



EQUALLOGIC

TECHNICAL SUPPORT INSTRUCTIONAL SERIES

Technical Report	Document ID: N/A
PS Series Groups	Revision: A02
Sizing Replication Space	Revision Date: August 28, 2009

PS Series Groups

Sizing Replication Space

Abstract

This Technical Report describes replication technology from EqualLogic and how to size the space needed on the primary and secondary groups. It describes how replication space is utilized and PS Series group behavior when replication space is consumed.

Copyright © 2005 EqualLogic, Inc.

April 2005

EqualLogic is a registered trademark of EqualLogic, Inc.

All trademarks and registered trademarks mentioned herein are the property of their respective owners.

Possession, use, or copying of the documentation or the software described in this publication is authorized only under the license agreement specified herein.

EqualLogic, Inc. will not be held liable for technical or editorial errors or omissions contained herein. The information in this document is subject to change.

This document is EqualLogic company confidential.

PS Series Firmware Version 2.0 or later.

EqualLogic, Inc.
9 Townsend West
Nashua, NH 03063

Tel: 603.579.9762
Fax: 603.579.6910

Table of Contents

Introduction	1
Replication Operation	1
Replication Space Utilization.....	3
Minimizing Replication Completion Times.....	6
Minimizing Replication Space.....	7
Local Reserve in Depth.....	8
Sizing the Local Reserve	9
Local Reserve Sizing Examples	10
Handling Depleted Local Reserve.....	11
Recovering Local Reserve.....	11
Affect of Volume Size Increase on Local Reserve	12
Replica Reserve and Delegated Space in Depth	12
Sizing Delegated Space	13
Sizing the Replica Reserve.....	13
Replica Reserve Sizing Examples	15
Handling Depleted Replica Reserve.....	16
Recovering Replica Reserve.....	17
Affect of Increasing the Volume Size on the Replica Reserve	17
Documentation and Customer Support	18

This Technical Report describes replication technology from EqualLogic and how to size the space needed on the two PS Series groups that are configured as replication partners. With replication, volume data can be automatically transferred across the groups, protecting volume data from serious failures—ranging from the destruction of the volume to a complete site disaster—with no impact on data availability or performance.

Replication requires you to reserve space on the **primary group**, where the base volume resides, and on the **secondary group**, where the replicas are stored. The amount of replication space needed depends on a number of factors:

- Volume size.
- Amount of changed volume data during each replication data transfer.
- Amount of changed volume data between consecutive replication operations.

To ensure successful and effective replication, you *must* reserve sufficient space on both groups. EqualLogic chose the default values for replication space to ensure that replication operations always complete successfully, regardless of data change rates. Decreasing the default replication space values can save disk space, but may cause replication operations to pause or fail due to lack of space.

The key to properly setting replication space values is to understand the volume change rate. However, this can be difficult because workloads vary over time. If you do not want to use the default values, you must calculate appropriate values, monitor the space utilization, and, if necessary, adjust the values as your workload changes.

There is no easy way to predict the amount of volume changes (writes) that will occur during a replication data transfer and between consecutive replications. The rate of volume changes depends on the application using the volume and the volume workload. Although some applications make a consistent number of volume writes, others have a workload that varies from day to day. Also, for volumes with mainly random writes, you may need more replication space than the actual volume changes suggest.

The information in this Technical Report can help you estimate the amount of replication space that you need to reserve in order to set up successful and effective volume replication.

Replication Operation

Similar to a snapshot, a **replica** represents the contents of a volume at the point in time that the replication started. However, unlike a snapshot, a replica is located in a group separate from the group where the **base volume** resides. Because the volume and the replicas are physically separated, a complete disaster at one site does not destroy the data at the other site.

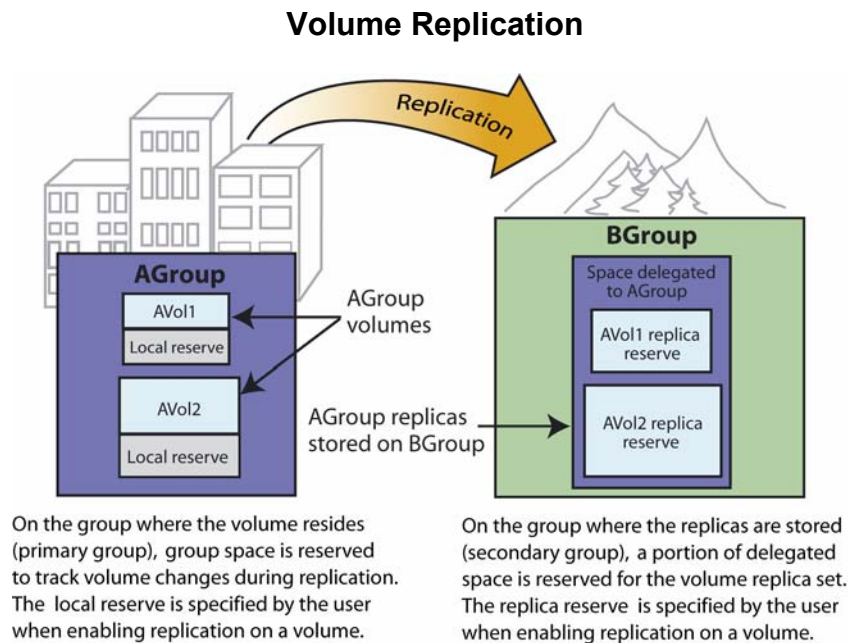
When you create a replica, either manually, by script, or by using a schedule, the first replica of a volume is a *complete transfer* of the volume data that existed when the replication started. For subsequent replicas, only the data that changed since the previous replica started (that is, the deltas or differences between the volume contents) is transferred to the secondary group. This ensures optimal use of network infrastructure between the primary and secondary group.

The **replica set** includes all the completed replicas for a volume. The **replica volume** (not visible to the user, but equivalent to the most recent complete replica) ensures that the secondary group *always preserves a complete, stable copy of the volume data*—the key goal of EqualLogic replication technology.

