



[12] 发明专利申请公开说明书

[21] 申请号 200510072316.2

[43] 公开日 2006年2月1日

[11] 公开号 CN 1728685A

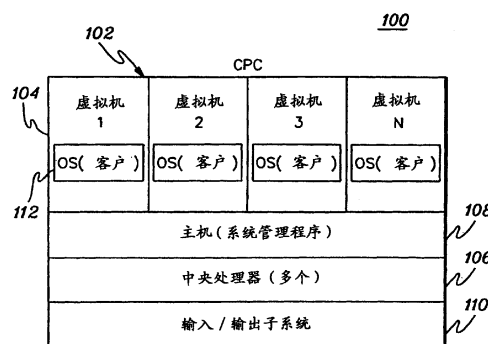
[22] 申请日 2005.5.26
 [21] 申请号 200510072316.2
 [30] 优先权
 [32] 2004.5.27 [33] US [31] 10/855,200
 [71] 申请人 国际商业机器公司
 地址 美国纽约
 [72] 发明人 J·R·伊斯顿 W·A·霍尔德
 B·纳兹 D·L·奥西赛克
 G·E·西特曼 R·P·陶尔曹
 L·W·梅姆恩

[74] 专利代理机构 北京市中咨律师事务所
 代理人 于静 李峥

权利要求书9页 说明书68页 附图31页

[54] 发明名称
 促进可分页模式虚拟环境中的数据传输的方法和系统

[57] 摘要
 来自可分页存储模式客户的输入/输出(I/O)操作请求在没有主机干预的情况下被解释。在可分页模式虚拟环境中，由可分页存储模式客户发出的请求被该环境的一个或多个处理器处理，而没有来自该环境的一个或多个主机的干预。所述请求的处理包括由至少一个处理器代表所述客户来操纵被存储在主机存储器内的缓冲区状态信息。所述操纵通过由所述客户启动并由一个或多个所述处理器处理的指令来执行。



1. 一种促进可分页模式虚拟环境中的数据传输的方法，所述方法包括：

由所述可分页模式虚拟环境的处理器获得将在数据传输中使用的缓冲区的指示，所述缓冲区被与运行在所述可分页模式虚拟环境的可分页模式虚拟机中的客户程序相关联；以及

由所述处理器执行与所述缓冲区关联的处理，以便允许数据传输，所述执行没有所述可分页模式虚拟环境的主机的干预。

2. 如权利要求1的方法，进一步包括在主机存储器中为所述缓冲区提供状态信息，并且其中所述执行包括由所述处理器操纵所述状态信息。

3. 如权利要求2的方法，其中所述操纵包括采用至少一个由所述客户程序发出并由所述处理器执行的指令来操纵所述状态信息。

4. 如权利要求3的方法，其中所述至少一个指令包括设置缓冲区状态指令和提取缓冲区状态指令中的至少一个指令。

5. 如权利要求1的方法，其中所述获得包括从所述客户程序接收所述缓冲区的所述指示。

6. 如权利要求5的方法，其中所述客户程序通过一个指令将所述缓冲区的所述指示传输给所述处理器。

7. 如权利要求1的方法，其中所述缓冲区被与由标记标识的设备相关联，并且其中所述执行包括：

检查所述标记的有效性；以及

响应于有效标记的指示，继续与所述缓冲区关联的处理。

8. 如权利要求1的方法，其中所述缓冲区是队列的缓冲区，并且其中所述执行包括锁定所述队列以便执行与所述缓冲区关联的处理。

9. 如权利要求8的方法，其中所述处理遇到异常，使得处理将由所述主机而不是所述处理器来完成，并且所述方法进一步包括将所述锁的所有权从所述处理器传递给所述主机。

10. 如权利要求1的方法，其中所述执行包括：

标识将在数据传输中使用的所述缓冲区的一个或多个页；以及

将所述一个或多个页钉在物理主机存储器中。

11. 如权利要求 10 的方法, 其中通过设置缓冲区状态指令来启动所述钉住。

12. 如权利要求 10 的方法, 其中将被钉住的所述一个或多个页中的一个页是无效的, 因为它没有驻留在物理存储器内, 并且其中所述执行进一步包括解决该页以使其有效。

13. 如权利要求 12 的方法, 其中所述解决包括使用被提供给所述处理器的一个或多个可用描述符的列表来动态地解决所述无效页, 所述一个或多个可用描述符指定了可用于支持物理存储器中的所述页的一个或多个可用页帧。

14. 如权利要求 10 的方法, 进一步包括:

判定所述缓冲区的一个或多个页是否将被解除钉住; 以及解除所述一个或多个页的钉住。

15. 如权利要求 14 的方法, 其中所述数据传输包括采用光纤通道协议, 并且其中所述判定包括检查输入队列上关于请求所述数据传输的 I/O 操作是否完成的指示, 并且其中当所述指示指出完成时执行所述解除钉住。

16. 如权利要求 14 的方法, 其中通过提取缓冲区状态指令来启动所述解除钉住。

17. 如权利要求 14 的方法, 进一步包括在解除所述缓冲区的一个或多个页的钉住之前, 由所述处理器提取所述缓冲区的状态。

18. 如权利要求 14 的方法, 进一步包括由所述处理器追踪用于一个或多个页的钉住和解除钉住操作中的至少一个操作。

19. 如权利要求 1 的方法, 其中所述执行包括更改所述缓冲区的状态。

20. 如权利要求 19 的方法, 其中所述更改状态包括将所述状态从程序拥有更改为适配器拥有。

21. 如权利要求 1 的方法, 其中所述获得包括获得将在数据传输中使用的所述缓冲区的页的标识, 并且其中所述执行包括:

判定所述页是否驻留在物理存储器中;

响应于指出所述页没有驻留的所述判定, 解决所述页以使所述页驻留; 以及

将所述页钉在物理存储器中。

22. 如权利要求 21 的方法，其中所述数据传输被与一个输入操作相关联，对于该输入操作而言，所述页的先前内容不再被需要，并且其中所述解决包括将所述页与先前被清除的页帧相关联。

23. 如权利要求 21 的方法，其中所述获得所述页的所述标识包括接收客户数据结构，该客户数据结构包括所述页的客户程序指定，并且其中所述执行进一步包括：

由所述处理器将所述客户数据结构的内容的至少一部分复制到主机数据结构；

从所述主机数据结构取回将被用于判定所述页是否驻留在物理存储器内的所述客户程序指定；

获得所述页的主机指定；以及

将所述页的所述主机指定置于所述主机数据结构内，所述主机指定可用于数据传输。

24. 如权利要求 23 的方法，进一步包括将所述缓冲区的状态更改为适配器拥有，使得所述可分页模式虚拟环境的适配器能够访问所述缓冲区。

25. 如权利要求 24 的方法，其中所述适配器使用所述主机指定来访问所述缓冲区的所述页。

26. 如权利要求 23 的方法，其中从所述客户程序来接收所述客户数据结构。

27. 如权利要求 23 的方法，进一步包括：

判定是否将执行解除钉住操作；以及

响应于指出将执行所述解除钉住操作的所述判定，执行所述解除钉住操作。

28. 如权利要求 27 的方法，其中所述缓冲区被与一个队列相关联，并且其中所述方法包括锁定所述队列以便执行对所述缓冲区的处理。

29. 如权利要求 28 的方法，其中所述缓冲区对应于由标记标识的设备，并且所述执行处理进一步包括验证所述标记。

30. 如权利要求 1 的方法，其中所述数据传输涉及通信适配器。

31. 如权利要求 1 的方法，其中所述数据传输涉及光纤通道协议适配

器。

32. 如权利要求 1 的方法, 进一步包括由所述客户程序询问所述缓冲区的状态。

33. 如权利要求 32 的方法, 其中通过提取缓冲区状态指令来进行所述询问。

34. 一种促进可分页模式虚拟环境中的输入/输出 (I/O) 操作的处理的方法, 所述方法包括:

由所述可分页模式虚拟环境的处理器来处理在所述可分页模式虚拟环境的虚拟机中运行的客户程序的 I/O 操作, 其中所述处理没有所述可分页模式虚拟环境的主机的干预。

35. 如权利要求 34 的方法, 其中所述 I/O 操作包括排队直接 I/O (QDIO) 操作。

36. 如权利要求 35 的方法, 其中所述 QDIO 操作涉及通信适配器。

37. 如权利要求 35 的方法, 其中所述 QDIO 操作涉及光纤通道协议适配器。

38. 一种促进可分页模式虚拟环境中的数据传送的系统, 所述系统包括:

一装置, 所述装置用于由所述可分页模式虚拟环境的处理器获得将在数据传输中使用的缓冲区的指示, 所述缓冲区被与运行在所述可分页模式虚拟环境的可分页模式虚拟机中的客户程序相关联; 以及

一装置, 所述装置用于由所述处理器执行与所述缓冲区关联的处理, 以便允许数据传输, 所述执行没有所述可分页模式虚拟环境的主机的干预。

39. 如权利要求 38 的系统, 进一步包括用于在主机存储器中为所述缓冲区提供状态信息的装置, 并且其中所述用于执行的装置包括用于由所述处理器操纵所述状态信息的装置。

40. 如权利要求 38 的系统, 其中所述用于获得的装置包括用于从所述客户程序接收所述缓冲区的所述指示的装置。

41. 如权利要求 38 的系统, 其中所述缓冲区被与由标记标识的设备相关联, 并且其中所述用于执行的装置包括:

一装置, 所述装置用于检查所述标记的有效性; 以及

一装置，所述装置用于响应于有效标记的指示，继续与所述缓冲区关联的处理。

42. 如权利要求 38 的系统，其中所述缓冲区是队列的缓冲区，并且其中所述用于执行的装置包括用于锁定所述队列以便执行与所述缓冲区关联的处理的装置。

43. 如权利要求 42 的系统，其中所述处理遇到异常，使得处理将由所述主机而不是所述处理器来完成，并且所述系统进一步包括用于将所述锁的所有权从所述处理器传递给所述主机的装置。

44. 如权利要求 38 的系统，其中所述用于执行的装置包括：

一装置，所述装置用于标识将在数据传输中使用的所述缓冲区的一个或多个页；以及

一装置，所述装置用于将所述一个或多个页钉在物理主机存储器中。

45. 如权利要求 44 的系统，其中将被钉住的所述一个或多个页中的一个页是无效的，因为它没有驻留在物理存储器内，并且其中所述用于执行的装置进一步包括用于解决该页以使其有效的装置。

46. 如权利要求 45 的系统，其中所述用于解决的装置包括用于使用被提供给所述处理器的一个或多个可用描述符的列表来动态地解决所述无效页的装置，所述一个或多个可用描述符指定了可用于支持物理存储器中的所述页的一个或多个可用页帧。

47. 如权利要求 44 的系统，进一步包括：

一装置，所述装置用于判定所述缓冲区的一个或多个页是否将被解除钉住；以及

一装置，所述装置用于解除所述一个或多个页的钉住。

48. 如权利要求 47 的系统，进一步包括用于在解除所述缓冲区的一个或多个页的钉住之前，由所述处理器提取所述缓冲区的状态的装置。

49. 如权利要求 47 的系统，进一步包括用于由所述处理器追踪用于一个或多个页的钉住和解除钉住操作中的至少一个操作的装置。

50. 如权利要求 38 的系统，其中所述用于执行的装置包括用于更改所述缓冲区的状态的装置。

51. 如权利要求 38 的系统，其中所述用于获得的装置包括用于获得

将在数据传输中使用的所述缓冲区的页的标识的装置，并且其中所述用于执行的装置包括：

一装置，所述装置用于判定所述页是否驻留在物理存储器中；

一装置，所述装置用于响应于指出所述页没有驻留的所述判定，解决所述页以使所述页驻留；以及

一装置，所述装置用于将所述页钉在物理存储器中。

52. 如权利要求 51 的系统，其中所述数据传输被与一个输入操作相关联，对于该输入操作而言，所述页的先前内容不再被需要，并且其中所述用于解决的装置包括用于将所述页与先前被清除的页帧相关联的装置。

53. 如权利要求 51 的系统，其中所述用于获得所述页的所述标识的装置包括用于接收客户数据结构的装置，该客户数据结构包括所述页的客户程序指定，并且其中所述用于执行的装置进一步包括：

一装置，所述装置用于由所述处理器将所述客户数据结构的内容的至少一部分复制到主机数据结构；

一装置，所述装置用于从所述主机数据结构取回将被用于判定所述页是否驻留在物理存储器内的所述客户程序指定；

一装置，所述装置用于获得所述页的主机指定；以及

一装置，所述装置用于将所述页的所述主机指定置于所述主机数据结构内，所述主机指定可用于数据传输。

54. 如权利要求 53 的系统，进一步包括用于将所述缓冲区的状态更改为适配器拥有，使得所述可分页模式虚拟环境的适配器能够访问所述缓冲区的装置。

55. 如权利要求 53 的系统，进一步包括：

一装置，所述装置用于判定是否将执行解除钉住操作；以及

一装置，所述装置用于响应于指出将执行所述解除钉住操作的所述判定，执行所述解除钉住操作。

56. 如权利要求 55 的系统，其中所述缓冲区被与一个队列相关联，并且其中所述系统包括用于锁定所述队列以便执行对所述缓冲区的处理的装置。

57. 如权利要求 56 的系统，其中所述缓冲区对应于由标记标识的设

备，并且所述用于执行处理的装置进一步包括用于验证所述标记的装置。

58. 如权利要求 38 的系统，进一步包括用于由所述客户程序询问所述缓冲区的状态的装置。

59. 一种促进可分页模式虚拟环境中的数据传输的系统，所述系统包括：

所述可分页模式虚拟环境的处理器，所述处理器适合于获得将在数据传输中使用的缓冲区的指示，所述缓冲区被与运行在所述可分页模式虚拟环境的可分页模式虚拟机中的客户程序相关联；并且

所述处理器还适合于执行与所述缓冲区关联的处理，以便允许数据传输，所述执行没有所述可分页模式虚拟环境的主机的干预。

60. 一件制品，所述制品包括：

至少一个计算机可用介质，所述介质具有计算机可读程序代码逻辑，以便管理促进可分页模式虚拟环境中的数据传输，所述计算机可读程序代码逻辑包括：

获得逻辑，以便由所述可分页模式虚拟环境的处理器获得将在数据传输中使用的缓冲区的指示，所述缓冲区被与运行在所述可分页模式虚拟环境的可分页模式虚拟机中的客户程序相关联；以及

执行逻辑，以便由所述处理器执行与所述缓冲区关联的处理以允许数据传输，所述执行没有所述可分页模式虚拟环境的主机的干预。

61. 如权利要求 60 的制品，进一步包括提供逻辑，以便在主机存储器中为所述缓冲区提供状态信息，并且其中所述执行逻辑包括操纵逻辑，以便由所述处理器操纵所述状态信息。

62. 如权利要求 60 的制品，其中所述获得逻辑包括接收逻辑，以便从所述客户程序接收所述缓冲区的指示。

63. 如权利要求 60 的制品，其中所述缓冲区被与由标记标识的设备相关联，并且其中所述执行逻辑包括：

检查逻辑，以便检查所述标记的有效性；以及

继续逻辑，以便响应于有效标记的指示，继续与所述缓冲区关联的处理。

64. 如权利要求 60 的制品，其中所述缓冲区是队列的缓冲区，并且

其中所述执行逻辑包括锁定逻辑，以便锁定所述队列以执行与所述缓冲区关联的处理。

65. 如权利要求 64 的制品，其中所述处理遇到异常，使得处理将由所述主机而不是所述处理器来完成，并且所述处理逻辑进一步包括传递逻辑，以便将所述锁的所有权从所述处理器传递给所述主机。

66. 如权利要求 60 的制品，其中所述执行逻辑包括：
标识逻辑，以便标识将在数据传输中使用的所述缓冲区的一个或多个页；以及

钉住逻辑，以便将所述一个或多个页钉在物理主机存储器中。

67. 如权利要求 66 的制品，其中将被钉住的所述一个或多个页中的一个页是无效的，因为它没有驻留在物理存储器内，并且其中所述执行逻辑进一步包括解决逻辑，以便解决该页以使其有效。

68. 如权利要求 67 的制品，其中所述解决逻辑包括用于使用被提供给所述处理器的一个或多个可用描述符的列表来动态地解决所述无效页的逻辑，所述一个或多个可用描述符指定了可用于支持物理存储器中的所述页的一个或多个可用页帧。

69. 如权利要求 66 的制品，进一步包括：

判定逻辑，以便判定所述缓冲区的一个或多个页是否将被解除钉住；
以及

解除钉住逻辑，以便解除所述一个或多个页的钉住。

70. 如权利要求 69 的制品，进一步包括提取逻辑，以便在解除所述缓冲区的一个或多个页的钉住之前，由所述处理器提取所述缓冲区的状态。

71. 如权利要求 69 的制品，进一步包括追踪逻辑，以便由所述处理器追踪用于一个或多个页的钉住和解除钉住操作中的至少一个操作。

72. 如权利要求 60 的制品，其中所述用于执行的执行逻辑包括更改逻辑，以便更改所述缓冲区的状态。

73. 如权利要求 60 的制品，其中所述获得逻辑包括用于获得将在数据传输中使用的所述缓冲区的页的标识的逻辑，并且其中所述执行逻辑包括：

判定逻辑，以便判定所述页是否驻留在物理存储器中；

解决逻辑，以便响应于指出所述页没有驻留的所述判定，解决所述页以使所述页驻留；以及

钉住逻辑，以便将所述页钉在物理存储器中。

74. 如权利要求 73 的制品，其中所述数据传输被与一个输入操作相关联，对于该输入操作而言，所述页的先前内容不再被需要，并且其中所述解决包括将所述页与先前被清除的页帧相关联。

75. 如权利要求 73 的制品，其中所述用于获得所述页的所述标识的获得逻辑包括接收逻辑，以便接收客户数据结构，该客户数据结构包括所述页的客户程序指定，并且其中所述执行逻辑进一步包括：

复制逻辑，以便由所述处理器将所述客户数据结构的内容的至少一部分复制到主机数据结构；

取回逻辑，以便从所述主机数据结构取回将被用于判定所述页是否驻留在物理存储器内的所述客户程序指定；

获得逻辑，以便获得所述页的主机指定；以及

放置逻辑，以便将所述页的所述主机指定置于所述主机数据结构内，所述主机指定可用于数据传输。

76. 如权利要求 75 的制品，进一步包括更改逻辑，以便将所述缓冲区的状态更改为适配器拥有，使得所述可分页模式虚拟环境的适配器能够访问所述缓冲区。

77. 如权利要求 75 的制品，进一步包括：

判定逻辑，以便判定是否将执行解除钉住操作；以及

执行逻辑，以便响应于指出将执行所述解除钉住操作的所述判定，执行所述解除钉住操作。

78. 如权利要求 77 的制品，其中所述缓冲区被与一个队列相关联，并且其中所述制品包括锁定逻辑，以便锁定所述队列以执行对所述缓冲区的处理。

79. 如权利要求 78 的制品，其中所述缓冲区对应于由标记标识的设备，并且所述用于执行处理的执行逻辑进一步包括验证逻辑，以便验证所述标记。

80. 如权利要求 60 的制品，进一步包括询问逻辑，以便由所述客户程序询问所述缓冲区的状态。

促进可分页模式虚拟环境中的数据传输的方法和系统

技术领域

本发明一般地涉及输入/输出 (I/O) 处理, 具体地说, 涉及促进对可分页客户的 I/O 操作请求的解释。

背景技术

高效的输入/输出 (I/O) 处理是一种不断增加的挑战, 并且正朝着提高效率不断地大步前进。作为一个例子, 为了获得明显高于以前所允许的带宽, 已经由国际商业机器公司设计出一种被称为排队直接输入/输出 (QDIO) 的增强的输入/输出体系结构。该体系结构被用于各种系统和产品, 包括由国际商业机器公司所提供的 zSeries 系统 (例如, z900) 和 z/VM 操作系统产品。这些系统和产品被增强以便允许多个操作系统映像在主机的 (例如, z/VM) 控制下运行, 以使用 TCP/IP 通信适配器和光纤通道协议 (FCP) 适配器来执行排队直接 I/O 操作。

当前, 由在可分页模式虚拟机中执行的客户操作系统使用 QDIO 体系结构需要主机的大量开销。例如, 主机需要创建并维护影子 QDIO 数据队列, 该队列对针对客户操作系统 (其创建用于访问连接到 QDIO 适配器的网络或 I/O 设备的 QDIO 队列) 采取的 QDIO 操作进行镜像。由于主机虚拟化客户操作系统实际存储器地址并透明地在客户虚拟机 (每个所述操作系统在其中执行) 之间共享系统的实际存储器页的固有方式, 造成存在这种 QDIO 建立影子 (shadowing) /镜像要求。

具体地说, 客户操作系统使用用于 QDIO 队列的虚拟化的实际存储器。然而, 适配器只访问主机实际存储器位置, 而不是虚拟化的客户实际存储器位置。这是因为适配器不象中央处理单元那样具有动态地址转换机制。因此, 由客户操作系统用于 QDIO 队列的虚拟化的实际存储器被主机透明地映射到实际或物理存储器中的不同存储器页中。因此, 系统管理程序必须执行各种操作以便允许客户操作系统使用 QDIO 适配器。

首先，它必须通过在主机实际存储器中创建客户 QDIO 队列的影子拷贝，来将客户实际存储器中由客户操作系统创建的 QDIO 队列进行镜像。这些影子队列由主机透明地创建并维护，并且依次由客户隐式地使用以访问 QDIO 适配器。只有影子队列被连接到 QDIO 适配器，并且只有 QDIO 影子队列中包含的主机实际存储器 I/O 数据区域地址被实际地用于向/从适配器和主存储器传送数据。

进而，主机必须使每个客户创建的 QDIO 队列与其相应的主机创建的影子 QDIO 队列（由 QDIO 适配器使用以执行客户操作系统指定的 QDIO 操作）保持同步。具体地说，客户实际存储器中的客户 QDIO 队列中的所有 QDIO 状态信息必须与主机实际存储器中的相应主机影子 QDIO 队列中的状态信息保持同步；并且客户 QDIO 队列中的操作系统指定的客户实际存储器 I/O 数据区域地址必须由主机分配相应的主机实际存储器数据区域，然后在由适配器使用的主机创建的影子队列中必须指定这些数据区域地址。

这种建立影子并维持同步的处理将要求每次客户操作系统执行 QDIO 操作时，主机都获得对中央处理器的控制。进而，与每次客户执行 QDIO 操作时从/向客户操作系统和主机传送控制相关联的处理周期显著地增大了 QDIO 处理所需的总体处理器开销（超过了当操作系统在本机环境中执行时所需的开销）。与本机执行相比较，在运行作为可分页客户的工作负荷所需的总的处理周期中，这种额外处理开销能够造成几乎 40% 的增加。

基于上述情况，需要一种增强输入/输出处理的能力。例如，需要一种减小与执行诸如 QDIO 操作之类的 I/O 操作关联的主机处理开销的能力。

发明内容

通过提供一种促进可分页模式虚拟环境中的数据传输的方法，克服了现有技术的缺陷并提供了其他优点。所述方法包括，例如，由所述可分页模式虚拟环境的处理器获得将被用在数据传输中的缓冲区的指示，所述缓冲区与在所述可分页模式虚拟环境的可分页模式虚拟机中运行的客户程序相关联；以及由所述处理器执行与所述缓冲区关联的处理以便启动数据传输，所述执行没有所述可分页模式虚拟环境的主机的干预。

在本发明的进一个方面中，提供了一种促进可分页模式虚拟环境中的输入/输出（I/O）操作的处理的方法。所述方法包括，例如，由所述可分页模式虚拟环境的处理器处理运行在所述可分页模式虚拟环境的虚拟机中的客户程序的 I/O 操作，其中所述处理没有所述可分页模式虚拟环境的主机的干预。

此处还描述并要求保护对应于上述方法的系统和计算机程序产品。

通过本发明的技术实现了其他特征和优点。此处详细描述了本发明的其他实施例和方面，并且它们被看作所要求保护的发明的一部分。

附图说明

在本说明书结尾处的权利要求书中特别指出并明确要求了被看作本发明的主题。通过以下结合附图的详细说明，本发明的上述和其他目标、特性和优点是显而易见的，这些附图是：

图 1a 给出了包含并使用本发明的一个和多个方面的计算环境的一个实施例；

图 1b 给出了根据本发明的一个方面的图 1a 的 I/O 子系统的进一步的细节；

图 2a - 2b 给出了根据本发明的一个方面的与客户配置相关联的逻辑的一个实施例；

图 3a 给出了根据本发明的一个方面的用于存储通道子系统特征命令的请求块的一个例子；

图 3b 给出了根据本发明的一个方面的用于存储通道子系统特征命令的响应块的一个例子；

图 4a 给出了根据本发明的一个方面的用于存储子通道 QDIO 数据命令的请求块的一个例子；

图 4b 给出了根据本发明的一个方面的用于存储子通道 QDIO 数据命令的响应块的一个例子；

图 4c 给出了根据本发明的一个方面的图 4b 的响应块的子通道 QDIO 描述块的一个例子；

图 5 给出了根据本发明的一个方面的 QDIO 队列的数据结构的一个实

施例；

图 6a 给出了根据本发明的一个方面的 QDIO 缓冲区信息控制块 (QBICB) 的一个例子；

图 6b 给出了根据本发明的一个方面的图 6a 的 QBICB 的队列存储器描述符的一个例子；

图 6c 给出了根据本发明的一个方面的光纤通道协议 (FCP) 操作块 (FOB) 的一个例子；

图 7a 给出了根据本发明的一个方面的用于设置通道子系统特征命令的请求块的一个例子；

图 7b 给出了根据本发明的一个方面的图 7a 的请求块的操作数据区域的一个例子；

图 7c 给出了根据本发明的一个方面的用于设置通道子系统特征命令的响应块的一个例子；

图 8a - 8c 给出了根据本发明的一个方面的用于和网络协议 I/O 请求一起使用的与设置 QDIO 缓冲区状态指令相关联的逻辑的一个实施例；

图 9 给出了根据本发明的一个方面的所使用的帧描述符的一个例子；

图 10 给出了根据本发明的一个方面的设置 QDIO 缓冲区状态指令的格式的一个例子；

图 11a 给出了根据本发明的一个方面的页状态表表项 (PGSTE) 的一个例子；

图 11b 给出了根据本发明的一个方面的页表表项 (PTE) 的一个例子；

图 12 给出了根据本发明的一个方面的与钉住 (pin) 功能相关联的逻辑的一个实施例；

图 13a - 13b 给出了根据本发明的一个方面的与解决 (resolve) 功能相关联的逻辑的一个实施例；

图 14a - 14c 给出了根据本发明的一个方面的用于和网络协议 I/O 请求一起使用的与提取 QDIO 缓冲区状态指令相关联的逻辑的一个实施例；

图 15 给出了根据本发明的一个方面的提取 QDIO 缓冲区状态指令的格式的一个例子；

图 16 给出了根据本发明的一个方面的与由虚拟地址解除主机页钉住

功能相关联的逻辑的一个实施例;

图 17 给出了根据本发明的一个方面的与由 PTE 地址解除主机页钉住功能相关联的逻辑的一个实施例;

图 18a - 18d 给出了用于和光纤通道协议 I/O 请求一起使用的与设置 QDIO 缓冲区状态指令相关联的逻辑的一个实施例; 以及

图 19a - 19e 给出了用于和光纤通道协议 I/O 请求一起使用的与提取 QDIO 缓冲区状态指令相关联的逻辑的一个实施例。

具体实施方式

根据本发明的一个方面, 来自可分页存储客户 (例如, V = V 客户) 的 I/O 操作请求在没有主机干预的情况下被解释。例如, 提供了一种能力, 其允许诸如 z/VM[®] 系统管理程序之类的主机以及诸如 zSeries 处理器之类的处理器的硬件/固件以一种受控的协作方式来彼此交互, 以便在不需要从 / 向客户操作系统和主机传递控制的情况下处理 V = V 客户操作系统 QDIO 操作。

进而, 在本发明的另一个方面中, 提供了可以在解释可分页存储模式客户的过程中被调用的主机页管理辅助功能。这些功能包括: 解决主机页功能, 其被用于动态地解决主机页无效状态; 钉住主机页功能, 其被用于指出主机页的内容被钉在主机页帧中; 以及由虚拟地址解除主机页钉住功能和/或由 PTE 地址解除主机页钉住功能, 它们被用于指出主机页的内容不再被钉在主机页帧中。

参照图 1a 说明了包含并使用本发明的一个或多个方面的计算环境的一个实施例。计算环境 100 基于例如由纽约阿蒙克的国际商业机器公司所提供的 z/Architecture。在标题为 “z/Architecture Principles of Operation” (“z/Architecture 运行原理”, IBM[®] 出版物 No.SA22-7832-02, 2003 年 6 月, 其全部内容在此引入作为参考) 的 IBM[®] 出版物中描述了所述 z/Architecture。 (IBM[®] 是美国纽约阿蒙克的国际商业机器公司的注册商标。此处所使用的其他名称可能是国际商业机器公司或其他公司的注册商标、商标或产品名称。) 在一个例子中, 基于所述 z/Architecture 的计算环境包括由纽约阿蒙克的国际商业机器公司所提供的 eServer zSeries。

作为一个例子，计算环境 100 包括中央处理器复合体 (CPC) 102，其提供虚拟机支持。CPC 102 包括例如一个或多个虚拟机 104、一个或多个中央处理器 106、至少一个主机 108 (例如，诸如系统管理程序之类的控制程序) 以及输入/输出子系统 110，以下对每一个进行描述。

CPC 的虚拟机支持提供了运行大量虚拟机的能力，每个虚拟机能够容纳诸如 Linux 之类的客户操作系统 112。每个虚拟机 104 能够用作单独的系统。即，每个虚拟机可以被独立地重置、容纳客户操作系统以及与不同的程序一起运行。在虚拟机中运行的操作系统或应用程序似乎在访问完整的和全部的系统，但事实上只有系统的一部分是可用的。

在此特定的例子中，虚拟机的模型是 V = V 模型，该模型中，由虚拟存储器而不是实际存储器来支持虚拟机的存储器。每个虚拟机具有虚拟的线性存储器空间。物理资源由主机 108 拥有，并且在需要时由主机向诸客户操作系统分派共享的物理资源，以便满足它们的处理要求。这种 V = V 虚拟机模型假设诸客户操作系统之间的交互和物理共享的计算机资源受主机控制，因为大量的客户通常使得主机简单地进行分区并向所配置的客户分配硬件资源成为不可能。在标题为“z/VM: Running Guest Operating Systems” (“z/VM: 运行客户操作系统”，IBM 出版物 No. SC24-5997-02，2001 年 10 月，其全部内容在此引入作为参考) 的 IBM 出版物中进一步描述了 V = V 模型的一个或多个方面。

中央处理器 106 是可以分配给虚拟机的物理的处理器资源。例如，虚拟机 104 包括一个或多个逻辑处理器，每个逻辑处理器代表可以被动态地分配给虚拟机的所有的或部分的物理处理器资源 106。虚拟机 104 由主机 108 来管理。作为例子，所述主机可以在处理器 106 上运行的微代码来实现，或者是在计算机上执行的主机操作系统的一部分。在一个例子中，主机 108 是诸如由纽约阿蒙克的国际商业机器公司所提供的 z/VM[®] 之类的 VM 系统管理程序。在标题为“z/VM: General Information Manual” (“z/VM: 通用信息手册”，IBM 出版物 No. GC24-5991-04，2001 年 10 月，其全部内容在此引入作为参考) 的 IBM 出版物中描述了 z/VM[®] 的一个例子。

输入/输出子系统 110 在诸设备与主存储器之间引导信息流。它被连接

到中央处理复合体，因为它可以是中央处理复合体的一部分或与其分离。I/O 子系统减轻了中央处理器与被连接到 CPC 的 I/O 设备直接通信的任务，并且允许数据处理与 I/O 处理同时进行。在一个实施例中，所述 I/O 子系统基于排队直接 I/O (QDIO) 体系结构，在标题为“Method And Apparatus For Providing Configuration Information Using A Queued Direct Input-Output Device” (“使用排队直接输入/输出设备来提供配置信息的方法和装置”，Markos 等人，美国专利 No.6,519,645B2，2003 年 2 月 11 日公开，其全部内容在此引入作为参考) 的美国专利中描述了所述体系结构的一个例子。参考图 1b 说明了关于 I/O 子系统 110 的其他细节。

I/O 子系统 110 包括被连接到多个 I/O 设备 122 的多个适配器 120。一个特定的适配器可以被连接到一个或多个 I/O 设备，并且一个 I/O 设备可以被连接到一个或多个适配器。

适配器的一个例子是 QDIO 适配器。QDIO 适配器提供了通常与输入/输出控制单元相关的功能，以及通常与输入/输出通道相关的功能。QDIO 适配器具有直接访问主存储器的能力，为的是与访问同一主存储器的程序交换数据。

QDIO 适配器的一种逻辑表示是 QDIO 控制单元。可以有多个 QDIO 控制单元表示同一 QDIO 适配器。

QDIO 适配器被通过在适配器与主存储器之间提供数据传输路径的通道连接到 CPC 中的其他单元。为了将从所连接的 I/O 设备或设备网络接收到的数据传输给程序，适配器使用位于主存储器中的 QDIO 输入队列。类似地，为了向 QDIO 适配器传输数据 (QDIO 适配器又可以向所连接的 I/O 设备或设备网络传输数据)，程序使用 QDIO 输出队列。

可以有多个 (例如，32 个) 输入队列和输出队列可用于一个特定的适配器。为了指出哪些队列可用于特定的适配器，采用了一种被称为 QDIO 子通道的逻辑结构。QDIO 子通道包括零个或多个输入队列以及零个或多个输出队列，并且一个或多个子通道被与一个特定的适配器相关联，使得所述适配器能够访问这些队列。

对于 QDIO 输入队列和 QDIO 输出队列，主存储器 (例如，主机存储器) 被用作在程序和适配器之间交换数据的介质。此外，这些队列为程序

和适配器提供了以可预知的和高效的异步方式彼此直接通信的能力，无需诸如操作系统输入/输出系统管理程序之类的集中式控制机制的服务，并且没有此类控制机制带来的结果开销。输入和输出队列都由程序在主存储器中构造，并且在 QDIO 适配器处被建立和激活。在 I/O 操作的处理过程中，这些队列被用于通信。

根据本发明的一个方面，为了在不需要从/向客户操作系统和主机传送控制的情况下处理 V = V 客户操作系统 QDIO 操作，主机（例如，z/VM）和处理器（例如，zSeries）硬件/固件以受控的协作方式彼此交互。客户 QDIO 操作可以利用一种被称为 QDIO 增强缓冲区状态管理（QEBSM）的工具，并且当这么做时，在没有主机干预的情况下，通过一种被称作 QDIO 增强缓冲区状态管理解释（QEBSMI）工具（其允许为可分页存储模式客户解释性地执行所选择的指令）的工具，操作可以直接经过主机到达 QDIO 适配器/从 QDIO 适配器到达主机。在 System/370 Extended Architecture/Interpretative Execution（“系统/370 扩展的体系结构/解释性执行”，IBM 出版物 No.SA22-7095-01，1985 年 9 月，其全部内容在此引入作为参考）中描述了解释性执行（例如，启动解释性执行（SIE））的一个例子。

