Urban Drainage and Flood Control District

ALERT System Loading and Data Loss Analysis

Prepared By

Don Van Wie
Telos Services

Edited By

Ilse Gayl
OneRain, Inc.

July 31, 2007

Final Draft

Executive Summary

Past performance of the UDFCD system has been adequate, although contention losses exceeding 30% have been experienced for brief periods during intense storms. Because of the inherent ability to infer rainfall from sequential accumulator values, the estimates of rainfall have exceeded 95% in all but a few situations. However, none of the storms studied was particularly severe; most likely, none had a recurrence interval greater than 2 years.

Data loss rates were determined for 10-minute periods during six events; the load/loss values plot along an exponential line that is in excellent agreement with the predictions of a Poisson function. The agreement through traffic loads of 2,500 messages per hour lends confidence that losses under higher loadings can be accurately modeled.

The loss model raises several points of concern regarding the ALERT network. First, the maximum message throughput cannot exceed 2,000 messages per hour, which will occur at a data loss rate of 63%. At input loadings greater than 5,400 messages per hour, throughput actually decreases and the system progressively collapses.

The "model storm" used to evaluate system performance under critical conditions was created directly from the time series radar record of a storm which occurred over south Denver and parts of Douglas and Arapahoe Counties on July 8, 2001. Using GIS tools, the storm was positioned over the Denver-Aurora area, and the 5-minute radar reflectivity values were translated into simulated ALERT rain reports.

The model storm resulted in a peak traffic load exceeding 7,400 reports per hour, with data losses that exceeded 80% through the Smokey Hill repeater. The rain estimation shortfall was in the range of 40-45%, and the implication of the loss rates is that the timeliness of weather data and stage alarms would be seriously affected. It is evident that a significant storm can stress the system past the point of maximum throughput and into the deteriorating phase of "contention collapse." Significant data losses occur on both the input and output frequencies.

Securing the desired performance of the UDFCD ALERT system will require changes that increase the data transmission capacity. The most cost-effective modification that can be implemented in the short term is to add another radio frequency to both the input and output sides of the repeater network, then repeat two channels from each existing repeater site. In addition to one new frequency, this change requires four new repeaters. Model storm peak data losses would be reduced from 75% to 55%, a substantial improvement but insufficient to bring the system to the desired level of performance.

Channel capacity could be increased much further if features of the new ALERT 2 protocol could be implemented. The new protocol will offer an increase in data efficiency of about 18-fold; some major benefits from ALERT 2 are a near-term possibility. Using message-concentrating ALERT 2 repeaters along with the second data channel, data losses during a model storm event would be comparable to the performance seen in historical events.

As demands and opportunities for the system evolve, the tools developed in this study offer some shorthand methods to evaluate loading impacts and the cost effectiveness of remediation alternatives.

Introduction

The Urban Drainage and Flood Control District (UDFCD) operates a network of ALERT gages that monitor rainfall, water level and related meteorological parameters in real time. The system has grown from an initial 9 sensors in 1978 to more than 325 at present. The ALERT protocol uses short, self-initiated transmissions from sites to convey data through one or more radio repeaters to base stations located in the Denver metro area.

Transmissions initiated at random have the potential to interfere with each other if they overlap in time, a problem known as "data contention." Contention increases with traffic loading and, at any given traffic loading, it increases with message duration. Data losses due to contention are a fact of life in ALERT systems, and to the extent possible, systems are designed to provide useful information in spite of high losses and uncorrected errors.

As the UDFCD system has grown, data losses due to contention have become increasingly evident, yet further growth of the system is expected. There have been no serious storms over the most densely gaged areas in the last decade, and so the system's ability to perform adequately under stress is not well understood.

The purpose of this study is to understand past data losses in the UDFCD ALERT network and to predict future performance under likely storm conditions. Our primary objectives are to understand the factors which affect contention losses, assess whether system design changes are needed, and have tools to evaluate the efficacy of alternative approaches.

This study has been conducted in four parts:

1. ALERT data from past storms were used to determine data losses over a range of conditions the system has experienced to date. The objective was to reconstitute missing data in order to estimate past losses, and in particular to establish an accurate relationship between input traffic loading and data losses.
2. The composite information from historical events was used to validate and calibrate a mathematical model of contention losses. The objective was to develop methods that permit us to confidently predict system performance under conditions that have not yet occurred, but under which the system is expected to fulfill its purpose.
3. With the assistance of the UDFCD, a model storm was developed that could be applied over a selected area of the network. The storm scenario is a "likely bad case," but far from the worst that could be encountered. Algorithms were developed based on historical performance to simulate the resulting traffic from rain gages, stage gages and weather sensors. The contention loss model was then applied to the model storm data to predict data losses and assess likely performance.
4. The tools developed to project contention losses were used to evaluate the effectiveness of several possible approaches to reducing them. The objective

was to determine what sorts of changes are the most practical and cost-effective, and to suggest general directions for the mitigation effort.

Analysis of past storms

Archived records for the period between 2000 and 2006 were scanned for days of peak activity, and several storms with the highest peak loading per 10-minute period were identified. Table 1 summarizes the events that were used in this analysis:

Date	~Rainfall Duration	Reports	Rain	Stage	Peak Thru	Peak Input	Max Rain	Gages > 2"
8/17/2000	6 hr	6866	2502	3966	1332	1586	4.04	5
7/23/2004	6 hr	5095	1900	2558	1182	1506	2.6	2
8/18/2004	4 hr	8790	4356	2964	1458	2058	4.8	33
6/3/2005	3 hr	2980	607	1682	1740	2955	2.32	1
8/4/2005	6 hr	13571	3798	6582	1194	1541	2.48	15
8/13/2006	2 hr	3797	1275	1616	1164	1675	1.47	0

Table 1. Rainfall events for peak loading study.

Note that peak input rates do not correlate with storm size. Short storms of modest rainfall can produce high loading depending on where they occur and the intensity of rainfall. Brief durations of intense rain, however, have minimal impact on overall system performance. Figure 1 shows a graph of the loading for one of the studied storms, below.

