

HIGH-AVAILABILITY SOLUTIONS FOR RAPID™ AND RAPID-PHARMA™

INTRODUCTION

The ability to protect and keep mission critical data and applications available has become a key issue for many companies. Keeping operational 24 hours a day 7 days a week (24/7) can determine the success or failure of entire businesses. As this dependence grows, the losses caused by computer and system downtime are staggering.

THE PROBLEM

Fiscal loss due to unavailability of mission critical systems (such as RAPID) can cost companies and organizations dearly. The most common causes of (unexpected) downtime are server hardware failure and network outages. In a manufacturing environment downtime of such systems can quickly run into millions of dollars in lost product. Couple this with the fact that in many manufacturing environments production data cannot be lost or the product cannot be released for sale. This is particularly true of the FDA regulated industries.

THE AUTOMSOFT SOLUTION

Automsoft's high-availability solutions for our RAPID' and RAPID-Pharma' products ensure the level of mission critical production data and system availability that 21st century businesses need. These high-availability solutions protect against both network segment failure (Network High-Availability) and server hardware failure (Server High-Availability) on the Windows 2000 server platforms. When both Server HA and Network HA are deployed together, they deliver a high-availability solution for the RAPID product family with zero data loss and a low total cost of ownership compared to expensive hardware only solutions.

This white paper describes how both software solutions, Server HA and Network HA work and illustrates some common deployment models (sometimes referred to as architectures). These models are real life solutions developed with our customers needs in mind. Automsoft's high-availability solutions are, of course, fully compatible with regulatory requirements, such as 21 CFR Part 11.

This document assumes the reader has some familiarity with the RAPID product family. Note that for the remainder of the white paper RAPID refers to both RAPID and RAPID-Pharma for brevity.

SERVER HIGH-AVAILABILITY

Server HA is designed to protect the RAPID software and databases (and allied programs) against critical server hardware failure (e.g. disk crash, chip fault, etc.). Specifically, as indicated in Figure 1 below, two Windows 2000 Sever (or Advanced Sever) computers are coupled together to form a 'cluster'. Each computer then acts as a hot spare for its counterpart and is also fully functional as a network server. The primary server computer is also the primary RAPID server. The cluster may be optionally connected to a shared storage device – usually a SAN. RAPID clients are connected over the network to the cluster.

SERVER HIGH-AVAILABILITY IN OPERATION

Normal Operation During normal operation the primary server computer is running the RAPID sever side software, serving client requests to retrieve and store data in its RAPID database(s). The standby computer has the RAPID server side software installed, but during normal operation it is not active. The primary and standby computers also must have the RAPID Server HA software installed and running. Both the primary and standby computers are active on the network backbone and can also perform network, file and print functions.

A continuous bi-directional mirroring process sends (RAPID) data across the dedicated network link ensuring that each computers hard drive(s) are kept up to date with the data from both computers. This real-time link mirrors the RAPID process databases (process, batch and A&E data), the HUSBO database and the Audit Trail database, as well as key housekeeping information from Windows itself.

