



研究报告

IBM SmartCloud Analytics: 实现系统自主分析

介绍

随着全新自动化“基于分析”运维管理方案的上市，信息系统管理市场发生了巨大的变化。这些程序能够查看信息系统活动；学习系统行为模式；找到所采集的系统数据中的异常并基于分析运维数据时确定的模式和问题预测未来系统行为。此外，这些方案的数据分析速度比人类快得多，从而加快了问题的解决、提高了准确性并且避免了问题。

多年来，系统供应商和管理软件市场专注于简化用户界面和集成管理组合。这种专注提高了信息技术 (IT) 管理效能，并且帮助降低了数据中心的管理成本。但今天，这些供应商正在将注意力转向对运维分析数据进行自动分析上，以进一步减轻 IT 经理/管理员的工作量并减少人为失误。这些新的运维分析产品正与传统的图形用户界面/集成管理软件产品一同使用，以自动排除系统/软件问题并且在这些问题发生前进行预测。

在新一代自动分析工具中，最令人喜欢的一点是，该工具可以使系统自己进行思考。在运维分析工具出现前，故障排除系统/存储/网络问题的负担一直落在已超负荷工作的信息系统经理和管理员身上。这些信息技术 (IT) 专业人员必须制定系统性能指标和阈值，并且他们不得不使用事件监控工具并分析日志文件，才能解决 IT 问题。新一代运维分析工具能够比人类操作员更快地帮助识别问题——从大型非结构化存储库中获得新的见解；通过对准相关结构化和非结构化数据，更快速地开始进行问题根源分析，从而帮助隔离问题；而且这些工具还能通过为操作员提供他们所需的信息，使他们无需手动筛选大量日志文件和其他运维数据，从而有助于更快速地修复问题。由于节省了时间，IT 经理和管理员能够腾出时间来关注价值更高的管理和调整工作。

在本《研究报告》中，Clabby Analytics 讨论了 IBM SmartCloud Analytics 的环境。我们描述了该工具如何与其他 IBM 运维分析工具（监控和日志分析）一同使用，以学习系统行为和识别潜在问题。我们讨论了多家 IBM 客户如何使用 SmartCloud 分析工具降低管理成本并改进服务。并且我们还介绍了 SmartCloud Analytics 与 Splunk（与合作伙伴 Prelert 研发）、Netuitive 和惠普（IT 运维分析工具）相比所具有的优势。读者可以发现它在配置（不再设置阈值）、服务模型建立（有了 SmartCloud Analytics 便不需要）以及问题确定和学习行为方面与其他产品有诸多显著差异。我们对本报告进行了总结并且强烈建议：**要想显著减少 IT 管理人力成本，企业就需要评估 IBM SmartCloud Analytics 的运维分析产品。**

大改变：使用分析工具快速解决问题并预测未来问题

之前，IT 经理和管理员确定性能问题和解决 IT 问题时，需要制定指标、检查事件监控工具器和浏览大量文档才能确定异常。然后围绕可用性、响应时间、延迟、完成时间和指令通路等元素制定指标，并用来这些指标分析系统性能。同时使用活动监控工具追踪系统行为，使用日志文件详细检查系统活动，从而找出解决问题的线索。这些方法共有的一个问题是需要耗费过多的人力：

- **系统调优** – 需要从多个来源收集各项指标，从而在系统调整前创建系统行为综合图；
- **事件监控** – 需要进行编程，从而优先化和追踪数据中心经理希望监控的事件；
- **日志文件** 已成为最大的挑战，因为它们是包含随机的结构和非结构化系统信息的大量记录（有些企业每天收集的日志文件信息超过了 1TB）。需要耗费大量人力在这些日志中寻找线索，才能排除系统和应用程序问题。

但在过去几年，机器分析工具开始进入系统管理市场。现已有多家供应商提供丰富的运维分析产品，包括 IBM、惠普 (HP)、Oracle、SAP 以及若干小公司，比如 Netuitive 和 Splunk。这些供应商中的大部分提供能够单独解决各类问题的软件产品，包括提前避免故障和加快问题解决、提高系统资源关系的可见性和洞察以及优化性能等产品。而 IBM 提供可执行所有这些功能的一整套产品（参见图 1）。

图 1 – 系统管理分析工具的使用



来源 – IBM Corporation, 2014 年 5 月

IBM 运维分析工具组合

IBM 提供一整套能够快速搜索大量结构化和非结构化数据（IBM 的非结构化流数据读取能力尤其强大）、帮助优化 IT 基础架构以及在问题发生前进行预测的运维分析工具。这一组合包括三大元素：