For example, after the first replication finishes (a complete volume data transfer), the replica volume is created. After the second replication finishes (only the changed volume data is transferred), the second replica becomes the most recent complete replica. Now, the replica volume is formed from the first replica and the changed data from the second replica. As additional replicas are created, this process is repeated.

The replica volume is *never* overwritten until all the data has been successfully copied from the primary group to the secondary group. This ensures that a network outage will not compromise the stable volume data represented by the replica volume. In other words, EqualLogic ensures that you go from one “good” replica to another “good” replica on the secondary group.

The following figure shows replication from one group to another. Note that you can also have reciprocal replication, in which AGroup replicates volumes to BGroup, as shown, but BGroup replicates volumes to AGroup.



To recover volume data, the secondary group administrator can **clone** a replica or **promote** a replica set. Even if you accidentally delete all the individual replicas for a volume, the replica volume will still exist, and you can still promote the replica set.

The only way to delete the replica volume (and recover replication space on the secondary group) is to delete the replica set, which also disables replication for this volume. If you later re-enable replication on the volume, the next replication will be a complete transfer of all the volume data.

Replication Space Utilization

Replication requires space on the primary group, where the base volume resides, and on the secondary group, where the volume data is transferred and the replicas are stored. The primary group administrator specifies replication space values when enabling replication on a volume and can modify these values at any time.

Replication space includes the following:

- **Local reserve.** On the primary group, the local reserve (5% to 100% of the volume size) is used to preserve the state of the volume while a replication data transfer is in progress.

If volume writes occur during a replication data transfer, the original (unmodified) volume data is stored in the local reserve, progressively consuming space. Once the data transfer completes, any volume data in the local reserve is released, and the local reserve is not used again until volume writes occur during the *next* replication data transfer.

How much local reserve you need for any given replication operation depends on the volume changes, which can vary for each replication and are based on the workload and the time it takes for the data transfer to complete. For example, if no volume writes occur during a replication data transfer, no local reserve is consumed (although the space is still reserved from free group space). As volume writes occur, the local reserve is progressively consumed. For example, if the entire volume is re-written, 100% of the volume size will be needed in the local reserve. When the data transfer completes, utilization of the local reserve returns to zero.

Because the local reserve requirement can vary for each replication, you may want to size the local reserve conservatively. If the local reserve is not big enough to store volume data during a replication data transfer, as needed, the in-progress replication will fail. Specifying a value of 100% for the local reserve ensures that a replication will never fail because of a lack of local reserve.

If you want to specify a local reserve that is less than 100%, you should enable the feature that allows free group space to be temporarily *borrowed* if the size of the local reserve is not adequate. Once the data transfer is complete, any utilized group space is freed. Note that if the size of the local reserve is insufficient *and* free group space is not available, the in-progress replication will fail.

To summarize the local reserve considerations:

- Set the local reserve to the default (100%) to ensure that the size of the local reserve will be sufficient and a replication will never fail because of lack of local reserve.
- Set the local reserve to less than 100% (minimum 5%) if you are certain that the volume changes during replication will be consistently less than 100% of the volume size. However, if the local reserve is insufficient, the in-progress replication will fail.
- Set the local reserve to less than 100% and enable the feature that lets you temporarily borrow free group space when the local reserve is too small. However, if the local reserve is insufficient and there is insufficient free group space, the in-progress replication will fail.

- **Replica reserve.** On the secondary group, the replica reserve is a portion of the space that the secondary group administrator delegated to the primary group. Delegated space is used to store the volume replica sets from the primary group.

You can think of the replica reserve for a volume as having two parts:

- The initial replication of a volume is a complete transfer of the base volume data and consumes 100% of the base volume size. This is the first part of the replica reserve, which stores the most recent complete replica (the replica volume), consuming 100% of the base volume size. At the completion of each replication, the replica volume will always represent a stable copy of the most recent volume data.
- Each replica subsequent to the first is a data transfer consisting only of the deltas (differences) between the volume contents at start of the previously completed replication and the volume contents at the start of the current replication. The second part of the replica reserve is used to store the deltas while the data transfer is on-going and also to store complete replicas other than the most recent.

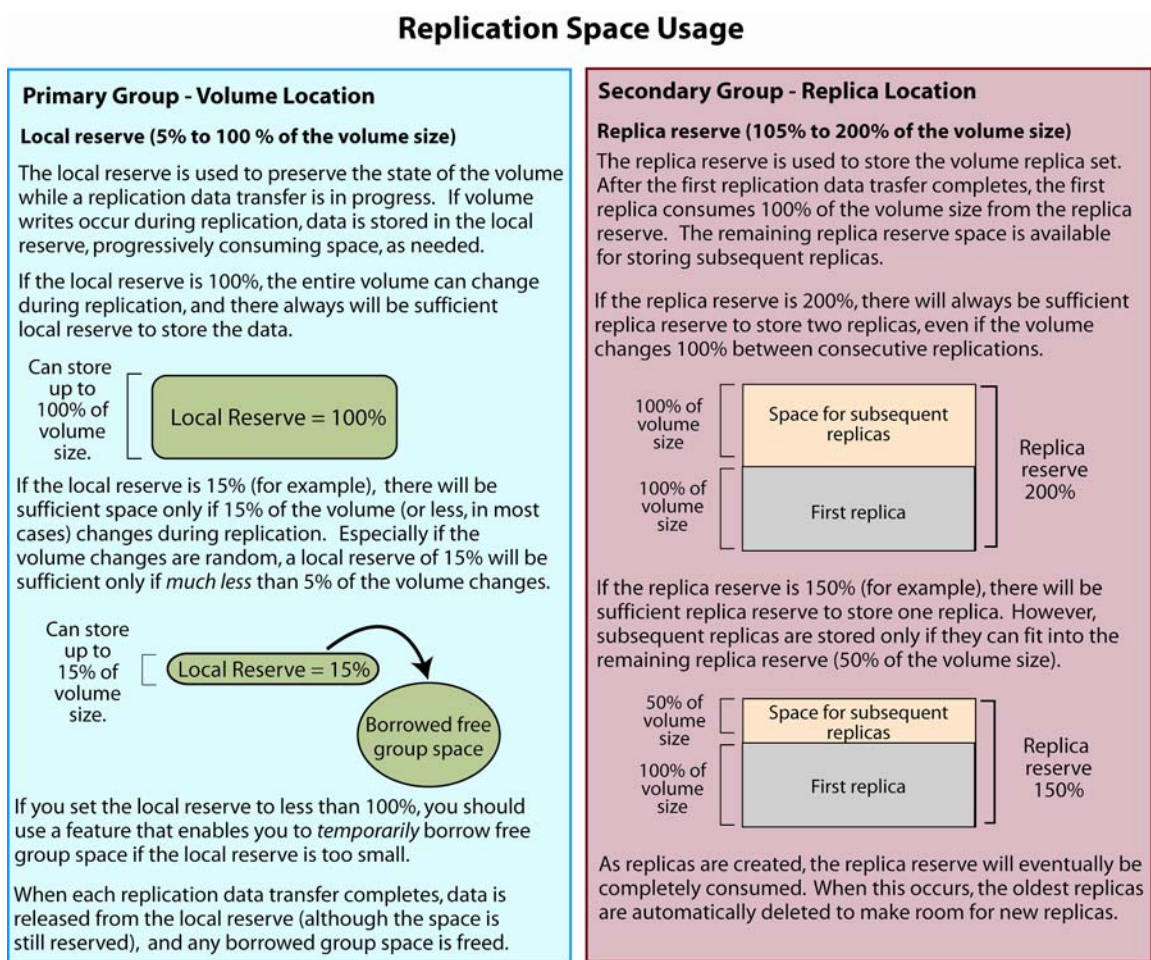
When the second part of the replica reserve is consumed, the oldest replicas are automatically deleted to make room for new replicas. If all the replicas except the most recent are deleted, but the size of the second part of the replica reserve is still not big enough to store the deltas for a replication operation, the data transfer will pause until the administrator increases the replica reserve. The goal is to keep replica reserve space needs as low as possible.