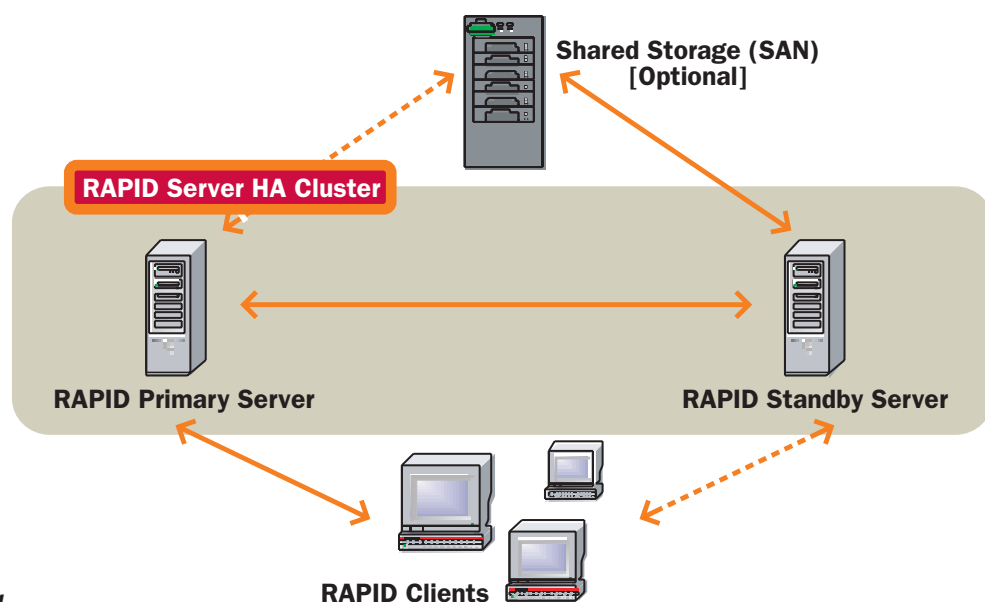


Figure 1.



Primary Server Failure Should the primary server fail for any reason, then the Server HA software transfers critical functions (switches over) from the primary server to the standby server. These critical functions include IP addresses, network shares, print queues, server names and any running applications (i.e. RAPID). Databases that were mirrored (over the dedicated network link) are now available to the standby RAPID computer and are now made available to the network. Once the switch over process is complete the standby server will effectively be the (new) primary server complete with the primary server's name, IP addresses, network shares, etc and, of course, running the RAPID server software. During the switch over process the Server HA software re-binds the primary servers IP address to the standby servers MAC address. The re-binding is broadcast over the network updating routers, bridges etc., ensuring that RAPID client software is uninterrupted.

As the standby sever effectively assumes the identity of primary after a switch over, there is minimal interruption to client applications on the network. In many cases, these clients will be unaware that a switch over from primary to secondary RAPID server has occurred! In particular, the Excel Add-in, RAPID Client and Client Designer reporting tools have complete retry logic built in, and in the event of the RAPID server being unavailable will retry the users query seamlessly. In this case, the user will be isolated from the fact that a switch over has completed.

There is a small window of downtime during the switch over process. Typically the downtime is limited to a couple of minutes, and often less than a minute. For those scenarios and customers where downtime is not an option then the Complete High-Availability solution from Automsoft may be deployed. See below for details.

After the switch over process has been completed the original primary server may be repaired or replaced as appropriate. The repaired or new server may then be reconnected into the cluster and the mirroring process restarted. In this case all disk and RAPID database updates that have occurred while this server was unavailable were queued on the (original) standby server will now be migrated to this server in the background until all databases and disks are synchronized.

Note that, at this point, the original primary server is now the standby server and vice versa.

Standby Server Failure Should the standby sever fail for any reason, there is no reason for the Server HA software to initiate a switch over as the RAPID software is running on the primary server as normal. In this instance the mirroring process will stop, and the network administrator alerted to the hardware failure. The primary sever will queue any disk updates for the standby sever. When the failed hardware is replaced or repaired the mirroring operation resumes and seamlessly transfers the queued disk updates (and new updates) from the primary to the standby server in the background, without loss of performance.

TECHNICAL FEATURES

Solution Architecture Server HA is based on a shared-nothing architecture. A shared-nothing architecture eliminates the possibility of any single hardware element (or software application) bringing down a server (or RAPID) within the two-server cluster. With this architecture each server maintains a copy of the clustered resources as well as a copy of the system registry database. These are synchronized via the dedicated network link as part of the mirroring process.

