

# Estrutura de gerenciamento e monitoramento de eventos para HPC Ambientes usando ServiceNow e Prometheus

Nitin Sukhija  
Departamento de Ciência da  
Computação Slippery Rock  
Universidade da  
Pensilvânia Slippery Rock,  
PA nitin.sukhija@sru.edu

Daniel Gens  
NERSC  
Lawrence Berkeley National  
Laboratory Berkeley,  
CA dygens@lbl.gov

Tony Quan  
NERSC  
Lawrence Berkeley National  
Laboratory Berkeley,  
CA twquan@lbl.gov

Elizabeth Bautista  
NERSC  
Lawrence Berkeley National  
Laboratory Berkeley,  
CA ejbautista@lbl.gov

Siqi Deng  
Laboratório Nacional NERSC  
Lawrence Berkeley  
Berkeley, CA  
siqideng@lbl.gov

Basil Lalli  
NERSC  
Lawrence Berkeley National  
Laboratory Berkeley,  
CA bdlalli@lbl.gov

Owen James  
NERSC  
Lawrence Berkeley National  
Laboratory Berkeley,  
CA o1james@lbl.gov

Yulok Lam  
NERSC  
Lawrence Berkeley National  
Laboratory Berkeley,  
CA yllam@lbl.gov

## ABSTRATO

O desafio de monitoramento e gerenciamento de resposta a eventos de uma instalação de computação de alto desempenho cresce significativamente à medida que as instalações empregam e orquestram sistemas e infraestrutura mais complexos e heterogêneos. À medida que os componentes computacionais que abrangem o sistema de instalação de HPC aumentam, as experiências da equipe computacional aumentam a fadiga de alerta devido aos falsos alarmes e ruídos relacionados a eventos semelhantes gerados pelas ferramentas de monitoramento. O Centro Nacional de Computação Científica de Pesquisa Energética (NERSC) no Laboratório Nacional Lawrence Berkeley (LBNL) começou a abordar as questões de duplicação de alertas e remediação de alertas. No entanto, mais automação e integração são necessárias para coletar, agregar, correlacionar, analisar, gerenciar e visualizar a escala de eventos que serão gerados pelas infraestruturas de computação híbrida emergentes. Neste artigo, apresentamos uma estrutura de gerenciamento e monitoramento de eventos que aborda as necessidades operacionais dos futuros sistemas pré-exaescala no Centro Nacional de Computação Científica de Pesquisa em Energia (NERSC) do Laboratório Nacional Lawrence Berkeley. A estrutura integra a infraestrutura de monitoramento e notificação de operações (OMNI) no NERSC com as plataformas Prometheus,

para ajudar a identificar, diagnosticar e resolver incidentes em tempo real, bem como realizar análises pós-incidente mais completas, habilitadas pelos painéis intuitivos que fornecem um único painel de console para um gerenciamento eficiente de operações e monitoramento proativo em tempo real.

## CONCEITOS DE CCS

• **Sistemas de informação** e **Data centers**; • **Geral e referência** Projeto; • **Metodologias computacionais** e Abordagens de aprendizado de máquina; • **Hardware** e **Recuperação de falhas, manutenção e autorreparo**;

## PALAVRAS-CHAVE

Resposta do Evento; Monitoramento de Dados; Serviço agora; Prometeu; Grafana; Visualização

### Formato de referência ACM:

Nitin Sukhija, Elizabeth Bautista, Owen James, Daniel Gens, Siqi Deng, Yulok Lam, Tony Quan e Basil Lalli. 2020. Estrutura de monitoramento e gerenciamento de eventos para ambientes HPC usando ServiceNow e Prometheus. Na 12ª Conferência Internacional sobre Gestão de Ecossistemas Digitais (MEDES '20), 2 a 4 de novembro de 2020, Evento Virtual, Emirados Árabes Unidos. ACM, Nova York, NY, EUA, 8 páginas. <https://doi.org/10.1145/3415958.3433046>

## 1. INTRODUÇÃO

À medida que avançamos em direção ao próximo marco da supercomputação exaescala, a complexidade e a heterogeneidade da infraestrutura do centro computacional que permite a capacidade de pico desses exaflops também aumentará substancialmente. Além disso, com esse crescimento da complexidade da infraestrutura, a escala de eventos gerados por esses sistemas tende a

A permissão para fazer cópias digitais ou impressas de todo ou parte deste trabalho para uso pessoal ou em sala de aula é concedida sem taxa, desde que as cópias não sejam feitas ou distribuídas com fins lucrativos ou vantagens comerciais e que as cópias contenham este aviso e a citação completa na primeira página. Os direitos autorais de componentes deste trabalho pertencentes a outros que não a ACM devem ser respeitados. Abstraindo com crédito é permitido. Para copiar de outra forma, ou republicar, postar em servidores ou redistribuir para listas, requer permissão específica prévia e/ou taxa. Solicite permissões de [permissions@acm.org](mailto:permissions@acm.org). MEDES '20, 2–4 de novembro de 2020, Evento Virtual, Emirados Árabes

Unidos © 2020 Association for Computing Machinery.  
ACM ISBN 978-1-4503-8115-4 / 20/11...\$  
15,00 <https://doi.org/10.1145/3415958.3433046>

crescer exponencialmente, aumentando assim os desafios de gerenciamento e segurança das instalações computacionais que tratam desses eventos [21]. Este desafio é exacerbado pelo número de imprevisíveis e fatores altamente dinâmicos, como 1) tamanho diversificado, processamento de dados, cargas de trabalho em nuvem, requisitos de virtualização do data center; 2) múltiplas considerações ambientais de um data center, como como, temperatura, ar condicionado, estado da unidade de distribuição de energia, umidade, fluxo de ar etc.; e 3) inúmeras métricas de servidor e sistema de armazenamento, como CPU, RAM, consumo de energia e muitas mais fatores. Para enfrentar os desafios acima mencionados, vários soluções de monitoramento foram propostas e implantadas com sucesso por centros computacionais modernos para gerenciar e monitorar esses complexos de computação distribuída de alto desempenho ambientes [24].

