

Gerenciamento de nível de serviço: White Paper de práticas recomendadas

Contents

[Introduction](#)

[Visão geral do gerenciamento de nível de serviço](#)

[Fatores críticos para o sucesso](#)

[Indicadores de desempenho](#)

[Fluxo de processo de gerenciamento de nível de serviço](#)

[Implementando o gerenciamento de nível de serviço](#)

[Definindo níveis de serviço de rede](#)

[Criação e manutenção de SLAs](#)

[Indicadores de desempenho do gerenciamento de nível de serviço](#)

[Contrato de nível de serviço documentado ou Definição de nível de serviço](#)

[Métricas de indicador de desempenho](#)

[Revisão do gerenciamento de nível de serviço](#)

[Resumo de gerenciamento de nível de serviço](#)

[Informações Relacionadas](#)

Introduction

Este documento descreve o gerenciamento de nível de serviço e os SLAs (acordos de nível de serviço) para redes de alta disponibilidade. Inclui fatores críticos de sucesso para o gerenciamento em nível de serviço e indicadores de desempenho para ajudar a avaliar o sucesso. O documento também oferece detalhes significativos para os SLAs que seguem as orientações de melhores práticas identificadas pela equipe de serviços de alta disponibilidade.

Visão geral do gerenciamento de nível de serviço

As organizações de rede sempre atenderam aos requisitos de rede em expansão, criando uma infraestrutura de rede sólida e trabalhando de forma reativa para lidar com problemas de serviço individuais. Quando ocorria uma interrupção, a organização desenvolvia novos processos, recursos de gerenciamento ou infra-estrutura para evitar que uma interrupção específica ocorresse novamente. No entanto, devido a uma taxa de alteração mais alta e aos requisitos de disponibilidade cada vez maiores, agora precisamos de um modelo aprimorado para evitar períodos de inatividade não planejados e reparar rapidamente a rede. Muitas empresas de provedores de serviços e corporações tentaram definir melhor o nível de serviço necessário para atingir as metas comerciais.

Fatores críticos para o sucesso

Os fatores críticos para o sucesso dos SLAs são usados para definir elementos-chave para a construção bem-sucedida de níveis de serviço e para a manutenção de SLAs. Para se qualificar como um fator de sucesso crítico, um processo ou etapa do processo deve melhorar a qualidade do SLA e beneficiar a disponibilidade da rede em geral. O fator de sucesso crítico também deve ser medido de modo que a organização possa determinar o sucesso relativo ao procedimento definido.

Consulte [Implementando gerenciamento em nível de serviço](#) para obter mais detalhes.

[Indicadores de desempenho](#)

Os indicadores de desempenho oferecem o mecanismo pelo qual uma organização mede os fatores de sucesso importantes. Normalmente, você os analisa mensalmente para garantir que as definições de nível de serviço ou SLAs estejam funcionando bem. O grupo de operações de rede e os grupos de ferramentas necessárias podem realizar as seguintes métricas.

Observação: para organizações sem SLAs, recomendamos que você execute definições de nível de serviço e revisões de nível de serviço além de métricas.

Os indicadores de desempenho incluem:

- Definição de nível de serviço ou SLA documentada que inclui disponibilidade, desempenho, tempo de resposta de serviço reativo, metas de resolução de problemas e escalonamento de problemas.
- Reunião mensal de análise de nível de serviço de rede para revisar a conformidade de nível de serviço e implementar melhorias.
- Métricas do indicador de desempenho, incluindo disponibilidade, desempenho, tempo de resposta do serviço por prioridade, tempo para resolver por prioridade e outros parâmetros SLA mensuráveis.

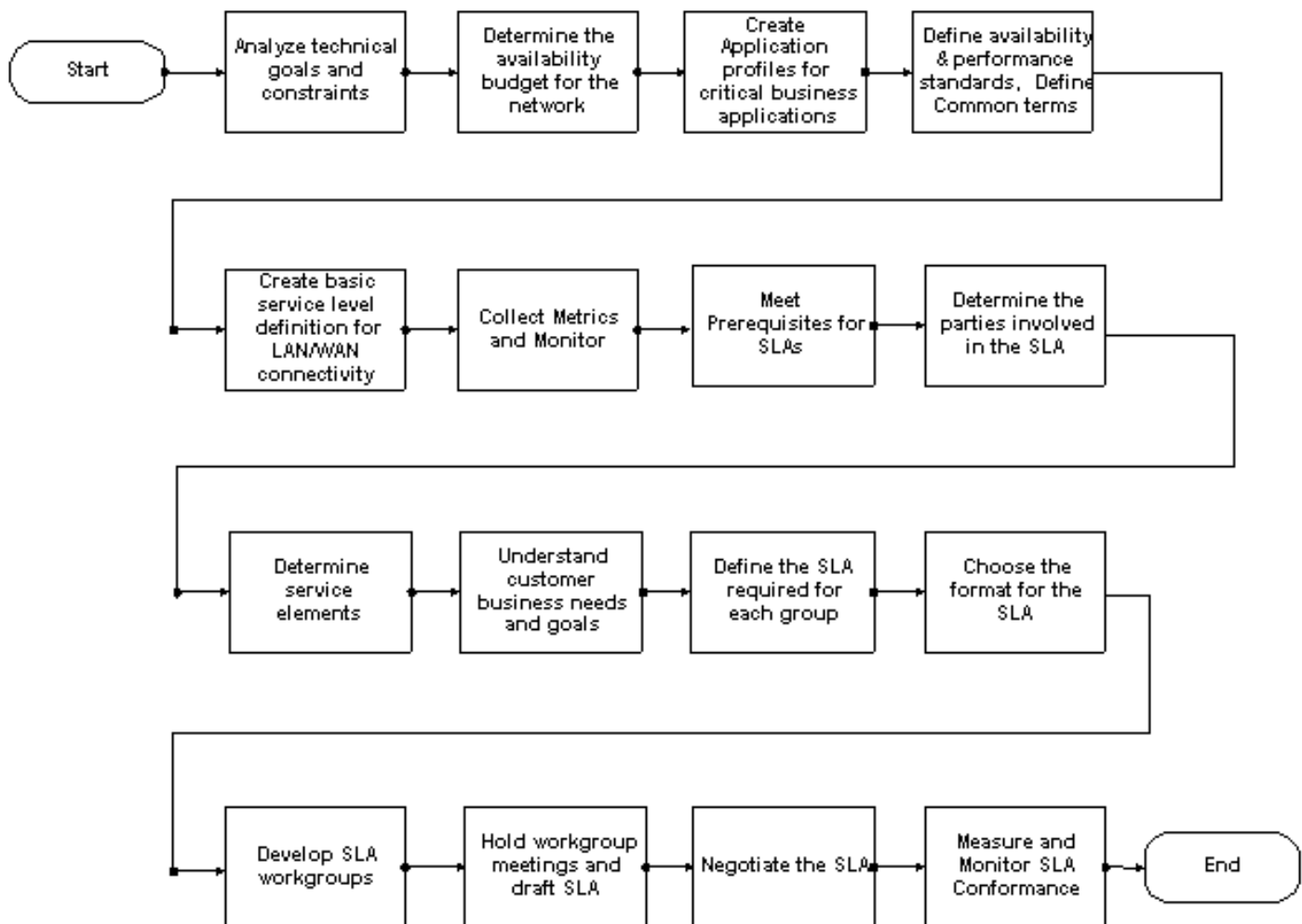
Consulte [Implementando gerenciamento em nível de serviço](#) para obter mais informações.

[Fluxo de processo de gerenciamento de nível de serviço](#)

O fluxo de processo de alto nível para o gerenciamento de nível de serviço contém dois grupos importantes:

1. [Definindo níveis de serviço de rede](#)
2. [Criação e manutenção de SLAs](#)

Clique nos objetos no diagrama a seguir para ver os detalhes dessa etapa.



Implementando o gerenciamento de nível de serviço

A implementação do gerenciamento de nível de serviço consiste em dezesseis etapas divididas nas duas categorias principais a seguir:

- [Definindo níveis de serviço de rede](#)
- [Criação e manutenção de SLAs](#)

Definindo níveis de serviço de rede

Os gerentes de rede precisam definir as principais regras pelas quais a rede é suportada, gerenciada e medida. Os níveis de serviço fornecem metas para todos os funcionários da rede e podem ser usados como métrica na qualidade do serviço geral. Você também pode nos fornecer definições de nível de serviço como uma ferramenta para fazer o orçamento de recursos de rede e como evidência para a necessidade de financiar um QoS mais alto. Também proporcionam uma forma de avaliar o desempenho de fornecedores e portadoras.

Sem uma definição e medição de nível de serviço, a organização não tem metas claras. A satisfação com o serviço pode ser regida por usuários com uma pequena diferenciação entre aplicativos, operações de servidor/cliente ou suporte de rede. O orçamento pode ser mais difícil porque o resultado final não é claro para a organização e, por fim, a organização de rede tende a ser mais reativa, e não proativa, na melhoria da rede e do modelo de suporte.

Recomendamos as seguintes etapas para criar e suportar um modelo de nível de serviço:

1. [Analisar metas e restrições técnicas.](#)
2. [Determine o orçamento de disponibilidade.](#)
3. [Crie perfis de aplicativos detalhando as características da rede de aplicativos críticos.](#)
4. [Defina os padrões de disponibilidade e desempenho e defina termos comuns.](#)
5. [Crie uma definição de nível de serviço que inclua disponibilidade, desempenho, tempo de resposta do serviço, tempo médio para resolver problemas, detecção de falhas, limiares de atualização e caminho de escalonamento.](#)
6. [Colete métricas e monitore a definição do nível de serviço.](#)

[Passo 1: Analisar objetivos técnicos e restrições](#)

A melhor maneira de iniciar a análise de restrições e objetivos técnicos é discutir ou pesquisar os objetivos e os requisitos técnicos. Algumas vezes, pode ajudar convidar outros colegas de técnicos de TI para participar dessa discussão porque esses indivíduos tem metas específicas relacionadas a seus serviços. As metas técnicas incluem níveis de disponibilidade, throughput, instabilidade, atraso, tempo de resposta, requisitos de escalabilidade, introduções de novos recursos, introduções de novos aplicativos, segurança, gerenciabilidade e até mesmo custo. A organização deve investigar as restrições para atingir essas metas fornecidas pelos recursos disponíveis. Você pode criar planilhas para cada objetivo, com uma explicação de restrições. Inicialmente, pode parecer que a maioria dos objetivos não é alcançável. Dê prioridade aos objetivos ou reduza as expectativas que ainda podem satisfazer os requisitos de negócios.

Por exemplo, você pode ter um nível de disponibilidade de 99,999 por cento, ou 5 minutos de inatividade por ano. Existem inúmeras limitações para alcançar este objetivo, como pontos únicos de falha no hardware, hardware quebrado em tempo médio ao reparo (MTTR), confiabilidade no portador, capacidades de detecção de falha pró-ativa, altas taxas de alteração e limitações de capacidade de rede atual. Como resultado, você pode ajustar o objetivo para um nível mais alcançável. O modelo de disponibilidade [na próxima seção](#) pode ajudá-lo a definir metas realistas.

Você também pode considerar fornecer maior disponibilidade em determinadas áreas da rede que tenham menos restrições. Quando a organização de rede publica padrões de serviço para disponibilidade, os grupos de negócios dentro da organização podem achar o nível inaceitável. Portanto, esse é um ponto natural para começar discussões sobre SLA ou modelos de financiamento/orçamentos que podem alcançar os requisitos comerciais.

Tente identificar todas as restrições e riscos envolvidos para alcançar o objetivo técnico. Priorize as restrições em termos do maior risco ou impacto sobre o objetivo almejado. Isso ajuda a organização a priorizar iniciativas de melhoria de rede e determinar com que facilidade a restrição pode ser tratada. Há três tipos de restrições:

- Tecnologia, resiliência e configuração de rede
- Práticas do ciclo de vida, incluindo planejamento, projeto, implementação e operação
- Carga de tráfego atual ou comportamento do aplicativo

As restrições de tecnologia, resiliência e configuração de rede são quaisquer limitações ou riscos associados à tecnologia, hardware, links, design ou configuração atuais. Limitações de tecnologia abrangem qualquer restrição imposta pela própria tecnologia. Por exemplo, nenhuma tecnologia atual permite tempos de convergência de menos de um segundo em ambientes de rede redundantes, que podem ser críticos para manter conexões de voz em toda a rede. Outro exemplo pode ser a velocidade bruta que os dados podem atravessar em links terrestres, que são aproximadamente 100 milhas por milissegundo.