Hardware and Software Requirements

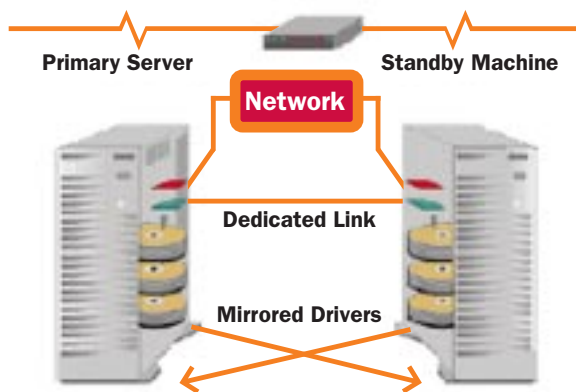


Figure 2. Figure 2 above illustrates the hardware components required to enable Server HA. The two server computers must be Intel based industry standard PC servers running either Windows 2000 Server or Advanced Server. We recommend that neither server be configured as a domain controller.

Two standard network cards per server must be installed, one for the connection to the LAN, and the other to connect the two servers together. We recommend that minimums of 1GB cards be used. This dedicated link is used to mirror disk writes between the servers and also prevents the possibility of a false switch over based on partial network failure. The TCP/IP protocol is used exclusively for both this heart beat purpose and for mirroring traffic.

Windows 2000 must be installed on a disk partition other than the partition where the mirrored RAPID databases reside. Server HA is compatible with all storage devices that are compatible with Windows 2000, including RAID, SCSI and even SAN and NAS environments. Note that in a SAN environment the mirroring feature is optional, as a SAN is a high-availability shared-storage system.

Finally, Server HA requires a RAPID Server HA software license for each of the RAPID server computers, and is sold as a RAPID Option.

Management of Server High-Availability Management, configuration and installation have been designed to be easy - no special knowledge of high-availability systems or architectures are required. The Server HA software is supplied as a separately licensed install program, and should be installed on both servers before installing the RAPID Server software. Graphical tools are provided to make the configuration of partitions and databases as straightforward as possible. Once installed the configuration can be inspected and changed with the Remote Management Console tool from either server computer or from anywhere on the network.

As promised by Automsoft's commitment to low TCO, on-going maintenance costs are as for both servers, as if they were not configured as a cluster running Server HA.

Summary Server HA is ideal for those scenarios where availability of the RAPID historical data is of utmost importance to the organization. For instance, in the case of the pharmaceutical sector historical production data must be available at all times. The automated Server HA switch over to the standby sever facilitates such organizational and business requirements. The simplicity of the solution featuring a shared-nothing architecture, standard hardware and ease of deployment and management, makes Server HA an ideal solution for any organizations using the RAPID system at a low TCO.

NETWORK HIGH-AVAILABILITY

COMPLEMENTARY TO SERVER HA IS NETWORK HA. WHEREAS SERVER HA IS DESIGNED TO ENSURE AVAILABILITY OF THE RAPID SOFTWARE AND DATABASES IN THE EVENT OF SERVER HARDWARE FAILURE, NETWORK HA IS DESIGNED TO PROTECT AGAINST LOSS OF DATA COLLECTION DUE TO NETWORK OUTAGES, OR SERVER DOWNTIME DURING A SERVER HA SWITCHOVER.

NETWORK OUTAGES

Network outages whether due to either scheduled maintenance or physical failure are a very common cause of system downtime. In the case of scheduled maintenance backup network paths may be available, but not always. Physical failure of the network can often interrupt systems for hours or even days while the fault is rectified. The situation is even worse in the case of geographically diverse networks such as a WAN, or over a public network via PSTN or ISDN telephone lines or even a satellite link. Network HA is designed to ensure collection of vital production data continues uninterrupted during a network fault.

OPC AND DCOM

The OPC family of protocols is the most common method for collecting data at a RAPID Server. The OPC protocols (Data Access and Alarms & Events) are implemented by OPC server software, and this may be local to the RAPID Server software on the same computer, or more commonly as OPC is based on Microsoft's DCOM technology, on another computer on the network.

Network High-Availability in Operation A common deployment model for RAPID systems is to use a single centralized RAPID Server system collecting data from multiple OPC servers (data sources) distributed around the network. Clients may then query and report on the centralized database(s). However, in the event of a network outage between the RAPID Server and one or more OPC servers, continuity of data collection is compromised, and hence valuable production data will be lost.