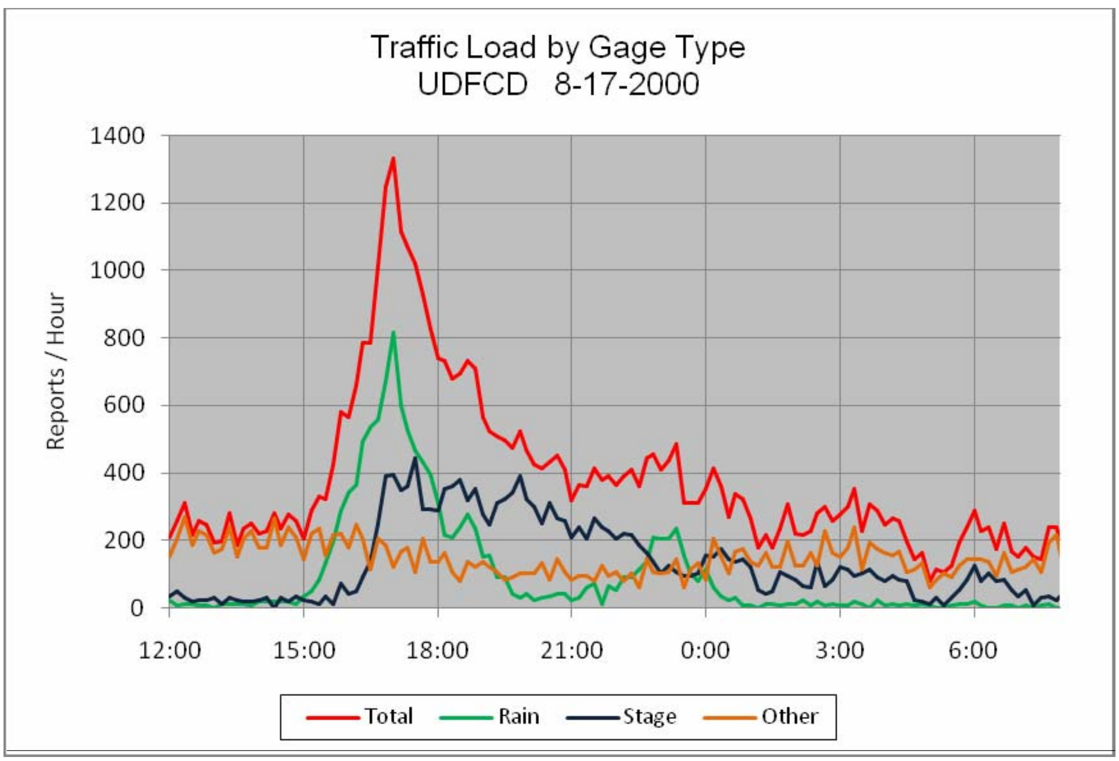

Figure 1. UDFCD traffic by gage type during 17 August 2000 rainfall event.

Background traffic ("Other" in Figure 1, above) accounts for 200 to 250 messages per hour. Rain activity sharply increases the input loading, followed within a few minutes by a rise in stage gage traffic. In the smaller basins where the rain gages are most densely distributed, the stage traffic makes a substantial contribution to the peak storm loading. In all cases studied, the peak system loading occurs at the time of peak rain activity. Note that the background traffic throughput decreases during the event, which is due at least in part to contention losses.

Insertion of missing rain reports

The rain data from the DIADvisor record at the OneRain offices were used for these analyses. The rain record from the active storm period was sorted in chronological order by gage, and extracted from Access to an Excel spreadsheet. Gaps in the accumulator sequence were deemed to be missing reports, and were filled by inserting one or more reports distributed at equal intervals between existing reports. This process was automated with an Excel macro. The reconstituted data set was returned to Access as a separate table. These files are available as support information to this report.

This method works well during periods of significant rain activity, when there are two or more tips per 10-minute analysis interval (a rain rate of ½ inch per hour). Below this rate, the assignment of a report to an interval becomes increasingly uncertain.

In addition, reports missed at the beginning or end of the event can't be detected, and no data contention can be identified at all if rain is not occurring. The method is well suited to analyzing periods of high activity, but will underestimate data loss rates during rain-quiet periods. Therefore, input loadings of less than 750 reports per hour were not used in determining the input/data loss relationship.

Effect of gage holdoff

Gages are limited to one report each 15 seconds on the majority of equipment used in the study area (a few units have a 20-second holdoff). The limiting report rate is therefore 4 per minute, or 40 per 10-minute analysis period. This is a rate in excess of 9 inches per hour. Although we know that gage holdoffs do occur during brief intense periods, they are uncommon and their impact on report rate is minimal.

The maximum number of modeled reports at any rain gage is 9 per five-minute interval or 4.25 inches/hour. Although it is likely that instantaneous rain rates would exceed 9 inches per hour for brief periods, again the effect on total reporting would be small and was ignored for the study.

Estimation of input load

The percentage of missing rain reports was determined for each 10-minute period as the ratio of missing rain reports to the sum of received and missing rain reports. Contention loss rates are independent of sensor type, so the total input load was estimated by dividing the total throughput for each period by the rate of loss for the rain gages. Only periods with a total input report rate exceeding 750 reports per hour were used. The analysis of 6 storms resulted in 127 10-minute periods with value pairs for input load and loss rate, which were combined into a single data set and are presented in Figure 2, below.

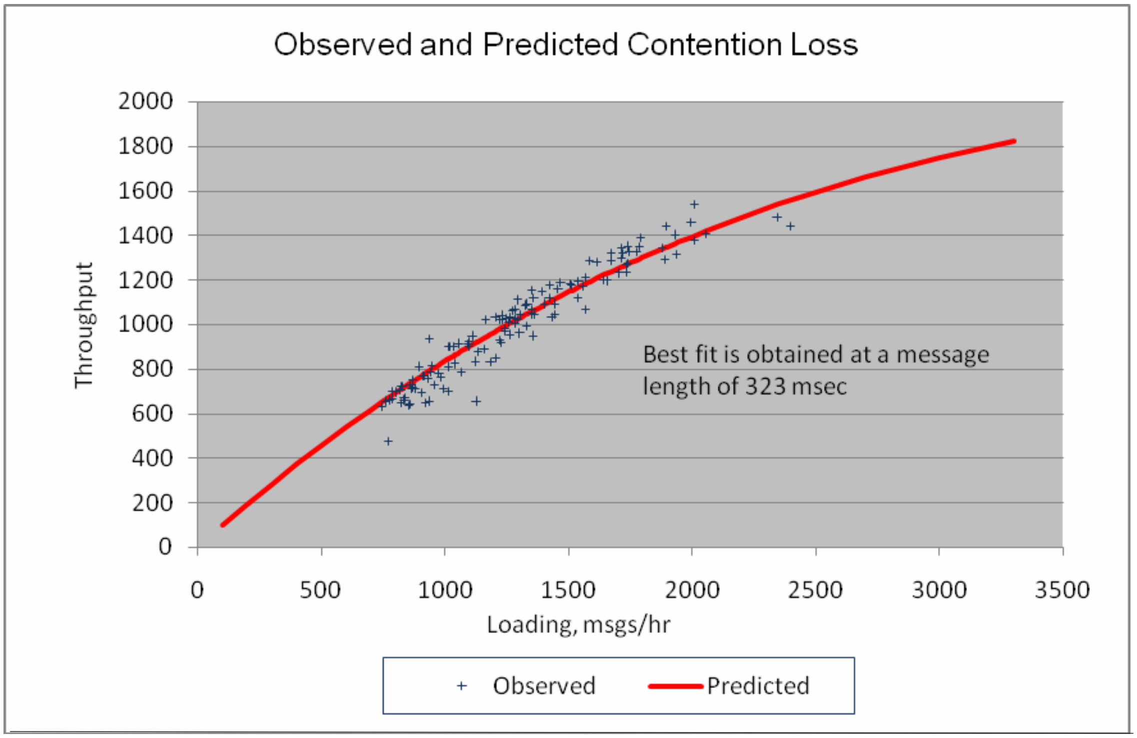

Figure 2. Observed and predicted throughput as a function of traffic load.

