Microsoft

微软 Azure的物联网参考架构

V版为 2.1

2018年9月26日

你想提供反馈或者你有一个想法或建议，基于与物联网的经验？我们很想听到您的声音！请发邮件 AzureIoTRefArcVoice@microsoft.com

内容

这里的信息是只供参考之用，并代表Microsoft Corporation的当前视图作为本公布之日起的。由于Microsoft必须适应不断变化的市场情况作出反应，它不应该被解释为对微软的承诺，并且Microsoft不能保证此演示文稿日期之后提供的任何信息的准确性。MICROSOFT不做任何明示，暗示或法律，以在此演示文稿中的信息。

1。 概观

连接的传感器，设备和智能操作可以改变企业，使事情的微软Azure联网（IoT）服务新的增长机会。

该文件的目的是提供推荐的体系结构和实现技术选择的概述 怎么样 打造物联网Azure的解决方案。这种架构描述术语，技术原理，常见的配置环境，和Azure的物联网服务，物理设备组成，与智能边缘设备。这份文件的主要目标是架构师，系统设计师，开发人员和谁正在建设物联网**解决方案等物联网技术决策者。构建，运行和维护物联网解决方案是一个显著的努力，我们建议客户评估物联网中心**[1]，**微软的物联网SaaS产品，当确定如何构建一个解决方案。看到** [附录5.3的SaaS，PaaS和IaaS的指导](#) **对使用物联网中心的指导。**

物联网应用程序可以被描述为 物联网（ 或装置），发送数据或用于生成事件 见解， 其用于生成 操作 以帮助提高业务或流程。一个例子是发动机（一物），发送该用于评估发动机是否正在执行如预期（洞察），其用于主动优先维护计划用于发动机（动作）的压力和温度的数据。本文将**重点介绍 怎么样 承担的业务洞察力，我们通过从资产收集数据，找到行动：建立一个物联网解决方案，但是它是认识的架构的最终目标是非常**重要的。

该文件包含四个部分：1）一个 概述 - 含有的IoT整体解决方案建议的体系结构（分成子系统），简要介绍的IoT应用子系统，每一个子系统默认技术建议，和的交叉问题为的IoT应用的讨论，2） 基本概念和原则 - 的概念和原理的核心构建可伸缩的IoT应用在本节中描述，3） 子系统的详细信息 - 每个子系统的子部分专用于描述用于执行子系统的职责和技术的替代方案，和4） 解决方案设计注意事项 - 描述解决方案和垂直行业架构的实现考虑的部分。

这个文件是一个活的文件，将被更新，云计算和技术环境发展。该文件将跟踪技术的变化，并为物联网Azure的解决方案架构和技术的最佳实践最新建议提供了。

每个组织都有独特的技能和经验，每一个物联网应用具有独特的需求和注意事项。根据需要为各该文献中所建议的参考架构和技术选择应该进行修改。

使用统一的标准产生每个子系统的技术建议。有些标准是所有子系统和替代技术常见; 例如，安全，简便，性能，规模和成本是关键不管子系统或技术。然而，有些标准是唯一的一个特定的子系统; 例如，查询功能

[1] https://azure.microsoft.com/en-us/services/iot-central/

温暖的存储解决方案。用于评估技术建议标准中详细子系统部分中描述。

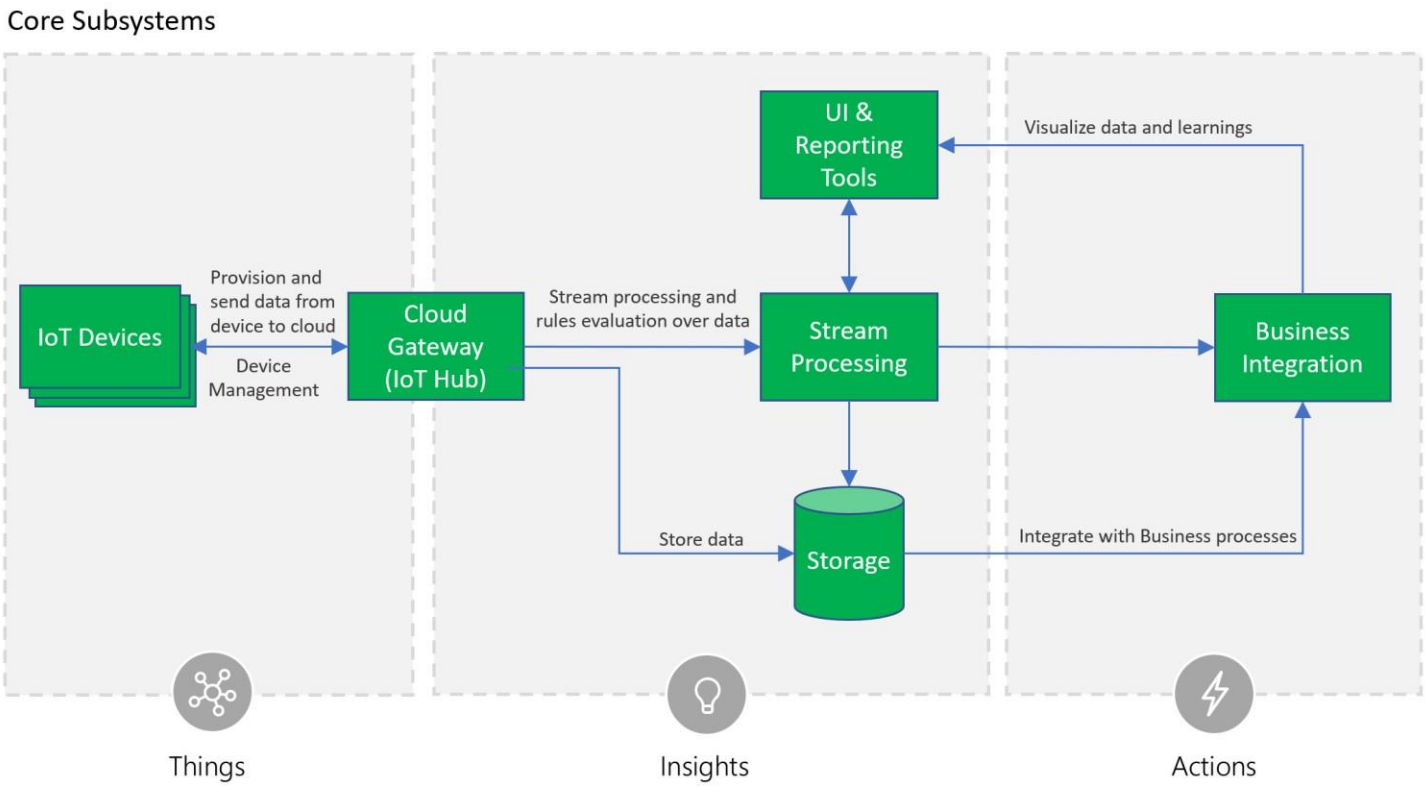
1.1 架构概述

我们推荐的物联网解决方案的架构是云本地人，微服务，以及无服务器为主。物联网解决方案的子系统应建为离散服务，这是独立部署，并能独立地扩展。这些属性使更大规模，在更新各个子系统更大的灵活性，并提供灵活选择合适的技术在每个子系统的基础。至关重要的是必须监控各个子系统，以及物联网应用作为一个整体的能力。我们建议子系统使用JSON在REST / HTTPS通信（因为它是人类可读）虽然二进制协议应当用于高性能的需求。该架构还支持混合云和边缘计算策略; 即一些数据处理，预计在前提下发生。我们推荐使用的Orchestrator（如

1.2 核心子系统

在芯IOT中应用由以下子系统组成：1）设备（和/或上边缘前提网关）具有安全地与云寄存器的能力，和连接选项，用于发送和与云接收数据，2）云网关服务，或 毅，以安全地接受数据，并提供设备管理功能，3）

流处理器 消耗数据，与整合 业务流程，并把数据转换成 存储，以及4）一个
用户界面 可视化遥测数据和促进设备管理。之后，这些子系统简要规定的技术建议描述。部分覆盖在深度这些子系统是在本文件的第3。



云网关提供了用于安全连接，遥测和事件摄取和设备管理（包括命令和控制）功能的云集线器。我们建议使用 Azure的物联网中心服务 作为云网关。物联网中心提供内置的安全连接，遥测和事件摄取，并具有双向通信

设备包括与命令和控制能力的设备管理。此外，物联网中心提供了可以用来存储设备的元数据的实体店。

对于注册和连接大套 设备 我们建议使用Azure的物联网中心设备供应服务（DPS）。DPS允许特定天青设备的分配和登记 物联网中心端点规模化。我们建议使用Azure的物联网中心的SDK来实现安全设备的连接和遥测数据发送到云中。

流处理过程的数据记录大流和评估为这些流规则。对于流处理，我们建议您使用Azure的数据流分析的，在规模需要复杂的规则处理物联网应用。对于简单的规则处理，我们建议对Azure的功能使用Azure的物联网中心路线。

业务流程整合有利于执行基于流处理过程中从设备的遥测数据囊括见解行动。整合可能包括参考消息，报警，发送电子邮件或短信，集成与CRM，以及更多的存储空间。我们建议使用Azure的功能和逻辑应用业务流程集成。

存储可分为暖路径（即需要为可用于报告，并从设备立即可视化数据），和冷路径（被存储较长的术语和用于批量处理数据）。我们建议使用Azure的宇宙DB温暖路径存储和Azure的Blob存储冷库。对于时间序列特定的报告需求的应用，我们建议使用Azure的时间序列见解。

一个的IoT应用程序的用户界面可以在宽的阵列的设备类型上被输送，在本机应用程序和浏览器。对于UI和报告整个物联网系统的需求是多种多样的，我们建议使用Power BI，TSI浏览器，本地应用程序和自定义Web UI应用程序。

Azure的物联网远程监控[2] 和连接工厂[3] 解决方案加速器是开源产品提供端到端的例子展示了使用Azure的技术与散装的上述技术的建议，让有兴趣的人士快速上手。

1.3 可选子系统

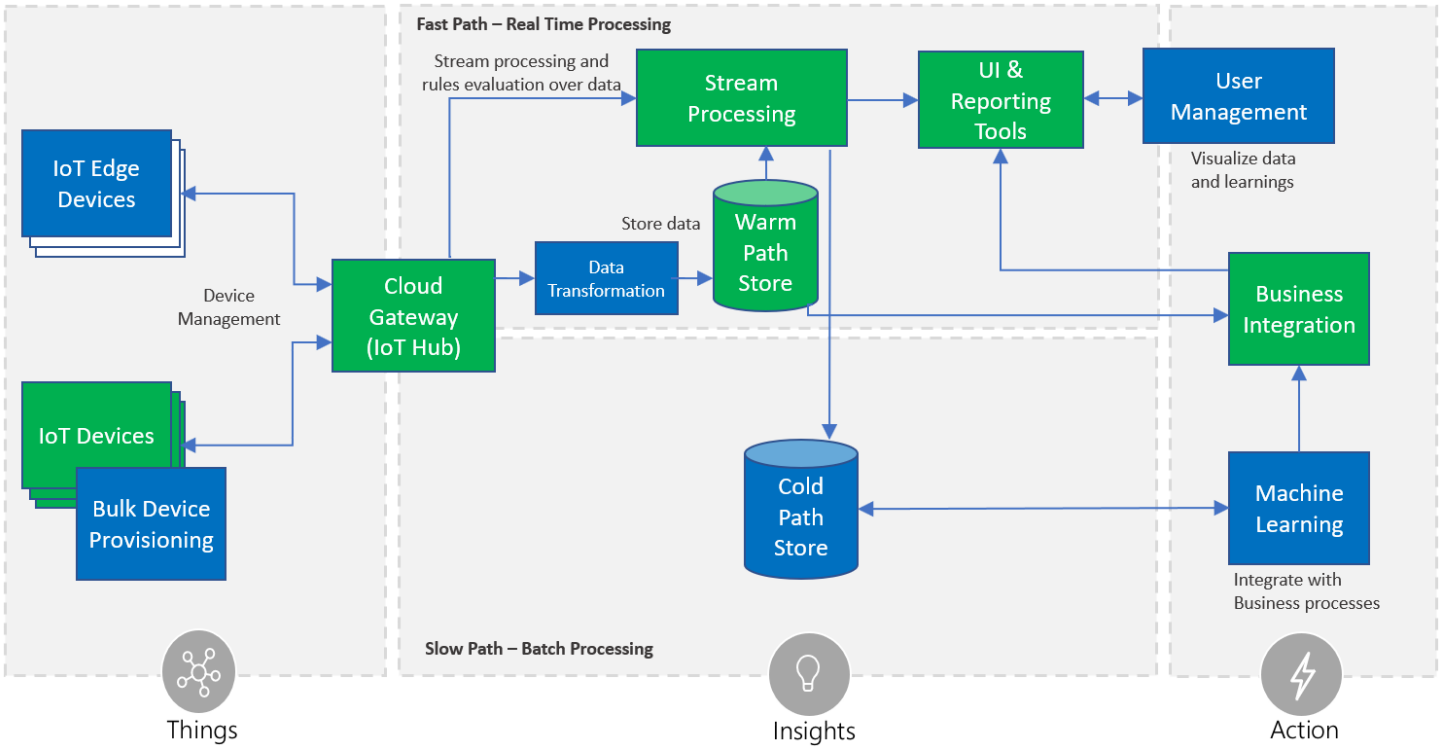
除了核心子系统很多物联网应用将包括子系统：5）智能边缘设备 其允许聚合或遥测数据的转化及对前提的处理，6）云遥测 数据转换 这允许重组，组合，或从设备发送的遥测数据的变换，7）

机器学习 这允许预测算法超过历史遥测数据来执行，从而实现场景，如预测维护，并 8）用户管理 它允许不同之间角色和用户的功能分裂。

[2] https://azure.microsoft.com/blog/getting-started-with-the-new-azure-iot-suite-remote-monitoring-preconfigured-solution

[3] https://github.com/Azure/azure-iot-connected-factory

All Subsystems – Lambda Architecture



智能前端设备 服务于管理访问和信息流的积极作用。它们可辅助设备配置，数据滤波，配料和聚集，数据的缓冲，协议转换，事件规则处理，并且更。我们建议Azure的物联网边缘[4] 用于这些内部部署的需求。Azure的物联网边缘还提供了通过边缘模块，实现自定义的功能可扩展模型。

数据转换 涉及操纵或遥测数据流的聚集之前或它是由云网关服务（的IoT集线器）接收之后。操纵可包括协议转换（例如转换二进制流式数据到JSON），结合的数据点，以及更多。对于遥测数据的转换之前已经通过物联网中心，我们收到 建议使用协议网关。对于数据的转换已经通过物联网中心收到后，我们建议使用Azure的与物联网的功能整合枢纽。

该 **机器学习（ML）子系统** 使系统从数据和经验学习，没有被明确地编程采取行动。场景如预测性维护是通过ML启用。我们建议使用Azure的机器学习的机器学习需求。

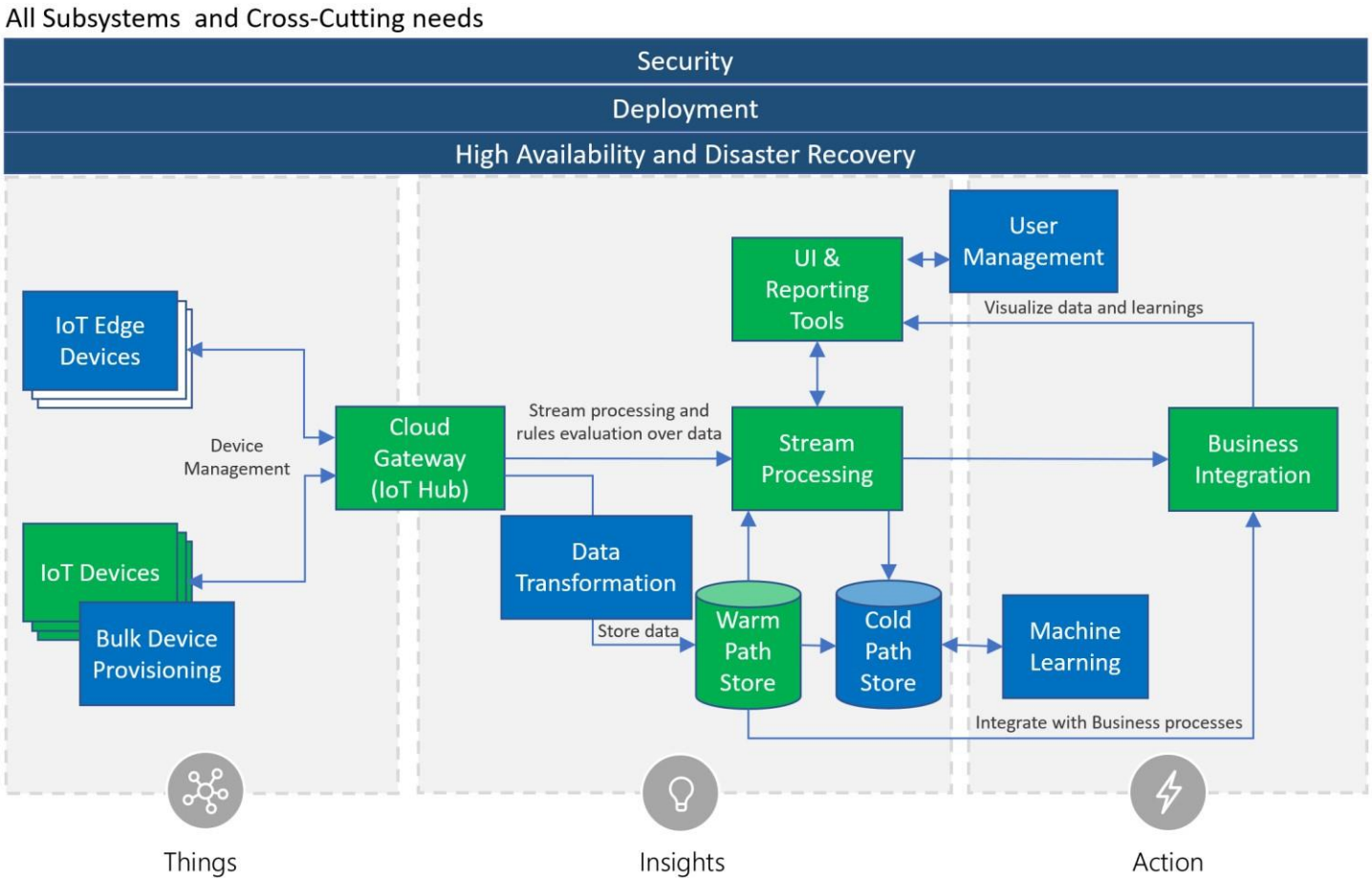
该 **用户管理子系统** 允许不同的功能规范的用户和组执行上的设备的动作（例如命令和作为升级固件用于设备控制，例如）在应用和功能的用户。它是作为下面的跨领域安全要求的一部分进一步讨论。

1.4 跨领域的物联网应用需求

存在用于的IoT应用的多个交叉的需求是成功的关键，包括：8） **安全**
要求; 包括用户管理和审核，设备连接，在途遥测，并在其余的安全性，9）
记录和监控 用于的IoT云应用是用于确定健康和故障排除故障临界

[4] https://azure.microsoft.com/en-us/services/iot-edge/

无论对于各个子系统和应用程序作为一个整体，和10） 高可用性和灾难恢复 其是用于快速地从全身故障中恢复。



安全 在每个子系统的一个重要的考虑因素。保护的IoT的解决方案需要的设备的安全的配置，设备之间的安全连接，边缘装置，和云，到后端的解决方案的安全访问，并且在云中时的处理和存储（加密静止时）的安全的数据保护。如前所述，我们建议使用Azure的物联网中心它提供了一个全托管服务，使通过使用每个设备的安全证书和访问物联网设备和Azure的服务，如Azure的机器学习和Azure中的分析数据之间的可靠和安全的双向通信控制。对于存储技术，我们建议使用Azure的宇宙DB温暖路径存储和Azure的Blob存储冷库两者都在休息支持加密。对于用户管理，如用户认证证书，用户UI功能的授权，报告和管理工具的用户可以访问和审计应用程序的活动，我们建议Azure的Active Directory中。Azure的Active Directory支持广泛使用的OAuth2授权协议，ID连接验证层，并提供系统活动的审计日志记录。

记录和监控 对的IoT的应用是至关重要的确定系统正常运行和故障排除故障。我们建议使用Azure的运营管理套件（OMS），应用程序映射和应用程式深入分析运营监控，日志记录和故障排除。

高可用性和灾难恢复（ HA / DR ）**重点是确保一个物联网系统始终可用，包括从灾难造成的故障。在物联网子系统采用的技术有不同的故障转移**和跨区域的支撑特性。对于物联网应用中，这可能会导致需要重复的服务托管和复制应用

跨取决于可接受的故障转移停机和数据丢失区域的数据。请参阅下的解决方案考虑高可用性和灾难恢复节上的HA / DR的讨论。

2。 基本原则和概念

2.1 原则

参考架构允许组装的安全，复杂的解决方案支持极端规模，但允许灵活性对于溶液场景。这促使整个架构的不同区域以下指导原则。

异质性。 该参考架构必须适应各种场景，环境，设备，处理模式和标准。它应该能够处理庞大的硬件和软件的异质性。

安全[5]。由于物联网解决方案代表了数字和物理世界之间的强有力的联系，构建安全系统是构建安全系统的必要基础。这个参考模型考虑在所有领域，包括设备和用户身份，认证和授权，在休息和数据动态数据数据保护，以及对数据的认证策略的安全性和私密性的措施。

超大规模的部署。 所提出的架构支持数百万联网设备。这将允许与少量设备的启动概念和试点项目证明是向外扩展的，以超大规模的尺寸。

灵活性。 物联网市场的异质需要必要的服务和组件的开放式组成。参考架构是在可组合的原理构建通过组合许多积木以允许的IoT溶液的创建和允许在不同的第一方或第三方技术用于各个概念组件的使用。一些扩展点允许与现有的系统和应用程序的集成。与经纪人的通信的比例高，事件驱动的架构是用于服务和处理模块中的一松耦合组合物中的主链。

2.2 数据概念

了解数据的概念是用于构建设备为中心的数据收集，分析和控制系统的一个关键的第一步。设备和数据模型，数据流，和编码的作用是在以下部分详细说明。

2.3 设备和数据模型

描述的设备，它们的属性和相关数据架构信息模型是实现解决方案的业务逻辑和处理的关键。

有许多不同的设备建模工作正在进行中不同行业，而这个参考架构需要一个中立的立场，以支持这些正在进行的建模和图式的努力。

例如，在一个工业的IoT情景的情况下，数据的语义和结构可以基于OPC UA信息建模框架。[6] 其他实现如家庭自动化和汽车应用可以使用完全不同的特定行业的建模和架构标准。

该架构采用数据流，其中流不需要设备和数据模型，航线，还是在核心平台组件存储信息的基本抽象。在液层，结构化数据将被防护

[5] https://docs.microsoft.com/en-us/azure/iot-fundamentals/iot-security-architecture#threat-modeling-the-azure-iot-reference-architecture

[6] https://opcfoundation.org

通过每当它由部件产生或消耗的数据模型和架构。开发人员使用模式对设备的客户机的开发，后端分析，或者所要求的溶液特定处理逻辑的选择。

2.4 数据记录和流

的IoT解决方案的设计考虑的设备的基本方面周期性地发送 数据记录， 这被表示，分析和作为多个和连续的存储 数据流。消息，事件，遥测，警报 和
摄取 是描述的IoT的数据流时的常用术语。

数据记录通常时间戳，按时间排序和关联到至少一个源。例如，遥测记录可包含测量和当接收到所述数据的时间的时间，并且可以关联到其中测量时，到遥测收集网关，集线器，其中所述设备的名称遥测被摄取等

***食入* 是上载的数据记录到存储，通过网关如天青的IoT Edge和天青的IoT集线器的过程。数据记录可以被摄取一次一个或批量。流的内容可以是重放实时数据或过去的流量。**

***消息* 和 *事件* 是指由连接的设备所产生的数据记录时使用的可互换的术语。术语 *telemetr*y被用于携带由设备传感器报告的数据消息具体地使用，例如来自温度传感器的设备上发送的当前温度。遥测记录可以携带一个或多个 *数据点*， 例如具有一个湿度和一个温度传感器的设备可能发送在相同的消息或在单独消息中的湿度和温度测量。**

设备可以安装多个传感器，并且可以发送记录与自上次遥测已更改所有传感器或仅值报三围被送往，例如，以减少传输的数据量。数据点的遥测记录的值将成为 *最后已知的状态。* 当仅发送差分的记录，设备偶尔也可以发送所有的传感器值的完整快照（称为 *关键帧*），为了一致性和同步的目的。

遥测记录通常进行了分析，在本地或云中，针对一组规则。一种不同类型的数据记录可以作为一个结果来产生，通常被称为一个 *警报。*

数据记录格式

数据记录没有规定的格式。每个数据流的假设是，所有的记录使用兼容的结构和语义。由设备制造商和解决方案，物联网所选择的格式取决于多种因素，如设备上运行的软件，CPU的，带宽，安全性等方面的能力，我们建议物联网解决方案采用JSON格式，由于要求其可读性和相对较低的空间，但有几个二进制格式，如Avro中，可以提高性能和降低成本。