Dada a tendência inexorável de exigir a utilização de sistemas de automação e orquestração para antecipar diversos problemas com componentes de sistemas de computação heterogêneos operando em ambientes de produção que mudam dinamicamente, o gerenciamento dessas facilidades computacionais está aumentando em complexidade. o novas soluções de monitoramento atualmente implantadas e propostas para os centros computacionais pós-peta-escala consideram a escalabilidade e manutenibilidade juntamente com baixa sobrecarga como os atributos de design mais importantes de uma solução de gerenciamento bem-sucedida. Além disso, uma solução de monitoramento eficaz requer a coleta de dados heterogêneos **informações dos sistemas complexos e fontes que abrangem** uma infraestrutura computacional, analisando esses dados de escala extrema, e compilar um plano de ação para responder a incidentes em condições quase reais tempo com tempo de resposta mínimo. Assim, a eficiência operacional em tais ambientes de alto desempenho em rápida mudança exigem equipe do centro computacional para coletar, gerenciar e analisar quantidades exponenciais de dados [28]. As instalações de computação agora enfrentam milhares ou mesmo milhões de eventos ocorrendo em sua infraestrutura por dia, o que se traduz em fadiga extrema da equipe de operações e dificuldades na priorização eficaz de eventos e na separação **sinais do ruído. Isso se traduz na necessidade de integração** de um componente de gerenciamento de eventos para o moderno monitoramento soluções para diminuir o ruído relacionado a eventos semelhantes gerado por várias ferramentas de monitoramento e para permitir a correlação dos eventos para facilitar alertas e incidentes acionáveis por meio do uso de algumas técnicas de modelagem preditiva [33] [27].

Com a proliferação dos modelos de computação híbrida e/ou a compactação de muitos serviços complexos que englobam as operações dos centros computacionais modernos, a coleta de dados e métricas de saúde das infraestruturas computacionais juntamente com o arquivamento dos dados coletados dados em tempo real é de suma importância para reduzir a resposta tempo ou tempo de inatividade devido às ameaças físicas e digitais. Atualmente, existem muitos sistemas de gerenciamento e monitoramento de data centers ferramentas integradas com uma variedade de complexidades e equipadas com recursos avançados de alerta. No entanto, existem muito poucas ferramentas que enfatizam significativamente uma visão holística do centro de computação monitoramento reunindo diversos dados em diferentes formatos de ambientes de computação e respondendo proativamente às ameaças e invasões em tempo real, fornecendo informações atualizadas por meio de visualizações de painel de todas as operações do centro. Portanto, com o crescimento contínuo na escala e complexidades do arquitetura e políticas híbridas do centro computacional, as abordagens de monitoramento existentes envolvendo Nagios [25], Spiceworks [14],

Icinga [5], ou Zabbix [16] logo se tornará obsoleto. As soluções abrangentes de monitoramento facilitam uma baixa sobrecarga, escalável **e infraestrutura integrada de coleta e análise de dados operacionais para ingerir diversos dados e métricas, correlacionar eventos e** fornecendo relatórios e gerenciamento de incidência em tempo real, em alcançar a excelência operacional nos centros computacionais. Recentemente, o Laboratório Nacional Lawrence Berkeley (doravante referido ao Laboratório de Berkeley) o Centro Nacional de Computação Científica de Pesquisa em Energia (NERSC) desenvolveu e implantou um **Infraestrutura de Monitoramento e Notificação, OMNI se reunirá próximo** dados e métricas em tempo real referentes à saúde da computação sistemas, bem como para arquivar os dados de escala extrema para facilitar **monitoramento dessa orquestração e automação altamente complexas** plataformas. A infraestrutura OMNI construída em tecnologias de código aberto, como o Elastic Stack [3][20] permite o armazenamento de dados de monitoramento de séries temporais de streaming em tempo real de várias fontes, incluindo sistemas de computação no NERSC e sua infraestrutura computacional de suporte, sensores ambientais, sistemas mecânicos e mais. Neste artigo, propomos uma resposta automática a eventos **framework de gestão que integra a infraestrutura OMNI** com as plataformas Prometheus [12], Grafana [4] e ServiceNow [13] para permitir o monitoramento proativo de alertas, inteligência preditiva para reduzir o ruído de eventos, análise de causa raiz, visualizações de dados e **painel de vidro único para gerenciamento de desempenho. O proposto** infraestrutura ajudará a reduzir o tempo médio de reparo (MTTR) entender a causa raiz dos problemas operacionais em tempo real através de alertas acionáveis e na melhoria da disponibilidade do serviço do centro computacional eliminando proativamente interrupções e outros questões críticas no NERSC. O projeto da estrutura abrangente de monitoramento e resposta a eventos proposta é motivado pela **tendências emergentes de monitoramento eficiente e gerenciamento escalável dos sistemas de computação heterogêneos atuais e futuros, como** como, Perlmutter em NERSC [11].

O resto do artigo está organizado da seguinte forma. Uma revisão do trabalho relacionado e gerenciamento de centro computacional de escala extrema desafios é apresentado na Seção 2. A concepção e organização **da estrutura abrangente de resposta e monitoramento de eventos** é descrito na Seção 3. O gerenciamento e monitoramento do evento os fluxos de trabalho estão resumidos na Seção 4, seguidos pelas conclusões e possíveis direções futuras na Seção 5.

## 2 ANTECEDENTES E TRABALHOS RELACIONADOS

Com o crescimento da eficiência dos data centers, o gerenciamento e monitoramento desses ambientes híbridos de alta complexidade é tornando-se cada vez mais difícil e penoso. Como a margem para erro nas instalações computacionais atuais e futuras está se tornando perigosamente finas, as soluções de monitoramento tradicionais são limitadas por sua capacidade de dimensionar e fornecer informações em tempo real, precisas e monitoramento de grãos da infraestrutura de computação [17]. Além disso, como a computação em nuvem e inovações e complexidades em hiperescala estão substituindo os sistemas computacionais tradicionais, uma solução de monitoramento para os data centers modernos emergentes requer integração da combinação certa de ferramentas e tecnologias que facilitam não apenas o monitoramento, mas também a identificação e priorização dos serviços de infraestrutura e eventos de segurança para responder proativamente