As investigações de risco de resiliência de hardware de rede devem se concentrar na topologia de hardware, na hierarquia, na modularidade, na redundância e no MTBF ao longo de caminhos definidos na rede. As restrições de enlace de rede devem se concentrar nos enlaces de rede e na conectividade da operadora para organizações corporativas. As restrições de link podem incluir redundância e diversidade de links, limitações de mídia, infraestrutura de cabeamento, conectividade de loop local e conectividade de longa distância. As restrições de projeto relacionam-se ao projeto físico ou lógico da rede e incluem tudo, desde o espaço disponível para equipamentos até a escalabilidade da implementação do protocolo de roteamento. Todos os projetos de protocolo e mídia devem ser considerados em relação à configuração, disponibilidade, escalabilidade, desempenho e capacidade. As restrições de serviço de rede, como Dynamic Host Configuration Protocol (DHCP), Domain Name System (DNS), firewalls, tradutores de protocolo e tradutores de endereços de rede também devem ser consideradas.

As práticas do ciclo de vida definem os processos e o gerenciamento da rede usados para implantar soluções, detectar e reparar problemas, prevenir problemas de capacidade ou desempenho e configurar a rede para consistência e modularidade. É necessário considerar essa área porque experiência e processos normalmente são os itens que mais contribuem para não-disponibilidade. O ciclo de vida da rede refere-se ao ciclo de planejamento, projeto, implementação e operações. Em cada uma dessas áreas, você deve entender a funcionalidade do gerenciamento de rede como o gerenciamento de desempenho, gerenciamento de configuração, gerenciamento de falhas e segurança. Uma avaliação do ciclo de vida da rede está disponível nos serviços de alta disponibilidade (HAS) do Cisco NSA, mostrando as restrições atuais de disponibilidade da rede associadas às práticas do ciclo de vida da rede.

A carga de tráfego atual ou as restrições de aplicativos simplesmente se referem ao impacto do tráfego e dos aplicativos atuais.

Infelizmente, muitos aplicativos possuem restrições significativas que exigem gerenciamento cauteloso. Os requisitos de variação de sinal, atraso, throughput e largura de banda para os aplicativos atuais normalmente têm muitas restrições. A maneira como o aplicativo foi escrito também pode criar restrições. A criação de perfis de aplicativos ajuda você a entender melhor esses problemas; [a próxima seção](#) abrange este recurso. Investigar a disponibilidade, o tráfego, a capacidade e o desempenho atuais em geral também ajuda os gerentes de rede a entender as expectativas e os riscos atuais do nível de serviço. Isso é normalmente realizado com um processo chamado linha de base de rede, que ajuda a definir o desempenho, a disponibilidade ou as médias de capacidade da rede para um período de tempo definido, normalmente em torno de um mês. Essas informações são normalmente usadas para planejamento de capacidade e tendências, mas também podem ser usadas para entender problemas de nível de serviço.

A planilha abaixo utiliza o método de objetivos/restrições acima mencionado, com o objetivo de exemplificar o impedimento de um ataque de segurança ou de um ataque de recusa de serviço (DoS). Você também pode usar esta planilha para ajudar a determinar a cobertura do serviço para minimizar os ataques à segurança.

Risco ou restrição	Tipo de restrição	Impacto potencial
As ferramentas de detecção de DoS disponíveis não podem detectar todos os tipos de ataques de DoS.	Tecnologia/elasticidade	Alto
Não tenha a equipe e o processo	Práticas do	Alto

necessários para reagir a alertas.	ciclo de vida	
As políticas de acesso à rede atuais não estão em vigor.	Práticas do ciclo de vida	Méd io
A conexão atual com a Internet de banda mais baixa pode ser um fator se o congestionamento da largura de banda for usado para o ataque.	Capacidad e da rede	Méd io
Atualmente, a configuração de segurança para ajudar a evitar ataques pode não ser completa.	Tecnologia/ elasticidad e	Méd io

Passo 2: Determine o orçamento de disponibilidade

Um orçamento de disponibilidade é a disponibilidade teórica esperada da rede entre dois pontos definidos. Informações teóricas precisas são úteis de várias maneiras:

- A organização pode usá-las como meta para disponibilidade interna e os desvios podem ser rapidamente definidos e remediados.
- As informações podem ser usadas pelos planejadores de rede ao determinar a disponibilidade do sistema, ajudando a garantir que o projeto atenda aos requisitos de negócios.

Os fatores que contribuem para o tempo de indisponibilidade ou paralisação incluem falha de hardware, falha de software, problemas de energia e ambiente, falha de link ou portadora, projeto de rede, erro humano ou falta de processo. Você deve avaliar bem cada um desses parâmetros ao avaliar o orçamento geral de disponibilidade para a rede.

Se a organização avaliar a disponibilidade no momento, talvez você não precise de um orçamento de disponibilidade. Use a medida de disponibilidade como linha de base para estimar o nível de serviço atual usado para uma definição de nível de serviço. No entanto, você pode estar interessado em comparar os dois para entender a disponibilidade teórica potencial em comparação com o resultado medido real.

Disponibilidade é a probabilidade de um produto ou serviço operar quando necessário. Veja as seguintes definições:

1. **Disponibilidade**¹ (tempo total de interrupção da conexão) / (tempo total de conexão em serviço)¹ - $[\text{Sigma}(\text{num connections affected in outage } i \times \text{duration of outage } i)] / (\text{num conns in service } \times \text{operating time})$
2. **Indisponibilidade**¹ - Disponibilidade, ou total interrupção do tempo de conexão devido a (falha de hardware, falha de software, problemas ambientais e de energia, falha de enlace ou de transporte, design da rede ou erro do usuário e falha no processo)
3. **Disponibilidade de hardware**A primeira área a ser investigada é a possível falha de hardware e o efeito na indisponibilidade. Para determinar isso, a organização precisa compreender o MTBF de todos os componentes de rede e o MTTR para problemas de hardware para todos os dispositivos em um caminho entre dois pontos. Se a rede for modular e hierárquica, a disponibilidade do hardware será a mesma entre quase dois pontos. As informações de MTBF estão disponíveis para todos os componentes da Cisco e estão disponíveis mediante

solicitação para um gerente de conta local. O programa Cisco NSA HAS também usa uma ferramenta para ajudar a determinar a disponibilidade de hardware ao longo de caminhos de rede, mesmo quando há redundância de módulo, redundância de chassi e redundância de caminho no sistema. Um fato principal de confiabilidade de hardware é o MTTR. As empresas devem avaliar a rapidez com que podem reparar o hardware quebrado. Se a empresa não tiver um plano sobressalente e depender de um contrato padrão Cisco SMARTnet™, o tempo médio de substituição em potencial será de aproximadamente 24 horas. Em um ambiente típico de LAN com redundância central e sem redundância de acesso, a disponibilidade aproximada é de 99,99% com um MTTR de 4 horas.

4. **Disponibilidade de software** A próxima área para investigação são falhas de software. Para fins de medição, a Cisco define falhas de software como o dispositivo pode ser iniciado devido a um erro de software. A Cisco fez um progresso significativo para compreender a disponibilidade do software; no entanto, versões mais recentes levam tempo para serem medidas e são consideradas menos disponíveis do que o software de implantação geral. O software de implantação geral, como IOS versão 11.2(18), foi medido em mais de 99,9999 por cento de disponibilidade. Isso é calculado com base nas partidas a frio reais nos Cisco routers usando seis minutos como o tempo de reparo (tempo para o roteador ser recarregado). As organizações que trabalham com várias versões devem obter uma disponibilidade ligeiramente inferior, devido à complexidade, à interoperabilidade e ao aumento de tempo no Troubleshooting. Espera-se que as organizações com as versões de software mais recentes tenham uma indisponibilidade maior. A distribuição para a não-disponibilidade também é muito ampla, o que significa que os clientes podem experimentar ou não-disponibilidade significativa ou disponibilidade fechada para uma versão de distribuição geral.
5. **Disponibilidade ambiental e de energia** Você também deve considerar problemas ambientais e de energia na disponibilidade. As questões ambientais dizem respeito à avaria dos sistemas de arrefecimento necessários para manter o equipamento a uma temperatura de funcionamento específica. Muitos dispositivos da Cisco serão simplesmente desligados quando estiverem consideravelmente fora da especificação, em vez de arriscar danos a todo o hardware. Para fins de orçamento de disponibilidade, a energia será usada porque é a principal causa de indisponibilidade nessa área. Embora as falhas de energia sejam um aspecto importante para determinar a disponibilidade da rede, essa discussão é limitada porque a análise teórica de energia não pode ser feita com precisão. O que uma organização deve avaliar é uma medida aproximada da disponibilidade de energia de seus dispositivos com base na experiência de sua área geográfica, de seus recursos de backup de energia e processo implementado para garantir energia com qualidade consistente a todos os dispositivos. Para uma avaliação conservadora, podemos dizer que uma empresa com geradores de backup, sistemas de UPS (Uninterruptible Power Supply [fonte de alimentação ininterrupta]) e processos de implementação de energia de qualidade podem apresentar seis noves de disponibilidade, ou 99,999%, enquanto as empresas sem esses sistemas podem ter disponibilidade de 99,99%, ou aproximadamente 36 minutos de inatividade por ano. É claro que você pode ajustar esses valores para valores mais realistas com base na percepção da empresa ou nos dados reais.
6. **Falha de link ou portadora** Falhas de link e portadora são fatores importantes relacionados à disponibilidade em ambientes WAN. Lembre-se de que os ambientes de WAN são simplesmente outras redes que estão sujeitas aos mesmos problemas de disponibilidade que a rede da organização, incluindo falha de hardware, falha de software, erro de usuário e falha de energia. Muitas redes de operadoras já executaram um orçamento de

disponibilidade em seus sistemas, mas obter essas informações pode ser difícil. Lembre-se de que as operadoras também frequentemente têm níveis de garantia de disponibilidade que têm pouca ou nenhuma base em um orçamento de disponibilidade real. Estes níveis de garantia são por vezes simples métodos de comercialização e de venda utilizados para promover a transportadora. Em alguns casos, essas redes também publicam estatísticas de disponibilidade que parecem extremamente boas. Lembre-se de que essas estatísticas podem se aplicar somente a redes de núcleo completamente redundantes e não ter em conta a indisponibilidade devido ao acesso de loop local, que é um dos principais contribuintes para a indisponibilidade em redes WAN. A criação de uma estimativa da disponibilidade para ambientes de WAN deve ser baseada nas informações reais da operadora e no nível de redundância para conectividade de WAN. Se uma organização tiver várias instalações de entrada do prédio, provedores de loop local redundantes, acesso local SONET (Synchronous-Optical-Network) e operadoras de longa distância redundantes com diversidade geográfica, a disponibilidade da WAN será consideravelmente aprimorada. O serviço de telefonia tem um orçamento de disponibilidade razoavelmente preciso para conectividade de rede não redundante em ambientes WAN. A conectividade fim-a-fim para telefones tem um orçamento de disponibilidade aproximado de 99,94% usando uma metodologia de orçamento de disponibilidade semelhante à descrita nesta seção. Essa metodologia era usada com êxito em ambientes de dados com apenas uma leve variação e, no momento, é usada como um destino na especificação do cabo do pacote para redes a cabo de provedor de serviços. Se aplicarmos esse valor a um sistema completamente redundante, podemos supor que a disponibilidade da WAN estará próxima de 99,9999 por cento disponível. É claro que muito poucas organizações têm sistemas de WAN completamente redundantes e dispersos geograficamente por causa da despesa e da disponibilidade, portanto, use o julgamento apropriado em relação a esse recurso. As falhas de link em um ambiente de LAN são menos prováveis. No entanto, os planejadores podem querer supor uma pequena quantidade de tempo de inatividade devido a conectores quebrados ou soltos. Para redes LAN, uma estimativa conservadora é de aproximadamente 99,999% de disponibilidade, ou aproximadamente 30 segundos por ano.

