Lenovo ThinkAgile CP
High Availability Architecture
(Technical Brief)

Models: CP 4000, CP 6000

Note

Before using this information and the product it supports, be sure to read and understand the safety information and the safety instructions, which are available at the following address:

http://thinksystem.lenovofiles.com/help/topic/safety_documentation/pdf_files.html

In addition, be sure that you are familiar with the terms and conditions of the Lenovo warranty for your solution, which can be found at the following address:

http://datacentersupport.lenovo.com/warrantylookup

First Edition (March 2020)

Contents

ThinkAgile CP High Availability Architecture

This technical brief describes the High-Availability (HA) Architecture of the ThinkAgile CP platform. The composable platform is designed to give you a true public cloud experience but inside your own datacenter. The ThinkAgile CPinnovation is in its software, which is designed to run on standard hardware; however, to ensure the very best customer experience, our offering includes both software and hardware. As will become clear, there are no single points of failure in ThinkAgile CP.

Download PDF

Refer to the following topics:

ThinkAgile CP platform overview

The ThinkAgile CP platform has three hardware blocks - the compute block, the storage block and the interconnect switches.

Hardware

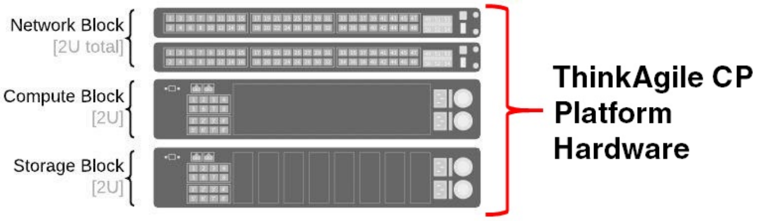


Figure 1. Platform Hardware

Platform Software

The ThinkAgile CP Platform has five software components, described as follows:

- Hypervisor and compute controller software running in compute nodes in the compute blocks (ThinkAgile CP OS)
- Network virtualization software running in compute nodes in the compute blocks (AON)
- Storage software running in storage controllers in the storage blocks (EBF)
- Network interconnect software running on interconnect switches
- Management portal software (ThinkAgile CP Cloud Controller) running in the public cloud

ThinkAgile CP Abstraction

In place of a cluster, ThinkAgile CP supports the concept of a Migration Zone. A migration zone is comprised of a set of compute nodes among which an application instance (or virtual machine) may migrate. Compute nodes (and application instances that run on these nodes) can be assigned categories and tags.

Note: In the ThinkAgile CP platform, we refer to virtual machines (VMs) as application instances.

Categories are used to categorize or classify compute nodes, for application instance placement, and for managing oversubscription. A node can have only one category assigned to it. Tags determine where an application can run and are also used to support application instance placement and migration. Unlike with categories, a node can have many tags assigned to it. Together, categories and tags are called compute constraints. Application instances have compute constraints and nodes have compute constraints. An application instance will only be allowed to run on a node with the specified compute constraints, that is, the specified category and specified tags.