Gerenciamento de eventos usando ServiceNow e Prometheus

ou evitar totalmente incidentes potenciais [31]. Assim, o gerenciamento superior dos centros computacionais exige o emprego de um plataformas de monitoramento e gerenciamento de eventos que incorporam a principais considerações de projeto, como:

**Coleta de Eventos em Tempo Real:** É imperativo coletar vários dados de eventos e métricas pertencentes não apenas à integridade do sistemas de computação e a infra-estrutura subjacente, mas também em relação aos fatores ambientais que abrangem um ambiente computacional. Centro. Esses diversos conjuntos de dados e métricas de eventos representam o volume, desafios de velocidade, variedade e veracidade para um centro computacional para fornecer uma solução eficiente de monitoramento em tempo real. Para abordar o questões acima mencionadas, há uma necessidade de mecanismos escaláveis para coletando eventos de alto volume provenientes de várias fontes, como como logs do sistema, estatísticas de trabalho e rede, dados do ambiente e mais. Assim, as plataformas de monitoramento e gerenciamento de eventos devem ser capaz de coletar eventos de escala extrema todos os dias, o que se traduz em taxas de ingestão de eventos excepcionalmente altas por segundo, o que apresenta desafios significativos de investimento e validação.

**Correlação e Análise de Eventos em Tempo Real:** Com o aumento na complexidade dos centros computacionais modernos, a produção de dados de eventos também atingiu uma escala exponencial. Além disso à recolha e recolha de eventos, a interpretação deste dados de eventos díspares são igualmente importantes para executar alertas, notificações e correlações para priorizar e responder aos falhas críticas potenciais ou exceções de nível de serviço. As notificações de eventos podem crescer a uma taxa de spam gerando vários alertas para o mesmo evento. Assim, os mecanismos de agrupamento e inibição precisam a ser construído na plataforma de monitoramento e gerenciamento de eventos para **reduzir o ruído gerado por múltiplos eventos e correlacionar o** eventos para agir em dados de eventos coletados para responder em tempo real com ações corretivas com base na gravidade dos alertas.

**Relatórios em tempo real:** Um dos principais objetivos de um monitoramento e plataforma de gerenciamento de eventos é facilitar e agilizar o processo de tomada de decisão da instalação ou organização computacional relatar informações precisas e em tempo real de maneira útil e significativa. No entanto, fornecer uma visão consolidada dos resultados da análise sumarizada com base na agregação e correlação de os principais alertas em tempo real para diferentes níveis organizacionais podem ser extremamente desafiador. Portanto, a plataforma de monitoramento e gerenciamento de eventos deve facilitar um único painel de painéis de vidro que ajudará a equipe computacional a transformar a infraestrutura eventos em alertas acionáveis e incidentes associados para realizar análise de causa raiz em tempo quase real, levando a decisões informadas, evitando assim erros, ineficiência, tempo de inatividade e desperdício de capacidade do recurso computacional.

**Escalabilidade:** Os centros computacionais de hoje são compostos por nós heterogêneos, rede, aplicativo, servidor e elementos de infraestrutura. Com o número de eventos gerados pelos diversos componentes que compõem o centro computacional crescendo exponencialmente, a necessidade de maior eficiência, portanto, as tecnologias de scale-out estão se tornando mais populares do que o scale-up abordagens. As abordagens de expansão atualmente oferecem maior volumes de aplicativos e recursos de armazenamento maiores a um custo reduzido despesas, maior disponibilidade e tarefas de gerenciamento reduzidas em comparação com as abordagens de scale-up envolvendo uma máquina monolítica para monitorar e gerenciar o centro computacional.

MEDES '20, 2 a 4 de novembro de 2020, evento virtual, Emirados Árabes Unidos

**Automação e Integração:** Com um volume explosivo de "big data" de eventos não estruturados gerados a partir de todos os componentes de infraestrutura, rede e aplicativos subjacentes dos centros computacionais modernos de hoje, há uma tremenda pressão sobre o equipe de operações do centro para coletar, normalizar, armazenar e correlacionar eventos para manter o centro de computação ágil e garantir a integridade a infraestrutura otimizada por meio de métricas de tomada de decisão superiores e regras de alerta. Assim, a automação de muitas operações de rotina atividades de gerenciamento são necessárias para reduzir a carga nas operações pessoal e alcançar um monitoramento e gerenciamento de eventos mais eficazes .

A demanda por agilidade e implantação em escala exige a plataformas de monitoramento e gerenciamento de eventos para empregar aplicativos interfaces de programa (API) para integrar instalações e operações de TI, facilitando assim a comunicação eficaz e a colaboração levando para desempenho, capacidade e disponibilidade otimizados.

**Alta Disponibilidade e Inteligência Operacional:** O propósito da plataforma de monitoramento e gerenciamento de eventos é auxiliar na alcançar inteligência operacional empregando várias métricas operacionais e abordagens de aprendizado de máquina para identificar e priorizar eventos do centro computacional e auxiliar proativamente na compreensão de questões de desempenho e no planejamento de capacidade. Além disso, além de resolver o desempenho da infraestrutura crítica questões, a plataforma de monitoramento e gerenciamento de eventos também em alcançar alta disponibilidade dos serviços do centro computacional **detectando os incidentes de causa raiz e acionando** soluções de remediação.

Ao longo dos anos, várias soluções de monitoramento foram propostas e implantados nos ambientes computacionais. Além disso, muitos estudos envolvendo soluções de gerenciamento de eventos têm sido pesquisados e relatado. No entanto, apenas alguns estudos publicados ilustram a integração do monitoramento em tempo real e gerenciamento de eventos para fornecer uma estrutura abrangente com base nas considerações acima . Ao longo de 15 anos, a ferramenta Nagios [8] tem sido utilizada por muitos organizações para monitorar o status dos dispositivos de rede e seus serviços e para auxiliar na resolução de questões críticas relacionadas a eles. Até embora o Nagios seja a solução líder para monitoramento, visualização, e alertando problemas críticos, a versão estática do Nagios é limitada **pela escalabilidade e sua gestão torna-se extremamente difícil** em ambientes dinâmicos. O monitoramento da rede Spiceworks tool [14] é como o Nagios em termos de escalabilidade limitada, apesar de apresentar painéis dinâmicos extraordinários. Além disso, Spiceworks A ferramenta também não tem suporte para o protocolo Simple Network Management Pro (SNMP) 3.0 [22]. Algumas outras principais ferramentas de monitoramento incluem Monitor de rede Paessler PRTG [10], Zabbix [16] e Icinga [5], **que vem equipado com muitas personalizações, como bom** Interface de usuário da Web e suporte de API, mas ainda são limitados em desempenho e eficiência quando empregados em implantações de grande escala composto por milhares de dispositivos e sensores.