The second part of the replica reserve is the most difficult to size because of workload variations and replication scheduling. The minimum space is 5% of the volume size.

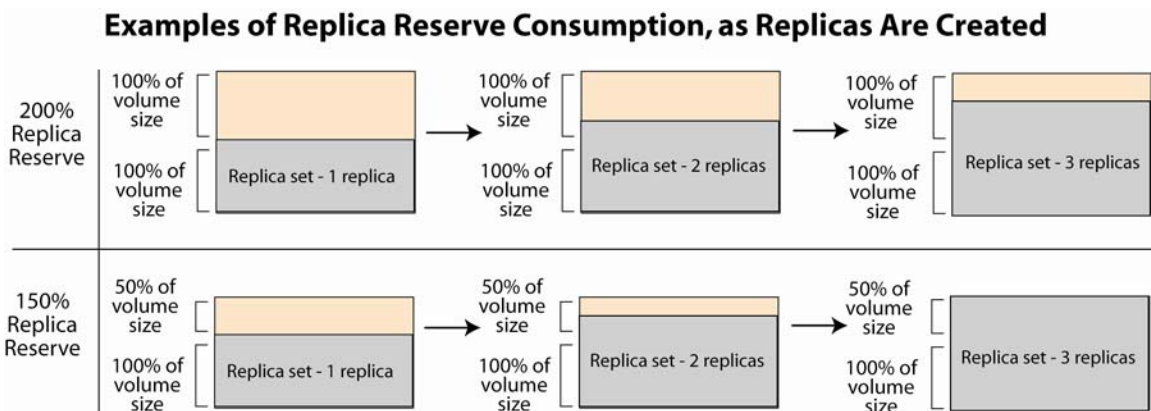
These two parts added together comprise the replica reserve, totaling a minimum of 105% of the volume size to a percentage limited only by the amount of delegated space. Essentially, the replica reserve should be big enough to store the initial complete data transfer, in addition to the volume changes for subsequent replicas. For example, specifying 200% for the replica reserve (the default) ensures that at least two complete replicas can be stored on the secondary group, with no user intervention. Specifying less than 200% percent ensures that there is enough space to store the initial replica, but subsequent replicas may or may not be able to fit in the remaining space, depending on the amount of volume changes that are transferred.

You may want to be conservative when sizing the replica reserve. Insufficient replica reserve will result in the replication operation being paused and a message informing the administrator that the replica reserve must be increased to a specific value in order to complete the data transfer. If you cannot increase the replica reserve, the paused replication operation will not be able to complete, and you may have to manually cancel the paused replication operation and re-plan your overall replication design.

The following figure shows how replication space is used on the primary and secondary groups.



The following figure shows how replica reserve is consumed as replicas are created.



Minimizing Replication Completion Times

Replication moves volume data from one location to another for disaster protection. It is important to design a replication environment that enables you to reliably perform replication operations.

A key performance factor in replication completion times is the available network bandwidth between the primary group and the secondary group. A successful replication design will take into account a number of factors, including volume sizes, change rates, replication frequency, and network bandwidth to ensure predictable and reliable operations.

For the first replication of a volume, the time that it takes to complete the data transfer from the primary group to the secondary group depends on the amount of volume data being replicated. The more data that needs to be transferred, the longer the replication will take.

The time that it takes to complete the data transfer for *subsequent* replicas depends on the amount of volume data that changed since the previous replication (that is, the deltas or differences between the volume contents at the time the last successful replication started and the volume contents at the time the current replication started). For example, if all the volume data is rewritten, the replication will take as long as the initial data transfer. If no data was changed, the replication will be near instantaneous because no data needs to be transferred.

If a replication transfer takes a long time (for example, because of slow network links), the amount of volume changes during the replication will probably increase; therefore, more local reserve on the primary group may be needed.

If a replication is still in progress when the next replication of the volume is scheduled to start, the scheduled replication will not start, and you will see a message stating the problem. When a replication is skipped, more data will probably need to be transferred in the next replication; therefore, more replica reserve on the secondary group may be needed.