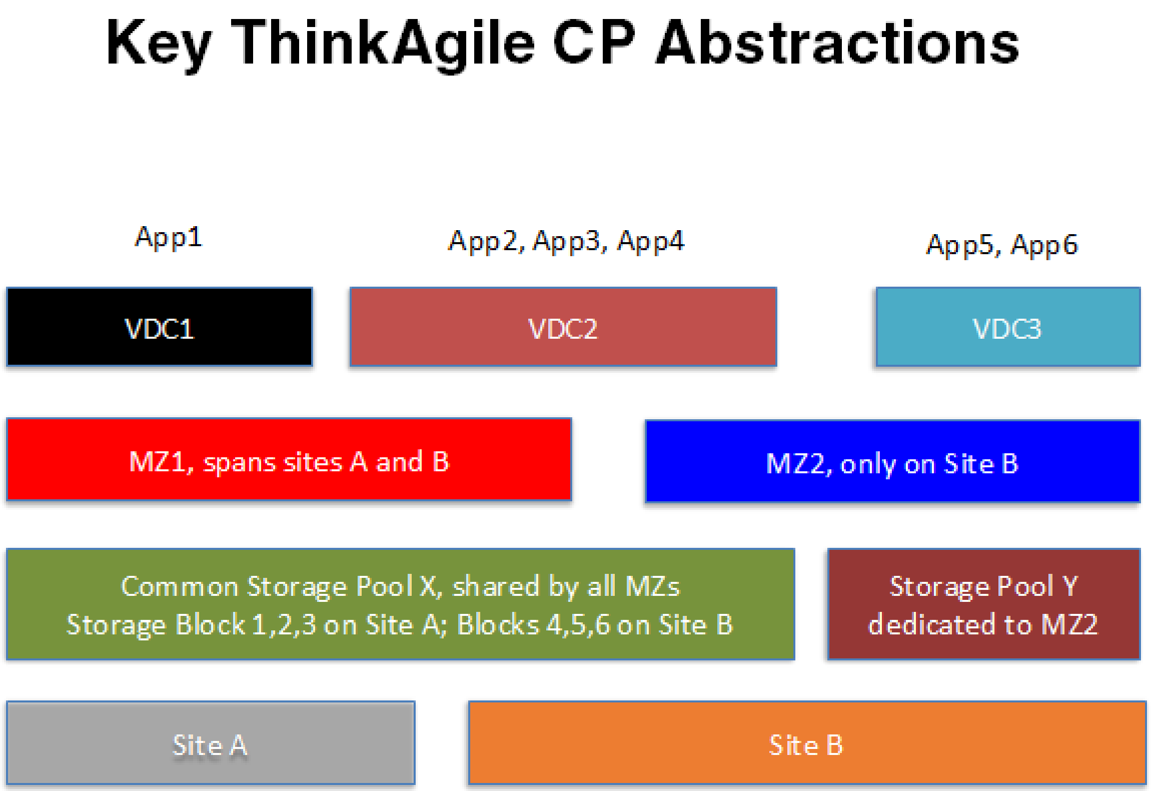
Storage is grouped into storage pools for administration purposes. There is no limit to the size of a storage pool.[1] Storage pools can be shared across all migration zones, but we can also choose to limit the storage pools that are available and accessible to a migration zone. Storage pools consist of storage blocks.[2] A storage block is a pair of storage controllers and all storage accessed through those two storage controllers.

The last abstraction of interest is the notion of a Virtual Datacenter, which is a pool of logical resources assigned to a tenant of the ThinkAgile CP Platform. The logical resources for a virtual datacenter may come from one or more migration zones and one or more storage pools. Application instances are provisioned inside virtual datacenters.

By using migration zones and storage pools instead of clusters, our approach is more scalable, higher performing, more resource-efficient, and easier to manage than traditional approaches than using clustering.

A pictorial view of some of the key abstractions is shown below.

[1] Theoretical limit is 9000 PBs. Practical values range from 32 TB to 256 PB.
[2] This is different than the industry-wide use of the term storage block to mean a 512 byte sector on a storage device.



Key ThinkAgile CP Abstractions

App1

App2, App3, App4

App5, App6

VDC1

VDC2

VDC3

MZ1, spans sites A and B

MZ2, only on Site B

Common Storage Pool X, shared by all MZs
Storage Block 1,2,3 on Site A; Blocks 4,5,6 on Site B

Storage Pool Y
dedicated to MZ2

Site A

Site B

Figure 2. Pictorial View of Key Abstractions

Hardware scaling

ThinkAgile CP supports non-disruptive hardware upgrades and non-disruptive scaling of both compute and storage.

The minimum system is a compute block with two compute nodes, a storage block with two storage controllers and eight SSDs, and either one (non-redundant) or two interconnect switches.[3] Additional compute nodes may be added one node at a time without disruption to existing workloads. ThinkAgile CP automatically detects new nodes. Once the new nodes are connected to the interconnect switches and brought online, they are added to a migration zone. Either new workloads may be initiated on the node, or existing workloads from ThinkAgile CP or other non-ThinkAgile CP platforms may be non-disruptively migrated to the node. When a compute block is full (either four or eight nodes, depending on the hardware used), a new compute block is first added and racked to hold the new node, before following the steps previously described for bringing workloads to the new node.