- 7. Projeto de rede** O projeto de rede é outro grande contribuinte para a disponibilidade. Projetos não escaláveis, erros de projeto e tempo de convergência da rede afetam negativamente a disponibilidade. **Observação:** para os fins deste documento, erros de projeto ou projeto não escaláveis estão incluídos na seção a seguir. O projeto de rede é então limitado a um valor mensurável com base na falha de software e hardware na rede que causa o reroteamento de tráfego. Esse valor é normalmente chamado de "tempo de switchover de sistema" e é um elemento dos recursos do protocolo de autocorreção no sistema. Calcule a disponibilidade simplesmente usando os mesmos métodos para cálculos do sistema. No entanto, isso não é válido a menos que o tempo de switchover de rede atenda aos requisitos de aplicativos de rede. Se o tempo de switchover for aceitável, remova-o do cálculo. Se o tempo de switchover não for aceitável, você deverá adicioná-lo aos cálculos. Um exemplo pode ser a voz sobre IP (VoIP) em um ambiente onde o tempo de switchover estimado ou real é de 30 segundos. Neste exemplo, os usuários simplesmente desligarão o telefone e possivelmente tentarão novamente. Os usuários certamente verão esse período de tempo como uma indisponibilidade, porém isso ainda não foi estimado na provisão de disponibilidade. Calcule a não disponibilidade devido ao tempo de switchover do sistema observando a disponibilidade teórica de software e hardware ao longo de caminhos redundantes, pois a comutação ocorrerá nessa área. Você deve saber o número de dispositivos que podem falhar e causar switchover no caminho redundante, o MTBF desses dispositivos e o tempo

de switchover. Um exemplo simples seria um MTBF de 35.433 horas para cada um dos dois dispositivos redundantes idênticos e um tempo de switchover de 30 segundos. Dividindo 35.433 por 8766 (média de horas por ano para incluir anos bissextos), vemos que o dispositivo falhará uma vez a cada quatro anos. Se usarmos 30 segundos como tempo de switchover, podemos então supor que cada dispositivo terá, em média, 7,5 segundos por ano de indisponibilidade devido à comutação. Como os usuários podem estar atravessando qualquer caminho, o resultado é então dobrado para 15 segundos por ano. Quando isso é calculado em termos de segundos por ano, a quantidade de disponibilidade devido ao switchover pode ser calculada como 99,99999785 por cento de disponibilidade neste sistema simples. Esse número pode ser maior em outros ambientes devido ao número de dispositivos redundantes na rede em que o switchover é possível.

8. **Erro e processo do usuário** Erro de usuário e problemas de disponibilidade de processo são as principais causas da não-disponibilidade em redes de empreendimento e de portadora. Aproximadamente 80% da indisponibilidade ocorre devido a problemas como não detecção de erros, falhas de alteração e problemas de desempenho. As organizações simplesmente não desejam usar quatro vezes todas as outras indisponibilidades teóricas para determinar o orçamento de disponibilidade, embora evidências sugiram constantemente que esse é o caso em muitos ambientes. A [próxima seção](#) discute este aspecto da não disponibilidade de maneira mais completa. Como você não pode, teoricamente, calcular a quantidade de indisponibilidade devido a erros e processos do usuário, recomendamos que você remova essa quantidade do orçamento de disponibilidade e que as empresas lutem pela perfeição. A única ressalva é que as empresas precisam entender o risco atual à disponibilidade em seus próprios processos e níveis de experiência. Após entender melhor esses riscos e inibidores, os planejadores de rede podem fatorar certa quantidade de indisponibilidade devido a essas questões. O programa Cisco NSA HAS investiga esses problemas e auxilia as organizações a compreender a indisponibilidade potencial devido a problemas de processamento, erro do usuário ou de conhecimento especializado.
9. **Determinando o orçamento de disponibilidade final** Você pode determinar o orçamento de disponibilidade geral multiplicando a disponibilidade de cada uma das áreas definidas anteriormente. Geralmente, isto é feito para ambientes homogêneos em que a conectividade é semelhante entre quaisquer dois pontos; por exemplo, um ambiente LAN modular hierárquico ou um ambiente WAN padrão hierárquico. Neste exemplo, o orçamento de disponibilidade é feito para um ambiente LAN modular hierárquico. O ambiente usa geradores de backup e sistemas UPS para todos os componentes da rede e gerencia a energia corretamente. A organização não usa VoIP e não deseja fatorar no tempo de switchover de software. As estimativas são: Disponibilidade de caminho de hardware entre dois pontos finais = disponibilidade de 99,99% Disponibilidade do software usando o software GD como referência = disponibilidade de 99,9999 por cento Disponibilidade ambiental e de energia com sistemas de backup = disponibilidade de 99,9999% Falha de link no ambiente de LAN = disponibilidade de 99,9999% Tempo de comutação do sistema não contabilizado = 100% de disponibilidade Previsão perfeita para erros de usuário e disponibilidade do processo = 100 por cento de disponibilidade O orçamento de disponibilidade final que as organizações devem se esforçar para obter é igual a $0,9999 \times 0,999999 \times 0,999999 \times 0,999999 = 0,999896$ ou 99,9896 por cento de disponibilidade. Se considerarmos a possível indisponibilidade devido a erro de usuário ou processo e assumirmos que a indisponibilidade é 4X devido a fatores técnicos, podemos supor que o orçamento de disponibilidade é de 99,95%. Esta análise de exemplo indica que a disponibilidade da LAN recairia em média entre 99,95 e 99,989 por cento. Esses números podem ser utilizados agora como uma meta

de nível de serviço para a organização da rede de comunicação. Um valor adicional pode ser obtido pela medição da disponibilidade no sistema e da definição da porcentagem de indisponibilidade devido a cada uma das seis áreas acima. Isso permite que a organização avalie fornecedores, portadoras, processos e equipes corretamente. O número também pode ser utilizado para definir expectativas dentro da empresa. Se o número for inaceitável, faça um orçamento de recursos adicionais para obter os níveis desejados. Pode ser útil para os gerentes de rede entender a quantidade de tempo de inatividade em qualquer nível de disponibilidade específico. O tempo de interrupção em minutos para um período de um ano, com qualquer nível de disponibilidade, é: $\text{Minutos de inatividade em um ano} = 525600 - (\text{Nível de disponibilidade} \times 5256)$. Se você usar o nível de disponibilidade de 99,95%, isso é igual a $525600 - (99,95 \times 5256)$ ou 262,8 minutos de inatividade. Para a definição de disponibilidade acima, é igual à quantidade média de tempo de inatividade para todas as conexões em serviço dentro da rede.

Passo 3: Criar perfis de aplicação

Os perfis de aplicativos ajudam a organização de rede a entender e definir os requisitos de nível de serviço de rede para aplicativos individuais. Isso ajuda a garantir que a rede suporta requisitos de aplicativos individuais e serviços de rede em geral. Os perfis de aplicativos também podem servir como uma linha de base documentada para o suporte de serviço de rede quando grupos de aplicativos ou servidores apontam para a rede como o problema. Em última análise, os perfis de aplicativos ajudam a alinhar objetivos de serviço de rede com requisitos de aplicativos e de negócios comparando requisitos de aplicativos como desempenho e disponibilidade a objetivos de serviço de rede realísticos ou limitações atuais. É importante não apenas para gerenciamento em nível de serviço, mas também para o design da rede geral de cima para baixo.

Crie perfis de aplicativos sempre que introduzir novos aplicativos na rede. Talvez seja necessário um acordo entre o grupo de aplicativos de TI, os grupos de administração de servidor e a rede de comunicação para ajudar a aplicar a criação do perfil de aplicativo de serviços novos e existentes. Preencha os perfis de aplicativos para aplicativos comerciais e aplicativos de sistema. Os aplicativos comerciais podem incluir e-mail, transferência de arquivos, navegação na Web, imagens médicas ou fabricação. Aplicativos de sistema podem incluir distribuição de software, autenticação de usuário, backup de rede e gerenciamento de rede.

Um analista de rede e um aplicativo ou aplicativo de suporte de servidor devem criar o perfil do aplicativo. Novos aplicativos podem exigir o uso de um analisador de protocolo e emulador de WAN com emulação de retardo para caracterizar corretamente os requisitos do aplicativo. Isso ajuda a identificar a largura de banda necessária, o retardo máximo para utilização do aplicativo e os requisitos de variação de sinal. É possível executar essa tarefa em um ambiente de laboratório desde que você tenha os servidores necessários. Em outros casos, como no VoIP, os requisitos de rede, incluindo instabilidade, atraso e largura de banda, são bem publicados e testes de laboratório não serão necessários. Um perfil de aplicativo deve incluir os seguintes itens:

- Nome do aplicativo
- Tipo de aplicação
- Novo aplicativo?
- Importância comercial
- Requisitos de disponibilidade
- Protocolos e portas usados
- Largura de banda estimada do usuário (kbps)

- Número e localização dos utilizadores
- Requisitos de transferência de arquivos (incluindo tempo, volume e endpoints)
- Impacto da interrupção da rede
- Requisitos de atraso, instabilidade e disponibilidade

O objetivo do perfil de aplicativo é entender requisitos de negócios para o aplicativo, a crítica de negócio e os requisitos de rede como largura de banda, retardo e jitter. Além disso, a organização de rede deve entender o impacto do tempo de inatividade da rede. Em alguns casos, você precisará de reinicializações de aplicativos ou servidores que aumentem significativamente o tempo de inatividade geral do aplicativo. Quando você completa o perfil do aplicativo, você pode comparar todos os recursos de rede e ajudar a alinhar os níveis de serviço de rede com os requisitos de negócios e de aplicativos.

Passo 4: Definir padrões de disponibilidade e desempenho

Os padrões de disponibilidade e desempenho definem as expectativas de serviço para a empresa. Estes podem ser definidos para diferentes áreas da rede ou de aplicativos específicos. O desempenho também pode ser definido em termos de atraso de ida e volta, instabilidade, throughput máximo, compromissos de largura de banda e escalabilidade geral. Além de definir as expectativas de serviço, a organização também deve ter cuidado para definir cada um dos padrões de serviço para que os usuários e grupos de TI que trabalham com redes compreendam completamente o padrão de serviço e como ele se relaciona com seus requisitos de administração de aplicativos ou servidores. Usuário e grupos de TI também devem entender como o padrão de serviço deve ser medido.

Resultados a partir de passos de definição de nível de serviço prévio auxiliarão a criar o padrão. Neste ponto, a organização de rede deve ter uma compreensão clara dos riscos e restrições atuais na rede, uma compreensão do comportamento do aplicativo e uma análise teórica de disponibilidade ou linha de base de disponibilidade.

1. Defina as áreas geográficas ou de aplicação onde os padrões de serviço serão aplicados. Isto pode incluir áreas como LAN de campus, WAN doméstica, extranet ou conectividade de parceiro. Em alguns casos, a organização pode ter diferentes metas de nível de serviço em uma área. Isso é comum para organizações provedoras de serviço ou empreendimentos. Nesses casos, não seria incomum criar padrões de nível de serviço diferentes com base em requisitos de serviço individuais. Eles podem ser classificados como padrões de serviço ouro, prata e bronze dentro de uma área geográfica ou de serviços.
2. Defina os parâmetros do padrão de serviço. A disponibilidade e o atraso de ida e volta são os padrões de serviço de rede mais comuns. Throughput máximo, compromisso de largura de banda mínima, instabilidade, taxas de erro aceitáveis e recursos de escalabilidade também podem ser incluídos conforme necessário. Tenha cuidado ao revisar o parâmetro de serviço para métodos de medição. Independentemente de o parâmetro seguir para um SLA, a organização deve pensar sobre como o parâmetro de serviço poderá ser medido ou justificado quando ocorrerem problemas ou desacordos de serviço.

Depois de definir as áreas de serviço e os parâmetros de serviço, use as informações das etapas anteriores para criar uma matriz de padrões de serviço. A organização também precisará definir áreas que podem ser confusas para usuários e grupos de TI. Por exemplo, o tempo máximo de resposta será muito diferente para um ping round trip do que para pressionar a tecla Enter em um local remoto para um aplicativo específico. A tabela a seguir mostra os objetivos de desempenho nos Estados Unidos.

Área de rede	Destino de disponibilidade	Método de medição	Tempo médio de resposta da rede	Tempo máximo de resposta aceito	Método de Medição do Tempo de Resposta
LAN	99.99%	Minutos de usuários afetados	Menos de 5 ms	10 ms	Resposta de ping round trip
WAN	99.99%	Minutos de usuários afetados	Abaixo de 100 ms (ping round trip)	150 ms	Resposta de ping round trip
WAN e extranet essenciais	99.99%	Minutos de usuários afetados	Abaixo de 100 ms (ping round trip)	150 ms	Resposta de ping round trip

[Passo 5: Definir serviço de rede](#)

Esta é a última etapa para o gerenciamento de nível de serviço básico; ele define os processos reativos e proativos e os recursos de gerenciamento de rede que você implementa para atingir metas de nível de serviço. O documento final é geralmente chamado de plano de suporte operacional. A maioria dos planos de suporte a aplicativos inclui apenas requisitos de suporte reativo. Em ambientes de alta disponibilidade, a organização também deve considerar processos de gerenciamento pró-ativo que serão usados para isolar e resolver problemas de rede antes que as chamadas de serviço do usuário sejam iniciadas. Acima de tudo, o documento final deve:

- Descrever o processo reativo e pró-ativo usado para atingir o objetivo do nível de serviço
- Como o processo de serviço será gerenciado
- Como o objetivo e o processo de serviço serão medidos.