To minimize replication times, follow these recommendations:

- Use a reliable, adequately sized network link between the primary and secondary groups.
- Consider using network prioritization to ensure adequate network throughput if the network is shared with other users. This would entail adjusting the priority of iSCSI traffic (port 3260) to meet your overall needs.
- Adjust volume replication schedules to better match data protection needs with available network bandwidth. This could include staggering replication operations so that volumes with high data change rates are provided with sufficient network bandwidth.

The following table describes the possible network bandwidth between groups.

Network Link Bandwidth for Different Network Configurations

	T1	T3	OC3	OC12	GbE	OC48	OC192	10 GB
Mb/sec	1.5	45	155	622	1000	2488	9952	10000
Usable Megabits transferred in one second	1.5	45	140	564	1000	2254	9017	10000
Usable Megabytes transferred in one second	0.19	5.63	17.55	70.44	125.00	281.77	1127.06	1250.00
Gigabytes transferred in one hour	0.6	17.8	57.0	228.8	439	915.1	3962	4394

Minimizing Replication Space

Correctly sizing replication space is crucial to successful and effective volume replication.

- Insufficient local reserve in the primary group (and free group space, if the “borrow free group space” feature is enabled) will cause the replication to fail. As stated in *Minimizing Replication Completion Times*, long replication transfer times increase the possibility that volume writes will occur during the data transfer and thus increase the need for local reserve.
- Insufficient replica reserve in the secondary group will result in the replication pausing and a message informing the administrator that replica reserve space must be increased in order to complete the replication. As stated in *Minimizing Replication Completion Times*, if a replication is cancelled or fails, the next replication will likely require the transfer of more data, which will increase the amount of replica reserve needed to store the data.

A simple approach for reliable operations is to use the default space values (100% for the local reserve and 200% for the replica reserve). These values guarantee the following:

- Replication will not fail due to lack of local reserve or replica reserve space.
- At least the two most recent replicas will be retained on the secondary group, without user intervention.

If you want to use less than the default values, determining adequate sizes for the local reserve and the replica reserve requires you to use a monitor and adjust approach. This is because it can be difficult to predict the amount of volume data changes during a replication data transfer (a local reserve consideration) and between consecutive replications (a replica reserve consideration). The volume’s workload and application utilization and the replication completion time may vary daily.

When sizing the local reserve and replica reserve, you must consider:

- **Rate of changes.** The rate of volume changes is specific to the application using the volume and the workload on the volume. Although some applications make a consistent number of volume writes, other applications have a workload that varies. For example, a volume may incur 10GB of changes one day and 2GB the next day.
- **IT operations.** Some disk operations can affect replication. For example, defragmenting a disk or performing a database reorganization on a volume that *already* has been replicated may require a large transfer of data for the next replication. However, subsequent replications may transfer data more efficiently than before the disk operation.
- **Locality of change.** There may be no direct correspondence between changes to the base volume and the amount of replication space needed. Especially for random volume writes, you may find that you need relatively more local reserve and replica reserve than expected, based on the actual volume changes. For example, 2GB of volume writes could require as much as 3GB of replication space or as little as 1GB, depending on how data is distributed on the file system.

In some cases, you can adjust IT operations and replication schedules to meet your overall replication goals within the existing infrastructure, using scripts if necessary. For example, you could temporarily adjust replication space to handle replication after a defragmentation or reorganization, a large workload, or a restore.

The following table summarizes local reserve considerations and provides an example for a 10GB volume. See *Local Reserve in Depth* for more information on sizing the local reserve.

Primary Group Local Reserve Considerations	10GB Volume Example
Used only if volume writes occur <i>during</i> a replication data transfer. The exact space needed depends on the amount and type of volume changes, which may not be consistent for each replication.	<p>If the base volume is 10GB, and the maximum amount of volume changes during data transfer is 5% of the volume size, the local reserve must be able to store .5GB of data.</p> <p>Set the local reserve to a minimum of 5% and enable the “borrow free group space” feature. Monitor and adjust space as needed.</p>

The following table summarizes replica reserve considerations and provides an example for a 10GB volume. The example assumes that a goal of your disaster recovery plan is to retain two of the most recent replicas on the secondary group, and volume changes are *consistent* between consecutive replications. See *Replica Reserve and Delegated Space in Depth* for more information.

Secondary Group Replica Reserve Considerations	10GB Volume Example
The most recent replica (replica volume) always consumes 100% of the base volume size from the replica reserve.	If the base volume is 10GB, the first part of the replica reserve must be able to store 100% of the volume size (10GB).
<p>Additional space is needed to store replicas other than the most recent. Note that only the deltas are transferred for replication operations other than the first. The exact space needed depends on the amount of volume changes between consecutive replications, which may not be consistent.</p> <p>To <i>roughly</i> determine how much space you need, multiply the amount of volume changes between consecutive replications by the number of replicas you want to retain. If volume changes are not consistent, you <i>may</i> be able to use an average to start.</p>	<p>Assume the amount of volume changes between consecutive replications is consistently 25% of the volume size (2.5GB).</p> <p>To <i>roughly</i> determine the space needed to store replicas other than the most recent (the second part of the replica reserve), use this formula:</p> $25\% \times 2 = 50\% (5GB)$
Add the above space requirements to determine the total size of the replica reserve.	<p>Set the value of the replica reserve to at least:</p> $100\% + 50\% = 150\% (15GB)$ <p>Monitor and adjust space as needed.</p>

Local Reserve in Depth

On the primary group, where the base volume resides, the local reserve (5% to 100% of the volume size) is used to preserve the state of the volume while a replication data transfer is in progress. You can think of the local reserve as space that only needs to store one snapshot. Note that this is in addition to the regular snapshot space you reserve for volume snapshots.

If volume writes occur during a replication data transfer, the original (unmodified) volume data is stored in the local reserve (and possibly free group space if the “borrow free group space” feature is enabled), progressively consuming space. Once the data transfer completes, any volume data in the local reserve (and any utilized group space) is released, and the local reserve is not used again until volume writes occur during the *next* replication data transfer.