Existem apenas alguns projetos recentes que envolvem integração de gerenciamento de eventos e infraestrutura de monitoramento. O serviço Management Project [18] é um dos projetos que substituiu o Rem edy pela infraestrutura ServiceNow para o gerenciamento de eventos e monitoramento dos servidores disponíveis e serviços oferecidos no Centro de Informática CERN IT. Os pesquisadores em [29] propuseram uma sistema de monitoramento de eventos, chamado Trumpet, que utiliza os recursos da CPU e a programação do host final para monitorar eventos de rede,

relatar e identificar a causa raiz do congestionamento na rede em escalas de tempo de milissegundos . No entanto, os autores fornecem apenas o design de um gerenciador de eventos com sua aplicabilidade em ambientes reais de data center. Além disso, um novo processamento de eventos de nó multicamada (MNEP) para melhorar a eficiência energética do data center junto com o gerenciamento de eventos o monitoramento da infraestrutura física foi apresentado em [28]. Juntamente com os estudos acima mencionados e projetos, vários projetos, tais como, infraestrutura Trinity Monitoring [23], OpenLorenz [9] e outros [30] também foram implantados para facilitando soluções de monitoramento escaláveis. Em nosso trabalho anterior [34], propusemos uma solução abrangente de monitoramento escalável. Para melhor do nosso conhecimento, em comparação com os mencionados soluções, a estrutura de monitoramento e gerenciamento de eventos proposta reúne dados de uma fonte de dados Prometheus, coleta os dados no OMNI e cria um mecanismo de rastreamento no ServiceNOW. este oferece uma solução completa de gerenciamento de eventos e monitoramento em tempo real escalável , permitindo inteligência operacional ao coletar dados de eventos diversos, normalizando alertas, armazenando métricas operacionais e de integridade, correlacionando eventos e utilizando aprendizado de máquina em abordagens para descobrir as causas-raiz do problema computacional centralizar questões críticas e para a tomada de decisão proativa. Nossa proposta de monitoramento e estrutura de gerenciamento de eventos facilitará inteligência operacional em tempo real, auxiliando assim na otimização do desempenho, disponibilidade e planejamento de capacidade dos atuais infra-estrutura de sistemas de computação Petascale e também ajudará na planejando nossas futuras implantações de sistemas.

**3 DESIGN E ARQUITETURA DO EVENTO GESTÃO E MONITORAMENTO ESTRUTURA**

Nesta seção detalhamos o design e a arquitetura do escalável infra-estrutura que será empregada para fornecer monitoramento e gerenciamento de eventos do sistema de computação heterogêneo atual e futuro , rede e aplicativos no NERSC computacional instalação. O gerenciamento e monitoramento abrangentes de eventos estrutura proposta neste artigo (como mostrado na Figura 1) integra várias plataformas de código aberto, como Prometheus [12], Grafana [4], Kafka [1] e ServiceNow [13] com a estrutura OMNI existente para permitir a coleta automática de eventos, correlação de eventos, previsão de alertas, redução de ruídos de alertas, análise de causa raiz, relatórios de métricas operacionais e painéis de exibição de painel único para aprimorar a inteligência operacional. A estrutura proposta ajuda a reduzir o tempo médio de reparo (MTTR) fornecendo alertas análise e resolução de insights por meio de técnicas de aprendizado de máquina, assim facilita a alta disponibilidade de serviço do centro computacional priorizando alertas e resolvendo proativamente problemas críticos em NERSC.

Com o crescimento do número de plataformas e componentes abrangendo a moderna petascale e pós-petascale computacional centros, a escala de eventos gerados e relatados em seus infra-estrutura por dia aumentará exponencialmente. Isso se traduz em crescimento em incidentes e problemas de gerenciamento de software levando a fadiga da equipe de operações e dificuldades em priorizar efetivamente eventos, na filtragem de alertas do ruído e outros desafios de correção de alertas, degradando assim a resposta a problemas computacionais críticos questões de instalações. Portanto, para enfrentar os desafios acima, há

necessidade de integração de sistemas e infraestrutura de automação que permite a automação do monitoramento de eventos, redução de ruídos relacionados a eventos semelhantes, correlação dos eventos gerados para facilitar previsão automática e prevenção do desempenho operacional questões de degradação. Por esse motivo, investigamos soluções automatizadas como as plataformas Prometheus, Grafana, Kafka e ServiceNow para facilitar a orquestração dos serviços e implantações do sistema, automação e análise de correlação de dados de streaming, e adaptação e resolução dos alertas para identificação do core problemas de ineficiência operacional.

Dada a dinamicidade dos ambientes computacionais, o NERSC está investigando plataformas que serão mais adequadas para monitorar os sistemas de computação e infraestrutura em rápida mudança ecossistemas. Uma dessas ferramentas é o Prometheus, que é de código aberto e altamente adaptável para monitorar escalas e dinâmicas tão extremas centros computacionais. A ferramenta suporta SNMP, alerta avançado mecanismos melhores que o Nagios e projetados especificamente para monitorar e métricas de integridade da infraestrutura de computação híbrida construída em contêineres e serviços sem despesa de qualquer computação Recursos.