Esta seção contém exemplos para definições de serviços reativos e definições de serviços proativos a serem consideradas para muitas empresas de provedores de serviços e empresas. O objetivo de desenvolver as definições do nível de serviço é criar um serviço que irá satisfazer as metas de disponibilidade e desempenho. Para conseguir isso, a organização deve criar o serviço tendo em mente as restrições técnicas, o orçamento disponível e os perfis de aplicativos atuais. Especificamente, a organização deve definir e criar um serviço que identifique e resolva problemas de forma consistente e rápida nos tempos alocados pelo modelo de disponibilidade. A organização também deve definir um serviço que possa identificar e resolver rapidamente os problemas potenciais de serviços que afetarão a disponibilidade e o desempenho, se ignorados.

Você não atingirá o nível de serviço desejado da noite para o dia. Deficiências, como pouca experiência, limitações do processo atual ou nível inadequado dos profissionais, podem impedir que a organização alcance seus padrões desejados ou atinja suas metas, mesmo após realizar os passos de análise de serviço anteriores. Não existe método preciso para fazer a

correspondência exata do nível de serviço necessário com os objetivos desejados. Para acomodar isso, a organização deve medir os padrões de serviço e os parâmetros de serviço usados para suportar os padrões de serviço. Quando a organização não está atendendo às metas de serviço, ela deve procurar métricas de serviço para ajudar a entender o problema. Em muitos casos, podem ser feitos aumentos no orçamento para melhorar os serviços de suporte e fazer melhorias necessárias para atingir os objetivos de serviço desejados. Com o tempo, a organização pode fazer vários ajustes, tanto na meta de serviço quanto na definição do serviço, a fim de alinhar os serviços de rede com os requisitos comerciais.

Por exemplo, uma empresa pode atingir 99% de disponibilidade quando a meta era muito maior com 99,9% de disponibilidade. Ao examinar métricas de serviço e suporte, representantes da organização descobriram que a substituição do hardware estava demorando cerca de 24 horas, o que era muito mais do que a estimativa original, pois a organização havia orçado apenas quatro horas. Além disso, a empresa descobriu que os recursos de gerenciamento pró-ativo estavam sendo ignorados e que os dispositivos de rede redundantes não estavam sendo reparados. Também descobriram que não tinham pessoal para fazer melhorias. Como resultado, após considerar a redução das metas de serviço atuais, a organização orçou recursos adicionais necessários para atingir o nível de serviço desejado.

As definições de serviço devem incluir definições de suporte reativo e definições proativas. As definições reativas definem como a organização reagirá aos problemas depois que eles forem identificados a partir de reclamações do usuário ou recursos de gerenciamento de rede. As definições proativas descrevem como a organização identificará e resolverá possíveis problemas de rede, incluindo reparo de componentes de rede "em espera" quebrados, detecção de erros e limites de capacidade e atualizações. As seções a seguir fornecem exemplos de definições de nível de serviço reativo e proativo.

Reativar definições do nível de serviço

As seguintes áreas de nível de serviços são tipicamente medidas usando as estatísticas de um banco de dados de help desk e auditorias periódicas. Esta tabela mostra um exemplo da gravidade do problema para uma organização. Observe que o gráfico não inclui como lidar com solicitações de um novo serviço, que podem ser tratadas por um SLA ou perfil de aplicativo adicional e análise de "e se" de desempenho. Normalmente a gravidade 5 pode ser uma solicitação para o novo serviço, se controlado por meio do mesmo processo de suporte.

Severidade 1	Severidade 2	Severidade 3	Severidade 4
<i>Impacto comercial grave no segmento de usuário ou servidor da LAN abaixo do local crítico</i>	<i>Alto impacto comercial por perda ou degradação, possível solução alternativa no local da LAN do campus; 5 a 99 usuários afetaram o local da WAN doméstica em um local da WAN internacional e reduziram o impacto crítico no</i>	<i>Algumas funcionalidades específicas da rede são perdidas ou degradadas, como perda de redundância</i> Desempenho da LAN do campus afetou a perda de redundância da	<i>Uma consulta ou falha funcional que não tem impacto comercial</i>

da WAN	desempenho	LAN	<i>para a organização</i>
--------	------------	-----	---------------------------

Após a definição da gravidade do problema, defina ou investigue o processo de suporte para criar definições de resposta de serviço. Em geral, as definições de resposta de serviço exigem uma estrutura de suporte hierárquico, juntamente com um sistema de suporte de software de help desk para rastrear problemas por meio de tíquetes de problemas. As métricas também devem estar disponíveis no tempo de resposta e no tempo de resolução para cada prioridade, número de chamadas por prioridade e qualidade de resposta/resolução. Para definir o processo de suporte, é útil definir as metas de cada camada de suporte na organização e suas funções e responsabilidades. Isso é útil para que a organização compreenda os requisitos de recurso e os níveis de especialização para cada nível de suporte. As tabelas a seguir fornecem um exemplo de uma organização de suporte alinhado com diretrizes de resolução de problemas.

Nível de suporte	Responsabilidade	Metas
Tier 1 Suporte	Suporte de help desk em tempo integral Atenda chamadas de suporte, coloque tíquetes de problemas, trabalhe com problemas em até 15 minutos, registre o tíquete e encaminhe-o para o suporte adequado de nível 2	Resolução de 40% das chamadas recebidas
Tier 2 Suporte	Monitoramento de fila, gerenciamento de rede, monitoramento de estação Colocar tíquetes de problemas para problemas identificados por software Implementar Fazer chamadas de nível 1, fornecedor e escalonamento de nível 3 Suponha posse de chamada até a resolução	Resolução de 100% das chamadas no nível 2
Tier 3 Suporte	Deve fornecer suporte imediato ao nível 2 para todos os problemas de prioridade 1 Concordo para ajudar com todos os problemas não resolvidos pelo nível 2 no período de resolução do SLA	Sem propriedade direta de problemas

A próxima etapa é criar a matriz para a resposta do serviço e a definição do serviço de resolução de serviços. Isto define metas para o modo como os problemas são solucionados rapidamente, inclusive substituição de hardware. É importante definir metas nessa área porque o tempo de resposta e o tempo de recuperação do serviço afetam diretamente a disponibilidade da rede. Os horários de resolução de problemas também devem estar alinhados ao orçamento de disponibilidade. Se um grande número de problemas de alta gravidade não for contabilizado no

orçamento de disponibilidade, a organização poderá então trabalhar para entender a origem desses problemas e uma possível correção. Veja a seguinte tabela:

Severidade do problema	Resposta do Help Desk	Resposta da camada 2	Nível 2 no local	Substituição de hardware	Resolução de problemas
1	Escalamento imediato para nível 2, gerenciador de operações de rede	5 minutos	2 horas	2 horas	4 horas
2	Escalamento imediato para nível 2, gerenciador de operações de rede	5 minutos	4 horas	4 horas	8 horas
3	15 minutos	2 horas	12 horas	24 horas	36 horas
4	15 minutos	4 horas	3 dias	3 dias	6 dias

Além da resposta do serviço e da resolução do serviço, crie uma matriz para encaminhamento. A matriz de escalonamento ajuda a garantir que os recursos disponíveis se concentrem em problemas que afetam gravemente o serviço. Em geral, quando os analistas estão focados na solução de problemas, eles raramente se concentram em trazer recursos adicionais para o problema. Definir quando recursos adicionais devem ser notificados ajuda a promover a conscientização sobre problemas no gerenciamento e pode, em geral, ajudar a levar a futuras medidas proativas ou preventivas. Veja a seguinte tabela:

Tempo decorrido	Severidade 1	Severidade 2	Severidade 3	Severidade 4
5 minutos	Gerente de operações de rede, suporte de nível 3, diretor de rede			
1 hora	Atualização para gerenciador de operações de rede, suporte de nível 3, diretor de	Atualização para gerenciador de operações de rede, suporte de nível 3, diretor de rede de comunicação		

	rede de comunicação			
2 horas	Expandir para VP, atualizar para diretor, gerente de operações			
4 horas	Notificação do CEO sobre a razão principal de consultas não resolvidas para o VP, o diretor, o gerente de operações, o suporte de nível 3	Expandir para VP, atualizar para diretor, gerente de operações		
24 horas			Gerente de operações de rede	
5 dias				Gerente de operações de rede

Até aqui, as definições de nível de serviços enfatizavam a forma como a organização de suporte a operações reage a problemas depois que eles são identificados. As empresas de operações têm criado, há anos, planos de suporte operacionais com informações similares às acima. No entanto, o que está faltando nesses casos é como a organização identificará problemas e quais problemas identificará. Organizações de rede mais sofisticadas tentaram resolver esse problema simplesmente criando metas para a porcentagem de problemas que são identificados proativamente, ao contrário dos problemas identificados de forma reativa pelo relatório de problemas ou reclamação do usuário.

A tabela a seguir mostra como uma organização pode querer medir capacidades de suporte proativo e o total de suporte proativo.

Área de rede	Razão de identificação do problema proativo	Taxa de identificação do problema reativo
LAN	80 %	20 %

WAN	80 %	20 %
-----	------	------

Esse é um bom começo para definir definições de suporte mais proativas, pois é simples e bastante fácil de medir, especialmente se ferramentas proativas geram automaticamente tíquetes de problemas. Isso também ajuda a concentrar as ferramentas/informações de gerenciamento de rede na solução proativa de problemas, em vez de ajudar na causa básica. No entanto, o principal problema com esse método é que ele não define requisitos de suporte pró-ativo. Isto geralmente cria intervalos nas capacidades de gerenciamento de suporte proativo e resulta em risco adicional à disponibilidade.

Definições niveladas de serviço proativo

Uma metodologia mais abrangente para criar definições de nível de serviço inclui mais detalhes sobre como a rede é monitorada e como a organização operacional reage aos limites definidos da NMS (network management station, estação de gerenciamento de rede) em uma base 7 x 24. Essa tarefa poderá ser considerada impossível devido ao número precipitado de variáveis de Base de Informação de Gerenciamento (MIB) e à quantidade de informações de gerenciamento de rede que são pertinentes ao funcionamento da rede. Também poderia ser extremamente caro e usar muitos recursos. Infelizmente, essas objeções impedem que muitas pessoas implementem uma definição de serviço proativa que, por natureza, deve ser simples, bem fácil de seguir e aplicável apenas à melhor disponibilidade ou a riscos de desempenho na rede. Se uma empresa perceber o valor nas definições básicas de serviço pró-ativo, mais variáveis podem ser adicionadas ao longo do tempo sem impacto significativo, desde que você implemente uma abordagem em fases.

Inclua a primeira área de definições de serviço pró-ativo em todos os planos de suporte às operações. A definição do serviço simplesmente declara como o grupo de operações identificará e responderá proativamente às condições de rede ou link inativo em diferentes áreas da rede. Sem esta definição (ou suporte de gerenciamento), a organização pode esperar suporte variável, expectativas não realísticas de usuários e, enfim, uma disponibilidade mais baixa de rede.

A tabela a seguir mostra como uma organização criaria uma definição de serviço para as condições de link/dispositivo desativado. O exemplo mostra uma organização corporativa que pode ter diferentes requisitos de notificação e resposta com base na hora do dia e área da rede.

Dispositivo de rede ou link desativado	Método de detecção	Notificação 5 x 8	Notificação 7 x 24	Resolução 5 x 8	Resolução 7 x 24
LAN de núcleo	Dispositivo de SNMP e apuração de link,	NOC cria tíquete e de problema, pager de serviço	pager de serviço de LAN de página autom	Analista de LAN atribuído em 15 minutos pelo NOC, reparo conforme definição de resposta de serviço	Prioridades 1 e 2 investigação e resolução imediata Prioridades 3 e 4 para resolução

	armadilhas	o de LAN de página	ática, pessoa de serviço de LAN criação de ticket e de problema para a fila de LAN central		matinal
WAN doméstica	Dispositivo de SNMP e apuração de link, armadilhas	NOC criação de ticket e de problemas, pager de serviço de WAN de página	pager de serviço de WAN de página automática, pessoa de serviço de WAN criação de ticket e de problema para a fila de WAN	Analista de WAN atribuído dentro de 15 minutos por NOC, reparo como definição de resposta por serviço	Prioridades 1 e 2 investigação e resolução imediata Prioridades 3 e 4 para resolução matinal
Extranet	Dispositivo de SNMP e apuração de link, armadilhas	NOC criação de ticket e de problema, pager de obrigação do parceiro de	pager de obrigação do parceiro de página automática, pessoa de	Analista de parceiro designado em 15 minutos pelo NOC; faça o reparo de acordo com a definição de resposta de serviço	Investigação e resolução imediata das prioridades 1 e 2; Fila de prioridades 3 e 4 para resolução matinal

		página	dever do parceiro criativo e de problema para a fila do parceiro		
--	--	--------	--	--	--

As definições de nível de serviço pró-ativo restantes podem ser divididas em duas categorias: erros de rede e problemas de capacidade/desempenho. Apenas uma pequena porcentagem das organizações de rede possui definições de nível de serviço nessas áreas. Como resultado, esses problemas são ignorados ou tratados esporadicamente. Isso pode ser adequado para alguns ambientes de rede, mas ambientes de alta disponibilidade geralmente exigirão de gerenciamento proativo de serviços.