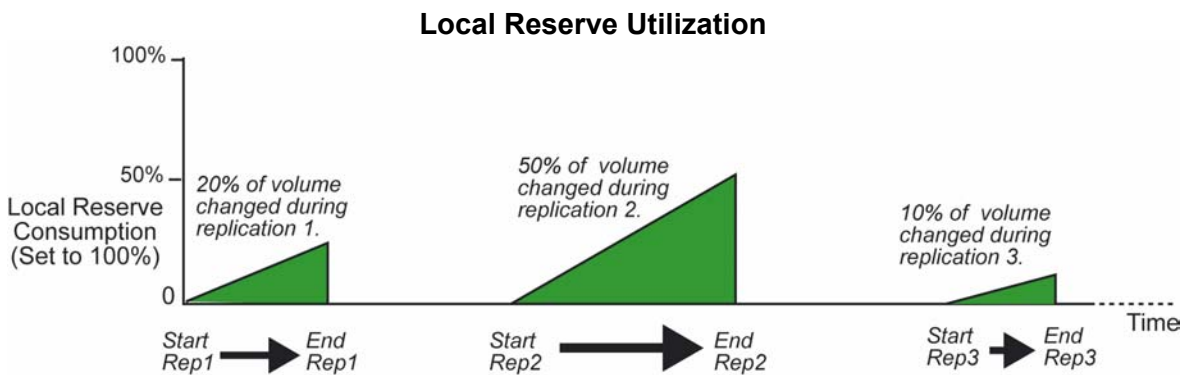
How much local reserve you need for any given replication operation depends on the volume changes, which can vary for each replication and are based on the workload and the time it takes for the data transfer to complete. If there is insufficient local reserve (and no free group space, if this feature is enabled), the replication will fail, data transfer will stop, and an event message will be logged. In addition, if free group space falls below a threshold, all replication data transfers that are relying on borrowed group space will fail.

The primary group administrator specifies the local reserve when enabling replication on a volume, and the space is immediately deducted from free primary group space. The administrator can increase or decrease the local reserve as needed, if the primary group has sufficient free space.

Sizing the Local Reserve

You can specify 5% to 100% (the default) of the volume size for the local reserve. The appropriate value for the local reserve depends on the amount by which you expect a volume to change during replication data transfer.

However, as the following figure shows, the amount of volume writes during each replication may not be consistent. This is because the volume change rate can vary due to workload and IT operations. Also, there may be no direct correspondence between changes to the base volume and the amount of local reserve needed. Especially in cases of random volume writes, you may find that you need relatively more local reserve than expected, based on the actual volume changes.



To *roughly* determine how much local reserve you need for a volume, you may want to initially size the local reserve so that it can handle the largest amount of volume changes that will occur during a replication data transfer. In the previous figure, this would be 50% of the volume size. If free group space is available, also enable the “borrow free group space” feature. Then, as needed, adjust the value up or down before arriving at the appropriate size for the local reserve.

The following table describes the benefits and disadvantages of setting the local reserve to 100%.

Local Reserve Set to 100%	
Benefit	Replication data transfer will never fail due to lack of local reserve, even if all the volume data is re-written during the replication.
Disadvantage	Replication uses more space than may be necessary on the primary group.
Solution	Experiment with decreasing the local reserve. Be sure to use the option that enables the borrowing of free group space if the size of the local reserve is insufficient.

The following table describes the benefits and disadvantages of setting the local reserve to less than 100%.

Local Reserve Set to Less Than 100%	
Benefit	Replication uses less space on the primary group.
Disadvantage	If many volume writes occur during a replication (for example, if the volume is under a heavy workload or the replication takes a long time to complete), the size of the local reserve may be insufficient and the replication will fail.
Solutions	<ul style="list-style-type: none"> • Size the local reserve so that it can handle the largest amount of volume changes that will occur during a replication data transfer. You may need to experiment with different values. See <i>Local Reserve Sizing Examples</i> for information. • Be sure to use the option that enables the borrowing of free group space if the size of the local reserve is insufficient. • Use a high-performance, reliable network link between the primary and secondary groups. The longer a replication takes, the greater the chance that volume changes will occur during the replication, increasing the need for local reserve. • Stagger replication operations and schedule them during times of light workloads, so that free group space can be efficiently used.

Local Reserve Sizing Examples

Only by using a local reserve of 100% will you ensure that a replication does not fail due to lack of local reserve. If you want to set to the local reserve to less than 100%, be sure to enable the “borrow free group space” feature.

As an example, assume that the local reserve is set to 10% on the primary group, and the “borrow free group space” feature is *not* enabled. In this case, the local reserve itself will be sufficient only if the volume writes during the replication data transfer equal 10% of the volume size (or less in most cases, because there may be no direct correspondence between changes to the base volume and the amount of local reserve needed). If more than 10% of the volume changes, replication will fail.

Again, assume that the local reserve is set to 10%, but the “borrow free group space” feature is enabled. In this case, if more than 10% of the volume changes during the data transfer, the replication will succeed if there is enough free group space to store the data above what the local reserve can hold. For example, if 15% of the volume changes and free group space equal to 5% or more of the volume size is available, the replication will succeed.

You may need to experiment with different values, adjusting the local reserve up or down, before finding the appropriate value. In addition, local reserve needs can be minimized by ensuring that replication operations complete in a timely manner. See *Minimizing Replication Completion Times* for more information.

Handling Depleted Local Reserve

If the local reserve is insufficient, there are several scenarios to consider:

- If there is insufficient local reserve and the “borrow free group space” feature is enabled (and free group space is available), the replication will continue and you will receive the following INFO message:

```
Replication for volume VolD has exceeded the local replica reserve space of 25 percent. The volume is allowed to exceed the local replica reserve space. Group free pages will be used for local replica reserve till the group free page threshold is reached. At that time replication will be cancelled. To prevent replication from being cancelled, please increase the local replica reserve space for the volume.
```

- If there is insufficient local reserve and the “borrow free group space” feature is not enabled, the replication will fail, data transfer will stop, and an ERROR message will be logged, as follows:

```
Replication for volume VolB has been cancelled because the local replica reserve space has been exceeded. The volume is currently not allowed to exceed the local replica reserve.
```

- If there is insufficient local reserve and the “borrow free group space” feature is enabled but there is not enough free group space, the replication will fail, data transfer will stop, and an ERROR message will be logged, as follows:

```
Replication for volume VolC has been cancelled because the local replica reserve space has been exceeded. The volume is currently allowed to exceed the local replica reserve but the group free space is below the threshold.
```

If a replication fails due to lack of local reserve space, you may be able to resolve the problem, as follows:

- Reevaluate your local reserve value and increase it as needed. Free group space must be available in order to increase the local reserve.
- Enable the feature that allows you to borrow free group space when the size of the local reserve is insufficient.
- If free group space is limited and multiple replication operations are “borrowing” group space, you may want to stagger the replications so they are not all competing for group space at the same time.