SSDs may be added to storage blocks in groups of eight SSDs at a time. When the second or later group of eight SSDs are being added to a storage block, ThinkAgile CP automatically restripes data across the additional drives added, and this slightly increases the background load until the restripe completes. Other than the additional background activity,[4] the addition of new SSDs is non-disruptiveto existing workloads.

[3] Each interconnect has 48x10 GbE ports and 4x40 GbE ports.
[4] Performed at a lower priority than user reads and writes.

When a storage block is full (with 24 SSDs or three groups of eight SSDs[5]), a new storage block with eight (or other allowed multiple of eight) SSDs may be added and racked and made part of a storage pool. When storage for new workloads is assigned to this storage pool, space in this new storage block may be used.[6]

Compute nodes and storage blocks are added until all the downlink ports (44x10 Gb) on the internal interconnects are used up. A pair of interconnect switches and all the compute and storage that can be attached to it is called a *stack*. After the stack is full, we will need to add another pair of interconnect switches that starts a new stack, before we can add more compute nodes or storage blocks and SSDs. Stack-to-stack communication occurs using one or more of the 40 GbE uplink ports on the interconnects.

Software upgrades

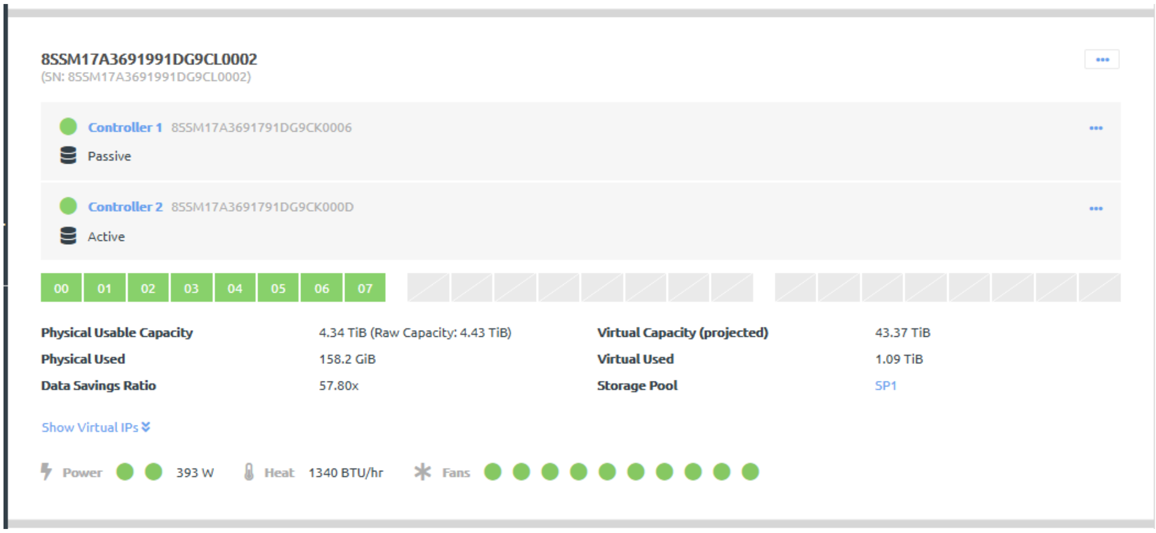
ThinkAgile CP has a software push upgrade capability, therefore, most of our upgrades to new releases of our software can be done remotely and non-disruptively. In some cases, the upgrade will require maintenance windows, but this still can be done remotely.

In rare instances, such as when the disk format or switch configuration changes, the upgrade will be performed on-site as a service offering. Firmware upgrades (for example, BIOS) will be done manually, but most can still be done with minimal disruption, since we have HA across storage, network and hypervisor, as described below. For example, if a compute node needs to be brought down to apply a firmware update, VMs running on that compute node will be migrated to another compute node and migrated back after the update is applied.[7]

We support rolling software upgrades. All nodes in a migration zone can be updated to a new release of software, while other migration zones remain at the previous release.

Storage high availability

The two storage controllers in a storage block operate in active-passive mode, with automatic failover when the failure of a storage controller is detected.