为了配置一种能够使用 QEBSM 和 QEBSMI 的系统，执行了各种配置任务。如参照图 2a - 2b 所说明的，由程序（例如，客户程序）200、主机（例如，z/VM）202 以及所述客户和主机运行于其上的机器 204（例如，处理器）执行这些任务。

在一个实施例中，在步骤 206，客户程序 200（图 2a）判定程序是否能够使用 QEBSM 来管理 QDIO 连接。此判定是通过例如由程序向主机发出的通道子系统调用（CHSC）存储通道子系统特性命令来作出的。在步骤 208，主机又向硬件发出 CHSC 存储通道子系统特性指令，以便获得机器上已安装的工具。在步骤 210，所述机器处理 CHSC 指令，并且将信息传递回主机。

存储通道子系统特性命令被用于存储与 I/O 子系统的特性有关的信息。存储通道子系统特性命令的执行并不改变包含在 I/O 子系统内的任何信息。存储通道子系统特性命令被同步地执行，并且不是被解释地执行。

参照图 3a 说明了用于存储通道子系统特性命令的请求块的一个实施例。命令请求块 300 包括，例如：

- (a) 长度区 302: 该区指定命令请求块的长度。
- (b) 命令代码 304: 该区指定存储通道子系统特性命令。
- (c) 映像 Id (IID) 308: 该 IID 区包括多个 0 (作为一个例子)。
- (d) 格式 310: 命令请求格式区包括一个无符号整数，该无符号整数的值指定了命令请求块的布局。

参照图 3b 说明了用于存储通道子系统特性命令的响应块的一个实施例。在一个例子中，响应块 320 包括：

- (a) 长度区 322: 该区指定作为尝试执行存储通道子系统特性命令的结果被存储的命令响应代码的长度。

如果除了成功指示之外的响应代码被存储在响应代码区中，则没有特性信息作为尝试执行所述命令的结果被存储，并且该长度区为所述命令响应块指定例如 8 字节的长度。

如果指示成功的响应代码被存储在响应代码区中，则总体特性和 CHSC 命令特性都被存储在命令响应块中。因此，长度区指定例如 4080 字节的命令响应块长度。

- (b) 响应代码 324: 该区包括一个无符号二进制整数，该整数说明尝试执行存储通道子系统特性命令的结果。

- (c) 格式 326: 命令响应格式区包括一个无符号整数，该整数的值指定了命令响应块的布局。

(d) 总体特性 328: 如果指示成功的响应代码被存储在响应代码区中，则总体特性区包括指定各种工具中的哪些工具由 I/O 子系统来提供的信息。一些工具包括例如 CHSC 子通道；动态 I/O 配置工具；消息处理器通道配置工具；QDIO 增强缓冲区状态管理工具；QDIO 增强缓冲区管理解释 (QEBSMI) 工具等。虽然此处描述了多种工具，但是也可以存在其他的或不同的工具。当对应于特定工具的总体特性区的位是 1 时，表示提供该工具。

- (e) CHSC 特性 330: 如果成功响应代码被存储在响应代码区中，则 CHSC 特性区包括指定某些 CHSC 命令是否由 I/O 子系统来提供的信

息。这些命令的例子包括，例如，改变通道路径配置，改变控制单元配置，改变 I/O 设备配置，存储配置组件列表，存储域配置组件列表，以及设置域属性等。当对应于特定命令的 CHSC 特性区的位是 1 时，将提供该命令。

返回图 2a，在步骤 212，响应于主机接收到已安装工具的列表，该主机判定 QDIO 增强缓冲区状态管理解释 (QEBSMI) 工具是否已被安装在机器上。假如已安装了 QEBSMI 且子通道具有 QDIO 能力，则主机打开 CHSC 响应数据中的 QEBSM 可用指示器，指出 QEBSM 工具可用。因此，在步骤 214，以指出 QEBSM 工具可用来结束该命令。

接着，在步骤 216，作出关于被指定的子通道号的 QDIO 特性的判定。在一个例子中，通过由程序向主机发出 CHSC 存储子通道 QDIO 数据命令来作出此判定。然后，在步骤 218，主机将该指令发给实际的子通道，以便获得队列特性。在步骤 220，机器处理该指令。

存储子通道 QDIO 数据命令被用于获得与指定的子通道范围相关联的 QDIO 适配器的自述信息。

存储子通道 QDIO 数据命令的执行不改变包含在 I/O 子系统中的任意信息。在一个例子中，它被同步执行。

用于存储子通道 QDIO 数据命令的命令请求块 400 (图 4a) 包括例如：

- (a) 长度区 402: 该区指定命令请求块的长度。
- (b) 命令代码 404: 该区指定存储子通道 QDIO 数据命令。
- (c) 映像 ID (IID) 406: 在此实施例中，IID 区包括 0 值。
- (d) 格式 (FMT) 408: 命令请求格式区包括一个无符号整数，该整数的值指定命令请求块的布局。
- (e) 第一子通道号 410: 该区指定标识为其请求 QDIO 适配器信息的第一子通道的子通道号。
- (f) 最终子通道号 412: 该区指定标识为其请求 QDIO 适配器信息的最终子通道的子通道号。

参照图 4b 说明了用于所述存储子通道 QDIO 数据命令的响应块的一个实施例。在一个例子中，响应块 420 包括例如：

- (a) 长度区 422: 该区指定命令响应块的长度。该长度取决于作为尝试执行存储子通道 QDIO 数据命令的结果被存储的响应代码。

如果除了成功指示之外的响应代码被存储在响应代码区中，则没有子通道 QDIO 描述块作为尝试执行所述命令的结果被存储，并且该长度区为所述命令响应块指定例如 8 字节的长度。

如果指示成功的响应代码被存储在响应代码区中，则至少一个子通道 QDIO 描述块作为执行所述命令的结果被存储在命令响应块中，并且该长度区为被存储的每个子通道 QDIO 描述块指定例如 8 字节加 32 字节的命令响应块长度。

(b) 响应代码 424: 该区包括一个无符号二进制整数，该整数说明了尝试执行存储子通道 QDIO 数据命令的结果。

(c) 压缩 (C) 426: 该区指定子通道 QDIO 描述块是否被压缩。

当是 1 时，C 位指定描述块被压缩。即，只有子通道被指定为具有 QDIO 能力的那些描述块被存储。不具有 QDIO 能力的子通道的描述块不被存储在响应块中。

当是 0 时，C 位指定描述块没有被压缩。即，描述块被存储在响应块中而没有考虑它们是否描述了具有 QDIO 能力的子通道。对于与被请求的子通道范围内所存储的块同样多的描述块，由一个描述块来表示每个可能的子通道号。

(d) 格式 (FMT) 428: 命令响应格式区包括一个无符号整数，该整数的值指定命令响应块的布局。该区的值是，例如，0。

(e) 子通道 QDIO 描述块 (多个) 430: 当一个成功响应代码被存储时，多至 127 个 32-字节 (作为一个例子) 子通道 QDIO 描述块被存储在命令响应块的响应数据区域中。所存储的子通道 QDIO 描述块的数量取决于在命令请求块中指定的范围内的子通道的数量、当执行 CHSC 时 I/O 子系统内的情况以及 I/O 子系统模型。

至少一个子通道 QDIO 描述块被存储，并且通过从命令响应块的大小中减去 8 个字节并将剩余部分除以 32 来确定所存储的块的实际数量。

从由第一子通道号区所指定的子通道开始，为一些或所有被指定的子通道存储子通道 QDIO 描述块。按照子通道号升序来存储子通道 QDIO 描述块。

参照图 4c 描述了子通道 QDIO 描述块的一个例子。子通道 QDIO 描

述块 430 包括例如:

(a) 标志 440: 该区包括与子通道 QDIO 描述块的内容有关的信息。在一个例子中, 位 0-7 的含义如下:

位	含义
0	具有 QDIO 能力: 当是 1 时, 位 0 指出与指定的子通道关联的 I/O 设备能够执行 QDIO 操作。当是 0 时, 位 0 指出与指定的子通道关联的 I/O 设备不能够执行 QDIO 操作。
1	有效性: 当是 1 时, 位 1 指出子通道 QDIO 描述块有效。当是 0 时, 位 1 指出子通道 QDIO 描述块无效: 除了位 0、1 和 2, 以及子通道号区之外, 子通道 QDIO 描述块的内容没有意义。对于没有被指定为具有 QDIO 能力 (即当位 0 是 1 时) 的子通道, 位 1 是 0。
2	最后的 QDIO 子通道: 当是 1 时, 位 2 指出没有具有 QDIO 能力的更高编号的子通道。当是 0 时, 位 2 指出存在具有 QDIO 能力的更高编号的子通道。

(b) 子通道号 442: 该区指定向其应用子通道 QDIO 描述块中的信息的子通道。

(c) QDIO 队列格式 (QFMT) 444: 该区包括一个无符号二进制整数, 该整数标识了可被用在指定子通道上的 QDIO 队列的格式。

(d) 参数 446: 该区包括与由子通道 QDIO 描述块所表示的子通道或设备有关的参数信息。

(e) QDIO 适配器特性 1 (QDIOAC1) 448: 该区描述与指定子通道相关联的 QDIO 适配器的模型相关的特性。在一个例子中, 位 0-7 的含义如下:

位	含义
0	<p>未经请求的输入队列中断: 当是 1 时, 位 0 指出当缓冲区已转换到已准备好状态, 同时子通道处于 QDIO 有效状态时, QDIO 适配器可以动态地启动向程序给出未经请求的输入队列进展报告中断。这些中断独立于经请求的中断被给出, 所述经请求的中断作为程序已在输入队列 SBAL 中设置了输入队列中断请求指示器(如下面所述)的结果被给出(当支持输入队列中断请求时)。</p> <p>当是 0 时, 位 0 指出 QDIO 适配器没有启动向程序给出未经请求的输入队列进展报告中断。</p>
1	<p>SIGA 启动输入队列: 当位 1 是 1 并且位 7 是 0 时, QDIO 适配器要求程序执行 SIGNAL ADAPTER - 由子通道读 (SIGA-r)。当位 1 是 1 并且位 7 是 1 时, QDIO 适配器要求程序执行 SIGNAL ADAPTER - 由标记读 (SIGA-rt)。在 2001 年 12 月 18 日公布的标题为“Self-contained Queues With Associated Control Information For Receipt And Transfer Of Incoming And Outgoing Data Using A Queued Direct Input-output Device”(用于使用排队直接输入-输出设备来接收和传输进入和外出数据的具有关联控制信息的自足队列)的美国专利 No.6,332,171 中; 在 2002 年 2 月 5 日公布的标题为“Method And Apparatus Of Simulation Of Data In A Virtual Environment Using A Queued Direct Input-Output Device”(在虚拟环境中使用排队直接输入-输出设备来模拟数据的方法和装置)的美国专利 No.6,519,645 中; 以及在 2003 年 2 月 11 日公布的标题为“Method And Apparatus For Providing Configuration Information Using A Queued Direct</p>

Input-Output Device”（使用排队直接输入-输出设备来提供配置信息的方法和装置）的美国专利 No.6,345,241 中（所有这些专利的全部内容在此引入作为参考）说明了 SIGA 的一个例子。

当是 0 时，位 1 指定 QDIO 适配器不要求程序执行 SIGA-r 或 SGIA-rt。

- 2 SGIA 启动或处理输出队列：当位 2 是 1 并且位 7 是 0 时，出于指令的说明中所指出的原因，QDIO 适配器要求程序执行 SIGNAL ADAPTER - 由子通道写（SIGA-w）。当位 2 是 1 并且位 7 是 1 时，QDIO 适配器要求程序执行 SIGNAL ADAPTER - 由标记写（SIGA-wt）。

当是 0 时，位 2 指定 QDIO 适配器不要求程序执行 SIGA-w 或 SIGA-wt。

- 3 SIGA 同步：当位 3 是 1 并且位 7 是 0 时，QDIO 适配器要求程序执行 SIGNAL ADAPTER - 由子通道同步（SIGA-s）。当位 3 是 1 并且位 7 是 1 时，程序可以执行 SIGNAL ADAPTER - 由标记注销（SIGA-lt）。

当是 0 时，位 3 指定 QDIO 适配器不要求程序执行 SIGA-s。

- 4 适配器中断队列同步辅助有效：当位 4 是 1 时，适配器中断队列同步辅助有效。当该辅助有效且 SIGA-s 被请求（位 3 是 1 并且位 7 是 0）时，与适配器（其与指定的子通道关联）相关联的输入和输出队列将在中断之前被动态地同步。因此，不需要程序在适配器中断设备（AIF）中断之后立即发出 SIGA-s 指令。

当位 4 是 0 时，适配器中断队列同步辅助无效。

- 5 PCI 中断队列同步辅助有效：当位 5 是 1 时，PCI

中断队列同步辅助有效。当该辅助有效且 SIGA-s 被请求（位 3 是 1 并且位 7 是 0）时，与指定子通道相关联的输出队列将在中断之前被动态地同步。因此，不需要程序在 PCI 中断之后立即发出 SIGA-s 指令。

当位 5 是 0 时，PCI 中断队列同步辅助无效。

QDIOAC1 区的位 4 和 5 由主机来设置。

- 6 QEBSM 可用：当是 1 时，位 6 指出 QDIO 增强缓冲区状态管理可用于允许指定的子通道。

当是 0 时，位 6 指出 QDIO 增强缓冲区状态管理不可用于允许指定的子通道。

基于响应 CHSC 存储通道子系统特性命令而被存储的总体特性的指定位的设置，由主机来设置这个区。

- 7 QEBSM 允许：当是 1 时，位 7 指出指定的子通道被允许 QDIO 增强缓冲区状态管理，结果，以下为真：

- 与指定的子通道相关联的 QDIO 缓冲区状态不能在程序存储中的在下面所述的 SLSB 中被访问。而是，程序使用设置 QDIO 缓冲区状态和提取 QDIO 缓冲区状态指令。

- 子通道标记区包括子通道标记。

当是 0 时，位 7 指出指定的子通道没有被允许 QDIO 增强缓冲区状态管理，并且指出在程序存储（其可以由程序直接操作）中的 SLSB 中可以访问 QDIO 缓冲区状态。

基于指定的子通道是否被允许 QDIO 增强缓冲区状态管理解释（QEBSMI），由主机来设置这个区。当主机完全模拟适配器时（例如，在客户跟踪期间），位 7 也可以由主机来设置。（要了解 QDIO 子通道如何被允许 QEBSMI 的说明，请参考下面描述的

CHSC 设置通道子系统特性命令。)

(f) 类 450: 该区包括一个无符号二进制整数, 该整数标识了与指定子通道相关联的 I/O 设备的类。类的一个例子是通信控制器。

(g) 输入队列计数 (IQCNT) 452: 该区包括一个无符号二进制整数, 该整数指定了由 QDIO 适配器使用以执行 I/O 操作的输入 QDIO 队列的数量。

(h) 输出队列计数 (OQCNT) 454: 该区包括一个无符号二进制整数, 该整数指定了由 QDIO 适配器使用以执行 I/O 操作的输出 QDIO 队列的数量。

(i) 最大缓冲区链计数 (MBCCNT) 456: 当这个区不是 0 时, 它包括适配器所支持的链接缓冲区 (即, 下面所述的链接的 SBAL) 的最大数目的无符号整数计数。当这个区是 0 且 QFMT 区指出格式 1 队列 (即, FCP) 被用在指定子通道上时, 最多 36 个缓冲区可以被链接。当这个区是 0 且 QFMT 区指出格式 1 队列没有被用在指定子通道上时, 适配器不支持缓冲区链接。

(j) QDIO 适配器特性 2 (QDIOAC2) 458: 该区描述与指定子通道相关联的 QDIO 适配器的模型相关的特性。在一个例子中, 位 0-3 的含义如下:

位	含义
0	输入队列中断请求支持: 当是 1 时, 位 0 指出适配器支持下面所述的输入队列 SBAL 的 SBALF 0 区中的中断请求位。当位 0 是 1 并且位 3 是 0 时, 报告适配器在关联输入队列上的进展的中断严格地遵循程序的请求。当位 0 是 1 并且位 3 是 1 时, 报告适配器在关联输入队列上的进展的中断遵循程序的请求, 并且还可以发生未经请求的报告进展的中断。当是 0 时, 位 0 指出适配器不支持用于输入队列的 SBAL 的 SBALF 0 区中的中断请求位。
1	输出队列中断请求支持: 当是 1 时, 位 1 指出适配器支持输出队列 SBAL 的 SBALF 0 区中的中断请求

位。报告适配器在关联输出队列上的进展的中断严格地遵循程序的请求。

当是 0 时，位 1 指出适配器不支持用于输出队列的 SBAL 的 SBALF 0 区中的中断请求位。

2 单缓冲区预取有效：当是 1 时，位 2 指出适配器在该适配器正在处理的当前缓冲区之前最多预取一个缓冲区。当是 0 时，位 2 指出适配器可以预取多于一个的缓冲区。

3 未经请求的输入队列中断无效：当是 1 时，位 3 指出未经请求的输入队列中断对于适配器是无效的。当是 0 时，位 3 指出未经请求的输入队列中断对于适配器是有效的。当位 3 是 0 时，QDIOAC1 位 0 是 1。

(k) 子通道标记 460：当 QDIOAC1 区的位 7 是 1 时，该区包括与指定的子通道相关联的标记。

子通道标记区由主机来设置。如果指定的子通道允许 QEBSMI，则主机将子通道标记区设置为与该子通道关联的 QDIO 缓冲区信息控制块 (QBICB) 的主机实际或绝对地址的值。(如下所述，当 QDIO 子通道允许 QEBSMI 时，通过 CHSC 设置通道子系统特性命令将 QBICB 与该子通道相关联。)当主机完全模拟适配器时，也可以由主机来设置子通道标记。如果指定的子通道不允许 QEBSMI 且主机没有模拟适配器，则将 0 作为子通道标记值返回。

子通道标记被用于为 QEBSM 操作指定 QDIO 子通道。当发生导致从子通道解散 QDIO 队列的活动或事件时，子通道标记变成是过时的，并且在程序可以再次使用 QEBSM 操作之前，要获得新的子通道标记。先前分配的标记变成无效，并且可以被随后分配给另一个子通道。

在存储子通道 QDIO 数据命令被执行的时刻，存储在子通道 QDIO 描述块中的适配器描述信息是最新的。然而，作为由指定的适配器成功完成以下所述的建立 QDIO 队列 CCW 命令的结果，某些适配器描述信息可能会改变。因此，在成功完成建立 QDIO 队列 CCW 命令之后，要重新发出

存储子通道 QDIO 数据命令，以便获得以下内容的最新说明：

- QDIOAC1 位 1 (SIGA 启动输入队列)
- QDIOAC1 位 2 (SIGA 启动或处理输出队列)
- QDIOAC1 位 3 (SIGA 同步)
- QDIOAC1 位 7 (QEBSM 允许)
- QDIOAC2 位 2 (单缓冲区预取有效)
- QDIOAC2 位 3 (未经请求的输入队列中断无效)
- 子通道标记

剩余的适配器描述信息是最新的，与建立 QDIO 队列命令的成功完成无关。

返回图 2a, 步骤 222, 响应 CHSC 存储子通道 QDIO 数据命令的完成, 在步骤 224 主机判定 QEBSMI 工具是否在机器上可用。如果 QEBSMI 可用, 则子通道具有 QDIO 能力, 并且主机支持用于这种连接的 QEBSM, 然后在返回给程序的 CHSC 响应数据中作出 QEBSM 可被用于该子通道的指示。在步骤 226, 程序接收具有子通道 QDIO 数据的结束操作。

在步骤 228, 采用在 CHSC 响应数据中被返回的信息, 在主存储器中由程序来创建 QDIO 队列。QDIO 队列包括描述队列的数据结构以及用于数据传输的缓冲区存储器块。在一个例子中, 共同描述队列的特性并提供控制, 以便允许在程序和适配器之间交换数据的多个单独的数据结构 (被称作队列组件) 包括, 例如:

1. 队列信息块 (QIB) 500 (图 5): 每个 QDIO 子通道定义一个 QIB。QIB 提供了与子通道相关联的输入和输出队列的集合的有关信息。通过此处说明的建立 QDIO 队列命令, QIB 的绝对地址被提供给 QDIO 适配器。QIB 具有多个区, 包括例如下述部分, 其中一些部分在图 5 中示出:

(a) 队列格式 (QFMT): 该区包括一个无符号二进制整数, 其标识了 QDIO 队列的格式。

(b) 实现相关的参数格式 (PFMT): 该区包括一个无符号二进制整数, 其标识了实现相关的参数的内容的格式。

(c) 请求标志 (RFlags): 该区包括请求标志。一个标志包括 QDIO 增强缓冲区状态管理 (QEBSM) 允许。当这个位在发出建立

QDIO 队列 CCW 命令的时刻为 1 时，则指出为 QEBSM 启动子通道的请求。

(d) 适配器特性 (AC)：该区包括与适配器的能力有关的信息（就与子通道相关联的输入和输出队列的集合而言）。在建立 QDIO 队列 CCW 命令被发出但未被适配器检查之前，由程序将该 AC 区设置为 0。

在建立 QDIO 队列 CCW 命令的执行期间，由适配器来设置该区，并且该区被设置为指出在成功完成该命令之前适配器所支持的值。在一个例子中，位 0-1 的含义如下：

位	含义
0	输入队列中断请求支持：当是 0 时，位 0 指出适配器不支持用于输入队列的 SBAL 的 SBALF 0 区（下面所述）中的中断请求位。 当是 1 时，位 0 指出适配器支持 SBAL 的 SBALF 0 区中的中断请求位。
1	输出队列中断请求支持：当是 0 时，位 1 指出适配器不支持用于输出队列的 SBAL 的 SBALF 0 区中的中断请求位。不给出报告适配器在关联的输出队列上的处理的中断。 当是 1 时，位 1 指出适配器支持 SBAL 的 SBALF 0 区中的中断请求位。

(e) 第一输入存储器列表信息块 (SLIB) 地址 502：该区包括用于 QIB 被与其关联的子通道的最高优先权输入队列 SLIB 的逻辑地址。

(f) 第一输出 SLIB 地址 504：该区包括用于 QIB 被与其关联的子通道的最高优先权输出队列 SLIB 的逻辑地址。

(g) 适配器标识符：该区包括程序分配给用于子通道 (QIB 被与其关联) 的 QDIO 适配器的名称。

(h) 实现相关的参数：该区包括一个或多个实现相关的参数。

2. 存储器列表信息块 (SLIB) 506：为每个队列定义一个 SLIB。SLIB

提供了与队列以及在保存数据中可用的每个队列缓冲区有关的信息。通过建立 QDIO 队列命令，SLIB 的绝对地址被提供给 QDIO 适配器。

在一个例子中，存储器列表信息块包括以下区，其中的一些区在图 5 中示出：

(a) 下一个存储器列表信息块地址 508：如果对于同一子通道，存在用于相同类型的 QDIO 队列（输入或输出）的其他更低优先权的 SLIB，则该区包括该 SLIB 的逻辑地址。

(b) 存储器列表地址 510：该区包括被称作存储器列表的控制块的逻辑地址。存储器列表包括存储器块地址列表（SBAL）的绝对地址。每个 SBAL 包括最多例如 16 个存储器块（例如，页）的绝对地址。能够使用单个 SBAL 中的地址被定位的存储器块的集合被称作 QDIO 缓冲区。

(c) 存储器列表状态块地址（SLSB）512：当 QIB RFlags 区中的增强 QDIO 缓冲区状态管理允许指示器是 0 时，该区包括存储器列表状态块（SLSB）的逻辑地址。

3. 存储器列表（SL）514：为每个队列定义一个 SL。SL 包括例如 128 个表项，与队列关联的每个 QDIO I/O 缓冲区一个表项。SL 提供了关于 I/O 缓冲区在主存储器中的位置的信息。每个表项包括存储器块地址列表 516 的绝对地址，并且每个 SBAL 包括共同构成与该队列相关联的 128 个数据缓冲区中的一个缓冲区的存储器块的绝对地址列表。

通过建立 QDIO 队列命令，存储器列表的绝对地址被提供给 QDIO 适配器。

4. 存储器列表状态块（SLSB）或存储器列表状态注销（SLSL）518：当 QDIO 子通道被允许增强的 QDIO 管理操作时，为每个队列定义一个 SLSL，并且对于程序存储之外的每个队列，还存在相应的 SLSB。SLSL 长度为 128 字节，并且包括 SLSB 的注销（无论何时 Signal Adapter-由标记注销（SIGA - It）被执行时或当队列被从适配器解散时作出）。SLSB 包括 128 个表项 520，被称作 SQBN（队列缓冲区 N 的状态），与队列相关联的 128 个 I/O 缓冲区中的每个缓冲区一个表项。每个表项提供了关于其关联的 I/O 缓冲区的状态信息。本文中说明的设置 QDIO 缓冲区状态和

提取 QDIO 缓冲区状态指令被用于访问 SLSB 中的状态信息。作为例子，所述状态指出缓冲区是程序拥有还是适配器拥有的，以及缓冲区是输出缓冲区还是输入缓冲区。

通过建立 QDIO 队列命令，存储器列表状态块的绝对地址被提供给 QDIO 适配器。

5. 存储器块地址列表 (SBAL) 522: 为与每个队列关联的 128 个 I/O 缓冲区中的每一个缓冲区定义一个 SBAL。SBAL 为 I/O 缓冲区提供关于存储器块 (SB) 位置的信息。

存储器块地址列表包括 16 个被称为存储器块地址列表表项 (SBALE) 524 的表项。每个 SBALE 包括存储器块 526 的绝对存储地址。由单个 SBAL 的表项寻址的存储器块共同构成 QDIO 队列的 128 个可能的 QDIO 缓冲区中的一个缓冲区。

在一个例子中，存储器块地址列表表项包括：

- (a) 标志：该区包括关于 SBALE 内容的信息。
- (b) 虚拟化标志：当安装了 QDIO 增强缓冲区状态管理解释工具时，该区包括虚拟化信息。对该区的更改应在 QSD 锁定字的控制下被串行化。在一个例子中，该区的位 0 表示已钉住页指示器。当位 0 是 1 时，对应于数据地址区中的数据地址的主机页为此 SBALE 被钉住。当位 0 是 0 时，对应于数据地址区中的数据地址的主机页没有为此 SBALE 被钉住。
- (c) SBAL 标志 (SBALF)：该区包括与包含该 SBALE 的 SBAL 相关联的缓冲区有关的信息，并不只是关于与该 SBALE 相关联的存储器块。在一个例子中，位 0 指出当缓冲区处于程序拥有状态，或是接着返回程序拥有状态时，在该缓冲区执行下一个 SQBS 操作之前将为该缓冲区执行 EQBS 操作。位 1 指出当缓冲区处于程序拥有状态，或是接着返回程序拥有状态时，作为为该缓冲区执行的下一个 EQBS 操作的一部分，该 SBAL 及其对应的客户 SBAL 将被同步。缓冲区同步可以包括从对应的主机 SBALE 来更新客户 SBALE，并且还可以包括解除主机页的钉住。

- (d) 数据计数: 该区包括一个无符号二进制整数, 该整数指定了要在数据地址区中指定的地址开始处读或写的数据的字节数。
- (e) 数据地址: 该区包括存储器块 (包括将被写 (输出队列) 的数据) 的绝对地址或将被读 (输入队列) 的数据的地址。

6. 存储器块 (SB) 526: 定义了从 1 到 16 个单独存储器块 (例如页), 它们共同定义了单个 I/O 缓冲区。

在标题为 “Method And Apparatus For Providing Configuration Information Using A Queued Direct Input-Output Device” (“使用排队直接输入/输出设备来提供配置信息的方法和装置”, Markos 等人, 2003 年 2 月 11 日公开, 其全部内容在此引入作为参考) 的美国专利 No.6,519,645 以及在标题为 “Method And Apparatus For Simulation Of Data In A Virtual Environment Using A Queued Direct Input-Output Device” (“在虚拟环境中使用排队直接输入-输出设备来模拟数据的方法和装置”, Brice 等人, 2002 年 2 月 5 日公开, 其全部内容在此引入作为参考) 的美国专利 No.6,345,241 中提供了关于 QDIO 队列的其他详细信息。

返回图 2a, 在创建队列之后, 在步骤 230, 如果子通道支持 QEBSM, 则 QIB 中的 QEBSM 允许位被设置, 以便使用 QEBSM 来管理连接。

此后, 在步骤 232, 在适配器中建立 QDIO 队列。作为一个例子, 这通过程序发出带有建立 QDIO 队列通道命令字的启动子通道指令来完成, 在标题为 “Self-contained Queues With Associated Control Information For Receipt And Transfer Of Incoming And Outgoing Data Using A Queued Direct Input-output Device” (“用于使用排队直接输入-输出设备来接收和传输进入和外出数据的具有关联控制信息的自足队列”, Baskey 等人, 2001 年 12 月 18 日公布, 其全部内容在此引入作为参考) 的美国专利 No.6,332,171 中说明了进行这种处理的一个例子。

建立 QDIO 队列命令被用于将队列描述符记录 (QDR) 写到 I/O 设备。与 I/O 设备相关联的控制单元 (例如, 适配器) 使用队列描述符记录中的信息来定位被用于建立具有优先权的数据队列的 QDIO 控制块。随后这些队列被用于向 I/O 设备或从 I/O 设备直接传输数据。

在步骤 234, 建立过程使主机为由程序建立的队列在主机存储器中创

建一组影子 QDIO 队列。即，主机建立某些 QDIO 控制结构的影子版本，这些 QDIO 控制结构包括例如存储器列表和存储器块地址列表。同时，由于 SLSB 被定义为程序和适配器协同处理的控制机制，虽然它不包含地址，SLSB 也被在主存储器中分配。与建立 QDIO 队列命令相关联的队列描述符记录包括如程序所见的 QDIO 队列组件的主存储器地址。主机在建立其自己的队列描述符记录和影子 SL 和 SBAL 的拷贝时，对这些地址以及 SL 和 SBAL 中的地址进行转换。关于与 QDR 相关联的建立影子和处理的其 他详细信息被包括在其全部内容在此引入作为参考的上述美国专利中。

在创建影子队列之后，在步骤 236，如果在 QIB 中 QEBSM 允许位被设置，则主机在主存储器中创建一个被称为 QDIO 缓冲区信息控制块 (QBICB) 的结构，以便将客户和主机的影子 QDIO 队列映射到机器。QBICB 包括描述客户和主机结构的指针和信息。其与 QDIO 子通道相关联，以便当该子通道被标识为 SQBS 和 EQBS 的运算对象时(如下面所述)，客户和主机结构可以被定位和使用。参考图 6a - 6b 说明了关于 QBICB 的其他详细信息。

QDIO 缓冲区信息控制块 (QBICB) 600 (图 6a) 包括关于 QDIO 数据队列集合的缓冲区存储和缓冲区状态的信息，并且被用于将缓冲区存储和缓冲区状态与给定的 QDIO 子通道相关联。例如，在主机的实际或绝对存储器中，它是一个以双字边界开始并具有 64 字节的最小长度的变长控制块。如其队列计数区所指定的那样，QBICB 的大小由其描述的队列的数目来确定 (即，由它包括的队列描述符的数目来确定)。

在一个例子中，QBICB 600 包括以下区：

(a) 队列格式 (QFMT) 602: 该区包括一个无符号二进制整数代码值，该代码值标识了队列描述符的格式和由该队列描述符的内容寻址的控制块的格式。作为例子，格式 0 表示 OSA 直接表达队列，格式 1 表示光纤通道协议 (FCP) 队列，并且格式 2 表示内部排队直接通信队列。

(b) 队列计数 (QCNT) 604: 该区包括一个无符号整数，该整数指定了由 QBICB 描述的输入和输出队列的总的计数。

(c) 子系统 ID 606: 该区包括与 QBICB 相关联的 QDIO 子通道

的子系统 ID。

(d) 队列描述符 (多个) 608: 队列描述符是例如一个 16 字节的结构, 并且在一个实施例中 QBICB 可以包括多至 64 个队列描述符。每个队列描述符包括队列存储器描述符的地址, 所述队列存储器描述符被用于为一个 QDIO 输入或输出队列来描述缓冲区状态和缓冲区数据存储器。

用于输入队列 (如果有) 的队列描述符被连续地列出。用于输出队列 (如果有) 的队列描述符在用于输入队列的队列描述符的列表之后被连续地列出。QCNT 区中的值指定了队列描述符的总数。

QBICB 队列描述符包括队列存储器描述符起始地址区, 该区以及被附加在右侧的 7 个 0 为相应的队列指定了队列存储器描述符的 64 位主机实际或绝对地址。

队列存储器描述符是例如一个 128 字节的结构, 其包括与用于相应的队列的主机和客户缓冲区存储器有关的信息。在一个例子中, 队列存储器描述符 620 (图 6b) 包括下面的区:

(a) 锁定字 622: 锁定字区是例如一个 32 位的区, 它是用于队列存储器描述符和有效请求头部 FOB 列表 (在下面说明) 的互锁控制。取决于其预定的值, 锁定字值可以具有下面含义中的一种含义: 可用: 与锁关联的结构可由机器或主机使用; 为机器锁定: 机器已经获得了锁和对相关联的结构的所有权; 由机器为主机锁定: 机器遇到了需要来自主机的服务的情况, 并且已经将对锁和相关联的结构的所有权传递给了主机; 以及由主机为主机锁定: 主机已经获得了锁和对相关联的结构的所有权。

(b) 队列标志 (QFlags) 624: 该区包括例如 8 位的标志区。
标志的例子包括:

位	含义
0	队列/缓冲区类型: 当位 0 是 0 时, 队列是输入队列。当位 0 是 1 时, 队列是输出队列。

1

未决请求组件 **FOB** 列表完成: 当位 1 是 1 时, 由第一个有效请求头部 **FOB** 地址区标识的 **FOB** 是一个多缓冲区请求的头部 **FOB**, 对于它来说, 请求组件 **FOB** 列表是不完整的, 并且仍然在被构造。

(c) **SLSB** 地址 626: 该区包括与相应的队列关联的存储器列表状态块 (**SLSB**) 的主机实际或绝对地址。

(d) **主机 SL** 地址 628: 该区包括与相应的队列关联的主机存储器列表 (**SL**) 的主机实际或绝对地址。由该地址指定的 **SL** 可以被适配器观察到。

(e) **客户 SL** 地址 630: 该区包括与相应的队列关联的客户存储器列表 (**SL**) 的客户绝对地址。由该地址指定的 **SL** 不能被适配器观察到。

(f) **客户 SLSL** 地址 632: 该区包括与相应的队列关联的存储器列表状态注销 (**SLSL**) 区域的客户绝对地址。

(g) **状态转换计数** 634: 该区包括对由程序观察到的 (通过提取 **QDIO** 缓冲区状态指令) 从适配器拥有状态到程序拥有状态的缓冲区状态转换数目的无符号整数计数。

(h) **字节传输计数** 636: 该区包括对由适配器传输的字节数目的无符号整数计数。对于光纤通道协议 (**FCP**) 队列, 字节传输计数不是实际传输的字节的计数。而是对指定将被传输的字节的计数。

