Apparent NETWORKS™

Technical White Paper

A Guide to AppCritical™ Core Technologies

Table of Contents

Overview of AppCritical

AppCritical is an enterprise-class, network performance management solution. It gives customers unique, real-time views of how network performance impacts application delivery and their users' quality of experience. AppCritical understands the complex dynamics that exist between applications and network resources in today's converged environments. Through continuous, real-time monitoring and expert system analytics, it enables fast and effective resolution of even the most difficult network problems.

Unlike traditional, device-centric monitoring tools, AppCritical takes a radically different approach. It focuses exclusively on 'network paths' – the entire string of devices that application traffic 'hops' along as it transits a network – whether across an office or around the world. By focusing on paths instead of individual devices, AppCritical assesses network infrastructure from the perspective of the applications that run on it.

AppCritical automatically learns about, monitors and measures certain key characteristics of every network path it encounters and the elements that comprise it – even those owned and controlled by service providers or other third parties. It understands a network path's potential and it can immediately see when, where and why performance is falling short. AppCritical gives customers complete, unbroken views of their networks, delivering visibility and insight into network performance and application delivery that is unmatched in the industry today.

The AppCritical solution delivers business value throughout the enterprise:

- For NOC technicians, AppCritical provides definitive, actionable network performance information in real time, enabling them to find and fix problems fast – or avoid problems entirely.
- For NOC managers, it delivers roll-up reports and trend analyses on network performance and application delivery that help them focus their initiatives and manage their operations more effectively.
- For senior technology leaders, it enables them to show executive management how well IT is supporting the organization's strategic objectives and business goals in reports with plain language and easy-to-understand graphics.
- For application owners and business line managers, AppCritical shows them how well IT is meeting their application delivery requirements, and by extension, customer satisfaction requirements.
- For managed service providers, it lets them create new premium offerings around quality of service not just availability.

In summary, AppCritical delivers the real-time network visibility and performance information that organizations need today. It gives IT professionals and their business-side colleagues a common, coherent and effective way to manage network performance and dynamic application delivery in their converged environments.

This document provides a concise summary of each of the major aspects of the AppCritical implementation including:

- Packet Sampling – high-precision, non-intrusive active probing
- Measurement – quantifying the network performance
- Analytics – detecting, locating, and identifying the source of degradations
- Monitoring – continuously evaluating network paths in real-time
- Application Performance – assessing the Readiness of the network path for a specific application

Packet Sampling

AppCritical's key advantages derive from the unique methodology it uses. Very short sequences of packets are transmitted using commonly available IP network mechanisms (such as ICMP Echo) to arbitrary IP end-hosts. Common used sequence lengths are 1, 10, and 50 packets. In some cases, as many as 400 packets may be sent at once. The packets are carefully controlled to maintain a high degree of precision in the departure and arrival timings. Different sizes of packets are sent. Since the sequences are quite short, the load on the network is kept very low – typically only a few hundred Kbps. Overall, per path loads tend to average 2 Kbps with short periods of deeper analysis averaging between 50 to 100 Kbps.

AppCritical sends sets of distinct sequences. They represent the range of different traffic conditions that a network path might experience due to application use. By probing the path repeatedly with the set of sequences, a statistically significant collection of responses for each type is collected. If the period of the sampling is relatively short compared to the rate at which the traffic conditions are changing, then the sampled response represents a snapshot of the conditions at the time of testing.

Most traffic conditions are known to change over time – sometimes as fast as minute-by-minute or hour-by-hour. For example, routes may change, capacities may be altered by interfaces resetting, or traffic levels may significantly rise or fall. This would be typical for LAN, WAN, and Internet paths. In some cases, such as mobile or wireless usage, the circumstances may be changing more rapidly, on the order of many seconds to minutes or faster.

AppCritical can build up a complete set of statistics very quickly – in many cases, under a minute. However, it is possible for the transmitted sequences to interfere with each other, distorting the results. To avoid this issue, AppCritical usually takes samples over a much longer timescale to ensure that the statistics are clean.