为了简化反序列化，非破坏性变更和版本流的分离应该被允许。最好的做法是为解决方案的IoT使用的消息属性，指定格式和版本包括在每个记录中的元数据，例如。有了一个版本模型，解决方案开发人员可以适当地解决语义或类型方面记录字段的潜在冲突; 例如，如果特定设备的固件的变化，然后所述装置以不同的格式的版本发送数据的记录将允许溶液显影剂到数据流之间的歧义。

在Azure平台的物联网服务，有效载荷无关，不需要任何特定的领域存在于一个消息。消息完整性和兼容性设备和解决方案的开发人员的责任。

2.5 设备交互

参考模型采用服务辅助沟通的原则[7] 方法用于与潜在地部署在不可信物理空间的设备建立可信赖的双向通信。以下原则适用：

- 设备不接受不请自来的网络连接。所有的连接和路由建立在 ***仅出站*** **时尚。**
- **一般设备** ***只连接或建立路由知名服务网关*** **他们正在用凝视。在情况下，他们需要将信息反馈给或从多种服务接收命令，设备被窥视**与下游需要路由信息的关心，并确保命令从授权方他们路由到设备之前只接受一个网关。

- **设备和服务或设备和网关之间的通信路径是** ***在传输和应用协议层固定，*** **相互认证的设备向服务或网关，反之亦然。设备应用程序不**信任链路层的网络。

- **系统级授权和认证，应根据** ***每个设备的身份，*** **并获得证书和权限应该是近乎即时的设备滥用的情况下撤销。**

- 对于被零散地连接由于电源或连接性关系的设备的双向通信可以通过，直到它们连接到挑选那些保持向上命令和通知设备来促进。

- 应用有效载荷数据可以通过网关单独地固定受保护转变到特定的服务。

注意： **用于通知重要的命令功率受限设备的常见图案而断开是通过使用出的带外通信信道，如蜂窝网络协议和服务的。例如，**SMS消息可以被用于“唤醒”的设备，并指示它建立一个出站网络连接到它的“原籍”网关。一旦连接，设备将接收的命令和消息。

2.6 通信协议

今天有可用于设备的场景许多通信协议和数字还在不断增长。从那些用于以确保安全运行具有超大规模系统的使用选择，同时提供由所选择的协议承诺的能力和保证，需要在建立了分布式系统显著的专业知识。然而，对于该协议的选择已经作出，并且这些设备必须集成到解决方案的现有设备的大量。

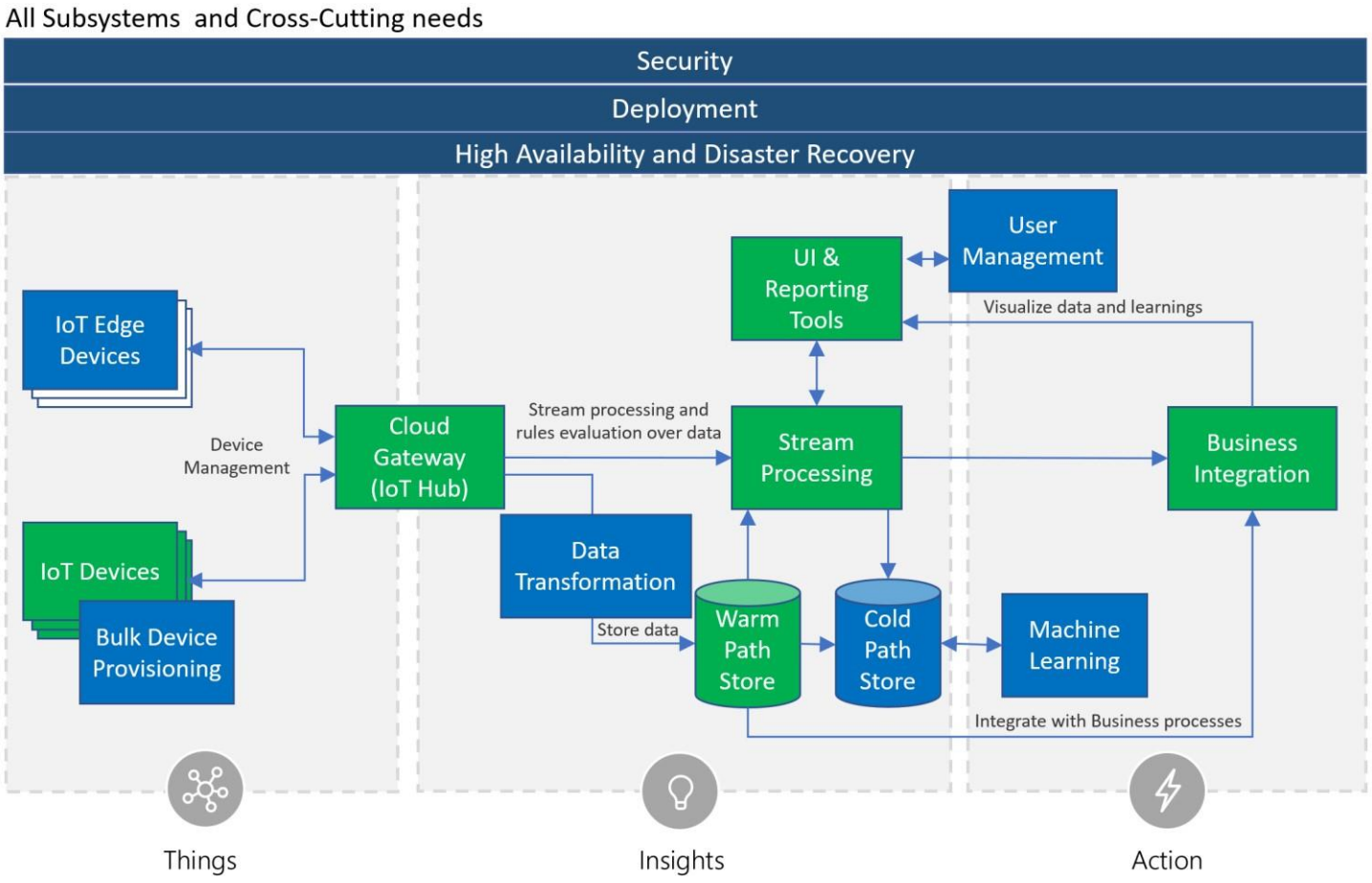
这个参考模型讨论优选的通信协议的选择，说明潜在的折衷与这些选择，并且还明确地允许自定义协议的可扩展性，适应性和本地处理在外地网关噪比（IoT EDGE），基于云的协议网关，或者在流处理。

请注意，通信协议定义的有效负载是如何移动和携带，可用于分派/路由和解码净荷的元数据，但通常没有定义有效载荷形状或格式。例如，该通信可以由AMQP协议被启用，但数据编码可以是Apache的阿夫罗，或JSON或AMQP的本地编码。

[7] http://blogs.msdn.com/b/clemensv/archive/2014/02/10/service-assisted-communication-for-connected-devices.aspx

3。 建筑子系统详细

在这一节中的每个建筑子系统进行了详细讨论，包括子系统的宗旨，为实现技术方案，并建议实施选择。



3.1 设备，设备连接，现场网关（边缘设备），云网关

设备可直接或间接地通过一个字段网关噪声比（IoT边缘装置）来连接。这两款器件和场网关可以实现边的情报能力。这使得汇聚和传输到后端之前减少原始遥测数据，并在边缘地方决策的能力。

以下是对物联网解决方案，不同的设备连接选项的概念表示。该图中的数字对应于四个主要连接模式，定义如下：

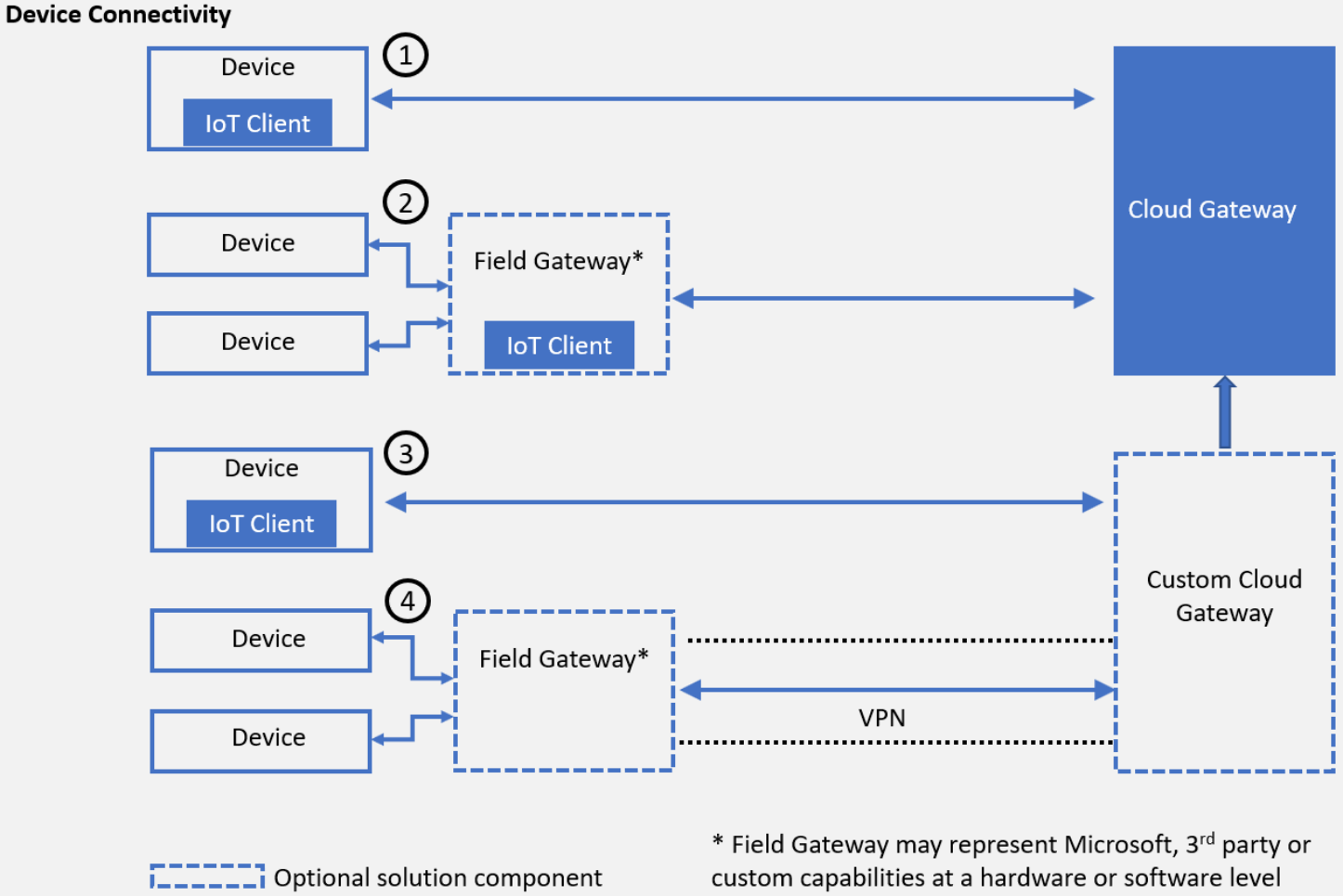
1。 直接设备连接到云网关：对于IP功能的设备，可以通过建立安全连接互联网。

2。 通过现场网关噪声比（IoT边缘设备）连接：对于使用行业的具体标准的设备（如约束应用协议[CoAP协议[8]或OPC UA），短距离通信技术（如蓝牙或ZigBee），以及用于不能够托管一个TLS / SSL堆的资源受限的设备，或设备不暴露于因特网。当转移到云之前，在现场执行网关流和数据的汇总，此选项也非常有用。

[8] http://en.wikipedia.org/wiki/Constrained_Application_Protocol

3。 **通过自定义云网关连接：对于需要协议转换或某种形式的定制的设备**
到达云网关通信端点之前的处理。

4。 **通过现场网关和一个自定义的云网关连接：到以前的模式类似，外地网关**
场景可能需要在云计算方面的一些协议适配或定制，因此可以选择连接到在云中运行一个自定义的网关。有些场景需要使用独立的网络隧道，或者使用VPN技术或者使用应用程序级别的中继服务领域和云网关的集成。



设备的连接的概念上的表示

直接的设备到设备的通信可实现本地网络控制活动和信息流，或其中多个装置进行协调操作的协同操作。单纯的局部作用是这种架构的范围和行业标准，如AllJoyn，经UPnP / DLNA，以及其他覆盖。

理解的术语，并用于在天青的IoT参考架构来描述设备的连接的关键部件是很重要的。以下部分提供了更详细的描述。

设备

异构设备支持。 我们的目标是使几乎任何类型的设备和云网关之间的安全，高效和可靠的通信。这可以直接或通过网关来完成。

目标设备。 在焦点的设备是业务线的资产，从简单的温度传感器到复杂的生产线厂房与数百个组件，里面还传感器。

这些设备的目的将决定他们的技术设计，以及所需的生产资源量和预定寿命操作。这两个因素的组合将定义可用的操作能量和物理空间，因此可用的存储，计算，和安全功能。参考架构是大体朝向运行时，平台，操作系统中性，并执行所述设备的功能。

现场网关（边缘设备）

字段网关，或边缘装置，是一个专门设备器具或通用软件充当通信引擎，并且潜在地，作为本地设备控制系统和设备的数据处理中心。字段网关可以为设备上执行的本地处理和控制功能，并且可以过滤或聚集设备遥测，从而减少数据量被转移到云后端。

场网关的范围包括外地网关本身及其连接的所有设备。顾名思义，现场网关行动外专用数据处理设备，通常和设备搭配。

场网关是在这一领域的网关仅仅流量路由器的不同可以在管理访问和信息流的积极作用，即它们是“智能”设备。字段网关可以辅助设备配置，**数据滤波，配料和聚集，数据的缓冲，协议转换，和事件的规则处理。我们建议Azure的物联网边缘[9]用于场网关中的IoT的解决办法; Azure的**物联网边缘作为一个智能边缘，便于在执行的前提下路由，协议转换，机器学习，人工智能，流处理，等等。Azure的物联网边缘提供了通过边缘模块，实现自定义的功能可扩展模型。

云网关

云网关能够向和从设备，边缘设备，这潜在地驻留在多个不同的网站的远程通信。云网关要么是通过公共因特网或网络虚拟化覆盖（VPN），或专用网络连接到数据中心天青可达，绝缘从其他网络流量云网关和它的所有连接的设备或边缘装置。

它通常管理通信的所有方面，包括传输协议级连接管理，该通信路径的保护，设备认证和授权朝系统。它强制连接和吞吐量配额，并收集用于计费，诊断和其他监视任务的数据。虽然云网关通过一个或多个应用级消息协议执行的数据从该装置流出。

为了支持事件驱动架构云网关通常提供撮合通信模型。遥测和从设备的其它消息输入到云和消息交换由网关促成。数据被持久缓冲，这不仅解耦来自接收器的发送器，还能够使多个数据消费者。从服务后端设备（如通知和命令）的流量是通过一个“收件箱”图案共同地实现。即使当设备处于脱机状态，发送给它的消息将被持久地坚持在商店或队列（表示一个设备的收件箱），并且一旦设备连接递送。由于事件可能timedelayed消耗，提供了一个时间生存（TTL）值是非常重要的，尤其是对时间敏感的命令，如“开汽车或家庭门” 或“启动汽车或机器。” 收件箱模式将这些消息存储在给定的时间TTL，在此之后，消息将到期的持久存储。

[9] https://azure.microsoft.com/en-us/services/iot-edge/

通过所描述的图案经纪的通信允许去耦从云组件的边缘相对于运行时间相关性，处理的速度，和行为的合同。这也使出版商和需要建立高效，规模化，事件驱动的解决方案，消费者可组合。

技术选项

Azure的物联网中心。 天青的IoT集线器是比例高服务从各种设备实现安全的双向通信。天青的IoT集线器连接百万的设备，并支持高容量遥测摄取到云后端以及命令和控制流量的装置。Azure的物联网中心支持云摄入多的消费者，以及对设备的收件箱模式。Azure的物联网中心提供了AMQP 1.0可选的WebSocket支持[10] 支持，MQTT 3.1.1[11]，和本地HTTP 1.1通过TLS协议。

Azure的活动中心。 Azure的事件集线器是一个大规模高摄入，仅在非常高的吞吐率收集来自并发源遥感数据服务。事件集线器也可以在物联网方案中使用，除了的IoT集线器，对二次遥测流（即，非设备遥测），或收集来自其他系统数据源的数据（例如天气饲料或社会流）。事件集线器不提供每个设备身份或指挥和控制能力，所以它可能只为可与在后端设备遥测相关，但不作为连接设备的主网关的附加数据流是适合。Azure的事件集线器提供了AMQP 1.0可选WebSocket的支持，和本地HTTPS协议的支持。

对于超出AMQP，MQTT和HTTP附加协议的支持可以使用协议网关适配模型来实现。CoAP协议是可以使用该模型的协议的一个例子。

自定义云网关

自定义云网关使得协议适配和/或到达云网关通信端点之前某种形式的定制加工。这可以包括由设备（或字段网关）所需的相应的协议实现，同时将消息转发到用于进一步处理的云网关和从所述云网关发送命令和控制信息返回到设备。此外，定制处理，诸如消息转换或压缩/解压缩也可以被实现为定制的网关的一部分。然而，这需要，因为在一般情况下，这有利于数据尽可能快地摄取到云网关，然后再从摄取脱钩云后台执行转换进行仔细评估。

定制网关帮助连接多种定制或专有要求的设备和规范的云底边缘业务。具体的解决方案定制网关通常会充当一个传递设备，可以实现自定义的身份验证或依靠云网关的认证和授权功能。

> 注定制网关可以在边缘部署以及。在某些情况下，可能存在一设备和所述云网关之间多个网关。

技术选项

自定义网关通常是建造和运营，以满足特定解决方案的要求。他们可以依靠内置在合作与系统集成商（SI）和独立软件供应商（ISV）社区共享的开放式源代码。

[10] http://en.wikipedia.org/wiki/WebSocket
[11] http://mqtt.org

天青的IoT协议网关。 天青的IoT协议网关是一种开源框架 定制网关和协议适配。天青的IoT协议网关便于高规模，设备和天青的IoT集线器之间的双向通信。它包括MQTT，展示了实现自定义协议的技术，使MQTT协议行为的定制，如果需要，协议适配器。协议网关还允许额外的处理，例如自定义身份验证，消息转换，压缩/解压缩，或加密/解密。

物联网的客户端

与设备，边缘设备云的通信必须通过安全通道向云网关端点（或云托管定制网关）发生。

除了安全通信信道，该装置通常需要遥测数据传送到云网关，并允许用于接收消息并执行动作或调度那些在客户端适当的处理程序。如前所述，所有设备和网关的连接应在出站唯一的方式建立。

有在物联网系统中使用的客户端连接三个主要模式：

- 从设备的应用程序/软件层的直接连接
- 通过代理连接
- 使用集成在设备或网关的应用程序/软件层客户端组件

直接连通。 在这种情况下，以云网关端点的通信在设备或现场网关软件层使用期望协议本身编码。这就要求所需的协议和消息交换模式的知识，但提供了包括电线上的数据的格式实现全面控制。

代理。 的试剂是安装在该代表另一个程序或管理组件的执行动作的装置或场网关的软件组件。在空间的IoT剂通常控制并作用为在云后端运行的组件。例如，在发送给设备的命令的情况下，代理将接收命令，并且可以在设备上直接执行它。

剂可以是专有的试剂，具体地用于特定的软件解决方案写入，或基于标准的代理实现特定的标准，如OMA LWM2M。在这两种情况下，它的方便设备开发人员集成，并依靠代理商的封装能力; 但是，也有一定的局限性。典型地，代理代表一个封闭的系统，受限于由代理为一组的支持平台提供的功能。移植到其它平台或定制和扩展超出了所提供的功能通常是不可能的。

客户端组件。 客户端组件提供了一组可被集成在设备上运行，以简化连接到后端软件代码的能力。它们典型地提供为库或可以链接或编译成该装置的软件层的SDK。例如，如果一个云后端发送命令到一个装置中，所述客户端组件将简化接收到该命令，虽然执行将在应用程序/软件层的范围中进行。

相比于代理，客户端组件需要集成精力投入到设备的软件，但它们提供了可扩展性和可移植性最大的灵活性。

技术选择Azure的物联网设备的SDK。 天青的IoT设备的SDK代表一组能够在设备上或网关被用于简化连接到天青的IoT集线器客户端组件。该器件的SDK可以用于实现一个客户端的IoT促进所述连接到云。它们提供了跨平台一致的客户端开发经验，并帮助抽象分布式系统从设备开发短信的复杂性。这些库启用设备和现场网关的异构范围的到基于天青的IoT溶液的连通性。他们通过抽象底层协议和消息处理模式的细节简化常见的连接任务。

Azure的物联网设备的SDK是一个开源框架，与天青物联网平台功能一致。虽然这些库简化了连接到物联网Azure的集线器，它们是可选的，如果设备开发商选择连接到使用现有的框架和支持的协议标准，物联网中心端点不是必需的。

3.2 设备身份店

设备身份授权。 该设备身份存储是所有设备身份信息的权限。它还商店和允许对设备客户端身份验证的目的密码秘密的验证。身份存储通常不提供任何索引或超出由设备标识符直接查找搜索工具; 功能性角色由另一家商店，保持应用程序特定的域模型（详见下节）承担。这些店一次分离出于安全原因; 在设备上查找不应该允许披露的加密材料。此外，限制了身份店里一套系统控制属性的最小有助于提供快速反应行动，而在另一方面，领域模型存储的架构是由解决方案的要求来确定。

云网关依赖于身份存储信息的设备认证和管理的目的。身份存储可能被包含在云网关，或可替代地云网关可以外部使用单独的设备标识。

供应。 设备配置使用的身份存储系统的范围，以创造新的设备身份或从系统中删除设备。设备也可以启用或禁用。当他们是残疾人，他们无法连接到系统，但所有访问规则，钥匙，和元数据留在地方。

上的设备身份存储改变应该通过调配制成，在第3.4节中描述。

技术选项

我们建议使用Azure的物联网中心，其中包括一个内置的设备身份商店是注册设备的授权，并提供每个设备的安全证书。

当使用自定义的云网关，它也可以依靠物联网中心的身份商店，其验证和授权功能。万一有具体的解决方案的需求使得需要的身份存储的自定义实现，这将是一个单独的部件，其将主要执行标识符的唯一性，存储用于设备的所有必需的安全密钥，并且将有可能保持一个“启用/禁用”状态。如果它包括透射密码短语，这些应存放在盐腌散列的形式。请记住，需要进行适当的保护身份存储的自定义实现，因为它存储的证书信息。

身份店应该只允许访问该系统作为必要的特权部分; 定制网关将查找所需的认证材料这家店。

如果不使用Azure的物联网中心，对外实现可以通过Azure的宇宙DB，天青表，Azure的SQL数据库，或第三方解决方案来实现：

- **Azure的宇宙DB：** 在Azure的宇宙DB，[12] 每个装置由文件表示。系统级的设备标识符直接对应于所述文档的"ID"。所有其他属性保持旁边的文档中的"ID"。

- **Azure的表：** 在Azure的表，身份存储映射到表。每个设备由行表示。系统级的设备标识符在PartitionKey和RowKey的组合，它们一起提供唯一性保持。所有其他属性在列举行; 如果需要复杂的数据可以存储为JSON，。跨越这些领域中的标识信息的具体拆分是应用特定的和应遵循的服务规模的指导。[13]

- **SQL数据库：** 在SQL中，身份存储也映射到一个表，每个设备由行表示。系统级的设备标识符在群集索引主键列保持。所有其他的属性存储在列; 复杂的数据或者需要扩展数据可以根据需要被存储为JSON，。

- **第三方选项：** 可通过Azure的Marketplace或直接部署到Azure的计算节点的第三方解决方案可以被使用。例如，在卡桑德拉，每个设备可以通过在列族一行表示。这家商店将被分割，并根据需要建立索引以便快速访问。

3.3 拓扑和实体店

设备和应用模型。 设备和应用程序模型基础构建的应用程序的业务逻辑。实例包括定义和配置的业务规则，对于装置或应用实体的一个子集进行检索，建立用户界面和仪表板，以及以确保整个溶液和其他后端系统的不同组件的一致性的能力。

设备型号常常形容：

- 该架构为关于包括的特性和/或装置的功能的设备元数据。元数据模式和价值观很少改变。

  设备元数据的实例是设备类型，制造商，型号，序列号，容量等
- 数据模式用于由所述装置发射的数据，它定义了遥测与他们的数据类型和允许的范围的属性。

  例如，环境监测设备将发射温度，定义为属性名：温度，数据类型：小数，测量单位：华氏，和数据范围[10 - 110]，且被定义为属性名湿度：湿度，数据类型：十进制，测量单元：百分比，数据范围[0-100]。