(i) **第一个有效请求头部 FCP 操作块 (FOB) 地址** 638: 当 **QBICB** 中的 **QFMT** 区指定队列是格式 1 队列 (即, **FCP**) 并且该区的预定位 (带有被附加在右侧的 8 个 0) 不形成该区的第一个字节的主机实际或绝对地址时, 作为一个例子, 该区的预定位 (带有被附加在右侧的 8 个 0) 形成有效 **FCP** 请求列表中的第一个 **FCP**

请求的头部 FOB 的主机实际或绝对地址（即，有效请求头部 FOB 列表中的第一个 FCP 请求的头部 FOB 的地址）。当 QBICB 中的 QFMT 区指定队列是格式 1 队列并且预定位（带有被附加在右侧的 8 个 0）形成该区的第一个字节的主机实际或绝对地址时，作为一个例子，不存在有效 FCP 请求并且有效请求头部 FOB 列表为空。

(j) 最后的有效请求头部 FOB 地址 640: 当 QBICB 中的 QFMT 区指定队列是格式 1 队列并且该区的预定位（带有被附加在右侧的 8 个 0）不形成区 638 的第一个字节的主机实际或绝对地址时，作为一个例子，该区的预定位（带有被附加在右侧的 8 个 0）形成有效 FCP 请求列表中的最后一个 FCP 请求的头部 FOB 的 64 位主机实际或绝对地址（即，有效请求头部 FOB 列表中的最后一个 FCP 请求的头部 FOB 的地址）。当 QBICB 中的 QFMT 区指定队列是格式 1 队列并且该区的预定位（作为一个例子，带有被附加在右侧的 8 个 0）形成区 638 的第一个字节的主机实际或绝对地址时，不存在有效 FCP 请求并且有效请求头部 FOB 列表为空。

光纤通道协议操作块（FOB）描述了在 FCP I/O 操作期间被钉在主机页帧中的多至例如 16 个的主机页。在一个例子中，FCP 操作块（FOB）648（图 6c）包括下面的区：

(a) 下一个请求头部 FOB 地址 650: 如果 FOB 是有效请求头部 FOB，该区以及被附加在右侧的 8 个 0 形成了有效请求头部 FOB 列表中的下一个 FCP 请求的头部 FOB 的主机实际或绝对地址，或者如果该 FOB 是有效请求头部 FOB 列表中的最后的头部 FOB，则形成 QSD 中的第一个有效请求头部 FOB 地址区的头部 FOB 的主机实际或绝对地址。如果该 FOB 在可用 FOB 列表中并且该区不是 0，则该区以及被附加在右侧的 8 个 0 形成可用 FOB 列表中的下一个 FOB 的主机实际或绝对地址。如果该 FOB 在可用 FOB 列表中并且该区是 0，则该 FOB 是可用 FOB 列表中的最后的 FOB。如果该 FOB 是除头部 FOB 之外的请求组件 FOB，则该区的内容没有意义。

(b) 前一个请求头部 FOB 地址 652: 如果该 FOB 是有效请求头部 FOB，则该区以及被附加在右侧的 8 个 0 形成了有效请求头部 FOB

列表中的前一个 FCP 请求的头部 FOB 的主机实际或绝对地址，或者如果该 FOB 是有效请求头部 FOB 列表中的第一个头部 FOB，则形成 QSD 中的第一个有效请求头部 FOB 地址区的头部 FOB 的主机实际或绝对地址。如果该 FOB 在可用 FOB 列表中或是除头部 FOB 之外的请求组件 FOB，则该区的内容没有意义。

(c) 下一个请求组件 FOB 地址 654: 如果该 FOB 是请求组件 FOB 并且该区不是 0，则该区以及被附加在右侧的 8 个 0 形成了请求组件 FOB 列表中的下一个 FOB 的主机实际或绝对地址。如果该 FOB 是请求组件 FOB 并且该区是 0，则该 FOB 是请求组件 FOB 列表中的最后的 FOB 或是仅有的 FOB。

(d) 队列存储器描述符(QSD)地址 656: 如果该 FOB 是头部 FOB，则该区以及被附加在右侧的 7 个 0 形成了输出队列的队列存储器描述符的主机实际或绝对地址，用于请求的头部 FOB 被从所述输出队列排队。

(e) QDIO 缓冲区信息控制块(QBICB)地址 658: 如果该 FOB 是头部 FOB，则该区以及被附加在右侧的 3 个 0 形成与子通道(FCP 操作对于其有效)关联的 QDIO 缓冲区信息控制块的主机实际或绝对地址。

(f) 客户请求 ID 660: 如果 FOB 是头部 FOB，则该区包括由其 SBAL 中的客户指定的 FCP 请求 id。

(g) 标志 662: 当 FOB 是头部 FOB 时，该区包括例如一个 8 位的标志区。这些位和它们的含义的例子包括，例如：

- 1) 位 0 - 请求有效: 位 0 指出该头部 FOB 在有效请求头部 FOB 列表中，并且代表一个有效的 FCP 请求；
- 2) 位 1 - 构造完成: 位 1 指出包括 FCP 请求的缓冲区已经被处理，并且完成了所述 FCP 请求的构造；
- 3) 位 2 - 请求 ID 已投递: 位 2 指出适配器已经投递了用于所述请求的 FCP 请求描述符；
- 4) 位 3 - 4 存储器块类型: 位 3 - 4 为由 FOB 及其相应 SBAL 所描述的格式 1 输出缓冲区指定存储器块类型。位 3 - 4 与相应

的 SBAL 中的 SBALF 0 区的位 3-4 具有相同的值和含义。

(h) 页表表项 (PTE) 计数 (PTECNT) 664: 该区包括有意义的 PTE 地址数组表项的计数。

(i) PTE 地址 666: 该区包括一个 0-16 个 PTE 地址区的数组。每个 PTE 地址区的位以及被附加在右侧的 3 个 0 形成 PTE 的主机实际或绝对地址, 其为 FOB 所属的请求被钉住。

各 FOB 可以彼此链接以形成各种列表。一种列表是可用 FOB 列表。当 FOB 已被由主机分配并且未被用于 QDIO 操作时, 它驻留在可用 FOB 列表 (AFOBL) 中。在一个例子中, 为每个处理器提供一个单独的可用 FOB 列表。可用 FOB 列表中的各 FOB 被使用下一个请求头部 FOB 地址区来单独地链接。列表中的最后一个表项在其下一个请求头部 FOB 地址区中被指定为 0 值。可用 FOB 列表的起始地址和可用 FOB 列表中的各 FOB 的计数由 CPU 前置区域中的各区来指定。

另一种列表是请求组件 FOB 列表, 它包括构成单个有效 FCP 请求的一个或多个 FOB。请求组件 FOB 列表中的 FOB 的数目对应于 FCP 请求中的 SBAL 的数目。请求组件 FOB 列表中的第一个 FOB 是头部 FOB, 并且被用于表示有效请求头部 FOB 列表中的有效请求的存在。

对于单个 SBAL (即, 单个缓冲区) FCP 请求, 请求组件列表由头部 FOB 组成。在头部 FOB 被构造之后, 它被放置在有效请求头部 FOB 列表的开头, 并且 FOB 的标志区中的构造完成和请求有效指示器都被设置为 1, 以便指出完成了整个 FCP 请求的构造, 并且请求是有效的。

对于一个多 SBAL (即多缓冲区) FCP 请求, 相应于包括该请求的每个 SBAL 来构造一个 FOB。被构造的第一个 FOB 成为头部 FOB。头部 FOB 被添加到有效请求头部 FOB 列表的开头, 并且在它被构造之后, FOB 的标志区中的请求有效位被设置为 1, 以便指出请求是有效的。其余的请求组件 FOB 的列表的锚 (anchor) 保存在头部 FOB 中。

其余的 FOB 被添加到请求组件 FOB 列表, 并且从头部 FOB 开始被使用下一个请求组件 FOB 地址区来单独地链接。列表中的最后的 FOB 在其下一个请求组件 FOB 地址区中被指定为 0 值。在最后的 FOB 被添加到列表之后, 头部 FOB 的标志区中的构造完成指示器被设置为 1, 以便指出

完成了整个 FCP 请求的构造，并且请求是有效的。

包括头部 FOB 的请求组件 FOB 列表中的各 FOB 被看作请求组件 FOB。

另一种列表是有效请求头部 FOB 列表。当 FCP 请求是有效的时，该请求的头部 FOB 在有效请求头部 FOB 列表中。有效请求头部 FOB 列表是使用下一个请求头部 FOB 地址和前一个请求头部 FOB 地址区的双重链接列表。有效请求头部 FOB 列表的锚是用于该队列的队列存储器描述符中的第一个有效请求头部 FOB 地址和最后的有效请求头部 FOB 地址区。该列表中的第一个表项由以下两个值来指定：头部 FOB 的前一个请求头部 FOB 地址区中的队列存储器描述符的第一个有效请求头部 FOB 地址区的地址值，以及该队列存储器描述符的第一个有效请求头部 FOB 地址中的第一个表项的头部 FOB 的地址值。该列表的最后一个表项由以下两个值来指定：头部 FOB 的下一个请求头部 FOB 地址区中的队列存储器描述符的第一个有效请求头部 FOB 地址区的地址值，以及队列描述符的最后的有效请求头部 FOB 地址中的最后表项的头部 FOB 的地址值。

回到图 2a，在创建 QBICB 之后，在步骤 238，由主机向机器发出启动子通道指令（带有建立 QDIO 队列通道命令字），以便将用于正在被建立的连接的影子 QDIO 队列的位置通知给适配器。在步骤 240，机器将建立 QDIO 队列命令发送到适配器，并且在步骤 242，主机等待完成（图 2b）。在步骤 244，机器发出 I/O 完成中断，在步骤 246，由主机接收该中断，其反应了虚拟 I/O 完成中断。在步骤 248，由程序 200 接收用于建立 QDIO 队列命令的 I/O 完成中断。

此后，在步骤 250，程序获取用于 QEBSM 连接的子通道标记和管理该连接所需的（或所希望的）SIGA 的类型。在一个例子中，这是通过发出上面所述的 CHSC 存储子通道 QDIO 数据命令来完成的。

在步骤 252，主机将 QBICB 与所述子通道关联。在一个例子中，通过发出设置通道子系统特性命令来完成这种关联。CHSC 设置通道子系统特性命令基于用于该命令的命令请求块的操作码来设置所选择的通道子系统控制。

参考图 7a 说明了命令请求块 700 的一个例子。该请求块包括，例如：

- (a) 长度区 702: 该区指定了命令请求块的长度。
- (b) 命令码 704: 该区指定了设置通道子系统特性命令。
- (c) 操作码 (OC) 706: 该区包括一个值, 该值指定了要被执行的操作。作为一个例子, 为 0 的 OC 指定设置适配器设备控制操作。执行该设置适配器设备控制操作要求为指定的子通道允许并设置指定的适配器设备控制, 或为子通道禁止指定的适配器设备控制。适配器设备控制是, 例如:

- 设备状态改变指示器的地址;
- 用于设备状态改变指示器的存储器保护关键字;
- 适配器本地概要指示器的地址;
- 用于适配器本地概要指示器的存储器保护关键字;
- 中断子类码;
- 客户中断子类码; 以及
- QBICB 的起始地址。

(d) 格式 (FMT) 708: 命令请求格式区包括一个无符号整数, 该整数的值指定了命令请求块的布局。该区的值的一个例子是 0。

(e) 操作数据区域 710: 命令请求块的这个区域可以被操作使用。操作数据区域的数据格式特定于每个操作码。

作为一个例子, 设置适配器设备控制操作的操作数据区域包括, 例如, 将参考图 7b 说明的下面的区:

- (a) 适配器本地概要指示器 (ALSI) 地址 720: 该区包括在用于指定子通道的存储器中的适配器本地概要指示器字节的绝对地址。
- (b) 设备状态改变指示器 (DSCI) 地址 722: 该区包括在用于指定子通道的存储器中的设备状态改变指示器字节的绝对地址。
- (c) 适配器本地概要指示器关键字 (KS) 724: 该区包括存储器访问关键字, 所述存储器访问关键字用于访问指定子通道的适配器本地概要指示器字节。
- (d) 设备状态改变指示器关键字 (KC) 726: 该区包括存储器访问关键字, 所述存储器访问关键字用于访问指定子通道的设备状态改变指示器字节。

(e) 中断子类码 (IS) 732: 该区包括一个无符号整数, 其指定了将被用于控制允许适配器中断的中断子类码。

(f) 子系统标识 (SID) 734: 该区包括一个无符号整数, 其指定了一个子系统 ID, 该子系统 ID 标识了为其设置适配器本地概要指示器和设备状态改变指示器地址的子通道。

(g) QDIO 缓冲区信息控制块 (QBICB) 起始地址 736: 该区被用于形成 QBICB 的绝对地址。当该区指定了一个不是 0 的有效地址, 并且 QBICB 没有被指定跨 4K 字节边界时, 指定的 QBICB 被与指定的子通道相关联, 并且该子通道被允许 QDIO 增强缓冲区状态管理解释。当该区包括 0 并且指定的子通道被与 QBICB 关联时, 子通道被从该 QBICB 解除关联, 并且该子通道被禁止 QDIO 增强缓冲区状态管理操作。当该区包含无效地址时, 可以存储一个响应码。当该区指定了一个不是 0 的有效地址, 并且 QBICB 被指定跨 4K 字节边界时, QBICB 说明是无效的, 并且可以存储一个响应码。

当在操作数据区域中提供了一个有效的非 0 QBICB 起始地址时, 设备还被与指定的 QBICB 关联, 并被允许 QDIO 增强缓冲区状态管理解释。当在操作数据区域中提供为 0 的 QBICB 起始地址时, 如果设备与一个 QBICB 相关联, 则它被从该 QBICB 解除关联。

参考图 7c 说明了用于设置通道子系统特性命令的命令响应块 740 的一个例子。该响应块包括, 例如:

(a) 长度区 742: 该区指定了命令响应块的长度。

(b) 响应码 744: 该区包括一个无符号整数, 该整数描述了尝试执行设置通道子系统特性命令的结果。

成功的响应码指出指定的通道子系统特性已经如命令请求块中的操作码和操作数据区域确定的那样被设置。当除了成功之外的响应码被存储时, 该命令被禁止, 并且指定的通道子系统特性没有被设置。

(c) 格式 746: 该区包括一个无符号整数, 该整数的值指定了命令响应块的布局。

返回图 2b, 在步骤 254, 机器处理设置通道子系统特性指令, 并且信息被传递回主机。

此外，在步骤 256，主机获取程序所需的最新的 QDIO 特性。在一个例子中，这是使用此处说明的存储子通道 QDIO 数据命令来完成的。在步骤 258，由机器处理存储子通道 QDIO 数据指令，并且将信息传递回主机。

在步骤 260，主机指出 QEBSM 被允许，并且返回如子通道标记和 SIGA 在程序的 CHSC 响应数据中所要求的 QBICB 的主机绝对地址。在步骤 262 程序接收用于连接的子通道标记和特性。

此后，在步骤 264，程序借助返回的子通道标记使用 SQBS 指令将 QDIO 缓冲区初始化为适当的启动状态。在步骤 266，由机器处理在下面说明的 SQBS 指令，并且将信息传递回程序。

此外，在步骤 268，由程序发出启动子通道命令（带有有效的 QDIO 队列通道命令字），以便激活用于数据传输的队列。在标题为“Self-contained Queues With Associated Control Information For Receipt And Transfer Of Incoming And Outgoing Data Using A Queued Direct Input-output Device”（“用于使用排队直接输入-输出设备来接收和传输进入和外出数据的具有关联控制信息的自足队列”，Baskey 等人，2001 年 12 月 18 日公布，其全部内容在此引入作为参考）的美国专利 No.6,332,171 中说明了与该命令相关联的处理的一个例子。在步骤 270，主机接收此启动子通道命令并发出启动子通道命令（带有有效的 QDIO 队列通道命令字）以通知适配器，以便激活用于数据传输的队列。然后在步骤 272，机器将有效的 QDIO 队列发送给适配器。在步骤 274，当请求被启动时，可以开始 QDIO 数据传输。

作为 QDIO 数据传输的一部分，客户程序管理 QDIO 缓冲区状态。根据本发明的一个方面，为了执行这个任务实现了两个指令。一个指令被称为设置 QDIO 缓冲区状态（SQBS）指令，另一个被称为提取 QDIO 缓冲区状态（EQBS）指令。这些指令允许将 V = V 虚拟机（客户）QDIO 队列与这些队列的主机影子拷贝动态地同步。由机器在客户执行期间在没有主机干预的情况下执行动态队列同步。因此，客户对设置 QDIO 缓冲区状态和提取 QDIO 缓冲区状态指令的使用消除了以前为保持客户 QDIO 队列和主机影子 QDIO 队列同步而产生的环境转换开销。

参考图 8a - 8c 说明了使用设置 QDIO 缓冲区状态指令的一个实施例。

在调用设置 QDIO 缓冲区状态指令之前，将进行某些初始化。例如，在步骤 800（图 8a），主机执行各种清除任务，包括例如删除和处理已处理帧描述符列表上的帧。

已处理帧描述符列表（PFDL）是帧描述符的列表，其描述了在客户解释过程中已经被用于解决主机页无效状态的主机页帧。由 PFDL 描述的主机页帧被分配给为客户提供存储的主机页。由在指定的主机实际地址处的 PFDL 起始地址（PFDLO）来指定 PFDL。PFDLO 的内容以及被附加在右侧的 5 个 0 指定了 PFDL 上的第一个帧描述符的主机本地空间虚拟地址。0 值指出该列表为空。

PFDLO 由主机来初始化，并且可以通过主机或主机页管理辅助功能（此处说明的）来更改。在一个实施例中，借助维护列表完整性的双字并发互锁更新操作来更改 PFDLO。

帧描述符描述主机页帧，并且多个帧描述符可以彼此链接以形成诸如 PFDL 或可用帧描述符列表（AFDL）之类的列表。帧描述符由主机来分配、取消分配和初始化，并且可以由主机页管理辅助功能（此处说明的）来更新。在一个例子中，帧描述符 900（图 9）是例如驻留在主机本地空间虚拟存储器中 32 字节边界上的 32 字节块，并且包括下面的区：

(a) 下一个帧描述符地址 902: 在一个例子中，该区的内容以及被附加在右侧的 5 个 0，指定了该列表上的下一个帧描述符的主机本地空间虚拟地址。为 0 的值指出该帧描述符是列表上的最后一个帧描述符。

该区由主机来初始化，并且可以由主机或由主机页管理辅助功能来更改。

(b) 页帧实际地址或 PTE 拷贝 904: 当帧描述符在可用帧描述符列表（AFDL）内时，该区的内容以及被附加在右侧的 12 个 0，指定了可用于分配以便提供主机存储器的主机页帧的第一个字节（字节 0）的主机实际地址。

当帧描述符在已处理帧描述符列表（PFDL）中时，如在主机页被解决之前所见的，该区包括由页表表项地址区指定的页表表项（PTE）的拷贝。

该区由主机初始化，并且可以由主机或由主机页管理辅助功能更改。

(c) 页表表项地址 906: 当帧描述符在已处理帧描述符列表中时，该区的内容以及被附加在右侧的 3 个 0，指定了用于主机虚拟页的页表表项的主机实际或主机绝对地址。

该区由主机初始化，并且可以由主机或由主机页管理辅助功能更改。

回到图 8a，除了执行清除之外，在步骤 802，主机确保在可用帧描述符列表上有足够的帧以便为接着几个主机存储器请求提供服务。可用帧描述符列表 (AFDL) 是帧描述符的列表，它描述了主机使得可用于进行主机页分配的主机帧。AFDL 由指定主机实际地址处的 AFDL 起始地址 (AFDLO) 来指定。

AFDLO 的内容以及被附加在右侧的 5 个 0，指定了 AFDL 上的第一个帧描述符的主机本地空间虚拟地址。为 0 的值指出该列表为空。

AFDLO 由主机来初始化，并且可以由主机或由主机页管理辅助功能来更改。AFDLO 借助非互锁的更新操作来更改。

帧描述符存在于两种列表中的一种列表中：已处理帧描述符列表 (PFDL) 或可用帧描述符列表 (AFDL)。为每个 CPU 提供单独一对这些列表。借助 CPU 前置区域中的区来指定用于 CPU 的 AFDL 和 PFDL 的起始地址。

此后，在步骤 804，主机可以启动程序的解释执行 (SIE)。响应于启动解释执行，程序开始执行，并且在执行过程中，在步骤 806，选择要传输给适配器的下一个 QDIO 缓冲区。在步骤 808，程序以将在数据传输中使用的存储器块的程序绝对地址来填充 SBAL。进而，在步骤 810，程序通过发出设置 QDIO 缓冲区状态指令将 QDIO 缓冲区的状态从程序拥有改为适配器拥有，来将缓冲区的所有权传输给适配器。参考图 10 说明了设置 QDIO 缓冲区状态指令的格式的一个例子。设置 QDIO 缓冲区状态指令 1000 包括操作码 1002，在这个例子中，它是一个分裂操作码，指定了设置 QDIO 缓冲区状态指令；多个通用寄存器 1004，诸如 R_1 和 R_3 ；以及由 B_2 、 DL_2 和 DH_2 指定的第二运算对象 1006。

使用设置 QDIO 缓冲区状态指令，通用寄存器（由 R_1 和 R_3 区指定）所指定的一个或多个 QDIO 输入队列或输出队列缓冲区的状态被设置为由

第二运算对象指定的状态。操作继续进行直到指定的缓冲区状态被设置，或是直到由 CPU 所确定的数目的缓冲区状态已经被设置为止（无论哪个首先发生）。结果在状态码中指出。

通用寄存器 1 包括子通道标记，该标记指定允许 QDIO 增强缓冲区状态管理（QEBSM）并且将对其一个或多个缓冲区状态进行设置的 QDIO 子通道。（在 QDIO 子通道被允许 QEBSM 之后，通过 CHSC 存储子通道 QDIO 数据命令来获得子通道标记。）

由 R_1 区指定的通用寄存器包括对 QDIO 队列和第一个缓冲区（将其状态进行设置）的说明。当该指令被发出时，由 R_1 区指定的通用寄存器包括下面的区：

(a) 队列索引（QIX）：该区包括一个无符号整数索引值，该值指定了包含将对其状态进行设置的缓冲区的队列。如果该队列是输入队列，则队列索引等于队列的输入队列号。如果队列是输出队列，则队列索引等于队列的输出队列号加上已建立的输入队列的总数的和。该区的值将指定现有的队列（即，队列索引值不超过已建立的输入队列的总数加上已建立的输出队列的总数的和减 1）；否则，在一个例子中，状态码和状态码限定符被设置。

(b) 缓冲区号（BNUM）：该区包括一个无符号整数，该整数指定了将对其状态进行设置的第一个缓冲区的号码。

由 R_3 区指定的通用寄存器包括将对其状态进行设置的缓冲区的数目的计数。 R_3 寄存器还包括一个状态码限定符区，当指令完成时，该区与所述状态码的设置一起被设置。当指令被发出时，由 R_3 区指定的通用寄存器包括下面的区：

(a) 状态码限定符（CCQ）：当指令被发出时，该区的内容没有意义。当指令完成时，该区包括一个当状态码被设置时对其进行设置的状态码限定符（CCQ）码。下面的列表包括可能的 CCQ。

- 所有指定的缓冲区状态被设置。
- 当前指定的缓冲区处于适配器拥有状态。该缓冲区的状态没有改变。
- 自从状态最后从适配器拥有改变为程序拥有状态以来，当前指定

的缓冲区状态没有被提取 QDIO 缓冲区状态指令提取。该缓冲区的状态没有改变。

- 用于当前指定的队列的存储器列表 (SL) 的地址是无效的。
- 用于当前指定的缓冲区的存储器缓冲区地址列表 (SBAL) 的地址是无效的。
- 用于当前指定的缓冲区的存储器块 (SB) 的地址是无效的。
- 在指定的缓冲区中存在 SBAL 序列错误。
- 指定的子通道不可运行。
- 指定的子通道标记没有指定一个允许 QEBSM 的有效子通道。
- 没有为指定的子通道建立指定的队列。
- 指定的缓冲区号无效。
- 指定的缓冲区计数无效。
- 指定的缓冲区状态既不是适配器拥有也不是程序拥有状态，或者指定的缓冲区状态既是适配器拥有又是程序拥有状态。
- CPU 确定的数目的缓冲区状态已经被设置。

(b) 计数: 该区包括一个无符号整数, 该整数指定了将对其状态进行设置的缓冲区的计数。在一个例子中, 该区的值要大于 0 并且小于或等于 128; 否则, 将设置状态码和状态码限定符。

作为一个例子, 缓冲区号空间在 0 到 127 的范围内。计数值结合 R_1 BNUM 值, 可以指定缓冲区号空间中的绕回 (wrap around)。

第二运算对象地址没有被用于地址数据。而是, 第二运算对象地址的最右侧 8 个位以存储器列表状态块 (SLSB) 缓冲区状态指示器的形式指定了将被设置的 QDIO 缓冲区状态。如果第二运算对象地址指定了一个既不是适配器拥有也不是程序拥有状态的状态, 或者第二运算对象指定了一个既是适配器拥有也是程序拥有状态的状态 (即, 如果所指定的状态的位 0 和位 1 相等), 则设置状态码, 并且以预先指定的值来设置 R_3 状态码限定符。

用于设置 QDIO 缓冲区状态的操作单元包括设置一个缓冲区状态、将 R_3 寄存器中的计数区递减 1 以及将 R_1 寄存器中的缓冲区号区递增 1。如果递增缓冲区号区使得缓冲区号超过最大缓冲区号 127, 则在缓冲区号空间

中发生绕回，并且缓冲区号被重置为 0。

当下面任一情况先发生时，指令操作完成：

- 当 R_3 计数区中指定的缓冲区状态的计数已经被设置，并且没有遇到错误或异常状态时。在这种情况下，状态码和 R_3 状态码限定符区被设置。
- 当遇到异常状态并且没有、某些或全部指定的缓冲区状态已经被设置时。在这种情况下，状态码被设置并且使用一个值来设置 R_3 状态码限定符区以便指出异常。以已经被设置的缓冲区状态的数目来递减 R_3 寄存器中的计数区，并且以相同的数目来递增 R_1 寄存器中的缓冲区号 (BNUM) 区，且绕回通过 0 (如果需要)。因此，如果 R_3 寄存器中的计数区是 0，则指定的缓冲区状态已经被设置。如果该计数区不是 0，则 R_1 队列索引 (QIX) 和存储器号 (BNUM) 区标识出当遇到异常状态时正被处理的缓冲区。
- 当遇到错误状态，并且没有指定的缓冲区状态已经被设置时。在这种情况下，设置状态码，并且使用一个值来设置 R_3 状态码限定符区以便指出错误。
- 当 CPU 确定的数目的指定缓冲区状态已经被设置，而没有设置缓冲区状态的总的指定计数时。在这种情况下，状态码和 R_3 状态码限定符区被设置。以已经被设置的缓冲区状态的数目来递减 R_3 寄存器中的计数区，并且以同样的数目来递增 R_1 寄存器中的缓冲区号 (BNUM) 区，且绕回通过 0 (如果需要)。

在指定的缓冲区状态被设置之前和之后，执行串行化和检查点同步功能。

设置 QDIO 缓冲区状态指令的执行不改变包含在子通道中的任何信息，也不使 I/O 子系统向适配器发送信号，并且也不改变程序存储器中的任何缓冲区状态信息。

参考图 8a - 8c 说明了与执行 SQBS 指令相关联的处理。除非发生异常状态，有利地，由机器执行指令。如果发生了这种异常状态，则由主机完成所述处理，其中主机执行被列于机器之下的剩余的步骤。

参考图 8a，响应于发出 SQBS 指令，在步骤 812，机器执行主机允许

检查以便判定主机是否允许客户使用 QEBSM 工具。在步骤 814, 如果 QEBSMI 允许控制 (例如, 在客户状态中) 被设置为 0, 指出不允许, 则在步骤 816, 将由主机而不是机器来完成所述处理。当在设置 QDIO 缓冲区状态解释的过程中遇到一个状态, 需要来自主机的某些活动或服务, 或是指令阻断或是部分执行阻断 (PEI) 被识别, 使用设置 QDIO 缓冲区状态指令文本来设置阻断参数, 并且对于 PEI, 在客户状态中设置 QEBSM - PEI 参数。

接着, 在步骤 818, 机器执行子通道判定。在一个例子中, SQBS 子通道标记被用于定位包含有标识所述子通道的 SID 的 QBICB。作出关于子通道标记是否与存储在由 QBICB 指定的子通道中的标记相匹配的判定。如果不匹配, 则在步骤 820 由主机完成处理。否则由机器继续处理。

更详细地, 当设置 QDIO 缓冲区状态指令 (以及向适配器发送信号指令或提取 QDIO 缓冲区状态指令) 被执行时, 在通用寄存器 1 中的子通道标记被用于指定子通道。该标记被用于判定指定的子通道是否允许 QDIO 增强缓冲区状态管理解释, 并且被用于确定所述子通道的子系统 id。在一个例子中, 用于作出这些判定的步骤包括:

- (1) 当子通道被允许 QDIO 增强缓冲区状态管理解释时, 寄存器 1 中的子通道标记被用作与该子通道相关联的 QDIO 缓冲区信息控制块 (QBICB) 的主机实际或绝对地址。如果当引用该 QBICB 时存在访问异常状态, 则客户指令被禁止并且指令阻断被识别; 否则处理继续。
- (2) QBICB 的子系统 id 区指定了要使用的子系统。如果 QBICB 的子系统 id 区不包含有效的子系统 id, 则客户指令被禁止并且指令阻断被识别; 否则处理继续。这允许例如主机处理模拟设备的操作或处理追踪能力。
- (3) 如果由 QBICB 子系统 id 区指定的子通道没有被允许 QEBSMI 或是没有与由所述子通道标记指定的 QBICB 相关联, 则客户指令被禁止并且指令阻断被识别; 否则处理继续。

接着, 在步骤 821, 由机器作出关于 SQBS 运算对象是否有效的判定。作为例子, 缓冲区号被验证。在一个例子中, 如果 R_1 缓冲区号区指定了大

于 127 的值，则指令完成，并且状态码和状态码限定符被设置。此外，计数值被验证。如果 R_3 计数区指定了 0 值或大于 128 的值，则指令被再次完成，并且状态码和状态码限定符被设置。此外，对指定的缓冲区状态的有效性进行检查。如果由第二运算对象地址指定的缓冲区状态指定了一个同时为适配器拥有和程序拥有的状态，或指定了一个既不是适配器拥有也不是程序拥有的状态，则该指令完成，并且状态码和状态码限定符被设置。进而，用于指定队列的队列存储器描述符被定位。如果 R_1 队列索引区大于或等于 QBICB 队列计数区中的值，则没有为所述子通道建立指定的队列并且指令完成，并且状态码和状态码限定符被设置。否则， R_1 队列索引区的值被用作 QBICB 队列描述符数组中的索引以便确定队列描述符地址。

假如上面的任何一个测试失败，则处理完成，并且在步骤 822 状态码和状态码限定符被设置。否则处理继续。

在客户状态中的间隔完成指示器被设置为 1，并且执行串行化和检查点同步功能。在机器故障的情况下，该指示器通知主机主机控制结构可能处于有问题的状态。

接着，在步骤 824 为指定的队列进行获得队列存储器描述符锁的尝试。为指定的队列从 QBICB 队列描述符中的队列存储器地址区中获取队列存储器描述符地址。使用互锁更新操作进行尝试以便将 QSD 锁定字值从可用改变为被机器锁定。如果 QSD 地址为 0 或当尝试设置该锁定字时存在访问异常状态，则一个有效性阻断被识别并且设置有效性阻断原因码。

如果互锁更新操作失败，可以对这个操作进行取决于模型的次数的重试。如果 QSD 锁定字不能被设置为由机器值锁定，则在步骤 826 部分执行阻断被识别，返回给主机程序的 QEBSM - PEI 参数项中的部分执行原因码被设置为（例如）1，并且由主机完成处理。如果互锁更新操作成功，则处理继续。

下面的逻辑被重复由 SQBS 指令上的计数区指定的 QDIO 缓冲区的数目。初始地，在步骤 828，用于要处理的 QDIO 缓冲区的 SQBN 和 SBAL 被定位。在一个例子中，用于指定队列的主机存储器列表被定位。即，从队列存储器描述符中的主机存储器列表地址区取回主机存储器列表地址。然后，定位指定的队列存储器列表状态块。从队列存储器描述符中的 SLSB

地址区取回该地址。要设置的第一个 SLSB 表项 (SQBN) 被定位。R₁ 缓冲区号 (BNUM) 区的值被用作 SLSB 中的索引以便确定要设置的第一个 SQBN 的地址。如果在访问 SQBN 时存在访问异常状态, 则有效性阻断被识别并且有效性阻断原因码被设置。

此后, 在步骤 830, 作出关于指定的缓冲区是否当前被程序拥有的判定。在一个例子中, 通过检查 SQBN 作出该判定。如果不是, 则在步骤 832, 处理完成并且在设置状态码和状态码限定符。否则, 在步骤 834, 处理继续, 机器作出关于指定的缓冲区的状态是否需要被提取的判定 (图 8b)。如果状态需要被通过例如 EQBS 指令提取, 则在步骤 836, 处理完成并且设置状态码和状态码限定符。如果指定缓冲区的状态不需要被提取, 则在步骤 838, 程序 SBAL 的内容被复制到主机的 SBAL。此外, 在步骤 840, 在主机的 SBAL 中设置缓冲区同步和 EQBS 操作所需的位。

进而, 在步骤 842, 对于具有有效存储器块的每个有意义的主机 SBALE, 执行下面的功能: (1) 在步骤 844, 从主机 SBALE 取回程序绝对 (主机虚拟) 存储器块地址; (2) 在步骤 846, 调用下面说明的 HPMa 钉住功能以便在存储器中转换并钉住页。在步骤 848, 如果 HPMa 钉住功能没有成功, 则在步骤 850, 由主机完成处理。否则由机器继续处理; (4) 在步骤 852, 由 HPMa 钉住返回的存储器块的主机绝对地址被移动到主机的 SBALE 中; 以及 (5) 在步骤 854, 主机的 SBAL 中的已钉住页指示器被设置 (图 8c)。

此后, 在步骤 856, 缓冲区的 SQBN 被设置为由 SQBS 指令指定的适配器拥有状态。在一个例子中, 该状态由第二运算对象指定。另外, 在步骤 858, SQBS 缓冲区号 (BNUM) 递增 1, 并且在步骤 860, SQBS 计数区递减 1。

在步骤 862, 作出关于 SQBS 计数是否等于 0 的判定。假如 SQBS 计数不等于 0, 则在步骤 828 处理继续, 定位要被处理的 QDIO 缓冲区的 SQBN 和 SBAL (图 8a)。然而, 如果 SQBS 计数等于 0, 则在步骤 864, 处理继续, 解锁 QSD 锁 (图 8c)。如果成功, 在步骤 866, 成功指示被转发给程序。