As organizações de rede tendem a lidar com definições de serviços proativos por vários motivos. Isso ocorre principalmente porque não executaram uma análise de requisitos para definições de serviço proativo com base nos riscos de disponibilidade, orçamento de disponibilidade e problemas de aplicativo. Isso leva a requisitos pouco claros para definições de serviços proativos e benefícios não claros, especialmente porque podem ser necessários recursos adicionais.

O segundo motivo envolve balancear a quantidade de gerenciamento proativo que pode ser feito com recursos existentes ou recém-definidos. Gere somente os alertas que tenham um possível impacto importante na disponibilidade ou no desempenho. Você também deve considerar o gerenciamento ou os processos de correlação de eventos para garantir que vários tíquetes de problemas proativos não sejam gerados para o mesmo problema. O último motivo com o qual as organizações podem implicar é que criar um novo conjunto de alertas muitas vezes pode gerar uma inundação inicial de mensagens que anteriormente passaram despercebidas. O grupo de operações deve estar preparado para essa inundação inicial de problemas e recursos adicionais de curto prazo para corrigir ou resolver essas condições não detectadas anteriormente.

A primeira categoria de definições de nível de serviço pró-ativo é erros de rede. Os erros de rede podem ser subdivididos em erros de sistema que incluem erros de software ou de hardware, erros de protocolo, erros de controle de mídia, erros de precisão e avisos ambientais. O desenvolvimento de uma definição de nível de serviço começa com uma compreensão geral de como essas condições de problema serão detectadas, quem as verá e o que acontecerá quando ocorrerem. Adicione mensagens ou problemas específicos à definição do nível de serviço se houver necessidade. Também pode ser necessário um trabalho adicional nas áreas a seguir para assegurar o sucesso:

- Responsabilidades de suporte de nível 1, nível 2 e nível 3
- Equilibrar a prioridade das informações de gerenciamento de rede com a quantidade de trabalho pró-ativo que o grupo de operações pode lidar com eficiência
- Requisitos de treinamento para garantir que a equipe de suporte técnico possa tratar de modo eficaz os alertas definidos
- Metodologias de correlação de eventos para garantir que vários tíquetes de problemas não sejam gerados para o mesmo problema de causa raiz

- Documentação sobre mensagens ou alertas específicos que ajudam na identificação de eventos no nível de suporte de nível 1

A tabela a seguir mostra um exemplo de definição de nível de serviço para erros de rede que permite entender claramente quem é responsável pelos alertas de erro de rede proativos, como o problema será identificado e o que acontecerá quando ele ocorrer. A organização pode ainda precisar de esforços adicionais, conforme definido acima, para garantir o sucesso

s.

Categoria de erro	Método de detecção	Limite	Ação realizada
Erros de Software (travamentos causados por software)	Revisão diária das mensagens de syslog usando o visualizador de syslog Concluído pelo suporte de nível 2	Qualquer ocorrência para prioridade 0, 1 e 2 Sobre 100 ocorrências de nível 3 ou superior	Revise o problema, crie bilhete de problema e envie caso uma nova ocorrência ou problema exija atenção
Erros de hardware (travamentos forçados pelo hardware)	Revisão diária das mensagens de syslog usando o visualizador de syslog Concluído pelo suporte de nível 2	Qualquer ocorrência para prioridade 0, 1 e 2 Sobre 100 ocorrências de nível 3 ou superior	Revise o problema, crie bilhete de problema e envie caso uma nova ocorrência ou problema exija atenção
Erros de protocolo (somente protocolos de roteamento IP)	Revisão diária das mensagens de syslog usando o visualizador de syslog Concluído pelo suporte de nível 2	Dez mensagens por dia de prioridades 0, 1 e 2 Mais de 100 ocorrências de nível 3 ou superior	Revise o problema, crie bilhete de problema e envie caso uma nova ocorrência ou problema exija atenção
Erros de controle de mídia (somente FDDI, POS e Fast Ethernet)	Revisão diária das mensagens de syslog usando o visualizador de syslog Concluído pelo suporte de nível 2	Dez mensagens por dia de prioridades 0, 1 e 2 Mais de 100 ocorrências de nível 3 ou superior	Revise o problema, crie bilhete de problema e envie caso uma nova ocorrência ou problema exija atenção
Mensagens ambientais	Revisão diária das mensagens de syslog usando	Qualquer mensagem	Criar tíquete de problemas e enviar para novos

(energia e temperatura)	o visualizador de syslog Concluído pelo suporte de nível 2		problemas
Erros de precisão (erros de entrada de link)	Pesquisa SNMP em intervalos de 5 minutos Eventos de limite recebidos pelo NOC	Erros de entrada ou saída Um erro em qualquer intervalo de 5 minutos em qualquer link	Crie um tíquete de problemas para novos problemas e envie para o suporte de nível 2

A outra categoria de definições de nível de serviço pró-ativo se aplica ao desempenho e à capacidade. O desempenho real e o gerenciamento de capacidade incluem gerenciamento de exceções, avaliação comparativa e tendência e análise what if. A definição do nível de serviço simplesmente define os limiares de desempenho e de exceção de capacidade e os limiares médios que iniciarão a investigação ou a atualização. Estes limiares podem, então, ser aplicados aos três processos de gerenciamento de desempenho e de capacidade de alguma maneira.

As definições de nível de serviço de capacidade e desempenho podem ser divididas em várias categorias: links de rede, dispositivos de rede, desempenho de ponta a ponta e desempenho de aplicativos. O desenvolvimento de definições de nível de serviço nessas áreas exige um conhecimento técnico aprofundado sobre aspectos específicos da capacidade do dispositivo, capacidade da mídia, características de QoS e requisitos de aplicativos. Por esse motivo, recomendamos que os arquitetos de rede desenvolvam definições de nível de serviço relacionadas ao desempenho e à capacidade com a entrada do fornecedor.

Como os erros de rede, o desenvolvimento de uma definição de nível de serviço para capacidade e desempenho começa com uma compreensão geral de como essas condições de problema serão detectadas, quem as observará e o que acontecerá quando ocorrerem. Você pode adicionar definições de eventos específicos para a definição de nível de serviço, caso haja necessidade. Também pode ser necessário um trabalho adicional nas áreas a seguir para assegurar o sucesso:

- Uma compreensão clara dos requisitos de desempenho do aplicativo
- Investigação técnica detalhada sobre os valores-limite que fazem sentido para a organização com base nos requisitos comerciais e nos custos gerais
- Requisitos de ciclo orçamental e atualização fora de ciclo
- Responsabilidades de suporte de nível 1, nível 2 e nível 3
- A prioridade e criticalidade das informações de gerenciamento da rede equilibradas com a quantidade de trabalho proativo que o grupo de operações pode manipular efetivamente
- Requisitos de treinamento para garantir que a equipe de suporte compreenda as mensagens ou alertas e possa lidar efetivamente com a condição definida
- Metodologias ou processos de correlação de eventos para garantir que vários tíquetes de problemas não sejam gerados para o mesmo problema de causa raiz
- Documentação sobre mensagens ou alertas específicos que ajudam na identificação de eventos no nível de suporte de nível 1

A tabela a seguir mostra um exemplo de definição de nível de serviço para utilização de link que fornece uma compreensão clara de quem é responsável por alertas de erro de rede proativos,

como o problema será identificado e o que acontecerá quando o problema ocorrer. Conforme definido anteriormente, podem ser necessários esforços adicionais da organização para garantir o sucesso.

Área/ Mídia de Rede	Método de detecção	Limite	Ação realizada
Links de distribuição e backbone de LAN de campus	Pesquisa SNMP em intervalos de 5 minutos de armadilhas de exceção RMON no núcleo e links de distribuição	50% de utilização em intervalos de 5 minutos 90% de utilização através de armadilha de exceção	Notificação de e-mail para o grupo de alias de e-mail de desempenho para avaliar o requisito de QoS ou planejar atualização para problemas recorrentes
Links WAN domésticos	Eleição SNMP em intervalos de 5 minutos	75% de utilização em intervalos de 5 minutos	Notificação de e-mail para o grupo de alias de e-mail de desempenho para avaliar o requisito de QoS ou planejar atualização para problemas recorrentes
Links de WAN da Extranet	Eleição SNMP em intervalos de 5 minutos	60% de utilização em intervalos de 5 minutos	Notificação de e-mail para o grupo de alias de e-mail de desempenho para avaliar o requisito de QoS ou planejar atualização para problemas recorrentes

A tabela a seguir define definições de nível de serviço para limiares de capacidade e desempenho do dispositivo. Assegure-se de criar limiares significativos e úteis para evitar problemas de rede ou problemas de disponibilidade. Essa é uma área muito importante pois problemas de recurso do plano de controle do dispositivo não verificados podem ter um impacto grave sobre a rede.

Cisco 7500	CPU, memória, buffers	SNMP polling em intervalos de 5 - 5	CPU em 75% durante intervalos de 5 minutos, 99% via memória de notificação	Notificação por e-mail para o grupo de alias de e-mail de desempenho
------------	-----------------------	-------------------------------------	--	--

		minutos notificação RMON para CPU	RMON a 50% durante intervalos de 5 minutos Buffers a 99% de utilização	e capacidade para resolver problemas ou planejar atualização de CPU RMON em 99%, colocar tíquete de problema e pager de suporte de nível 2 da página
Cisco 2600	CPU, memória	Eleição SNMP em intervalos de 5 minutos	CPU a 75% durante intervalos de 5 minutos Memória a 50% durante intervalos de 5 minutos	Notificação por e-mail para o grupo de alias de e- mail de desempenho e capacidade para resolver problemas ou planejar atualização
Catalyst 5000	Utilização do backplane, memória	Eleição SNMP em intervalos de 5 minutos	Backplane com 50% de utilização Memória com 75% de utilização	Notificação por e-mail para o grupo de alias de e- mail de desempenho e capacidade para resolver problemas ou planejar atualização
Switch ATM LightStream® 1010	CPU, memória	Eleição SNMP em intervalos de 5 minutos	CPU com 65% de utilização Memória com 50% de utilização	Notificação por e-mail para o grupo de alias de e- mail de desempenho e capacidade para resolver problemas ou planejar atualização

A próxima tabela estabelece as definições de nível de serviço para desempenho e capacidade de ponta a ponta. Esses limiares geralmente se baseiam nos requisitos de aplicativo, mas também podem ser usados para indicar algum tipo de problema de desempenho de rede ou capacidade. A maioria das organizações com definições de nível de serviço para desempenho cria apenas um

punhado de definições de desempenho, pois medir o desempenho de cada ponto da rede para todos os outros pontos exige recursos significativos e cria uma grande quantidade de sobrecarga de rede. Esses problemas de desempenho de ponta a ponta também podem ser obtidos em limiares de recurso de dispositivo ou enlace. Recomendamos definições gerais por área geográfica. Alguns sites ou links críticos podem ser adicionados, se necessário.

Área/ Mídia de Rede	Método de medição	Limite	Ação realizada
LAN de camp us	Nenhum Nenhum problema esperado Difícil medir toda a infraestrutura de LAN	Tempo de resposta de round trip de 10 milissegundos ou menos todas as vezes	Notificação por e-mail para o grupo de alias de e-mail de desempenho e capacidade para resolver problemas ou planejar atualização
Links WAN domé stico s	Medição atual de SF para NY e SF para Chicago usando apenas eco ICMP do Internet Performance Monitor (IPM)	Tempo médio de resposta de ida e volta de 75 milissegundos no período de 5 minutos	Notificação de e- mail para o grupo de alias de e-mail de desempenho para avaliar o requisito de QoS ou planejar atualização para problemas recorrentes
San Fran cisco até Tóqui o	Medição atual de São Francisco a Bruxelas usando IPM e eco ICMP	Tempo de resposta de ida e volta de 250 milissegundos , em média, durante um período de 5 minutos	Notificação de e- mail para o grupo de alias de e-mail de desempenho para avaliar o requisito de QoS ou planejar atualização para problemas recorrentes
São Fran cisco a Brux elas	Medição atual de São Francisco a Bruxelas usando IPM e eco ICMP	Tempo de resposta de ida e volta de 175 milissegundos , em média, durante um período de 5 minutos	Notificação de e- mail para o grupo de alias de e-mail de desempenho para avaliar o requisito de QoS ou planejar atualização para problemas

			recorrentes
--	--	--	-------------

A área final das definições de nível de serviço é para desempenho do aplicativo. As definições de nível de serviço de desempenho de aplicativos são normalmente criadas pelo grupo de administração de aplicativos ou servidores porque o desempenho e a capacidade dos próprios servidores são provavelmente o maior fator no desempenho dos aplicativos. As organizações de rede podem obter um enorme benefício criando definições de nível de serviço para o desempenho de aplicativos de rede, pois:

- medidas e definições de nível de serviço podem ajudar a eliminar conflitos entre grupos.
- as definições de nível de serviço dos aplicativos individuais são importantes quando o QoS está configurado para aplicativos chave e outro tráfego é considerado opcional.

