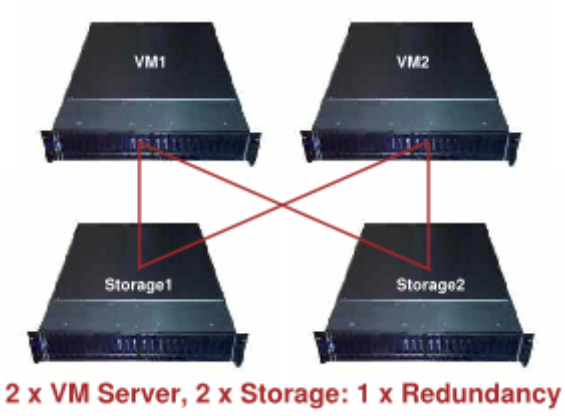
Conventional cluster concepts offer little security against failure.

By way of introduction, we should state that a cluster is a grouping of servers. For a failsafe configuration (high availability), this means that if one or more components fail, the computer grouping (cluster) can continue running without interruptions. We are also talking here about redundancy (each component is doubled up).

Before we present the new SwissRocket cluster concept, I would like to introduce some existing cluster concepts, which I have observed in KMU companies in the last two years. Even in smaller companies, there is the problem that compromises are often made on cost grounds and these later have consequences, sometimes fatal, in operation or for further expansion.

With virtualization, a physical server hosts a number of instances (virtual computers). CPU cores and hard disk(s) are also split. The individual instances also exist as an image on the physical server. Instances can be backed up and also played back again. Only in the event of a failure are all virtualized instances affected. This is why concepts that include a failsafe are often chosen. In other words, several servers are set up for virtualization (VM servers) and data storage (storage servers). If a server is "struggling", then the operation continues without any restrictions. Moreover, the watchword is often "Failover". The remaining computers detect failure of a member and share the load over the remaining computers.

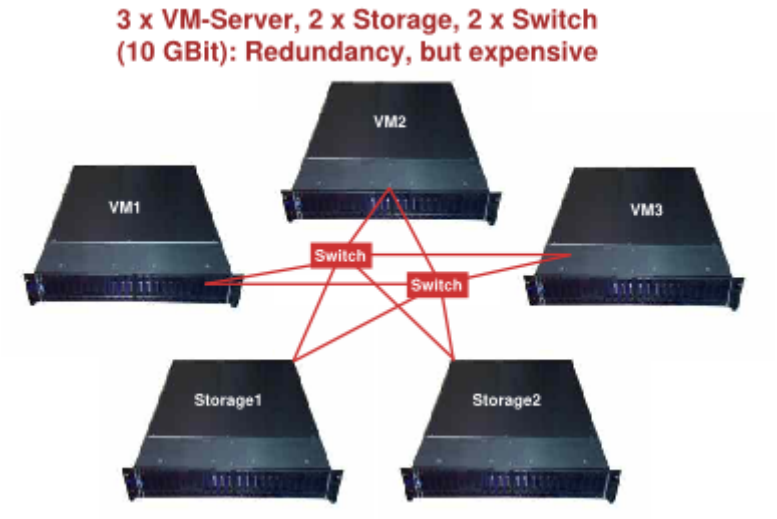
Because with virtualization, the machines in operation have to share the capacity of the hard disks, separate servers are usually configured for virtualization and data storage. Servers for virtualization contain as many CPU cores and as much RAM as possible, with the storage servers (storage, NAS, SAN etc.) providing fast hard disk links. However, what appears fully evident at first glance is not considered complete. Let us consider an implementation with 3 servers. Two servers work for virtualization, the third computer stores the data in two hard disk groupings with two separate controllers. We have a problem with this configuration, because there is again no redundancy. Should there be a problem with the data storage server for, the entire cluster fails.



Therefore, at least two servers for virtualization and two servers for data storage should be implemented. Both VM servers store the data redundantly on the storage servers. If a VM server fails, the second VM server takes over the task; if a storage server fails, the second "storage slave" takes over the task. We have therefore achieved a failsafe situation. However, with this concept, it should be noted that conventional network cards (1 GB) can transfer only

approx. 100 MB of data per second. In other words, the fastest storage servers are of no avail if the data is queued on the line. With this concept, you must therefore provide fast 10 GB network cards for adequate bandwidth (approx. 1 GB/s).

10 GB network cards are no longer expensive, but the corresponding switches are. With 2 VM servers and 2 storage servers, dual port network cards with crossover cables can be used. In this way, it is possible to avoid using switches. However, such a cluster cannot be expanded without further equipment. Essentially, four servers are required to make a computer failure more difficult.



Formerly, there were four non-virtualized servers in the company, but now there are always four [virtualized] servers in the server room. However, if a server fails, then we can continue operations without interruption on the three remaining servers. OnlOnly without virtualization, we could install additional new servers as required at any time; with a cluster for virtualization, this is no longer so simple. Clusters are therefore often implemented with 3 or more VM servers.