如果这是在此刻要传输的最后的 QDIO 缓冲区, 则在步骤 868, 发出

适当的 SIGA 指令功能，以便向适配器发信号：I/O 请求未决，例如，对于输出操作为 SIGA-wt，或者当输入可供适配器后续使用时为 SIGA -rt。

注意，如果 SQBS 处理必须由主机完成，则主机（而不是机器）向程序给出适当的结束操作。

在对 SQBS 指令的说明中，引用了主机页管理辅助（HPMA）。根据本发明的一个方面，主机页管理辅助包括可以由其它 CPU 工具在可分页存储客户的解释过程中调用的 4 个辅助功能。这些功能包括钉住功能、解决主机页功能、由虚拟地址解除主机页钉住功能以及由 PTE 地址解除主机页钉住功能，下面将说明每一个功能。

从 SQBS 指令调用的 HPMA 钉住功能为 CPU 提供了这样的能力：指出主机页的内容被钉在主机页帧内。当页需要驻留在单个指令的执行之外（例如，在 I/O 操作期间）时，页可以被钉住。当页处于被钉住状态时，这是给主机的指示，指示该页无资格进行页失效。当页被成功地由钉住主机页功能钉住时，与该页关联的已钉住页计数器可以被递增。

在一个例子中，增量（delta）已钉住页计数器数组（DPPCA）是一个例如 31 位有符号整数已钉住页计数器的数组，当页被成功地钉住时，所述计数器之一可以通过钉住主机页功能来递增。进而，当页被成功地解除钉住时，DPPCA 计数器之一可以通过由虚拟地址解除主机页钉住功能或者由 PTE 地址解除主机页钉住功能（在下面说明）来递减。在一种实现中，DPPCA 中的一个计数器被与驻留在给定大小（例如 2GB）之下的主机实际存储器中的页相关联，并且其它计数器被与驻留在该大小或高于该大小的主机实际存储器中的页相关联。

主机将每个页分类为类 0 页或类 1 页，并且相应地为该页在页状态表表项(PGSTE)中设置页分类指示器。参考图 11a 说明了页状态表表项 1100 的一个例子，页状态表表项 1100 包括例如下列内容：

- (a) Acc 1102: 访问控制关键字；
- (b) FP 1104: 取回保护指示器；
- (c) 页控制互锁(PCL)1106: 这是用于串行化对页表表项(PTE)和相应 PGSTE(除被标记为保留的 PGSTE 状态区域和 PGSTE 位之外)的更新的互锁控制。

- (d) HR 1108: 主机引用备份指示器;
- (e) HC 1110: 主机改变备份指示器;
- (f) GR 1112: 客户引用备份指示器;
- (g) GC 1114: 客户改变备份指示器;
- (h) 页内容逻辑 0 指示器 (Z) 1116: 当为 1 时, 由该 PGSTE 和相应的 PTE 描述的页的内容被看作 0. 该页的任何先前内容不必被保留, 并且可以被是 0 的页替代。

当为 0 时, 由该 PGSTE 和相应的 PTE 描述的页的内容不被看作 0. 该页的内容将被保留并且不应被替换。

当相应的 PTE 页无效 (PTE.I) 位为 1 时, 这个位有意义。

1 当 Z 位是 1 并且相应的 PTE.I 位 (在下面说明) 是 1 时, 页内容可以由主机或主机页管理辅助功能来替换。

2 当 Z 位是 1, 相应的 PTE.I 位是 1, 并且页内容被替换时, 应通过将页与被设置为 0 的帧关联来替换页。

(i) 页分类 (PC) 1118: 当为 0 时, 由 PGSTE 和相应的 PTE 描述的页为类 0 页, 并且用于类 0 页的增量已钉住页计数数组 (DPPCA) 被用于计数该页的钉住和解除钉住操作。当是 1 时, 由 PGSTE 和相应的 PTE 描述的页为类 1 页, 并且类 1 页的 DPPCA 被用于计数该页的钉住和解除钉住操作。

(j) 钉住计数溢出 (PCO) 1120: 当为 1 时, 钉住计数区处于溢出状态。当为 0 时, 钉住计数区没有处于溢出状态。在这种情况下, 总的钉住计数由主机保存在不被机器访问的另一个数据结构中。

(k) 已处理帧描述符列表 (FPL) 上的帧描述符 1122: 当为 1 时, 由 PGSTE 和相应的 PTE 描述的页的帧描述符在已处理帧描述符列表中。帧描述符标识了由用于页的 HPMA 解决主机页功能所使用的主机页帧。

(l) 已请求页内容替换 (PCR) 1124: 当是 1 时, 当为由 PGSTE 和相应的 PTE 表示的页调用 HPMA 解决主机页功能时, 页内容替换被请求。

(m) 钉住计数 1126: 为无符号二进制整数, 用于指出由 PGSTE 和相应的 PTE 表示的主机虚拟页的内容是否被钉在由 PTE 的页帧实际地

址区所指定的实际主机页帧中。当该区的值大于 0 或页计数溢出 (PCO) 位是 1 时, 相应的页被认为是被钉住的。当该区的值是 0 并且 PCO 位为 0 时, 相应的页被认为没有被钉住。

在页被主机或 CPU 钉住时, 该区应递增 1。在页被主机或 CPU 解除钉住时, 该区应递减 1。

当钉住计数区的值大于 0 或 PCO 位为 1 时, 相应的 PTE.I (页无效) 位要为 0。否则, 可能发生不可预知的结果。

当页被钉住时, 主机程序不应更改 PTE 页帧实际地址 (PRFA) 区的内容、PTE 页无效 (I) 位的设置或者 PTE 或段表表项 (STE) 中的页保护 (P) 位的设置。否则, 可能发生不可预知的结果。

PGSTE 对应于一个页表表项 (PTE), 参考图 11b 说明了它的一个例子。页表表项 1150 包括, 例如:

(a) 页帧实际地址 1152: 该区提供了实际存储器地址的最左边的位。当这些位被与右侧的虚拟地址的字节索引区连接成串时, 可以获得实际地址。

(b) 页无效指示器 1154: 该区控制与页表表项相关联的页是否可用。当该指示器为 0 时, 通过使用页表表项进行地址转换。当该指示器为 1 时, 页表表项不能被用于转换。

(c) 页保护指示器 1156: 该区控制是否允许对页进行存储访问。

在标题为“z/Architecture Principles of Operation” (“z/Architecture 运行原理”, IBM®出版物 No. SA22 - 7832 - 02, 2003 年 6 月, 其全部内容在此引入作为参考) 的 IBM®出版物中提供了关于页表表项和页表以及此处提到的段表表项的进一步的详细信息。

继续对 DPPCA 的讨论, 单独的 DPPCA 被与每个用于类 0 钉住页计数的虚拟配置相关联, 并且由所述虚拟配置的每个描述中的类 0 DPPCA 起始地址区来指定。存在用于类 1 钉住页计数的单独的 DPPCA, 并且由前置区域中的类 1 DPPCA 起始地址区来指定。

DPPCA 不受关键字控制的保护或低地址保护。在一个例子中, 它是一个在双字边界上被分配的 8 字节控制块, 并且包括, 例如, 用于小于特定大小 (例如 2GB) 的页帧实际地址的增量已钉住页计数, 和用于大于或

等于该大小的页帧实际地址的增量已钉住页计数。互锁更新操作被用于改变这个区，以便维护其内容的准确性。当页成功地由钉住主机页功能钉住时，增量已钉住页计数器可以被递增。

当钉住主机页功能被调用时，下面的数据被输入该功能：将被钉住的主机虚拟地址（页）；标识分配给客户的地址空间的有效主机地址空间控制元素（ASCE）和要使用的转换表；对该页是否被为取回或存储操作钉住的指示；以及页内容替换指示器。

参考图 12 说明了与钉住主机页功能关联的逻辑的一个实施例。初始地，在步骤 1200，定位用于页的 PTE/PGSTE 对。例如，使用 ASCE 和将被钉住的虚拟地址（页），主机存储器中的用于将被钉住的该虚拟地址的页表表项和页状态表表项的位置被确定。

在步骤 1202，被定位的 PTE/PGSTE 对被使用例如互锁更新操作来锁定，以便将 PGSTE 中的页控制锁指示器从 0 更改为 1。在询问 1204，如果所述对不能被锁定，或者如果页控制互锁（PCL）已经为 1，则在步骤 1206 指出失败。否则处理继续，在询问 1208，判定用于页的钉住计数在下一个钉住时是否将溢出（即，如果 PGSTE 中的钉住计数区的值为一个预定的值（例如，255））。如果是的，则在步骤 1206，处理以失败结束。否则在步骤 1210，用于将被钉住的虚拟页的 PTE 被取回。

接着，在询问 1212，作出关于该页是否被保护并且正在被为存储操作钉住的判定。即，在步骤 1206，如果输出指示用于存储操作，并且用于该页的段表表项或者页表表项中的页保护位是 1，则指出失败。否则，处理继续，在询问 1214 判定该页是否无效。如果该页无效，则在步骤 1216 调用解决主机页功能（如下所述）。如果在询问 1217，解决主机页功能失败，则在步骤 1206 钉住功能失败。然而，如果解决成功，或如果该页不是无效的，则在步骤 1218，钉住计数被递增。

随后，在询问 1220，作出关于钉住状态是否正在被从解除钉住更改为钉住状态的判定。如果是的，在步骤 1222，关联的增量已钉住页计数（即，基于页分类和 PFRA 所选择的计数）递增 1。即，如果钉住计数现在是 1 并且钉住计数不处于溢出状态（PGSTE.PCL = 0），则页状态已经从解除钉住状态转换为钉住状态。在这种情况下，增量已钉住页计数被递增 1。

此后，或者如果页状态没有正在被从解除钉住状态更改为钉住状态，则在步骤 1224，PTE/PGSTE 对被解锁。例如，PGSTE 中的页控制互锁位被设置为 0，并且处理成功完成。

当钉住主机页功能退出且指示成功时，指定的页被钉住，并且还可能会已经通过调用解决主机页功能而被解决。在这种情况下，用于该页的页计数已经在状态表中由机器递增，并且主机转换表可能也已经由机器更新，先前的对传递控制给主机以便进行页钉住和解决的要求被减轻了。

当页主机钉住功能退出且指示失败时，对转换和状态表没有作出永久的更改，并且当识别到失败时，由钉住功能设置的任何页控制的互锁位被重置。在这种情况下，机器就像没有调用该功能一样，并且主机将被给予控制以便处理所述失败。

如上所述，如果将被钉住的页无效，则调用解决主机页功能。当条件允许时，解决主机页功能利用主机提供的可由 CPU 使用的可用主机页帧的列表，以便动态地解决主机页无效状态。该功能还提供了一种手段，借助该手段，通过被指配给特定的主机页，已经被用于解决主机页无效状态的任何主机页帧被通知给主机。

如果主机页管理辅助被允许，则为主机页调用解决主机页功能，并且描述该页的页表表项中的页无效位 (PTE.I) 为 1，然后主机已使其成为可用的主机页帧中的一个主机页帧被用于动态地解决主机页无效状态。

解决主机页功能的输入包括可用帧描述符列表、将被解决的虚拟地址 (页)、标识所述地址空间的有效地址空间控制元素 (ASCE) 和将被使用的转换表、页内容替换指示器、对间隔完成指示器是否已经由调用者设置的指示、以及对 PTE/PTSTE 对的访问是否已经被调用者锁定的指示。

参考图 13a - 13b 说明了与解决功能相关联的逻辑的一个实施例。初始地，在询问 1300，作出关于是否存在任何可用帧的判定 (图 13a)。通过例如检查可用帧描述符列表是否为空来作出该判定。如果 AFDLO 指定了为 0 的地址 (即，对于本地处理器没有可用的主机页帧)，则没有可用的帧，并且在步骤 1302，处理退出并指出失败。

然而，如果存在可用的帧，则在步骤 1304，处理器继续，为所述页定位 PTE/PGSTE 对。在一个例子中，ASCE 和将被解决的虚拟地址 (页)

被用于为将被解决的虚拟地址确定主机存储器中的页表表项和页状态表表项的位置。

此后，在步骤 1306，从可用帧描述符列表中的第一个帧描述符取回页帧实际地址。同时，如果尚未被指示为已经由调用者建立，则间隔完成指示器被标记，以便指出解决主机页功能已经开始，并且在步骤 1308，串行化和检查点同步功能也被执行。

随后，在步骤 1310，PTE/PGSTE 对被有条件地锁定。如果解决主机页功能表项状态指出对相应于将被解决的主机虚拟地址的 PTE 和 PGSTE 的访问尚未被串行化，则尝试使用互锁更新操作来将 PGSTE 中的页控制互锁指示器 (PGSTE.PCL) 从 0 更改为 1。如果在询问 1312，PTE/PGSTE 对不能被锁定或者如果 PGSTE.PCL 被发现已经为 1，则在步骤 1302，处理以失败退出。否则，在步骤 1314，处理继续，为将被解决的虚拟页取回 PTE。在一个例子中，页表表项被从主机存储器同步地以块取回。

此外，在询问 1316，作出关于页是否需要被解决的判定。在一个例子中，通过检查 PTE 中的有效指示器来作出此判定。如果 PTE 指出页是有效的 (PTE.I=0)，则该页被解决并且处理完成。然而，如果页将被解决，则在询问 1318，作出关于该页能否被解决的进一步的判定。例如，如果用于将被解决的虚拟页的 PGSTE 指出页内容逻辑上不是 0 (PGSTE.Z=0) 并且输入页内容替换指示器指出该页不能被替换，则在步骤 1302，处理以失败退出。然而，如果 PGSTE 中的页内容逻辑 0 指示器为 1 或解决主机页功能表项状态指出将被解决的主机虚拟页的内容可以被替换，则在步骤 1320，可以作出解决该页的尝试并且处理继续。

为了解决该页，如参考图 13b 所说明的执行各个步骤。一个步骤 (步骤 1330) 包括为页设置存储器关键字。例如，用于由帧描述符的页帧实际地址区所标识的主机帧的存储器关键字被通过下面的步骤设置：设置来自 PGSTE.ACC 的访问控制位；设置来自 PGSTE.FP 的取回保护位；将引用指示器设置为 0；以及将更改指示器设置为 0。

此外，在步骤 1332，帧描述符被从可用帧描述符列表中删除。作为一个例子，使用非互锁更新操作来从可用帧描述符列表中删除所述帧描述符。该帧描述符的下一个帧描述符地址区的内容替换 AFDLO 中的旧值。

接着，在步骤 1334，帧分配被记录。通过将 PTE 复制到帧描述符的 PTE 拷贝区中，主机虚拟页的 PTE 被保留。此外，用于主机虚拟页的 PTE 的地址被存储器在帧描述符的页表表项地址区中。

接着，在询问 1338，作出关于页的内容是否能够被替换的判定。作为一个例子，通过检查输入页内容替换指示器作出此判定。如果输入页内容替换指示器指出页内容可以被替换，则在步骤 1340，PGSTE 中的页内容替换指示器 (PGSTE.PCR) 被设置为 1。这允许主机持续统计允许页内容替换的操作的数目。

此后，或者如果内容没有被替换，用于主机虚拟页的 PTE 被更新以便删除页无效状态。例如，并发存储器操作使用所述帧描述符的页帧实际地址区的内容来替换页表表项的页帧实际地址区的内容，并且在步骤 1342，将该页表表项中的页无效位设置为 0。PTE 中的其它位保持不变。此外，在步骤 1344，该帧描述符被添加到已处理列表中。这是通过互锁更新操作执行的。这完成了解决处理。

回到图 13a，在解决该页之后，在步骤 1322，PTE/PGSTE 对被有条件地解锁。例如，如果页控制互锁指示器被由该功能设置为 1，则指示器被设置为 0。

此外，如果由此功能设置了间隔完成指示器，则在步骤 1324 执行串行化和检查点同步功能，并且间隔完成指示器被重置。这完成了所述解决功能的处理。

当解决功能退出且指示成功时，所指定的页已经被解决。在这种情况下，主机转换表已经由机器更新，并且先前的对传递控制给主机以便进行页解决的要求被减轻了。

当解决功能退出且指示失败时，没有对转换表作出更改。在这种情况下，机器就像没有调用该功能那样，并且主机被给予控制以便处理所述失败。

根据本发明的一个方面，除了 SQBS 指令之外，还实现了 EQBS 指令。作为例子，当客户从适配器接收到通知时，客户发出提取 QDIO 缓冲区状态指令，或是当缓冲区状态将被提取时，客户主动发出该指令。参考图 14a - 14c 说明了与提取 QDIO 缓冲区状态指令相关联的逻辑的一个实施例。

在运行该程序之前，进行某些初始化。例如，在步骤 1400，主机执行各种清除任务（图 14a），包括例如删除和处理已处理帧描述符列表上的帧。进而，在步骤 1402，主机确保在可用帧描述符列表上存在足够的帧。此后，在步骤 1404，主机开始程序的解释执行。

在执行过程中，在步骤 1406，程序选择它正等待适配器结束处理的下一个输出 QDIO 缓冲区。然后在步骤 1408，程序发出 EQBS 指令以便取得所选择的缓冲区的当前状态。

参考图 15 说明了 EQBS 指令的格式的一个例子。提取 QDIO 缓冲区状态指令 1500 包括指定 EQBS 指令的操作码 1502；通用寄存器 R_1 区 1504；通用寄存器 R_3 区 1506；以及通用寄存器 R_2 区 1508；下面说明它们中的每一个。

借助提取 QDIO 缓冲区状态指令，由 R_1 和 R_3 区指定的通用寄存器所指定的一个或多个相同状态的 QDIO 输入队列或输出队列缓冲区的状态被提取，并且以 QDIO 缓冲区状态指示器的形式被装入由 R_2 区指定的通用寄存器。可选地，处于输入缓冲区准备好状态的指定的缓冲区可以被更改为输入缓冲区已确认状态。操作继续，直到指定的缓冲区使其公共状态被提取为止，或是直到至少某些指定的缓冲区已使其公共状态被提取并且下一个将被提取的缓冲区状态不同于先前提取的那些状态时为止，或是直到 CPU 确定的数目的缓冲区状态已使其公共的状态被提取为止，以先发生的为准。结果在状态码中指出。

通用寄存器 1 包括子通道标记，该标记指定了被允许 QDIO 增强缓冲区状态管理 (QEBSM) 并且将使其一个或多个缓冲区状态被提取的 QDIO 子通道。

由 R_1 区指定的通用寄存器包括对 QDIO 队列和第一个缓冲区（其状态将被提取）的说明。当发出该指令时，由 R_1 区指定的通用寄存器包括如下的区：

- (a) 队列索引 (QIX)：该区包括一个无符号整数索引值，该值指定了包含其状态将被提取的缓冲区的队列。如果队列是输入队列，则队列索引等于该队列的输入队列号。如果队列是输出队列，则队列索引等于该队列的输出队列号加已建立的输入

队列的总数的和。该区将指定存在的队列（即，队列索引值不超过已建立的输入队列的总数加已建立的输出队列的总数的和减1）。

(b) 缓冲区号 (BNUM)：该区包括一个无符号整数，该整数指定了其状态将被提取的第一个缓冲区的号码。在一个例子中，该区的值小于 128。

由 R_3 区指定的通用寄存器包括对其状态将被提取的缓冲区的数目的计数。 R_3 寄存器还包括当该指令完成时随着状态码的设置而被一起设置的状态码限定符。在一个例子中，当该指令被发出时，由 R_3 区指定的通用寄存器包括下面的区：

(a) 状态码限定符 (CCQ)：当所述指令被发出时，该区没有意义。当所述指令完成时，该区包括一个当状态码被设置时被设置的无符号整数状态码限定符 (CCQ) 码。状态码限定符的例子包括：

- 所有指定的缓冲区状态被提取。
- 当前指定的缓冲区的状态与在该指令的执行期间先前提取的那些状态不同。该缓冲区的状态不被提取。
- 用于当前指定的队列的存储器列表 (SL) 的地址无效。
- 用于当前指定的缓冲区的存储器缓冲区地址列表 (SBAL) 的地址无效。
- 用于当前指定的缓冲区的存储器块 (SB) 的地址无效。
- 指定的子通道不可操作。
- 指定的子通道标记没有指定有效的允许 QEBSM 的子通道。
- 没有为指定的子通道建立指定的队列。
- 指定的缓冲区号无效。
- 指定的缓冲区计数无效。
- CPU 确定的数目的缓冲区状态已经被提取。
- CPU 确定的数目的缓冲区存储器块已经被处理，但是没有完整的缓冲区已经被处理并且没有缓冲区状态已经被提取。

(b) 计数：该区包括一个无符号整数，该整数指定了其状态将被提取的缓冲区的计数。该区的值应大于 0 并且小于或等于 128；否则，

状态码和 R_3 状态码限定符区被设置。缓冲区号空间在 0 到 127 的范围内。所述计数值结合 R_1 BNUM 值，可以指定缓冲区号空间中的绕回。

在发出该指令之前，由 R_2 区指定的通用寄存器包括这样的指示：在缓冲区状态被提取后，是否应将要提取的和被发现处于输入缓冲区准备好状态的指定缓冲区状态更改为输入缓冲区已确认状态。当该指令完成时，结果状态码和状态码限定符指出，例如，由 R_2 区指定的寄存器是否包括一个或多个缓冲区的状态。当该指令被发出时，由 R_1 QIX 和 BNUM 区的值标识出所述状态将应用到其的第一个缓冲区。所述状态将应用到其的被顺序编号的缓冲区的计数是当指令发出时 R_3 计数区与指令完成时递减的 R_3 计数区之间的差。

由 R_2 区指定的通用寄存器包括例如下面的区：

- (a) 确认输入缓冲区控制 (I)：当该指令被发出且该区为 1 时，在缓冲区状态被提取之后，被发现处于输入缓冲区准备好状态的指定缓冲区将被更改为输入缓冲区已确认状态。当该指令被发出并且该区为 0 时，没有输入缓冲区状态被改变。
- (b) 缓冲区状态 (BSTATE)：当该指令被发出时，该区没有意义。当该指令完成并且 R_3 计数区已经被递减时，该区包括从存储器列表状态块 (SLSB) 缓冲区状态指示器取回的值。当该指令完成并且 R_3 计数区没有被递减时，该区的内容保持不变并且没有意义。

用于提取 QDIO 缓冲区状态的操作单元包括提取一个缓冲区状态、将该状态装入 R_2 BSTATE 区、可选地更改该缓冲区状态、将 R_3 寄存器中的计数区递减 1 以及将 R_1 寄存器中的缓冲区号区递增 1。如果递增缓冲区号区使缓冲区号超过了最大例如 127 的缓冲区号，则在缓冲区号空间内发生绕回，并且缓冲区号被重置为例如 0。

当下面之一首先发生时指令操作完成：

- 当 R_3 计数区中指定的缓冲区的计数已经使其公共状态被提取到 R_2 缓冲区状态 (BSTATE) 区中，并且没有遇到错误或异常情况时。在这种情况下，设置状态码和 R_3 状态码限定符。
- 当遇到异常状态并且没有或某些指定缓冲区使其公共状态被提取

时。在这种情况下，使用一个值来设置状态码和 R_3 状态码限定符以便指示该异常。 R_3 寄存器中的计数区被递减已经被提取的缓冲区状态的数目，并且 R_1 寄存器的缓冲区号 (BNUM) 区被递增相同的数目，且绕回通过 0 (如果需要)。指定的缓冲区状态已经被提取。如果所述计数区不是 0，则 R_1 队列索引 (QIX) 和缓冲区号 (BNUM) 区标识出当遇到异常情况时正在被处理的缓冲区。

- 当遇到错误状态并且没有指定的缓冲区使其状态被提取时。在这种情况下，使用一个值来设置状态码和 R_3 状态码限定符区以便指示该错误。
- 当 CPU 确定的数目的缓冲区已使其公共状态被提取而没有提取缓冲区状态的全部指定计数时。在这种情况下，设置状态码和 R_3 状态码限定符区。 R_3 寄存器中的计数区被递减已经被提取的缓冲区状态的数目，并且 R_1 寄存器中的缓冲区号 (BNUM) 区被递增相同的数目，且绕回通过 0 (如果需要)。
- 当 CPU 在完成单个操作单元之前决定结束该指令时。在这种情况下，设置状态码和状态码限定符。 R_3 寄存器中的计数区和 R_1 寄存器中的缓冲区号 (BNUM) 区保持不变，并且没有状态信息被放置到 R_2 寄存器的缓冲区状态 (BSTATE) 区内。

在指定的缓冲区状态被提取之前和之后，执行串行化和检查点同步功能。

提取 QDIO 缓冲区状态的执行既不改变包含在子通道中的任何信息，也不会使 I/O 子系统询问或发信号给适配器。

参考图 14a - 14b 说明了与执行 EQBS 指令相关联的处理。除非发生异常状态，有利地，该指令由机器执行。如果发生了这种异常状态，则由主机完成该指令的处理。

参考图 14a，响应于发出 EQBS 指令，在步骤 1410，机器执行主机允许检查以便确定主机是否允许客户使用 QEBSM 工具。如果客户不是分页存储模式客户或者 QEBSMI 工具允许控制是 0，则在步骤 1412，将由主机完成处理。否则由机器继续处理。

接着，在步骤 1414，机器执行子通道判定。在一个例子中，通用寄存

器 1 中的子通道标记被用于定位 QBICB, 所述 QBICB 包括标识所述子通道的 SID。作出关于该子通道标记是否与存储在由 QBICB 指定的子通道中的标记相匹配的判定。如果否, 则由主机完成处理; 否则由机器继续处理。

在步骤 1416, 作出关于 EQBS 运算对象是否有效的判定。在一个例子中, 缓冲区号被验证。如果 R_1 缓冲区号 (BNUM) 区指定了一个大于例如 127 的值, 则在步骤 1418, 指令完成, 并且状态码和状态码限定符被设置。进而, 所述计数值被验证。如果 R_3 计数区指定了一个 0 值或大于例如 128 的值, 则指令完成, 并且状态码和状态码限定符被设置。进而, 用于指定队列的队列存储器描述符被定位。如果 R_1 队列索引区大于或等于 QBICB 队列计数区中的值, 则没有为子通道建立指定的队列, 指令完成, 并且状态码和状态码限定符被设置。否则, R_1 队列索引区的值被用作 QBICB 队列描述符地址数组中的索引, 以便确定队列描述符地址。

此后, 在步骤 1420, 由机器尝试为指定的队列获得队列存储器描述符锁。为指定的队列从 QBICB 队列描述符中的队列存储器地址区取回 QSD 地址。使用互锁更新操作作出尝试, 以便将 QSD 锁定字值从可用更改为由机器锁定。如果当尝试设置锁定字时, QSD 地址是 0 或存在访问异常状态, 则识别出异常。如果互锁更新操作失败, 该操作可以被重试 (次数取决于模型)。如果 QSD 锁定字不能被设置为预定的值, 则在步骤 1422, 识别出部分执行阻断, 并且由主机继续处理。然而, 如果互锁更新操作成功, 则由机器继续处理。

接着, 在步骤 1424, 用于第一个 QDIO 缓冲区的 SQBN 被定位 (图 14b)。该缓冲区的状态是当指令完成时被在 STATE 中返回的状态。具体地说, 通过从队列存储器描述符中的主机存储器列表地址区取回主机存储器列表地址, 用于指定队列的主机存储器列表被定位。然后通过从队列存储器描述符中的 SLSB 地址区取回 SLSB 地址, 指定队列的存储器列表状态块被定位。进而, 将被提取的第一个 SLSB 表项 (SQBN) 被定位。 R_1 缓冲区号区的值被用作 SLSB 中的索引, 以便确定将被提取的第一个 SQBN 的地址。

下面的逻辑被重复由 EQBS 指令上的计数区指定的 QDIO 缓冲区的数

目。初始地,在步骤 1426,用于将被处理的 QDIO 缓冲区的 SQBN 和 SBAL 被定位。接着在步骤 1428,作出关于该 SQBN 是否与将被返回的 STATE 相匹配的判定。即,如果至少一个缓冲区状态已经被提取,并且用于当前缓冲区的 SQBN 包括对不同于先前提取的状态的指定,则在步骤 1430,操作单元被禁止, EQBS 指令完成并且状态码和状态码限定符被设置。否则,处理继续。

如果在步骤 1431,指定的缓冲区当前由程序拥有,并且需要被提取(主机中的缓冲区同步 SBAL = 1),则为 SBAL 中的所有 SBALE (例如 16 个)执行:

- (1) 在步骤 1432,将主机 SBALE 中的 SBALF 区(除 SBALF 2 以外)复制到程序 SBALE 中的相应区。
- (2) 如果主机 SBAL 中的页钉住指示器为打开,则:
 - a) 在步骤 1434,使用主机虚拟地址(来自程序 SBALE)和主机实际地址(来自主机 SBALE)来调用 HPMA 解除钉住功能。
 - b) 在步骤 1436,作出关于解除钉住功能是否成功的判定。如果解除钉住功能没有成功,则在步骤 1438 由主机完成处理。否则由机器继续处理。
 - c) 在步骤 1438,在主机 SBALE 中将页钉住指示器设置为 0。
 - d) 如果队列是输入队列,则在步骤 1440,主机 SBALE 计数被复制到程序 SBALE (图 14c)。
 - e) 如果是格式 2 的输入队列,则在步骤 1442,主机 SBALE 的缓冲区数据偏移(即,缓冲区页中的数据偏移)被复制到客户 SBALE 中,保持缓冲区地址的其余部分不变。
- (3) 此后,在步骤 1444,缓冲区同步所需的以及 EQBS 操作所需的指示器被在主机 SBAL 中关闭,并且 QSD 中的状态转变计数被递增 1。

如果已确认输入缓冲区控制为打开,并且 SQBN 中的值指出输入缓冲区已准备好,则在步骤 1446, SQBN 被设置为输入缓冲区已确认状态。在步骤 1448, EQBS 缓冲区号也递增 1,并且在步骤 1450, EQBS 计数区递

减 1。

接着，在步骤 1452，作出关于 EQBS 计数是否等于 0 的判定。如果否，则处理继续，在步骤 1426，为将被处理的 QDIO 缓冲区定位 SQBN 和 SBAL（图 14b）。如果 EQBS 计数等于 0，则处理完成且指示成功。然后在步骤 1454，程序可以处理已经从适配器拥有转换为程序拥有状态的缓冲区。例如，可以处理已经到达的输入，或处理输出请求的完成（例如，收回缓冲区）。

如果 EQBS 处理必须由主机完成，则主机而不是机器向程序给出正确的状态码和状态码限定符。

如上所述，EQBS 指令可以调用由虚拟地址解除 HPMA 钉住功能。该功能为 CPU 提供了这样的能力：代表 I/O 请求来指出主机页的内容不再需要被钉在主机页帧内。将被解除钉住的页由主机虚拟地址指定。当页被成功地解除钉住时，与该页相关联的一个或多个已钉住页计数器可以被递减。

在一个例子中，该功能的输入包括将被解除钉住的虚拟地址（页）；标识分配给客户的地址空间和将被使用的转换表的有效地址空间控制元素（ASCE）；以及将被解除钉住的页的主机实际地址。

为了便于主机数据结构有效性判定，由虚拟地址解除主机页钉住功能的调用者负责确保间隔完成指示器在该功能被调用的整个检查点间隔中被设置。由虚拟地址解除主机页钉住功能不检查或修改间隔完成指示器。

参考图 16 说明了与由虚拟地址解除主机页钉住功能相关联的逻辑的一个实施例。初始地，在步骤 1600，用于页的 PTE/PGSTE 对被定位。使用将被解除钉住的主机虚拟地址和主机地址空间控制元素（ASCE），相应的页表表项和页状态表表项的主机实际地址被确定。一旦被定位，在步骤 1602，使用例如互锁更新操作（其将页控制互锁指示器从 0 更改为 1）来锁定所述 PTE/PGSTE 对。如果在询问 1604，PTE/PGSTE 对不能被锁定，或是如果 PGSTE.PCL 已经为 1，则在步骤 1606 处理以失败结束；否则处理继续。

在询问 1608，作出关于该页的钉住计数对于下一个解除钉住是否下溢的判定（即，PGSTE 中的钉住计数区的值是否为 0）。如果是，则在步骤 1606，处理再一次以失败结束；否则在步骤 1610，处理继续，取回将被

解除钉住的虚拟页的 PTE。在一个例子中，从主机存储器取回（块并发地）页表表项。

如果在询问 1612，页是无效的或者页无效位是 1，则在步骤 1606，处理再次以失败结束。否则在询问 1614，作出关于是否已经定位了一个不正确的 PTE 的判定。即，如果 PTE 中的页帧实际地址区没有指定与将被解除钉住的页的输入主机实际地址相同的页，则已经定位了一个不正确的 PTE，并且在步骤 1606 处理结束。否则处理继续。在步骤 1616，PGSTE 中的钉住计数区递减 1，并且在询问 1618，作出关于页状态是否被从钉住更改为解除钉住的判定。如果页状态被从钉住更改为解除钉住状态，则在步骤 1620，关联的增量已钉住页计数递减 1。即，如果钉住计数现在是 0，并且钉住计数没有处于溢出状态（PGSTE.PCO = 0），则页状态已经从钉住转换为解除钉住状态，并且增量已钉住页计数递减 1。

此后，或者如果页状态没有从钉住被更改为解除钉住，则在步骤 1622，通过例如将页控制互锁指示器设置为 0 来解锁 PTE/PGSTE 对。这完成了由虚拟地址解除钉住功能的处理。

当解除钉住功能退出且指出失败时，没有作出对状态表的改变。在这种情况下，机器就像没有调用该功能，并且主机被给予控制以便处理所述失败。当该功能退出且指出成功时，指定的页已经被解除钉住。在这种情况下，主机状态表已经由机器更新，并且先前对传递控制给主机以便进行页钉住的要求被减轻了。

另一个可以被调用的功能是由 PTE 地址解除主机页钉住功能。该功能为 CPU 提供了这样的能力：指出主机页的内容不再被钉住在主机页帧内。将被解除钉住的页由主机页表表项的主机实际或绝对地址来指定。当页被成功地解除钉住时，与页表相关联的一个或多个已钉住页计数器被递减。

在一个例子中，由 PTE 地址解除主机页钉住功能的输入包括将被解除钉住的页的主机实际地址或主机页表表项。

参考图 17 说明了与由 PTE 地址解除主机页钉住功能相关联的逻辑的一个实施例。初始地，在步骤 1700，相应于输入页表表项的页状态表表项被定位。然后在步骤 1702，PTE/PGSTE 对被锁定。在一个例子中，使用互锁更新操作来尝试将 PGSTE 中的页控制互锁指示器从 0 更改为 1。如果

在询问 1704, PTE/PGSTE 对不能被锁定, 或是如果 PGSTE.PCL 已经是 1, 则在步骤 1706 处理以失败退出。

另一方面, 如果锁定成功, 则在询问 1708, 作出关于页的钉住计数对于下一个解除钉住是否将下溢的进一步判定(即, PGSTE 中的钉住计数区的值是否为 0)。如果是的, 则在步骤 1706, 处理再次以失败退出。否则在步骤 1710, 用于将被解除钉住的虚拟页的 PTE 被取回。在一个例子中, 它被从主机存储器取回(块并发地)。

此后, 在询问 1712, 作出关于该页是否无效的判定。具体地说, 检查页无效指示器。如果它是 1, 则该页是无效的, 并且在步骤 1706 处理以失败退出。如果该页有效, 则在步骤 1714, PGSTE 中的钉住计数区递减 1。