- 模式配置参数的控制装置的行为。
  例如，一些环境监测设备的行为的可以是由参数，如取样频率，遥测发送间隔，和操作模式控制。

- **操作和参数的控制动作的设备可以执行。例如，具有连接的致动器的设备可以暴露远程操作，例如** *左转（度），turn_right（度），* **和** *flash_warning_light（NUMBER_OF_TIMES）。*

- 代表域模型设备的拓扑结构，如设备和其他实体，并为企业的经营范围内的语义之间的连接关系的丰富。

[12] https://azure.microsoft.com/services/cosmos-db
[13] http://msdn.microsoft.com/library/azure/hh508997.aspx

例如，建筑物管理系统可使用的域模型包括实体，诸如校园（或建筑群），建筑物，楼层，房间，资源，设备，和传感器。拓扑模型定义了实体的属性（如属性，操作，等等）以及所述实体之间的关系。

该应用模型的复杂性高度依赖于特定领域的需求。在某些情况下，分层拓扑模型将会被使用（例如，用于建模的校园/建筑/楼层/房间/对建筑物管理系统资源和设备），而在其他情况下，图布局可能更合适（如运输物流公司一种操作船队可能需要更灵活的关系中，具有例如一个车辆与多个车队组相关联，并与常动态变化的关系。

设备被表示为在整个应用程序拓扑节点。在众多的解决方案从业务的角度感兴趣的实体是不一样的设备本身。一个公司的主资源可以是机器或产品，其具有嵌入的一个或多个设备。在楼宇管理解决方案的情况下，建筑物和房间实体，将有一个大量的应用，并与他们相关的业务逻辑的，而设备提供配套的监控功能和远程控制。在车队管理解决方案中，连接的车辆，例如，具有与其相关联的多个设备和这些设备的一个子集与所述通信到和从车辆帮助。该应用程序的业务逻辑主要集中在车辆的车辆组，

定义和功能。 拓扑结构和实体店是一个包含应用程序的实体和实体之间的关系的数据库。它也包含设备元数据和用于调配的设备（由设备实体在整体拓扑表示）属性。

拓扑结构和实体店包含应用程序模型的“运行时”表示。天青的IoT参考架构不强加对于设备元数据的任何特定实体或设备模型，架构或结构。它假定这些是在“设计时间”的具体的IoT溶液定义的，通过适当的建模工具的系统的配置期间发育过程中或动态即。它可以定义自己的应用程序模型或选择一些垂直行业标准模型。

在配置中，每个装置与在拓扑和实体存储元数据记录（设备实体的实例），其可包含结构化的和/或非结构化的元数据，基于所定义的模型（在设计时）注册。

设备身份登记与拓扑结构和实体店。 虽然设备身份商店只包含systemcontrolled属性和加密材料，拓扑结构和实体店有设备的完整表示，包括其关联到其他实体，如产品，资产，或机器。在身份存储在记录确定装置是否被注册，并可以与系统进行认证。出于安全考虑，这是一个很好的做法，以保持与安全相关的信息从设备实体分开。实体店不能存储与该设备的任何键或其他加密信息。

设备标识存储表示设备身份（主要用于认证目的）的权威列表。与此相反的拓扑结构和实体店，有关系，其他应用实体，用于执行其业务功能的应用程序所需之间的全套设备的元数据（设备属性，属性，操作等）。这家商店是一个用于设备发现，以及其他应用实体的发现，并提供覆盖索引和强大的搜索功能。

拓扑结构和实体店是实体及其对物联网解决方案关系的权威商店，确保整个系统一致的看法。由于技术原因，的包含的信息预测可以存储或快速访问其他组件缓存。然而，真理的实体及其关系的来源是这家店。

更改它可能需要根据需要进行同步或传播到其他组件。例如，一些设备的元数据属性可能会在设备管理业务流程的物联网中心设备双必要。在这种情况下，该装置实体的那些属性的变化需要被施加到的IoT集线器设备双胞胎。反之亦然，如果属性值是在装置的双（从设备推出）改变，这种改变将传播到的拓扑结构和实体存储设备的实体。在其他情况下，具体的高级分析任务可能需要在一个特定的存储组件或格式的设备基准数据的副本。

元数据。 描述反映了设备或它的操作环境的状态的装置本身和业务数据的元数据之间的区别是重要的，因为它直接影响该设备的信息如何可以被使用，高速缓存，和在整个系统中分布。元数据通常是缓慢变化的数据，而操作数据预计将快速变化。

例如，交通灯极的地理位置的元数据，但在车辆的当前地理位置的位置被认为是操作数据。车辆识别号，型号，并会的元数据。对道路的某一地段的所有交通信号灯发现可作为对拓扑和实体店查询，同时寻找所有车辆当前行驶的道路的某一地段将超过运营数据的分析任务来执行。在拓扑结构和实体店的元数据可以帮助作为参考数据寻找的道路上特定型号的所有车辆，但是。

技术选项

拓扑结构和实体店提供有关的实体和设备应提供与提供快速查找的目标丰富或自由索引能力的描述信息。

注册表商店可以在下列技术之一之上实现：

- **Azure的宇宙DB：** Azure的宇宙DB是一个全面管理图形数据库，并允许物联网设备，实体和它们作为图形拓扑结构的建模。在天青波斯菊DB，每个设备可以通过一个文件来表示。系统级的设备标识符直接对应于所述文档的“ID”。所有其他属性保持旁边的“ID”宇宙DB是非常适合的拓扑结构和实体店的功能，因为它接受任意结构化数据，并自动创建索引（除非禁用特定的属性）。这允许快速和**灵活的查询[14] 整个注册设备和其他实体。它也允许使用宇宙DB API图形拓扑的轻松导航。**

- **Azure的SQL数据库：** 在SQL中，每个器件可通过在表中的行来表示。系统级的设备标识符在群集索引主键列保持。所有其他的属性存储在列; 复杂的数据或需要扩展数据可以根据需要被存储为JSON或XML。基于查询模式相应的列需要建立索引。其他应用实体可以通过SQL表来表示。实体之间的关系可以用SQL数据库的关系数据库功能来表示。

- **第三方选项：** 除了管理Azure服务，可通过Azure的Marketplace或直接部署到Azure的计算节点的第三方数据服务可以被使用。在这种情况下，实际模式取决于所使用的产品，但结构将是类似于用于Azure的宇宙DB或SQL Azure的数据库之一。根据需要快速访问基于适当的设备或实体的属性分区和索引将被应用。例如，对于在Apache的卡桑德拉数据库代表器件[15]，

[14] https://docs.microsoft.com/azure/cosmos-db/documentdb-sql-query

[15] http://cassandra.apache.org

Cassandra的列族可以具有设备标识符作为分区键和可以在该装置的其它性质定义附加的索引。

3.4 设备配置

定义。 配置表示该设备的生命周期的步骤中，当一个设备要进行系统已知的。配置API是共同的外接口，用于改变如何到后端的内部组件，特别是设备标识存储和拓扑和实体商店制成。它提供了具有共同的手势的抽象接口，并且对于该设备的身份和拓扑结构和实体存储了抽象接口的实现。实现可扩展到包括其他组件和系统。

设备配置在后端典型地发起，通过在系统中登记的设备他们变得可操作之前。在一些情况下，这可以在制造的设备（包括在设备标识和连接到的IoT后端所需的凭据燃烧）的过程中发生。在其他情况下，供应可以立即装置在安装过程中该装置被接通之前对使用情况，例如进行。第一时间的装置试图建立到后端的连接，可能会执行附加步骤，以完成其配置（或"引导"的话）以供使用。

供应工作流程。 中的溶液的供应工作流程负责处理从个人和散装请求用于注册新的设备和更新或移除现有的设备。它也将处理激活，并有可能暂时中止访问和最终的访问重新开始。这也可能包括与外部系统的交互，如移动运营商的M2M API来启用或禁用SIM卡，或与业务系统如计费，支持或客户关系管理解决方案。供应工作流程保证，该设备与需要了解其身份和其他元数据的属性所需的所有后端系统注册。

引导工作流程。 当一个设备想要连接到首次后端系统，可以执行附加的步骤，以完成其配置。这可能包括配置或软件更新设备的第一次使用前，应用。在引导步骤中，该设备可能被分配到与它的新凭证的新"家"端点。这是多租户系统或全球分布式部署尤为重要。关于谁是将要使用的设备（租户），或者该设备将被用于（地理位置）影响的决定在哪里"家"的设备（即，云网关负责连接与设备信息）。在许多情况下，供应组件实现为一个"全球性"服务，代表设备的登记入区域部署，引导过程中以及策划设备到他们的"家"端点。在这种情况下，全球性的服务可以实现更高级别的工作流程配置，使用相同或类似的外部接口的区域组成，并委托操作区域配置组件为宜。

技术选项

我们建议使用Azure的物联网集线器设备供应服务[16]（DPS）为设备配置。DPS是支持登记和配置在多个的IoT集线器设备的全球供应服务（即引导）。DPS简化设备配置的自动化到设备身份存储噪声比（IoT集线器的一部分），同时提供灵活性以控制设备的分配。DPS提供用于设备登记后端系统的设备配置（自举）的API，以及API。DPS可以在供应工作流被用于跨的IoT集线器，设备的分配自动化一起与其它步骤在其它后端组**件和系统（诸如拓扑和实体存储注册设备，或3-RD 党提供商系统）。对于分布在全球的部署，每一个步骤，可能需要使用一个全球性或区域性系**统注册。

[16] https://azure.microsoft.com/roadmap/azure-iot-hub-device-provisioning

Azure的API应用程序[17] 可用于外部供应API的实现。API应用程序提供了构建，托管和分发在云和内部部署的API的平台。API应用天青逻辑应用无缝集成，[18] 其可以用来在溶液的IoT组件和外部业务系统一个总体供应工作流程的执行。

设置接口是一个简单的姿势集，用于管理设备的生命周期。设置接口（API）应在装置的身份和拓扑结构和实体的注册表存储和任选的其它内部溶液成分主要的API如果需要来实现。它不仅从溶液UI（例如，设备管理门户）使用，但也可以用作用于较高级别的工作流的接口，也可以与外部系统，如移动运营商的M2M API用于管理SIM卡交互或用于激活计费服务后端业务系统。

安全密钥可以在API的外部产生并传递作为参数或可以创建和由服务分配作为配置API调用的一部分。

生成一个安全令牌可以在配置API中使用所需的签名密钥来进行。颁发给设备令牌将在范围被限制在一个特定的端点（例如，在集线器的IoT或**事件集线器发行者策略的情况下的设备端点）。在返回的数据 *寄存器* 和 *ResetCredentials* 操作包含必须被转移到设备所需的安全令牌。可替换**地，可以在设备上生成或外部安全令牌，并传递给设备。

对于自定义网关，可以在外部产生所需的凭据并传递到API用于存储，或API可以扩展到创建密钥。

3.5 存储

定义。 物联网解决方案取决于有多少设备是在溶液中，他们经常发送数据产生显著的数据和有效载荷的在从设备发送的数据记录的大小。数据常常是时间序列数据和需要被存储在何处它可以在可视化中使用和报告，以及用于附加的处理以后访问。这是常见的有数据拆分成“暖”和“冷”数据存储。温暖的数据存储认为，需要与低延迟访问最新数据。存储在冷库数据通常是历史数据。最常选择的冷库数据库解决方案将是成本便宜，但提供了比温暖的数据库解决方案更少的查询和报告功能。

存储常见的做法是保持最近在温暖的存储遥测数据的范围（如最后一天，一周或一个月），并储存在冷库的历史数据。使用这种实现，应用程序可以访问到最新的数据，并可以快速观察近期的遥测数据和趋势。检索历史信息的设备可以使用冷库来完成，通常比如果数据是在温暖的存储较高的延迟。对于一般用途的情况下，我们建议Azure的宇宙DB温暖存储和Azure的Blob存储冷库。如果解决方案需要涉及在许多事件和在许多设备上聚集频繁的查询，我们建议的时间序列见解温暖的存储。

暖储藏

为预先确定的最近间隔A暖储藏数据库存储设备状态，并且还可以存储每个设备易于存取的最后已知状态。这些数据必须在数据库中提供快速（内之事最好

[17] https://azure.microsoft.com/documentation/articles/app-service-api-apps-why-best-platform

[18] https://azure.microsoft.com/documentation/articles/app-service-logic-what-are-logic-apps

从当数据被摄取到云网关从设备）和容易地查询用于简单的场景，如可视化当前的设备的传感器值或在最近的时间内可视化值秒。常见的查询模式包括：对于最近的日期和时间范围为设备数据，一个或多个设备聚集的数据，以及用于遥测点为特定设备的最后已知值。存储在数据库中温暖的数据可以是原始数据，聚合数据，或两者。见第3.6节 - 数据流和数据流处理上的流处理的信息。

如果解决方案具有摄取率很高（几十万或数百万每秒消息的顺序），可能需要专门的高摄入的数据库。高摄入的数据库的建议是即将到来的。

评估标准

基于以下标准温暖的存储解决方案进行了评估。这些标准将不适用于所有的物联网解决方案，但设计是最普遍适用于整个物联网解决方案。

1. **安全。 该解决方案提供成熟和强大的功能，如在休息，身份验证和加密**
   授权和网络安全。
2. **简单。 该溶液是有据可查的，并具有良好定义的架构。发展任务**
   据可查，通过软件开发工具包（“软件开发工具包”）的支持，并在本地开发环境测试。部署和运营任务由文档，工具和用户界面的支持。
3. **性能。 读取和写入到数据库是快速和扩展到许多并发读取和写入。询问**
   性能也快。
4. **可扩展性。 数据库支持存储千兆字节到万亿字节的数据。向外扩展时不需要停机。**
   理想的解决方案可自动适应成本和计算能力来提供的负载。
5. **查询功能。 该数据库有必要对整体解决方案的查询功能。**
6. **价钱。 该数据库是负担得起的两个存储容量和吞吐量的需求。**

技术选项

Azure服务与第三方选项

所有的Azure服务是安全的，使安装和维修简单的维修。除了Azure服务，第三方，在Azure中托管的自我管理选项可用于存储为好; 如卡桑德拉。由于增加了复杂性和较高的开发商运营成本（导致拥有更高的总拥有成本）的自我管理的服务，我们建议使用，而不是自我管理服务的托管服务。自我管理服务需要存储和计算需求，以及运营团队来管理资源规划。应当指出的是不过自我管理服务提供了极大的灵活性和控制能力使它们有可能适合您的方案。

一般用途

推荐：天蓝色的宇宙DB。 我们建议Azure的宇宙DB作为通用温暖的存储解决方案。天青宇宙DB是一个安全的，高度可扩展的，低延迟的NoSQL数据库（没有数据存储或吞吐量限制）。这是最好的，可以从中受益的数据集 灵活，模式无关，自动 索引和丰富的查询界面。Azure的宇宙DB有5个API类型和数据模型，SQL，MongoDB的，图，表和卡桑德拉，提供灵活选择基于解决方案的数据需求的数据模型。宇宙DB允许多区域的读写，并支持除了自动故障转移手动故障转移。另外，宇宙DB允许用户设置自己的数据，这使得即将到期的旧数据的自动时间生存（TTL）。定价是基于存储和使用单位申请

供应[19]。宇宙DB是最好的，不需要涉及超过大集在许多设备上数据的汇总查询，这些查询需要比一个基本的查询更请求单元，如一个设备的最后一个事件的情况。

Azure的SQL数据库。 Azure的SQL数据库最适合需要关系存储和查询功能的数据集。Azure的SQL数据库也提供了数据管理，保护和安全以及业务连续性高级功能。定价基于供应的存储配置和数据库交易单元或弹性数据库交易单元的组合[20]。

手动缩放，以增加存储空间没有停机时间。SQL数据库还具有内置的复制和自动故障转移的区域，以确保数据在断电不丢失。我们建议Azure的宇宙DB在Azure的SQL数据库由于关系存储的需求，需要手动缩放数据库，并写摄取规模和吞吐量的限制。

Azure的时间序列数据分析（ TSI）。天青TSI是时间序列数据的分析，存储和可视化的服务，提供的功能，包括类似SQL的过滤和聚集，减轻对用户定义的函数的需要。在Azure的TSI所有数据都存储在内存和固态硬盘，保证数据被快速支持交互式分析。Azure的TSI还提供可视化等不同的时间序列，仪表盘比较，访问表格视图和热图的叠加。Azure的TSI提供数据资源管理器可视化和查询数据以及REST API的查询。此外，它暴露了 的JavaScript控件库 使嵌入时间序列图到自定义应用程序。Azure的TSI适合于需要建立在可视化服务，并且不需要对数据立即报告（TSI拥有用于查询30-60秒的数据记录的近似等待时间）的解决方案。TSI是非常适合于需要查询聚集在大的数据集，如TSI允许任意数量的用户进行查询的数量不受限制无需支付额外费用的解决方案。今天，TSI为400天，最大保留和3 TB的最大存储限制，因此使用TSI解决方案将需要使用冷藏库（可能为需要查询交换数据为TSI），以及如果顾客需要较大的量存储或更长的保留的。TSI是我们的时间序列数据的存储和分析推荐

第三方：Apache的卡桑德拉。 阿帕奇Cassandra是一个线性缩放，高度可用的NoSQL数据库，可以跨越不同地理区域集群。它使用CQL查询语言，这是SQL蓝本。它提供了认证，加密和防火墙功能，以及数据复制。此外，它在writeheavy场景表现良好（它可以实现每秒超过100万**的写入[21]，这使得它非常适合的物联网解决方案，具有遥测摄入高水平。**

比较网格通用

		Azure的宇宙DB[22]	Azure的SQL数据库[23]	Apache的卡桑德拉
安全	加密在休息	是	是	是
	认证的主密钥，主动目录集成		SQL身份验证，Azure的Active Directory验证	JMX用户名和密码，基于角色的访问
简单	支持防火墙	是	是	是
	数据模型	多模型	相关的	宽列

[19] https://azure.microsoft.com/en- 我们/价格/信息/宇宙 -D b/
[20] https://azure.microsoft.com/en-us/pricing/details/sql- 数据库/
[21] https://medium.com/netflix-techblog/benchmarking-cassandra-scalability- 上 -aws-过一百万写入每秒，39f45f066c9e
[22] https://azure.microsoft.com/en-us/services/cosmos - D b/
[23] https://azure.microsoft.com/en-us/services/sql- 数据库/

	开发者的SDK .NET，Java和Node.js的，Python中所有的API。	。NET，Java和Node.js的，Python和Ruby，和更多	。NET，C / C ++，Java和Node.js的，PHP，Python和更
	可用性	99.99％的SLA	99.99％的SLA	高可用性，无单点故障
	地区供应	30+ Azure的地区	30+ Azure的地区	N / A
	数据复制	自动本地复制。Georeplication可用。	自动本地复制。Georeplication可用。	可以实现与多个节点的复制。
	灾难恢复	自动故障转移。	自动故障转移。	需要配置。可以传播集群跨越多个区域。
	数据限制	没有	最多4 TB	没有
	产量限制	没有	最大4000的DTU / eDTUs[24]每个数据库/弹性普尔	没有
	单写入性能	<1秒	<1秒	<1秒
	单读取性能	<1秒	<1秒	<1秒
PERF /可扩展性	单一的简单查询性能[25]	<1秒	<1秒	<1秒
	汇总查询性能[26]	>1秒[27]	1秒	>1秒[28]
	查询语言（S）	SQL（先前DocumentDB），MongoDB的，表，卡桑德拉，格拉夫（阿帕奇TinkerPop有关，的Gremlin）	T-SQL	CQL
查询功能	定价模型	存储使用和RU供应。可扩展的资源单元上或根据需要向下。	存储和DTU / eDTU供应。	依赖于安装
价钱				

[24] https://docs.microsoft.com/en-us/azure/sql-database/sql-database-what-is-a-dtu

[25] 例如：最后一分钟选择一台设备的数据

[26] 防爆。对于最后一小时设备数据返回百分位数（假设3600个数据点，计算P25，P50，P75，P90，P99）

[27] 聚合函数本身不支持在宇宙DB。用户定义的功能是必需的。

[28] 如SUM函数SQL的支持，但百分位值和更复杂的计算都没有。用户定义的功能是必需的。

冷库

相反，保持所有数据与低延迟，高吞吐量和完整的查询功能，一个温暖的数据存储，数据可以被分成冷暖存储路径。这可以提供更低的存储成本，同时仍保留历史数据。冷库数据库保存不需要尽快和/或频繁地暖存储的数据，但仍可能有必要在未来的访问报告，分析，机器学习使用，等等。

评估标准

基于以下标准冷库解决方案进行评估。这些标准将不适用于所有的解决方案，但设计是最普遍适用的一个物联网解决方案。

1. **安全。** **该解决方案提供成熟和强大的功能，如在休息，认证和授权，以及网络安全加密。**

2. **简单。** **该溶液是有据可查的，并具有良好定义的架构。发展的任务是记录，由软件开发工具包（“软件开发工具包”）的支持，**并在一定程度上可测试在本地开发工作站。部署和运营任务由文档，工具和用户界面的支持。

3. **可扩展性。** **数据库支持存储大量的数据。向外扩展时不需要停机。该数据库具有非常长期或无限期保留（两年左右）。理想的**解决方案可自动适应成本和计算能力来提供的负载。

4. **价钱。** **该数据库是负担得起的大量数据。**

技术选择

一个解决方案的最佳冷藏库是高度依赖于什么目的数据库将有助于。下面的两个数据存储解决方案是专为低价高档次，但每个都有不同情况下的长处。我们建议Azure的Blob存储在一般的情况下，因为它比Azure的数据湖便宜，尤其是在写入请求的条件，目前已在多个地区提供，并有更好的灾难恢复。但是，如果解决方案需要冷藏的数据分析（用Hadoop，Azure的数据分析等），或者需要带U-SQL查询，数据湖设计时考虑到该方案中，可能是更好的选择。

推荐：Azure的Blob存储。 Azure的Blob存储是一种简单，廉价的文件存储数据库。斑点可用于存储原始设备数据。使用网页的斑点，而不是块或追加斑点应视写入操作的频率被认为是[29]。Azure的Blob存储有完整的安全功能，本地或地理冗余存储选项，并在所有Azure的地区。它是高度可伸缩的，所述最大存储限制为500 TB和每个帐户的最大请求速率是每秒20,000个请求[30]。

Azure的数据湖。 Azure的数据湖是一个分布式数据存储，能坚持大量的关系和非关系数据的不进行改造或模式定义。它是用于存储数据库的一个很好的选择，如果需要大数据分析和/或无限的存储空间。它是略高于Azure的Blob存储（特别是在写操作方面）更昂贵，但它是为大数据分析工作负载而优化。该数据库可以从Hadoop的通过WebHDFS兼容的REST API或使用U-SQL语言进行访问。它拥有本地冗余存储和在美国的一些地区天青以及北欧可用。我们建议Blob存储在数据湖，由于在价格上的小幅贴水，一些小区域的可用性，以及缺乏地理冗余存储的。

[29] https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blobs-introduction

[30] https://docs.microsoft.com/en-us/azure/azure-subscription-service-limits#storage-limits

比较网格

		Azure的Blob存储[31]	Azure的数据湖[32]
安全	加密在休息	是	是
	认证	共享秘密，HMAC	Azure的Active Directory中，OAuth 2.0用户
	支持防火墙	是	是
简单	审计	是	是
	数据模型	与平面命名空间对象存储	分层文件系统
	API	通过HTTPS / HTTP REST API	REST API在HTTPS，U-SQL
	服务器端API	Azure的Blob存储REST API	WebHDFS兼容的REST API
	开发人员SDK	。NET，使用Java，Python，Node.js的，C ++，Ruby的，PHP，围棋，安卓，iOS版	。NET，使用Java，Python，Node.js的
	可用性	之间99.9〜99.99％SLA用于读（取决于复制），99.9％SLA用于写	99.9％SLA用于读/写
	地区供应	Azure的所有地区	东美2，美国中部，北欧
	数据复制	默认情况下，本地冗余。也可以只读地理冗余	本地冗余
	灾难恢复	如果有地理复制的数据：可以从中断的情况下，二次阅读，并在紧急情况下故障转移数据。	能否处理瞬时硬件故障。建议在区域故障的情况下手动数据到辅助区域复制。
可扩展性	数据限制	500 TB	没有
	产量限制	20,000个请求/秒	没有
价钱	定价模型	数据使用和请求计数[33]	数据使用和请求计数[34]

食入高

即将节

3.6 数据流和流处理

当数据被摄取到的IoT后端，了解如何数据处理的流程可能会有所不同是很重要的。取决于情景和应用程序，数据记录可以流过不同的阶段，在组合不同的顺序，并经常通过并发并行任务处理。