8SSM17A3691991DG9CL0002
(SN: 8SSM17A3691991DG9CL0002)

● Controller 1 8SSM17A3691791DG9CK0006
Passive

● Controller 2 8SSM17A3691791DG9CK000D
Active

Physical Usable Capacity	4.34 TiB (Raw Capacity: 4.43 TiB)	Virtual Capacity (projected)	43.37 TiB
Physical Used	158.2 GiB	Virtual Used	1.09 TiB
Data Savings Ratio	57.80x	Storage Pool	SP1

Power ● ● 393 W Heat 1340 BTU/hr Fans ● ● ● ● ● ● ● ● ● ●

Figure 3. Storage controllers shown in the ThinkAgile CP Cloud Controller

[5] 32 SSDs or four groups of 8 SSDs with Dell hardware.
[6] In the near future, we will also allow for migration of data from existing workloads to enable better load balancing.
[7] Please see our other companion technical brief, *Resource Reservation in ThinkAgile CP*, on how to make sure adequate resources have been reserved to ensure the migrated VMs have enough resources to run on one of the other nodes.

The storage controllers send heartbeats every four seconds to the internal interconnect, which acts as the quorum node. If the internal interconnect does not receive a heartbeat for 32 seconds from the active storage controller, it triggers a storage controller failover. The secondary controller takes over in 60-90 seconds, with no data loss, and no rebuilds needed. Any new or in-flight I/O operations will hang until the failover completes because the iSCSI timeouts are set to be greater than the failover time.[8]

RAID 50 and hot-sparing are used for storage redundancy of the Flash SSDs. Each group of eight SSDs is organized as a 6+P+S RAID-5 array, and data is striped across the multiple[9] RAID-5 arrays in a storage block to make it a RAID-50 array. This protects against the failure of flash drives. When a drive fails, the data from that failed drive is automatically rebuilt and written to the spare drive. The rebuild operation will impose some small background load on that group of eight drives, but user reads and writes are always prioritized higher than the background rebuild operations.

It is important to replace the failed drive as soon as possible.[10] The replaced drive will then become the new hot spare drive for the RAID group. If the failed drive is not replaced, the array is operating without a hot spare - this is not recommended as there will be nowhere to rebuild data to if a second drive fails in this group of drives. If a second drive fails during the multi-hour rebuild process, or if a drive sector is unreadable during the rebuild process, which is a double failure, there will be a data loss. In this case, to recover the lost data, the customer will need to go back to a prior quick DR backup on a different storage pool.

In addition to ECC, each Flash drive further uses erasure coding to handle uncorrectable errors within the drive. If an error is not correctable by either the ECC or the erasure coding on the drive, we will use the RAID to recover that data.

The platform supports application-consistent backups[11] and clones. Both scheduled backups and user-generated backups are supported on the platform. If a VM gets damaged or corrupted due to any reason, such as a software crash or bug or a virus or other form of security attack, it is possible to simply revert back to a previous backup; or, you can start a new VM against a backup. Of course, any updates since the most recent backup will be lost.

If an operator error causes a single file to be mistakenly deleted, we allow a single file to be restored from a backup.[12]

The ThinkAgile CP platform supports disaster recovery (DR). VMs provisioned in virtual datacenters with access to more than one storage pool can turn on disaster recovery. This will allow the VM backups to be sent to a storage pool (possibly at a remote physical site) different than the one that holds the VM vDisks.

If the DR storage pool is at the same site, it can be used to recover from the double failures described previously. If the DR storage pool is at a secondary site, then it can be used to recover from disasters such as earthquakes, hurricanes, floods, and so on, which affect everything at a primary site. In this case, the backups at the secondary site can be used to recover and continue the IT operations. The time to recover at the secondary site, called the Recovery Time Objective (RTO) can be small; the amount of data lost when disaster strikes, called the Recovery Point Objective (RPO), can be as small as the last minute's worth of data updates.

For more information about setting up quick DR backups, see the following topic:

https://thinkagile.lenovofiles.com/help/topic/thinkagile_cp/manage-quick-dr-backups.html

[8] Specifically, they have been set to 20 minutes, far greater than the 1-2 minutes failover time.

[9] Three for Lenovo ThinkAgile CP.

[10] In many, but not all, cases, ThinkAgile CP Support can tell that a drive has failed and ship the customer a replacement drive automatically.