Prediction of losses due to contention

We would like to be able to accurately predict the data losses that will occur due to message contention as traffic increases beyond levels we have seen historically. A theoretical analysis of the probability of collision is relatively straightforward; the problem was studied extensively in the telecommunications domain in relation to packet collision in Ethernet environments. These studies assume that packets of equal length are sent independently from multiple sources at random intervals.

ALERT deviates from these criteria in that some transmissions may consist of several concatenated messages, and others are produced on timed intervals. However, these exceptions are generated by the background traffic, while storm traffic consists almost entirely of single reports. The inaccuracies introduced by the exceptions are not large enough to be detectable with the data available and make virtually no contribution to the storm-generated traffic that challenges system capacity.

Thus the probability of collision in the UDFCD ALERT system can be predicted using the Poisson equation. This equation yields the probability (P) that any given number of events (N) will occur in a unit period of time, given the average frequency of events, λ.

$$P = (e^{-\lambda} * \lambda^{N})/N!$$

In the ALERT contention case, we are interested in knowing the probability that no messages will initiate during the time it takes for a message to insert into the stream (N=0). This simplifies the equation:

$$P = e^{-\lambda}$$

The minimum time required to successfully insert a message is two message lengths, since after initiation, a period of one message length follows before quiet returns, and another message length of quiet must occur to permit unobstructed insertion. The ALERT message consists of 40 bits transmitted at 300 bits per second, which occupies 133 milliseconds. The data segment is preceded by a tone (preamble) that is typically 200 milliseconds, so that the total message length is approximately 333 milliseconds. The required time between event initiations to avoid contention is therefore 666 milliseconds.

There are 5,400 insertion periods per hour. The variable λ is the ratio of the input load (L) to the insertion capacity:

$$\lambda = L/5400$$

A characteristic of the ALERT contention process is that there is a point of diminishing return as the input traffic load increases. This point occurs when $\lambda = 1$, or 5,400 messages per hour; beyond this input loading, the number of messages successfully received per unit time declines as loading increases. This condition is called contention collapse.

The expected data losses due to collision are presented in relation to input loads in the UDFCD system below in Table 2:

Traffic loading and probability of collision		
Messages /hour	Probability of collision	UDFCD context
250	4.4 %	Base loading
1000	16.9 %	Typical summer event
2000	31.0 %	Seen 1-2 times annually
3500	47.7 %	Peak observed in study
7500	75.1%	Model storm
12000	89.2%	Possible in UDFCD

Table 2. Probability of message collisions as a function of traffic loading.

The data loss predicted by the Poisson function is co-plotted with the observed values on the graph above. The curve was fitted to the empirical data by optimizing the denominator in the Lambda relationship; the best fit was achieved for a message length of 323 milliseconds. Since the data portion of the message is 133 msec, this implies a mean preamble length of 190 milliseconds, which in fact, is the preamble setting on the majority of the transmitters used in the study area.

Preamble lengths vary widely across ALERT hardware, with 200 msec often used as a working average. All modeling in this study was based on the assumption of a 333 msec message (200 msec preamble). If all UDFCD gages have a 190 msec preamble, the estimates of loss would be high by about 1% at the peak of the model storm.

Effects of repeaters on contention losses

Repeaters can have major effects on contention losses. The loss model shown here is based on a single channel environment with no repeater, while the measured losses occurred in a two-channel, multiple repeater network. The effects of various repeater options and configurations were examined in order to provide better insight into their impacts on contention.

Any single frequency store-and-forward repeater greatly reduces channel capacity, since the required insertion period is more than doubled. When the repeater receives a message, the message is decoded to its digital form and stored. The receiver is then turned off and the transmitter is turned on; the processing and hardware switching occupy a few tens of milliseconds of "turnaround time." Finally, the digital message is re-encoded as an audio signal and retransmitted on the same frequency. Contention collapse begins at about 2,500 messages per hour, and losses at the peak of analyzed storm would have exceeded 70%.

Dual-frequency repeaters can avoid this problem by transmitting on the output channel while continuing to monitor the input frequency. The current generation of repeaters from both major ALERT manufacturers supports this capability, as do the repeaters in the UDFCD system. For this type of repeater, the insertion time is the same as for a no-repeater system. It is important that this feature be available on any dual-frequency repeaters used in the UDFCD system.

Prior generations of store-and-forward repeater – the majority of those in service today – do single-threaded processing. These repeaters cannot monitor the input frequency while transmitting, so that even though the message is not repeated back on the same channel, any new message initiating during the repeat process is lost.

Repeater transmit delays are used in some systems; for example, to avoid output side contention between two repeaters which may hear the same input site. A simplex or single-threaded repeater is unable to monitor the input frequency during the transmit delay and, since these may be more than a second long, the effect on channel capacity is enormously negative.

Three single-frequency repeaters are used on the input side of the UDFCD system. Their effect is to increase contention on the input side of the system, but their effect is limited since they pass only a limited number of gages each. None of the storms studied heavily involved any of these repeaters' areas, although a storm centered over the Evergreen area could produce loadings on the Chokecherry repeater that significantly increase losses in the critical area.

Estimating losses across input and output frequencies

There is clear evidence that losses occur on both the input and output frequencies. For example, the Smokey Hill rain gage has a direct input to the repeater, so there are no losses on the input side. However, report losses from sensor ID 740 increase with loading, though at a lower rate than any other gage; they are clearly occurring on the output channel. No contention can occur in the output stream of any single repeater, but Poisson function losses do occur among the streams from multiple repeaters.

This study estimates system-wide losses and assumes that a model for a single-frequency, no-repeater system applies to the UDFCD system. The validity of this approach was evaluated by modeling the contention probabilities in the UDFCD system under equal and widely disparate traffic distributions among repeaters.

Depending on the loading and the distribution of the load across repeaters, differences between simplex and the UDFCD dual-frequency system vary by 2-3%, an amount too small to be considered significant given the other data limitations.