The sampling usually uses ICMP Echo – a Layer 3 protocol that is common to almost all IP devices. The vast majority of IP addresses respond to an Echo request with an Echo response. This provides for a widespread, predictable mechanism for soliciting responses from any IP-based network host. In some situations, ICMP has been disabled at the target host, or is being blocked or shaped by some mid-path network element. Under these circumstances, the end-host can be instrumented with software or hardware that is constructed to provide an equivalent response (e.g. UDP Echo/Responder).

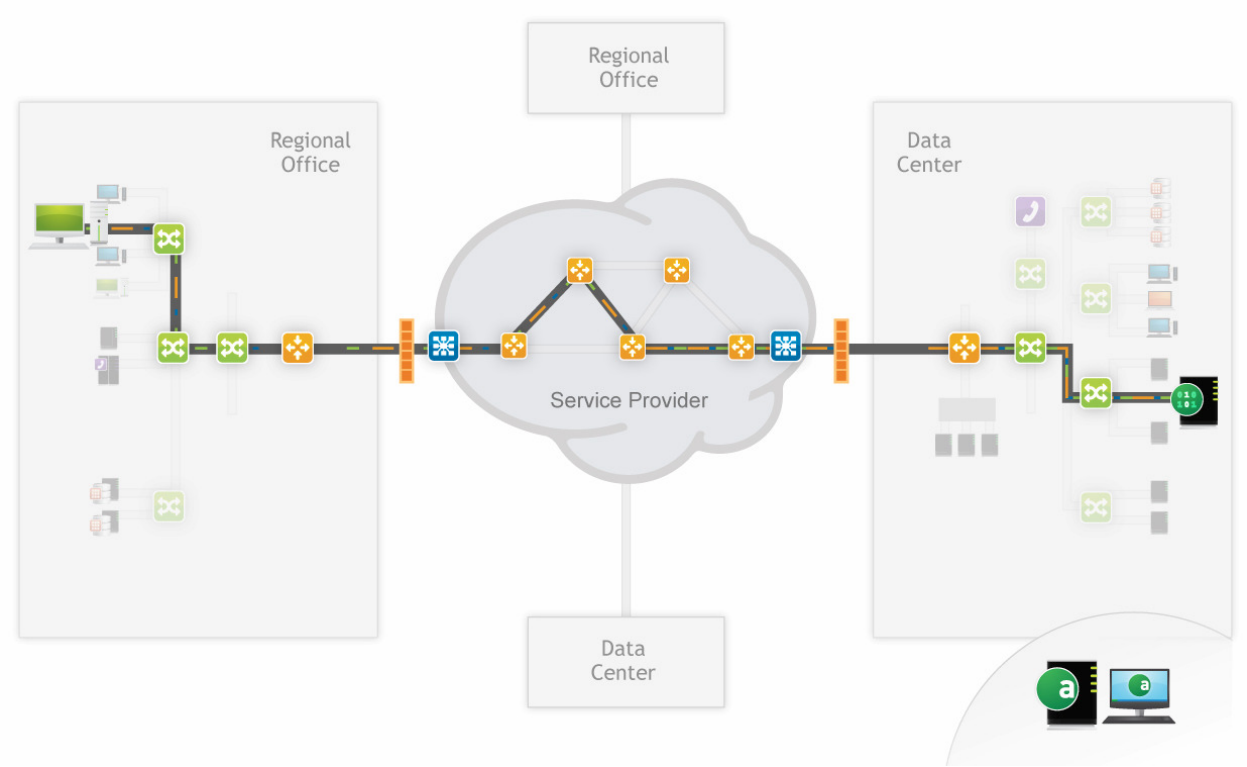


Figure 1 – Various sequences of packets are sent along the entire network path

NOTE: Almost any mechanism that provides a predictable response can be used. For this methodology, the payload of the packets and the protocol employed is not important. The critical requirement for AppCritical analytics is to extract packet timings from the end-to-end network - almost any packet will do.