Recovering Local Reserve

Circumstances may arise in which you need to decrease the local reserve to return space to the free group space pool. You can decrease the local reserve for a volume *while* replication is in progress only if the new value is *more* than what is currently being consumed.

Affect of Volume Size Increase on Local Reserve

Because the local reserve is based on the volume size, the size of the local reserve will automatically increase if you increase the size of the volume, thus consuming more primary group space. For example, if the size of a volume is increased to 10GB, and the local reserve is set to 50% of the volume size, the local reserve will increase to 5GB. If you increase the size of the volume to 16GB, the local reserve will increase to 8GB.

Before increasing the size of a replicated volume, be sure that the primary group has sufficient space to handle the automatic local reserve increase. If you use the GUI to increase the volume size, the Space Utilization table will show the new value for the local reserve and indicate if there is not enough primary group space. The resize operation will fail if there is insufficient space.

Replica Reserve and Delegated Space in Depth

The secondary group stores volume replica sets from the primary group in the space that the secondary group administrator delegated to the primary group. This delegated space can *only* be used to store replicas from the primary group. Delegated space is immediately deducted from free space on the secondary group. If there is sufficient free space, the secondary group administrator can increase the delegated space at any time. You cannot decrease delegated space below the amount that is currently being used to store replica sets.

When configuring a volume for replication, the primary group administrator allocates a portion of the delegated space, the replica reserve, to be used to store the volume's replica set. The replica reserve can range from 105% to an unlimited percentage of the volume size. The default is 200%.

Replica reserve space utilization proceeds as follows:

1. After the first replication finishes (a complete volume data transfer), the replica volume is created, consuming 100% of the volume size from the replica reserve. Note that the replica volume is always a stable version of the most recent complete replica.
2. After the second replication finishes (only the volume data that changed between the start of the first replication and the start of the second replication is transferred), the second replica becomes the most recent complete replica. Now, the replica volume is formed from the first replica and the changed data from the second replica.

Note that the replica volume is *never* overwritten until all the volume data has been successfully copied from the primary group to the secondary group. This ensures that a network outage will not compromise the stable volume data represented by the replica volume.

3. The procedure described in Step 2 is repeated. As replicas are created, the replica reserve is gradually consumed. Once the replica reserve is exhausted, the oldest replicas are automatically deleted to make room for new replicas. However, the most recent complete replica (and thus the replica volume) is never deleted.

If there is insufficient replica reserve to complete a replication—even if *all the replicas but the most recent are deleted*—you will receive a message that the replication has been paused and the amount to which you must increase the replica reserve to complete the data transfer.

If there is free delegated space, the primary group administrator can increase the replica reserve at any time. You cannot decrease the replica reserve below the amount that is currently being used to store replicas.

Sizing Delegated Space

The amount of space that the secondary group administrator must delegate to a primary group depends on the replica reserve requirement for each replicated volume. First, calculate the replica reserve needed for each volume. Then, add up the replica reserve requirements for the volumes to determine approximately how much delegated space is needed.

Keep in mind that you should always delegate *more* space than is minimally required. This will enable you to increase a volume's replica reserve if this action is needed to complete a replication.

Sizing the Replica Reserve

Sizing the replica reserve can be difficult. As described previously, it can be useful to think of a volume's replica reserve as consisting of two parts:

- The initial replication of a volume is a complete transfer of the base volume data and consumes 100% of the base volume size. This is the first part of the replica reserve, which stores the most recent complete replica (the replica volume), consuming 100% of the base volume size. At the completion of each replication, the replica volume will always represent a stable copy of the most recent volume data. For example, when the first replica completes, the replica volume is formed from the first replica. When the second replica is created, it replaces the first replica as the most recent replica, and thus the second replica becomes the replica volume.

Note that 100% of the volume size will be consumed for the initial replication data transfer, even if the file system on the volume is not completely populated with data. However, the amount of data in a volume will affect the data transfer time. Deleting files from a volume will not decrease the consumption of space in the replica reserve.

- Each replica subsequent to the first is a data transfer consisting only of the deltas (differences) between the volume contents at start of the previously completed replication and the volume contents at the start of the current replication. The second part of the replica reserve is used to store the deltas while the data transfer is on-going and also to store complete replicas other than the most recent. The minimum space you can specify for the second part of the replica reserve is 5% of the volume size.