Contention Losses in a Dual Frequency System, Varied Load						
Repeater	Input	In Loss	Thru	Out Loss	Thru	Net Loss
Smokey	5000	60.38%	1981	50%	990	**80.2%**
Blue	2000	30.95%	1381	50%	690	**65.5%**
Gold	200	3.64%	193	50%	96	**51.9%**
Lee	200	3.64%	193	50%	96	**51.9%**
	7400	49%	**3747**	50%	**1872**	**74.7%**
All Simplex	7400	75%			**1880**	**75%**

Contention Losses in a Dual Frequency System, Equal Load						
Repeater	Input	In Loss	Thru	Out Loss	Thru	Net Loss
Smokey	1850	29.01%	1313	62%	496	**73.2%**
Blue	1850	29.01%	1313	62%	496	**73.2%**
Gold	1850	29.01%	1313	62%	496	**73.2%**
Lee	1850	29.01%	1313	62%	496	**73.2%**
	7400	29%	**5253**	62%	**1986**	**73.2%**
All Simplex	7400	75%			**1880**	**75%**

Table 3. Contention losses in dual-frequency system with varied (top) and equal (bottom) traffic loads by repeater.

Data loss variation among rain gages

Losses were analyzed by gage for 4 storms and only gages estimated to have generated more than 100 reports were included. The amount of variability in data loss among gages by storm decreases as the number of reports in the storm increases, which is to be expected. The variability also decreases as the number of storms included in the composite increases, but persistent, non-random differences in gage performance emerge. For example, gages 400 and 850 emerge as consistently poorer performers. In the composite analysis of losses, the mean loss was 25.8%, with a range from 12% to 41%. The losses fit a normal distribution with a standard deviation of 6.9%. The maximum standard deviation was 2.2, with 95% of all values falling within 2 standard deviations or less.

This performance difference is not expected in a random data contention model and indicates other processes are involved; it probably relates largely to received signal strength at the repeater. This is not demonstrated by path models, but the RF model lacks the resolution to include the effects of micro-topographic conditions that are important. It may also be caused by path degradation from intense rainfall along the radio path; path length and storm position will cause variation among gages. There may also be storm-induced hardware effects, such as nearby lightning strikes that cause brief disruptions in processor operation. Although not analyzed statistically, inspection of the data showed sequential gaps in data from gages that were extremely improbable by random contention alone.

Whatever the cause, the higher losses at some gages greatly increases the probability that these gages will suffer longer sequences of missing reports. When gaps in accumulator value exceed the established limits, the data collection validation procedures reject the reports as erroneous and rain is underestimated. Further examination of this would help us understand and better estimate the deterioration of rainfall estimates during intense activity. The software rain validation process overcomes much of the data loss, but the correction is not as complete as would be expected if the data contention losses were purely random.

The model storm and modeling data losses

The purpose of the study was to understand how the UDFCD system would perform under critical conditions. The ideal test scenario would be an intense, localized storm that results in the need to identify the specific area and severity of problems that require immediate public safety and infrastructure protection responses. In order to assure the credibility of the event, the study sponsors decided to apply a storm that had actually occurred in the District during the study period, but was located a few miles outside the primary distribution of gages.

The selected storm occurred on July 8, 2001, over southern Denver and parts of Arapahoe and Douglas Counties. The 5-minute NWS radar record had previously been reviewed and validated by Henz Meteorological Services. Chad Kudym of UDFCD used GIS tools to place the storm over a dense concentration of gages in the eastern Denver and Aurora area. Each gage was paired with a radar pixel and a 5-minute radar reflectivity value was assigned to each one for the 5-hour storm period. Rainfall was then assigned to each gage in 5-minute increments based on the following Z-R relationship, Table 4:

Reflectivity dBZ	Incremental Rainfall, inches
< 25	0.00
25	0.04
30	0.08
35	0.16
40	0.24
45	0.28
>45	0.36

Table 4. Radar reflectivity to rainfall conversion algorithm.

Rain reports

Rain reports were generated for each gage by distributing the number of 0.04-inch tips equally across the 5-minute period. A synthetic data record was created using sensor ID and a sequential accumulator value. The model storm produced the rainfall distribution shown in Table 4, below. The peak rainfall total was 6.64 inches at Iliff pond, with a system average rain of 2.23 inches. The storm's design led to a maximum 5-minute rainfall of 0.36 inches, or 4.3 inches per hour. The historical data record suggests peak rainfall rates would have been greater, which would likely cause a higher peak loading and data loss under real-world conditions. Table 5, below shows the rainfall distribution for the model storm:

Rainfall Distribution	
Amount	Number of Gages
No rain	7
0.01 thru 0.99 inch	58
1.00 thru 1.99 inches	43
2.00 thru 2.99 inches	29
3.00 thru 3.99 inches	29
4.00 thru 4.99 inches	17
5.00 thru 5.99 inches	15
6.00 or more inches	3

Table 5. Model storm rainfall storm totals distribution.

The storm was centered over the highest concentration of gages, which produced a highly skewed distribution of reports among repeaters; as can be seen in Table 6, below, almost half the rain reports were directed to Smokey Hill.

Model Storm Rain Reports			
Repeater	Reports	Inches	Percent
Smokey Hill	4701	188.04	49%
Blue Mountain	3827	153.08	40%
Gold Hill	728	29.12	8%
Lee Hill	361	14.44	4%
Total	**9617**	**384.68**	**100%**

Table 6. Model storm rain reports by repeater.

Estimation of stage gage traffic

ALERT reports are generated in response to changes in stage, and the peak reporting frequency occurs at the peak rate of change in level. If a unit hydrograph applies to a stage site, then a "unit reporting graph" could, in principle, be developed for the ALERT gage at that site, based on the derivative of the unit hydrograph. The report rate will peak during the rising limb, fall off during the crest, increase again in the early phase of the falling limb, then decline slowly over time. In practice, turbulence adds significantly to reporting frequency at higher levels. Nonetheless, the historical records show that peak reporting frequency occurs on the rising limb of the unit hydrograph, with the implication that gage reporting peaks well ahead of the flow. For small catchments, the stage gage loading may add significantly to rainfall loading while rain is still occurring, even in short storms. In the UDFCD system, significant simultaneous contributions of rain and stage reports will always occur if the duration of intense rain exceeds 30 minutes.

The distribution of reports was not broken down by gage for the model. Rather, we developed a set of criteria to predict the composite loading and distribution of stage gage reports based on patterns seen in the historical analysis.

At a single repeater, the peak stage reporting lags the peak rain by only 30 minutes, but when viewed on a system-wide basis, the peak stage reporting lags the peak rain by slightly more than one hour. About half of all storm-generated stage reports occur in the first 5 hours, and 90% occur in the first 8 hours. (The period considered in the model storm is 5 hours). Stage reports begin to increase from the start of the rain and reach cumulative totals of 20% at the rain peak and 40% at the stage report peak.