NORMAL OPERATION

Network HA, as illustrated in Figure 3 below, addresses the data collection problem by collocating a basic RAPID Server on each of the data source computers, and thus the connection between the OPC servers and this extra RAPID Server is local. The RAPID Server at the data source collects the process data from the local OPC server, and via the RAPID Data Replication eXtension (DRX) sends this data over the network to the centralized RAPID Server.

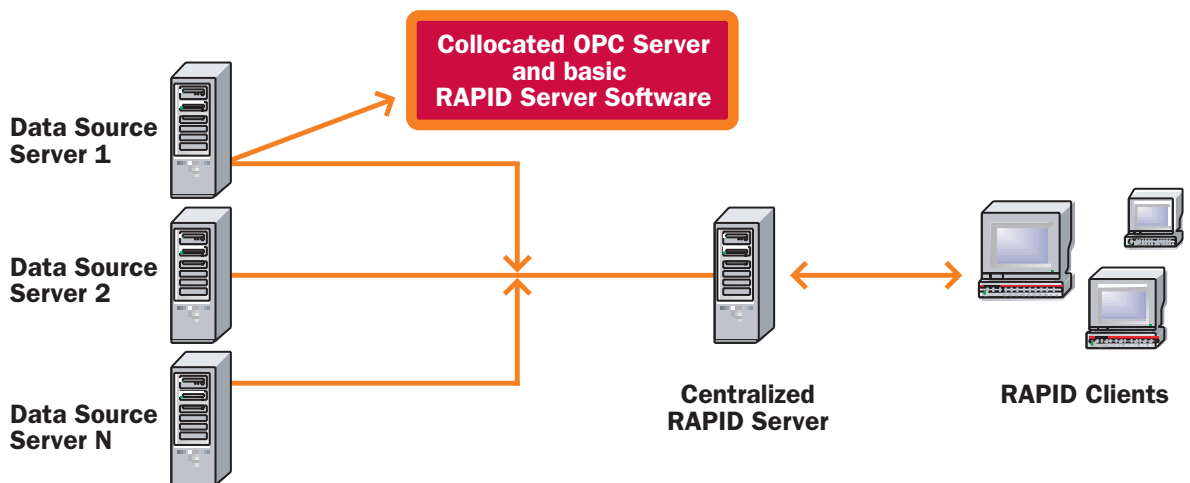


Figure 3.



The basic RAPID Server (or RAPID Data Source) on the data source computer is also a fully functioning RAPID system with a single RAPID client license. Data replicated to the centralized RAPID Server is also written to the local RAPID database(s). Therefore the RAPID Data Source may also be used to run local reports and queries on the local database(s). This is particularly useful when the Data Source computer is physically sited elsewhere from the central server, e.g. at a different manufacturing facility.

Hardware and Software Requirements Network HA requires a minimum of two server class computers – one to act as the Data Source and the other as the Central Server. Up to 25 Data Sources may be replicated into a single Central Server. In addition to the RAPID software, each Data Source computer must have Microsoft's Internet Information Server software installed to transmit the priority queue of XML files.

Network HA requires a RAPID Network HA software license for the Central Server computers, and is sold as a RAPID Option.

Management of Network High-Availability

Management, configuration and installation have been designed to be easy - no special knowledge of high-availability systems or networking is required. The Network HA software is supplied as two installation programs – an Express Install for the Data Source computers and a Network HA install for the Central Server. The Express Install installs the exact components required by each Data Source computer, and the Network HA install installs and configures the Central Server. All licensing of the Data Source computers is handled by the Central Server and is based on a pooled license count.

As promised by Automsoft's commitment to low TCO, on going maintenance costs are minimal.

Summary Network HA is ideal for those scenarios where loss of production data collection is of utmost importance to the organization. The open standards such as HTTP used by this solution allow Network HA to bridge the Internet and pass production data through firewalls. Network HA may also be deployed with Server HA to ensure no loss of critical data ever.