这些阶段可分为四类： 存储，路由分析 和 动作/显示：

- *存储* 包括内存缓存，临时队列和永久存档。

31 https://azure.microsoft.com/en-us/services/storage/blobs/

32 https://azure.microsoft.com/en-us/services/data-lake-store/

33 https://azure.microsoft.com/en-us/pricing/details/storage/blobs/

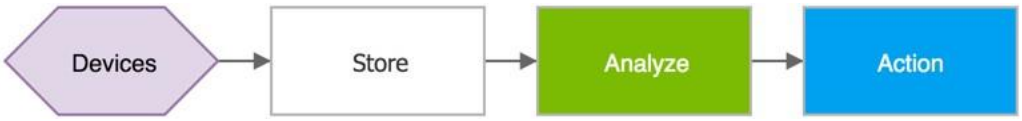
34 https://azure.microsoft.com/en-us/pricing/details/data-lake-store/

- *路由* 允许的数据记录分派给一个或多个存储端点，分析过程和操作。
- *分析* 被用于通过一组条件来运行输入数据记录，并且可以产生不同的输出数据记录。例如，在编码阿夫罗输入遥测可以返回JSON格式编码的输出遥测。
- 原始输入数据的记录和分析输出记录通常被存储和提供给显示器，并且可以触发诸如电子邮件，即时消息，事件票据，CRM任务，设备命令等的操作

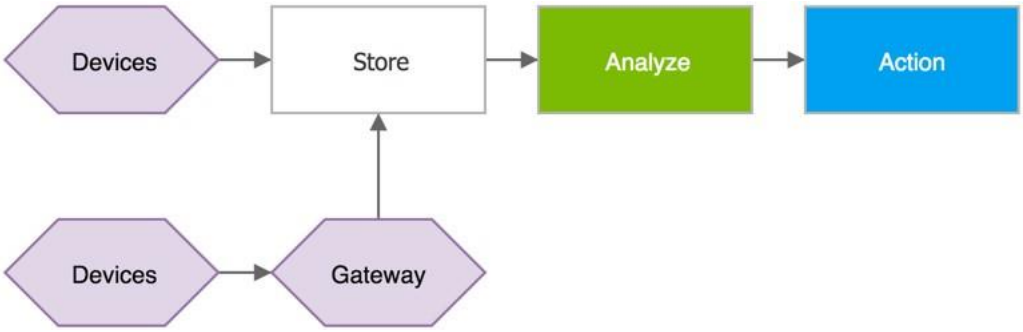
这些方法可以以简单的组合 *图*，例如，以实时显示接收到的原始遥测，或更复杂的图形执行多个和高级任务，例如更新仪表盘，触发报警，并启动业务整合流程等

例如，下面的图表示一个简单的场景中哪些设备发送被暂时存储在天青的IoT集线器，然后立即显示在图形屏幕上可视化的遥测记录：

下图代表了另一种常见的情况，在这种设备发送遥测，存储它短期内天青物联网中心，分析数据以发现异常后不久，然后触发如电子邮件，短信，即时消息等操作：



物联网体系结构也可以由多个摄入点。例如，可以在发生前提一些遥测存储和/或分析，设备和场/边缘网关内; 或协议转换可能需要约束设备连接到云。而得到的曲线图是更复杂的，逻辑构建块是相同的：



推荐数据流

该参考架构假设一个企业运行多个并行的流处理器，可以通过分割摄入流，或通过转发数据记录到多条管线。我们建议划分可以基于消息属性（例如，设备ID，设备位置，有效载荷格式，等等），以避免路由之前有效载荷反序列化，但该文档包括能够路由的基于JSON消息的内容的解决方案。

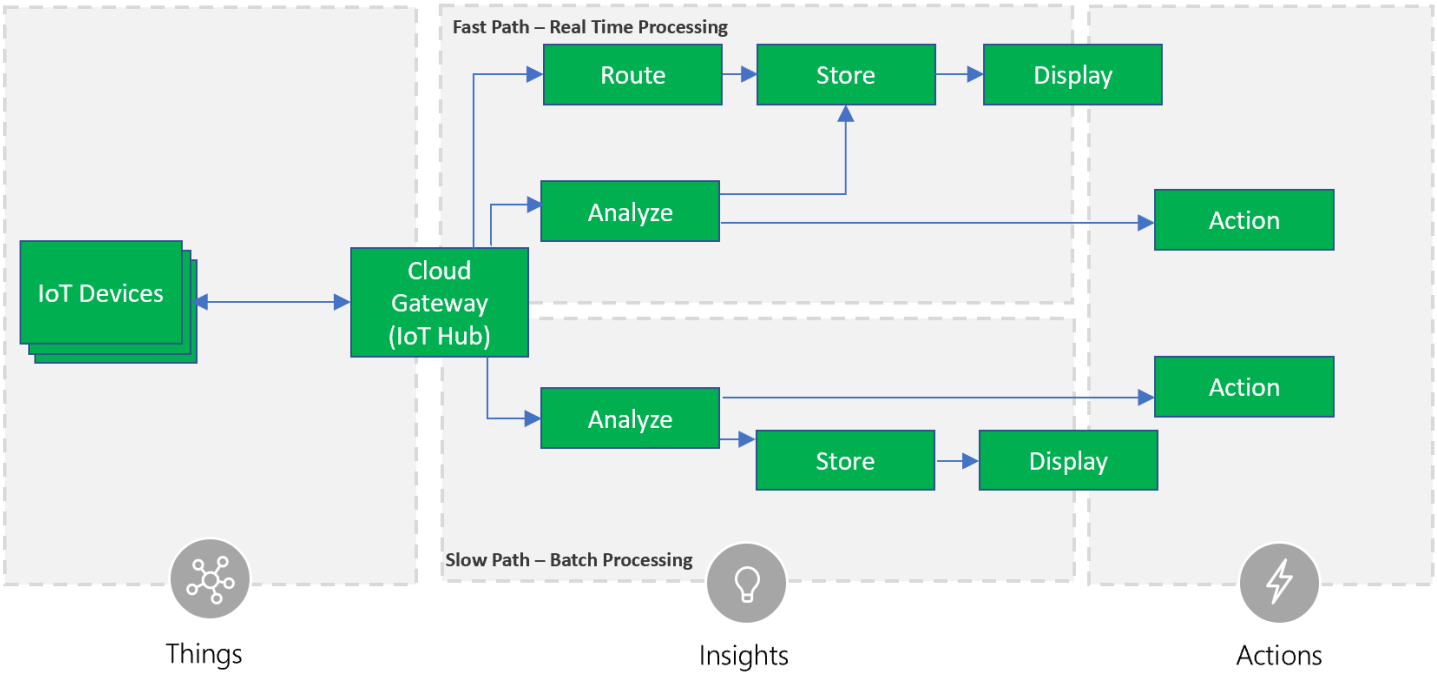
下图，也被称为 *LAMBDA架构*，示出了设备到云的消息和事件的在IOT中溶液推荐流动。数据记录流经两个不同的路径：

1。 一个快速的过程，档案，并显示收到的消息，并分析这些记录生成短期
短期关键信息和行动，例如报警。

2。 **缓慢处理管线中执行复杂的分析，例如来自多个源和组合数据**

在较长的时间段（例如数小时或数天），并形成新的信息，例如报告，机器学习模型等

Recommended data flow - Lambda architecture



在lambda架构中，快速的数据流由延迟要求的限制，所以有在该分析可能的复杂性的限制。通常，这需要有利于数据和分析的准确度一定程度上是准备尽快的权衡。例如，平均化的功能和趋势分析只能在有限的数据量被执行，典型地在几秒钟的次序。

数据流入慢速路径，在另一方面，是不是受到相同的等待时间要求，并允许在大型数据集的高精确度的计算，这是非常耗费时间。还值得注意的是慢速路径分析的结果可以通过快速路径分析加以利用; 例如，一个解决方案可能需要计算运行的平均收益数据超过一个星期，并提供平均值作为基准数据，快速路径计算。

技术选项

有迹象表明，可以使用和组合构建可靠和可扩展的物联网架构的几个Azure和第三方服务，但是，选择部署该服务时，某些方面应首先考虑：

- **无状态VS状态：** 在可能的情况，溶液应该实现无状态的处理器，以降低运行成本和提高可用性。在另一方面，有状态处理允许更丰富的分析，并实现更高层次的功能通常需要。

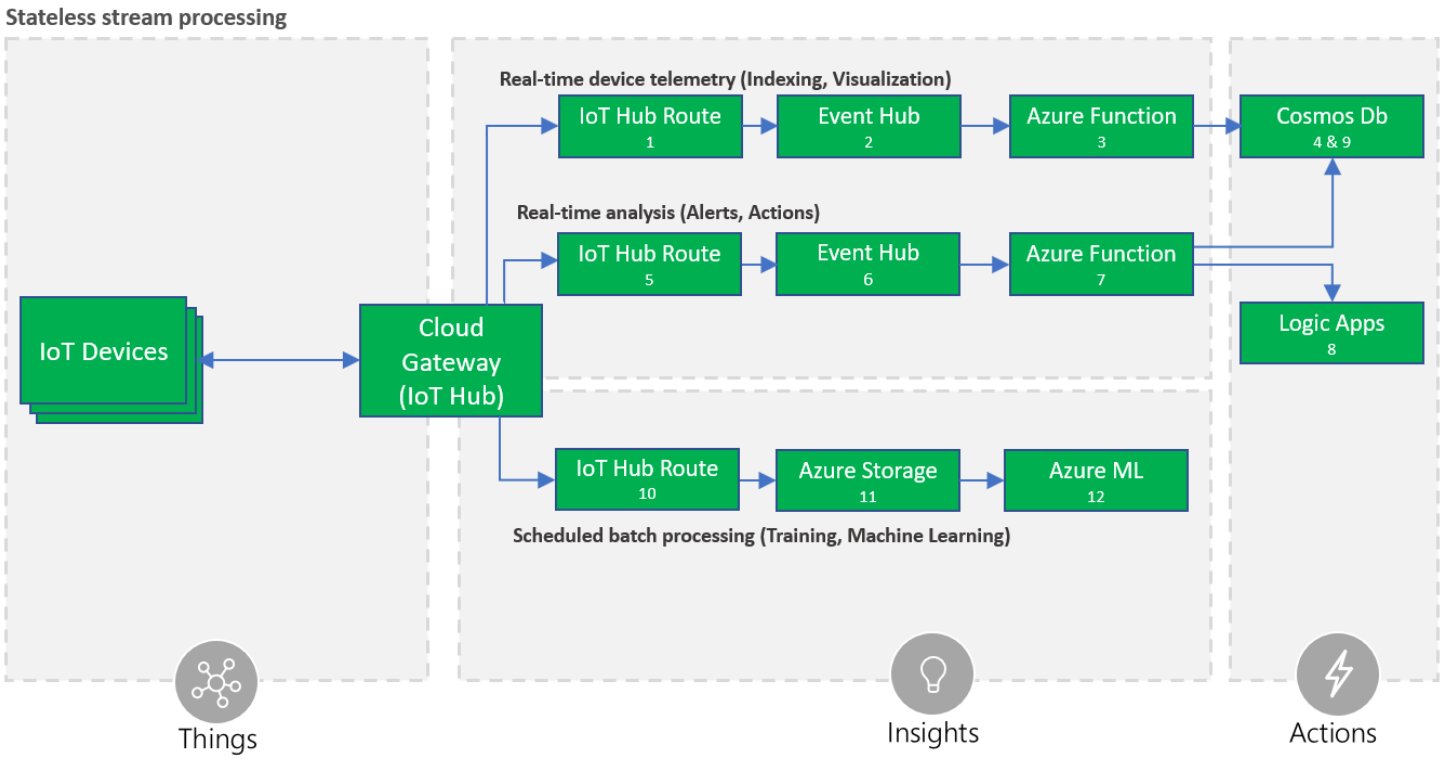
- **静态VS动态规则：** 如果分析规则不改变，不引用改变外部数据，有可能以较低的成本来选择更简单的技术。在需要更大的灵活性，以支持可变负载情况下，经常改变流处理逻辑，和可变的外部基准数据可用的技术是更复杂和昂贵部署。

该文件提出了两个方案，一个解决简单的场景与无国籍处理器和相当静态的规则，一个复杂的情况，例如，用动态分析逻辑和参考数据状态的处理器。

下面的解决方案都带有假设Azure的管理服务，从而提高了整个系统的安全性，并降低安装和维护的成本。在另一方面，解决方案开发人员可以创建异构系统，通过利用其他Azure中提供像Azure的虚拟机，Azure的集装箱服务和Azure的HDInsight结合了专有的，第三方或开源组件，如Spark和卡桑德拉管理服务。

无状态流处理

下面架构提供了用于在需要只无国籍分析，使用小组简单的逻辑规则的摄取数据记录快速和可扩展的实时分析，在场景中的溶液。此外，也有慢的路径，允许更复杂的分析的执行，例如机器学习工作，没有快速路径的速度限制。



该架构被推荐用于其中输入数据记录在JSON序列化方案中，和处理规则采取在输入一个消息的时间，而不考虑历史数据。该架构利用来定义有效载荷条件的能力（＃5）中天青的IoT集线器，以便仅转发特定的消息并触发经由逻辑应用连接服务的动作（＃8）。

一个天青的IoT集线器路线也可用于所有的遥测（＃1）转发到天青功能（＃3），可以将其转化为不同的格式，例如加入外部信息，并将其存储到天青宇宙DB（＃4）暖存储消耗，例如，显示在仪表板上。

另一个Azure的物联网中枢航线（＃10），用于将所有传入的数据记录复制到Azure存储的斑点（＃11），用于冷库，它可以无限期地以低成本存档，并进行批量处理方便，如Azure的机器学习数据的科学任务（＃12）。

该架构的优点：

1。 高可用性由于地理冗余和Azure服务的快速灾难恢复功能。
2。 成本低：大部分组件的自动调整，适应变化的工作量，最大限度地降低成本
每当有没有要处理的数据。
3。 最低的运营成本，因为所有的组件管理Azure服务。
4。 灵活性：Azure的功能和Azure的宇宙DB允许摄入的数据转换到任何需要的模式，
支持多种访问模式和API，如MongoDB的，卡桑德拉和图形的API。
5。 操作和业务集成：集成了多种可供选择通过逻辑应用程序和Azure的ML是可用的。

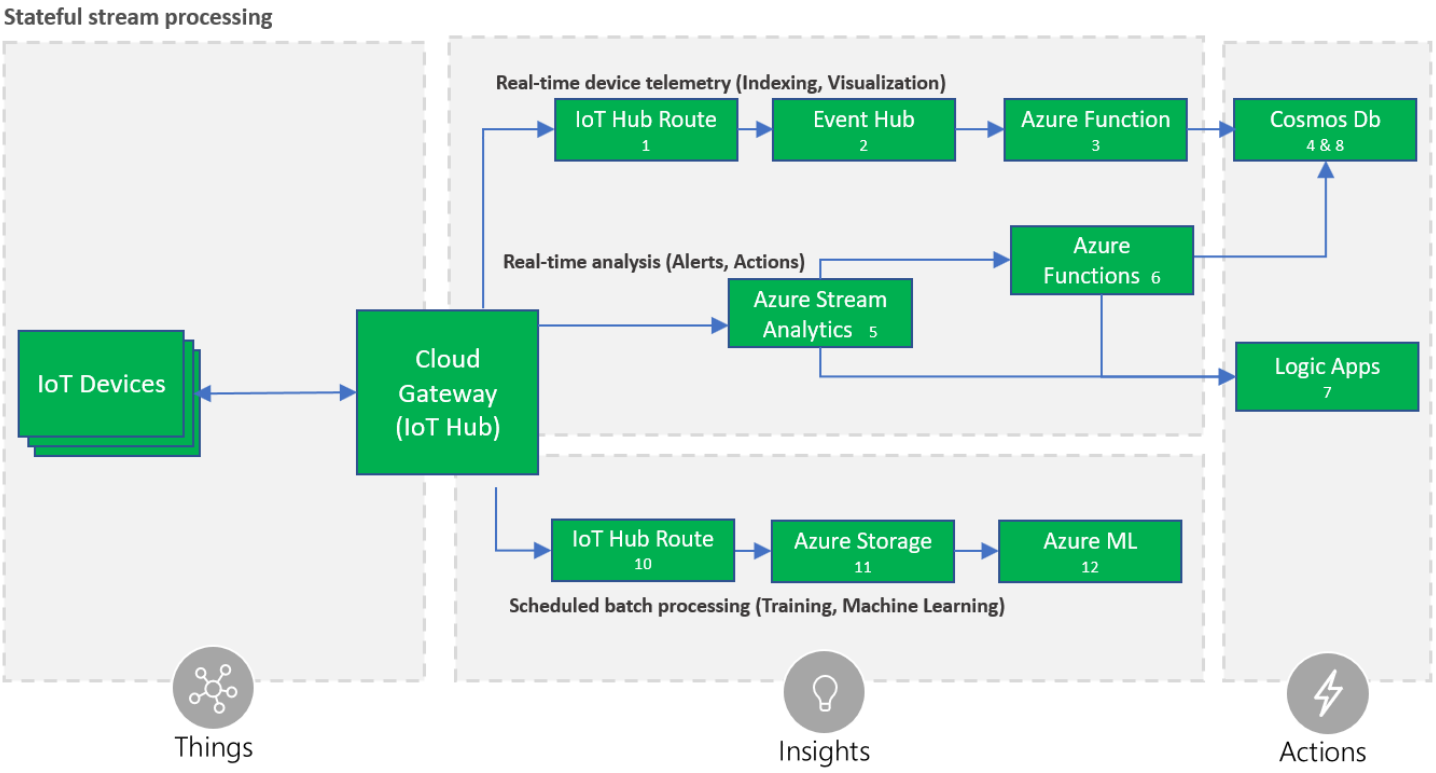
当使用这种架构：

1。 输入数据记录的序列化JSON格式。
2。 需要的规则一个小数目。目前，物联网Azure的集线器支持多达100条路线。
3。 数据记录可以分析一次一个; 即不存在需要聚集在多个数据点数据
（例如平均）或数据流（例如来自多个设备的合并数据）。

有状态流处理

以下体系结构描述了用于以多种格式摄取数据记录状态的实时分析一种快速，灵活和可扩展的解决方案，与参考外部数据，而不先前架构的局限，以更大的运营成本为代价的能力。

该架构包括在之前的架构与机器学习和其他复杂的分析无法在快速路径使用看到了同样的慢速路径。



该架构是类似推荐用于无状态处理仅分析路径中的溶液被替换为天青流分析（ASA）（＃5）。

ASA是专为超大规模的分析和数据记录路由，在有状态的方式，随着时间的推移时段多流应用复杂查询的能力。查询使用类似SQL的语言，**它允许变换和计算定义。服务容忍后期（21天）和乱序（长达一小时）活动，当被处理 *报名时间*[35]，在这种情况下输出因此由时间差延迟。**

ASA也保证 *正好一次交付* 到支持的目的地，很少有记载[36] 异常可能会产生重复。查询语言允许优化的性能分析通过并行，并打破查询到的步骤。

ASA还支持Avro的格式，以减少延迟和带宽成本的紧凑二进制格式的数据记录。

除了ASA表演流处理，在该体系结构一个天青的IoT集线器路由用于所有遥测（＃1）转发到天青功能（＃3），可以将其转化为不同的格式，例如加入的外部信息和存储它至宇宙DB（＃4），用于消耗，例如在仪表板上显示。

一个单独的Azure的物联网中枢航线（＃10），用于复制所有传入的数据记录到Azure存储的斑点（＃11），它可以无限期地以低成本存档，并进行批量处理方便，如Azure的机器学习数据科学任务（＃12）。

该架构的优点：

1. **高可用性由于地理冗余和Azure服务快速灾难恢复功能。**
2. **最低的运营成本，因为所有的组件管理Azure服务。**
3. **Azure的数据流分析，在规模，执行复杂的分析，例如利用翻滚/滑动/跳跃的能力**
   窗，流聚合，以及外部数据源连接。
4. **灵活性：Azure的功能和宇宙DB允许摄入的数据转换到任何需要的模式，**
   支持多种访问模式和API，如MongoDB的，卡桑德拉和图形的API。
5. **操作和业务集成：集成了多种可供选择通过逻辑应用程序和Azure的ML是可用的。**
6. **性能：二进制数据流的支持，以减少等待时间。**

当实现这个架构：

1. 输入数据记录需要复杂的分析，例如时间窗，流聚合，或与外部数据源，这是不可能的无状态架构联接。

2. **处理逻辑由多个规则或逻辑单元，这可能会在时间的增长。**
3. 输入遥测像Avro的二进制格式进行序列化。

3.7 解决方案的用户界面

该解决方案的用户界面（UI）通常包括网站和报告，而且还可以包括网络服务和移动或桌面应用程序。

[35] https://docs.microsoft.com/azure/stream-analytics/stream-analytics-out-of-order-and-late-events

[36] https://msdn.microsoft.com/azure/stream-analytics/reference/event-delivery-guarantees-azure-stream-analytics

该解决方案的用户界面可以提供访问和设备数据和分析结果，通过注册表，命令和控制功能设备的发现和配置的工作流的可视化。在许多情况下，最终用户将收到通知警报，警报条件，或需要通过推送通知采取必要的行动。

该解决方案的用户界面还可以提供或与活和交互式仪表盘，其是用于可视化的IoT场景人口众多设备中的合适的形式集成。

物联网解决方案通常包括地理位置和地理感知服务和用户界面将需要提供适当的控制和功能。

如在本文件的开头所述，安全性是关键的，并且需要与由用户角色分化并且取决于授权访问控制适当地固定在溶液用户界面，在所述系统和设备提供控制。

技术选项

Azure的应用服务是构建Web和移动应用为许多平台和移动设备的强大功能的管理平台。Web应用程序和移动应用程序允许开发者建立网络和使用像.NET，Java和的NodeJS，PHP或者Python语言的移动应用程序。此外，Azure的API允许应用程序的API，可以通过手机或Web客户端进行访问的易暴露和管理。

天青时间序列数据分析（TSI）包括时间序列数据优化的用户界面，包括图表，热图，用于比较的可视化的透视图，和与底层数据相关联的商业智能的统计信息。TSI还提供便于存储在TSI自定义应用程序数据的可视化集成JavaScript控件库。

该TSI JavaScript控件库能够与现有应用程序开发者嵌入图表为他们的物联网数据的可视化。如果用户想创建一个新的应用程序，开始与Azure的物联网解决方案加速器（远程监控或连接厂）是推荐的方法。这些开源参考架构实现使用可视化的TSI JavaScript控件库开箱和缓解需要从头建立一个自定义Web应用程序。注意，这些解决方案加速器是端到端的，物联网解决方案的生产就绪的例子; 即，它们包括用于流处理，存储等的实现

天青通知集线器使得能够发送推送通知到个人移动设备（如智能手机和片剂）。它支持的iOS，安卓，Windows和Kindle的平台，同时提炼了不同的平台通知系统（PNS）的细节。与单个API调用，通知可以针对单独的用户或观众段具有大数量的用户。

除了传统的用户界面，因为它们提供了聚合视图以自然的方式，并帮助可视设备的大量仪表盘在物联网的情况非常重要。电力BI是一个基于云的服务，它提供了一种简单的方法来创建可视化和分析丰富的交互式仪表盘。电力BI还提供现场仪表板，它允许用户监视数据和指标的变化。电力BI包括桌面和移动设备的本地应用程序。

另一种合适的技术物联网的可视化是Azure的地图。[37] **Azure的地图API包括地图控件以及可用于集成Azure的地图在应用程序和网站服务。除**了交互和静态

[37] https://azure.microsoft.com/en-us/services/azure-maps/

地图，所述API提供访问的地理空间的功能，如地址解析，路由和业务数据，而且可以被用来存储空间数据源和具有空间分量的查询数据，如设备的位置。

Web和移动应用程序可以与身份验证和授权控制Azure中的Active Directory（AAD）的集成。这些应用将依赖于用户身份的管理和AAD能够提供应用程序功能的基于角色的访问控制。在许多情况下会出现的IoT的设备和用户之间（或设备的组和用户的组之间的）逻辑关联。例如，一个设备可以由一个人所拥有，由别人使用，并且安装或由另一个用户修复。类似的例子可以为设备和用户组真实的。权限和基于角色的访问控制可以通过AAD管理设备身份（保持在设备标识存储）和用户身份之间的关联矩阵的一部分进行管理。该矩阵的具体设计，权限的粒度，和控制水平将取决于具体的解决方案的要求。该矩阵可以在设备注册表的顶部来实现，或者可以使用不同的技术使用一个单独的存储区。例如，设备注册表可以用宇宙DB来实现，而联想和权限矩阵可以使用关系SQL数据库建立。请注意，由于用户认证和授权浮出水面作为UX的一部分，这个话题在本节讨论; 然而，实际的实现将在多个基础组件被传播，包括设备注册表和应用程序的后端，在接下来的部分讨论。设备注册表可以用宇宙DB来实现，而联想和权限矩阵可以使用关系SQL数据库建立。请注意，由于用户认证和授权浮出水面作为UX的一部分，这个话题在本节讨论; 然而，实际的实现将在多个基础组件被传播，包括设备注册表和应用程序的后端，在接下来的部分讨论。设备注册表可以用宇宙DB来实现，而联想和权限矩阵可以使用关系SQL数据库建立。请注意，由于用户认证和授权浮出水面作为UX的一部分，这个话题在本节讨论; 然而，实际的实现将在多个基础组件被传播，包括设备注册表和应用程序的后端，在接下来的部分讨论。