1. *预测*：IBM SmartCloud Analytics - Predictive Insights 技术；
2. *搜索*：IBM SmartCloud Analytics - Log Analysis 产品；
3. *优化*：IBM SmartCloud Monitoring 和 IBM SmartCloud Virtual Storage Center (SCVSC – 有关本产品的详细信息，请参见本 [Clabby Analytics 报告](#))

以下章节将更详细地描述这些产品。

IBM SmartCloud Analytics Predictive Insights

IBM SmartCloud Analytics Predictive Insights 环境是在 IBM 的 Watson 认知计算环境的开发过程中诞生的。Watson（这一认知计算环境由 IBM 实验室开发，用于分析海量结构化和非结构化数据）系统可以表明语言认知计算环境是否能与人类智力相媲美（Watson 在一个系统与人类对决的知识测试活动中击败了兩名“Jeopardy”电视游戏竞赛节目的冠军。）IBM SmartCloud Analytics 建立在 Watson 分析基础之上，使用自定义算法在性能数据中探测系统异常，并预测未来系统/应用程序行为结果。

Clabby Analytics 是 Watson 认知计算环境的忠实“信徒”。在我们看来，Watson 对其所在环境的认识（认知）将越来越深刻，并且将能够使用丰富的高级算法自行进行大量分析。基于 Watson 的系统自行采取纠正措施的能力也将越来越强，同时还能自己发现问题并且在如何纠正这些问题上寻求人类的指导，从而成为对于系统经理和管理员而言极其宝贵的一种工具。Watson 将能够从这些互动中学习，并且最终能够承担越来越多的沉闷的系统管理工作，从而使人能够腾出时间专注于为应用程序开发、业务流程编排以及其他高价值活动做贡献。

在结构上，IBM SmartCloud Analytics Predictive Insights 包含两个部分：

1. 一个合并探测到的异常情况的事件浏览器。
2. 可查看事件起因、导致事件发生的情况以及事件相关性能指标的深入分析视图

本产品的运维原理是通过比较一段时间的系统行为方式（分析时间序列数据）“学习”系统/应用程序行为模式。通过对比“好的”系统行为和“坏的”系统行为，IBM SmartCloud Analytics 能够在问题发生前找出问题，为 IT 经理提供一个识别潜在问题并且在它们变得更加严重前修复这些问题的方法。

IBM SmartCloud Analytics 环境的真正特别之处在于它能够充分利用 IBM InfoSphere Streams 环境。该环境可以将数据传送到应用程序，从而实现结构化和非结构化数据的实时读取。由于使用数据流技术，环境无需对所有被传送的数据进行存储和分析。相反，数据经过一个系统，而系统仅捕获和分析相关信息（比如异常）。这意味着只会捕获相关数据并根据这些数据执行操作，因此在寻找异常时，宝贵的系统循环不浏览日志和其他数据源。此外，这还可以节省数 TB 的存储空间（同样是因为只采集和保存相关数据）。

与主要竞争对手相比，IBM 在数据流技术方面尤其强大。凭借其数据流传输能力和一大套预测分析算法，IBM 在预测分析工具市场具有强大的竞争优势。

IBM SmartCloud Predictive Insights 的另一个特别之处在于它无需长期维护或操作员设置阈值，从而减轻了部署和管理工作负担。此外，它也无需设置预测、描述和决策服务模型。换言之，IBM SmartCloud Predictive Insights 的部署和运维无需大量的操作员干预和长期维护。

IBM SmartCloud Analytics – Log Analysis

IBM 将其 SmartCloud Analytics - Log Analysis 产品描述为一个“采集大量模糊的非结构化数据并且通过分析将其转化成可用智能”的环境。该产品可通过日志和指标简化异常检测。

产品本身包含三大组成部分：

1. 启动各项活动的仪表板；
2. “洞察套件”（提供一系列展示工具，帮助经理和管理员查看系统问题的专用环境扩展插件）；
3. 通过集成高级文本分析技巧的智能支持文件集成和专家建议端口集成。

这一产品的出色之处在于可用于分析丰富的数据类型，包括：

1. *监控和指标* – 比如关联日志搜索结果中的指标；
2. *问题异常/检测* – 比如基于从日志中观察到的趋势进行检测和提醒系统经理；
3. *事件* – 比如搜索事件数据和检测季节性趋势
4. *发现和应用程序性能管理* – 比如使用拓扑结构和配置内容优化要搜索的日志和文档范围，从而帮助找出性能问题；
5. *服务台解决方案* – 比如搜索和分析服务工单或使用工单内容搜索事件日志和文档，以寻找特定类型的活动。