Se você optar por criar e medir o desempenho do aplicativo, provavelmente é melhor não medir o desempenho para o próprio servidor. Então, isso ajuda a distinguir entre problemas de rede e problemas de aplicativo ou servidor. Use sondas ou o software de agente de disponibilidade de sistema em execução nos Cisco routers e o Cisco IPM que controla o tipo de pacote e a frequência de medição.

A tabela a seguir mostra uma definição de nível de serviço simples para desempenho do aplicativo.

Aplicativo	Método de medição	Limite	Ação realizada
Aplicativo Enterprise Resource Planning (ERP) Porta TCP 1529 Bruxelas para SF	Bruxelas a São Francisco usando o IPM medindo a porta 1529 desempenho de ida e volta, gateway Bruxelas para o gateway SFO 2	Tempo de resposta de ida e volta de 175 milissegundos, em média, durante um período de 5 minutos	Notificação por e-mail para o grupo de alias de e-mail de desempenho para avaliar problemas ou planejar atualização para problemas recorrentes
Aplicativo ERP Porta TCP 1529 Tokyo para SF	Bruxelas a São Francisco usando o IPM medindo a porta 1529 desempenho de ida e volta, gateway Bruxelas para o gateway SFO 2	Tempo de resposta de ida e volta de 200 milissegundos, em média, durante um período de 5 minutos	Notificação por e-mail para o grupo de alias de e-mail de desempenho para avaliar problemas ou planejar atualização para problemas recorrentes
Porta TCP do aplicativo de	Sydney a São Francisco usando a porta de medição	Tempo de resposta de ida e volta de	Notificação por e-mail para o grupo de alias de e-mail de

suporte ao cliente 1702 Sydney para SF	IPM de desempenho de ida e volta 1702 Gateway Sydney para gateway SFO 1	250 milissegundos, em média, durante um período de 5 minutos	desempenho para avaliar problemas ou planejar atualização para problemas recorrentes
--	---	--	--

Passo 6: Coletar métricas e monitorar

as definições de nível de serviço, por si só, são inúteis, a menos que a organização obtenha êxito de métricas e monitores. Ao criar uma definição de nível de serviço crítico, defina como o nível de serviço será medido e informado. A medição do nível de serviço determina se a organização está atendendo aos objetivos e também identifica a causa principal dos problemas de disponibilidade ou desempenho. Considere também o objetivo ao escolher um método para medir a definição do nível de serviço. Consulte [Criação e manutenção de SLAs](#) para obter mais informações.

O monitoramento dos níveis de serviço exige a realização de uma reunião de revisão periódica, normalmente todos os meses, para discutir o serviço periódico. Discuta todas as métricas e se elas estão em conformidade com os objetivos. Se não estiverem em conformidade, determine a causa principal do problema e implemente melhorias. Você também deve abranger as iniciativas e o progresso atuais no sentido de melhorar situações individuais.

Criação e manutenção de SLAs

as definições de nível de serviço são um excelente bloco de construção já que ajudam a criar um QoS consistente em toda a organização e ajuda a aprimorar a disponibilidade. A próxima etapa são os SLAs, que são uma melhoria porque alinham os objetivos comerciais e os requisitos de custo diretamente com a qualidade do serviço. O SLA bem construído serve então como um modelo para eficiência, qualidade e sinergia entre a comunidade de usuários e o grupo de suporte, mantendo processos e procedimentos claros para problemas ou problemas de rede.

Os SLAs oferecem vários benefícios:

- Os SLAs estabelecem responsabilidade em dois sentidos para serviço, o que significa que os usuários e os grupos de aplicativo também são responsáveis pelo serviço da rede. Se eles não ajudarem a criar um SLA para um serviço específico e a comunicar o impacto comercial com o grupo de rede, eles podem ser realmente responsáveis pelo problema.
- Os SLAs ajudam a determinar as ferramentas e os recursos padrão necessários para atender às necessidades dos negócios. Decidir quantas pessoas e quais ferramentas usar sem SLAs é frequentemente uma estimativa orçamentária. O serviço pode estar com excesso de engenharia, o que leva ao gasto excessivo ou à subengenharia, que faz com que os objetivos comerciais não sejam atendidos. Ajustar os SLAs ajuda a alcançar esse nível ótimo de equilíbrio.
- O SLA documentado cria um veículo mais explícito para a configuração das expectativas de nível de serviço.

Para o desenvolvimento de SLAs após a criação de definições de nível de serviço, recomendamos os seguintes passos: Para o desenvolvimento de SLAs após a criação de definições de nível de serviço, recomendamos os seguintes passos:

7. [Atenda aos pré-requisitos dos SLAs.](#)
8. [Determine as partes envolvidas no SLA.](#)
9. [Determine os elementos de serviço.](#)
10. [Compreender os Objetivos e as Necessidades de Negócios dos Clientes](#)
11. [Defina o SLA necessário para cada grupo.](#)
12. [Escolha o formato do SLA](#)
13. [Desenvolver grupos de trabalho de SLA](#)
14. [Realizar reuniões de grupo de trabalho e preparar o SLA.](#)
15. [Negocie o SLA.](#)
16. [Meça e monitore a conformidade do SLA.](#)

Passo 7: Atenda aos pré-requisitos dos SLAs

Os especialistas em desenvolvimento de SLA de TI identificaram três pré-requisitos para um SLA bem-sucedido. Infelizmente, organizações que não alcançam esses objetivos podem esperar problemas com o processo SLA e devem considerar os possíveis problemas envolvidos com o processo SLA. A não implementação de SLAs não é prejudicial se a organização de rede puder criar definições de nível de serviço que atendam aos requisitos gerais da empresa. Os seguintes são pré-requisitos para o processo de SLA:

- Sua empresa deve ter uma cultura orientada para serviços. A organização deve colocar as necessidades dos clientes em primeiro lugar. É preciso um comprometimento completo com o serviço, resultando em um entendimento total das necessidades e percepções do cliente. Realizar pesquisas de satisfação do cliente e iniciativas de serviço voltadas para o cliente. Outro indicador de serviço pode ser que a organização declare a satisfação do serviço ou do suporte como um objetivo corporativo. Isto não é raro, pois as organizações de TI estão fortemente ligadas ao êxito geral da empresa. A cultura do serviço é importante pois o processo SLA é fundamentalmente sobre como fazer melhorias com base nas necessidades do cliente e nos requisitos comerciais. Se as empresas não tiverem feito isso no passado, o processo de SLA será difícil.
- As iniciativas do cliente/empresa devem orientar todas as atividades de TI. A visão ou a declaração de missão da empresa deve estar alinhada às iniciativas dos clientes e dos negócios, que impulsionam todas as atividades de TI, incluindo os SLAs. É muito comum que se crie uma rede para atender a um determinado objetivo, ainda que o grupo de rede de comunicação perca de vista esse objetivo e os requisitos de negócios subsequentes. Nesses casos, um determinado orçamento é atribuído à rede, o que pode reagir em excesso às necessidades atuais ou subestimar grosseiramente o requisito, resultando em falha. Quando as iniciativas de clientes/empresas são alinhadas a iniciativas de TI, a organização da rede pode se ajustar mais facilmente às implementações de novos aplicativos, novos serviços e outros requisitos comerciais. A relação e o foco em visão geral comum no atendimento de metas corporativas estão presentes e todos os grupos funcionam como uma equipe.

- Você deve se comprometer com o processo e contrato do SLA. Primeiro, deve haver compromisso de aprender o processo de SLA para desenvolver acordos eficazes. Segundo, você deve respeitar os requisitos de serviço do contrato. Não espere criar SLAs poderosos sem informações e comprometimento significativos de todos os indivíduos envolvidos. Esse compromisso também deve vir da gerência e de todos os indivíduos associados ao processo de SLA.

Passo 8: Determine as partes envolvidas no SLA

Os SLAs de rede de nível corporativo dependem muito de elementos de rede, elementos de administração do servidor, suporte de help desk, elementos de aplicativos e requisitos comerciais ou de usuário. Normalmente, o gerenciamento de cada área será envolvido no processo de SLA. Este cenário funciona bem quando a organização está construindo SLAs básicos de suporte reativos. Organizações corporativas com requisitos de maior disponibilidade podem precisar de assistência técnica durante o processo de SLA para ajudar com problemas como orçamento de disponibilidade, limitações de desempenho, criação de perfis de aplicativos ou recursos de gerenciamento pró-ativo. Para aspectos de SLA de gerenciamento mais proativos, recomendamos uma equipe técnica de arquitetos de rede e arquitetos de aplicativos. A assistência técnica pode aproximar consideravelmente os recursos de disponibilidade e desempenho da rede e o que é necessário para alcançar objetivos específicos.

Os SLAs de provedores de serviços normalmente não incluem a entrada do usuário porque são criados com o único objetivo de obter vantagem competitiva em outros provedores de serviços. Em alguns casos, o gerenciamento superior criará esses SLAs em níveis de alta disponibilidade ou alto desempenho para promover seus serviços e fornecer metas internas para funcionários internos. Outros provedores de serviço irão se concentrar nos aspectos técnicos da melhoria de disponibilidade, criando definições de nível de serviço forte que são medidas e gerenciadas internamente. Em outros casos, ambos os esforços ocorrem simultaneamente mas não necessariamente em conjunto ou com os mesmos objetivos.

A escolha das partes envolvidas no SLA deve então se basear nos objetivos do SLA. Alguns objetivos possíveis são:

- Atender aos objetivos comerciais de suporte reativo
- Fornecendo o mais alto nível de disponibilidade definindo SLAs proativos
- Promovendo ou vendendo um serviço

Etapa 9: Determinar elementos de serviço

Geralmente, os principais SLAs de serviço/suporte possuem vários componentes, incluindo o nível de suporte, como esse nível será medido, a progressão do caminho para a reconciliação do SLA e questões gerais de orçamento. Os elementos de serviço para ambientes de alta disponibilidade devem incluir definições de serviços proativos e objetivos reativos. Os detalhes adicionais incluem:

- Suporte no local durante o horário de trabalho e procedimentos para obter suporte fora desse horário
- Definições de prioridade, incluindo tipo de problema, tempo máximo para iniciar o trabalho no problema, tempo máximo para resolver o problema e procedimentos de escalonamento
- Produtos ou serviços a serem suportados, classificados por ordem de gravidade de negócio

- Suporte para expectativas de experiência, expectativas de nível de desempenho, relato de status e responsabilidades de usuário para a solução de problemas.
- Problemas e requisitos do nível de suporte da unidade de negócios ou geográfica
- Metodologia e procedimentos de gerenciamento do problema (sistema call-tracking)
- Objetivos do Help Desk
- Detecção de erros de rede e resposta ao serviço
- Medição e relatório da disponibilidade da rede
- Medição e geração de relatórios de capacidade e desempenho da rede
- Procedimentos de resolução de conflitos
- Financiamento do SLA implementado

Os SLAs de aplicativos ou serviços em rede podem ter necessidades adicionais com base nos requisitos do grupo de usuários e na importância comercial. A organização de rede deve escutar atentamente esses requisitos de negócios e desenvolver soluções especializadas que se encaixem na estrutura de suporte geral. Integrar-se na cultura de suporte geral é fundamental porque é importante não criar um serviço premier destinado apenas a alguns indivíduos ou grupos. Em muitos casos, esses requisitos adicionais podem ser colocados em categorias de "solução". Um exemplo pode ser uma solução de platina, ouro e prata com base na necessidade da empresa. Consulte os exemplos de requisitos de SLA a seguir para obter mais informações sobre necessidades de negócios específicos.