**A. Prometeu:**

Prometheus desenvolvido em 2012, é uma solução de monitoramento de código aberto que foi lançada em 2016 pela SoundCloud. O Prometheus facilita um modelo de dados multidimensional, coleta todos os dados de vários fontes de dados como fluxo de valores com carimbo de data/hora e armazena a hora dados de série que podem ser identificados pelo nome da métrica e chave/valor pares. Contador, Medidor, Histograma e Resumo são quatro tipos de métricas que são oferecidas pelas bibliotecas cliente do Prometheus. Além disso, em comparação com a maioria das ferramentas de monitoramento como OpenTSDB, Nagios, Sensu, a plataforma Prometheus também oferece um poderoso e consultas flexíveis e alertas em tempo real empregando PromQL, uma linguagem de consulta funcional que utiliza o modelo pull sobre HTTP e permite que o usuário selecione e agregue métricas em tempo real do banco de dados de séries temporais. Além disso, os dados consultados podem ser visualizados como gráficos ou tabelas no navegador usando a API HTTP. Além disso, o componente alertmanager da plataforma Prometheus vem útil com muito mais recursos do que apenas silenciar alertas como realizada pela ferramenta Nagios. Além disso, o Prometheus permite a inspeção das entranhas dos aplicativos (monitoramento de caixa branca) em comparação com ferramentas como Nagios, Sensu que só permitem inspecionar o estado do host ou serviço não ajuda a entender como o host/serviço atingiu esse estado atual (monitoramento de caixa preta ). O ecossistema Prometheus compreende os seguintes principais componentes: 1) Servidor Prometheus que possibilita a raspagem do métricas dos trabalhos e coletando os dados multidimensionais conforme séries temporais numéricas; 2) Bibliotecas de cliente que implementam os quatro tipos de métrica do Prometheus e permite a exposição de métricas internas via endpoint HTTP para a instância do aplicativo correspondente; 3) Alertmanager que facilita o manuseio de alertas enviados pelo Prometheus servidor e ajuda no agrupamento (notificação única para alertas semelhantes), Inibição (suprimindo a notificação de alerta que não está relacionada ao real problema), Silenciar (silenciar alertas por um período de tempo especificado) e de duplicação (reduzir o número dos mesmos alertas disparados ao longo de um único incidente e evitar fadiga de alerta); 4) Os exportadores permitem a disseminação de métricas de séries temporais existentes de terceiros sistemas como HAProxy, StatsD, Grafite como métricas do Prometheus ; e 5) Grafana facilita uma plataforma de análise puxando o

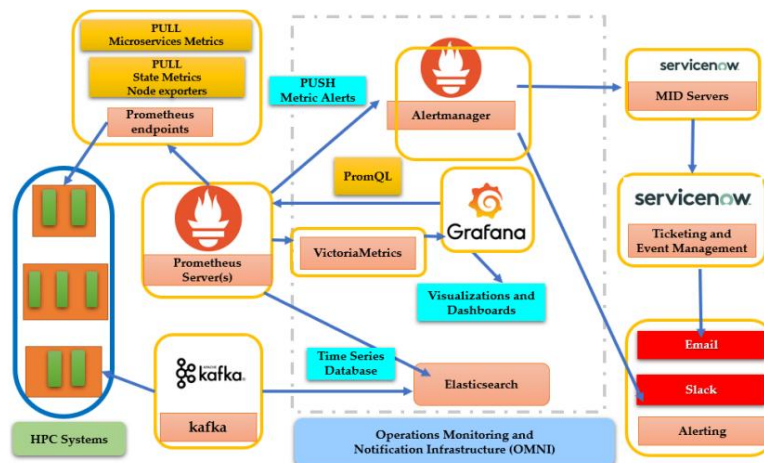


Figura 1: Design e Arquitetura da Estrutura de Gerenciamento e Monitoramento de Eventos.

métricas consultando o Prometheus e permitindo a visualização de dados e alertas por meio de painéis intuitivos. A série temporal de monitoramento dados relativos aos eventos de infraestrutura computacional NERSC coletados e agregados pelo Prometheus serão armazenados nos OMNI's Plataforma Elasticsearch e será usada para fornecer tendências históricas e métricas para planejamento de capacidade e para exibir painéis contendo gráficos e visualizações para obter insights operacionais e potencializando a tomada de decisão.

#### B. Infraestrutura de Monitoramento e Notificação de Operações (OMNI)

OMNI é uma solução flexível de big data, um armazém para coletar, gerenciar e analisar dados relacionados ao monitoramento de computação de escala extrema sistemas [19]. Essa infraestrutura facilita um único local para armazenar os conjuntos de dados heterogêneos e é apoiado por um escalável e banco de dados paralelo de séries temporais, Elasticsearch [3]. Exemplos de dados operacionais incluem dados de séries temporais do ambiente (por exemplo, temperatura, potência, níveis de umidade e níveis de partículas), monitorando dados (por exemplo, velocidades de rede, latência, perda de pacotes, utilização ou aqueles que monitoram o sistema de arquivos para velocidades de gravação de disco, E/S, CRC erros) e dados de eventos (por exemplo, logs do sistema, logs do console, eventos de falha, eventos de energia essencialmente qualquer coisa que tenha um início e hora de término). Obtendo dados dos vários sistemas e sensores no Elasticsearch [3] ocorre via RabbitMQ [35], um agente de mensagens que suporta vários protocolos de mensagens e enfileiramento. Dados de várias fontes podem ser enfileirados diretamente em um enfileiramento RabbitMQ sistema com um formato json. Alternativamente, pode ser coletado primeiro de um sistema via Collectd então analisado pelo Logstash [7] antes de ser enfileirado no RabbitMQ. Implementamos vários RabbitMQ fluxos para diferenciar diferentes conjuntos de dados, cada tipo de conjunto de dados tendo sua própria fila. Usando RabbitMQ, Logstash e Elasticsearch, O OMNI é capaz de ingerir mais de 25.000 mensagens por segundo de fontes heterogêneas e distribuídas dentro e fora do data center.

Uma vez ingerido, o Elasticsearch indexa os dados quase em tempo real recuperação e consulta. Os dados podem ser consultados diretamente do Elastic search usando as APIs RESTful nativas ou usando visualização e ferramentas de descoberta de dados, como Kibana [6] ou Grafana. Além disso, esta solução tem alta disponibilidade com operações de manutenção mínimas

interromper o funcionamento da infraestrutura de coleta de dados. Como sistemas tornam-se maiores, mais heterogêneos e mais complexos, métodos de monitoramento existentes se tornarão incontrolláveis. Perlmutter, o sistema 2020 terá mais de 3 vezes o poder de computação do sistema atual, Cori, e contém uma mistura de apenas CPU e nós acelerados por GPU. A instalação tem agora mais de 20.000 sensores com centenas de fontes de ambiente, computação e dados de eventos de fluxos de água e leituras de umidade, temperaturas em diferentes áreas dos racks e centro, sensores no subestação, os diferentes níveis de PDU's disponíveis no prédio, configuração de UPS/generador, nós de computação, servidores de sistema de arquivos, login nós, nós de gateway, nós de suporte, sistemas de arquivos Lustre e GPFS [32], syslogs e outros logs.