SmartCloud Analytics Log Analysis 使 IT 经理和管理员能够以分析工具代替人类脑力，更快速地隔离和修复问题。系统能够筛选大量结构化和非结构化信息，并且指导经理/管理员解决问题。在对其他日志分析产品进行评估后，我们总结出 IBM SmartCloud Log Analysis 产品的两个最大不同之处在于与其他 IBM 运维分析工具/系统管理产品的集成度以及与其他故障排除工具协作进行根源分析的能力，包括指标、配置、信息、事件、日志、痕迹和拓扑产品。其他不同之处包括定价模式（基于平均数据消费模型而非高峰时间的使用量）、其高级文本分析能力（业内唯一一个使用高级文本分析工具从非结构化来源获取洞察力的产品）以及与 IBM 大数据平台的链接。

IBM SmartCloud Monitoring 和 SmartCloud Virtual Storage Center

云平台管理牵涉到系统、网络知识以及构成既定云环境的存储设备。为了实现高效协作，这些资源需要加入到一个可通过轻松的分析追踪性能或其他问题根源的互连交叉系统云架构。IBM SmartCloud 管理工具（包括 IBM SmartCloud Application Performance Management 环境、SmartCloud Monitoring、SmartCloud Provisioning 以及 SmartCloud Control Desk）可协同工作，提供统一的云端应用程序活动视图，从而进行应用程序故障排除和调整，并且在必要时提供采取纠正措施所需的工具。SmartCloud Virtual Storage Center 实现了异构存储的集中管理，同时提供存储虚拟化功能，因而可以根据工作要求按需集中、调配存储资源和取消存储资源调配。

IBM SmartCloud Monitoring

云基础架构能够给用户带来诸多益处，比如灵活性、扩展性、更高的效率和更低的成本。有效的容量规划是实现这些益处并挖掘云平台全部潜能的关键之一。容量规划工具有助于减少硬件和软件许可证成本并且节省能源成本和楼层空间要求。优化资源和工作负载分配可提高工作负载性能。

IBM SmartCloud Monitoring 管理物理和虚拟云基础架构，并且提供性能分析功能和容量规划工具以优化工作负载性能和可用性。该工具可通过监视面板监视主机性能和资源的可用性。这款工具内含用户可自行配置的政策型分析工具，从而可进行智能工作负载分配，以确保快速、高效地处理工作负载。IBM SPSS 预测分析工具包含在 IBM SmartCloud Monitoring 中，可支持非线性预测和数据建模，从而通过检查性能数据的复杂关系更好地预测趋势、提高预测的准确度并且制定更明智的决策。

SmartCloud Monitoring 的“假设”容量分析可以对基础架构改变所产生的影响建模，从而降低风险并有助于提高可用性。优化利用当前容量和规划增量资源提高可降低能源和服务成本。IBM SmartCloud Monitoring 通过实现最高效的虚拟机密度和最高的虚拟化工作负载性能，为企业的云技术投资保驾护航。

IBM SmartCloud Virtual Storage Center

IBM SmartCloud Virtual Storage Center 包含存储管理程序、集中管理以及用于即时备份和恢复的应用程序感知快照。管理程序隔离底层硬件，并支持（比如输入/输出缓存、精简配置、重复数据消除、压缩等）在各种异构存储上交付和管理服务，从而将几乎所有供应商（EMC、惠普、戴尔、Sun、日立、NEC、IBM 等）的所有存储平台的存储池化并且显著提高存储利用率。

无论选择何种存储阵列，都能一致提供输入/输出缓存、精简配置、压缩、自动分层、应用程序感知快照/镜像以及避免流动性驱动的中断等服务，从而提高效率和简化管理。通过简单易用的图形用户界面实现的集中管理除存储管理程序之外还拥有多项宝贵功能。分析简化了管理以及自动配置、存储分层和数据迁移。

比较预测分析产品

我们认为与 IBM SmartCloud Predictive Insights 产品构成直接竞争关系的产品包括：

1. Splunk 的 Prelert
2. 带 IT 运维分析功能的惠普工具
3. Netuitive。

Splunk/Prelert

Splunk 是一款简单易用、便于安装的软件平台，可从几乎任何来源实时采集和索引任何 IT 流媒体机器和历史数据，包括事件日志、网络服务器日志、应用程序日志、实时应用程序日志、网络馈送、归档文件、移动设备、传感器等。最新版本还提供数据关联、视觉化、报告和分析功能。预测分析面板（来自 PRELERT）搜索 Splunk 数据中的异常，以提供结果统计信息，并且基于预设阈值确定可能表示潜在问题的异常。