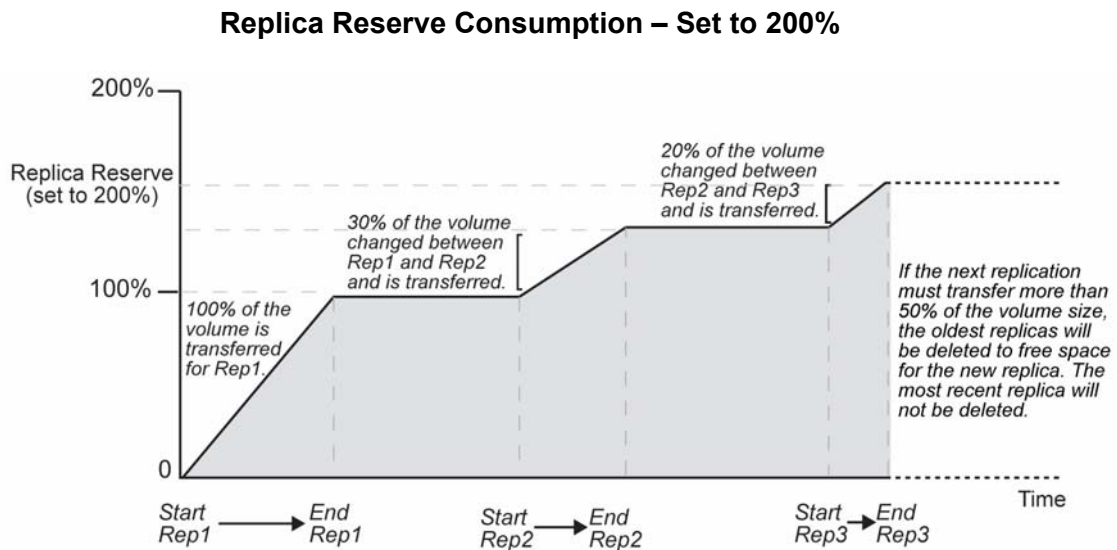
When the second part of the replica reserve is consumed, the oldest replicas (except the most recent) are automatically deleted to make room for new replicas. If the size of the second part of the replica reserve is still not big enough to store the deltas for a replication operation, even if all the replicas except the most recent are deleted, the data transfer will pause until the administrator increases the replica reserve to a value indicated in an event message.

The second part of the replica reserve is the most difficult to size because of the following:

- Rate of volume changes between consecutive replications. The rate of changes is specific to the application using the volume and the workload on the volume and is difficult to predict. Some applications make a consistent number of volume writes, while other applications have a workload that varies. For example, a volume may have 10GB of changes one day and 2GB the next day.

- IT operations. Some disk operations can affect replication. For example, defragmenting a disk or performing a database reorganization on a volume that *already* has been replicated may require a large transfer of data for the next replication. However, subsequent replications may transfer data more efficiently than before the disk operation.
- Locality of change. There may be no direct correspondence between changes to the base volume and the amount of replication space needed. Especially in cases of random volume writes, you may find that you need relatively more local reserve and replica reserve than expected, based on the actual volume changes. For example, 2GB of volume writes could require as much as 3GB of replication space or as little as 1GB, depending on how data is distributed on the file system.

The following figure shows how an example of how replica reserve, set to 200%, is gradually consumed as replicas are created.

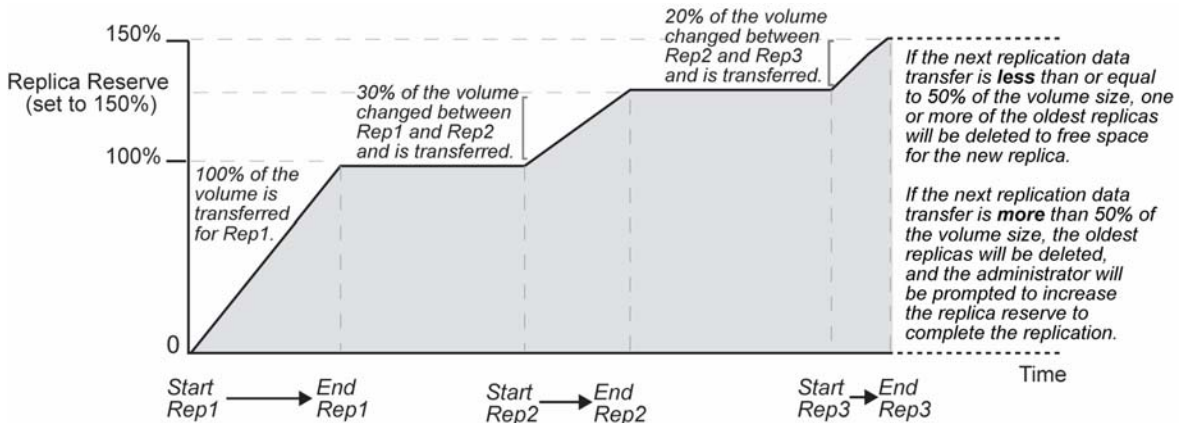


The following table shows the benefits and disadvantages of setting the replica reserve to the default of 200%.

Replica Reserve Set to 200%	
Benefit	Replication data transfer will never fail due to lack of replica reserve space. At least two complete replicas will always be retained, without user intervention.
Disadvantage	Replication uses more space than may be necessary on secondary group.
Solution	Experiment with a replica reserve that is less than 200%.

The following figure shows an example of how replica reserve, set to 150%, is gradually consumed as replicas are created.

Replica Reserve Consumption – Set to 150%



The following table shows the benefits and disadvantages of setting the replica reserve to less than 200%.

Replica Reserve Set to Less than 200%	
Benefit	Replication uses less space on the secondary group.
Disadvantage	If the replica reserve is not big enough to store volume data, you will need to increase the replica reserve (as directed) or manually cancel the replication. Canceling a replication will probably result in a relatively larger data transfer for the next replication, because it is likely that more data will have changed.
Solution	<ul style="list-style-type: none"> • Size the replica reserve so that it can handle the average number of volume writes that occur between consecutive replications. You may need to experiment with different values. See <i>Replica Reserve Sizing Examples</i> for more information. • Use a high-performance, reliable network link between the primary and secondary groups. This will ensure that replication completes in a timely manner. If a replication cannot start because the in-progress replication has not completed, relatively more data will be transferred for the next replication, increasing the need for replica reserve. • Replicate more often. Because relatively less data will need to be transferred, more replicas can be retained in the replica reserve. • Retain fewer replicas. The more replicas you want to retain, the more replica reserve you need.

Replica Reserve Sizing Examples

Only by using a replica reserve of 200% will you ensure that the secondary group can store the two most recent replicas. There are several ways to *roughly* predict how much replica reserve a volume needs, based on the volume changes between consecutive replications and the number of replicas that you want to retain. In all cases, you may need to adjust the value up or down before arriving at the appropriate size for the replica reserve. If space is tight, you may want to prioritize your volumes, and provide extra space to those volumes deemed important.