接着, 在询问 1716, 作出关于页状态是否被从钉住更改为解除钉住的判定。如果钉住计数现在是 0 并且钉住计数没有处于溢出状态, 则页状态已经从钉住转换到解除钉住状态。因此, 在步骤 1718, 增量已钉住页计数被递减 1。此后, 或是如果页状态没有从钉住更改为解除钉住, 则在步骤 1720, PTE/PGSTE 对被解锁。在一个例子中, 这包括将页控制互锁指示器设置为 0。这结束了由 PTE 地址解除页钉住功能的处理。

当解除钉住功能退出且指出失败时, 没有对状态表作出改变。在这种情况下, 机器就像没有调用该功能, 并且主机被给予控制以便处理所述失败。当该功能退出且指出成功时, 指定的页被解除钉住。在这种情况下, 主机状态表已经由机器更新, 并且先前对将控制传递给主机以便进行页钉住的要求被减轻了。

上面的详细说明是在没有主机干预的情况下与来自可分页存储客户的解释 I/O 操作请求相关联的处理。所述请求可以是基于网络协议的通过 QDIO 适配器或类似适配器对网络协议(例如, 网际协议)或网络设备的请求。作为进一步的例子, 所述请求可以是基于光纤通道协议(FCP)的对连接到光纤通道的设备的请求。

与诸如 QDIO 适配器之类的非 FCP 适配器类似, FCP 适配器通过适配器与主存储器之间的通道路径来连接到 CPC 中的其它元件。然而, 与 FCP 适配器关联的处理在某些方面和与非 FCP 适配器关联的处理不同。例如, 数据输入和输出是通过 QDIO 输出队列执行的, 并且 QDIO 输入队

列仅用于适配器发送信号通知 I/O 操作的完成。进而，当 FCP 适配器将输出队列缓冲区置于程序拥有状态时，包括数据传输的 I/O 操作可能没有全部完成。（这与用于非 FCP 适配器的 QDIO 体系结构相反。）即使缓冲区处于程序拥有状态，程序也不对缓冲区页执行任何活动，直到适配器通过输入队列发送信号通知 I/O 操作完成为止。然而，程序可以在与所述缓冲区相关联的 QDIO 数据结构上执行活动，只要所述缓冲区页未被改变。

当 FCP 操作被启动时，一个唯一请求 id 被从程序传递给适配器。是此在输入队列中被返回的请求 id 发送信号通知操作完全结束并且程序可以与缓冲区页交互。

因此，借助 FCP，可以具有将缓冲区页从 QDIO 数据结构断开的能力。换言之，QDIO 数据结构中的缓冲区页的地址的拷贝可以被保存在另一个位置，并且可以分配新的缓冲区页，它们的地址被置于该 QDIO 数据结构中，具体地说，置于希望的 SBALE 内，并且启动新的 I/O 操作 - 始终等待来自适配器（在缓冲区页的最初集合上的操作已经完成）的请求 id（谁的地址，复制了哪里）。

为了便于保持已断开连接的缓冲区页的描述，设计了 FCP 操作块（FOB）。FOB 是一种结构，该结构当由指定 FCP 请求的客户发出 SQBS 时被建立，并且在机器解决和钉住页时，它在 FOB 中记录缓冲区页的主机页表表项地址。机器还在 FOB 中记录客户的请求 id，并使用等于该 FOB 的地址的请求 id 来替换所述客户的请求 id。然后，当稍后适配器返回输入队列中的该请求 id 并且客户发出 EQBS 以便询问输入队列的状态时，机器可以定位该 FOB、解除页的钉住、使用先前记录的初始客户值来替换所述请求 id。

参考图 18a - 18d 和 19a - 19e 说明了关于 FCP 的处理的进一步的详细信息。具体地说，参考图 18a - 18d 说明了与用于 FCP 协议的 SQBS 指令相关联的处理的一个例子，并且参考图 19a - 19e 说明了与用于 FCP 协议的 EQBS 指令相关联的处理的一个例子。

首先参考图 18a，在为 FCP 调用设置 QDIO 缓冲区状态指令之前，执行某些初始化。例如，在步骤 1800，主机执行各种清除任务（图 18a），包括例如删除和处理已处理帧描述符列表上的帧。除了执行清除之外，在

步骤 1802, 主机确保在可用帧描述符列表和可用 FOB 列表上有足够的帧。

在执行清除和/或确保在相关的列表上具有足够的帧之后, 在步骤 1804, 主机可以启动程序的解释执行。在步骤 1806, 程序开始执行, 并且在执行过程中程序选择将被传输给适配器的下一个 QDIO 缓冲区。在步骤 1808, 程序以将被用于数据传输的存储器块的程序绝对地址来填充 SBAL。进而, 在步骤 1810, 通过发出设置 QDIO 缓冲区状态指令将 QDIO 缓冲区的状态从程序拥有更改为适配器拥有, 来将缓冲区的所有权传递给适配器。

响应于发出 SQBS 指令, 在步骤 1812, 机器执行主机允许检查以便确定主机是否允许客户使用 QEBSM 工具。如果在步骤 1814, QEBSNI 允许控制 (例如, 在客户状态中) 被设置为 0, 指出不允许, 则在步骤 1816, 由主机而不是机器来完成处理。当在设置 QDIO 缓冲区状态解释过程中遇到需要来自主机的某些活动或服务状态时, 指令阻断或者部分执行阻断 (PEI) 被识别, 使用设置 QDIO 缓冲区状态指令文本来设置阻断参数, 并且对于 PEI, 在客户状态中设置 QEBSM - PEI 参数。

接着, 在步骤 1818, 机器执行子通道判定, 其中作出关于子通道标记是否与存储在由 QBICB 指定的子通道内的标记相匹配的判定。如果不匹配, 则在步骤 1820 由主机完成处理。否则, 由机器继续处理。

此外, 在步骤 1822, 由机器作出关于 SQBS 运算对象是否有效的判定。如果任意运算对象无效, 则处理结束, 并且在步骤 1824 设置状态码和状态码限定符。否则处理继续。

接着, 在步骤 1826, 尝试获得用于指定队列的队列存储器描述符锁。如果没有获得 QSD 锁, 则在步骤 1828 由主机完成处理。否则由机器继续处理。

下面的逻辑被重复由 SQBS 指令的计数区指定的 QDIO 缓冲区的数目。初始地, 在步骤 1830, 用于将被处理的 QDIO 缓冲区的 SQBN、程序 SBAL 和相应的主机 SBAL 被定位。此后, 在步骤 1832, 作出关于指定的缓冲区当前是否由程序拥有的判定。如果该缓冲区不是由程序拥有, 则在步骤 1834, 处理以状态码和状态码限定符结束。否则在步骤 1836, 处理继续, 判定指定缓冲区的状态是否将被使用 EQBS 指令来提取 (图 18b)。如果是, 则在步骤 1838, 处理再次以状态码和状态码限定符结束。否则,

在步骤 1840, 处理继续, 且将客户 SBAL 的内容复制到主机 SBAL。此外, 在步骤 1842, 缓冲区同步和 EQBS 操作所需的位在主机的 SBAL 中被设置。

进而, 如果 SBAL 缓冲区序列位指出这是一个新的 FCP 请求 (与跨多个 SBAL 的请求的继续相反), 则在步骤 1844, 以前的 FCP 请求 (如果有) 被标记为已完成。例如, 为新的请求的指示检查主机 SBALE 0 表项: 如果客户 SBALE 0 的 SBAL - Flags (SBALF) 区中的缓冲区序列类型指出一个命令被指定, 则 QSD 地址的队列标志区中的未决请求组件 FOB 列表完成位是 1, 并且有效请求头部 FOB 列表中存在 FOB, 该列表中的第一个头部 FOB 的标志区中的构造完成位被设置为 1, 并且 QSD 地址的队列标志区中的未决请求组件 FOB 列表完成位被设置为 0。

接着, 在步骤 1846, 作出关于在可用 FOB 列表上是否存在 FOB 的判断。如果没有, 在步骤 1848 由主机完成处理。否则由机器继续处理。例如, 在步骤 1850, 可用 FOB 列表中的第一个可用 FOB 被从该列表上删除, 并且可用 FOB 计数递减 1。然后, 在步骤 1852, 基于 SBAL 中的缓冲区序列位, 新的 FOB 被添加到适当的列表。例如, 如果这是第一个或唯一的 SBAL, 则这是一个新的 FCP 请求。在这种情况下, 该 FOB 是头部 FOB 并且被插入到有效请求 FOB 列表, 并且该请求在该 FOB 中被标记为有效。然而, 如果这是一个中间或最后的 SBAL, 则这是多个 SBAL FCP 请求的继续或是最后的 SBAL。因此, 该 FOB 是非头部组件 FOB, 并且被添加到被在队列存储器描述符中锚定的有效请求 FOB 列表上的第一个头部 FOB 的请求组件 FOB 列表。

对于头部 FOB, 在步骤 1854, 发生下面的处理:

- 1) 在步骤 1856, 在 QSD 标志中的未决请求组件完成指示器被设置;
- 2) 在步骤 1858, SBALE 0 中由程序指定的请求标识符被保存在客户请求 id 区中;
- 3) 在步骤 1860, QSD 的地址被存储在 FOB 中的队列存储器描述符地址区内, 以便标识与该请求相关联的特定队列 (图 18c);
- 4) 在步骤 1862, QBICB 的地址被存储在 QDIO 缓冲区信息控制块地址区内, 以便将子通道与该请求相关联;

5) 在步骤 1864, 头部 FOB 的地址被如主机 SBAL 的 SBALE 0 中的请求 id 区那样设置; 以及

6) 在步骤 1866, 来自 SBALF 0 的 FCP 存储器块类型被存储在 FOB 内的存储器块类型中, 以便定义将被执行的存储操作的类型。

进而, 在步骤 1868, 对于具有有效存储器块地址的每个有意义的主机 SBALE, 下面的功能被执行:

1) 在步骤 1870, 从主机 SBALE 取回程序绝对(主机虚拟)存储器块地址;

2) 在步骤 1872, 调用 HPMA 钉住功能以便转换和锁定存储器中的页。在一个例子中, HPMA 功能被调用并且被传递给下面的输入数据项:

- 将被钉住的主机虚拟地址: 来自主机 SBALE 的客户存储器块数据地址是将被钉住的虚拟地址。

- 页将被钉住在其中的地址空间的有效主机地址空间控制元素 (ASCE): 主地址空间控制元素 (PASCE) 被指定为有效 ASCE。

- 取回或存储访问指示: 如果指定的队列是输入队列, 则指出存储访问操作。如果指定的队列是输出队列, 并且 QBICB 队列格式 (QFMT) 区指出该队列不是格式 1 的输出队列, 则指出取回-访问操作。如果指定的队列是输出队列, 并且 QBICB 队列格式 (QFMT) 区指出该队列是格式 1 的输出队列, 则指出取回访问操作 (如果主机 SBALE 指定了写缓冲区存储器块); 否则指出存储访问操作。

- 对页内容是否可以被替换的指示: 如果指定的队列是输入队列, 则指出页内容替换。如果指定的队列是输出队列, 并且 QBICB 队列格式 (QFMT) 区指出该队列不是格式 1 的输出队列, 则不指出页内容替换。如果指定的队列是输出队列, 并且 QBICB 队列格式 (QFMT) 区指出该队列是格式 1 的输出队列, 则指出页替换(如果主机 SBALE

指定了状况读缓冲区或读缓冲区存储器块，并且 SBALE 中的数据计数是例如 4K 字节）。

3) 在步骤 1874, 如果 HPMa 钉住功能没有成功, 则在步骤 1876 由主机完成处理。否则由机器继续处理。例如, 如果 HPMa 钉住主机页功能完成且指示成功, 则由钉住主机页功能返回的页帧实际地址被复制到主机 SBALE 的数据地址区内, 并且当前 SBALE 的虚拟化标志 (VFlags) 区中的已钉住页指示器位被设置为 1。如果 HPMa 钉住主机页功能完成且指出应识别一个有效性阻断, 则被指出的有效性阻断被识别。如果 HPMa 钉住主机页功能完成且指示失败, 则部分执行阻断被识别, 并且根据 HPMa 钉住主机页功能失败原因, 设置客户状态中的 QEBSM - PEI 参数区中的部分执行阻断原因码; 否则处理继续。

4) 在步骤 1878, 由 HPMa 返回的 PTE 地址被移动到 FOB PTE 地址数组中的下一个可用表项内, 并且 PTE 计数区递增 1 (图 18d); 以及

5) 在步骤 1880, 来自主机 SBALE 的数据计数区被添加到 QSD 中的字节传输计数。

如果这是多个 SBAL 的最后一个 SBAL, 或是 FCP 请求的唯一 SBAL, 则该请求在 FOB 中被标记为完成, 并且未决请求 FOB 列表完成指示器被设置为 0。

此后, 在步骤 1884, 缓冲区的 SQBN 被设置为由 SQBS 指令指定的适配器拥有状态。此外, 在步骤 1886, SQBS 缓冲区数目递增 1, 并且在步骤 1888, SQBS 计数区递减 1。

在步骤 1890, 作出关于 SQBS 计数是否等于 0 的判定。假如 SQBS 计数不等于 0, 则在步骤 1830, 处理继续, 为将被处理的 QDIO 缓冲区定位 SQBN、程序 SBAL 和主机 SBAL (图 18a)。然而, 如果 SQBS 计数等于 0, 则在步骤 1892, 处理继续且解锁 QSD 锁定。如果成功, 在步骤 1894, 成功指示被转发给程序。如果这是将在此时传输的最后一个 QDIO 缓冲区, 则在步骤 1896 发出 SIGA-wt 指令。

除了与处理用于 FCP 协议的 SQBS 指令相关联的上述更改之外，还存在与 EQBS 指令相关联的更改。参考图 19a - 19e 说明了与用于 FCP 请求的 EQBS 指令相关联的逻辑的一个例子。

在运行程序之前，进行某些初始化。例如，在步骤 1900，主机执行各种清除任务，包括例如删除和处理已处理帧描述符列表上的帧。进而，在步骤 1902，主机确保在可用帧描述符列表和可用 FOB 列表上有足够的帧。此后在步骤 1904，主机开始该程序的解释执行。

在执行过程中，在步骤 1906，程序选择下一个输出 QDIO 缓冲区或程序正在等待适配器结束处理的缓冲区。然后在步骤 1908，程序发出 EQBS 指令以便获取所选缓冲区（多个）的当前状态。

响应于发出 EQBS 指令，在步骤 1910，机器执行主机允许检查以便确定主机是否允许客户使用 QEBSM 工具。如果在询问 1912，客户不是分页存储模式客户，或是 QEBSMI 工具允许控制为 0，则在步骤 1914 由主机完成处理。否则由机器继续处理。

接着，在步骤 1916 机器执行子通道判定。如果子通道标记与被存储在由 QBICB 指定的子通道中的标记不匹配，则在步骤 1914 由主机完成处理；否则由机器继续处理。

进而，在步骤 1918，作出关于 EQBS 运算对象是否有效的判定。如果一个或多个运算对象无效，则处理结束，并且在步骤 1920 设置状态码和状态码限定符。否则处理继续。

在步骤 1922，由机器尝试为指定的队列获得队列存储器描述符锁。如果没有获得 QSD 锁，则部分执行阻断被识别，并且在步骤 1924 由主机继续处理。然而，如果获得了锁（例如，互锁更新操作成功），则由机器继续处理。

接着，在步骤 1926，用于第一个 QDIO 缓冲区的 SQBN 被定位。此缓冲区的状态是当指令完成时将在 STATE 中被返回的状态。

下面的逻辑被重复由 EQBS 指令的计数区指定的 QDIO 缓冲区的数目。初始地，在步骤 1928，用于将被处理的 QDIO 缓冲区的 SQBN、程序 SBAL 和相应的主机 SBAL 被定位。接着，在步骤 1930，作出关于 SQBN 是否与将被返回的 STATE 相匹配的判定（图 19b）。如果否，则处理结束，

并且在步骤 1932 设置状态码和状态码限定符。否则处理继续。

如果队列是 FCP（例如，格式 1）输出队列，并且 SQBN 指定了程序拥有状态，则在步骤 1934，缓冲区同步所需的和 EQBS 操作所需的指示器在主机 SBAL 中被关闭。此外，在步骤 1936，各 SBALF 区（除 SBALF 2 以外）被复制到程序 SBALE 中的相应区内。

如果指定的缓冲区当前由程序拥有，并且它是格式 1 的输入队列，则在步骤 1938，为在主机 SBAL 中 SBALE 请求标识符区不是 0 的有意义的 SBALE 执行下面的步骤：

- 1) 代表已完成的 FCP 请求的头部 FOB 的绝对地址是在主机 SBALE 中被返回的请求标识符。在步骤 1940，作出关于该请求标识符是否指定了有效的 FOB 的判定。例如，作出关于由请求 id 所指的 FOB 中指定的 QBICB 中的子通道标记是否与 SQBS 指令上指定的标记相匹配的检查。如果否，在步骤 1942 由主机完成处理；否则由机器继续处理。
- 2) 通过验证 FOB 中指定的 QBICB 中的子通道标记与 SQBS 指令上指定的标记相匹配，为此用户验证由适配器返回的头部 FOB。如果在步骤 1944，所述 FOB 并非用于此用户，则在步骤 1946 由主机完成处理；否则由机器继续处理。
- 3) 尝试获得用于在 FOB（在 FCP 请求启动时被建立）中指定的输出队列的 QSD 锁。如果在步骤 1948，在合理的时间内不能获得该锁（图 19c），则在步骤 1950 由主机完成处理；否则由机器继续处理。在这个步骤以及在随后的步骤（其中在机器已经获得了一个或多个锁之后，将由主机完成处理）中，这些锁的状态首先被更改为“由机器为主机锁定”。
- 4) 验证 FOB 当前在有效请求列表上并且已被完成。如果在步骤 1952，请求有效和构造完成指示器都是打开（on），并且在 FOB 标志中请求 id 已投递指示器为关闭（off），则在步骤 1954 由主机完成处理；否则由机器继续处理。
- 5) 在步骤 1956，头部 FOB 标志区中的请求 id 已投递指示器被设置为打开。

6) 在步骤 1958, 对于代表所述请求的头部 FOB 以及可以被与该头部 FOB 关联和断开与该头部 FOB 链接的组件 FOB (所述头部 FOB 被最后处理), 发生下面的逻辑:

a) 在步骤 1960, 对于在由 FOB PTE 计数指定的范围内的 FOB 中的每个非 0 PTE 地址数组表项, 调用 HPMA 由 PTE 地址解除主机钉住功能。如果解除钉住功能没有成功, 则在步骤 1962 由主机完成处理; 否则由机器继续处理。

b) 如果这是一个组件 FOB, 则在步骤 1964, 其被从与所述头部 FOB 解除锚定的请求组件 FOB 列表删除 (图 19d)。

c) 如果这是一个头部 FOB, 则在步骤 1966, 将来自该头部 FOB 的客户请求 id 复制到程序 SBALE 的请求 id 区内, 将主机 SBALE 中的请求 id 置 0 以便指出该请求 id 已经被处理, 并且从与输出队列 QSD 解除锚定的有效请求 FOB 列表删除所述头部 FOB。

d) 在步骤 1968, 清除该 FOB 的预先指定的字, 将其插入可用 FOB 列表的开始处, 并且将可用 FOB 计数递增 1。

7) 在步骤 1970, 释放输出队列 QSD 锁。

8) 在步骤 1972, 缓冲区同步所需的以及 EQBS 操作所需的指示器被在主机 SBAL 中关闭, 并且状态转换计数递增 1。

如果已确认的输入缓冲区控制为 1, 并且 SQBN 中的值指出输入缓冲区准备好状态, 则在步骤 1974, SQBN 被设置为输入缓冲区已确认状态。

此后, 在步骤 1976, EQBS 缓冲区号被递增 1, 并且在步骤 1978, EQBS 计数区被递减 1。如果在步骤 1980, EQBS 计数区不等于 0 (图 19e), 则在步骤 1928 处理继续, 为将被处理的 QDIO 缓冲区定位 SQBN、程序 SBAL 和相应的主机 SBAL (图 19a)。否则在步骤 1982, 程序结束。然后在步骤 1984, 程序可以处理已经从适配器拥有转换为程序拥有状态的缓冲区。

如上所述, QDIO 体系结构被扩展, 以使可分页存储模式客户可以体验到接近本地 QDIO 的性能。这种 V = V QDIO 通过能力可以接近与处理器效率相同的水平, 如在没有虚拟机操作环境的情况下本地地得到的那样。这是通过将来自可分页存储模式客户执行地址空间的 QDIO 缓冲区状态信

息重新定位到主机（例如，VM 主机）存储器内，并且实现指令以允许客户操纵缓冲区状态并相应地管理主机转换和影子结构来完成的。引入了两个指令，所述指令赋予程序通过队列号/缓冲区号指定（而不是通过在程序存储器内直接寻址缓冲区状态）来管理 QDIO 缓冲区状态的能力。这种缓冲区状态指定的抽象形式允许以提高可分页模式虚拟机 QDIO 操作的性能的方式来实现这些指令的解释版本。它还允许减少 SIGA 开销（更少的 SIGA）。此外，子通道标记表示允许为在程序中分配的子通道进行 SIGA 的解释，从而在大多数情况下消除了 SIGA 处理中牵涉到主机。

当指定对 QDIO 增强缓冲区状态管理允许的请求的建立 QDIO 队列 CCW 命令成功地完成时，QDIO 子通道被允许进行 QDIO 增强缓冲区状态管理。

当没有指定对 QDIO 增强缓冲区状态管理允许的请求的建立 QDIO 队列 CCW 命令成功地执行，或是发生了使队列从子通道解散的活动或事件时，QDIO 子通道被禁止进行 QDIO 增强缓冲区状态管理。

是 CHSC 存储子通道 QDIO 数据命令响应块的一部分的子通道 QDIO 数据块中的 QDIO 适配器特性 1 (QDIOAC1) 区的位 7 指出对于特定的子通道是否允许 QDIO 增强缓冲区状态管理。

当 QDIO 子通道被允许 QDIO 增强缓冲区状态管理时，对于下面的指令，由通用寄存器 1 中的子通道标记来指定子通道：

- 发信号通知适配器 (SIGA) - 具体地说，SIGA-rt (由标记读) 和 SIGA-wt (由标记写)，它们是 SIGA-r 和 SIGA-w 的替代物。SIGA-rt 和 SIGA-wt 指令分别执行与 SIGA-r 和 SIGA-w 相同的功能；然而，使用标记而不是子系统 id 来指定子通道。在标题为 “Self-contained Queues With Associated Control Information For Receipt And Transfer Of Incoming And Outgoing Data Using A Queued Direct Input-output Device” (“用于使用排队直接输入-输出设备来接收和传输进入和外出数据的具有关联控制信息的自足队列”，Baskey 等人，2001 年 12 月 18 日公布，其全部内容在此引入作为参考) 的美国专利 No.6,332,171 B1 中说明了 SIGA-r 和 SIGA-w 的例子；
- 设置 QDIO 缓冲区状态 (SQBS)；以及

- 提取 QDIO 缓冲区状态 (EQBS)。

从是对 CHSC 存储子通道 QDIO 数据命令 (在子通道已被允许 QEBSM 之后执行) 的响应的一部分的子通道 QDIO 数据块的子通道标记区来获得子通道标记。因此, 为了实现 QEBSM 允许并且为特定子通道获得子通道标记, 程序执行如下序列:

1. 执行 CHSC 存储子通道 QDIO 数据以便验证 QEBSM 可用于子通道。

2. 执行建立 QDIO 队列 CCW 命令以便允许子通道使用 QEBSM。

3. 执行 CHSC 存储子通道 QDIO 数据以便验证 QEBSM 被允许, 以获得子通道标记并检查适配器的特性, 作为成功执行建立 QDIO 队列 CCW 命令的结果, 某些适配器特性可能已改变。

在 QDIO 子通道已经被允许 QEBSM 之后, 使用设置 QDIO 缓冲区状态 (SQBS) 指令来改变在子通道上有效的队列缓冲区的状态。适时地使用提取 QDIO 缓冲区状态 (EQBS) 指令来询问在子通道上有效的队列缓冲区的状态允许主机页被以这样的方式解除钉住: 即不会不适当地对存储器造成与大量可分页存储模式客户的管理相关联的过度压力。可替代地, 如果程序没有良好地运行或是对主机存储器的要求高, 则主机可以解除钉住在程序拥有状态中的任何已钉住的缓冲区页, 而无需等待客户发出 EQBS。

一旦缓冲区状态已经通过设置 QDIO 缓冲区状态被从程序拥有状态更改为适配器拥有状态, 该状态将由提取 QDIO 缓冲区状态询问并且在另一个设置 QDIO 缓冲区状态被再次用于该缓冲区之前, 被发现已回到程序拥有状态。概括地说, 客户准备数据缓冲区并且执行设置 QDIO 缓冲区状态 (SQBS) 指令。SQBS 指令尝试确保所有指定的缓冲区页被驻留。如果缓冲区页尚未被引用, 则使用主机页管理辅助功能来动态地在页表中分配新的帧、钉住页以及更新主机影子 QDIO SBAL 表项。然而, 如果缓冲区页存在且当前被换出 (page out), 并且是输入队列缓冲区页, 则使用主机页管理辅助功能来动态地在页表中分配新的帧、钉住页以及更新主机影子 QDIO SBAL 表项。

适配器启动的从适配器拥有到程序拥有状态的状态改变可以由客户通过提取 QDIO 缓冲区状态 (EQBS) 指令直接观察到, 所述提取 QDIO 缓冲区状态 (EQBS) 指令指定了程序先前使其置于适配器拥有状态并且返

回当前缓冲区状态的缓冲区。

除了上述之外，主机页管理辅助工具（HPMA）可以由其它 CPU 工具在可分页存储模式客户的解释期间来调用。HPMA 工具的一个或多个方面可以与 QDIO 体系结构一起使用或不与 QDIO 体系结构一起使用。HPMA 包括例如 4 种 HPMA 辅助功能，诸如解决主机页功能、钉住主机页功能、由虚拟地址解除主机页钉住功能以及由 PTE 地址解除主机页钉住功能。

如果在指令取回或指令运算对象访问上存在主机页无效状态，并且主机页管理辅助被允许，则可以调用解决主机页功能。解决主机页功能可以在转换处理期间使用，以便动态地解决主机页无效状态，并且防止页转换异常被识别。如果解决主机页功能成功地完成，在一个例子中，主机页无效状态存在的操作单元被无效并重试。如果解决主机页功能没有成功地完成，则如主机页管理辅助没有被允许那样处理最初的主机页无效状态。

QEBSM 结合 HPMA 提供了下面的优点和性能改进：更少的 SIE 表项/退出开销和更少的主机环境切换和路由开销（用于生产性工作的更多的周期）；消除了 QDIO 缓冲区状态的主机影子拷贝以及关联的同步开销；动态更新页状态表和转换表；动态防止主机页故障；改进的 QDIO 建立影子；以及在无主机干预下的解释性执行模式中由 CPU 对主机影子结构的动态维护。

虽然此处说明了各种例子和实施例，但是这些仅是例子，并且许多变形被包括在本发明的范围内。例如，此处描述的计算环境仅是一个例子。许多其他环境，包括其他类型的通信环境，可以包括本发明的一个或多个方面。例如，可以采用不同类型的处理器、客户和/或主机。此外，其他类型的体系结构可以采用本发明的一个或多个方面。

进而，在此处提供的数据结构和流程的例子中，不同区的创建和/或使用可以包括许多变化，诸如不同数量的位；不同顺序的位；更多、更少或不同于此处说明的位；更多、更少或不同的区；不同顺序的区；不同大小的区等。再次地，这些区仅被作为例子提供，并且可以包括许多变化。进而，此处说明的指示器和/或控制可以具有许多不同形式。例如，可以以除了位以外的方式来表示它们。此外，虽然此处使用了术语地址，但是可以使用任何名称。

此外，虽然以互联网协议和光纤通道协议说明了本发明的各方面，再

次地，这些只是例子。其他协议可以从本发明的一个或多个方面获益，包括但不限于其他网络协议。

再次地，虽然参考 QDIO 和此处说明的例子描述了 HPMA，但是 HPMA 的一个或多个方面可自身使用。与 QDIO 一起使用仅是一个例子。

如此处所使用的，术语“页”被用于指固定大小或预定大小的存储器区域。页的大小可以变化，虽然此处提供的一个例子中页是 4K。类似地，缓冲区是保存数据的一种结构，并且可以有許多结构类型和结构大小。类似地，存储器块是存储器的块，并且如此处所使用的，等价于一个存储器页。然而，在其他实施例中，可以有不同大小的存储器块和/或页。许多其他替代物也是可能的。进而，虽然此处使用了诸如队列、列表、表之类的术语，但是可以使用任何类型的数据结构。再次地，此处提到的这些仅是例子。

本发明可以被包括在具有例如计算机可用介质的制品（例如，一个或多个计算机程序产品）中。所述介质中包含例如计算机可读程序代码装置或逻辑（例如，指令、代码、命令等），以便提供和促进本发明的功能。所述制品可以被包括为计算机系统的一部分或单独销售。

此外，可以提供至少一个计算机可读的程序存储装置，该装置包含至少一个计算机可执行的指令程序以执行本发明的功能。

此处示出的流程图只是实例。在不偏离本发明的精神的情况下，此处所述的这些图或步骤（或操作）可以有多种变化。例如，可以按不同顺序执行步骤，或者可以添加、删除或修改步骤。所有这些变化都被视为要求保护的发明的一部分。

虽然此处给出并详细说明了优选实施例，对相关领域的技术人员显而易见的是，可以做出各种修改、增加、替代等而不偏离本发明的精神，并且因此它们被看作在如以下权利要求限定的本发明的范围之内。