[11] For some legacy operating systems, we support crash-consistent, but not application-consistent backups.

[12] This is done by cloning and attaching a disk from the backup to another VM.

Compute node or hypervisor high availability

Each ThinkAgile CP compute node, storage controller, and network interconnect communicates with the management portal every 60 seconds, sending regular metadata updates on statistical information, such as CPU usage, memory usage, and power usage by node and application instance.

The portal can also use this to compute usage by virtual datacenter. The statistical information is aggregated and averaged to provide data points for graphical display. Information reported by a storage controller, on a per compute node basis, includes the iSCSI active session counts. The portal detects a node failure when all the iSCSI active session counts for a particular node go to zero. It can then initiate the action to restart application instances from that node onto one or more other nodes in the same migration zone that satisfy the compute constraints. This is called hypervisor high availability (HA).

Network High Availability

The network part of the ThinkAgile CP solution consists of redundant SDN-enabled ONIE (Open Network Install Environment) interconnects, each with 48x10 GbE and 6x40 GbE ports, running the Pica8 NOS (PicOS network operating system). We call these internal interconnects, since only servers and storage internal to ThinkAgile CP can connect directly to these ports.

These two internal interconnects are configured as MLAG peers. Forty-four of the 10 GbE ports are used to connect to our compute nodes and storage controllers. The remaining 10 GbE ports are used for the management network that ties all the storage compute and network blocks together. Two of the 40 GbE ports are used to connect the two internal routers to each other for MLAG traffic. At least one of the 40 GbE ports is used to connect into the customer data center. The remaining three may be used as needed for handling traffic between one interconnect pair and another interconnect pair, or for additional connections into the customer data center.

Each compute node and each storage controller are connected to both interconnects using separate NIC cards. This design protects against the failure of a NIC or of an interconnect.

The two interconnect switches operate in a primary and secondary role. If there is a need to fail over from the primary interconnect switch to the secondary interconnect switch, see the following topic:

https://thinkagile.lenovofiles.com/help/topic/thinkagile_cp/make_secondary_interconnect_primary.html

High availability support

ThinkAgile CP provides proactive support with ThinkAgile CP Early Insights. Our health monitoring support system automatically alerts the ThinkAgile CP support team of potential issues and failures. We can often see failures and initiate replacement service requests before the customer even knows there is a problem. A chat interface allows a customer to connect directly with the support team. Support mode allows us to help the customer directly without needing a technician to go on-site. The customer can easily disable the support mode at any time.

Non-disruptive migration of existing application instances

For customers interested in non-disruptively migrating their current applications from other hypervisors, such as VMware into ThinkAgile CP, we provide the ThinkAgile CP Migration Manager.

For more information about Migration Manager, see Introduction to the Migration Manager.

Minimizing catastrophic failures in ThinkAgile CP

Instead of clusters, we use the notion of a migration zone (MZ) in ThinkAgile CP. A migration zone is a set of compute nodes among which an application (VM) may migrate.

Unlike clusters, the number of nodes in a migration zone can be in the 1000s. Adding a node to a migration zone is very simple, simpler than adding a node to a cluster[13].

We have designed migration zones in such a way that the likelihood of a catastrophic failure that takes out an entire migration zone is even smaller than the probability of losing an entire cluster in clustered designs. Metadata is distributed across the storage controller pairs and the SaaS portal. Loss of both storage controllers can result in the localized outage of that storage block, but not a migration zone-wide outage as the other storage blocks are still operational. Portal metadata about an migration zone is stored across multiple availability zones and backed up every four hours. It would take a catastrophic failure across two availability zones of a public cloud provider like AWS for ThinkAgile CP to lose portal metadata related to migration zones. Even if that happens, we can still recover migration zone metadata from the backup. We may lose up to four hours worth of UI actions (such as create a VM, create a vDisk, create a new firewall rule, and so on), but we would not lose all customer metadata even in this catastrophic scenario, only those related to vDisks created in the last four hours.

[13] When a node is added to a migration zone, other nodes do not need to know. In contrast, cluster implementations require that every node in a cluster needs to know about every other node in the cluster, and high frequency heartbeats between nodes in a cluster, typically using a private cluster network, is used to maintain the cluster.

Lenovo