

表3-11 Multiport Syslog Source配置

参数	默认值	描述
<b>ports</b>	-	以空格分隔的要绑定的端口列表
<b>portHeader</b>	-	接收事件的端口包含在 header 中，作为这个 header 指定的键的值。如果该参数没有设置，端口信息不包括在内
<b>charset.default</b>	UTF-8	设置要使用的默认的字符
<b>charset.port.&lt;port&gt;</b>	-	为特定端口设置字符
<b>eventSize</b>	2500	一个事件以字节为单位的最大容量
<b>batchSize</b>	100	写入 Channel 之前在内存缓冲的事件的数量
<b>readBufferSize</b>	1024	底层服务器使用的缓冲区大小
<b>numProcessors</b>	-	使用的处理器的数量。这使得 Source 能提高并行的程度

Multiport Syslog Source 可以接收多个端口的数据，这些端口要以空格分隔罗列作为 **ports** 参数的值。每个接收到的事件可以带上端口号注释，表明事件是从哪个端口号接收的。配置参数 **portHeader** 的值用作 Flume 的事件中 header 的键，该 header 的值是端口号。

Source 还允许 syslog 消息将每个端口接收的数据编码为不同的字符集。为每个端口配置字符集，可以使用 **charset.port.** 前缀后面跟上端口号作为配置参数，值为字符集的名字。使用 **charset.default** 参数来设置默认的字符。当一个端口没有指定字符集则使用这个参数的值。

前面提及过，Source 假设每个事件是由一个换行符分隔的。有时，重要的是每个事件并不是固定大小的。每个事件最大的大小可以使用 **eventSize** 参数设置，以字节表示。如果事件的大小大于这个值，事件就被截断为该值的长度，并将 **flume.syslog.status** 的键的值设为 **Incomplete** 插入到 header。Multiport Syslog Source 还在内存中缓冲事件，以避免影响 Channel 的性能。批处理大小可以使用 **batchSize** 参数指定。

61

Source 使用一个名为 *Apache MINA* 的框架来接收消息。MINA 服务器使用一个内部缓冲区读取网络中的数据。这个缓冲区的大小是可配置的，通过配置参数 **readBufferSize** 指定字节数。MINA 对并行性支持良好。配置并行的程度，可以通过使用 **numProcessors** 参数设置使用的处理器的数量。如果没有定义，使用的处理器数量的值是自动检测来的。对于每个可以使用的处理器，MINA 将产生两个线程。如果要减少使用的线程（不经常需要），则减小 **numProcessors** 指定的值。