Also, as described in *Minimizing Replication Completion Times*, you may want to ensure that replication operations complete in a timely manner. This prevents a replication from being cancelled because the current replication is still in progress.

The first part of the replica reserve stores the most recent replica (replica volume), which always consumes 100% of the volume size, as described in *Sizing the Replica Reserve*. To size the second part of the replica reserve (minimum 5% of the volume size), you need to understand the rate of volume changes between consecutive replications.

If the volume changes are *consistent*, you may be able to determine an appropriate replica reserve value by adding 100% (space for the most recent replica) to the amount of volume changes multiplied by the number of replicas to retain. For example:

- If consistently 20% of the volume changes between consecutive replications, specify at least 160% to keep three replicas: $100\% + (20\% \times 3)$

If the volume size is 10GB, the replica reserve will consume 16GB from delegated space.

- If consistently 10% of the volume changes between consecutive replications, specify at least 140% to keep four replicas: $100\% + (10\% \times 4)$

If the volume size is 10GB, the replica reserve will consume 14GB from delegated space.

If volume changes are *not consistent*, you may want to determine both the minimum amount of volume changes that occur between consecutive replications and the maximum amount of volume changes. You can then choose the replica reserve value from the range of values between the minimum and the maximum.

For example, if the minimum amount of volume changes between consecutive replications is 16% and you want to keep three replicas, using the previous formula, the minimum replica reserve needed will be $100\% + (16\% \times 3) = 148\%$. If the maximum amount of volume changes is 30%, the maximum replica reserve will be $100\% + (30\% \times 3) = 190\%$. Therefore, replica reserve values can range from 148% to 190%, with an average of 169%.

After choosing a replica reserve value, monitor the replication and adjust the value as needed.

Handling Depleted Replica Reserve

If there is insufficient replica reserve to complete a replication data transfer, the oldest replicas will be automatically deleted to make room for the new replica. However, the most recent replica will never be automatically deleted. Even if you manually delete the most recent replica (not allowed in some PS Series firmware releases), you will not free all the replica reserve. See *Recovering Replica Reserve* for more information.

If there is still insufficient replica reserve to complete a replication—even if all the replicas but the most recent one will be deleted—the replication will be paused. You will receive a message, similar to the following, stating the amount to which you must increase the replica reserve in order to complete the data transfer:

```
Replication for volume testvol has been paused because the replica on partner
psa2 does not have enough replica reserve for all the current changes to be
replicated. After increasing the replica reserve for the volume to at least
140 percent, volume replication will resume automatically.
```

Note: If you receive the previous message stating that you must increase the replica reserve to complete a replication data transfer, you cannot free space and resume the replication by manually deleting replicas. The action specified in the message assumes the automatic deletion of replicas other than the most recent. If you cannot increase the replica reserve as directed in the message, you can cancel the replication.

If a lack of replica reserve is affecting replication, reevaluate your replica reserve value and increase it as needed. See *Sizing the Replica Reserve* for more information.

In order for the primary group administrator to increase the replica reserve, free delegated space must be available. The secondary group administrator can increase the space delegated to the primary group if the secondary group has free space available.

Recovering Replica Reserve

Circumstances may arise in which you need to decrease the replica reserve to return space to free delegated space. You can decrease the replica reserve for a volume while replication is in progress only if the new value will be *more* than what is currently being used to store replicas.

To free all the replica reserve space for a volume, you must delete the replica set (effectively disabling replication on the volume). If you later re-enable replication on the volume, the next replication will be a complete transfer of all the volume data.

Note: You cannot free all the replica reserve for a volume, even if you delete all the individual replicas for a volume. This is because the replica volume will still exist. Instead, you must delete the replica set to free all the replica reserve.

Affect of Increasing the Volume Size on the Replica Reserve

Because the replica reserve is based on the volume size, the size of the replica reserve will automatically increase if you increase the size of the volume, thus consuming more delegated space. For example, if the size of a volume is 10GB, and the replica reserve is set to 150% of the volume size, the replica reserve will be 15GB. If you increase the volume to 16GB, the replica reserve will increase to 24GB.

Before increasing the size of a replicated volume, be sure that the delegated space is large enough to handle the automatic replica reserve increase. If you use the GUI to increase the volume size, the Space Utilization table will show the new value for the replica reserve. However, the volume resize operation will not fail if there is insufficient delegated space.

You will not know if there is sufficient free delegated space to handle the change in volume size until the next replication occurs. If there is not enough delegated space, the replication data transfer will pause and with a status of `remote-resize-failed`. Also, an event message similar to the following will be logged:

```
Replication for volume VolA has been paused because partner
BGroup does not contain enough free delegated space. After
delegating additional space for the volume, you can resume
the volume replication.
```

To solve the problem, the secondary group administrator can increase the amount of space delegated to the primary group. The primary group administrator can also delete any unnecessary replica sets on the secondary group to free delegated space.

Documentation and Customer Support

Visit the EqualLogic Customer Support website where you can download the latest documentation and firmware. You can also view FAQs, the Knowledge Base, and Tech Reports and submit a service request.

EqualLogic PS Series storage array documentation includes the following:

- *Release Notes.* Provides the latest information about PS Series storage arrays and groups.
- *QuickStart.* Describes how to set up the hardware and start using a PS Series storage array.
- *Group Administration.* Describes how to use the Group Manager GUI to manage a PS Series group and provides comprehensive information about product concepts and procedures.
- *CLI Reference.* Describes how to use the Group Manager command line interface to manage a group and individual arrays.
- *Hardware Maintenance.* Provides information on maintaining the storage array hardware.

To access the Customer Support website, from the EqualLogic website (www.equallogic.com), click **Support** and log in to a support account. If you do not have an account, create one by clicking the link under the login prompt.

To contact customer support, send e-mail to supportnp@equallogic.com. If the issue is urgent, call 1-877-887-7337 to speak with a member of the customer support team.

Document Control and History Information

Revision	Change Author/Organization	Reason	Approved By/Organization	Date
A02	M. Cynewski / GCSS	ISO Compliance	M. Gauthier / GCSS	August 27, 2009
A01	Unknown	Initial Release	Unknown	April 2005