#### C. Serviço Agora

A ServiceNow, fundada em 2003, fornece uma plataforma unificada de computação em nuvem para automatizar e transformar processos de TI, como operações de segurança, serviço de TI, ativos, operações de negócios e gerenciamento de eventos e muito mais. ServiceNow é uma plataforma como serviço provedor (PaaS) [26] com foco principal na promoção da gestão de os eventos de incidentes, problemas e mudanças referentes à infraestrutura de TI e serviços. ServiceNow implementa uma configuração banco de dados de gerenciamento (CMDB), que ajuda a fornecer registros precisos e atualizados dos ativos de TI da organização para auxiliar no gerenciamento de ativos, conformidade e configuração. O banco de dados CMDB auxilia na análise de impacto do serviço, pois armazena informações atuais e precisas sobre todos os serviços técnicos em um Item de configuração (CI) correspondente a esses serviços. Além disso, o mapeamento de serviço é fortemente acoplado ao banco de dados CMDB para torná-lo ciente do serviço, onde os mapas de serviço empregam descoberta e informações de infraestrutura no CMDB para criar uma mapa completo baseado em tags de todos os aplicativos, sistemas virtuais, rede subjacente, bancos de dados, servidores e outros componentes de TI que suporta o serviço. Além disso, o mapeamento de serviço automatizado permite não apenas uma interface de usuário que mostra o nível de serviço preciso relacionamentos, mas também atualiza os mapas de serviço em tempo real, ajuda a evitar dados de infraestrutura irrelevantes e a obter mais rapidamente insights e histórico da topologia de serviço. A gestão de eventos

plataforma habilitada pelo ServiceNow permite uma fácil integração com várias ferramentas de monitoramento, como Prometheus para receber eventos dados para análise e resposta. As métricas operacionais capturadas pelo ServiceNow é usado para reduzir o tempo médio de resolução (MTTR) empregando aprendizado de máquina para análise de causa raiz e definir regras de alerta, limites e ações de correção. Além disso, Os painéis do ServiceNow fornecem visualização de cada serviço que ajuda a priorizar problemas de serviço com base no impacto e na criticidade.

A facilidade computacional NERSC está agora em transição para integrar a plataforma de gerenciamento e monitoramento de eventos em produção ambiente para se preparar para o próximo sistema de supercomputação de próxima geração. Em preparação para o próximo sistema, um novo está sendo desenvolvida uma infraestrutura de rede de coleta de dados que facilita a conectividade VXLAN . Além disso, o NERSC também investigou muitas versões do Kubernetes, para permitir orquestração e gerenciamento de contêineres e para investigar o melhor ajuste de gerenciamento no modelo de coleta de dados OMNI. O objetivo é testar e aprofundar o conhecimento sobre a implementação do NERSC-9 (N9) antes de dar um passo em direção ao supercomputador da próxima geração . Além disso, os colecionadores e exportadores do Prometheus têm já sendo implantado no NERSC Global Filesystems (NGF) sistema de desenvolvimento, nós de login ES de Cori e gabinetes de GPU de Cori [2]. Mais esforços de colaboração estão atualmente sendo focados em estender isso para o resto dos nós externos do Cori, o novo HPSS sistemas, o novo sistema de sistema de arquivos da comunidade e outros sistemas de apoio. Junto com a plataforma Prometheus, o Grafana e a plataforma ServiceNow também estão sendo implantadas e integradas na infraestrutura OMNI. Vários fluxos de trabalho de monitoramento e gerenciamento de eventos utilizando o abrangente plataforma foram implementados e testados para automatizar o gerenciamento de alertas e processos de correção, bem como para fornecer painéis intuitivos para aprimorar a inteligência operacional e eficiência.

4 GESTÃO E MONITORAMENTO DE EVENTOS

FLUXOS DE TRABALHO

Como o número de componentes que compõem a infraestrutura de computação e seus incidentes e problemas de gerenciamento correspondentes evoluem, a integração e a automação do sistema estão se tornando mais importantes. Há uma necessidade crescente de implementar e simplificar fluxos de trabalho utilizando a plataforma proposta de gerenciamento e monitoramento de eventos para simplificar o gerenciamento de vários processos, reduzir custos operacionais e aumentar a eficiência da equipe. Devido aos avanços na tecnologia de automação, esses esforços são focados não apenas na eficiência de triagem e solução de problemas, mas também nos fluxos de trabalho de remediação. A correção de alertas está se tornando uma importante parte do monitoramento operacional e gerenciamento de eventos do NERSC-9 a infraestrutura. Tanto o gerenciamento de eventos quanto a correção de alertas ser implementado usando ServiceNow - uma solução ITSM e ITOM com um backend de banco de dados MySQL. Na NERSC, ServiceNow é usado como uma plataforma de Gerenciamento de Incidentes para usuários e funcionários. Dado o seu versatilidade e funcionalidade de API personalizada, a plataforma ServiceNow é um componente central do layout de monitoramento NERSC-9.

Aproveitando a plataforma ServiceNow para gerenciamento de dados usando Event Management e CMDB (Configuration Management Database), nosso fluxo de trabalho automatizado facilita: 1) Auto Discovery -

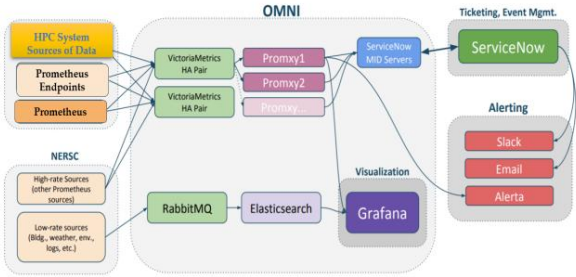


Figura 2: Resposta automatizada de eventos e fluxo de trabalho de monitoramento: várias fontes fornecem dados de endpoints do Prometheus ou Kafka, onde os dados são transformados e podem ir para o Elasticsearch, onde podem ser consultados pelo Kibana ou pode ir ao vmagent que lida com o envio das informações para várias áreas para alertas por meio do ServiceNow MID servidor. Os dados passam automaticamente por um módulo de gerenciamento de eventos no ServiceNow para lidar com a abertura de um ticket ou notificação de várias chamadas do alerta. Mais longe, por meio do alertmanager, os dados podem ir para o slack para processamento.