There are a number of significant advantages to using an ICMP-based sampling methodology:

1. Single-sided – no need to deploy agents or appliances at target end points
2. Ubiquitous – broadly available on most IP devices; no special hardware needed
3. Non-intrusive – only very small sequences are sent; low load on the network
4. Public – no permissions or authorizations required
5. Protocol independent – testing possible regardless of presence of application
6. Real time – reports results within a very short period relative to network change

Unlike other forms of active testing, this kind of sampling can generate almost all possible network responses without having to wait for disruptive circumstances to arise.

Measurement

AppCritical actively probes the specified network path and generates one or more packet timing distributions for that path. A number of different groupings of packets are sent, ranging from single packets to small groupings to short streams. Various sizes and, in some cases, various protocols are used. This process typically takes a few minutes but can be shortened to less than a minute under some conditions. Packet sequences are sent at an average of 30 Kbps – this is considered "non-intrusive" and should not affect networks operating above 1 Mbps. Slower sampling rates are used on slower paths. AppCritical controls its own test rate to insure proper sampling.

From the distributions of packet sequence timings that AppCritical captures - including loss and various forms of network error – it extracts critical performance data through sophisticated analysis. The numbers produced exactly reflect the response of the end-to-end path. Based on precise network models, the accuracy of results depends only on the quality of the timing distributions generated. For example, under loss conditions, additional iterations may be needed to produce statistically accurate results.

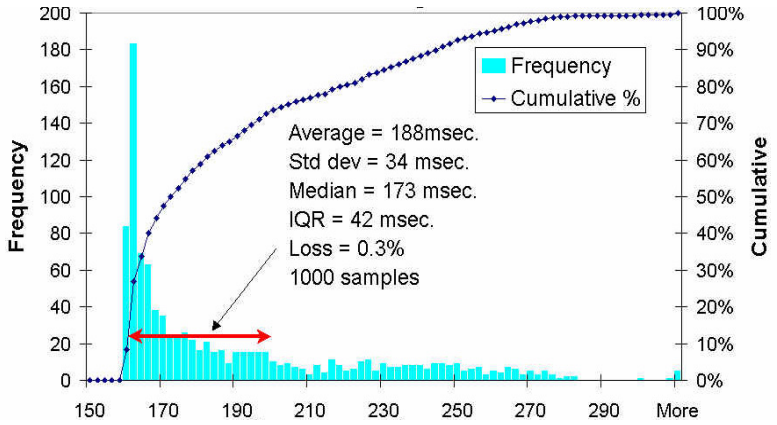


Figure 2 – Packet timing distributions are analyzed to extract various measures

The values extracted by AppCritical include maximum available bandwidth and utilization or currently available bandwidth, latency and jitter, loss and reordering, and variants of these quantities. One of the key measures is maximum bandwidth which is the upper limit on the data transfer capacity of the end-to-end network path. Like looking through a series of keyholes of varying dimension, this path's bandwidth is constrained by the smallest bandwidth on all the intervening links. This limiting value also constrains the performance of all applications using this path.

Among other techniques, AppCritical uses a form of analysis referred to as *packet train dispersion*. It notes how certain packet sequences are affected by the presence of a bottleneck. In particular, the bottleneck causes the distance between packets in a packet train to be increased. That separation exactly reflects the size of the bottleneck. In related but distinct ways, various other analyses conducted by AppCritical extract the other measures that appear at the Operations Console.