3.8 监控和日志记录

物联网解决方案记录和监控系统用于确定预期的解决方案是否能正常工作，并帮助解决什么是错的解决方案。监测和回答以下操作性的问题记录系统的援助：

- 处于错误状态的设备或系统？
- 是设备或系统配置是否正确？
- 正在生成准确的数据，设备或系统？
- 被系统满足商务和终端客户的期望是什么？

监视和记录系统帮助回答这些问题，当回答是“不”，他们表面的运营团队，以帮助减轻问题的相关信息。

物联网解决方案记录和监控系统往往比那些业务线的标准应用更加复杂。在复杂的事实是出现物联网解决方案涵盖：

- 物理传感器与环境交互。
- 在智能边缘应用执行数据整形，协议转换等。
- 基础设施组件，诸如预置网关，防火墙和开关。
- 食入和短信服务。
- 持久性机制。
- 洞察和报告应用程序。
- 在云中运行，扩展独立子系统。

的溶液的IoT复杂进一步组件源于这样的事实，有一组不同的利益相关者，其中包括：

内部利益相关者
IT和运营安全团队
现场技术人员与服务人员应用程序开发者的数
据科学家
业务团队与高管

外部利益相关者
供应商与合作伙伴的客户
合规和审计专家

一个监测和记录解决方案可以包括众多专业软件应用程序和库针对物联网解决方案的每个子系统。日志记录和监控工具通常由以下四个部分组成：

- 系统性能和时间表可视化工具 - 用于监视系统和基本故障排除。
- 缓冲的数据摄取-to缓冲器日志数据（其可以是详细）。
- 持久性存储-to存储日志数据。
- 搜索和查询功能-to在详细的故障排除使用日志数据。

大规模的物联网解决方案，可以由许多较小的子系统。它往往是合理部署记录的多个实例和监视组件为每个这些系统中，具有较高的电平实例的聚集数据，并从下级系统的分析。例如，在远程监控溶液加速器多个子系统噪比（IoT集线器，波斯菊Db的，天青流分析，定制微服务，等等）用于为设备的IoT提供操作者的能力。日志记录子系统是在个体水平完成，并且然后可以被聚合以提供一个端到端的视图中的溶液。

弹性和冗余必须考虑设计记录和监测系统时。增加记录和监测性能可以通过共同定位系统旁边的溶液来获得; 然而，该战略带有系统性中断的**风险**[38] **并造成负面影响的物联网核心解决方案本身的性能和扩展日志记录/监控系统的风险。被监测系统"失败的风险"的早期评估将确保强大的解决方案的发展。**[39]

监控和可视化

监控系统提供分析上市公司健康，安全与稳定，以及物联网解决方案的性能。在较高的水平，监控系统提供如预期的端到端解决方案是否能正常工作的快速视图。监控系统也可以提供一个更详细的视图，记录组件配置更改和提供能够表面潜在的安全漏洞提取记录数据，提高事件管理过程[40]，

并帮助系统的所有者解决问题。综合监控解决方案包括以查询特定的子系统或多个子系统的聚合信息的能力。

监控系统的发展应该通过定义健康运作，合规性和审计要求开始。收集到的指标可能包括：

[38] https://medium.com/@adrianco/who-monitors-the-monitoring-systems-715a333f97fc
[39] https://turbonomic.com/blog/on-technology/thinking-like-an-architect-understanding-failure-domains/
[40] https://www.axelos.com/news/blogs/september-2014/service-monitoring-strategic-opportunity

- 物理设备，边缘设备和基础设施组件报告配置的变化; 例如开放的网络端口，贴剂，服务和用户（审计＆顺应性），与一般的操作参数，诸如功耗，CPU，存储器，和磁盘使用沿。

- 应用程序报告配置更改，安全审计日志，请求速率，响应时间，错误率和垃圾收集统计信息管理的语言。

- 数据库，持久性存储和缓存报告查询和写入性能，架构更改，安全审计日志，锁或死锁，指数的表现，CPU，内存和磁盘使用情况。

- 管理服务（IaaS的，PaaS的，SaaS和FAAS）汇报健康指标和配置更改的影响取决于系统运行状况和性能。

应注意平衡收集和储存指标与提供的见解价值的性能开销。

收集指标的策略往往是由特定的物联网解决方案，增强的安全环境复杂。设备上的应用程序可以从主机的操作系统，它通过抑制部署到它们的应用程序收集装置的指标，需要额外的软件解决方案，以方便收集来沙盒。额外的系统安全性设计压力从审计和合规性要求出现。的SaaS，PaaS和IaaS的监控组件往往与认证[41] 这将缩小审计范围，以满足法规遵从要求。

监测指标警报利益相关者系统的不稳定性和促进事件响应的可视化。可视化应该是定制利益相关者的角色，并提供可扩展性，以适应溶液生长和成熟。呈现给操作者的数据应限于高冲击，可以被链接到系统状态的变化可操作度量，同时提供的能力深入了解具体问题需要。可视化时间表应该呈现相关数据，诸如部署，配置改变或先前事件。该相关数据的添加范围内提高决策和加快响应时间。

跟踪遥测

跟踪遥测允许运营商跟进系统一片遥测从创建之旅。跟踪是用于调试和故障排除重要。对于物联网解决方案，利用物联网Azure的枢纽，并在Azure SDK的设备，追踪数据报可起源，特设，如 *云到设备*[42] 消息和包括在遥测流。追踪ID和邮件标志允许跟踪数据报流过系统的处理链，同时提供沿途操作见解。

在只有差动遥测（例如，仅制冷温度的变化）被发送到摄取服务的IoT的解决方案，用于跟踪消息类似的方法也可被利用来建立心跳电路[43]。心跳电路用于确保低短信费用设备仍然活跃，不经常使用的通讯联系仍然活着，并表明在长时间运行的任务向前进步。心跳遥测可利用的建筑性能可视化，服务水平协议（SLA）报告，并且可以折叠成一个物联网解决方案的监测工具链。

41 https://azure.microsoft.com/overview/trusted-cloud

42 从集线器的IoT发送云到设备的消息（ https://docs.microsoft.com/azure/iot-hub/iot-hub-devguide-messages-c2d ）

43 https://docs.microsoft.com/system-center/scom/manage-agent-heartbeat-overview

记录

记录系统是了解什么行为或活动的解决方案已经完成，已发生的故障，并且可以在固定这些故障提供帮助不可或缺。日志可以进行分析，以帮助了解和纠正错误条件，提升性能，并确保遵守规则治理和法规。

基于文本和结构化的日志信息的两种主要模式。[44] 基于文本的日志记录应用程序的开发过程中通常实施了调试的目的，并为开发者提供的应用程序行为的叙事风格记录。结构化记录中加入情景语境和元数据与意图日志将被解析的叙述日志信息建立在基于文本的方法。在结构化记录方法，属性成为一等公民格式为键/值对，或者用一个固定的模式，以提高搜索和查询功能。

文本日志例如：

2017-12-18T08：15：30 [INFO] - ClientAuthed，偏差/ opcpub，35f634d7 2017-12-18T08：15：30 [INFO] - 对设备dev新设备连接/ opcpub 2017-12-18T08：15：30 [信息] - 对设备dev绑定设备代理/ opcpub 2017-12-18T08：15：30 [INFO] - 绑定消息信道设备dev / opcpub

结构化日志示例：

{“@t”：” 2017-12-18T08：15：30”，” @升”：”信息”，‘设备ID’：” opcpub”，” @米”：”。ClientAuthed，35f634d7”} {“@吨”：” 2017-12-18T08：15：30”，” @升”：”信息”，‘设备ID’：” opcpub”，” @ M”：”新设备连接”} {‘@t’： “2017-12-18T08：15：30”，” @升”：”信息，‘设备ID’：” opcpub”，” @ M”：”绑定设备代理”} {‘@t。’：” 2017-12 -18T08：15：30”，” @升”：”信息”，‘设备ID’：” opcpub”，” @ M”：”结合消息信道。 “}

虽然纯文本的日志记录是在前期开发成本较低的影响，这是更具挑战性的一台机器来解析/读取。我们建议构建日志被使用，收集的信息既机器可解析和人类可读。

评价标准监测

监控和可视化解决方案具有以下标准进行评价。虽然这些标准可能并不适用于所有的物联网解决方案，他们被选为最普遍适用。

1。 **完整的解决方案：解决方案提供了缓冲的数据摄入，很容易地集成数据持久化，搜索和查询**
能力，以及一套易于浏览的可视化。监控解决方案应该提供应用性能监控（APM）功能，一个报警和通知规则系统，可扩展性的问题跟踪系统集成，以及潜在的反馈回路，使解决方案自动缩放。

2。 **安全性：监测和相关的可视化系统提供了成熟和强大的安全选项，如**
传输层安全（TLS）的通信，资源集成的授权和认证，并在休息的数据加密。

3。 **可视化：解决方案提供了一套预建的可视化，便于定制的可视化，并提供建设**
搜索和数据的查询。到底层数据源和生成仪表板访问被授权和认证机制的保护。

[44] https://www.thoughtworks.com/radar/techniques/structured-logging

4。 **配置和机器监控：解决方案必须扫描和监测港口机械变化的能力，**
包装，服务，用户等应用程序或服务级别配置监测和检查整个部署或机器集群的一致性。该解决方案在Windows和Linux操作系统system s.45工作

5。 **提醒：解决方案具有集成的规则引擎映射到一个集成的报警和通知系统**
或提供易于配置挂钩第三方通知和售票系统。

6。 **合规性：解决方案应该满足合规性规范要求审核部署。**

7。 **服务模式：该解决方案可供自托管或与SLA SaaS解决方案。**

8。 **规模：此外，监控系统应该很容易与物联网解决方案的增长规模。**

技术选择

下面的监控和可视化解决方案进行了评估。运营管理套件是Azure的物联网解决方案推荐的解决方案。

运营管理套件（推荐）

运营管理套件（OMS）是推荐的监控和可视化解决方案。OMS提供物联网解决方案的运营管理工具链（云和内部部署）。机日志，服务和应用**程序，不管部署位置，可以直接推日志消息到采用OMS内置连接OMS或通过其数据收集器API。**[46] **OMS提供定制的日志分析，以促进事件和记录到的索引和搜索各个字段的分解。**[47]

监控日志搜索通过OMS日志搜索功能，或者定制的分析服务，它与OMS集成，并提供了先进的可视化选项可用。

警报是通过警报规则的支持。警报使用监视事件在指定的时间范围内发生的查询所定义。的阈值，频率选项和警报抑制可避免打草惊蛇洪水。**警报引擎透过WebHook支持集成：这使OMS与通讯解决方案，如微软的团队或松弛发送消息。警报仪表板也可提供活动警报的总体视图。**[48]

OMS数据保留策略是可配置的，并在日志分析服务执行的所有活动将被记录并可以进行审核。[49]

弹性栈（Elasticsearch，Logstash，Kibana和节拍）

共同地，弹性堆栈促进摄取，解析和数据的可视化。弹性堆叠可通过部署选项多种多样，包括自托管，在公共云或私有云，或者通过Azure的**市场。弹性堆栈具有广泛的社区，积极促进双方的核心部件和附加解决方案高度可定制的。弹性提供了成功案例的不同行业和使用案例库**[50]。

建立在成熟的Apache Lucene的图书馆，Elasticsearch组件提供了所收集的数据搜索引擎的能力和分析。它是专为性能和规模，分发跨节点碎片和自动地检测故障来重新平衡数据。一个与Elasticsearch的挑战之一是查询在数据

45 https://www.upguard.com/product/core#Features
46 https://docs.microsoft.com/azure/log-analytics/log-analytics-data-collector-api
47 https://docs.microsoft.com/azure/log-analytics/log-analytics-data-sources-custom-logs
48 https://docs.microsoft.com/en-us/azure/log-analytics/log-analytics-alerts
49 https://docs.microsoft.com/azure/log-analytics/log-analytics-data-security
50 https://www.elastic.co/use-cases

索引源，所以它被索引重要的是要加入数据[51]。经由Logstash摄取和原木的解析将进一步在下面的日志记录部分进行讨论。

Kibana是一个数据可视化工具，它支持基本图表和图形，地理空间特征，以及时间序列。仪表板可以嵌入在应用程序中，共享或导出为各种格式。

正如监测旨在断言系统的健康，一定要知道的监控解决方案本身保持健康是非常重要的[52]。弹性堆叠X-Pack附件为核心，以解决这个问题，以及基于角色的访问控制下降到文档字段级别，传输加密，集成的审计日志，警报和自动报告分布。X-包，可以将数据发送到一个单独的集群，使监控时可用生产节点都失败了。

Splunk的

Splunk的是通用的，实时的数据分析平台具有强大的可扩展性模型，在这两个提供托管和自托管的配置。Splunk的支持的输入数据格式的阵列[53]并具有使用无代理扫描或与通用代理来收集机遥测和日志的能力[54]剂。摄入的原始数据转化为时间序列的事件和随后编入索引的搜索和查询。Splunk的的通用信息模型（CIM）[55]有助于标准化数据以匹配共享语义模型，其作为在索引数据的检索时间架构。Splunk的的搜索处理语言（SPL）[56]结合SQL数据管道能力查询，丰富和变换（成CIM）索引的数据。SPL促进数据可视化，机器学习和异常检测和趋势分析。

Splunk的一般部署审计和法规遵从解决方案，因为它支持广泛的监测和报告工具，用于金融服务，医疗保健和公共部门组织的。Splunk的还通过集成的预测性维护，定制的警报和通知，以及设备性能管理有利于物联网的工作负载。

比较网格

		运营管理套件（OMS）	弹性堆栈	Splunk的（托管）
解决方案组件	食入缓冲	是	是	是
	集成的持久性	是	是	是
	搜索和查询	是	是	是
	可视化	是	是	是
安全	系统更新评定	是	没有	是
	反恶意软件	是	可能	是
	加密通信	是	可能	是
	数据加密在休息	是	通过加密文件系统	是

51 https://www.opsview.com/resources/elasticsearch/blog/using-elk-stack-business-intelligence
52 https://medium.com/@adrianco/who-monitors-the-monitoring-systems-715a333f97fc
53 http://dev.splunk.com/view/dev-guide/SP-CAAAE3A
54 https://docs.splunk.com/Documentation/Forwarder/7.1.0/Forwarder/Abouttheuniversalforwarder
55 http://docs.splunk.com/Documentation/CIM/4.10.0/User/Overview
56 https://docs.splunk.com/Documentation/Splunk/7.1.0/SearchReference/UnderstandingSPLsyntax

可视化	交错系统活动	是	是	是
	热映射	定制	是	是
	数据体积测量&通过综合查询切片语言	是	是	是
配置：机监控	远程计算机扫描	是	手册	是
	更改跟踪	是	手册	是
	事件推送数据关联	是	是	是
警报	集成的通知	是	是	是
	第三方扩展	是	是	是
服务模式	可靠性/ SLA	是	没有	是
	SaaS的	是	可选的	是
	自托管	没有	是	是
	可扩展性	是	是	是

记录

记录解决方案具有以下标准进行评价。虽然这些标准可能并不适用于所有的物联网解决方案，他们被选为最普遍适用。

1。 **安全：** 日志记录系统提供成熟和强大的安全选项，如TLS通信，资源集成的授权和认证，并在休息的数据加密。

2。 **编程模型：** 日志记录系统提供进出流程运作模式的，采用的是结构化的记录方式，已经有据可查的，易于使用的API或库。

3。 **存储和持久性：** 记录系统提供了灵活性以利用其能够在运行时被改变的多个持久存储/输出机制。这种灵活性满足了应用开发的不同阶段记录的需求：开发，质量控制，整合，分期和生产。该解决方案还应该提供对自动到期或推到冷库旧的日志信息的能力。

4。 **搜索和查询：** 日志系统支持搜索和查询机制，通过原产地或时间片通过日志信息筛选。

5。 **缓冲食入：** 日志记录系统提供了异步日志提交和缓冲层摄入。

6。 **可视化：** 记录系统提供日志记录度量的可视化的有限量的诸如摄取，储存消耗等可视化率不是必需的组件。

7。 **服务模式：** 该解决方案可供自托管或与SLA的SaaS。

8。 日志记录系统应与物联网解决方案的增长规模。

技术选择

下面记录的解决方案进行评估。应用见解和Serilog是Azure的物联网解决方案建议的解决方法。

Serilog / SerilogJ（推荐作为在应用日志框架）

Serilog（J）是用于与聚焦结构化测井的IoT的解决方案所推荐的诊断日志库。Serilog具有广泛的开源社区的支持，很容易配置，并支持广泛的**“汇”，它允许开发人员能够轻松调整如何以及在何处登录信息发送的范围。[57] Serilog包括日志数据富集，从而开发者可通过在横切的信息，如环境或处理信息的折叠延伸日志事件的概念。[58] Serilog可以用作在任何子系统，用于自定义代码的碱性记录解决方案。**

应用洞察（推荐云的日志和指标数据采集和持久性）

应用见解是一个可扩展的应用性能监控（APM）的服务。[59] 应用程序的见解是推荐的APM解决方案，它提供了集成的异常检测，简化了应用程序的性能和指标收集，并提供带的DevOps和SDLC（软件开发生命周期）工具链集成。开发人员可以使用仪器的Serilog应用洞察水槽他们的软件。

需要注意的是作为本文件的发布，应用见解还没有在休息热路径数据提供数据加密，但所有的冷数据被加密存储是非常重要的。

log4j的/ log4net的

Apache软件基金会的log4j™（此后，LOG4 *）日志框架是一种流行的Java日志库及其架构已经被移植到12种不同的语言，包括基于.NET的运**行时间。[60] 基于文本的日志框架实现分层记录器，这便于日志语句控制的粒度的概念。该LOG4 *架构非常适用于基于组件的应用开发[61] 和可扩展**的配置功能支持在运行时记录粒度的修改。

弹性栈（Elasticsearch，Logstash Kibana，和节拍）

弹性堆栈是四个OSS产品由工具的集合：

- Elasticsearch，分布式搜索和分析engine62
- Logstash，用于收集，分析和存储logs63的工具
- Kibana，用于记录，时间序列和APM使用cases64数据可视化和探索工具
- 节拍，剂，其发送度量，日志，心跳，包或其他自定义数据转换成Elasticsearch或Logstash65

经由Elasticsearch和Kibana搜索，分析和可视化数据监视部分中先前所讨论的。Logstash和节拍是支承传送数据的到弹性系统的弹性堆栈的组件。

Logstash是用于数据的摄取和处理管线，支持在一次至少-经由永久队列递送。数据可以来自任何数目的来源将被变换并存储在任何数量的接收器。有提供的滤光器的一个库，可以在转换管道被利用来执行任务，如：解析是结构

57 https://github.com/serilog/serilog/wiki/Provided-Sinks
58 https://github.com/serilog/serilog/wiki/Enrichment
59 https://docs.microsoft.com/azure/application-insights/app-insights-overview
60 马尔斯，汤姆和戴维斯，斯科特。“JBoss的工作中。”奥赖利，2005年“附录B：记录和JBoss”，pp.254 61 https://logging.apache.org/log4net/release/features.html#Hierarchical-logging-architecture

62 https://www.elastic.co/products/elasticsearch
63 https://www.elastic.co/products/logstash
64 https://aws.amazon.com/elasticsearch-service/kibana
65 https://www.elastic.co/products/beats

或非结构化数据，汇总事件，发生变异的字段，以及节流阀事件[66]。超过200个插件可用来创建自定义的管道，如果一个插件缺失，它可以建立和贡献[67]。

节拍是最新加入到弹性堆栈[68]。每个单用途代理跨生产节点上运行执行轻型收集和相关数据的运输到Elasticsearch或Logstash。弹性提供了节拍为日志文件，指标，网络数据，事件日志，审计数据和心跳。还有是社会的资源库节拍，开发者可以继续作出贡献[69]。

而弹性栈拥有自己的一套全面的数据摄取的API，该工具集可与增强工具链性能，提高开发人员的生产率汇Serilog（J）或log4j的（净）配对。自托管的弹性堆栈部署将需要运营，维护和安全性托管解决方案不会追加投资。

比较网格

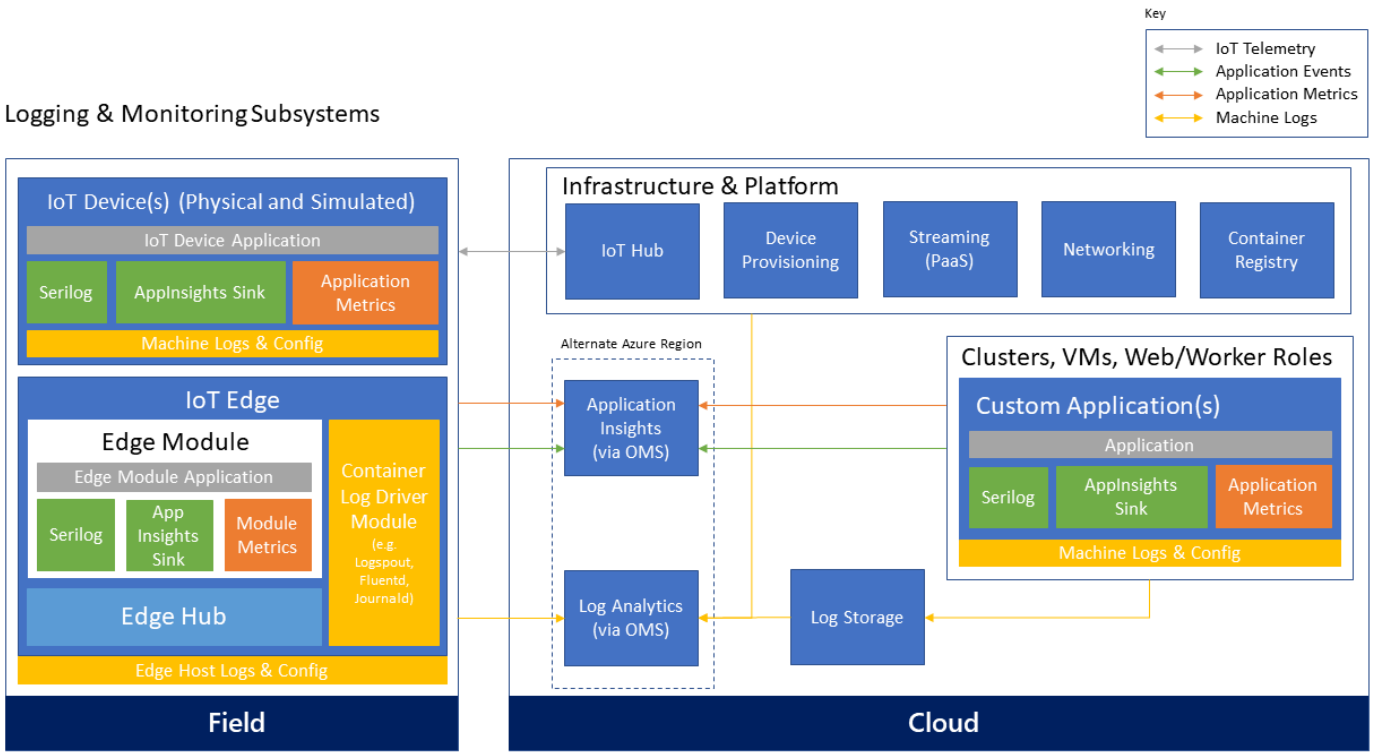
		Serilog / SerilogJ	log4j的/ log4net的	AppInsights	弹性堆栈
安全	安全通讯			是	可选的
	加密在休息	通过加密文件系统	通过加密文件系统	Q2 / 2018	通过加密文件系统
	集成的授权认证			是	没有
编程模型	输入/输出过程的	在过程	在过程	在过程	API - 在过程
	基于文本	没有	是	没有	没有
	结构化的	是	解决方法	是	是
存储和持久性	基于文件	是	是	没有	没有
	相关的	可选水槽	可选水槽	没有	没有
	柱状	可选水槽	可选水槽	是	没有
	核心价值	可选水槽	可选水槽	没有	没有
	文献	可选水槽	可选水槽	没有	是
	冷存储/归档	手动W /滚动文件的Appender	手动W /滚动文件的Appender	连续出口[70]	是[71]
食入缓冲	排队	在记忆中	在记忆中	集成	集成
	节流			32K事件/秒平均超过1分钟[72]	索引合并限制
	可扩展的服务层			是	在SaaS模式

66 https://www.elastic.co/guide/en/logstash/current/filter-plugins.html
67 https://www.elastic.co/guide/en/logstash/current/contributing-to-logstash.html
68 https://www.elastic.co/elk-stack
69 https://www.elastic.co/guide/en/beats/libbeat/current/community-beats.html
70 https://docs.microsoft.com/azure/application-insights/app-insights-data-retention-privacy#how-long-is-the-data-kept
71 https://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html
72 https://docs.microsoft.com/azure/application-insights/app-insights-pricing#data-rate

可视化数据关联				是	是
	定制的仪表板和下钻			是	是
服务模式	可靠性/ SLA			是	是（SaaS的部署）
	SaaS的			是	可选的
	自托管			没有	是
	定价模型			消费	根据计算与消费