通过 Splunk 框架和图形用户界面，Prelert Anomaly Detective 的自学技术可根据在海量数据（通过 Splunk 聚合的）中观察到的正常行为建立模型。模式检测算法将历史和实时数据分成离散活动，然后使用统计分析计算既定关键绩效指标 (KPI) 值的正常范围。范围外的值为异常值，相关结果可以表明可能造成性能问题/安全威胁的范围外数值的严重性或影响。通过确定活动的因果链，Prelert 还能识别问题根源。

优势:

- Splunk 支持多种计算机数据；
- 自学型 Prelert 分析工具；
- 一些客户将 Splunk 视为“同类中的最佳”解决方案。

劣势:

- 两家不同公司的两款产品延长了销售和实施流程
- 无法关联多种类型的指标并预测异常行为
- 集成要求可能产生新功能访问的滞后
- Prelert 的规模（20 人）使其容易成为收购目标
- 点解决方案没有与更广泛的管理组合相集成
- Prelert 与 CA Technologies 的原设备制造商关系（Prelert 技术是 CA Application Behavior Analytics 的基础）
- 不支持第三方监控工具

HP Operations Analytics

HP Operations Analytics 是 HP Business Service Management 的一项新功能，可提供结构化和非结构化数据的高级分析智能。该智能包括：计算机数据、日志、活动、拓扑结构和性能统计。HP Operations Analytics 采集和分析来自各种惠普产品的运维数据，包括 HP ArcSight Logger 和 HP Business Service Management 组合以及 Splunk 等第三方数据资源，并且包含以下功能：

- *IT 搜索* – IT 搜索功能使 IT 运维部门能够使用简单的查询语言在采集的所有 IT 运维数据中搜索。该语言可方便和加快在日志、指标、事件和拓扑的大数据存储数据集中搜索并且通过分析面板显示内容相关的信息。
- *可视化分析* - 使用户能够轻松检查具体时段异常或按地点或部门检查异常。可按照角色在视图中显示数据：
 - 运维经理视图可分析事件数据、日志和指标；
 - 应用程序、系统和网络管理团队视图可通过引导分析进行引用程序/基础架构问题隔离，并有 IT 搜索功能，可搜索隔离问题的特定使用案例；
 - 数据分析师和专家视图可创建针对指标、事件、拓扑和日志的专门查询
- *问题排除指引* - 问题排除指引轻松回顾历史服务性能，以识别和隔离确切的时间和故障源。一旦出现性能问题，就会启动环境相关分步流程。首先会建议相关搜索主题。根据所选择的主题，将列出优先的潜在问题，并且在确定问题后，提出修复建议。

HP Service Health Analyzer 是一款能够关联 HP Business Service Manager (BSM) 数据以检查各项性能指标的预测分析工具。这一关联加上高级分析和算法，可预测问题并且根据对业务的影响确定问题的优先性。该工具具有自学能力，还可以分析历史数据，从而自动创建实际阈值，然后结合异常数据预测潜在问题。HP SHA 目前未与 HP Operations Analytics 集成。

优势:

- 基于角色的视图和根据环境提出的专家建议。

劣势:

- 只有 HP SHA 提供自学功能
- 必须手动设置阈值/服务模型
- 整个管理产品系列的集成度低
- 对第三方监控工具的支持有限
- HP SHA 和 HP Operations Analytics 均以 HP BSM 为前提。使用 BSM 拓扑映射建立关联
- 预测准确性取决于 BSM 模型版本的新旧

Netuitive

Netuitive 是预测运维分析工具的先驱。*Clabby Analytics* 于 2010 年 7 月首次评估了该公司并且在 2012 年 2 月发布了更新（参见[此处](#)）。**Netuitive** 策略一直从公司现有的监控工具采集数据，包括服务器、存储、网络和来自 CA Technologies、IBM、BMC 等公司的应用程序，然后对这些数据进行合并、关联和分析。**Netuitive** 还能与 **Splunk Enterprise** 相整合，提供丰富的计算机数据源。

Netuitive 使用数学相关性和统计引擎“自学”IT 基础架构，从而采集和关联数据并且建立行为规范、区分任何既定情况中的“正常”和“异常”、实现端对端前瞻性管理。所采集的数据越多，**Netuitive** 就会变得越智能。基于这一分析，**Netuitive** 可以提前数日确定异常并且预测潜在的性能问题、加快根源分析、问题解决并减少误报。**Netuitive** 可充分利用各项性能指标，应用关联度算法并且评出综合健康分数。通过这一方式，这些性能指标一起回答了“我的 IT 环境是否健康”这一问题。**Netuitive** 包括适当行为规范，可显示一周中的当日，一日中的这一时刻和周期性活动的正常行为规范，从而管理环境中的所有关键指标。根据业务要求，还可以设置 **Netuitive**，使特定性能属性权重高于其他性能属性。