Estimating the total number of stage gage reports for the model storm presented some challenges. Over the composite of the smaller storms analyzed, each rain gage report generated approximately 1.2 stage reports over the ensuing 8 hours, although significant variations occurred due to storm location and intensity. However, there is a limit to which this relationship can be scaled. Most UDFCD stage gages are set to report on change, but not more frequently than once per minute. Therefore the theoretical maximum hourly report rate for the 91 gages in the system is probably about 5,460, and we concluded it is unlikely that even 60% of this rate would be achieved under the conditions of the model storm. We modified the distribution pattern to limit the reporting rate, applying constraints that increase progressively with traffic. This constrained the peak stage reporting rate in the model storm to 2,762 reports per hour, and affected the 5-minute values according to the reporting rate. The stage gage reports were distributed among repeaters in proportion to the rain gage reports received at each repeater.

Estimation of base level traffic

On a non-storm day, the UDFCD ALERT traffic averages about 250 reports per hour. This is generated by the twice-daily rain gage reports, the weather stations, stream gage activity under quiet conditions, and a few other regular status reports. For the model storm, we assumed the background level of all but the stage reports at a rate of 200 per hour. In addition, we included 400 reports per hour as an elevated background stage reporting generated by a recent prior event. This traffic was included in estimating the limiting number of stage gage reports during the peak of the model event.

Results of the Model Storm

Rain, stage and base traffic loads were summed by 5-minute period. During the peak of the model storm, traffic loading reached a rate of 7,400 reports per hour, three times what was observed in the historical storms studied. The peak loading is strongly driven by rain reports; the impact of stream reports is diminished by time limitation on reporting frequency. It becomes evident that background traffic, even when elevated, has little effect on the system during critical periods. The loading is shown in Figure 3, below.

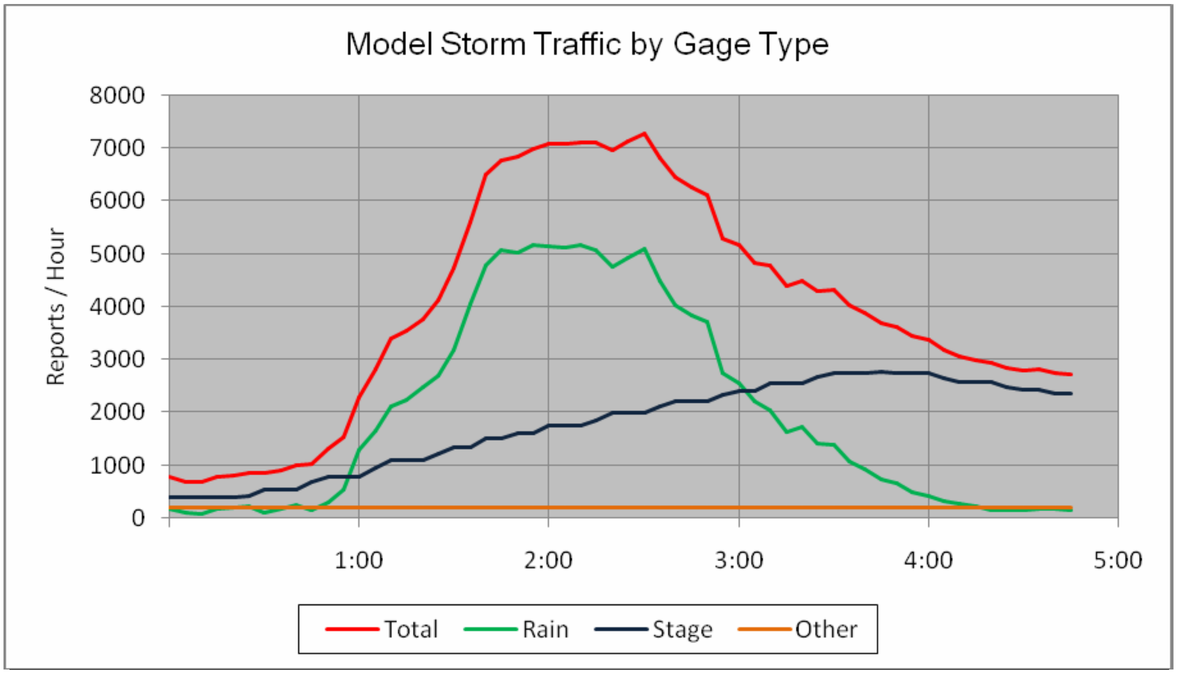

Figure 3. Model storm traffic by gage type.

Overall contention losses were estimated by the Poisson model and are graphed below in Figure 4. The beginning of contention collapse – 5,400 reports per hour – is reached 1 ½ hours into the event and throughput reaches its peak of slightly less than 2,000 reports per hour.

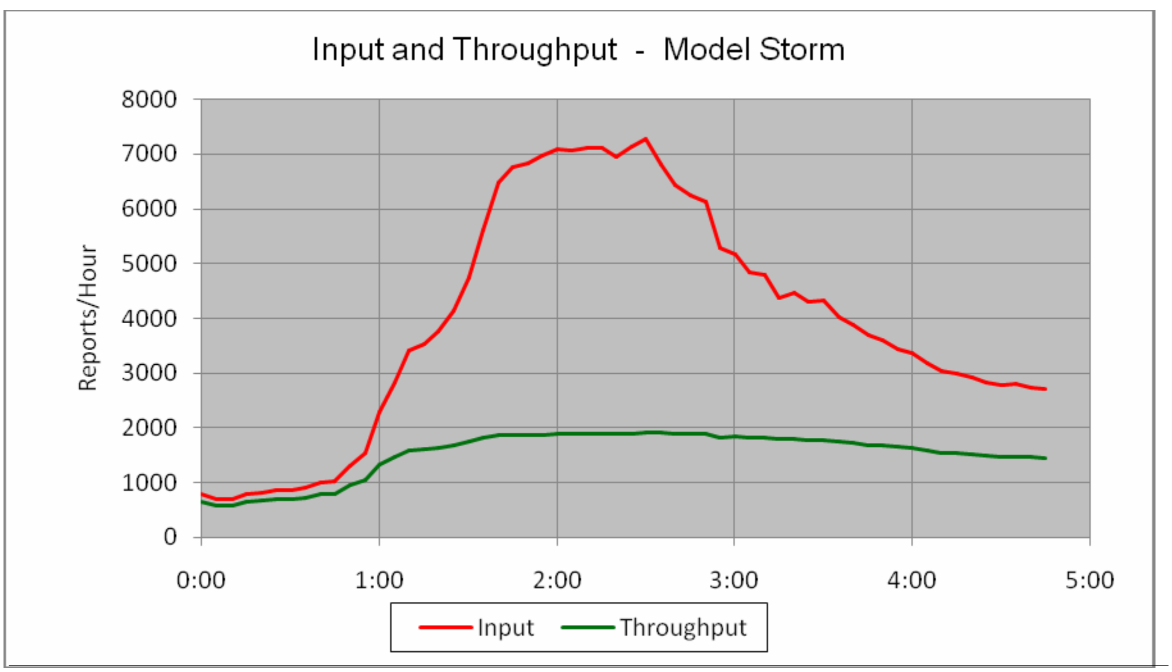

Figure 4. Model storm input traffic and throughput.