日志记录和监控体系

以下简化记录和监视体系结构示出了典型的IoT溶液组分以及它们如何利用上面详述的推荐技术的示例。



3.9 业务系统集成和后端应用程序处理

物联网的应用可以与整个组织的技术系统集成。设备，设备，业务规则和操作，以及访问和关联设备和用户之间的组进行控制。应用程序后端的重要部分是该溶液中，设备发现和可视化，设备状态管理和命令执行的“自定义”的控制逻辑，以及设备管理，其控制设备的生命周期，使得能够配置和软件更新的分布，并允许设备的远程控制。业务整合往往是从后端处理系统驱动; 即整合物联网环境下的成下游业务系统，如CRM，ERP和线路的业务（LOB）应用程序。

应用后端处理

不同于传统的业务系统，物联网的解决方案的业务逻辑可能在整个系统中的不同组件传播。溶液设备管理将通常使用的计算节点，而该溶液的分析法部分将在很大程度上直接相应的分析系统内实现。

在某些情况下，简单的解决方案可能没有独立部署和管理的“商业逻辑”的应用后端，但核心逻辑可以作为流处理承载的规则表达式，某些分析能力，和/或作为业务流程的一部分，连接器组件。

技术选项

有针对后端逻辑的几个实施方案。如上所述，一些逻辑将在系统的事件处理器和分析组件来实现。对于那些部件实现选择在各自的部分都淹没了。本节特别侧重于业务逻辑的后端。

编程不支持超大规模的技术。 **许多建筑模式和编程技术，已经流行了几十年的适用于物联网解决方案，但可能会在大量设备的可扩展性面临挑战。**因此，对于大型物联网的部署，这些模型只应与可扩展的计算节点上运行的应用程序无国籍后端使用。向外扩展一个应用层的状态代表了传统建筑的一个难题。在这些情况下，适当的规模计算模型，例如演员框架或批量处理可以被使用，如在接下来的章节中描述。

演员框架。 **演员框架代表了物联网的场景一个非常适合的计算模型。演员编程模型非常适合那里有大量的与特定行为和独立的本地数据/状态“独立”的单位。**演员框架为需要与后端服务通信设备良好的抽象模型。A（物理的）设备可以被建模为具有确定的行为和局部状态的演员，将在后端上运行。的演员成为物理设备的虚拟表示。一个演员可以代表管理自己的状态有状态计算单元。不同于传统的编程技术，在其中创建对象的实例，国家需要从外部加载的，有状态的演员有立即的内在状态。用1：一种装置和后端“代码之间1的关系，

此外，演员模型提供一种方式来创建角色的层次结构，以代表的设备或设备组之间的关系。例如，它很容易在建筑物作为演员的层次结构中所有传感器型号：建筑物可以是由一组地面演员的演员，地板演员被定义为一组房间的演员，每个房间的演员可以控制一组在室内的传感器。这样一来，很容易编写复杂的规则和逻辑进行迭代演员的层次结构。层次结构中的每一个元件提供作为或上级集合信息所需要的正确的行为和状态。

一个演员可以从设备处理消息，执行计算，并且当满足某些条件在后端发送命令或通知给设备。从一个抽象的角度来看，开发者可以专注于需要管理一个设备，这导致简单的编程模型的代码。大多数演员的框架使用一个基于消息的架构，以及通信和演员之间是由框架管理。当一个或多个消息是可用的并且需要被处理的行动者仅调用; 也就是说，演员由框架激活有要完成的工作时。有没有必要有任何“辅助角色”类型的组件，需要活下去，以检查是否有要完成的工作架构。演员框架调度负责调度演员与优化资源利用率的目标执行。在该架构的上下文中，一个

当接收到的设备事件，或从后端，基于从业务逻辑和规则，或业务线对的一系统即将发生的事件的演员可以被激活。

有几个可用的演员框架和开发人员可以选择最适合自己的编程背景和场景的要求之一。以下几段介绍三种流行的男主角框架：Azure的服务织物可靠的演员，阿卡和Akka.NET。

Azure的服务织物可靠的演员

Azure的服务织物[73]使开发人员能够构建和管理的密度非常高的机器上的共享池中运行的微服务组成的可扩展和可靠的应用程序，通常被称为一个服务织物集群。它提供了一种构建分布式，可扩展性，无界和有微服务和全面的应用管理功能，配置，部署，监控，升级/补丁，并删除部署的应用程序复杂的运行环境。在服务织物状态的服务提供具有可以直接通过该服务依赖外部工具，如高速缓存系统或存储中使用，而不需要完全复制本地数据的好处。

服务织物提供了可靠的演员编程模型。它是使用服务织物运行基础设施的实力提供了一个可扩展的和可靠的模型开发人员使用面向对象的编程背景会发现很熟悉的基于角色的编程模型。可靠的演员编程模型非常相似，黄蜂，和开发人员所熟悉的黄蜂可以轻松地迁移到可靠的演员，也可以继续使用运行时奥尔良。

除了可靠的演员，服务织物还提供了一个低级编程模型，可靠的服务[74]具有并发，分区和通信方面的简单性和灵活性之间不同的权衡[75]。

在这种模式下可靠的收藏[76]可以用来存储和管理设备的状态。

阿卡

阿卡[77]是一个Java虚拟机（JVM）上运行的知名演员的编程模型。它使用Scala编程语言开发的，但提供的Java API为好。基于阿卡，后端应用程序可以在Azure中托管，并且可以使用Azure的物联网服务，同时实现对那些已经使用Java或斯卡拉作为他们的首选语言的开发人员熟悉的编程模型。

Akka.NET

Akka.NET[78]是阿卡编程模型到.NET运行时的一个端口并支持C＃和F＃。随着阿卡，它为开发者使用阿卡编程模型，但在.NET运行时上运行的代码的方式。

Azure的批

批处理是非常适合于那些需要运行的自动化任务大量的工作负载，如执行定期（如每月或每季度）处理，或风险计算。Azure的批[79]是云规模的作业调度和计算管理服务，使用户能够运行高度并行计算负载。在Azure批次调度器可用于调度和监视整个大型计算集群的工作负荷的执行。它需要开始计算虚拟机池，安装处理作业和分段数据，运行作业的照顾，

[73] https://azure.microsoft.com/services/service-fabric
[74] https://azure.microsoft.com/documentation/articles/service-fabric-reliable-services-introduction
[75] https://azure.microsoft.com/documentation/articles/service-fabric-choose-framework
[76] https://azure.microsoft.com/documentation/articles/service-fabric-reliable-services-reliable-collections
[77] http://akka.io
[78] http://getakka.net
[79] https://azure.microsoft.com/services/batch

识别故障，并根据需要重新排队工作。它还能自动按比例缩小的资源作为工作完成了游泳池。

业务系统的集成

业务集成层负责一体化的IoT环境进入下游业务系统，诸如CRM，ERP，和线的行业（LOB）应用程序。典型的例子包括业务计费，客户支持，经销商和服务站，零件供应，第三方数据源，运营商配置文件和转移计划，时间和作业跟踪，等等。

物联网解决方案关系进入业务线，现有的通过业务连接器或EAI / B2B网关功能的应用程序和标准的软件解决方案。在B2B或B2C场景最终用户将通过该层的设备数据和特殊用途的IoT设备进行交互。在许多情况下，终端用户会使用个人移动设备访问该功能。那些个人移动设备比的IoT设备概念上不同的，尽管在某些情况下会出现在最终用户的移动设备和的IoT设备之间的关联或映射。例如，在一个家庭自动化的情况下，手机可能会充当网关领域，连接到物联网设备和促进这些通信。从授权透视端用户，个人移动设备之间的关联，

技术选项

Azure的逻辑应用程序提供了一个可靠的方法来实现业务流程自动化。横跨在Azure中托管的不同的系统的服务支持长时间运行流程编排，本地，或在第三方云。逻辑应用程序允许用户通过一个易于使用的可视化设计自动化业务流程执行和工作流程。该工作流程从触发启动并执行一系列步骤，每个调用接口或API，同时服用认证，检查点，耐用执行的照顾。有一个非常丰富的可用连接器的多项第一方和第三方系统，如数据库，消息传递，存储，ERP和CRM系统，以及支持EAI和EDI服务和先进的集成能力。

对于API集成，Azure的API管理提供了揭露和管理的API，包括终端到终端的管理功能，如一个全面的平台：安全和保护，使用计划和配额，用于将有效载荷，以及分析，监控策略，并警报。

集成在数据层可以通过天青数据工厂，这可提供用于构建数据管道用于转化和数据的移动的协调层启用。数据工厂可以跨本地和云环境中阅读，转换和发布数据。它允许用户以可视化的数据管道和监测数据管道健康之间的谱系和依赖。

3.10 机器学习（静止数据分析）

静止数据分析是在所收集的设备的遥测数据进行的，并且通常该数据被混合与其他企业数据，或从其它系统或组织遥测的次级源。分析和预测设备的操作数据和行为，基于与环境参数和遥测设备相关遥测，是一个强大的图案。

还有的情景时，为什么，以及如何后处于静止来分析数据的显著数量，而这个参考架构文档不打算提供这些选项或静止数据分析的深入的解释。物联网环境下，对于这些功能的通用指导直接应用于物联网解决方案，但并不局限于此。先进的分析和大数据解决方案可在这些情况下使用。

技术选项

对于数据的科学家与算法的基础上认识，Azure的机器学习提供了一个托管的机器学习能力。它提供了方便与直接集成运用到使用生成的Web服务接口的解决方案。

随着HDInsight，Azure平台提供了一个托管实现了Apache的Hadoop[80] 平台，提供的Apache蜂巢，[81] 阿帕奇亨利马乌，[82] MapReduce的，[83] 猪，[84] 和Apache风暴[85] 作为分析的能力。

双向电力支持创建的模型，关键绩效指标，并通过交互式仪表盘的可视化。它提供了监控的过程或操作的性能强大的分析解决方案，可以帮助识别趋势和发现有价值的见解。

其他选项包括Apache的火花，可用于运行大数据的工作，而且还提供了图形分析和机器学习模块。

4.解决方案设计注意事项

智能设备

连接的设备的目标是形成智能系统。一个关键问题是智能设备应该如何与该系统作为一个整体应该如何聪明是。答案将根据具体设备的目的，设计，可用计算资源和权力是不同的; 然而，有一组常见的取舍物联网系统的考虑。

设计一个装置发生在其生命周期的开始。在这个阶段设计的失败几乎是不可能的或非常昂贵的设备制造完毕后再纠正，虽然有些设备的行为可以通过固件/软件更新或通过配置改变而改变。软件的变化是不是改变或更换硬件，因此对于设备设计远程软件更新功能更容易是有帮助的。

即使设备具有软件更新功能，管理的数百万潜在边缘组件的更新比更新的集中解决方案后端复杂得多。在一般情况下，在边缘更智慧等同于在更高的频率边缘分量的可能更多的软件更新，同时对解决方案的后端更智能意味着可以以集中的方式进行维护。毫无疑问，其在后端更智能将最有可能增加边缘组件的相关性，但如果设计适当，他们应该能够即使没有到后台在线连接自主进行。

无论设备的功能能力，集中的软件操作的安全性在后端通常允许整个系统（尤其是当设备处于不可信区域）更好的安全控制。

IOT中溶液的寿命期间，不同世代和版本的多种设备类型将可能连接到该系统。即使一个物联网解决方案与一个设备类型开始，被部署设备人口的异质性应该可以预期的。随着越来越多的异质性的边缘组件的维护工作，预计

---

80 http://hadoop.apache.org
81 http://hive.apache.org
82 http://mahout.apache.org
83 http://en.wikipedia.org/wiki/MapReduce
84 http://en.wikipedia.org/wiki/Pig_(programming_tool)
85 http://storm.incubator.apache.org

显著增加，而后端软件的维护不应该影响到相同的程度。维护设备和后端之间的简单，稳定的界面将有助于从长远来看。

在一般情况下，从设备硬件移动时，向设备/边缘软件，到云后端的变化是逐渐更容易。出于这个原因，它总是一个很好的做法，开始在这个序列中，这是设计，为设备的第一设计。设备上的可用功率，计算资源，以及通信技术的选择将影响如何以及何时设备与服务进行通信。在许多情况下的某些处理需要发生在边缘，需要保证的响应时间时，例如，或者以执行发送到后端数据的滤波。具有设备上少的情报可能会增加在云后端的依赖，而且有利于提高系统的灵活性，并降低了维护和运营成本。

这些权衡应该在物联网解决方案的具体情况和业务需求考虑，并可以从场景到场景变化。

遥测设备

的类型和待收集遥测数据的频率是一个的IoT溶液的基本方面。这一决策过程应该由业务需求来驱动。决定收集哪些信息之前，商业动机和目标，如转化向服务提供商的商业模式，增加新的服务，提高客户的参与，或优化操作和维护，应加以澄清。约需要什么遥测的要求应该从业务目标的。

有收集的数据量和成本之间的一个关键的权衡。未收集的数据不能被分析，但在性能和成本方面支付收集的数据。试图收集尽可能多的数据可能并不总能保证在需要时正确的商业问题都可以回答。此外，收集过多或不必要的数据使它更难以区分有用信息“噪音”，也是影响运营和管理成本。在许多情况下，了解收集到的数据的价值可能是一个反复的过程。

一个可能的策略是将器件进行编程以发射的遥测数据的不同粒度，然后根据需要控制从云这一水平。然后，将配置改变命令可以用于指示该装置改变该收藏品轮廓，并开始发送不同级别的遥测数据。

此外，不同类别的数据可以被区别对待。设备可能拆分热路径的数据处理 - 实时发送到云和冷遥测，它可以在本地收集和延迟的基础上转移。例如，使用网络状态检测装置可以在移动网络发送热路径数据和无线网络连接后传输冷遥测数据，或者被建立有线连接。

在包含多个子组件（例如工业设备装置）的复杂设备的情况下，该装置遥测最有可能将需要对每个子部件分开处理（和由该溶液在逻辑上视为一个单独的设备）。如前所述，那些遥测数据流可以通过使用协议报头属性（诸如“的stream_id”），以允许分化和适当的处理在后端被隔离。

另一个要考虑的方面是如何将数据的设备，设备的拓扑结构，部件和系统之间的相互关联。遥测流程应包含相应的属性，使连接在后端的整个系统的整体见解的信息。

边连通

用于直接或间接设备连接不同的拓扑结构进行了前面所讨论的。当使用天青的IoT集线器作为云网关，边缘连接选项显示在图1。

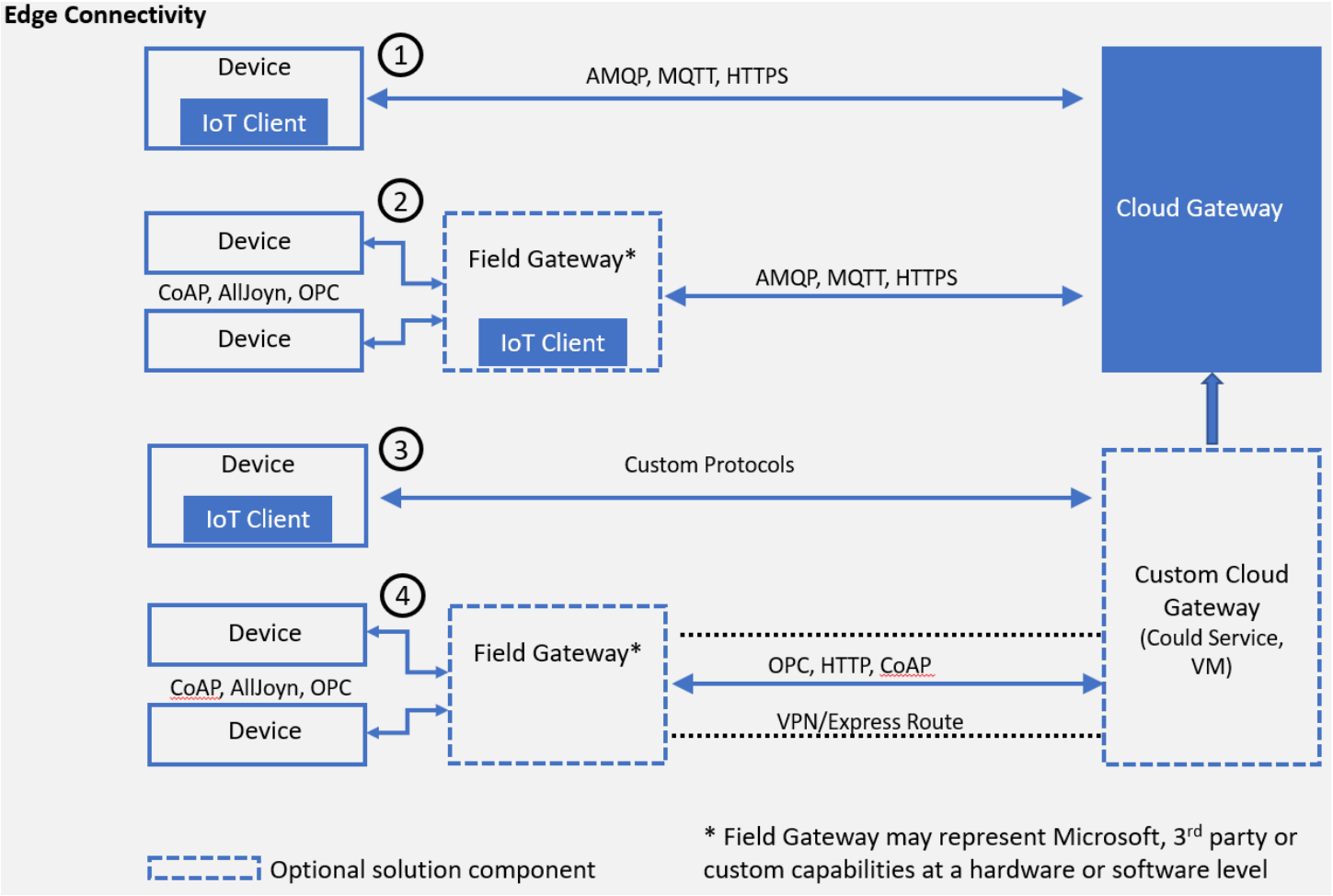


图1 - 与物联网中心边缘连接

除了拓扑结构，需要许多的其他方面被设计为解决方案，如在下面的部分中描述。

传输协议

在该架构中的传输和消息传递协议的重点是在的上下文中基于IP的设备（包括那些作为字段网关）和云的网关之间的通信。对等网络和本地网络的通信标准，链路层协议和物理数据传输（有线，无线）超出范围。

注意： 有（包括本地网络的交互涉及与AllSeen联盟，开发了AllJoyn标准的本地网络通信微软参与更多的努力

一个可选的云网关）- 和工业物联网设备的scenarios-与周围的OPC统一架构OPC基金会（OPC UA）微软参与。这两种

标准涵盖了本地网络发现和设备交互。

物理和链路层的考虑。 虽然物理转移，链路层和本地网络的技术选择和使用指导，超出范围的这个文件，了解它们的使用对在通信和传输层之上的潜在影响是很重要的。底层通信技术不仅会影响服务的质量，而且将决定由设备使用的频率和通信模式。

建立在无线和电力线技术网络很容易受到干扰和其它信号质量问题，这可能会导致数据帧和数据包损坏和丢失。电池供电的设备将优化发送/接收模式用于降低功耗，并且经常会使用无线电点播，这意味着设备将不会在任何时候都可达。移动运营商网络（例如，GSM，3G，4G）补偿许多的基于无线电的技术的影响，但更高的分组延迟和分组丢失仍然是常见的。随着国内漫游，即设备被允许在不同运营商网络上漫游，设备可能会相当频繁地切换连接和IP地址。移动车辆可以在基站之间每两分钟，这取决于特定的频带的范围进行切换。

这些实施例为共同的需要设计基于底层通信技术的设备的通信模式提供上下文。一种装置，能够检测连接的类型被使用（网络状态检测），并切换通信模式。例如，大的二进制传输可以通过Wi-Fi或有线连接来执行，而一个蜂窝或无线电网络上的装置将实现减小的通信简档。

用于实现在2.5节设备交互对于不总是可到达的功率受限或电池供电的设备所描述的服务辅助通信原理的另一个常见的模式，是使用带外通信信道（例如，移动运营商SMS）到“唤醒”的设备，并指示它建立出站网络连接到它的‘原籍’网关，当时间关键的指令需要被传播出去。

虚拟专用网（VPN）技术可以整合和隔离的网络，建立一个单一的地址空间，功能上等同于一个本地网，而在现实中跨多个底层网络。它提供了机制，在一个孤立的网络安全地加入和参与，但不保证网络内部的流量。如果没有一个像每个端点防火墙其它组件，它故意不限制如何在虚拟网络的参与者可以互相通信。在场景中参与VPN设备的用户或潜在的未知入侵者的物理控制，虚拟网络环境必须被视为敌对的互联网环境。

VPN可以提供稳定的寻址装置; 但即使有地址固定分配给每个设备的VPN上下文内，地址是仅当设备正在积极连接是有用的。更常见的是，设备将被分配在VPN动态的地址，通过DHCP，然后会在DNS中注册的发现。这种模式是信息设备一样，但通常会导致显著的管理负担个人物联网设备，尤其是当设备是移动，并经常删除连接。VPN具有建立和重新建立连接，以及为所有通信显著网络传输量和计算开销。因此，它更适合与（场）网关和功能强大的设备使用。

虽然VPN是推荐的技术选择整合现有的数据中心资产注入Azure的物联网解决方案，不建议用于集成移动或无线连接的设备，或异常大量设备，到Azure云。在工业自动化等环境稳定和可靠的连接，其中相对较少的设备（几十个，甚至几百个）或环境需要连接到云中，使用网络

融入Azure中应考虑：天蓝色的VPN，[86] ExpressRoute[87]。对于更高的带宽，可靠性和更低的延迟需要ExpressRoute应予以考虑。

在生产网络的边缘场网关（边缘设备）应具有用于两个环境单独的访问路径。边缘设备应促成在应用程序级别两种环境之间的信息交换，通过一个消息接发应用程序协议。网关可以连接成天青点到站点或站点到站点VPN，[88] 然后会也从云解决方案中的寻址和访问。在云计算方面，VPN和现场终端必须通过云服务或天青VM托管定制的云网关解决方案集成。

> *注意：* 在撰写本文时，Azure平台支持IPv4外部（与本文件中的指导假定的IPv4为基础），但不依赖于它，一旦它可用于微软Azure网络和网络边缘将直接转化向IPv6的。IPv4地址空间是从使用多种转换机制的IPv6地址空间内到达。[89]

传输协议的选项。 本文介绍了两种最常用的传输层协议：TCP和UDP。在这个水平，就像SCTP（IETF RFC4960）或径TCP（IETF实验RFC 6824），或UDP的高带宽应用（如UDT）其他协议可能在特殊情况下或现有的协议支持选择自定义云网关应用的作用但超出范围本文档。

TCP（IETF RFC793[90]）提供流完整性，流顺序，和流量控制两个网络端点之间，并且是除了那些在UDP部调出所有场景的默认传输选项。

UDP（IETF RFC768[91]）是一种简单的数据报（字节帧）传输模型作为薄层通过IP和具有最小开销。因此，它是受限的设备应用的热门人选。UDP不处理数据包的顺序或数据包丢失和不具有基于反馈的流量控制方案。这些是用于其中需要与非常小的端至端延迟，其中损失是可接受的待传输的信号的场景所需的性质，并且其中所述信号分量的顺序可以在接收器侧被重构时必要的。

音频和视频信号通常被组织在能够经由UDP或具有潜在的数据丢失任何其他单向传输方法被转移流容器格式（例如MPEG传输流）。

UDP路线应被固定在根据重叠的应用协议的规则，最常使用DTLS（IETF RFC6347[92]）。对于其中它是不能接受的招致的时间持续期间的数据丢失，并且其中延迟是不具有最高优先级，基于TCP的通信一般应外部本地网络应用优选的应用程序。内的本地网络，UDP可以是一个有用的选项，以限制为极其受限设备上的计算和存储器足迹并且与在下一节中讨论的CoAP协议协议组合的可行选择。

---

86 http://azure.microsoft.com/services/virtual-network
87 http://azure.microsoft.com/services/expressroute
88 https://docs.microsoft.com/azure/vpn-gateway/vpn-gateway-howto-site-to-site-resource-manager-portal
89 http://en.wikipedia.org/wiki/IPv6_transition_mechanisms
90 http://tools.ietf.org/html/rfc793
91 http://tools.ietf.org/html/rfc768
92 http://tools.ietf.org/html/rfc6347

设备和服务，积极听取UDP包很容易发生洪水的袭击，其中包括DTLS握手。在这些情况下提供额外的保护，例如在受信任的网络关系隔离网络隧道，应适用。

消息协议

超文本传输协议（HTTPS）。 HTTP（IETF RFC7230，RFC7231，RFC7232，RFC7233，RFC7234，RFC7235）是网页，用于请求/响应的交互优化的核心协议。HTTP使用TLS在IETF RFC2818中定义的结合（HTTPS）保护。HTTP 1.1协议是纯粹基于文本的和简单的。指定的HTTP / 2后继协议是更简洁，支持所有HTTP 1.1能力，并提供了相当复杂的成帧和连接管理的解决方案，允许复用和用于双向数据流进行建模。HTTP / 2实现必须支持TLS 1.2的保护。