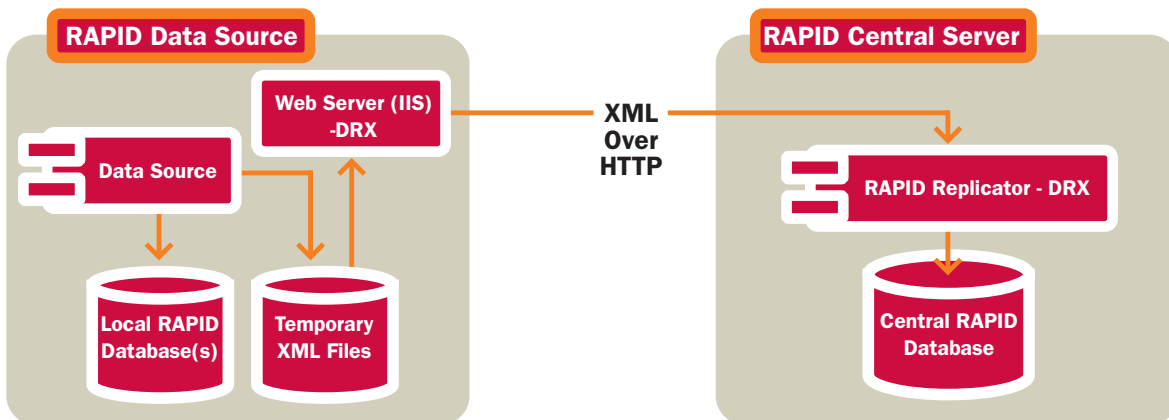


Figure 4.

COMPLETE HIGH-AVAILABILITY

AS WE HAVE SEEN ABOVE BOTH SERVER HA AND NETWORK HA ARE POWERFUL SOLUTIONS TO CUSTOMERS AND ORGANIZATIONS NEEDS FOR AVAILABILITY OF BOTH RAPID SYSTEMS AND DATA. BY COMBINING BOTH SOLUTIONS TOGETHER AS COMPLETE HIGH-AVAILABILITY, THEN BOTH THE RAPID SYSTEMS AND DATABASES MAY BE SAFELY DEPLOYED AT AN UNPRECEDENTED LEVEL OF CONTINUOUS AVAILABILITY.

As both Network HA and Server HA are separately licensed this allows for great flexibility in deploying both solutions together. The next section illustrates some common deployment models (architectures).

DEPLOYMENT MODELS

'Star Model' with Network HA and Server HA The deployment model illustrated in Figure 5 below represents the most common RAPID high-availability solution employed by Automsoft's customers, and is most prevalent in the pharmaceutical and biotechnology industries. This is a very effective solution where a centralized RAPID database is required and where the data sources are sited elsewhere e.g. at another facility.

In this model the primary function of Data Source computers is to collect data only. All reporting is conducted from central cluster with the RAPID Client tools, and reports may of course span multiple Data Sources. The central cluster collects data from the Data Source computers only – there are no OPC servers or connections present on the cluster.

Network HA is employed to protect against network outages as expected, however adopting this model in the event of a Server HA switch over in the cluster has another benefit – no loss of data collection during the switch over! This is because during the switch over process the Data Source computers will continue to buffer data, as if a network fault had occurred, until the switch over process completes.

This is the canonical example of how coupling Network HA and Server HA together protects against loss of data and ensures availability of the RAPID databases.

'Inverted Star Model' with Server HA and Network HA The deployment model illustrated in Figure 6 represents another common RAPID high-availability solution employed by Automsoft's customers. The model is often adopted in the oil and gas industries, where each data source is located on an offshore drilling platform, and the Data warehouse cluster is on shore. In this scenario, because of the harsh weather conditions and environment the Data Source computer is more likely to fail than the central server computers.