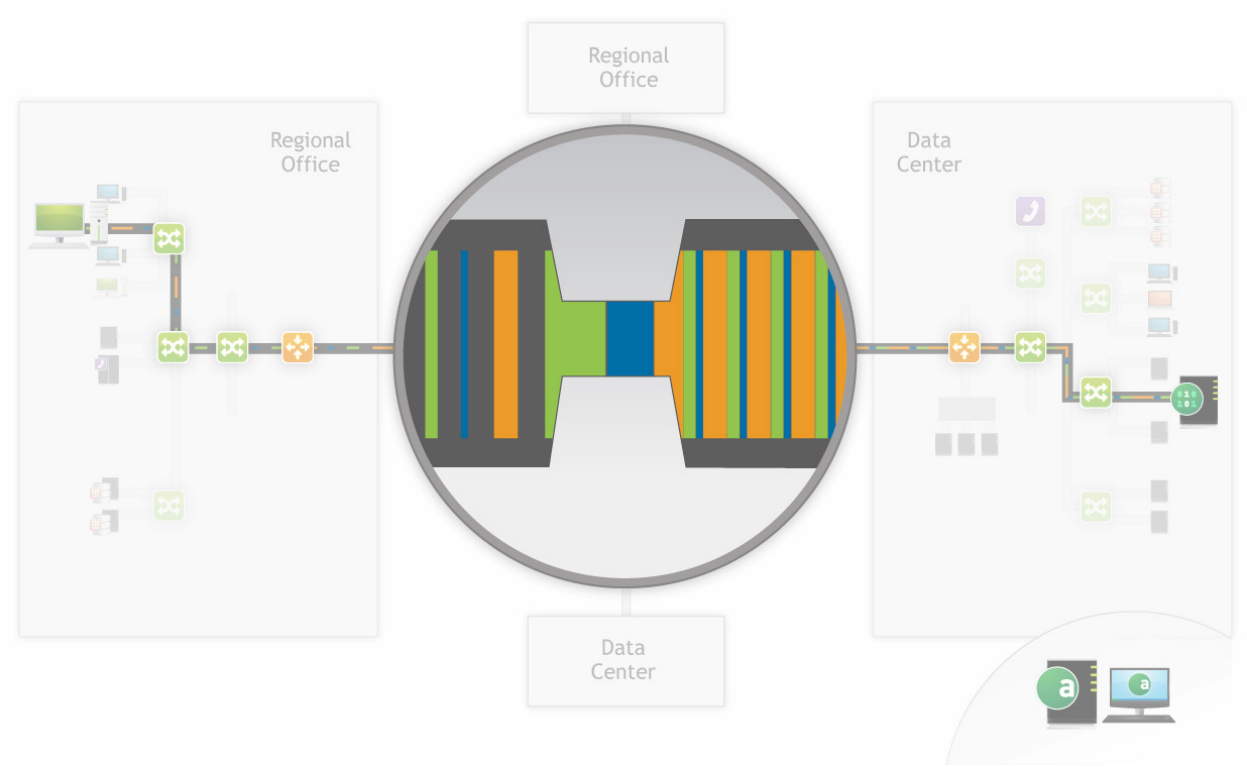


Figure 3 – Packet trains entering a bottleneck retain the effect of reduced bandwidth

Besides the values displayed in the AppCritical Operations Console, there are many more measurements made that are not published. They are utilized by other aspects of AppCritical, such as Analytics, Monitoring, and Application Performance analyses.

Analytics

AppCritical analyzes network paths in two ways - a "functional network model" and a "dysfunctional network model". "Functional" implies that the path is performing according to normal network design – in that case, the measurements made represent its capacities and usage.

Being "dysfunctional" implies behaviors that are outside design norms. The simplest example of this is packet loss. A perfect functional IP network should never lose packets. Once traffic levels have exceeded capacity, it is possible to have congestion loss – however that means that the network is then operating outside of design specification. Besides congestion, there are many other "dysfunctional" conditions that can cause loss or other behaviors that degrade performance.

When AppCritical detects degradation symptoms, it automatically performs diagnostic analyses against models of network dysfunction. These models isolate and identify characteristics that are specific to a particular source. Each type of degradation has a unique 'signature' that distinguishes it from any other. AppCritical's patented analytics perform a form of *pattern recognition* on the packet timings, loss, and other network errors to assess which known type of degradation may be present.