Peak system-wide data losses were greater than 74%, with 71% loss for the central 90 minutes of the storm. Because the storm affected mainly the Smokey and Blue repeater areas, traffic losses were heavier in these areas of greatest concern. Gages routed through Smokey lost 78.5% of all reports during the peak 5 minutes, and more than 74% during the central storm period.

At these loss rates, the impact on timeliness of data can be severe. Many weather values are reported at 15-minute intervals; during the critical part of the storm, the temperature and dew-point values from about one third of the weather stations would be more than an hour old. During the same time, the chance is one-in-four that a stream level alarm would be delayed by thirty minutes or more for a gage reporting at a 5-minute frequency.

In addition, gages were grouped by their repeater, and 5-minute loadings and input losses were determined for each repeater. The results are graphed in Figure 5, below. Because the storm was localized, Smokey and Blue Mountain input loadings vastly exceeded those of Gold Hill or Lee Hill. This resulted in much higher contention losses on the input side of the busier repeaters. Since the rate of loss is equal across all repeaters on the output frequency, losses are higher in the areas most affected by localized storms.

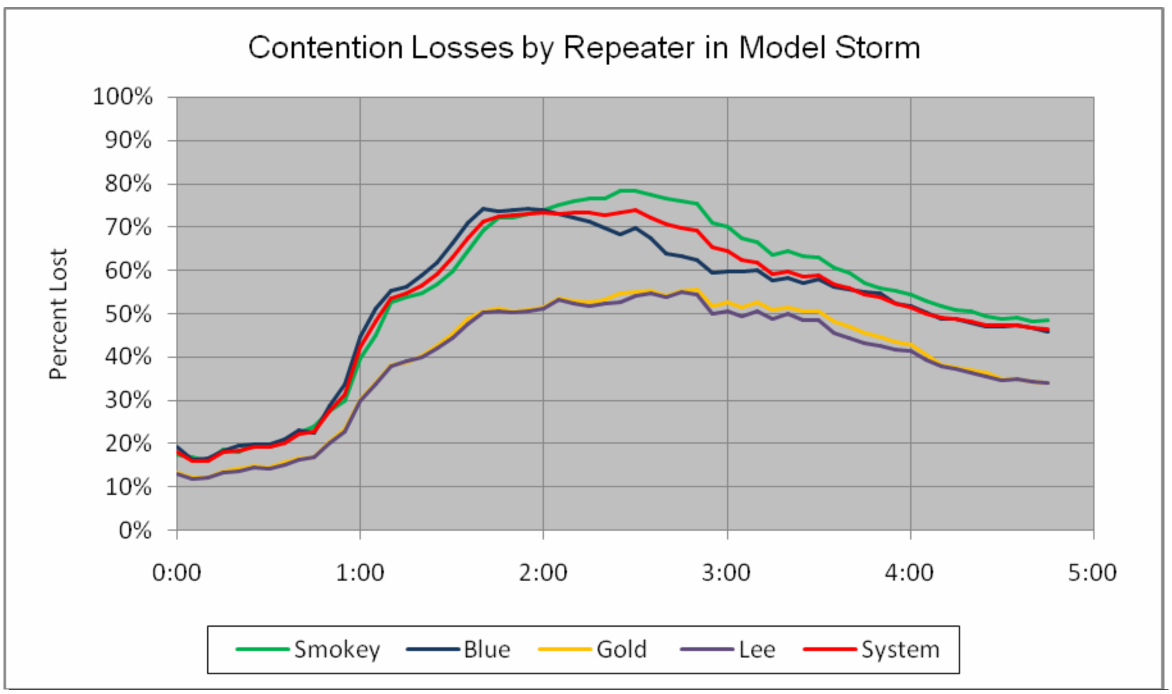

Figure 5. Model storm contention losses by repeater.

Figure 6, below, shows the repeater input losses for the traffic generated by the model storm.

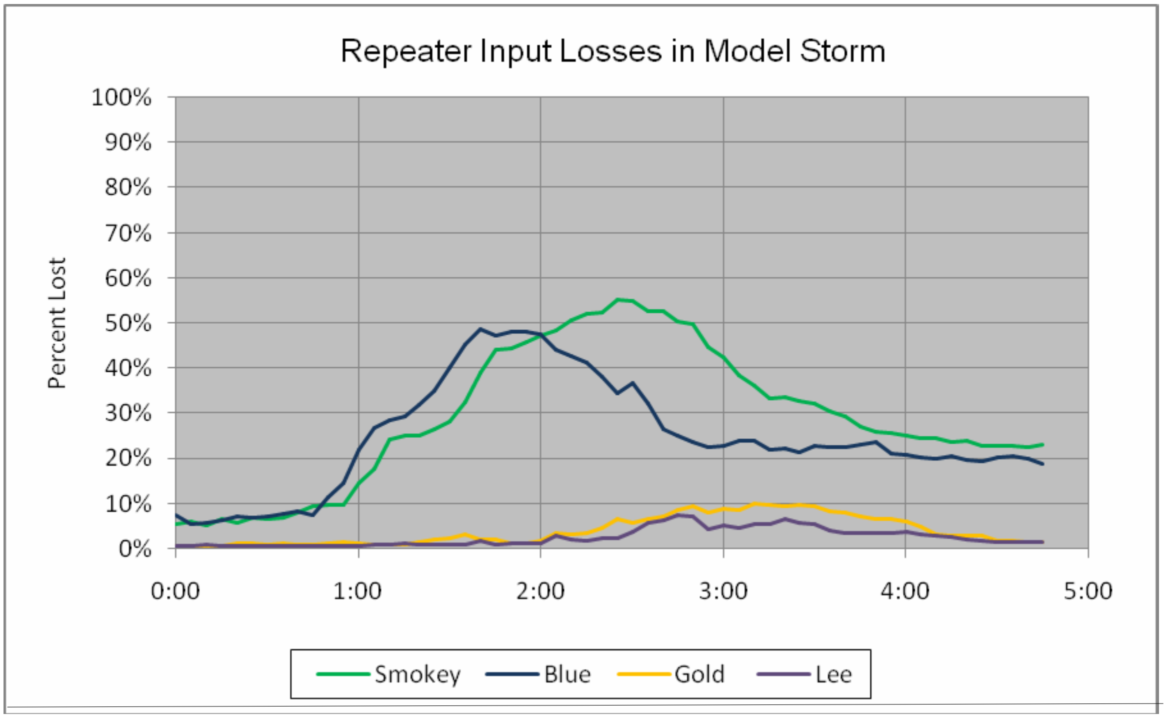

Figure 6. Model storm repeater input losses.