Figura 3: Fontes de dados de amostra visualizadas no Kibana ou Grafana.

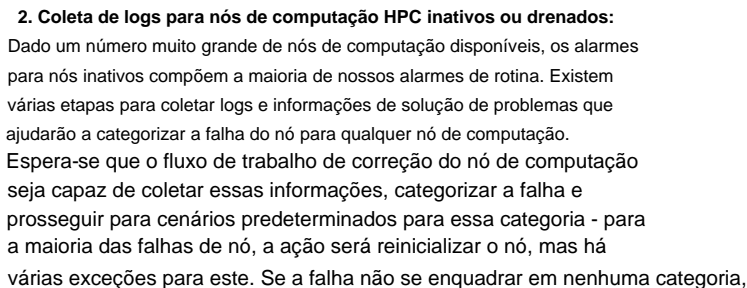
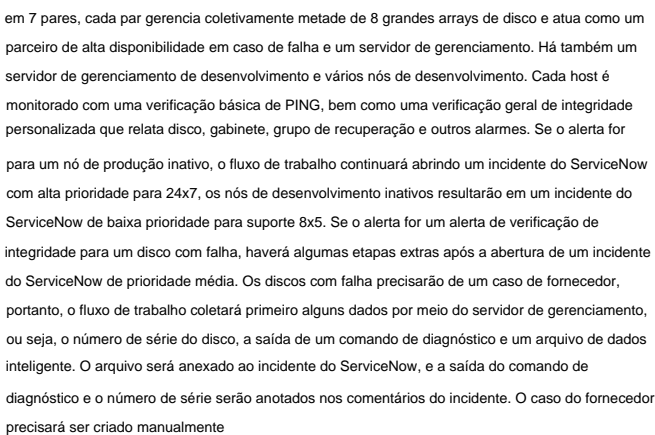


Figura 4: Fontes de dados adicionais do Apache Kafka visualizadas.

em conjunto com Netbox; 2) Relatórios Visuais Intuitivos - de incidentes (bilhetes); e 3) Integração do Prometheus com o Promxy - um proxy Prometheus, um único endpoint de API voltado para o usuário. o gerenciamento automatizado de resposta a eventos e fluxo de trabalho de monitoramento (como mostrado na Figura 2) implementa as seguintes etapas:

- Armazenamento de alto volume e alta taxa de dados: alta disponibilidade (HA) pair fornece redundância, realiza a deduplicação de várias fontes. Além disso, VictoriaMetrics [15] facilita uma





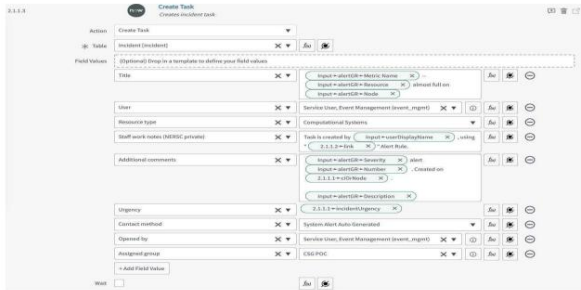


Figura 9: Criação automática de um ticket no ServiceNow assim que um alerta é identificado.

o fluxo de trabalho enviará um incidente do ServiceNow, atribuído ao grupo apropriado. Todos os logs coletados serão anexados ou anotados no Incidente para consideração de um engenheiro.

3. Correção de espaço em disco, carga e integridade do serviço verifica vários servidores NERSC:

Muitos desses alertas são repetitivos e terão diretrizes claras para resolver o problema, por exemplo, um host em que /var/log está sempre enchendo ou um serviço do sistema que geralmente precisa ser reiniciado após uma falha. Esses tipos de problemas exigem muito pouco trabalho de triagem, portanto, sua correção certamente pode ser automatizada no ServiceNow.

5. CONCLUSÃO

A estrutura de gerenciamento e monitoramento de eventos apresentada neste documento facilita o monitoramento proativo, coleções e correlação de eventos, visualizações de alertas e métricas, orquestração e remediação e análise de causa raiz automatizada em tempo real e gerenciamento de respostas nas implementações e serviços de cluster heterogêneos altamente mutantes usando uma solução integrada composta pelas plataformas OMNI, Prometheus, Grafana e ServiceNow. Além disso, a correção de alertas habilitada pelos fluxos de trabalho automatizados habilitados pela estrutura proposta reduzirá potencialmente drasticamente o número de incidentes que exigem intervenção da equipe operacional ou precisam ser redirecionados para os grupos apropriados. Além disso , a estrutura integrada acelerará consideravelmente o tempo de resolução dos possíveis problemas críticos do centro computacional com as implantações de fluxos de trabalho de correção automatizada propostas, bem como para outros problemas que exigem que os dados e os logs sejam coletados automaticamente.

6 AGRADECIMENTOS

Esta pesquisa utilizou recursos do National Energy Research Sci entific Computing Center (NERSC), um Departamento de Energia dos Estados Unidos Office of Science User Facility operado sob o Contrato N° DEAC02-05CH11231.