HTTP在本文档的上下文中通常是指HTTPS，即，HTTP 1.1 + TLS 1.2（RFC7230ff + RFC2818）。HTTP / 2已经为物联网方案和它的使用和采用物联网空间非常实用的功能进行监控。HTTP / 2不是此对话的焦点。

HTTP连接的管理模式进行了优化周围相对较短的客户端和服务器交互。通过HTTP支持的交互模式是一个请求/响应模型以与由流顺序请求的响应。有许多用于建模的附加交互模式（例如通知），或通过HTTP异步消息传递，如技术“长轮询”。

HTTPS可以被认为是其中的各个设备将数据发送到云网关偶尔场景和作为单个消息或多记录一个很好的选择“上载”，并且其中，不要求低等待时间的双向通信。“偶尔”是指该设备发送数据很少，以至于维持设备和云网关之间正在进行的连接是不经济或技术上是可行的。

安全，高通量事件流入使用HTTPS的基于Azure的解决方案是原生支持在Azure上物联网枢纽和事件中心。一个设备可以接收使用在一个定义的IoT集线器终端定期HTTPS查找命令或其他信息。如果设备需要以最小的延迟接收远程命令瞬间，需要具有可容易获得的网络路由到该装置的持久双向连接。

高级消息队列协议（AMQP）。 AMQP 1.0（ISO / IEC 19464：2014，OASIS[93]）是一个强大的，面向连接的，双向的，多路传输信息传输与固有的，紧凑的数据编码协议。它提供了用于连续地连接的设备中，高吞吐量通信优化，并且集成了流量控制来保护发送方和接收来自“过载”彼此。

对AMQP 1.0库可用于不同的语言和运行时，跨多种操作系统。

AMQP 1.0是其中的各个设备保持长寿命的连接，与正在进行的基础上云网关通信，并潜在地传输大量数据的情况下的好选择。

WebSocket协议。 WebSocket协议（IETF RFC6455[94]）是通过HTTP / HTTPS在具有协商TCP的双向层。它允许共享（复用）的HTTP / HTTPS基础设施和港口与运行在TCP等协议，即使这些协议及其实现需要的WebSockets的明确支持。

93 http://docs.oasis-open.org/amqp/core/v1.0/os/amqp-core-overview-v1.0-os.html

94 http://tools.ietf.org/html/rfc6455

对于WebSocket的最常见的用例场景是通过HTTP / HTTPS基础设施和端口使得在HTTP / HTML web上下文和隧道其他应用协议如AMQP 1.0双向通信。所述AMQP 1.0协议具有明确的用于通过HTTP / HTTPS基础设施WebSocket协议为防火墙穿越的目的结合。WebSocket协议的所有直接的应用，如直接通过WebSocket的帧上述数据编码之一的流动框架，被认为是自定义的协议，因为他们没有解决或元数据框架是AMQP或其他通讯协议提供一个标准的方式。

MQ遥测传输（MQTT）。 MQTT 3.1.1（ISO / IEC 20922，OASIS MQTT 3.1.1[95]）**为消息的轻量级客户端 - 服务器传输协议。MQTT为受限**设备有吸引力的，因为它是一个非常小的足迹在设备上极其致密的，并为消息帧（以及相应的网络带宽）。

一种设计折衷要注意的是，使用MQTT一个非常紧凑的报头格式，但具有用于消息元数据，不支持如自定义内容类型报头，要求出的带外发送机和接收机之间的协议。

有几个MQTT功能在大规模分布式，高可用性的基础设施，物联网使用时是一个挑战。在多节点的消息系统中，QoS2“正好一次”的传递保证将要求在任何时候都完全一致的系统（跨多个节点）。尽管这在技术上是可行的，这样的实现将是非常复杂的，并会影响整个系统（详细的延迟**和可用性指的是CAP定理[96]）。因此，不建议大物联网的部署QoS2的使用。恰好传递一次常见的替代方案是在接收器重复数据删除，或使用幂等**操作。例如，对于被建模来交换状态变化的系统，“至少一次”语义是足够的，因为在接收到相同的状态多于一次将导致相同的结果（假设消息传递顺序被保留，这是通常的情况下，使用MQTT当包括）。

另一个挑战表示“保留”消息的使用。这给服务器（如持续时间和消息数），这与highscale系统资源管理要求相冲突，并提供通过强制资源枯竭潜在的拒绝服务攻击向量无界状态管理的要求。适当的授权模式可以考虑和应用，以减轻这些风险。

如果MQTT是在其足迹优势特定场景的候选人，推荐是“最多一次分娩”或QoS 1的使用限制在QoS 0“至少一次的传递”，避免了“保留”功能的使用。

约束应用协议（CoAP协议）。 该受限应用协议（CoAP协议，IETF RFC7252[97]）是可以在UDP或任何其他数据报传输来实现，包括GSM短消息服务（SMS）一个datagrambased协议。CoAP协议是在数据报传输的原理和HTTP的方法非常紧凑的改动。开放移动联盟（OMA）轻量级M2M协议（LWM2M）在CoAP协议的顶层（见OMA LWM2M下面的管理协议部分有详细介绍）。

CoAP协议的优点是它的紧凑相比，HTTP等协议。因为它是基于数据报，也没有必要立即建立或保持连接或跨越多个任何状态的数据报，直到安全被DTLS的形式，这引入节点亲和性的安全上下文，并添加其具有的属性连接。支持基于UDP的协议和DTLS没有与云网关一起微不足道的，因为当流量从整个开放的网络承认所有通信各方都容易受到洪水袭击

95 http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.html
96 http://www.julianbrowne.com/article/viewer/brewers-cap-theorem
97 http://tools.ietf.org/html/rfc7252

（和大的解决方案可以很容易地成为这种攻击的目标）。此外，在拥挤的路线丢包可以显著。TCP和TLS应考虑长途传输更加稳健。

有关的移动运营商网络中支持LWM2M / CoAP协议或连接到现场设备，所述CoAP协议流量可以被分离，可靠地虚拟化通过VPN或ExpressRoute，既可以在其中设备驻留或到移动运营商的“私人的APN的位点的VPN网关。”（ *注意：* CoAP协议被示出在图1中在字段网关和定制云网关之间的VPN隧道来实施。）

OPC统一架构（OPC UA）。 开放平台通信（OPC）基金会的统一架构（OPC UA）包括一个数据模型，安全和一组传输协议的映射。的OPC UA服务器是需要为了获得数据将被连接到的可寻址实体。一个OPC UA场网关可以通过本地网络读取OPC服务器的数据，并将其转发到云网关通过使用OPC UA的PubSub，一个JSON有效载荷超过MQTT或AMQP一个TLS保护的路径，像我们的开源OPC出版商。对于单独的机器，它也可以使用我们的OPC代理，安全地“桥”的工厂网络上的OPC UA服务器和嵌入式在Azure云一个网络应用程序中，OPC UA客户端之间的通讯，而无需打开工厂防火墙。

安全

值得信赖和安全的通信。 如果有任何依赖于从信息接收和发送到设备的所有信息必须是值得信赖的。 *值得信赖的通信* 意味着信息是可验证的起源，是正确的，不变的，及时的，并且不能被未授权方以任何方式被滥用。

即使从一个报告房间的温度每五分钟不应该留给不安全简单的传感器遥测。如果任何控制系统响应该输入，或者从它绘制任何其他的结论，所述设备和从和给它的通信路径必须是可信的。

许多物联网设备，如数字服务能力，丰富廉价的传感器或普通民用或工业产品，将成本，这通常会导致交易的计算能力和内存节省成本进行优化。然而，这也意味着交易掉加密能力以及更普遍，对应变能力的潜在攻击。

除非一个设备可以支持以下密钥加密的功能，它的使用应该被限制在本地网络和互联网络的所有通信应通过现场网关简化：

- 数据加密用可证明安全的，公开了分析，并广泛实施对称密钥加密算法，如AES具有至少128位的密钥长度。
- 数字签名具有可证明安全的，公开了分析，并广泛实施对称密钥签名算法，如SHA-2具有至少128位的密钥长度。
- **对于任一TLS 1.2（IETF RFC5246支持[98]）用于TCP或其它基于数据流的通信路径或DTLS 1.2（IETF RFC6347），用于基于数据报的通信路径。**的X.509证书处理TLS-典型的支持是可选的，可以通过用于TLS（越计算高效和线高效预共享密钥模式“TLS / PSK，”IETF RFC4279所取代[99]），**其可以与用于上述AES和SHA-2算法的支持来实现。**
- 更新的密钥存储和每个设备的密钥。每个设备必须具有唯一的密钥材料或标识，并向系统令牌。设备应该能够安全地存储该密钥的设备上（例如，使用安全密钥存储）。该装置应该能够更新密钥或令牌周期性地或反应性地在

[98] http://tools.ietf.org/html/rfc5246
[99] http://tools.ietf.org/html/rfc4279

紧急情况系统都违反合同的。密钥更新可能会出现在空中或通过一些其他手段，但需要更新能力。

- 设备上的固件和应用软件必须允许更新，使发现的安全漏洞的修复。

作为一个基本原则， ***与装置或场网关所有云通信必须通过安全通道发生***
当设备直接与微软Azure平台提供的服务端点。

如果（传统）设备必须使用不安全的或非标准的和专有的通信路径进云系统，它们应当通过单独托管的自定义协议网关或本地网关字段被连接。

还有，将是很多情况下，对本地网络设备的访问控制已经通过网络级的访问控制已经完全实现，而网络中的所有加入的成员可以在没有任何自**由沟通，或天真，身份验证和授权。现有在这样的网络设备** ***必须通过现场网关进行通信*** **在不安全的网络的边缘。**

物理防篡改的安全
传感器和设备能够而且必须经常被放置在公共场所，任何人都可以潜在地对他们进行物理访问。此外，在设备的篡改是没有操纵设备的硬件或软件的只是行为。一种数字值得信赖的传感器可以被欺骗通过拆卸和重新定位它报告误导性数据。或者攻击者可以在设备周围影响环境，在设备的紧邻制作误导身体状况，推高了整体系统进入一个错误的反应。一个打火机点燃附近的烟雾或温度传感器可能保持，例如，诱骗数字楼宇控制系统到泛滥的宾馆走廊里的自动喷水灭火系统。

物联网提出了安全的一个新的层面，因为物联网设备在广泛的个人，商业和工业应用中使用，不仅威胁环境的各个场景中的环境有所不同，这也取决于设备相关的条件不同它的环境。例如，在运动的车辆具有不同的威胁环境比车辆在交通灯的前面空转，和另一个从停泊的车辆。确保车辆的数字部件，因此比固定一个“经典”的软件应用程序 - 要复杂得多，这适用于物联网的许多场景以类似的方式。

随着物联网的空间模糊数字和物理的担忧，这也模糊了安全性与安全性。突然，安全威胁成为安全威胁。如果出现“出错”使用自动或远程控制设备，物理缺陷，控制逻辑缺陷故意非法入侵和操纵生产批次可能会被破坏，建筑物可以被抢劫或烧毁，人们可能会受伤或死亡。这是一个不同的类比别人杏被盗的信用卡限额的损害。为使事情动了，也为传感器的数据，最终导致引起事物的运动命令，命令中的安全栏，必须比任何电子商务或银行场景更高。

即使显然已经超出了一个基于云的系统的控制，因此它是 ***强烈建议装置的设计采用了抵御物理操纵的企图特征*** **以确保整个系统的安全完整和**可信性。

可采取提高物理设备的安全性的一些措施是：

- **选择微控制器/微处理器或提供安全的储存和使用密码密钥材料的辅助硬件，诸如可信平台模块（TPM）[100] 积分。**

- 安全引导加载器和安全软件的加载，在TPM固定。
- 使用传感器来检测入侵企图并试图与警报操纵设备环境和设备的潜在的“数字自毁”。

数据编码

有一个大的和越来越多的可用数据编码格式。最佳的数据编码的选择从用例的不同使用情况，有时甚至由因素的制约等多少空间可用于在设备上的额外覆盖区。

XML和JSON是无处不在的服务器和客户端多平台。既享受非常广泛的图书馆或platforminherent的支持，但有非常显著线的足迹，因为它们基于文本的性质。

CSV是简单的，可互操作和紧凑（文本），但它的结构约束，以简单的值columns-，然而，是经常足够的时间序列数据行。

BSON和MessagePack是倚在JSON模型，但有很大的编码规模优势高效的二进制编码。这都需要自己的图书馆，有一些独特的选择，像BSON的情况下，缺乏一流的阵列支持。

谷歌的协议缓冲器（“的Protobuf”）和Apache节俭产生非常小的编码大小，但需要外部架构（甚至代码）分配给所有潜在的消费者，这代表与多个读者/消费者的平凡的组成复杂系统的挑战。

阿帕奇阿夫罗通常是一样有效-或更有效-比现有选项，同时本身支持层状上压缩。随着Avro的，该模式被嵌入作为一组记录的前导。该前同步码要求，使在阿夫罗相比MessagePack或BSON小型或高度结构化的有效载荷与最小的结构重复的缺点。

数据布局

作为编码同样重要的数据布局，这也可以对编码数据的大小大的影响。其中遥测数据是在对象的数组，其中每个对象携带对于所有的值显式属性的形式发送的幼稚JSON编码方法，具有极大更大元数据开销比数据布局模拟CSV，接着用阵列报头的共享列表携带该行的数据。

数据布局惯例定义了如何将数据的结构的解决方案的范围内的制约，从而使数据能够在整个系统，包括设备，后端处理，分析，和用户界面来处理。所有这些组件都需要依赖于一个通用模型/模式。

在事件驱动的系统的一个重要原理是，处理和加工在一个模型的上下文中的数据单元是一个记录。一则消息，一个存储块，或一个文件可以包含一个或多个数据记录（或“事件”）。的记录的序列可以跨越多个消息或存储单元。

[100] http://www.trustedcomputinggroup.org/developers/trusted_platform_module

CSV的行/列结构提供了天然组用于布局约束的，并且允许未明确地限定行（每个等同于一个记录）的列表中，与未明确地有界的一组列，其中每个列的值是的原始类型。

对于通过JSON，阿夫罗，AMQP支持地图/阵列/值的结构模型，并MessagePack数据编码，有以下共同的布局选项：单独的记录，记录序列，或记录序列的元数据前导码，类似于CSV，包含首标描述列，后面是表示该记录顺序中的行。

下面的矩阵可以帮助选择适当的编码。在布局列中，“平坦”是指由基本数据类型的单独记录。“复杂”是指其中记录的结构超出原始类型的数据。

布局	JSON	CSV	Avro公司	AMQP	MsgPack
单记录，平面数据	++	+	-	+++	+++
单记录，复杂的数据	++	N / A	-	+++	+++
记录序列，平面数据	+	++	+++	+	+
记录序列，复杂的数据+		N / A	+++	++	++
记录序列w /元序言	++	+++	+++	+++	+++

表1的数据布局编码的比较

注：（ - ）较差的；（ + ）好；（ ++ ）更好；（ +++ ）最好

边缘处理

在许多应用场景，特别是那些其中的各个设备经由计量网络他们的云后端系统通信时，它是不理想的通过通信链路发送原始传感器读数或状态信息到云因为放置在云相关联的成本和负载的系统中，当很多未经处理的数据流必须并行处理。

通常情况下，解决方案的IoT明确要求信号的数据流的评价，与视频和音频覆盖特定的信号的形状和频谱，通过数字信号处理算法或模式匹配或发现应用，所以需要在第一方来治疗这些类型的信号类时尚。

温度传感器提供周期性读数，也许在1Hz，在每读出一个数字，明显。上的通风风扇与确定在工业环境中设备的健康帮助的振动传感器提供周期性读数，也许在500Hz，在每读取一个数字，明显。音频传感器-传声器 - 提供了在每一个读取数表现的周期性读数，但在44千赫。视频传感器提供了在每读出一个非常大的矩阵表现周期性读数和在60赫兹或50赫兹这样做。

所有这些信号从预处理和压缩受益转移前，根据其种类信号有可能已被广泛接受和应用的如何编码，封装，并带有信号行业标准。如果是这样的话，像MPEG标准的音频和视频信号的压缩和编码，它们应该是首选，并通过不具备解释它们不变系统的所有部件通过。

可应用于任何时间序列最琐碎聚合是将若干个点在实时记录插入用于任一特定的时间段中的单个记录，其中读数已通过提供的平均或中值保持稳定，或者对于固定的周期对于读数。

此外，许多设备都相当能够预分析的原始数据，并使用当地的计算能力，是否应该通常优于跨越计量网络发送大量的数据。

在许多情况下设备可以使用网络状况检测和应用不同的预分析，聚合和基于所使用的连接类型的压缩算法。例如，如果需要购买静止分析原始数据，混合模型可以被使用，其中所需要即刻和近实时通过移动网络传输和所述原始数据被局部地保持并传送在数据在稍后的时间点通过有线（或Wi-Fi）连接。

边缘处理通常再加上后端，它们负责下游处理之前解释接收到的数据的适当组件。例如，压缩数据需要被解压缩的和编码的需要加以适当的解码。在许多情况下，这是在流处理器完成，从云网关读取数据。如果有数据的多个消费者，一个数据流处理器可以专用于解释输入的数据（例如，解压缩或反串行化），并输出转换后的数据到内部的流动缓冲液（如天青事件集线器）。通过这种方式，它将作为传入流量的所有其他事件流处理器的主要来源。

在某些情况下，这种类型的处理可以在自定义网关来完成，达到云网关的摄入点之前。天青的IoT协议网关展示如何这种类型的定制加工可以使用处理流水线，其中，不同的模块可以插入在将数据传递到下一个之前执行专门的处理的概念来实现。

管理协议

还有就是在行业设备管理协议一组不断发展。使用预定义的设备模型的使上的IoT设备有效地利用网络，处理和功率资源。OMA LWM2M（或轻型M2M）是由使用了紧凑资源模型和设备和服务器之间的交互，以支持非常受限设备的开放移动联盟定义的标准。OMA LWM2M提供传输与CoAP协议为受限设备绑定。

其他设备管理协议，包括OMA DM，TR-069，和COMI定义器件模型和交互使用的设备。OMA DM，在移动设备管理和一些的IoT实现中使用，使用XML（通过SyncML的定义），以使设备管理和因此比OMA LWM2M更详细。TR-069是由宽带论坛发布的技术规范，使用双向SOAP / HTTP为基础的协议来管理设备。除了大量的设备管理标准，存在许多定制设备管理协议，其中设备供应商已经需要提供设备和服务器/服务之间的系统的功能。的那些设备管理协议不同的层会影响此参考架构的实施方式。例如，

设备管理

物联网设备景观异质性极大地考虑到各种硬件选项，环境，操作系统和编程语言，以及设备和服务之间的通信手段。使用的设备管理协议提供了一个抽象通过定义一个协议，从下简化这种复杂性

传输层到较高应用层，使得服务可以提供必要的信息的装置，以保证设备的运行状况。

服务供应商和企业需要登记和探索，使连接，远程配置，以及由定义的策略和业务流程规定的方式设备更新软件。例如，根据不同的行业，会有下哪些设备可以远程配置和改变，审批链条，监管审计要求，物理的保护的存在，更多的情况下，大大政策别共。

每一个物联网系统，以确保设备和相关业务流程的健康提供了一套设备管理功能。一个设备注册表的概念是在远程设备上实现设备管理功能和启用服务侧接口，用于云应用使用由远程设备提供的功能是至关重要的。以下是可以通过物联网系统中启用设备管理功能的列表：

1。 **设备配置和发现**
2。 **设备访问管理**
3。 **遥控**
4。 **远程管理和监控**
5。 **远程配置**
6。 **远程固件和软件更新**

设备配置和发现。

大多数的IoT设备的生命周期表明，器件制造和部署到全世界各种位置。部署位置可能不会在制造时是已知的; 因此，它可能是重要的，以使多相自举过程，其中设备与自举服务，这稍后提供连接细节的知识制成的。当该装置展开时，组织中部署该设备提供进一步的信息，包括设备位置和任何其他所需的信息到引导服务。自举服务被配置为与将要使用该设备的云网关响应。该设备可能需要重复设置过程，例如在场景中设备所有权的变化。

为了使云应用程序执行设备管理活动，该装置可以当它创建了云网关的会话描述自己的云。有相关的设备是如何描述到系统三个核心概念：

- **自定义的设备模型**
  （使用设备模拟器或显影剂）的设备工程师通过对设备的能力进行迭代，因为它们建立该装置的过程中使用的自定义的设备模型。一种设备工程师可以通过创建一个具有几个属性和支持的命令以后添加更多的设备启动。类似地，设备工程师可以具有许多设备，**其中的每一个提供了独特的能力; 使用自定义的模型中，不需要在设备工程师要注册的设备模型的结构。**

- **预定义的设备型号**
  ，根据网络和电源/处理约束操作的制造的IoT部署从其中使用最小限度地使用该设备的处理和功率消耗的预定义的设备模型大大受益。类似地，最小的网络流量使设备通过使用有限的和昂贵的基础设施尤其是当（诸如一个卫星）异构网络（无线网络，2G / 3G / 4G，BLE，Sat等）来传输。当实现预先定义的设备型号，设备工程师可能会发送一个或编码的设备信息

两个字节充当密钥到预定义的设备模型。这种方法的简洁导致级一到两个数量相比，自定义设备模型的效率。

- **预定义的主模型**
  与所述设备相关的设备型号和元数据被存储并保持在云一侧，但设备将仍然不知道的那些。该图案是在棕方案是特别有用的，其中所述装置的固件不能被修改，或该设备不应该存储的元数据。

设备访问管理。 设备（由多方潜在管理）可以执行自己的属性和命令，包括创建，读取控制，和写入访问权限的设备性能和设备的命令执行权限。根据物联网应用，多权限级别，可能需要以适当控制访问设备的资源存在。

遥控。 在IT场景中，远程控制通常用来协助远程用户或远程配置的远程服务器。在物联网的情况，大多数设备不具备从事的用户，因此远程控制是一个方案，使远程配置和诊断。遥控器可以使用两种不同的模式来实现：

- **互动式连接**
  为了使通过（在Windows平台的Linux或远程桌面，例如，SSH）直接连接到设备的远程控制您需要创建到设备的连接。由于设备暴露在开放的互联网的安全风险，建议使用中继服务（如Azure的服务总线中继服务），以实现从设备到/连接和通信。由于继电器连接从设备的出站连接，它有助于限制的设备上打开的TCP端口的攻击面。

- **设备命令**
  通过设备命令遥控器使用的设备和天青的IoT集线器之间建立的现有连接和通信信道。为了使设备基于命令的遥控器，以下要求必须实现：

    - 物联网后台意识到设备上可用的设备命令。这通常被定义为所述设备模型的一部分。
    - 该设备上运行的软件需要实现远程控制命令。这些设备命令应遵循的请求（从后端的IoT到设备）和（到的IoT后端从装置）模式的响应。

物联网服务后端可以保持历史响应消息的纪录从设备命令以供审核。更新设备状态是通过设备的命令作出。更改设备元数据和状态需要分别被推动到设备注册表和状态存储。更新设备状态可以通过请求从所述的IoT后端被强制到设备，或该装置可以在识别的状态变化自动更新后端。从设备后端的自动更新应该尽量少做，因为它可能产生的网络流量，提高了设备的处理器和可用功率的使用。

远程管理和监控。 由于大多数的IoT装置不具有在溶液中部署后的直接用户，远程管理是经验，管理员可以通过使用设备的命令的监视他们的设备的状态和远程更新设备的状态或配置。

设备的状况可以通过监测他们发送到后端的数据来确定。这可以包括操作数据和元数据。

远程配置。 远程更改设备的配置是在设备的生命周期中几个阶段的要求：配置，诊断，或与业务流程的集成。

远程固件和软件更新。 软件缺陷可能是安全漏洞，这使得固件或软件的更新，以修复缺陷或提供新的功能每一个物联网系统的关键能力。在设备上远程更新固件和软件是一个分布式的，长时间运行的过程，通常涉及到业务流程的一个例子。例如，更新控制高功率燃料泵可能需要在相邻的系统的步骤为在执行所述更新和验证重新路由燃料的设备上的固件。

支持固件和软件更新装置通过该设备模型中定义（或通过与设备模型相关联的设备类型）。设备更新是在物联网后台启动，设备在通过设备命令一个适当的时候通知。当设备明确支持固件或软件的远程更新，物联网后台应该提供基于定义的业务流程和政策更新的命令。一旦接收到设备命令来更新，设备需要下载更新包，部署该更新包，重新启动到新部署（在固件更新的情况下）或启动新的软件包，并验证新的固件或软件如预期的那样运行。在整个多步骤的过程，

提供更新包可以通过像Azure存储或者通过CDN存储服务来完成。验证下载包的完整性，重要的是要确保包装起源于预期的来源。

完成固件更新后，设备必须能够验证并确定了良好的状态。如果设备没有顺利进入好状态，在设备上的软件应该启动回滚到已知的良好状态。已知的良好状态可能是最后已知的良好状态或称为存储在存储分区中的“金州”的设备固件映像。