Information is extracted from the packet timings to construct a test signature that is unique to that path at the time of testing. The test signature is compared to all the known signatures to determine which one is the most likely match. AppCritical uses probabilistic analysis to indicate what problem the current behavior most closely matches. This can analogously be compared to face recognition – a clear photograph of a face can be uniquely compared to sample photographs to generate a match, even if not identical.

The only obstacle to precise diagnostics is the quality of the information being analyzed. If the photograph is blurred, taken at a great distance, or otherwise indistinct, making a solid match is difficult. Similarly, with AppCritical's path analysis, insufficient iterations or high traffic noise may hamper successful diagnosis. Further, new sources of dysfunction occasionally appear and are not recognized – or may be confused with a known cause incorrectly. Apparent Networks routinely works with customers to identify the cause of unresolved diagnostics and then subsequently adds this information to AppCritical's Expert system.

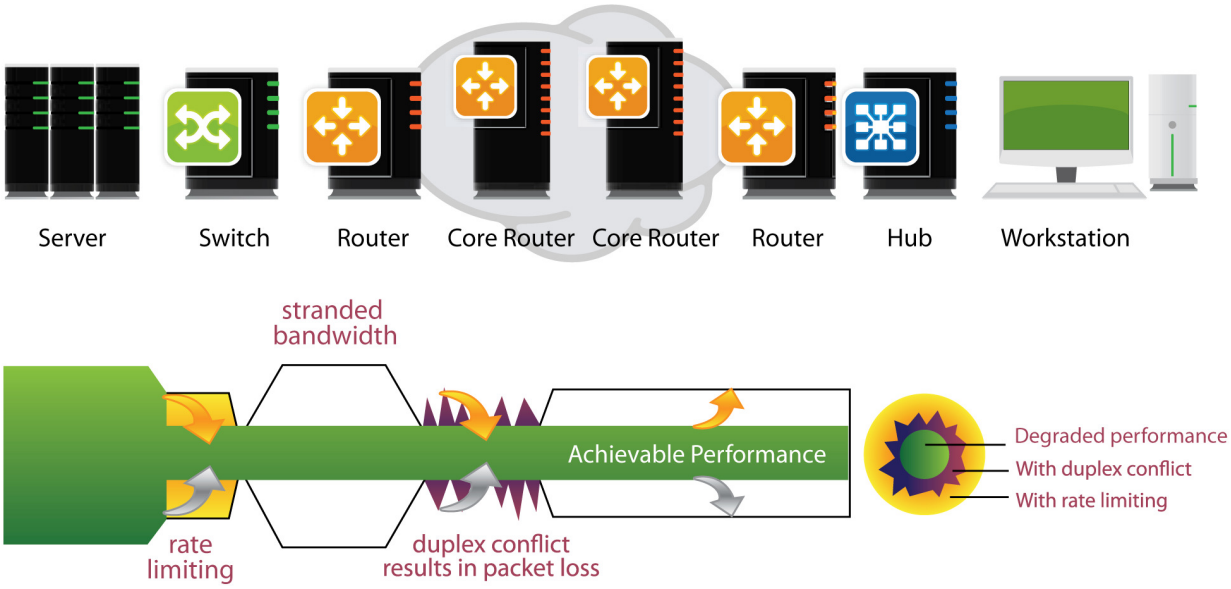


Figure 4 – A duplex conflict between a switch and a workstation results in dramatic packet loss

Thus, AppCritical can identify the common sources of significant degradation like duplex conflicts and distinguish them from others like congestion or media errors. AppCritical's Diagnostic system produces the various flags and statements that appear at the Operations Console. It also produces the certainty measures that reflect how closely a particular signature has been matched.