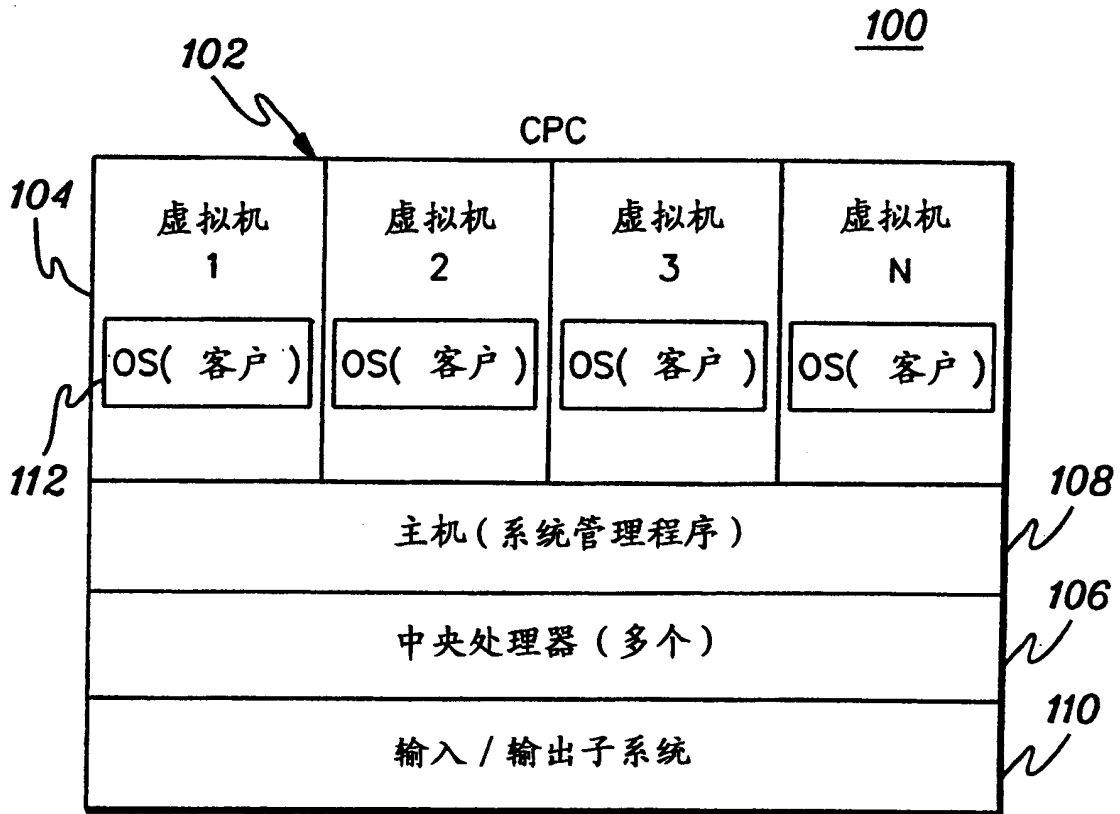


图 1A

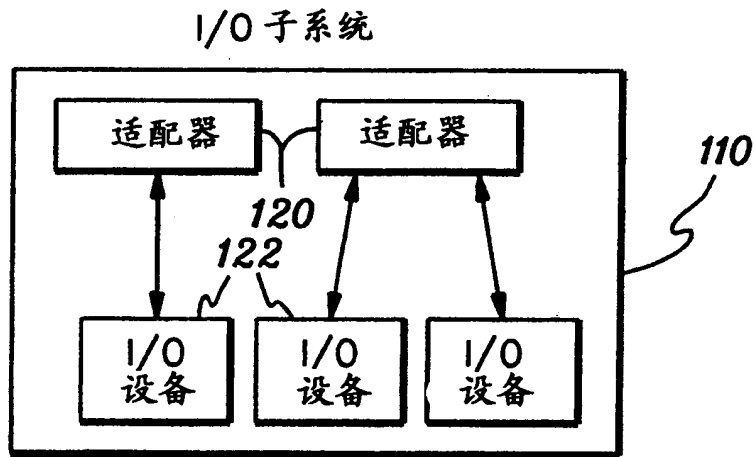


图 1B

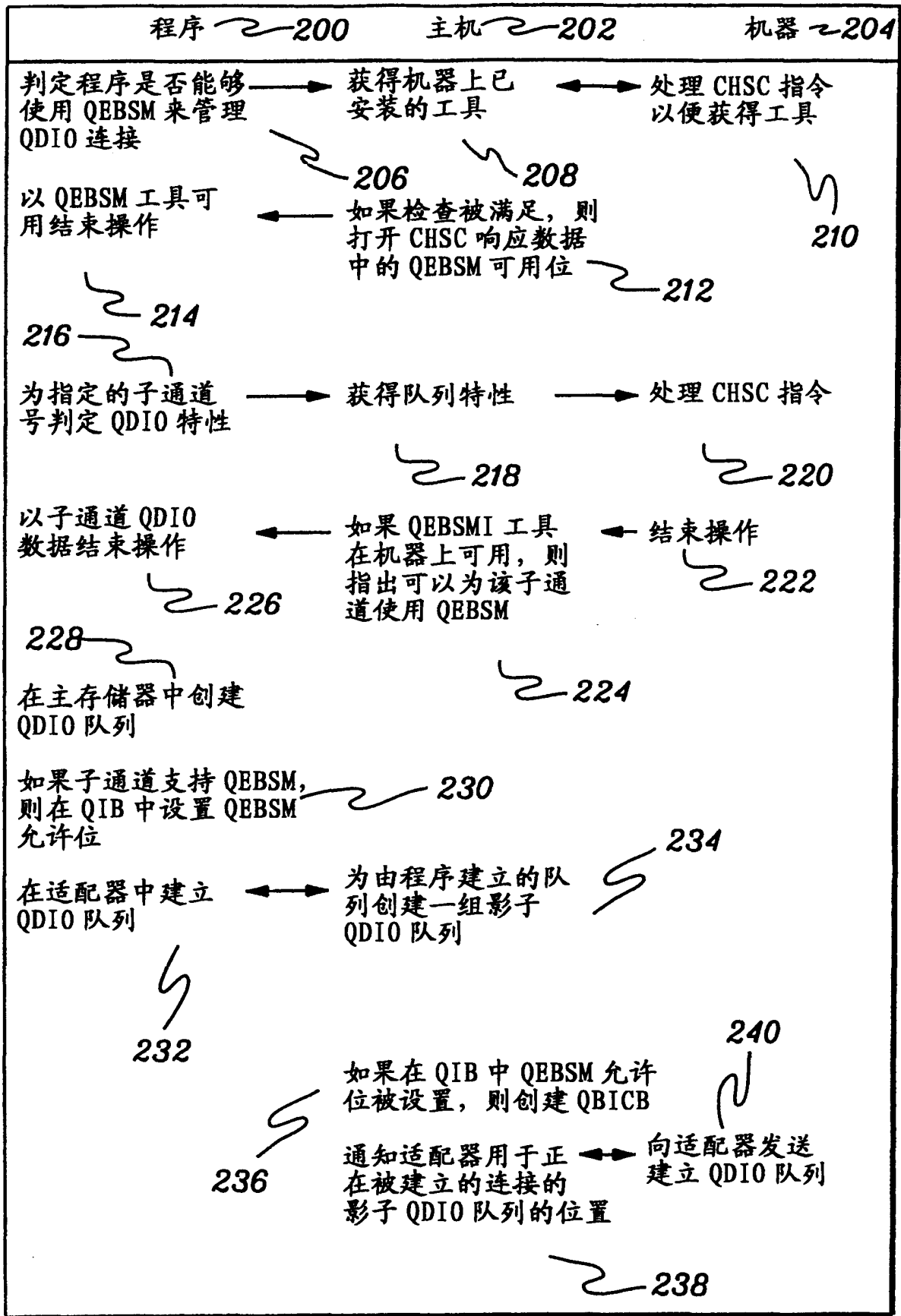


图 2A

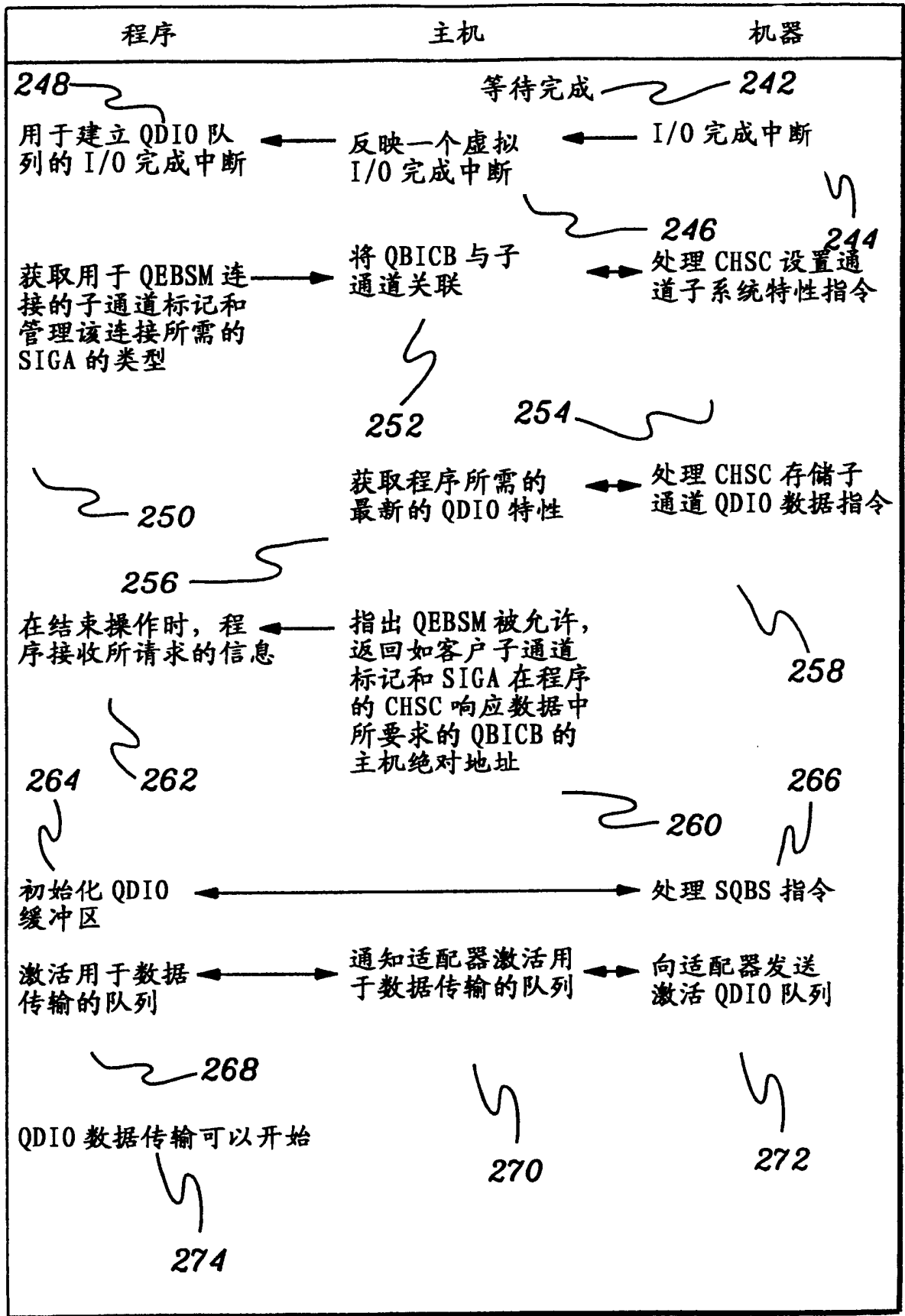


图 2B

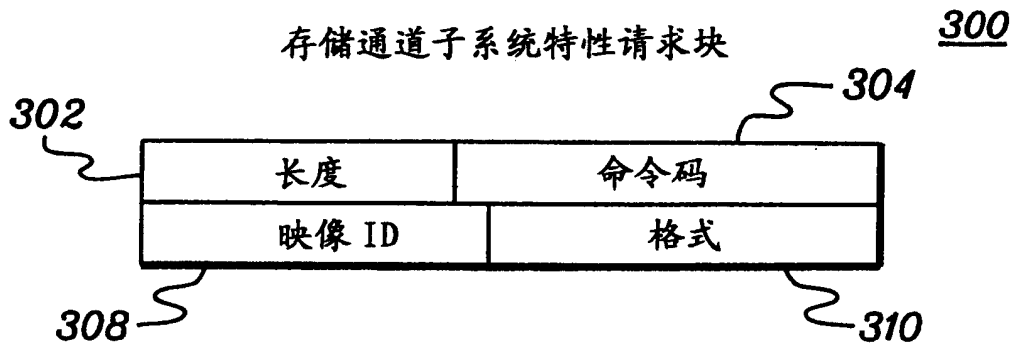


图 3A

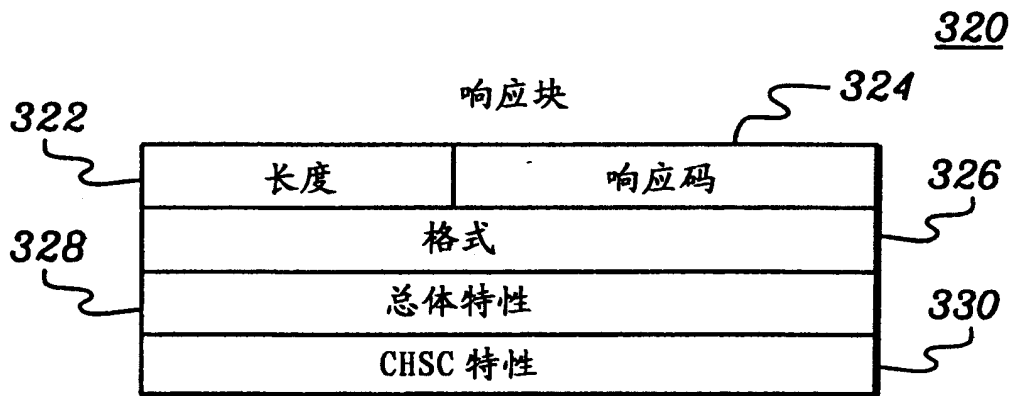


图 3B

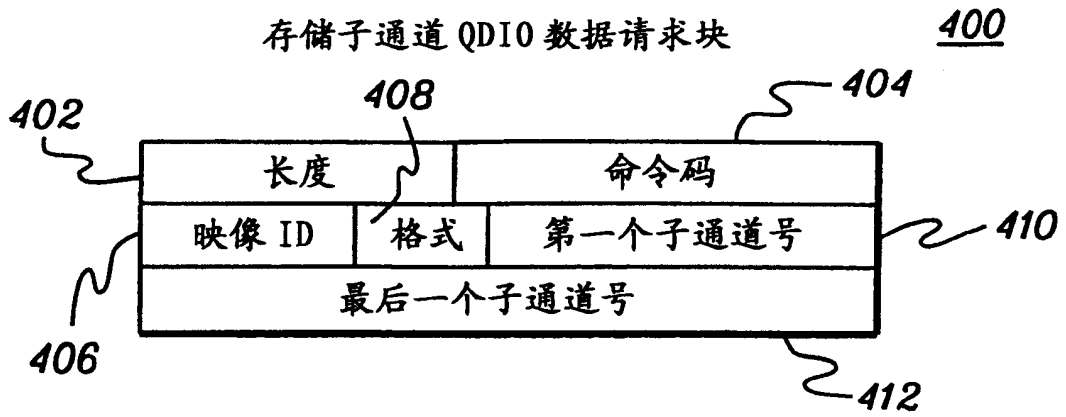


图 4A

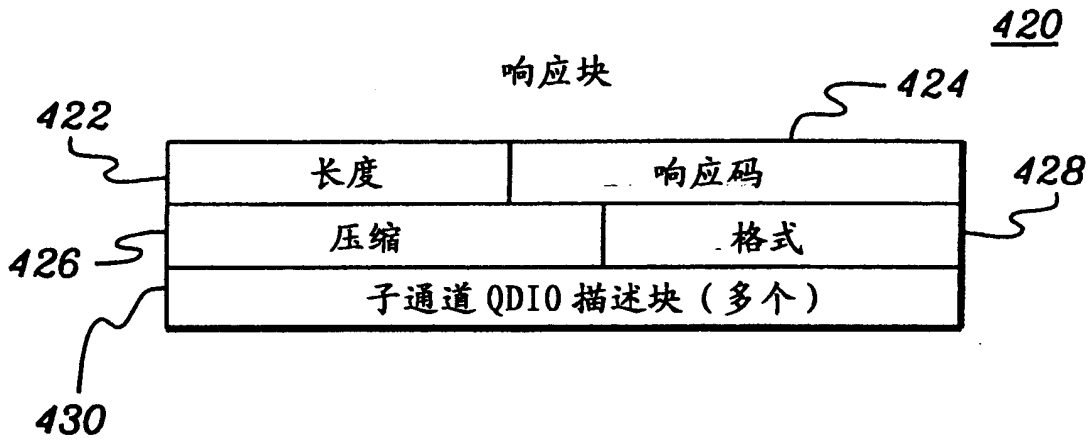


图 4B

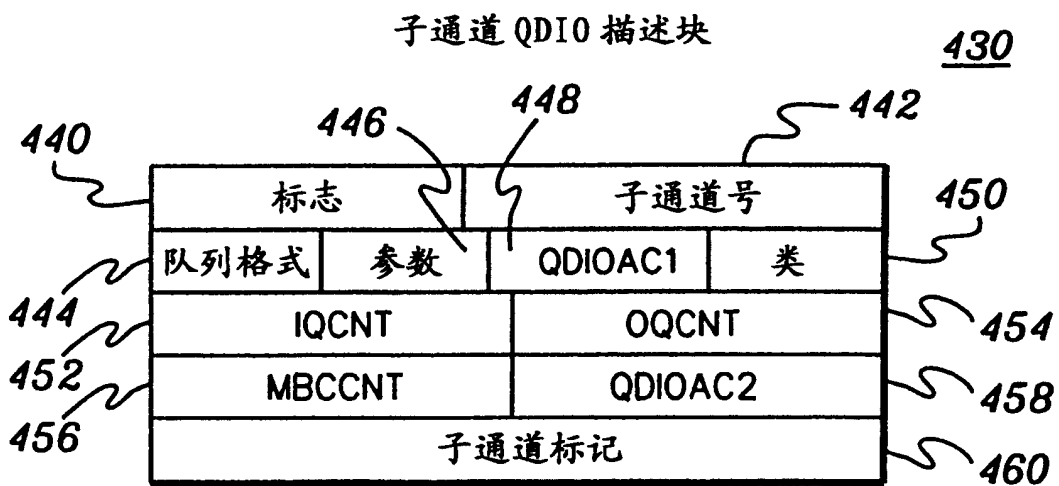


图 4C

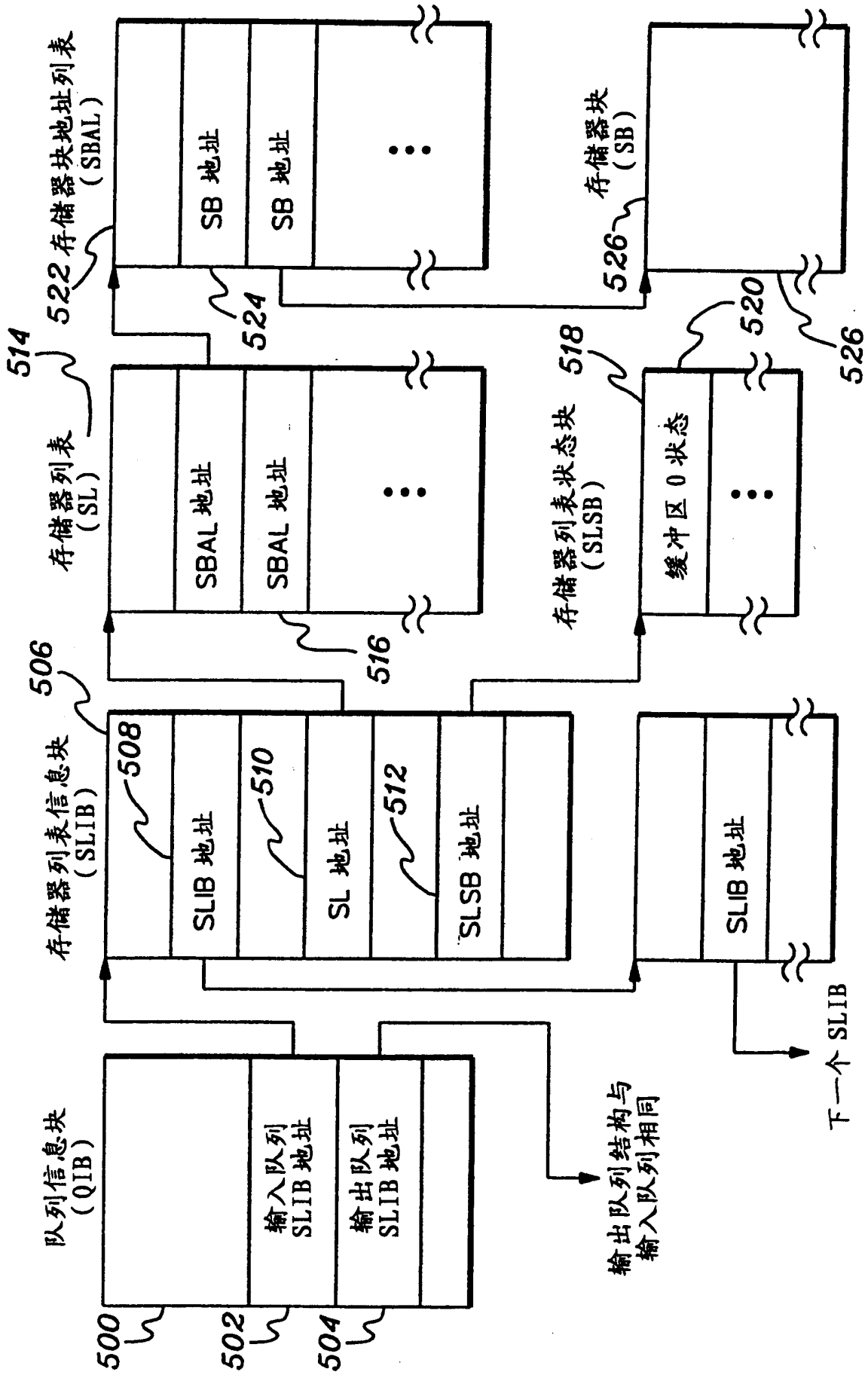


图 5

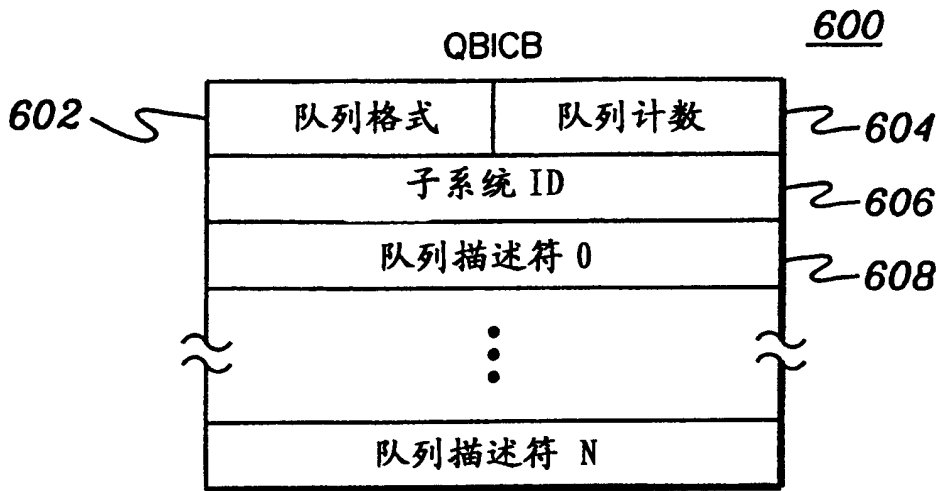


图 6A

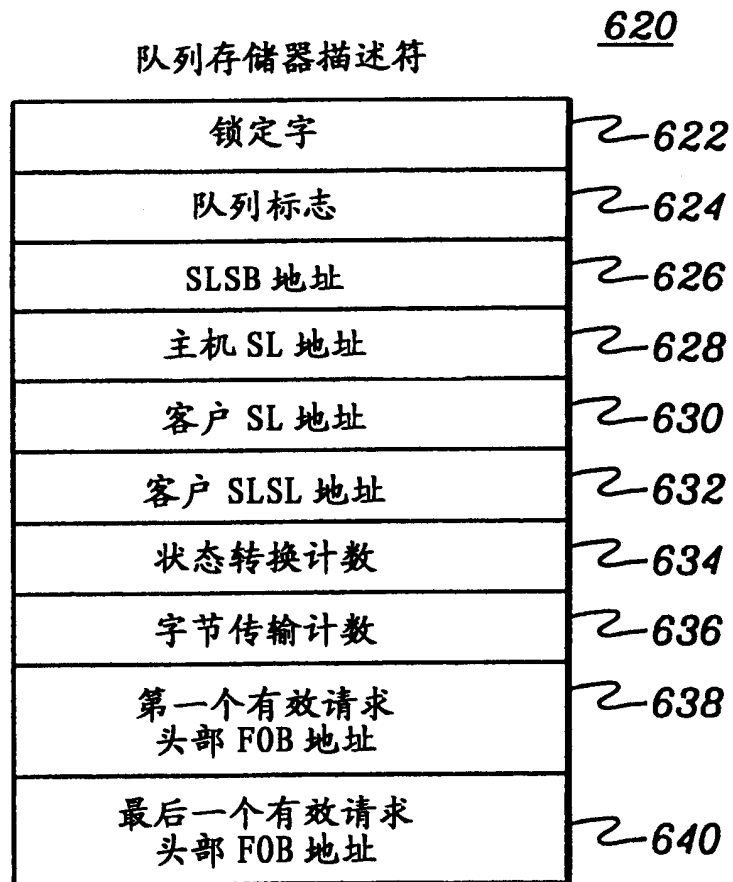


图 6B

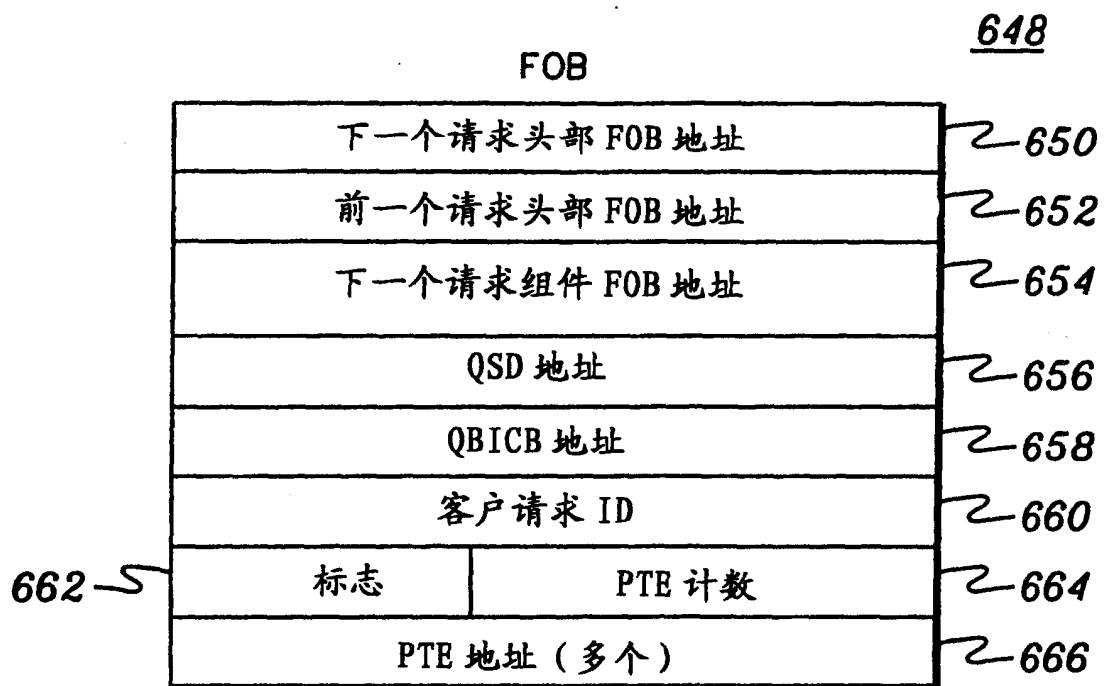


图 6C

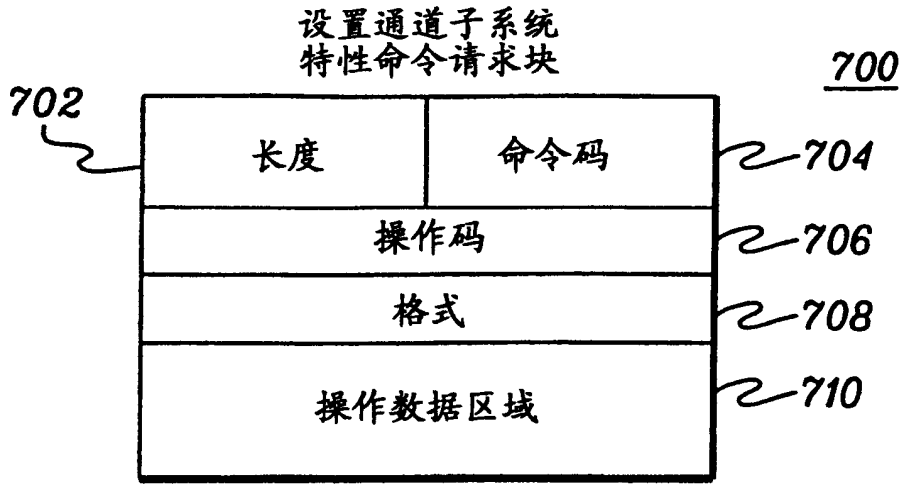


图 7A

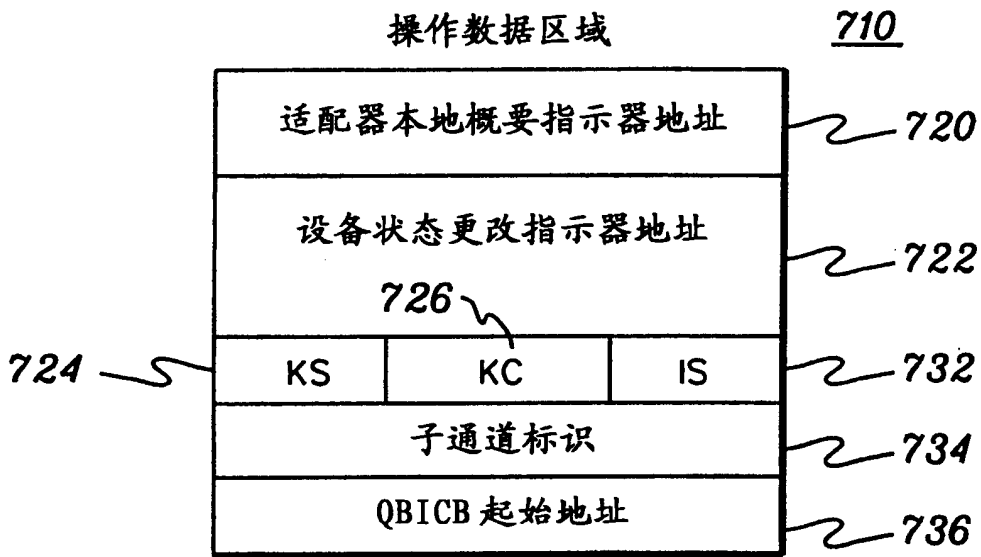


图 7B

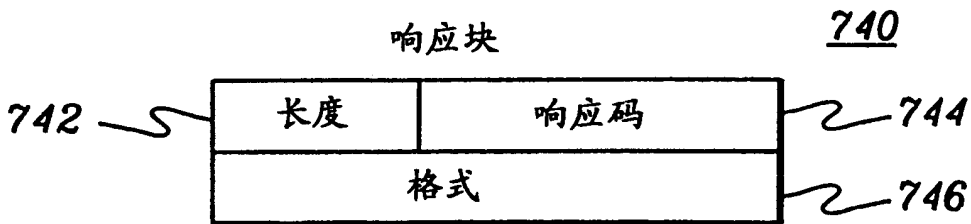


图 7C