Observação: a estrutura de suporte, o caminho de encaminhamento, os procedimentos de help desk, a medição e as definições de prioridade devem permanecer os mesmos para manter e melhorar uma cultura de serviço consistente.

- Requisitos e capacidades de largura de banda para intermitência
- Requisitos de desempenho
- Definições e requisitos de QoS
- Requisitos de disponibilidade e redundância para criar uma matriz de solução
- Requisitos, metodologia e procedimentos de acompanhamento e comunicação de informações
- Critérios de atualização para elementos de aplicativo/serviço
- Financiamento de requisitos fora do orçamento ou metodologia de cobrança cruzada

Por exemplo, você pode criar categorias de solução para a conectividade do local da WAN. A solução de platina seria fornecida com serviços T1 gêmeos para a estação. Uma operadora diferente forneceria cada linha T1. A estação deveria ter dois roteadores configurados de forma que, se qualquer T1 ou roteador falhasse, ela não sofreria nenhuma interrupção. O serviço gold teria dois roteadores, mas o Frame Relay de backup seria usado. Essa solução pode ter largura de banda limitada pela duração da interrupção. A solução de prata teria apenas um roteador e um serviço de portadora. Qualquer uma dessas soluções seria considerada para níveis de prioridade diferentes para bilhetes com problemas. Algumas organizações podem exigir uma solução de platina ou ouro se uma prioridade 1 ou 2 for necessária para uma falha. As organizações do cliente podem então financiar o nível de serviço necessário. A tabela a seguir exibe um exemplo de organização que oferece três níveis de serviço, dependendo da necessidade de negócios para conectividade extranet.

Solução	Platinum	Gold	Prata
Dispositivos	Roteadores redundantes para	Roteador redundante para backup no site	Sem redundância de

	conectividade e de WAN	central	dispositivo
<i>WAN</i>	Conectividade e T1 redundante, várias operadoras	Conectividade T1 com backup de Frame Relay	Sem redundância de WAN
<i>Requisitos de largura de banda e surto</i>	T1 redundante com compartilhamento de carga para intermitência	Compartilhamento sem carga, backup de Frame Relay somente para aplicativos críticos; Frame Relay 64K somente CIR	UP para T1
<i>Desempenho</i>	Tempo de resposta de ida e volta de 100 ms consistente ou menos	Tempo de resposta de 100 ms ou menos. Esperado 99,9%	Tempo de resposta igual ou inferior a 100 ms esperado 99%
<i>Requisito de disponibilidade</i>	99.99%	99.95%	99.9%
<i>Prioridade do help desk quando inativo</i>	Prioridade 1: business-critical service down	Prioridade 2: serviço com impacto sobre os negócios inativo	Prioridade 3: conectividade empresarial desativada

[Etapa 10: Compreender os Objetivos e as Necessidades de Negócios dos Clientes](#)

Esse passo dá uma grande credibilidade ao programador de SLA. Ao entender as necessidades dos vários grupos de negócios, o documento SLA inicial estará muito mais próximo dos requisitos de negócios e do resultado desejado. Tente entender o custo do tempo de inatividade para o serviço do cliente. Estimativa em termos de perda de produtividade, receita e boa vontade do cliente. Lembre-se de que mesmo conexões simples com algumas pessoas podem afetar seriamente a receita. Nesse caso, certifique-se de ajudar o cliente a entender os riscos de disponibilidade e desempenho que podem ocorrer para que a organização entenda melhor o nível de serviço de que precisa. Se você perder esta etapa, pode ter muitos clientes que simplesmente exigem disponibilidade de 100%.

O desenvolvedor do SLA também deve compreender os objetivos de negócios e o crescimento da organização para acomodar as atualizações de rede, a carga de trabalho e o orçamento. Também é útil entender os aplicativos que serão usados. Espera-se que a organização tenha perfis de aplicação em cada aplicativo, mas, se não, considere fazer uma avaliação técnica do aplicativo para determinar problemas relacionados à rede.

[Etapa 11: Definir o SLA necessário para cada grupo](#)

Os SLAs de suporte principais devem incluir unidades de negócios críticas e representação de grupo funcional, como operações de rede, operações de servidor e grupos de suporte a aplicativos. Esses grupos devem ser reconhecidos com base nas necessidades comerciais, bem como sua parte no processo de suporte. A representação de muitos grupos também ajuda a criar uma solução de suporte global equitativa sem preferência ou prioridade de grupo individual. Isso pode levar uma organização de suporte a fornecer o serviço principal a grupos individuais, um cenário que talvez determine a cultura global de serviços da organização. Por exemplo, um cliente pode insistir que seu aplicativo é o mais crítico na corporação quando, na realidade, o custo do tempo de inatividade para esse aplicativo é significativamente menor do que outros em termos de perda de receita, perda de produtividade e perda de boa vontade do cliente.

Diferentes unidades de negócios dentro da organização terão requisitos diferentes. Uma meta do SLA de rede deve ser o acordo quanto a um formato geral que acomode diferentes níveis de serviço. Esses requisitos são geralmente disponibilidade, QoS, desempenho e MTTR. No SLA de rede, essas variáveis são tratadas priorizando aplicativos de negócios para ajustes potenciais de QoS, definindo prioridades de help desk para MTTR de diferentes problemas que afetam a rede e desenvolvendo uma matriz de soluções que ajudará a lidar com diferentes requisitos de disponibilidade e desempenho. Um exemplo de uma matriz de solução simples para uma empresa de manufatura corporativa pode parecer com a tabela a seguir. Você pode adicionar informações sobre disponibilidade, QoS e desempenho.

Unidade de negócios	Aplicativos	Custo do tempo de inatividade	Prioridade do problema quando desativado	Requisito de servidor/rede
Fabricação	ERP	Alto	1	Maior redundância
Suporte ao cliente	Atendimento ao cliente	Alto	1	Maior redundância
Engenharia	Servidor de arquivos, design ASIC	Médio	2	redundância de núcleo de LAN
Marketing	Servidor de arquivos	Médio	2	redundância de núcleo de LAN

[Etapa 12: Escolha o formato do SLA](#)

O formato do SLA pode variar de acordo com os desejos do grupo ou com os requisitos organizacionais. A seguir está um exemplo de destaque recomendado para o SLA de rede:

1. Finalidade do acordo
Participantes do acordo
Objetivos e objetivos do acordo
2. Serviços fornecidos e produtos suportados
Serviço de help desk e rastreamento de chamadas
Definições de seriedade de problema com base no impacto comercial para as definições MTTR
Prioridades de serviço essenciais para os negócios para definições de QoS
Categorias de soluções definidas com base nos requisitos de disponibilidade e

desempenho Requisitos de formação Requisitos de planejamento de capacidade Requisitos de escalonamento Relatórios Soluções de rede fornecidas Novas solicitações de solução Produtos ou aplicativos não suportados

3. Políticas de negócios Suporte durante o horário comercial Definições de suporte após a hora Cobertura de férias Números de telefone de contato Previsão de carga de trabalho Resolução de ressentimentos Critério de qualificação de serviço Responsabilidades de segurança de usuário e grupo
4. Procedimentos de gerenciamento de problemas Início de chamada (usuário e automatizado) Resposta de primeiro nível e taxa de reparo de chamadas Acompanhamento de chamadas e manutenção de registros Responsabilidades do chamador Diagnóstico de problemas e requisitos de fechamento de chamada Detecção de problemas de gerenciamento de rede e resposta a serviços Categorias ou definições da solução de problemas Tratamento de problemas crônicos Tratamento de chamadas com exceção/problema crítico
5. Objetivos de qualidade de serviço Definições de qualidade Definições de medição Metas de qualidade Tempo médio para iniciar a resolução de problemas por prioridade de problema Tempo médio para resolver o problema por prioridade de problema Tempo médio para substituir o hardware por prioridade de problema Disponibilidade e desempenho da rede Gerenciamento da capacidade Gerenciamento do crescimento Relatório de Qualidade
6. Equipe e orçamentos Modelos de pessoal Orçamento das operações
7. Manutenção do Contrato Agenda de revisão de conformidade Relatórios e revisão de desempenho Reconciliação de métricas de relatório Atualizações periódicas de SLA
8. Aprovações
9. Anexos e exposições Diagramas de fluxo de chamadas Matriz de escalação Matriz de soluções de rede Exemplos de relatório

[Passo 13: Desenvolver grupos de trabalho de SLA](#)

O próximo passo é identificar os participantes no grupo de trabalho do SLA, inclusive o líder do grupo. O grupo de trabalho pode incluir usuários ou gerentes de unidades de negócios ou grupos funcionais ou representantes de uma base geográfica. Esses indivíduos comunicam problemas de SLA para seus respectivos grupos de trabalho. Os gerentes e tomadores de decisões que podem concordar com os principais elementos do SLA devem participar. Esses indivíduos podem incluir ser indivíduos administrativos e técnicos capazes de ajudar a definir questões técnicas relativas ao SLA e tomar decisões no nível de TI (por exemplo, gerente de helpdesk, gerente de operações de servidor, gerentes de aplicativos e gerente de operações de rede).

O grupo de trabalho SLA de rede também deve ser formado por uma ampla representação de aplicativos e negócios de forma a chegar a um acordo acerca de uma SLA de rede que abranja vários aplicativos e serviços. O grupo de trabalho deve ter autoridade para classificar processos e serviços essenciais para os negócios para a rede, bem como requisitos de disponibilidade e desempenho para serviços individuais. Estas informações serão usadas para criar prioridades para diferentes tipos de problemas que causam impacto aos negócios, priorizar tráfego crítico aos negócios na rede e criar soluções futuras de rede padrão com base nos requisitos de negócios.

[Passo 14: Realizar reuniões do grupo de trabalho e rascunhar o SLA](#)

O grupo de trabalho deveria criar inicialmente um estatuto do grupo de trabalho. O estatuto deve expressar os objetivos, as iniciativas e os cronogramas para o SLA. Em seguida, o grupo deverá

elaborar planos de ação específicos e definir calendários e calendários para o desenvolvimento e a implementação do SLA. O grupo deve igualmente desenvolver o processo de apresentação de relatórios para medir o nível de apoio em função dos critérios de apoio. A etapa final é criar o projeto de acordo de SLA.

Inicialmente, o grupo de trabalho do SLA de rede deve se encontrar uma vez por semana para desenvolver o SLA. Depois que o SLA tiver sido criado e aprovado, o grupo poderá se reunir mensalmente ou até trimestralmente para atualizações de SLA.

[Etapa 15: Negociar o SLA](#)

Os últimos passos na criação de uma SLA são a negociação final e a desconexão. Esta etapa inclui:

- Revisão do rascunho
- Negociando o conteúdo
- Edição e revisão do documento
- Obtenção da aprovação final

Este ciclo de revisão do rascunho, negociação do conteúdo e criação de revisões pode levar múltiplos ciclos antes da versão final ser enviada ao gerenciamento para aprovação.

Do ponto de vista do gerente de rede, é importante negociar resultados alcançáveis que possam ser medidos. Tente apoiar os acordos de desempenho e disponibilidade com os de outras organizações relacionadas. Isso pode incluir definições de qualidade, definições de medição e metas de qualidade. Lembre-se de que o serviço adicionado equivale a despesas adicionais. Certifique-se de que os grupos de usuários entendam que níveis adicionais de serviço custarão mais e permita que eles tomem a decisão se for um requisito essencial para os negócios. Você pode executar facilmente uma análise de custo em diversos aspectos do SLA, como tempo de substituição de hardware.

[Passo 16: Medir e monitorar a conformidade do SLA](#)

Medir a conformidade com SLA e relatar resultados são aspectos importantes do processo de SLA que ajudam a garantir a consistência e os resultados a longo prazo. Geralmente, recomendamos que qualquer componente principal de um SLA seja mensurável e que uma metodologia de medição seja implementada antes da implementação do SLA. Em seguida, realize reuniões mensais entre os grupos de usuários e suporte para revisar as medidas, identificar as causas principais do problema e propor soluções para atender ou exceder o requisito de nível de serviço. Isso ajuda a tornar o processo de SLA semelhante a qualquer programa de melhoria de qualidade moderno.

A seção a seguir fornece detalhes adicionais sobre como o gerenciamento em uma organização pode avaliar seus SLAs e seu gerenciamento de nível de serviço geral.

[Indicadores de desempenho do gerenciamento de nível de serviço](#)

Os indicadores de desempenho do gerenciamento de Nível de Serviço fornecem um mecanismo para monitorar e melhorar os níveis de serviço como uma medida de sucesso. Isso permite que a organização reaja mais rapidamente aos problemas de serviço e compreenda mais facilmente

problemas que afetam o serviço ou o custo de tempo ocioso em seu ambiente. Não medir definições de nível de serviço também nega qualquer trabalho pró-ativo positivo feito porque a organização é forçada a uma postura reativa. Ninguém vai ligar para dizer que o serviço está funcionando bem, mas muitos usuários vão ligar dizendo que o serviço não atende aos seus requisitos.