Since it is very difficult for end-users to interpret ambiguous or conflicting information, the AppCritical Operations Console presents a conservative analysis of the output of the analytics. Although the analytics always evaluate dysfunctional behaviors, the Operations Console will default away from showing unclear or misleading matches and instead, recommend steps to improve the testing.

Active Monitoring

AppCritical's Active Monitoring takes advantage of the variability of 'statistical resolution' to provide an automated mechanism for continuously monitoring thousands of network paths simultaneously. This approach is distinct from RMON/SNMP techniques which only monitor the state of individual elements. AppCritical uses low resolution- tens of packets per minute - to generate very low-quality estimates of various measurements. These imprecise but suggestive values are referred to as "critical indicators" and are pre-cursors to the measurements produced by AppCritical's full analysis.

When critical indicators vary, AppCritical responds by increasing statistical resolution to improve their accuracy and confirm the variation as an undesirable change in network conditions. If the improved estimates also indicate an undesirable trend, AppCritical further increases resolution until the minimum needed for an accurate measurement is achieved. Alternatively, the resolution drops back down to a normal, low resolution monitoring level if found unnecessary. AppCritical's ability to vary resolution means that the monitoring system can operate without human intervention, scaling from very light touch probing for most paths to comprehensive measurement and diagnostics where and when it is needed.

Figure 5 – Greater detail becomes apparent at each level of resolution, supporting recognition

It also means that AppCritical can scale easily to monitor thousands of paths. AppCritical can spread its attention very widely, focusing down as needed on the few paths that indicate deviation from performance norms. Once degradation has been confirmed by full resolution Measurement and Diagnostics, AppCritical generates alarms to inform the appropriate individual so that timely remediation efforts can begin. Any operator responding to an alarm is presented with full detail including the precise measurements and a completed diagnosis of the degradation.

AppCritical is an effective solution for proactive performance management. It delivers significant improvements over traditional SLA monitoring. Since AppCritical's methodology works across third party networks and segments a network path to show the boundaries, it provides a vastly more thorough and

accurate view of a network provider's quality of service. Active monitoring generates a continuous representation of a range of network behaviors over time such as bandwidth, loss, jitter, and latency.

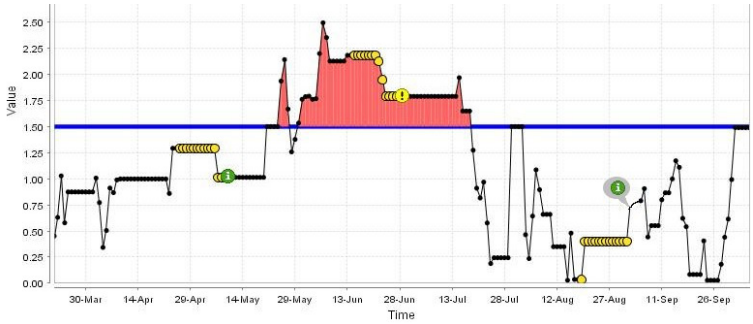


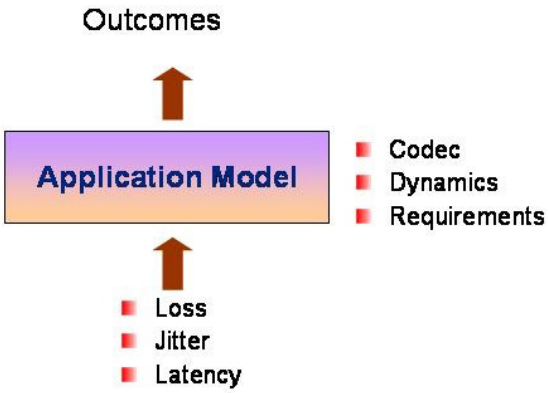
Figure 6 – Monitoring various critical indicators provides for automation and scalability

AppCritical is delivered preconfigured with comprehensive network model defaults that define 'normal' behavior. AppCritical also provides user-customizable criteria using Service Quality Definitions or SQDs. Since there are many different critical indicators, the end-user can choose which ones are important on a path-by-path basis, and what thresholds are significant. SQDs can be constructed to fit particular needs, network implementations, or application requirements.