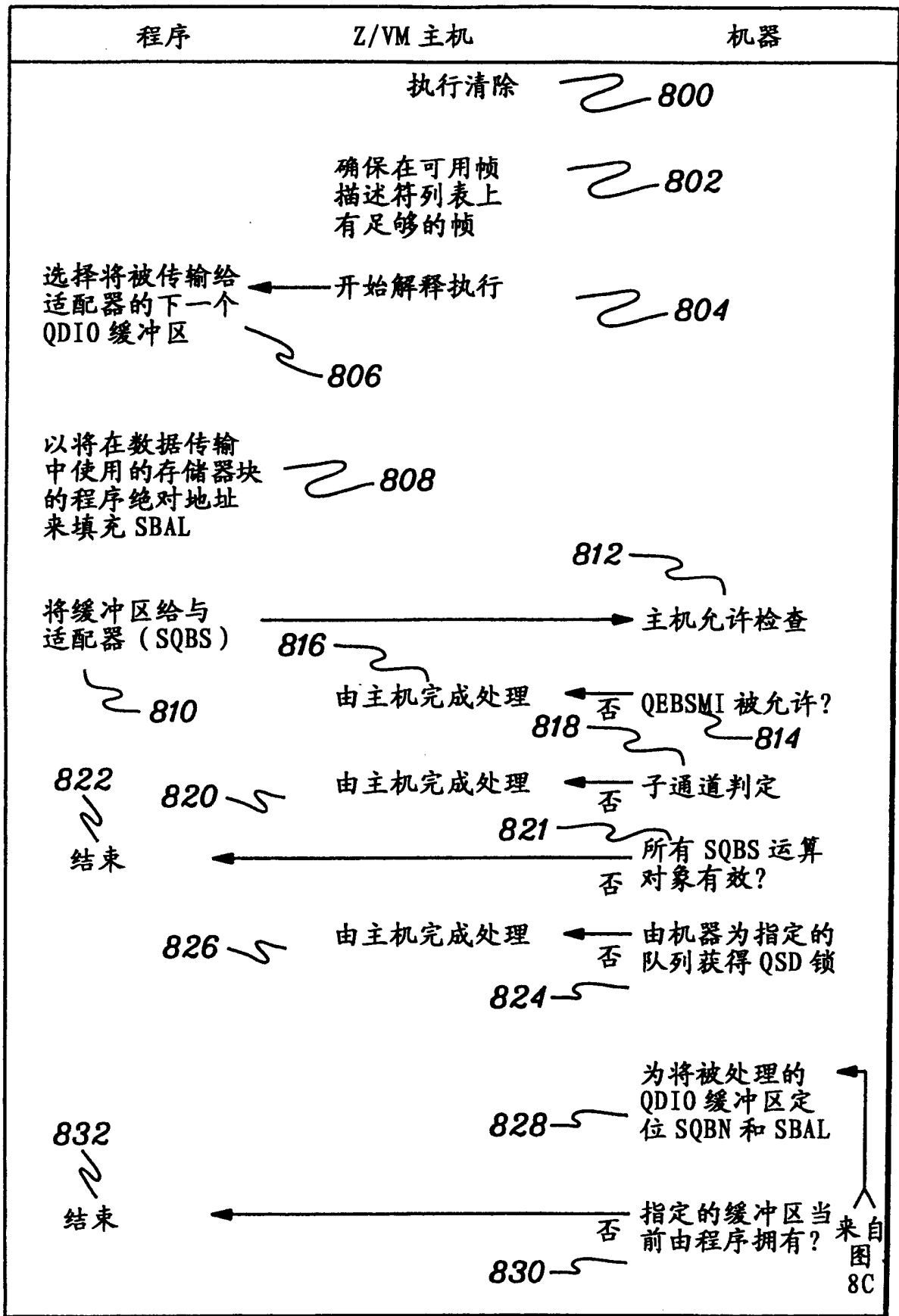


图 8A

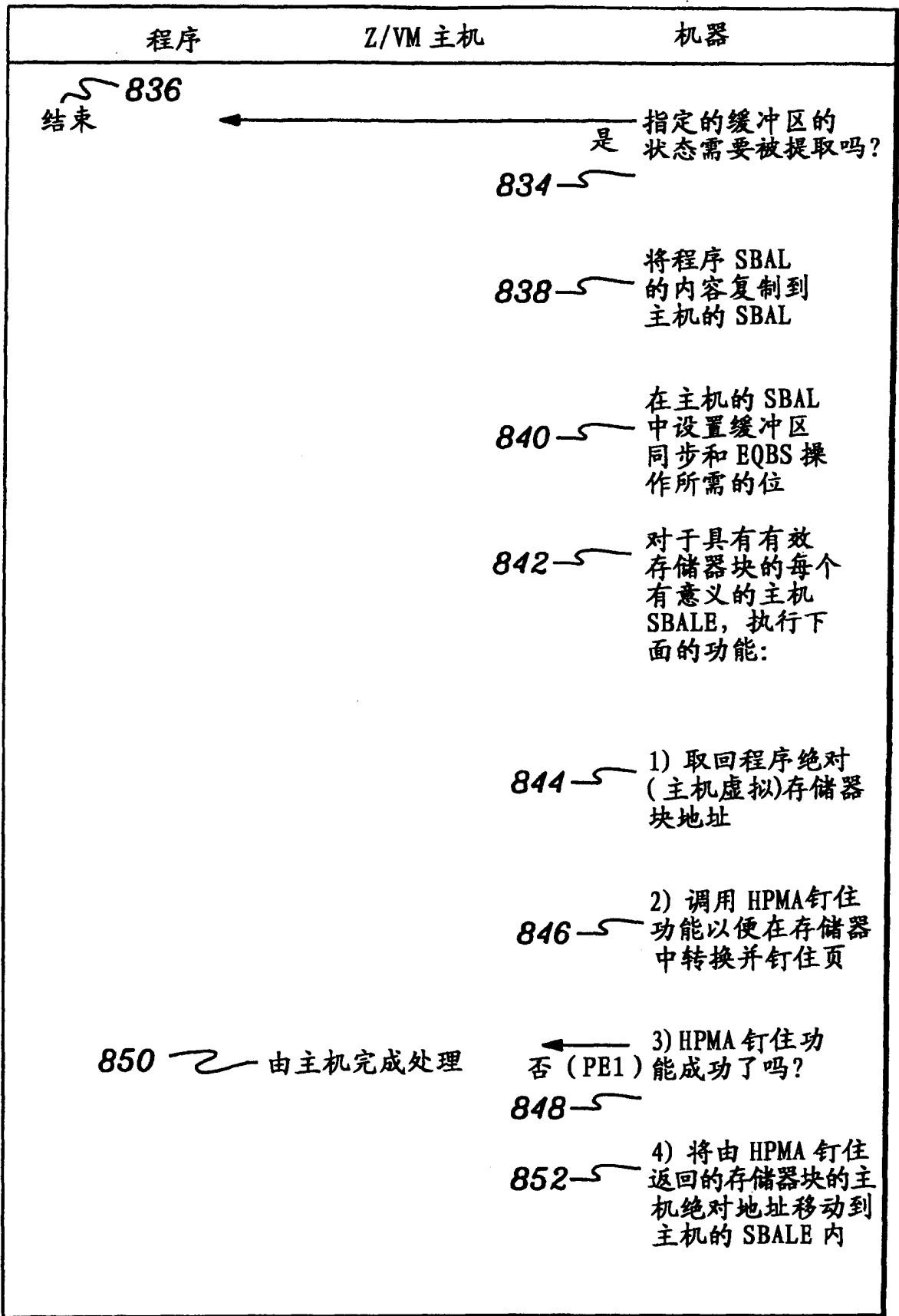


图 8B

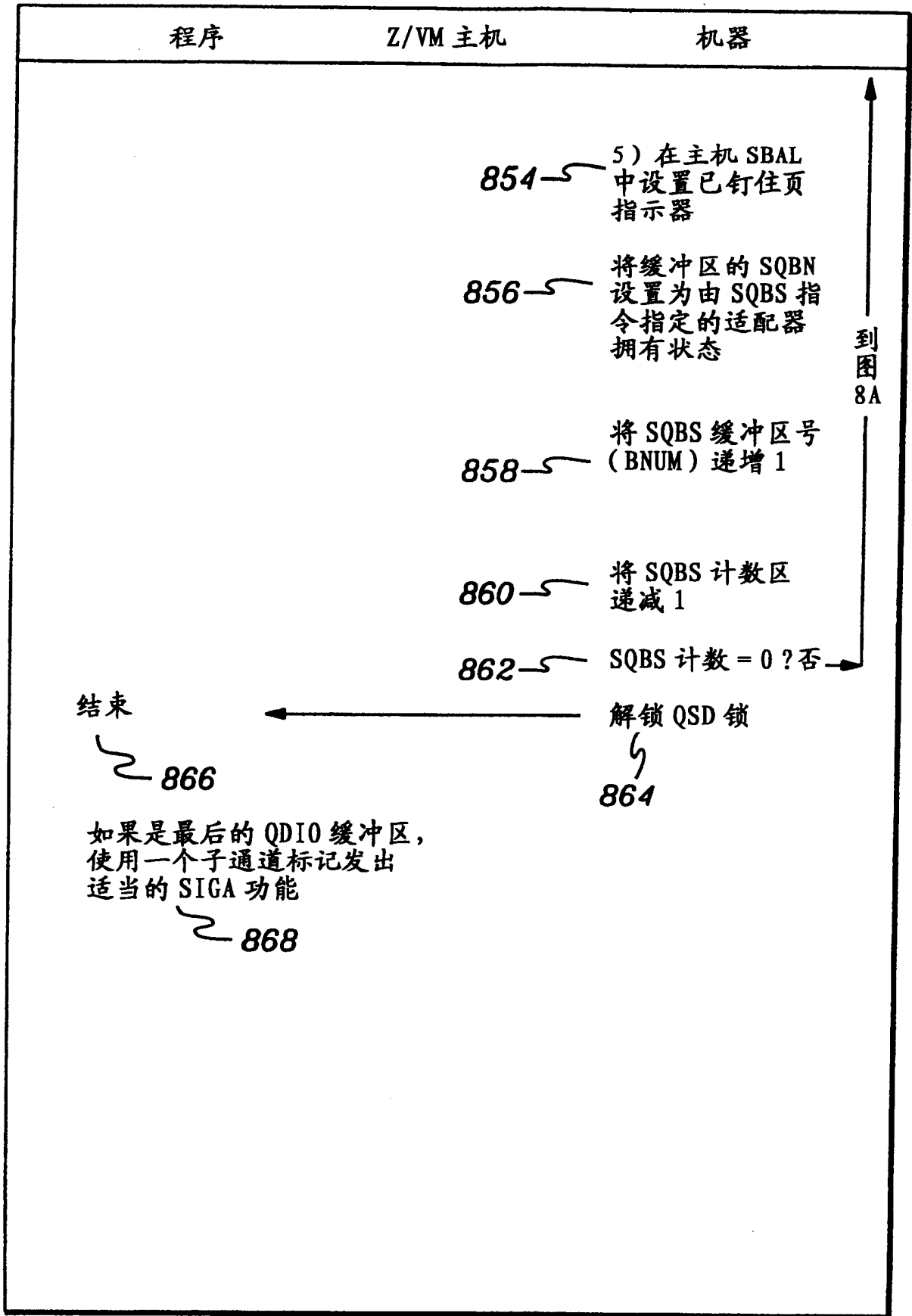


图 8C

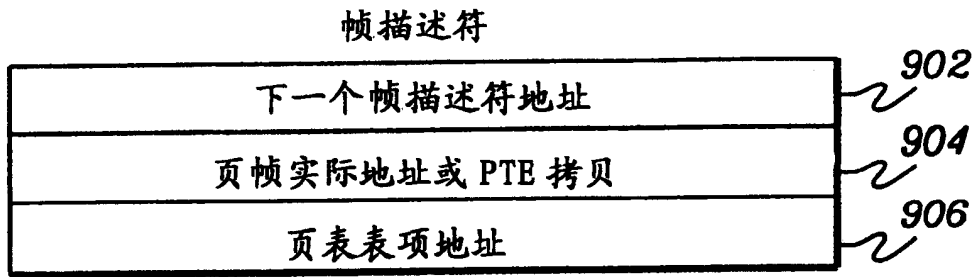


图 9

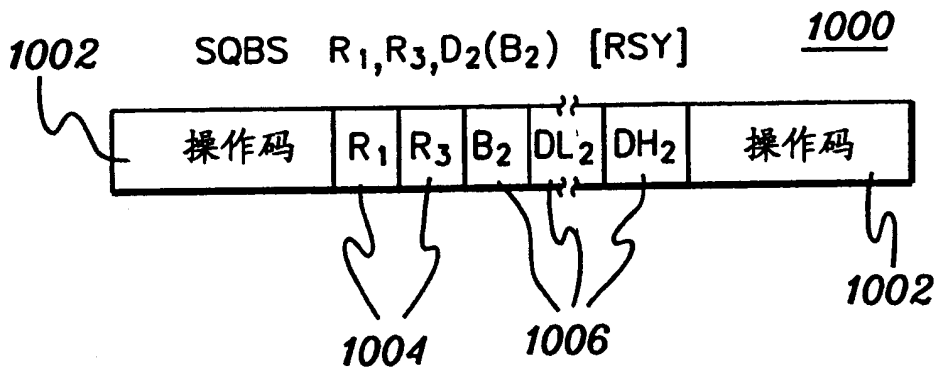


图 10

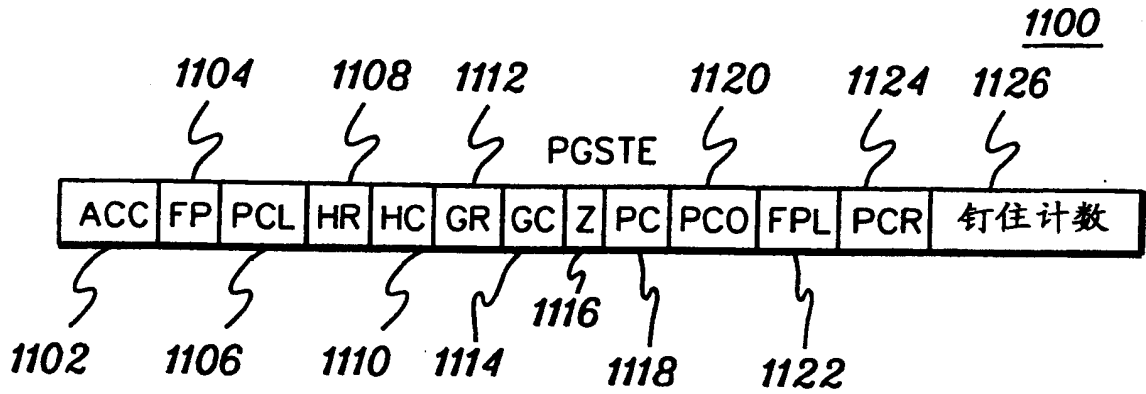


图 11A

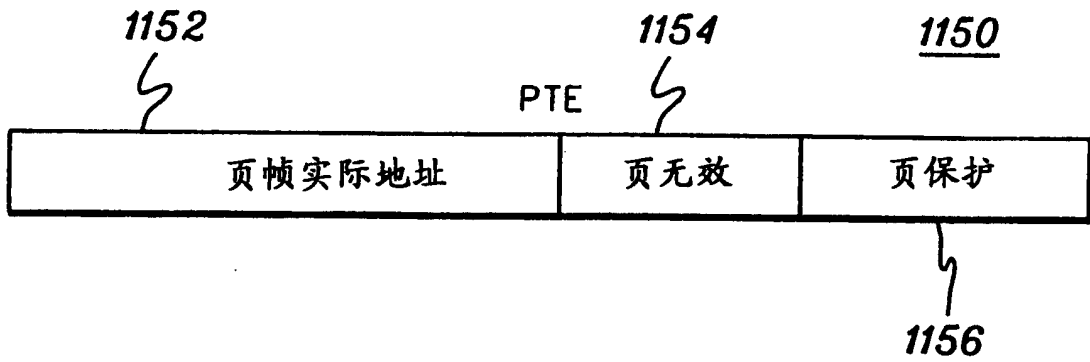


图 11B

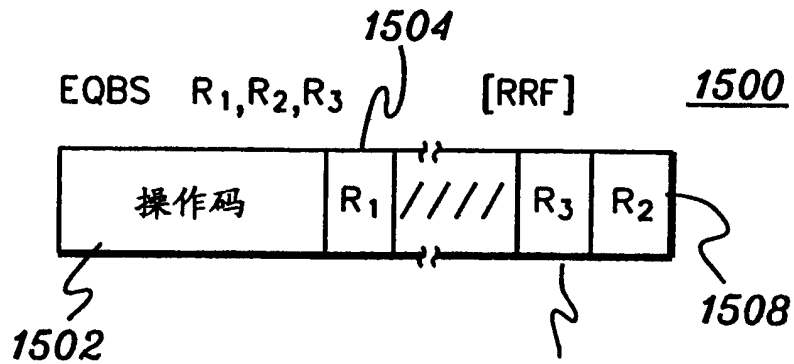


图 15

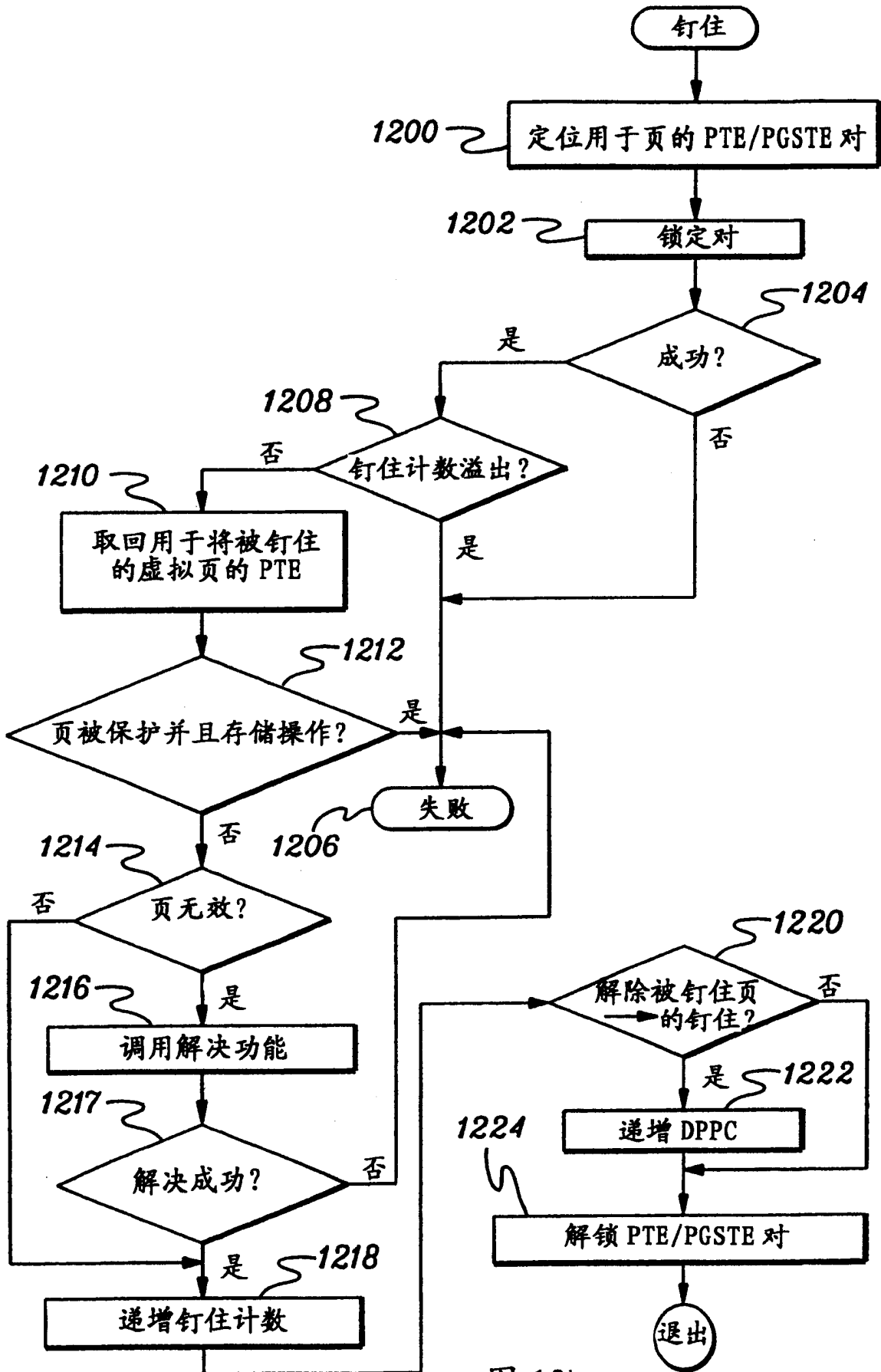


图 12

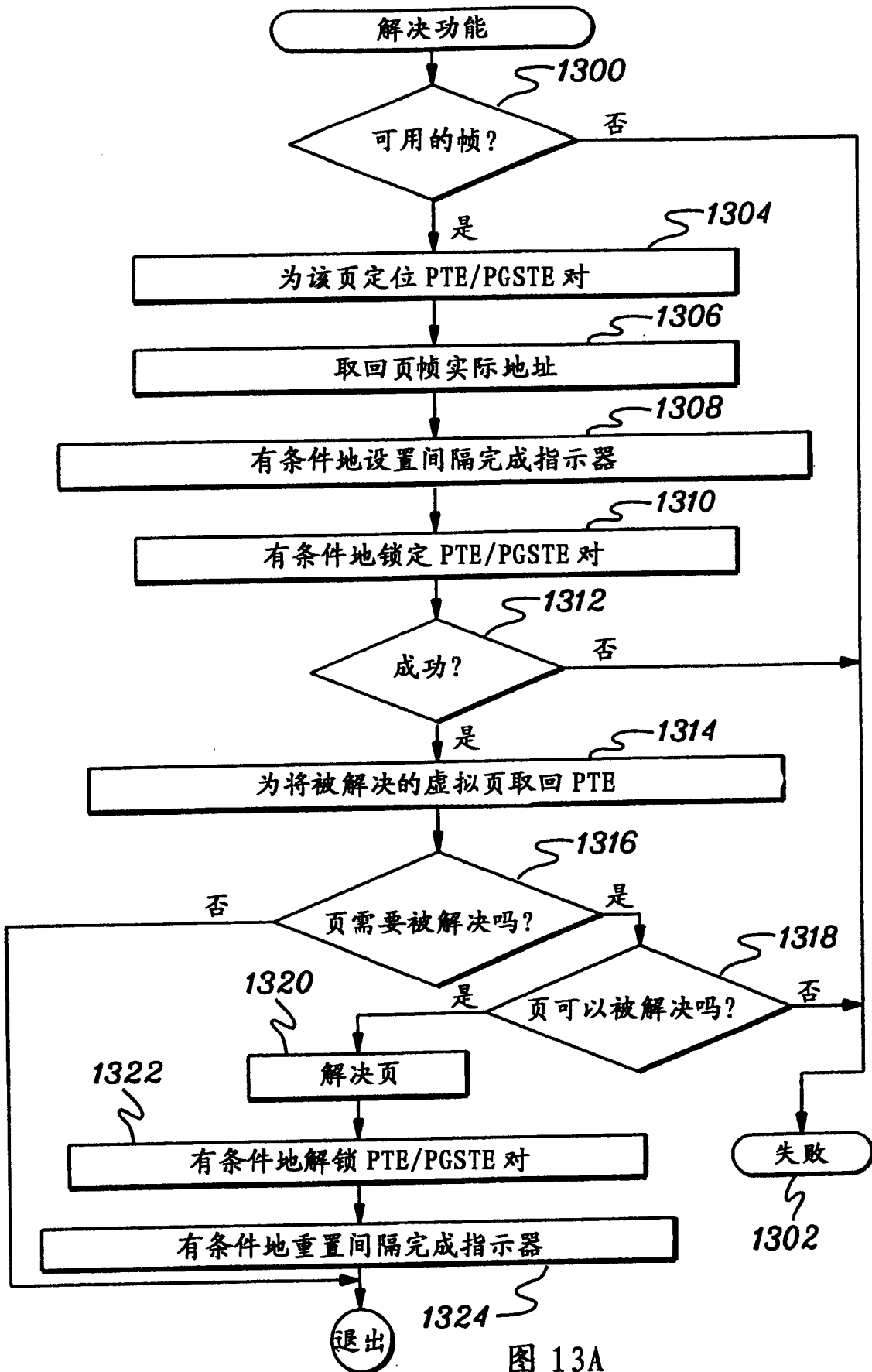


图 13A

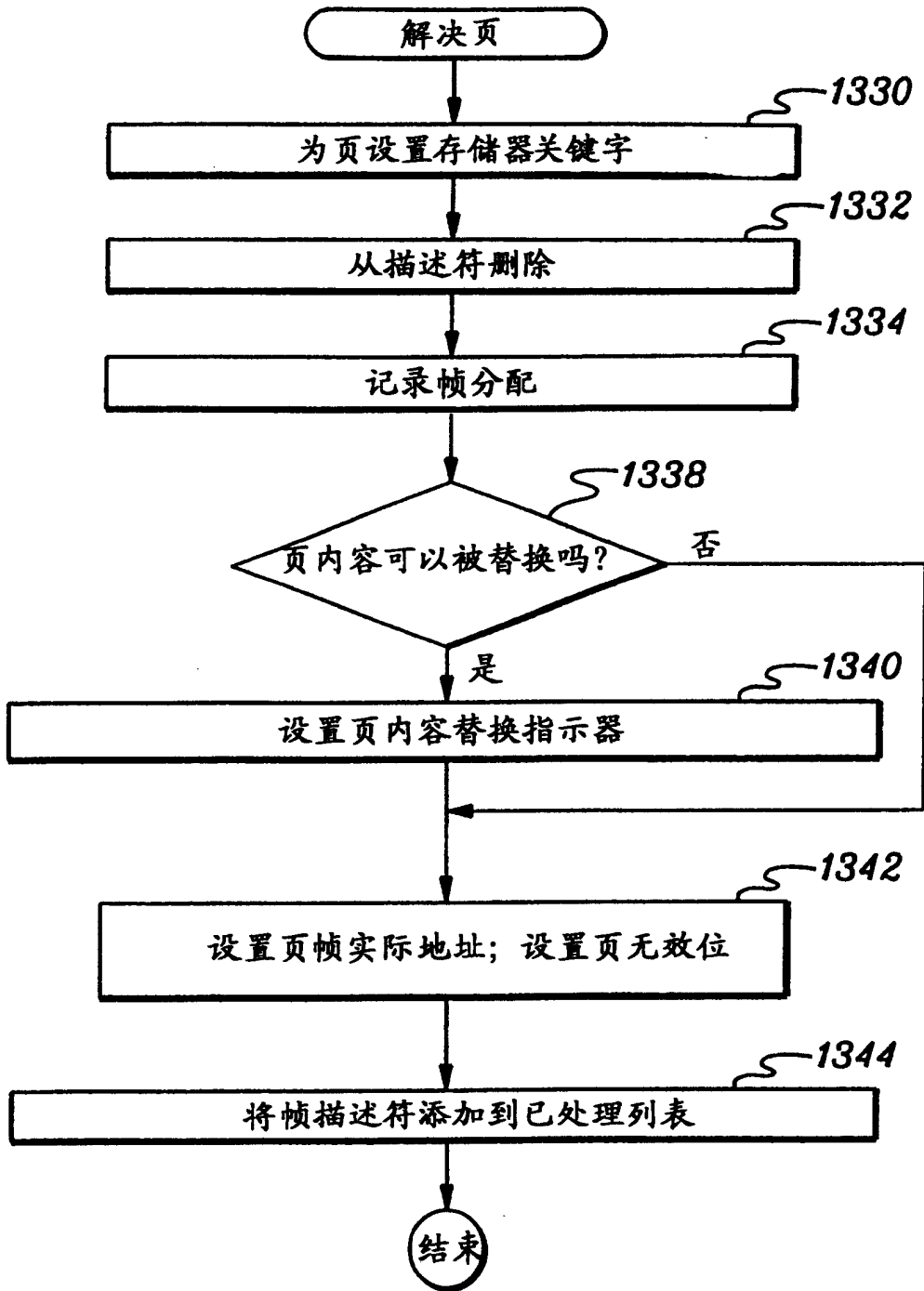


图 13B

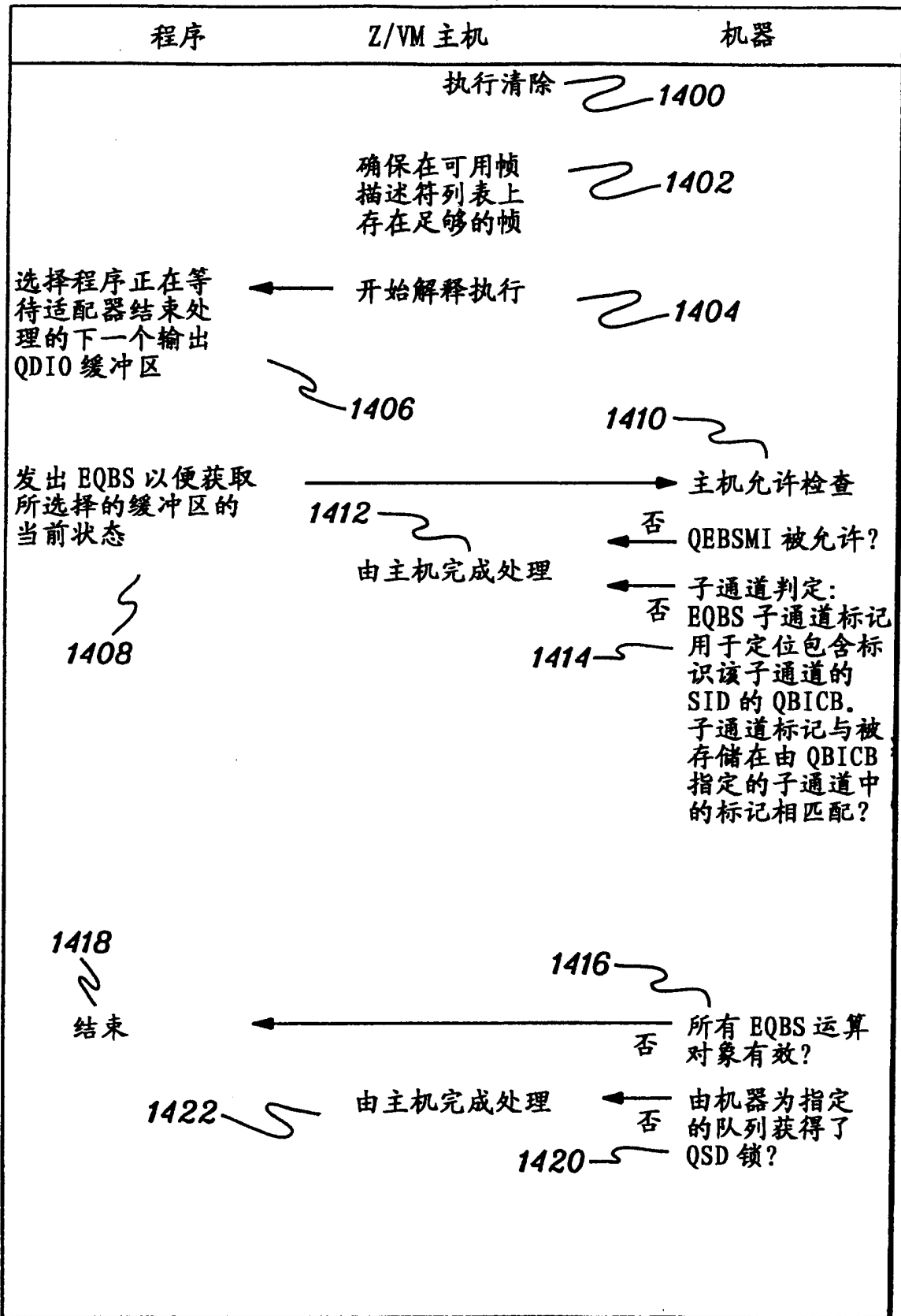


图 14A

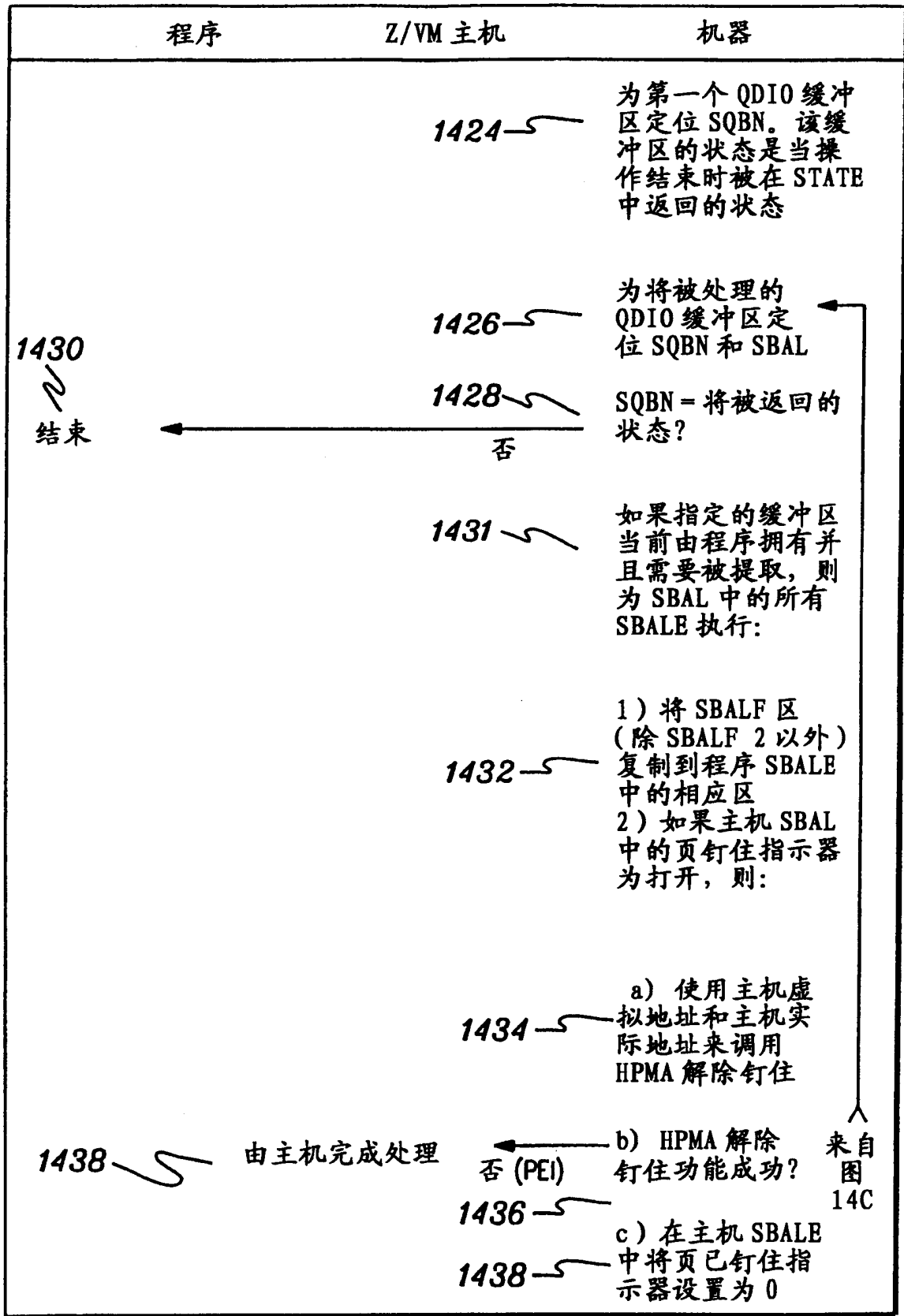


图 14B

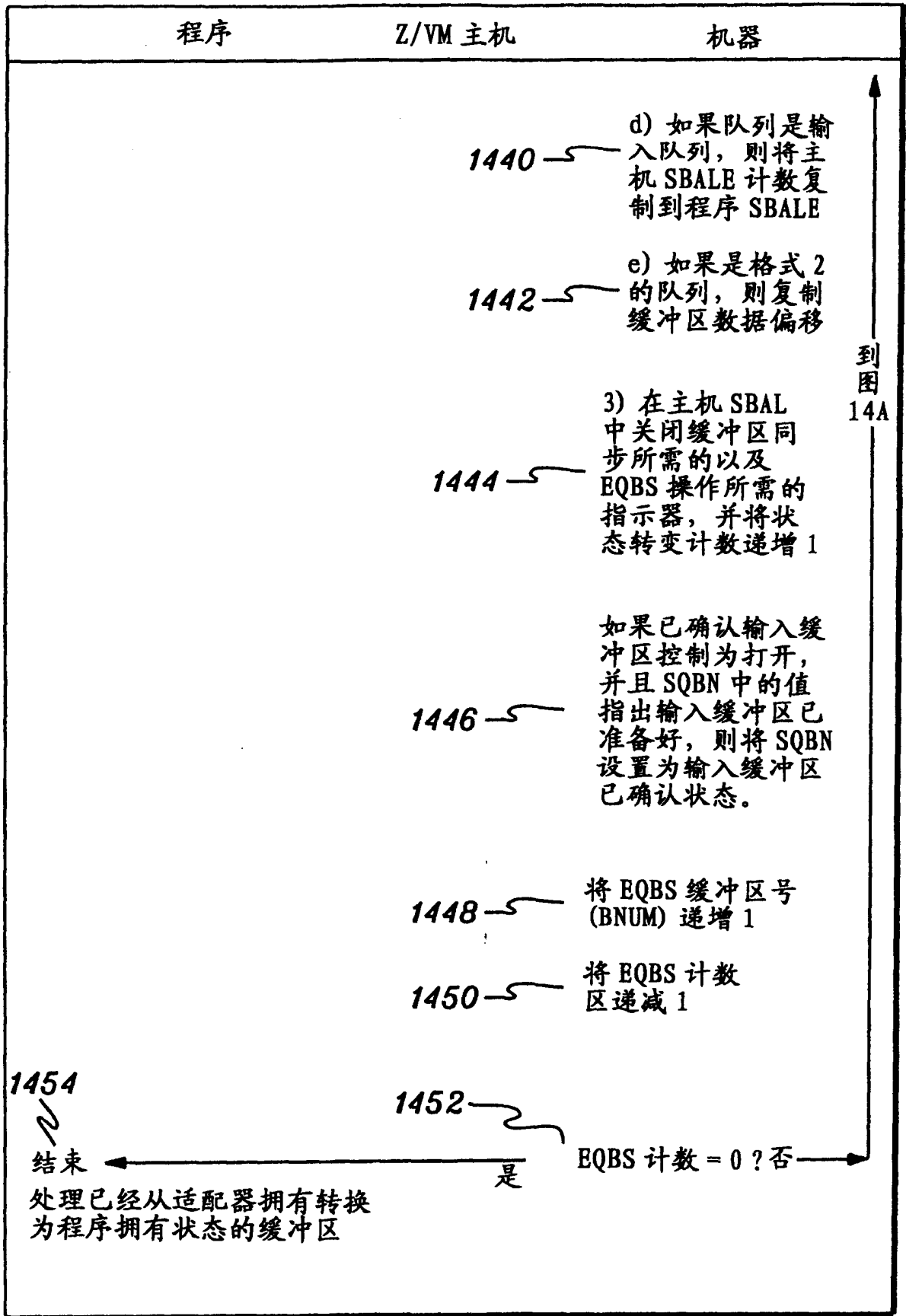


图 14C

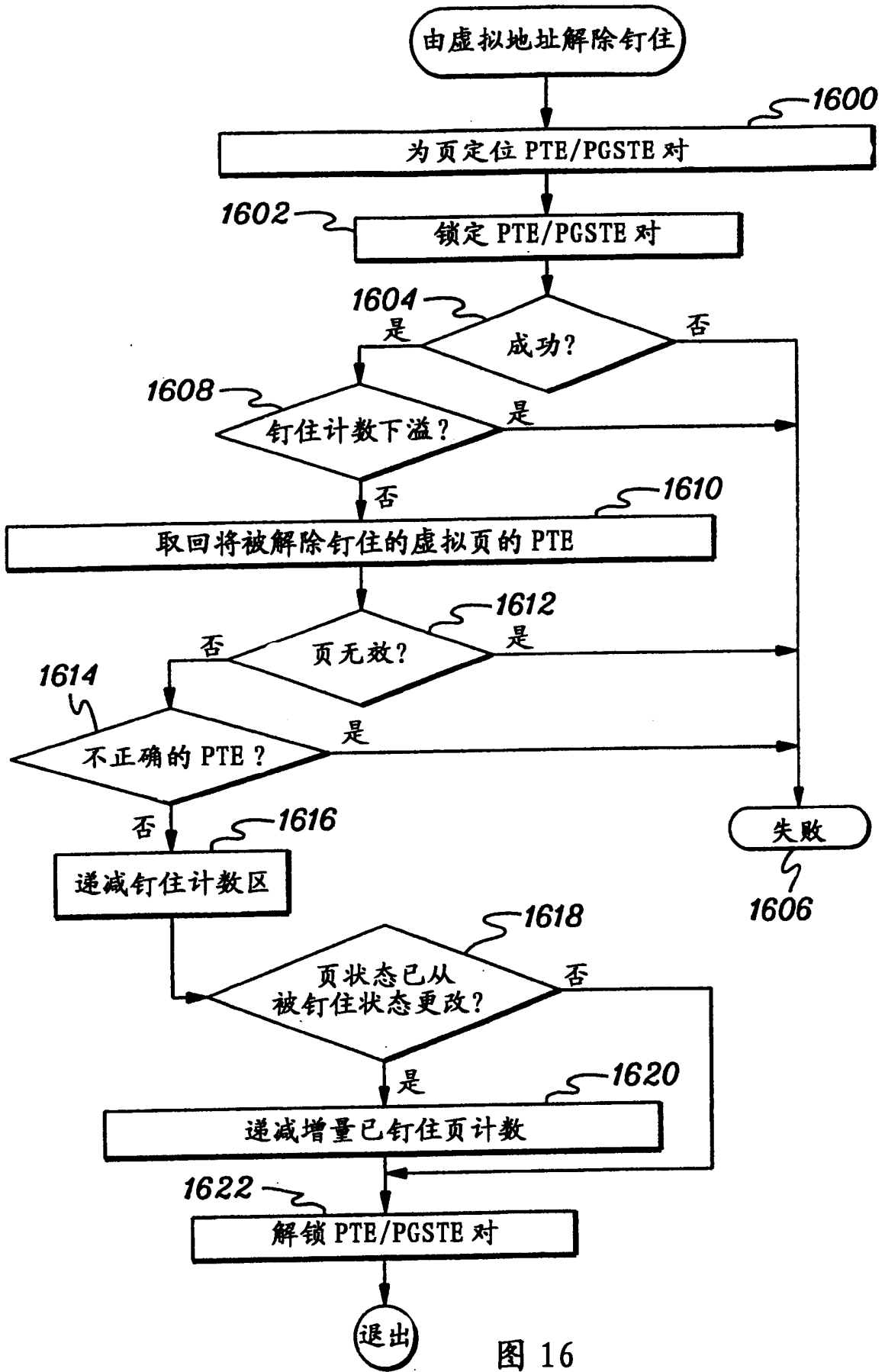


图 16

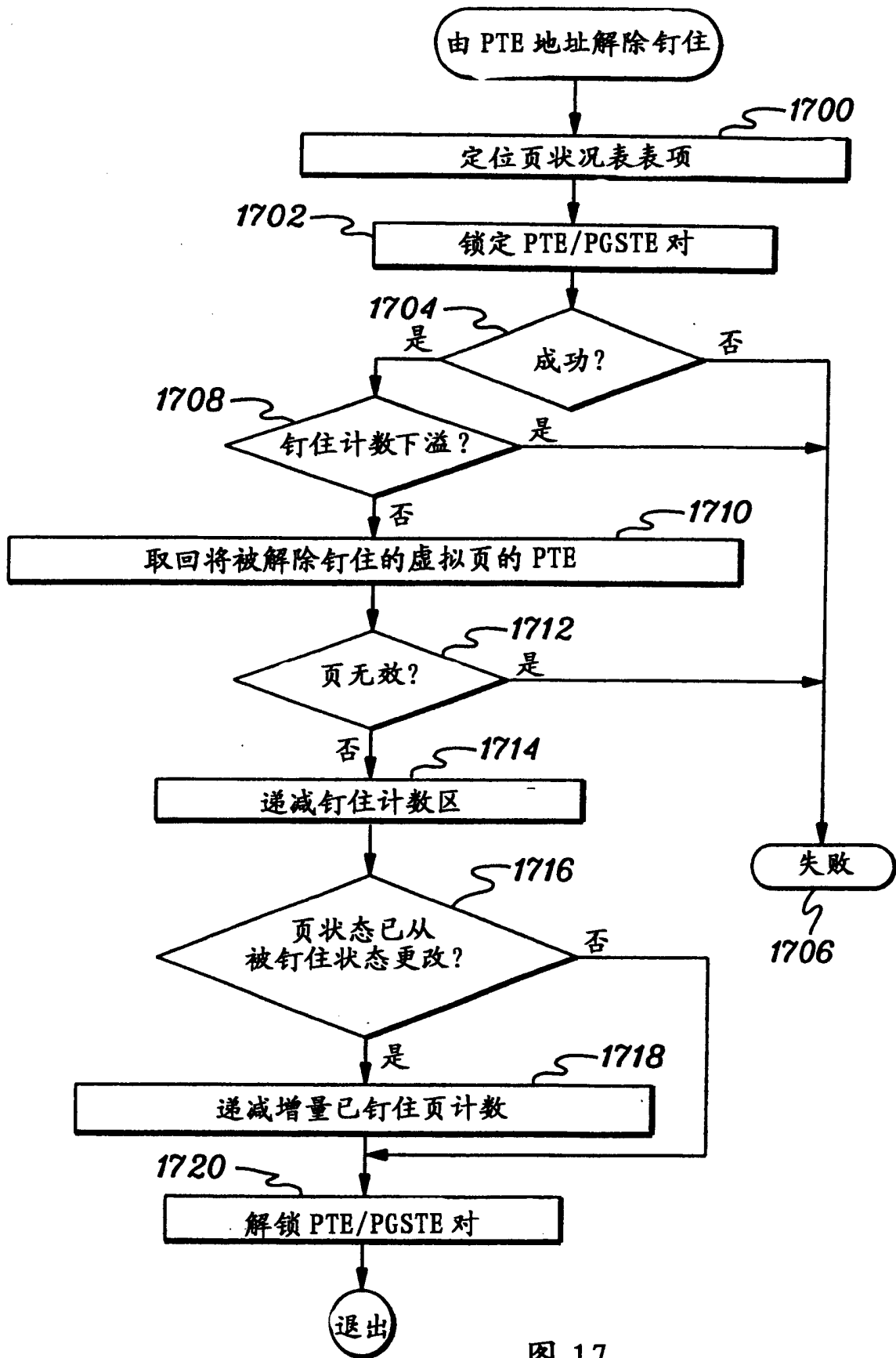


图 17