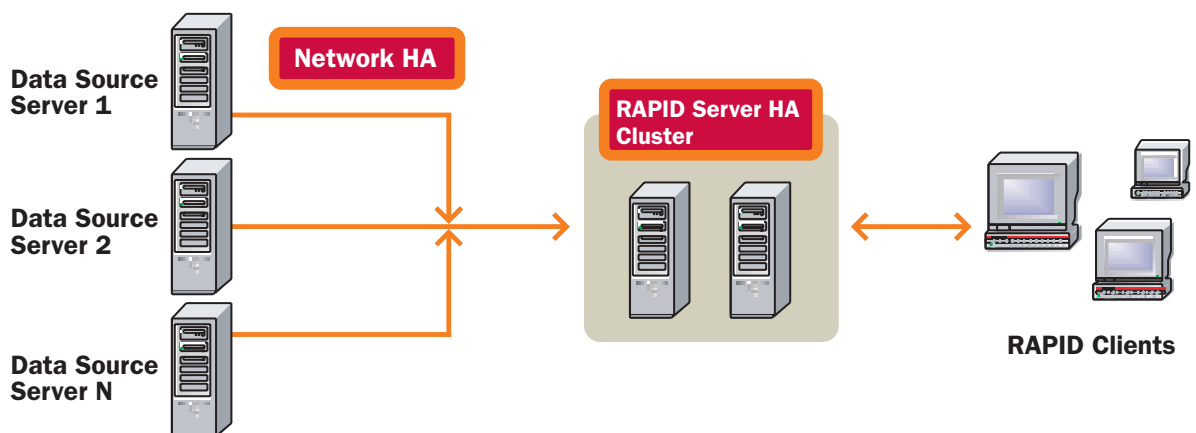
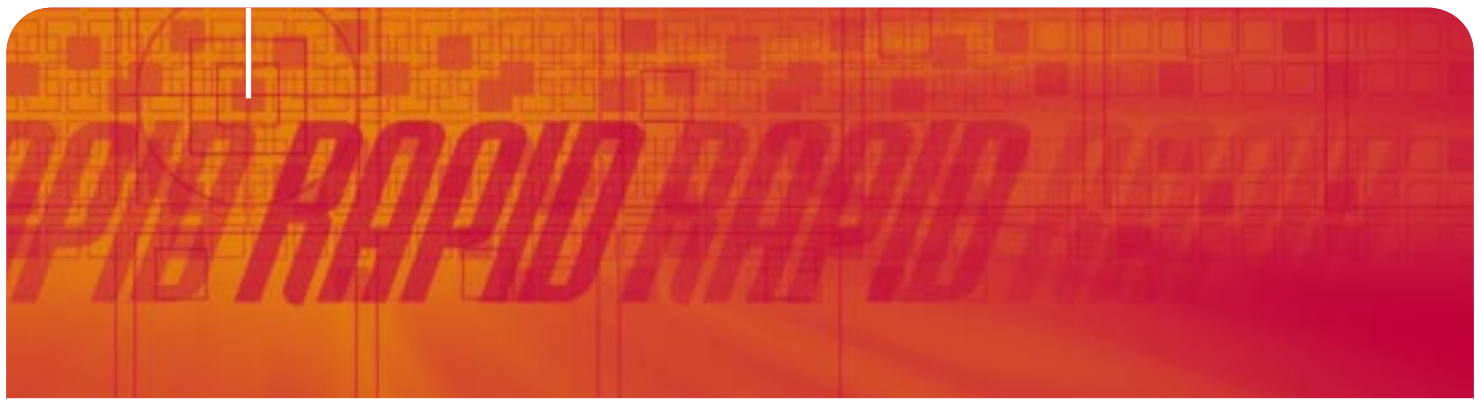


Figure 5.



Therefore to prevent Data Source failure, the Data Source is also configured as a Server HA cluster. There are also additional reporting requirements at the Data Source and this is configured to allow multiple clients to access these local databases. To allow multiple clients to query the Data Source computers and to facilitate switch over, these computers have the complete RAPID Server software installed rather than the Express Install described above.

High-level analysis and reporting across all Data Sources is conducted from the Data Warehouse.

This solution has all of the features of the previous model but allows for failure of the Data Source computers and also facilitates the customer's requirements of additional reporting capabilities at the Data Source computers.