Software rainfall estimates

The rain validation process was modeled by assigning a probability of successful transmission to every rain gage report throughout the storm. This probability took into account the loss on the input and output frequencies at the time of the report, as well as the variability in loss shown historically by each gage. Each report was then assigned a random number between zero and one, and if that number was smaller than the report's probability of success, it was assigned a "missing" status.

DIADvisor and NovaStar software use different rain validation algorithms and produce slightly different rainfall estimates. The list of "received" rain reports was processed using both the DIADvisor and Novastar rain validation algorithms. The performance for the model storm is shown in Table 7, below, using the limits that are presently in use by UDFCD for each software package. The rain total is the sum of the rain at all gages.

Rainfall Estimates in Model Storm			
Reports Generated		100.00 %	384.68 inches
Reports Received		29.66 %	135.60 inches
DIADvisor	Limit 5	61.1 %	235.00 inches
Novastar	Limit 3,6,9	52.1 %	200.28 inches

Table 7. Model storm input traffic and throughput.

Rainfall estimation by the base station software suffered severely. During the central part of the storm, the estimated rainfall was about half of the actual amount, and the overall storm capture was only slightly better.

Both software products permit adjustment of the acceptance criteria for sequence gaps, and three values were tested for each, as shown in Table 8, below.

Results from Alternative Rain Validation Limits			
Software	Limit	% Captured	Inches
DIADvisor	5	61.1 %	235.00
DIADvisor	7	74.2 %	285.32
DIADvisor	10	85.8 %	330.24
Novastar	3,6,9	52.1 %	200.28
Novastar	4,8,12	64.4 %	247.80
Novastar	5,10,15	72.5 %	279.00

Table 8. Model storm base station results for three different levels of base station software acceptance criteria parameters.

Relaxing the criteria increases rain capture in each case, but this is done at the expense of accepting more erroneous reports that can skew rainfall values and create other problems. The difference in performance of the two algorithms is not indicative of the superiority of one algorithm over the other – only that the selected acceptance criteria are not equivalent. The point of the analysis is that neither algorithm is capable of overcoming the gaps in sequence that will occur under model storm conditions.

Remediation options

In order to perform as expected, the UDFCD system should be configured to receive at least 7,500 reports per hour, while total losses through the most heavily affected repeater should not exceed 30%.

Several options that might improve system performance during critical periods were identified and evaluated. The tools developed to predict losses and quantify system performance are very helpful in assessing possible solutions, and the results were sometimes non-intuitive.

Add data channels

The most effective first step, both in time and cost, is to increase capacity by adding new channels. To be effective, channels must be added simultaneously on the input and output sides of a repeater; model analysis shows there is little gain in addressing either one alone.

Our recommendation for a new system configuration is to license a second output frequency. In Boulder County, one of the two repeaters would be moved to the new frequency, and the other left unchanged. Initially, the Boulder system could continue to operate with just two repeaters, largely because rain gages are more geographically intermingled in their repeater destination. In the Denver area, a second repeater on the new output frequency could be co-located with each of the existing Smokey and Blue repeaters.

Two input frequencies are presently used in the UDFCD system: One is used predominantly for Boulder gages, and the other for the Denver area. Better use could be made of these frequencies by intermingling their assignment throughout the system; in any localized event, the input loading would be distributed across two channels. This change can be accomplished at little cost with most transmitters requiring only a switch change or a radio programming update.

At Smokey and Blue, the new repeater would monitor the "new" input frequency, and the pass lists on both repeaters would be adjusted accordingly. The result would be that each output channel would carry traffic from the entire area, but each would carry half the loading. In addition to reducing contention, the impact of a repeater failure is lessened because no area is completely blacked out.

This system change would bring the peak contention losses down to 55% during the model storm. This is a significant improvement, but it falls well short of the goal of 30% losses. Adding a third channel to the system is an option, but implementation would be much more costly than the second. Space and power constraints would be significant at the existing sites, and two new frequencies would be required. We believe other options should be considered first.

Repeater carrier detect before transmit

To the extent that a repeater can hear the output traffic of the other repeaters, contention can be reduced by waiting for a clear channel before transmitting stored messages. This requires that the repeater receive and monitor the output frequency, a capability that has not been implemented in the existing repeaters.

This feature should be considered in new repeaters. However, its benefits may break down at very high traffic loadings, since the holding time may expire before a clear channel is available. The efficacy of this approach should be modeled as part of establishing the design criteria for this feature. Its availability depends on the hardware manufacturer's willingness to implement it as part of their product.

Phase 1 implementation of ALERT 2 protocol

As presently planned, the new ALERT 2 protocol will operate at 4800 baud with an error-corrected data payload of 256 bits and a packet duration of 200 milliseconds. The new protocol has a data throughput improvement of 18-fold, if the 32-byte data payload is fully utilized. Even without considering the packet contents, the shorter message duration reduces channel occupancy by 40%; maximum throughput increases to 3,300 messages per hour and contention collapse begins at 9,000 instead of 5,400 messages per hour.

Figure 9, below, compares calculated throughput by input traffic load for a single-threaded ALERT repeater system, an optimized dual-frequency-based ALERT repeater system and an ALERT 2-capable repeater system.

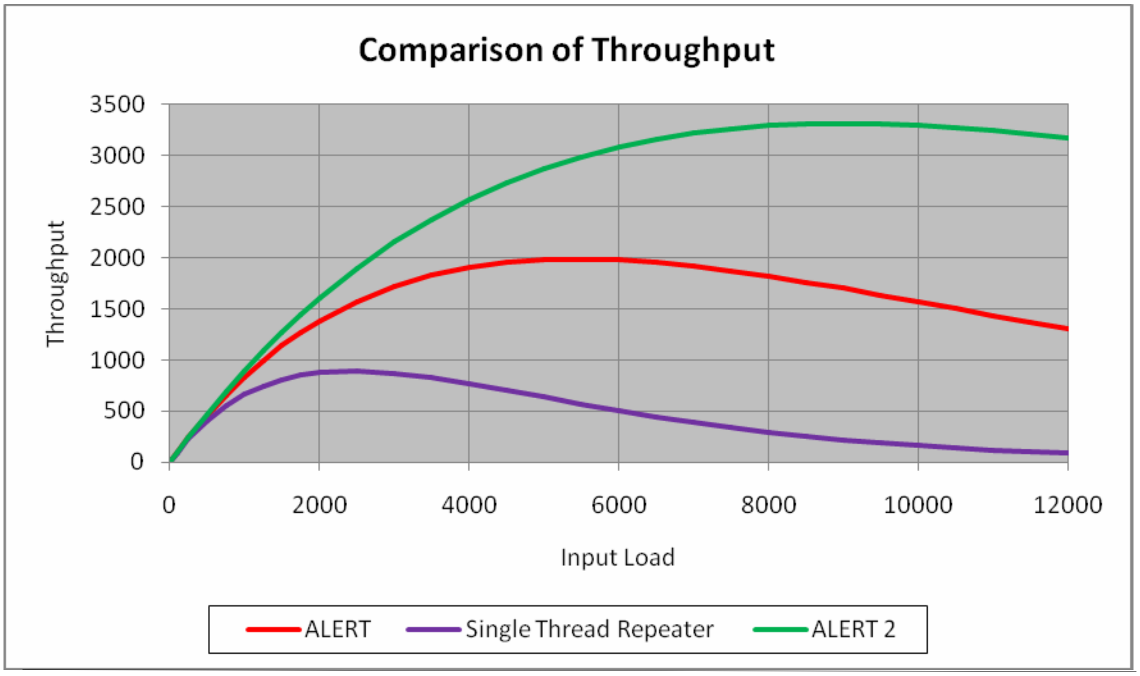

Figure 9. Comparison of throughputs by input load for three different system architectures and protocols.