高可用性和灾难恢复（HA / DR）部署拓扑

物联网的资产和设备通常形成分布式环境。它们可以是静止的或移动的，分散的或搭配，有时与本地站点相关的。基于解决方案的需求，这些设备可以连接到一个单一的集中式或分布式部署后端。

在几个云后端部署拓扑和在不同地点工作的分配方案：

- **单站点。** 这是最简单的模型，在这种情况下，云网关（S）和所有与设备相关的商店并置在一个单一的数据中心区域，而所使用的服务的高可用性和灾难平台的固有支持倾斜复苏。因为它的简单，这种拓扑结构通常是大多数解决方案的起点。

- **区域故障转移。** 在一个地区的故障切换模式，解决后端将在一个数据中心的位置作为单点模式主要是跑步，但解决方案的云网关和后端将在故障转移目的的额外数据中心的区域进行部署，万一云网关主数据中心遭受服务中断或从设备通过网络连接到主数据中心以某种方式中断。这些设备将需要使用每当主网关无法达到二级服务端点。用交叉区域故障转移能力，该解决方案可以可用性超过单个区域的高可用性得到改善。灾难恢复和地理概念，故障切换将更加深刻本节后面的覆盖。

- **多站点。** 在多站点拓扑，该解决方案同时，并在很大程度上独立运行在多个网站，但它在概念上是单一的解决方案。多个站点可以在同一数据中心区域并置以形成“尺度单位”的量，整个数据处理支柱可以应力测试，以最大容量，然后

更多的容量可以安全地通过加入另外的刻度单位添加。该系统的位点也可以位于在不同的数据中心区对于各种各样的原因，包括接近用于降低延迟到周围数据的位置的装置或政策的担忧。这些网站还可以有一个区域故障切换站点。在多点模式，设备注册，因此在网站的一个“宿主”。

- **漫游多点会议。** 在该变型的多站点模型，设备驻留在网站（刻度单位）中的一个，而是可以连接到基于某种形式的接近度估计的上的最近数据中心位置。所收集的信息发送到该设备的“家”网站。

- **多站点，多宿主。** 在该变型中，装置可跨站点漫游和捕获的数据存储在该设备连接到不同的位点，并且可以被收集，并根据需要合并。

拓扑的这份名单并不详尽，但有助于说明关键模式和权衡规划的物联网部署时。有时，特定拓扑结构可以应用到服务和组件的一个子集，而解决方案的其他部分可能基于特定的解决方案要求使用不同的部署拓扑。

从设备的角度有三种可能的设备或现场网关如何与服务后端通信：到一个单一的“家”的端点，到初级或次级端点（用于地理故障转移），或一组端点多站点，多家乡的情况。这些端点的配置可以是静态的（例如，配置在设备上设置）或管理使用从溶液后端命令作为动态的设备配置。

还有使用令牌服务设备的附加选项。如果设备无法到达目的地端点，它可以联系令牌服务获得新的端点和它相应的令牌。这种机制提供了用于在需要时（在对比到主动设备上保持的预定义配置的变化）装置的动态无功重定向。它可以在除了管理设备上的端点配置来施加。令牌服务可以智能地管理网站的地图，但也可以重新配置，作为一个例子，维护的目的。

除了管理该设备用来连接端点，域名系统（DNS）条目和相关的服务，如Azure的流量管理器可用于流量重定向到所需的后端端点。需要注意的是这种技术依赖于设备的使用DNS的能力，而且其精度是由DNS主机条目的时间 - 生存期（TTL）值驱动保存在本地DNS缓存是很重要的。

跨区域的可用性

应用从Azure中所提供的基本服务的高可用性（HA）在Azure中运行效益。对于许多Azure的服务和解决方案，高可用性是通过在Azure的区域级使用冗余提供。此外，Azure中提供了许多功能，有助于在必要时建立与灾难恢复（DR）功能或crossregion可用性解决方案。解决方案需**要设计和准备充分利用这些功能，以提供全球性的，跨区域的高可用性，设备或用户。文章“Azure的业务连续性技术指导”[101] 介绍了内置的业务连续性和灾难恢复功能天青。论文“灾难恢复和高可用性的Azure应用程序”[102] 提供关于Azure应用程序来实现HA / DR策略架构的指导。**

由于云解决方案是由多种服务，它要考虑的是要实现HA / DR的解决方案的各个服务或组件，而不是想着为整个解决方案的一个方法是很重要的。决定要应用技术之前，它定义了解决方案的子服务/组件的需求和预期的可用性是非常重要的。通常情况下，子都会有不同的要求，

101 https://msdn.microsoft.com/library/azure/hh873027.aspx

102 https://msdn.microsoft.com/library/azure/dn251004.aspx

可用性，可扩展性，性能和一致性。例如，设备遥测，命令设备，后端分析，LOB系统的交易，和最终用户UI将都具有不同的可用性，延迟时间，和一致性的目标。甚至不同遥测流或命令类型将具有不同的要求（例如，用于车辆的信息娱乐系统的遥测流具有比遥测不同的加工要求来自发动机的）。在发生灾难的情况下，一些组件可以在降级模式下运行，或者他们中的一些甚至可能不会需要一定的时间周期。灾难恢复技术**需要被设计为每个类别/类型的服务或单独的系统功能的基础上的具体要求。总有可用性之间的权衡，**[103]，**实施和运营成本。**

对于基于地理位置的拓扑结构要考虑的一个重要因素就是状态存储，如果服务进行无状态或有状态处理。状态处理可以被重定向（或故障），确保相应的服务（如计算节点，网站）到另一个刻度单位，网站或区域的供应存在。这些可以是积极运行所有的时间，可用但不活跃（即在待机模式），或者可以根据需求，作为灾难恢复过程的一部分来提供。有状态的服务，但是，代表了一个更大的挑战，因为，除了服务运行时，状态和数据需要被复制和同步。依赖于所需的一致性水平，状态和数据可以同步或异步，其中最终一致性是充分被复制。在某些情况下，数据可能不需要被复制到每个站点，如果它足够的收集和在以后巩固数据到一个集中的位置。此眷上的数据，具体的解决方案的需要的量。

利用所提出的IoT参考架构，国家相关被保持在以下组件和应为每个类别被定义为状态复制适当的技术：

- **设备身份店。** **该设备的身份和相关的安全材料需要在哪个设备预计将建立一个连接每个站点是已知的。这包括辅助站点进行故障转移**或任何其他现场设备可以连接到在多站点场景。通常情况下，身份慢慢改变通过供应工作流管理的数据。配置API提供了封装配置操作，是一个自然的地方延伸，并根据需要管理跨网站身份的一种很好的抽象层。例如，创建一个新的设备时，所述身份记录可以被立即写入到辅助站点。对于待机或按需部署灾难恢复站点，这可能是也足以进行一次的常规导出/导入到二级存储。出口之间的时间间隔将确定恢复点目标。

  在许多情况下，设备将被调配长才有效尝试连接到端点。在这些情况下，身份存储之间最终一致性是可以接受的。批量置备操作可能在对多个位置并行执行，甚至。利用技术，如检查点，并相应的错误处理应确保一致的状态在网站建立的。

- **拓扑店。** **拓扑商店用作设备发现性的指标，并为广大的情况下，实现可以假定为最终一致，一个众所周知的时间到现场进行复制的记**录。这意味着，元数据的更新可能最多到了这个时候生存极限复制。该DNS基础设施采用了类似的策略。如果至少对于每个设备的初始记录插入通过用于标识存储相同的机构设备注册表这将是有益的（即，配置API）。属性和元数据的变化可以通过异步系统中复制。

[103] https://en.wikipedia.org/wiki/CAP_theorem

在许多情况下，设备注册表中只包含缓慢变化的数据，并定期导入/导出可能是一个足以替代项的连续复制。

- **状态存储。** 设备的操作数据通常表征为高容量和高的速度数据。如先前所讨论的，该数据将被在不同的存储基于需求和访问模式分开。对于需要转移或复制的各数据类别进行分析。原始遥测数据很可能并不需要可在辅助站点。聚集的数据将代表这可能是更容易，如果需要复制的减小的数据量。

通常情况下，历史或以前的状态可能不是必要的应用程序后台逻辑。在许多情况下，警报，通知，或甚至指令和控制事件的装置可以只根据设备元数据被应用于（如类型的设备，组类别）或属性（例如，在遥测消息中接收的状态）。

一旦它的决定哪些类型的操作数据将需要被复制，对多个选项中的一个可以被应用：

一个）使用底层存储服务（内置功能Azure存储[104]和SQL数据库，[105]对于
例如，已经提供了内置的地理复制功能）。

b）使用专用的事件处理器拾取的相关信息，并将其传送到事件中心
远程站点（这表示，将需要由另一个事件处理器在远程站点上处理并转化成期望的存储格式的特殊数据的复制信道）。

C）使用内置的应用层（例如一部分一些其它机制，写操作到远程
存储账户或定期安排出口和进口分别远程站点，它可以定期或根据需要）进行的。

- **撮合消息。** 云网关和用于去耦的溶液的组分其他后端内部队列，主题或事件集线器，持久地存储的消息。典型地，消息从云网关接受后将由溶液后端处理，并且也没有必要复制这样的消息到另一个站点。在代理中断的情况下，这些消息仍然受到保护，但无法被读取。如果“飞行”这些消息被认为是跨区域故障转移绝对关键的，那么他们可能需要被复制到备用站点。然而，典型地，消息被促成的时间很短的时间，然后消耗并转移到所描述的持久性数据存储中的一个。在持久性存储的数据保护是典型的策略，因为复制的消息在飞行中的总延迟时间将是几乎相同的延迟保护持久性数据存储。因此，在飞行中防止消息通常不会显著影响RPO / RTO目标。

- **热路径分析状态。** 分析和复杂事件处理引擎保持在内存中状态的聚合或某时间段。有没有简单的方法来恢复这些引擎的内存状态没有成熟的事件重演的机制。如果关键的业务逻辑依赖于这种状态的基础上，坚持历史数据替代计算可能是必要的。

[104] https://azure.microsoft.com/documentation/articles/storage-redundancy
[105] https://docs.microsoft.com/azure/sql-database/sql-database-business-continuity

- **演员的状态。** 演员状态通过，如果需要应复制到远程站点的持久存储通常备份。在恢复的情况下，演员们可以在远程站点上重新装载它们的状态。
- **系统配置。** 变化到该溶液中的配置（例如，改变的阈值限制或业务规则）将需要根据需要被传播到辅助站点。

独立的个体设计选择，在分布式计算环境中，它总是使用幂等操作，以尽量减少副作用，不仅事件的最终一致的分布，而且从重复或乱序传递事件的一个很好的做法。此外，应用程序逻辑的设计应容忍，因为它需要为系统"治愈"或基于恢复点目标（RPO）的额外时间潜在的不一致或"稍微"外的最新状态。下面的文章提供了有关此主题的更多指导："故障安全：指导弹性云架构。"[106]

数据保护和隐私

随着物联网的方案得到越来越多的关注，从消费者保护团体和各国政府的数据保护监管机构，预计数据收集以及遥控情景受到加强监管。

解决方案制造商必须预见到什么数据收集默认情况下，当车主或设备运营商不得不选择退出的权利，或者他们必须选择在数据收集，甚至被允许允许不同区域调控。他们也应该预见到任何数据采集和远程控制功能必须允许暂时不和业主或设备运营商可能想要删除收集的数据为过去的周期。匿名数据收集可能是这方面的一个选项。

尽管设备制造商，保险公司，租赁公司和其他企业驱动数据收集行动，目前尚不清楚是否从车辆收集的数据被收集公司合法拥有。解决方案制造商必须预见到调控，通过不同的管辖权，将授权业主和设备操作员有其数据的使用权，并保持持续时间完全控制。

假设车辆的例子，有很多情况下，车辆在任何特定时间的地理位置可能是当前的驱动程序不想让知道任何人一个非常私人的事情。这意味着，地理位置不能出现，与车辆相关的，只要有与车辆驾驶员相关的方式，也不能出现被以任何永久记录驱动程序关联。

但是，地理位置可能是重要的，应在驾驶员进入事故可能采取行动，因为伤害他/她，潜在的乘客，和其他第三方可能会构成一个压倒一切的优先事项。

解决方案构建还需要预见选入和退出的方案和（潜在的监管授权）选择退出行动追溯长数据被收集后出现。在系统中保留一些数据也可以不通过合法的收集方所拥有和退出后可能因此不存在于系统中。

因此，建议在系统中任何地方存在的保留认证，并链接到其源，并可能有资格在被删除的数据立即任何关联方非聚集的信息。为了执行数据，并通过散装盗窃的敏感信息保护的强偏析，它可以进一步被需要来加密上的源逐源基础的信息。

它也应当预见到物联网系统和数据收集可能在调查的情况，以及在发生意外或其他事故的分析起着至关重要的作用，并有可能成为诉讼的理由。因此，

106 https://docs.microsoft.com/aspnet/aspnet/overview/developing-apps-with-windows-azure/building-real-world-cloud-apps-with-windows-azure/more-patternsand-guidance

强大的属性，包括真实性的证明，不允许数据源抗抵赖性将具有很高的价值，并在管制情况下可能需要。

在构建物联网解决方案，它由一层考虑合规性和认证要求层是非常重要的。为了实现遵从的整体的IoT解决方案，每个下层将必须满足特定的要求。通常情况下，并不是所有的解决方案和平台组件或服务具有相同的要求或完成相同的认证资格。例如，不是所有的Azure服务具有相同的认证（例如ISO 27001，SSAE 16）和解决方案制造商应考虑到他们需要使用，使他们能够实现预期的解决方案认证资格哪些。

5。 附录

5.1术语

本节提供本文档中通篇使用的一些术语的定义范围的。

设备。 有几种类别的设备：个人设备，专用设备或工业设备，仅举几例。个人电脑，手机和平板电脑主要是信息交互设备。从系统的角度来看，这些信息技术设备在很大程度上是作为对人的代理。他们是“人驱动器”，建议行动和“以人传感器”收集有关设备使用直接输入或输入。这些装置被称为在文件“个人移动设备”。

特殊用途的装置，从简单的温度传感器，以复杂的工厂生产线与成千上万它们内部的部件，是不同的。这些设备更作用域的目的，即使他们提供了一个用户界面（与人互动）的一定程度上，他们主要是与范围的接口或在物理世界纳入资产。他们衡量和报告环境的情况下，转阀，控制伺服系统，声音报警器，开关灯，以及做很多其他任务。他们帮助做了哪些信息设备是太普通了，太贵了，太大或太脆工作。这些器件的实际目的将决定他们的技术设计，以及所需的生产资源量和预定寿命操作。这两个关键因素的组合将定义可用的操作能量，物理尺寸，并因此可用存储，计算，和安全功能。特殊用途的装置，尤其是工业设备装置，也可以是复杂的系统，并在它们的多个子组件或子系统。

这些特殊用途的装置中，被称为“装置”，是主要的焦点为了本讨论，而信息设备（即，个人移动设备）仅仅播放朝向本文档中讨论的场景人类演员一个代理的作用。

设备环境。 该设备环境是物理访问和/或“本地网络”对等网络，数字对设备的访问是可行的设备周围的直接物理空间。

本地网络。 A“本地网络”被假定为一个网络，是不同的且从绝缘，但是潜在地桥接至公共因特网，并包括任何短距离无线的无线电技术，其允许设备的对等网络通信。**的“本地网络”这个概念呢 *不* 包括网络虚拟化技术建立这样一个本地网络的错觉，它也做 *不* 包括需要任意两个设备通过公**共网络空间连通，如果他们进入一个对等网络通信关系公共运营商网络。

字段网关（边缘装置）。 字段网关是一个专门的设备，或充当通信引擎和潜在的，作为设备控制系统和设备数据处理毂一些通用服务器的计算机软件。

本场网关的范围包括外地网关本身及其连接的所有设备。顾名思义，现场网关行动外专用数据处理设备，通常是位置的约束。

他们可能受到物理入侵，并有可能限制操作冗余。

字段网关是从单纯的交通路由器不同之处在于它在管理访问和信息流的积极作用，这意味着它是一个应用程序被寻址的实体和网络连接或会话终端。NAT设备或防火墙，相反，没有资格作为场网关，因为他们不明确的连接或会话终端，而是路线（或块）连接或会话通过他们做。

场网关有两个不同的表面区域。一面对被附着到它的设备和代表区段的内部，而另一面的外部各方（例如，云网关），并且是带的边缘。

云网关。 云网关是一个系统，使远程通信从和到器件或场网关，潜在地驻留在几个不同的位点，通过公共网络连接的空间。

云网关处理设备之间以及基于云的后端系统这样的系统的一个联合入站和出站通信，或。

在这里讨论的上下文中，“云”是指到未绑定到相同的位点连接的设备或现场网关，并且其中运行措施防止目标物理访问一个专用的数据处理系统，但不一定是“公共云”的基础设施。

云网关可潜在地映射到一个网络虚拟化覆盖绝缘从任何其他网络流量云网关和它的所有连接的设备或现场网关。

云网关 本身既不的设备控制系统，也不是设备数据的处理或存储设施; 这些设施接口与云网关。云网关的范围包括云网关本身与所有字段网关和设备直接或间接地连接到它沿。

云网关有两个不同的表面区域。一个面向设备和连接到IT领域的网关，而其它的脸部后端服务和潜在的外部各方。

服务。 在本文档的上下文中，服务被定义为任何软件组件或模块，其与设备通过场网关或云网关用于数据收集和分析，以及用于命令和控制的相互作用的接口连接。服务是调解员。他们的行为在他们自身的网关和其他子系统，存储身份和分析数据，自主颁发基于数据的见解或时间表设备的命令和揭露信息和控制能力，以授权的最终用户。

解。 一种用于一个特定的IoT场景的解决方案是系统的构建块，包括所有的用户提供的规则，扩展和码的组合物。它包括所有的数据存储和分析能力特定于溶液的已知范围。

该解决方案进行交互，并与存在的共享企业资源，如CRM或ERP系统或业务线，其他解决方案的其他系统集成。作为工作售票系统，专门为预测性维护解决方案出台将在解决方案范围，但很多时候CRM系统已经到位的客户服务技术支持人员CRM系统。在这种情况下，新的解决方案将与现有的支持工作售票系统，而不是引入一个新的整合。

5.2参考

要了解更多关于Azure的物联网， 访问我们的网站 。

下面的Microsoft产品支持Azure的物联网方案：

Azure的物联网解决方案加速器

Azure的物联网中心

Azure的物联网边缘

Azure存储

Azure的数据湖

Azure的宇宙DB

Azure的SQL数据库

Azure的HDInsight

Azure的数据流分析

Azure的服务总线

Azure的事件集线器

Azure的Web应用程序

Azure的移动应用

Azure的应用程序的逻辑

Azure的通知集线器

Azure的机器学习

Azure的机器学习工作室

双向电力

Azure中的Active Directory

Azure的主要库

欲了解更多的参考和信息支持这个文件，请阅读：

服务辅助通信	http://blogs.msdn.com/b/clemensv/archive/2014/02/10/service-assistedcommunicationfor-connected-devices.aspx
TCP	http://tools.ietf.org/html/rfc793
UDP	http://tools.ietf.org/html/rfc768
DTLS	http://tools.ietf.org/html/rfc6347
AMQP AMQP核心	http://www.amqp.org http://docs.oasis-open.org/amqp/core/v1.0/os/amqp-core-overview-v1.0os.html http://mqtt.org
MQTT	http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.html
CoAP协议	http://en.wikipedia.org/wiki/Constrained_Application_Protocol
OPC基金会	https://opcfoundation.org http://en.wikipedia.org/wiki/OPC_Foundation
的WebSockets	http://en.wikipedia.org/wiki/WebSockets
TLS	http://tools.ietf.org/html/rfc5246 http://tools.ietf.org/html/rfc4279
TPM整合	http://www.trustedcomputinggroup.org/developers/trusted_platform_module
Azure的VPN	http://azure.microsoft.com/services/virtual-network
ExpressRoute	http://azure.microsoft.com/services/expressroute
VPN网关和安全CrossPremises连接	https://docs.microsoft.com/azure/vpn-gateway
Azure的API应用程序	https://azure.microsoft.com/documentation/articles/app-service-apiapps-why-bestplatform
Azure的搜索	https://azure.microsoft.com/documentation/articles/search-what-isazure-search
Azure的地图	https://azure.microsoft.com/en-us/services/azure-maps/
服务织物	https://azure.microsoft.com/services/service-fabric

阿卡	http://akka.io
Akka.Net	http://getakka.net
Azure的批	https://azure.microsoft.com/services/batch
MapReduce的	http://en.wikipedia.org/wiki/MapReduce
猪	http://en.wikipedia.org/wiki/Pig_(programming_tool)
阿帕奇风暴	http://storm.incubator.apache.org
Apache的HBase的	http://hbase.apache.org
Apache的Hadoop的	http://hadoop.apache.org
Apache的蜂巢	http://hive.apache.org
阿帕奇亨利马乌	http://mahout.apache.org
CAP定理	http://www.julianbrowne.com/article/viewer/brewers-cap-theorem https://en.wikipedia.org/wiki/CAP_theorem
Azure的业务连续性技术指导	https://msdn.microsoft.com/library/azure/hh873027.aspx
HA / DR的Azure应用程序	https://msdn.microsoft.com/library/azure/dn251004.aspx
Azure存储复制	https://azure.microsoft.com/documentation/articles/storageredundancy
Azure的SQL数据库业务连续性	https://msdn.microsoft.com/library/hh852669.aspx
弹性云架构	https://msdn.microsoft.com/library/azure/jj853352.aspx
设计一个可扩展的分区策略 Azure的表存储	http://msdn.microsoft.com/library/azure/hh508997.aspx

5.3的SaaS，PaaS和IaaS的指导

客户对如何建立自己的物联网解决方案的多种选择。主要选择客户从建立自己的解决方案时，选择是：1）物联网中心的SaaS - 他们利用Azure的物联网中心，微软的SaaS产品，抽象所有技术的选择，使他们能够专注于他们的解决方案专，2）Azure中的PaaS直接 - 他们

利用Azure的PaaS的组件（物联网中心，CosmosDb，RedisCache，天青流分析等）的直接和整合自己的这些组件形成溶液。这通**常是通过使用端到端天青的IoT解决方案加速器提供的实施例的完成的，和3）IaaS的W / OSS - 他们使用开源软件组件（例如SMA**CK栈，星火，Mesos，阿卡，卡桑德拉和卡夫卡）来引导他们的系统和其托管在IaaS的虚拟机。

每个选项提供不同水平的控制，可定制性/可扩展性和简单性。这些属性都有不同程度的针对不同客户的重要性; 例如，有些客户需要的是高度定制的而另一些可能能够使用什么是“盒子”，仅此而已的解决方案。

我们如下评价每个选项的属性：

选项	控制 （1-5级）	可定制	简单
1。 **物联网中心的SaaS** - 使用物联网中心（我们的SaaS产品）来构建他们的解决方案。	1	2	五
2。 Azure中的PaaS直接 - 使用解决方案加速器，直接建立在 Azure PaaS的物联网服务（例如物联网中心，物联网边缘）的解决方案，因为需要利用其他Azure中的PaaS服务（宇宙，ASA等）。	4	4	2
3。 **IaaS的W / OSS** - 建立在IaaS的，杠杆SMACK或其他OSS组件的解决方案	五 控制的高水平	五 定制	1 复杂

这些选项在快速发展; 即物联网中心在不断增加新的功能，包括更好的定制选项，PaaS的解决方案加速器都从功能的角度和开发者体验的角度，以及控制的IaaS组件添加大幅面功能的能力的变化。

指导

客户需要确定他们是否最好使用我们的SaaS产品或建立自己的解决方案（使用在Azure的PaaS或IaaS的产品）提供服务。我们建议客户开始与物联网中心，然后，如果需要的话，移动到的，因为定制或控制需求的PaaS或IaaS产品。是顶级的问题，可以帮助指导这个决定如下：

- 客户是否云开发经验，和/或需要的东西内置定制他们的公司吗？

  Ø **如果答案是“否”，我们建议客户使用物联网中心化（SaaS）。**
  Ø **如果答案是“是”两个问题 - 客户应该使用的PaaS或IaaS的选项。如果**
  客户没有云的开发经验，我们建议用SI或ISV合作，构建定制的解决方案。

  - WRT使用PaaS的或完全控制通过IaaS的选择，客户应该问他们的解决方案是否需要控制和定制的高水平？
    - 如果不需要控制或定制的较高水平，我们建议客户使用PaaS的组成部分，从物联网Azure的解决方案加速器。

- 如果需要控制或定制的较高水平，我们推荐平台直接通过的IaaS上的建筑。

我们对客户的全面指导是要始终与评估启动 物联网中心 因为它提供了简单的体验，会得到大多数客户推向市场最快的。对于那些需要更严格的控制，并有较大的可定制需求的客户，我们建议使用 通过解决方案加速器Azure中的PaaS服务（ 远程监控，连接工厂，...），然后（在客户需要为云无关的或OSS为中心的）极端定制/控制情况 的IaaS / OSS。