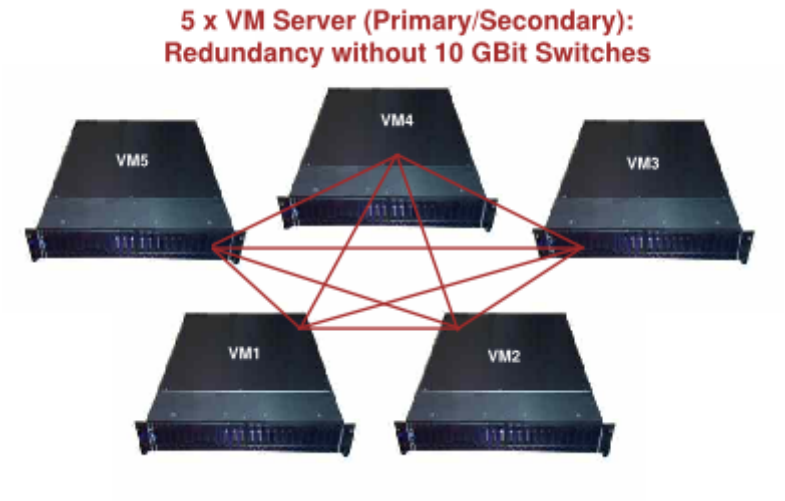
Below, we shall assume a scenario with 3 VM servers and 2 storage computers. In addition, we can use the full capacity of two VM servers for operations and the third VM server for test purposes. In order to keep these VM servers highly available, redundant 10-GB switches are needed, because a single switch can also fail and in this case the complete solution would come to a halt. 10-GB switches are not cost-effective – a 6-port switch still costs between CHF 6000 and 8000 and, with two switches, you should be thinking about costs of between CHF 12000 and 15000. To put this in perspective, connecting a maximum of 6 servers with a 6-port switch and a 12-port or 24-port switch with 10 GB capability soon costs CHF 20000 per switch.

Archivista SwissRocket: Failsafe cluster with ArchivistaVM

In this case, the concept uses our Archivista SwissRocket cluster. The 3-way cluster is already failsafe (2 computers in operation, 1 computer on standby). If there is a failure of one of the three servers, the standby computer is switched in. Yet how can we keep the data from the two productive servers available with 3 servers? Quite simply, all our SwissRocket servers are designed so that any server automatically stores the data from another computer with its own in a second hard disk grouping. Each SwissRocket server is at once primary (productive) and Secondary (standby). In a hard disk grouping with Raid10, therefore, approx. 2 to 3 percent of a CPU is needed for computing performance, in a Raid5 grouping, this figure is approx. 40 to 50 percent of a CPU.

High availability is always relative. The question of whether a server is available even during data backup must be distinguished from availability at the virtualization stage. In Archivista SwissRocket clusters, the secondary server disengages, completes the data backup and then cuts back in downstream. In order not to hazard data integrity, each instance on the primary server is either restarted or set to standby for a short time. The servers are then available for further work without any restrictions. With a restart, to reach failsafe status takes less than 1 minute, and on standby this is a few seconds.

With the SwissRocket model series, we offer a maximum of 24 CPUs and 24 TB of data per server. There are at least 6 CPUs and 3 TB of hard disk capacity; ultimately, not every KMU undertaking needs 24 CPUs in its basic scope (corresponds to approx. 12 to 24 servers). In a quintuple cluster, up to 120 CPUs and 120 TB can be achieved. Obviously, clusters can be expanded downstream. Redundancy is also possible already with 2 servers (1 × productive, 1 × standby). Such a dual cluster can be converted at any time to a triple, quadruple or quintuple cluster, just as a quintuple cluster can be expanded to a maximum of 24 clusters (24 × 48 = 1152 CPUs).



The Archivista SwissRocket cluster is completely subject to GPL licensing. With hardware, we rely on branded components, but we intentionally assemble the servers in house. Experience with bought-in servers has taught us that we achieve far better concepts if we use standard components. However, I should have mentioned here that the concept could also be implemented with branded servers, if the hard disks can be activated and 10 GB network cards are available. The **key technical data for our Archivista SwissRocket cluster can be taken from the price list,** just as you can experience the SwissRocket cluster live in action with us. You will be impressed by the performance; as an example, we quote here the time that is needed to write 1 TB of data on to a SwissRocket server:

```
dd if=/dev/zero of=/var/lib/vz/1TB.img bs=256M count=3760
oflag=direct
3760+0 records in
3760+0 records out
1009317314560 bytes (1.0 TB) copied, 861.507 s, 1.2 GB/s
```

For those who would prefer to calculate the computer performance, 861.5 seconds equals 14 minutes and 21 seconds. The author remembers the days when 1 minute was needed for 300 MB. Admittedly, this was a long time ago (if I remember correctly, it was at about the turn of the millennium). However, the comparison, converted to 14 minutes 21 second, produces approx. 4.3 GB. In other words, the SwissRocket server is 232 times faster.

Key data and prices for the new Archivista SwissRocket server

The Archivista SwissRocket clusters cover a range from 12 CPUs to 120 CPUs and there is nothing to prevent a further expansion (up to 24 clusters with 1152 CPUs). Clusters can be configured automatically. Depending on the expansion classification and performance class, rack and/or desktop housings can be used. When idle, 24 CPU machine with 24 hard disks need approx. 150 watts and under full load approx. 300

watts; in other words, we need between 6 and 12 watts (±) per CPU. The dual cluster starts at under CHF 7000, the triple cluster costs just under CHF 10000, a quintuple cluster with 120 CPUs and 120 TB memory (maximum usable, 96 CPUs and 110 TB) is somewhere between CHF 50000 and 60000. To put this in perspective, **prices include all the necessary hardware and software.** We will be **pleased to advise you** on the choice of cluster best for you.