REFERÊNCIAS

[1] [sd]. Apache Kafka. <https://kafka.apache.org/> [2] [nd]. Cori: o mais novo supercomputador do NERSC. <https://www.nersc.gov/users/systems-computacionais/cori/>  
[3] [sd]. Elasticsearch: Mecanismo RESTful distribuído. <https://www.elastic.co/ produtos/pesquisa elastica>  
[4] [nd]. Grafana. <https://grafana.com/> [5] [nd]. Icinga. <https://icinga.com/> [6] [nd]. Kibana: Sua janela para o Elastic Stack. <https://www.elastic.co/ produtos/kibana>

[7] [sd]. Logstash: Centralize, transforme e armazene seus dados. <https://www.elastic.co/products/logstash> [8] [nd]. Nagios. <https://www.nagios.org/> [9] [nd]. OpenLorenz: Painel HPC baseado na Web e muito mais. <https://software.llnl.gov/repo/#hpc/OpenLorenz> [10] [nd]. Monitor de rede Paessler PRTG. <https://www.paessler.com/prtg> [11] [nd]. Perlmutter: o próximo supercomputador do NERSC. <https://www.nersc.gov/ systems/perlmutter/> [12] [nd]. Prometheus. <https://prometheus.io/> [13] [nd]. Serviço agora. <https://www.servicenow.com/> [14] [nd]. Software de gerenciamento de monitoramento de rede Spiceworks. <https://www.spiceworks.com/> [15] [nd]. VictoriaMetrics. <https://victoriametrics.com/> [16] [nd]. Zabbix. <https://www.zabbix.com/> [17] Erika Abraham, Costas Bekas, Ivona Brandic, Samir Genaim, Einar Broch Johnsen, Ivan Kondov, Sabri Pilana e Achim Streit. 2015. Preparando aplicativos HPC para exascale: desafios e recomendações. Em 2015 18ª Conferência Internacional sobre Sistemas de Informação Baseados em Rede. IEEE, 401-406.

[18] R Alvarez Alonso, G Arneodo, O Barring, E Bonfillou, et al. 2014. Migração do sistema de suporte do data center cern it para servicenow. Em Journal of Physics: Conference Series, Vol. 513. 062032.  
[19] Elizabeth Bautista, Melissa Romanus, Thomas Davis, Cary Whitney e Theodore Kubaska. 2019. Coleta, monitoramento e análise de dados de instalações e sistemas no Centro Nacional de Computação Científica de Pesquisa Energética. Em 2019 Conferência Internacional sobre Processamento Paralelo. ACM, no prelo.  
[20] Elizabeth Bautista, Cary Whitney e Thomas Davis. 2016. Big data por trás de big data. Na conquista de Big Data com computação de alto desempenho. Springer, 163-189.  
[21] Keren Bergman, Shekhar Borkar, Dan Campbell, William Carlson, William Dally, Monty Denneau, Paul Franzon, William Harrod, Kerry Hill, Jon Hiller, et al. 2008. Estudo de Computação Exascale: Desafios Tecnológicos na Conquista de Sistemas Exascale. Gabinete de Técnicas de Processamento de Informação da Agência de Projetos de Pesquisa Avançada de Defesa (DARPA IPTO), Tech. Rep. 15 (2008).  
[22] J Davin, JD Case, M Fedor e ML Schoffstall. 1989. Gerenciamento de rede simples protocolo (SNMP). (1989).  
[23] Adam DeConinck, A Bonnie, K Kelly, S Sanchez, C Martin, M Mason, James M Brandt, Ann C Gentile, Benjamin A Allan, Anthony Michael Agelastos, et al. 2016. Projeto e Implementação de um Sistema de Monitoramento Escalável para Trinity. Relatório Técnico . Sandia National Lab. (SNL-NM), Albuquerque, NM (Estados Unidos).  
[24] Gabriel Iuhasz e Dana Petcu. 2019. Monitoramento do processamento de dados Exascale. Em 2019 IEEE International Conference on Advanced Scientific Computing (ICASC). IEEE, 1-5.  
[25] David Josephsen. 2007. Construindo uma infraestrutura de monitoramento com o Nagios. Aprendiz São Paulo.  
[26] Michael J Kavis. 2014. Arquitetando a nuvem: decisões de design para modelos de serviços de computação em nuvem (SaaS, PaaS e IaaS). John Wiley & Filhos.  
[27] Antonio Libri, Andrea Bartolini e Luca Benini. 2018. Dig: Habilitando monitoramento de alta resolução escalável fora de banda para análise, automação e controle de data center. No 2º Workshop Internacional da Indústria/Universidade sobre Automação, Análise e Controle de Data Centers (DAAC 2018). Automação, análise e controle de data center (DAAC).  
[28] Vojko Matko e Barbara Brezovec. 2018. Eficiência e disponibilidade de energia aprimoradas do data center com processamento de eventos de nó multicamadas. Energias 11, 9 (2018), 2478.  
[29] Masoud Moshref, Minlan Yu, Ramesh Govindan e Amin Vahdat. 2016. Trompette: Acionamentos oportunos e precisos em data centers. Em Anais da Conferência ACM SIGCOMM 2016. 129-143.  
[30] Dmitry A Nikitenko, Sergey A Zhumatiy e Pavel A Shvets. 2016. Tornar sistemas de grande escala observáveis - outro passo inevitável em direção à exascale. Supercomputing Frontiers and Innovations 3, 2 (2016), 72–79.  
[31] Sam Sanchez, Amanda Bonnie, Graham Van Heule, Conor Robinson, Adam DeConinck, Kathleen Kelly, Quellyn Sneed e J Brandt. 2016. Projeto e implementação de um sistema de monitoramento hpc escalável. Em 2016 IEEE International Parallel and Distributed Processing Workshops (IPDPSW). IEEE, 1721-1725.  
[32] Frank B Schmuck e Roger L Haskin. 2002. GPFS: Um Sistema de Arquivos em Disco Compartilhado para grandes clusters de computação. Em FAST, Vol. 2.  
[33] John Shalf, Sudip Dsanjh e John Morrison. 2010. Desafios da tecnologia de computação Exascale . Em Conferência Internacional sobre Computação de Alto Desempenho para Ciência Computacional. Springer, 1-25.  
[34] Nitin Sukhija e Elizabeth Bautista. 2019. Rumo a uma estrutura para monitorar e analisar ambientes de computação de alto desempenho usando Kubernetes e Prometheus. Em 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People e Smart City Innovation (Smart World/SCALCOM/UIC/ATC/CBDCom/IOP/SCI) . IEEE, 257-262.  
[35] Álvaro Videla e Jason JW Williams. 2012. RabbitMQ em ação: mensagens distribuídas para todos. Tripulação.