Summary The ability to protect and keep mission critical data and applications available is a key issue for many companies and none more so than manufacturing where fiscal loss due to unavailability of mission critical systems can incur staggering losses. Automsoft specializes in developing software for process industries. In this context, delivery of solutions that take account of industry and client needs is an intrinsic element of our design and development processes. At Automsoft we know that being operational 24/7 can determine the success or failure of entire businesses. At Automsoft we also know that the most common causes of unexpected downtime are server hardware failure and network outages. Automsoft designed and developed our High-Availability solutions specifically to address these needs. For Automsoft customers this translates into - protection of mission critical production data, guaranteed system availability, reduction in ongoing IT operational costs and ultimately increased productivity.

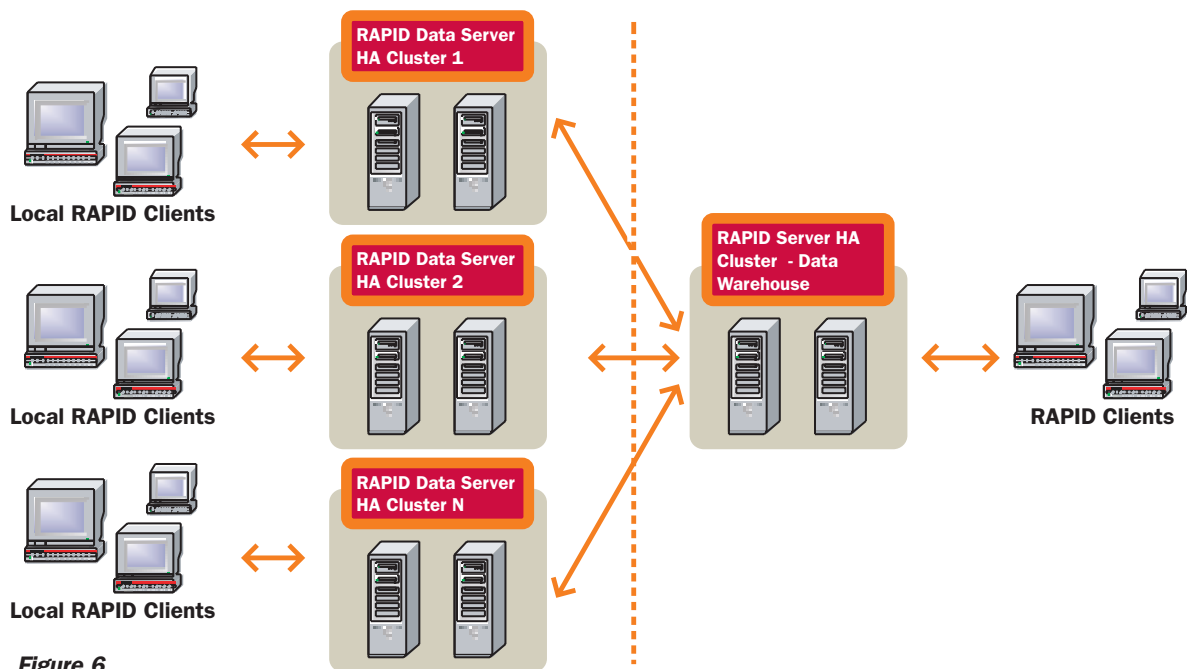


Figure 6.

HOW TO CONTACT US

FOR FURTHER INFORMATION OR ADVICE, PLEASE CONTACT OUR SALES TEAM IN ANY OF THE FOLLOWING LOCATIONS

North America

Automsoft
Sales, North America
330 South Service Road
Suite 120, Melville
NY 11747
USA

T +1 631 574 4922
F +1 631 574 4923

Automsoft
Professional Services
5 Independence Way
Suite 300, Princeton
NJ 08540
USA

T +1 609 919 6010
F +1 609 720 9000

Europe

Automsoft
Sales, Europe
18 Lansdowne Road
Dublin 4
Ireland

T +353-1-449 1100
F +353-1-449 1199

E info@automsoft.com
W www.automsoft.com