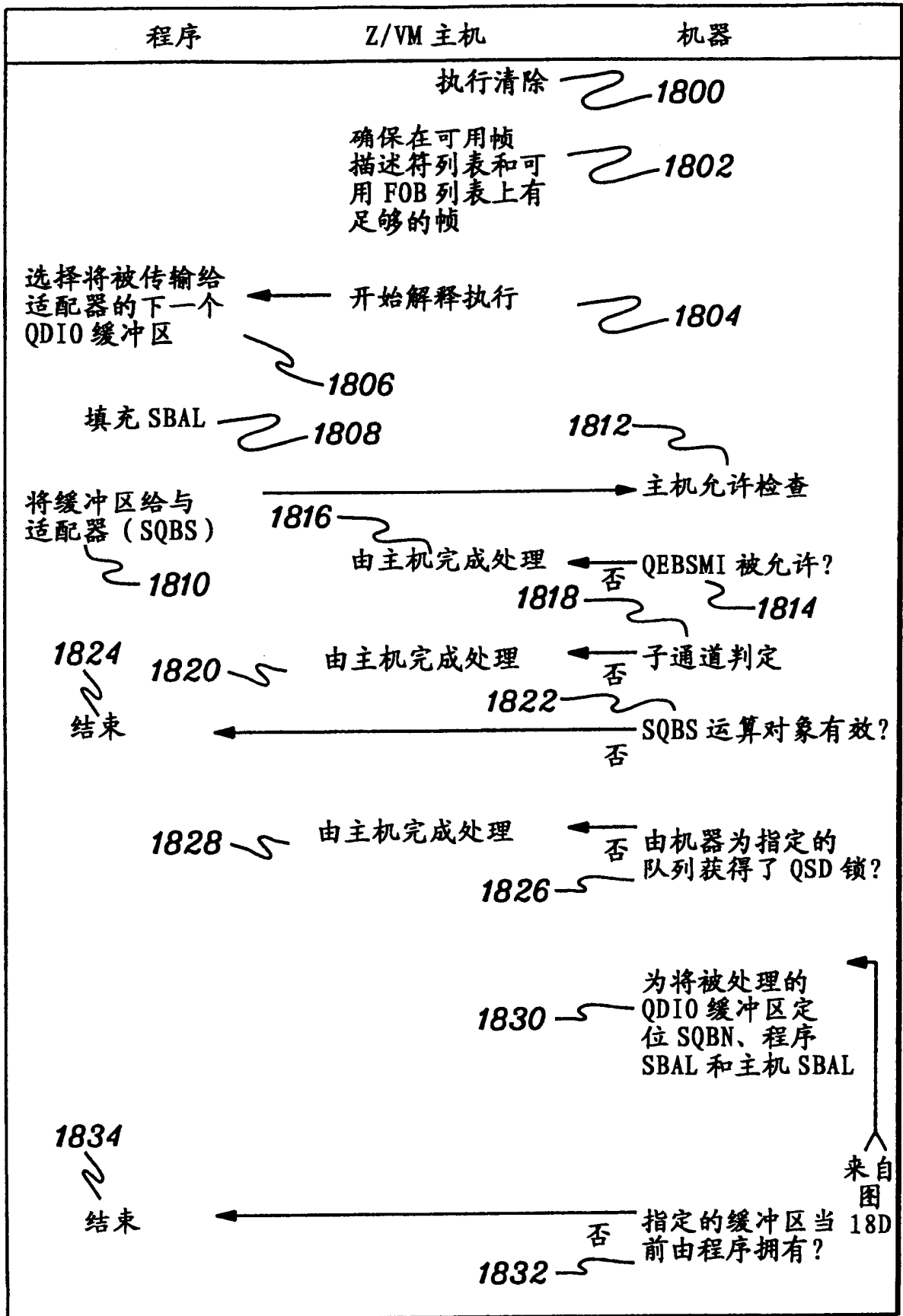


图 18A

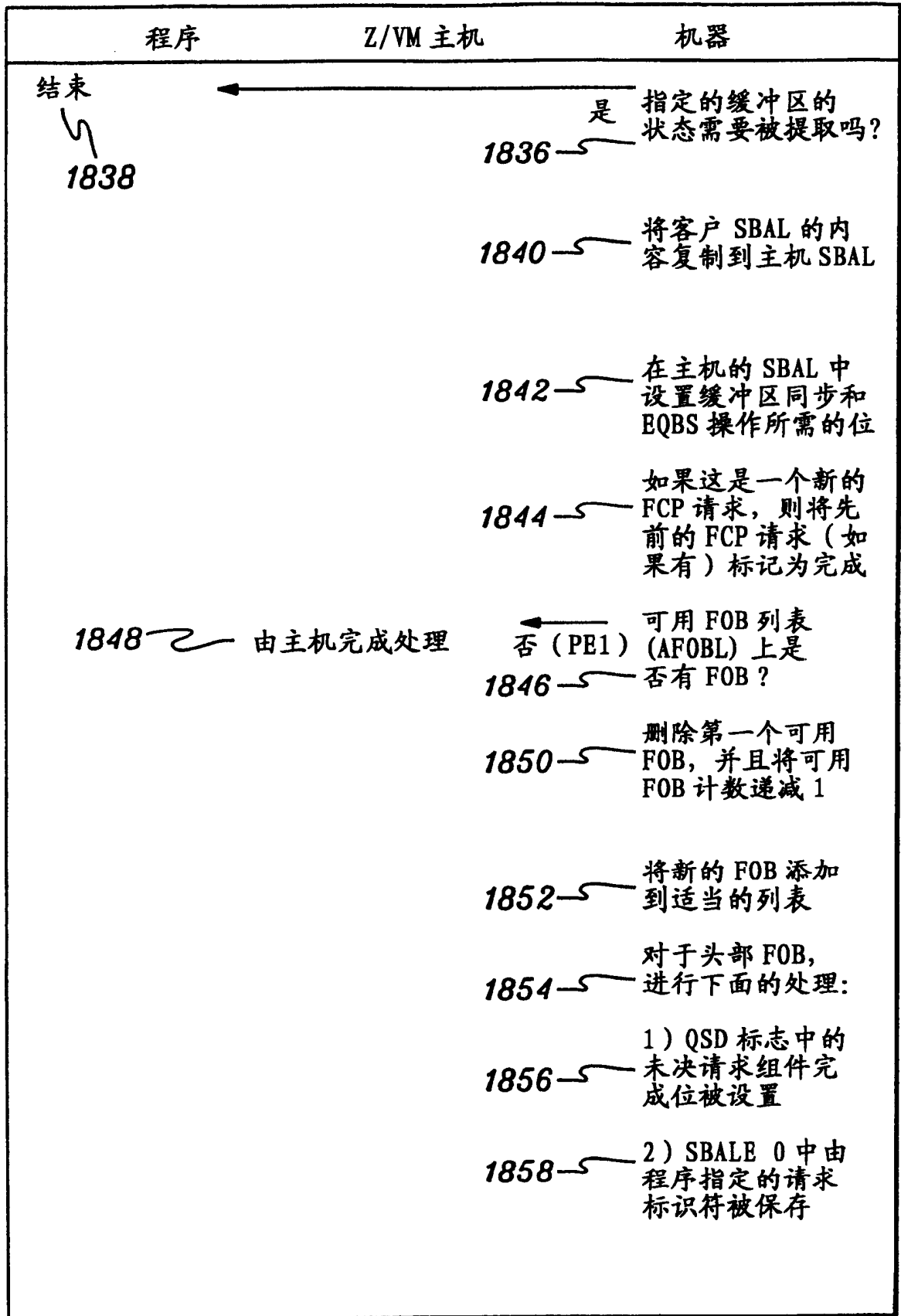


图 18B

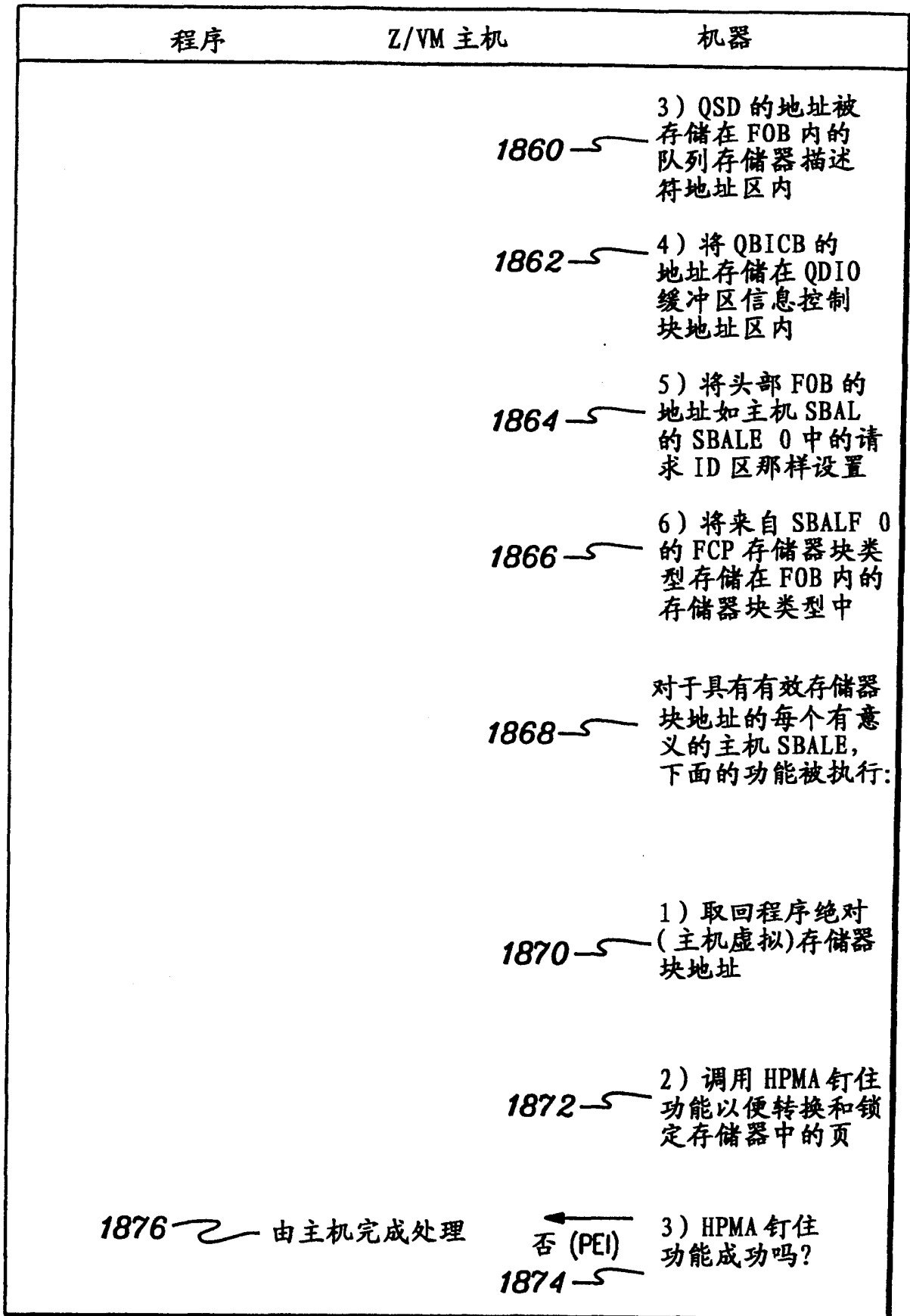


图 18C

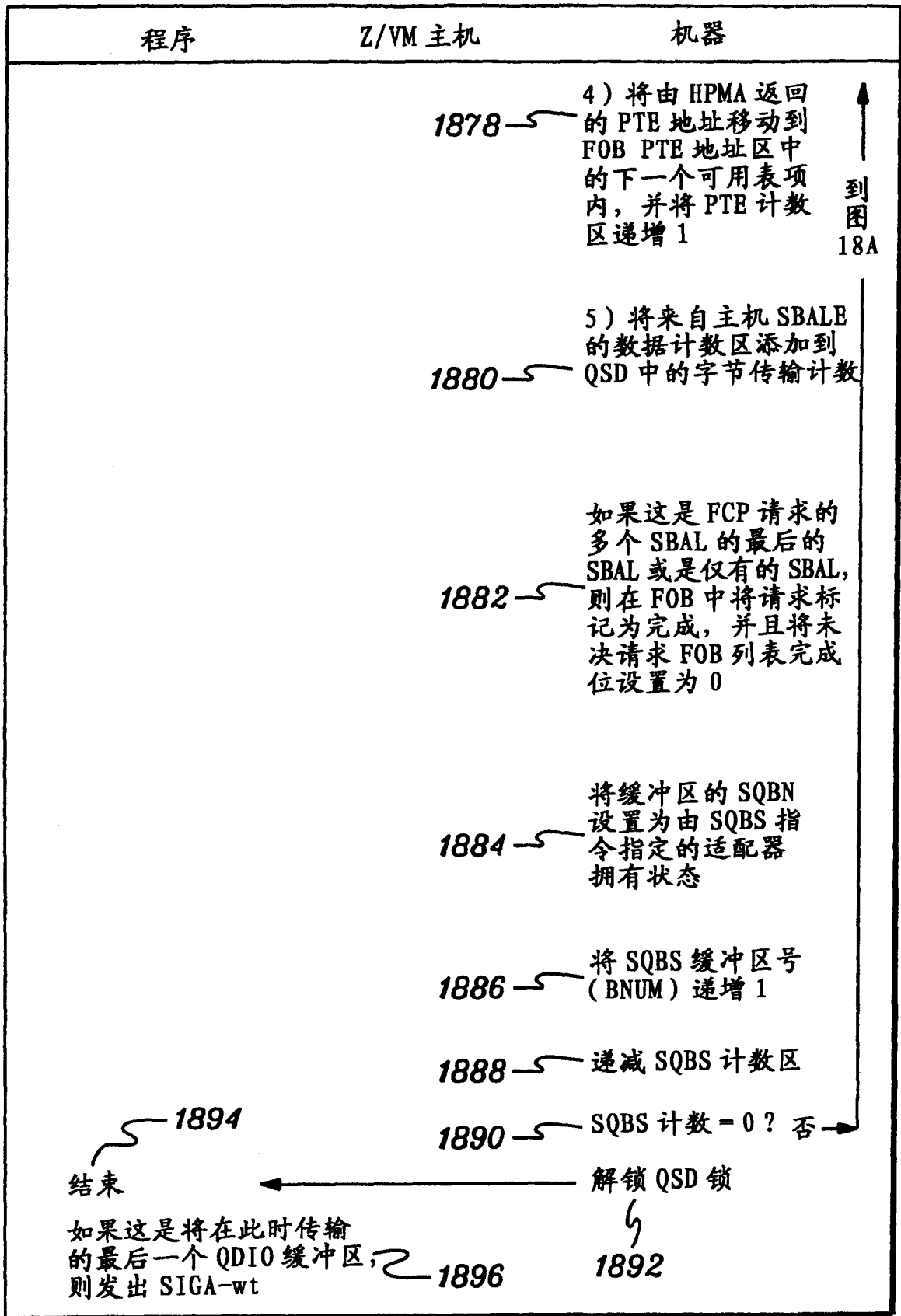


图 18D

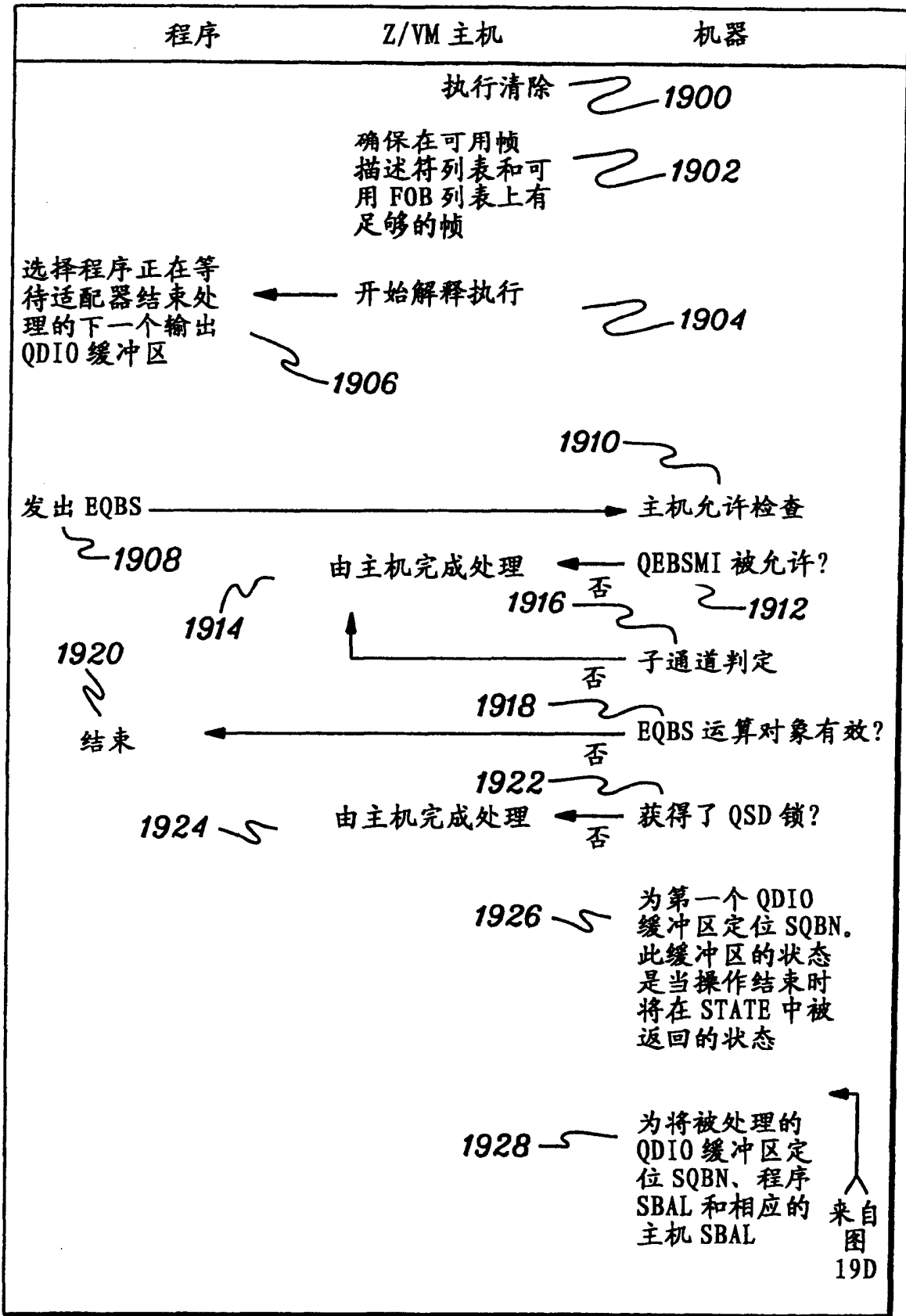


图 19A

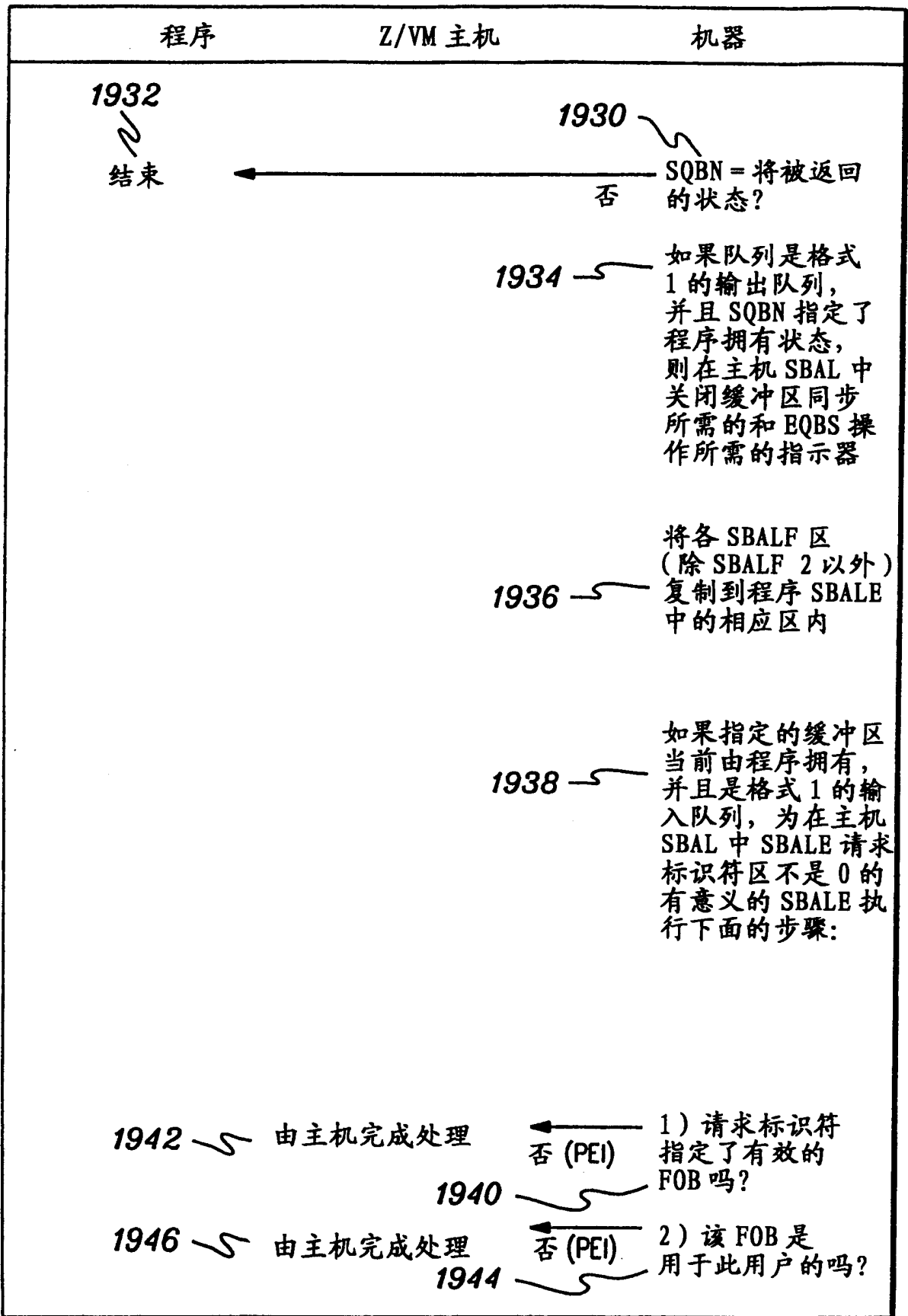


图 19B

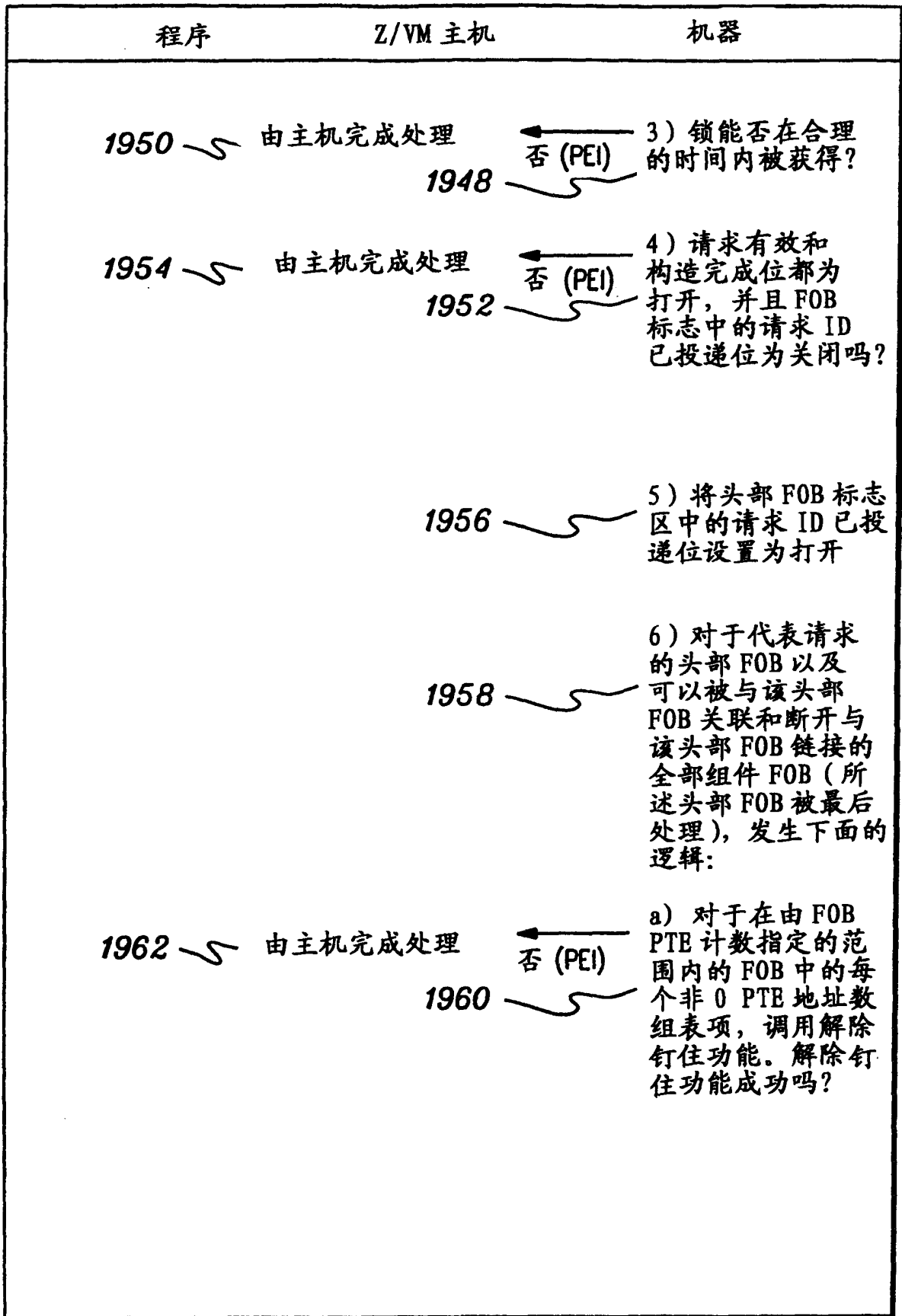


图 19C

程序	Z/VM 主机	机器
		b) 如果这是一个组件 FOB, 将其从与所述头部 FOB 解除锚定的请求组件 FOB 列表删除
	1964	
		c) 如果这是一个头部 FOB, 则执行操作
	1966	
		d) 清除 FOB 的预先指定的字, 将其插入可用 FOB 列表 (AFOBL) 的开始处, 并且将可用 FOB 计数递增 1.
	1968	
		7) 释放输出队列 QSD 锁
	1970	
		8) 将缓冲区同步所需的以及 EQBS 操作所需的指示器在 SBAL 中关闭, 并且将状态转换计数递增 1.
	1972	
		如果已确认的输入缓冲区控制为打开, 并且 SQBN 中的值指出输入缓冲区准备好状态, 则将 SQBN 设置为输入缓冲区已确认状态。
	1974	
		递增 EQBS 缓冲区号 (BNUM)
	1976	
		递减 EQBS 计数区
	1978	

图 19D

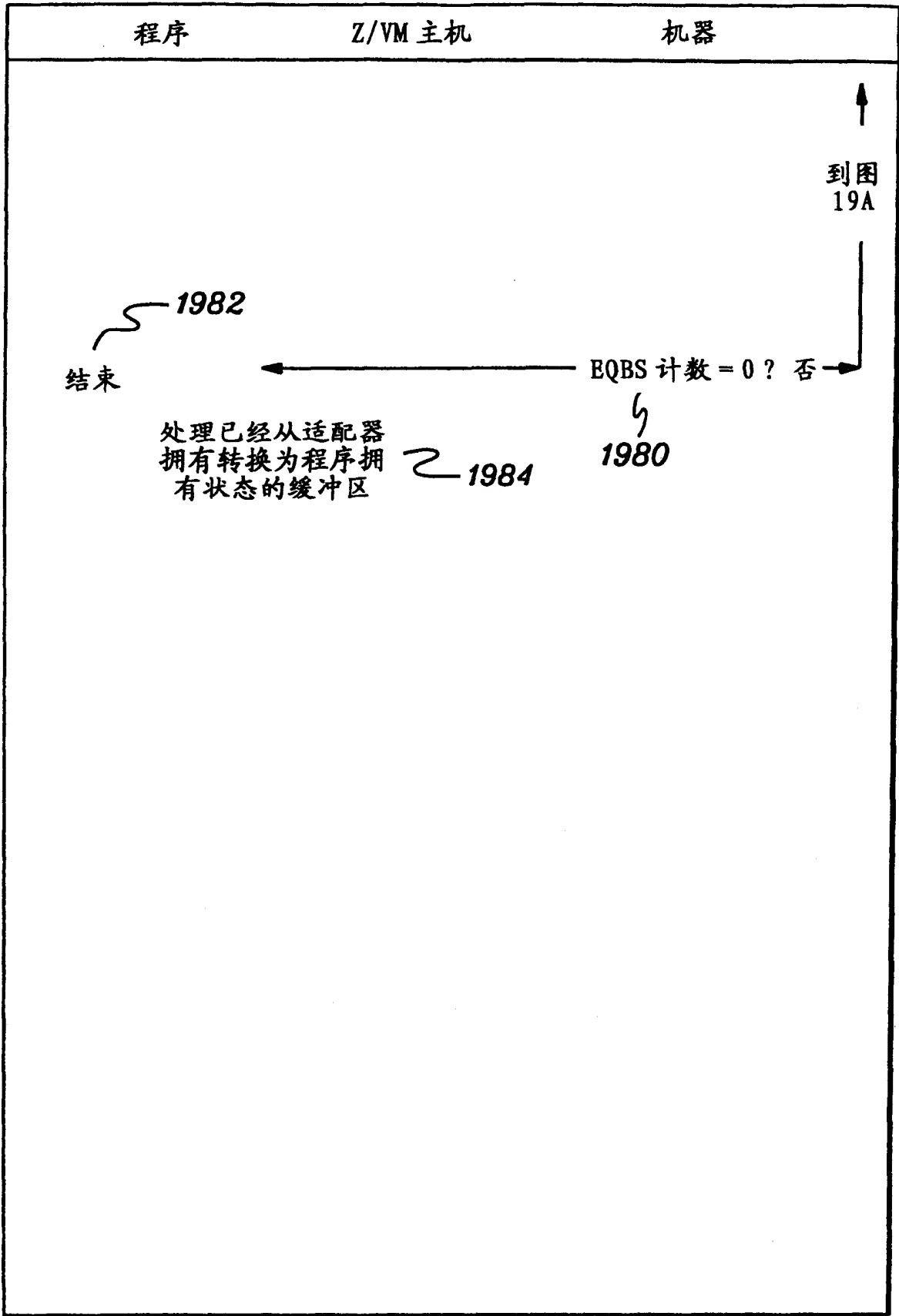


图 19E