Moving fully to ALERT 2 will eventually resolve the loading issues in the UDFCD system, but this is a long-term process: It will require replacing all gage transmitters in the system, the hardware is not yet available, and the protocol is not yet finalized. Still, the promise of the new protocol is so great that we believe focusing on its implementation will be more effective than a middle-term patchwork of marginally effective fixes using hardware that will soon become obsolete.

We expect that the UDFCD transition to ALERT 2 will occur in two phases. In Phase 1, the output channel will operate on the new protocol, with base station receiver-decoders converted to ALERT 2. Repeaters will be able to receive either protocol, and will repeat all traffic in ALERT 2 format. This will yield some significant benefits while requiring hardware changes only at repeaters and base stations.

With a 32-byte payload available, the best way to transport 4-byte ALERT messages would be to combine several of them into one ALERT 2 packet at the repeater. For example, each repeater could wait up to15 seconds to transmit a message. Six or more ALERT messages could be sent in a single 200 millisecond packet and when loading reaches 1,500 reports per hour, most packets would be full. At the receiver, the packets would be resolved back into discrete 4-byte ALERT messages. During critical times, two seconds of ALERT data would be sent in 200 milliseconds; this would fully address contention issues on the output channel. However, there would be no change in contention on the input channel.

Development of an ALERT 2 concentrating repeater is technically feasible in the near term. It is a logical early implementation of the new ALERT 2 physical layer and need not await finalization of the ALERT 2 application layer. It should be implemented in addition to adding a second channel, and we believe it could be ready soon enough to combine both changes. This would eliminate the need to purchase new repeaters constrained to the legacy ALERT format.

Figure 10, below, shows the data loss that would occur in the model storm under a variety of system configuration alternatives. The existing network is shown in red. Adding a second ALERT input and output channel would reduce peak losses from 74% to 55% (green line). If the concentrating ALERT 2 repeaters were used in the dual channel system, peak losses would reach only 31% (violet line).

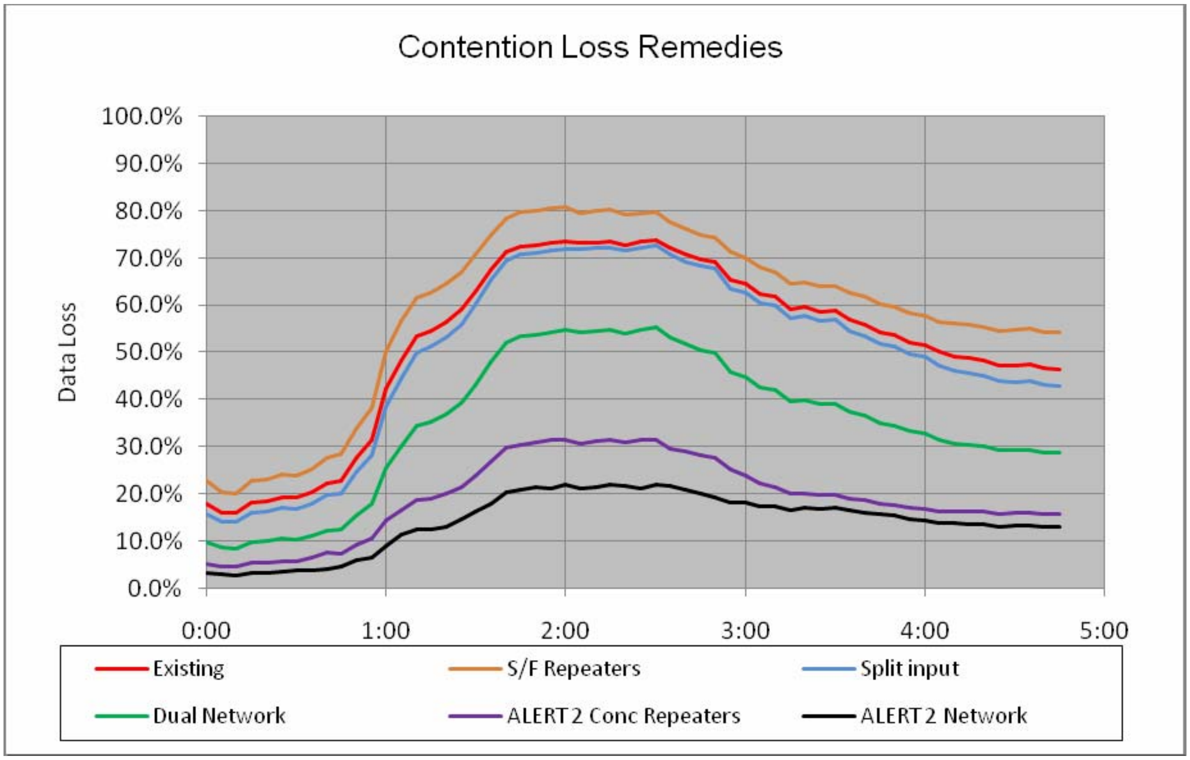

Figure 10. Comparison of contention-based data losses for various system remedies.

Phase 2 implementation of ALERT 2 protocol

As gages are replaced with ALERT 2 equipment, full advantage can be taken of the new protocol. Contention will be reduced on the input channel because of the shorter message length, but this is only one of the benefits. Other advantages are:

- The 32-byte message capacity will permit all sensors at a site to report in a single message, so co-located rain and stage sites will generate substantially fewer messages during peak activity.
- Error correction will permit us to remove the software limit now placed on jumps in rain accumulator value, so rain amounts will be estimated with virtually 100% accuracy. It will therefore be possible to limit the transmission rate to once per minute per site, further reducing system traffic without sacrificing data timeliness.

In a fully ALERT 2, dual-channel system, UDFCD data losses in the model storm would peak at a little over 20%. About 4% of the data would have gaps of 1 report, and 1%

would have gaps of 2 reports. During peak periods of model storm conditions, 99% of all rain and stage data would be within 4 minutes of current. Storm events much larger than the model storm could be monitored with system performance better than that seen in recent historical events.