Application Performance

Measurement, Analytics, and Active Monitoring operate primarily at Layer 3 of the IP network. While the primary goal of AppCritical is to provide visibility into network performance, the effectiveness of the network is defined by the applications (and users) that employ it. Poor network Quality of Service (QoS) only matters when it impacts Quality of Application (QoA) and subsequently the end-user experience or Quality of Experience (QoE). Low QoS does not necessarily mean low QoE. It is possible for an application to compensate for network conditions so that the user is not affected. For example, a VoIP phone uses a jitter buffer to minimize the effects of packet delivery variability up to the limit of the buffer.

AppCritical translates the outcomes of Measurement and Analytics into terms specific to different application types. The potential application categories include data-intensive (Data), real-time including VoIP (Voice) and Video, transactional, and best-effort.



AppCritical implements *application models* that capture the essential aspects of each application type. The models accept Layer 3 network conditions as inputs along with any specifics provided by the endpoint device (e.g. codec). The effects of the network conditions are translated into terms that correspond to the end-user experience of the application. For example, high jitter translates to packet discards in a VoIP device, leading to loss of data and subsequent impact on voice quality.

Using these models, AppCritical generates three kinds of results: (1) estimates of application performance in terms of standard metrics like Mean Opinion Score (MOS) for VoIP, Network Capacity for Data, and Video Quality Metric (VQM) for Video; (2) an assessment of the readiness of the network path; and (3) application-specific measures. Each of these outcomes reflects different aspects of AppCritical's analysis.

Application Performance

The IT industry is rapidly adopting metrics that simplify and standardize the assessment of application performance. VoIP led the way with the widespread introduction of MOS (ITU standard P.800). This standard has been in use for over 20 years in the telecommunications industry and has been adapted for digital IP telephony. As other critical applications have been introduced, the need for and benefit of metrics has become apparent. Although not yet standardized, similar metrics to MOS are appearing for Data, Video, and Transactional applications.

Network Readiness

One of AppCritical's key objectives is to answer questions like "Is my network ready to support that new application?", or "If I add a new office or facility, does my current infrastructure have the capacity to support it?" For example, the measurement of a metric like MOS for VoIP provides some indication of the answer. However, measuring only MOS falls short. For example, some problems such as duplex conflicts only appear on a transient basis and MOS may be good at one moment and bad at another. And MOS alone does not lead to remediation. Regardless of the current MOS, the Readiness of the network is evaluated in light of the Measurement and Analytics outcomes. AppCritical's Readiness value is an aggregate assessment of all the factors for that path and definitively answers the aforementioned questions.

Application-specific Measures:

AppCritical sends many different types of packet sequences. In certain cases, some of them experience high loss and others none at all. Consider the self-evident case of small packets as an example. If a significant percentage of small packets tend to be lost, but not large packets, it is apparent that a small-packet-based application like VoIP will be strongly affected, while a bandwidth-intensive backup application using large packets will not be. Measures such as loss that appear at the Ops Console reflect the typical values that might be measured by the applications under test – different measures as they are experienced by different applications.

Conclusion

The five main technological aspects of AppCritical – packet sampling, measurement, analytics, monitoring and application performance – combine to give customers unique and powerful new ways to monitor and manage their networks and the critical business applications that run on them. When performance on the network is not what it should be, AppCritical provides operators with faster, more effective and less expensive ways to locate, identify and remediate networks problems. AppCritical is *the* network performance and application delivery management solution for today's converged environments.

❖ ❖ ❖

Contact Us

For more information about the AppCritical solution, please visit www.ApparentNetworks.com, or contact one of our sales professionals at +1 (800) 508- 5233

© 2008 Apparent Networks, Inc. All rights reserved.