Portanto, os indicadores de desempenho do gerenciamento de nível de serviço são um requisito importante para o gerenciamento de nível de serviço, pois fornecem meios de compreender totalmente os níveis de serviço já existentes e de realizar ajustes com base nas questões atuais. Essa é a base para o fornecimento de um suporte pró-ativo e realização de melhorias qualitativas. Quando a organização faz uma análise de causa de raiz sobre os problemas e faz melhorias na qualidade, essa pode ser a melhor metodologia para melhorar a disponibilidade, o desempenho e a qualidade dos serviços disponíveis.

Por exemplo, considere o seguinte cenário real. A empresa X estava recebendo inúmeras reclamações de usuários de que a rede estava frequentemente inoperante por longos períodos. Ao medir a disponibilidade, a empresa descobriu que o problema principal eram alguns locais de WAN. Uma investigação mais detalhada desses pontos revelou que a maioria dos problemas estava em alguns locais da WAN. A causa raiz foi encontrada e a organização resolveu o problema. A organização define em seguida as metas de nível de serviço para disponibilidade e acordos realizados com o grupo de usuários. Medidas futuras identificaram problemas rapidamente devido à não conformidade com o SLA. O grupo de rede foi então visto como tendo um maior profissionalismo, experiência e um ativo geral para a organização. O grupo foi efetivamente alterado de reativo para proativo em sua natureza e ajudou os resultados da empresa.

Infelizmente, a maior parte das organizações de rede atuais tem definições de nível de serviço limitadas e nenhum indicador de desempenho. Como resultado, eles gastam a maior parte do tempo reagindo a reclamações ou problemas do usuário, em vez de identificar proativamente a causa básica e criar um serviço de rede que atenda aos requisitos comerciais.

Use os seguintes indicadores de desempenho do SLA para determinar o êxito do processo de gerenciamento de nível de serviço:

- Definição de nível de serviço documentada ou SLA que inclui disponibilidade, desempenho, tempo de resposta de serviço reativo, metas de resolução de problemas e escalonamento de problemas
- Métricas do indicador de desempenho, incluindo disponibilidade, desempenho, tempo de resposta do serviço por prioridade, tempo para resolver por prioridade e outros parâmetros SLA mensuráveis
- Reuniões mensais de gerenciamento de nível de serviço de rede para revisar a conformidade de nível de serviço e implementar melhorias

[Contrato de nível de serviço documentado ou Definição de nível de serviço](#)

O primeiro indicador de desempenho é um mero documento que detalha o SLA ou a definição do nível de serviços. As metas principais da definição do nível de serviço devem ser disponibilidade e desempenho porque são os requisitos principais para o usuário.

Objetivos secundários são importantes, pois ajudam a definir como os níveis de disponibilidade ou desempenho serão obtidos. Por exemplo, se a organização tiver metas agressivas de disponibilidade e desempenho, será importante evitar que ocorram problemas e corrigir

problemas rapidamente quando ocorrerem. As metas secundárias ajudam a definir os processos necessários para alcançar os níveis desejados de disponibilidade e desempenho.

Metas secundárias reativas incluem:

- Tempo de resposta do serviço reagente por prioridade de chamada
- Objetivos de resolução de problemas ou MTTR
- Procedimentos de escalada de problemas.

As metas secundárias proativas incluem:

- Detecção de dispositivo inativo ou link inativo
- Detecção de erros na rede
- Detecção de problemas de capacidade ou desempenho.

A definição de nível de serviço para objetivos principais, disponibilidade e desempenho deve incluir:

Objetivo

- Como o objetivo será medido
- Partes responsáveis pela medição da disponibilidade e do desempenho
- Partes responsáveis por alvos de disponibilidade e desempenho
- Processos de não conformidade

Se possível, recomendamos que os responsáveis pela medição e os responsáveis pelos resultados sejam diferentes para evitar um conflito de interesses. De tempos em tempos, também pode ser necessário ajustar os números de disponibilidade devido a erros de adição/movimentação/alteração, erros não detectados ou problemas de medição de disponibilidade. A definição de nível de serviço também pode incluir um processo para modificar resultados de forma a ajudar a melhorar a precisão e impedir ajustes impróprios. Consulte a [próxima seção](#) para obter informações sobre metodologias para medir disponibilidade e desempenho.

A definição de nível de serviço para objetivos secundários reativos especifica como a organização responderá a problemas de rede ou de IT depois de identificados, incluindo:

- Definições de prioridade de problemas
- Tempo de resposta do serviço reagente por prioridade de chamada
- Objetivos de definição de problemas ou MTTR
- Procedimentos de escalada de problemas

Em geral, estes objetivos definem quem será responsável pelos problemas a qualquer momento e em que medida os responsáveis devem abandonar as suas tarefas atuais para trabalhar nos problemas definidos. Como outras definições de nível de serviço, o documento de nível de serviço deve detalhar como as metas serão medidas, as partes responsáveis pela medição e os processos de não conformidade.

A definição de serviço para objetivos secundários pró-ativos define como a organização fornece suporte pró-ativo, incluindo a identificação de rede inativa, enlace inativo ou condições de inatividade do dispositivo, condições de erro de rede e limites de capacidade de rede. Defina metas que promovam o gerenciamento pró-ativo porque o gerenciamento pró-ativo de qualidade ajuda a eliminar problemas e ajuda a corrigir problemas mais rapidamente. Normalmente, isso é feito definindo-se uma meta de quantos casos proativos serão criados e resolvidos sem notificação do usuário. Muitas organizações configuram um sinalizador no software de help desk

para identificar casos proativos versus casos reativos para essa finalidade. O documento no nível de serviço deve conter também informações sobre como o objetivo será medido, as partes responsável pela medição e os processos de não conformidade.

Métricas de indicador de desempenho

Recomendamos sempre que toda meta de nível de serviço definido seja mensurável, permitindo à organização: medir os níveis de serviço, identificar problemas de serviço de causa principal que estejam impedindo o objetivo principal de disponibilidade e desempenho e fazer melhorias direcionadas a alvos específicos. Em geral, as métricas são simplesmente ferramentas que permitem aos gerentes de rede gerenciar a consistência do nível de serviço e fazer melhorias de acordo com os requisitos de negócios.

Infelizmente, muitas organizações não coletam dados de disponibilidade, desempenho e outras medições. As organizações atribuem isso à incapacidade de fornecer precisão completa, custo, sobrecarga de rede e recursos disponíveis. Estes fatores podem ter um impacto na habilidade de medição de níveis de serviço, mas a organização deveria focar os objetivos gerais para administrar e melhorar os níveis de serviço. Muitas organizações foram capazes de criar métricas de baixo custo e baixo custo, que podem não fornecer precisão completa, mas que satisfazem esses objetivos principais.

Medir a disponibilidade e o desempenho é uma área frequentemente negligenciada nas métricas de nível de serviço. As organizações bem-sucedidas com essas métricas usam dois métodos bastante simples. Um método é enviar os pacotes de ping do Protocolo de mensagens de controle da Internet (ICMP) de um local central na rede para as extremidades. Você também pode obter desempenho usando esse método. As organizações bem-sucedidas com esse método também agrupam dispositivos semelhantes em "grupos disponíveis", como dispositivos de LAN ou escritórios de campo domésticos. Isso também é atraente porque as empresas geralmente têm diferentes metas de nível de serviço para diferentes áreas geográficas ou essenciais para os negócios da rede. Permite que o grupo de métrica calcule a média de todos os dispositivos do grupo de disponibilidade para obter um resultado razoável.

O outro método bem-sucedido de calcular a disponibilidade é usar bilhetes de problemas e uma medida chamada IUM (usuários afetados por minuto). Este método tabula o número de usuários afetados por uma interrupção e multiplica esse número pelo número de minutos da interrupção. Quando expresso como uma porcentagem dos minutos totais no período de tempo, isso pode ser facilmente convertido para disponibilidade. Em ambos os casos, também pode ser útil identificar e medir a causa raiz do tempo de inatividade para que a melhoria possa ser mais facilmente direcionada. As categorias de causa de raiz incluem problemas de hardware, software, link ou portador, potência ou ambiente, falhas de alteração e erro de usuário.

Os objetivos mensuráveis de suporte reativo incluem:

- Tempo de resposta do serviço reagente por prioridade de chamada
- Objetivos de definição de problemas ou MTTR
- Tempo de escalada de problema

Meça as metas de suporte reativas gerando relatórios de bancos de dados de help desk, incluindo os seguintes campos:

- O horário em que uma chamada foi inicialmente reportada (ou informada no banco de dados)
- A hora em que a chamada foi aceita por alguém trabalhando no problema
- O tempo em que o problema foi escalonado

- A hora em que o problema foi fechado

Essas métricas podem exigir influência de gerenciamento para a entrada consistente de problemas no banco de dados e para atualizá-los em tempo real. Em alguns casos, as organizações podem gerar automaticamente tíquetes de problemas para eventos de rede ou solicitações de e-mail. Isso ajuda a fornecer precisão para identificar a hora de início de um problema. Os relatórios gerados com base nesse tipo de métrica geralmente classificarão os problemas por prioridade, grupo de trabalho e separadamente para ajudar a determinar problemas em potencial.

A medição dos processos de suporte pró-ativo é mais difícil porque exige que você monitore o trabalho pró-ativo e calcule alguma medida de eficácia. Pouco trabalho foi feito nesta área. No entanto, é evidente que apenas uma pequena percentagem de pessoas irá realmente comunicar problemas de rede a um serviço de assistência e, quando comunicarem o problema, será claramente necessário tempo para explicar o problema ou isolar o problema como sendo relacionado com a rede. Nem todos os casos proativos terão um efeito imediato sobre a disponibilidade e o desempenho devido à falha de dispositivos redundantes ou links terão pouco impacto sobre os usuários finais.

As organizações que implementam definições ou acordos de nível de serviço proativos o fazem por força de requisitos de negócios e potenciais riscos de disponibilidade. A medição é feita em termos de quantidade ou porcentagem de casos proativos, em vez de casos reativos gerados pelos usuários. É uma boa ideia medir a quantidade de casos proativos em cada área também. Essas categorias devem incluir dispositivos inativos, links inativos, erros de rede e violações de capacidade. Também é possível trabalhar um pouco utilizando a modelagem de disponibilidade e os casos proativos para determinar o efeito na disponibilidade obtido com a implementação de definições de serviço proativo.

Revisão do gerenciamento de nível de serviço

Outra medida de sucesso no gerenciamento do nível de serviço é a análise do gerenciamento do nível de serviço. Isso deve ser feito independentemente de os SLAs estarem ou não ocorrendo. Execute a revisão de gerenciamento do nível de serviço em uma reunião mensal com os responsáveis por medir e fornecer os níveis de serviço definidos. Os grupos de usuários também podem estar presentes quando os SLAs estiverem envolvidos. O objetivo da reunião é revisar o desempenho das definições do nível de serviço medido e fazer aperfeiçoamentos.

Cada reunião deve ter uma agenda definida que inclua:

- Revisão dos níveis de serviço medidos para o período especificado
- Revisão das iniciativas de aprimoramento definidas para áreas individuais.
- Métricas de nível de serviço atuais
- Uma discussão sobre quais melhorias são necessárias com base no conjunto atual de métricas.

Com o tempo, a empresa também pode apresentar tendências de conformidade de nível de serviço para determinar a eficácia do grupo. Este processo não é diferente de um círculo de qualidade ou processo de melhoria de qualidade. A reunião ajuda a abordar problemas individuais e determinar soluções com base na causa raiz.

Resumo de gerenciamento de nível de serviço

Em resumo, o gerenciamento de nível de serviço permite que uma organização mude de um modelo de suporte reativo para um modelo de suporte pró-ativo, onde a disponibilidade da rede e os níveis de desempenho são determinados pelos requisitos da empresa, e não pelo último conjunto de problemas. O processo ajuda a criar um ambiente de melhoria contínua de nível de serviço e de maior competitividade de negócios. O gerenciamento de nível de serviço também é o componente de gerenciamento mais importante para o gerenciamento pró-ativo da rede. Por esse motivo, o gerenciamento de nível de serviço é altamente recomendado em qualquer fase de planejamento e projeto de rede e deve começar com qualquer arquitetura de rede recém-definida. Isto permite que a organização implemente soluções corretamente na primeira vez, com um mínimo de ociosidade ou retrabalho.

[Informações Relacionadas](#)

- [Suporte Técnico - Cisco Systems](#)