其他功能包括：

- 当发现性能问题时，性能辅助自动进行诊断和根源隔离，并且可以追溯到各指标；
- 关联辅助追踪商业服务组件和性能指标之间的关系，以帮助进行根源分析以及确定潜在性能问题如何影响用户或应用程序；
- 智能搜索能力能够查询数百万个数据点，根据元素属性或性能级别识别系统；
- 最佳实践模板使用专家知识和统计技巧帮助确定哪些指标是特定基础架构组件的最佳性能指标。

优势:

- 第三方监控工具支持 – 与客户所拥有的工具合作；
- 支持多种数据来源；
- 自学分析的历史悠久；
- 通过多项指标的关联性确定业务影响

劣势

- 运维分析工具的点解决方案；无法与更广泛的组合集成
- 要求构建和维护商业服务模型，以建立各指标间的长期关联
- 需要时间和资源来构建和维护服务模型；不是“开箱即用”的解决方案

所有此类产品与 **IBM SmartCloud Predictive Analytics** 的并排比较请参见下一页的图 2。

图 2: 一流预测分析环境的比较

	IBM Predictive Insights	HP Operations Analytics	Splunk/Prelert	Netuitive
问题确定	有	有	有	有
根源分析	有, 通过 SmartCloud Analytics Log Analysis 实现	有	部分产品	有
设置阈值/服务模型	自动设置	必须由管理员设置	自动通过 Prelert 设置	需要服务模型和配置
第三方监控工具支持	有, 通过嵌入的中间工具和可选的 IBM Data Stage ETL 工具实现。	没有	没有	有, 通过适配器和 API 实现
搜索诊断	有, 与 SmartCloud Analytics - Log Analysis 集成	有, 通过 IT 搜索实现	没有	有, 通过智能搜索实现
自学	有	没有	有	有
根据环境提出的专家建议	有, 通过 Expert Advice 应用程序实现	有, 通过 Guided Troubleshooting 实现	没有	没有
附加优化	广泛支持多项技术 (Netcool Operations Insight、应用程序性能管理)	与其他产品的集成性不佳 (比如 HP Service Health Analyzer)	点解决方案	点解决方案
其他	非常全面的用户文件广泛, 深入			

来源: Clabby Analytics, 2014 年 4 月

总结观察

IT 经理和管理员现在可以使用新一代分析工具了。新一代分析工具可确保应用程序和服务的可用性, 可以比人类更快地发现和纠正问题, 可以帮助找出解决问题所需的行动, 并且可以为经理/管理员增强系统资源关系的可见性。此外, 这些工具还能够进行性能上的自我调整, 自动监控、控制和修复系统错误, 并且还能够能够在故障发生前预测硬件和软件层面的系统故障, 然后采取补救措施。云端的工作负载、基础架构性能和基础架构可用性通过预测性容量规划工具以及虚拟化、分层、快照和压缩技术等存储管理功能加以优化。

IBM SmartCloud Analytics: 自身便可实现系统分析

运维分析工具可以大幅提高需要耗费大量人力的事件监控/日志搜索方案的效率，并且可以非常显著地帮助降低管理成本，同时提高服务等级（因为加快了故障排除和高级预测性维护的速度）。不仅如此，预测性分析工具可以支持容量规划以及工作负载和基础架构优化，实现了资源利用的最大化并且降低了硬件、软件、能源和楼层空间成本。那些将前瞻性预测管理作为实现成本、可用性和性能优势的手段进行评估的企业，应仔细调查 IBM 的整套 [SmartCloud Analytics 产品](#)。IBM 对于一系列分析工具的支持及其深入、专业的分析使其在这一领域的供应商中脱颖而出。

Clabby Analytics
<http://www.clabbyanalytics.com>
电话: 001 (207) 846-6662

©2014 Clabby Analytics
保留所有权利
2014年6月

Clabby Analytics 是一家独立的技术研究和分析机构。与诸多其他研究机构不同，我们会提出特定的观点并鼓励我们的读者找到相反意见，然后对两种观点加以平衡，最后决定行动方案。

Clabby Analytics 开展的其他研究和分析请参见：
www.ClabbyAnalytics.com。